INTEGRACIÓN DE MLOPS EN LA ADMINISTRACIÓN DE RIESGO DE MODELOS: UN ENFOQUE INNOVADOR PARA LA FIABILIDAD Y ROBUSTEZ EN MODELOS PREDICTIVOS

Integration of MLOps in Model Risk Management: An Innovative Approach to Reliability and Robustness in Predictive Models

JOSE LUIS CASTEÑADA RIOS

Tesis

Asesor
Juan David Ospina Arango

UNIVERSIDAD EAFIT

ESCUELA DE CIENCIAS APLICADAS E INGENIERÍA

MAESTRÍA EN CIENCIAS DE LOS DATOS Y LA ANALÍTICA

MEDELLÍN

2024

Abstract

Este documento explora la propuesta de integración de metodologías de Operación de Modelos de Aprendizaje de Maquina (Machine Learning Operations - MLOps) en la Administración de Riesgo de Modelo (Model Risk Management - MRM) para promover la gestión efectiva de riesgo. A través de una revisión sistemática de la literatura, se examinan definiciones y prácticas actuales de MLOps y MRM, destacando su aplicabilidad en diversas industrias, y no exclusivamente para el sector financiero. El estudio identifica los desafíos clave de la MRM, como la validación de modelos y la calidad de los datos, y propone soluciones basadas en principios de MLOps y sus herramientas. Esta propuesta sugiere que la temática desarrollada en este documento puede reportar algún grado de novedad, y que además MLOps puede abordar efectivamente los retos de MRM, mejorando la administración del ciclo de vida de los modelos y asegurando el cumplimiento regulatorio, lo que es crucial para la estabilidad y eficiencia operativa de las instituciones.

1 Introducción

La inteligencia artificial (IA) y el aprendizaje automático (ML o Machine Learning) están transformando múltiples industrias, impulsando el uso de modelos predictivos. Sin embargo, este crecimiento también plantea preocupaciones sobre la fiabilidad y robustez de estos modelos, especialmente en la administración de sus riesgos de uso y operación. En este sentido, este tema se extiende en la práctica a las decisiones que se toman utilizando las salidas de modelos predictivos, expuestos riesgos en su desarrollo y mantenimiento en producción, que de no ser administrados, se incurren en pérdidas económicas (y no económicas) derivadas en las decisiones erradas influenciadas por los cambios inadvertidos y adversos sobre las predicciones de un modelo.

Las metodologías de Operaciones de Aprendizaje Automático (MLOps o Machine Learning Operations) ofrecen soluciones para integrar prácticas de desarrollo y operación, y su aplicación en la administración de riesgo de modelo (MRM o Model Risk Management) puede ser provechosa, mas sin embargo requiere una investigación y explicación detallada.

La meta principal de este documento por lo tanto, es responder a la pregunta: "¿Cómo pueden las metodologías de MLOps integrarse efectivamente en las prácticas de administración de riesgo de modelo para mejorar la fiabilidad y la robustez de los modelos de aprendizaje automático?". Para esto, se realiza una revisión sistemática de la literatura para recopilar una base de conocimiento actualizada y objetiva. Además, se plantean dos objetivos secundarios: ofrecer una versión actualizada de las definiciones y referencias teóricas sobre MLOps y MRM, y mostrar con algunas analogías que el riesgo de modelo es una preocupación universal para todas las empresas, no solo para el sector financiero. Los efectos de una mala administración operativa de un modelo son transversales a cualquier industria.

2 Objetivo

El objetivo principal de este documento es resolver la pregunta de investigación: "¿Cómo pueden las metodologías de MLOps integrarse efectivamente en las prácticas de administración de riesgo de modelo para mejorar la fiabilidad y la robustez de los modelos de aprendizaje automático?"

Además se establecen dos objetivos secundarios que ambientan el desarrollo de este documento. Primero, ofrecer una versión actualizada de las definiciones y referencias teóricas para las Operaciones de Modelos de Aprendizaje Automático (MLOps) y la Administración del Riesgo de Modelo (MRM). Esto incluye una revisión sistemática de la literatura existente y una clarificación de los conceptos clave relacionados con estos campos.

También se pretende proponer que el riesgo de modelo es una preocupación universal para todas las empresas, y no es exclusivo del sector financiero. Se utilizará la analogía de que los efectos indeseados de una mala administración operativa de un modelo son transversales a cualquier industria, demostrando que el riesgo de modelo afecta de manera similar a diversas organizaciones, independientemente de su sector.

3 Metodología

La metodología empleada en este documento comienza con una revisión sistemática destinada a explorar la relación entre MLOps (Machine Learning Operations) o bien Operaciones de Aprendizaje Automático y MRM (Model Risk Management) o bien la Administración de Riesgo de Modelos (MRM). Las revisiones sistemáticas son una forma rigurosa y estructurada de revisión de la literatura que permite la síntesis comprensiva de investigaciones existentes, garantizando un alto nivel de fiabilidad y validez en los hallazgos. Este enfoque es particularmente ventajoso para identificar brechas en el conocimiento actual, sintetizar evidencias a través de múltiples estudios y desarrollar preguntas de investigación robustas.

El proceso de revisión sistemática empieza con una estrategia de búsqueda exhaustiva diseñada para identificar estudios relevantes sobre la relación mencionada. La base de datos utilizada fue Scopus [1], una base de datos de resúmenes y citas para revistas académicas, libros y actas de conferencias, ampliamente utilizada para la búsqueda y evaluación de investigaciones.

La primera ecuación de búsqueda tiene la intención de reunir documentos sobre la relación explícita entre los temas de estudio, y es la siguiente:

```
(
    (
        TITLE-ABS-KEY(
            ( "machine learning operations" OR "mlops")
            AND
            ("model risk management" OR "mrm")
        )
    ) AND (
        LANGUAGE (english)
    ) AND (
        PUBYEAR > 2012
    )
) AND NOT (
    SRCTYPE ("Conference proceeding")
)
```

La búsqueda apunta a la presencia de los términos “machine learning operations” o “mlops” combinados con “model risk management” o “mrm” en el título, resumen o palabras clave de los documentos. Estos deben estar escritos en inglés, y solo se incluyen aquellos publicados después de 2012 en los resultados de búsqueda. Esta consulta excluye los clasificados como “Actas de conferencias” para centrarse en artículos de revistas y otros tipos de fuentes. Esta búsqueda arrojó cero resultados, lo que no solo muestra evidencia de una posible brecha de investigación a desarrollar, sino que también sugiere un cambio en la estrategia de búsqueda.

Los temas objetivo se buscaron en consultas similares a la primera, pero se separan de la siguiente manera:

```
(
     (
        TITLE-ABS-KEY(
            ("machine learning operations" OR "mlops")
            AND
            ("state of the art" OR "review")
        )
     )
     AND
     (LANGUAGE(english))
     AND
     (PUBYEAR > 2012)
)
AND NOT
(SRCTYPE("Conference proceeding"))
```

Y también se consultó:

```
(
     (
        TITLE-ABS-KEY (
            ("banking" OR "financial institutions")
            AND
            ("model risk management" OR "mrm")
        )
     )
     AND
     (LANGUAGE(english))
     AND
     (PUBYEAR>2012)
)
AND NOT
(SRCTYPE("Conference proceeding"))
```

La búsqueda relacionada con MLOps arrojó 69 artículos, mientras que la búsqueda relacionada con el tema de administración de riesgo de modelo solo arrojó 16 (ver Anexos). Esta última se modificó para incluir términos relacionados con la banca y las instituciones financieras, ya que sin ellos los resultados eran ampliamente diversos y no relacionados con el tema al revisar. Por lo tanto el alcance poblacional de esta elaboración se acota entonces a instituciones financieras y bancos, mencionando de antemano que el riesgo de modelo también existe por analogía en otras industrias, asunto que se desarrollará mas adelante. Un detalle a mencionar es que a pesar de excluir de manera explicita los items de conferencias, el resultado incluyó elementos relacionados a lo que se presume que son artículos con carencias en el etiquetado. Estos se evaluaran según los criterios de inclusión.

Una vez identificados los estudios relevantes en el contexto del buscador, se procedió a revisar el resumen para incluir solo aquellos relacionados con el objetivo de este documento, lo que resultó en una base de documentos final de 33 elementos. Sin embargo, se menciona que fueron necesarios otros artículos más allá de esta base para elaborar secciones adicionales para hacer referencia a regulaciones, metodologías y herramientas.

Los conocimientos obtenidos del proceso de revisión sistemática fueron fundamentales para formular la pregunta de investigación de este estudio. La revisión revela una brecha significativa en la literatura respecto a la integración de metodologías de MLOps para resolver los desafíos de la administración de riesgo de modelo. Esta brecha por lo tanto, motivó el desarrollo de la pregunta de investigación: "¿Cómo pueden las metodologías de MLOps integrarse efectivamente en las prácticas de administración de riesgo de modelo para mejorar la fiabilidad y la robustez de los modelos de aprendizaje automático?"

Tomando como base la pregunta de investigación, se desarrolló el tema de este documento para explorar el impacto teórico (y actualizado) de estrategias prácticas al integrar MLOps en la administración de riesgos de modelos. Esto implica examinar evaluar marcos existentes y referir a nuevos modelos para operacionalizar la administración de riesgos en el aprendizaje automático.

A continuación se propondrán los fundamentos teóricos y definiciones necesarias para desarrollar el objetivo del documento.

4 Operaciones de aprendizaje automático (MLOps)

El aprendizaje automático (ML) ha emergido como una tecnología transformadora, permitiendo el desarrollo de sistemas de predicción sofisticados en múltiples industrias. Sin embargo, el despliegue rápido de sistemas de ML a menudo incurre en una deuda técnica significativa y oculta, lo que puede complicar el mantenimiento y la evolución de los mencionados con el tiempo. Este concepto, discutido extensamente por Sculley et al. (2015) [2] en su artículo seminal, describe los desafíos únicos que presentan los sistemas de ML que van más allá de los encontrados en la ingeniería de software tradicional.

4.1 El origen: la deuda técnica en los sistemas de ML

La deuda técnica, una metáfora introducida por Ward Cunningham[2], se refiere a los costos a largo plazo incurridos debido a decisiones expeditivas pero subóptimas en el desarrollo de software. En los sistemas de ML, esta deuda se manifiesta en múltiples formas, principalmente debido al enredo de dependencias de datos, dependencias de modelos e interacciones a nivel de sistema.

- **Modelos Complejos Erosionan los Límites**: A diferencia del software tradicional, los sistemas de ML no tienen límites de encapsulación estrictos. Los cambios en una parte del modelo, como la distribución de datos de entrada, pueden propagarse y afectar el rendimiento del sistema en su totalidad de manera impredecible. Esta erosión de los límites lleva a costos de mantenimiento incrementados a medida que el sistema evoluciona. Un termino asociado regularmente a este fenómeno es la des calibración del modelo, que no es mas que la degradación de su rendimiento con respecto a una métrica.

- **Enredo**: Los sistemas de ML mezclan inherentemente varias señales, lo que dificulta aislar y mejorar componentes individuales sin afectar a otros. Este principio, conocido como "Cambiar Cualquier Cosa Cambia Todo" (CACE, por sus siglas en inglés), complica la estabilidad del modelo y la consistencia del rendimiento. Otra manera de ver esto es que no hay independencia entre los componentes de un modelo, y es difícil hacer un cambio particular sin afectar lo general.

- **Consumidores No Declarados**: Las predicciones de los modelos de ML a menudo son utilizadas por múltiples sistemas, productos o clientes. Sin controles de acceso adecuados, estos consumidores no declarados crean acoplamientos estrechos, haciendo que cualquier modificación del modelo sea arriesgada y potencialmente cause efectos secundarios no deseados en otros sistemas. Visto de otra manera, un consumidor no declarado puede verse afectado si no está al tanto de los cambios de fondo o de forma de los resultados de un modelo.

- **Dependencias de Datos**: Las dependencias de fuentes de datos, que pueden cambiar con el tiempo o ser controladas por diferentes equipos, introducen inestabilidad. Estas dependencias son más difíciles de gestionar en comparación con las dependencias de código debido a la falta de herramientas robustas para el análisis estático de flujos de datos.

4.1.1 Mitigación de la deuda técnica

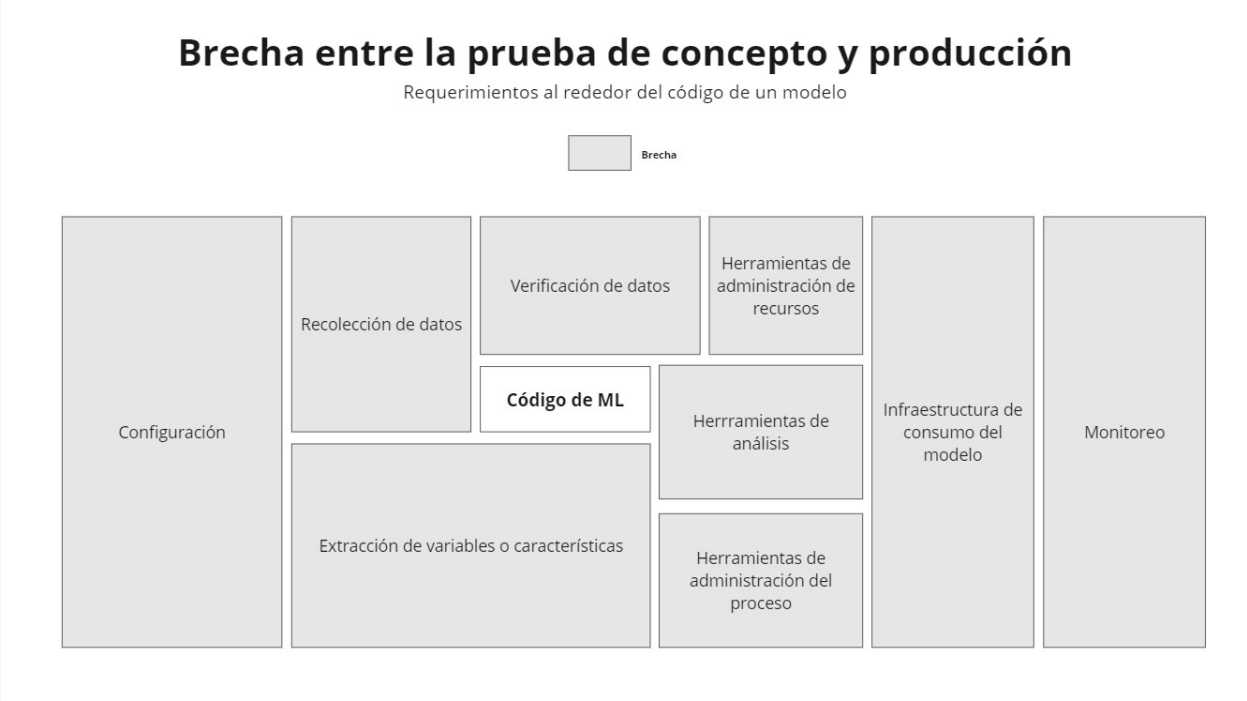
Para abordar estos problemas, Sculley et al.[2] proponen varias estrategias:

- **Aislar Modelos**: Servir modelos como productos puede ayudar a gestionar el enredo aislando subproductos. Este enfoque, sin embargo, requiere una administración cuidadosa para evitar nuevas formas de dependencia y enredo.

- **Dependencias de Datos Versionadas**: Crear copias versionadas de fuentes de datos puede mitigar los riesgos asociados con dependencias de datos inestables. Esta estrategia asegura que las actualizaciones de fuentes de entrada no afecten inesperadamente el rendimiento del modelo.

- **Monitoreo y Pruebas**: El monitoreo en tiempo real y los sistemas de respuesta automatizados son críticos. Estos sistemas deben rastrear métricas clave como sesgo de predicción y límites de acción para asegurar que el comportamiento del modelo permanezca consistente y fiable. A esto se debe agregar que deben existir otros objetivos de monitoreo propios del desarrollo de software, tales como consumo de recursos, velocidad, latencia, entre otros.

De esta manera, y utilizando estos conceptos se entiende que existen otros asuntos al rededor del código de ML necesarios para llevar un modelo a operaciones productivas. Estos componentes adicionales al código de entrenamiento y predicción también son conocidos como "POC to Production Gap" o brecha entre la prueba de concepto y producción. Esta brecha es la que fundamenta las necesidades de la existencia de una metodología para llevar modelos a producción[2].



4.2 Definición de MLOps

La noción presentada sobre la deuda técnica oculta establecen entonces una comprensión fundamental para el campo de MLOps. Este abarca prácticas que buscan agilizar el despliegue, monitoreo y mantenimiento de modelos de ML en entornos de producción. De manera consecuente, estas metodologías están diseñadas para:

- **Mejorar la Reproducibilidad**: Asegurar que los experimentos puedan ser replicados de manera fiable es crucial para mantener la integridad del modelo. Las técnicas para gestionar algoritmos aleatorios y condiciones iniciales son esenciales para este propósito.
- **Mejorar la Administración de Configuración**: Los sistemas de configuración efectivos que minimizan errores manuales y permiten una fácil especificación de cambios son vitales. Esto reduce la sobrecarga asociada con la administración de sistemas de ML complejos.
- **Facilitar el Monitoreo Continuo**: Implementar herramientas de monitoreo robustas que proporcionen información en tiempo real sobre el rendimiento del modelo y automatizar las respuestas a problemas detectados asegura la fiabilidad a largo plazo de los sistemas de ML.

Por lo tanto, las Operaciones de Aprendizaje Automático son un conjunto de prácticas destinadas a unificar los aspectos de desarrollo (Dev) y operación (Ops) de los sistemas de aprendizaje automático (ML)[3]. Ayuda a extender los principios de DevOps para incluir el aprendizaje automático y la ingeniería de datos[4], abordando los desafíos específicos de desplegar y mantener modelos de ML en producción. Este enfoque integral involucra integración continua, entrega continua (CI/CD) y entrenamiento continuo (CT) para asegurar que los modelos de ML sean fiables, escalables[5] y puedan actualizarse eficientemente a medida que nuevos datos estén disponibles [6][7][8][9][10].

4.3 Componentes y principios clave

4.3.1 Integración y entrega continua (CI/CD)

La Integración Continua (CI) y la Entrega Continua (CD) son principios fundamentales tomados de DevOps que son esenciales en MLOps. CI implica la integración frecuente de nuevos cambios de código y datos en

la rama principal de un proyecto (en términos del versionado de código). Esta práctica ayuda a detectar y abordar problemas de manera temprana al probar continuamente los cambios contra el sistema [6][7]. Las canalizaciones de CI/CD en MLOps automatizan el flujo de trabajo de ML de extremo a extremo[10], desde la preparación de datos hasta el despliegue del modelo. Esta automatización reduce errores manuales, acelera el proceso de despliegue y asegura que los modelos de ML puedan actualizarse y mantenerse de manera fiable [6][8].

El Entrenamiento Continuo (CT) extiende CI/CD al incorporar el reentrenamiento de modelos de ML a medida que nuevos datos están disponibles. Esto asegura que los modelos permanezcan actualizados y continúen desempeñándose bien a medida que los datos subyacentes evolucionan [8]. CT automatiza todo el proceso de reentrenamiento de modelos y despliegue de versiones actualizadas, lo cual es crucial para mantener la relevancia y precisión de los modelos de ML con el tiempo[7].

4.3.2 Colaboración y comunicación

MLOps enfatiza también la importancia de la colaboración entre científicos de datos, ingenieros de ML y equipos de operaciones. Este enfoque multidisciplinario asegura que los procesos de desarrollo y despliegue estén alineados, llevando a sistemas de ML más robustos y escalables [8]. La comunicación efectiva entre estos equipos ayuda a abordar las complejidades de los sistemas de ML, que involucran tanto componentes de software como de datos. Esta colaboración se facilita a través de herramientas, prácticas y plataformas compartidas que respaldan todo el ciclo de vida de ML[7][9].

4.3.3 Automatización y orquestación

La automatización es una piedra angular de MLOps, y tiene como objetivo reducir las intervenciones manuales y agilizar el ciclo de vida de ML. Esto incluye la automatización de la recopilación y preprocesamiento de datos, el entrenamiento, la validación y el despliegue de modelos[7]. Herramientas de orquestación de flujos de trabajo como Apache Airflow[11], Kubeflow Pipelines[12] y SageMaker Pipelines[13][14] se utilizan comúnmente para gestionar estos flujos de trabajo complejos, asegurando que cada paso se ejecute de manera eficiente y fiable [8].

La orquestación no solo automatiza tareas repetitivas sino que también asegura que las dependencias entre diferentes tareas se gestionen de manera efectiva. Esto conduce a una utilización más eficiente de los recursos y reduce la probabilidad de errores que pueden ocurrir cuando las tareas se realizan manualmente [15][7].

4.3.4 Reproducibilidad y versionado

La reproducibilidad es crítica en MLOps para asegurar que los experimentos y resultados puedan ser reproducidos de manera fiable por diferentes miembros del equipo de manera independiente o en diferentes momentos. Esto involucra el versionado de datos, código y modelos para rastrear cambios y permitir la reversión si es necesario [8]. Los sistemas de control de versiones como Git[16] se utilizan para gestionar el código, mientras que herramientas como DVC [17] (Control de Versiones de Datos) se utilizan para versionar datos y artefactos de modelos[7]. Últimamente herramientas como MLFlow[18] y Weights And Biases[19] facilitan de manera específica la validación, versionado y disposición de modelos.

Asegurar la reproducibilidad facilita una mejor colaboración y transparencia dentro de los equipos. Permite a los científicos de datos y a los ingenieros construir sobre el trabajo de otros, verificar resultados y mantener un historial claro de cambios, lo cual es esencial para la depuración y mejora de los modelos [8].

4.3.5 Monitoreo y bucles de retroalimentación

El monitoreo continuo de los modelos en producción es esencial para detectar degradación del rendimiento, desvío de conceptos (drift) y otros problemas que puedan surgir con el tiempo. Las herramientas de monitoreo rastrean métricas clave de rendimiento y configuran alertas para notificar a los equipos sobre cualquier anomalía o disminución en el rendimiento del modelo[7]. Los bucles de retroalimentación automatizados utilizan estos datos de monitoreo para desencadenar reentrenamientos u otras acciones correctivas, asegurando que los modelos permanezcan precisos y efectivos [8].

Los bucles de retroalimentación son cruciales para mantener la salud de los sistemas de ML[20]. Permiten la mejora continua al proporcionar información en tiempo real sobre cómo los modelos están desempeñándose en producción y permitir que los equipos respondan rápidamente a cualquier problema que surja[7].

4.3.6 Madurez de una metodología de MLOPS

El marco de MLOps típicamente involucra tres canalizaciones (pipelines) principales: la canalización de datos, la canalización de modelos y la canalización de despliegues (release o deployment). Cada canalización automatiza tareas específicas como el preprocesamiento de datos, el entrenamiento de modelos y el despliegue. Un modelo de madurez de MLOps describe las etapas a través de las cuales una organización evoluciona en su adopción de estas prácticas[21], que van desde procesos manuales hasta sistemas completamente automatizados[8][9].

- **Recolección de Datos Automatizada**: Transición de la administración manual de datos a procesos automatizados de recolección de datos.
- **Despliegue de Modelos Automatizado**: Implementación de canalizaciones de CI/CD para un despliegue de modelos sin problemas.
- **Monitoreo Semiautomatizado de Modelos**: Primeros pasos hacia la automatización del monitoreo de modelos desplegados.
- **Monitoreo Totalmente Automatizado de Modelos**: Automatización completa de los procesos de monitoreo y retroalimentación, permitiendo la mejora continua de los modelos[7].

5 Administración de riesgo de modelo (MRM)

La administración de riesgos de modelos (MRM, por sus siglas en inglés en Model Risk Management) se refiere al proceso sistemático de gestionar los riesgos (posibilidades de pérdidas económicas y no económicas) asociados el conjunto de posibles consecuencias adversas derivadas de decisiones basadas en resultados e informes incorrectos de modelos, o de su uso inapropiado[22]. Estos modelos, que aplican teorías, técnicas y suposiciones estadísticas, económicas, financieras o matemáticas, procesan datos de entrada para producir estimaciones cuantitativas[23].

El término “modelo” abarca métodos cuantitativos que transforman entradas en estimaciones a través de un sistema de tres componentes: un componente de entrada de información, un componente de procesamiento y un componente de reporte[24].

Aquí entonces es posible afirmar que considerando la definición base de modelo y riesgo de modelo, existen desafíos análogos en otras industrias[25][26][27] pues, aquellas que usan modelos de ML se ven sujetas a los mismos riesgos[28] y por lo tanto a las mismas posibilidades de pérdida al materializarse alguno de ellos.

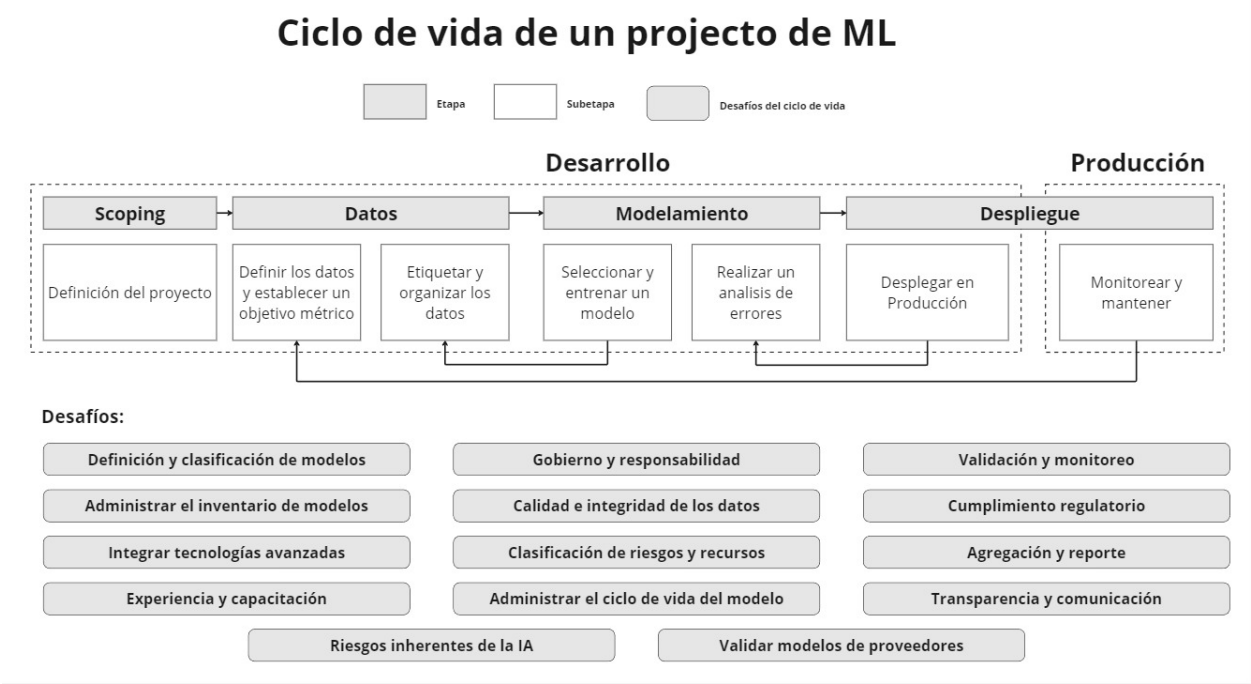


Figure 1: Ciclo de vida y desafíos de un proyecto de aprendizaje automatizado [29]

5.1 Relevancia

La importancia de MRM ha crecido sustancialmente debido al uso extensivo de modelos en varios aspectos de las industrias y la toma de decisiones con impacto financiero. Los modelos son fundamentales en el sector financiero para tareas como la evaluación de créditos, la valoración de exposiciones, la medición de riesgos y la determinación de la adecuación de capital[22]; no siendo esto el límite, también se utiliza para tomar decisiones sobre poblaciones de consumidores, criterios de seguridad y fraude, distribución de productos, entre otros. A medida que se expande el uso de estos modelos, también aumenta el riesgo de pérdidas financieras, decisiones incorrectas y daños reputacionales resultantes de salidas de modelos deterioradas o mal utilizadas[30][24]. De manera complementaria se puede confirmar que la normatividad asociada a la administración de este riesgo es más común para el sector financiero, ya que los resultados de la búsqueda sistemática están completamente asociados a este tipo de compañías[31].

5.2 Componentes clave

Desarrollo, implementación y uso de modelos:

1. **Desarrollo**: El desarrollo efectivo de modelos comienza con un propósito claro y alineación con el uso previsto. Incluye documentación rigurosa del diseño, teoría, metodologías y componentes de procesamiento. Los desarrolladores deben asegurar que los modelos sean conceptualmente sólidos y validados contra teorías y enfoques alternativos. El proceso de desarrollo también debe incluir pruebas y validaciones exhaustivas para asegurar precisión y fiabilidad bajo diversas condiciones[24].

2. **Implementación**: La implementación adecuada integra los modelos en sistemas de información más amplios, asegurando la integridad de los datos y la integración efectiva del sistema. Esta fase requiere pruebas continuas para verificar que los modelos funcionen como se espera bajo diversas condiciones. La implementación debe ser monitoreada para adaptarse a las cambiantes condiciones del mercado y necesidades comerciales[30].

3. **Uso**: El uso de modelos implica una evaluación regular del rendimiento y ajustes basados en retroalimentación y condiciones cambiantes. Los usuarios deben comprender las limitaciones del modelo y asegurar que los modelos se apliquen adecuadamente. También implica generar reportes claros y comprensibles para la toma de decisiones, considerando la incertidumbre e inexactitud del modelo[24][30].

5.2.1 Validación de modelos:

La validación de modelos es un proceso crítico que verifica si los modelos están funcionando como se espera y se si encuentran alineados con sus objetivos de desarrollo[31]. La validación efectiva incluye:

1. **Solidez conceptual**: Evaluar la calidad del diseño del modelo y la evidencia empírica que lo respalda. Esto incluye documentación rigurosa de todas las elecciones de modelo y pruebas de robustez y estabilidad[24].

2. **Monitoreo continuo**: Revisar regularmente el rendimiento del modelo y hacer ajustes necesarios. Esto incluye verificación de procesos, comparación con puntos de referencia y revisiones periódicas para asegurar que los modelos sigan siendo relevantes y efectivos[30].

