Sahinur Rahman Laskar

DOB: 20/11/1986 sahinur_rs@cse.nits.ac.in/sahinurlaskar.nits@gmail.com
https://www.linkedin.com/in/sahinur-rahman-laskar-0683ab1b0/
+91-7002460508

Education

2019 – present **Ph.D., CSE** National Institute of Technology, Silchar, India
Field of Study: Machine Translation, *Natural Language Processing.*
Current Status: Thesis submission by 30th October 2022.
Supervisor: Dr. Partha Pakray, Assistant Professor, Co-Supervisor: Prof. Sivaji Bandyopadhyay, Professor & Director, NIT Silchar, India.

2013 – 2015 **M.Tech., IT (7.76)** Assam University Silchar, India
Field of Study: Information Retrieval, *Natural Language Processing.*

2007 – 2011 **B.E., CSE (65.32%)** Assam Engineering College, Guwahati, India

Research Interest

Machine Translation, Natural Language Processing, Deep Learning, Machine Learning.

Programming Skills

Python, Java, C, C++, HTML, JavaScript, PHP, MySQL.

Work Experience and Responsibility

- Worked as a Guest Faculty/Lecturer on contractual position (3 years 4 months) at Assam University (CSE Department), Silchar, India and M.H.C.M Science College, Algapur (CS Department), Silchar, India
- Worked as a Junior Research Fellow (JRF) at Assam University, Silchar, India from 03/08/2015 to 31/01/2016.
- Guest speaker for demonstration of Deep Learning for Machine Translation on 29/06/2021 at Faculty development Program on Deep Learning for Natural Language Processing (DL-NLP 2021), Department of CSE and IT, Jaypee Institute of Information Technology, Noida, India.
- Worked as a student volunteer in a conference on "International Conference on Big Data, Machine Learning and Applications (BigDML 2019)" held from 16/12/2019 to 19/12/2019.
- Worked as a student volunteer in a workshop on "Deep Learning Techniques and Tools: An Academic and Industrial Approach" held from 08/04/2019 to 12/04/2019.

Research Participation and Achievement

- Participated in Shared Task: Propaganda Detection in Arabic at WANLP, EMNLP-2022, Our team name: CNLP-NITS-PP and achieved 3rd rank position in subtask-1. Result Link: https://sites.google.com/view/propaganda-detection-in-arabic/results?authuser=0
- Participated in Shared Task: Code-mixed Machine Translation (MixMT), EMNLP 2022 Sixth Conference on Machine Translation (WMT22). Our Team: CNLP-NITS-PP.
- Participated in Shared Task: WAT2022 The 9th Workshop on Asian Translation - “Indic-task: MultiIndicMT”. Our Team name: CNLP-NITS-PP.

- Participated in Shared Task: WAT2022 The 9th Workshop on Asian Translation - “Multi-Modal Translation Task for English to Bengali”. Achieved best results. Our Team name: CNLP-NITS-PP. Link: http://lotus.kuee.kyoto-u.ac.jp/WAT/evaluation/list.php?t=235&o=7.
- Participated in Shared Task: WAT2022 The 9th Workshop on Asian Translation - “Multi-Modal Translation Task for English to Hindi”. Achieved Second best results. Our Team name: CNLP-NITS-PP. Link: http://lotus.kuee.kyoto-u.ac.jp/WAT/evaluation/list.php?t=96&o=7
- Achieved best paper award presenter at 10th International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA 2022) June 18-19, 2022. Paper Title: "Improving English-Assamese Neural Machine Translation using Transliteration-based Approach" Link:https://ficta.org/assets/doc/FICTA-2022_List%20of%20Best%20Paper%20Awards.pdf
- Participated in Shared Task: Similar Language Translation, EMNLP 2021 Sixth Conference on Machine Translation (WMT21). Our Team: CNLP-NITS, achieved 4th rank for Tamil to Telugu and 6th rank for Telugu to Tamil.
- Participated in Shared Task: WAT2021 The 8th Workshop on Asian Translation - “Multi-Modal Translation Task for English to Hindi”. Our Team name: CNLP-NITS-PP.
- Participated in NLP HACK 2021 organized by CIE IIIT Hyderabad, India, April 2-3, 2021. Team Name: CNLP- NITS-PP.
- Participated in Shared Task: Fake News Detection in the Urdu Language (UrduFake) – FIRE 2020. Our Team name: CNLP-NITS and achieved 2nd rank.
- Participated in Shared Task: Similar Language Translation, EMNLP 2020 Fifth Conference on Machine Translation (WMT2020). Our Team: NITS-CNLP, achieved 10th rank for Hindi to Marathi and 15th rank for Marathi to Hindi.
- Participated in Shared Task: WAT2020, The 7th Workshop on Asian Translation - “Multi-Modal Translation Task for English to Hindi”. Our Team name: CNLP-NITS.
- Participated in Shared Task: 3rd Workshop on Technologies for MT of Low Resource Languages (LoResMT 2020) on zero-shot NMT: Russian-Hindi.
- Participated in Shared Task: Similar Language Translation, ACL 2019: Fourth Conference On Machine Translation (WMT2019), August 1-2, 2019 Florence, Italy – Our Team: NITS-CNLP, achieved the best result for Nepali-Hindi language pair, achieved first rank for both Hindi-to-Nepali and Nepali-to-Hindi translation.
- Participated in Shared Task: WAT2019 The 6th Workshop on Asian Translation – “Multi-Modal Translation Task”, Evaluation Week: Aug 03-10, 2019, and November 3-4, 2019: WAT2019 takes place. Our Team name: 683, achieved first rank.

