Mir Tafseer Nayeem

🏠 Ph.D. Student in Computing Science, University of Alberta

📮 tafseer.nayeem@gmail.com 🖥 Website 🎓 Scholar GitHub 🐦 Twitter in LinkedIn

📚 Education

University of Alberta   **Edmonton, Canada**
Ph.D. in Computing Science   Sept 2021 – *Current*
Thesis Title: *"User-Centric Modeling and Synthesis in Language Technologies"*
Supervisor: Prof. Davood Rafiei
Committee members: Prof. Osmar Zaiane, Prof. Lili Mou

University of Lethbridge   **Lethbridge, Canada**
M.Sc. in Computer Science
Thesis Title: *"Methods of Sentence Extraction, Abstraction and Ordering for Automatic Text Summarization"*
Supervisor: Prof. Yllias Chali
Committee members: Prof. Wendy Osborn, Prof. John Anvik

Islamic University of Technology   **Dhaka, Bangladesh**
B.Sc. in Computer Science & IT
Thesis Title: *"Designing a Probabilistic Path Selection Model for Vehicular Networks"*
Supervisor: Prof. Muhammad Mahbub Alam

🏆 Awards & Scholarships

Best Short Paper Award at VIS 2025
Best Resource Paper Award at EMNLP 2024
Huawei Ph.D. Fellowship [2022 – 2027]
Nominated for PhD Early Achievement Award
University of Alberta Graduate Recruitment Scholarship
Area Chair Favorite Paper Award at COLING 2018
ACM SIGIR Student Travel Grant
S.G.S Research Dissemination Travel Award
Graduate Student's Association (GSA) Travel Award
School of Graduate Studies (S.G.S) Dean's Scholarship
School of Graduate Studies (S.G.S) International Tuition Award
IUT-OIC Full Free Scholarship

🎓 Selected Publications

Full publication list can be accessed from **Google Scholar**. [* = Equal Contribution]

[1] **Mir Tafseer Nayeem***, Sawsan Alqahtani*, Md Tahmid Rahman Laskar, Tasnim Mohiuddin, and M Saiful Bari, *"Stop Taking Tokenizers for Granted: They Are Core Design Decisions in Large Language Models"*, 19th Conference of the European Chapter of the Association for Computational Linguistics (**EACL 2026**). [↻ Link]

[2] **Mir Tafseer Nayeem** and Davood Rafiei, *"OpinioRAG: Towards Generating User-Centric Opinion Highlights from Large-scale Online Reviews"*, The Second Conference on Language Modeling (**COLM 2025**). [↻ Link]

[3] **Mir Tafseer Nayeem***, Sawsan Alqahtani*, Md Tahmid Rahman Laskar, Tasnim Mohiuddin, and M Saiful Bari, *"Beyond Fertility: Analyzing STRR as a Metric for Multilingual Tokenization Evaluation"*, The Thirty-ninth Annual Conference on Neural Information Processing Systems (**NeurIPS 2025 Workshops**). [↻ Link]

[4] Tong Bao, **Mir Tafseer Nayeem**, Davood Rafiei, and Chengzhi Zhang, *"SurveyGen: Quality-Aware Scientific Survey Generation with Large Language Models"*, The 2025 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2025**). [⟲ Link]

[5] Ridwan Mahbub*, Mohammed Saidul Islam*, **Mir Tafseer Nayeem**, Md Tahmid Rahman Laskar, Mizanur Rahman, Shafiq Joty, and Enamul Hoque, *"From Charts to Fair Narratives: Uncovering and Mitigating Geo-Economic Biases in Chart-to-Text"*, The 2025 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2025**). [⟲ Link]

[6] Md Tahmid Rahman Laskar, Mohammed Saidul Islam, Ridwan Mahbub, Mizanur Rahman, Amran Bhuiyan, Israt Jahan, **Mir Tafseer Nayeem**, Shafiq Joty, Enamul Hoque, and Jimmy Huang, *"Deploying Tiny LVLM Judges for Real-World Evaluation of Chart Models: Lessons Learned and Best Practices"*, The 2025 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2025**) [Industry Track]. [⟲ Link]

[7] Ridwan Mahbub, MS Islam, MTR Laskar, Mizanur Rahman, **Mir Tafseer Nayeem**, and Enamul Hoque, *"The Perils of Chart Deception: How Misleading Visualizations Affect Vision-Language Models"*, The 2025 IEEE Visualization and Visual Analytics (VIS) Conference (**IEEE VIS 2025**). [⟲ Link] 🏆 Best Short Paper Award