3. **Análisis de resultados**: Comparar las salidas del modelo con los resultados reales para evaluar precisión y fiabilidad. Esto implica pruebas estadísticas y juicio experto para asegurar que los modelos continúen desempeñándose conforme a los objetivos de diseño[24].

5.2.2 Gobierno, políticas y controles:

Las estructuras de gobernanza robustas son esenciales para un MRM efectivo. Los aspectos clave incluyen:

1. **Marco de gobierno**: Establecer roles y responsabilidades, políticas y procedimientos para gestionar el riesgo de modelos. Esto incluye supervisión de la junta y la alta dirección y asegurar el cumplimiento de los estándares regulatorios[24].

2. **Documentación**: Mantener registros detallados del desarrollo de modelos, actividades de validación y monitoreo continuo. La documentación adecuada asegura transparencia, facilita el cumplimiento y apoya la mejora continua[30].

3. **Controles internos**: Implementar controles para asegurar el cumplimiento de políticas y procedimientos, respaldados por revisiones y auditorías independientes. Esto incluye definir procesos claros para el desarrollo, validación y uso de modelos[24].

5.2.3 Regulación:

Las instituciones financieras deben adherirse a estándares específicos de validación y documentación para cumplir con las expectativas regulatorias. Esto incluye actualizaciones y revisiones regulares para asegurar que los modelos sigan siendo relevantes y efectivos. La Autoridad de Regulación Prudencial (PRA, por sus siglas en inglés) y otros organismos también han establecido altos estándares para marcos de riesgo de modelos, particularmente para modelos de pruebas de estrés[30].

Algunos organismos reguladores han establecido directrices para asegurar prácticas efectivas de MRM. Estas incluyen:

1. **Marco de Basilea III**: El Acuerdo de Basilea III introdujo un marco comprensivo para la administración de riesgo de modelo que requiere que los bancos implementen procesos robustos para validar la precisión y fiabilidad de sus modelos internos [22]. Este marco es parte de un esfuerzo más amplio para mejorar la estabilidad del sistema financiero al asegurar que los modelos internos de los bancos sean sólidos y efectivos en la evaluación de riesgos [23].

2. **Guía de supervisión de la Reserva Federal**: Documentos como la SR 11-7 de la Reserva Federal de los Estadus Unidos de Norte América proporcionan orientación comprensiva sobre el desarrollo, implementación, uso y validación de modelos. Esta guía destaca la importancia de una validación rigurosa de modelos y prácticas de gobernanza sólidas[24].

3. **Principios de la PRA para la administración de riesgos de modelos**: La Autoridad de Regulación Prudencial (PRA - Reino Unido) ha desarrollado principios específicos para la administración efectiva de MRM, aplicables a bancos, sociedades de construcción y firmas de inversión incorporadas en el Reino Unido. Estos principios enfatizan la implementación proporcional de marcos de MRM, la responsabilidad de la alta dirección y la necesidad de gestionar los riesgos asociados con la IA y el ML en modelos. Las directrices de la PRA aseguran que las prácticas de MRM sean robustas, transparentes y alineadas con estándares internacionales [23].

5.3 Desafíos de la gestión

5.3.1 Definición y clasificación de modelos:

Un desafío significativo en la MRM es la definición y clasificación consistente de modelos. Las instituciones financieras a menudo luchan por establecer una definición clara y estandarizada de lo que constituye un modelo y de su clasificación de riesgo[32], lo cual es esencial para crear un inventario de modelos comprensivo. Esto incluye diferenciar entre modelos y otras herramientas o procesos cuantitativos utilizados dentro de la organización. Una definición robusta de modelos asegura que todos los modelos relevantes sean capturados y gestionados adecuadamente[30].

5.3.2 Gobierno y responsabilidad:

Los marcos de gobierno efectivos son cruciales para la MRM. Asegurar roles y responsabilidades claros en toda la organización es un desafío, especialmente en instituciones grandes y complejas. Las estructuras de gobierno deben definir quién es responsable del desarrollo, validación, uso y supervisión de modelos. Establecer y mantener esta responsabilidad es esencial para mitigar el riesgo de modelo[30][33].

5.3.3 Validación de modelos y monitoreo de rendimiento:

La validación regular y exhaustiva de modelos es una piedra angular de la MRM. Sin embargo, realizar una validación efectiva de modelos que incluya pruebas retrospectivas, comparaciones con puntos de referencia y análisis de sensibilidad puede ser intensivo en recursos y complejo. El monitoreo de rendimiento es igualmente desafiante, requiriendo una evaluación continua para asegurar que los modelos funcionen como se espera y sigan siendo relevantes bajo condiciones cambiantes del mercado[30][31].

5.3.4 Administración del inventario de modelos:

Mantener un inventario de modelos comprensivo y actualizado es otro desafío crítico. Este inventario debe incluir información detallada sobre el propósito de cada modelo, sus suposiciones, limitaciones e historial de validación. La administración eficiente de este inventario es crucial para rastrear el estado y el rendimiento de todos los modelos dentro de la organización[33].

5.3.5 Calidad e integridad de los datos:

Los datos de alta calidad son cruciales para la precisión y fiabilidad de los modelos. Asegurar la integridad y consistencia de los datos de entrada puede ser difícil, particularmente cuando los modelos dependen de grandes conjuntos de datos de múltiples y diversas fuentes[27]. La mala calidad de los datos puede llevar a salidas incorrectas de modelos (Garbage in Garbage Out) y aumentar el riesgo de modelo. Por lo tanto, son esenciales prácticas rigurosas de administración de datos[30].

5.3.6 Cumplimiento regulatorio:

Las instituciones financieras deben navegar un complejo panorama de requisitos regulatorios relacionados con la MRM. Regulaciones como la SR 11-7 de la Reserva Federal y las directrices de la Autoridad Bancaria Europea (EBA por sus siglas en ingles) imponen estándares estrictos para el desarrollo, validación y documentación de modelos. Mantenerse al día con estos requisitos en evolución y asegurar el cumplimiento es un desafío continuo[30][33].

5.3.7 Integración de tecnologías avanzadas:

La integración de tecnologías avanzadas como la inteligencia artificial (IA) y el aprendizaje automático (ML) en modelos introduce nuevos desafíos. Estas tecnologías pueden mejorar la precisión y eficiencia de los modelos, pero también añaden complejidad a los procesos de validación y monitoreo[31]. Las instituciones financieras deben desarrollar habilidades y metodologías especializadas para gestionar los riesgos asociados con estos modelos avanzados[30].

5.3.8 Clasificación de riesgos y asignación de recursos:

EL MRM efectivo requiere priorizar recursos según el riesgo que representan diferentes modelos. Desarrollar e implementar un marco robusto de clasificación de riesgos de modelos que evalúe y categorice con precisión los modelos según sus niveles de riesgo es un desafío. Este marco debe ser transparente, consistente y lo suficientemente flexible como para adaptarse a nuevos modelos y condiciones cambiantes[33].

5.3.9 Agregación y reporte de riesgo de modelos:

Agregar el riesgo de modelos en toda la organización y presentarlo de manera significativa a la alta dirección es esencial para la toma de decisiones efectiva. Esto implica sintetizar datos complejos de múltiples modelos en información clara y procesable. Crear un marco de reporte efectivo que resalte los riesgos clave y apoye la planificación estratégica es un desafío significativo[30].

5.3.10 Experiencia y capacitación:

Asegurar que el personal tenga la experiencia y capacitación necesarias para desarrollar, validar y gestionar modelos es elemental. La evolución rápida de las técnicas de modelado y los requisitos regulatorios significa que el desarrollo profesional continuo también lo necesario. Construir y mantener un equipo con las habilidades y conocimientos adecuados es un desafío constante para las compañías[30].

5.3.11 Administración del ciclo de vida de los modelos:

El MRM también requiere una administración comprensiva del ciclo de vida de los modelos, desde el desarrollo hasta el desmonte. Esto incluye asegurar que cada fase del ciclo de vida del modelo, como la validación, implementación y monitoreo continuo, esté gestionada y documentada adecuadamente[34].

5.3.12 Transparencia y comunicación:

La transparencia y la comunicación clara dentro de la organización y con las partes interesadas externas son críticas para el éxito de la MRM. Esto incluye actualizaciones regulares sobre el estado, rendimiento y riesgo de los modelos a la alta dirección y los organismos reguladores. La comunicación clara ayuda a asegurar que todas las partes interesadas estén al tanto de los riesgos potenciales y las acciones que se están tomando para mitigarlos[34].

5.3.13 Abordar los riesgos de los modelos de IA y ML:

Los modelos de IA y ML introducen riesgos únicos[35], incluidos problemas relacionados con la ética de los datos, la equidad y la explicabilidad[36]. Las compañías deben desarrollar marcos robustos para abordar estos riesgos, incluyendo la implementación de estructuras de gobernanza que aseguren transparencia y responsabilidad en el uso de modelos de IA y ML[34].

5.3.14 Validación de modelos de proveedores:

La validación de modelos de proveedores (o de terceros) presenta desafíos únicos debido a la falta de transparencia y control sobre el proceso de desarrollo del modelo. Las instituciones deben desarrollar procesos rigurosos de validación para modelos externos, para asegurar que cumplan con los estándares internos y los requisitos regulatorios. Esto incluye el monitoreo y evaluación regular del rendimiento del modelo y la idoneidad para su uso previsto[37].

6 Relación entre MLOps y la administración de riesgo de modelos (MRM)

A continuación se plantea una relación de desafío - solución entre algunos de los retos mencionados de la administración de riesgo de modelo y los beneficios ya conocidos de las metodologías de MLOps, con el ánimo de buscar soluciones probadas en metodologías conocidas[35] y evitar así "re inventar la rueda".

6.1 Sobre el desafío de la definición y clasificación de modelos

Definir y clasificar modelos dentro de una organización es un desafío significativo debido a las complejidades y variabilidades en el desarrollo y uso de estos, por lo tanto abordar este reto de manera efectiva puede lograrse aprovechando varios componentes y principios clave de MLOps. La **Colaboración y comunicación** fomentan un enfoque multidisciplinario que involucra a científicos de datos, ingenieros de ML y equipos de operaciones, asegurando que los criterios para definir y clasificar modelos sean claros y se apliquen consistentemente en toda la organización. Las herramientas y plataformas compartidas que respaldan todo el ciclo de vida de ML facilitan esta colaboración, permitiendo el desarrollo de una definición unificada y criterios de clasificación con los que todos los interesados estén de acuerdo[7][9]. Las canalizaciones (Pipelines) de **Integración y entrega continua (CI/CD)** juegan un papel fundamental al automatizar la integración de nuevos cambios de código y datos en la rama principal de un proyecto[10]. Esta integración frecuente y pruebas automatizadas ayudan a detectar y abordar cualquier inconsistencia o desviación de los criterios de clasificación establecidos temprano en el proceso de desarrollo[8]. Además, la **Reproducibilidad y el Versionado** aseguran que todos los cambios en modelos, datos y código se rastreen y documenten meticulosamente [38], proporcionando una clara pista de auditoría que ayuda en el proceso de clasificación. Las herramientas de versionado mencionadas aseguran que cada iteración de un modelo esté registrada, permitiendo a los equipos mantener un inventario de modelos comprensivo y preciso[7].

6.2 Sobre el desafío de gobierno y responsabilidad

Un gobierno efectivo requiere roles claros, responsabilidades y supervisión en todas las etapas del ciclo de vida del modelo, lo que puede facilitarse a través de **Integración y entrega continua (CI/CD)**[39]. Las canalizaciones (Pipelines) de CI/CD aseguran que las actualizaciones de modelos[14], datos y políticas de gobierno se integren, prueben y desplieguen automáticamente de manera controlada [38]. Esta automatización reduce el riesgo de errores humanos y asegura que todos los cambios cumplan con los estándares de gobierno establecidos. Por ejemplo, los procesos de CI/CD pueden incluir revisiones de código automatizadas, análisis de seguridad [40] y validaciones de cumplimiento antes de que cualquier cambio se fusione en la rama de producción[8], adicionalmente estas herramientas permiten validaciones por roles independientes. Además, las herramientas de **Colaboración y comunicación** son fundamentales para fomentar una cultura de transparencia y responsabilidad. Al usar plataformas compartidas para la comunicación y administración de proyectos, todos los interesados, incluidos los científicos de datos, ingenieros de ML, auditores y oficiales de cumplimiento, pueden mantenerse informados sobre los desarrollos, cambios y problemas del modelo. Las actualizaciones regulares y la documentación clara facilitadas por estas herramientas aseguran que cada miembro del equipo entienda sus roles y responsabilidades dentro del marco de gobierno[7][9].

Además, las herramientas de **Automatización y orquestación** pueden simplificar los procesos de gobierno al automatizar tareas rutinarias, como el seguimiento del uso del modelo, el monitoreo del rendimiento y la generación de informes de cumplimiento. Estas herramientas pueden hacer cumplir las políticas de gobierno consistentemente en todos los modelos y asegurar que cualquier desviación se detecte y aborde rápidamente. La **Reproducibilidad y el Versionado** también juegan un papel importante en el mantenimiento del gobierno al asegurar que todos los cambios en modelos, datos y código estén controlados por versiones y sean auditables. Esta transparencia permite revisiones exhaustivas y responsabilidad en cada paso del ciclo de vida del modelo[39], permitiendo a las organizaciones mantener estándares rigurosos y abordar rápidamente cualquier problema de cumplimiento.

6.3 Sobre el desafío de la validación de modelos y el monitoreo del rendimiento

Automatización y orquestación son críticos para agilizar el proceso de validación y asegurar verificaciones consistentes y repetibles en todos los modelos. Las herramientas de orquestación de flujos de trabajo pueden automatizar las tareas complejas de entrenamiento, validación y despliegue de modelos. Estas herramientas aseguran que cada paso se ejecute de manera eficiente y fiable, minimizando los errores humanos y asegurando la consistencia en las evaluaciones del rendimiento del modelo[8]. **Integración y entrega continua (CI/CD)** mejora aún más el proceso de validación al permitir pruebas automáticas e integración de nuevos cambios de código y datos en la rama productiva del proyecto. Las canalizaciones (Pipelines) de CI/CD pueden incorporar pruebas de validación que se ejecutan automáticamente cada vez que se realizan cambios[14], asegurando que cualquier problema se detecte y aborde temprano en el ciclo de desarrollo. Este proceso de validación continua ayuda a mantener altos estándares de precisión y fiabilidad del modelo[7][41].

Además, **Entrenamiento Continuo (CT)** extiende los principios de CI/CD al incorporar el reentrenamiento de modelos a medida que nuevos datos están disponibles. Esto asegura que los modelos permanezcan actualizados y continúen desempeñándose bien a medida que los datos subyacentes evolucionan. Los procesos automatizados de reentrenamiento y despliegue ayudan a mantener la relevancia y precisión del modelo con el tiempo, abordando uno de los desafíos centrales del monitoreo del rendimiento [8]. **Monitoreo y bucles de retroalimentación** son esenciales para la evaluación continua del desempeño de los modelos en producción[41]. Las herramientas de monitoreo rastrean métricas importantes y configuran alertas para cualquier anomalía o disminución en el desempeño del modelo. Estos bucles de retroalimentación automatizados utilizan estos datos de monitoreo para desencadenar reentrenamientos u otras acciones correctivas, asegurando que los modelos permanezcan precisos y efectivos en aplicaciones en tiempo real.

Algunos elementos de monitoreo que son de interés para MRM y a la vez son facilitados por las herramientas pueden ser: cambios en las distribuciones de datos de entrada y de salida de los modelos [38], consumo de recursos (procesamiento, latencia, memoria RAM, almacenamiento), entre otros. Este mecanismo de retroalimentación continua permite entonces generar respuestas rápidas a los problemas de rendimiento, manteniendo la robustez y fiabilidad de los modelos en entornos de producción[7][41].

De manera detallada, entre los elementos relevantes a monitorear se encuentra la precisión del modelo, que debe evaluarse continuamente para detectar posibles degradaciones en su capacidad predictiva. Esto implica comparar las predicciones del modelo con los resultados reales y ajustar el modelo si se observa una disminución en su precisión. Además, es fundamental supervisar la drift en los datos, es decir, los cambios en la distribución de los datos de entrada que pueden afectar la capacidad del modelo para generalizar adecuadamente. El sesgo y la equidad del modelo también deben ser vigilados para garantizar que el modelo no perpetúe o amplifique discriminaciones presentes en los datos. Otro aspecto relevante en el rendimiento computacional, incluye el tiempo de respuesta y el uso de recursos, para asegurar que el modelo se ejecute de manera eficiente y no genere cuellos de botella en el sistema. Finalmente, la seguridad del modelo debe ser un foco constante, protegiendo el sistema contra ataques adversariales y asegurando la integridad de los datos y el modelo.

Reproducibilidad y versionado también son necesarios para una validación y monitoreo efectivos de modelos. Asegurar que todos los cambios en modelos, datos y código estén controlados por versiones permite un seguimiento preciso y una reversión si es necesario. Las herramientas como Git[16] para la administración de código y DVC[17] para el versionado de datos y MLFlow[18] o Weights and Biases[19] para el versionado de artefactos de modelos proporcionan una clara pista de auditoría, facilitando revisiones exhaustivas y la reproducibilidad de los resultados. Esta transparencia asegura que cualquier cambio pueda ser rastreado y comprendido, ayudando en el proceso de validación y monitoreo continuo del rendimiento[7][41].

6.4 Sobre el desafío de la administración del inventario de modelos

Uno de los aspectos fundamentales es la **Reproducibilidad y el Versionado**, que aseguran que todos los cambios en modelos, datos y código se rastreen y documenten de manera detallada [38]. Las herramientas de versionado mencionadas para artefactos de modelos proporcionan una pista de auditoría comprensiva, permitiendo a los equipos mantener un inventario preciso y actualizado de todos los modelos. Este versionado meticuloso facilita el seguimiento de la línea de modelos y la capacidad de revertir a versiones anteriores si es necesario, asegurando que el inventario de modelos esté siempre actual y sea fiable[7]. Además, la **Automatización y orquestación** juegan un papel necesario en la administración de inventarios de modelos

al automatizar los procesos de documentación y catalogación. Las herramientas de orquestación de flujos de trabajo también pueden configurarse para actualizar automáticamente el inventario de modelos con nuevos modelos y cambios en los modelos existentes, asegurando que el inventario refleje el estado más reciente de los modelos de la organización sin requerir actualizaciones manuales[8].

Las canalizaciones (Pipelines) de **Integración y entrega continua (CI/CD)** mejoran aún más la administración del inventario de modelos al integrar actualizaciones continuas y verificaciones en el ciclo de vida del modelo[41]. Cada vez que se desarrolla, prueba o modifica un modelo, las canalizaciones de CI/CD pueden automatizar el proceso de actualización del inventario de modelos, asegurando que todos los cambios se capturen en tiempo real. Esta automatización reduce el riesgo de errores humanos y asegura que el inventario de modelos permanezca comprensivo y actualizado [8]. Las herramientas de **Colaboración y comunicación** también contribuyen significativamente a una administración efectiva del inventario de modelos. Al proporcionar una plataforma compartida para documentación y comunicación, estas herramientas permiten una colaboración fluida entre científicos de datos, ingenieros de ML y equipos de operaciones. Este enfoque multidisciplinario asegura que todos los interesados estén al tanto del estado y los cambios en el inventario de modelos, promoviendo la transparencia y la responsabilidad[7][9].

Además, los **Monitoreo y bucles de retroalimentación** aseguran una supervisión continua del inventario de modelos, con alertas y reportes automatizados que destacan discrepancias o modelos desactualizados. Este monitoreo continuo ayuda a identificar y abordar problemas rápidamente, asegurando que el inventario permanezca preciso y fiable.

6.5 Sobre el desafío de la calidad e integridad de los datos

Uno de los aspectos más críticos es la **Automatización y orquestación**, que pueden mejorar significativamente la administración de la calidad de los datos [38]. Las herramientas de orquestación de flujos de trabajo pueden automatizar los procesos de recopilación, preprocesamiento y validación de datos [38]. Estas herramientas aseguran que las canalizaciones (Pipelines) de datos se ejecuten consistentemente y que las verificaciones de integridad de los datos se apliquen automáticamente en cada etapa del ciclo de vida de los datos. Los procesos automatizados reducen el riesgo de errores humanos y aseguran que los datos se limpien, transformen y validen consistentemente antes de ser utilizados para el entrenamiento y evaluación de modelos[8]. Las canalizaciones (Pipelines) de **Integración y entrega continua (CI/CD)** apoyan aún más la calidad de los datos al integrar pruebas automatizadas y pasos de validación en la canalización de datos. Estas canalizaciones pueden incluir pruebas unitarias, verificaciones de consistencia de datos[41] y mecanismos de detección de anomalías que se ejecutan automáticamente cada vez que se ingieren nuevos datos o se actualizan datos existentes. Esta validación continua ayuda a detectar y abordar problemas de calidad de datos temprano, asegurando que solo datos de alta calidad se utilicen en el desarrollo y despliegue de modelos[7] [38].

Además, la **Reproducibilidad y el versionado** juegan un papel vital en el mantenimiento de la calidad y la integridad de los datos. Al versionar datos con herramientas como DVC[17] (Control de Versiones de Datos), las organizaciones pueden rastrear cambios en los conjuntos de datos a lo largo del tiempo, mantener un historial claro de las modificaciones de datos y revertir fácilmente a versiones anteriores si es necesario. Esto asegura que la procedencia de los datos esté bien documentada y que cualquier problema de calidad de los datos pueda rastrearse hasta su origen y corregirse rápidamente[7]. Las herramientas de **Colaboración y comunicación** también contribuyen significativamente a la administración de la calidad de los datos al facilitar la comunicación y documentación transparentes entre los científicos de datos, ingenieros de ML y administradores de datos. Las plataformas compartidas para la colaboración aseguran que los estándares y prácticas de calidad de los datos se apliquen consistentemente y que cualquier problema de datos se informe y aborde rápidamente por los equipos relevantes[9].

Además, el **Monitoreo y los bucles de retroalimentación** son esenciales para la supervisión continua de la calidad de los datos[41]. La implementación de herramientas de monitoreo para rastrear métricas clave de calidad de los datos, como la integridad, precisión y puntualidad, permite a los equipos configurar alertas para cualquier desviación de los estándares esperados de calidad de los datos [27][38]. Los bucles de retroalimentación automatizados pueden utilizar estos datos de monitoreo para desencadenar acciones correctivas, como volver a ejecutar pipelines de datos o reentrenar modelos con conjuntos de datos corregidos[14], asegurando que los problemas de calidad de los datos se resuelvan rápidamente y no afecten el rendimiento del

modelo[7].

6.6 Sobre el desafío del cumplimiento regulatorio

El primer elemento a mencionar es la **Integración y entrega continua (CI/CD)**, que asegura que los requisitos regulatorios se integren y prueben consistentemente a lo largo del ciclo de vida del desarrollo del modelo[39]. Los pipelines de CI/CD automatizan el proceso de verificación del cumplimiento con los estándares regulatorios cada vez que se realizan cambios de código o datos. Esto incluye la incorporación de verificaciones de cumplimiento automatizadas, análisis de seguridad [40] y pruebas de validación que se alinean con las directrices regulatorias, asegurando que cualquier problema de no cumplimiento se identifique y aborde temprano en el proceso de desarrollo[8]. La **Reproducibilidad y el Versionado** también son necesarias para el cumplimiento regulatorio, ya que proporcionan una pista de auditoría transparente y verificable de todos los cambios en modelos, datos y código. Estas herramientas facilitan la reversión fácil a versiones anteriores si surgen discrepancias o problemas de cumplimiento, manteniendo así un historial claro que cumple con el escrutinio regulatorio[7].

Las **Automatización y orquestación** apoyan aún más el cumplimiento al automatizar los procesos de documentación e informes requeridos para la adherencia regulatoria. Las herramientas de orquestación de flujos de trabajo pueden configurarse para generar y actualizar automáticamente la documentación relacionada con el desarrollo, validación y despliegue de modelos, asegurando que toda la documentación regulatoria se mantenga consistentemente y esté actualizada. Esta automatización reduce la carga sobre los equipos de rastrear y documentar manualmente el cumplimiento, minimizando el riesgo de errores humanos y asegurando que los requisitos regulatorios se cumplan de manera sistemática [8]. Las herramientas de **Colaboración y comunicación** son esenciales para facilitar la comunicación transparente entre diferentes equipos, incluidos los científicos de datos, ingenieros de ML, oficiales de cumplimiento y gestores de riesgos. Estas herramientas aseguran que todos los involucrados estén al tanto de los últimos requisitos regulatorios y cambios, fomentando un enfoque colaborativo para mantener el cumplimiento en toda la organización[9].

Además, el **Monitoreo y bucles de retroalimentación** juegan un papel importante en asegurar el cumplimiento regulatorio continuo. Al implementar herramientas de monitoreo robustas para rastrear métricas y indicadores de rendimiento relacionados con el cumplimiento, las organizaciones pueden configurar alertas para notificar a los equipos relevantes sobre cualquier desviación de los estándares regulatorios. Estos bucles de retroalimentación permiten la administración proactiva de problemas de cumplimiento, permitiendo a los equipos abordar y rectificar rápidamente cualquier problema de no cumplimiento detectado durante la operación del modelo. Los mecanismos de retroalimentación automatizados también pueden desencadenar el reentrenamiento o el redepliegue de modelos para asegurar que sigan cumpliendo con los requisitos regulatorios en evolución[7].

6.7 Sobre el desafío de la clasificación de riesgos y asignación de recursos

Enfrentar este reto implica priorizar efectivamente los modelos según sus niveles de riesgo y asignar recursos en consecuencia, lo cual puede mejorarse significativamente mediante la aplicación de componentes y principios clave de MLOps. Las canalizaciones de **Integración y entrega continua (CI/CD)** son instrumentales en este contexto, ya que facilitan la evaluación continua y la integración de actualizaciones de modelos. Al incorporar pruebas automáticas de evaluación de riesgos dentro de las canalizaciones de CI/CD, las organizaciones pueden evaluar continuamente los niveles de riesgo de diferentes modelos basándose en sus métricas de rendimiento y cumplimiento. Esta evaluación continua asegura que los modelos se clasifiquen correctamente según sus perfiles de riesgo, permitiendo una asignación de recursos más informada[8]. Las herramientas de **Automatización y orquestación** agilizan aún más este proceso al automatizar los flujos de trabajo asociados con la evaluación de riesgos y la asignación de recursos. Las herramientas como Apache Airflow[11] y Kubeflow Pipelines[12] pueden gestionar flujos de trabajo complejos, asegurando que los modelos de alto riesgo reciban validaciones y monitoreo más frecuentes, mientras que los modelos de menor riesgo se manejan con menos escrutinio intensivo, optimizando el uso de recursos[7].

Los **Monitoreos y bucles de retroalimentación** son esenciales para rastrear continuamente el rendimiento y los niveles de riesgo de los modelos en producción. Al configurar sistemas de monitoreo automatizados, las organizaciones pueden rastrear indicadores clave de rendimiento y métricas de riesgo en tiempo real. Estos

sistemas de monitoreo pueden generar alertas y reportes, proporcionando información sobre cuáles modelos están teniendo un rendimiento deficiente o mostrando signos de aumento de riesgo (según la métrica de riesgo que se defina). Los bucles de retroalimentación automatizados pueden desencadenar validaciones adicionales o procesos de reentrenamiento para modelos de alto riesgo, asegurando que reciban la atención y los recursos necesarios para mantener su fiabilidad y rendimiento[7]. La **Reproducibilidad y el Versionado** aseguran que todos los cambios en modelos, datos y código se documenten meticulosamente, proporcionando una pista de auditoría clara que respalda las decisiones de evaluación de riesgos y asignación de recursos. Las herramientas de versionado permiten a las organizaciones rastrear la evolución de los modelos y conjuntos de datos, asegurando que las evaluaciones de riesgos se basen en datos históricos comprensivos y precisos[7].

Además, las herramientas de **Colaboración y comunicación** facilitan la coordinación y alineación de los diferentes equipos involucrados en el proceso de MRM. Al proporcionar una plataforma compartida para la documentación y la comunicación, estas herramientas aseguran que los científicos de datos, ingenieros de ML, gestores de riesgos y oficiales de cumplimiento estén todos en la misma página respecto a los niveles de riesgo de los modelos y las estrategias de asignación de recursos. Este enfoque colaborativo asegura que las decisiones se tomen con la aportación de todos los interesados relevantes, llevando a una distribución de recursos más equilibrada y efectiva[9]. Finalmente, aprovechar el **Entrenamiento Continuo (CT)** extiende los beneficios de CI/CD al asegurar que los modelos se actualicen continuamente con nuevos datos, manteniendo su relevancia y rendimiento con el tiempo. Los procesos de entrenamiento continuo pueden priorizarse según los niveles de riesgo de los modelos, asegurando que los modelos de alto riesgo se reentrenen con mayor frecuencia para abordar cualquier riesgo emergente [8].

6.8 Sobre el desafío de asegurar suficiente experiencia y capacitación

Las herramientas de **Colaboración y comunicación** son importantes para fomentar un ambiente de aprendizaje continuo y asegurar que el conocimiento se comparta entre todos los equipos involucrados en el proceso de MRM. Las plataformas como Slack[42], Microsoft Teams[43], Confluence[44], Notion[45] o plataformas dedicadas a MLOps facilitan la comunicación y colaboración regular entre científicos de datos, ingenieros de ML, gestores de riesgos y oficiales de cumplimiento. Esta interacción asegura que los miembros del equipo estén consistentemente actualizados sobre las mejores prácticas, nuevas tecnologías y cambios regulatorios, promoviendo una cultura de mejora continua y experiencia compartida[9]. Las canalizaciones de **Integración y entrega continua (CI/CD)** mejoran aún más esto al integrar módulos de capacitación automatizados y actualizaciones de documentación en el flujo de trabajo de desarrollo. Las canalizaciones de CI/CD pueden incluir pasos para desplegar recursos de capacitación, ejecutar pruebas educativas o de desarrollo y asegurar que todos los miembros del equipo estén familiarizados con las últimas herramientas y técnicas que se están utilizando. Esta integración ayuda a mantener un alto nivel de experiencia en el equipo al asegurar que todos tengan acceso a materiales de capacitación actualizados y oportunidades de aprendizaje práctico[8].

