Jaishankar Harshit Geddam

(703) 509-9178 | jg3682@drexel.edu| linkedin.com/in/harshitgeddam| github.com/harshitgeddam

EDUCATION

MS in Data Science| Drexel University | **GPA: 3.9/4.0** **(**Expected graduation: June 2021)
Coursework: Machine Learning, Deep Learning, Natural Language Processing, Statistical Inference and Modeling, Data Mining, Data Visualization, Artificial Intelligence, Time Series Analysis, Database Management, Marketing Analytics
Teaching Assistant: Database Management, Data Mining, Applied Machine Learning

BTech in Computer Science| Vellore Institute of Technology| India | **GPA: 3.8/4.0** June 2018
Coursework: Object Oriented Programming, Operating System, Computer Networks, Algorithms and Paradigms, Web Development

TECHNICAL SKILLS

Programming: Python (Advanced) [Pandas, NumPy, Scikit-Learn, Matplotlib], SQL(Advanced), R, C/C++, Java, HTML, CSS, Angular JS
Tools, DBs & Technologies: Tableau, MS Excel, MySQL, PostgreSQL, NoSQL, PySpark, Hadoop, MapReduce Hive, GitHub, AWS
ML Algorithms: Supervised Learning (Regression-Linear, Ridge, Lasso, Logistic | Decision Trees | SVM | Neural Networks | Random Forest | Gradient Boosting, XGBoost), Unsupervised Learning (K-Means | Hierarchical Clustering | Dimensionality Reduction | PCA | Cluster Analysis | Anomaly Detection), ANOVA, Market Basket Analysis, A/B and hypothesis testing, ARIMA
Certifications: AWS certified Machine Learning Specialist, GCP Data engineering, Oracle Sales Cloud Implementation specialist
Functional: Business Intelligence (Dashboard), Storyboarding, Project Management and Design Thinking

WORK EXPERIENCE

Data Science/Quantitative Analyst Intern | IsmileTechnologies| Chicago, IL Sept 2020 - Jan 2021

- Predicted food freshness of food items by running multinomial logistic regression models using 1 million data points from 145 restaurants. Results were used by the restaurants stock management team to analyze the next stock order.
- Increased freshness accuracy prediction by 27% by incorporating features from IoT sensors.
- Improved model robustness by creating linear splines, using feature selection methods, by reducing multicollinearity.

Data Science Intern | Verif-y| Philadelphia, PA June 2020 - Sept 2020

- Performed predictive analysis and anomaly detection to identify fraudulent behavior in KYC (Know your customer) process using Decision Trees and Naïve Bayesian Classifier in Python to decrease fraud in KYC by at least 25%.
- Automated the face verification process used for KYC thus reducing the manual work by 30 hrs./week. Designed interactive dashboards using Tableau to effectively portray insights on two year's aggregated KYC process trends.

Data Analyst| Oracle| India July 2018 - July 2019

- Optimized report generation for various clients by designing personalized parameters and using advanced SQL queries in Oracle BI. Reduced redundancy in the number of reports by 65%.
- Designed and developed interactive dashboards using Oracle BI to help a client in analyzing and assessing their sale's incentive compensation, thereby reducing manual work of calculating sale's incentive by 50 hrs./week.
- Developed yearly compensation plans by speaking with the client's key sales leaders regarding design, implementation and also model compensation plan designs against future sales metrics for cost estimations and financial accruals to accomplish the speeding of calculation of incentive's process by 80%.

Data Analyst Intern| Electronics Corporation of India Limited (ECIL) | India April-2017 to July-2017

- Analyzed requirements and developed the data models (Conceptual, Logical and Physical), schema designs (Star and Snowflake) and loaded it into databases (Oracle) using Informatica as ETL tool and Erwin as Data Modeler.
- Designed and developed business intelligence dashboards, analytical reports and data visualizations using Tableau by creating multiple measures using LOD expressions for different user groups, operations and finance team needs.
- Cleaned unstructured data by processing the data using Python Pandas and NumPy, developed database objects, including tables, views, and materialized views to store data in a structured format using Oracle SQL.

ACADEMIC PROJECTS

Amazon Go

- Leveraged **TensorFlow** object detection framework to detect human faces and images of the products such as Coca-Cola, Evian Water etc. with less than 0.1% error to accelerate the checkouts at the grocery store.
- Developed an application using Python shell to link the products to their corresponding prices to facilitate faster checkout thereby saving the customer's time by 90% and also interaction of humans to bill the items in the cart is reduced to 0%.

Heart Stroke Predictions

- Performed data wrangling, adjusted missing values, identified and dealt with outliers, anomalies to predict, heart stroke.
- Employed machine learning models such as **support vector machine, logistic regression, random forest, decision tree** and performed feature engineering using **backward elimination** process to accurately model if a person would have a heart stroke or not and achieved accuracy score of 0.91 by optimizing the models using **GridSearch**.

COVID-19 Analysis

- Explored and filtered data using **Spark SQL** to analyze the confirmed cases and fatalities of Covid-19 provided by John Hopkin's University and implemented exploratory data analysis by building multiple dashboards using **Tableau.**
- Explored the range of regression models such as **Linear Regression, Random Forest Regression** and **Decision Tree Regression** to achieve a r2 score of 0.85 and RMSE of 2.36, built the pipeline using **PySpark**, **Spark ML.**

Identifying Customer Segments for a Grocery Store

- Clustered the customer's features using **MiniBatchKMeans** algorithm and **PCA** to identify potential customers and improve efficiency to structure their delivery service to meet the needs of each customer.
- Performed exploratory data analysis on the features and created visualizations on the clusters. Used **Silhouette Coefficient** to evaluate the models.

RESEARCH PAPER/PUBLICATIONS

Review Paper on Cricket Predictions Using Various Machine Learning Algorithms, International Journal for Research in Applied Science & Engineering Technology (IJRASET), (SYNERGY-2018 conference) Link: https://www.ijraset.com/fileserve.php?FID=17099