

Sistema de Machine Learning para préstamos integrados

Monitoreo y despliegue estandarizado de modelos de Machine Learning en producción.

Industria: **Banca**

*Imágen con fines ilustrativos

Resumen

Modernización y mejora de la infraestructura de Machine Learning para un cliente Fintech, atendiendo la falta de visibilidad sobre las predicciones en producción.

Principales desafíos

- Falta de visibilidad y monitoreo de las predicciones en producción de los modelos de Machine Learning.
- Ausencia de infraestructura para pruebas A/B de modelos y parámetros en producción.
- El proceso de despliegue de modelos no estaba estandarizado ni bien documentado.

Enfoque

El equipo de Novacomp se integró al equipo nativo de la compañía, participando en sus daily stand-ups, Sprint Planning y sincronizaciones, en colaboración con otros equipos de Novacomp como Ingeniería de Datos y Analítica. Como servicio de Staff Augmentation, se diseñó e implementó una arquitectura de Machine Learning de punta usando FastAPI, Split.io, Weights & Biases, Docker y Apache Pulsar.

Resultados

- Desarrollo de un sistema de monitoreo basado en eventos para las predicciones de Machine Learning usando Apache Pulsar.
- Documentación del proceso de despliegue, permitiendo al equipo de ciencia de datos implementar cambios de forma autónoma.
- Mejora en la latencia de los servicios de Machine Learning mediante refactorización de código y llamadas asíncronas.

Tecnologías utilizadas

Python	Catboost	Split.io
Weights & Biases	Apache Pulsar	
XGBoost	Scikit-learn	Flask
FastAPI	Docker	

Conclusión

Implementación de una arquitectura de Machine Learning basada en microservicios que permitió despliegues paralelos de modelos para pruebas A/B.