

OLIVER GONZÁLEZ BAGATELLA

 <https://github.com/Olivergb98>  <https://www.linkedin.com/in/oliver-gb/> 
olivergb98@gmail.com  Puebla, Puebla, México

Data Scientist — Machine Learning Engineer — NLP Specialist (Entry-level)

FORMACIÓN

Ingeniería en Agronomía <i>Universidad Popular Autónoma de Puebla (UPAEP), Puebla, México</i>	<i>Junio 2018 - Mayo 2023</i>
Bootcamp de Ciencia de Datos <i>TripleTen</i>	<i>Junio 2024 - Junio 2025</i>

PERFIL PROFESIONAL

Científico de Datos con formación en Ingeniería, y ahora con especialización en Machine Learning, NLP y análisis estadístico. Experiencia en proyectos aplicados con datos reales, implementando modelos de regresión, clasificación y procesamiento de lenguaje natural (NLP). Fuerte dominio en Python, SQL y bibliotecas como Scikit-learn, TensorFlow, NLTK y BERT. Capacidad para traducir problemas de negocio en soluciones técnicas con impacto. Me encuentro en búsqueda de oportunidades como Data Scientist donde pueda aportar soluciones basadas en datos y seguir desarrollando mi potencial técnico.

HABILIDADES TÉCNICAS

Lenguajes:	Python, SQL
Bibliotecas:	Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch. Machine Learning: Regresión, Clasificación, Random Forest, LGBM, Tuning de Hiperparámetros, NLP: NLTK, SpaCy, BERT, Tfidf Vectorizer. Visualización: Matplotlib, Seaborn
Idiomas:	Español: Nativo — Inglés: Avanzado

EXPERIENCIA TÉCNICA (DISPONIBLE EN GITHUB)

- Optimización de Planes Telefónicos con ML (Proyecto 9 del bootcamp)**
- Analicé patrones de consumo de clientes para determinar si el plan contratado se ajustaba a su comportamiento real. Entrené y comparé modelos de regresión lineal, árboles de decisión y Random Forest. Ajusté hiperparámetros con GridSearch para maximizar el rendimiento predictivo.
 - Logros: Mejoré en 23% la precisión del modelo sugerido para la empresa. Las recomendaciones generadas impactaron en la estrategia de fidelización y rentabilidad.
- Clasificación de Sentimiento en Reseñas de Películas (NLP) (Proyecto 15 del bootcamp)**
- Implementé modelos de clasificación de texto para identificar sentimientos positivos o negativos. Utilicé técnicas de vectorización (Tfidf, BERT) y modelos como LGBM y Regresión Logística. Evalué con métricas de desempeño (F1-score) y tiempos de computación.
 - Logros: Alcancé F1-score del 89% en validación con reducción de tiempo de entrenamiento en un 30%. Proyecto con potencial de aplicación en sistemas de recomendación y atención al cliente.

EXPERIENCIA LABORAL

- Data annotator - DataAnnotation** *Abril 2025 - Actual*
- Anotación y validación manual de datos multimodales (texto, imágenes y audio) para el entrenamiento y mejora de modelos de lenguaje de gran escala (LLM).
 - Aplicación de criterios técnicos rigurosos para garantizar precisión semántica y consistencia en conjuntos de datos destinados a inteligencia artificial.
 - Contribución directa a la calidad de sistemas de aprendizaje automático mediante la curación de datos relevantes, diversos y libres de sesgos.
- Asistente de Investigación Territorial – Agrova Avalúos** *2018 - Actual*
- Apoyo en procesos de investigación y recolección de datos para avalúos de predios agropecuarios en distintos municipios.
 - Captura y análisis de información geográfica, catastral y socioeconómica relevante para la valuación de

terrenos rurales.

- Realización de visitas técnicas a predios para levantamiento de datos, entrevistas y validación de información en campo.
- Consulta de bases institucionales y fuentes oficiales para enriquecer los estudios de valuación.

Colaborador en Rancho Familiar – Experiencia Práctica

2024

- Participación en actividades generales de un rancho lechero, con enfoque en la comprensión de procesos productivos y operativos del sector agropecuario.
- Aprendizaje práctico sobre manejo de ganado, logística de ordeña, mantenimiento básico y rutinas diarias propias de la industria rural.
- Exposición directa a dinámicas de trabajo físico, colaboración en equipo y resolución de problemas cotidianos en un entorno agrícola.

Quality Assurance Specialist en Scale AI

Dic 2021 - Mayo 2023

- Validación y clasificación de información para modelos de machine learning, asegurando altos estándares de calidad en datasets utilizados en entrenamientos supervisados.
- Ejecución eficiente de tareas bajo métricas de rendimiento, cumpliendo tiempos de entrega exigentes en entornos ágiles.
- Actualización dinámica de parámetros de proyecto según lineamientos internos y retroalimentación recibida por los líderes de equipo.