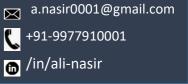
ALI NASIR



[Jul'22- Present]

EDUCATION

IIT Bombay- M. Tech (2019-2021)

Geoinformatics and Natural Resource Engineering – Image Processing 9.0 C.G.P.A

SPA Bhopal- B. Plan (2015-2019)

Urban and Regional planning-7.45 C.G.P.A

AREA OF INTEREST

Machine learning | Computer Vision | Data Science | Statistics | Deep Learning

SKILLS

Toolkit

GPT-4, Heroku, Docker, Apache-airflow, postman, GIT, MS Office, RAG, AWS

Implemented Models

Predictive & Statistical modelling, Transfer learning, Few-shot learning, NLP, Hugging Face, LLM

Frameworks

FastAPI, Scikit-learn, keras, TensorFlow, Pytorch, Pandas, OpenCV, NumPy, Panda, NLTK, SciPy, spaCy, Word2Vec, PySpark OpenAi, LangChain

Languages

Python, SQL

CONCEPTS

Generative Ai, LLM

Machine Learning: SVM, LR, RF, Naive Bayes, XGBoost, Logistic Regression, Decision trees, clustering algorithms

Deep Learning: ANN CNN, RNN, LSTM

INTERESTS

Charcoal Sketching, Painting, Music, Snooker, Hiking

PROFESSIONAL EXPERIENCE

Physics Wallah | | DataScientist-2

Project: Smart Doubt Recommendations

Impact: To reduce work of SMEs by providing a funnel to answer doubts asked by the students in any video lectures.

- o Developed clusters of doubts based on contextual meaning for 200 chapters using Ada-003 embeddings
- Integrated a funnel of doubt classification models built using SVM and Neural Network with hyper parameter optimization to achieve an overall F-1 score of 90% and delivered API endpoint using FastAPI.
- o NCERT data is chunked and inserted into Vector Store (AstraDB) using Langchain and Cassandra.
- o Integrated AstraDB by DataStax for our RAG application and built the appropriate prompt which considers subject, exam and chapter metadata to extract the top 5 relevant document from the **RAG** using "mmr".
- o Extracted documents are fed to LLM (Azure OpenAI GPT 3.5) as context and the result is obtained.

Project: Personalized Test-series using Adaptive Learning: User-Engagement through 'Infinite Practice'

Impact: To give personalized test questions to every user in ratios of Easy, medium and hard category depending upon his/her calculated knowledge using statistical modelling and engagement.

- O Conceptualized state-of-art and innovative solutions, **item response theory (IRT)** to estimate the ability level of a person based on their performance on a test using difficulty, discrimination, speed and guessing parameters.
- O Formulated an algorithm from scratch for the **3Parameter Logistic model** (a statistical model) to fit parameters with optimized using **Bayesian estimation** to calculate **knowledge parameters** (θ)
- O Created 2 APIs using FastAPI to return questions & optimized query to reduce latency from 40 secs to 530 ms.

Project- Topper Identification Using Academic Score (PATENT PROJECT)

Impact: Personalized Engagement Scores nurturing Potential Toppers using their test and level of engagement.

o Performed Queries to generate **41 KPIs** from 60 raw parameters and trained Regression model for around 37K users of JEE With an **RMSE**: **0.11** and 20k users of NEET exam with an **RMSE 0.09**.

Vedantu Innovation Pvt. Ltd. | | DataScientist-1

Project: VGyan (Doubt Solving through Ai)

Impact: Built an automatic doubt solving chatbot for Ai Live sessions to help in solving academic/Non-academic doubts of the students on the platform itself without help of assistant teachers.

oAutomated the process of classifying academic doubts & providing relevant answers for each subject and grade. oCreated sentence embeddings using distilroberta and BERT- transformer, built a classifier with **92% accuracy** that can determine whether a doubt is academic or not and test its performance.

oBuilt a regressor with **r2 score 0.978** to predict the number of doubts asked in a respective session based on the attendance of students, so that the teachers can be allocated to the session for efficient resource utilization.

Project: Post Session Comments

Impact: Developed an **ETL pipeline** to get **dispositions** & **sentiments** of **students' comments** to automate work of Students' Account Manager in designating the students' issues received after a session on platform

oPerformed multi-class classification on the text-comments with an accuracy of 81%.

oCreated and deployed a cronjob-script from scratch in production using Apache-airflow

Ai India | | Computer Vision (R & D) (Summer Internship)

oProvided Solution for Route Optimization problem with land suitability analysis of Pune District oTested different ML classifiers for supervised LULC classification with **XGBoost giving 93%** accuracy

oDesigned a CNN architecture to outperform the existing ImageNet DL classification models

MAJOR PROJECT

Urban Growth Predictive Modelling Using Machine Learning (Master's Thesis)

[Jul'20- June'21]

[Jul'20-Sep'20]

[Jul'21- June'22]

- o Performed time series analysis of 20 years on satellite images to predict the urban land expansion.
- o Implemented Dynamic simulation models: *cellular automata (CA)*-MARKOV chain & Advanced ANN to achieve an overall model accuracy of 92.3 % for prediction modelling.

NATIONAL ACHIEVEMENTS

o Presented paper titled 'Optimal Location of Smart Water-Meters & HOTSpot Analysis of Excessive Water Consumption' at National seminar on recent advances in geospatial technology at IIRS (ISRO)

oAcquired 1st place at National NOSPlan convention 2017 serving as **TEAM LEADER** of 100+ delegates.

POSITIONS OF RESPONSIBILITY

o Teaching Assistant at IIT [Aug'19-Jun'20] PR and Logistics Coordinator- PG Cult IITB [Sep'19-Apr'20]

o President- Rotaract (SPA Bhopal) [Jul'17-Feb'18] Lead-Coordinator- Annual fest SPA Bhopal [Sep'17-Mar'18]