GHAIDA ALSHIDDI

gsa200000@utdallas.edu| https://galshiddi-coder.github.io/CS-4395-NLP-/

Skills And Technologies

Technical Skills: Python, SKlearn, Information Extraction, TF-IDF, Text processing and parsing, Ngrams, Bigrams, Sdntence parsing, WordNet, NLTK, , part of speech tagging, training and testing dataset, Vector space (TF, TF-IDF, cos similarity) BeautifulSoup, WebCrawler, numpy, pandas, seaborn, SKlearn, Naïve Bayes, log regression, sequential model, RNN model, CNN model, and embedding layer and Neural networks and Text Classification

Soft Skills: communication, time management, teamwork, problem-solving, creativity, and leadership

Some Academic Projects

Sports Chatbot Application (Python, Text processing and parsing, cos similarity, term frequency)

- Designed a chatbot that that uses Cosine Similarity, term frequency, and text processing and parsing to assist and responds to user request
- Designed a knowledge base the chatbot uses to provide information to the use, the program gets the best results to user by using the cosine similarity and term frequency to analyze user text using vector space and information extraction

Sports Sports web crawler (Python, Text processing and parsing, term frequency, BeautifulSoup)

- Designed a program that that uses term frequency, and text processing and parsing to create a knowledge base using BeautifulSoup and a starter website
- The program cleans the data it got from the websites and highlight important terminologies and uses dictionary where the term is the key and the data that includes the term is the value

Text classification programs (Python, machine learning algorithms, deep learning algorithms)

- Designed a program that uses machine learning algorithms to train and test large dataset and then evaluate the results. The algorithm used: numpy, pandas, seaborn, SKlearn, Naïve Bayes, log regression, and Neural networks
- Designed a program that uses Deep Leanring algorithms to train and test large dataset and then evaluate the results. The algorithm used: sequential model, RNN model, CNN model, and embedding layer