La **Reproducibilidad y el Versionado** juegan un papel relevante en la capacitación al asegurar que todos los cambios en modelos, datos y código estén documentados y puedan ser fácilmente entendidos y replicados por nuevos miembros del equipo. Las herramientas de versionado de datos y código ayudan a mantener un historial claro de todas las modificaciones, proporcionando a los nuevos miembros del equipo documentación comprensiva y ejemplos de los que aprender. Esta transparencia es crucial para fines de capacitación, ya que permite a los nuevos empleados ponerse al día rápidamente y comprender el contexto y la razón detrás de los modelos y procesos existentes[7]. Las herramientas de **Automatización y orquestación** también apoyan los esfuerzos de capacitación al automatizar tareas rutinarias y flujos de trabajo, liberando tiempo para que los miembros del equipo se concentren en el desarrollo de habilidades y la capacitación avanzada. Las herramientas de orquestación de flujos de trabajo pueden automatizar tareas repetitivas, permitiendo a los miembros experimentados del equipo entrenar a los nuevos empleados y concentrarse en actividades estratégicas de nivel superior[7].

Los **Monitoreo y bucles de retroalimentación** contribuyen a asegurar una experiencia adecuada al proporcionar información continua sobre el rendimiento del equipo e identificar áreas donde se puede necesitar capacitación adicional. Al rastrear métricas de rendimiento y configurar bucles de retroalimentación, las organizaciones pueden identificar brechas de habilidades y abordarlas proactivamente a través de programas de capacitación específicos. Este mecanismo de retroalimentación continua asegura que la experiencia del

equipo se mantenga alineada con las demandas cambiantes de la administración de riesgos de modelos y los requisitos regulatorios[7].

6.9 Sobre el desafío de la administración del ciclo de vida de los modelos

Las canalizaciones de **Integración y entrega continua (CI/CD)** son fundamentales en este proceso, ya que automatizan la integración de nuevos cambios de código, datos y actualizaciones, asegurando que los modelos se prueben y validen consistentemente en cada etapa de su ciclo de vida[41].

Las canalizaciones de CI/CD facilitan la transición fluida de los modelos desde el desarrollo hasta la producción, asegurando que cualquier problema se identifique y aborde temprano, manteniendo así la integridad y el rendimiento del modelo a lo largo de su ciclo de vida[8]. Las herramientas de **Automatización y orquestación**, mejoran aún más la administración del ciclo de vida al automatizar tareas rutinarias asociadas con el entrenamiento, validación, despliegue y monitoreo de modelos. Estas herramientas aseguran que cada fase del ciclo de vida del modelo se ejecute de manera eficiente y fiable, reduciendo la probabilidad de errores y liberando recursos para tareas más estratégicas[7].

La **Reproducibilidad y el Versionado** son criticas para mantener un historial claro y preciso de todos los cambios realizados en los modelos, datos y código a lo largo de su ciclo de vida. Las herramientas como Git para el control de versiones y DVC (Control de Versiones de Datos) permiten a las organizaciones rastrear y documentar cada modificación, asegurando que todos los cambios sean transparentes y puedan ser auditados si es necesario. Este nivel de documentación es esencial para el cumplimiento regulatorio y para mantener la reproducibilidad de los resultados a través de todo el ciclo de vida, lo cual es vital para depurar y perfeccionar modelos con el tiempo[7]. El **Entrenamiento Continuo (CT)** extiende los beneficios de CI/CD al asegurar que los modelos se actualicen continuamente con nuevos datos, manteniendo su relevancia y precisión a medida que los datos subyacentes evolucionan. Los procesos automatizados de reentrenamiento y redespliegue aseguran que los modelos siempre desempeñen su mejor versión, adaptándose a nuevas tendencias y patrones en los datos [8].

Los **Monitoreo y bucles de retroalimentación** proporcionan supervisión continua de los modelos en producción, rastreando métricas clave de rendimiento y alertando a los equipos sobre cualquier problema que pueda surgir. Estas herramientas de monitoreo son esenciales para detectar la degradación del rendimiento, el desvío de conceptos y otras anomalías que puedan afectar la precisión y fiabilidad del modelo. Los bucles de retroalimentación automatizados utilizan estos datos de monitoreo para desencadenar el reentrenamiento u otras acciones correctivas, asegurando que los modelos permanezcan efectivos y alineados con su propósito previsto a lo largo de su ciclo de vida[7]. Además, las herramientas de **Colaboración y comunicación** facilitan la interacción fluida entre los diversos equipos involucrados en el ciclo de vida del modelo, incluidos los científicos de datos, ingenieros de ML, gestores de riesgos y oficiales de cumplimiento. Las plataformas compartidas para la documentación y la comunicación aseguran que todos los interesados estén informados sobre el estado, cambios y rendimiento del modelo, promoviendo la transparencia y la responsabilidad desde la concepción hasta la puesta en producción y eventual desmonte del modelo[9].

6.10 Sobre el desafío de la transparencia y comunicación

Aprovechar las herramientas de **Colaboración y comunicación** es fundamental para este esfuerzo. Las plataformas de comunicación o documentación mencionadas y plataformas especializadas en MLOps facilitan la comunicación y colaboración fluida entre científicos de datos, ingenieros de ML, gestores de riesgos y oficiales de cumplimiento. Estas herramientas permiten el intercambio en tiempo real de actualizaciones, conocimientos y decisiones, asegurando que todos los involucrados en el proceso de desarrollo y despliegue del modelo tengan acceso a la información más reciente. Esta comunicación en tiempo real reduce los malentendidos y asegura que todos los miembros del equipo estén al tanto de sus roles y responsabilidades, promoviendo la responsabilidad y la cohesión[9].

Las canalizaciones de **Integración y entrega continua (CI/CD)** mejoran la transparencia al automatizar la integración de nuevos cambios de código y datos, asegurando que todas las modificaciones se documenten y prueben sistemáticamente antes de ser desplegadas. Esta automatización proporciona una pista de auditoría clara de los cambios, lo cual es esencial para comprender la evolución de los modelos y para realizar revisiones y auditorías exhaustivas. Las canalizaciones de CI/CD también aseguran que cualquier

problema se identifique y comunique rápidamente, permitiendo a los equipos responder de manera oportuna y mantener altos estándares de calidad del modelo y cumplimiento[8]. La **Reproducibilidad y el Versionado** apoyan aún más la transparencia al mantener un historial comprensivo de todos los cambios en modelos, datos y código. Las herramientas de versionado permiten a los equipos rastrear cada modificación, asegurando que todos los cambios se documenten y puedan ser fácilmente revisados o revertidos si es necesario. Este nivel de documentación es crucial para la transparencia, ya que permite a todos los interesados comprender el historial de desarrollo y modificación de los modelos, facilitando una mejor toma de decisiones y cumplimiento[7].

Las herramientas de **Automatización y orquestación** juegan un papel vital en asegurar que la comunicación y la documentación sean consistentes y fiables. Estas herramientas automatizan tareas repetitivas y flujos de trabajo, asegurando que todos los procesos se ejecuten de manera estandarizada y que la documentación relevante se actualice automáticamente. Esta automatización reduce la carga sobre los miembros del equipo de documentar manualmente cada cambio, asegurando que todas las actualizaciones se registren con precisión y estén accesibles para todos los interesados[7]. Los **Monitoreo y bucles de retroalimentación** también contribuyen a la transparencia al proporcionar información continua sobre el rendimiento del modelo y alertando a los equipos sobre cualquier problema que surja. Estas herramientas de monitoreo rastrean indicadores de rendimiento y generan reportes automatizados, asegurando que todos los interesados estén informados sobre el estado y rendimiento actuales de los modelos. Los bucles de retroalimentación automatizados facilitan respuestas rápidas a cualquier problema detectado, manteniendo la integridad y fiabilidad de los modelos en producción[7].

6.11 Sobre el desafío de los riesgos de los modelos de IA y ML

Las canalizaciones de **Integración y entrega continua (CI/CD)** son fundamentales en este sentido, ya que permiten la integración fluida de nuevos cambios de código, actualizaciones y datos en los modelos de IA y ML. Al incorporar pasos de prueba y validación automatizados dentro de las canalizaciones de CI/CD, las organizaciones pueden asegurar que los modelos se evalúen continuamente en términos de rendimiento, sesgo y cumplimiento con los estándares regulatorios. Este proceso de validación continua ayuda a identificar y rectificar problemas temprano, manteniendo así la integridad y fiabilidad de los modelos[8]. La **Automatización y orquestación** mejoran aún más la administración de riesgos al automatizar los procesos de flujo de trabajo asociados con el entrenamiento, despliegue y monitoreo de modelos. La automatización asegura que todos los modelos se sometan a pruebas rigurosas y consistentes antes del despliegue, mitigando así los posibles riesgos[7].

El **Entrenamiento Continuo (CT)** es otro componente crítico que extiende los principios de CI/CD al incorporar el reentrenamiento de modelos a medida que nuevos datos están disponibles. El entrenamiento continuo asegura que los modelos permanezcan actualizados con las últimas tendencias y patrones de datos, manteniendo así su precisión y relevancia con el tiempo. Esto es particularmente importante para los modelos de IA y ML que pueden sufrir desvío de concepto, donde las propiedades estadísticas de la variable objetivo cambian con el tiempo. Los procesos automatizados de reentrenamiento y redespliegue ayudan a adaptar los modelos a estos cambios, mitigando así los riesgos asociados con modelos desactualizados o irrelevantes [8]. La **Reproducibilidad y el Versionado** son esenciales para gestionar los riesgos de los modelos de IA y ML, ya que aseguran que todos los cambios en modelos, datos y código se documenten y puedan ser reproducidos.

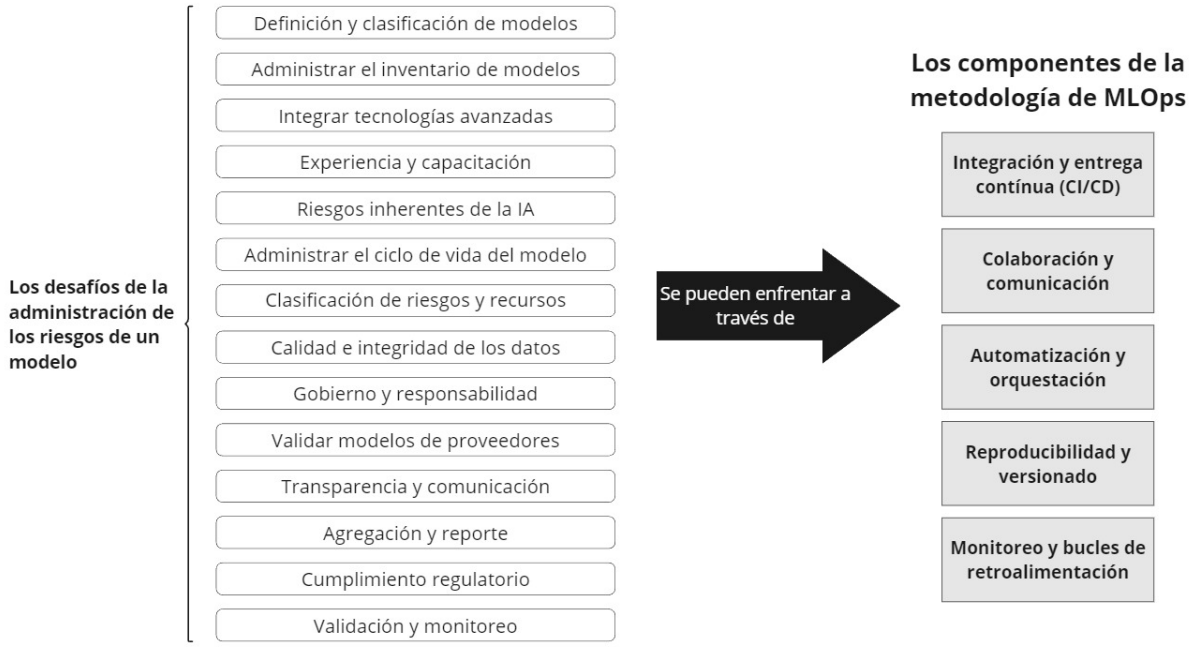
Las herramientas de **Colaboración y comunicación** juegan un papel vital en la administración de los riesgos de los modelos de IA y ML al facilitar la comunicación transparente y eficiente entre todos los interesados. Las plataformas compartidas para la documentación y la comunicación aseguran que los científicos de datos, ingenieros de ML, gestores de riesgos y oficiales de cumplimiento estén alineados e informados sobre los últimos desarrollos y posibles riesgos asociados con los modelos. Este enfoque colaborativo asegura que los riesgos se identifiquen y aborden proactivamente, fomentando una cultura de mejora continua y responsabilidad[9].

Sobre el desafío de la validación de modelos de proveedores

Resolver este desafío implica asegurar que los modelos de terceros cumplan con los estándares internos de precisión, fiabilidad y cumplimiento. Las canalizaciones de **Integración y entrega continua (CI/CD)** son fundamentales para integrar los modelos de proveedores en el sistema existente. Al incorporar pasos automatizados de prueba y validación en las canalizaciones de CI/CD, las organizaciones pueden asegurar que los modelos de proveedores se evalúen exhaustivamente en términos de rendimiento y cumplimiento antes de su despliegue. Esto incluye la ejecución de conjuntos de pruebas comprensivas que cubren varios escenarios para evaluar la robustez de los modelos[8]. Las herramientas de **Automatización y orquestación**, mejoran aún más el proceso de validación al automatizar la ejecución de estas pruebas y asegurar flujos de trabajo de validación consistentes y repetibles. Esta automatización reduce el riesgo de errores humanos, garantizan evaluaciones independientes y asegura que los modelos de proveedores se sometan a los mismos estándares rigurosos de validación que los modelos desarrollados internamente[7].

La **Reproducibilidad y el Versionado** son esenciales para mantener la transparencia y responsabilidad en el proceso de validación. Esto es particularmente importante para los modelos de proveedores, ya que permite a las organizaciones verificar la procedencia de los modelos y reproducir los resultados de validación de manera independiente. Este nivel de documentación y reproducibilidad es critico para el cumplimiento regulatorio y para construir confianza en los modelos de proveedores[7]. El **Monitoreo y bucles de retroalimentación** también son importantes para la validación continua de los modelos de proveedores en producción. Estas herramientas aseguran que los modelos de proveedores continúen cumpliendo con los estándares de rendimiento a lo largo del tiempo, y los bucles de retroalimentación automatizados pueden desencadenar el reentrenamiento, reclamos al tercero o ajustes según sea necesario para mantener la precisión y fiabilidad del modelo[7].

Las herramientas de **Colaboración y comunicación** proporcionan plataformas compartidas para la documentación y la comunicación aseguran que los científicos de datos, ingenieros de ML, gestores de riesgos y oficiales de cumplimiento puedan revisar y evaluar colaborativamente los modelos de proveedores. Este enfoque colaborativo asegura que todos los interesados relevantes estén involucrados en el proceso de validación y que cualquier problema se aborde rápidamente[9].



7 Caso aplicado para ilustrar algunas soluciones

La motivación de esta sección es proporcionar ejemplos prácticos de cómo la metodología de MLOps ofrece soluciones ya existentes en el desarrollo de software a los retos asociados con la gestión de riesgos de modelos (MRM). Para ilustrar esto, se ha recreado el flujo de trabajo clásico de entrenamiento y despliegue de un modelo, destacando cómo los desafíos típicos de MRM son abordados. Este flujo de trabajo se encuentra disponible en un repositorio [46] de GitHub de acceso público, donde se puede consultar el código en detalle y reproducir el ejemplo si se desea. Cabe destacar que aspectos específicos del modelado, como la selección de variables y la selección de hiperparámetros, no se abordan en el alcance de este ejercicio. Como se ha mencionado, el objetivo principal es enfocar la atención en los riesgos operativos del modelo y cómo la metodología de MLOps puede resolverlos.

7.1 Interpretación como caso de negocio

De manera hipotética, se propone una institución financiera que ofrece prestamos para vehículos, la cual utiliza análisis predictivo para identificar posibles incumplimientos de pago en financiamiento de vehículos. Analizando un conjunto de datos completo que incluye características de los prestatarios y de los préstamos, la institución busca predecir qué clientes probablemente no podrán pagar su primera cuota mensual (EMI). Entre los datos clave se encuentran el ratio de préstamo a valor (LTV), el tipo de empleo, la puntuación crediticia y el tiempo desde la entrega del préstamo. Por ejemplo, el modelo podría señalar a un nuevo solicitante, Juan Pérez, como un riesgo moderado debido a un alto ratio LTV y una puntuación crediticia inferior a la media, lo que lleva a la institución a considerar acciones de mitigación como ajustar los términos del préstamo o exigir garantías adicionales. Este enfoque proactivo permite a esta institución financiera gestionar el riesgo de manera más efectiva, asegurando mejores resultados financieros tanto para la empresa como para sus clientes.

7.2 Metodología

La metodología para la ejecución de este caso, y de cualquier proyecto de ML abarca una serie de pasos sistemáticos que garantizan la robustez y eficacia del modelo desarrollado [29]. El proceso comienza con la definición del problema, que bien se menciona en al principio de este caso aplicado, donde también se establece los objetivos específicos del caso y determina las métricas de éxito. Posteriormente, se realiza la recolección y preparación de datos que para esta ocasión proviene de un dataset curado, donde se obtienen los datos relevantes y se llevan a cabo tareas de limpieza y transformación para asegurar su calidad y adecuación para el modelo. La siguiente etapa es la exploración y análisis de datos, que incluye el análisis exploratorio para identificar patrones, anomalías y relaciones entre variables (aspecto documentado en detalle en el repositorio de código de este trabajo). Con los datos preparados, se procede a la selección y entrenamiento del modelo, en la cual se eligen los algoritmos más adecuados y se ajustan sus parámetros para optimizar el rendimiento. Tras el entrenamiento, es esencial realizar una evaluación rigurosa del modelo mediante técnicas de validación y pruebas para verificar su capacidad predictiva y evitar el sobreajuste. Una vez que el modelo ha sido validado, se pasa a la implementación y despliegue, donde el modelo se integra en el entorno de producción y se pone a disposición de los usuarios finales, que bien para este trabajo no aplica, pues el fin ilustrativo de este caso no tiene como alcance un ambiente productivo. Finalmente, la monitorización y mantenimiento continuos son cruciales para garantizar que el modelo se adapte a los cambios en los datos y siga cumpliendo con los objetivos establecidos.

7.3 Arquitectura del caso de aplicación

En el Flujo de Información Conceptual (Figura 2) se presentan las interacciones típicas que se abordarán en el ejercicio, incluyendo artefactos, procesos, etapas (ambientes) y roles. Los roles son responsables de ejecutar cada etapa. En la etapa de desarrollo, por ejemplo, se ilustran las interacciones clásicas entre los procesos y los resultados (artefactos en sus múltiples versiones), con el desafío particular de garantizar la trazabilidad del proceso y la comprensión de la participación de los roles. En la etapa de producción, se destaca que el modelo se somete a pruebas independientes y al monitoreo de variables de interés, tareas que corresponden a otros roles. El proceso de retroalimentación representa la información que llega al equipo de

desarrollo a partir de las revisiones y los hallazgos metodológicos o tecnológicos. Aunque existen limitaciones (recursos) para representar todas las interacciones entre los roles y los ambientes, se ha incluido este flujo para resaltar la eventual complejidad en la comunicación.

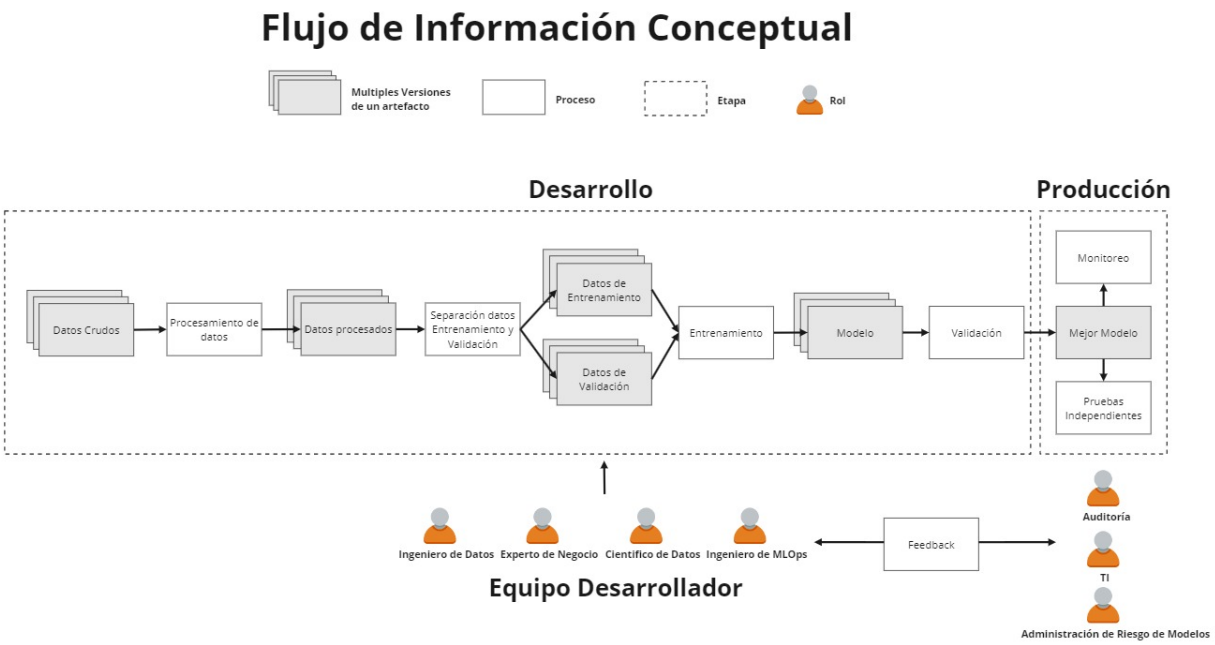


Figure 2: Flujo de Datos Conceptual

De manera particular se describe como se pasa desde la ingesta de datos crudos hasta la puesta en producción del modelo, contando con el preprocesamiento de la información, la separación en conjuntos de datos de entrenamiento y validación, búsqueda de parámetros (experimentación), validación y posterior selección del mejor modelo. Aunque dar una descripción metodológica exhaustiva para cada una de estas etapas no es el alcance de esté documento, se abordará cada una de ellas de manera circunstancial en la sección de retos solucionados en el contexto de esta aplicación.

7.3.1 Descripción de los componentes de la arquitectura

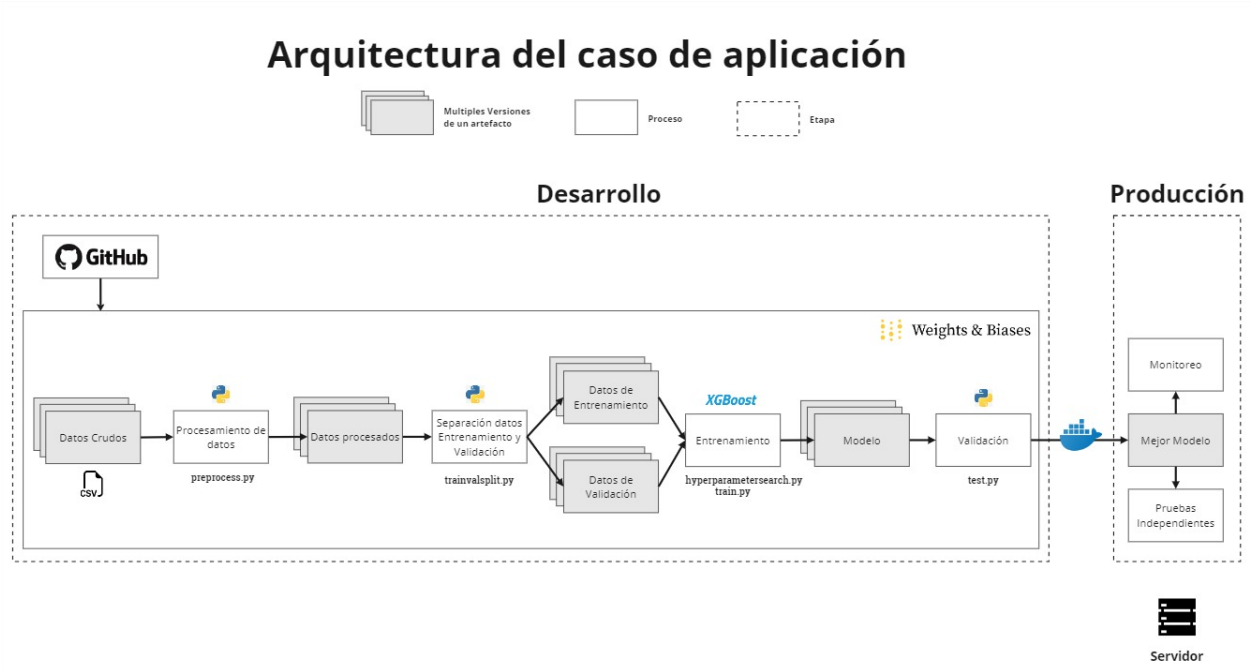


Figure 3: Flujo de datos sobre la arquitectura

1. **Conjunto de datos**: Es un conjunto estructurado de información que incluye características relevantes de los solicitantes de crédito y su comportamiento. En este caso, se ha utilizado un set de datos[47] para predecir el incumplimiento de pagos de préstamos para vehículos, específicamente para identificar a los clientes que no podrán pagar la primera cuota mensual (EMI).

   Las instituciones financieras enfrentan pérdidas debido al incumplimiento de pagos de préstamos para vehículos, lo que ha llevado a adaptaciones constantes en las políticas de aprobación de estos préstamos y un aumento en las tasas de rechazo. Estas instituciones buscan mejores modelos de evaluación del riesgo crediticio para predecir de manera precisa la probabilidad de incumplimiento de esa primera cuota.

   De manera general, el set de datos incluye la siguiente información:

   - **Información del prestatario:** Datos demográficos como identidad, estado de empleo, etc.
   - **Información del préstamo:** Detalles del desembolso, relación préstamo-valor, etc.

2. **Python** [48]: Es un lenguaje de programación ampliamente utilizado en aprendizaje de máquina por su simplicidad y extensa biblioteca de herramientas. Facilita el desarrollo, la implementación y el análisis de modelos, permitiendo a los científicos de datos escribir y probar algoritmos de manera eficiente.

3. **Librería XGBoost** [49]: Es un paquete optimizado de aprendizaje de maquina que implementa algoritmos de boosting, especialmente útil para problemas de clasificación y regresión. En el contexto de préstamos para crédito de vehículos, XGBoost mejora la precisión del modelo al manejar grandes conjuntos de datos y detectar patrones complejos en las características de los solicitantes.

4. **GitHub** [50]: Es una plataforma de control de versiones y colaboración que permite a los desarrolladores almacenar y gestionar su código fuente. Facilita la colaboración entre equipos, el seguimiento

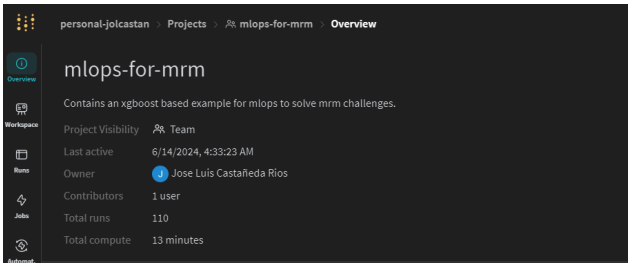
de cambios en el código y la integración de nuevas características de manera controlada en el desarrollo del modelo de crédito.

5. **Weights And Biases** [19]: Es una plataforma de gestión de experimentos y monitoreo que ayuda a rastrear, comparar y reproducir experimentos. Permite a los desarrolladores de modelos de crédito para vehículos analizar el rendimiento de diferentes versiones del modelo, optimizar hiperparámetros y mantener un registro detallado de los experimentos realizados.

6. **Docker** [51]: Es una herramienta de contenedores que permite empaquetar aplicaciones y sus dependencias en un contenedor que puede ejecutarse de manera consistente en cualquier entorno. En el desarrollo de modelos de crédito, Docker asegura que el entorno de ejecución del modelo sea el mismo en desarrollo, pruebas y producción, reduciendo problemas de compatibilidad y facilitando el despliegue continuo.

7.4 Retos solucionados en el contexto de la aplicación

7.4.1 Solución al desafío de gobierno y responsabilidad



El proyecto de elaboración del proyecto se asignó a un equipo (de una sola persona en este caso, el autor de este documento) en el que se configuran accesos y responsabilidades, este se extiende a los diferentes roles y actividades que pueden existir entre los desarrolladores, los equipos de TI, auditoría y Administración de Riesgo de modelo definiendo claramente quién accede y que acciones puede tomar.

7.4.2 Solución al desafío de validación y monitoreo

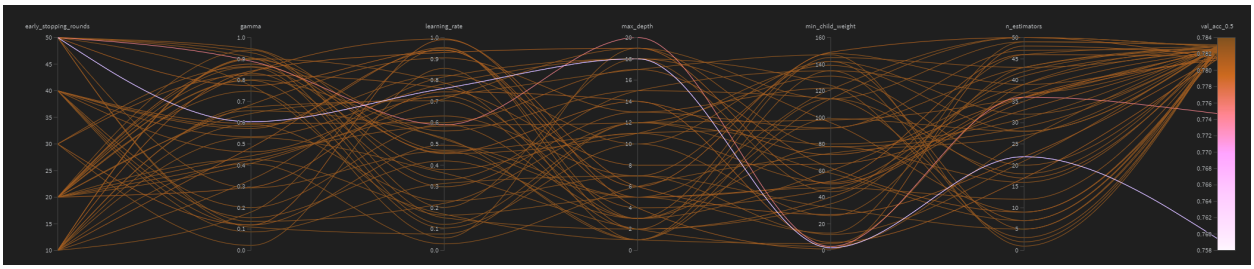


Figure 4: Evidencia de búsqueda de hiperparámetros

En el gráfico anterior, de manera general se exponen todas las búsquedas de hiperparámetros realizadas, y en cada una de ellas se encuentra el quién, con qué criterio y que método la realizo. Lo anterior absolutamente asociado al contexto del modelo y el proyecto, dando detalle y trazabilidad para la validación metodológica del desarrollo de un modelo. Para su interpretación se presenta que las líneas (corridas) cruzan cada eje en el nivel paramétrico que seleccionó y llegan al extremo derecho mostrando el nivel de precisión alcanzado.

