Mahendra Singh

 +91 999999999 [](mailto:nishantsinha00@gmail.com) Maehndra11@gmail.com [](https://www.linkedin.com/in/nishant-sinha-201885191/) [LinkedIn [](https://github.com/nishantsinha00) [Github](file:///C:\Users\xps\Downloads\hhh)](https://www.linkedin.com/in/mahendra-singh-%F0%9F%9A%80-83699485/)

# Education

**MS Institute of Technology, Kharagpur** **2019 - 2024**

*Dual Degree in Civil Engineering/Artificial Intelligence Machine Learning & Applications*

# Internships

**XM Health May 2023 – July 2023**

*Artificial Intelligence Engineer Remote*

* Created a system that adeptly handles diverse formats: audio (Whisper API), images (Azure Vision OCR), and PDFs
* Developed a **Langchain** Q&A 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

**Smart AI Apr 2023 – Jun 2023**

*Data Science Intern Remote*

* 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

**MS Data Science Solutions Apr 2022 – Jun 2022**

*Machine Learning Engineer Remote*

* 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

# Projects

**Term Project Jan 2023 – Apr 2023**

*Design Lab IIT Kharagpur*

* Performed topic modeling methods like **BERTopic**, **LDA** on NLP literature dataset, Concluded BERTopic’s superiority
* Used similarity techniques like **Cosine Similarity**, **KL Divergence** and **Jaccard index**, uncovering NLP trends
* Explored temporal trends using a 5-year date filter on the extracted topics, revealing recent research patterns

**Self Project May 2023**

*Semantic Search*

* 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 user interface using **Gradio**, facilitating query input, and retrieving up to 10 most relevant questions

**Self-Project Mar 2024**

*AI-Enhanced Resume Analyzer*

* Developed an ATS application with Django, employing OpenAI’s API for advanced CV parsing in JSON format
* Used **OpenAI** sentence embeddings to query relevant CV data given a job description based on cosine similarity
* Incorporated custom sentence embedding transformers (**all-MiniLM-L6-v2**) to provide relevancy score for CV sections

# Technical Skills

**Languages**: Python, C++, SQL

**Backend Frameworks**: Django, Flask, FastAPI, Gradio **Clouds & Databases**: Azure, PostgreSQL, Pinecone **Web Technologies**: Docker

**Developer Tools**: Postman, VS Code, GitHub

**AI Skills**: NLP, LLM Finetuning, Computer Vision, Statistical ML, Time-Series Forecasting, Deep Learning

**ML/DL Frameworks**: Pytorch, PyTorch Lightning, NLTK, Spacy, HuggingFace, Scikit-learn

**Python Frameworks**: Numpy, Pandas, Matplotlib, Seaborn, Plotly, PySpark, BeautifulSoup

# Competitions/Achievements

**Warehouse Demand Forecasting — Nihilent Ltd Mar 2022**

*Inter Hall Data Analytics IIT Kharagpur*

* Clustering the dataset into 3 clusters of similar time series using **Kmeans** with dynamic time warping distance metric
* Utilized EDA techniques and time series forecasting models like **Auto-ARIMA**, **Facebook Prophet**, and **LSTM**
* Transformed the data from time-series to regression via data melting, achieved MAPE of **0.34** with **LightGBM**