AKSHATA SALUNKHE

Tampa, FL | +1(305)-502-4285 | salunkhe10@usf.edu | LinkedIn | GitHub

EDUCATION

Master's (MS) in Artificial Intelligence and Business Analytics (Major: Data Science)   **GPA: 3.67/4.0**
University of South Florida, Tampa, FL   **Aug 2023 – Dec 2024**

Relevant Coursework: Data Mining (Python), Advanced Database Management (SQL Server), Data Science Programming, System Design, Data Structure and Algorithms, Applied Machine Learning, Data Warehousing, Big Data Technology, Cloud Computing.

EXPERIENCE

Eaton India Innovation Center
Associate Data Engineer   **Aug 2021 – Aug 2023**

- Directed a dedicated team of data engineers to implement data acquisition strategies, establishing robust ETL data pipelines to extract valuable insights from historical datasets. Utilized statistical analysis techniques for predictive modeling and data forecasting to empower data-driven decision-making.
- We leveraged Spark, Python, SQL playing a pivotal role in applying analytics for takeouts insights for guidance to maintain a healthy workflow and simplify the management of everyday testing.
- Played a key role in driving process improvement initiatives encompassing key elements for scalability, security, and high availability featuring by applying Six Sigma methodologies, achieving a 60% reduction in manual testing efforts.
- Achieved Six Sigma Green Belt certification post DMAIC training; successfully deployed data-driven approaches leading to a 60% decrease in testing time and a 70% decline in defects.
- Championed the adoption of Agile Development Methodology, enabling a consistent flow of new product releases annually and reducing time-to-market by 25%.

IOTIOT
Data Engineer Intern   **Apr 2020 – Aug 2020**

- Collaborated with data scientists and fellow data engineers to develop data-driven solutions in the geospatial domain, utilizing AWS infrastructure and Big Data technologies such as Apache Spark.
- Managed & orchestrated CI/CD processes for real time data ingestion, while leveraging Bitbucket data pipelines, and SQL procedures and helped streamlining development workflows.
- Identified ETL workflow bottlenecks, executing optimized algorithms to streamline processes, resulting in a 40% reduction in data processing time, significantly improving accuracy and timeliness.
  .

PROJECTS

PII Redaction from vCons – Tad Hack – Company STORLID - Link   **Mar 2024 - Mar 2024**

- Developed a specialized Named Entity Recognition (NER) model fine-tuned for identifying entities in human conversations, leveraging advanced Natural Language Processing (NLP) techniques.
- Implemented the model to accurately identify entities within conversations, subsequently replacing them with the [REDACTED] tag to ensure data privacy and confidentiality.
- Enhanced NER model accuracy in NLP, integrating sentiment analysis and text summarization for diverse input formats, optimizing entity identification in conversational data.

State of Tampa Region | Graduate Assistant at USF (Research Project)   **Sep 2023 - Jan 2024**

- Elevated role to lead the Tampa Bay e-insights report, focusing on themes of affordability and talent pipeline comparing with other 19 metropolitan statistical areas in the United States.
- Extracted data from public sources (Google Trends, Census, FRED), expertly cleaning and transforming raw datasets using Python, Excel, and R.
- Created Business Intelligence (BI) solutions using Tableau, improving data interpretability for stakeholders, revealing deeper insights into underlying dynamics and trends.

CloudWeatherETL: Automating Weather Data Pipelines on AWS   **Jun 2022 – Sept 2022**

- Developed project utilizing a variety of cutting-edge technologies on Amazon Web Services (AWS), including Apache Airflow, Python, Apache Spark, AWS Glue, and Redshift.
- Executed a robust data pipeline capable of seamlessly extracting weather data from the open weather map API, employin Apache Spark for efficient data transformation, and orchestrating the end-to-end ETL process through Airflow.
- Implemented integration with multiple AWS services, leveraging the power of AWS Glue for data cataloging and Redshift for scalable and high-performance data warehousing, resulting in a fully automated and scalable solution for handling weather data in a cloud environment.

TECHNICAL SKILLS

Core Skills: C++, Python, R, SQL, C#, JavaScript, Unix-shell scripting.
ML/DL: Scikit Learn, TensorFlow, Matplotlib, NumPy, Pandas, Statistical Methods (ANOVA, T-test), SVC, Random Forest, Regression.
Databases: MSSQL, Oracle, SQL Server Management Studio, Azure Data Studio, PostgreSQL, NOSQL,
Big Data Technologies: Apache Hive, Apache Spark, HDFS, Impala, MapReduce, AWS (Lambda, Glue, S3, Redshift), IBM db2, Cloudera.
Version Control: Git Repository, Scrum, Jira, Confluence, Jenkins, Docker
Data Visualization: Tableau, Power BI, Excel, Microsoft Power Point, ERD
Project Management: Scheduling, Leadership, Collaboration, Active Listening, Problem Solver, Communication.
Others: Visio, Virtual machines, Agile practices, Windows, Linux, MS Office, Task Management, Minitab, Six-Sigma Methodologies, Ranorex

PUBLICATION AND AWARDS

- Winner **Tad Hack 2024** - STORLID vCon PII Redaction Pitch and Podcast
- Presented a Research Paper on **"State of Tampa – E-Insights",** in University of South Florida 2024 publication.
- Presented a Research Paper on **"Software Defect Prediction using Cutting Edge Clustering Module"** in IEEE 2022 publication.