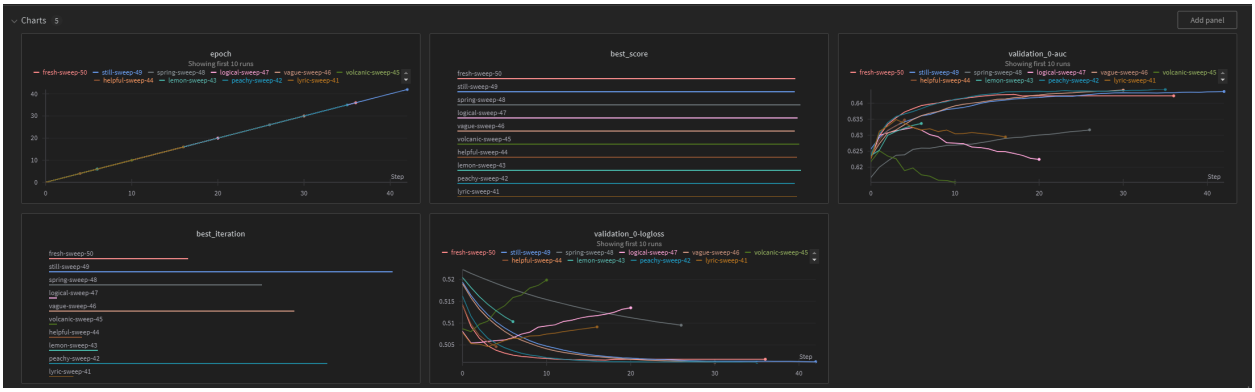


Figure 5: Desempeño de cada corrida de hiperparámetros en términos de una métrica de pérdida y de precisión

En el mismo ejercicio de búsqueda paramétrica también se puede brindar a los interesados la evidencia de que combinación fué exitosa con criterios métricos, de esta manera se puede demostrar el por qué se seleccionó un modelo en particular sobre los otros.

En el contexto de un modelo de clasificación utilizando XGBoost, las métricas Log Loss y AUC (Área Bajo la Curva ROC por sus siglas en ingles) son esenciales para evaluar el rendimiento y la efectividad del modelo. Estas métricas proporcionan información clave sobre la precisión y la capacidad de discriminación del modelo, lo cual es fundamental para garantizar predicciones confiables y útiles.

Log Loss, también conocida como pérdida logística, mide la incertidumbre de las predicciones de probabilidad del modelo. Específicamente, penaliza las predicciones que están más lejos de la etiqueta real, asignando una mayor pérdida cuanto más errónea sea la predicción. Un menor valor de Log Loss indica un modelo con mejores probabilidades de clasificación, ya que las predicciones están más cerca de las etiquetas verdaderas. Esto es particularmente relevante en problemas de clasificación donde las probabilidades predichas son tan importantes como la clasificación final.

Por otro lado, el AUC evalúa la capacidad del modelo para discriminar entre clases positivas y negativas, independientemente del umbral de decisión utilizado. Se basa en la curva ROC, que grafica la tasa de verdaderos positivos frente a la tasa de falsos positivos a diferentes umbrales. Un AUC cercano a 1 indica una excelente capacidad de discriminación, mientras que un AUC de 0.5 sugiere que el modelo no tiene mejor desempeño que un clasificador aleatorio. En un modelo de clasificación con XGBoost, un alto AUC es deseable ya que demuestra que el modelo puede separar correctamente las clases en la mayoría de los casos, lo cual es crítico en aplicaciones donde la clasificación precisa es crucial, como en la detección de fraudes o el diagnóstico médico o bien en este caso la clasificación de riesgo crediticio.

Para este caso aplicado de ilustración, Log Loss y AUC son métricas complementarias que proporcionan una visión integral del rendimiento del modelo, ayudando a los investigadores y profesionales a evaluar y comparar diferentes modelos para seleccionar el más adecuado para sus necesidades específicas.

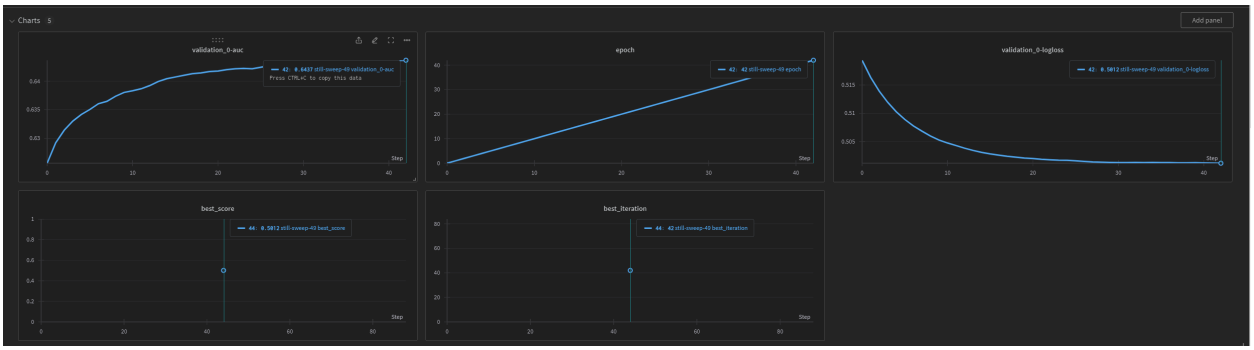


Figure 6: Detalle mejor corrida de la búsqueda de hiperparámetros

Para realizar un detalle particular del nivel de monitoreo de los modelos, se presenta la imagen anterior con un zoom sobre el desempeño particular de esa corrida, revisando el detalle sobre la métrica seleccionada. Ahora bien, no solo es de interés revisar los temas cuantitativos, pues como se mencionó en la elaboración del documento, es crucial conocer otros aspectos del modelo a postular, como el monitoreo sobre su consumo en los recursos computacionales, para lo cual también se provee el siguiente detalle.

De manera particular, estas métricas proporcionan información valiosa sobre el uso de los recursos del sistema y ayudan a identificar posibles cuellos de botella o ineficiencias que podrían afectar la ejecución del modelo.

Network Traffic (bytes): El tráfico de red mide la cantidad de datos que se transfieren entre los nodos en un clúster o entre la aplicación y el almacenamiento externo. En el caso de modelos que dependen de grandes volúmenes de datos, un monitoreo adecuado del tráfico de evita latencias significativas que puedan retrasar el tiempo de respuesta del modelo y afectar la experiencia del usuario.

Disk I/O Utilization (MB): La utilización de entrada/salida de disco mide la cantidad de datos leídos y escritos en el disco. XGBoost puede requerir acceso a grandes conjuntos de datos almacenados en disco, especialmente durante el entrenamiento y la inferencia. Un alto uso de I/O puede indicar un posible cuello de botella, lo cual podría ralentizar el proceso y aumentar los tiempos de ejecución.

Disk Utilization: Esta métrica se refiere al espacio de almacenamiento utilizado en el disco. Es importante monitorearla para asegurarse de que haya suficiente espacio disponible para almacenar datos temporales y resultados intermedios durante el procesamiento. La falta de espacio en disco puede llevar a errores inesperados y detener la ejecución del modelo.

Process CPU Threads in Use: XGBoost es conocido por su capacidad de paralelización, aprovechando múltiples hilos de CPU para acelerar el procesamiento. Monitorear los hilos de CPU en uso permite evaluar si el modelo está utilizando eficientemente los recursos de la CPU. Una subutilización podría indicar una configuración ineficaz del modelo, mientras que una sobreutilización podría afectar negativamente a otros procesos en el mismo sistema.

Process Memory Available (non-swap) (MB): La memoria disponible no intercambiada indica la cantidad de RAM disponible para el proceso sin recurrir a la memoria de intercambio en disco. XGBoost puede consumir una cantidad significativa de memoria, especialmente cuando se entrenan modelos complejos o se manejan grandes volúmenes de datos. Un monitoreo adecuado de la memoria disponible evita que el sistema se quede sin memoria y comience a utilizar el disco (swap), lo que podría degradar severamente el rendimiento.

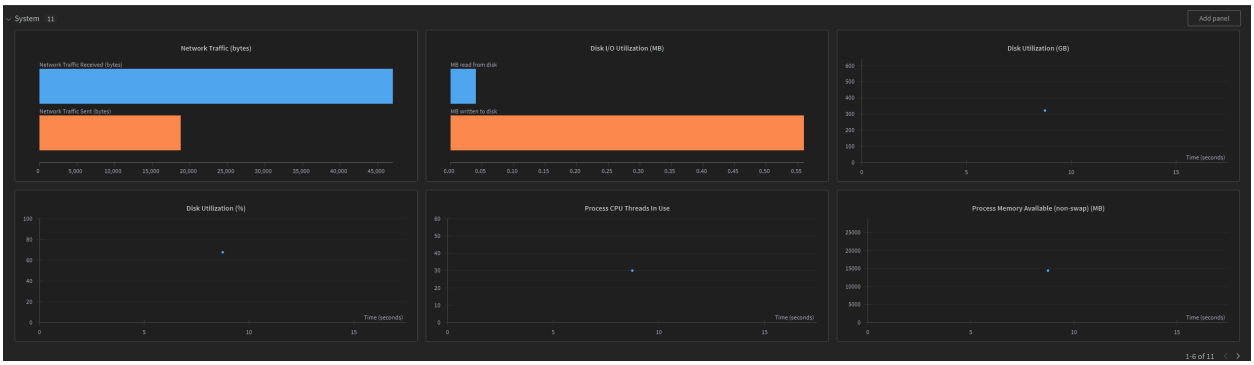


Figure 7: Monitoreo del desempeño computacional de la mejor corrida de hiperparámetros

Uno de los elementos mas relevantes para el monitoreo y la validación es saber las condiciones y características en las que se lleva a cabo esa corrida ganadora, que eventualmente se postulará como modelo. Es posible documentar entonces, como se muestra a continuación quien realizó este experimento, con qué código, sistema operativo, lenguaje, repositorio de código y hasta mediante los comandos de ejecución que se realizó.

still-sweep-49

Clone run with Launch

Description	What makes this run special?
Tags	+
Author	jolcastan
State	Finished
Start time	June 13th, 2024 at 11:09:29 pm
Sweep	1rt58ij9
Duration	8s
Run path	personal-jolcastan/mlops-for-mrm/aonu8f74
Hostname	EN2010410
OS	Windows-10-10.0.22621-SP0
Python version	3.9.13
Python executable	C:\Users\jcastaneda\OneDrive - ENDAVA\Documents\Personal\tdg\mlops_for_mrm\venv\Scripts\python.exe
Git repository	`git clone https://github.com/jolcast/mlops_for_mrm.git`
Git state	`git checkout -b "still-sweep-49" 77a73b65975c1d124193fadfce71dbb8cc0a555e`
Command	`C:\Users\jcastaneda\OneDrive - ENDAVA\Documents\Personal\tdg\mlops_for_mrm\modelling\hyperparametersearch.py`
System Hardware	CPU count 4
W&B CLI Version	0.17.1
Job Type	sweep

Config

View raw data

Config parameters are your model's inputs. ***Learn more***

Search keys with regex

Config parameters: {} 8 keys

- early_stopping_rounds: 20
- gamma: 0.7984604290463463
- **learner:** {} 6 keys
- learning_rate: 0.10494861223409212
- max_depth: 10
- min_child_weight: 26
- n_estimators: 43
- **version:** [] 3 items
  - 0: 2
  - 1: 0
  - 2: 3

Summary

View raw data

Summary metrics are your model's outputs. ***Learn more***

Search keys with regex

Summary metrics: {} 11 keys

- best_iteration: 42
- best_score: 0.5011677979115936
- epoch: 42
- Feature Importance_table: "table-file"
- train_auc: 0.6969915531303242
- train_log_loss: 0.4826276624018003
- val_acc_0.5: 0.7828578290929677
- val_auc: 0.6436618688471446
- val_log_loss: 0.5011677995989637
- validation_0-auc.max: 0.6436618688471446
- validation_0-logloss.min: 0.5011677979115936

Artifact Inputs

This run consumed these artifacts as inputs. ***Learn more***

Search 1-1 of 1

Type	Name	Consumer count
train-val-dataset	vehicle_...	106

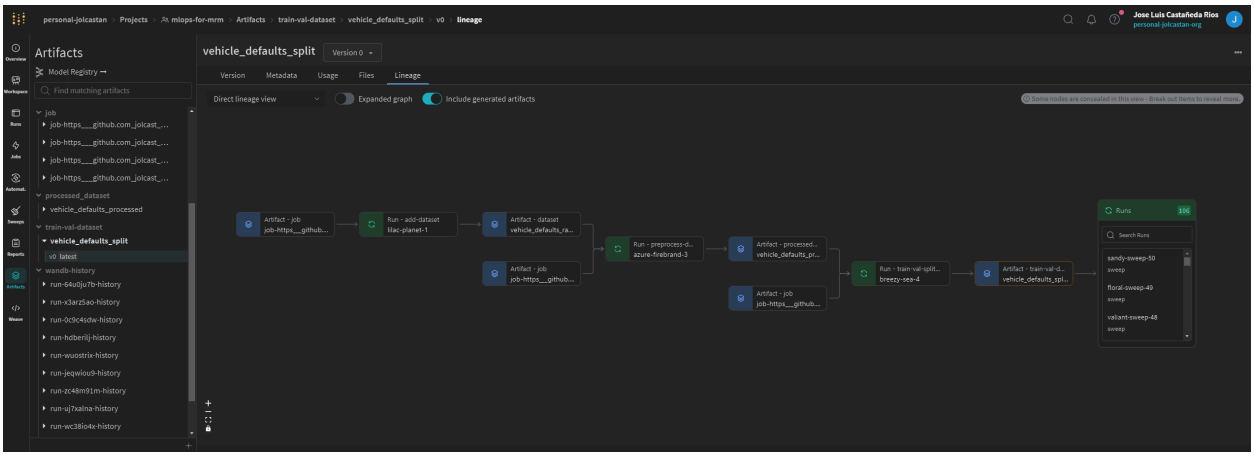
Artifact Outputs

This run produced these artifacts as outputs. ***Learn more***

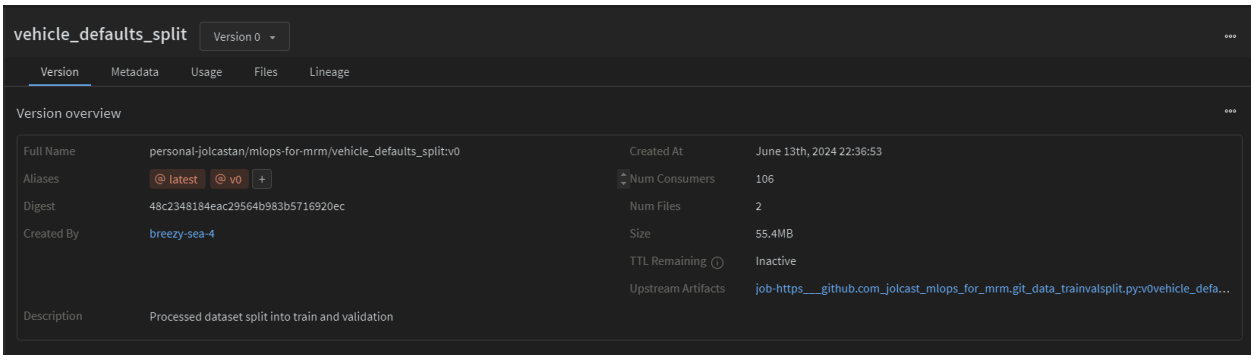
Search 1-2 of 2

Type	Name	Consumer count
run_table	run-aonu8f74-FeatureI...	0

7.4.3 Solución al desafío de calidad e integridad de los datos



Para cada conjunto de datos existe completa trazabilidad desde su origen, sus versiones, quien lo generó y con qué versión de código del repositorio fué generado. Esto permite constantemente validar el linaje y documentar de manera sistemática los datos utilizados en cada experimento. El ejemplo en la imagen muestra que uno de los artefactos registrados set de validación y de entrenamiento, cuyo origen es el primer job que genero la transformación de los datos crudos, luego un job que los procesa, luego un job que los divide, generando un artefacto v0 que contiene los archivos generados.



De manera particular se documenta que versión se está utilizando del artefacto, cuando fue creada, que script la generó y el autor. Además se cuenta con una descripción del contenido y la metadata paramétrica que se utilizo para esta versión en particular.

7.4.4 Solución al desafío de administración del inventario de modelos

Si bien se ha insistido sobre la trazabilidad en la metodología de MLOps, es también crucial ofrecer un inventario centralizado y documentado de los modelos. Para este caso en la figura inferior se muestra la tarjeta de un modelo y un repositorio de sus múltiples versiones, entre ellas la mas reciente (latest) y la declarada como mejor (champion) en el contexto de su relativo mejor desempeño métrico con respecto a su versión anterior.

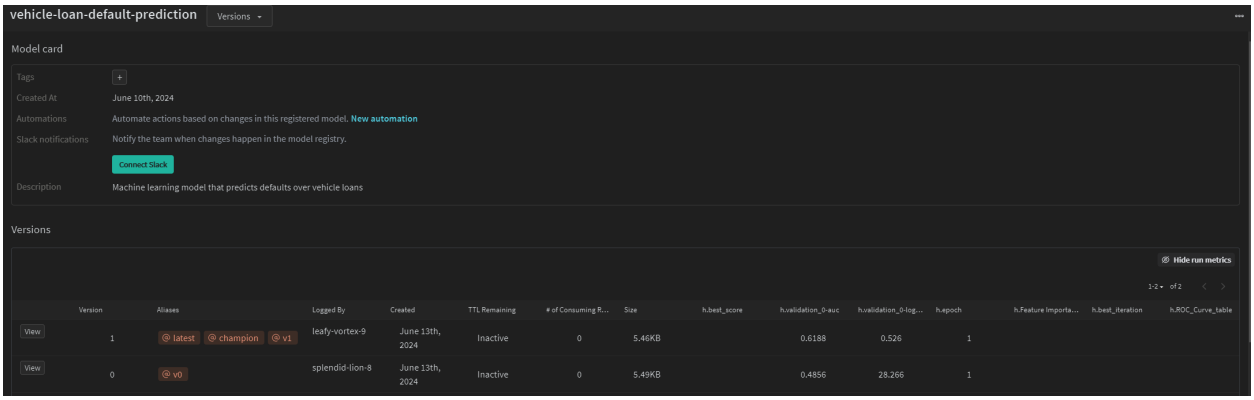


Figure 9: Tarjeta del modelo y registro como artefacto

Este no es solo un inventario descriptivo, sino también un repositorio de objetos que permiten consumir un modelo y disponerlo de manera regulada a sus eventuales consumidores, brindando un detalle de su elaboración y las partes que intervinieron en su desarrollo y aprobación.

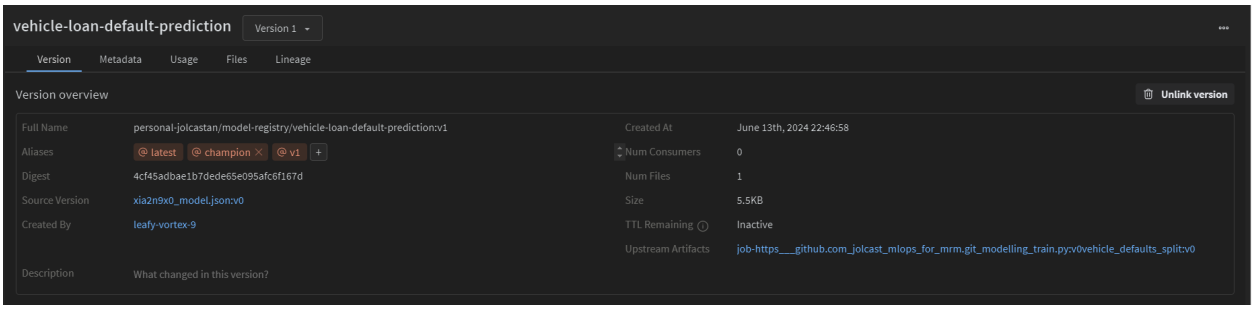


Figure 10: Detalle del inventario del mejor modelo registrado y aprobado

En el detalle anterior se puede evidenciar como llego allí ese modelo, quién lo aprobó, cuándo, usando qué código, la metadata y los parámetros utilizados, además de quién lo está consumiendo actualmente.

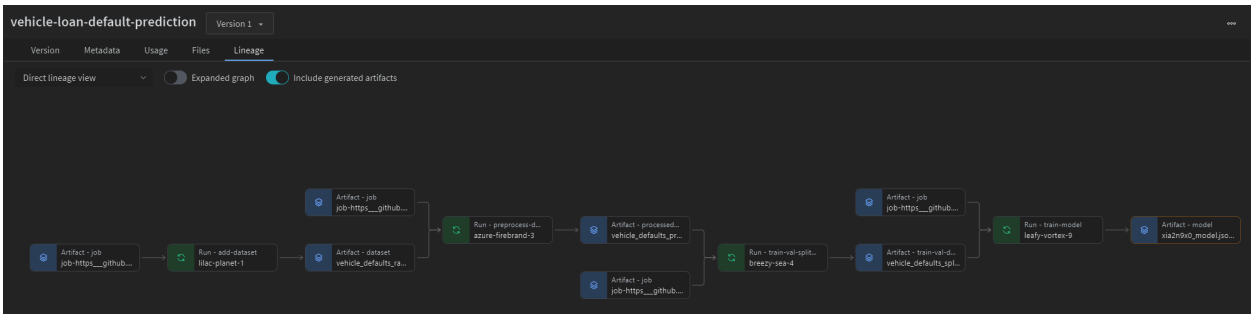


Figure 11: Linaje mejor modelo seleccionado

Como detalle adicional, y como con cualquier otro artefacto del proyecto, se puede evidenciar el linaje del modelo, permitiendo al validador conocer su origen de principio a fin, los roles que participaron en su elaboración, el código que se utilizó, el criterio de selección que se documentó, entre otros.

7.5 Aspectos éticos del caso aplicado

En la elaboración de un modelo predictivo para identificar posibles incumplimientos de pago en el financiamiento de vehículos, , se deben considerar varios aspectos éticos relevantes para asegurar la justicia, la transparencia y la responsabilidad en su implementación y uso:

- **Transparencia y Explicabilidad**: Los solicitantes de crédito deben entender cómo y por qué se toman decisiones que afectan sus solicitudes. Esto implica proporcionar explicaciones claras y comprensibles de los factores que influyen en la decisión del modelo, especialmente cuando se rechaza un préstamo o se imponen condiciones más estrictas.

- **No Discriminación y Equidad**: Se debe garantizar que el modelo no perpetúe sesgos ni discrimine a los solicitantes de crédito en función de características protegidas, como raza, género, edad, orientación sexual, o estado civil. Deben realizarse evaluaciones exhaustivas para identificar y mitigar cualquier sesgo en los datos o en los resultados del modelo, asegurando un trato justo y equitativo para todos los solicitantes.

- **Privacidad y Protección de Datos**: El manejo de datos personales y financieros sensibles exige un alto nivel de responsabilidad. Es necesario garantizar la protección y la confidencialidad de la información de los solicitantes, cumpliendo con las leyes y regulaciones de privacidad de datos aplicables, como el GDPR[52] o la CCPA[53]. Además, los datos deben ser utilizados únicamente para los fines para los que fueron recolectados y con el consentimiento explícito de los individuos.

- **Impacto en los Solicitantes**: Las decisiones basadas en modelos predictivos pueden tener un impacto significativo en la vida de los solicitantes, como el acceso a financiamiento o la imposición de términos más estrictos. Se deben tener en cuenta las implicaciones éticas de estas decisiones y, cuando sea necesario, proporcionar mecanismos para que los solicitantes puedan apelar o cuestionar las decisiones automáticas.

- **Responsabilidad y Supervisión**: Las instituciones financieras deben asumir la responsabilidad por las decisiones tomadas por sus modelos predictivos. Esto incluye la supervisión continua del modelo para garantizar que funcione según lo previsto y que no cause daños injustificados a los solicitantes. También implica estar preparados para realizar ajustes en el modelo si se identifican problemas éticos o de desempeño.

- **Minimización de Daños**: Debe haber un enfoque en minimizar cualquier daño potencial causado por la implementación del modelo, como la exclusión financiera o la imposición de condiciones de crédito injustas. Esto puede incluir la provisión de opciones alternativas o programas de apoyo para los solicitantes identificados como de alto riesgo.

Discusión

El futuro de la Administración del Riesgo de Modelos (MRM) está destinado a evolucionar significativamente a medida que las instituciones financieras buscan integrarla más completamente en sus marcos de administración de riesgos. Un área crítica será el avance y la automatización de los procesos de MRM, especialmente con la adopción de técnicas de modelado cuantitativo y algoritmos de aprendizaje automático[30][33]. Este cambio requiere el desarrollo de marcos robustos y adaptativos capaces de gestionar las complejidades introducidas por estas tecnologías, en donde, como se construyó en el artículo, las metodologías proveerán herramientas para facilitar su implementación.

Además, se espera que el panorama regulatorio se vuelva más estricto. Los reguladores enfatizan la necesidad de documentación integral y estructuras de gobernanza claras en torno al riesgo de modelos, lo que impulsa a las instituciones a adoptar prácticas de validación y monitoreo más rigurosas[34]. Las directrices de la Reserva Federal y la Autoridad Bancaria Europea subrayan la importancia de mantener altos estándares para el desarrollo y la validación de modelos, e integrar la evaluación del riesgo de modelos en estrategias de administración de riesgos más amplias. De esta manera, las herramientas de documentación y versionado de modelo de datos tomarán un papel aún mas importante en el cumplimiento regulatorio.

En el contexto actual, los cambios estructurales del mercado, las técnicas de aprendizaje automático y las iniciativas de administración del riesgo climático presentarán algunos de los desafíos y oportunidades más interesantes para la administración del riesgo de modelos[54]. La evolución económica y los cambios en las tasas de interés y la inflación requerirán que los equipos de modelado adapten sus herramientas a condiciones económicas desconocidas y la administración de riesgos que aún hoy representan un reto metodológico, fortaleciendo los vínculos entre los factores económicos y los resultados del riesgo de mercado y crédito.

Las discusiones futuras en MRM probablemente se centrarán en los desafíos de gestionar un número creciente de modelos con ingredientes de alta complejidad, especialmente a medida que las instituciones financieras amplían su uso de estos para la toma de decisiones en diversos dominios. La necesidad de una asignación eficiente de recursos a través de la clasificación de riesgo de modelos es crítica, ya que asegura que los esfuerzos de validación sean proporcionales al riesgo que presenta cada modelo[33]. Este enfoque ayuda a priorizar los modelos que tienen un mayor impacto en la estabilidad financiera y operativa de la institución.

Finalmente, el papel de la tecnología y la administración de datos no puede subestimarse y se considera evidente su relación con el riesgo de modelo. El futuro de MRM verá una mayor dependencia de herramientas automatizadas para la administración de datos, la validación de modelos y el seguimiento del rendimiento. Esta automatización no solo mejorará la eficiencia de los procesos de MRM, sino que también reducirá el riesgo de error humano, asegurando resultados de modelos más precisos y confiables[34].

Conclusiones

La principal contribución de este trabajo académico es la de proponer que las metodologías Operaciones de Aprendizaje Automático (MLOps) ofrecen soluciones de hecho a los complejos retos de la Administración de Riesgo de Modelo, a través de métodos y herramientas que resuelven problemas ya existentes en la gestión del ciclo de vida de un modelo de Aprendizaje de Automático (ML). Adicionalmente, este documento ofrece esta contribución con un tono novedoso, pues como bien se mostró en la revisión sistemática de literatura, no se encuentran trabajos académicos que traten esta sinergia de metodologías. En términos particulares, la anterior afirmación se puede extender así:

- Se ha identificado una brecha significativa en la literatura sobre la integración de metodologías de MLOps en la administración de riesgo de modelo. La falta de estudios específicos que aborden esta relación resalta la necesidad de desarrollar investigaciones adicionales para explorar y validar las propuestas y enfoques presentados en este documento.

- La integración de metodologías de Operaciones de Aprendizaje Automático (MLOps) en la administración de riesgos de modelos (MRM) puede mejorar significativamente la fiabilidad y la robustez de los modelos. Esta integración permite automatizar procesos, asegurar el cumplimiento regulatorio, el gobierno, las revisiones independientes y facilitar una gestión más efectiva del ciclo de vida de los modelos.

- Se espera que el panorama regulatorio se vuelva más estricto. Los reguladores enfatizan la necesidad de documentación integral y estructuras de gobierno claras en torno al riesgo de modelos, lo que impulsa a las instituciones a adoptar prácticas de validación y monitoreo más rigurosas. Las herramientas de documentación y versionado de modelo de datos tomarán un papel aún más importante en el cumplimiento regulatorio.

- El futuro de la Administración del Riesgo de Modelos (MRM) está destinado a evolucionar significativamente a medida que las instituciones financieras buscan integrarla más completamente en sus marcos de administración de riesgos. Un área crítica será el avance y la automatización de los procesos de MRM, especialmente con la adopción de técnicas de modelado cuantitativo y algoritmos de aprendizaje automático.

- El riesgo de modelo no es exclusivo del sector financiero; es una preocupación universal para todas las industrias que utilizan modelos de aprendizaje automático. Los efectos adversos de una mala administración operativa de un modelo pueden impactar diversas organizaciones, independientemente de su sector, destacando la necesidad de prácticas robustas de gestión de riesgos en todos los campos.

Referencias

Operaciones de modelos de aprendizaje automático

[2] D Sculley et al. "Hidden Technical Debt in Machine Learning Systems". In: ed. by C Cortes et al. Vol. 28. Curran Associates, Inc., 2015. URL: https://proceedings.neurips.cc/paper_files/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf.

