

Anurag Mallick

Phone : +1 732-801-8730
Email: anurag.mallick@rutgers.edu

Linkedin: linkedin.com/in/anurag-mallick
Github: github.com/mallick20

WORK EXPERIENCE

• Applied Data Finance

2021 - 2024

Senior Data Scientist

- Utilized NLP Techniques to analyze and mitigate operational risks across various modes of communication. Developed models to identify risks with **70% precision** to make the process efficient.
- Built Topic Models using BERTopic to segment customer requests. Effectively grouped **up to 80% of customer requests** into Valid Topics for more efficient resolution.
- Led the foundation for improved resource management by categorizing and prioritizing cases for manual review, leading to more efficient review and response times.
- Worked on finetuning LLM for better classification of customer complaints.

Data Scientist

- Built a classifier model using customer's repayment behavior to predict the success rate of payment retries and successfully segmented low-risk accounts, having a **37% higher success rate** than average.
- Enhanced the Best Time to Call strategy to call customers according to their preferences and call history. Achieved a mean **increase of 20% RPCs** over six months.
- Conducted in-depth analysis of returning customers to fix defects in the data and explore trends according to issues in disbursal plans.
- Designed campaigns to engage delinquent customers and used statistical analysis to test the outcomes among various segments and improve collections.

EDUCATION

Masters in Data Science

- Rutgers, The State University of New Jersey 2026
Courses: Prob. & Statistical Inference for Data Science, Regression and Time Series, Data Structure and Algorithm,

B.E in Electrical and Electronics Engineering

- Birla Institute of Technology and Science, Pilani 2021

RESEARCH EXPERIENCE

• Real-Time Tweet Classification

BITS Pilani

2021

- Implemented the foundational framework for classifying tweets based on their Non-Profit status.
- Pre-processed raw tweets using regular expressions and developed the dataset processing pipeline.
- Used BERTweet to build and integrate the classification pipeline's Level 1 (L1) and Level 2 (L2) stages.

• Indian Language Machine Translation

IIT Roorkee

2020

- Utilized deep learning framework to improve the accuracy of machine translation for Indian Languages.
- Implemented word-level translation of OOV words, achieving a **BLEU Score of 73.5%**.

SKILLS

- **Data Analysis and Visualization**
- **Machine Learning:** Linear Regression, Logistic Regression, Gradient Boosting
- **NLP:** Text Processing, Document Classification, Topic Modelling, LLM Finetuning
- **ML Libraries:** Pytorch, Scikit-Learn, NLTK, Spacy
- **Technical Skills:** Python, R, C++, SQL, Redshift, Git, Microsoft Office Tools
- **Soft Skills:** Concise Presentation, Time, and Resource Management

AWARDS

- Secured 2nd position at ADF Data Science Hackathon – 2022
Built an XGBoost Model to predict customer risk using Tradeline Attributes.