Shalin Parikh

Evanston, IL | 312-973-6102 | shalin.parikh@northwestern.edu | linkedin.com/in/shalinsparikh | github.com/Shalinparikh

Education

Northwestern University Sep 2021 – Dec 2022
Master of Science in Artificial Intelligence, GPA: 3.88/4.0 *Evanston, IL, US*

- **Relevant Coursework:** Machine Learning, Natural Language Processing, Data Science Seminar, Artificial Intelligence, Deep Learning, Computer Vision, Deep Reinforcement Learning

Pandit Deendayal Petroleum University Aug 2016 – Jul 2020
Bachelor of Technology in Information and Communication Technology, , GPA: 3.5/4.0 *Gandhinagar, India*

- **Relevant Coursework:** Software Engineering, Data Structures and Algorithms, Object Oriented Programming

Technical Skills

Programming Languages: Python, C++, SQL, R, JavaScript, HTML, CSS
Databases: MySQL, PostgreSQL, MongoDB
Frameworks and Libraries: TensorFlow, Keras, PyTorch, Caffe, Scikit-learn, OpenCV, Librosa, PySpark, MapReduce
Developer Tools: Amazon Web Services (AWS), Linux, Git, Tableau, Docker, Apache Spark, Hadoop, Hive, REST API

Experience

Omdena Mar 2021 – Jul 2021
Junior Machine Learning Engineer *Ahmedabad, India*

- Directed **Exploratory Data Analysis (EDA)** and **Data Visualization** of the Linguistic, Cognitive, and Logical exercise parameters of memory, speed, picture recognition, and reasoning deploying **Tableau** and **Python**.
- Created **Collaborative Filtering** and **Content-Based Recommendation Systems** to recommend people ideal mental exercises based on skill level and similar users with a **Mean Average Precision (MAP)** of **73.2%**.

Pandit Deendayal Petroleum University Sep 2020 – Dec 2020
Data Science Intern *Gandhinagar, India*

- Extracted and visualized the audio features executing **Fast Fourier Transform (FFT)**, **Speech Synthesis**, and **Natural Language Processing** techniques on the VoxCeleb and LibriSpeech audio datasets.
- Created an audio translation system to produce novel audio in the same user's voice in different languages.
- Built **CNN-RNN** models to compare model-generated audio and original audio and attained an accuracy of **76%**

V2Solutions Jun 2019 – Aug 2019
Data Science Intern *Navi Mumbai, India*

- Designed an end-to-end solution leveraging ML clustering algorithms **DBSCAN** and **Agglomerative Clustering** to cluster legal summary documents in different categories attaining a baseline evaluation of **82.7%** accuracy.
- Improved performance of primary prediction pipeline by **45%** through parallelization and code deduplication.

Projects

Abstractive Text Summarization | *Python, TensorFlow, PyTorch, Transformers, NLTK* Jan 2022 – Mar 2022

- Implemented an **abstractive** text summarizer on the **CNN/ Dailymail** Dataset from **HuggingFace** working on over 300k news articles to give a concise and coherent summary of the articles.
- Built a **seq-to-seq model** utilizing **RNNs** as encoder and **LSTMs** as decoder and pretrained simple **T5 (Text-to Text Transfer Transformer)** model to achieve high **ROUGE-1** and **ROUGE-2** scores on the articles.

Chicago Police Database Analysis | *Python, PostgreSQL, Tableau, D3.js, PySpark* Sep 2021 – Dec 2021

- Analyzed Chicago Police Department Database on **PostgreSQL** and produced interactive visualizations using **Tableau** and **D3.js** to get insights on the number and categories of complaints about police officers.
- Conducted **Sentiment Analysis** on Police complaint report statements and mapped police officers with sustained complaints by doing **Graph Analytics** on data leveraging **Graphframes**, **PageRank** algorithm, and **PySpark**.

Sentiment Analysis using LSTMs | *Python, Keras, Pandas, Numpy, Matplotlib* Jan 2020 – Jun 2020

- Devised an algorithm to analyze sentiments by incorporating **word2vec** and **stacked bi-directional LSTM**.
- Utilized **Continuous Bag of Words(CBOW)** to obtain embeddings and programmed a binary **Softmax Classifier** accomplishing a high accuracy of **91%** on the prediction of sentiment orientation.

Twitter Emoticon Predictor | *Python, Gensim, Spacy, NLTK, Scikit-learn, Matplotlib* Jan 2020 – May 2020

- Implemented multi-class classification to forecast emoticons for Tweet sentiments operating Neural Networks.
- Formulated **SMOTE analysis** to eliminate data imbalance and generated embeddings using **Stanford gloVe** as input to **bi-directional LSTMs** to achieve an accuracy of **82.3%**.