NITISH KUMAR

nitish.kumar.mec14@iitbhu.ac.in ● +91-9452688058 ● github.com/nitishsingh41 ● Bangalore, India

SUMMARY

Experienced data scientist with 6 years specializing in LLMs, NLP, and deployment, focusing on large-scale solutions. Proficient in Python, TensorFlow, AWS, REST API, and Docker, with hands-on experience developing and deploying AI solutions. Successfully executed projects in chatbots, NLP, image processing, and quantitative finance.

EDUCATION

- **Indian Institute of Technology (BHU), Varanasi** *July2014–June2018*
  B.Tech in Mechanical Engineering

SKILLS AND INTERESTS

Languages: Python, R, C++
Technologies: OOPS, Tensorflow, Git, MATLAB, Pandas, Numpy, Pytorch, OpenCV, Linux, GCP, ONNX, Streamlit, Langchain, AWS EC2, FastAPI, Retrieval augmented generation(RAG), Docker
Courses: Topics in Mathematics with Applications in Finance(MIT OCW)(**Derivative Pricing, VaR, Monte Carlo SImulation, Stress Testing, Volatility Modeling**), Stochastic Calculus, Introduction to Financial Engineering and risk management(Coursera), **Neural Networks** and Deep Learning(Coursera), Data Structure and Algorithms, Practical Deep Learning for Coders(Fast.ai), Data Analysis using **Pyspark**(Coursera), **Generative AI** with Large Language Models(Coursera), Linear Algebra, Probability and Statistics

WORK EXPERIENCE

Model Risk Quant at Acuity Knowledge Partners (formerly known as Moody's Analytics) from May 2023 to Current
Data Scientist at AbInbev from Jan 2022 to May 2023
Data Scientist at Karza Technologies from May 2020 to Jan 2022
Intern- Data Scientist at Karza Technologies from Jan 2020 to April 2020
Associate Consultant at Larsen & Toubro Infotech Ltd. (L&T Infotech) from Jul 2018 to Jan 2020

PROJECTS

Model Validation **Acuity Knowledge Partners** ***May2023-Current***

- Validated credit risk and machine learning models around credit default, **stress testing** and **identity resolution** for one of the top 3 banks in Hawaii. Prepared model review documentation.
- Worked on validating multiple risk profile asset portfolios, based on **Modern Portfolio Theory** , for one of the top 10 US banks using univariate **Monte Carlo Simulation** based on forward looking **CMAs** along with **Latin Hypercube Sampling** and **Variance reduction** technique. Prepared model review documentation.

NLP Projects **Self** ***May2022-Current***

- Created a **chatgpt** chatbot based on **GPT4** to work as AI tutor. Used **pinecone** vector database to store the indexed documents. Used langchain to efficiently implement it as a **streamlit** app.
- Made a streamlit app on top of OpenAI **GPT3** to extract transcript from any Youtube video and summarize in 10 points using **prompt engineering**.
- Made a streamlit app based on a transformer based sentiment analysis model and deployed on **AWS EC2** using **Docker**. Sentiment Analysis Streamlit app with Transformers using Docker on AWS EC2 | Complete Tutorial
- Developed a **Llama2 RAG** based chatbot which uses **retrieval augmented generation** (**RAG**) using any webpage with support for chat history in **Langchain**. Further served it as RESTapi using **FastAPI**.
  Langchain Llama2 RAG chatbot with history | FastAPI

TAX Analytics **AB InBev** ***Jan2022-May2023***

- Transform the existing manual process by codifying business rules and metrics to automate tax analytics. Used Azure devops to prepare the repository and make deployable and scalable solutions. It saved **$4M** in Taxes.
- Further Made **PowerBI** dashboard based on weekly data refresh and encompassing 5 geographies for it.

FX hedging using deep learning **AB InBev** ***Jun2022-Sept2022***

- Provided a solution for making the decision of which derivative to use for hedging **USD/BRL currency pair** against the raw material price fluctuations over year. Gained 0.3% in value compared to the prior solution.
- Used tabular learning to choose between **Options and Futures** over a weekly period based on market and macroeconomic variables.

Face Matching Algorithm Implementation **Karza Technologies** ***Jan2020-Jan 2022***

- Implemented **few-shot** learning for face recognition using **ResNet100** and cosine-based loss, reducing training time by half on GCP.
- Developed a **reviewNeeded** tag for low-quality images, extending the evaluation dataset for better match quality.
- Implemented a **low-latency, API-servable** solution deployed on **AWS Lambda**. It produced results in 50ms.

Face Detection Implementation **Karza Technologies** **May*2020-April2021***

- Implemented **retinaface** face detection based on wider face dataset and prepared whole codebase for AWS lambda based l**ow latency api servable** solution.
- Further made the model robust towards **rotated faces,** which isn't tackled on the open source solutions.
- Retrained the model for small faces with poor quality in KYC domain images and removal of holographic faces.

APTOS 2019 Blindness Detection by Kaggle

- Created a model based on **EfficientNet-B4** to determine diabetic retinopathy severity, ranking in the top 9% (248/2943) on the private leaderboard and got a **bronze** medal for it.
- Preprocessed and cropped eye images using **OpenCV** to reduce black areas, leveraging pretrained models on ImageNet and Diabetic Retinopathy datasets for fine-tuning.
- Treated the problem as regression, using MSE and Kappa score metrics, and implemented test-time augmentation to enhance generalization.

Jigsaw Unintended Bias in Toxicity classification by Kaggle

- Made an ensemble of **BERT** models for toxicity classification, ranking in the top 10% (309/3167) on the public leaderboard.
- Applied multi-task learning and transfer learning with PyTorch to improve generalization and performance.
- Developed a custom binary cross-entropy loss function, a custom data loader to reduce training time, and various learning rate reduction strategies.