SAVANKUMAR GOSWAMI

Chandler, AZ, 85248 | savankumargoswami@gmail.com | (608)-320-9384 |LinkedIn Profile|
GitHub|Portfolio

SUMMARY

- **AWS** Cloud Practitioner **and Tableau** Desktop Certified professional Looking forward to work in the areas **of Big Data Analytics**, **ETL platform development**, **Reporting**, **Data Management Solutions** and **Project development**.
- Extremely comfortable using Python, especially skilled in **Python** and worked on packages like **PySpark, MRJob, Airflow**, Keras, TensorFlow, NumPy, Pandas, NLTK, Scikit-Learn, BeautifulSoup, OpenCV, Tox, Flask.
- Knowledgeable in **Big Data technologies** like Hadoop, Kafka, HDFS, Hive, Cassandra, MapReduce, Spark, ElasticSearch; and open source platforms like Databricks, Hue and Jupyter notebooks.

TECHNICAL SKILLS

Languages: Python, SQL, BigQuery, R, C/C++, Scala
Databases: RDBMS (MySQL, Teradata, Postgres)
NoSQL (MongoDB, DynamoDB, S3)
Big Data/Cloud: Spark, Kafka, Hadoop, MapReduce, AWS, GCP, AZURE
Libraries/Frameworks: Tensorflow, PyTorch, Keras, Pandas, NumPy, XG Boost
Tools: Looker, Tableau, PowerBI, Data Studio, Docker, Kubernete

EDUCATION

- **Master of Data Analytics (MS)** **(Jan 2022 – May 2024)**
  Long Island University – Long Island, NY GPA: 3.89/4.0
- **Bachelor of Engineering in Mechanical Engineering (BSME)** **(Aug 2011- Aug 2015)**
  Gujarat Technological University, India

PROFESSIONAL WORK EXPERIENCE

Data Engineer Intern, Pop Business Solutions, Tempe, AZ, USA **(Aug 2023 – Dec 2023)**

- Designed Python-based SQL wrappers for data analysis to negate client's **>10 yrs** roadblock and synchronization of **~350M rows** of time-series dataset
- Enhanced execution and query retrieval time from a few days to ~40 minutes into Teradata Vantage Database using Transcend on-premise ecosystem and Docker, resulting in **~95%** latency **improvement**
- Worked with multi-functional teams to design **ETL pipelines** to parse HDF5 (Hierarchical Data Format) datasets into Teradata data frames for **feature engineering** and multi-cloud predictive analytics
- **Led** initiation to design product prototypes and apply analytic workflows into cloud-native architecture (VantageCloud) following Agile methodology
- Deployed client specific automated ModelOps workflows and Git models for governance and model lifecycle management leading to **30%** improvement in model drift monitoring metrics

BUSINESS INTELLIGENCE ANALYST, Dodder Engineering Consulting, Pune, India **(Dec 2020 – Dec 2021)**

- Conducted exploratory data analysis, discovering insights that led to a **25% improvement** in operational processes.
- Implemented data profiling techniques, identifying new avenues for optimization that resulted in a **10% reduction in operational costs**.
- Leveraged machine learning algorithms to analyze historical maintenance data, resulting in a **25% reduction in unplanned downtime** with **12% improvement** in overall delivery efficiency.
- Helped Developed real-time data pipelines delivering data into HDFS and SQL Server using Stream sets data collectors, contributing to a **25% reduction in data delivery time**.
- Performed end-to-end data analysis using Python including data quality assessment, and multivariate analysis, and built an interactive dashboard using **Power BI and Tableau** to provide the client with insights on potential customers to increase their profit. The system included data cleaning, exploratory data analysis, and data visualization using Power BI and Tableau.

MECHANICAL ENGINEER, Oasis Food Processing Pvt Ltd, Indore, INDIA **(Jan 2018– Dec 2020)**

- Led the New Product development (**NPD**) team to **develop and improve** crystallization reactor & automatic fluid control systems.
- By optimizing the process windows **(JMP)** and re-using the parts wherever necessary increased **annual production** volume of food process equipment by **over 0.25M parts**.
- Conducted detailed analysis of supplier quotes proposing new product ranges using **SAS and SQL** to **increase the delivery rate by 10%.**
- Added new **Key Performance Indicators (KPIs)** utilizing data collection and analysis tools **(SQL)** for daily monitoring. This initiative resulted in a **yield improvement of 98%.**
- By offering **engineering support** to both suppliers and customers, we **expedited the decision-making process** for order confirmations, ensuring swift outcomes.

MECHANICAL ENGINEER, Rajesh Machines India Pvt. Ltd., Ahmedabad, IND **(Dec 2015 – Aug 2017)**

- **Designed, developed, and analyzed** customized prototypes of Hydraulic, and pneumatic power press machines using **SolidWorks.**
- Used manufacturing strategies (DFM, DFA) to improve Hydraulic Press drift issue by improving **automatic crowning** and **angle correction techniques.**
- Using **SolidWorks** prepared **3D & 2D** drawings from analytical calculations or handmade drawings or per requirement of customers and effectively delivered developed models to customers and suppliers for next step.
- Introduced the **DFMEA tool**, **and other Six Sigma tools** for **reducing material waste and reworks of machine components** (**Stroke Counter, Table & Slide (RAM) by 40%**.
- Gathered customer data of failed parts through in-person visits and performed **root cause analysis (RCA)** for improved product outcomes.

ENTHUSIASTIC PROJECTS

NYC Data Engineering - Modern Data Analytics: Designed and implemented a high-impact **data analytics pipeline** using **Google Cloud Platform (GCP)** services. Processed a staggering 2 crore data points annually from NYC Taxi Trip Records (2020-2023). Utilized cutting-edge technology to extract actionable insights and enable informed decision-making. Proficient in **Python, GCP Storage, Compute Engine, BigQuery, and Looker Studio.**

Telecom Customer Churn Prediction Model: Identified and gathered factors influencing customer contract renewals for a financial services company using **Python, Pandas, Seaborn, scikit-learn**. Conducted data preprocessing, feature engineering, and **implemented a Decision Tree classifier** to predict customer churn. Achieved an **88.2% accuracy** rate in predicting customer contract renewals. Provided valuable insights to reduce customer churn and improve retention strategies.