Srivathsan (Sri) Badrinarayanan

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

Education

Carnegie Mellon University — Pittsburgh, PA
M.S. in Artificial Intelligence Engineering - Chemical Engineering, GPA: 4.0 — Dec 2024
Relevant Coursework: Generative AI, Large Language Models, Deep Learning, Trustworthy AI

Indian Institute of Technology Madras — Chennai, India
B.Tech in Chemical Engineering (Minor in Bioprocess Engineering) — Jul 2023

Publications

- **Multi-Peptide: Multimodality Leveraged Language-Graph Learning**, S. Badrinarayanan et al. *ACS JCIM*, arXiv:2407.03380
- **Multimodal Language and Graph Learning in Catalysis**, J. Ock, S. Badrinarayanan et al. *Nature Machine Intelligence*, Nature MI

Projects

Multimodal Language and Graph Learning of Adsorption Configuration in Catalysis
- Developed a novel generative-predictive pipeline to predict adsorption energies without atomic spatial coordinates.
- Achieved $\sim 10\%$ reduction in MAE by introducing a transferable model integrating large language models and graph-assisted multimodal pretraining, significantly enhancing adsorption energy prediction accuracy for catalyst screening.

Multi-Peptide: Multimodality Leveraged Language-Graph Learning of Peptide Properties
- Developed a multimodal-model for peptide property prediction with applications in drug development and bioinformatics.
- Achieved state-of-the-art results by implementing a contrastive learning framework (CLIP) between protein sequence data (transformer language models) and AlphaFold generated structures (GNN), culminating in a publication.

MOFGPT (Ongoing)
- Designing a GPT-2-based generative model with policy gradients to create novel, viable Metal-Organic Frameworks (MOFs), revolutionizing materials discovery for gas storage, separation, and catalysis.

AI Agent for Drug Discovery (Ongoing)
- Developing an autonomous pipeline with generative models and verifiers to accelerate de novo drug design and molecular optimization. Aiming to reduce the timeline for discovery by incorporating agents with decision-making capabilities.

Character consistent diffusion for visual storylines
- Developed and evaluated novel character-consistent diffusion models, integrating Prompt-to-Prompt techniques, caption-based dynamic conditioning, and latent space alignment to generate coherent and visually consistent sequential storylines.

THALA - Thermodynamics Helper and Analysis Library for Applications
- Developed and released a Python library for thermodynamic analysis, showcasing expertise in end-to-end project execution, version control, and package deployment.

Experience

Graduate Research Assistant, Mechanical and AI Lab (MAIL), Carnegie Mellon University — May 2024 - Present
- Leading pioneering AI-driven research, producing high-impact publications in cutting-edge AI4Science.
- 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 reduced preparation time.
- Mentored 10 undergraduates, advancing their expertise in AI-protein research and guiding impactful projects.

MITACS Research Intern, University of Alberta — May 2022 - Aug 2022
- 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.

ML Intern, Tirios AI — Jun 2021 - Jul 2021
- Engineered a ML pipeline for object recognition, deployed using AWS Sagemaker, for accurately identifying home appliances needing repair from photos, reducing service wait time by half.

Skills

Programming: Python, C++, MATLAB, R **Chem/Bio-informatics:** RDKit, ASE, AlphaFold, Biopython, Pymol
ML Tools: PyTorch, TensorFlow, Keras, sklearn, SparkML, PySpark, HuggingFace, Transformers, Diffusers, Pandas, Numpy, Seaborn, Matplotlib, Scipy, BeautifulSoup, OpenCV, Pillow, Git
Cloud & DevOps: AWS, GCP, Docker, CUDA, PostgreSQL, DynamoDB, Apache Kafka