[3] Tsakani Mboweni, Themba Masombuka, and Cyrille Dongmo. "A Systematic Review of Machine Learning DevOps". In: Institute of Electrical and Electronics Engineers Inc., 2022. ISBN: 9781665470872. DOI: 10.1109/ICECET55527.2022.9872968.

[4] Christian Haertel et al. "MLOps in Data Science Projects: A Review". In: Institute of Electrical and Electronics Engineers Inc., 2023, pp. 2396–2404. ISBN: 9798350324457. DOI: 10.1109/BigData59044.2023.10386139.

[5] Samridhi Jain and Puneet Kumar. "Cost Effective Generic Machine Learning Operation: A Case Study". In: Institute of Electrical and Electronics Engineers Inc., 2023. ISBN: 9798350301595. DOI: 10.1109/ICDSNS58469.2023.10245408.

[6] Matteo Testi et al. *MLOps: A Taxonomy and a Methodology*. 2022. DOI: 10.1109/ACCESS.2022.3181730.

[7] Georgios Symeonidis et al. "MLOps - Definitions, Tools and Challenges". In: Institute of Electrical and Electronics Engineers Inc., 2022, pp. 453–460. ISBN: 9781665483032. DOI: 10.1109/CCWC54503.2022.9720902.

[8] Meenu Mary John, Helena Holmstrom Olsson, and Jan Bosch. "Towards MLOps: A Framework and Maturity Model". In: Institute of Electrical and Electronics Engineers Inc., Sept. 2021, pp. 334–341. ISBN: 9781665427050. DOI: 10.1109/SEAA53835.2021.00050.

[9] Dominik Kreuzberger, Niklas Kuhl, and Sebastian Hirschl. "Machine Learning Operations (MLOps): Overview, Definition, and Architecture". In: *IEEE Access* 11 (2023), pp. 31866–31879. ISSN: 21693536. DOI: 10.1109/ACCESS.2023.3262138.

[10] "The pipeline for the continuous development of artificial intelligence models—Current state of research and practice". In: *Journal of Systems and Software* 199 (May 2023). ISSN: 01641212. DOI: 10.1016/j.jss.2023.111615.

[14] Nathalie Rauschmayr et al. "Profiling Deep Learning Workloads at Scale using Amazon SageMaker". In: Association for Computing Machinery, Aug. 2022, pp. 3801–3809. ISBN: 9781450393850. DOI: 10.1145/3534678.3539036.

[15] Tiago Yukio Fujii et al. "A Digital Twin Architecture Model Applied with MLOps Techniques to Improve Short-Term Energy Consumption Prediction". In: *Machines* 10 (1 Jan. 2022). ISSN: 20751702. DOI: 10.3390/machines10010023.

[20] Ask Berstad Kolltveit and Jingyue Li. "Operationalizing Machine Learning Models - A Systematic Literature Review". In: Institute of Electrical and Electronics Engineers Inc., 2022, pp. 1–8. ISBN: 9781450393195. DOI: 10.1145/3526073.3527584.

[21] Sini Vänskä et al. "Continuous Software Engineering Practices in AI/ML Development Past the Narrow Lens of MLOps: Adoption Challenges". In: *E-Informatica Software Engineering Journal* 18 (1 2024). ISSN: 20844840. DOI: 10.37190/e-Inf240102.

[25] Tim Raffin et al. "Qualitative assessment of the impact of manufacturing-specific influences on Machine Learning Operations". In: vol. 115. Elsevier B.V., 2022, pp. 136–141. DOI: 10.1016/j.procir.2022.10.063.

[26] Eduardo e Oliveira et al. "Unlabeled learning algorithms and operations: overview and future trends in defense sector". In: *Artificial Intelligence Review* 57 (3 Mar. 2024). ISSN: 15737462. DOI: 10.1007/s10462-023-10692-0.

[27] Eyad Kannout, Michal Grodzki, and Marek Grzegorowski. "Considering various aspects of models' quality in the ML pipeline - application in the logistics sector". In: Institute of Electrical and Electronics Engineers Inc., 2022, pp. 403–412. ISBN: 9788396589712. DOI: 10.15439/2022F296.

[28] Rakshith Subramanya, Seppo Sierla, and Valeriy Vyatkin. "From DevOps to MLOps: Overview and Application to Electricity Market Forecasting". In: *Applied Sciences (Switzerland)* 12 (19 Oct. 2022). ISSN: 20763417. DOI: 10.3390/app12199851.

[29] Andrew Ng. *Introduction to Machine Learning in Production*. https://www.coursera.org/learn/introduction-to-machine-learning-in-production. Available at https://www.coursera.org/learn/introduction-to-machine-learning-in-production. 2024.

[35] Josu Diaz-De-Arcaya et al. "A Joint Study of the Challenges, Opportunities, and Roadmap of MLOps and AIOps: A Systematic Survey". In: *ACM Computing Surveys* 56 (4 Oct. 2023). ISSN: 15577341. DOI: 10.1145/3625289.

[36] Qinghua Lu et al. "Responsible AI Pattern Catalogue: A Collection of Best Practices for AI Governance and Engineering". In: *ACM Computing Surveys* 56 (7 Apr. 2024). ISSN: 15577341. DOI: 10.1145/3626234.

[38] Prerna Singh. "Systematic review of data-centric approaches in artificial intelligence and machine learning". In: *Data Science and Management* 6 (3 Sept. 2023), pp. 144–157. ISSN: 26667649. DOI: 10.1016/j.dsm.2023.06.001.

[39] *The machine learning life cycle and the cloud: implications for drug discovery*. 2021. DOI: 10.1080/17460441.2021.1932812.

[40] Alain Hennebelle et al. "Secure and privacy-preserving automated machine learning operations into end-to-end integrated IoT-edge-artificial intelligence-blockchain monitoring system for diabetes mellitus prediction". In: *Computational and Structural Biotechnology Journal* 23 (Dec. 2024), pp. 212–233. ISSN: 20010370. DOI: 10.1016/j.csbj.2023.11.038.

[41] Ningxin Su and Baochun Li. "MLOps in the Metaverse: Human-Centric Continuous Integration". In: *IEEE Journal on Selected Areas in Communications* 42 (3 Mar. 2024), pp. 737–751. ISSN: 15580008. DOI: 10.1109/JSAC.2023.3345385.

Administración de riesgo de modelo

[22] Basel Committee on Banking Supervision. *Basel III: International Regulatory Framework for Banks*. https://www.bis.org/bcbs/basel3.htm. Bank for International Settlements. 2011.

[24] Board of Governors of the Federal Reserve System Office of the Comptroller of the Currency. *Supervisory guidance on model risk management contents*. 2011.

[30] M. Garro. "The evolution of model risk management processes". In: *Journal of Risk Management in Financial Institutions* 13 (1 2020), pp. 16–23.

[31] Jingwei Gan et al. "Automated Counterfactual Generation in Financial Model Risk Management". In: Institute of Electrical and Electronics Engineers Inc., 2021, pp. 4064–4068. ISBN: 9781665439022. DOI: 10.1109/BigData52589.2021.9671561.

[32] J. P. Florens, C. Gourieroux, and A. Monfort. "Model risk management: Limits and future of Bayesian approaches". In: *Annals of Economics and Statistics* (136 2019), pp. 1–26. ISSN: 19683863. DOI: 10.15609/annaeconstat2009.136.0001.

[33] Nick Kiritz, Miles Ravitz, and Mark Levonian. "Model risk tiering: An exploration of industry practices and principles". In: *Journal of Risk Model Validation* 13 (2 2019), pp. 47–77. ISSN: 17539587. DOI: 10.21314/JRMV.2019.205.

[34] KPMG International. *Increasing challenges for model risk management (MRM) within financial institutions*. Retrieved from https://assets.kpmg.com/content/dam/kpmg/es/pdf/2023/01/model-risk-management.pdf. 2023. URL: https://assets.kpmg.com/content/dam/kpmg/es/pdf/2023/01/model-risk-management.pdf.

[37] Jon R Hill. *The top 14 challenges for today's model risk managers: Has the time come to think about going beyond SR11-7?* 2019.

[52] European Parliament and Council of the European Union. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)*. Official Journal of the European Union, L 119, 1-88. 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[53] California State Legislature. *California Consumer Privacy Act of 2018 (CCPA)*. California Civil Code Title 1.81.5, Section 1798.100-1798.199. 2018. URL: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375.

Herramientas

[1] *Scopus: Access and Use Research Publications*. https://www.elsevier.com/solutions/scopus. Accessed: 2024-05-26. 2024.

[11] Apache Software Foundation. *Apache Airflow*. https://airflow.apache.org/. Version 2.5.0. 2024.

[12] Kubeflow Community. *Kubeflow Pipelines*. https://www.kubeflow.org/docs/components/pipelines/. Version 1.8.1. 2024.

[13] Amazon Web Services. *SageMaker Pipelines*. https://aws.amazon.com/sagemaker/pipelines/. Version 2.0. 2024.

[16] Git Community. *Git*. https://git-scm.com/. Version 2.41.0. 2024.

[17] Iterative. *DVC: Data Version Control*. https://dvc.org. Version 2.0. 2024.

[18] Databricks. *MLFlow*. https://mlflow.org/. Version 2.4.1. 2024.

[19] Weights & Biases. *Weights & Biases*. https://wandb.ai/. Version 0.15.5. 2024.

[42] Slack Technologies. *Slack*. https://slack.com/. Version 4.23.0. 2024.

[43] Microsoft Corporation. *Microsoft Teams*. https://www.microsoft.com/en/microsoft-teams/group-chat-software. Version 1.4.00.4167. 2024.

[44] Atlassian. *Confluence*. https://www.atlassian.com/software/confluence. Version 7.13.0. 2024.

[45] Inc. Notion Labs. *Notion*. https://www.notion.so/. Version 2.0.17. 2024.

[46] Castañeda Rios Jose Luis. *MLOps for Model Risk Management*. https://github.com/jolcast/mlops_for_mrm.git. GitHub repository. 2024. URL: https://github.com/jolcast/mlops_for_mrm.git.

[47] Sneha Shinde. *LTFS Data for Vehicle Loan Default Prediction*. https://www.kaggle.com/datasets/sneharshinde/ltfs-av-data. Kaggle Dataset. 2024. URL: https://www.kaggle.com/datasets/sneharshinde/ltfs-av-data.

[48] Python Software Foundation. *Python Programming Language*. https://www.python.org. Version 3.9. 2024. URL: https://www.python.org.

[49] Tianqi Chen and Carlos Guestrin. *XGBoost: A Scalable Tree Boosting System*. https://xgboost.readthedocs.io. Python package version 1.4.2. 2016. URL: https://xgboost.readthedocs.io.

[50] Inc. GitHub. *GitHub*. https://github.com. Version Control Platform. 2024. URL: https://github.com.

[51] Inc. Docker. *Docker*. https://www.docker.com. Software. 2024. URL: https://www.docker.com.

Anexos

Autores	Titulo	Abstract	Keywords	DOI	Acción
Lu, Qinghua and Zhu, Liming and Xu, Xiwei and Whittle, Jon and Zowghi, Didar and Jacquet, Aurelie	Responsible AI Pattern Catalogue: A Collection of Best Practices for AI Governance and Engineering	Responsible Artificial Intelligence (RAI) is widely considered as one of the greatest scientific challenges of our time and is key to increase the adoption of Artificial Intelligence (AI). Recently, a number of AI ethics principles frameworks have been published. However, without further guidance on best practices, practitioners are left with nothing much beyond truisms. In addition, significant efforts have been placed at algorithm level rather than system level, mainly focusing on a subset of mathematics-amenable ethical principles, such as fairness. Nevertheless, ethical issues can arise at any step of the development lifecycle, cutting across many AI and non-AI components of systems beyond AI algorithms and models. To operationalize RAI from a system perspective, in this article, we present an RAI Pattern Catalogue based on the results of a multivocal literature review. Rather than staying at the principle or algorithm level, we focus on patterns that AI system stakeholders can undertake in practice to ensure that the developed AI systems are responsible throughout the entire governance and engineering lifecycle. The RAI Pattern Catalogue classifies the patterns into three groups: multi-level governance patterns, trustworthy process patterns, and RAI-by-design product patterns. These patterns provide systematic and actionable guidance for stakeholders to implement RAI. © 2024 Copyright held by the owner/author(s).	AI engineering; AI governance; best practice; ethical AI; MLOps; pattern; Responsible AI; software architecture; software engineering; trustworthy AI	10.1145/3626234	Incluido
Cohen, Robert	Digital Strategy, Machine Learning, and Industry Survey of MLOps	As part of a digital strategy, machine learning (ML) has become a common toolset and capability across many businesses. However, the operational aspects of machine learning (MLOps) are often overlooked for ML projects until they are already installed and being executed in the business environment. This chapter provides a review of MLOps products and vendors to give data scientists the ability to set up the appropriate ML infrastructure in a proactive manner. © 2023 by World Scientific Publishing Co. Pte. Ltd.	machine learning, ml, mlops	10.1142/9789811271984_0008	Excluido
Symeonidis, Georgios and Nerantzis, Evangelos and Kazakis, Apostolos and Papakostas, George A.	MLOps - Definitions, Tools and Challenges	This paper is an concentrated overview of the Machine Learning Operations (MLOps) area. Our aim is to define the operation and the components of such systems by highlighting the current problems and trends. In this context we present the different tools and their usefulness in order to provide the corresponding guidelines. Moreover, the connection between MLOps and AutoML (Automated Machine Learning) is identified and how this combination could work is proposed. The novelty of our approach relies on the combination of state-of-the-art topics such as AutoML, explainability and sustain-ability in order to overcome the current challenges in MLOps identifying them not only as the answer for the incorporation of ML models in production but also as a possible tool for efficient, robust and accurate machine learning models. © 2022 IEEE.	AutoML; Deployment; Explainability; Fairness; Machine learning; MLOps; Monitoring; Re-training; Robustness; Sustainability	10.1109/CCWC54503.2022.9720902	Incluido
Paul, Souma Kanti and Riaz, Sadia and Das, Suchismita	A Conceptual Architecture for AI in Supply Chain Risk Management	A high level conceptual architecture is proposed for Artificial Intelligence (AI) in Supply Chain Risk Management (SCRM) based on literature review and qualitative research. It incorporates architecture components across data management, model development, and model deployment. © 2022 IEEE.	AI; AutoML; Building Blocks; Components; MLOps; SCRM	10.1109/TENCON55691.2022.9978154	Excluido
Patel, Dhaval and Shrivastava, Shrey and Gifford, Wesley and Siegel, Stuart and Kalagnanam, Jayant and Reddy, Chandra	Smart-ML: A System for Machine Learning Model Exploration using Pipeline Graph	In this paper, we describe an overarching ML system with a simple programming interface that leverages existing AI and ML frameworks to make the task of model exploration easier. The proposed system introduces a new programming construct namely pipeline graph (a directed acyclic graph) consisting of multiple machine learning operations provided by different ML repositories. End user uses the pipeline graph as a common interface for modelling different ML tasks such as classification, regression, and time-series prediction, while enabling efficient execution on different environments (Spark, Celery and Cloud). We further annotated the pipeline graph with a hyper-parameter grid and an option to try-out a wide range of optimization strategies (i.e., Random, Bayesian, Bandit, AutoLearn, etc). Given a large pre-defined pipeline graph along with its hyper-parameters, we provided a general-purpose, scalable and efficient pipeline-graph exploration technique to provide the automated solutions to a variety of ML tasks. We compare our automated approach to several state-of-the-art automated AI systems and find that we achieve performance comparable to the best results, while often producing simpler pipelines using off-the-shelf components. Our evaluation suite consists of experiments on 60+ classifications and regressions datasets. © 2020 IEEE.	automated, different, exploration, graph, hyper	10.1109/BigData50022.2020.9378082	Excluido
Cob-Parro, Antonio Carlos and Lalangui, Yerhard and Lazcano, Raquel	Fostering Agricultural Transformation through AI: An Open-Source AI Architecture Exploiting the MLOps Paradigm	As the global population is expected to reach 10 billion by 2050, the agricultural sector faces the challenge of achieving an increase of 60% in food production without using much more land. This paper explores the potential of Artificial Intelligence (AI) to bridge this "land gap" and mitigate the environmental implications of agricultural land use. Typically, the problem with using AI in such agricultural sectors is the need for more specific infrastructure to enable developers to design AI and ML engineers to deploy these AIs. It is, therefore, essential to develop dedicated infrastructures to apply AI models that optimize resource extraction in the agricultural sector. This article presents an infrastructure for the execution and development of AI-based models using open-source technology, and this infrastructure has been optimized and tuned for agricultural environments. By embracing the MLOps culture, the automation of AI model development processes is promoted, ensuring efficient workflows, fostering collaboration among multidisciplinary teams, and promoting the rapid deployment of AI-driven solutions adaptable to changing field conditions. The proposed architecture integrates state-of-the-art tools to cover the entire AI model lifecycle, enabling efficient workflows for data scientists and ML engineers. Considering the nature of the agricultural field, it also supports diverse IoT protocols, ensuring communication between sensors and AI models and running multiple AI models simultaneously, optimizing hardware resource utilization. Surveys specifically designed and conducted for this paper with professionals related to AI show promising results. These findings demonstrate that the proposed architecture helps close the gap between data scientists and ML engineers, easing the collaboration between them and simplifying their work through the whole AI model lifecycle. © 2024 by the authors.	Agriculture; Artificial Intelligence; Deep Learning; DevOps; Internet of Things; Machine Learning; MLOps	10.3390/agronomy14020259	Excluido
Pelekis, Sotiris and Karakolis, Evangelos and Silva, Francisco and Schoinas, Vasilelos and Mouzakitis, Spiros and Kormpakis, Georgios and Amaro, Nuno and Psarras, John	In Search of Deep Learning Architectures for Load Forecasting: A Comparative Analysis and the Impact of the Covid-19 Pandemic on Model Performance	In power grids, short-term load forecasting (STLF) is crucial as it contributes to the optimization of their reliability, emissions, and costs, while it enables the participation of energy companies in the energy market. STLF is a challenging task, due to the complex demand of active and reactive power from multiple types of electrical loads and their dependence on numerous exogenous variables. Amongst them, special circumstances such as the COVID-19 pandemic can often be the reason behind distribution shifts of load series. This work conducts a comparative study of Deep Learning (DL) architectures namely Neural Basis Expansion Analysis Time Series Forecasting (N-BEATS), Long Short-Term Memory (LSTM), and Temporal Convolutional Networks (TCN) with respect to forecasting accuracy and training sustainability, meanwhile examining their out-of-distribution generalization capabilities during the COVID-19 pandemic era. A Pattern Sequence Forecasting (PSF) model is used as baseline. The case study focuses on day-ahead forecasts for the Portuguese nation a 15-minute resolution net load time series. The results can be leveraged by energy companies and network operators (i) to reinforce their forecasting toolkit with state-of-the-art DL models; (ii) to become aware of the serious consequences of crisis events on model performance; (iii) as a high-level model evaluation, deployment, and sustainability guide within a smart grid context. © 2022 IEEE.	COVID-19; Deep Learning; Forecasting; LSTM; MLOps; NBEATS; Out-of-Distribution Generalization; Short-Term Load Forecasting; Smart Grid; Sustainability; Temporal Convolution	10.1109/IISA56318.2022.9904363	Excluido
Dzieciolowski, Krzysztof and Zhao, Ningsheng	TOWARDS EXPLAINABLE MACHINE LEARNING OPERATIONS (MLOPS)	In the dynamically expanding AI applications, there is a need for an effective operational framework for managing ever increasing number of predictive models. A new field of Machine Learning Operations (MLOps) has recently emerged to address the issue of efficient operations and management of machine learning models. However, enterprises still struggle to develop effective MLOps systems due to lack of expertise and limited experience in AI operational needs. Attempts of an ad-hoc administration of ML algorithms have not been successful to deal with the challenges of fast proliferation of AI models. At the same time, existing concepts of MLOps frameworks don't adequately address an important question of accurate machine learning models diagnostics and explainability. In this reflection paper, we review commonly used MLOps approaches and explainability methods and suggest novel methods to address challenges of current model explainability methods. We discuss the need for university curriculum to address the issues facing management of AI models, their diagnostics and explainability. © ICT 2023.All rights reserved.	Importance Sampling; Machine Learning Operations (MLOps); Model Diagnostics; Model Explanations	N/A	Excluido

John, Meenu Mary and Olsson, Helena Holmstrom and Bosch, Jan	Towards MLOps: A Framework and Maturity Model	The adoption of continuous software engineering practices such as DevOps (Development and Operations) in business operations has contributed to significantly shorter software development and deployment cycles. Recently, the term MLOps (Machine Learning Operations) has gained increasing interest as a practice that brings together data scientists and operations teams. However, the adoption of MLOps in practice is still in its infancy and there are few common guidelines on how to effectively integrate it into existing software development practices. In this paper, we conduct a systematic literature review and a grey literature review to derive a framework that identifies the activities involved in the adoption of MLOps and the stages in which companies evolve as they become more mature and advanced. We validate this framework in three case companies and show how they have managed to adopt and integrate MLOps in their large-scale software development companies. The contribution of this paper is threefold. First, we review contemporary literature to provide an overview of the state-of-the-art in MLOps. Based on this review, we derive an MLOps framework that details the activities involved in the continuous development of machine learning models. Second, we present a maturity model in which we outline the different stages that companies go through in evolving their MLOps practices. Third, we validate our framework in three embedded systems case companies and map the companies to the stages in the maturity model. © 2021 IEEE.	Framework; GLR; Maturity Model; MLOps; SLR; Validation Study	10.1109/SEAA53835.2021.00050	Incluido
Nájera, Vanessa Landero and Ortega, Joaquín Pérez and Collazos Morales, Carlos Andrés and Rodríguez Aguilar, Sandra Silvia	From Selecting Best Algorithm to Explaining Why It is: A General Review, Formal Problem Statement and Guidelines Towards to an Empirical Generalization	It has been observed on solution algorithms for problems as sorting, forecasting, classification, clustering, constraint satisfaction, decision, optimization from several disciplines (computational complexity theory, data mining, artificial intelligence, machine learning, operations research) that algorithm performance is better in certain problem instances than other. This paper describes how has been the way for trying to reach the empirical generalization for this phenomenon existing in the experimental relation problem – algorithm. For each understanding level, research questions, problem description were formulated, using the same Rice's nomenclature and supplementing it; as well as, influence indexes and analysis approaches were described. A diagram about this long trajectory and a reflection is performed, highlighting contributions and scope. It shows that up to now the problem of explaining formally why an algorithm is the best for solving an instance set had remained open. A formal problem statement for describing this phenomenon and a general framework were proposed as a guide for working in adequate way toward generation of theories; which could contribute to build generalized indexes and self-adaptive algorithms to give the best solution to problems. © 2023, The Author(s), under exclusive license to Springer Nature Switzerland AG.	algorithm, algorithms, best, indexes, phenomenon	10.1007/978-3-031-36805-9_45	Excluido
e Oliveira, Eduardo and Rodrigues, Marco and Pereira, João Paulo and Lopes, António M. and Mestric, Ivana Ilic and Bjelogrlic, Sandro	Unlabeled learning algorithms and operations: overview and future trends in defense sector	In the defense sector, artificial intelligence (AI) and machine learning (ML) have been used to analyse and decipher massive volumes of data, namely for target recognition, surveillance, threat detection and cybersecurity, autonomous vehicles and drones guidance, and language translation. However, there are key points that have been identified as barriers or challenges, especially related to data curation. For this reason, and also due to the need for quick response, the defense sector is looking for AI technologies capable of successfully processing and extracting results from huge amounts of unlabelled or very poorly labelled data. This paper presents an in-depth review of AI/ML algorithms for unsupervised or poorly supervised data, and machine learning operations (MLOps) techniques that are suitable for the defense industry. The algorithms are divided according to their nature, meaning that they either focus on techniques, or on applications. Techniques can belong to the supervision spectrum, or focus on explainability. Applications are either focused on text processing or computer vision. MLOps techniques, tools and practices are then discussed, revealing approaches and reporting experiences with the objective of declaring how to make the operationalization of ML integrated systems more efficient. Despite many contributions from several researchers and industry, further efforts are required to construct substantially robust and reliable models and supporting infrastructures for AI systems, which are reliable and suitable for the defense sector. This review brings up-to-date information regarding AI algorithms and MLOps that will be helpful for future research in the field. © The Author(s) 2024.	Artificial Intelligence; Machine learning algorithms; Machine learning operations	10.1007/s10462-023-10692-0	Incluido
Akkineni, Arunkumar and Koohborfardhaghighi, Somayeh and Singh, Shailesh	Centrality of AI Quality in MLOPs Lifecycle and Its Impact on the Adoption of AI/ML Solutions	Despite the challenges around incorporating Artificial Intelligence into business processes, AI is revolutionizing the way companies are doing business. The biggest business and social challenge in the adoption of AI solutions is achieving the end users' trust in the models and scaling prototypes to production ready models in an enterprise environment. Scaling trustworthy AI in a more transparent, responsible, and governed manner could facilitate widespread adoption of AI solutions in an enterprise environment. After conducting an extensive literature review on different aspects of AI quality, we have developed an integrated AI Quality-MLOps framework which enables the development and deployment of AI solutions in an enterprise environment. AI Quality is the center of the proposed framework, and it guides businesses towards putting a complete set of quality metrics, tests, approaches, and algorithms together to ensure conformance with business objectives. This approach improves the delivery efficiency of the solution both during the design and production phase while conforming to the regulatory guidelines adopted by an organization. © 2023, The Author(s), under exclusive license to Springer Nature Switzerland AG.	AI Ethics; AI Quality; AI Risk Assessment; Causality; Data Privacy; Explainability; MLOps; Model Security; Reliability	10.1007/978-3-031-35510-3_42	Excluido
Jesus, Rodrigo M. and Brito, Miguel A. and Duarte, Duarte N.	Machine Learning on Insurance Premium Prediction	The insurance field is going through a phase of great transformation due to the growth of new technologies and techniques that are causing a change in the way data is handled and analyzed. The main perpetrator of this phenomenon is the introduction of Machine Learning (ML) in financial decision-making due to their efficiency and productivity. However, there is a new intervenient in the room, which will automate and support all steps of ML system development. Machine Learning Operations (MLOPs) will reduce technical friction, so that the model may move from an idea into production, in the shortest amount of time, and subsequently to market with the least possible risk. In this paper, a detailed review of the impacts of ML on insurance premium forecasting and the influence that MLOPs can have on forecasting outcomes is provided. Furthermore, a comprehensive summary is presented of crucial principles in the insurance industry, which are essential for comprehending the role that MLOPs will play in tailoring and individualizing insurance policies and premiums. © 2023 ACM.	Insurance Premium Prediction; Machine Learning; MLOPs	10.1145/3605423.3605450	Excluido
Kreuzberger, Dominik and Kuhl, Niklas and Hirschl, Sebastian	Machine Learning Operations (MLOps): Overview, Definition, and Architecture	The final goal of all industrial machine learning (ML) projects is to develop ML products and rapidly bring them into production. However, it is highly challenging to automate and operationalize ML products and thus many ML endeavors fail to deliver on their expectations. The paradigm of Machine Learning Operations (MLOps) addresses this issue. MLOps includes several aspects, such as best practices, sets of concepts, and development culture. However, MLOps is still a vague term and its consequences for researchers and professionals are ambiguous. To address this gap, we conduct mixed-method research, including a literature review, a tool review, and expert interviews. As a result of these investigations, we contribute to the body of knowledge by providing an aggregated overview of the necessary principles, components, and roles, as well as the associated architecture and workflows. Furthermore, we provide a comprehensive definition of MLOps and highlight open challenges in the field. Finally, this work provides guidance for ML researchers and practitioners who want to automate and operate their ML products with a designated set of technologies. © 2023 IEEE.	CI/CD; DevOps; machine learning; MLOps; operations; workflow orchestration	10.1109/ACCESS.2023.3262138	Incluido
Vänskä, Sini and Kemell, Kai-Kristian and Mikkonen, Tommi and Abrahamsson, Pekka	Continuous Software Engineering Practices in AI/ML Development Past the Narrow Lens of MLOps: Adoption Challenges	Background: Continuous software engineering practices are currently considered state of the art in software engineering (SE). Recently, this interest in continuous SE has extended to ML system development as well, primarily through MLOps. However, little is known about continuous SE in ML development outside the specific continuous practices present in MLOps. Aim: In this paper, we explored continuous SE in ML development more generally, outside the specific scope of MLOps. We sought to understand what challenges organizations face in adopting all the 13 continuous SE practices identified in existing literature. Method: We conducted a multiple case study of organizations developing ML systems. Data from the cases was collected through thematic interviews. The interview instrument focused on different aspects of continuous SE, as well as the use of relevant tools and methods. Results: We interviewed 8 ML experts from different organizations. Based on the data, we identified various challenges associated with the adoption of continuous SE practices in ML development. Our results are summarized through 7 key findings. Conclusion: The largest challenges we identified seem to stem from communication issues. ML experts seem to continue to work in silos, detached from both the rest of the project and the customers. © 2024 The Authors. Published by Wrocław University of Science and Technology Publishing House.	artificial intelligence; continuous software engineering; continuous star; machine learning; multiple case study	10.37190/e-Inf240102	Incluido
Tripathy, B.K. and Reddy Maddikunta, Praveen Kumar and Pham, Quoc-Viet and Gadekallu, Thippa Reddy and Dev, Kapal and Pandya, Sharnil and Elhalawany, Basem M.	Harris Hawk Optimization: A Survey onVariants and Applications	In this review, we intend to present a complete literature survey on the conception and variants of the recent successful optimization algorithm, Harris Hawk optimizer (HHO), along with an updated set of applications in well-established works. For this purpose, we first present an overview of HHO, including its logic of equations and mathematical model. Next, we focus on reviewing different variants of HHO from the available well-established literature. To provide readers a deep vision and foster the application of the HHO, we review the state-of-the-art improvements of HHO, focusing mainly on fuzzy HHO and a new intuitionistic fuzzy HHO algorithm. We also review the applications of HHO in enhancing machine learning operations and in tackling engineering optimization problems. This survey can cover different aspects of HHO and its future applications to provide a basis for future research in the development of swarm intelligence paths and the use of HHO for real-world problems. © 2022 B. K. Tripathy et al.	applications, future, fuzzy, hho, literature	10.1155/2022/2218594	Excluido

