Keerthika Pujari

BS in Data Science and Applications
Indian Institute of Technology, Madras, India

keerthikapujari25@gmail.com
Website
linkedin.com/in/keerthika-pujari
github.com/keerthika-pujari
leetcode.com/keethu22

Education

Degree	Institute/Board	CGPA/Percentage
B.Tech	SRM Institute of Science and Technology, Ramapuram	9.98(5th semester)
BS	Indian Institute of Technology, Madras	Ongoing

Experience

Research Intern @ IITM | ML in IBN Ongoing
Guide: Dr. Krishna Moorthy Sivalingam *Chennai*

- Researching Machine Learning and Intent-Based Networking (IBN).
- Completed data scraping and preprocessing of Kubernetes documentation.
- Fine-tuned lightweight language models (LLMs) using Hugging Face transformers library.
- Analyzed and summarized numerous survey papers on network performance and intent-based networking and assisted in the publication of a paper.

Research Intern @ IITM | ML in Computational Chemistry Ongoing
Guide: Dr.Chaitanya Sharma Yamijala *Chennai*

- Reviewed survey papers on the use of machine learning models in environmental chemistry.
- Currently reviewing related articles to identify valuable research concepts.
- Investigated the classification of PFAS sources from PFAS fingerprints in fish tissue.
- Conducted data analysis and preprocessing to replicate an existing repository with modifications.
- Applied machine learning techniques to identify the source of PFAS contamination.

Bachelor Of Sciences | Data Science and Applications Ongoing
University: Indian Institute of Technology Madras *Chennai*

- Completed the Foundational level courses. Relevant Courses include Statistics, Linear Algebra, Computational thinking, Python Programming
- Currently pursuing the Diploma in Data Sciences. Relevant courses include Machine Learning Foundations, Tools in Data Science, Machine Learning practices, Buisness analytics, Buisness Data management, Database management systems, Data Structures and Algorithms, Modern Application development.

Hackathon | Top 50 Finalist, Smart India Hackathon Ongoing
Representing University: SRM Institute Of Science and Technology *Chennai*

- As one of the two primary developers on a team of six from SRM Institute of Science and Technology, contributed to the development of a SAR image colorization model for the pan-India SIH competition, with over 1000 participants. Advanced through three rounds, including idea pitching and prototype demonstration, to become one of the top 50 teams.
- Collaborated on a semi-supervised deep learning approach for defense and environmental monitoring applications, currently awaiting results for potential final presentation in New Delhi.

Technical Proficiency

- **Programming Languages**
  Python, C, C++, Java, HTML, SQL, Latex
- **Technologies**: React.js, Django, TensorFlow, PyTorch, Bootstrap, Flask,Node.js, Android SDK, Excel, Tableu
- **Concepts**: Calculus, Probability and Statistics, Operating System, Virtual Memory, Cache Memory, Encryption, Decryption, Artificial Intelligence, Machine Learning, Neural Networks, API, Database Normalization, Intent Based Networking, Natural Language processing,Data Structures and Algorithms,Big Data Tools and techniques, Object oriented programming, Formal language and automation

Certifications

Certificate	Organisation	Year
Foundational level - BS in Data Science	IIT Madras	May 2024
Python for Data Science	NPTEL	September 2023
Database Management Systems	NPTEL	March 2024
Introduction to Machine Learning	NPTEL	October 2024
Demystifying Networks	NPTEL	October 2024
Google Data Analytics Professional Certification	Google	January 2024
E-cell(Entrepreneurship)	IIT Madras	March 2024
Summer Analytics 2024	IIT Guwahati	May 2024
AI-ML Virtual Internship	AICTE - Eduskills	December 2024

Projects

Automated Text Generation of Kubernetes Data Using LLM Models | *Transformers, Kubernetes, GPT Models*

Guide: Dr. Krishna Moorthy Sivalingam, IIT Madras

- Scraped extensive data from Kubernetes documentation, creating a high-quality dataset for training purposes.
- Performed domain-specific preprocessing on the Kubernetes data to ensure relevancy and accuracy for model training.
- Fine-tuned lightweight GPT models using the preprocessed Kubernetes dataset, achieving efficient and relevant text generation.
- Evaluated and optimized the models for performance, ensuring they meet the desired standards for accuracy and efficiency.

ML in PFAS source Identification, Classification | *Google colab, Python*

Guide: Dr.Chaitanya Sharma Yamijala, IIT Madras

- Reviewed survey papers on the application of Machine Learning in PFAS source classification and identification.
- Replicated and analyzed ideas from existing research papers to understand PFAS fingerprinting in fish.
- Conducted deep analysis to draw insights from the replicated study..
- Currently developing and refining a problem statement for PFAS source classification.

Election Data Analysis for Kerala's Thiruvalla Constituency | *Python, Pandas, Excel*

Guide: Dr. Anand.S, IIT Madras

- Conducted an in-depth analysis of voting patterns and trends for Kerala's Thiruvalla constituency using historical election data.
- Extracted and cleaned data to ensure accuracy, removing noise and inappropriate records.
- Calculated key electoral metrics such as vote share, margin percentage, and identified instances of deposit loss and absolute majority.

- Utilized Python and Pandas for data manipulation and Microsoft Excel for final data formatting and presentation.

