

Vamsi Deeduvanu

(765)-694-9091 | vamsi10010@gmail.com | [linkedin.com/in/vamsideeduvanu/](https://www.linkedin.com/in/vamsideeduvanu/) | github.com/vamsi10010

EDUCATION

Purdue University

Aug. 2022 – May 2025 (exp.)

Bachelor of Science in Computer Science and Data Science

West Lafayette, IN

- *Coursework:* Object Oriented Programming, Multivariate Calculus, Linear Algebra, Discrete Math, Probability, Statistics, Data Structures and Algorithms, Computer Architecture
- *Clubs and Extracurriculars:* ML@Purdue, Purdue Astronomy Club, Intramural Sports
- *Honors:* Dean's List and Semester Honors (Fall 2022, Spring 2023)
- *GPA:* 4.0/4.0

Resonance Junior College

June 2020 – May 2022

Intermediate

Hyderabad, India

- *JEE Mains:* 99.49th percentile all India

EXPERIENCE

Undergraduate Data Science Researcher

Aug. 2022 – May 2023

The Data Mine, Purdue University

West Lafayette, IN

- Partnered with Battelle on researching hyper-parameter optimization procedures for NLP models resulting in significant improvement in performance.
- Fine-tuned a BERT model from the HuggingFace library to accurately identify adverse drug events in electronic health records.
- Applied hyperband and population-based training algorithms from RayTune to tune hyperparameters and improved overall f1 score by more than 20%.
- Delivered updates in weekly sprint meetings with client and documented a standard operating procedure as a reference for future projects.
- Presented research poster at The Data Mine Symposium and demonstrated entity recognition on a live document.

PROJECTS

cgrad | C, cmocka, Deep Learning

Aug. 2023 – Present

- Designed a lightweight backpropagation and artificial neural network library in C to optimize memory consumption during training processes.
- Implemented a directed acyclic graph to represent computational graph of a neural network and perform backpropagation to calculate gradients.
- Utilized dynamic memory allocation to constantly allocate and free heap space for the graph during runtime for efficient memory usage.
- Applied cgrad to train an ANN to classify handwritten digits from MNIST dataset with 96% accuracy, using less than 2GB of memory during training.
- Created unit tests for library using cmocka framework to automate testing process and ensure functionality.

YourCollege | Python, Scikit-Learn, Pandas, Streamlit

Jan. 2023 – Present

- Developed a college recommender application to assist high school students in college search.
- Collected data from multiple sources and performed data cleaning and feature engineering for classification.
- Trained an unsupervised learning algorithm to create a unique classification of colleges and recommendations for every user tailored to personal preferences.
- Built and deployed application on web through Streamlit to make it accessible to users.

Time Series Forecasting | Python, Statsmodels, Pandas, Keras, Matplotlib

Sep. 2022 – May 2023

- Collaborated with members of ML@Purdue to forecast PM-10 pollution levels in California.
- Implemented an ARIMA model from statsmodels library and an LSTM model from Keras to predict PM-10 levels achieving high accuracy rates and providing valuable insights.
- Researched causes and trends of pollution levels in California and presented results at club standup meetings.

SKILLS

Languages: Python, C, Java, SQL (Postgres), R, LaTeX

Developer Tools: Git, Bash, VS Code, Jupyter, PyCharm, IntelliJ, Agile Methodologies

Libraries: PyTorch, HuggingFace, Keras, RayTune, Scikit-Learn, Streamlit, Pandas, NumPy, Matplotlib