Animesh Joshi

858-226-7898 | animeshjoshi88@gmail.com | LinkedIn | GitHub

Education

Purdue University, West Lafayette — West Lafayette, Indiana

Bachelors of Science in Mathematics and Statistics, Certificate in Applied Data Science — *Expected Dec 2025*

Courses: Statistical Theory*, Operations Research and Optimization*, Signals and Systems*, Probability Theory, Regression Analysis, Statistical Machine Learning, Programming in C, International Relations, U.S. Foreign Policy, Ethics of Data Science

Research and Internship Experience

Data Science Research Assistant — Sep 2024 - Present

Responsible Data Science Lab; Purdue University — *West Lafayette, In*

- Conducting research on responsible data management and algorithmic fairness in machine learning systems
- Developed logistic regression classification system for credit rating scores and computed 9 algorithmic bias metrics (equalized odds, fairness through awareness, counterfactual fairness, etc.) to test the fairness of the system
- Implementing estimations of influence functions and shapley values to quantify the marginal bias of training data points on overall bias in the system

Credit Analytics Intern — Jun 2024 - Aug 2024

Credit Analytics Department; Axos Bank — *San Diego, Ca*

- Compiled datasets of 4,000+ loan originations and 37,000+ active portfolio loans and built trended summary tables to track origination and asset quality risk metrics (FICO, Risk Rating, Utilization, etc.) using T-SQL queries
- Visualized trends in credit risk metrics for over $12.5 billion in loan originations and over $20 billion in active portfolio loans and analyzed trends to determine drivers of over $500 million in delinquent loan balances using Tableau reports and dashboards
- Maintained accurate analytics reports used for credit risk monitoring, regulatory reporting, and portfolio management using Microsoft SQL Reporting Services

Other Roles: Machine Learning Research Assistant at CLAN Lab, Computer Vision Researcher at The Data Mine

Projects

Pass Utility Optimization via Utility Functions

NFL Big Data Bowl 2025

- Constructed utility function to measure expected reward of a pass play at each point in time and computed marginal utility of each player involved in the play
- Automated data analysis pipeline which transformed player tracking data into structured datasets with spatial variables to enable prediction of pass completion probability and yardage gain after the catch
- Built interactive data visualizations to show play animations, trends in play utility over time under different coverage schemes, trends in marginal utility of a position over time under different coverage schemes, and to validate predictive power of the utility metric

Recidivism Risk Forecasting

National Institute of Justice Data

- Collected data from the National Institute of Justice (NIJ) and performed exploratory data analysis and data cleaning to build a dataset of prisoner attributes and recidivism status
- Deployed statistical learning methods (Random Forests, Gradient Boosting, Regression) to forecast the probability of recidivism for a prisoner given attributes measured at time of incarceration
- Validated model results using cross validation and classification performance metrics and explored predictive power of each prisoner attribute to find drivers of recidivism

Technical Skills

Programming Languages: Python, R, SQL, Matlab
Tools: Jupyter, R-Studio, Tableau, Microsoft SQL Server, Git, VSCode, Microsoft Excel, Plotly, Linux
Libraries: Pandas, Numpy, Matplotlib, Sci-Kit Learn, Statsmodels, Tensorflow, Scipy, Gurobi, Dplyr, MASS, GGPlot
Competencies: Inferential Statistics, Probability, Statistical Modeling, Regression, Machine Learning, Operations Research, Linear Programming, Mathematical Modeling, Computational Mathematics, Scientific Programming, Epidemiology, Causal Inference, Econometrics, Optimization