

Hiten Mandaliya

Kandivali W, Mumbai | hitenmandaliya562@gmail.com | [LinkedIn](#) | [Github](#) | +91 8450904594

ACADEMIC DETAILS

NMIMS University

Master of Technology, Data Science (Business Analytics) | CGPA: 3.81/4

August 2022 – April 2024

Mumbai University

Bachelor of Engineering, Electronics and Telecommunication | CGPA: 6.87/10

July 2015 – May 2019

SKILLS

Languages: Python, R, SQL

Frameworks/Libraries: langchain, nltk, pandas, numpy, scikit-learn, keras, tensorflow, seaborn, matplotlib, plotly, streamlit, presidio, nemogaurdrails, streamlit, chroma, FAISS, pytorch

Skills: Data Analysis, Data Visualization, Data Cleaning, Machine Learning, Deep Learning, Natural Language Processing (NLP), Time Series, EDA, LLM, Statistics.

Tools: Git Lab, SAS, VS Code, Pycharm, Jupyter Notebook, Google Collab, Microsoft Excel, Microsoft PowerPoint, Tableau

PROFESSIONAL EXPERIENCE

Tata Communications Limited | Data Science Project Trainee (Internship)

January 2024 – July 2024

- Developed a sophisticated chat-bot using a **Large Language Model (LLM)** with **Retrieval Augmented Generation (RAG)**, enabling users to upload documents and interactively **extract pertinent information**.
- Designed a robust data processing pipeline with **data cleaning**, **embedding generation**, and efficient retrieval through a **vector store** database.
- Implemented intent classification using **SVM classification model** to improve understanding of relevant questions, achieving over 88% accuracy in user queries based on uploaded documents which were evaluated using the Ragas framework, BLEU score, ROUGE score, and Cosine Similarity.
- Presented demos to stakeholders, answering technical questions and illustrating business scenarios in meaningful ways.

rFund.ai | Quant Research Intern

July 2023 – October 2023

- Developed a system for replicating trades from a master account to multiple slave accounts with proportional allocation.
- Created a **dashboard**-based brokerage calculator for stock and derivatives using **Streamlit**, which obtains trade execution data and computes brokerage based on that data

Science Kidz | Senior Technical Associate and STEM Educator

June 2019 - April 2022

- Conducted customer and student data analysis using **Python** and **Excel** to enhance targeted marketing, optimize course offerings, which led to around 30% increase in enrollment and profitability over 6 months.
- Performed **sentiment analysis** on customer feedback to improve the quality of courses and student satisfaction, resulting in enhanced course content and better customer engagement.
- Built educational robotics projects in field of healthcare, agriculture, road safety & transportation. Mentored over 500 students, ranging from 3rd grade to engineering graduates in robotics and **python** programming.
- Led a team of 5 to design and create 4 online courses for Hindustan Times Schools in 3 months.

PROJECTS

Built and Published Library [NLPWash](#) on PyPi:

- A python library that streamlines text pre-processing, saving significant time on cleaning tasks such as normalization, hyperlink and HTML tag removal, punctuation elimination, tokenization, stop word removal, stemming, and lemmatization. All these processes can be executed effortlessly with a single command.

Turnover Guard - Prediction for Employee Retention:

- Implemented employee turnover prediction using machine learning algorithms, including Logistic Regression and Random Forest, and Preprocessed relevant data, conducted extensive EDA to uncover new relationships, patterns, and trends.

RespiroSense - Advancing Early Health Assessment:

- Applied deep learning techniques to the Respiratory Sound Database for early health assessment, incorporating data cleaning, audio file processing, and handling class imbalance. Extracted features from audio signals and implemented a dense Convolutional Neural Network (CNN) with the integration of early stopping, ensuring robustness and scalability.

OPEN SOURCE CONTRIBUTIONS

- Contributed to Microsoft's Presidio Library by developing and integrating a feature for Personally Identifiable Information (PII) detection. This features helps to detect Indian Passport Numbers which is a sensitive information in many use cases. Here is the [Link](#).