MLOps aplicado en Gobierno y asuntos públicos

Sergio Guillén
Escuela de Electronica
Instituto Tecnologico
Cartago, Costa Rica
guillen_guillen@hotmail.com

***Abstract*—MLOps es un subconjunto de ModelOps es una práctica de colaboración y comunicación entre científicos de datos y profesionales de operaciones para ayudar a administrar el ciclo de vida de ML de producción. Es importante tomar en cuenta que DevOps ayuda a romper la barrera entre desarrollo (Cambios constantes) y operaciones (la necesidad de mantener estable el sistema) pero no es lo mismo que Machine Learning. Hay una diferencia fundamental entre ML y el software tradicional: ML no es solo código, es código y datos. Esto implica que el comportamiento del modelo también depende de los datos que recibirá en el momento de la predicción, los cuales no se pueden conocer a priori. En el presente documento se muestra un ejemplo de aplicación de MLOps en el ámbito gubernamental sobre los asuntos públicos.**

***Index Terms*—DevOps, MLOps, Dataflow.**

I. Introducción

Es importante, como punto de partida, tener claro las diferencias fundamentales entre MLOps y DevOps: 1) Versionar datos es diferente a Versionar código: en MLOps se trata de controlar versiones de conjuntos de datos mientras estos cambian su esquema y orígenes 2) Los requerimientos de una auditoría digital son diferentes cuando se maneja el código y los datos del cliente al mismo tiempo 3) La reutilización del modelo es diferente a la reutilización del software, ya que los modelos deben sintonizarse según los datos de entrada y cada escenario. 4) Para reutilizar un modelo en MLOps se necesita tener conocimiento sobre las transferencias que éste realizará en un pipeline de entrenamiento. 5) Los modelos tienden a deteriorarse a lo largo del tiempo y se necesita reentrenarlos según la demanda para garantizar que sigan siendo útiles en un contexto de producción.

Antes de abordar el tema de gobernanza y asuntos públicos también es importante recolectar los puntos claves sobre los cuales radica la importancia de utilizar MLOps en éste y cualquier ejemplo práctico existente: 1) Reproducibilidad y control de versiones del modelo: Significa el seguimiento, captura y manejo de datos y código (insumos) utilizados para el modelo. 2) Explicabilidad y Auditabilidad del modelo: Significa el mantenimiento de la integridad de los insumos y controles de acceso. 3) Empaquetamiento y Validación del modelo: Implica darle portabilidad a través de una variedad de plataformas. El rendimiento del modelo debe ser certificado para cumplir con requerimientos funcionales y de latencias. 4) Despliegue y monitoreo del sistema: Se debe mantener una trazabilidad sobre las señales de posibles desfasamientos que implique re-entrenamiento del sistema.

II. Azure Government

El primer ejemplo viene dado por un desarrollo originalmente orientado a solventar las necesidades del DevOps como lo es Azure. Sin embargo, su derivación posterior en Azure Machine Learning es una colección administrada de servicios en la nube, relevante para el aprendizaje automático, que se ofrece en forma de un espacio de trabajo y un kit de desarrollo de software (SDK). Un registro de modelos centralizado para ayudar a realizar un seguimiento de dichos modelos y experimentos, independientemente de dónde y cómo se hayan creado. Esto precisamente es la definición del tema que nos atañe.

