Diego Aarón Moreno Galván

Nationality: Mexican | XXXXXXXXX | Phone number: XXXXXX | Email: diego.moreno@cimat.mx or diego.amg97@gmail.com

Background

I am a graduate student of the "Centro de Investigación en Matemáticas (CIMAT)," where **I got my master's degree in Computer Science** (June 2022). My research thesis was related to deep learning and Natural Language Processing topics, specifically using multimodal data with a proposed architecture based on transformers. I am a graduate student of the Universidad de Guanajuato, where I achieved **my bachelor's degree in Mathematics** (June 2020). Obtaining my bachelor's degree, I did my research thesis about the benefits in the human health of a novel agave product, doing statistical analysis and computer modeling. My interests include Data Analysis with Machine Learning, Deep Learning, Language models, Transformers, and Multimodal Classification.

Teamwork

- **CIMAT 2022 Industrial Problem Solutions Workshop**. We were a 10-member team working to solve a company's problem: auto-detecting and counting the number of mosquito eggs from pictures.
- **CIMAT 2021 Industrial Problem Solutions Workshop**. We were a 10-member team working to find a dynamic index for modeling COVID cases in Mexico.
- **Bachelor's Degree Thesis 2020**. Working on my thesis, I gained expertise as a research assistant and as a team with other scientists of different areas, such as doctors and biologists.
- **University of Michigan BDSI 2019**. A Big Data summer research program. We were a 4-member team working to develop a new neural network architecture for time-series ICU data (proposed by us).

Research and Expertise

- **Master's Thesis 2022.** I developed a novel NN architecture with PyTorch for multimodal data (data that contains channels (text, video, and audio) with the same information) and making robust classification decisions. This architecture involves the BERT model, Multimodal Transformers, and Gated Multimodal Units for combining relevant information. We have achieved the **SotA** scores in classifying movie genres and emotions. A **paper** will be published in Aug 2022.
- **Technological Project 2021**. Research current problems of **multimodal language classification** (optimal use of text, video, and audio) to obtain a robust system and achieve the task. For example, on Facebook, we can detect aggression or emotions in videos or posts.
- **Bachelor's Thesis 2020**: data from a new agave product called agavins was statistically analyzed. Agavins have a significant impact on health, for example, on diabetes. We classified patients' improvement in consuming agavins. We performed data mining and information retrieval to fit a logistic regression model.
- **University of Michigan BDSI 2019**. We had hospital data of patients who had entered to ICU. There was data from 12k patients during the first 48 hours, such as blood pressure, beats per minute, etc. We analyzed patterns in time-series data proposing a novel architecture considering if data was imputed using deep learning with PyTorch to predict if a newly hospitalized person was at risk of death.

Teaching experience

- **Natural Language Processing teacher assistant 2022.** As a teacher assistant for **master students**, you give classes or help sessions and check students' homework.
- **Computer Science teacher assistant 2020.** As a teaching assistant for **undergraduate students**, you give classes or help sessions and check students' homework.

- **Voluntary/Social work in Olympiads 2015 and 2016.** I collaborated on evaluating contest tests and giving training workshops to participants of the **mathematics and informatics Olympiad**.

Math and CS Major Projects

Stenosis Detection (2021):

A project to detect arterial stenosis using an object detection neural network.
- YOLO pretrained network was used with the Roboflow tool for object detection.
- Images were preprocessed (clean noise and split into patches) to simulate examples of objects with and without stenosis cases.

Dual-Parabolic detection (2020):

A Project for detecting two parabolas models the arteries of a retina's images.
- Implemented in C++, based on ideas from the paper, Dual-Parabolic Modeling, to apply the Hough transform to images and detect parabolic shapes; then, we stay with the best.
- It is implemented through a cross-platform graphical interface in C++ with the help of the wxWidgets library.

Basketball Game (2019):

A virtual reality game where you can throw a ball and make as many baskets as possible.
- Implemented in unity using C# and mounted into oculus rift for virtual reality visualization.
- Design of the scene. I used physics for the ball, rim, and throwing movement in the virtual reality controller.