Su, Ningxin and Li, Baochun	MLOps in the Metaverse: Human-Centric Continuous Integration	The metaverse is a virtual world that exists entirely in a computer-generated environment, and it offers a new frontier for machine learning. One of the major challenges for using machine learning in the metaverse is MLOps (Machine Learning Operations), an emerging field that focuses on deploying and managing machine learning models in production. It has been widely acknowledged that machine learning models require a large amount of data to learn and make accurate predictions, and such data is generated progressively in real time as human users interact with the metaverse. Due to the human-centric nature of the metaverse, it goes without saying that, once deployed, models need to be able to adapt to the constantly changing interactive environment and still make accurate predictions. Borrowing a page from software engineering, in this paper, we explore the design space of human-centric continuous integration in metaverse environments, where labeled data samples accumulated with explicit human interactive behavior (e.g., using virtual reality or augmented reality headsets) are used for fine-tuning a deployed deep learning model over a sustained period of time. We propose SPIN, a new mechanism that efficiently utilizes data samples collected from a large number of participating human users over time to fine-tune a deployed model that is shared across all the users. In an extensive array of experimental results using image classification and state-of-the-art YOLOv8 object detection models as case studies, we show that SPIN outperforms FedBuff, a state-of-the-art asynchronous FL mechanism from conventional federated learning, by a substantial margin. © 2023 IEEE.	Continuous integration; federated learning; metaverse; MLOps	10.1109/ISAC.2023.3345385	Incluido
Lima, Anderson and Monteiro, Luciano and Furtado, Ana Paula	MLOps: Practices, Maturity Models, Roles, Tools, and Challenges - A Systematic Literature Review	Context: The development of machine learning solutions has increased significantly due to the advancement of technology based on artificial intelligence. MLOps have emerged as an approach to minimizing efforts and improving integration between those who are in the process of deploying the models in the production environment. Objective: This paper undertakes a systematic literature review in order to identify practices, standards, roles, maturity models, challenges, and tools related to MLOps. Method: The study is founded on an automatic search method of selected digital libraries that applies selection and quality criteria to identify suitable papers that underpin the research. Results: The search initially found 1,905 articles of which 30 papers were selected for analysis. This analysis led to findings that made it possible to achieve the objectives of the research. Conclusion: The results allowed us to conclude that MLOps is still in its initial stage, and to recognize that there is an opportunity to undertake further academic studies that will prompt organizations to adopt MLOps practices. Copyright © 2022 by SCITEPRESS - Science and Technology Publications, Lda. All rights reserved.	Continuous Delivery; Machine Learning; MLOps	10.5220/0010997300003179	Excluido
Fujii, Tiago Yukio and Hayashi, Victor Takashi and Arakaki, Reginaldo and Ruggiero, Wilson Vicente and Bulla, Romeo and Hayashi, Fabio Hirotsugu and Khalil, Khalil Ahmad	A Digital Twin Architecture Model Applied with MLOps Techniques to Improve Short-Term Energy Consumption Prediction	Using extensive databases and known algorithms to predict short-term energy consumption comprises most computational solutions based on artificial intelligence today. State-of-the-art approaches validate their prediction models in offline environments that disregard automation, quality monitoring, and retraining challenges present in online scenarios. The existing demand response initiatives lack personalization, thus not engaging consumers. Obtaining specific and valuable recommendations is difficult for most digital platforms due to their solution pattern: Extensive database, specialized algorithms, and using profiles with similar aspects. The challenges and present personalization tactics have been researched by adopting a digital twin model. This study creates a different approach by adding structural topology to build a new category of recommendation platform using the digital twin model with real-time data collected by IoT sensors to improve machine learning methods. A residential study case with 31 IoT smart meter and smart plug devices with 19-month data (measurements performed each second) validated Digital Twin MLOps architecture for personalized demand response suggestions based on online short-term energy consumption prediction. © 2021 by the authors. Licensee MDPI, Basel, Switzerland.	Digital twin; IoT; Machine learning; MLOps; Prediction	10.3390/machines10010023	Incluido
Subramanya, Rakshith and Sierla, Seppo and Vyatkin, Valeriy	From DevOps to MLOps: Overview and Application to Electricity Market Forecasting	In the Software Development Life Cycle (SDLC), Development and Operations (DevOps) has been proven to deliver reliable, scalable software within a shorter time. Due to the explosion of Machine Learning (ML) applications, the term Machine Learning Operations (MLOps) has gained significant interest among ML practitioners. This paper explains the DevOps and MLOps processes relevant to the implementation of MLOps. The contribution of this paper towards the MLOps framework is threefold: First, we review the state of the art in MLOps by analyzing the related work in MLOps. Second, we present an overview of the leading DevOps principles relevant to MLOps. Third, we derive an MLOps framework from the MLOps theory and apply it to a time-series forecasting application in the hourly day-ahead electricity market. The paper concludes with how MLOps could be generalized and applied to two more use cases with minor changes. © 2022 by the authors.	continuous software engineering; DevOps; electricity market; Machine Learning; MLOps; time-series analysis	10.3390/app12199851	Incluido
Weiher, Karsten and Rieck, Sebastian and Pankrath, Hannes and Beuss, Florian and Geist, Michael and Sender, Jan and Ruegge, Wilko	Automated visual inspection of manufactured parts using deep convolutional neural networks and transfer learning	Most manufacturing processes involve some form of visual quality control of the produced parts. Automated solutions can reduce the required manual work significantly while increasing reliability. However, common obstacles to the construction of smart visual inspection systems are the complexity of the inspected parts, varying types of defects, and small datasets. In this study, we apply state-of-the-art convolutional neural networks to classify infrared images of thermal conductive components manufactured in a real factory setting. Typically, training deep neural architectures requires very large datasets, but this effect is mitigated by using transfer learning. The dataset consists of 6,000 images with 4,200 defect samples and 1,800 intact samples, including different types of flaws and component models. We present a concept for implementing the automated visual inspection system, including dataset preparation, model training, and the inline application. The goal is to establish a Human-in-the-Loop approach, that maximizes accuracy and safety while keeping the required human work at a minimum. A key finding of our research is that dataset preparation and cleaning had a greater impact on the classification accuracy than the optimal choice of the model or training parameters. © 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0) Peer-review under responsibility of the scientific committee of the 56th CIRP International Conference on Manufacturing Systems 2023.	Deep Learning; Human-in-the-Loop (HITL); Machine Learning Operations (MLOps); Quality Control; Transfer Learning	10.1016/j.procir.2023.09.088	Excluido
Wasielewska-Michniewska, Katarzyna and Ganzha, Maria and Paprzycki, Marcin and Pawłowski, Wiesław	Review and Critical Analysis of Ontologies for Artificial Intelligence Systems	With the rising popularity of artificial intelligence-based solutions, it is becoming important not only to deploy machine learning models/pipelines with a good accuracy, but also to be able to control and manage their documentation and information related to monitoring, performance tracking, etc. Moreover, crucial aspects of data that is to be used by said applications, needs to be captured and represented in an organized and standardized way. These include, among others, provenance, access restrictions, usage limitations, format. Being able to (pseudo-)formally represent (and communicate) properties of data and applications should enhance efficient MLOps and facilitate governance of artificial intelligence-based systems. This is of particular importance for multi-stakeholder ecosystems, such as Internet of Things deployments, in which collaboration is required. The question thus arises, how such meta-level description of artificial intelligence system components and data that they consume, process and produce, can be realized. In this context, we investigate state-of-the-art of related ontologies, as well as several non-ontological, but important, approaches to represent knowledge about artificial intelligence models/components, related processes and data, to provide transparency into models development and system performance. We analyze for which elements of such overall description there are ontologies for potential reuse and where there are gaps that need to be filled to capture the ontological representation of all pertinent aspects of artificial intelligence-based system. © The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024.	AI/ML pipelines; edge-cloud continuum; IoT; Ontologies	10.1007/978-981-97-0855-0_70	Excluido
Kirchhof, Jörg Christian and Kusmenko, Evgeny and Ritz, Jonas and Rumpe, Bernhard and Mohn, Armin and Badii, Atta and Günnemann, Stephan and Challenger, Moharram	MDE for Machine Learning-Enabled Software Systems: A Case Study and Comparison of MontiAnna and ML-Quadrat	In this paper, we propose to adopt the MDE paradigm for the development of Machine Learning (ML)-enabled software systems with a focus on the Internet of Things (IoT) domain. We illustrate how two state-of-The-Art open-source modeling tools, namely MontiAnna and ML-Quadrat can be used for this purpose as demonstrated through a case study. The case study illustrates using ML, in particular deep Artificial Neural Networks (ANNs), for automated image recognition of handwritten digits using the MNIST reference dataset, and integrating the machine learning components into an IoT-system. Subsequently, we conduct a functional comparison of the two frameworks, setting out an analysis base to include a broad range of design considerations, such as the problem domain, methods for the ML integration into larger systems, and supported ML methods, as well as topics of recent intense interest to the ML community, such as AutoML and MLOps. Accordingly, this paper is focused on elucidating the potential of the MDE approach in the ML domain. This supports the ML-engineer in developing the (ML/software) model rather than implementing the code, and additionally enforces reusability and modularity of the design through enabling the out-of-The-box integration of ML functionality as a component of the IoT or cyber-physical systems. © 2022 ACM.	artificial intelligence; domain specific modeling; machine learning; model-driven engineering; tools	10.1145/3550356.3561576	Excluido

Rauschmayr, Nathalie and Kama, Sami and Kim, Muhyun and Choi, Miyoung and Kenthapadi, Krishnaram	Profiling Deep Learning Workloads at Scale using Amazon SageMaker	With the rise of deep learning (DL), machine learning (ML) has become compute and data intensive, typically requiring multi-node multi-GPU clusters. As state-of-the-art models grow in size in the order of trillions of parameters, their computational complexity and cost also increase rapidly. Since 2012, the cost of deep learning doubled roughly every quarter, and this trend is likely to continue. ML practitioners have to cope with common challenges of efficient resource utilization when training such large models. In this paper, we propose a new profiling tool that cross-correlates relevant system utilization metrics and framework operations. The tool supports profiling DL models at scale, identifies performance bottlenecks, and provides insights with recommendations. We deployed the profiling functionality as an add-on to Amazon SageMaker Debugger, a fully-managed service that leverages an on-the-fly analysis system (called rules) to automatically identify complex issues in DL training jobs. By presenting deployment results and customer case studies, we show that it enables users to identify and fix issues caused by inefficient hardware resource usage, thereby reducing training time and cost. © 2022 Owner/Author.	Amazon SageMaker; mlops; performance profiling	10.1145/3534678.3539036	Incluido
Jain, Samridhi and Kumar, Puneet	Cost Effective Generic Machine Learning Operation : A Case Study	In this research, we have proposed a mechanism to implement a typical Mops pipeline for small scale organization who cannot afford the operational expenditures to bring the pipeline at Cloudera, Horton works platform or cloud premises like AWS, GCP or Azure. This paper gives a very detailed understanding of operationalization of a typical ML pipelines to adhere all the elements and artifacts without even using any Docker, Kubernetes or even any API generating platforms like Flask or FastAPI. Using the combination of a simple Python/R along with SQL and Shell scripts we can manage the entire workflow at on premises with a very low cost approach. From some angle this mechanism would not be comparable with the architectures like market ready MLOps platforms like Azure Devops, MLflow, Kubeflow, Apache Airflow, Databricks with Data factory or Sagemaker Studio workflow but from conceptual point of view, suffice almost 90% of the requirements with efficient manner. We have also done a latest review related to MLOps in recent past and also listed out the several research gaps that can be solved in future research. © 2023 IEEE.	cloud platform; Low Budget Architecture; ML Engineering; ML pipeline; MLOps	10.1109/ICDSNS58469.2023.10245408	Incluido
Conference Proceedings	Proceedings - Workshop on Software Engineering for Responsible AI, SE4RAI 2022	The proceedings contain 8 papers. The topics discussed include: operationalizing machine learning models - a systematic literature review; robustness testing of a machine learning-based road object detection system: an industrial case; towards trusting the ethical evolution of autonomous dynamic ecosystems; the concept of ethical digital identities; challenges in machine learning application development: an industrial experience report; non-functional requirements for machine learning: an exploration of system scope and interest; augur: a step towards realistic drift detection in production ML systems; and MLOps: a guide to its adoption in the context of responsible AI.	detection, ethical, industrial, learning, machine	N/A	Excluido
Khalid, Nur Hidayah Mohd and Ismail, Amelia Ritahani and Aziz, Normaziah A.	Interpretation of Machine Learning Model Using Medical Record Visual Analytics	The state of the art of medical application that being implemented are mostly based on common machine learning model. Nevertheless, one of the drawbacks of the practice of medical diagnosis is the lack of explanation on the proposed solution, which is also known as a black box, without knowing the internal decision process between the input and output. It will lead to untrustworthiness and difficult to understand by the medical expert. They are questioning how the complexity of machine learning methods decide on the output without clear and understandable explanations. Moreover, in machine learning field the characteristic of a black box model may lead to biased data analysis and incorrect output decisions. There is work that uses visual analytics techniques to interpret the machine learning output to ease the understanding of medical experts. However, the functionality of existed and combined visual analytics techniques is not sufficient to visualized and interpreted the output of machine learning operation. Other visual analytic techniques faced the same problem, unreliability to produce strong reason on the output when working with complex machine learning models. This paper analyzed several visual analytics approaches instantiated in machine learning algorithm for medical record analytics. The motivation of this paper is to allow medical experts to understand the interpretation of a black box machine learning model in predicting medical outcome. This paper studied on the effectiveness of visual analytics techniques to identify the appropriate technique to be instantiated to the machine learning algorithm to further elaborate the results obtained by demonstrating transparency, interpretability and explainability of the machine learning algorithm. The visual analytics that are been studied are Local Interpretable Model-agnostic Explanations (LIME) and Shapley Additive exPlanations (SHAP). Based on the comparison of LIME and SHAP methods, this paper found that SHAP has consistent interpretability as compared to LIME. © 2022, The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd.	Interpretability; Machine learning; Visual analytics	10.1007/978-981-16-8515-6_48	Excluido
Greco, Salvatore and Cerquitelli, Tania	Drift Lens : Real-time unsupervised Concept Drift detection by evaluating per-label embedding distributions	Despite the significant improvements made by deep learning models, their adoption in real-world dynamic applications is still limited. Concept drift is among the open issues preventing the widespread exploitation of deep learning models in real-life settings. The dynamic world changes very quickly, and the collected data drifts accordingly. Prediction models, usually trained on static historical data, should be promptly re-trained in case of new real-time drifted data distributions. Although some drift detection methodologies have been proposed over the years, different issues are still open since state-of-the-art solutions show limited effectiveness and efficiency. This paper proposes Drift Lens, a novel real-time unsupervised per-label drift detection methodology based on embedding distribution distances in deep learning models. The preliminary experiments performed on a transformer-based model fine-tuned for topic text classification show promising results in drift detection accuracy, drift characterization, and efficient execution time to support real-time concept drift detection. © 2021 IEEE.	Data drift; deep learning; MLOps; text topic classification; transformer-based models	10.1109/ICDMW53433.2021.00049	Excluido
Testi, Matteo and Ballabio, Matteo and Frontoni, Emanuele and Iannello, Giulio and Moccia, Sara and Soda, Paolo and Vessio, Gennaro	MLOps: A Taxonomy and a Methodology	Over the past few decades, the substantial growth in enterprise data availability and the advancements in Artificial Intelligence (AI) have allowed companies to solve real-world problems using Machine Learning (ML). ML Operations (MLOps) represents an effective strategy for bringing ML models from academic resources to useful tools for solving problems in the corporate world. The current literature on MLOps is still mostly disconnected and sporadic. In this work, we review the existing scientific literature and we propose a taxonomy for clustering research papers on MLOps. In addition, we present methodologies and operations aimed at defining an ML pipeline to simplify the release of ML applications in the industry. The pipeline is based on ten steps: business problem understanding, data acquisition, ML methodology, ML training & testing, continuous integration, continuous delivery, continuous training, continuous monitoring, explainability, and sustainability. The scientific and business interest and the impact of MLOps have grown significantly over the past years: the definition of a clear and standardized methodology for conducting MLOps projects is the main contribution of this paper. © 2022 IEEE.	continuous delivery; continuous integration; continuous monitoring; continuous training; MLOps; sustainability; XAI	10.1109/ACCESS.2022.3181730	Incluido
Ali, Altiaz and Pasha, Muhammad Fermi and Guerrieri, Antonio and Guzzo, Antonella and Sun, Xiaobing and Saeed, Aamir and Hussain, Amir and Fortino, Giancarlo	A Novel Homomorphic Encryption and Consortium Blockchain-Based Hybrid Deep Learning Model for Industrial Internet of Medical Things	Securing Electronic Medical Records (EMRs) is one of the most critical applications of cryptography over the Internet due to the value and importance of data contained in such EMRs. Although blockchain-based healthcare systems can provide security, privacy, and immutability to EMRs, several outstanding security and latency issues are associated with existing schemes. For example, some researchers have used the blockchain as a storage tool which increases delay and adversely affects the blockchain performance since it stores a copy of each transaction. A distributed ledger also requires appropriate space and computational power with increased data size. In addition, existing healthcare-based approaches usually rely on centralized servers connected to clouds, which are vulnerable to denial of service (DoS), distributed DoS (DDoS), and collusion attacks. This paper proposes a novel hybrid deep learning-based homomorphic encryption (HE) model for the Industrial Internet of Medical Things (IIoMT) to cope with such challenges using a consortium blockchain. Integrating HE with the proposed IIoMT system is a vital contribution of this work. The use of HE while outsourcing to the cloud the storage provides a unique facility to perform any statistical and machine learning operation on the encrypted EMR data, hence providing resistance to collusion and phishing attacks. Our proposed model uses a pre-trained hybrid deep learning model in the cloud and deploys the trained model into blockchain-based edge devices in order to classify and train local models using EMRs. This is further conditioned on the private data of each edge and IoT device connected with the consortium blockchain. All local models obtained are aggregated to the cloud to update a global model, which is finally disseminated to the edge nodes. Our proposed approach provides more privacy and security than conventional models and can deliver high efficiency and low end-to-end latency for users. Comparative simulation analysis with state-of-the-art approaches is carried out using benchmark performance metrics, which show that our proposed model provides enhanced security, efficiency, and transparency. © 2013 IEEE.	blockchain; electronic medical records; homomorphic encryption; hybrid deep learning; industrial internet of medical things; Machine learning; privacy; security	10.1109/TNSE.2023.3285070	Excluido

Conference Proceedings	30th International Working Conference on Requirements Engineering: Foundation for Software Quality, REFSQ 2024	The proceedings contain 22 papers. The special focus in this conference is on Requirements Engineering: Foundation for Software Quality. The topics include: Assessing the Understandability and Acceptance of Attack Defense Trees for Modelling Security Requirements; learning to Rank Privacy Design Patterns: A Semantic Approach to Meeting Privacy Requirements; a New Usability Inspection Method: Experience-Based Analysis; governance-Focused Classification of Security and Privacy Requirements from Obligations in Software Engineering Contracts; what Impact Do My Preferences Have?; Candidate Solutions for Defining Explainability Requirements of AI Systems; Opportunities and Limitations of AI in Human Centered Design a Research Preview; A Tertiary Study on AI for Requirements Engineering; Exploring LLMs' Ability to Detect Variability in Requirements; Designing NLP Based Solutions for Requirements Variability Management: Experiences from a Design Science Study at Visma; Natural2CTL: A Dataset for Natural Language Requirements and Their CTL Formal Equivalents; Towards a Comprehensive Ontology for Requirements Engineering for AI-Powered Systems; Operationalizing Machine Learning Using Requirements-Grounded MLOps; unveiling Competition Dynamics in Mobile App Markets Through User Reviews; exploring the Automatic Classification of Usage Information in Feedback; channelling the Voice of the Crowd: Applying Structured Queries in User Feedback Collection; requirements Information in Backlog Items: Content Analysis; Requirements Engineering for No-Code Development (RE4NCD); behavior Driven Specification in Practice: An Experience Report; the Return of Formal Requirements Engineering in the Era of Large Language Models.	ai, classification, design, engineering, experience	N/A	Excluido
Rocha, Atslands and Monteiro, Matheus and Maitos, César and Dias, Madson and Soares, Jorge and Magalhães, Regis and Macedo, José	Edge AI for Internet of Medical Things: A literature review	The Internet of Things (IoT) consists of heterogeneous devices such as wearables and monitoring devices that collect data to provide autonomous decision-making and smart applications. IoT technologies, such as the Internet of Medical Things (IoMT), have become gradually popular for medical purposes, combining IoT and medical devices to achieve good health and well-being. However, IoMT devices are often tight and have resource constraints, which leads to limited local data processing in the device. Edge computing provides access to additional computation and storage resources for IoMT devices, bringing intelligent processing closer to the data sources. This technology opens up great possibilities for IoMT applications, especially when combined with Artificial Intelligence (AI). Edge AI runs AI computations close to the IoT devices and users instead of centralized services such as cloud servers. This paper investigates the potential of Edge AI and IoMT. In this sense, this survey is the first work to further detail Edge AI and Machine Learning Operations in IoMT domains and wearable technology, thus contributing to the literature by comprehensively exploring the potential of ML strategies and operations at the network's edge and intelligence distribution. This study also presents a case study on heart anomaly detection. © 2024 Elsevier Ltd	Artificial intelligence; Edge computing; Edge intelligence; Internet of medical things; Machine learning operation; Smart health	10.1016/j.compeleceng.2024.109202	Excluido
Conference Proceedings	Proceedings - 2023 30th Asia-Pacific Software Engineering Conference, APSEC 2023	The proceedings contain 90 papers. The topics discussed include: exploring loss scenarios of STPA with reinforcement learning: a case study of platform screen door; sentiment analysis for requirements elicitation from app reviews: a systematic mapping study; understanding developers' contribution motivation in stack overflow: a systematic review; on the impact of development frameworks on mobile apps; investigating multi-and many-objective search for stability-aware configuration of an autonomous delivery system; detecting bot on GitHub leveraging transformer-based models: a preliminary study; exploring trade-offs in MLOps adoption; an empirical study on regression techniques for software defect number prediction; architectural design decisions for data communication of cyber-physical systems; and interpreting CodeBERT for semantic code clone detection.	exploring, offs, overflow	N/A	Excluido
Verbiest, Joeri R. and Bonnechère, Bruno and Saeys, Wim and Van de Walle, Patricia and Truijen, Steven and Meyns, Pieter	Gait Stride Length Estimation Using Embedded Machine Learning	Introduction. Spatiotemporal gait parameters, e.g., gait stride length, are measurements that are classically derived from instrumented gait analysis. Today, different solutions are available for gait assessment outside the laboratory, specifically for spatiotemporal gait parameters. Such solutions are wearable devices that comprise an inertial measurement unit (IMU) sensor and a microcontroller (MCU). However, these existing wearable devices are resource-constrained. They contain a processing unit with limited processing and memory capabilities which limit the use of machine learning to estimate spatiotemporal gait parameters directly on the device. The solution for this limitation is embedded machine learning or tiny machine learning (tinyML). This study aims to create a machine-learning model for gait stride length estimation deployable on a microcontroller. Materials and Method. Starting from a dataset consisting of 4467 gait strides from 15 healthy people, measured by IMU sensor, and using state-of-the-art machine learning frameworks and machine learning operations (MLOps) tools, a multilayer 1D convolutional float32 and int8 model for gait stride length estimation was developed. Results. The developed float32 model demonstrated a mean accuracy and precision of 0.23 ± 4.3 cm, and the int8 model demonstrated a mean accuracy and precision of 0.07 ± 4.3 cm. The memory usage for the float32 model was 284.5 kB flash and 31.9 kB RAM. The int8 model memory usage was 91.6 kB flash and 13.6 kB RAM. Both models were able to be deployed on a Cortex-M4F 64 MHz microcontroller with 1 MB flash memory and 256 kB RAM. Conclusions. This study shows that estimating gait stride length directly on a microcontroller is feasible and demonstrates the potential of embedded machine learning, or tinyML, in designing wearable sensor devices for gait analysis. © 2023 by the authors.	embedded machine learning; gait analysis; gait stride length; healthcare; IMU; inertial measurement unit; machine learning; MCU; microcontroller; neural network; regression; tinyML; wearable sensors	10.3390/s23167166	Excluido
Das, Saikat and Shiva, Sajjan	Machine Learning application lifecycle augmented with explanation and security	We have developed a Distributed Denial of Service (DDoS) intrusion detection framework that employs ML ensembles of both supervised and unsupervised classifiers that are complementary in reaching a corroborated classification decision. Our work has been limited to DDoS attack detection techniques. We propose to extend our framework to general ML system development, based on our review of current ML system development life cycles. We also propose to augment the general life cycle model to include security features to enable building security in as the development progresses and bolt security on as flaws are discovered after deployment. Most ML systems today operate in a black box mode, providing users with only the predictions without associated reasoning as to how the predictions are brought about. There is heavy emphasis now to build mechanisms that help the user develop higher confidence in accepting the predictions of ML systems. Such explainability feature of ML model predictions is a must for critical systems. We also propose to augment our lifecycle model with explainability features. Thus, our ultimate goal is to develop a generic ML lifecycle process augmented with security and explainability features. Such an ML lifecycle process will be of immense use in ML systems development for all domains. © 2021 IEEE.	Explanation augmented ML life cycle; Machine learning life cycle; MLOps; Security augmented ML life cycle	10.1109/UEMCON53757.2021.9666619	Excluido
Kiranyaz, Serkan and Avci, Onur and Abdeljaber, Osama and Ince, Turker and Gabbouj, Moncef and Inman, Daniel J.	1D convolutional neural networks and applications: A survey	During the last decade, Convolutional Neural Networks (CNNs) have become the de facto standard for various Computer Vision and Machine Learning operations. CNNs are feed-forward Artificial Neural Networks (ANNs) with alternating convolutional and subsampling layers. Deep 2D CNNs with many hidden layers and millions of parameters have the ability to learn complex objects and patterns providing that they can be trained on a massive size visual database with ground truth labels. With a proper training, this unique ability makes them the primary tool for various engineering applications for 2D signals such as images and video frames. Yet, this may not be a viable option in numerous applications over 1D signals especially when the training data is scarce or application specific. To address this issue, 1D CNNs have recently been proposed and immediately achieved the state-of-the-art performance levels in several applications such as personalized biomedical data classification and early diagnosis, structural health monitoring, anomaly detection and identification in power electronics and electrical motor fault detection. Another major advantage is that a real-time and low-cost hardware implementation is feasible due to the simple and compact configuration of 1D CNNs that perform only 1D convolutions (scalar multiplications and additions). This paper presents a comprehensive review of the general architecture and principals of 1D CNNs along with their major engineering applications, especially focused on the recent progress in this field. Their state-of-the-art performance is highlighted concluding with their unique properties. The benchmark datasets and the principal 1D CNN software used in those applications are also publicly shared in a dedicated website. While there has not been a paper on the review of 1D CNNs and its applications in the literature, this paper fulfills this gap. © 2020 The Author(s)	Arrhythmia detection and identification; Artificial Neural Networks; Condition monitoring; Convolutional neural networks; Deep learning; Fault detection; Machine learning; Structural damage detection; Structural health monitoring	10.1016/j.ymssp.2020.107398	Excluido
Steffen, Bruno and Zohren, Jonas and Pazarci, Utku and Kullmann, Fiona and Weißenfels, Hendrik	Continuous Integration of Neural Networks in Autonomous Systems	The perception of the autonomous driving software of the F5223, a low-level sensor fusion of Lidar and Camera data requires the use of a neural network for image classification. To keep the neural network up to date with updates in the training data, we introduce a Continuous Integration (CI) pipeline to re-train the network. The network is then automatically validated and integrated into the code base of the autonomous system. The introduction of proper CI methods in these high-speed embedded software applications is an application of state-of-the-art MLOps techniques that aim to provide rapid generation of production-ready models. It further serves the purpose of professionalizing the otherwise script-based software production, which is re-done almost completely every year as the teams change from one year to the next. © 2024, The Author(s), under exclusive license to Springer Nature Switzerland AG.	Continuous Integration; ML Ops; Neural Networks	10.1007/978-3-031-49252-5_21	Excluido

