SASANK YADAV DALIBOYINA

+1 8573813570 | daliboyina.s@northeastern.edu | LinkedIn |Boston, MA

EDUCATION

Northeastern University **Massachusetts, US**
Masters in Applied Machine Intelligence | 3.90 GPA Dec 2024
Predictive Data Management and Big Data, Applications of AI, Python and Analytics System

Gandhi Institute of Technology and Sciences (GITAM). **Andhra Pradesh, India**
Bachelor of Technology in Computer Science. Aug 2021

EXPERIENCE

Data Science Researcher | CLOUDPORT(Start-up) **|** Boston, MA. Oct 2023 – Jan 2024

- Analyzed Medicare data, enhancing prediction accuracy by 15% with statistical modeling, driving healthcare strategies.
- Excelled in data visualization, boosting stakeholder engagement by 20% through insightful dashboards.
- Refined machine learning models like Random Forest and XGBoost, advancing analytical precision and insight extraction.
- Streamlined predictive analytics workflows, harnessing Big Query to manage and analyze healthcare datasets, significantly enhancing data-driven policy formulation.
- Engaged in continuous learning, exploring emerging technologies in AI and machine learning, including neural networks with TensorFlow and natural language processing with spaCy, to keep pace with cutting-edge analytical methodologies.

Big Data Engineer | TATA CONSULTANCY SERVICES(TCS) **|** Hyderabad, India Jun 2021 – Dec 2022

- Led Tableau dashboards creation, yielding a 25% boost in decision-making efficiency through data insights.
- Executed SQL and Python analyses to craft predictive models, reducing telecom downtime by 30%.
- Headed Hadoop and Spark Big Data projects, translating telecom data into network improvements.
- Deployed TensorFlow and scikit-learn for telecom analytics, enhancing predictive maintenance.
- Enhanced team knowledge and project delivery by incorporating data modeling techniques and best practices in data architecture.
- Established and maintained real-time data processing workflows with Kafka, improving data throughput and latency.
- Managed code versioning and team collaboration using Git, ensuring high standards of code quality and project tracking.

Data Engineer – Business Intelligence | IBM **|** Bangalore, India Nov 2019 – May 2021

- Enhanced data pipelines and ETL efficiency by 20% with Apache NiFi and Talend, aligning with IBM's emphasis on data integration and quality improvement.
- Managed complex SQL, MySQL, and Oracle databases, leveraging Tableau and Power BI alongside IBM Cognos for comprehensive data visualization, leading to a 25% improvement in analytics reporting speed.
- Spearheaded the deployment of a cloud-based data warehousing solution using IBM Cloud Pak for Data, boosting analytical processing capabilities by 30% and supporting enterprise-wide decision-making.

SKILLS

- **Languages**: Python, R, C/C++, SQL, HTML/CSS, Java Script, MATLAB.
- **Technologies/Tools**: MS Excel, AWS, Azure, Tableau, Keras, Power BI, JIRA, Git, OpenCV, Jupyter, Big Query, Hadoop, Snowflake, SAS.
- **ML Libraries**: NumPy, Pandas, Keras, TensorFlow, Plotly, OpenCV, Scikit-learn, IBM Watson Studio.
- **Databases**: MySQL server, NoSQL, MongoDB, IBM Db2, Azure.
- **Additional:** Strong analytical and problem-solving abilities, excellent communication skills, cross-functional relationships.

PROJECTS

- **MEDICARE HEALTHCARE INSURANCE PREDICTION** – RESEARCH PROJECT(Statistics, ML Models, Big Query) Oct 2023 – Jan 2024

Harnessed the power of statistical modeling and machine learning, including Random Forest and XGBoost, to cut through complex healthcare data. This approach yielded a 14.41% decrease in Mean Squared Error and a 22% reduction in prediction errors, setting a new standard for accuracy in our predictive analytics. The project's success laid the groundwork for innovative, data-informed healthcare insurance policies, showcasing the potential of data science to revolutionize patient care strategies.

- **SALES FORECASTING AND INVENTORY OPTIMIZATION** – (Statistics, ARIMA, Time-series) Apr 2023 – Sept 2023

Utilizing Python, Matplotlib, and Seaborn enabled comprehensive exploratory data analysis that unveiled pivotal sales trends, seasonal patterns, and correlations, resulting in an 18% increase in revenue. Additionally, intricate time-series forecasting models were engineered, merging ARIMA and exponential smoothing techniques, which led to a 15% reduction in forecasting errors, thereby enhancing the precision of sales predictions and aiding strategic decision-making.

- **NEW DATA PIPELINE DEVELOPMENT FOR IBM's WATSON AI PLATFORM**(IBM Cloud, Watson Studio, Python, SQL)

A highly efficient data pipeline was implemented to support IBM's Watson AI platform, resulting in a 40% enhancement in performance and scalability, which in turn accelerated and refined the accuracy of data analysis. Collaboration with a team of seasoned engineers was key in the meticulous refining of the pipeline's architecture, seamlessly integrating ETL processes and effectively resolving 98% of complex data-related challenges.

ACHIEVEMENTS

- Azure Fundamentals and DP – 203 Associate Data Engineer Certificate from Microsoft Azure, Tableau certified Data Analyst from Tableau.