

ANSHUL SHIVHARE

M.Tech - Artificial Intelligence

@ shivhareanshul78@gmail.com
G Google Scholar

@ anshulshivha@iisc.ac.in

+91 6266331149

GitHub

Medium

in LinkedIn

EXPERIENCE

Mango Sciences

Data Scientist

Nov 2023 – Present

Bangalore

Project: Discharge Summary Automation from Healthcare Data

- Used Large Language Models (LLMs) to analyze and extract fields from cancer patient notes and chemo orders.
- Fine-tuned LLAMA2 13B on patient notes for retrieval and deployed the weights endpoint.
- Enhanced evaluation metrics, notably BLEU and BERTScore (0.3 to 0.8 F1 Score), post-fine-tuning compared to the base model.

Ernst & Young (EY)

Artificial Intelligence Engineer

Aug 2022 – Nov 2023

Bangalore

Project : Synthetic Data Generation as a Service

- Used open-source library SDV (Synthetic Data Vault) for synthetic data generation.
- Incorporated privacy checks and data protection measures.
- Evaluated synthetic data using metrics for quality assurance.

Project : Spend Classification Engine

- Multiclass text classification on spend data. Used various classical algorithms: Logistic Regression, Random Forest.
- Deployed the model on Azure ML with Databricks.

WinZO

Data Science Intern

Jun 2021 – Aug 2021

Bangalore

Project : Graph Analysis for Recommendation System

- Graph analysis on 4M users (out of 40M) for recommendation system, Used few accelerated data science techniques like RAPIDS (for faster processing).
- Implemented Community detection, Centrality algorithm (to find influential users), Page Rank algorithm on graph. Formulated a measure to detect fraud using graphs.
- Used Graph Convolutional Networks (GCN) to predict whether the new user will be a depositor or not. Got around 64 % accuracy.

RESEARCH EXPERIENCE

Project : Classification and Visualization of Medical Images

Mar 2021 – Aug 2022

Bangalore

- The retinal optical coherence tomography (OCT) data are used for the diagnosis of multiple retinal diseases.
- Confocal Endomicroscopy (CONVIVO) images consist of brain tissue in real-time. Hence, we have to detect various types of brain tumor.
- Removed speckle noise from both types of images using convolutional neural networks. Generated synthetic data for the same.
- Multi-label Medical image Classification, Visualization with the help of Grad-CAM and Score-CAM technique.

ACADEMIC PROJECTS

Image Denoising using U-Net

GitHub

- Residual learning technique is used to denoise the image.
- Experimented on 2 types of noises, Gamma and Poisson.
- Achieved 6 DB increase in PSNR metric and SSIM upto 0.93 on predicted image compared to noisy image.

Link/Edge Prediction

GitHub

- Predicting K important non-existing edges, based on ranking the set of non-existent edges in the graph.
- Implemented three different ways to predict edges (Jaccard's coefficient, Katz's score, hitting time).

Graph Representation Learning with Generative Adversarial Networks

- Embed each vertex in a graph into a low-dimensional vector space.
- Softmax is replaced by graph softmax.
- The accuracy of Discriminator and Generator were 76 % at the end of training.

Movie Recommendation System

- Content-based recommendation, Finds other movies with similar content. Then ranks movies according to their similarity scores.
- Collaborative filtering using Matrix Factorization.
- The data set used is Movie Lens with 25M movie ratings.

PERSONAL PROJECTS

- Mosaic implementation: Merging many overlapping images to form single image (image stitching)
- Segmentation
- Modeling the spread of COVID-19 in Karnataka
- Duckworth-Lewis method for predicting target runs in limited overs cricket

SKILLS

Python

C

MATLAB

C++

PyTorch

Keras

Scikit-learn

Pandas

PUBLICATIONS

- MDPI Journals-2023 -"Deep-Learning-Based Visualization and Volumetric Analysis of Fluid Regions in Optical Coherence Tomography Scans". ↗
- Med-NeurIPS-2022 -"Denoising Enhances Visualization of Optical Coherence Tomography Images". ↗

COURSES

- Computational Methods of Optimization, Computational Linear Algebra, Stochastic Models and Application.
- Pattern Recognition and Neural Networks, Data Structures and Algorithms, Machine Learning and Signal Processing, Computer Vision.
- Advanced Deep Learning, Data Analytics, Digital Video: Perception and Algorithms.

EDUCATION

M.Tech Artificial Intelligence

Indian Institute of Science

 2020 - 2022  Bangalore

CGPA: 7.9/10

B.E Computer Science Engineering


Acropolis Institute of Technology and Research

 2016 - 2020  Indore

CGPA: 7.63/10

XII & X

Prestige Public School

 2014 - 2016  Indore

Grades: 86% & 8.8 CGPA