Kumar Tanmay

Research Fellow
Microsoft Research India

kr.tanmay147@gmail.com | t-ktanmay@microsoft.com
 | Google Scholar | Website
+91-9572127367

Education

Indian Institute of Technology, Kharagpur, India	2018 - 2022

Bachelor of Technology (Hons.) in Instrumentation Engineering
Department: Electrical Engineering | Minor: Computer Science
CGPA: 8.97/10.0, **Department Rank 3**

Rajendra Vidyalaya, Jamshedpur, India	2018

Indian School Certificate (ISC) - 12th Grade - English, Science, Maths, and Computer Science
Percentage score: 95.6%, **Ranked 1st in school**

Rajendra Vidyalaya, Jamshedpur, India	2016

Indian Certificate of Secondary Education (ICSE) - 10th Grade
Percentage score: 96.5%, **Ranked 1st in school**

Talks/Presentations

DUBLIN: Visual Document Understanding By Language-Image Network [Slides, Video]	Dec 2023

EMNLP 2023

A Case and Framework for In-Context Ethical Policies in LLMs [Slides]	Dec 2023

EMNLP 2023

Publications (Google Scholar)

*equal contribution

Kumar Tanmay*, Aditi Khandelwal*, Utkarsh Agarwal*, Monojit Choudhury. "Probing the Moral Development of Large Language Models through Defining Issues Test". *NAACL/COLM*. 2024. (To be submitted, pdf)

Utkarsh Agarwal*, **Kumar Tanmay***, Aditi Khandelwal*, Monojit Choudhury. "Ethical Reasoning and Moral Value Alignment of LLMs Depend on the Language we Prompt them in". *LREC-COLING*. 2024. (Under Review, pdf)

Aditi Khandelwal*, Utkarsh Agarwal*, **Kumar Tanmay***, Monojit Choudhury. "Do Moral Judgment and Reasoning Capability of LLMs Change with Language? A Study using the Multilingual Defining Issues Test". *EACL*. 2024. (Under Review, pdf)

Kumar Tanmay*, Aditi Khandelwal*, Utkarsh Agarwal*, Monojit Choudhury. "Probing the Moral Development of Large Language Models through Defining Issues Test". *Workshop on AI meets Moral Philosophy and Moral Psychology (MP^2) - Neurips*. 2023. (Accepted, Workshop, pdf)

Abhinav Rao*, Aditi Khandelwal*, **Kumar Tanmay***, Utkarsh Agarwal*, Monojit Choudhury. "Ethical Reasoning over Moral Alignment: A Case and Framework for In-Context Ethical Policies in LLMs". *Findings of EMNLP*. 2023. (Accepted, Poster, pdf)

Kriti Aggarwal*, Aditi Khandelwal*, **Kumar Tanmay***, Owais Mohammed Khan, Qiang Liu, Monojit Choudhury, Hardik Hansrajbhai Chauhan, Subhojit Som, Vishrav Chaudhary, Saurabh Tiwary. "DUBLIN: Visual Document Understanding By Language-Image Network" *EMNLP Industry Track*. 2023.(Accepted, Poster, pdf)

Kumar Tanmay*, Kumar Ayush*. "Augmented Reality Based Recommendations based on Perceptual Shape Style Compatibility with Objects in the Viewpoint and Color Compatibility with the Background". *Advances in Image Manipulation Workshop (AIM) - ICCV*. 2019. (Accepted, Workshop, pdf)

Kumar Ayush*, Burak Uzkent*, **Kumar Tanmay**, Marshall Burke, David Lobell, Stefano Ermon. "Efficient Poverty Mapping from High Resolution Remote Sensing Images". *AAAI*. 2021. (Accepted, Oral, pdf)

Shuvam Chakraborty, Kumar Ayush*, Burak Uzkent*, **Kumar Tanmay**, Evan Sheehan, Stefano Ermon. "Efficient Conditional Pre-training for Transfer Learning". *Workshop on Learning with Limited Labelled Data for Image and Video Understanding (L3D-IVU) - CVPR*. 2022. (Accepted, Workshop, pdf)

Kumar Ayush*, Burak Uzkent*, Chenlin Meng*, **Kumar Tanmay**, Marshall Burke, David Lobell, Stefano Ermon. "Geography-Aware Self-Supervised Learning". *ICCV*. 2021. (Accepted, Poster, Webpage, pdf,)

