Shrestha Kumar Agarwal

(+91) 9193911757 | shrestha.k.agarwal@gmail.com | linkedin.com/in/shresthakrag | github.com/shresthakrag

EDUCATION

Sri Sri University, Odisha, India — ***July 2021 – April 2025***

Bachelor of Technology (Computer Science with specialization in Data Science) — *9.3 CGPA*

EXPERIENCES

Jio Saavn, Saavn Media Private Limited — ***Hybrid, Navi Mumbai, India***

Engineer Intern — *May 2024 – July 2024*

- Leveraged SQL to process and analyse large datasets, streamlining insights into key online metrics.
- Improved homepage personalization by analyzing feature distributions and optimizing top-k recommendation modules, enhancing relevance for new and returning users.
- Conducted feature distribution analysis by comparing training and inference data across different modules, identifying significant anomalies in historical impressions and user interactions.
- Developed and automated performance anomaly detection using z-score techniques, reducing manual effort and improving real-time metric monitoring by an estimated 30%.
- Created automated anomaly reports and visualizations for stakeholders, automating daily performance reviews and flagging irregularities in real-time metrics, reducing report generation time by ~60% and enabling quicker decision-making by 30%.

PROJECTS

Sentiment Analysis – Amazon E-commerce Product Review *[GitHub]*

- Developed a sentiment classification system to identify positive, neutral, and negative reviews.
- Engineered a robust preprocessing pipeline (tokenization, lemmatization, TF-IDF vectorization) and addressed class imbalance using SMOTE.
- Built and evaluated multiple classifiers (Logistic Regression, SVM, Random Forest,K-means); Logistic Regression achieved 95.27% accuracy.
- Built deep learning models (Bi-LSTM) for sentiment analysis of e-commerce product reviews to understand user sentiment and feedback patterns, attaining 82% accuracy.
- Used Python to implement, tune, and validate models with cross-validation, ROC, and confusion matrix analysis.

Hybrid Recommendation System – Article/Book Personalized Recommendation *[GitHub]*

- Designed a system to deliver personalized book recommendations by addressing limitations in traditional recommendation methods, improving accuracy and response times, and scaling to support large user bases.
- Integrated user and book datasets as sources, performed data preprocessing and feature engineering, applied matrix factorization techniques (SVD for CF, embeddings for CBF), and engineered a hybrid recommendation system achieving a 10–15% improvement in accuracy.
- Reduced response times to 200ms per query and scaled the system to handle 10,000 users efficiently.
- Validated the system by outperforming baseline models on RMSE (0.88) and MAE (0.65) metrics.

Online Fraud Detection (Credit Cards) *[GitHub]*

- Built a system to identify fraudulent transactions and minimize financial losses for credit card companies.
- Developed ML models (Random Forest, XGBoost) for real-time fraud detection in credit card transactions, achieving 93% precision to reduce financial risk.

Employee Attrition Prediction *[GitHub]*

- Predicted employee turnover to help organizations identify high-risk employees and address attrition drivers.
- Analyzed organizational workforce data, applied machine learning models (XGBoost, AdaBoost) for prediction, achieving 82% AUC accuracy, and visualized insights via Power BI dashboards.

TECHNICAL SKILLS | TOOLS | PLATFORMS

- **Languages/Libraries/Frameworks:** Python, SQL, NumPy, Pandas, Matplotlib, Seaborn, Plotly, Scikit-learn, TensorFlow, PyTorch, Keras, NLTK, Hugging Face, Natural Language Processing (NLP)
- **Concepts:** Data Pipelines, Statistical Hypothesis Testing, Business Metrics Optimization, Agile Methodology, Machine Learning, Deep Learning, Recommendation Systems, Ranking Algorithms, Time Series Forecasting, Fraud Detection, Sentiment Analysis, Product Optimization, Demand Prediction, Cloud Platforms (GCP, AWS)
- **Model Evaluation:** AUC-ROC, Confusion Matrix, Precision/Recall, F1 score, RMSE, MAE
- **Tools:** Git, Google Colab, Jupyter Notebooks, VS Code, Google Sheets, MS Excel, Power BI, Tableau, MySQL

CERTIFICATES

Data Analytics using Excel and Tableau (Deloitte), Data Science: Visualization (HarvardX), Data Analytics on AWS, IBM - Introduction to Data Science, Machine Learning & Data Science Program (GFG), Empowering Business with Effective Insights using Power BI (TATA Group)