Varun **Gumma**

Graduate Research Assistant, IIT Mardas | AI4Bharat

@ varun230999@gmail.com github.com/VarunGumma Google Scholar
E-34 Shantiniketam, HWP Colony, 507116 twitter.com/VarunGumma23

Education

Present – Aug 2021 | **Indian Institute of Technology (IIT) Madras** — **Chennai, India**
M.Tech. in Computer Science & Engineering
CGPA: 9.49/10
Teaching Exp: Pattern Recognition & Machine Learning, Linear Algebra & Random Processes, Fundamentals of Deep Learning

Jun 2021 – Aug 2017 | **Birla Institute of Technology and Science (BITS) Pilani** — **Hyderabad, India**
B.E. in Computer Science & Engineering with Minor in Physics
CGPA: 9.14/10
Teaching Exp: Introductory Physics, Computer Programming, Discrete mathematics, Data Structures & Algorithms, Foundations of Data Science, Machine Learning

Experience

Present – Aug 2022 | **Indian Institute of Technology, Madras | AI4Bharat** [🌐] — **Chennai, India**
Graduate Research Assistant | Advisors: Dr. Raj Dabre, Dr. Pratyush Kumar, Dr. Mitesh Khapra
Working on understanding the effects of Knowledge Distillation on Multilingual Neural Machine Translation models, specifically for Indian languages.

Jul 2022 – Jun 2022 | **National University of Singapore** [🌐] — **Kent Ridge, Singapore**
Visiting Research Scholar | Advisor: Dr. Aditya Karanam
Worked on developing Machine Learning and Deep Learning models to extract *suggestions* and identify *suggested-features* from noisy user comments using Name-Entity-Recognition.

May 2021 – Jan 2021 | **Birla Institute of Technology and Sciences, Pilani** [🌐] — **Hyderabad, India**
Undergraduate Research | Advisor: Dr. N.L. Bhanu Murthy
Worked on developing a Deep Learning model to automatically generate comments for Java code snippets.

May 2021 – Jan 2021 | **Birla Institute of Technology and Sciences, Pilani** [🌐] — **Hyderabad, India**
Undergraduate Research | Advisor: Dr. Barsha Mitra
Worked on analyzing the performance of various Machine Learning models on ABAC datasets for *Policy Generalization & Augmentation*.

May 2020 – Jan 2020 | **Birla Institute of Technology and Sciences, Pilani** [🌐] — **Hyderabad, India**
Undergraduate Research | Advisor: Dr. N.L. Bhanu Murthy
Worked on fine-tuning BERT for Name-Entity-Recognition in Telugu.

July 2019 – May 2019 | **Defense Research & Development Organization-Research Centre Imarat** [🌐] — **Hyderabad, India**
Student Intern
Worked on building an Object Detection Model to identify cars in a given image.

Publications

S=In Submission, C=Conference, W=Workshop, P=Poster/Demo, J=Journal, A=Arxiv (∗ = equal contribution)

[A.1] IndicTrans2: Towards High-Quality and Accessible Machine Translation Models for all 22 Scheduled Indian Languages
AI4Bharat, Jay Gala*, Pranjal A. Chitale*, Raghavan AK, Sumanth Doddapaneni, Varun Gumma, Aswanth Kumar, Janki Nawale, Anupama Sujatha, Ratish Puduppully, Vivek Raghavan, Pratyush Kumar, Mitesh M. Khapra, Raj Dabre, Anoop Kunchukuttan
e-Print Archive [**Arxiv**]

[C.2] An Empirical Study of Leveraging Knowledge Distillation for Compressing Multilingual Neural Machine Translation Models
Varun Gumma, Raj Dabre, Pratyush Kumar
24^{th} *Annual Conference of The European Association of Machine Translation* [**EAMT'23**]

[C.1] PAMMELA: Policy Administration Methodology using Machine Learning
Varun Gumma, Barsha Mitra, Soumyadeep Dey, Pratik Shashikantbhai Patel*, Sourabh Suman*, Saptarshi Das, Jaideep Vaidya
19^{th} *International Conference on Security and Cryptography* [**SECRYPT'22**]

Select Research Projects

Emperical Investigation of Knowledge Distillation for MNMT models Aug'22 - Present
Advisors: Dr. Raj Dabre, Dr. Pratyush Kumar, Dr. Mitesh Khapra

- Researched to explore the application of end-to-end Knowledge Distillation to MNMT models, followed by a comprehensive analysis of available KD methods for NMT. The insights were then applied to distill IndicTrans.
- Investigated the performance of extreme parameter shared MNMT models and compared the performance of wider-vs-deeper models. Demonstrated that fine-tuning with High-Quality translation pairs can improve the model's performance.
- Analyzed the effect of N-way parallel translation pairs in many-to-one translation scenarios and concluded that they could have a detrimental impact. The study was expanded by experimenting with different data scales and model sizes to provide more comprehensive insights.

Suggestion Mining from Noisy User Comments Jun'22 - July'22
Advisor: Dr. Aditya Karanam

- Conducted a performance analysis of multiple Machine Learning and Deep Learning models, including vanilla Conditional Random Fields (CRF) and BERT-CRF, to identify *suggested features* from user comments data with a high noise level.
- Created a hierarchical pipeline that involves *suggestion classification* using TF-IDF vectors and SVMs, followed by *suggested features* extraction with BERT-CRF.

Automatic Code Comment Generation Jan'21 - May'21
Advisor: Dr. N.L. Bhanu Murthy

- Implemented an LSTM encoder-decoder model that inputs source code and Abstract Syntax Trees and generates comments for the corresponding code snippets. This model performs similarly to a vanilla Transformer trained on the same dataset.
- The model utilizes a *copy-Attention* mechanism and a *pointer-generator network* to reduce the number of unknowns in the target comment by directly copying tokens from the source code snippet. Additionally, the model is enhanced with a *coverage* regularization technique to minimize the repetition of tokens when generating the target sequence.

Policy Administration using Machine Learning Jan'21 - May'21
Advisor: Dr. Barsha Mitra

- Conducted an analysis of multiple Machine Learning models to address the ABAC Policy-Inference-Problem (ABAC-PIP), which involves extracting a new set of attribute-based rules from an existing policy.
- The models were trained on a predetermined set of *University* and *Project-Management* policies and then evaluated on a similar but slightly different set of policies to assess their generalization ability.

Relevant Coursework

M.Tech. Advanced Data Structures & Algorithms, Pattern Recognition & Machine Learning, Fundamentals of Deep Learning, Natural Language Processing, Reinforcement Learning, Linear Programming & Combinatorial Optimization
B.E. Software Engineering, Foundations of Data Science, Machine Learning, Deep Learning, Quantum Information and Computing

Skills

Programming Languages: Python, Java, C++, C, SQL, LaTeX
Libraries & Frameworks: Fairseq, PyTorch-Lightning, PyTorch, TensorFlow, Scikit-Learn, Weights & Biases, Unix, GIT

Awards and Achievements

IIT Madras STAR TA, 2022 For outstanding contributions as a Graduate Teaching Assistant

GATE CS&IT, 2021 Secured an All India Rank of 159 with a score of 816/1000

BITS Merit Scholarship, 2018 For meritorious academic performance in the year 2017-2018