Satwik **Bhattamishra**

PhD student, University of Oxford

satwikb.com | satwik55@gmail.com | github.com/satwik77 | Google Scholar

Education

Present | **University of Oxford** | **Oxford, UK**
Oct 2021 | Ph.D. in Computer Science | DeepMind Scholarship
Advisors: Prof. Phil Blunsom and Prof. Varun Kanade

May 2019 | **Birla Institute of Technology and Science Pilani** | **Pilani, India**
Aug 2014 | B.E. (Hons.), Computer Science and Int. M.Sc. (Hons), Biological Science

Experience

Oct 2023 | **Cohere** | **London, UK**
June 2023 | *Research Intern | Foundations team*
Worked on building large language models with linear recurrent architectures. Implemented attention-free architectures and worked on optimizing various components to improve performance. Trained a 7B linear recurrent base model that performed competitively with a 7B Transformer.

July 2021 | **Microsoft Research** | **Bangalore, India**
July 2019 | *Research Fellow | Advisors: Dr. Navin Goyal, Dr. Monojit Choudhury*
Worked on analyzing the abilities of Transformers and LSTMs in modelling different behaviours and properties relevant to modelling language. Also worked on building robust and interpretable methods for improving compositional generalization in semantic parsing problems.

Jan 2019 | *Research Intern | Advisors: Dr. Navin Goyal, Dr. Monojit Choudhury*
Worked on Semantic Parsing problems, particularly on NL-to-SQL problem. Explored the usage of graph neural networks to encode sentence as well as the table schema to improve generalization performance.

Dec 2018 | **Machine and Language Learning Lab, IISc** | **Bangalore, India**
June 2018 | *Research Intern (Senior Thesis) | Advisor: Prof. Partha P. Talukdar*
Worked on increasing diversity in paraphrase generation while retaining fidelity (quality). To incorporate diversity in the paraphrase generation framework, we introduced a submodular optimization-based approach for generating diverse meaning-preserving paraphrases.

Aug 2017 | **Agency of Science Technology & Research** | **Singapore**
May 2017 | *Summer Research Intern | Advisor: Dr. Anders Skanderup*
Worked on improving methods for learning low-level representations of gene expression data based on various optimization and regularization methods of non-negative matrix factorization.

Aug 2016 | **Google Summer of Code 2016 [🌐]** | **Remote**
May 2016 | *Student Developer | Mentors: Dr. Chris Mungall, Dr. Dan Keith*
Worked on the Phenopacket scraper project which extracts information from texts scraped from life sciences websites, analyzes them and generates a phenopacket based on the correct ontology references.

Publications

Understanding In-Context Learning in Transformers by Learning to Learn Discrete Functions [pdf][code]
Satwik Bhattamishra, Arkil Patel, Phil Blunsom, Varun Kanade
2024 International Conference on Learning Representations **[Oral]** [ICLR'24]

Simplicity Bias in Transformers and their Ability to Learn Sparse Boolean Functions [pdf] [code]
Satwik Bhattamishra, Arkil Patel, Varun Kanade, Phil Blunsom
2023 Annual Meeting of the Association for Computational Linguistics [ACL'23]

MAGNIFICo: Evaluating the In-Context Learning Ability of Large Language Models to Generalize to Novel Interpretations [pdf]
Arkil Patel, Satwik Bhattamishra, Siva Reddy, Dzmitry Bahdanau
2023 Conference on Empirical Methods in Natural Language Processing [EMNLP'23]

DynaQuant: Compressing Deep Learning Training Checkpoints via Dynamic Quantization [pdf]
A. Agrawal, S. Reddy, Satwik Bhattamishra, V. Prabhakara Sarath Nookala, V. Vashishth, K. Rong, A. Tumanov
2023 Arxiv (Under Review) [Preprint'23]

Revisiting the Compositional Generalization Abilities of Neural Sequence Models [pdf] [code]
Arkil Patel, Satwik Bhattamishra, Phil Blunsom, Navin Goyal
2022 Annual Meeting of the Association for Computational Linguistics [ACL'22]

Are NLP Models really able to Solve Simple Math Word Problems? [pdf] [code]
Arkil Patel, Satwik Bhattamishra, Navin Goyal
2021 Conference of North American Chapter of the Association for Computational Linguistics [NAACL'21]

On the Ability and Limitations of Transformers to Recognize Formal Languages [pdf] [code]
Satwik Bhattamishra, Kabir Ahuja, Navin Goyal
2020 Conference on Empirical Methods in Natural Language Processing [EMNLP'20]

On the Practical Ability of Recurrent Neural Networks to Recognize Hierarchical Languages [pdf] [code]
Satwik Bhattamishra, Kabir Ahuja, Navin Goyal
2020 International Conference on Computational Linguistics **[Best Short Paper Award]** [COLING'20]

On the Computational Power of Transformers and Its Implication in Sequence Modeling [pdf] [code]
Satwik Bhattamishra, Arkil Patel, Navin Goyal
2020 Conference on Computational Natural Language Learning [CoNLL'20]