Patents

Kumar Tanmay, Monojit Choudhury, Subhojit Som, Vishrav Chaudhary, Saurabh Tiwary. "A Highly Optimized Variable Input Resolution Training Strategy for Vision Transformer Models". 2023. (Filed)

Kumar Tanmay, Monojit Choudhury, Subhojit Som, Vishrav Chaudhary, Saurabh Tiwary. "A Template-based Multimodal Instruction Tuning Method to Enhance Visual Document Understanding". 2023. (Filed)

Work Experience

Research Fellow, Turing Team , Microsoft Research, India
Visual Document Understanding, Image QA, Table QA   Jul 2022 - Aug 2023
Advisors: Dr. Subhojit Som, Vishrav Chaudhary and Prof. Dr. Monojit Choudhury

- Co-led the development of DUBLIN (*accepted in EMNLP Industry Track 2023*), a large-scale transformer-based encoder-decoder model for visual document understanding (VDU) that can analyze both text and visual elements in document images (including infographics, webpages, tables, forms, and natural images) achieving state-of-the-art-performance on diverse downstream tasks such as question-answering, information extraction, bounding box prediction, summarization, image captioning, and classification.
- Standardized 20 multimodal datasets into a unified benchmark format, implementing metrics such as CIDEr and ANLS for comprehensive model evaluation.
- Developed a novel and efficient method to handle variable input resolution images, enabling DUBLIN to process documents with varying aspect ratios resulting in significant performance gains for long documents.
- Introduced a novel training strategy as part of curriculum learning and incorporated synthetic table data derived from private logs to enhance the model's ability to understand tables.
- Developed a template-based multimodal instruction tuning method to handle a wide range of diverse tasks by creating templates for various tasks thus eliminating the need for external task-specific layers.
- Achieved significant performance improvements, with an average 5% enhancement across 20 benchmarking datasets. Notable gains were observed in AI2D (21%), InfographicsVQA (7.5%), and DocVQA (5.4%) compared to existing pixel-only and specialized pipeline based state-of-the-art models.
- Curated multimodal datasets for TableQA and ImageQA to train DUBLIN to support real-world production scenarios.

Ethical Generalisation and Value Pluralism in Large Language Models (LLMs)   Dec 2022 - Oct 2023
Advisor: Prof. Dr. Monojit Choudhury

- Developed a psychometric assessment tool inspired by the Kohlbergian Moral Development and Defining Issues Test frameworks to evaluate the moral judgment and reasoning capabilities of LLMs, including GPT-4, LLaMA, PaLM, and GPT3.x series. *Accepted in MP[2] Neurips 2023.*
- Established a comprehensive framework designed to facilitate the infusion of ethical policies for moral alignment in LLMs levraging in-context learning to address complex social dilemmas characterized by conflicting values. *Accepted in Findings of EMNLP 2023.*
- Conducted an in-depth examination of the "Foreign Language Effect" in LLMs, with a specific focus on understanding the moral reasoning abilities of these models within multilingual contexts. *Two papers under review in EACL 2024 and LREC-COLING 2024.*

Cryptographically Secure Large Language Models   Oct 2023 - Present
Advisor: Dr. Nishant Chandra

- Engaged in instruction training the Turing Language Model to differentiate instructions from data, employing cryptographic delimiters to prevent man-in-the-middle threats. Focus is on bolstering model security against jailbreak attacks, addressing challenges unaddressed by prompt engineering or post-processing.

Multilingual Large Language Modeling   Sep 2023 - Present
Advisors: Vishrav Chaudhary and Prof. Dr. Monojit Choudhury

- Leading efforts to improve the reasoning capabilities of Turing Language Generation Models in multilingual settings, focusing on transformer architectures with 6.7B and 13.6B parameters.
- Established a robust 2.2M multilingual instruction dataset using GPT-4 and GPT-3.5 Turbo, employing Chain-of-Thought Prompting.
- Analyzed tokenizer fertility across 100+ languages to address tokenization impact, and implemented a systematic pipeline for filtering undesirable samples from the multilingual dataset.
- Conducted extensive instruction tuning experiments, resulting in a remarkable 25% average performance improvement over non-instruction-finetuned models.

Faster Inference Architectures   Oct 2023 - Present

Advisor: Dr. Tejas Indulal Dhamecha

- Currently, I am engaged in the process of adapting the Early-Exit optimization approach, initially implemented in Tensorflow, to the PyTorch framework. This adaptation aims to enhance the efficiency of Turing's in-house language generation models by minimizing latency and conserving computational resources. The primary goal is to achieve these improvements without relying on the complete model for the generation process.