Ahora bien, en campo específico de aplicación gubernamental, tenemos su mas reciente desarrollo para clientes gubernamentales: Azure Government. Los nuevos servicios de inteligencia artificial y aprendizaje automático disponibles en Azure Government incluyen Azure Cognitive Search, QnA Maker y Azure Machine Learning. Azure Cognitive Search (anteriormente conocido como Azure Search) es un servicio de búsqueda en la nube con capacidades de inteligencia artificial integradas que enriquecen todo tipo de información para identificar y explorar fácilmente contenido relevante a escala. La búsqueda como servicio completamente administrada lo ayuda a reducir la complejidad y escalar más rápido. Utiliza el mismo lenguaje integrado que Microsoft, Bing y Office y servicios de inteligencia artificial en el campo de visión de computadoras, lenguaje y habla. QnA Maker es un servicio de API basado en la nube que le permite crear una capa conversacional de preguntas y respuestas sobre sus datos existentes. Puede crear una base de conocimientos extrayendo preguntas y respuestas de contenido semiestructurado, incluidas preguntas frecuentes y documentos. Además responde automáticamente las preguntas de los usuarios con las respuestas de su base de conocimientos.

III. Iniciativas específicas

En diciembre de 2018 se adoptó el Plan Coordinado sobre el Desarrollo y Uso de la Inteligencia Artificial Made in Europe, con el fin de desarrollar acciones conjuntas para una cooperación más estrecha y eficiente entre los Estados miembros, Noruega, Suiza y la Comisión Europea en cuatro áreas

clave: aumentar la inversión, hacer más datos disponibles, fomentando el talento y asegurando la confianza. En este contexto, el paper sobre Inteligencia Artificial -A European approach to excellence and trust [COM(2020)]- va más allá al incluir una sección sobre Promoción de la adopción de la IA por parte del sector público, donde se menciona que es esencial que las administraciones públicas, los hospitales, los servicios públicos y de transporte, los supervisores financieros y otras áreas de interés público comiencen rápidamente a implementar productos y servicios que se basan en la IA en sus actividades, con un enfoque específico en el área de la salud y el transporte

A. SATIKAS [3]

En la Junta de Información y Registros Agrícolas de Estonia (ARIB, en inglés), la IA se utiliza para detectar si los pastizales agrícolas se han cortado o no mediante el reconocimiento de imágenes. Este sistema, llamado SATIKAS32, utiliza métodos de aprendizaje profundo y enfoques de redes neuronales convolucionales para analizar los datos satelitales provenientes del programa europeo COPERNICUS para detectar automáticamente si se ha cortado el césped en los pastizales de Estonia. Las imágenes de satélite óptico de Sentinel 1 y 2 se analizan junto con datos de referencia de campos de agricultores, registros de inspección históricos y datos meteorológicos del Servicio Meteorológico de Estonia. Este sistema de inteligencia artificial ahora se considera una de las primeras aplicaciones de inteligencia artificial utilizadas por el gobierno de Estonia. El sistema SATIKAS aún está en desarrollo y ampliará sus características y capacidades en un futuro próximo. Si bien por el momento el sistema es capaz de detectar el corte de césped, en el futuro también se utilizará para identificar diferentes tipos de cultivos y árboles.

B. Predictive System

En 2014, la Agencia Flamenca para la Infancia y la Familia (Kind en Gezin) desarrolló un sistema de inteligencia artificial que permite realizar predicciones más precisas para detectar los servicios de guardería que requieren una mayor inspección. Estas inspecciones permiten a las agencias mantener alta la calidad de los servicios de guardería y mejorar el bienestar de los niños. Durante la fase de desarrollo, la Agencia trabajó en estrecha colaboración con el equipo de ciencia de datos del Departamento de Bienestar, Salud Pública y Familia porque ya había algo de experiencia en minería de textos en el mismo. El sistema de IA ha tenido un mantenimiento y mejora constante del modelo para asegurar su precisión y confiabilidad: si se ignorara este mantenimiento, la precisión del modelo podría disminuir, lo que reduciría la confianza en el modelo y en otros proyectos futuros relacionados con los datos. Caso típico de MLOps en plena operación para un desarrollo implicado en gobernanza y asuntos públicos.

