Ratish Puduppully

ratishpuduppully@gmail.com
+65 86170655
https://ratishsp.github.io
http://www.linkedin.com/in/ratishsp/

Research Interests

I am a dedicated and accomplished Research Scientist with a PhD in Informatics, and a specialization in Natural Language Processing (NLP) and Deep Learning. My academic and research journey at prestigious institutions has equipped me with a robust understanding of various projects in NLP and Machine Learning. I have made innovative contributions in the field of Natural Language Generation including planning for long document generation. In multilingual research, I have contributed to developing several notable systems and models, including Brahmi-Net and IndicBART.

Education

Sep 2017 - Nov 2021 — PhD in Informatics, University of Edinburgh.
Advisor: Prof. Mirella Lapata

Dec 2014 - Feb 2017 — MS in CSE by Research.
IIIT Hyderabad
Grade: 9.33 CGPA

Jul 2001 – Jun 2005 — B.E in Electronics and Telecommunications.
Mumbai University.
Grade: equiv. to 9 CGPA.

Research Publications

Preprint

1. Mundra, N., Doddapaneni, S., Dabre, R., Kunchukuttan, A., **Puduppully**, **R.**, & Khapra, M. M. (2023). *A comprehensive analysis of adapter efficiency*. arXiv: 2305.07491 [cs.CL]

Journal Articles

1. **Puduppully**, **R.**, Fu, Y., & Lapata, M. (2022). Data-to-text generation with variational sequential planning. *Transactions of the Association for Computational Linguistics (TACL) (to appear), abs/2202.13756.* Retrieved from https://arxiv.org/abs/2202.13756
2. **Puduppully**, **R.**, & Lapata, M. (2021). Data-to-text generation with macro planning. *Transactions of the Association for Computational Linguistics (TACL).* doi:https://doi.org/10.1162/tacl_a_00381

Conference Proceedings

1. **Puduppully**, **R.**, Jain, P., Chen, N., & Steedman, M. (2023). Multi-document summarization with centroid-based pretraining. In *Proceedings the 61st annual meeting of the association for computational linguistics (to appear).* arXiv: 2208.01006 [cs.CL]
2. Kumar, A., Shrotriya, H., Sahu, P., Mishra, A., Dabre, R., **Puduppully**, **R.**, … Kumar, P. (2022). IndicNLG benchmark: Multilingual datasets for diverse NLG tasks in Indic languages. In *Proceedings of the 2022 conference on empirical methods in natural language processing* (pp. 5363–5394). Abu Dhabi, United Arab Emirates: Association for Computational Linguistics. Retrieved from https://aclanthology.org/2022.emnlp-main.360

3. Dabre, R., Shrotriya, H., Kunchukuttan, A., **Puduppully**, **R.**, Khapra, M., & Kumar, P. (2022). IndicBART: A pre-trained model for indic natural language generation. In *Findings of the association for computational linguistics: Acl 2022* (pp. 1849–1863). doi:10.18653/v1/2022.findings-acl.145

4. **Puduppully**, **R.**, Dong, L., & Lapata, M. (2019a). Data-to-text generation with entity modeling. In *Proceedings of the 57th annual meeting of the association for computational linguistics* (pp. 2023–2035). doi:10.18653/v1/P19-1195

5. **Puduppully**, **R.**, Dong, L., & Lapata, M. (2019b). Data-to-text generation with content selection and planning. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*. doi:https://doi.org/10.1609/aaai.v33i01.33016908

6. **Puduppully**, **R.**, Zhang, Y., & Shrivastava, M. (2017). Transition-based deep input linearization. In *Proceedings of the 15th conference of the European chapter of the association for computational linguistics: Volume 1, long papers* (pp. 643–654). Valencia, Spain: Association for Computational Linguistics. Retrieved from https://www.aclweb.org/anthology/E17-1061

7. **Puduppully**, **R.**, Zhang, Y., & Shrivastava, M. (2016). Transition-based syntactic linearization with lookahead features. In *Proceedings of the 2016 conference of the north American chapter of the association for computational linguistics: Human language technologies* (pp. 488–493). doi:10.18653/v1/N16-1058

