


# 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

|                       |   |                  |
|-----------------------|---|------------------|
| June 2023<br>Aug 2021 | <b>Indian Institute of Technology (IIT) Madras</b><br>M.Tech. in Computer Science & Engineering<br>CGPA: 9.62/10<br><b>Teaching Exp:</b> Pattern Recognition & Machine Learning, Linear Algebra & Random Processes, Fundamentals of Deep Learning   | Chennai, India   |
| Jun 2021<br>Aug 2017  | <b>Birla Institute of Technology and Science (BITS) Pilani</b><br>B.E. in Computer Science & Engineering with Minor in Physics<br>CGPA: 9.14/10<br><b>Teaching Exp:</b> Introductory Physics, Computer Programming, Discrete mathematics, Data Structures & Algorithms, Foundations of Data Science, Machine Learning | Hyderabad, India |

## Experience

|                       |   |                       |
|-----------------------|---|-----------------------|
| Present<br>May 2023   | <b>Microsoft Research</b> <br>Research Intern / Advisor: <a href="#">Dr. Sunayana Sitaram</a> ,<br>Working on debiasing of Large Language Models (LLMs).   | Bangalore, India      |
| May 2023<br>Aug 2022  | <b>Indian Institute of Technology, Madras   AI4Bharat</b> <br>Graduate Research / Advisors: <a href="#">Dr. Raj Dabre</a> , <a href="#">Dr. Pratyush Kumar</a> , <a href="#">Dr. Mitesh Khapra</a><br>Working on understanding the effects of Knowledge Distillation on Multilingual Neural Machine Translation models, specifically for Indian languages. | Chennai, India        |
| Jul 2022<br>Jun 2022  | <b>National University of Singapore</b> <br>Visiting Research Scholar / Advisor: <a href="#">Dr. Aditya Karanam</a><br>Worked on developing Machine Learning and Deep Learning models to extract <i>suggestions</i> and identify <i>suggested-features</i> from noisy user comments using Name-Entity-Recognition.                                       | Kent Ridge, Singapore |
| May 2021<br>Jan 2021  | <b>Birla Institute of Technology and Sciences, Pilani</b> <br>Undergraduate Research / Advisor: <a href="#">Dr. N.L. Bhanu Murthy</a><br>Worked on developing a Deep Learning model to automatically generate comments for Java code snippets.   | Hyderabad, India      |
| May 2021<br>Jan 2021  | <b>Birla Institute of Technology and Sciences, Pilani</b> <br>Undergraduate Research / Advisor: <a href="#">Dr. Barsha Mitra</a><br>Worked on analyzing the performance of various Machine Learning models on ABAC datasets for <i>Policy Generalization &amp; Augmentation</i> .  | Hyderabad, India      |
| May 2020<br>Jan 2020  | <b>Birla Institute of Technology and Sciences, Pilani</b> <br>Undergraduate Research / Advisor: <a href="#">Dr. N.L. Bhanu Murthy</a><br>Worked on fine-tuning BERT for Name-Entity-Recognition in Telugu.   | Hyderabad, India      |
| July 2019<br>May 2019 | <b>Defense Research &amp; Development Organization-Research Centre Imarat</b> <br>Student Intern<br>Worked on building an Object Detection Model to identify cars in a given image.  | Hyderabad, India      |

- [A.1] **IndicTrans2: Towards High-Quality and Accessible Machine Translation Models for all 22 Scheduled Indian Languages**  
 AI4Bharat, Jay Gala\*, Pranjali 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<sup>th</sup> 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<sup>th</sup> International Conference on Security and Cryptography [SECRYPT'22]

## Select Research Projects

### Empirical 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,  $\text{\LaTeX}$
- Libraries & Frameworks:** HuggingFace Transformers, 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