Girishkumar Ponkiya

Data Scientist, ISS India
Ph.D, IIT Bombay, Mumbai; M.Tech, ISI Kolkata

✉ girishp@cse.iitb.ac.in | girish.isical@gmail.com
☎ +91-9004009268 / 7021387332
⌂ http://www.cse.iitb.ac.in/~girishp
⌽/girishponkiya | in/in/girishponkiya | ≋/users/1726226

OBJECTIVE

To delve into the problem of intelligent information access – especially with natural language, learn continuously in the process and make contributions in developing valuable solutions.

EDUCATION

Qualification	Board/University	Year	CPI / %
Ph.D (NLP; Computer Science)	IIT Bombay, Mumbai	2023	7.7 CPI
M.Tech (Computer Science)	Indian Statistical Institute, Kolkata	2013	78.67 %
B.E. (Computer Engineering)	Saurashtra University, Rajkot	2009	66.77 %

Relevant Course Work

IIT Bombay: Foundation of ML, NLP and the Web, AI, Topics in NLP (Machine Translation).

ISI Kolkata: Pattern Recognition, Algorithms, Graph Theory, Probability and Stochastic Process.

EXPERIENCE/INTERNSHIP

- **Data Scientist (AVP)**, ISS ESG, Mumbai *Mar 2021 to present*
  Developed an efficient news filtering pipeline to track ESG-related controversies in a timely manner. Developed (and deployed) multiple NLP services, such as NER, entity-linking, and article grouping.

- **Machine Learning Consultant**, UnFound, Mumbai *Dec 2018 to Jun 2019*
  Responsibilities: Advise on ML pipeline and assist fellow team members in ML-related tasks.
  Outcome: Re-defined question answering and stance detection pipeline using BERT; Developed multi-task learning approach for stance detection.

- **Research Intern**, TRDDC, Pune *May 2014 to July 2014*

- **Research Intern** (Theoretical Computer Science), IMSc, Chennai *May 2012 to July 2012*

- **Trainer & Software Developer** (Java and PHP), Ishu InfoNet, Rajkot (GJ) *Sep 2009 to May 2010*

- **Software Developer Trainee** (Java/J2EE), HCL CDC, Ahmedabad *Jan 2009 to Apr 2009*
  (Best Trainee award) .

TECHNICAL SKILLS

- **Expertise**: NLP (Natural Language Processing), ML (Machine Learning).

- **Relevant Libraries/Tools**: Python, NumPy, Pandas, scikit-learn (sklearn), SciPy, NLTK, CoreNLP, Theano, PyTorch, TensorFlow, Keras.

PUBLICATIONS

[1] Siddhesh Pawar, Shyam Thombre, Anirudh Mittal, Girishkumar Ponkiya, and Pushpak Bhattacharyya. "Tapping BERT for Preposition Sense Disambiguation". In: *arXiv preprint arXiv:2111.13972* (2021). URL: https://arxiv.org/abs/2111.13972.

[2] Girishkumar Ponkiya, Diptesh Kanojia, Pushpak Bhattacharyya, and Girish Palshikar. "FrameNet-assisted Noun Compound Interpretation". In: *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. 2021, pp. 2901–2911. URL: https://aclanthology.org/2021.findings-acl.256.

[3] Girishkumar Ponkiya, Rudra Murthy, Pushpak Bhattacharyya, and Girish K Palshikar. "Looking inside Noun Compounds: Unsupervised Prepositional and Free Paraphrasing using Language Models". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings (Findings of EMNLP)*. 2020. URL: https://www.aclweb.org/anthology/2020.findings-emnlp.386.

[4] Girishkumar Ponkiya, Kevin Patel, Pushpak Bhattacharyya, and Girish K Palshikar. "Treat us like the sequences we are: Prepositional Paraphrasing of Noun Compounds using LSTM". In: *The 27th International Conference on Computational Linguistics (COLING)*. Santa Fe, New-Mexico, USA, 2018, pp. 1827–1836. URL: https://www.aclweb.org/anthology/C18-1155.

[5] Girishkumar Ponkiya, Kevin Patel, Pushpak Bhattacharyya, and Girish K Palshikar. "Towards a Standardized Dataset for Noun Compound Interpretation". In: *Language Resources and Evaluation Conference (**LREC**)*. Miyazaki, Japan, 2018, pp. 3092–3097. URL: https://www.aclweb.org/anthology/L18-1489.

[6] Girishkumar Ponkiya, Pushpak Bhattacharyya, and Girish K Palshikar. "On Why Coarse Class Classification is a Bottleneck for Noun Compound Interpretation". In: *13th International Conference on Natural Language Processing (**ICON**)*. 2016, pp. 293–198. URL: http://www.aclweb.org/anthology/W16-6336.

INVITED TALKS

- Talk on *Entities and Relations* at AICTE-ISTE approved STTP program on NLP Novice-to-Pro, Shah & Anchor Kutchhi Engineering College, Mumbai. (Jan 7, 2020)
- Talk on *Noun Compound Interpretation* at Workshop on Natural Language Processing, VIVA Institute of Technology, Virar, June 20-24, 2016.

THESIS AND PROJECT

- **Ph.D Thesis: Noun Compound Interpretation**. *Jul 2013 to Mar 2021*
  Under the guidance of Prof. Pushpak Bhattacharyya and Mr. Girish K. Palshikar (TRDDC, Pune).

  Description: A noun compound is a sequence of two or more nouns acting as a single noun, e.g., *student protest, colon cancer*, etc. The information about how individual nouns in the compound are related (*student(s)* is a <u>doer</u> of the *protest*, whereas *colon* is a <u>location</u> of *cancer*), is missing at the surface level. Interpretation of the noun compound is the task of uncovering such information automatically.

  The thesis proposed (1) an LSTM-based supervised model for prepositional paraphrasing, (2) unsupervised methods for prepositional paraphrasing and free paraphrasing using pre-trained language models, (3) FrameNet-based relations, and (4) zero-shot learning using FrameNet – a lexical resource.

  *Awarded **Best Poster (Technical)*** at Research-Scholar Mela 2016 at CSE Department of IIT Bombay.

- **Priority Search Tree for Secondary Memory and its Application** *Aug 2012 to July 2013*
  M.Tech dissertation under the guidance of Prof. Subhas C. Nandy at ISI Kolkata.

  Objective: Given a very large set of points in R^2, create an external data-structure (on a secondary device) to efficiently report MER (Maximum-area Empty Rectangle) around a given query point.

  Result: (1) Proposed an algorithm to create an external priority search tree with linear space in log-linear time; (2) Show that all the standard queries can be answered efficiently; (3) Proposed an algorithm to solve MER problem with extra $O(\log n)$ space in main memory. *(Nominated for Best Dissertation Award)*

POSITION OF RESPONSIBILITY

- Served as elected **General Secretary** of Hostel-1, IIT Bombay for the year 2016-17. (Acting General Secretary for the year 2017-18)
- **IBM Campus Ambassador** (at ISI Kolkata) for the year 2012-2013.

OTHER ACHIEVEMENTS

- Passed **GATE-2013**. Percentiles: **99.78**; All India Rank: 495.
- Passed **JEST-2013** with 65th rank.
- Cleared C2C Java Exam, conducted by Government of Gujarat and TCS, with percentage of 100%.
- Stood 1st **in Junk Yard War**, part of Integration-2012 (Tech-fest at ISI Kolkata).

Date: September 10, 2023
Place: Mumbai

Girishkumar M. Ponkiya