PRITOM SAHA AKASH

⚲ Siebel School of Computing and Data Science, Urbana, IL 61801

✉ pakash2@illinois.edu | ⌂ pritomsaha.github.io

EDUCATION

University of Illinois at Urbana-Champaign	IL, USA
Ph.D. in Computer Science	*Aug. 2020 – Present*
Institute of Information Technology, University of Dhaka	Dhaka, Bangladesh
B.S and M.S. in Software Engineering	*Jan. 2014 – June 2019*

RESEARCH INTERESTS

NLP (i.e., understanding and using LLMs), **Text Mining** and **Recommender Systems**.

UNIVERSITY EMPLOYMENTS

Research Assistant — Aug. 2024 – Present
University of Illinois at Urbana-Champaign — *IL, USA*
- **Previously**: Aug.'20 – May'22, Aug.'22 – May'23

Co-Instructor — Aug. 2024 – Present
Dept. of CS, University of Illinois at Urbana-Champaign — *IL, USA*
- **Previously**: Jan.'24 – May'24
- **Courses**: CS 598: Understanding and Using LLMs

Teaching Assistant — Aug. 2023 – Dec. 2023
Dept. of CS, University of Illinois at Urbana-Champaign — *IL, USA*
- **Courses**: STAT 107- Data Science Discovery

Short Course Instructor — Jan. 2018 – Sep. 2019
Institute of Information Technology, University of Dhaka — *Dhaka, Bangladesh*
- **Courses:** Web Programming, Web design and Office application

INDUSTRY EXPERIENCES

Applied Scientist Intern (Up-coming) — May 2025 – Aug. 2025
Amazon Prime Video & Studios CoreTech, Amazon — *Seattle, WA, USA*

PhD Student Researcher — Sep. 2022 – Dec. 2024
Co-op of CS@UIUC and eBay Inc. — *IL, USA*
- **Topic**: Recommender System

Applied Scientist Intern — May 2024 – Aug. 2024
Amazon Prime Video & Studios CoreTech, Amazon — *Seattle, WA, USA*
- **Topic**: LLM in Prime Video Search

Research Scientist Intern — May 2023 – Aug. 2023
Almaden Lab, IBM Research — *San Jose, CA, USA*
- **Topic**: Conversational AI and Long-form QA

Software Engineer — Nov. 2019 – Aug. 2020
Samsung R&D Institute Bangladesh — *Dhaka, Bangladesh*
- **Topic:** Samsung Wearable Application Development

Software Engineer Intern — Jan. 2017 – June 2017
Datasoft System Bangladesh Limited — *Dhaka, Bangladesh*
- **Topic:** Website Development

PUBLICATIONS

Google Scholar: scholar.google.com/citations?user=mBDDbmEAAAAJ (*equal contribution, ✉ corresponding author)

14. **Pritom Saha Akash**, Kevin Chen-Chuan Chang. Enhancing Short-Text Topic Modeling with LLM-Driven Context Expansion and Prefix-Tuned VAEs. Findings of the 2024 Conference on Empirical Methods in Natural Language Processing **(EMNLP Findings)**. 2024.

13. Kashob Kumar Roy, **Pritom Saha Akash**, Lucian Popa, and Kevin Chen-Chuan Chang. ConTReGen: Context-driven Tree-structured Retrieval for Open-domain Long-form Text Generation. Findings of the 2024 Conference on Empirical Methods in Natural Language Processing **(EMNLP Findings)**. 2024.

12. Lam Do, **Pritom Saha Akash**, Kevin Chen-Chuan Chang. Unsupervised Open-domain Keyphrase Generation. The 61st Annual Meeting of the Association for Computational Linguistics **(ACL)**. 2023.

11. **Pritom Saha Akash**, Jie Huang, Kevin Chen-Chuan Chang. Coordinated Topic Modeling. The 2022 Conference on Empirical Methods in Natural Language Processing **(EMNLP)**. 2022

10. **Pritom Saha Akash**, Jie Huang, Kevin Chen-Chuan Chang, Yunyao Li, Lucian Popa, ChengXiang Zhai. Domain Representative Keywords Selection: A Probabilistic Approach. Findings of the 60th Annual Meeting of the Association for Computational Linguistics **(ACL Findings)**. 2022.

9. Md. Eusha Kadir, **Pritom Saha Akash**, Sadia Sharmin, Amin Ahsan Ali, and Mohammad Shoyaib. A Proximity Weighted Evidential k Nearest Neighbor Classifier for Imbalanced Data. In Pacific-Asia Conference on Knowledge Discovery and Data Mining, **(PAKDD)** pp. 71-83. Springer, Cham, 2020.

8. Tarek, M. H., Kadir, M. E., Mahbub, M., **Pritom Saha Akash**, Ali, A. A., & Shoyaib, M. (2020, August). Mutual Information based feature selection for nurse care activity recognition. In 2020 Joint 9th International Conference on Informatics, Electronics & Vision (ICIEV) and 2020 4th International Conference on Imaging, Vision & Pattern Recognition (icIVPR) (pp. 1-6). IEEE.