Research Publications

Patent (2)

- “SYSTEM ZUR VORBEREITUNG UND UNTERSUCHUNG EINES ENGLISCH-MIZO- KORPUS DURCH AUFFINDEN VON TONALEN WÖRTERN” ("System for Preparing and Investigating an English-Mizo corpus by encountering Tonal Words") – German Utility Patent, 03 August 2022 (Status: Filed)
- “A System on Answering Questions and Recommending Recipes based on Ingredients”, – German Utility Patent, 08 September 2022 (Status: Filed)

Journal Articles (3 SCIE, 3 SCOPUS)

1. Khenglawt, V., **Laskar, Sahinur Rahman**, Pakray, P., & Khan, A. K. (2022). Machine translation for low-resource english–mizo pair encountering tonal words. *Computación y Sistemas (SCOPUS), 26*(3). Retrieved from https://www.cys.cic.ipn.mx/ojs/index.php/CyS/article/viewFile/4358/3432

2. **Laskar, Sahinur Rahman**, Khilji, A. F. U. R., Pakray, P., & Bandyopadhyay, S. (2022). Improved neural machine translation for low-resource english-assamese pair. *Journal of Intelligent Fuzzy Systems (SCIE) (Impact Factor:* 1.851*), 42*, 4727–4738. doi:https://doi.org/10.3233/JIFS-219260

3. **Laskar, Sahinur Rahman**, Pakray, P., & Bandyopadhyay, S. (2022). Investigation of negation effect for english–assamese machine translation. *Indian Academy of Sciences (Sadhana) (SCIE) (Impact Factor:* 1.188*)*. (Accepted) (In press).

4. **Laskar, Sahinur Rahman**, Paul, B., Pakray, P., & Bandyopadhyay, S. (2022a). English-assamese multimodal neural machine translation using transliteration-based phrase augmentation approach. *ICMLDE-22, Procedia Computer Science Journal, Elsevier (SCOPUS)*. (Accepted) (In press).

5. Khilji, A. F. U. R., Manna, R., **Laskar, Sahinur Rahman**, Pakray, P., Das, D., Bandyopadhyay, S., & Gelbukh, A. (2021). CookingQA: Answering questions and recommending recipes based on ingredients. *Arabian Journal for Science and Engineering (SCIE) (Impact Factor:* 2.334*), 46*(4), 3701–3712. doi:https://doi.org/10.1007/s13369-020-05236-5

6. Rahman Khilji, A. F. U., Manna, R., **Laskar, Sahinur Rahman**, Pakray, P., Das, D., Bandyopadhyay, S., & Gelbukh, A. (2020). Question classification and answer extraction for developing a cooking qa system. *Computación y Sistemas (SCOPUS), 24*(2), 927–933. doi:https://doi.org/10.13053/cys-24-2-3445

Conference/Workshop Proceedings

1. Adhikary, P. K., Manna, R., **Laskar, Sahinur Rahman**, & Pakray, P. (2022). Ontology-based healthcare hierarchy towards chatbot. In *Computational intelligence in communications and business analytics, crc press* (pp. 326–335). doi:https://doi.org/10.1007/978-3-031-10766-5_26

2. Khenglawt, V., **Laskar, Sahinur Rahman**, Manna, R., Pakray, P., & Kumar Khan, A. (2022a). Mizo visual genome 1.0: A dataset for english-mizo multimodal neural machine translation. In *Artificial intelligence, data science computing, ieee silcon 2022*. (Accepted) (In press).

3. Khenglawt, V., **Laskar, Sahinur Rahman**, Manna, R., Pakray, P., & Kumar Khan, A. (2022b). Recent trends on low-resource neural machine translation and research scope for english-mizo pair. In *International conference on intelligent computing systems and applications, springer, 2022*. (Accepted) (In press).

