

# OLIVER GONZÁLEZ BAGATELLA

🌐 <https://github.com/Olivergb98> in <https://www.linkedin.com/in/oliver-gb/> ✉ [olivergb98@gmail.com](mailto:olivergb98@gmail.com)

📍 Puebla, Puebla, Mexico

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

## FORMACIÓN

### Ingeniería en Agronomía

*Junio 2018 - Mayo 2023*

*Universidad Popular Autónoma de Puebla (UPAEP), Puebla, México*

### Bootcamp de Ciencia de Datos

*Junio 2024 - Junio 2025*

*TripleTen*

## PERFIL PROFESIONAL

Científico de Datos con formación en Ingeniería y 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, TfidfVectorizer. 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 Sentimientos 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 en DataAnnotation

*Abril 2025 - Actual*

- Realización de anotación manual y validación de datos para proyectos de machine learning, con el objetivo de mejorar la calidad general del LLM, incluyendo imágenes, texto y/o audio.

### Quality Assurance Specialist en Scale AI

*Dic 2021 - Mayo 2023*

- Catalogar información de machine learning, con un fuerte énfasis en calidad. Completar tareas en tiempo y forma mientras se trabaja de la mano con el cliente y sus requerimientos