ABHILASH NEOG

🌐 Website 💼 LinkedIn 🎓 Scholar 🐙 Github ✉ abhilash22@vt.edu

Research Interests

- Foundation Models • Physics-Informed ML • LLMs & Multimodal Models • Time-Series Modeling

Education

Virginia Tech **Aug 2022 - Present**
Ph.D., Computer Science. Advisor: *Dr. Anuj Karpatne*. GPA: 4.0/4.0 Blacksburg, USA

Virginia Tech **Aug 2022 - Dec 2024**
M.S., Computer Science. Advisor: *Dr. Anuj Karpatne*. GPA: 4.0/4.0 Blacksburg, USA

Birla Institute of Technology and Science (BITS), Pilani **July 2016 − July 2020**
Bachelor of Engineering (B.E.), Computer Science, GPA: 8.08/10 Pilani, India

Research Experience

KGML Lab, Virginia Tech | Graduate Research Assistant **Jan 2023 − Present**

- Built a model-agnostic approach for **robust time-series** modeling in the presence of partially observed data
- Building a framework for **knowledge distillation** of **LLMs** into light-weight interpretable models for time-series modeling
- Developed a **Foundation Model** for Aquatic Sciences to learn effective representations of variables within a dynamical system, inter-variable interactions and correlations between different aquatic systems
- Empirical evaluation of **fine-grained semantic segmentation** performance of vision models for a scientific dataset
- Benchmarked zero-shot effectiveness & reasoning ability of SOTA Vision-Language Models (**VLMs**) like **GPT-4, LlaVa** through Visual Question Answering tasks on a diverse species dataset.

Industry Experience

ThinkSense Inc. | Machine Learning Intern **May 2023 − Aug 2023**

- Developed an outlier detection model for denoising sensor-based Human Activity Recognition (HAR) **time series** data
- Built & deployed a CNN-based HAR model achieving 82% F-1 score on an android app using Keras & TensorFlow Lite

Oracle | Data Scientist **Sep 2020 − July 2022**

- Built & deployed Machine Learning applications into ETL pipelines, leveraging **Spark** systems, **MLOps** & **CI/CD** pipelines
- Designed and deployed a *Demand Prediction* application for time series forecasting using the DeepAR model
- Developed an unsupervised classification algorithm (utilizing HuggingFace, FastText models, NLP techniques like NER, POS tagging) achieving 40% higher accuracy than then SOTA **LLMs** on a 71k-label dataset.

VMware | Software Development Engineer Intern **Jan 2020 − June 2020**

- Streamlined the process of fetching & filtering raw data from Workspace ONE Cloud using Spring Boot REST APIs
- Contributed to an end-user federation app on Workspace ONE Cloud, and wrote unit tests using JUnit and Mockito

Samsung Research Institute | Summer Intern **May 2019 − July 2019**

- Performed a feasibility study of Multi-frame Noise Reduction solutions' deployment in Live Focus for Low light conditions
- Optimized the existing HAL call flow, in C++, with considerable noise reduction in the first phase of live focus capture

Publications

1. KS Mehrab, M. Maruf, Arka Daw, **Abhilash Neog**, HB Manogaran, et al. "Fish-Vista: A Multi-Purpose Dataset for Understanding Identification of Traits from Images". *CVPR 2025*

2. **Abhilash Neog**, Arka Daw, Sepideh Fatemi, Anuj Karpatne. "Masking the Gaps: An Imputation-Free Approach to Time Series Modeling with Missing Data". *NeurIPS 2024 Workshop*

3. M. Maruf, Arka Daw, KS Mehrab, HB Manogaran, **Abhilash Neog**, M. Sawhney, et al. "VLM4Bio: A Benchmark Dataset to Evaluate Pretrained Vision-Language Models for Trait Discovery from Biological Images". *NeurIPS 2024*

4. Baviskar, A., Ramanathan, K., **Abhilash, N.**, Pawar, D. and Bangalore, K., Oracle International Corp, 2024. "Machine Learning Based Spend Classification." *U.S. Patent Application 17/903,161.*

5. R. Ladwig, A. Daw, E.A. Albright, C. Buelo, A. Karpatne, M.F. Meyer, **A. Neog**, P. C. Hanson, and H. A. Dugan. "Modular Compositional Learning Improves 1D Hydrodynamic Lake Model Performance by Merging Process-Based Modeling With Deep Learning." *Journal of Advances in Modeling Earth Systems (JAMES) 16, no. 1 (2024)*

6. Lavika Goel, **Abhilash Neog**, Ashish Aman, and Arshveer Kaur. "Hybrid Nature-Inspired Optimization Techniques in Face Recognition." *Transactions on Computational Science XXXVI, Springer LNCS, 2020.*

Selected Projects

Can Large Vision Language Models Ground Fine-grained Attribute? ⬀PDF   **Aug '24 – Dec '24**
- Developed a novel dual-scale attention framework for fine-grained attribute localization in Large Vision-Language Models (**LLaVa**), incorporating entropy-based head selection, maximally connected component filtering, and hierarchical constraints

Evaluating Model Reasoning & Hallucinations in Medical LLMs ⬀Code ⬀PDF   **Jan '24 – April '24**
- Analyzed and evaluated factual error propagation in open-source medical LLMs such as BioMistral, Asclepius, Alpacare, and PMC-LLaMA to identify variations in their efficacy and ensure reliable information dissemination in medical settings.

Convergence analysis of PINN for solving inverse PDEs ⬀Code ⬀PDF   **Aug '23 – Dec '23**
- Performed adaptive weighing of physics-based and data-driven loss terms in Physics-informed Neural Networks
- Achieved 50% average error reduction in PDE (Partial Differential Eq.) parameter estimation of Burgers & Allen-Cahn eq.

Mathematical Reasoning in Large Language Models (LLMs) ⬀Code ⬀PDF   **Aug '23 – Dec '23**
- Worked on the problem of numerical headline generation and numeral masked-fill as part of NumEval @ SemEval 2024
- Adapted **Llama**, **T5**, **BART** & **RoBERTa** models by Direct **fine-tuning** & **prompt engineering** for the respective tasks

Text Summarization of Electronic Theses and Dissertations (ETD) ⬀PDF   **Sept '22 – Dec '22**
- Developed a text summarization pipeline, integrating both Transformer-based abstractive algorithms (pre-trained Pegasus & RoBERTa) and traditional extractive algorithms like TextRank, LexRank & LSA, within an ETD Info. Retrieval system

Technical Skills

Languages: Python, Java, C++, SQL, R
Frameworks: PyTorch, Tensorflow Keras, Git, Spark

Miscellaneous

- Reviewer - ICML 2025, ICLR 2025
- Received NSF NAIRR (National AI Research Resource) Pilot Award, 2024
- Graduate Teaching Assistant, CS 5805 Machine Learning, Spring 2024
- Gave a talk on Transfer Learning in Lake Ecosystems at "NSF Macrosystems Biology Meeting", 2024.
- Gave a Lightning Talk at the "Frontiers in Ecological Forecasting" event at Virginia Tech, 2023.
- Awarded "Star of the Month (Dec 2021)" within the Oracle Analytics Cloud Organization, Oracle India