Sanket Zanwar

201-238-9319 | snaketzanwar1128@gmail.com | Atlanta, GA | LinkedIn | GitHub

Education

Stevens Institute Of Technology, USA — **May 2024**
Master Of Science in Data Science — **CGPA: 3.99/4.0**

National Institute of Technology, Goa, India — **June 2022**
Bachelor of Technology in Electronics and Communication Engineering — **CGPA: 9.35/10.00**

Skills

Programming Languages: Python, R, PySpark, SQL, C++

Packages and Tools: GitHub, Docker, Power BI, Tableau, Pytorch, Databricks, Hadoop, Spark, Hive, AWS (S3, EC2, IAM, Data Lake, Auto ML, Athena, Sagemaker), Google Cloud(BigQuery), TensorFlow, Pandas, Numpy, GitLab, Keras, Excel, Salesforce, Redshift, Streamlit, Matplotlib, Yellowbrick

Experience

Data Scientist, Cox Communications, Atlanta, (GA) *#ML #MLOPS* — May 2023 - Ongoing

- Spearheaded the development of propensity-to-buy models for the Next Best Action product, increasing target customer identification by 64% and acquiring 14,400 new customers within 12 months, surpassing initial 6-month expectations
- Improved predictive accuracy by 13% by leveraging ANOVA and Odds Ratio to identify the top 100 influential variables from 7,500+ features, enabling targeted campaigns that enhanced customer engagement and boosted sales
- Enhanced model reliability by 18% (95% CI) through A/B testing and calibration of XGBoost and CatBoost algorithms, achieving a 31% reduction in log-loss and saving over 12 hours in manual recalibration efforts
- Increased Market Share(MS) by 8.7% YoY by analyzing competitor strategies, customer flow leading to a 16% improvement in customer acquisition and $2.5M in annual revenue growth
- Optimized customer retention by 7% in high-competition markets by designing A/B testing frameworks for MS campaigns, resulting in $4.8M in saved revenue and strengthening customer loyalty through engagement strategies
- Reduced customer churn by 14% by analyzing complex churned customer datasets using AWS Athena and delivering actionable insights via AWS Quicksight that improved campaign ROI by 22% and enhanced customer satisfaction
- Developed a monitoring system to track model performance, data quality, and drift over time, reducing model downtime by 40% and automating alerts that saved 15 hours per week in manual monitoring, ensuring consistent model accuracy
- Streamlined model deployment and team collaboration by implementing Git version control and CI/CD pipelines, reducing deployment time by 30% and improving overall team productivity by 20%, while ensuring seamless integration of machine learning models

Data Science Intern, Dataperformers *#ETL #Data Engineering #EDA* — April 2021 - June 2021

- Engineered Python pipelines optimizing data retrieval from AWS S3 and Snowflake; integrated class functions to enhance modularity. Resulted in a 25% memory usage reduction, enabling user-ordered data views for purchasing analysis and price anomaly detection in shopping carts.
- Boosted the performance of the product recommendation analytics engine by 23% through targeted feature engineering, along with implementing an advanced customer segmentation strategy resulting in an improvement of 18% in balanced accuracy metrics with the aid of the MLflow library for log-loss error tracking.

Academic Projects

Non-Seasonal GARCH Time Series Analysis (Link) *#Forecasting* — January 2023 - April 2023

- Developed a highly accurate Cardano cryptocurrency price prediction model utilizing the GARCH method, achieving a mean absolute percentage error (MAPE) of 3.8%, exceeding industry benchmarks by 8%
- Applied advanced statistical techniques, including ARIMA, GARCH, and ACF/PACF analysis, to effectively model and forecast the volatility and collinearity of cryptocurrency market data with a confidence level of 95%

Quora Question Pair Similarity (Link) *#NLP* — November 2022 - April 2023

- Implemented transfer learning techniques by fine-tuning the BERT model for question pair classification, leveraging its powerful contextualized word embeddings and the Hugging Face's Transformers library
- Achieved an accuracy of 76% surpassing the baseline score by 6%, highlighting the effectiveness of fine-tuning BERT

Leadership Positions

Teaching Assistant and Grader for MA541: Statistical Methods — January 2023 - May 2023

- Provided comprehensive instruction on fundamental statistical concepts, including p-values and hypothesis testing, during weekly review sessions.