

# Ameer Sohail Shaik

+1 (813)-817-1935 | [skameersohail148@gmail.com](mailto:skameersohail148@gmail.com) | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

## SUMMARY

---

Aspiring Data Scientist and MS Data Science student at University of Maryland with hands-on experience building production-level ML systems that delivered significant portfolio performance improvements. Combines rigorous academic foundation in Machine Learning and Statistical Analysis with real-world application of Python, TensorFlow.

## EXPERIENCE

---

### Machine Learning Engineer

*Procedit - (Remote)*

August 2024 – August 2025

*Barcelona, Spain*

- Designed and implemented advanced time-series forecasting models using ARIMA, SARIMAX, LSTM models achieving a 98% accuracy rate for corrected target prices. This enabled actionable financial insights and supported strategic decision-making, enhancing predictive capabilities and informing investment strategies.
- Engineered automated data pipeline for trading ML models that reduced critical data update times from 60+ minutes to 5 minutes, improving operational efficiency by 50%.

## EDUCATION

---

### Master of Science in Data Science

*University of Maryland - College Park*

Aug 2025 - Present

*College Park, MD*

- Relevant coursework: Machine Learning, Statistical Analysis, Big Data Systems

### Bachelor of Technology in Computer Science and Engineering

*Vellore Institute of Technology, AP*

2020 - 2024

*AP, India*

## PROJECTS

---

### AutoML-ify | *Python, Streamlit, Machine Learning*

March 2025

- Designed and implemented end-to-end automated machine learning application streamlining data preprocessing, model selection, and evaluation without coding knowledge.
- Engineered scalable data pipeline architecture supporting multiple data formats and real-time model training capabilities.
- Integrated 5+ machine learning algorithms with automated hyperparameter tuning for non-technical users.

### Dynamic SQL Assistant | *Python, Streamlit, LangChain, Groq API, SQLite, NLP*

Feb 2025

- Developed a Text-to-SQL application using Streamlit, integrating LLM capabilities via Llama 3, Groq API and LangChain achieving a 95%+ success rate in translating natural language to SQL queries. This improved data accessibility for nontechnical users and improved query accuracy.
- Engineered a dynamic database creation system that automatically converts CSV datasets into queryable SQLite databases, streamlining data integration and reducing setup time by up to 90%. Additionally, implemented robust error handling for SQL queries, reducing query failures by 60% and enhancing system reliability.

## TECHNICAL SKILLS

---

**Programming Languages:** Python (Advanced), Java (Intermediate), R (Basic), SQL (Advanced)

**Databases:** MySQL, PostgreSQL, MongoDB

**Tools:** Jupyter Notebook, Google Colab, Git, Jira, Airflow, Spark, Matlab, Visual Studio, Microsoft Excel, Confluence

**Data Science Libraries:** Pandas, NumPy, Matplotlib, Scikit-learn, TensorFlow

**Data Analysis:** EDA (Exploratory Data Analysis), Statistical Analysis, Data Visualization

**Cloud Platforms:** AWS S3, AWS Redshift