



Degree	University/Institute	Year	CGPA/Marks(%)
M.Tech Computer Science & Engineering	Indian Institute of Technology, Hyderabad	2025	8.41
B.Tech Computer Science & Engineering	St. Thomas' College of Engineering and Technology	2021	9.21
XII (CISCE)	M.P Birla Foundation Higher Secondary School	2017	95.00%
X (CISCE)	M.P Birla Foundation Higher Secondary School	2015	94.20%

WORK EXPERIENCE / INTERNSHIPS

Research Assistant at Indian Institute of Technology, Hyderabad

Jul'22 - Present

- Specialization in the intersection of machine learning and systems, actively addressing key challenges in the field as part of my ongoing master's degree.
- Current projects include enhancing algorithms for the binary similarity problem, utilizing embeddings for automated program classification, and applying reinforcement learning techniques to optimize function inlining.

Systems Engineer at Tata Consultancy Services

Aug'21 - Jul'22

- Worked as a software developer on PearlChain project under Canon Europe client for around a year after immediate completion of my bachelor's degree.
- Day to day responsibilities included analysing various defects, implementing code changes and bug fixes.
- Led another parallel project responsible for upgrading the versions of JDK, Spring, Hibernate, Jackson and other libraries for the entire codebase of 10M+ LOC.

Software Developer Intern at Thinking Stack

Jun'20 - Jul'20

- Contributed to building a food delivery website as part of a team of 5 members.
- Added the cart feature for the website, which displays the items in cart as well as the net payable amount.
- Customized the search bar to enable searching by food items along with restaurant names.

PROJECTS

Binary Similarity using ML framework

Mar'23 - Jul'24

- Introduces a binary similarity method using VEX- IR, leveraging architecture neutral Intermediate Representation to assess functional similarities between binary programs.
- Extracts embeddings from basic block sequences (peepholes) through random walks on the control- flow graph, normalized using compiler-inspired transformations to reduce architectural variations.
- Utilizes VexNet, a feed- forward Siamese neural network, to derive function embeddings for similarity assessment, effectively positioning similar functions closely in an n-dimensional space while addressing Out-Of-Vocabulary (OOV) challenges.

Algorithmic Classification using Embedding based approach with Branch Probability Information

Apr'24 - Jul'24

- Generate robust embeddings from source codes with emphasis on semantic similarity to classify and group programs.
- The model was built on top of [IR2Vec](#) framework for the embedding generation, with added Branch Probability Info pass from LLVM.
- [POJ-104](#) dataset was used for this experiment to calculate the accuracy of program classification using this approach.

Analysis and Transformation Passes using [LLVM](#)

Jan'24 - May'24

- Developed an inter- procedural, context- insensitive, flow- sensitive constant propagation analysis pass in LLVM-14.x, and subsequent transformation pass to substitute constant values, eliminating redundant computations.
- Implemented an intra-procedural, flow sensitive May Alias analysis pass, identifying the May Alias relationship between each pointer variable pair at the end of the program.
- Published the implementation on [GitHub](#) with comprehensive code comments and provided examples in repository for easy understanding and replication of the optimization technique.

CineScope-GPT - AI-Powered Movie Discovery Redefined

Jan'24 - Mar'24

- [CineScope-GPT](#) is an advanced web app created using React, Redux and Firebase that provides personalized movie recommendations, surpassing traditional OTT platforms in search customization and complexity.
- Leveraged **OpenAI** API keys to utilize the *GPT-3.5 Turbo* model for intuitive natural language queries, enhancing user interaction.
- Designed a simple and user-friendly interface, seamlessly integrating with the **TMDB** movie database for real-time movie data retrieval.

Brand Perception using Twitter Sentiment Analysis

Jan'21 - Jun'21

- Extracting tweets from social media website Twitter to analyze the sentiments behind them and have a general perception among the users about a particular brand.
- Enabling customers to choose wisely among multiple brands for a certain category of products.

- Hosted a Python-based brand perception and comparison system on [GitHub](#), complete with detailed instructions on how to run the code both locally and on Google Colab, with clear examples and a user-friendly interface provided through Anvil.

