Arkajyoti Chakraborty

9818064803 | arkajyoti09062001@gmail.com

Education

B.Tech Engineering Physics	2019-2023	Delhi Technological University	8.3 GPA(till 4th sem)
AISSCE (CBSE – Class XII)	2019	Kendriya Vidyalaya	90%
AISSE (CBSE – Class X)	2017	Kendriya Vidyalaya	10 CGPA

Experience

Vigilance AI Illinois, US(Remote)

Data Science Intern Sep 2021 – Dec 2021

- Worked on activity recognition problem on video-based data building a model to classify different activities.
- Used **CNN** feature extractor, and fine-tuned the transformer encoder model to enhance the accuracy, and tested the models in real-time.
- Worked on abnormal breathing detection pipeline, using **YOLO** for masking the video, and used the activity pipeline to detect abnormal breathing.

Calibre DTU New Delhi, In

Research Intern Jun 2021-Dec 2021

- Worked on sentiment analysis by collecting tweets over covid-19, omicron, and recently sports events like Olympics, Euro Cup, and World-t20.
- Used different vectorization(Tf-Idf, Countvectorizer) and word embeddings(Word2Vec, Glove) integrating them with machine learning and deep learning algorithms.
- Trained different models, LSTM performed the best among all with an accuracy of **89%**.

Projects

Hate Speech Classification

- Collected Twitter data, youtube comments, and Reddit data related to various events and balanced the dataset that was classified into three classes.
- Trained machine learning and deep learning models with **TF-IDF** vectorizer.
- Achieved an improved accuracy of 4% using **RNN** as compared to **SVM**.
- Github: ArkajyotiChakraborty/Hate-Speech-Sentiment-Analysis-NLP (github.com)

Olympics Sentiment Analysis

- Collected a large amount of text data from Twitter using the tweepy library.
- Used **Text-Blob**, to calculate the polarity score of the data and divided the data into three classes.
- Build machine learning and deep learning models, using **TF-IDF** vectorizer, **Word2Vec,** and **Glove** embeddings.
- Recorded an accuracy of **89%** with RNN and LSTM models and an accuracy of **86%** using Random Forest.
- Github: ArkajyotiChakraborty/Paper (github.com)

Skills

- Proficient Language: Java, Python
- Frameworks and Tools: Machine Learning, Deep Learning, Pytorch, Tensorflow, Keras

Achievements

- Research paper naming "Olympics Retweets! An analysis of public sentiments towards Olympics'20" is being submitted at Entertainment Computing.
- Achieved a global rank of 242 solving 5 out of 10 questions in Codechef August Long challenge.

I hereby declare that the details furnished above are true and correct to the best of my knowledge