Internship Experience

qure.ai, Mumbai
Classification and Localization of Tubes and Catheters in Chest X-Ray Images   Summer 2021
Advisors: Tarun Raj and Dr. Pooja Rao   [qXR-BT]

- Developed an end-to-end pipeline for classification and segmentation of medical entities (tubes & catheters) in Chest X-Rays, addressing class imbalance through a weighted-binary cross-entropy loss function.
- Developed a two-step strategy for precise tube tip localization using deep learning-based segmentation and advanced image processing.
- Conducted extensive experiments with various CNN architectures, surpassing baseline metrics and contributing to the successful integration of models into qure.ai's deep learning stack for the qXR-BT product.

Sustainability and AI Lab, Stanford University, CA
Machine Learning for Socioeconomic, Sustainability and Computer Vision Tasks   Fall 2020
Advisor: Prof. Stefano Ermon

- **Geography-Aware Self-Supervised Learning (ICCV 2021)**: Helped develop and validate a novel contrastive learning method for remote sensing data, leveraging spatio-temporal structures and geo-location showing improvements in image classification, object detection, and semantic segmentation.
- **Efficient Poverty Mapping from High Resolution Remote Sensing Images (AAAI 2021)**: Helped develop and validate a novel reinforcement learning method to optimize acquisition of high-resolution satellite images based on analysis of free low-resolution imagery, enhancing poverty prediction in Uganda.
- **Efficient Conditional Pre-training for Transfer Learning (CVPRW 2022)**: Helped develop and validated efficient and adaptable methods for pre-training dataset filtering to reduce high training costs, emphasizing performance, adaptability, and flexibility.

Undergraduate Research Projects

Medical Tube Abnormality Detection in Chest-X Rays using Deep Learning   Spring 2022
Advisors: Prof. Pabitra Mitra   [Bachelor's Thesis]

Developed DL algorithms for abnormal tube position detection in Chest X-Rays, achieving a 24% improvement via knowledge distillation. Created a proprietary dataset, incorporating segmentation masks for anatomical regions and ideal tip positions, leading to a 35% avg. improvement in abnormalities detection across tube types. Developed CNN-based segmentation models and applied image processing techniques for optimal detection.

Typologically Diverse QA - Zero-Shot and Few Shot Language Jackknifing   Fall 2021
Self-Project   [Project]

Investigated cross-lingual generalization of multilingual BERT (mBERT) using TYDI QA dataset, exploring its performance on QA tasks across various languages. Identified linguistic feature impacts on zero-shot transfer, finding that fine-tuning mBERT as a language model on QA questions significantly improved zero-shot cross-lingual understanding, addressing question interpretation challenges.

Effect of Negative Labels on Conditional Generative Adversarial Networks   Spring 2021
Self-Project   [Project]

Proposed a novel data augmentation technique for stabilizing the training of class-conditional Generative Adversarial Networks (CGANs), addressing issues such as non-convergence and mode-collapse. Additionally, introduced a new evaluation metric based on the classification ability of the Discriminator, demonstrating its positive correlation with visual inspection, surpassing traditional metrics like Inception score and FID.

Augmented Reality Based Context Aware Recommendations   Summer 2019
Self-Project   [Project]

Developed a novel consumer targeting system through modeling AR-based data. Created persuasive recommendations by using statistical modeling, an exemplar part-based 2D-3D alignment method to find the best matching 3D models of furniture present in the user's preferred purchase viewpoint in the AR app and a combination of 3D style compatibility and color compatibility. *Work published in ICCV Workshop 2019.*

Other Selected Projects

MusicAnalyst, OpenAnalytics (intra-collegiate tech competition), IIT Kharagpur   Spring 2022

> Led a hostel team, developing a neural network based model to predict popularity (highest possible revenue) of music tracks based on features associated with the music. Incorporated a meticulously designed loss function to address class imbalance challenges. *Received Second Prize.*

EthPhoto, OpenSoft (inter-hostel tech competition), IIT Kharagpur   Spring 2022

> Co-led the hostel team, developing a decentralized location & tag based photo-sharing app using Ethereum blockchain technology and IPFS storage. *Received Third Prize.*

Pet Projects ⌂
implemented the following deep learning algorithms in Python/PyTorch

> (a) Knowledge Distillation, (b) Deep Convolutional Generative Adversarial Networks (DCGAN), (c) Global Wheat Detection: Detection of Wheat Heads using Faster-RCNN algorithm; confidence scores calculated using weighted boxes fusion (better than Non-Maximum Suppression), (d) Sentimal Analysis using Naive Bayes and Logistic Regression, (e) Character-based model (using LSTM) for Sonnet Generation.

