Avishek Das

Linkedin: *linkedin.com/in/avi539*
GitHub: avishekdas539

Email : *avishekdas539@gmail.com*
Mobile : +91-833-785-6650
LeetCode : avishekdas539

Education

- **Indian Institute of Engineering Science and Technology, Shibpur** — Howrah, Westbengal
  Bachelor of Technology In Civil Engineering *; CGPA: 9.15/10.0* — *2018 - 2022*
  Key Courses Taken: Engineering Mathematics, Uncertainty Quantification, Structural, and Geotechnical Engineering

Professional Experience

- **Tata Consultancy Services** — Kolkata, India
  System Engineer — *Aug 2022 - Currently Working*
  Developer @ DTS (Digital Transformation Services)
  - Working on a search result relevance classifier to improve search results for e-commerce website using multiple **pre-trained Encoder-Decoder models and LLMs e.g. BERT/ Sentence-Transformer/ Universal Sentence Encoder/ LlaMa2/ VertexAI-textbison**.
  - Worked on the development of an automatic sandbox for processing unorganized data and uploading to SQL server using Python.
    Developed the core part for formatting the unorganized data and convert it to organized and processable data frames using pandas. Applied different string manipulation techniques to improve the data quality.
    Optimized the sandbox core and was able to **reduce the processing time by one-sixth** which has removed the blockers at this stage for the client.
  - Worked on text analyzing and summarizing of financial reports using token **dependency mapping, Named Entity Recognition** and Natural Language Processing. Prepared word cloud to visualize the important words depending on frequency.
  - Worked on Natural Language Processing for HSE sector to predict upcoming incidents based on historical data. Performed topic modeling on free text data to identify different causes of incidents. Used sentence-transformer to obtain similarities between many incidents and predicted upcoming incidents. Developed front and back-end of web app using Flask framework to use those models and statistics coming out from the historical data to inform users beforehand about any predicted incident.

Projects

- **STAAD-Bot a Gemini-Pro Powered RAG Based Chat Bot for STAAD.Pro Solutions** — January 2024
  - Extracted and pre-processed texts from 7000+ pages of STAAD documentation and guide books. Performed chunking and semantic vectorization using embedding-001 model by Google and stored into FAISS vector database for Retrieval Augmented Generation. Leveraged prompt engineering techniques to optimize output from Gemini-Pro model.
  - Developed an suitable user interface using streamlit for a user friendly experience to the users where they can ask question from documentation, ask question related to their own staad files.
  - Used Git and Github for version control and hosted on Streanlit hosting platform for public usage.
- **Real time Driver Alertness Monitoring Using Computer Vision** — August 2023
  - Prepossessed the image-labeled data for the data pipeline using TensorFlow Dataset API. Used custom layers on top of pre-trained MobileNetV2 model for faster classification using depth and point-wise convolution. Trained the custom layers to achieve **99.4% training accuracy & 96.8% validation accuracy.**
- **Face Detection Using VGG16 With Custom Output Layers** — June 2023
  - Generated labeled custom dataset and augmented the labeled data. Trained custom output layers for object detection over **VGG16** model. Achieved training and testing **IoU as 97% and 91%** respectively.
- **Keyword Extraction Using TextRank Algorithm** — June 2023
  - Applied TextRank algorithm to calculate the weights of each word in a given text after removing the stopwords and unwanted words based on the POS given.
- **Text Emotion Prediction Using Machine Learning** — May 2023
  - Prepossessed textual data by tokenizing, removing stop word, removing numbers, expanding short contractions, and lemmatization. Also converted text to vector using Tf-IDF and word embeddings.
  - Created classification model using DecisionTree, RandomForest algorithm with 82% and 84.1% accuracy. Trained Embedding and LSTM-based deep learning model with a training accuracy of 92.9% and validation accuracy of 89.6%.
- **Automation in Structural Engineering** — Jun 2022 - Ongoing
  - Developed a deep learning-based regression model inspired by wave-net to predict the response of buildings due to wind and gravity load.
  - Developed model to predict optimum structural section sizes for a given condition using the KNN-Regression technique with 99.5% success rate of being the optimum section.

Skills Summary

- **Programming Languages**: Python, Java, JavaScript
- **Machine Learning/ Deep Learning:**: Natural Language Processing, Computer Vision, Deep Neural Network, Clustering Based Algorithm, LLMs API/Local with Prompt Engineering
- **Frameworks/ Libraries**: tensorflow, nltk, spacy, scikit-learn, Gensim, BeautifulSoup4, numpy, pandas, pyspark, scipy, matplotlib, plotly, seaborn, flask, streamlit, SQLite3, HTML, CSS
- **Tools**: Dataiku (Core Designer Certified), Advanced Microsoft Excel, Microsoft Word, Microsoft PowerPoint
- **Version Control**: Git, GitHub

Internship

- **Summer Research Intern - IIEST, Shibpur** May 2020 - July 2020
  - Worked on vehicle detection and tracking using YOLO pre-trained model to extract traffic-related data and movement curve of traffic on some given road.

Positions of Responsibility

- **Manager In-Charge, BEing Civil - IIEST, Shibpur** January 2020 - May 2022
  - BEing Civil is a website designed to simplify the lives of engineering students. It serves as a repository for various engineering-related materials. As part of my role in the organization, I manage the website, manage the team, and keep the website updated.

Awards

- **WatsonX Assistant Challenge for GSI - IBM - 4th Position** 2023
  - Developed a chat bot using WatsonX Assistant on IBM Cloud platform by defining custom actions and their steps for banking specific industries.
- **National Award Competition (Civil) - Institute for Steel Development & Growth - 1st Position** 2021
  - Designed “Iconic Steel Roof Structure over An Open-Air Theatre and Stage" with a team of 4 people. Awarded the first position at the national level
- **Bentley Future Infrastructure Star Challenge - Bentley Systems - Selected in top 20 projects** 2021
  - Innovative Idea Presentation - Energy extraction from the turbulence of wind due to high-speed vehicle. Worked in a team of 2 people. Selected In Top 20 Projects.
- **Build O’ Innovate (Smart Traffic) - Indian Institute of Technology, Patna - 3rd Position** 2021
  - Development of an automatic flow meter to measure traffic parameters using YOLO pre-trained models. Participated with a team of 3. Awarded second place at the national level.

Courses/Certifications

Dataiku - Core Designer. | Natural Language Processing with Python. | Neural Networks and Deep Learning. | Improving Deep Neural Networks: Hyper-parameter Tuning, Regularization and Optimization. | Data Structures & Algorithms - Java.