C. Automated Public Services

En el municipio de Trelleborg, Suecia, las tecnologías de inteligencia artificial se utilizan para automatizar varias decisiones de asistencia social desde 2016. Ese fue el primer municipio en utilizar la automatización robótica de procesos (RPA) para manejar diversas aplicaciones de asistencia social. Por el momento, el sistema automatizado de toma de decisiones puede procesar solicitudes de asistencia domiciliaria, prestaciones por enfermedad, prestaciones por desempleo e impuestos, y se ha considerado un ejemplo exitoso a seguir por otros. Varios otros municipios suecos están explorando cómo implementar el modelo de Trelleborg para obtener acceso a los mismos tipos de beneficios. Sin embargo, los informes de casos mencionan la gran necesidad de hacer que el proceso de automatización sea confiable. Si no hay confianza en el uso de la IA, el personal se verá obligado a verificar todos los procesos, lo que podría conducir a una disminución de la eficiencia y la eficacia. Además, algunos observadores expresaron su preocupación por el riesgo de excluir a algunos ciudadanos más vulnerables cuando todos los procesos están automatizados en línea, ya que esto dificulta la evaluación de las necesidades individuales. Precisamente es por esto que en la actualidad, ya se utiliza MLOps para generar esta confianza en el proceso de evolución del sistema de IA.

D. Chatbot UNA [4]

UNA está disponible tanto en el sitio web del Registro de Empresas como en la página de Facebook y como parte de la aplicación Messenger. UNA es capaz de responder preguntas frecuentes sobre el registro y liquidación de empresas, comerciantes, empresas y organizaciones. Si los ciudadanos ya tienen una solicitud en curso, también pueden pedir información al respecto. El alto compromiso de los recursos organizacionales dedicados a responder el mismo tipo de preguntas pudo reducirse fácilmente mediante el uso de la inteligencia artificial, especialmente las técnicas de procesamiento del lenguaje natural. Según algunos indicadores de desempeño, el 44% de las preguntas formuladas en UNA se consideran de carácter general y fáciles de resolver por el Chatbot.

E. Tengai [5]

El municipio sueco de Upplands-Bro ha comenzado a experimentar con el robot Tengai en sus procesos de reclutamiento desde junio de 2019. Tengai es uno de los primeros robots entrevistadores desarrollados con el objetivo de hacer que el proceso de reclutamiento sea menos sesgado que las prácticas tradicionales de entrevista. El robot es adoptado por la agencia de contratación y dotación de personal del municipio, lo que ya ha hecho que sus procesos de contratación sean menos sesgados. Los primeros resultados tras la adopción del robot Tengai se consideraron exitosos y atrajeron una atención significativa de los medios de comunicación al municipio debido al enfoque innovador de la contratación. Según uno de los directores del municipio, el robot Tengai ha hecho que los procesos de selección y contratación sean más rápidos, económicos e imparciales, liberando recursos cruciales para gastar en otras tareas.

F. SyRi [6]

Varios municipios de los Países Bajos han estado utilizando el sistema SyRi para detectar el fraude a la asistencia social de manera más eficaz. SyRi ha sido desarrollado por el gobierno holandés y utiliza varios indicadores de riesgo de los sistemas gubernamentales existentes, como impuestos, seguro médico, residencia, educación y muchos más, para detectar qué direcciones tienen un mayor riesgo de fraude o uso indebido de los beneficios sociales. SyRi se desarrolló en 2014 después de que varios municipios crearan sus propios sistemas para detectar fraudes. Para permitir el intercambio de diferentes elementos de información que son relevantes para SyRi, el sistema opera sobre una base legal que indica claramente qué tipo de datos se pueden capturar, almacenar y compartir. Sin embargo, no todo a sido color de rosa para este sistema. En los municipios donde el sistema realmente dio recomendaciones sobre posibles conductas fraudulentas, la tasa de éxito fue muy baja. Como los costos de SyRi se han estimado en más de 325.000 euros por año, esto ha generado que muchos se hayan preguntando si el sistema valió la pena por los costos financieros y de privacidad. Nótese que este sistema ha empezado también a gestionarse bajo la consigna de MLOps, lo que promete una reducción de costos hacia el futuro.