4. Khenglawt, V., **Laskar, Sahinur Rahman**, Pal, S., Pakray, P., & Khan, A. K. (2022). Language resource building and english-to-mizo neural machine translation encountering tonal words. In *Proceedings of the wildre-6 workshop @lrec2020, marseille, european language resources association (elra)* (pp. 48–54). Retrieved from http://www.lrec-conf.org/proceedings/lrec2022/workshops/WILDRE6/pdf/2022.wildre6-1.9.pdf

5. **Laskar, Sahinur Rahman**, Dadure, P., Manna, R., Pakray, P., & Bandyopadhyay, S. (2022). English to bengali multimodal neural machine translation using transliteration-based phrase pairs augmentation. In *9th workshop on asian translation wat2022, colling*. (Accepted) (In press).

6. **Laskar, Sahinur Rahman**, Manna, R., Pakray, P., & Bandyopadhyay, S. (2022). Investigation of multilingual neural machine translation for indian languages. In *9th workshop on asian translation wat2022, colling*. (Accepted) (In press).

7. **Laskar, Sahinur Rahman**, Paul, B., Pakray, P., & Bandyopadhyay, S. (2022b). Improving english-assamese neural machine translation using transliteration-based approach. In *Proceedings of the international conference on frontiers of intelligent computing: Theory and applications, ficta 2022*. (Accepted) (In press).

8. **Laskar, Sahinur Rahman**, Singh, R., Karim, M. F., Manna, R., Pakray, P., & Bandyopadhyay, S. (2022). Investigation of english to hindi multimodal neural machine translation using transliteration-based phrase pairs augmentation. In *9th workshop on asian translation wat2022, colling*. (Accepted) (In press).

9. **Laskar, Sahinur Rahman**, Darsh, A. F. U. R. K., Pakray, P., Bandyopadhyay, S. et al. (2021). Enkhcorp1. 0: An english–khasi corpus. In *Proceedings of the 4th workshop on technologies for mt of low resource languages (loresmt2021), mtsummit* (pp. 89–95). Retrieved from https://aclanthology.org/2021.mtsummit-loresmt.9

10. **Laskar, Sahinur Rahman**, Khilji, A. F. U. R., Kaushik, D., Pakray, P., & Bandyopadhyay, S. (2021). Improved english to hindi multimodal neural machine translation. In *Proceedings of the 8th workshop on asian translation (wat2021), ijcnlp* (pp. 155–160). ACL. doi:https://doi.org/10.18653/v1/2021.wat-1.17

11. **Laskar, Sahinur Rahman**, Pakray, P., & Bandyopadhyay, S. (2021a). Neural machine translation for low resource assamese–english. In *Proceedings of the international conference on computing and communication systems: I3cs 2020, nehu, shillong, india* (Vol. 170, p. 35). Springer. doi:https://doi.org/10.1007/978-981-33-4084-8_4

12. **Laskar, Sahinur Rahman**, Pakray, P., & Bandyopadhyay, S. (2021b). Neural machine translation: Assamese–bengali. In *Modeling, simulation and optimization: Proceedings of comso 2020* (pp. 571–579). Springer. doi:https://doi.org/10.1007/978-981-15-9829-6_45

13. **Laskar, Sahinur Rahman**, Paul, B., Adhikary, P. K., Pakray, P., & Bandyopadhyay, S. (2021). Neural machine translation for tamil–telugu pair. In *Proceedings of the sixth conference on machine translation (wmt), emnlp* (pp. 289–292). ACL. Retrieved from https://aclanthology.org/2021.wmt-1.29

14. **Laskar, Sahinur Rahman**, Paul, B., Paudwal, S., Gautam, P., Biswas, N., & Pakray, P. (2021). Multimodal neural machine translation for english-assamese pair. In *2021 international conference on computational performance evaluation (compe)* (pp. 387–392). IEEE. doi:https://doi.org/10.1109/ComPE53109.2021.9752181

15. **Laskar, Sahinur Rahman**, Khilji, A. F. U. R., Pakray, P., & Bandyopadhyay, S. (2020a). EnAsCorp1.0: English-Assamese corpus. In *Proceedings of the 3rd workshop on technologies for mt of low resource languages, aacl* (pp. 62–68). ACL. Suzhou, China. Retrieved from https://aclanthology.org/2020.loresmt-1.9

16. **Laskar, Sahinur Rahman**, Khilji, A. F. U. R., Pakray, P., & Bandyopadhyay, S. (2020b). Multimodal neural machine translation for English to Hindi. In *Proceedings of the 7th workshop on asian translation, aacl* (pp. 109–113). ACL. Suzhou, China. Retrieved from https://aclanthology.org/2020.wat-1.11

