SRIVATHSAN BADRINARAYANAN

srivathsanb14@gmail.com • srivathsanb14.github.io/ • linkedin.com/in/srivathsan-badrinarayanan/

EDUCATION

Carnegie Mellon University *(GPA - 4.0)* Pittsburgh, PA
Master of Science in Artificial Intelligence Engineering - Chemical Engineering December 2024
Coursework: Generative AI, Large Language Models, Deep Learning, Scientific Software Engineering, Trustworthy AI

Indian Institute of Technology Madras Chennai, India
Bachelor of Technology in Chemical Engineering (Minor in Bioprocess Engineering) July 2023
Coursework: Data-driven Modeling and Optimization of Bioprocesses, Mathematical Foundations of Data Science

EXPERIENCE

Graduate Research Assistant January 2024 - Present
Mechanical and AI Lab (MAIL), Carnegie Mellon University Pittsburgh, PA

Project: Multimodal Language and Graph Learning of Adsorption Configuration in Catalysis

- Developed a novel generative-predictive pipeline integrating large language models and graph-assisted multimodal pretraining, significantly enhancing adsorption energy prediction accuracy for catalyst screening.
- Achieved a 10% reduction in MAE by introducing a transferable model that aligns with domain knowledge, predicting adsorption configurations without reliance on atomic spatial coordinates, leading to a publication.

Project: Multi-Peptide - Multimodality Leveraged Language-Graph Learning of Peptide Properties

- Developed a multimodal model, combining transformer-based language models and GNNs to enhance protein property prediction for applications in drug development and bioinformatics.
- Achieved state-of-the-art results by implementing a contrastive learning framework (CLIP) between protein sequence data (transformer) and AlphaFold generated structures (GNN), culminating in a publication.

Project: Generative AI Podcasts (hosted on Spotify @dreampods) and automated lectures

- Led the development of an innovative prompt-to-audioform generative AI pipeline by integrating Large Language Models with Text-to-Speech models, enhancing public knowledge accessibility through automated podcast creation.
- Leading a cross-functional team in the development and pilot of a pdf-to-lecture video automated ML pipeline, assisting universities and professors in creating educational content with 50% reduced preparation time.

MITACS Globalink Research Intern May 2022 - August 2022
University of Alberta Edmonton, Canada

Project: Data-Driven Optimization of Refinery Process Operations

- Developed deep multivariate models with 97% accuracy to predict process outputs in a hydro-processing plant, using time-series models on real-time refinery data to optimize parameters and improve operational efficiency.

Machine Learning Intern June 2021 - July 2021
Tirios AI Austin, TX (Remote)

- Engineered a robust machine learning pipeline for object recognition, deployed using AWS Sagemaker, for accurately identifying home appliances needing repair from photos, reducing service wait time by 50%.

SKILLS

Programming Languages: Advanced - Python, C++, MATLAB; Basic - R, Julia, SQL
Machine Learning & Data Science: PyTorch, TensorFlow, Keras, sklearn, SparkML, PySpark, HuggingFace, Transformers
Data & Image Processing: Pandas, Numpy, Seaborn, Matplotlib, Scipy, BeautifulSoup, OpenCV, Pillow
Cheminformatics & Bioinformatics: RDKit, ASE, Pymol, Biopython, AlphaFold
Cloud & DevOps: AWS, GCP, Docker, CUDA, PostgreSQL, DynamoDB, Apache Kafka
Simulation: COMSOL, Aspen Plus, Simulink, Fusion360, SuperPro Designer **Version Control:** Git

PUBLICATIONS

Multi-Peptide: Multimodality Leveraged Language-Graph Learning of Peptide Properties, S. Badrinarayanan et al.
(Under review in ACS JCIM, arXiv preprint: https://doi.org/10.48550/arXiv.2407.03380)
Multimodal Language and Graph Learning of Adsorption Configuration in Catalysis, J. Ock et al.
(Accepted in Nature Machine Intelligence, arXiv preprint: https://doi.org/10.48550/arXiv.2401.07408)