Uber Data Analysis | *GCP, BigQuery, Google Data Studio, Python, SQL*

Guide: Dr. Karthik Elangovan, SRM Ramapuram

- Developed a comprehensive data visualization dashboard using Google Data Studio, providing stakeholders with real-time insights into Uber ride patterns, demand forecasting, and user behavior.
- Leveraged Google BigQuery for efficient querying and analysis of large datasets, and Google Cloud Storage for data management, streamlining data operations.
- Implemented a data ingestion and processing pipeline using Python, enhancing the accuracy and speed of data analysis by 30%.
- Optimized decision-making processes by presenting key operational metrics and trends through interactive and dynamic visualizations, contributing to data-driven strategy adjustments.

SAR Image Colorization | *Google Colab, Tensorflow, OpenCV, Pytorch*

Guide: Dr.Aarthi B, SRM Ramapuram

- Developed a SAR (Synthetic Aperture Radar) image colorization model using a deep learning approach to enable critical insights for defense and environmental monitoring.
- The project involved the MC-GAN model, which used discriminators and generators in a cyclic process to colorize grayscale SAR images by enforcing consistency through cyclic loss.
- This model supports real-time applications by enabling colorized imagery without the need for paired datasets, which is vital for defense applications like border surveillance and environmental uses such as oil spill detection in low-light conditions.

Predictive Modeling in Lung Cancer Analysis | *Google Colab, TensorFlow, PyTorch, scikit-learn, Transformers*

Guide: Mrs.M.Ramaprabha, SRM Ramapuram

- Developed machine learning models to address lung cancer detection, survival prediction, and treatment response prediction, aimed at enhancing data-driven decision-making in oncology.
- Implemented a CNN-based detection model trained on the IQ-OTH/NCCD dataset to identify lung cancer in CT scans.
- Utilized Cox Proportional Hazards and SVM models for survival prediction, evaluated with precision, recall, F1 score, and ROC AUC metrics to assess patient survival based on clinical data.
- Employed multimodal transformers to predict treatment response, using clinical and demographic features for binary classification and assessment of treatment outcomes.

Areas Of Interest

- Research interests in Data Sciences and analytics, Artificial Intelligence, Machine Learning, Computer Networks, Intent based networking, Natural Language processing
- Personal interests include teaching methodologies and Fashion Designing.

Achievements

- Awarded Topper badges in Python, English 1, and English 2 at the foundational level and in Tools in Data Science, Business Data Management at the diploma level in the BS program at IIT Madras.
- Secured Rank 1 in the CSE-BDA department in B.Tech at SRM Institute of Science and Technology for three consecutive semesters.
- Achieved proficiency in Abacus, completing all necessary levels.
- Awarded 1st and 2nd ranks in English and Mathematics by the National Science Olympiad (NSO).

- Earned Hindi Proficiency certification from the Dakshin Bharat Hindi Prachar Sabha (DBHPS/India), having completed all required courses.
- Attained beginner-level proficiency in Sanskrit, awarded by Samskrita Bharati (India).
- Possess 7+ years of experience in Classical Dance (Bharatnatyam) and Carnatic Music, demonstrating significant artistic skills.

Languages

English (Full professional proficiency), **Hindi** (Full professional proficiency), **Tamil** (Bilingual proficiency), **Telugu** (Native proficiency), **Sanskrit** (Elementary proficiency)

Relevant Courses

Computational Thinking, Probability and Statistics, Programming in Python, Machine learning foundations and Techniques, Tools in Data Science, Data structures, Design and analysis of algorithms(C), Operating Systems(C++), Linear Algebra with Computational Applications, Database management systems, Object Oriented Programming (Java), Business Data Management, Computer Networks, Artificial Intelligence Machine learning for data analytics

Social Engagements

- **Contigent Leader**: Of SRM University, representing 20 plus students at E-Summit IITM, a three-day entrepreneurship event organized by E-Cell IITM.
- **Volunteer**: at Leonard Cheshire Disability, a home for differently abled Children and Elders, for a month.
- **Sports-Engagements**: Kho-Kho(District level), Badminton(Beginner level), Chess(Intermediate level)

University Clubs

- **Research and Content Member, Veritas Oratory Society**: Contributed research and content for a technical club at IIT Madras, focusing on oratory and technical communication skills.
- **Content Creator, Google Developer Groups (GDG) IIT Madras**: Member of the content creation team, helping organize and promote technical events, workshops, and resources as part of IIT Madras's GDG chapter.

Conferences and Publications

- **IEEE International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics**
  Paper Title: Towards a Comprehensive Deep Learning-Based System for Lung Cancer Nodule Detection, Prognosis, and Treatment Response Prediction (Invited for poster presentation)
  Authors: Pujari Keerthika, Padmalakshmi.S, Souvik Ruhidas, Ramaprabha Marimuthu
  Abstract: This study presents a robust deep learning framework for the early detection of lung cancer nodules, integrating CNNs and U-Net for nodule segmentation and employing Cox proportional hazards models and SVMs for survival prediction. Multi-modal transformers further enhance treatment response prediction by integrating imaging, clinical, and genomic data. The efficacy is evaluated using multiple metrics to improve lung cancer management.
  Domain: SST24: AI and Machine Learning Application in Healthcare for Personalized Disease Prediction and Management.