7. Md. Eusha Kadir*, **Pritom Saha Akash***, Sadia Sharmin, Amin Ahsan Ali, and Mohammad Shoyaib. Can a Simple Approach Identify Complex Nurse Care Activity? In Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers **(UbiComp/ISWC Adjunct)**. 2019

6. **Pritom Saha Akash**, Md. Eusha Kadir, Amin Ahsan Ali and Mohammad Shoyaib. Inter-node Hellinger distance based Decision Tree, In Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, **(IJCAI)**. 2019.

5. **Pritom Saha Akash**, Md. Eusha Kadir, Amin Ahsan Ali, Md. Nurul Ahad Tawhid and Mohammad Shoyaib. Introducing Confidence as a Weight in Random Forest. 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST).[Best Presenter Award]

4. **Pritom Saha Akash**, Ali Zafar Sadiq and Ahmedul Kabir. An Approach of Extracting God Class Exploiting Both Structural and Semantic Similarity. 14th International Conference on Evaluation of Novel Approaches to Software Engineering **(ENASE)**. 2019

3. Jeba, Tahmim, Tarek Mahmud, **Pritom Saha Akash**, and Nadia Nahar. God Class Refactoring Recommendation and Extraction Using Context based Grouping. 2019.

2. Md. Eusha Kadir, **Pritom Saha Akash**, Amin Ahsan Ali, Mohammad Shoyaib and Zerina Begum. Evidential SVM for binary classification. ICASERT-2019: 1st International Conference on Advances in Science, Engineering and Robotics Technology. 2019.

1. Ali Zafar Sadiq, Ahmedul Kabir, **Pritom Saha Akash** and Md. Jubair Ibna Mostafa. Analyzing Corrective Maintenance using Change Coupled Clusters at Fix-inducing Changes. 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE). 2019.

PREPRINTS AND UNDER REVIEW

7. Youan Cong, Cheng Wang, **Pritom Saha Akash**✉, and Kevin Chen-Chuan Chang. Query Optimization for Parametric Knowledge Refinement in Retrieval-Augmented Large Language Models. arXiv preprint arXiv:2411.07820 (2024).

6. Qihao Wang*, **Pritom Saha Akash***, Varvara Kollia, Biwei Jiang, Vadim Von Brzeski, Kevin Chen-Chuan Chang. On Category Recommendation: A Cascading Approach. [On Review]

5. **Pritom Saha Akash**, Jie Huang, and Kevin Chen-Chuan Chang. Let the Pretrained Language Models "Imagine" for Short Texts Topic Modeling. arXiv preprint arXiv:2310.15420 (2023).

4. **Pritom Saha Akash***, Kashob Kumar Roy*, Lucian Popa, and Kevin Chen-Chuan Chang. Long-form Question Answering: An Iterative Planning-Retrieval-Generation Approach. arXiv preprint arXiv:2311.09383 (2023).

3. **Pritom Saha Akash***, Trisha Das*, and Kevin Chen-Chuan Chang. TopicAdapt-An Inter-Corpora Topics Adaptation Approach. arXiv preprint arXiv:2310.04978 (2023).

2. **Pritom Saha Akash**, and Kevin Chen-Chuan Chang. Exploring variational graph auto-encoders for extract class refactoring recommendation. arXiv preprint arXiv:2203.08787 (2022).

1. **Pritom Saha Akash**, Wei-Cheng Lai, and Po-Wen Lin. Online aggregation based approximate query processing: A literature survey. arXiv preprint arXiv:2204.07125 (2022).

MASTER'S THESIS

- **Pritom Saha Akash**, An Evidential Inter-node Hellinger Distance based Tree Classifier, *Institute of Information Technology, University of Dhaka, 2019.*

SENIOR PROJECT

Context Sensitive Bengali Spell Checker   Jul. 2017 - Nov. 2017
- A Bengali Spell Checker for detecting and recommending corrections of typographic and contextual errors
- Developed using Python 3.5 and Django 1.11.2

AWARDS & GRANTS

1. IIT Academic Excellence Gold Medal Award, 2019

2. 2019 IEEE Computer Society Bangladesh Chapter's Award for Academic & Research Excellence in recognition of Outstanding Research Accomplishments

3. Winner in a Human activity recognition challenge called *NURSE CARE ACTIVITY RECOGNITION CHALLENGE* as a part of HASCA workshop at Ubicomp 2019

4. Awarded for *IJCAI-AIJ 2019 TRAVEL and ACCESSIBILITY GRANT*

5. Awarded as the best presenter in an international conference, ICREST 2019

6. Awarded as an ICT fellow (2018-19) under ICT Division Bangladesh

ADITIONAL INFORMATION

Reviewer: EMNLP'23,24; ACL'23,24; EACL'24, ARR Oct.'23,Dec.'24
Co-Reviewer: KDD'21,22,23,24; SIGIR'21,22,23,24; CIKIM'21; ICDM'21; WSDM'21

Technical Skills

Programming: Python, Java, Matlab, C/C++, MySQL, JavaScript, R
Software and Tools: PyTorch, scikit-learn, Git, LaTex, IntelliJ, Eclipse, Android Studio