

Graduate Teaching Assistant Graduate Teaching Assistant Graduate Student | Actively seeking Full Time Opportunities | Data Science & Big Data Enthusiast Dallas, TX ? 3+ years' experience in Data Analytics with strong analytical and problem solving skills ? Excellent programming skills in Python, SQL, UNIX, R and SAS ? Certified professional in SQL, Hadoop and Spark ? Hands on experience in -

- * Data Preprocessing using pandas
- * Exploratory Data Analysis (Dimensionality Reduction, Clustering and Principal Component Analysis)
- * Datawarehousing & Data Modeling
- * Big Data technologies (Hadoop MapReduce, Spark)
- * Hypothesis Testing
- * Machine Learning (Decision Trees, Naive Bayes, OLS Regression, Logistic Regression, Neural Networks, Support Vector Machines, Ensemble Methods, Singular Value Decomposition)
- * Data Visualization (Tableau and Oracle Business Intelligence Suite)
- * Agile Methodologies

? Team player with excellent communication skills and leadership abilities ? Volunteered for a few social and knowledge imparting causes Authorized to work in the US for any employer Work Experience Graduate Teaching Assistant University of Texas at Arlington - Arlington, TX June 2017 to Present ? Prepared coursework for Data Visualization using Tableau and delivered lectures to a class of 30 achieving 98% pass rate ? Presented a demo about E-commerce sales forecasting using Arima models in R ? Tutored 60+ students on Analysis with SAS and R for various analytics projects Programmer Analyst Cognizant Technology Solutions - Bangalore, Karnataka June 2014 to July 2016 ? Enhanced ETL code by writing complex SQL scripts and stored procedures to improve performance of processes from 60% to 85% ? Boosted loyalty programs by developing dashboards by analyzing hotel stay data using Advanced Excel techniques (pivot tables, VLOOKUP) to discover insights based on guest behavior and preferences ? Developed an ETL pipeline using Spark to ingest hospitality data from multiple data sources (xls, csv, XML, text files etc.) and ensured data consistency for downstream applications ? Took the initiative to learn Oracle Business Intelligence suite to automate project reporting thereby reducing manual effort by 90% ? Developed UNIX scripts to create trigger files, send failure notifications and to backup ETL processes ? Enabled quick retrieval of data from the database by creating indexes and views on large tables ? Organized team building events. ? Volunteered for the Outreach Program. Python Developer Intern

Defense Research & Development Organization - Bangalore, Karnataka December 2013 to May 2014 Developed a Network Intrusion Detection System using statistical frameworks like Na ve Bayes and Hidden Markov model in python. Mitigated Distributed Denial of Service attacks on the network traffic. Relevant Experience ? Estimated probability of customer retention by analyzing service contracts using Maximum Likelihood Estimation in R ? Developed a predictive model in SAS to predict the Net Revenue generated by internet marketing for Air France ? Developed an Anime Recommendation System using collaborative filtering algorithms in Spark to enhance user experience ? Performed clickstream analysis of 1 million users on msnbc.com using core python to provide competitive advantage ? Created an inverted index of authors and their publications from 250+ journals using Hadoop MapReduce ? Predicted Walmart sales during a weather event using regression models in sklearn and forecasted sales during stormy weather ? Derived insights about prevalent unemployment in the US and forecasted future unemployment rate in Tableau dashboard ? Developed an E-commerce sales forecast using Arima models for Q4 of 2017 and Q1 of 2018 for an online retailer ? Performed Market Basket analysis on retail store data of 1500 transactions to find the top 10 rules (frequent item list) ? Performed sentiment analysis on 20000 tweets about NFL players' protest by calculating sentiment score for each US state and presented the findings in Tableau Education MS in Information Systems University of Texas at Arlington - Arlington, TX August 2014 to May 2018 BE in Computer Science and Engineering Visvesvaraya Technological University - Bangalore, Karnataka September 2010 to June 2014 Skills Informatica (2 years), Oracle Business Intelligence Suite (2 years), Python (4 years), Shell Scripting (5 years), Spark (4 years), Hadoop (4 years), R (2 years), Tableau (2 years), SAS (2 years), Advanced Excel (4 years) Links <http://www.linkedin.com/in/padmaksekar> Awards Rising Star 2015-06 Presented the Rising Star Award for emerging as the next generation achiever. Certifications/Licenses Hadoop April 2018 to Present Spark April 2018 to Present MySQL March 2018 to Present SQL September 2017 to Present Network Analysis using Python June 2017 to Present Additional Information Programming Languages: SQL, Python(sklearn, numpy, pandas, matplotlib), R, SAS, Scala, UNIX, Java, C, C++, HTML, CSS, JavaScript Databases: MySQL, Oracle, IBM DB2, SQL Server, NoSQL (MongoDB,

Cassandra) Tools: Google Analytics, AWS (Redshift), SAP, Informatica, Microsoft Office Suite
(Word, PowerPoint), PowerBI, AngularJS, Docker, Linux, Git Statistical Methods: Time Series
Forecasting, Regression Models, Hypothesis Testing, Principal Component Analysis Operating
Systems & Virtual Machines: Windows, Linux, macOS, VirtualBox, VMware

Name: Michelle Keith

Email: erin36@example.org

Phone: 001-907-774-4861x09442