

Pranitha Velusamy Sundararaj

✉ pranithavs2310@gmail.com

☎ (313) 414-4989

🌐 linkedin.com/in/pranitha-v-s

SUMMARY

Data Scientist with expertise in **predictive modeling**, **cloud-native solutions**, and **large-scale data processing**. Proficient in **Python**, **SQL**, **AWS** (Lex, Lambda, Rekognition), and **ML frameworks** (TensorFlow, Scikit-Learn). Skilled in **ETL workflows**, **Spark-driven analytics**, and end-to-end **deployment of AI tools**. Passionate about transforming raw data into actionable insights for scalable, real-world impact.

SKILLS

- **Programming Languages:** Python, SQL (MySQL, PostgreSQL), R Programming, PySpark
- **Frameworks & Tools:** Pandas, Hadoop, Apache Spark, Docker, NLP
- **Cloud Platforms:** AWS (S3, Lambda, Rekognition, API Gateway, DynamoDB, Lex, IAM, Kinesis, VPC), Snowflake
- **Data Visualization:** Tableau, Power BI, Matplotlib, Seaborn
- **Machine Learning:** Scikit-Learn, TensorFlow, Keras
- **Developer Tools:** Jupyter Notebook, GitHub, RStudio, Google Colab, VS Code, MS Word, Excel, PowerPoint

WORK EXPERIENCE

Project Intern at SAMSUNG PRISM, India

Jul 2021 - Mar 2022

- Developed **Bi-Directional LSTM RNN** with **Manhattan Distance**, achieving **81.62% similarity accuracy** on Quora dataset.
- Preprocessed **417K+ sentence pairs** using **NLTK** for tokenization, cleaning, and lemmatization.
- Trained model with **Keras embeddings**, deployed using **TensorFlow Lite** for on-device efficiency.
- Integrated model into Android app using **Stanford CoreNLP** for dynamic MWUC sentence handling.
- Achieved **95.65% accuracy** in semantic similarity and improved voice assistant usability via MWUC reformulation.

PROJECT EXPERIENCE

DEEP AUDIO CLASSIFIER – CAPUCHIN BIRD CALL DETECTOR

Jan 2025 - Apr 2025

- Built a binary **CNN in Python/TensorFlow** to detect bird calls, achieving **96% accuracy** and **75% precision**.
- Preprocessed rainforest audio using **Librosa** with spectrograms, noise reduction, and data augmentation.
- Applied **early stopping**, **learning rate scheduling**, **L2 regularization**, and **dropout** to reduce overfitting.

STUDENT PERFORMANCE PREDICTION

Aug 2024 - Dec 2024

- Engineered high-accuracy **MLR model (98.8% R²)** using **Python/Scikit-Learn** for performance prediction.
- Identified key predictors via **EDA/correlation analysis** leveraging **Matplotlib** and **Tableau** visualizations.
- Implemented robust **data preprocessing (Pandas)** and model validation (**VIF**), ensuring reliable results.

DINING CONCIERGE CHATBOT

Aug 2024 - Dec 2024

- Deployed serverless chatbot using **AWS Lex (NLP)**, **Lambda (Python)**, and **API Gateway** for restaurant suggestions.
- Integrated **Yelp API** and **DynamoDB (NoSQL)**, enabling real-time data storage and dynamic retrieval.
- Developed Lambda functions orchestrating conversational AI flow and backend API resource interactions.

INTELLIGENT PHOTO ALBUM SEARCH APPLICATION

Aug 2024 - Dec 2024

- Built intelligent photo search using **AWS Rekognition**, **Lex (NLP)**, and **ElasticSearch** for backend indexing.
- Implemented **serverless Lambda (Python)** triggered by **S3** events for automated photo metadata indexing.
- Enabled natural language text/voice search via **API Gateway** querying **ElasticSearch** photo indices.

SMART HEALTH DISEASE PREDICTION

Jan 2024 - Apr 2024

- Developed **Naive Bayes** classifier using **Python/Scikit-Learn** for disease prediction based on symptom data.
- Executed **data preprocessing** and feature scaling on the symptom dataset utilizing **Pandas** library functions.
- Achieved **92.7% prediction accuracy** by evaluating model performance with Scikit-Learn metrics effectively.

RECOMMENDATION SYSTEMS

Aug 2023 - Dec 2023

- Implemented **Popularity** and **Collaborative Filtering (ALS)** recommenders using **Apache Spark/MLlib** framework.
- Processed **7.8M Amazon ratings (Big Data)** utilizing Spark's efficient distributed computing capabilities.
- Achieved optimized ALS model performance (**RMSE 0.19**) through evaluation within the **Spark** ecosystem.

EDUCATION

Master of Science in Data Science

University of Michigan- Dearborn, USA

Aug 2023 - Apr 2025

GPA: 3.74

Bachelor of Engineering in Computer Science

PSG College of Technology- Coimbatore, India

Aug 2019 - Jul 2023

GPA: 3.33

PUBLICATIONS

ML, NLP, Capsule Development | Bixby Utterance Reformulation (BUttr)

3rd International Conference on Engineering, Social Sciences, and Humanities

30th - 31st Mar 2023