17. Khilji, A. F. U. R., **Laskar, Sahinur Rahman**, Pakray, P., & Bandyopadhyay, S. (2020). Urdu fake news detection using generalized autoregressors. In *Ceur workshop proceedings, forum for information retrieval evaluation 2020, december 16-20, 2020, hyderabad, India*. CEUR-WS. Retrieved from http://ceur-ws.org/Vol-2826/T3-3.pdf

18. Khilji, A. F. U. R., **Laskar, Sahinur Rahman**, Pakray, P., Kadir, R. A., Lydia, M. S., & Bandyopadhyay, S. (2020). Healfavor: Dataset and a prototype system for healthcare chatbot. In *2020 international conference on data science, artificial intelligence, and business analytics (databia)* (pp. 1–4). IEEE. doi:https://doi.org/10.1109/DATABIA50434.2020.9190281

19. **Laskar, Sahinur Rahman**, Khilji, A. F. U. R., Pakray, P., & Bandyopadhyay, S. (2020c). Hindi-marathi cross lingual model. In *Proceedings of the fifth conference on machine translation* (pp. 396–401). ACL. Retrieved from https://aclanthology.org/2020.wmt-1.45

20. **Laskar, Sahinur Rahman**, Khilji, A. F. U. R., Pakray, P., & Bandyopadhyay, S. (2020d). Zero-shot neural machine translation: Russian-hindi@ loresmt 2020. In *Proceedings of the 3rd workshop on technologies for mt of low resource languages, aacl* (pp. 38–42). ACL. Retrieved from https://aclanthology.org/2020.loresmt-1.5

21. **Laskar, Sahinur Rahman**, Dutta, A., Pakray, P., & Bandyopadhyay, S. (2019). Neural machine translation: English to hindi. In *2019 ieee conference on information and communication technology* (pp. 1–6). IEEE. doi:https://doi.org/10.1109/CICT48419.2019.9066238

22. **Laskar, Sahinur Rahman**, Pakray, P., & Bandyopadhyay, S. (2019). Neural machine translation: Hindi-nepali. In *Proceedings of the fourth conference on machine translation (volume 3)* (pp. 202–207). ACL. doi:https://doi.org/10.18653/v1/W19-5427

23. **Laskar, Sahinur Rahman**, Singh, R. P., Pakray, P., & Bandyopadhyay, S. (2019). English to hindi multi-modal neural machine translation and hindi image captioning. In *Proceedings of the 6th workshop on asian translation, emnlp* (pp. 62–67). ACL. doi:https://doi.org/10.18653/v1/D19-5205

24. Bhagawati, R., **Laskar, Sahinur Rahman**, & Swain, B. (2016). Documents clustering using quantum clustering algorithm. In *2016 international conference on microelectronics, computing and communications (microcom)* (pp. 1–4). IEEE. doi:https://doi.org/10.1109/MicroCom.2016.7522562

25. **Laskar, Sahinur Rahman**, & Swain, B. (2015). Analyzing quantum probability ranking principle with the concept of hyperspace analogue to language (hal). In *2015 international symposium on advanced computing and communication (isacc)* (pp. 266–271). IEEE. doi:https://doi.org/10.1109/ISACC.2015.7377353

Book Chapters

1. **Laskar, Sahinur Rahman**, Khilji, A. F. U. R., Pakray, P., Kadir, R. A., Lydia, M. S., & Bandyopadhyay, S. (2022). *Healfavor: Machine translation enabled healthcare chat based application.* doi:https://doi.org/10.1201/9781003138013-6

2. Khilji, A. F. U. R., **Laskar, Sahinur Rahman**, Pakray, P., Kadir, R. A., Lydia, M. S., & Bandyopadhyay, S. (2021). *Healfavor: A chatbot application in healthcare.* doi:https://doi.org/10.1201/9781003146810-3

Reviewer

- Journal of Experimental and Theoretical Artificial Intelligence (TETA)
- Transactions on Asian and Low-Resource Language Information Processing (TALLIP)
- International Conference on Intelligence Computing Systems and Applications (ICICSA-2022) (Springer)
- IEEE, SILCON-2022

Visibility: Google Scholar and ResearchGate

Google Scholar Link:
https://scholar.google.com/citations?hl=en&authuser=1&user=fbqD9i8AAAAJ
Citations: 119; h-index:7; i10-index:4 (Accessed on 24/09/2022)
ResearchGate Link: https://www.researchgate.net/profile/Sahinur-Laskar
Citations: 71; h-index: 5; (Accessed on 24/09/2022)

Membership

IEEE Student and Young Professionals.