Data Engineer Data Engineer Graduate Student at University of Texas at Dallas Dallas, TX I am currently pursuing my Masters in Business Analytics with specialization in Data Science and Marketing from University of Texas at Dallas. Expected to graduate in December 2019. I have 3 years of industry experience working with data and handling them to derive better business insights. I am a fast paced learner with an ability to deliver valuable insights via analytics and data-driven methods. SKILLSET: Analytical Tools: Python, R, SQL, Tableau, PowerBI, Stata, PL/SQL, SAS, MS Excel, Hadoop, HIVE, Shell scripting Statistical and Machine Learning Techniques: Linear Regression, Logistic Regression, Clustering, Random Forest, SVM, Decision Trees, Ensemble Methods, Forecasting Techniques, Predictive Analytics, Unsupervised Learning (K Means, Hierarchical Clustering ), Time Series Forecasting, Dimensionality Reduction using Principal Component Analytics Work Experience Data Engineer Hewlett Packard Enterprise September 2016 to May 2018 Developed visualizations, interactive reports and dashboards using Tableau/ PowerBI to track company performance, KPI metrics and contribution of individual units over time. Devised and implemented Business name swap model using Naive Bayes Algorithm to tackle problems where names were wrongly tagged as businesses and vice versa. Used Logistic regression and random forest to model the probability of attrition with an accuracy of 80%. Strategized insights to clients by mining customer data to increase the customer satisfaction score by 12% Applied k-means clustering on channel transactions, card transactions, product holdings and demographics to categorize customers into 16 segments. Analyzing the segments helped the team to conduct targeted online campaigns to migrate 16% of the customers from low frequency buyers to repeated buyers. Internship, IIM Lucknow June 2017 to August 2017 Applied statistical concepts to solve and analyze business cases published by Harvard Business School. Analyzed the pricing of premium economic tickets relative to regular ones using Linear regression and built a predictive model to predict the factors that affect this difference in pricing using RStudio Database Management (UT Dallas) November 2018 - December 2018 Performed data cleaning, normalization, designed data flow diagrams, schema mapping on Oracle database. Applied Machine Learning (UT Dallas) February 2019 - March 2019 Implemented linear regression using

gradient descent and Machine learning classification algorithms (SVM, Decision tree, Boosting) using Python on UCI Facebook Comment Dataset and experimented with the performance of various kernels and algorithms to decide on the best suitable algorithm. Service Information Developer February 2015 to August 2016 Created Oracle Stored procedures and performed unit testing on all stored procedures to ensure successful deployment in production Developed the ETL Solutions to move data from the in-house custom-grown application to the reporting suite. Analyzed business needs, interviewed stakeholders, compiled functional system requirements and modeled workflow processes for internal BI projects Database Administrator September 2014 to February 2015 Performed database activities on SQL/ Oracle production databases such as enhancing database performance, performing restores, implementing recovery procedures and conducting regular system backups. Built efficient SSIS packages for processing fact and dimension tables with complex transforms and implementing best practices to maintain optimal performance using SQL Server 2012. ANALYTICAL PROJECTS Education M.S in Analytics The University of Texas at Dallas - Dallas, TX December 2019 B.E in Electronics and Communication Cambridge Institute of technology, Visvesvaraya Technological University June 2014 Skills Decision trees, Linear regression, Logistic regression, Machine learning, Pca Links http://linkedin.com/in/meghana1911

Name: Andrea Jones

Email: brianarodriguez@example.org

Phone: +1-608-237-6202x53261