My motivation for obtaining a PhD is to contribute to **research in large-scale machine learning (ML)**. More specifically, working at the intersection of gradient-based methods, probabilistic modeling, and derivative-free optimization, I want to develop new approaches for designing fast and efficient algorithms to perform automated ML and meta learning. My ambition after completing my PhD is to advance fundamental artificial intelligence (AI) research, apply it to real-world challenges and increase its approachability and usability globally.

My research interests revolve around investigating how we can adapt zero/first/higher order information to design modern methods efficiently suited for ML tasks. For instance, looking beyond stochastic gradient descent for training deep neural networks or developing optimization-based meta learning algorithms to automatically design deep learning architectures.

My research and industry experiences have well-prepared me for the challenging undertaking of graduate studies. In my undergrad thesis, I studied the main principles behind stochastic optimization. The goal was to develop novel algorithms for object tracking in video. I analyzed and implemented color-based particle filters and incorporated a dynamic optimization step to enhance the state estimation over time. I evaluated this approach in benchmarks and showed that the heuristic mechanism improved the accuracy of predictions with little computational cost added overall. However, this algorithm was not efficient enough for practical scenarios like real-time tracking. This experience helped me appreciate the difficulty of developing approaches that ideally advance in both efficiency and accuracy and also motivated my interest in acquiring different perspectives in optimization and learning. Now I am excited to continue fostering my knowledge in ML where similar trade-off challenges exist such as overparametrized models and neural architecture search, as in the case of deep learning.

In my master's thesis, I conducted research in black-box optimization with the goal of developing algorithms for continuous domains. I analyzed population-based optimizers built upon the framework of Information Geometry, a mathematical field that studies the geometric structure of families of probability distributions. Standard approaches for global optimization iteratively minimize a function’s expected value over a set of probability distributions and update the densities by following the natural gradient direction. Instead, I proposed to search for the optimum by minimizing the Kullback-Leibler divergence of the probability densities w.r.t. the Boltzmann distribution associated with the objective function. I chose a univariate gaussian model to reduce the complexity of computing the inverse Fisher information matrix and assembled this idea into a practical algorithm. I experimented with the Boltzmann temperature parameter to explore controlling the variance, and thus the convergence of the algorithm. However, it was too complicated to come up with an auto-adjustment procedure, due to a problem-dependent behavior. This difficulty fueled my passion to investigate self-adaptive mechanisms from first principles and contribute towards more automatic algorithms and hyperparameter-free ML models. Through this experience I also got a fascination for going deeper into mathematical areas such as information theory, geometry, and probabilistic modeling. During my PhD studies, I would like to analyze how to derive algorithms with theoretical guarantees by combining the complementary strengths of gradient-based optimization theory and simulation-based methods.

In addition, I have the goal of **applying these algorithms in practice**. In the financial industry, I built tools for pricing financial derivatives using stochastic models that fit observable market prices. In the tech sector, I designed models capable of processing telemetry in real-time. Both contexts required training large-scale models on terabytes of data, demanding computational resources, typically accessible only to major enterprises. These industrial-scale challenges helped me appreciate the importance of developing effective approaches that not only advance the state-of-the-art in ML but also contribute to making them widely accessible, democratizing its use on a global scale. Since this effort requires highly innovative work in a foundational field I believe the academic environment is the best place to conduct it. This experience catapulted my enthusiasm for pursuing a PhD.

My goal is to pursue a world-class education in the United States (US) by working in a machine learning group with a breadth of work in both theoretical and applied research. Computer science departments with a strong tradition of faculty members who maintain dual affiliations in academia and industry are of my particular interest. I find the structure of US PhD programs particularly well-suited to my learning style, as it closely aligns with the system I experienced at CIMAT, especially the balance between coursework and research. Relevant coursework will enable me to strengthen my understanding of core machine learning, then following years will allow me to focus on fully developing my proposal.

My experiences have reinforced my passion and commitment to pursuing a career as a researcher. With the support of the Fulbright-Garcia-Robles program, I aim to advance the next generation of ML methods and ensure their broad applicability.