SAURABH CHOPDA

458-254-9250•saurabhchopda1997@gmail.com•linkedin.com/in/saurabh-chopda•github.com/saurabhchopda

EDUCATION

University of Texas at Dallas, Richardson, USA Dec. 22
M.S.,Information Technology and Management GPA:3.6/4
Relevant Coursework: Business Analytics with R, Data Visualization, Data Management with SQL, Python, Statistics and Data Analysis, System Analysis and Project Management, Predictive Analytics with SAS, Spreadsheet Modeling

Vishwakarma Institute of Technology, Pune, India May. 19
Bachelor of Technology, Electronics and Telecommunications Engineering GPA:7.1/10
Relevant Coursework**:** Data Structures and Algorithms, Linear Algebra, Probability, Calculus, Linux, Digital Electronics, Robotics

TECHNICAL SKILLS

Programming: Python, JavaScript, R, SQL, Java, C++, HTML, CSS, Shell Scripting, Scala
Technologies: Azure, AWS, React, Express JS, Databricks, Apache Pinot, Kubernetes, Apache Spark, Tableau, Git, Docker
Databases: PostgreSQL, MySQL, Oracle, MongoDB, CosmosDB
Certifications: Machine Learning (Coursera), Python (Udemy), R (Data Camp), Tableau(LinkedIn Learning)

PROJECTS

Descriptive Data Visualization Fall 21
- Designed multi-maps in Tableau to visualize and narrate the top 5 natural disasters across the various states in the USA.

Data Mining Summer 21
- Developed a Netflix like movie recommendation system using Collaborative Filtering and K-means clustering algorithms.
- Tuned and created a logistic regression model to predict whether a person would default on debt or not based on the type of debt.

Dallas Animal Shelter Analysis Spring 21
- Analyzed actual animal shelter data and designed normalized entity relationship diagrams to create tables in a Oracle database.
- Generated reports and SQL queries to analyze which breed of cats and dog survived and intake type versus the outcome of the breed.

Non-Intrusive Load Monitoring Spring 19
- Developed a Markov learning algorithm using Python to monitor and detect patterns in energy usage of a house.

Human Activity Recognition Spring 18
- Developed a Convolutional Neural network(CNN) using Keras and TensorFlow to identify basic human activities like standing, sitting and walking from a video.

EXPERIENCE

7-Eleven , Irving TX, USA Jan. 23 - Present
Data Engineer
- Collaborated towards designing a framework for large-scale Enterprise ML and Data Engineering solutions.
- Created Spark jobs for data partitioning resulting in increased ingestion pipeline performance and reduced memory footprint.
- Developed Spark pipelines utilizing Pyspark and Spark-SQL to extract, transform, and aggregate data for analysis to reveal insights into consumer usage patterns.

7-Eleven , Irving TX, USA Feb. 22 - Dec .22
Data Engineer Co-op
- Developed and launched a full stack React app with Express JS backend to perform integrations with Kubernetes API and Jupyter Hub.
- Conceptualized and implemented a specialized ETL process in DataBricks using SQL that handled over a million transactions per day.
- Migrated pinot data ingestion jobs from Azure Databricks to Spark on Kubernetes.

University of Texas at Dallas, Richardson, USA Jan. 22 - Feb. 22
Graduate Teaching Assistant
- Graduate Teaching Assistant for the Quantitative Business Analysis class.

Chopda Trading Company, Pune, India Jun. 19 - Dec. 20
Business Analyst
- Designed a central warehouse for customers, orders, and payments data for systematic storage and analysis using various tools like SQL, Python, Apache Airflow and AWS.
- Analyzed metrics using Python and R for sales and financial location analysis resulting in 15 % increase in sales.
- Developed dashboards, visualizations and analysis reports using Tableau to project growth in sales.

Unithing Technologies LLP, Pune, India Aug. 18 - Dec. 18
Machine Learning Intern
- Developed machine learning models for energy monitoring and conservation resulting in savings of up to 12 % of total energy consumed.
- Automated ETL pipeline for efficient data processing and local storage using Python, Raspberry Pi and AWS.
- Created REST APIs of different learning models using Flask in Python and deployed to AWS and Digital Ocean using Nginx.