# Abhilash Neog

♥ Website in LinkedIn ► Scholar ♥ Github ■ abhilash22@vt.edu

#### Research Interests

• Time-Series Modeling • AI for Science • Foundation Models • LLM Reasoning • ML Trusthworthiness

#### Education

Virginia Tech Aug 2024 - Present Blacksburg, USA

Ph.D., Computer Science. Advisor: Anuj Karpatne. GPA: 4.0/4.0

Virginia Tech Aug 2022 - May 2024

M.S., Computer Science. Advisor: Anuj Karpatne. GPA: 4.0/4.0

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

#### Virginia Tech | Graduate Research Assistant

Jan 2023 - Present

Blacksburg, USA

- Proposed a robust Time-series modeling approach for handling partially observed data
- · Working on effective knowledge-guided LLM alignment with continuous unbounded time-series data
- Developing a Foundation Model for lake ecosystems to generalize across unseen time-series distributions, while also incorporating cross-frequency learning & capturing time-invariant entity characteristics
- Building ML models with Modular Compositional Learning for 2D Lake Hydrodynamics prediction
- Benchmarked zero-shot effectiveness & reasoning ability of SOTA Vision-Language Models (VLMs) like GPT-4, LlaVa

## Industry Experience

#### ThinkSense Inc. | Machine Learning Engineer 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

- End-to-end ML model development, from data extraction to deployment on ETL pipeline, leveraging Spark systems
- Built an Accounts Receivable Delay Prediction app using GBT Regression (SparkML), deployed to the NetSuite Warehouse
- 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

#### PASS Consulting Group | Summer Intern

May 2018 - July 2018

• Developed a MLP-based model using Keras and sklearn, to automate motor valve open/close ops for a water SCADA system

## Publications

- 1. Abhilash Neog, Arka Daw, Sepideh Fatemi, Anuj Karpatne. "Masking the Gaps: An Imputation-Free Approach to Time Series Modeling with Missing Data". Workshop on Time-Series in the Age of Large Models, NeurIPS 2024
- 2. 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
- 3. 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". (Under Review)
- 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.

- 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? ZReport 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

## 

• 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 CCode CReport 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) CCode CReport 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 tuning for the respective tasks

## 

• 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, HTML, C

Frameworks: PyTorch, Tensorflow Keras, Git, Spark, Spring Boot

## Talks & Awards

2024 February: Presented a Poster on Transfer Learning in Lake Ecosystems at the "NSF Macrosystems Biology Meeting".

2023 May: Gave a Lightning Talk at the "Frontiers in Ecological Forecasting" event at Virginia Tech.

2021 December: Awarded "Star of the Month" within the Oracle Analytics Cloud Organization, Oracle India