NISHANT SINHA | 19CE3AI09

CIVIL ENGINEERING/ARTIFICIAL INTELLIGENCE MACHINE LEARNING AND APPLICATIONS

EDUCATION

Year	Degree/Exam	Institute	CGPA/Marks
2024	M.TECH Dual Degree 5Y	IIT Kharagpur	7.64 / 10
2018	CBSE	Abhinav Public School	78.2%
2016	Senior Secondary	DPS Sushant Lok	9.8 / 10

INTERNSHIPS

Artificial Intelligence Engineer | Evva Health **[May. 2023 - Jul. 2023]**
- Created a system that adeptly handles diverse formats: audio (**Whisper API**), images (**Azure Vision OCR**), and PDFs
- Developed a **Langchain** QnA chatbot integrated with OpenAI models for providing contextual assistance to users
- Engineered backend using Computer Vision & NLP, enhancing user documents to actionable insights and analytics
- Utilized **Pinecone** for efficient text embeddings storage and **Postgres** to store document keywords and summaries

Data Science Intern | FutureSmart AI **[Apr. 2023 - Jun. 2023]**
- Engineered a robust **Whisper** ASR endpoint, Incorporated **S3 bucket** integration to ensure efficient audio storage
- Built **Streamlit** app for model explainability, Adaboost outperformed other models with top metrics (**F1-score: 0.98**)
- Trained a deepfake model and engineered an api, **ResNet50** achieved **99.6%** accuracy, surpassing other models
- Built a **OCR** endpoint for handwritten and printed text detection using open-source libraries and vision models

Machine Learning Engineer | Anicca Data Science Solutions **[Apr. 2022 - Jun. 2022]**
- Researched and tested various image classifiers. Utilized **VGG16**, **ResNet50**, **Inceptionv3**, and **EfficientNet** models
- Conducted research and tested various Time Series, ML, and DL algorithms for **15-day** stock price forecasting
- Developed a **Flask** endpoint, contributing to the development of the web application for stock price prediction

Machine Learning Intern | Feynn Labs **[Dec. 2021 - Feb. 2022]**
- Ideated ML/AI based solutions for small and medium businesses enhancing sales, operations, and marketing
- Analyzed the EdTech Market in India using segmentation analysis and came up with a strategy to enter the market
- Utilized diverse **clustering** techniques to identify 4 essential market segments for an EdTech startup's success

PROJECTS

Term Project | Design Lab **[Jan. 2023 - Apr. 2023]**
- Performed topic modeling methods like **BERTopic**, **LDA** on NLP literature dataset, Concluded BERTopic's superiority
- Utilized various similarity techniques like **Cosine**, **KL Divergence**, and **Jaccard index**, uncovering evolving NLP trends
- Explored temporal trends using a 5-year date filter on the extracted topics, revealing recent research patterns

Self Project | Semantic Search **[May. 2023]**
- Spearheaded Semantic Search project, integrating NLP and ML to achieve context-aware information retrieval
- Used **Huggingface** sentence transformers to convert text to vector embedding, Utilized **Pinecone** to store vectors
- Utilized **Cosine** similarity as a metric, ensuring retrieval of the most relevant questions in response to user queries
- Developed a userinterface using **Gradio**, facilitating query input, and retrieving upto 10 most relevant questions

COMPETITION/CONFERENCE

Natural Language Processing with Disaster Tweets | Kaggle **[Ongoing]**
Finetuned a Large Language Model to distinguish disaster-related tweets, empowering rapid emergency response teams
- Secured a rank of **16** out of **950** participants, placing in the top **2%**, demonstrating strong competitive performance
- Employed rigorous text preprocessing, fine-tuned **DistilBERT** on the tweets data, yielding a perfect F1 score of **1.0**

Warehouse Demand Forecasting Inter Hall Data Analytics | Nihilent Ltd **[Mar. 22]**
Predicted fan manufacturing inventory one month in advance, proactively mitigating potential warehouse low fill rate challenges
- Clustering the dataset into 3 clusters of similar time series using **Kmeans** with **dynamic time warping** distance metric
- Utilized various EDA techniques and time series forecasting models like **Auto-ARIMA, Facebook Prophet** and **LSTM**
- Expanded the data by transforming time-series to regression via **data melting**, achieved MAPE of **0.34** with **LightGBM**

SKILLS AND EXPERTISE

Azure | AWS | Machine Learning | Deep Learning | OpenAI API | HuggingFace | NLP | Computer Vision | Time Series Modelling | Statistical Modelling | Data Structures and Algorithms | Web Scraping | ML/DL Algorithms | Data Visualizations | Clustering | Pinecone | Langchain | ASR | Classification | Pytorch | Sklearn | FastAPI | Docker
Programming Languages: Python | C | C++ | PostgreSQL | HTML | MATLAB | SCALA

COURSEWORK INFORMATION

Linear Algebra for AI and ML | Machine Learning Foundations and Applications | Deep Learning Foundations and Applications | Statistical Foundation for Artificial Intelligence and Machine Learning | Probability & Statistics

!Self declared by the student, CDC could not verify the relevant documents