Satwik **Bhattamishra**

@ t-satbh@microsoft.com   ○ github.com/satwik77   in linkedin.com/in/satwik77

○ Microsoft Research, VIGYAN, Lavelle Road, Bangalore, KA, India 560001

Education

May 2019 **August 2014**	**Birla Institute of Technology and Science Pilani, Pilani, India** B.E. (Hons.), Computer Science and Int. M.Sc. (Hons), Biological Science CGPA : 8.1/10

Experience

Present **July 2019**	**Microsoft Research, Bangalore, India** *Research Fellow	Advisor: Dr. Navin Goyal* Working on analyzing the abilities of Transformers and LSTMs in modeling different behaviour and properties relevant for modeling language. Analyzed the expressiveness and learning abilities of the networks in modeling context-free and counter languages. Prior work also includes analysis of components necessary for Turing-completeness of Transformers.
January 2019	*Research Intern	Advisor: Dr. Navin Goyal* Worked on Semantic Parsing problems, particularly on NL-to-SQL problem. Explored the possibility of using graph networks to encode sentence as well as the table schema (context) to improve the model.
December 2018 **June 2018**	**Machine and Language Learning Lab, IISc, Bangalore, India** *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.
August 2017 **May 2017**	**Agency of Science Technology & Research, Singapore** *Summer Research Intern	Advisor: Dr. Anders Skanderup* Worked on improving methods for learning lower-level representations of gene expression data. Work involved analysis of gene expression data and exploring various optimization and regularization methods of non-negative matrix factorization for better deconvolution and clustering of gene expression data.
August 2016 **May 2016**	**Google Summer of Code 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 extracted information and correct external ontology references.

Publications

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

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

On the Practical Ability of Recurrent Neural Networks to Recognize Hierarchical Languages [Link]
Satwik Bhattamishra, Kabir Ahuja, Navin Goyal
2020 International Conference on Computational Linguistics   [`COLING'20`]

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

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

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`]

Selected Projects

Machine Learning Contests
[Kaggle Profile]

> Kaggle Level : **Competitions Expert**. Global Rank: **5270**th / 140,000+ active users
> **Silver** medal in Kaggle Satander Value Prediction Challenge, Rank: Pvt. 185th | Pub. 189th/4484
> **Bronze** medal in Kaggle Instacart Competiton, Rank: Pvt. 195th | Pub. 74th/2623
> Qualified for **Zonal** Round in India Hacks Machine Learning Competition by Hackerearth, Rank: 29th/860

libNMF
numpy, scipy, cvxopt | [Github]

> An easy to use python library with implementations of a set of tested optimization and regularization methods of NMF.
> Implemented Algorithms include graph regularized NMF, probabilistic NMF, a first-order primal-dual algorithm ...etc
> Most of the algorithms in the library are not found in popularly used libraries of NMF such as nimfa and scikit-learn.

pyDPP
numpy, scipy | [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.

Review Miner | Microsoft Code.Fun.Do
nltk, Naive Bayes, Sentiment analysis | [Github]

> A cross-platform application which analyses reviews from commercial websites.
> Extracts features, analyzes sentiments and classifies associated keywords as features using Naive Bayes classifier based on opinion.
> Generates interactive visualizations based on the analysis for the users.
> **Winner** of Microsoft Code.Fun.Do at BITS Pilani, Rank 1st/90+

Open Source Contributions

> **Monarch Initiative:** Phenopacket scraper project - *commits*
> **Honeynet:** Thug's Rumal Project - *commits*
> **Systers:** Webapp Project - *commits*

Teaching and Leadership Roles

May 2018 **Jan 2018**	**Neural Networks and Fuzzy Logic - Spring 2018** *Teaching Assistant* Responsible for designing and evaluating 3 programming tests and a course project. Conducted tutorials on python and numpy.
November 2016 **June 2016**	**Oasis, BITS Pilani** *Joint-Coordinator, DVM* 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.

Skills

Languages	Python • C/C++ • Java • LaTeX • HTML
Frameworks	Django • Flask
Machine Learning	Numpy • PyTorch • scikit-learn • Keras • cvxopt • scikit-image
Cloud	Azure • DigitalOcean
Spoken Languages	English, Hindi, Odia