[8] Md Tahmid Rahman Laskar, Mohammed Saidul Islam, Ridwan Mahbub, Ahmed Masry, Mizanur Rahman, Amran Bhuiyan, **Mir Tafseer Nayeem**, Shafiq Joty, Enamul Hoque, and Jimmy Huang, *"Judging the Judges: Can Large Vision-Language Models Fairly Evaluate Chart Comprehension and Reasoning?"*, 63rd Annual Meeting of the Association for Computational Linguistics (**ACL 2025**) [Industry Track]. [⟲ Link]

[9] Luis Antonio Gutiérrez Guanilo, **Mir Tafseer Nayeem**, Cristian López, and Davood Rafiei, *"eC-Tab2Text: Aspect-Based Text Generation from e-Commerce Product Tables"*, The 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (**NAACL 2025**) [Industry Track]. [⟲ Link]

[10] **Mir Tafseer Nayeem** and Davood Rafiei, *"KidLM: Advancing Language Models for Children – Early Insights and Future Directions"*, The 2024 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2024**). [⟲ Link] 🏆 Best Resource Paper Award

[11] Mohammed Saidul Islam, Raian Rahman, Ahmed Masry*, Md Tahmid Rahman Laskar*, **Mir Tafseer Nayeem***, and Enamul Hoque, *"Are Large Vision Language Models up to the Challenge of Chart Comprehension and Reasoning"*, The 2024 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2024 Findings**). [⟲ Link]

[12] **Mir Tafseer Nayeem***, Faisal Tareque Shohan*, Samsul Islam, Abu Ubaida Akash, and Shafiq Joty, *"XL-HeadTags: Leveraging Multimodal Retrieval Augmentation for the Multilingual Generation of News Headlines and Tags"*, 62nd Annual Meeting of the Association for Computational Linguistics (**ACL 2024 Findings**). [⟲ Link]

[13] Mohsinul Kabir*, Mohammed Saidul Islam*, Md Tahmid Rahman Laskar, **Mir Tafseer Nayeem**, M Saiful Bari, and Enamul Hoque, *"BenLLMEval: A Comprehensive Evaluation into the Potentials and Pitfalls of Large Language Models on Bengali NLP"*, Joint International Conference on Computational Linguistics, Language Resources and Evaluation (**COLING 2024**). [⟲ Link]

[14] Ali Naeim abadi, **Mir Tafseer Nayeem**, and Davood Rafiei, *"Product Entity Matching via Tabular Data"*, 32nd ACM International Conference on Information and Knowledge Management (**CIKM 2023**). [⟲ Link]

[15] **Mir Tafseer Nayeem** and Davood Rafiei, *"On the Role of Reviewer Expertise in Temporal Review Helpfulness Prediction"*, 17th Conference of the European Chapter of the Association for Computational Linguistics (**EACL 2023 Findings**). [⟲ Link]

[16] **Mir Tafseer Nayeem***, Abu Ubaida Akash*, Faisal Tareque Shohan, and Tanvir Islam, *"Shironaam: Bengali News Headline Generation using Auxiliary Information"*, 17th Conference of the European Chapter of the Association for Computational Linguistics (**EACL 2023**). [⟲ Link]

[17] **Mir Tafseer Nayeem***, Susmoy Chakraborty*, and Wasi Ahmad, *"Simple or Complex? Learning to Predict Readability of Bengali Texts"*, Thirty-Fifth AAAI Conference on Artificial Intelligence (**AAAI 2021**). [⟲ Link]

[18] **Mir Tafseer Nayeem***, Radia Rayan Chowdhury*, Tahsin Tasnim Mim, Md. Saifur Rahman Chowdhury, and Taufiqul Jannat, *"Unsupervised Abstractive Summarization of Bengali Text Documents"*, 16th Conference of the European Chapter of the Association for Computational Linguistics (**EACL 2021**). [🔗 Link]

[19] **Mir Tafseer Nayeem**, Tanvir Ahmed Fuad, and Yllias Chali, *"Neural Diverse Abstractive Sentence Compression Generation"*, 41st European Conference on Information Retrieval (**ECIR 2019**). [🔗 Link]

