

Roy Quillca Pacco

Data Scientist

Piura - Piura, Perú
(+51) 997277973
roquidata@gmail.com
[Linkedin](#) - [Github](#)

Data Scientist con background en Estadística e Ingeniería Civil que me ayudaron a desarrollar una sólida base en el análisis, resolución de problemas complejos, y aplicar esta habilidad en el mundo de la programación con éxito. Conocimientos en metodologías ágiles, flujos de trabajo, estructura de datos, evaluación de modelos de machine learning, extracción, transformación y análisis de grandes volúmenes de datos. Tengo experiencia en desarrollo de ETL automatizado, generación de reportes y refactorización del pipeline de modelos de machine learning con Python, SQL, Power BI, FastAPI, Scikit-learn, Streamlit y otras librerías de procesamiento de Python. Con pensamiento creativo, mentalidad ágil y sobre todo autodidacta.

PROYECTOS

Experiencia académica.

Henry Bootcamp, Buenos Aires, Argentina.

Data Science - “Análisis de inversión en el mercado de taxis para lanzar una flota de vehículos eléctricos en la ciudad de Nueva York - DaTaxi” **abr. 2023**

- Desarrollar las actividades siguiendo la metodología scrum.
- Desarrollar ingeniería de datos, data warehouse, modelo de machine learning e interfaz de usuario (DaTaxi) para realizar predicciones sobre métricas en particular.

Herramientas: Spark, Streamlit, Scikit-learn, Pyspark, Power BI, Plotly, Pandas, Python, Trello. **Base de Datos:** Azure SQL. **Plataformas:** Azure.

Data Engineering - “Análisis de inversión en empresas del índice SP500” **feb. – mar. 2023**

- Desarrollar el ETL automatizado para extraer datos de las inversiones en empresa del índice SP500 desde yfinance de los últimos 23 años.
- Desarrollar un dashboard interactivo con el análisis de inversiones por sector y compañía que permitan seleccionar las tres mejores compañías para invertir en los próximos meses.

Herramientas: Python, Pandas, Numpy, Plotly, Streamlit. **Base de Datos:** Archivos parquet. **Plataformas:** Streamlit Community.

MLOps & Data Engineering - “Sistema de recomendación en una startup de agregación de plataformas de streaming” **feb. – mar. 2023**

- Recopilación, procesamiento y carga de datos brindados por el usuario/cliente en Google Cloud Storage.
- Desarrollo e implementación del modelo de recomendación basado en el filtro colaborativo de los usuarios usando el algoritmo **Singular Value Decomposition (SVD)**.
- Desarrollo de la API en **Deta Space** de acuerdo con los requerimientos del usuario/cliente y el modelo de machine learning.

Herramientas: Python, Pandas, scikit-learn, Docker, FastAPI, Github, DVC. **Base de Datos:** Cloud Storage. **Plataformas:** Google Cloud Platform, Deta Space, Gradio.

TECNOLOGÍAS:

Lenguajes de Programación: Python, SQL, JavaScript, HTML, CSS **Herramientas:** Streamlit, Pandas, Numpy, scikit-learn, FastAPI, Plotly, Git, Github, Docker, Power BI, Pyspark. **Base de Datos:** MySQL, PostgreSQL. **Plataformas:** GCP, Azure, Deta Space. **Sistemas operativos:** Windows, Linux (Ubuntu).

EDUCACIÓN PROFESIONAL

- Data Science. Henry Bootcamp. 2022.
- Cuarto año de Ingeniería Civil. Universidad de Piura. 2019.

EDUCACIÓN COMPLEMENTARIA

- Curso para crear servicios web con Python y FastAPI. CódigoFacilito. Octubre, 2022.
- Certificate of Achievement: SQL. 365 Data Science. Octubre, 2022.
- Proyecto del PlatziChallenge Datacademy 2022. Platzi. Mayo, 2022.
- Python Challenge. Platzi. Noviembre, 2021.

IDIOMAS

- Inglés B1 – Intermedio.