PUBLICATIONS

- VexIR2Vec** - An Architecture-Neutral Embedding Framework for Binary Similarity Mar'23 - Jul'24

Authors : S. VenkataKeerthy, **Soumya Banerjee** , Sayan Dey, Yashas Andaluri, Raghul PS, Subrahmanyam Kalyanasundaram, Fernando Magno Quintão Pereira, Ramakrishna Upadrasta

Abstract Summary: VexIR2Vec proposes a novel approach for binary similarity. It uses VEX- IR to represent binaries, normalizes them to mitigate architectural and compiler variations, and learns a vocabulary of IR entities. Finally, it uses a Siamese neural network to compare function embeddings for similarity.

Publication Link: <https://arxiv.org/pdf/2312.00507>, submitted in [ACM Transactions on Software Engineering and Methodology](#), 2024 (A* journal).

- ML-based Inlining using Reinforcement Learning** Aug'24 - Present

Research Summary : To optimize function inlining decisions, learning from historical data and adjusting in real-time, reinforcement learning models can intelligently balance the trade-offs between execution speed and resource overhead. The goal is to surpass traditional heuristics by making more adaptive, data- driven inlining decisions that lead to more efficient and performant compiled code.

Publication : To be submitted in future conference.

RELEVANT COURSES

- Masters:** Foundations of Machine Learning, Deep Learning, Advanced DSA, Advanced Compiler Engineering, Program Analysis and Compiler Optimizations , Linear Optimization, Advanced Compiler Optimizations
- Bachelors :** Data Structures & Algorithms, Operating Systems, Database Management Systems, Computer Networks, Compiler Design, Object Oriented Programming

TECHNICAL SKILLS & INTERESTS

- Programming:** C/C++, Java, Python, SQL, HTML, CSS, JavaScript, Shell Script
- Software Development Tools/ Libraries :** Git, LLVM, Docker, Clang, React, Redux, Firebase, Vim, MySQL, Linux, Visual Studio
- Data Science & Machine Learning Tools/ Libraries :** PyTorch, Pandas, NumPy, SciPy, Scikit-learn, Matplotlib, NLTK, Google Colab
- Interests:** Algorithmic problem solving, Software Development, Systems, Data Science, Deep Learning, NLP

CERTIFICATIONS & WORKSHOPS

- Attended **Intel Unnati** workshop at IITH in August 2024, focused on their GPU accelerators and AI/ deep learning in architecture, enhancing industry-ready skills in the domain of AI and GPUs.
- Problem Solving certification issued by **HackerRank** for clearing their Intermediate and Basic levels of programming contests with focus on data structures and algorithms.
- Attended Machine Learning with Python Bootcamp organized by MyWbut and Nitish Singh (**CampusX**) and completed a project on [Movie Recommendation System](#) participating in a team of 3 members.
- NPTEL** Online Certification on Programming in Java issued by IIT Kharagpur for attending the online lecture series and clearing the final examination with **silver** medal.
- Completed an online course on Responsive Web Design focused on software development and [hands-on sessions](#) organized by **freeCodeCamp.org**.

POSITIONS OF RESPONSIBILITY & ACHIEVEMENTS

- Teaching Assistant for the following course at IITH: **CS1023** (Software Development Fundamentals), enrolled by 70+ students, responsible for creating 30+ quizzes, correcting answers and designing the final project for the course. Jan'24 - Jun'24
- Teaching Assistant for the following courses at IITH: **CS3423** (Advanced Compilers), responsible for designing the assignments posted at regular intervals during the course tenure and correcting the coding solutions. Fall'22 and '24
- Core member of the [Institution of Engineers](#) (IEI), India during my bachelors, responsible for organizing events to promote scientific research, engineering workshops and conferences. Jul'17 - Aug'21
- Secured a percentile of **97.66** in the Graduate Aptitude Test in Engineering (**GATE**) held in February, 2021 during my final year of bachelor's, ranking among the top 2.34% of 101,922 candidates.
- Secured a global rank of **955** among 1,30,000+ candidates in **TCS Codevita** Season 9 held in Aug, 2020, which is an algorithmic contest and hackathon, secured a full time offer for the role of Systems Engineer from this.
- Co- founded and spearheaded the **STCET Coding Club** during my bachelor's, fostering a competitive programming culture by organizing practice contests and sessions to support internships and placements for juniors and peers. Jul'19 - Aug'21