

Hariharavarshan Nandakumar

236 995 3198 | nhariharavarshan@gmail.com | [LinkedIn](#) | [Huggingface](#) | [Kaggle](#) | Canada

EDUCATION

University of British Columbia, Vancouver, Canada

September 2022 – June 2023

Master of Data Science - Computational Linguistics

Projects: [Fine Tuning Google LLM\(FlanT5\)](#), [LightGBM classification](#), [House price prediction hosted in FlaskAPI](#)

Anna University, India

June 2012 - May 2016

Bachelors of Engineering – Electrical Engineering

TECHNICAL SKILLS

Languages & Tools: Python, SQL, R, PySpark, Pandas, Numpy, SciPy, Statsmodels, Hugging face, Scikit-learn, Keras, TensorFlow, Pytorch, Altair, MongoDB, Postgres, Docker, Github, Power BI, Linux GCP, BigQuery, VertexAI, Jira.

Techniques: Machine learning, Deep learning, A/B testing, NLP, T5, Flan T5, BERT, NLU, stochastic gradient descent, LSTM, Large Language Models (LLM), CNN, NER, Sentiment analysis, Transformers (Attention), clustering, Trees and Graphs.

WORK EXPERIENCE – 7 years

Prompt Engineer - Meta, Vancouver

June 2023 – Aug 2023

- Worked on **Llama2 personas** that acts as an AI assistant models to Qualitatively measure the security and robustness of the models. Prompting for future Reinforcement ideas that needs to be incorporated.

NLP Research Intern - Seasalt.ai, Vancouver

Apr 2023 – June 2023

- Finetuned LLM** models such as **LLAMA, Alpaca, ChatGPT API** to generate Chinese data set for training Nemo Punctuation toolkit to add Punctuation from ASR raw data with the help of **Langchain** package.
- Used GCP(Bigquery,VertexAI) to design, train and evaluate models on GPUs. Hosted the final fine-tuned models on GCP for productionizing.

Data Scientist Consultant – HTC Global

Mar 2022 – Nov 2022

- To read the Ingredients and Nutrition values from food items package, developed a **Transformer architecture** with finetuning **BERT** and **OCR** techniques. **Deployed models using IAAS such as AWS Lambda, EC2 instance.**
- Developed sentiment classification for customer surveys using **seq2seq with attention**, LSTM models.
- Deploying the model in **ML Ops** services to provide end-to-end services for frictionless training, validation, deployment and monitoring of ML models, including a **CI/CD component for ML.**

Associate Data Scientist – Royal Dutch Shell

Jan 2021 – Jan 2022

- Enabling the UK retail team to take action on promotions and statistically (A/B testing) understanding if the promotion was successful understating grievances mentioned by the customers in the shell GO+ app.
- Crunched text data from app to classify sentiments and recognize entity using **Goole's Flan T5** based Algorithm and **NLP** augmentation. **Deployed using Flask API.** Thus increased the customer loyalty score by 8%
- Utilized ARIMA, SARIMA, exponential smoothing, and **LSTM-based** models to generate accurate short-term and long-term sales forecasts, resulting in a 10% reduction in forecast error. Deployed using Flask API.

Senior Data Analyst – Access Health Care Services

Nov 2018 – Dec 2020

- Implemented Dynamic Work Allocation frame work by classifying the claims and employees pertaining to fraud by allocating claims according to their expertise and constraints using **LightGBM and XG-Boost. (Supervised and Unsupervised)**
- With Oversampling using SMOTE to balance the classes increased the performance productivity metric of the workforce by 13.8 % for an employee count of 11,000. The model is containerized using **Docker.**

Data Analyst – Infosys Ltd

June 2016 – Oct 2018

- In operations TAT, developed Neural Network models like LSTM to help the ticket assignment/Ticket Classification faster avoiding human bias. Created an all-in-one Power BI Dashboard for KPI management system.

EXTERNAL PINPOINT

- Business Analytics & Intelligence Program** from Indian Institute of Management-Bangalore
[Supply chain Analytics](#) scored 83% from Massachusetts Institute of Technology (MITx)

2019-2020

Aug 2021