RISHABH SRIVASTAVA

+1 (929) 358-5907 | rs4489@columbia.edu | linkedin.com/in/rishabhsrivastava6 | github.com/RishabhS66

Education

Columbia University — New York, NY
MS in Computer Science (Machine Learning Track), GPA: 3.99/4.00 — Expected Dec 2024
Relevant Courses: Natural Language Processing, Machine Learning, High-Performance ML, Databases
TA for: Topics in Software Engineering, Advanced Software Engineering
Recipient of Data Science Institute Scholarship (Fall 2024)

Indian Institute of Technology Guwahati — Assam, IN
BTech in Electronics and Electrical Engineering, Minor in Computer Science — Jul 2021
Relevant Courses: Computer Vision, Probability, Data Structures and Algorithms
Recipient of Samsung Fellowship Award

Technical Skills

Languages: Python, CUDA, C++, Java, MySQL, MongoDB, MATLAB, React, NodeJS, TypeScript
Technologies/Frameworks: PyTorch, Scikit-learn, TensorFlow, OpenCV, vLLM, Wandb, AWS, GCP, Kubernetes, Docker

Work Experience

Rubicon Robotics Inc. — New York, NY
Software Engineer Intern — May 2024 – *Present*

- Developed and implemented CV algorithms for swimmer detection by SwimBot, attaining a **90%** accuracy rate. Employed OpenPose model for comprehensive posture analysis.
- Created Django backend interfacing with AWS RDS, deployed site using AWS EC2 behind Application Load Balancers and Route53 for custom domain assignment.
- Established CI/CD pipeline using GitHub Actions, boosting development efficiency and site reliability.

Adobe Inc. - Adobe Experience Manager (AEM) Assets — Noida, IN
Software Development Engineer Level II — Jul 2021 – Aug 2023

- Spearheaded enhancement of AEM Assets Search by utilizing Lucene indexing for efficient information retrieval, Hugging Face's BLIP APIs for asset auto-captioning and GPT-4 for query pre-processing.
- Led end-to-end implementation of Smart Tags Block-list in AEM Assets Essentials, empowering users to manage and block inappropriate smart tags for assets, ensuring content appropriateness and brand compliance.
- GenAI Hackathon - integrated Adobe Firefly to improve search experience for AEM Assets Essentials, allowing customers to generate custom images if search results are irrelevant; selected to be presented at Adobe EMEA Summit 2023.

Research Experience

Advanced Research in Software Engineering (ARISE) Lab, Columbia University — New York, NY
Research Assistant under Prof. Baisakhi Ray — May 2024 – Aug 2024

- Fine-tuned DeepSeek-Coder-V2-Lite-base using custom-built PYX dataset to get SemCoder-S, a semantic-aware CodeLLM.
- Conducted experiments comparing SemCoder-S with other CodeLLMs , achieving superior performance with F1 score of **0.678** for code correctness and **62.4%** accuracy for execution prediction on HumanEval-based dataset.

Adobe Inc. — Noida, IN
Media and Data Science Research Intern — Apr 2020 – Jul 2020

- Implemented Reinforcement Learning-based algorithms to extract top relevant patterns from temporal, sequential datasets.
- Trained Deep Q-Network using TF-Agents and extracted patterns ranked by user-specified measure of interest.
- Proposed algorithm allowed monitoring and improving user-targeting based on certain Key Performance Indicators.

Hanyang University — Ansan, KR
Research Intern under Prof. Frank Rhee — May 2019 – Jul 2019

- Designed a new algorithm Adaptive Shadowed C-Means (ASCM), to cluster data using fuzzy and shadowed sets.
- Reduced impact of noise in clustering by keeping outliers concentrated in shadow region.
- Implemented algorithm on Iris dataset and Breast Cancer Wisconsin data set, and demonstrated its use for image segmentation.

Publication

- Rishabh Srivastava, Addrish Roy, "Abstract Art Interpretation Using ControlNet," arXiv preprint, 2024 [arXiv:2408.13287]

Projects

FOMC Statement Hawkish-Dovish Analysis Using LLMs | *Transformers, BeautifulSoup* | Jun 2024 - Aug 2024

Supervisor: Prof Ali Hirsa, and associated with Morgan Stanley

- Used CentralBankRoBERTa to predict market dovishness/hawkishness from FOMC statements and meeting minutes from Jan 2019 to July 2024.
- Web scraped, cleaned, and extended the dataset, then prompt-engineered GPT-4 to classify text into pre-defined categorical labels, benchmarking results against the CentralBankRoBERTa model and MacroMicro AI Hawkish-Dovish index.
- Stress tested models by modifying prompts, data size, and order of inputs, and analyzed label consistency across models using Kendall's W.

Inference Acceleration of Stable Diffusion | *PyTorch Lightning, Transformers, Wandb* | Apr 2024 - May 2024

GitHub: RishabhS66/Inference-Acceleration-of-Stable-Diffusion

- Devised Time-step calibrated quantization for Stable Diffusion, achieving the lowest FID score and highest CLIP score compared to other quantization techniques.
- Conducted L1-unstructured pruning and combined quantization, compressing the model by **20%** and reducing inference time by **5%** without significant performance loss.

Abstract Art Interpretation Using ControlNet | *PyTorch Lightning, Transformers, BLIP* | Apr 2024

GitHub: RishabhS66/Abstract-Art-Interpretation-Using-ControlNet

- Leveraged ControlNet and Stable Diffusion to enhance spatial control over image composition and enable interpretation of abstract art through detailed geometric conditions.
- Developed a custom dataset of **14,279** image pairs to train model, achieving high-quality image generation with innovative artistic representations.

Lexical Substitution Task with WordNet, Word2Vec Embeddings, and BERT | *NLTK, Transformers* | Nov 2023

GitHub: RishabhS66/Lexical-Substitution-using-BERT

- Devised a novel fusion strategy, combining BERT's contextual understanding with Word2Vec's semantic similarity and WordNet's semantic relations, to improve lexical substitution accuracy and suggest contextually fitting word replacements.
- Attained a precision of **0.189** and recall of **0.189** on 206 attempted instances with mode-specific scoring.