Anurag Mallick

Phone : +1 732-801-8730   Linkedin: linkedin.com/in/anurag-mallick
Email: anurag.mallick@rutgers.edu   Github: github.com/mallick20

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, Statistical Learning for Data Science

B.E in Electrical and Electronics Engineering
- Birla Institute of Technology and Science, Pilani   2021

  Courses: Operating Systems, OOP, Optimization, Data Mining

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 with **70% accuracy** to make the process efficient.
  - Built Topic Models using BERTopic to segment customer requests. Effectively grouped **upto 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.

  Data Scientist
  - Conducted in-depth analysis of returning customers to fix issues in the data and explore trends in disbursal plans.
  - Built a classifier model using customer's repayment behavior to predict the success rate of payment retries and successfully segmented low-risk accounts having **37% higher success rate** than average.
  - Designed campaigns to engage delinquent customers and used statistical analysis to test the outcomes among various groups and improve collections.
  - Enhanced the Best Time to Call strategy to call customers according to there preference and call history. Achieved a mean **increase of 20% RPCs** over a period of 6 months.

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 for the project.
  - Used BERTweet to build and integrate Level 1 (L1) and Level 2 (L2) stages of the classification pipeline.

- **Indian Language Machine Translation**

  IIT Roorkee   2020
  - Utilized deep learning framework to significantly 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:** Data Cleaning, Statistical Tests
- **Machine Learning and NLP:** Linear Regression, Logistic Regression, Gradient Boosting, Sentence Transformers
- **ML Libraries:** Pytorch, Scikit-Learn, NLTK, Spacy
- **Technical Skills:** Python, R, SQL, Git, Microsoft Office Tools
- **Soft Skills:** Concise Presentation, Time and Resource Management, Communication

AWARDS

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