SHREERAM VENKATESH

svenkateshiitc@gmail.com | **Chicago, IL** | 401-541-1455 | LINKEDIN | GitHub | PORTFOLIO

PROFESSIONAL EXPERIENCE

Data Science Intern, Labelmaster, Chicago, IL, USA **Jan 2024 – May 2024**

- Spearheaded a team of **6** to achieve **95%** accuracy in a real-time **Machine Learning** (ML) **price guidance system**, leveraging **7** years of client data (over **10 million** records) enhancing revenue
- Implemented **statistical hypothesis testing** and **Unsupervised/Supervised Learning (Clustering and Regression algorithms)** driving increased customer engagement by **20%** and boosting annual revenue by $X million
- Enabled decision-making with insights via a **Power BI** dashboard and **Streamlit** UI, improving operational efficiency by **25%**
- Collaborated with Pricing and Analytics team to validate ML models and integrate the system, resulting in a **12%** revenue increase
- Analysed data for key performance indicators (**KPIs**), driving targeted business strategies, increasing customer retention by **11%**

Graduate Teaching Assistant, Illinois Institute of Technology, Chicago, IL, USA **Aug 2023 - May 2024**

- Facilitated discussions during TA office hours, graded assignments, and provided feedback on SQL, MySQL, PostgreSQL, DBMS, OLAP, Excel and Python for Database Organization course, boosting student engagement by **40%**

Data Analytics Engineer, KPIT Technology Limited, Bangalore, India **Feb 2020 - May 2022**

- Orchestrated the streamlining of **ETL** processes in AUTOSAR BSW and RTE layer by applying **AWS EMR, S3, Glue, Lambda, Apache Spark, and SQL,** enhancing data synchronization and system performance by **80%**
- Deployed Machine learning models to process AUTOSAR diagnostic data over X gigabytes, achieving a **90%** boost in diagnostic accuracy. Leveraged **Tableau** for creating dashboards, empowering stakeholders with actionable insights, and accelerating delivery
- Transformed AUTOSAR Ethernet Time Synchronization by crafting intricate Python scripts with over **2000** lines of code, leading to a **60%** increase in project automation and facilitating orderly delivery of deliverables
- Consolidated data from CAN, LIN, and Ethernet teams, ensuring quality and consistency with **AWS Glue** and **Redshift**, resulting in a significant reduction in integration time
- Conducted A/B tests to optimize features, modules, and products/tools, resulting in a **24%** increase in user engagement
- Engaged in **Agile** and **Scrum** methodologies, optimizing project management efficiency through adept utilization of **JIRA**

Intern, KPIT Technology Limited, Bangalore, India **Mar 2019 - Jan 2020**

- Developed an end-to-end pipeline to drive ADAS (LKAS, ACC, TPMS, ABS) insights at KPIT, employing Azure Data Lake, **SQL, Python, Databricks, Power BI**, and **Teradata**, resulting in a **90%** increase in customer satisfaction
- Coordinated projects using vehicle telemetry data, integrating **XG Boost, Clustering** with **Azure Databricks** and **SQL Server**, to boost predictive accuracy by **60%**
- Defined KPIs aligned with stakeholder requirements and business objectives, using **Tableau** to create interactive reports and dashboards, resulting in a 25% improvement in data visualization effectiveness and informed decision-making

PROJECTS

E-Commerce-Market Analysis

- Analyzed cosmetic e-commerce data (**10 MIL** rows) in R, identifying key user behavior patterns and popular product categories
- Applied **K-means, RF, PCA, cross-validation**, and **hypothesis testing** to build accurate predictive models (**R-squared: 0.945, 0.847**)
- Presented the findings from EDA and feature engineering in a concise **PowerPoint**, aiding in decision-making for stakeholders

Sentimental Analysis of COVID-19 tweets

- Leveraged **Natural Language Processing** techniques to analyze sentiments expressed in tweets related to the COVID-19 pandemic
- Trained machine learning models including **XGBoost, Random Forest, SVM,** and **Stochastic Gradient Descent,** achieving **86%** accuracy with **SGD**

Netflix Movie Recommendation System

- Crafted a Netflix Recommendation System with **KNN, SVD, and SVD++,** yielding **60%** higher user satisfaction and retention
- Achieved **RMSE (0.8916 - 0.9995)** and **MAE (0.6859 - 0.7798)** on Movie dataset, optimizing Hit Rate with content and collaborative filtering

Telecom Churn Rate Data Warehousing with AWS Redshift

- Implemented **AWS Redshift** and Machine Learning (ML) algorithms to predict telecom churn rate with an accuracy score of **85%** (**AWS Redshift**, **SQL**, **ML**, **EDA**, **XGBoost, Python, boto3 SDK**

TECHNICAL SKILLS

Programming Language - Python, C++/C Programming, R, Embedded C, SQL, Bash, DAX, Java, MongoDB, T-SQL
Cloud Platform- AWS (S3, Redshift, Sagemaker, Lambda), Azure SQL, Data Lake, Databricks, Hadoop
Frameworks & Libraries- NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow, PySpark, Keras, PyTorch, NLTK, Seaborn, ggplot2
Tools- Power BI, Tableau, Looker, Microsoft Excel, SAP, SaaS, Jira, MySQL, PostgreSQL, MATLAB, GitHub, Streamlit

EDUCATION

Masters, Data Science- Illinois Institute of Technology, Chicago **Aug 2022 - May 2024**
Bachelors of Technology, Electronics and Computer Engineering- SRM University, Chennai **Jun 2015 - May 2019**

COURSEWORK

Machine Learning, Big Data, Database Organization, Applied Statistics, Statistical Learning, Algorithms, Project Management, Data Science Practicum, Data Structures, Object Oriented Programming