

# Osvaldo Velazquez Gonzalez

+52 961 192 6030 | [osvaldodvego@gmail.com](mailto:osvaldodvego@gmail.com) | [linkedin.com/in/osvaldovg](https://www.linkedin.com/in/osvaldovg) | [github.com/valdolab](https://github.com/valdolab)

Languages: English (professional proficiency), Spanish

## EXPERIENCE

### Senior Data Scientist

Feb 2023 – Present

*EnhanceIT*

*Atlanta, USA*

- Development end-to-end of a churn predictive model applying XGBoost and AutoPilot in AWS SageMaker using demographic information about the customer.
- Development PoC of a sentiment analysis model applying transformers and transfer learning. Also, the development of a text Generative model to assist in the generation of product descriptions applying LLM, such as GPT3.5.
- Collaborate with ML engineers, data engineers, and stakeholders to develop data and models in an accurate manner.
- Deploying and monitoring ML models in order to report and track model performance during the time.

### Machine Learning Engineer

Aug 2022 – Feb 2023

*Tata Consultancy services (Goldman Sachs Group, Inc. Client)*

*Guadalajara, Mexico*

- Developed a product's demand forecasting model using Time-Series models, such as ARIMA's and LSTM.
- Migrated, and deployed ML models to identify high-risk clients and opportunities using XGBoost, and Random Forest.
- Development of pipelines of machine learning models, ensuring reliable model deployment, monitoring, and management.
- MLOps applications using ML and Cloud computing tools.

### Machine Learning Research Assistant

Sep 2020 – July 2022

*Centro de Investigación en Computación, IPN*

*Mexico City, Mexico*

- Designed a *novel* Machine Learning algorithm to deal efficiently with class imbalance, with higher performance of various state-of-the-art models.
- Implemented several ML models, like k-NN, RandomForest, C4.5, MLP, Naive Bayes, and SVM in order to predict and pre-diagnosis several respiratory diseases.
- Proposed and conducted ML research experiments applying post-hoc methods like Friedman and Holm to evaluate designed models against state-of-the-art models.

### Data Analyst

Dec 2019 – Sep 2020

*Subsecretaría de Educación Federalizada de Chiapas (SEF)*

*Tuxtla Gutiérrez, Mexico*

- Cleaned data, conducted analyses, and created reports using SQL, Tableau, Power Bi, and Excel.
- Played a role in gathering ETL requirements and translating them into functional needs.

### Undergraduate AI Research Assistant

Jun 2019 – Nov 2019

*Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE)*

*Tonantzintla, Mexico*

- Developed a pre-diagnosis interface utilizing kNN and neuro-fuzzy models. Achieved prediction of infant hearing loss-related diseases and identification of reasons behind an infant's cry. Applied audio signal processing techniques.
- Developed and deployed a classification model into a medical software tool to support the early detection of infant diseases.

## EDUCATION

### MSc in Computer Science (Cum Laude)

*Centro de Investigación en Computación, Instituto Politécnico Nacional (CIC-IPN)*

*Mexico City, Mexico*

*Thesis: Computational Intelligence Algorithms for Medical Pre-diagnosis of Respiratory Diseases*

### BEng in Computational Systems

*Instituto Tecnológico de Tuxtla Gutiérrez (TecNM)*

*Tuxtla Gutiérrez, Mexico*

## TECHNICAL SKILLS

**Languages:** proficiency in Python, SQL, Matlab. Good understanding of R, PHP, JavaScript, Bash.

**Frameworks and ML Tools:** Flask, FastAPI, Git, Docker, GCP, AWS, VS Code, Databricks, PySpark, LangChain, AWS SageMaker, Google VertexAI, DataBricks, OpenIA libs, TensorFlow, Keras, scikit-learn, NLTK.

**Libraries:** Pandas, NumPy, Matplotlib, Seaborn, SciPy, Scikit-learn, TensorFlow, Keras, NLTK, Stanza.

**Data Science and ML Specialities.** *Natural Language Processing (NLP):* transfer learning, text embedding, universal sentence encoders, BERT, roBERTa. *Computer Vision:* EfficientNet, ResNet50, YOLO. *classic classifiers:* kNN, SVM, Neive Bayes, MLP, Random Forest, XGBoost, linear regression. *Time Series:* ARIMAs families, NAR, LSTM. *Clustering:* KMeans.

**Analytics:** feature selection, and engineering, Bayesian modeling, PCA, L1 and L2 regularization, and hyperparameter tuning.

**Optimization:** metaheuristic, evolutionary, and genetic algorithms.

## CERTIFICATIONS

### Database Desing and Programming with SQL

Jan 2017 – Jun 2017

- Imparted by Oracle Academy.