

Mandeep

Vancouver, BC, Canada | mandeepkhokhar1110@gmail.com | +1 (236) 862-1674 | [LinkedIn](#) | [Google Scholar](#) | [GitHub](#)

As a driven machine learning engineer with a passion for advancing AI in diverse fields, I excel in blending technical skills with creative solutions to drive technological breakthroughs.

EDUCATION

University of British Columbia

Master of Data Science in Computational Linguistics

Vancouver, BC, Canada

September 2023 - June 2024

Thapar Institute of Engineering and Technology

Bachelor of Technology in Computer Engineering

Patiala, Punjab, India

August 2019 - July 2023

SKILLS

Programming Languages: Python, R, MATLAB

Machine Learning: Pandas, NumPy, SciPy, PyTorch, PyTorch Lightning, TensorFlow, Keras, Hugging Face, Numba, NLTK, spaCy

Deep Learning: Transformers, Large Language Models (LLM), LSTMs, RNNs

Visualization: Altair, Plotly, Plotly-Dash, Datapane, Matplotlib, Seaborn

Databases: Relational database design, PostgreSQL, MongoDB

Others: GitHub, A/B testing, Docker, Weights & Biases (Wandb)

WORK EXPERIENCE

Betterdata

Singapore (Remote)

Machine Learning Intern

February 2022 - July 2023

- Spearheaded research in tabular synthetic data generation, implementing cutting-edge generative model architectures that ensure data diversity, fidelity, and privacy.
- Developed and integrated differential privacy into generative models using PyTorch, ensuring compliance with data protection standards.
- Adapted generative model code for versatile deployment, both cloud-based and on-premises, optimizing training and evaluation pipelines using the Numba library for parallel processing.
- Engineered an evaluation system measuring relevance, accuracy, and privacy of generated data, utilizing TensorFlow for model prototyping, Scikit-learn for statistical analysis, and PyTorch for implementing deep learning-based metrics.

SpaceML

USA (Remote)

AI Researcher

February 2021 - February 2022

- Collaborated with NASA's IMPACT team to develop an AI system utilizing Self-Supervised and Active Learning techniques, enabling swift identification of pertinent satellite imagery from petabyte-scale datasets.
- Spearheaded the creation of a scalable Active Learning pipeline, integrating multiple cloud scalability layers for cost efficiency using AWS and Google Cloud Platform (GCP).
- Achieved a drastic reduction in manual image labelling time, decreasing it from 7000 hours to 52 minutes for a five million-image climate dataset, utilizing libraries such as TensorFlow and PyTorch.
- Open-sourced a high-quality package on [GitHub](#) that is already used for labelling in [NASA's Phenomenon portal](#).

- Secured a NASA Science Mission Directorate grant with the proposal ranking in the top 5 out of 79 research initiatives.

Thapar University, Adviser: Dr. Husanbir Pannu

India

Undergraduate Student Researcher

January 2020 - July 2022

- Researched convolutional neural networks and the effect of hyperparameter tuning and architectures on explainable AI techniques like XAI: LIME and SHAP, using libraries such as TensorFlow and Keras.
- Explored the application of generative adversarial networks in enhancing deep learning detection systems by generating synthetic COVID-19 CT and MRI scans utilizing PyTorch and OpenCV.
- Studied the efficacy of machine learning methods in forecasting the social media traction of various topics, employing libraries such as Numpy, Pandas, and SQL for data analysis and manipulation.

Bournemouth University

UK (Remote)

Research Assistant

June 2021 - August 2021

- Developed a real-time camera-based grape leaf disease diagnosis iOS app using Swift, enabling deep learning-based disease detection and remedy suggestions on iPhone 8 (2017) and above.
- Researched Convolutional Neural Network architectures to maximize accuracy while minimizing floating-point operations and memory requirements, ensuring efficient performance on resource-constrained mobile devices, utilizing TensorFlow Lite and Core ML.
- Optimized the base model for real-time performance using techniques such as model quantization and pruning, enabling the deep learning model to run natively on the device with reduced latency and power consumption.

PUBLICATIONS

- Deep learning-based explainable target classification for synthetic aperture radar images *March 2020*
- SpaceML Worldview Search - Learnings from an AI citizen scientist team building a NoCode Data Curator from Unlabeled Petabyte Scale Imagery *December 2021*
- Smartphone Based Grape Leaf Disease Diagnosis and Remedial System Assisted with Explanations *April 2022*
- Machine Learning Based Explainable Financial Forecasting *June 2022*

PROJECTS

Project Title: Biomedical Lay Summarization using Large Language Models

March 2024 - June 2024

Project Description: This project focuses on generating accessible summaries of biomedical research articles for non-expert audiences. The goal is to bridge the gap between complex biomedical knowledge and public understanding, facilitating wider dissemination of scientific findings.

Methodology:

- Led the development of innovative NLP models using domain-specific LLMs like BioMistral7b, focusing on zero-shot and few-shot learning techniques for summary generation.
- Explored advanced techniques such as retrieval-augmented generation (RAG) and representation engineering to enhance model performance and alignment with human values.
- Implemented control vectors to dynamically adjust model outputs, optimizing for readability and factual accuracy while maintaining technical relevance.

Outcomes: Developed an innovative approach by integrating control vectors with the BioMistral7b LLM, achieving a breakthrough in generating accessible biomedical lay summaries with unparalleled accuracy and readability.