Mukhachev, Petr	Architecture of Modern DataOps Pipeline for Spacecraft Health Monitoring	Many publications show that it is possible to apply data driven techniques and specifically machine learning in aerospace for control, health monitoring, data compression and other applications. Despite these sucess stories, many developed models still do not make their way into production. DataOps an MLOps pratices were recently introduced in order to structure the data processing and eventually automate many routine processes. In this work we review typical MLOps and DataOps processes and provide several reference architectures for teams of different sizes in order to help them ensure model quality and smooth way of their models into production. Copyright © 2023 by the authors. All rights reserved.	architecture; machine learning operations; pipeline; workflow	N/A	Excluido
Brik, Boutiane and Boutiba, Karim and Ksentini, Adlen	Deep Learning for B5G Open Radio Access Network: Evolution, Survey, Case Studies, and Challenges	Open Radio Access Network (O-RAN) alliance was recently launched to devise a new RAN architecture featuring open, software-driven, virtual, and intelligent radio access architecture. O-RAN architecture is based on (1) disaggregated RAN functions that run as Virtual Network Function (VNF) and Physical Network Function (PNF); (2) the notion of RAN controller that runs centrally RAN applications such as mobility management, users' scheduling, radio resources allocation, etc. The RAN controller is in charge of enforcing the application decisions by using open interfaces with the RAN functions. One important feature introduced by O-RAN is the heavy usage of Machine Learning (ML) techniques, particularly Deep Learning (DL), to foster innovation and ease the deployment of intelligent RAN applications that are able to fulfill the Quality of Service (QoS) requirements of the envisioned 5G and beyond network services. In this work, we first give an overview of the evolution of RAN architectures toward 5G and beyond, namely C-RAN, vRAN, and O-RAN. We also compare them based on various perspectives, such as edge support, virtualization, control and management, energy consumption, and AI support. Then, we review existing DL-based solutions addressing the RAN part. We also show how they can be integrated/mapped to the O-RAN architecture since these works were not initially adapted to the O-RAN architecture. In addition, we present two case studies for DL techniques deployment in O-RAN. Furthermore, we describe how the main steps of deployed DL models in O-RAN can be automated, to ensure stable performance of these models, introducing ML system operations (MLOps) concept in O-RAN. Finally, we identify key technical challenges, open issues, and future research directions related to the Artificial Intelligence (AI) enabled O-RAN architecture. © 2020 IEEE.	B5G networks; deep learning; MLOps; open RAN architecture; RAN; RAN intelligent controller	10.1109/OJCOMS.2022.3146618	Excluido
Mitash, Chaitanya and Wen, Bowen and Bekris, Kostas and Boularias, Abdeslam	Scene-level Pose Estimation for Multiple Instances of Densely Packed Objects	This paper introduces key machine learning operations that allow the realization of robust, joint 6D pose estimation of multiple instances of objects either densely packed or in unstructured piles from RGB-D data. The first objective is to learn semantic and instance boundary detectors without manual labeling. An adversarial training framework in conjunction with physics-based simulation is used to achieve detectors that behave similarly in synthetic and real data. Given the stochastic output of such detectors, candidates for object poses are sampled. The second objective is to automatically learn a single score for each pose candidate that represents its quality in terms of explaining the entire scene via a gradient boosted tree. The proposed method uses features derived from surface and boundary alignment between the observed scene and the object model placed at hypothesized poses. Scene-level, multi-instance pose estimation is then achieved by an integer linear programming process that selects hypotheses that maximize the sum of the learned individual scores, while respecting constraints, such as avoiding collisions. To evaluate this method, a dataset of densely packed objects with challenging setups for state-of-the-art approaches is collected. Experiments on this dataset and a public one show that the method significantly outperforms alternatives in terms of 6D pose accuracy while trained only with synthetic datasets. © 2019 CoRL. All Rights Reserved.	Computer Vision; Pose estimation; Synthetic data training	N/A	Excluido
Spjuth, Ola and Frid, Jens and Hellander, Andreas	The machine learning life cycle and the cloud: implications for drug discovery	Introduction: Artificial intelligence (AI) and machine learning (ML) are increasingly used in many aspects of drug discovery. Larger data sizes and methods such as Deep Neural Networks contribute to challenges in data management, the required software stack, and computational infrastructure. There is an increasing need in drug discovery to continuously re-train models and make them available in production environments. Areas covered: This article describes how cloud computing can aid the ML life cycle in drug discovery. The authors discuss opportunities with containerization and scientific workflows and introduce the concept of MLOps and describe how it can facilitate reproducible and robust ML modelling in drug discovery organizations. They also discuss ML on private, sensitive and regulated data. Expert opinion: Cloud computing offers a compelling suite of building blocks to sustain the ML life cycle integrated in iterative drug discovery. Containerization and platforms such as Kubernetes together with scientific workflows can enable reproducible and resilient analysis pipelines, and the elasticity and flexibility of cloud infrastructures enables scalable and efficient access to compute resources. Drug discovery commonly involves working with sensitive or private data, and cloud computing and federated learning can contribute toward enabling collaborative drug discovery within and between organizations. Abbreviations: AI = Artificial Intelligence; DL = Deep Learning; GPU = Graphics Processing Unit; IaaS = Infrastructure as a Service; K8S = Kubernetes; ML = Machine Learning; MLOps = Machine Learning and Operations; PaaS = Platform as a Service; QC = Quality Control; SaaS = Software as a Service. © 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.	Artificial intelligence; drug discovery; life cycle; machine learning	10.1080/17460441.2021.1932812	Incluido
Chen, Yujun Johnny and Gulati, Nikhil and Annamalai, Subbu and Delgoshaie, Amir and Jin, Zhaoyang Larry and Murasdov, Riyad	Cloud-Based Machine Learning System for Large Scale Implementation of Virtual Meters for Production Optimization	This paper proposes a virtual metering (VM) system using machine learning and well production data for estimation of emulsion, produced gas, and vapors for steam-assisted gravity drainage (SAGD) production system. The system consists of four types of virtual meter models for every well: Emulsion Rate VM, Produced Gas Rate VM, Steam Vapor Rate VM, and Flash Vapor Rate VM. Emulsion Rate VM is used to estimate the oil rate output, while the other three VM types are used as the risk level indicators. For optimal well production and well management, operations need to consider both oil production and risks at the same time. Providing both well rate estimation and ability to assess risk at near real time is a challenging task for large production fields, especially when the data is sparse and there is uncertainty in the well data. We propose a comprehensive system that utilizes a scalable framework to develop and deploy data-driven but process-informed virtual meters at scale and at near real time. Specifically, in this paper, we describe a) a robust well data preprocessing pipeline developed to continuously process the stream of input and target data for online learning. b) a state of the art modeling framework designed to combine and segment large amounts of training data from the production filed. c) a scalable and flexible framework to ensure stable model performance with low data latency. d) a streamlined domain-specific model monitoring and MLOps process to enable tuning and retraining of the models by monitoring the performance and data drift. Copyright © 2022, Society of Petroleum Engineers.	data, emulsion, framework, process, production	10.2118/210983-MS	Excluido
Kannout, Eyad and Grodzki, Michal and Grzegorowski, Marek	Considering various aspects of models' quality in the ML pipeline - application in the logistics sector	The industrial machine learning applications today involve developing and deploying MLOps pipelines to ensure the versatile quality of forecasting models over an extended period, simultaneously assuring the models accuracy, stability, short training time, and resilience. In this study, we present the ML pipeline conforming to all the above-mentioned aspects of models' quality formulated as a constrained multi-objective optimization problem. We also provide the reference implementation on state-of-the-art methods for data preprocessing, feature extraction, dimensionality reduction, feature and instance selection, model fitting, and ensemble blending. The experimental study on the real data set from the logistics industry confirmed the qualities of the proposed approach, as the successful participation in an international data competition did. © 2022 Polish Information Processing Society.	Dimensionality reduction; Ensemble blending; Feature extraction; Feature selection; Logistics; MCDA; MOO; XGBoost	10.15439/2022F296	Incluido
Almeida, Bruna and David, João and Campos, Felipe S. and Cabral, Pedro	Satellite-based Machine Learning modelling of Ecosystem Services indicators: A review and meta-analysis	Satellite-based Machine Learning (ML) modelling has emerged as a powerful tool to understand and quantify spatial relationships between landscape dynamics, biophysical variables and natural stocks. Ecosystem Services indicators (ESI) provide qualitative and quantitative information aiding the assessment of ecosystems' status. Through a systematic meta-analysis following the PRISMA guidelines, studies from one decade (2012–2022) were analyzed and synthesized. The results indicated that Random Forest emerged as the most frequently utilized ML algorithm, while Landsat missions stood out as the primary source of Satellite Earth Observation (SEO) data. Nonetheless, authors favoured Sentinel-2 due to its superior spatial, spectral, and temporal resolution. While 30% of the examined studies focused on modelling proxies of climate regulation services, assessments of natural stocks such as biomass, water, food production, and raw materials were also frequently applied. Meta-analysis illustrated the utilization of classification and regression tasks in estimating measurements of ecosystems' extent and conditions and findings underscored the connections between established methods and their replication. This study offers current perspectives on existing satellite-based approaches, contributing to the ongoing efforts to employ ML and artificial intelligence for unveiling the potential of SEO data and technologies in modelling ESI. © 2024 The Authors	Biodiversity; Data fusion; Environmental modelling; MLOps; Natural capital; Remote sensing	10.1016/j.apgeog.2024.103249	Excluido

Chatterjee, Ayan and Ahmed, Bestoun S. and Hallin, Erik and Engman, Anton	Testing of machine learning models with limited samples: an industrial vacuum pumping application	There is often a scarcity of training data for machine learning (ML) classification and regression models in industrial production, especially for time-consuming or sparsely run manufacturing processes. Traditionally, a majority of the limited ground truth data is used for training, while a handful of samples are left for testing. In that case, the number of test samples is inadequate to properly evaluate the robustness of the ML models under test (i.e., the system under test) for classification and regression. Furthermore, the output of these ML models may be inaccurate or even fail if the input data differ from the expected. This is the case for ML models used in the Electroslag Remelting (ESR) process in the refined steel industry to predict the pressure in a vacuum chamber. A vacuum pumping event that occurs once a workday generates a few hundred samples in a year of pumping for training and testing. In the absence of adequate training and test samples, this paper first presents a method to generate a fresh set of augmented samples based on vacuum pumping principles. Based on the generated augmented samples, three test scenarios and one test oracle are presented to assess the robustness of an ML model used for production on an industrial scale. Experiments are conducted with real industrial production data obtained from Uddeholms AB steel company. The evaluations indicate that Ensemble and Neural Network are the most robust when trained on augmented data using the proposed testing strategy. The evaluation also demonstrates the proposed method's effectiveness in checking and improving ML algorithms' robustness in such situations. The work improves software testing's state-of-the-art robustness testing in similar settings. Finally, the paper presents an MLOps implementation of the proposed approach for real-time ML model prediction and action on the edge node and automated continuous delivery of ML software from the cloud. © 2022 Owner/Author.	data augmentation; data decomposition; machine learning; mlops; software testing; vacuum pumping	10.1145/3540250.3558943	Excluido
Kollitveit, Ask Berstad and Li, Jingyue	Operationalizing Machine Learning Models - A Systematic Literature Review	Deploying machine learning (ML) models to production with the same level of rigor and automation as traditional software systems has shown itself to be a non-trivial task, requiring extra care and infrastructure to deal with the additional challenges. Although many studies focus on adapting ML software engineering (SE) approaches and techniques, few studies have summarized the status and challenges of operationalizing ML models. Model operationalization encompasses all steps after model training and evaluation, including packaging the model in a format appropriate for deployment, publishing to a model registry or storage, integrating the model into a broader software system, serving, and monitoring. This study is the first systematic literature review investigating the techniques, tools, and infrastructures to operationalize ML models. After reviewing 24 primary studies, the results show that there are a number of tools for most use cases to operationalize ML models and cloud deployment in particular. The review also revealed several research opportunities, such as dynamic model switching, continuous model monitoring, and efficient edge ML deployments. CCS CONCEPTS • General and reference → Surveys and overviews; • Computing methodologies → Machine learning; • Software and its engineering → Software development techniques. © 2022 ACM.	Deployment; Machine learning; MLOps; Operationalization; Systematic literature review	10.1145/3526073.3527584	Incluido
Díaz-De-Arcaya, Josu and Torre-Bastida, Ana I. and Zárate, Gorka and Miñón, Raúl and Almeida, Aitor	A Joint Study of the Challenges, Opportunities, and Roadmap of MLOps and AIOps: A Systematic Survey	Data science projects represent a greater challenge than software engineering for organizations pursuing their adoption. The diverse stakeholders involved emphasize the need for a collaborative culture in organizations. This article aims to offer joint insights into the role of MLOps and AIOps methodologies for raising the success of data science projects in various fields, ranging from pure research to more traditional industries. We analyze the open issues, opportunities, and future trends organizations face when implementing MLOps and AIOps. Then, the frameworks and architectures that promote these paradigms are presented, as are the different fields in which they are being utilized. This systematic review was conducted using an automated procedure that identified 44,903 records, which were filtered down to 93 studies. These articles are meant to better clarify the problem at hand and highlight the future areas in both research and industry in which MLOPs and AIOps are thriving. Our findings indicate that AIOps flourish in challenging circumstances like those presented by 5G and 6G technologies, whereas MLOps is more prevalent in traditional industrial environments. The use of AIOps in certain stages of the ML lifecycle, such as deployment, remains underrepresented in scientific literature. © 2023 held by the owner/author(s). Publication rights licensed to ACM.	AI; AIOps; architectures; challenges; frameworks; future trends; ML; MLOps; opportunities; SLR; systematic survey	10.1145/3625289	Incluido
Heydari, Mohammad and Rezvani, Zahra	Challenges and Experiences of Iranian Developers with MLOps at Enterprise	Data is becoming more complex, and so are the approaches designed to process it. Enterprises have access to more data than ever, but many still struggle to glean the full potential of insights from what they have. This research explores the challenges and experiences of Iranian developers in implementing the MLOps paradigm within enterprise settings. MLOps, or Machine Learning Operations, is a discipline focused on automating the continuous delivery of machine learning models. In this study, we review the most popular MLOps tools used by leading technology enterprises. Additionally, we present the results of a questionnaire answered by over 110 Iranian Machine Learning experts and Software Developers, shedding light on MLOps tools and the primary obstacles faced. The findings reveal that data quality problems, a lack of resources, and difficulties in model deployment are among the primary challenges faced by practitioners. Collaboration between ML, DevOps, Ops, and Science teams is seen as a pivotal challenge in implementing MLOps effectively. © 2023 IEEE.	Machine Learning Operations; MLOps; MLOps Challenges; MLOps Solutions; MLOps Tools	10.1109/ICAEA60387.2023.10414442	Excluido
Kaoudi, Zoi and Quiané-Ruiz, Jorge-Arnulfo	Unified Data Analytics: State-of-the-art and Open Problems	There is an urgent need for unifying data analytics as more and more application tasks become more complex: Nowadays, it is normal to see tasks performing data preparation, analytical processing, and machine learning operations in a single pipeline. Despite this need, achieving this is still a dreadful process where developers have to get familiar with many data processing platforms and write ad-hoc scripts for integrating them. This tutorial is motivated by this need from both academia and industry. We will discuss the importance of unifying data processing as well as the current efforts to achieve it. In particular, we will introduce a classification of the different cases where an application needs or benefits from data analytics unification and discuss the challenges in each case. Along with this classification, we will also present current efforts known up to date that aim at unifying data processing, such as Apache Beam and Apache Wayang, and emphasize their differences. We will conclude with open problems and their challenges. © 2022, VLDB Endowment. All rights reserved.	analytics, apache, application, current, data	10.14778/3554821.3554898	Excluido
Psaromanolakis, Nikos and Theodorou, Vasileios and Laskaratos, Dimitris and Kalogeropoulos, Ioannis and Vlontzou, Maria-Eleftheria and Zarogianni, Eleni and Samaras, Georgios	MLOps meets Edge Computing: An Edge Platform with Embedded Intelligence towards 6G Systems	The evolution towards more human-centered 6G networks requires the extension of network functionalities with advanced, pervasive automation features. In this direction, cloud-native, softwarized network functions and integration of extreme/far edge devices shall be supported by even more distributed and decomposable systems, such as Edge Cloud environments, while building on AI/ML data-driven mechanisms to improve their performance and resilience for the stringent requirements of next generation applications. In this work, we propose an intelligence-native Edge Management Platform coupled with MLOps functionalities the π-Edge Platform which encompasses automation features for cloud-native lifecycle management of Edge Services. Our introduced architecture incorporates MLOps services and processes, operating as integrated micro-services with the rest of the π-Edge architectural components, ensuring the reliable operation and QoS of Edge network and application services. We experimentally validate our approach with a prototypical implementation of key π-Edge features, including the incorporation of state-of-the-art ML models for performance prediction and anomaly detection, on a multi-media streaming use case based on scenarios from real production environment. © 2023 IEEE.	6G; Artificial Intelligence; Edge Computing; Edge Intelligence; Machine Learning; MLOps; QoS Monitoring	10.1109/EuCNC/6GSummit58263.2023.10188244	Excluido
Recupito, Gilberto and Pecorelli, Fabiano and Catolino, Gemma and Moreschini, Sergio and Nucci, Dario Di and Palomba, Fabio and Tamburri, Damian A.	A Multivocal Literature Review of MLOps Tools and Features	DevOps has become increasingly widespread, with companies employing its methods in different fields. In this context, MLOps automates Machine Learning pipelines by applying DevOps practices. Considering the high number of tools available and the high interest of the practitioners to be supported by tools to automate the steps of Machine Learning pipelines, little is known concerning MLOps tools and their functionalities. To this aim, we conducted a Multivocal Literature Review (MLR) to (i) extract tools that allow for and support the creation of MLOps pipelines and (ii) analyze their main characteristics and features to provide a comprehensive overview of their value. Overall, we investigate the functionalities of 13 MLOps Tools. Our results show that most MLOps Tools support the same features but apply different approaches that can bring different advantages, depending on user requirements. © 2022 IEEE.	DevOps; Machine Learning; MLOps; Multivocal Literature Review	10.1109/SEAA56994.2022.00021	Excluido
Baudzus, Arn and Li, Bin and Jadid, Adnane and Müller, Emmanuel	The Good, The Bad, and The Average: Benchmarking of Reconstruction Based Multivariate Time Series Anomaly Detection	Reconstruction-based algorithms offer state-of-the-art performance in multivariate time series anomaly detection. But as always: there is no single best algorithm. To find the optimal solution, one has to compare different methods and tune their hyperparameters. This paper introduces a lightweight modular benchmarking framework for data scientists and researchers in the field. The framework can be easily set up and automatically create a visual summary of the relevant performance indicators and automatically selected examples to give insight into the behavior of the model and aid during the development. © 2023, The Author(s), under exclusive license to Springer Nature Switzerland AG.	Anomaly Detection; Autoencoder; Benchmark; Experiment tracking; MLOps; Multivariate Time Series; Reconstruction-based Models; Visualisation	10.1007/978-3-031-43430-3_30	Excluido

Hennebelle, Alain and Ismail, Leila and Materwala, Huned and Al Kaabi, Juma and Ranjan, Priya and Janardhanan, Rajiv	Secure and privacy-preserving automated machine learning operations into end-to-end integrated IoT-edge-artificial intelligence-blockchain monitoring system for diabetes mellitus prediction	Diabetes Mellitus, one of the leading causes of death worldwide, has no cure to date and can lead to severe health complications, such as retinopathy, limb amputation, cardiovascular diseases, and neuronal disease, if left untreated. Consequently, it becomes crucial to be able to monitor and predict the incidence of diabetes. Machine learning approaches have been proposed and evaluated in the literature for diabetes prediction. This paper proposes an IoT-edge-Artificial Intelligence (AI)-blockchain system for diabetes prediction based on risk factors. The proposed system is underpinned by blockchain to obtain a cohesive view of the risk factors data from patients across different hospitals and ensure security and privacy of the user's data. We provide a comparative analysis of different medical sensors, devices, and methods to measure and collect the risk factors values in the system. Numerical experiments and comparative analysis were carried out within our proposed system, using the most accurate random forest (RF) model, and the two most used state-of-the-art machine learning approaches, Logistic Regression (LR) and Support Vector Machine (SVM), using three real-life diabetes datasets. The results show that the proposed system predicts diabetes using RF with 4.57% more accuracy on average in comparison with the other models LR and SVM, with 2.87 times more execution time. Data balancing without feature selection does not show significant improvement. When using feature selection, the performance is improved by 1.34% for PIMA Indian and 0.02% for Sylhet datasets, while it is reduced by 0.89% for MIMIC III. © 2023 The Author(s)	Artificial intelligence (AI); Blockchain; Diabetes mellitus type 2; Diagnosis; Digital health; eHealth; Logistic regression (LR); Machine learning; Prognosis; Random forest (RF); Risk factors; Smart connected healthcare; Support vector machine (SVM)	10.1016/j.csbj.2023.11.038	Incluido
Conference Proceedings	ISE 2022 - Proceedings of the 1st International Workshop on Intelligent Software Engineering, co-located with 29th Asia-Pacific Software Engineering Conference, APSEC 2022	The proceedings contain 14 papers. The topics discussed include: MediaPipe-based LSTM-autoencoder sarcopenia anomaly detection and requirements for improving detection accuracy; the application of machine learning on the injury prediction of soccer players; real-time twist rebar detection system exploiting GAN-based data augmentation technique; performance evaluation of fire and smoke detection with object based DNN algorithm; business process model annotation techniques: identification, classification and analysis; automatic testing of OCE, a human-centered reinforcement learning system for automated software composition; towards automated open source assessment - an empirical study; load balancing in cloud computing using nature-inspired algorithms: a systematic literature review; and quality assurance in MLOps setting: an industrial perspective.	automated, based, detection, learning, papers	N/A	Excluido
Conference Proceedings	7th International Conference on Smart City Applications, SCA 2022	The proceedings contain 80 papers. The special focus in this conference is on Smart City Applications. The topics include: Using Practical Activities for Robotics E-Learning: Case of Study on Web-Based Robotics Remote Labs; smart Home and Machine Learning as a Sustainable Healthcare Solution: Review and Perspectives; MLOps: Overview of Current State and Future Directions; cloud Services for Smart Farming: A Case Study of the Veracruz Almond Crops in Portugal; nyon: A Ubiquitous Fall Detection Device for Elders; design of the Content Part of the Information System in Smart Cities from the Perspective of Regional Governments and Security of Residents; security Approaches for Smart Campus; assessing the Implementation of Smart Energy Efficient Street Lighting Projects Within Cities; environmental and Climate Risk Management in Public Procurement: A Proposed Decision Support Tool; marine and Climatic Impacts on the Hydrogeochemical Functioning of Lake Sidi Boughaba (Kenitra, North West, Morocco); Assessment of the Role of Micromobility in ITS by A'WOT Analysis; learning Healthcare Ecosystems for Equity in Health Service Provisioning and Delivery: Smart Cities and the Quintuple Aim; Comparative Study of Optimal Tuning PID Controller for Manipulator Robot; Real-Time Hand Gesture Recognition for Humanoid Robot Control Using Python CVZone; RobVR L@bs: Web-Based Application for Teaching Robotics; implementing Machine Learning-Based Simulation in Physics Virtual Laboratory; ALF - Ambient Assisted Living for Healthcare Framework Based on IoMT and Big Data; Leveraging Moroccan Arabic Sentiment Analysis Using AraBERT and QARiB; Evaluating Dimensionality Reduction Approaches on Erisivhile Hyperion and Newly Launched PRISMA Datasets; transfer Learning for Automated Melanoma Classification System: Data Augmentation; intelligent Multi-sensor Mobile System.	Proceedings	N/A	Excluido
Hardy, Corentin and Le Merrer, Erwan and Sericola, Bruno	Distributed deep learning on edge-devices: Feasibility via adaptive compression	A large portion of data mining and analytic services use modern machine learning techniques, such as deep learning. The state-of-the-art results by deep learning come at the price of an intensive use of computing resources. The leading frameworks (e.g., TensorFlow) are executed on GPUs or on high-end servers in datacenters. On the other end, there is a proliferation of personal devices with possibly free CPU cycles; this can enable services to run in users' homes, embedding machine learning operations. In this paper, we ask the following question: Is distributed deep learning computation on WAN connected devices feasible, in spite of the traffic caused by learning tasks? We show that such a setup rises some important challenges, most notably the ingress traffic that the servers hosting the up-to-date model have to sustain. In order to reduce this stress, we propose AdaComp, a novel algorithm for compressing worker updates to the model on the server. Applicable to stochastic gradient descent based approaches, it combines efficient gradient selection and learning rate modulation. We then experiment and measure the impact of compression, device heterogeneity and reliability on the accuracy of learned models, with an emulator platform that embeds TensorFlow into Linux containers. We report a reduction of the total amount of data sent by workers to the server by two order of magnitude (e.g., 191-fold reduction for a convolutional network on the MNIST dataset), when compared to a standard asynchronous stochastic gradient descent, while preserving model accuracy. © 2017 IEEE.	Machine Learning; Data Mining	10.1109/NCA.2017.8171350	Excluido
Singh, Prerna	Systematic review of data-centric approaches in artificial intelligence and machine learning	Artificial intelligence (AI) relies on data and algorithms. State-of-the-art (SOTA) AI smart algorithms have been developed to improve the performance of AI-oriented structures. However, model-centric approaches are limited by the absence of high-quality data. Data-centric AI is an emerging approach for solving machine learning (ML) problems. It is a collection of various data manipulation techniques that allow ML practitioners to systematically improve the quality of the data used in an ML pipeline. However, data-centric AI approaches are not well documented. Researchers have conducted various experiments without a clear set of guidelines. This survey highlights six major data-centric AI aspects that researchers are already using to intentionally or unintentionally improve the quality of AI systems. These include big data quality assessment, data preprocessing, transfer learning, semi-supervised learning, machine learning operations (MLOps), and the effect of adding more data. In addition, it highlights recent data-centric techniques adopted by ML practitioners. We addressed how adding data might harm datasets and how HoloClean can be used to restore and clean them. Finally, we discuss the causes of technical debt in AI. Technical debt builds up when software design and implementation decisions run into "or outright collide with" business goals and timelines. This survey lays the groundwork for future data-centric AI discussions by summarizing various data-centric approaches. © 2023 Xi'an Jiaotong University	Data management; Data preprocessing; Data-centric; Machine learning; MLOps; Semi-supervised learning; Technical debt	10.1016/j.dsm.2023.06.001	Incluido
Kiefakis, Spyridon and Mavrogiorgou, Argyro and Mavrogiorgos, Konstantinos and Kiourtis, Athanasios and Kyriazis, Dimosthenis	Digital Twin in Healthcare Through the Eyes of the Vitruvian Man	In recent years, worldwide, with the development of technology, a huge amount of data is collected in Electronic Health Records (EHRs). Although vast progress has been made with the use of artificial intelligence in various areas of health domain and for specific problems, it is a fact that to date there is no holistic approach to a patient's state of health using these technologies. Digital Twin refers to a complete physical and functional description of an item, product, or system, which includes pretty much all the information that could be useful in all—current and next—life cycle phases. This paper presents a platform that, using state of the art technologies such as Microservice Architecture (MSA), containerization (Docker), orchestration (Kubernetes) and Machine Learning Operations (MLOps), whereas it is inspired by Leonardo DaVinci's Vitruvian man, building the Digital Twin of Patient platform. To achieve that, the platform's architecture is designed with multiple clusters of Docker containers and Kubernetes orchestration. Specific parts or organs of the human body, are represented by clusters called "digital_twin_components"—DTCs. The set of those DTCs structure the "patient_digital_twin" cluster in which appropriate pipelines define and monitor in real time the "best" possible construction of the patient's digital twin. © 2022, The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd.	Containers; Digital twin; Electronic health records; Healthcare; Machine learning; Microservices; MLOps	10.1007/978-981-19-3440-7_7	Excluido
Das, Anup	Real-Time Scheduling of Machine Learning Operations on Heterogeneous Neuromorphic SoC	Neuromorphic Systems-on-Chip (NSoCs) are becoming heterogeneous by integrating general-purpose processors (GPPs) and neural processing units (NPUs) on the same SoC. For embedded systems, an NSoC may need to execute user applications built using a variety of machine learning models. We propose a real-time scheduler, called PRISM, which can schedule machine learning models on a heterogeneous NSoC either individually or concurrently to improve their system performance. PRISM consists of the following four key steps. First, it constructs an interprocessor communication (IPC) graph of a machine learning model from a mapping and a self-timed schedule. Second, it creates a transaction order for the communication actors and embeds this order into the IPC graph. Third, it schedules the graph on an NSoC by overlapping communication with the computation. Finally, it uses a Hill Climbing heuristic to explore the design space of mapping operations on GPPs and NPUs to improve the performance. Unlike existing schedulers which use only the NPUs of an NSoC, PRISM improves performance by enabling batch, pipeline, and operation parallelism via exploiting a platform's heterogeneity. For use-cases with concurrent applications, PRISM uses a heuristic resource sharing strategy and a non-preemptive scheduling to reduce the expected wait time before concurrent operations can be scheduled on contending resources. Our extensive evaluations with 20 machine learning workloads show that PRISM significantly improves the performance per watt for both individual applications and use-cases when compared to state-of-the-art schedulers. © 2022 IEEE.	hill climbing; interprocessor communication (IPC); neuromorphic system-on-chip (NSoC); self-timed execution; spiking deep convolution neural network (SDCNN)	10.1109/MEMOCODES57689.2022.9954596	Excluido

