SUMANTH MANDURU

+1-571-585-9446 ◇ sumanth.manduru@gmail.com ◇ sumanthmanduru ◇ Github ◇ GScholar

EDUCATION

George Mason University *August 2024 - Present*
PhD in Computer Science advised by ***Prof. Carlotta Domeniconi*** Fairfax, VA

George Mason University *August 2022 - May 2024*
Masters in Computer Science, ***Specialization*** in Machine Learning Fairfax, VA

WORK EXPERIENCE

Summer Research Fellowship - Generative AI May 2023 - August 2023
22^{nd} *Century Technologies, Inc.*

- Assembled keyword extraction and text summarization pipeline using KeyBERT and BART, boosting AI powered chatbot responsiveness and search efficiency by 30%.
- Integrated Llama Index, FastAPI and S3 Storage, enhancing document processing by 60%.

Graduate Research Assistant, George Mason University, USA September 2022 - Present
Data Mining & Machine Learning Lab and C5I Center

- Implemented the context-aware recommendation system using neural networks and factorization models, improving prediction accuracy by 5% through optimized contextual feature integration.
- Engineered a novel clustering algorithm for UAV categorization, visualizing traversal paths of 50+ unique drones using Plotly, enhancing data interpretation and analysis efficiency by 35%.
- Evaluated a PPO-based Reinforcement Learning agent for high-speed aerospace vehicle emergency descents, achieving robust performance across a wide range of initial conditions, ±10% to ±40%, in altitude and velocity.

Data Scientist May 2019 - August 2022
Jio AI-CoE

- Constructed REST APIs for validating Self Causality, Forecasting, and Clustering Pipelines on Demand Planning and Discount Optimization, boosting sales by 18% for AJIO.
- Achieved 93% accuracy in document classification using logistic regression and neural networks and built BERT for NER with 40 financial entities, saving $50K and 200 man-hours annually.
- Architected comprehensive AI framework: Interpretation Engine with 20 Explainable AI methods and Ensemble Engine with techniques like Bagging, Boosting etc, enhancing model interpretability and performance by 25%.
- Crafted LSTM model for ad-copy generation using custom Google Ads data, boosting impressions by 74% and conversions by 29% with efficient keyword seeding.
- Devised predictive model using similarity measures to classify real-time leads, blocking irrelevant placements and keywords, saving clients $100 per lead.
- Programmed EDA dashboard for real-time KPI tracking across multiple Ad platforms, exploiting PySpark scalability for seamless processing, optimizing daily budget adjustments, increasing ROI by 15%.

TECHNICAL SKILLS

Languages	Python, SQL, C/C++, HTML/CSS, Linux
Tools	Apache Spark, FastAPI, Flask, AWS (S3, SageMaker), GitHub
Frameworks	PyTorch, Transformers, HuggingFace, LangChain, Llama Index, TensorFlow, Keras

AWARDS

Distinguished Academic Achievement, George Mason University, May 2024

Data Science Research Contributor, PathCheck Foundation (MIT Spin-off), 2021.

PUBLICATIONS

Viswanatha Reddy G., Chaitanya B.S.N.V., Prathyush P., **Sumanth M.**, et al. DFW-PP: Dynamic Feature Weighting based Popularity Prediction for Social Media Content. *The Journal of Supercomputing, 2023.*

Mohsin Ali, Sai Teja Kandukuri, **Sumanth Manduru**, et al. PESTO: Switching Point Based Dynamic and Relative Positional Encoding for Code-Mixed Languages. **AAAI** 2022, *Proceedings of the* 36^{th} *AAAI Conference on Artificial Intelligence.*

RESEARCH PROJECTS

Neural Network(s) Pruning: One and Ensembles *Advisor - Prof. Carlotta Domeniconi*
Python, PyTorch, Matplotlib, SLURM

- Pioneered DropNet algorithm, pruning up to 80% of ResNet and VGG filters while preserving accuracy, demonstrating significant model compression.
- Synthesized pruned DNN ensembles using Hierarchical Pruning, boosting accuracy and reducing computational costs by 15% on CIFAR-10 dataset.

Modeling Long Documents Using Graph Neural Networks *Prof. Ziyu Yao*
PyTorch, HuggingFace, Googletrans, Pandas, Numpy

- Reproduced HiPool for long document classification, securing 53% accuracy on custom translated multilingual datasets using BERT Multilingual and XLM RoBERTa.
- Conducted a thorough comparative analysis to evaluate robustness of HiPool BERT with GNNs, leveraging techniques like Word Dropping, Jumbling, and Misspelling.