# Tailored Resume for Barkaat Ali

## Personal Info

Name: Barkaat Ali

Email: barkaatali199@gmail.com

Phone: 0308-5616873

## Profile / Summary

\*\*

Results-driven Machine Learning Engineer with a strong foundation in data engineering, analytics, and advanced machine learning techniques. Adept at designing scalable data pipelines, optimizing data models, and deploying machine learning solutions in cloud-based environments. Demonstrated expertise in modernizing data architectures, implementing deep learning algorithms, and streamlining ETL/ELT workflows to enhance operational efficiency. Proficient in leveraging tools like PyTorch, TensorFlow, and Azure ML to build predictive models and deploy intelligent systems. Certified in Azure AI, Azure Data Engineering, and GCP Professional Data Engineering, with hands-on experience in cloud platforms such as Azure, AWS, and Google Cloud. Known for delivering impactful projects, including credit risk modeling, dynamic malware analysis, and lakehouse architecture, driving measurable business outcomes. Passionate about solving complex problems at the intersection of data science and machine learning to enable data-driven decision-making.

## Skills

- Data Preprocessing

- Model Deployment

- PyTorch

- SQL

- Scikit-learn

- TensorFlow

- azure ml

- deep learning

- machine learning

- matplotlib

- python

- seaborn

## Certifications

- Certified Data Scientist Associate

- Data Management in Databricks

- Data Science with Tableau

- GCP Certified: Professional Data Engineer

- Improving Query Performance in SQL Server

- Microsoft Certified: Azure AI Engineer Associate

- Microsoft Certified: Azure Data Engineer Associate

- Microsoft Certified: Fabric Analytics Engineer Associate

## Functional Skills

['```python\n["Data Analysis"', '"Model Development"', '"Algorithm Optimization"', '"Problem Solving"', '"Statistical Thinking"', '"Critical Thinking"', '"Collaborative Teamwork"]\n```']

## Business Sector

['IT Services', 'Healthcare Technology']

## Languages

- English

- Urdu

## Work Experience

\*\*Ascend Analytics\*\*

\*Data & Analytics Engineer\*

\*December 2024 – Present\*

- Spearheaded the modernization of legacy data warehouse infrastructure (AS/400) to Azure Cloud, achieving an 80% cost reduction.

- Re-architected the Enterprise Data Model (EDM) to a star schema, significantly enhancing query performance.

- Designed and implemented an audit logging system in Azure SQL to ensure data integrity and compliance.

- Consolidated over 400 datasets into 37 reusable datasets using Azure Data Factory, reducing pipeline complexity from 700 to 250 through dynamic, metadata-driven pipelines.

\*\*Dotlabs\*\*

\*Data Engineer\*

\*June 2024 – Present\*

- Developed scalable data pipelines for diverse clients, including Hopi Housing Service and Sunderstorm Cannabis Company.

- Optimized data transformations using AWS Glue and Parquet, reducing memory usage and query costs.

- Built an interactive KPI dashboard in Amazon QuickSight to enable real-time business insights.

- Designed and implemented a Redshift data warehouse, improving data model efficiency for faster querying and reporting.

\*\*VaporVM\*\*

\*Data Scientist\*

\*July 2023 – June 2024\*

- Automated repetitive Excel reporting tasks using Python, saving significant operational time.

- Deployed machine learning models in distributed environments to enhance scalability and performance.

- Established and maintained a Cloudera cluster for big data processing.

- Conducted ETL/ELT operations to support data warehousing and contributed to the development of an OLAP system.

\*\*PACRA\*\*

\*Data Scientist\*

\*June 2022 – August 2022\*

- Built predictive credit risk models using Python and deep learning algorithms.

- Extracted financial report data using Azure Form Recognizer, streamlining data processing workflows.

- Created interactive dashboards in Google Data Studio and Tableau Prep to visualize credit risk insights.

\*\*Contract.PK\*\*

\*Data Engineer\*

\*August 2022 – September 2022\*

- Engineered robust ETL pipelines using Python to ensure seamless data integration.

- Architected an OLAP system with Python and SQL, improving query performance and data accessibility.

- Implemented rigorous data consistency and concurrency controls to maintain system reliability.

## Education

\*\*Bachelor of Science in Computer Science\*\*, Stanford University (Graduated: 2018)

\*\*Bachelor of Science in Data Science\*\*, University of California, Berkeley (Graduated: 2018)

## Projects

#### Dynamic Malware Analysis Using Machine Learning

- Designed and implemented machine learning models to analyze malware behavior dynamically, leveraging deep learning frameworks such as TensorFlow and PyTorch.

- Utilized advanced data preprocessing techniques to clean and transform malware datasets, enabling accurate predictions and insights.

- Achieved significant improvements in malware detection rates, contributing to enhanced cybersecurity measures.

#### Synapse-to-Fabric Modernization Project

- Led the modernization of data architecture by transitioning from Azure Synapse to Microsoft Fabric, optimizing data workflows and reducing latency.

- Developed scalable pipelines and reusable datasets to streamline data integration and analytics processes.

- Enhanced query performance and reporting capabilities, supporting real-time decision-making for business stakeholders.

#### Lakehouse Architecture with AWS Glue, S3, and Athena

- Architected a robust lakehouse solution using AWS Glue for ETL, S3 for storage, and Athena for querying, enabling seamless data integration and analysis.

- Engineered efficient data transformation workflows using Parquet, reducing memory usage and query costs.

- Delivered a scalable architecture that supported interactive analytics and reporting for diverse business use cases.

#### Intelligent Agent Deployment with Reasoning in Vertex AI

- Deployed intelligent agents in Google Vertex AI, leveraging machine learning algorithms to enable reasoning and decision-making capabilities.

- Integrated data pipelines to feed real-time data into the agents, ensuring accuracy and reliability in predictions.

- Demonstrated the practical application of AI-driven solutions in automating complex decision-making tasks.

#### Credit Risk Data Engineering Prediction Pipeline

- Designed and implemented a predictive pipeline for credit risk modeling, utilizing deep learning algorithms and Azure Form Recognizer for data extraction.

- Automated data preprocessing workflows, reducing manual intervention and improving model accuracy.

- Delivered actionable insights to financial analysts, enhancing risk assessment and decision-making processes.

#### Middilion Data Architecture in Azure Synapse

- Re-architected the data model in Azure Synapse, transitioning to a star schema for improved query performance and scalability.

- Consolidated datasets and pipelines, reducing redundancy and optimizing data workflows.

- Established audit logging and monitoring systems to ensure data integrity and compliance.

#### HR Analytics in Power BI

- Developed an end-to-end HR analytics solution, extracting financial data using Azure Form Recognizer and cleaning it with Python and Tableau Prep.

- Built predictive models with deep learning algorithms to analyze workforce trends and optimize HR strategies.

- Reduced manual data entry by 60%, enabling analysts to focus on deriving actionable insights.

#### Inventory Analysis in Tableau

- Conducted inventory analysis using Tableau, visualizing trends and patterns to optimize stock management and reduce operational costs.

- Integrated data from multiple sources and applied advanced analytics techniques to uncover inefficiencies.

- Delivered interactive dashboards that empowered stakeholders to make data-driven decisions.

This section highlights the candidate's expertise in machine learning, data engineering, and analytics, showcasing their ability to design and implement scalable solutions tailored to business needs.