ADITYA MAVLE

Pune, 411009, India| adityamavle@gmail.com| +91 7620469531 |
github.com/adityamavle|linkedin.com/adityamavle|sites.google.com/adityamavle|

EDUCATION

Pune Institute of Computer Technology (PICT), affiliated to Savitribai Phule Pune Univ. Expected: July 2024
Candidate for Bachelor of Engineering in Electronics and Telecomms., Honors in Data Science. **(GPA 9.06/10)**
Relevant Courses: *Engineering Mathematics, Data Structures, Programming and Problem-Solving, Data Science and Visualization, Statistics and Machine Learning, Object-Oriented Programming, Database Systems, Computer Networks, Digital Image Processing, Deep Learning*, Cloud Computing**

Massive Open Online Courses (MOOCs): *Machine Learning (Stanford Online), Algorithms (Stanford Online), Computer Architecture (Princeton University)*

PROFESSIONAL EXPERIENCE

Johns Hopkins University, Baltimore, USA *(Computational Statistics)* **Sept 2022 – Aug 2023**
Research Intern – Johns Hopkins University Baltimore-India Clinical Trials Unit, Pune, India

- Worked on devising simulation-based methods for evaluating the design requirements of Randomized Controlled Trials in the planning stage.
- Executed clinical trial simulations in Python to assess the effectiveness of risk factor interventions for TB treatment outcomes improvement. Deployed system as a FastAPI web app for easy experimentation with simulation functions.
- Performed Hypothesis Testing and Statistical power analysis on clinical datasets.
- Developed a recursive resampling algorithm that determines the ideal sample size required for optimal statistical power.

IIIT Hyderabad, Hyderabad, India *(Computer Vision)* **May 2023 – July 2023**
Summer Research Intern – Center for Visual Information Technology Lab (CVIT)

- Worked on OCR for Indian Languages, fine-tuned a ResNet-based OCR model on HPC GPU clusters, outperforming pre-trained variants, utilizing PyTorch, DocTR, and OpenCV.
- Evaluated and improved existing word detector models on Benchmark Indian Languages Datasets.
- Built an end-to-end pipeline for document annotation, pre-processing, and image augmentation.
- Balanced dataset classes using a recall-based document difficulty classifier and implemented weighted sampling from multiple training datasets.

IIT Bombay, Mumbai, India *(Web Development, 3D Graphics)* **Feb 2023 – May 2023**
Software Development Intern – FOSSEE, Osdag Team
Osdag is an Open-Source Software for the design and analysis of steel structures, developed by FOSSEE.

- Developed new features for Osdag and led the task of building a web-based application for Osdag, using Python and JavaScript.
- Developed the system's backend with Django, implementing Django REST APIs to enable session creation, input design values, CAD model rendering, and design report generation within the web application.
- Created a web-based CAD model and 3D graphics renderer using Three.js.
- Implemented cross-file format compatibility through a FreeCAD macro-based CAD file converter.

UST Global, Pune, India *(Computer Vision, Multimodal ML)* **Nov 2022 – Jan 2023**
Machine Learning Intern

- Delivered multiple Computer Vision and Multimodal ML-based projects using Intel's OpenVINO framework and implemented in Python.
- Trained multiple YOLOv5, MobileNet, and VisualBERT-based models for the applications of Face Recognition, Real-Time Scene Text Detection, VisualQA, and Multimodal Motion-pose detection.
- Developed an end-to-end attendance management system and successfully deployed it in a university's lab.
- Deployed models into the organization's workflows. Authored documentation for the projects executed.

Pune Institute of Computer Technology, Pune, India *(Natural Language Processing)* **Jun 2022 – Aug 2022**
Undergraduate Research Assistant

- Worked on the research and development of ML-integrated file systems.
- Established a Client-Server paradigm-based Backup System using Python Socket Programming for remote file backup in chunked format.

- Generated descriptive tags on the files using NLP and CV tasks of keyword extraction, topic modeling, name entity recognition, abstractive summarization, and image, and video classification.
- Developed a module for PII data detection using keyword extraction and cosine similarity based on BERT, for sensitive data protection.

PrimeNumerics Consulting Inc, Pune, India *(Natural Language Processing)* **Jan 2022 – March 2022**

Machine Learning Intern

- Performed semantic analysis of electronic health records (EHR) data to maintain patient profiles and assess health conditions using Natural Language Processing.
- Developed a model inference pipeline for constructing a tabular dataset of patient records from unstructured input text.
- Fine-tuned a DistilBERT-based model for multiclass text classification, enabling accurate analysis of patient data for case severity assessment based on EHR data.

PROJECTS

Domain-Specific Large Language Model for Financial Risk Analysis* *(Natural Language Processing)*

- Building a Large Language model tuned for financial risk analysis and question answering based on contracts, specifically addressing credit and liquidity risk.
- Generated instruction dataset for model training by processing a financial data corpus using a variant of the T5 model.
- Instruction tuning of LlaMa2, Falcon7b, and FlanT5 models on a financial knowledge base of annotated data and instruction dataset, using the Low-Rank Adaptation method of fine-tuning.
- Tech Stack: Python, PyTorch, LangChain, Huggingface.

Cohort Intelligence-based Hyperparameter Tuning of Neural Networks* *(Bio-Inspired optimization)*

- Implementing Cohort Intelligence algorithm for optimal hyperparameter tuning of Neural Network parameters of learning rate, epochs, and number of layers.
- Performed comparative analysis and implementation of grid search, Bayesian Optimization, and Ant-Colony optimization-based methods for Hyperparameter tuning.
- Tech Stack: Python, SciPy, MATLAB.

Electrical Equipment Failure Prediction Using Neural Networks *(Predictive Maintenance)*

- Performed as a project member of the Google Advance Data Center Infrastructure team.
- Analyzed the performance of the Data Center's equipment employing statistical and fault tree-based methods.
- Trained an Autoencoder architecture-based anomaly detection model for prediction of electrical equipment failure.
- Integrated ML-based predictive models with the team's standard physics-based models for efficient maintenance scheduling.
- Tech Stack: Python, TensorFlow, Pandas, Google Cloud.

PUBLICATIONS

- *Accepted for oral presentation*: **'Review of Language Models in the Financial Domain.'** Aarushi Wagh, Aditya Mavle. IEEE Intelligent Systems and Machine Learning Conference (ISML 2024)
- *Under Review*: **'RCT Simulations for Enhanced TB Clinical Trial Design'** Aditya Mavle, Aarushi Wagh, Atharva Nagmoti, Nikhil Gupte. 12th IEEE International Conference on Healthcare Informatics. (ICHI 2024) *(preprint)*

AFFILIATION/MEMBERSHIPS

- **Institute of Electrical and Electronics Engineers (IEEE):** Member of PICT IEEE Student Branch (PISB). Representative of the DataWiz (Data Science competitions) and Marketing teams. Part of the organizing team for Credenz 22'
- **Computer Society of India (CSI):** Member of PICT CSI Student Branch (PISB). Part of the Marketing Team.
- **TEDxPICT:** Member of PICT's TEDx Chapter. Representative of the Curations and Content teams. Part of the organizing team for our annual event of TEDxPICT's Marcato 22'.

HONORS/AWARDS

- Ranked 13th out of 330 students in the Department of Electronics and Telecommunications at PICT for the 2022/2023 academic year. Stood 1st for the subjects of Computer Networks, Database Systems, and Project Management.
- Secured 2nd Place in DataCup, a Data Science Competition, at PICT CSI's flagship event, Xenia 2022.