Ayush Oturkar

Linkedin: https://www.linkedin.com/in/ayushoturkarnitb/
Mobile: +1-551-344-8094
Email : ayush.oturkar@rutgers.edu
Github: https://github.com/Ayush1695
Location : Jersey City, NJ, United States

Summary: Seasoned Data Scientist with 5 years of experience in Predictive Modeling, Machine Learning, Advanced Analytics and Statistics. Proven track record in driving growth, revenue, and efficiency across diverse industries

Experience

- **Rocket Central** — Detroit, MI, United States
  Data Scientist - Summer'23 — *May 2023 - August 2023*
  - **Credit Card Propensity Model**: Built a predictive propensity Xgboost model to target the right audience for the company's flagship CC, aimed in reducing marketing costs and boosting revenue with targeted 2% EBIDTA improvements.
  - **Opt Out Modelling**: Engineered and executed a highly efficient end-to-end data pipeline on SQL Server to limit phone opt-out rates through explainable Logistic Regression Pipeline, aiming for significant improvements in customer retention.
- **Anheuser-Busch InBev** — Bangalore, India
  Data Scientist II - Permanent Full Time — *May 2021 - August 2022*
  - **Duplicate Payment Detection**: Designed and deployed a scalable duplicate payment detection system using NLP rules and the LightGBM model, achieving a 90% reduction in false positives & resulting in over $1 million in EBIDTA savings.
  - **MLOps & Solution Scaling**: Led end-to-end project delivery and scaled solutions using the MLOps lifecycle. Improved data visualization by leveraging Power BI and Tableau for impactful visual storytelling and consumption by stakeholders.
  - **Upsell Recommendation**:
    - Engineered an intelligent CX upsell Recommendation System using Collaborative Filtering, Apriori lift, and business rules for one of the biggest consumer clusters, leveraging volume sales data to replace the existing approach.
    - Led to a significant market share uplift, with a 2% boost in top-line sales during A/B testing post experimental design.
- **Fractal Analytics** — Bangalore, India
  Data Science Consultant - Permanent Full Time — *September 2019 - May 2021*
  - **Discount Optimization & Yield Management**:
    - Built an S&S discount pricing simulator for pet products on eCommerce platforms, including Amazon, by leveraging different ML models like LGBM, Random Forest, achieving a strategic 2.5% increase in top-line sales for fortune 50 client.
    - Orchestrated the solution by analyzing over 1 TB of Big Data using PySpark through Azure Data Factory. This streamlined operations, reducing manual efforts by nearly 99%.
  - **Pack Segment Modelling**: Developed an automated Bayesian Structural Time Series modeling pipeline using R over the client's pack segment Retail data, driving a $2 million revenue boost through strategic optimization in pricing & marketing.
  - **Automation**: Streamlined Excel-based business reports using Python and Automation Anywhere, eliminating manual efforts by 100% and reducing refresh time from weeks to minutes.
- **Trinity Life Sciences** — Bangalore, India
  Business Analyst - Permanent Full Time — *July 2018 - September 2019*
  - **Prescribers Decliners Prediction**:
    - Spearheaded the development of a Binary Classification model, achieving a 79.8% f1-score in detecting Decliners.
    - Mitigated a potential 10% decline in prescription volume by analyzing prescriber behavior and attributes, resulting in an additional 3% revenue growth for fortune 50 healthcare client.
  - **Patients Unmet Needs Estimation**: Implemented a text mining/NLP technique using LSTM to derive patient-level insights from cancer support forums, addressing over 20% of patients' unmet needs.
  - **Adhoc**: Collaborated closely with cross-functional teams to improve existing solutions and facilitate code migration, leading to a notable 5-10% improvement in overall efficacy.
- **Indian Institute of Technology (IIT-B)** — Mumbai, India
  Research Intern — *May 2017 - July 2017*
  - **Driver Risk Assessment over Horizontal Highway Curves**: Deployed unsupervised hierarchical clustering algorithm to assess driver demography and road geometry profile for research publication in Transportation Journal.

Education

- **Rutgers University** — New Brunswick, NJ
  Masters of Science in Statistics-Data Science; GPA: 3.857 — *September 2022 - December 2023*
- **Maulana Azad National Institute of technology (NIT)** — Bhopal, India
  Bachelors of Technology in Civil Engineering; GPA: 3.3 (7.83/10.0) — *August 2014 - April 2018*

Kaggle & Academic Projects

- **CIFAR Image Classification**: Designed a robust deep CNN image classification architecture using torch to accurately predict multiple classes from the CIFAR dataset, reaching a validation accuracy greater than 97%. Link
- **Credit Card Fraud Detection**: Performed EDA on a credit card transaction dataset, analyzing trends and developed optimized models (hyperopt optimization) with an F1-score exceeding 70% for fraud detection. Link
- **Diabetics Detection & Explainability Using SHAP**: Implemented a multi-class classification Xgboost model on publicly available diabetic data. Advanced techniques, including SHAP, for comprehensive local and global explainability, revealing underlying model mechanisms. Link
- **Starcraft Gamers Analytics**: Constructed a robust Random Forest model to predict player ranks in Starcraft by analyzing gamers' attributes. This model provides valuable insights for coaches and talent scouts by identifying essential qualities influencing players' placement in their respective leagues. Link
- **Spotify Song Popularity Analysis**: Conducted analysis of Spotify song popularity utilizing statistical models such as OLS, Ridge, and Lasso. Employed hypothesis testing to gain insights into the relationships among diverse song attributes. Link

Skills Summary

- **Core Domain Experience**: Data Mining, Applied Machine Learning & AI, Data Analytics, Deep Learning, Statistics, MLOPs
- **Languages**: Python, R programming, Apache Spark, SQL
- **Framework**: PyTorch, Tensorflow, Pandas, Numpy, Causal Inference, Scikit-learn, ML flow, Data pipelines, ETL
- **Tools**: Tableau, PowerBI, Azure, AWS, Redshift, MS SQL Server, Databricks, Git, JIRA, Devops Boards
- **ML Model Expertise**: Linear Regression, Logistic Regression, SVM, Naive Bayes, Decision Tree, KNN, Random Forest, Bayesian Regression, Adaptive Boosting, XGBoost, Clustering, Neural Networks, KMeans, CNN, RNN, Transformers LLM
- **ML Metrics**: AUC, Accuracy, Precision, Recall, F1, MAPE, SMAPE, RMSE, AIC, BIC
- **Soft Skills**: Intensive client-facing experience, Operations Research, Strategic thinking, Agile

Honors and Awards

- Ranked amongst the top 1% on Kaggle as a Kaggle Notebook Master
- Secured 1st prize in a Monthly Data Science competition amongst 100 participants by developing a robust fracture detection model at Trinity Life Sciences, improving accuracy by 10%
- Secured Rank amongst the top 1% amongst more than 1 Million JEE mains applicants in 2014