Dr. Yannik Pitcan

Contact Information

+1 407 267 8890
pitcany@outlook.com
pitcany.github.io
217 Delfino Ave. Richmond, CA 94801

Education

University of California, Berkeley, Berkeley, California, United States

Ph.D. in Statistics

- Dissertation: An Assortment of Analyses for Optimal Transport Inspired by Domain Adaptation.
- Proved two new sample complexity bounds on the generalization errors of statistical methods using techniques from empirical process theory and proposed another type of distance measure called Sliced Multi-Marginal Wasserstein.

Harvard University, Cambridge, Massachusetts, United States

A.B. ***cum laude*** *with Honors in Mathematics*

Professional Experience

Google, Mountain View, California

Data Scientist **2024 – Present**

- Applied machine learning and statistical modeling on the advertising team.

PerformanceStar LLC, Santa Clara, California

Contract Machine Learning Software Engineer **2022 – 2023**

- Head of the XAI (explainable AI) project that assists humans in understanding and interpreting the decisions made by AI.
- Spearheaded the end-to-end development and successful implementation of a Python-based regression library for XAI.
- Resulted in a 25% improvement in model interpretability and supporting data-driven business decision-making processes.
- Engineered an advanced technique for estimating Shapley values in LSTM and RNN time series models, affecting predictive model interpretation by uncovering temporal feature dynamics and facilitating improved decision-making.

Walgreens Boots Alliance, Chicago, Illinois

Senior Algorithms and Machine Learning Scientist **2020 – 2021**

- Developed advanced algorithms, including XGBoost and deep learning models, to revolutionize product development; increased revenue by 25% through increased sales forecasting and customer segmentation.

- Created feature store that reduced time other data scientists across the company spent on feature engineering by 30% based on metrics.
- Migrated model pipelines to distributed cloud clusters within Azure Databricks using Pyspark, resulting in a remarkable 10x increase in the workload capacity of the 'Return to Stock' project. This achievement earned company-wide recognition for the pivotal role played in the project's success.
- Overhauled hiring procedures and established strategic partnership with UC Berkeley to recruit underrepresented minorities in STEM fields for roles at Walgreens. This initiative resulted in a 20% increase in diverse STEM talent acquisition.

Notable Projects

Hidden Markov Models for Stock Return Analysis

Constructed two HMMs to model the stock returns for every 10-day period. First model used the Baum-Welch algorithm for inference about volatility, which regards volatility as hidden states and uses a mean zero Gaussian distribution as the emission probability for the stock returns. Second model uses a spectral algorithm to perform stock returns forecasting. Analyzed the tradeoffs of these two implementations as well.

Skills

Programming Languages

R, Python, SQL, C/C++, Java, Matlab, Scala, Mathematica

Frameworks and Platforms

AWS, Azure, PyTorch, MlFlow, Spark, Scikit-Learn, TensorFlow

Methodologies

Bayesian Inference, Time Series Analysis, Causal Inference, Deep Learning, Large Language Models

Other Activities

Tutored more than fifty students in mathematics and statistics at all levels including graduate level.