

Sukanta Sen

sukantasen10@gmail.com • <https://www.iitp.ac.in/~sukanta.pcs15>

EDUCATION

Ph.D., Computer Science & Engineering

2015 – Present

Indian Institute of Technology Patna

Thesis: Low-resource Neural Machine Translation

Advisors: Dr. Asif Ekbal and Prof. Pushpak Bhattacharyya

RESEARCH INTERESTS

Neural Machine Translation, Statistical Machine Translation, Natural Language Processing, Deep Learning for Natural Language Processing.

PUBLICATIONS

- [11] Sukanta Sen, Mohammed Hasanuzzaman, Asif Ekbal, Pushpak Bhattacharyya, Andy Way. Neural Machine Translation of Low-resource Languages using SMT Phrase Pair Injection. In **Journal of Natural Language Engineering** (accepted in 2020)
- [10] Sukanta Sen, Kamal Kumar Gupta, Asif Ekbal, Pushpak Bhattacharyya. Multilingual Unsupervised NMT using Shared Encoder and Language-Specific Decoders. In **ACL 2019**
- [9] Sukanta Sen, Asif Ekbal, Pushpak Bhattacharyya. Parallel Corpus Filtering based on Fuzzy String Matching. In **WMT 2019**
- [8] Sukanta Sen, Kamal Kumar Gupta, Asif Ekbal, Pushpak Bhattacharyya. IITP-MT System for Gujarati-English News Translation Task at WMT 2019. In **WMT 2019**
- [7] Sukanta Sen, Mohammed Hasanuzzaman, Asif Ekbal, Pushpak Bhattacharyya, Andy Way. Take Help from Elder Brother: Old to Modern English NMT with Phrase Pair Feedback. In **CICLing 2019**
- [6] Kamal Kumar Gupta, Sukanta Sen, Asif Ekbal, and Pushpak Bhattacharyya. Improving Low-Resource NMT with Parser Generated Syntactic Phrases. In **CICLing 2019**
- [5] Sukanta Sen, Kamal Kumar Gupta, Asif Ekbal, and Pushpak Bhattacharyya. IITP-MT at WAT 2018: Transformer-based Multilingual Indic-English Neural Machine Translation System. In **WAT 2018** (co-located with **PACLIC 2018**)
- [4] Md Shad Akhtar, Palaash Sawant, Sukanta Sen, Asif Ekbal and Pushpak Bhattacharyya. Improving Word Embedding Coverage in Less-resource Language through Multi-linguality and Cross-linguality: A Case Study with Aspect based Sentiment Analysis. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, Vol. 18, No. 2, Article 15, 2018.
- [3] Md Shad Akhtar, Palaash Sawant, Sukanta Sen, Asif Ekbal and Pushpak Bhattacharyya. Solving Data Sparsity for Aspect Based Sentiment Analysis using Cross-linguality and Multi-linguality. In **NAACL 2018**
- [2] Sabyasachi Kamila, Sukanta Sen, Mohammad Hasanuzzaman, Asif Ekbal, Andy Way, Pushpak Bhattacharyya. Temporality as Seen through Translation: A Case Study on Hindi Texts. In **Machine Translation Summit 2017**
- [1] Sukanta Sen, Debajyoty Banik, Asif Ekbal, Pushpak Bhattacharyya. IITP English-Hindi Machine Translation System at WAT 2016. In **WAT 2016** (co-located with **COLING 2016**)

TECHNICAL & LANGUAGE SKILLS

Programming Languages: Python, C/C++, Java, Shell script
MT Tools: Moses, Sockeye, OpenNMT, Nematus
Deep Learning Library: PyTorch
Web / Publishing: HTML, CSS, JS, JQuery, php, Flask, LaTeX
Platform: Linux
Natural Languages: English (fluent), Bengali (native), Hindi (fluent)

DEPLOYED MT SYSTEMS

Hindi-English Neural Machine Translation System for Judicial Domain
<http://www1.iitp.ac.in/~sukanta.pcs15/jud-mt> (Deployed with Sockeye and Flask)

Neural Machine Translation between Seven Indic Languages and English
<http://www.iitp.ac.in/~sukanta.pcs15/indic-mt> (Deployed with Sockeye and Flask)

TEACHING

Teaching Assistant, Dept. of Computer Science & Engineering,
Indian Institute of Technology Patna

2015 – Present

- OOP and Data Structures Lab (CS201)
- Programming Language Theory and Lab (CS331)
- Compilers Lab (CS347)
- Artificial Intelligence Theory and Lab (CS571)

GRANTS / AWARDS

Received travel support from Microsoft Research India to present our paper at the ACL 2019

OTHER ACTIVITIES

Reviewer ACM TALLIP, Sadhana (by the Indian Academy of Sciences), Computational Linguistics, ACL 2020

Student volunteer for the ACL 2019

Organized GIAN Workshop on Neural Machine Translation, 4-10 December, 2017 at Indian Institute of Technology Patna

System Administrator for AI-NLP-ML Group, Indian Institute of Technology Patna