Pranav Kuramkote Sudhir

Boston, MA • +1 857 465 9377• kuramkotesudhir.p@northeastern.edu• https://sites.google.com/view/pranavksudhir

EDUCATION

Northeastern University, Boston, MA | GPA: 3.741 Expected May 2025
Master of Science (M.S.) Data Analytics, Concentration – Data Science
Coursework: Foundation of Data Analytics, Computation and Visualization, Data Management, Data Mining
Siddaganga Institute of technology, Tumakuru, India| GPA: 3.89 May 2019
Bachelor of Engineering (B.E.) Mechanical Engineering
Activities: Core Committee organizer for University Fest, volunteered in student cultural organizations

SKILLS AND CERTIFICATIONS

Programming Technologies: Python, SQL, kafka, Tableau, PowerBI, Linux, Spark, Hadoop, pytorch, SPSS, SAS, Tableau, Microsoft Azure
Project Management Tools: Jira, GitHub, ServiceNow, Advanced Excel, AGILE Methodologies
Productivity Tools: Microsoft Suite, Microsoft office, SAP, Google Suite, Zoom, Canva, Discord
Certification: IBM Data Science Professional Certificate, Taming big data with Apache Spark and python(Udemy)

WORK EXPERIENCE

HCL Technologies limited | Bangalore, India | Member Technical Staff Feb 2021 – Jun 2023
•Engineered end-to-end data pipelines using PySpark, Kafka, and Python scripts, facilitating seamless ETL processes and storage of data in HDFS format
•Implemented real-time data streaming with Kafka, ensuring timely ingestion of data for immediate processing within the pipeline
•Designed and deployed REST APIs to monitor the health of HDFS nodes, enhancing system reliability and performance
•Collaborated with cross-functional teams, conducted performance tuning, and optimized data processing tasks for improved efficiency and reduced processing times
•Set up spark SQL, Used SQL to query the database, created stored procedure to capture changes, created dashboards with meaningful data-driven insights
•Worked as the team audit specialist, prepared KPI, IDM and FIR reports and attended monthly audit meetings.

QUEST Global Engineering |Bangalore, India | trainee Engineer Nov 2019 - Oct 2020
•Implemented ETL pipelines using Databricks and Azure data factory, streamlining extraction, transformation, and loading of data
•Collaborated with cross-functional teams to integrate diverse data sources, ensuring a unified approach to data processing
•Performed data extraction, transformation, and loading tasks directly within the database environment using Azure SQL database, to transform data before analysis.
•Employed visualization Power BI, to present data visualizations and contribute to data driven decision making.

PROJECTS

Yelp_review Real Time Data Streaming (Python, Kafka, GPT-3.5 Turbo, Streamlit, PostgreSQL, Docker Compose):
•Developed end-to-end Kafka-based pipeline for sentiment analysis of Yelp reviews, integrating GPT-3.5 Turbo model, PostgreSQL storage, and Tableau visualization for insights.
•Engineered Kafka consumer to process real-time data from reviews_db and sentiment_analysis, extracting review details and sentiment analysis results.
•Ensured data integrity and consistency by designing PostgreSQL schema and tables for storing review data and sentiment analysis outcomes.
•Created Tableau visualizations for analyzing sentiment trends, category distribution, user ratings based on the stored data.
Customer Segmentation using RFM Analysis (Data wrangling Data, validation, Dashboard creation):
•Led a team project leveraging Python, Jupyter Notebooks, and Pandas for comprehensive analysis of an eCommerce dataset
•Applied RFM analysis and K-Means clustering for customer categorization, enhancing insights with Tableau for visual representation.
• Utilized Tableau for impactful visualization of results, enhancing communication.
•Presented actionable marketing recommendations in a concise report, demonstrating proficiency in Python, Pandas, and Tableau for data analysis and visualization.
F1 Data Processing and Real-time Analysis:
•Leveraged Kaggle F1 dataset, processing it using Azure Databricks for seamless data integration and storage.
•Implemented a data broadcasting solution in Python for optimized data dissemination, showcasing technical proficiency in data manipulation.
•Conducted real-time analysis of F1 data using Power BI, demonstrating the ability to derive insights for immediate decision-making.
•Successfully executed end-to-end data processing and analysis, emphasizing skills in cloud computing, data integration, and realtime analytics.

Awards and Recognitions

• Awarded HCL ERS Champion certificate for the quarters AMJ’22, JAS’22