

Roberto Cholaky Mejía

Data Scientist

Teléfono

E-Mail

LinkedIn

GitHub

(+569) 9658 1200

rocholaky@gmail.com

linkedin.com/in/robertocholaky

github.com/rocholaky

EXPERIENCIA LABORAL

Data Scientist

Brain Food, an Artefact company

Julio 2023 - Actualidad

- Procesé, analicé y entrené modelos con grandes volúmenes de datos (>60 millones de registros) en entornos de **big data**, utilizando **PySpark y SQL** para optimizar modelos y algoritmos. Implementé **pipelines de aprendizaje automático en PySpark ML** para mejorar la escalabilidad y eficiencia en la construcción de modelos para empresas como VISA, Nestlé y Femsa.
- Desarrollé diversos **modelos de clasificación**, incluyendo modelos de propensión de compra, fuga y activación de tarjeta. Estos modelos predicen mensualmente la probabilidad de que un cliente incurra en un comportamiento específico, basándose en la actividad reciente de las últimas semanas.
- Desarrollé modelos de **segmentación de usuarios** aplicando técnicas de Machine Learning para comprender patrones internos de los clientes. Esto permitió definir perfiles para la planificación y optimización de campañas de Marketing.
- Colaboré con equipos de **producto e ingeniería** para integrar modelos predictivos en sistemas productivos, asegurando su correcta implementación y proporcionando información clave para la toma de decisiones estratégicas.
- Desarrollé y producticé **agentes de GenIA** en GCP y Azure, utilizando **Retrieval-Augmented Generation (RAG)** para la evaluación automatizada de profesores. Además, diseñé agentes de apoyo en la construcción de experimentos, empleando frameworks como **Langchain y Langraph**.
- Desarrollé e implementé **modelos de forecasting** para optimizar la predicción de demanda en la industria alimentaria (Sell-in/Sell-out), logrando un error menor al 10% para Nestlé. Diseñé modelos predictivos de ventas minoristas, mejorando la precisión en un 11% al evaluar Prophet, modelos de árbol y ARIMA.

Senior Data Engineer

Arkenco Marketing and Sales Analytics

Marzo 2020 - Junio 2023

- Lideré la modernización de la infraestructura de datos en Arkenco mediante el desarrollo de **pipelines escalables y automatizados**, integrando múltiples fuentes de datos y mejorando la disponibilidad de información en tiempo real para optimizar estrategias de marketing y ventas.
- Implementé **CRM Arkenco**, una solución robusta basada en **Docker, Python (Django) y React**, fortaleciendo la infraestructura backend para gestionar un creciente volumen de clientes y mejorar la eficiencia operativa.
- Innové en la interacción con clientes al integrar la plataforma con la **API de WhatsApp** y desarrollar **chatbots inteligentes**, automatizando la comunicación con clientes y contribuyendo a un **aumento del 15% en la captación de clientes**.

Ingeniero en Machine Learning - Práctica Profesional II

Asicom

Enero - Marzo 2020

Se diseñó un prototipo de chatbot capaz de extraer información a partir de la imagen de la cédula de identidad, además de documentos en PDF y Excel de manera automatizada. Además, se realizó un tutorial sobre el uso de Amazon Web Services para la capacitación de la empresa.

EDUCACIÓN

Diploma Making Data-Driven Decisions	Massachusetts Institute of Technology (MIT) IDSS
Magíster en Ciencias de la Ingeniería Mención Eléctrica (Inteligencia Computacional)	Universidad de Chile (Distinción Máxima)
Minor en Ciencia de Datos	Universidad de Chile
Ingeniería Civil Eléctrica (Mención Inteligencia Computacional)	Universidad de Chile (Distinción Máxima)

EXPERIENCIA EN INVESTIGACIÓN

Investigador proyectos aprendizaje reforzado y regresión simbólica

Universidad de Chile

2021 - 2023

Realicé investigaciones de cooperación de agentes mediante aprendizaje reforzado, diseñando y probando algoritmos Actor-Crítico centralizados en entornos de robots multiagente, lo que permitió reducir los tiempos de entrenamiento en al menos un 15% al incluir encoders que agregan nuevas funcionalidades progresivamente, modelo que fue desarrollado en Tensorflow. Asimismo, diseñé e implementé un innovador modelo de red neuronal en PyTorch para resolver problemas de regresión simbólica, superando las limitaciones de búsqueda de potencias de los métodos convencionales y logrando mejoras de precisión superiores al 25%.

HABILIDADES TÉCNICAS

Lenguajes de Programación

Python | SQL | Java

Software

Docker | Git | Microsoft Office

Idiomas

Español - Nativo | Inglés Avanzado - B2 (TOEFL)

Frameworks

PyTorch | TensorFlow | Spark | MLflow | Prophet | Pandas | Polars | Scikit-learn | statsmodels | Nixtla

Servicios de Nube (Cloud Services)

Amazon Web Services (AWS) | Google Cloud | Azure

Visualización

Matplotlib | Seaborn | Tableau | Looker Studio

CERTIFICACIONES

Natural Language Processing Specialization

Especialización en NLP, donde adquirió una comprensión profunda de modelos avanzados como Transformer, GPT y BERT, aplicando técnicas de deep learning para el análisis y comprensión del lenguaje natural. Desarrolló competencias en aplicaciones como generación de texto y análisis semántico.

Machine Learning Engineer Track

Diploma en Machine Learning Engineering, cubriendo el ciclo completo desde el desarrollo hasta el despliegue de modelos en producción. Profundizó en MLOps, CI/CD, versionado de modelos y gestión de problemas como data drift y concept drift para asegurar la confiabilidad y eficiencia de los modelos en entornos dinámicos.

DeepLearning.ai

Datacamp.com