# Isaac Gutierrez, Artificial Intelligence Engineer

**Phone:** +52 2224738808
**Email:** isaacgr2121@gmail.com
**LinkedIn:** LinkedIn Profile
**GitHub:** GitHub Profile

## PROFILE

As a Machine Learning Engineer with extensive experience in developing and deploying computer vision models, I have successfully built and optimized object detection and tracking systems using state-of-the-art architectures such as YOLO and SSD, with a strong focus on real-time performance for edge and embedded devices. My hands-on expertise includes quantizing and converting models for deployment with TensorFlow Lite and ONNX, as well as collaborating closely with embedded engineers to tune and optimize performance on Linux-based platforms. I have managed the full machine learning lifecycle, from data collection and labeling pipeline development to model training, evaluation, and continuous improvement, leveraging robust Python, PyTorch, and TensorFlow skills. My background includes optimizing data workflows on cloud and enterprise platforms, and I am adept at problem-solving, cross-team communication, and adapting to new technologies and evolving industry trends.

## EMPLOYMENT HISTORY

**Machine Learning Engineer,** Apex Systems (Nov 2024 — Present)

- Optimized data workflows and managed enterprise-level databases using Snowflake and AWS, significantly improving efficiency for AI/ML projects
- Developed a model to assess product usage and define performance metrics, which enhanced model accuracy and delivered actionable operational insights
- Created a self-supervised embeddings model to derive vector representations from time series interaction data, facilitating accurate usage tracking and anomaly detection

**Machine Learning Engineer,** Kosmos (Feb 2022 — Nov 2024)

- Designed and implemented advanced facial detection and identification systems using state-of-the-art AI models
- Created a counterfeit image detection model for banks, combining CNN-based spoof detection with effective OCR techniques
- Enhanced machine learning models in production and fine-tuned a Large Language Model (LLM) to boost code generation efficiency
-

**Junior Financial Analyst,** Vector Casa de Bolsa (Jan 2021 — Dec 2021)

- Collaborated with a Quant in Asset Management to develop AI solutions tailored for financial applications
- Designed a social volume monitor for high-volume stocks using NLP techniques, emphasizing Jaccard similarity and word count analysis of Bank of Mexico minutes
- Utilized a Large Language Model to predict monetary policy changes by extracting insights from these minutes

## EDUCATION

**Actuarial Sciences,** Universidad de las Américas - Puebla (Jan 2019 — Dec 2023)

During my studies I took classes in statistics, probability, econometrics and simulation. Learned to read academic papers and write essays.

**Honor Student,** Universidad de las Américas - Puebla (Jan 2019 — Dec 2023)

Honor student working under Dr. Luz Maria Garcia Avila in applications for Set and Ramsey theory.

## LANGUAGES

- **English:** C1
- **Spanish:** Native speaker

## SKILLS

- Object Detection
- Edge Deployment
- Model Quantization & Optimization
- Python
- PyTorch
- TensorFlow
- Linux Environments
- C++
- Data Pipeline Development
- Model Evaluation & Monitoring
- SSD
- YOLO

## PROJECTS

**Spoof Detection Model** (Feb 2022 — Feb 2024)

Develop a spoof detection model, using convolutional neural networks (CNN), the model is based on state-of-the-art techniques. The model first created a depth-map for the image and then made inferences on the image and the depth-map.

**Face detection** (Feb 2022 — Feb 2023)

Develop a small face detection model (SDD) that was deployed on edge using tf-lite and ONNX.

**Development of a Facial Attribute Detection Model for Mobile and Web** (Oct 2022 — Jul 2023)

Developed a facial attribute detection model that was efficient and small to be implemented on mobile and web using multi-task training techniques. The model made inferences on:

- Occlusions on faces
- Angle of the head
- Use of a hat
- Use of glasses (no glasses, glasses and tinted glasses)
- Whether or not the face was properly lit

**Document Cropping and Alignment Model** (Jul 2024 — Aug 2024)

Developed a recurrent convolutional neural network to detect the corners of a document in pictures to crop and align before feeding the output to an OCR system.

**Improvement and Development of OCR Service** (Jul 2023 — Present)

Worked on the data extraction process for documents such as tax status certificates and identifications, additionally developed a new system to extract data directly from the PDF using clustering techniques and graph neural networks (GNN).

**Finetuning for Code Generation LLM** (Mar 2024 — Present)

Worked on finetuning existing code LLM for code generation for a no code platform.

**Sentiment Analysis for Bank of Mexico Minutes** (Aug 2022 — Dec 2022)

Developed a Transformer language model by leveraging an existing pre-trained BERT model and training it with paragraphs from bank of Mexico minutes to make inferences on monetary policy.

**Sentiment Analysis about Feminism on Twitter** (Mar 2022 — Feb 2023)

A research project to measure the change in sentiment about feminism and women's rights in Mexico over the last 10 years on Twitter, using language models.

**Interest Curves Monitor** (Apr 2021 — Jun 2022)

Developed an interactive curves monitor in excel using XLwings. The monitor would bootstrap a curve out bond data and showed relevant markers while allowing the user to change price targets and create theoretical curves.

**Social Volume Monitor** (Jan 2021 — Apr 2021)

Developed a Web Page monitor for tracking the number of tweets and its correlations with the trade volume in the stock market using python. First we would pull tweets using the twitter API then we would use sentiment classification to count bullish and bearish tweets and then show how it contrasts with buy and sell volumes. The idea was to use the sentiment of tweets to supplement incomplete volume data.

**Word Analysis for Bank of Mexico Minutes** (Jun 2022 — Aug 2022)

I developed a python script that would scrape minutes from the bank of Mexico then would stem the words and do a word count for relevant terms while also checking jaccard similarity with previous documents.

**Development and Maintenance of Model to Solve Captchas** (Feb 2022 — Feb 2022)

Trained and maintained 2 models to solve captchas. Additionally, developed a system to monitor their performance.

**Development of a Data Collection Service** (Jan 2024 — Mar 2024)

Developed a service using SQLite and FAST API to collect images that passed an inference from my facial attribute detection model.

**Undergraduate Thesis** (2022)

I wrote my math thesis about an alternative proof to Ramsey's theorem using a modern method known as forcing. Writing my thesis helped me develop critical thinking and allowed me to familiarize myself with reading papers.

**SAT and IMSS Scraping Service** (Jan 2022 — Nov 2024)

Developed and maintained scraping system to download government documents from clients using selenium and BeautifulSoup.

**Implemented State of the Art Paper on Vector Databases** (Jan 2025 — Present)

Implemented "Contextual document embedding" a paper in text retrieval, I implemented the paper to allow me to efficiently search a database of monsters for D&D.

From implementing the paper I learned:

- Gradient caching techniques
- Modern best practices vector databases
- State of the art techniques for contrastive learning

## COURSES

**Machine Learning,** Stanford University (Dec 2020 — Feb 2021)

**Deep Learning Specialization,** DeepLearning.AI (Feb 2020 — May 2021)

**Solid Edge ST10 Certified Associate,** Siemens (Jan 2017 — Dec 2025)

## EXTRA-CURRICULAR ACTIVITIES

**Participant in the Research Congress CUAM-ACMor 2017**

Researched into methodologies to model the prey-predator interactions between the tilapia and the axolotl to validate estimates made by experts on the topic

**May 2017** Category: Math and Physics

**Finalist in the Research Congress CUAM-ACMor 2018**

Proposed a new route public transport route by systematically searching for the route that would maximize the convenience for the largest number of people.

**Mar 2018** Category: Math and Physics