On the Ability of Self-Attention Networks to Recognize Counter Languages [pdf]
Satwik Bhattamishra, Kabir Ahuja, Navin Goyal
5th Workshop on Representation Learning for NLP [Extended Abstract] [Repl4NLP@ACL'20]

Unsung Challenges of Building and Deploying Language Technologies for Low Resource Language Communities
P. Joshi, C. Barnes, S. Santy, S. Khanuja, S. Shah, A. Srinivasan, Satwik Bhattamishra, S. Sitaram, M. Choudhury, K. Bali
16th International Conference on Natural Language Processing [ICON'19]

Submodular Optimization-based Diverse Paraphrasing and its Effectiveness in Data Augmentation [pdf] [code]
Satwik Bhattamishra*, Ashutosh Kumar*, Manik Bhandari, Partha Talukdar (* = Equal Contribution)
2019 Conference of North American Chapter of the Association for Computational Linguistics **[Oral]** [NAACL'19]

Selected Research Projects

Analysis of RNNs on Context-free languages
Advisor: Dr. Navin Goyal — Jan'20 - Aug'20

- Worked on investigating the disparity between the results of previous works examining the ability of RNNs to model nested dependencies based on context-free languages and natural languages.
- Provided theoretical and empirical evidence to show that RNNs can recognize context-free languages with bounded depth which have closer connections to natural language. This led to a publication at **COLING'20**.

Semantic Parsing Problems
Advisors: Dr. Navin Goyal, Dr. Monojit Choudhury — Jan'19 - Aug'21

- Worked on building methods to solve Math Word Problems. Conducted various experiments to show that existing models rely on shallow heuristics to solve the problem. The work led to a publication at **NAACL'21**.
- Explored approaches to build more robust and interpretable model to solve math word problems.
- Previously explored graph neural networks based approaches for NL-to-SQL problem.

Analysis of Sequence Models on Formal Languages
Advisor: Dr. Navin Goyal — July'19 - Aug'20

- Initially worked on analyzing the computational power of complete Transformer architecture. Showed the components necessary for Turing-completeness of the network. This work led to a publication at **CoNLL'20**.
- Subsequently worked on analyzing the encoder-only version of Transformers. Showed that they can recognize certain counter languages but are limited in their ability to recognize regular languages. This work got accepted at **EMNLP'20**.
- Provided theoretical and empirical evidence to show that RNNs can recognize context-free languages with bounded depth which have closer connections to natural language. This led to a publication at **COLING'20**.

Diverse Paraphrase Generation
Advisor: Prof. Partha P. Talukdar — June'18 - Dec'18

- Explored the domain of diverse subset selection methods to build a method for diverse generation of paraphrases. Devised a method based on submodular optimization to generate diverse paraphrases while having constraints over the quality of the generated outputs.
- Also demonstrated that the paraphrases generated by our proposed method is more effective for data augmentation as compared to several other baselines. This work was accepted at **NAACL'19**.

Selected Dev Projects

LibNMF **[Github]**

An easy to use python library with implementations of a set of tested optimization and regularization methods of Non-Negative Matrix Factorization (NMF). Implemented Algorithms include graph regularized NMF, probabilistic NMF, a first-order primal-dual algorithm ...etc.

pyDPP **[Github]**

Developed a python package available in pip (Python packaging index) with modules for sampling from Determinantal Point Processes (DPP). Contains implementations of algorithms to sample from DPPs that encourage diversity in the selection of a subset of points from a grounded superset.

Machine Learning Contests **[Kaggle Profile]**

Kaggle Level : Competitions Expert. **Silver** medal in Kaggle Satander Value Prediction Challenge, Rank: Pvt. 185^{th} | Pub. 189^{th}/4484. **Bronze** medal in Kaggle Instacart Competiton, Rank: Pvt. 195^{th} | Pub. 74^{th}/2623. Qualified for **Zonal** Round in India Hacks Machine Learning Competition by Hackerearth, Rank: 29^{th}/860.

Review Miner | Microsoft Code.Fun.Do Hackathon **[Github]**

Developed a cross-platform application which analyses reviews from commercial websites and provides insights about products based on keyword extraction and sentiment analysis. **Winner** of Hackathon at BITS Pilani, Rank 1^{st}/90+.

Services, Teaching and Leadership Roles

Teaching Assistant $\times 2$ **Computational Learning Theory, Oxford** Fall 22 and Fall 23

Responsible for taking six classes on learning theory concepts and problem-solving as well as for marking the answer sheets of students. Topics include PAC-learning, SQ-learning, VC dimension, Rademacher complexity, etc.

Teaching Assistant **Neural Networks and Fuzzy Logic, BITS Pilani** Jan'18 - May'18

Responsible for designing and evaluating 3 programming tests and a course project. Conducted tutorials on python and numpy for over 70 students.

Joint-Coordinator, DVM **Oasis, BITS Pilani** June'16 - Nov'16

Head of the department of over 40 skilled enthusiasts. Responsible for building websites, mobile apps, videos and registration systems. Our systems managed over 1000 participants over the course of four days.

Reviewer **ICML 2024, 2023, Neurips 2023, 2022, ACL 2023, 2022, ICLR 2022, EMNLP 2022, 2021, 2020, NAACL 2021, ACL Rolling Review**