

USMAN JADOON

Acworth, GA 30102 | 404-798-7473 | usmanzki@gmail.com | [linkedin.com/in/usmanjadoon/](https://www.linkedin.com/in/usmanjadoon/)

OBJECTIVE

Data Scientist with 3+ years of enterprise analytics experience supporting Microsoft and PwC, and an M.S. in Data Science and Analytics (**GPA 3.96**). Experienced in designing, evaluating, and deploying machine learning models for predictive analytics and NLP applications. Proven record of automating large-scale workflows (**40,000+ hours saved annually**) and delivering data-driven solutions that drive measurable business impact.

TECHNICAL SKILLS

Programming & Data Tools: Python, SQL, R

ML: Scikit-learn, TensorFlow, PyTorch, Random Forest, Regression, Classification, Clustering, SARIMA, NLP, LLMs, CNNs

Modeling Techniques: Cross-Validation, Hyperparameter Tuning, Model Evaluation (RMSE, R^2 , ROUGE)

Data Processing: Pandas, NumPy, PySpark, Vaex

Visualization: Matplotlib, Plotly, Power BI, Tableau

Databases: MySQL, MongoDB, Neo4j

Workflow Tools: Alteryx

CAREER EXPERIENCE

Data Analyst – Tata Consultancy Services (Microsoft Client – Redmond, WA)

May 2020 – Feb 2021

- Analyzed **1M+** employee access logs using SQL & Python to detect anomalies, reducing GDPR compliance risk exposure by **20%**.
- Designed anomaly detection logic using statistical thresholds & historical baseline comparisons to identify abnormal access behavior.
- Enhanced data governance through risk-based access auditing, compliance reporting, and anomaly investigation.

Data Analyst – Tata Consultancy Services (PricewaterhouseCoopers Client – Tampa, FL)

Mar 2019 – Apr 2020

- Analyzed enterprise data storage patterns using Alteryx and Tableau to uncover inefficiencies and optimize resource allocation.
- Engineered automated data validation and workflow pipeline using Alteryx and SQL, eliminating **40,000+ manual hours** annually across **2,000 employees**.
- Improved end-user **efficiency by over 40%**, contributing key insights to align data strategy with business objectives.
- Developed interactive dashboards using Tableau and Alteryx for executive-level reporting.

Data Analyst Intern – Fulton County Government Center – Atlanta, GA

Aug 2018 – Dec 2018

- Analyzed criminal justice datasets to identify seasonal, operational, and demographic factors driving jail population fluctuations.
- Built ETL workflows and a centralized data warehouse, reducing data redundancy by **14%** and improving data accessibility.
- Performed exploratory and statistical analysis using R, Excel, Tableau, and Power BI to uncover insights and inefficiencies.
- Developed **dynamic dashboards** and executive-level reports, translating analytical findings into data-driven recommendations.

ACADEMIC PROJECTS (M.S.)

E-Commerce Chatbot (LLM & NLP | GPT-3.5, LangChain, Flask)

- Built a RAG pipeline using GPT-3.5 and Chroma Vector Database to answer domain-specific e-commerce queries.
- Embedded **~5,000 FAQ** documents and evaluated performance on **200 test queries**, achieving **~80% relevance accuracy**.
- Reduced response latency by **~20%** through optimized vector retrieval and prompt engineering.
- Developed and deployed a Flask-based web interface to serve model responses in real time.

Airbnb Price Prediction (Time Series & ML Models)

- Performed exploratory data analysis and feature engineering on **~10,000 Airbnb listings**.
- Built and compared SARIMA, Linear Regression, and Random Forest models.
- Performed k-fold cross-validation and hyperparameter tuning using GridSearchCV to optimize model performance.
- Random Forest achieved the best performance, capturing non-linear relationships across features with **$R^2 \sim 0.78$** .
- Deployed model via Flask app, enabling CSV-based automated price predictions.

News Text Summarization Using T5 Transformer (NLP)

- Fine-tuned T5 Transformer model using HuggingFace Transformers and PyTorch for abstractive news summarization.
- Preprocessed **~1,500 articles** with tokenization, lemmatization, and text normalization.
- Evaluated model performance using ROUGE metrics (**ROUGE-1: 0.42 | ROUGE-2: 0.28 | ROUGE-L: 0.38**).
- Optimized training through hyperparameter tuning and validation-based model selection.

EDUCATION

M.S. Data Science and Analytics (GPA 3.96) – Georgia State University, Atlanta, GA

Aug 2022 – Dec 2023

B.B.A. Computer Information Systems (GPA 3.36) – Georgia State University, Atlanta, GA

Aug 2014 – Dec 2018