Niraj Kumar

neeraj2681.github.io | neeraj2681@gmail.com | +919568753620 | linkedIn/neerajbhatt2 | github/neeraj2681

EDUCATION

M.Tech in Computer Science — Kolkata, India | 2020-22
Indian Statistical Institute

B.Tech in Computer Science and Engineering — Pauri, Uttarakhand | 2015-19
Govind Ballabh Pant Engineering College

WORK EXPERIENCE

Data Scientist — Gurgaon, Haryana | Since May 2024
@Accertify(Split from American Express in 2024)

- Worked on a new kind of product called **Auto-Review** model which led to annual savings of $250K per client and significantly reduced the operational costs
- Introduced a new set of features **(linkage and past-lookup)** further enhancing the model performance
- Skills: **Python, Pyspark, SQL, Feature Engineering, Case-Studies, Machine Learning Algorithms, Strategy Designing**

Data Scientist — Gurgaon, Haryana | Aug 2022 - Apr 2024
@American Express(Accertify)

- As part of **Accertify**, designed **Fraud Detection Models** for a variety of Industries like Money Transfer, Retail, Travel etc.
- Designed Strategy Rules over and above the base models to enhance the model performance and capture seasonal fraud trends.
- Was Awarded **Analyst of the Quarter** for Q1 2023 for delivering a fraud detection model for Money-Transfer Industry
- Skills: **Python, Pyspark, SQL, Feature Engineering, Case-Studies, Machine Learning Algorithms, Strategy Designing**

PROJECTS

DETECTING ANOMALIES IN SURVEILLANCE VIDEOS — Tensorflow, Deep Neural Networks, 3D CNNs

- Utilising Encoder-Decoder based **reconstruction and prediction techniques**
- Utilisation of different **Deep Neural Networks**
- Real-time applicable solution @30FPS

OPEN-SOURCE BASED CUSTOM LLM — Langchain, LLMs, Pinecone

- **Used different open source LLMs to train on custom dataset using Retrieval Augmented Generation(RAG)**
- Model answers the user queries based on a given proprietary document provided by the user
- **Model Evaluation using Giskard**
- **End to End model development and deployment on the cloud**

SKILLS

Languages: C, Python, Java, Bash, SQL

Domains: Algorithm Designing, Classical Machine Learning, Deep Learning, Image Processing and Computer Vision, NLP, LLMs

Tools and Frameworks: Pandas, MATLAB, Git, Pytorch, Tensorflow, LaTeX, Pyspark, Langchain, Streamlit, Fast API, MongoDB, AWS

PERSONAL INTERESTS

Testing different investing strategies, Mathematics, Learning about different cultures around the globe(By maintaining my Duolingo streak), Hiking.