Barrak, Amine and Petrillo, Fabio and Jaafar, Fehmi	Serverless on Machine Learning: A Systematic Mapping Study	Machine Learning Operations (MLOps) is an approach to managing the entire lifecycle of a machine learning model. It has evolved over the last years and has started attracting many people in research and businesses in the industry. It supports the development of machine learning (ML) pipelines typical in the phases of data collection, data pre-processing, building datasets, model training, hyper-parameters refinement, testing, and deployment to production. This complex pipeline workflow is a tedious process of iterative experimentation. Moreover, cloud computing services provide advanced features for managing ML stages and deploying them efficiently to production. Specifically, serverless computing has been applied in different stages of the machine learning pipeline. However, to the best of our knowledge, it is missing to know the serverless suitability and benefits it can provide to the ML pipeline. In this paper, we provide a systematic mapping study of machine learning systems applied on serverless architecture that include 53 relevant studies. During this study, we focused on (1) exploring the evolution trend and the main venues; (2) determining the researchers' focus and interest in using serverless on machine learning; (3) discussing solutions that serverless computing provides to machine learning. Our results show that serverless usage is growing, and several venues are interested in the topic. In addition, we found that the most widely used serverless provider is AWS Lambda, where the primary application was used in the deployment of the ML model. Additionally, several challenges were explored, such as reducing cost, resource scalability, and reducing latency. We also discuss the potential challenges of adopting ML on serverless, such as respecting service level agreement, the cold start problem, security, and privacy. Finally, our contribution provides foundations for future research and applications that involve machine learning in serverless computing. © 2013 IEEE.	FaaS; function as a service; machine learning; Serverless; SLR; SM; systematic literature review; systematic mapping	10.1109/ACCESS.2022.3206366	Excluido
Haertel, Christian and Staegemann, Daniel and Daase, Christian and Pohl, Matthias and Nahhas, Abdulrahman and Turowski, Klaus	MLOps in Data Science Projects: A Review	Data Science (DS) has gained increased relevance due to the potential to extract useful insights from data. Quite commonly, this involves the utilization of Machine Learning (ML). The challenging pursuit of developing and productionizing ML models can be supported and automated through MLOps, a specialization of the DevOps paradigm from software development. Therefore, MLOps offers significant potential for DS projects, which are suffering from notable failure rates. Accordingly, this literature review focuses on examining the current state-of-the-art of the publications in this area. Most importantly, the analysis showed that the current MLOps approaches in the literature predominantly emphasize model development and deployment, while organizational aspects (business understanding, evaluation) in a DS project are neglected. As DS project success is not exclusively dependent on technical matters, advancing the MLOps field by bridging the gap between business objectives and the modeling perspective through appropriate frameworks should be pursued in future research. © 2023 IEEE.	Data Science; DevOps; Literature Review; MLOps; Project Management	10.1109/BigData59044.2023.10386139	Incluido
Raffin, Tim and Reichenstein, Tobias and Kiier, Dennis and Kühl, Alexander and Franke, Jörg	Qualitative assessment of the impact of manufacturing-specific influences on Machine Learning Operations	Machine Learning Operations (MLOps) enables the streamlining of the development and deployment processes of machine learning models; thus, manufacturers can utilize the inherent flexibility and adaptability of Deep Learning at scale to further optimize their processes. This publication provides insights into the challenges that companies face while striving for the efficient operationalization of machine learning algorithms. Moreover, a mapping of capabilities and requirements is presented to provide a baseline for the qualitative analysis of the current state of the art. In conclusion, this article discusses the shortcomings of the existing literature and provides novel implications for MLOps systems in manufacturing. © 2022 Elsevier B.V.. All rights reserved.	Deep Learning; IIoT; Machine Learning; MLOps	10.1016/j.procir.2022.10.063	Incluido
Conference Proceedings	ICTC 2023 - 14th International Conference on Information and Communication Technology Convergence: Exploring the Frontiers of ICT Innovation	The proceedings contain 478 papers. The topics discussed include: enhancing fine-tuning in low data regime by increasing representation entropy during pre-training phase; QC-LDPC codes from various Golomb rulers; wireless embedded index coded transmissions with multiple antennas; blockchain and eGovernment convergence for the optimization of ODA activities in developing countries: a scoping review; trends and applications of robotics development in virtual environments; a micro service-based MLOps platform for efficient development of ai services in an edge-cloud environment; how technologies like ai affect the profit rate: empirical study for 6 capitalist countries; dual-band single-layer reflectarray antenna for LEO satellite communication; who can be your ai doctor?: evaluation for disease diagnosis on large language models; and mimicking GFSK modulation for Wi-Fi to BLE cross technology communication.	ai, communication, countries, development, phase	N/A	Excluido
Mboweni, Tsakani and Masombuka, Themba and Dongmo, Cyrille	A Systematic Review of Machine Learning DevOps	DevOps and Machine Learning (ML) on their own are insufficient in certain organisational context. To achieve the desired results, industries are now combining the benefits of DevOps automation and machine learning capabilities. This synergy is known as Machine Learning DevOps (ML-DevOps). However, there is currently no shared vision, strategy, or guidance across industry enterprises on how to use ML-DevOps for application development. While still in its infancy, ML-DevOps has showed some prospects to become a dynamic approach for developing and delivering secure, robust, and scalable applications. This article describes a systematic review of peer-reviewed conference papers and journal articles on ML-DevOps from 2015 to 2022 to ascertain the state-of-the-art and then identify any gaps in the current literature. Using empirical analysis, we systematically examined more than 60 studies that primarily deal with ML-DevOps (commonly referred to as MLOps). We combined meta-analysis, document analysis and triangulation to map how ML-DevOps is currently understood and how it compares and differs from related approaches such as DevOps. The articles used were accessed via credible scholarly databases. The results of our search strategy retrieved articles that primarily discuss the topics of ML and DevOps synergy. The review of the existing literature shows a paucity of studies that strictly investigate ML-DevOps/MLOps. In most cases, the focus of these studies is on clarifying the meaning of MLOps. Our study discovered that although there are common elements in the definitions of MLOps, there remains a lack of standardisation on how it is defined, which sometimes cause some misconceptions and ambiguity. Finally, although there is significant literature written on MLOps, the review did not find evidence that there is a common understanding amongst scholars and experts on how MLOps should be implemented and institutionalized across industry to create a common vision. © 2022 IEEE.	DevOps; Machine Learning; ML-DevOps; MLOps	10.1109/ICECET55527.2022.9872968	Incluido
Steidl, Monika and Felderer, Michael and Ramler, Rudolf	The pipeline for the continuous development of artificial intelligence models—Current state of research and practice	Companies struggle to continuously develop and deploy Artificial Intelligence (AI) models to complex production systems due to AI characteristics while assuring quality. To ease the development process, continuous pipelines for AI have become an active research area where consolidated and in-depth analysis regarding the terminology, triggers, tasks, and challenges is required. This paper includes a Multivocal Literature Review (MLR) where we consolidated 151 relevant formal and informal sources. In addition, nine semi-structured interviews with participants from academia and industry verified and extended the obtained information. Based on these sources, this paper provides and compares terminologies for Development and Operations (DevOps) and Continuous Integration (CI)/Continuous Delivery (CD) for AI, Machine Learning Operations (MLOps), (end-to-end) lifecycle management, and Continuous Delivery for Machine Learning (CD4ML). Furthermore, the paper provides an aggregated list of potential triggers for reiterating the pipeline, such as alert systems or schedules. In addition, this work uses a taxonomy creation strategy to present a consolidated pipeline comprising tasks regarding the continuous development of AI. This pipeline consists of four stages: Data Handling, Model Learning, Software Development and System Operations. Moreover, we map challenges regarding pipeline implementation, adaption, and usage for the continuous development of AI to these four stages. © 2023 The Authors	CI/CD for AI; Continuous (end-to-end) lifecycle pipeline for AI; Continuous development of AI; DevOps for AI; MLOps; Multivocal literature review	10.1016/j.jss.2023.111615	Incluido
Li, Peizheng and Thomas, Jonathan and Wang, Xiaoyang and Khalil, Ahmed and Ahmad, Abdelrahim and Inacio, Rui and Kapoor, Shipra and Parekh, Arjun and Doufexi, Angela and Shojaeifard, Arman and Piechocki, Robert J.	RLOps: Development Life-Cycle of Reinforcement Learning Aided Open RAN	Radio access network (RAN) technologies continue to evolve, with Open RAN gaining the most recent momentum. In the O-RAN specifications, the RAN intelligent controllers (RICs) are software-defined orchestration and automation functions for the intelligent management of RAN. This article introduces principles for machine learning (ML), in particular, reinforcement learning (RL) applications in the O-RAN stack. Furthermore, we review the state-of-the-art research in wireless networks and cast it onto the RAN framework and the hierarchy of the O-RAN architecture. We provide a taxonomy for the challenges faced by ML/RL models throughout the development life-cycle: from the system specification to production deployment (data acquisition, model design, testing and management, etc.). To address the challenges, we integrate a set of existing MLOps principles with unique characteristics when RL agents are considered. This paper discusses a systematic model development, testing and validation life-cycle, termed: RLOps. We discuss fundamental parts of RLOps, which include: model specification, development, production environment serving, operations monitoring and safety/security. Based on these principles, we propose the best practices for RLOps to achieve an automated and reproducible model development process. At last, a holistic data analytics platform rooted in the O-RAN deployment is designed and implemented, aiming to embrace and fulfill the aforementioned principles and best practices of RLOps. © 2013 IEEE.	data engineering; digital twins; machine learning; MLOps; O-RAN; reinforcement learning; RLOps	10.1109/ACCESS.2022.3217511	Excluido

Shetty, Shashank and Ananthanarayana, V.S. and Mahale, Ajit	Diagnostic Performance Evaluation of Deep Learning Based Medical Text Modelling to Predict Pulmonary Diseases from Unstructured Radiology Free Text Reports	The third most common cause of death worldwide is attributed to pulmonary diseases, making it imperative to diagnose them promptly. Radiology is a medical discipline that utilizes medical imaging to guide treatment. Radiologists prepare reports interpreting details and findings analysed from medical images. Radiology free-text reports are a rich source of textual information that can be exploited to enhance the efficacy of medical prognosis, treatment and research. Radiology reports exist in an unstructured format as are not suitable by themselves for mathematical computation or machine learning operations. Therefore, natural language processing (NLP) strategies are employed to convert unstructured natural language text into a structured format that can be fed into machine learning (ML) or deep learning (DL) models for information extraction. We propose a DL-based medical text modelling framework incorporating a knowledge base to predict pulmonary diseases from unstructured radiology free text reports. We make detailed diagnostic performance evaluations of our proposed technique by comparing it with state-of-the-art NLP techniques on radiology free-text reports extracted from two medical institutions. The comprehensive analysis shows that the proposed model achieves superior results compared to existing state-of-the-art text modelling techniques. © 2023 Prague University of Economics and Business. All Rights Reserved.	Deep learning; Natural language processing; Radiology reports; Unstructured data	10.18267/j.aip.214	Excluido
Kathole, Akash and Shinde, Sagar and Wadhwa, Lalitkumar	Integrated MLOps and EEG techniques for enhanced crime detection and prevention	The use of machine learning and artificial intelligence in crime detection has gained significant attention in recent years. This research paper explores the potential of MLOps techniques in identifying criminal behavior through the analysis of vehicle plate numbers (Tripathi et al 2021), person detection, and object behavior detection. This paper presents a literature review of studies that have investigated the use of MLOps in crime detection and highlight the potential of these techniques to be used in criminal investigations. Specifically, the use of machine learning models for vehicle plate number recognition, person detection through image and video analysis, and object behavior detection through image and video analysis has been discussed. The paper presents a framework for integrating MLOps techniques into criminal investigations, which involves a combination of data acquisition, data preprocessing, model development, model training and testing, and deployment. Additionally, the discussion includes the ethical implications of using MLOps techniques in criminal investigations and highlights the need for transparency and fairness in model development and deployment. The detection and prevention of criminal behavior is a critical issue for society. In recent years, there has been growing interest in the use of electroencephalogram (EEG) techniques to detect and predict criminal behavior. This research paper explores the potential of EEG techniques as a means of detecting crime by analyzing brainwave activity. Specifically, this paper examines the use of alpha and beta waves in identifying deceptive or abnormal behavior (Xie et al 2022). The paper presents a literature review of studies that have investigated the relationship between EEG signals and criminal behavior and highlight the potential of these techniques to be used in criminal investigations (Xie et al 2022). Furthermore, a framework is proposed for integrating EEG techniques into criminal investigations, which involves a combination of data acquisition, analysis, and interpretation. Finally, the ethical implications of using EEG techniques in criminal investigations and the need for further research in this area have been discussed. © 2024 Multidiscip. Sci. J. All rights reserved.	artificial intelligence; brainwave; criminal; machine learning	10.31893/multiscience.2024009	Excluido
Gürses-Tran, Gonca and Monti, Antonello	Advances in Time Series Forecasting Development for Power Systems' Operation with MLOps	Forecast developers predominantly assess residuals and error statistics when tuning the targeted model's quality. With that, eventual cost or rewards of the underlying business application are typically not considered in the model development phase. The analysis of the power system wholesale market allows us to translate a time series forecast method's quality to its respective business value. For instance, near real-time capacity procurement takes place in the wholesale market, which is subject to complex interrelations of system operators' grid activities and balancing parties' scheduling behavior. Such forecasting tasks can hardly be solved with model-driven approaches because of the large solution space and non-convexity of the optimization problem. Thus, we generate load forecasts by means of a data-driven based forecasting tool ProLoaF, which we benchmark with state-of-the-art baseline models and the auto-machine learning models auto.arima and Facebook Prophet. © 2022 by the authors.	congestion management; grid operation; MLOps; probabilistic load forecasting; uncertainty	10.3390/forecast4020028	Excluido
Kiritz, Nick and Ravitz, Miles and Levonian, Mark	Model risk tiering: An exploration of industry practices and principles	Model risk management continues to be an important area of risk management across financial services firms. Regulatory expectations for model risk management apply to an expanding number and variety of financial firms. In our conversations with clients, we find that many risk managers are unaware of standard practices in developing and implementing frameworks for model risk management. This paper - based on our experience as consultants, former banking supervisors and leaders of units focused on model risk management - seeks to shed light on one critical area of such frameworks: model risk tiering, or the rating of risk inherent in the use of individual models, which can benefit a firm's resource allocation and overall risk management capabilities. Although the range of practice in model risk tiering across financial services is broad and varied, our work has revealed a number of insights. • Most financial firms use explicit procedures and decision tools to assign models to risk tiers. • Nearly all firms assign models to several risk tiers rather than seeking to rank them individually from high to low. • Although most firms use model risk tiering tools, these tools serve primarily to systematize the application of judgment to the risk posed by models rather than to directly quantify risk. • A set of useful principles regarding the construction of such risk tiering tools can be gleaned from observing industry practices. • Tiering tools generally fall into one of two categories - scorecards or decision trees - with scorecard approaches being the more common of the two. • Tiering tools generally take into account both the risk and the impact of model failure; the building blocks for these two factors include a variety of descriptors regarding the construction and use of each model in the inventory. • Calibration of model risk tiering tools is largely judgmental and often depends on whether the firm has a preexisting allocation of models to risk tiers. • Governance over model risk tiering tools is an evolving process with rising standards. © 2019 Infopro Digital Risk (IP) Limited.	Model governance; Model risk rating; Model risk tier; OCC 2011-12; SR 11-7; Supervisory guidance on model risk management	10.21314/JRMV.2019.205	Incluido
Canals-Cerdá, José J.	Can we take the "stress" out of stress testing? Applications of generalized structural equation modeling to consumer finance	Statistical model infrastructures at financial institutions are often developed using a piecemeal approach to model building, in which different components of complex interrelated statistical models are developed and validated separately. We propose an empirical framework, spurred by recent developments in the implementation of generalized structural equation modeling (GSEM), which brings to bear a modular and all-inclusive approach to statistical model building. We illustrate the "game changing" potential of this framework with an application to the stress testing of credit risk for a representative portfolio of mortgages; we also extend it to the analysis of the allowance for credit loss under the novel Current Expected Credit Losses (CECL) accounting regulation. We illustrate how GSEM techniques can significantly enhance every step of the modelling management life cycle, from development and documentation to validation, production and redevelopment. We also illustrate how GSEM can be used to combine various risk management projects and tasks into a single framework; we specifically illustrate how to seamlessly integrate stress testing and CECL (or International Financial Reporting Standard 9) frameworks and champion, and challenger, modelling frameworks. Finally, we identify other areas of model risk management that can benefit from the GSEM framework and highlight other potentially fruitful applications of the methodology. © 2022. Infopro Digital Risk (IP) Limited	allowances; Comprehensive Capital Analysis and Review (CCAR); Current Expected Credit Losses (CECL); Generalized structural equation modeling (GSEM); mortgages; regulatory capital	10.21314/JRMV.2022.017	Excluido
Painman, K. and Bilijon, L.van and Haasbroek, L.J. and Schutte, W.D. and Verster, T.	Quantification of the estimation risk inherent in loss distribution approach models	Financial institutions rely heavily on risk models for operational and strategic purposes. These models are simplified representations of complex dynamics and are therefore prone to contain a quantum of "model risk". Although model risk management practices have materially improved in recent years, model risk quantification approaches, which can assist in further advancing model risk management, require more development. We propose a generic approach to quantifying the estimation risk of risk models using the error of a maximum likelihood estimator. We label our approach "simulate errors from assumed truth", or SEAT. To demonstrate SEAT, operational risk examples using the loss distribution approach (LDA) are examined. The simulation design is chosen to include sparse loss data and heavy tailed severity distributions. Model risk exposure is materially reduced when loss data samples are larger, the severity distribution exhibits light tails and the model is used for less-extreme quantiles of the aggregate loss distribution. In fact, model risk is reduced by up to 75% when the target percentile is lowered from the 99.9th to the 95th percentile for practical heavy tailed distributions. Further, the aggregate model risk across a portfolio of risk models is much smaller than that of the individual risk models – a reduction of 29.5% in some of our examples. Finally, we show that a firm is more likely to overestimate than underestimate capital requirements calculated from extreme quantiles in our example. © 2019 Infopro Digital Risk (IP) Limited	financial risk models; heavy tailed distributions; maximum likelihood estimation (MLE); model risk; parameter estimation uncertainty	10.21314/JRMV.2019.212	Excluido

Liu, Yang	Impact assessment of loss given default (LGD) models' risk on regulatory capital: A bayesian approach	"The model is wrong!" so it is determined. All of the estimated output using the model becomes un reliable immediately. And so is every other result calculated using the unreliable output. So what is the impact of the model being "wrong" in the later calculations? To address this question, this paper presents a Bayesian approach that provides a quantitative assessment for the impact on downstream results calculated using the unreliable estimates. Section 1 details the practical challenge in the financial industry and discusses why this is important. Section 2 starts the discussion with a description of the overall framework for this Bayesian approach, introducing and defining each individual component. Then Sections 3 and 4 carry on discussing the prior and likelihood distributions, respectively. Section 5 then obtains the target posterior distribution by applying the Bayesian posterior update using obtained prior and likelihood results. Then conditioning on value of the unreliable estimate already in place in the portfolio, the density distribution obtained can be used to update the output of the "wrong" model and assess the impact in further calculations. This approach bridges the practitioners' initial expectations with the model performance and provides an intuitive quantitative assessment for the impact in the follow up calculations which are largely affected by the unreliable estimate. The presented approach is the first in literature to raise the concern of uncertain impact caused by "wrong" models and propose a solution. The pioneer demonstration using uncertainty in the loss given default (LGD) models as an example and assessing the impact on the then calculated regulatory capital provides a timely assessment tool for model risk management in the current banking industry. Note that the abuse of the word wrong in quotation marks is an exaggeration of the uncertainty involved, in practice, impact analysis could be requested at any level of uncertainty. © 2018 WIT Press.	Bayesian LGD estimate; Best effort loss estimate; Expected loss ratio; Impact analysis; Loss given default (LGD); Model risk management; Observed loss rate; Post observation; Risk weighted assets	10.2495/RISK180101	Excluido
Garro, Maurizio	The evolution of model risk management processes	Model risk is a relatively new area that only recently separated from operational risk. This paper provides an overview of what needs to be done in model risk management versus the older concept of model validation. The works published by the Federal Reserve and later by the European Baking Authority and Bank of England have helped to establish the foundations for a model risk framework. The real challenge for model risk is to evolve from a pure cost to a value added function for financial institutions. This journey requires the application of many elements, including a consistent model definition and performance monitoring of model risk in line with the risk appetite and business model of financial institutions. The final step is to identify a coherent and clear way to aggregate and report model risk to allow senior management to fully understand it in detail and integrate it in the broader risk management strategy. © Henry Stewart Publications 1752-8887 (2020).	Bank of England; Federal Reserve; Financial crisis; Model performance monitoring; Model risk; Model risk aggregation and reporting	N/A	Incluido
Kim, Kyoung-Sup and Woo, Joong-Il and Kim, Jung-Eun and Park, Dong-Soo	Implementation of MCA rule mapper for cloud computing environments	The client-server system in the old finance and banking circle was operated in such a way as to respond to a relatively small number of devices, such as ATMs (automated teller machines) and personal computers. Due to the significant increase in the use of smart devices at the present time, however, the number of devices that need to be connected to financial transaction servers continually skyrockets. In this paper, a mapper that can analyze the information of the messages sent from newly added terminals and then link to the MCA was designed, which would allow new terminals to be added without correcting the programs. In addition, as myriads of terminals need to be added, the programs cannot be corrected on a case-to-case basis; in addition, it is actually impossible to correct all the programs. Thus, the mapper, an automatic mapping system, was developed to add terminals automatically. As a result, the mapper is likely to minimize the labor and money that need to be invested for mapping works and to establish a definitive standard in using various mapping tools, thus preventing losses accrued from the use of each different tool. © Springer Science+Business Media Singapore 2016.	Cloud computing; Delivery channel system; MCA; Metafield data; MRM	10.1007/978-981-10-0068-3_26	Excluido
Bennett, Dennis E.	Governance and organizational requirements for effective model risk management	Effective model risk management requires a combination of analytical skills, governance and organizational structure, as well as the ability to negotiate. In early studies of model risk management (MRM), analytics and model validation were the primary points of focus. Various writers led the way in discussing the possibility of model errors and the need for analytic review and skepticism toward models and results. Others addressed the risk of assuming models to accurately represent or measure risk within financial institutions. Losses due to model failure resulted in fresh regulations addressing the risk posed by inaccurate or inappropriate models. Very few writers, however, have addressed comprehensive model risk governance, organization or negotiation structures within MRM. This paper expands on the foundation of model risk analytics to address the governance, organizational and human behavior challenges associated with enterprise MRM. It proposes a comprehensive framework for model risk governance, organizational responsibilities and human behavior covering the Risk Committee of the Board to MRM. It expands the definition of models into decision-support tools (DSTs) so that end-user computing (EUC), big data analytics and machine learning DSTs may be included in model risk governance, organization, negotiation and control. © 2017 Infopro Digital Risk (IP) Limited.	Decision support tools; Model governance; Model risk; Negotiation; Risk committee; Risk culture	10.21314/JRMV.2017.188	Incluido
Fan, Lingling and Schneider, Alex and Joumaa, Mazin	A new automated model validation tool for financial institutions	We present a new automated validation tool to validate predictive models for financial organizations based on the regulatory guidance of the US Federal Reserve and the Office of the Comptroller of the Currency. This automated tool is designed to help validate linear and logistic regression models. It automatically completes validation processes for seven areas: data sets, model algorithm assumptions, model coefficients and performance, model stability, backtesting, sensitivity testing and stress testing. The tool is packaged as a PYTHON library and can validate models developed in any language, such as PYTHON, R and the SAS language. Further, it can automatically generate a validation report as a portable document format (PDF) file while saving all the generated tables and charts in separate EXCEL and portable network graphic (PNG) files. With this automated tool, validators can standardize model validation procedures, improve efficiency and reduce human error. The tool can also be used during model development. © 2023 Infopro Digital Risk (IP) Limited	automation; financial institutions; linear regression; logistic regression; model risk management; model validation	10.21314/JRMV.2023.006	Incluido
Gan, Jingwei and Zhang, Shinan and Zhang, Chi and Li, Andy	Automated Counterfactual Generation in Financial Model Risk Management	The increasing adoption of machine learning and AI models in finance and banking poses challenges for model testing and model risk management (MRM). Conventional testing has limitations in evaluating more complex models while the latest testing methodologies have not yet been widely productionized. We propose a model-agnostic testing algorithm based on counterfactual generation for machine learning models in finance. Our method provides actionable counterfactual samples that can guide future model analysis and improvement. To make our algorithm modeler friendly, we design a cloud native implementation that is easy to integrate with any new and existing models, and ready for production-level automation, scalability, and parallel processing. © 2021 IEEE.	counterfactual generation; financial risk management; model testing	10.1109/BigData52589.2021.9671561	Incluido
Florens, J.P. and Gourieroux, C. and Monfort, A.	Model risk management: Limits and future of Bayesian approaches	This paper discusses the use of Bayesian approaches when the models are misspecified (model risk). In particular we explore the limits and future of Bayesian approaches in order to provide answers to the following questions recently asked by the prudential supervision for Finance/Insurance : How to measure model risk? How to use in a coherent way the different misspecified models (as rating models) usually employed within and between financial institutions. © 2019 GENES (Groupe des Ecoles en Economie et Statistiques). All rights reserved.	Bayesian Approach; Model Choice; Model Risk; Validation	10.15609/annaeconstat2009.136.0001	Incluido
Gudiño, Jose Juan Chavez and Mora, Jose Antonio Nuñez	Machine Learning Models, Risk Management Current Regulation and Perspectives	Current machine Learning techniques spreading in banking and quantitative finance brings issues to be face by risk management and financial institutions regulation. The former increases models predictive power but at the same time makes models more complex and has to struggle to better explain the role, weight and direction of explanatory variables in the outcomes as well as the sensibility of the result to changes in the feature, different technics have been developed to deal with this problem but there are still important tasks to be completed. On the other side the later has found that the use of internal models as regulated in early Basel regulation stages (Basel II) foster the risk sensitivity for regulatory capital allocation but may lead to undue complexity and so reducing comparability among banks risk. The regulatory response in Basel IV regulation has been to discouraged the use of internal models to reduce the impact of undue complexity in models to improve regulatory capital comparability in the banking system. This is the environment that machine learning models for regulatory purposes have to deal with. © The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022.	Basel regulation; Machine learning; Risk management	10.1007/978-981-19-4695-0_3	Incluido

Rinalducci, Sara and Longo, Valentina and Ceci, Luigi R. and Zolla, Lello	Targeted quantitative phosphoproteomic analysis of erythrocyte membranes during blood bank storage	One of the hallmarks of blood bank stored red blood cells (RBCs) is the irreversible transition from a discoid to a spherocyte-like morphology with membrane perturbation and cytoskeleton disorders. Therefore, identification of the storage-associated modifications in the protein-protein interactions between the cytoskeleton and the lipid bilayer may contribute to enlighten the molecular mechanisms involved in the alterations of mechanical properties of stored RBCs. Here we report the results obtained analyzing RBCs after 0, 21 and 35 days of storage under standard blood banking conditions by label free mass spectrometry (MS)-based experiments. We could quantitatively measure changes in the phosphorylation level of crucial phosphopeptides belonging to β-spectrin, ankyrin-1, α-adducin, dematin, glycophorin A and glycophorin C proteins. Data have been validated by both western blotting and pseudo-Multiple Reaction Monitoring (MRM). Although each phosphopeptide showed a distinctive trend, a sharp increase in the phosphorylation level during the storage duration was observed. Phosphopeptide mapping and structural modeling analysis indicated that the phosphorylated residues localize in protein functional domains fundamental for the maintenance of membrane structural integrity. Along with previous morphological evidence acquired by electron microscopy, our results seem to indicate that 21-day storage may represent a key point for the molecular processes leading to the erythrocyte deformability reduction observed during blood storage. These findings could therefore be helpful in understanding and preventing the morphology-linked mechanisms responsible for the post-transfusion survival of preserved RBCs. Copyright © 2015 John Wiley & Sons, Ltd. Copyright © 2015 John Wiley & Sons, Ltd.	blood storage; membrane skeletal proteins; MS-based quantitation; phosphorylation; red cells	10.1002/jms.3531	Excluido
Cosma, Simona and Rimo, Giuseppe and Torluccio, Giuseppe	Knowledge mapping of model risk in banking	For years, bank management has relied on mathematical, statistical and financial models, which increasingly expose banks to model risk. The latter is also extended by the phenomenon of innovation: machine learning, artificial intelligence and big data make the models more and more sophisticated and difficult to manage. This study aims to clarify how the literature on model risk is evolving through a bibliometric survey to understand state of the art and identify the discussion topics, open questions and challenges for the future. The study results show that the literature on model risk is still quite young and sparse. The problems to be solved are conceptual, computational, and organizational. The considerations made lead to the question of whether adding further complexity to model risk management is a solution or whether, on the contrary, it creates new model risks. © 2023 Elsevier Inc.	Banking; Bibliometric review; Machine learning impact; Model risk; Risk management	10.1016/j.irfa.2023.102800	Incluido
De Jongh, Pieter J. Riaan and Larney, Janette and Mare, Eben and Van Vuuren, Gary W. and Verster, Tanja	A proposed best practice model validation framework for banks	Background: With the increasing use of complex quantitative models in applications throughout the financial world, model risk has become a major concern. The credit crisis of 2008–2009 provoked added concern about the use of models in finance. Measuring and managing model risk has subsequently come under scrutiny from regulators, supervisors, banks and other financial institutions. Regulatory guidance indicates that meticulous monitoring of all phases of model development and implementation is required to mitigate this risk. Considerable resources must be mobilised for this purpose. The exercise must embrace model development, assembly, implementation, validation and effective governance. Setting: Model validation practices are generally patchy, disparate and sometimes contradictory, and although the Basel Accord and some regulatory authorities have attempted to establish guiding principles, no definite set of global standards exists. Aim: Assessing the available literature for the best validation practices. Methods: This comprehensive literature study provided a background to the complexities of effective model management and focussed on model validation as a component of model risk management. Results: We propose a coherent 'best practice' framework for model validation. Scorecard tools are also presented to evaluate if the proposed best practice model validation framework has been adequately assembled and implemented. Conclusion: The proposed best practice model validation framework is designed to assist firms in the construction of an effective, robust and fully compliant model validation programme and comprises three principal elements: model validation governance, policy and process. © 2017. The Authors. Licensee: AOSIS.	best, effective, framework, governance, literature	10.4102/sajems.v20i1.1490	Incluido
Wong, Yan and Chui, Xiuying and Qing, Jin	A Model of Risk Management Strategy at Banks in their Operational Activities	By partnering with the business, the operational-risk discipline can create a more secure and profitable institution. Successful banks embrace risks while developing powerful mechanisms to prevent or manage risk and stay ahead. This study is devoted to the problems of improving the banking risk management, taking into account the new regulatory and technological requirements based on the use of modern technology and combining the latest achievements in numerical mathematics, statistics and information technology. Focusing on literature and by conducting a case study on operational risk management on one of largest retail banks, this paper aims on answering how operational risk management can be improved by implicating risk governance and internal control frameworks. It was concluded that although improvements have taken place in how operational risks are being managed, there is still room for improvements. An intersection of risk type and business/functional area of focus, prioritized based on the business environment is critical when fulfilling this life cycle of risk management activities. When too linear, the steps are not agile enough, so the cycle needs to be performed on a micro-level and aggregated for macro-level insights/risk decision-making. The suggested model outlines the understanding of the safe and prudent management of risks under the current prudential framework. It also demonstrates the ability of appropriate operations management techniques to reduce operational risk losses in the banking industry significantly. © Journal of Management World, 2023.	Capital Requirements; Model-risk-management (MRM); Operational Risk Management (ORM); Prudential Regulations	N/A	Incluido
Hassani, Bertrand K.	Model risk management: From epistemology to corporate governance	One of the main concerns and regulatory topics that financial institutions have to deal with is model risk. Senior managers are increasingly considering model risk to be one of the most severe vulnerabilities to which a financial institution can be exposed. While the concept seems relatively simple, the definition of model risk (both traditional and regulatory), its origins (from dogma to misuse) and ways to manage it (from engineering conservatism into models to having a proper governance process) are not necessarily handled well by practitioners, academics and regulators. Having a clear definition of model failure and understanding its root causes allow us to adopt the appropriate management style to deal with potential issues that could lead to dramatic failures. In this paper, we conduct an analysis of model risk in an attempt to understand the main issues that lead to failures and the best way to address such issues. © Infopro Digital Limited 2019. All rights reserved.	Corporate governance; Epistemology; Management; Model risk; Risk measurement	10.21314/JRMV.2019.208	Incluido