JULIAN VENEGAS

(248) 660-5369 | venegas5@msu.edu | linkedin.com/in/jdevenegas | Detroit, MI

EDUCATION

PhD Computational Mathematics, Science & Engineering
Michigan State University
2018 – 2024

BA Mathematics
Oakland University
2011 - 2015

SUMMARY

Data Scientist & Machine Learning Engineer with extensive experience in the design, development and optimization of computational methods for complex systems & problems. My experience spans various domains, e.g., finance/risk, drug-discovery, network analysis, image & natural language processing. My focus is on integrative learning techniques (transfer-multitask-multimodal learning), graph learning & interpretable machine learning.

SKILLS

Programming || Python | SQL | R | C++ | MATLAB ||
Machine Learning || Deep Learning | PyTorch | TensorFlow | XLA | JAX | DGL | PyG ||
Linux-macOS-Windows || **HPC** || Shell Scripting | Slurm ||
Data Analytics & Engineering || KNIME | Power BI | Relational Database Systems | Microsoft SQL Server | MySQL | ETL | AWS | Google Cloud ||
Productivity & Product Management || DevOps | Git | Docker | Jira | Slack ||

COURSEWORK

Applied Deep Learning || Mathematics of Deep Learning || Mathematical Foundations of Data Science || Optimization (convex and non-convex) || Probability Theory & Statistical Inference || Statistical Models (Linear & Mixed || Numerical Linear Algebra

WORK & RESEARCH EXPERIENCE

Research Consultant - Machine Learning & Data Science
Institute for Cyber-Enabled Research - Michigan State University
East Lansing, Michigan | September 2024 – present

- Consult and support researchers in using advanced computing (supercomputer & cloud), data systems and software to perform large-scale computational & data science research across a broad range of scientific disciplines.
- Developed LLM platform applications & workflows for users of the high-performance computer (HPC). Includes a no-code user friendly GUI with fine tuning and Retrieval-Augmented Generation capabilities.
- Organized the MSU AI Community of Practice to share idea, tools, and support for researchers integrating AI & ML in their work.

Machine Learning Research, Cheminformatics Co-op
Blueprint Medicines | Cambridge, Massachusetts | May 2023 – April 2024

- Research & development of machine learning methods for property prediction tasks. Focus on graph-based learning methods, transfer learning, and uncertainty quantification.
- Developed & deployed an automated modeling pipeline, which continually re-trains, tests and deploys a graph-neural network used for internal chemical profile prediction and selectivity estimation.

Graduate Research | NSF Graduate Research Fellow
Michigan State University | East Lansing, Michigan | 2018 – 2024

- Machine learning research with a focus on integrative learning, i.e. transfer, multi-task and multi-modal learning.
- Graph-based and sequential learning methods with multi-modal data integration. Applications in bioinformatics, e.g., spatial/bulk deconvolution, cell annotation, and segmentation.

AWARDS

2020 **NSF Graduate Research Fellowship**, National Science Foundation

2019 **NSF NRT IMPACTS Fellowship**, National Science Foundation

2018 **Engineering Distinguished Scholar**
Michigan State University College of Engineering

2015 **Louis R. Bragg Graduating Senior Award**
Oakland University Department of Mathematics & Statistics

PROJECTS

- DANCE: Deep graph-based learning library and benchmark for single-cell analysis.
- chemprop-BP: Deep graph-based learning library for chemical property prediction tasks.
- DeepCAT-PlantBio: Deep convolutional attention model for predicting plant stress response from raw DNA sequences.

Quantitative Research Intern
Wells Fargo | San Francisco, California | June 2021 – September 2021

- Research & development of monotonic and interpretable machine learning methods for operational, market, and credit risk modeling.
- Carried out a benchmarking experiment of monotonic-interpretable methods on internal & external credit risk data. Wrote generalized benchmarking scripts for continual assessments of internal models.

Engineer Analyst – Data Reporting & Analytics
Comerica, Inc | Auburn Hills, Michigan | June 2016 – June 2018

- Gather requirements, design, and implement models, data visualizations, and reports to enable data-driven decisions from middle-management to executive levels.
- Develop & manage data pipelines: loading, cleaning/pre-processing, transforming/summarizing and loading for end-users & analysis.

Data Analyst & Consultant
Urban Science Applications, Inc.| Detroit, Michigan | June 2015 – June 2016

- Employ statistical methods and operations research to optimize car dealership and servicing networks, operational capacity, and inventory levels.
- Develop & manage data pipelines.

Product Development Research Intern
PTC, Inc | Troy, Michigan | June 2014 – May 2015

- Research & develop system dynamics models of business and product development processes. Assist business transformation managers in business process, and product development strategy.
- Co-authored & published two papers with supervisor on dynamic systems modeling, simulation and analysis of product development & business processes.

PUBLICATIONS

- *SpatialCTD: a large-scale TME spatial transcriptomic dataset to evaluate cell type deconvolution for immuno-oncology*, Journal of Computational Biology, September 2024
- *DANCE: A Deep Learning Library and Benchmark Platform for Single-Cell Analysis*, Genome Biology, March 2024
- *Deep Learning in Single-Cell Analysis*, ACM Transactions on Intelligent Systems and Technology, March 2024
- *Changes in salivary biomarkers associated with periodontitis and diabetic neuropathy in individuals with type 1 diabetes*, Nature - Scientific Reports, June 2022
- *Composite modeling of leaf shape across shoots discriminates Vitis species better than individual leaves*, Applications in Plant Sciences, December 2020
- *Simulation-Based Process Reliability Design,* Society of Automotive Engineers World Congress, Detroit, MI, April 2015
- *Insights Into Process Reliability Through Simulation*, Reliability and Maintainability Symposium Conference, Palm Harbor, FL, January 2015