G. Unemployed profiling

El proceso de elaboración de perfiles automatizada divide a las personas desempleadas en tres categorías, teniendo en cuenta una serie de características individuales. La asignación a una categoría determinada determina para qué tipos de programas es elegible un beneficiario (por ejemplo, colocación laboral, formación profesional, aprendizaje, asignación de activación). El sistema se basa en datos recopilados durante una entrevista inicial (por ejemplo, edad, sexo, discapacidad y duración del desempleo) y una prueba posterior por computadora que califica 24 dimensiones diferentes. La asignación a uno de los tres grupos de perfiles indica el nivel necesario de apoyo y la carga de recursos. Es importante destacar que en este caso la categorización se traduce en decisiones binarias que van a impactar la vida de las personas de una forma decisiva, es decir, tendrán apoyo estatal o no. Sorprendentemente, como descubrió un estudio posterior, menos de 1 de cada 100 decisiones tomadas por el algoritmo han sido cuestionadas por los empleados responsables. Excluyendo la creencia en la extraordinaria precisión del algoritmo, otras de las razones encontradas por el estudio, para no cuestionar las decisiones automatizadas incluyen la falta de tiempo para considerar más detalles; miedo a las repercusiones de los supervisores; y una presunción de objetividad del proceso. De esta forma, lo que se suponía que era un mecanismo consultivo se terminó por convertir en el máximo responsable de la toma de decisiones.

H. VeriPol

Recientemente, la policía nacional española ha adoptado el sistema de inteligencia artificial VeriPol para detectar informes policiales falsos. El sistema se diseñó para integrarse en el sistema de información existente de la Policía Nacional española denominado SIDENPOL, lo que permite un uso más sencillo e integración en las prácticas laborales existentes. Su desarrollo fue el resultado de un proyecto de colaboración entre la Universidad de Cardiff, la Universidad Carlos III de Madrid y la Policía Nacional de España. La base de datos de los informes policiales se puso a disposición de los investigadores de las universidades con el fin de entrenar el sistema de IA. Para ello, se utilizaron 1122 informes, incluidos 534 informes verdaderos y 588 falsos. VeriPol explota una combinación de procesamiento de lenguaje natural y algoritmos de clasificación de aprendizaje automático, capaces de estimar la probabilidad de informes policiales falsos con una precisión significativa. Además de eso, el sistema también permite conocer las diferencias entre los informes policiales falsos y verdaderos. Por ejemplo, los estudios piloto encontraron que es más probable que los informes policiales falsos incluyan declaraciones más breves, centradas en los objetos robados y que carecen de detalles. Actualmente el sistema se ha implementado para que lo utilicen todos los departamentos de la Policía Nacional de España. El impacto que se espera del uso del sistema es que sea capaz de detectar informes falsos tempranamente, dejando más recursos policiales disponibles para engancharse en otras tareas e informes, mientras que al mismo tiempo disuada a las personas de presentar declaraciones falsas. Un beneficio adicional del sistema que obtiene más información sobre cómo la gente le miente a los agentes de policía, así como adquirir más conocimientos sobre la detección de informes policiales verdaderos y falsos.

IV. Bibliografía

[1] https://knowledge4policy.ec.europa.eu/ai-watch/topic/ai-public-sector_en

[2] https://azure.microsoft.com/es-es/blog/automated-machine-learning-and-mlops-with-azure-machine-learning/

[3] https://ec.europa.eu/jrc/sites/jrcsh/files/bleive.pdf

[4] https://oecd-opsi.org/innovations/una-the-first-virtual-assistant-of-public-administration-in-latvia/

[5] https://www.tengai-unbiased.com/

[6] https://www.loc.gov/law/foreign-news/article/netherlands-court-prohibits-governments-use-of-ai-software-to-detect-welfare-fraud/