MOHIT CHANDRA KANTI
Tampa, Florida 813-897-2353 mohitkanti55@gmail.com LinkedIn

PROFESSIONAL SUMMARY

Results-driven **Data Engineer** with over **4+ years of experience** in architecting, deploying, and refining sophisticated **data pipelines** and **ETL processes**. Proficient in leveraging advanced technologies including **Spark**, **AWS S3**, **Apache Kafka**, and robust analytics tools such as **Pandas** and **NumPy** to elevate data processing efficiency and system reliability. Demonstrated prowess in spearheading **agile projects**, optimizing **data storage**, and facilitating seamless **data integration**. Actively pursuing a **Master's in Business Analytics** to merge technical skills with strategic business insights. Aspiring to contribute to and lead innovative **data-driven solutions** in forward-thinking organizations.

EDUCATION

University of South Florida **Aug 2022 – May 2024**
Master of Science in Business Analytics and Information Systems ***GPA: 3.87/4.00*** *Tampa, Florida*

PROFESSIONAL EXPERIENCE

Accenture **Jun 2021 – Jul 2022**
Data Engineering Analyst *Hyderabad, India*

- Spearheaded **Spark**-based data processing on **AWS S3**, improving retail analytics by **30%** in inventory accuracy and **25%** in speed. Enhanced data retrieval and query efficiency on S3 by **35%** and **20%**, respectively, using optimized **Hive** queries.
- Orchestrated **Airflow**-driven **ETL** workflows, reducing manual data handling by **40%** and boosting reliability by **20%**. Achieved a **60%** decrease in manual effort and a **30%** gain in process efficiency through **Java** and **Python** automation.
- Developed **RESTful APIs** with **Flask** and **Spring Boot**, doubling data interchange and interoperability. Facilitated a **50%** increase in data science accessibility by integrating **Snowflake** data warehouse solutions, enhancing analytics workflows.
- Enhanced **Agile Scrum** and sprint planning productivity by **30%**, shortening project delivery by **20%**. Improved data flow efficiency by **40%** and reduced latency by **25%** with **Kafka Connects**, also elevating system reliability by **30%** through proactive monitoring via **App Dynamics** and **Splunk**.

Accenture **Jun 2019 – May 2021**
Data Engineering Associate *Hyderabad, India*

- Led data analytics on **AWS S3** using **Spark**, enhancing processing efficiency by **45%** and analytics accuracy by **30%**. Utilized **Java** and **Python** for data analysis, improving decision-making speed by **20%**.
- Improved data quality and integrity by **40%** through advanced profiling and **Java/Python ETL** processes, and streamlined **Kafka** data integration, boosting processing by **35%**. Implemented business intelligence and data modelling techniques.
- Developed scalable **ETL** processes with **Spark** to **Snowflake**, increasing data storage access by **50%**. Enhanced **Hive** schema efficiency by **30%** and implemented effective data lake strategies.
- Enhanced operational efficiency by automating data pipelines, reducing manual effort by **20%** and increasing workflow efficiency by **50%**. Led **CI/CD** with **Jenkins**, boosting deployment frequency by **50%** and cutting rollbacks by **35%**.

Team JEC **May 2018 – May 2019**
Data Engineering Intern *Visakhapatnam, India*

- Created and maintained interactive dashboards in **Tableau**, enabling stakeholders to easily monitor key performance indicators (KPIs) and make data-driven decisions.
- Produced detailed mapping spreadsheets, enhancing data transformation efficiency for the **ETL team** by **30%**.
- Conducted in-depth data analysis using **SQL scripts**, improving data quality and accuracy during loads and extracts by **20%**.
- Oversaw extensive dataset management, resulting in a **25% improvement in analytical efficiency** and crafted compelling visualizations using **Tableau** enhancing data presentation.
- Designed dynamic business solutions and dashboards using **Excel**, **Power Pivot**, **DAX**, and **Power Query**, contributing to a **15% increase in data-driven decision-making efficiency**.

SKILLS

Languages: Python, R, SQL, Java.
Packages and Platforms: Pandas, NumPy, Matplotlib, Power BI, Power Query editor, Databricks, MapReduce, Hadoop, Apache Spark, Azure ML studio, R studio, Kubernetes, Product lifecycle management (Agile), Tableau, Linux, Unix, Cloudera, GIT.
Databases: SQL, Snowflake, Hive, Impala, Pig, Spark SQL, Databases SQL-Server, MySQL, PostgreSQL, NoSQL, MS Access, HDFS, HBase, Oracle, Teradata, Netezza, Mongo DB, Cassandra.
Developer Tools: Eclipse (IDE), Code Blocks, PyCharm, Jupyter Notebook, Putty, WinSCP, GitHub, Kibana, Jenkins.
Cloud Technologies: AWS S3, AWS EMR, AWS EC2, AWS Lambda, Athena, GCP, MS Azure.

ACADEMIC PROJECTS

Car Evaluation | ***Python, R, SQL, Azure ML Studio, Tableau, Power BI***

- Led a team in analyzing used car prices through data mining techniques, employing Python and R for data processing, and Azure ML for implementing and evaluating regression models (Decision Forest, Boosted Decision Tree, Linear Regression).
- Visualized data correlations and model performance metrics using Tableau and Power BI, providing actionable insights for automotive companies and car resellers to accurately price used cars.

US Food Imports Analysis | ***Power BI, Tableau, Python***

- Analyzed the value and volume of US food imports using Power BI and Tableau, revealing trends and identifying top imported products and countries over a five-year period.
- Created predictive models with Python to forecast import trends, enhancing understanding of market dynamics and global trade's impact on the US food industry.