Awards and Achievements

OpenAnalytics Intra-collegiate Tech Competition (Captain) - Second Prize - IIT Kharagpur   2022
Led a team, developing a neural network based model to predict popularity of music tracks based on the features associated with the music. Competition featured 50 teams.

OpenSoft Inter-hostel Tech Competition (Captain) - Third Prize - IIT Kharagpur   2022
Co-led the hostel team, developing a decentralized photo-sharing application (EthPhoto) using Ethereum blockchain technology and IPFS storage. Competition featured 14 teams.

Honda Young Engineer & Scientist's (Y-E-S) Fellowship Finalist   2021
Amongst 20 finalists from all over India.

Kamalavati Syngal and Goralal Syngal Memorial Scholarship   2019
Awarded for academic excellence at IIT Kharagpur.

Kishore Vaigyanik Protsahan Yojana (KVPY) Scholarship   2017
Awarded to 1323 students from around 100,000 applicants by Dept. of Science and Technology, Govt. of India for exceptional aptitude in basic sciences.

Abhay Seva Sansthan Gold Medal (Class XII)   2018
Awarded for scoring 100% in Computer Science in ISC (12th grade) board examination.

C. D. Sinha Award   2018
Awarded for outstanding performance (95.6%) in ISC (12th grade) board examination.

Abhay Seva Sansthan Gold Medal (Class X)   2016
Awarded for scoring 100% in Computer Science in ICSE (10th grade) board examination.

S. P. Sinha Scholarship   2016
Awarded full scholarship for 11th & 12th grade due to outstanding performance (96.5%) in ICSE (10th grade) board examination.

Technical Skills

Languages: Python, C, C++, R, MATLAB, Octave
Libraries: PyTorch, Pytorch Lightning, Tensorflow, Keras, NumPy, Scikit-learn, Selenium, Pandas, OpenCV, Beautiful Soup, Scipy, NLTK, Flask, Tesseract, EasyOCR

Relevant Coursework

Programming & Data Structures (+Lab)	Signals & Networks (+Lab)	Data Analytics
Mathematics I & II	Theory and App. of Blockchain	Digital Signal Processing
Algorithms I (+Lab)	Probability and Stochastic Processes	Social Computing
Marketing & Market Research	Big Data Processing	Artificial Intelligence
Transform Calculus	Foundation of Educational Technology	

Online MOOCs | Coursera

Machine Learning - Certificate
Neural Networks and Deep Learning - Certificate
Structuring Machine Learning Projects - Certificate
Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization - Certificate
Convolutional Neural Networks - Certificate
Sequence Models - Certificate
Natural Language Processing with Classification and Vector Spaces - Certificate
Natural Language Processing with Probabilistic Models - Certificate
Natural Language Processing with Sequence Models - Certificate

Extra Curricular Activities

Sub-reviewer for ACL 2023 and AACL-IJCNLP 2023.

Teaching Assistant (2023) for a workshop course in Practical NLP and Large Language Models (LLMs) for bachelor's students at the Indian Institute of Science (IISc) as part of Kotak-IISc AI-ML Centre.

Volunteer Teacher at eVidyaloka (2023): Taught Maths and Science to underprivileged children in remote areas of Jharkhand via Skype.

Technovation Mentor for Youth Coding Initiative (2022): Mentored 5 high-school girls to build a business plan and mobile app for autistic children using the MIT app inventor that we piloted with 10 families.

Captain (2022) of the OpenAnalytics (intra-collegiate tech competition) team of my hall of residence, IIT Kharagpur.

Captain (2022) of the OpenSoft (inter-hostel tech competition) team of my hall of residence, IIT Kharagpur.

Student Academic Mentor, Student Welfare Group, IIT Kharagpur (2020 - 2022): Looked after the orientation and guidance of 5 new students to ease their transition into college life.

National Sports Organization (2018 - 2020): Involved in Health & Fitness/Athletics as part of NSO, IIT Kharagpur.