

NASHEED JAFRI

(260) 508-9096 | njafri@iu.edu | Webpage: nashjafri.github.io | linkedin.com/in/nasheedjafri

EDUCATION

Ph.D. Mathematics (Ph.D. Minor in Data Science) <i>Indiana University, Bloomington, IN</i>	GPA : 3.98 2020 - 2026
M.S. Applied Statistics <i>Indiana University, Bloomington, IN</i>	GPA : 4.00 2024 - 2026
M.S. Mathematics <i>Indian Institute of Technology, Delhi, India</i>	GPA : 3.70 2018 - 2020
B.S. Mathematics (Honors) <i>University of Delhi, India</i>	GPA : 3.82 2014 - 2017

PROJECTS

- Home Credit Loan Default Prediction** [\[Link\]](#) [\[GitHub\]](#)
- Collaborated with a team of data scientists to predict loan defaults for Home Credit using Machine Learning techniques in Python, achieving 92% test accuracy with XGBoost.
 - Conducted EDA, handled missing values, performed correlation analysis and feature engineering.
 - Built and optimized models (Logistic Regression, Decision Trees, Random Forests, Gradient Boosting, XGBoost, SVC and MLP neural networks) using PCA and ensemble methods like voting and stacking.
- Approximate Bayesian Computation for Disease Outbreak** [\[Link\]](#) [\[GitHub\]](#)
- Implemented Approximate Bayesian Computation in R to fit an epidemic model for influenza outbreaks.
 - Created custom functions for parameter sampling, data simulation, and ABC rejection sampler algorithm.
 - Performed model comparisons by estimating posterior probabilities to analyze variations in infection transmission rates across outbreaks of the same strain and different strains of the virus.

EXPERIENCE

- University of Illinois, Urbana - Champaign, IL* 2023 - 2024
- Data Science Trainee (Internship Network in Mathematical Sciences)**
- Completed training workshops on Python, Statistics and Machine Learning.
 - Analyzed housing data from Redfin, applied linear regression to predict house prices across different cities.
 - Performed EDA, hypothesis testing, and employed forward, backward, and stepwise model selection techniques to predict wine quality based on physicochemical attributes from the Vinho Verde wine dataset.
- Indiana University, Bloomington, IN* 2021 - 2026
- Associate Instructor in Linear Algebra for Data Science**
- Mentored students in key subjects including vector spaces, matrix operations, eigenvalues, least squares, singular value decomposition (SVD), and principal component analysis (PCA).
 - Planned engaging weekly group learning activities within the class setting.
- Assistant Instructor in Probability and Statistics for Data Science**
- Developed curriculum to apply data analysis to real-world problems in social and natural sciences.
 - Taught conditional probability, random variables, distributions, statistical inference and hypothesis testing.
- REU Mentor**
- Supervised a group of undergrad students in a graduate-level research project on Fourier Transform.
- Research**
- Conducting research in Functional Analysis with a focus on Linear Algebra and Matrix Theory, specifically studying the uniqueness of invariant subspaces for a class of nilpotent matrices.
 - Exploring applications of combinatorics to matrix theory via structures called puzzles and honeycombs.

SKILLS

Mathematics: Probability, Statistics, Linear Algebra, Differential Equations, Numerical Analysis, Graph Theory, Functional Analysis, Fourier Analysis, Econometrics, Markov Chains, Monte Carlo Simulations, Combinatorics

Data Science: Data Exploration, Linear Models, Inference, Regression, Classification, Predictive Models, Machine Learning Algorithms, Principal Component Analysis, Statistical Computing, Bayesian Inference

Programming: R, SQL, Python - Pandas, Numpy, Scikit-learn, TensorFlow, PyTorch; Git Version Control