People detection (2018):

An image analysis using machine learning to detect if there is a person or not in the picture.
- I implemented it using the Keras library. I used a fine-tuning of a convolutional NN with pre-trained layers (Xception).
- Around 1000 images of people and 1000 no-person were used for training and evaluation.

Relative movement (2018):

A simulation of an object's movement using physics laws and differential equations.
- It was implemented in Wolfram Mathematica for ODE solving and movement simulation.
- I studied, simulated, and dynamically visualized the movement of Earth around the Sun, Moon around Earth, etc.

Consume of electrical energy (2018):

Data analysis of daily electricity consumption before and after the daylight-saving time change.
- Implemented in R., I analyze electricity consumption data from 5 states of Mexico, six months before the daylight-saving time and six months after.
- My motivation was to realize if daylight-saving time serves to save electricity or not.
- The analysis was to compare the means of each set of data. They were giving formal and robust evidence of different means.

Queue simulation (2017):

A collaborative project which simulates different kinds of queues analyzing their probability distribution.
- It was a project developed for the probability course by two partners and me.
- We simulate a company's production queue. In the company, there are machines of production that, with some time distribution, broke down, and they must be sent to maintenance.
- We must determine how many machines the company needs to have in reserve.

Damorgal's Battle (2015):

The battle is a 2-player turn-based game where players can select their character and a place with different gravity.
- We implemented it in C++ with a library named SHELL. Dr. Jose Luis Marroquin created the library.
- Library SHELL doesn't support audio, but I found a creative way to add music to my game.
- Game buildings appear randomly at the scene. Things that a player throws are constantly spinning. The path of the thrown thing depends on the force and angle (selected by the player with the mouse); it also depends on the gravity

of the place chosen. These special effects need much mathematics like the rotation equation or the thrown thing's path.

Programming Languages and Libraries

- **Eight** years: **C/C++**. **Two** years: **STL**.
- **Five** years: **Python**. **Three** years: **PyTorch**. **One** year: **NLTK**.
- **Three** years: **R**. **Six** months: **tidyverse.**
- **One** year: **Mathematica.**
- **Six** months: **MATLAB.**
- **Platforms:** Unix/Linux.

Interesting Courses

- Natural Language Processing (2021)
- Optimization (2021)
- Biomedical Image Processing (2021)
- Computer vision (2021)
- Machine learning 2 (applications to vision and natural language) (2020)
- Data analysis (2020)
- Pattern Recognition course (2020).
- Data Visualization and Computer Graphs (2020).
- Real-Time Programming (2020).
- Artificial Intelligence in embedded systems (2020). (Speech-recognition on a Raspberry Pi).
- Probability and statistical inference course (2017, 2018).
- Stochastic simulation course (2018).
- Machine learning 1 course (2018).
- Efficient and competitive programming course (2018).
- Computer science and algorithms course (2016).
- Algorithm analysis course (2016).
- Embedded programming course (2017).
- Numerical methods (2017).
- Karel/C++ problem-solving in the computer science Olympiad (2011, 2012, 2014).

Awards and Prizes

- A novel research project at the University of Michigan (2019).
- 1st. Place in the mini-conference of statistical research posters presentation (2018).
- 3rd Place in the National Mathematics Olympiad (2010, 2012).
- 2nd Place in the National Mathematics Olympiad (2013).
- 1st Place in the Guanajuato State Olympiad of Computer Science (2011, 2012, 2014).
- 3rd Place in the State high schools Physics Olympiad (2013), (Competition between all State high schools).
- 3rd and 2nd place team in the National Basketball Olympiad (2011, 2013, 2014).

Hobbies and sports

- I like to play sports, but I used to participate in basketball Olympiads. My team and I won the cities competition yearly, so I was on the State team to compete against the entire country. We won the national Olympiad once and 2nd and 3rd place several times.
- At the university, I also played basketball in universities competition. Basketball is a team sport, so I have been **comfortable working in a team** since I was a child.
- I enjoy playing piano, ping pong, painting, running, and creating electronic music.