

Khaja Mujahiddin Mohammed

Email: khajamujahiddin@gmail.com | LinkedIn: linkedin.com/in/khajamujahiddin | GitHub: github.com/khajamujahiddin

PROFESSIONAL SUMMARY

- Data Scientist & Cloud Engineer with 3+ years of experience designing and deploying end-to-end data workflows, machine learning models, and cloud-based analytics solutions.
- Specialized in NLP, Computer Vision, MLOps, and real-time data engineering pipelines using AWS, Python, and Spark.
- Proven track record in delivering measurable business outcomes - automating 80% of manual data tasks, improving forecast accuracy by 15%, and reducing decision-making time by 30%.
- Collaborative team player adept at bridging business needs with technical execution to deliver cost-optimized, production-ready solutions.

CORE COMPETENCIES

- Technical: Python, SQL, PyTorch, TensorFlow, Scikit-learn, Hugging Face Transformers, OpenCV, Spark, Apache Airflow, dbt, MLflow, Docker, Kubernetes
- Cloud & Data: AWS (EC2, S3, RDS, Redshift, EMR), Snowflake, MySQL, Data Modeling, ETL Pipelines, Data Warehousing, Streaming Data Processing
- Analytics & BI: Power BI, Tableau, DAX, KPI Dashboards, Advanced Excel, Data Visualization, Predictive Analytics
- Soft Skills: Agile Development, Team Leadership, Cross-functional Collaboration, Problem-Solving, Communication

PROFESSIONAL EXPERIENCE

AWS Cloud Practitioner - EduBridge Learning Pvt. Ltd. | Feb 2023 - May 2023 | Remote

- Improved data processing efficiency by 25% by designing automated ETL pipelines in Python.
- Leveraged AWS EC2, S3, RDS, and Redshift to optimize data workflows, reducing system load by 20%.
- Developed cloud-native, scalable architectures to support large-scale analytics.
- Integrated Apache Airflow and dbt for automated, scheduled workflows.

Power BI Developer - SRIK Consulting Services Pvt. Ltd. | Sep 2020 - Feb 2023 | Hyderabad, India

- Created 20+ interactive Power BI dashboards integrated with MySQL and Excel.
- Automated data pipelines with SQL scripts and Power Query, reducing manual errors.
- Developed KPI dashboards improving operational visibility by 30% and enabling faster decision-making.
- Implemented advanced DAX functions for custom business metrics and time-series analysis.

Data Analytics Intern - ExcelR Solutions | Oct 2022 - Jan 2023 | Remote

- Designed interactive dashboards (Power BI, Tableau) for KPI tracking and performance trends.
- Improved data extraction accuracy by 30% using optimized SQL queries.
- Reduced reporting time by 25% through statistical and visual data analysis.

PROJECTS

Fake News Detection (Multimodal ML)

- Built a multimodal machine learning model combining NLP-based text features and metadata to classify news as real or fake.
- Achieved 85% classification accuracy using TensorFlow, Hugging Face Transformers, and Scikit-learn.
- Applied advanced NLP preprocessing techniques including tokenization, lemmatization, and stopword removal.
- Automated data preprocessing and retraining pipelines using Apache Airflow for scalable MLOps.
- Deployed the model to Hugging Face Spaces for real-time public inference.
- Skills: Python, TensorFlow, Hugging Face Transformers, Apache Airflow, MLOps, NLP, Scikit-learn, Power BI

Self-Driving Car: Pedestrian & Cyclist Segmentation

- Developed a U-Net semantic segmentation model using TensorFlow for real-time detection of pedestrians and cyclists in urban environments.
- Annotated large-scale dataset and performed image preprocessing with OpenCV.
- Enhanced precision and recall for safer autonomous vehicle navigation.

Khaja Mujahiddin Mohammed

Email: khajamujahiddin@gmail.com | LinkedIn: linkedin.com/in/khajamujahiddin | GitHub: github.com/khajamujahiddin

- Integrated outputs into evaluation pipelines for continuous performance monitoring.
- Skills: Python, TensorFlow, U-Net, Computer Vision, OpenCV, Data Annotation, Deep Learning

E-Commerce Sales Dashboard

- Designed an annual sales reporting dashboard in Power BI for 2017-2020 to analyze customer behavior and sales trends.
- Identified a 43% sales loss in 2020 compared to 2019, enabling strategic recovery planning.
- Built visualizations including Top/Bottom Customers, Top Products, and Profit by Category.
- Applied data modeling and warehousing techniques to track KPIs and trends.
- Skills: Power BI, Data Modeling, Advanced Excel, DAX, Data Warehousing, Business Intelligence

Customer Segmentation Using Machine Learning

- Implemented K-Means clustering for customer segmentation based on purchasing patterns.
- Conducted feature engineering and PCA for dimensionality reduction and better cluster visualization.
- Visualized segmented clusters using Power BI dashboards for actionable marketing insights.
- Improved targeting efficiency by 10% through data-driven segmentation.
- Skills: Python, Scikit-learn, Pandas, NumPy, K-Means Clustering, PCA, Power BI

Street View House Numbers Dataset Analysis

- Developed a CNN model in TensorFlow to classify digits from street images, achieving 93% accuracy.
- Applied image preprocessing including normalization and augmentation to improve robustness.
- Implemented object-oriented programming principles for modular and reusable code.
- Evaluated model performance using confusion matrices and classification reports.
- Skills: Python, TensorFlow, CNN, Deep Learning, Image Processing, OOP, Model Evaluation

EDUCATION

- Master's in Data Science - University of New Haven, CT, USA | Aug 2023 - May 2025
- Bachelor's in Mechanical Engineering - [Your University], India

CERTIFICATIONS

- AWS Cloud Practitioner - EduBridge
- Google Data Analytics - Coursera
- HackerRank SQL - 5 Star Certification
- Data Analytics & Visualization Virtual Experience - Forage