[20] **Mir Tafseer Nayeem***, Tanvir Ahmed Fuad*, Asif Mahmud, and Yllias Chali, *"Neural Sentence Fusion for Diversity Driven Abstractive Multi-Document Summarization"*, **Computer Speech & Language**, 2019, 58:216-230. [🔗 Link]

[21] **Mir Tafseer Nayeem**, Tanvir Ahmed Fuad, and Yllias Chali, *"Abstractive Unsupervised Multi-Document Summarization using Paraphrastic Sentence Fusion"*, 27th International Conference on Computational Linguistics (**COLING 2018**). [🔗 Link] 🏆 Area Chair Favorite Paper Award

[22] **Mir Tafseer Nayeem** and Yllias Chali, *"Paraphrastic Fusion for Abstractive Multi-Sentence Compression Generation"*, 26th ACM International Conference on Information and Knowledge Management (**CIKM 2017**). [🔗 Link] 💰 ACM SIGIR Student Travel Grant

[23] Yllias Chali, Moin Mahmud Tanvee, and **Mir Tafseer Nayeem**, *"Towards Abstractive Multi-Document Summarization Using Submodular Function-Based Framework, Sentence Compression and Merging"*, 8th International Joint Conference on Natural Language Processing (**IJCNLP 2017**). [🔗 Link]

[24] **Mir Tafseer Nayeem** and Yllias Chali, *"Extract with Order for Coherent Multi-Document Summarization"*, TextGraphs-11: The Workshop on Graph-based Methods for Natural Language Processing (**ACL 2017 Workshop**). [🔗 Link]

👨‍💻 Activities

📄 Services

👨‍💼 Area Chair (AC)
- **2026:** ACL, EACL, ARR
- **2025:** ACL, EMNLP, ARR

📝 Reviewer
- **2025:** NAACL
- **2024:** ACL, EMNLP, NAACL, EACL, SIGIR, WSDM
- **2023:** SIGIR, KDD, WSDM, CIKM
- **2022:** SIGIR, KDD, CIKM, ICDE

🏅 Outstanding Reviews
High-quality reviews recognized by *ACL venues as **"Great Reviews"**
- **1/3 reviews** recognized at NAACL 2025
- **2/4 reviews** recognized at ACL 2024

👨‍🏫 Invited Talks
- Talk on KidLM at Honors Seminar, University of Alberta (UofA), AB, Canada. [2024]
- Talk on Introduction to NLTK at University of Lethbridge (UofL), AB, Canada. [2017]
- Guest Lecturer in CPSC 3750 - Artificial Intelligence course, UofL, AB, Canada. [2017]

👨‍💼 Leadership
- Lab Manager, Data & Language Intelligence (DaLI) Lab, University of Alberta. [2022 – *Current*]
- Founding Admin of a Discord server comprising approximately 4,500+ members dedicated to assisting prospective students with grad school applications. [2021 - *Current*]
- Member, Logistics & IT Support Committee, ICCIT 2020 Conference.
- Organizing committee, Commonwealth of Learning (COL) TEL AUST Project [report].

- Organizer and Host, Intel Nervana AI Academy - Workshop.
- Convener, Project Showcasing, AUST CSE WEEK.
- Organizing committee member, AUST CSE Fest; BUBT CSE Fiesta; IUT ICT Fest.

💼 Experience

Graduate Teaching Assistant   **Sept 2021 – *Current***
University of Alberta, Computing Science Department   Edmonton, Canada
  Courses TAed:
    - CMPUT 101 - Introduction to Computing
    - CMPUT 291 - Introduction to File and Database Management

Faculty Member, CSE Department   **Oct 2018 – Jul 2021**
Ahsanullah University of Science and Technology (AUST)   Dhaka, Bangladesh
  Courses Taught:
    - CSE 4237 & 4238 - Soft Computing with Lab
    - CSE 4129 & CSE 4130 - Formal Languages & Compilers with Lab
    - CSE 3223 & 3224 - Information System Design and Software Engineering with Lab
    - CSE 3104 - Database Lab
    - CSE 1108 - Introduction to Computer Systems

Graduate Teaching Assistant   **Sept 2015 – Sept 2017**
University of Lethbridge, Computer Science Department   Lethbridge, Canada
  Courses TAed:
    - CPSC 1620 & 2620 - Fundamental of Programming I & II
    - CPSC 2720 - Practical Software Development

📒 References

Available on request.

[CV updated on February 6, 2026]