DHVANI MITESH KOTHARI

Buffalo, NY | (716)-936-4705 | dhvanikothari03@gmail.com | www.linkedin.com/in/dhvanik/ | github.com/dhvani-k

SUMMARY

A data wizard in making, with a wealth of experience as a Data Engineer and proficiency in Data Science, specializing in Machine Learning with a Master's degree. Managing big data and leveraging cutting-edge techniques to uncover insights that will revolutionize the way organizations approach data.

EDUCATION

University at Buffalo, The State University of New York — **Buffalo, NY**
*Master of Science, Computer Science & Engineering (**GPA:** 3.75)* — *August 2022 – December 2023*
Courses: Data Intensive Computing, Algorithms for Modern Computing Systems, Modern Networking Concepts, Machine Learning
Yeshwantrao Chavan College of Engineering — **Nagpur, India**
*Bachelor of Engineering, Computer Technology (**GPA:** 4.0)* — *July 2015 – May 2019*

SKILLS & LANGUAGES

Programming Languages: Python, SQL, PySpark, C/C++, Scala, MATLAB, Java
Big Data: Apache Spark, Hadoop, Hive, Pig, MapReduce, PySpark, Airflow, Kafka, Snowflake, Splunk, HBase
ETL: Informatica PowerCenter, Talend, Databricks, Azure Data Factory
Databases/Server: Oracle, PostgreSQL, SQL Server, MySQL, MongoDB, Cassandra, Redshift, SSIS, RDBMS
Machine Learning: CNN, RNN, Text Classification, Predictive Modeling, NLP, OpenCV, Statistical Analysis
Cloud: AWS, Google Cloud Platform, Microsoft Azure, AWS EC2, EMR
Data Visualization: Tableau, PowerBI, Google Data Studio, Microsoft Excel, VBA

PROFESSIONAL & EXPERIENCE

Solutions Architect Intern, BrainChip Inc., Laguna Hills, California — *June 2023 – Present*
- Built a Web SAP with FastAPI to facilitate scalable estimations for neural networks on neuromorphic edge-computing ICs
- Leveraged Pytest and Drone to establish an efficient automated user-centric testing infrastructure for an advanced deep-learning framework

Data Engineer, Walmart Global Tech, Bangalore, India — *May 2021 – July 2022*
- Developed a time-series ensemble forecasting solution for **599** Sam's Clubs by leveraging historical sales and calendar event data predicting net sales at club-month level diminishing annual operation planning cost by ~**$1.2** million
- Created real-time data and model monitoring alerting system using Python, detecting data quality anomalies and model drift, increasing accuracy of performance degradation predictions by **25%**
- Collaborated with tech professionals on **8+** projects including data requirement gathering, data collection, processing, model training/validation, and evaluation of performance metrics

Data Analyst, Capgemini, Navi Mumbai, India — *June 2019 – April 2021*
- Optimized Azure data pipelines by incorporating Deepomatic's Deep Learning Anonymization algorithm, increasing accuracy by **50%** over **100** pilot tests of autonomous car driving systems
- Integrated Hadoop and Spark into Data Lake, accelerating extraction, transformation, loading of unstructured data by **40%**
- Led a team of **6** to develop ETL mappings and data modeling, empowering data-driven insights for business derivatives

ACADEMIC PROJECTS

StockStream: Continuous Stock Market Data Flow and Insights — **Buffalo, NY**
Data Engineering: Python, Kafka, AWS EC2, Athena, Glue crawler — *May 2023 – June 2023*
- Implemented python script to automate the loading of stock market data from a CSV file into the producer stream of Kafka, enabling real-time data processing and analysis
- Built scalable Kafka cluster on EC2 instance with 99.9% uptime, ensuring continuous consumption of consumer data stream

Predicting Formula-1 Race Winner — **Buffalo, NY**
Data Science: Data Cleaning & Preprocessing, EDA, Regression, Classification, Neural Networks — *February 2023 –May 2023*
- Performed data preprocessing and explanatory data analysis to identify important features/relationships in historical race data
- Developing **5** machine learning models (SVM, Random Forest Classifier, Gaussian Naive Bayes, Decision Trees, and Neural networks) to predict Formula-1 race winners; achieved **85%** prediction accuracy

YouTube Trending Video Analysis — **Buffalo, NY**
Data Engineering: AWS Athena, Glue, S3, QuickSight — *January 2023 – February 2022*
- Architected a scalable data pipeline with AWS Glue ETL to transform over 100k records, stored securely in an S3 bucket
- Utilized QuickSight to analyze engagement metrics, identifying top-performing videos, and improve marketing strategies

ACTIVITIES & LEADERSHIP

President, **Association for Computing Machinery Student Chapter, YCCE, Nagpur, India** — *July 2018 – May 2019*

International Conference on Advancements in Computing & Management, Jaipur, India — *January 2018 – May 2019*
Paper publication on Monitoring Logistics Through Blockchain