Workshop/ Demonstrations Proceedings

1. Gehrmann, S., Bhattacharjee, A., Mahendiran, A., Wang, A., Papangelis, A., Madaan, A., … Hou, Y. (2022). *GEMv2: Multilingual NLG benchmarking in a single line of code*. Abu Dhabi, UAE: Association for Computational Linguistics. Retrieved from https://aclanthology.org/2022.emnlp-demos.27

2. **Puduppully**, **R.**, Mallinson, J., & Lapata, M. (2019). *University of Edinburgh's submission to the document-level generation and translation shared task*. doi:10.18653/v1/D19-5630

3. Kunchukuttan, A., **Puduppully**, **R.**, & Bhattacharyya, P. (2015). *Brahmi-net: A transliteration and script conversion system for languages of the Indian subcontinent*. doi:10.3115/v1/N15-3017

4. Bhingardive, S., **Puduppully**, **R.**, Singh, D., & Bhattacharyya, P. (2014). *Merging verb senses of Hindi WordNet using word embeddings*. Goa, India: NLP Association of India. Retrieved from https://www.aclweb.org/anthology/W14-5148

Patents

2016 Method and system for sharing content. US Patent 9,256,695. Willis, B. Natraj,S., Shinde, S., Agarwal,T., **Puduppully, R.**, Santhi Pulagala S. and Chang S.

Research Positions

September 2022 - Present — Scientist III, A*STAR Research
I work on research related to Large Language Models (LLMs) for machine translation, math language processing, and other applications.

Dec 2021 - July 2022 — Research Associate, University of Edinburgh.
I worked with Prof. Mark Steedman in the area of multi-document summarization.

Jun - Oct 2019 — Intern, Google Research London.
I interned with Ryan McDonald's text summarization team at Google Research London. I worked on researching recurrent neural network grammar based approaches to jointly generate summary and its parse tree.

Research Positions (continued)

Mar - Aug 2017 — Research Assistant, Singapore University of Technology and Design.
I was part of Prof. Yue Zhang's NLP lab. I worked in the area of financial market prediction using text data.

May - Dec 2014 — Research Engineer, Center for Indian Language Technology (CFILT) lab at IIT Bombay.
I worked in areas of Machine Translation and Word Sense Disambiguation under the guidance of Prof. Pushpak Bhattacharyya.

Teaching Experience

2021 — Tutor, Demonstrator and Marker for Natural Language Understanding, Generation, and Machine Translation

2018 — Tutor for Accelerated Natural Language Processing

Employment History

2005 – 2008 — Software Engineer, Infosys Technologies Ltd.

2008 – 2014 — Technical Architect, R&D Division, Saba Software, Mumbai.

Skills

Languages — Strong reading, writing and speaking competencies for English, Hindi, Marathi and Malayalam.

Programming Languages — Regular Use: Python. Less recent use: Java, C++, Javascript

Machine Learning Tools — Regular use: Pytorch. Less recent use: Tensorflow, Dynet.

Miscellaneous Experience

Awards and Achievements

2022 — Best PhD Dissertation in Scotland award from SICSA Scotland

2017 — Edinburgh Global Research Scholarship and Principal's Career Development Scholarship for pursuing PhD studies at University of Edinburgh.

Travel Grant from ACM India for presenting paper at EACL conference in Valencia, Spain.

2016 — Travel Grant from Microsoft for presenting paper at NAACL conference in San Diego, US

2002-04 — JRD Tata Scholarship for Academic Excellence for undergraduate studies.

Service

2018-2021 — **Reviewer**. ACL: Annual Meeting of the Association for Computational Linguistics.

2018-2020* — **Reviewer**. EMNLP: Conference on Empirical Methods in Natural Language Processing
* Outstanding reviewer

Volunteering

2018-2020 — Digital Ambassador at University of Edinburgh.
Volunteered as a Digital Ambassador to help improve digital literacy amongst people, mainly elderly persons in community. The project won the 2019 University of Edinburgh Social Responsibility and Sustainability Community Partnership Award.