Prannay Khushalani, B.Tech | M.Sc.

469-545-6814 | prannay.khushalani5@gmail.com |Dallas, TX |Linkedin Profile | Portfolio

Professional Summary

Data Science professional with **advanced expertise in Data Analytics, Predictive Modeling, Machine Learning, and Strategic Decision-Making**. Skilled in leveraging **SQL, Python, Power BI, and Agile methodologies** to drive operational excellence and deliver insightful analytics solutions. Demonstrated success in **optimizing processes, enhancing business performance,** and **implementing innovative data-driven strategies** in fast-paced environments. Seeking impactful roles as a **Data Analyst, Business Analyst, AI Engineer or Data Scientist.**

Technical Skills

Programming Languages: Python (Intermediate), SQL (Advanced), R (Intermediate)
Data Science & Machine Learning Libraries: pandas, NumPy, Matplotlib, SciPy, Scikit-learn, TensorFlow, Keras, PyTorch, Seaborn, Plotly, NLTK, SpaCy, XGBoost, PySpark
Data & Visualization Tools: Power BI (DAX, Power Query), Tableau, Alteryx, Excel, Hadoop, Hive, Impala, Sqoop, Spark, Snowflake, Databricks
Cloud Technologies: Azure ML, Azure Databricks, Azure Synapse Analytics, AWS SageMaker, AWS Glue, AWS Athena, AWS EMR,
Generative AI Skills: Prompt Engineering, Large Language Models (LLMs), OpenAI API, Fine-tuning Models, Chatbot Development, Retrieval-Augmented Generation (RAG) Pipelines
Key Competencies: Data Analysis, Machine Learning, Data Mining, ETL/ELT, A/B Testing, Time Series Forecasting, Database Management, Data Visualization, Agile & Scrum Methodologies

Work Experience

Business Analyst Intern Jan 2023 - Jul 2023

Global Dreamz Edutech (Top Educational Consultancy with 5000+ students) *Pune, IN*

- Automated **SQL-based data extraction** and transformation pipelines from fragmented application sources, reducing weekly processing time by ~7 hours and enabling **scalable downstream analytics**
- Developed interactive Power BI dashboards to monitor student application trends and acceptance success rates, improving client satisfaction scores by ~25%
- Collaborated cross-functionally to identify bottlenecks in data workflows and implemented **modular ETL scripts**, enhancing agility in reporting and analysis for **5,000+ student records**

Business Analyst Intern Jul 2022 - Dec 2022

Mangala Industries (Automobile parts Manufacturer with 50+ Million USD Annual Revenue) *Pune, IN*

- Engineered supply chain analytics pipelines using **SQL** and **Power BI** to monitor procurement and delivery KPIs, boosting operational efficiency by ~15% YoY.
- Led the automation of inventory management logic via rule-based **SQL** processes, reducing component reorder time by ~40% and saving $15,000 from a $150,000 annual inventory spend.
- Built a reporting framework to track vendor performance and part delays, enabling proactive issue resolution and improving supplier reliability metrics.

Academic Projects

JSOM Chatbot - OWLIE *NLP Engineer | Retrieval-Augmented Generation (RAG), Web Scraping, LLM, Sentence Transformers* Jan 2025 - May 2025

- Engineered an NLP-driven chatbot utilizing Retrieval-Augmented Generation (RAG), Sentence Transformers, and LLAMA3-70B LLM using Groq API to deliver precise, context-aware responses,improving response accuracy by **~85%** and reducing average response latency by **~95%** to 3 seconds per answer
- **Designed** scalable web-scraping pipelines and **implemented** advanced text preprocessing and FAISS vector indexing, increasing relevant context retrieval rates by **~36%** across over **5,000+ JSOM webpages**
- Collaborated directly with the **Senior Assistant Dean** for Graduate Programs at UT Dallas JSOM, enhancing chatbot accuracy, scalability, and real-time response quality

Credit Risk Evaluation Model *Machine Learning, Neural Networks, Data Analytics* Jan 2024 - May 2024

- Engineered a machine learning-based credit risk model in Python, **projecting reduction** in loan default rates **from 25.8% to 3%** based on historical data, an **88.4%** improvement in risk assessment accuracy using XGBoost and Neural Networks
- Validated data pipelines using SQL to ensure high data quality during feature engineering and model deployment

Research Projects

ASR for Regional languages using Fine-Tuned Wav2Vec2 Model *Python,NLP* Feb 2022 - Jul 2022

- Developed an **Automatic Speech Recognition system** for Tamil, a low-resource language, **reducing Word Error Rate from 85.9% to 61.3 %** using pre-training and custom tokenization
- Executed SQL queries to efficiently manage and retrieve audio data from structured databases, enabling cross-lingual learning and self-supervised training using **XLSR Wav2Vec2** to optimize speech-to-text accuracy
- Published findings in ”Advances in Intelligent Systems and Computing” (Springer Nature Singapore, 2023)

Predictive Analysis of multiple diseases using ensemble learning Aug 2021 - Dec 2021

- Implemented ensemble learning techniques (SVC, Naive Bayes, Decision Trees) to predict diseases**, increasing accuracy from 95% to 99%** through bagging and boosting
- Extracted and transformed medical record data using SQL, **reducing data processing time by ~42%**, and presented actionable insights on disease prediction model performance to the stakeholders
- Published research in IJRTE that demonstrated the application of ensemble techniques for multi-disease prediction; **findings revealed a ~25% increase in predictive accuracy,** paving the way for future healthcare innovations

Education

University of Texas, Dallas Aug 2023 - May 2025
Masters of Science in Business Analytics and AI
Graduate Certificate in Applied Machine Learning
Graduate Certificate in Business Decision Analytics

Vishwakarma Institute of Technology, Pune Aug 2019 - May 2023
Bachelors of Technology in Information Technology