Samarth Garg

samarthgarg0904@gmail.com | +91-95189-76525 | samarthgarg.vercel.app

linkedin.com/in/samarth-garg-a16145202 \mid github.com/SamarthGarg09

Education

| Integrated M.Tech in Information Technology, ABV-IIITM, Gwalior, India • CGPA: 7.96 | Sep 2020 – Present |
|---|--|
| • Relevant Coursework: Machine Learning, Deep Learning, NLP, Computer Vision, Soft Computing, Linear Algebra, Probability, and Statistics, Data mining, Big Data | Multi-Objective Optimization, Analytics |
| 12th Grade, Sir Padampat Singhania School, Kota, IndiaPercentage: 89.2% | Sep 2019 – 2020 |
| Research Experience | |
| Research Intern at GREYC CNRS lab, UNICAEN – Certificate Link | Jan 2024 – June 2024 |
| • Integrated Dempster-Shafer theory with language models like BERT,GPT-2 models modeling. | for uncertainty-aware |
| • Achieved superior performance on all nine tasks of the GLUE Benchmark compared | d to the plain BERT model. |
| • Tech Stack Used: Pytorch, HuggingFace, Slurm, Matplotlib, Numpy, Scipy | |
| Research Intern at LTRC lab, IIIT Hyderabad – Certificate Link | Jan 2024 – June 2024 |
| Improved models for toxicity detection by desensitizing it towards certain concepts Reported an F1-score and ROC-AUC score that are 6% and 1% respectively, better Tech Stack used: Hugging Face, BitsAndBytes, Pytorch, Pandas, Captum | s like implicit hate speech. than the baseline model. |
| Research Intern at NLP lab, University of Tartu, Estonia – Certificate Link | May 2023 – Oct 2023 |
| Worked on neutralizing biases in low-resource languages like Hindi and Bengali.Fine-tuned models like Indic-BERT, multilingual-T5 on IndicCorp v2. | |
| Data Science Intern, BarRaiser – Certificate Link | May 2022 – Aug 2022 |
| • Worked with GPT-3 models such as DaVinci, Babbage, and Curie to extract essentia transcripts. | al questions and context from |
| Optimized the existing algorithm to incorporate lemmatization, stop-word removalImplemented various string and pattern-matching algorithms to improve the mode | l, and insertion techniques. l's accuracy and efficiency. |
| Publications | |
| Interpretability Based on Concepts for Detecting Toxicity – GitHub Link | Under Review |
| Samarth Garg*, Deeksha Varshney, Rui Mao, Erik Cambria, Ranjan Satapathy | |
| Introduced a concept-based interpretability approach for improving toxicity detection | n models. |
| KTCR: Improving Implicit Hate Detection with Knowledge Transfer-driven Concept Refinement – GitHub Link, Paper Link | Under Review |
| Samarth Garg*, Vivek Hruday Kavuri, Gargi Shroff, Rahul Mishra | |
| Proposed a method for enhancing implicit hate detection using knowledge transfer a techniques. | nd concept refinement |
| MixRevDetect: Aiming to Identify AI-Generated Material in Hybrid Peer Reviews | Under Review |
| Samarth Garg*, Sandeep Kumar*, Sagnik Sengupta, Tirthankar Ghosal, Asif Ekbal | |
| Developed a method for detecting AI-generated content in peer reviews to maintain a | academic integrity. |

* Lead author

Technical Projects

| Text Detoxification – GitHub Link | Aug 2023 – Sept 2023 |
|---|---|
| • Developed a system using GANs to convert hate speech to neutral text with 80% style a | accuracy. |
| • Employed cyclic consistency and adversarial losses between generator and discrimina detoxification. | ator for improved |
| • Implemented contrastive search decoding and curriculum learning to stabilize GAN | training. |
| • Utilized the Yelp dataset for training and evaluation. | |
| Protein Secondary Structure Prediction Using Contact Maps – GitHub Link Developed a multi-modal model combining Distil-ProtBERT and Relational Graph And Relational Graph And Relational Graph And | Apr 2023 – July 2023 nvolutional Networks. |
| • Achieved improvement over current SOTA protein language models; findings communic Journal. | cated at Soft Computing |
| • Created the training pipeline using PyTorch, PyTorch-Geometric, and HuggingFace lib | oraries. |
| Long Document Summarization – GitHub Link | Nov 2022 – Dec 2022 |
| • Implemented extractive summarization using TextRank and abstractive summarization by fine-tuning BART-large-CNN . | |
| • Utilized embeddings from sentence-transformers for relationship representation. | |
| Visual Question Answering (VQA) Task – GitHub Link | May 2022 – Oct 2022 |
| • Developed models combining ResNet-52 with LSTM and using BEiT with BERT transfer | ormers. |
| • Achieved an accuracy of 51% and an F1 score of 61% with the first model. | |

• Explored late fusion techniques for modality integration.

Image Processing and Compression Library – GitHub Link

• Developed a library for basic image processing operations and image compression using the **k-means** algorithm.

April 2021 – Oct 2021

• Provided CLI support for image editing, compression, and batch operations.

Technical Skills

Domains: Natural Language Processing, Explainable AI, Generative AI, Deep Learning, Computer Vision, Machine Learning

Language, Frameworks & Libraries: Python, Letta, Langchain, Pinecone, PyTorch, Hugging Face, Datasets, Captum, Scikit-learn, SciPy, spaCy, NLTK, NumPy, Matplotlib, Seaborn, PyTorch Geometric, SLURM Tools: Git, Linux, Docker

Certifications

- Machine Learning
- Neural Networks and Deep Learning
- Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization
- Sequence Models
- Convolutional Neural Networks
- Natural Language Processing with Probabilistic Models
- Natural Language Processing with Classification and Vector Spaces
- Natural Language Processing in TensorFlow

Academic Honors

- Ranked 300th among 18,500 teams in the Amazon ML Challenge 2024.
- Selected for UNICAEN research internship.
- **GDSC ML Campus Facilitator** for 2021 and 2022. Conducted a 7-day session teaching ML algorithms to nearly 100 students.
- B.Tech Thesis project score: A on Text Detoxification using pre-trained language model.