

Enamul Hassan

PhD Candidate, Department of Computer Science • Stony Brook University, Stony Brook, NY 11794

ehassan@cs.stonybrook.edu • enamcse.github.io • Google Scholar • ORCID 0000-0001-8020-5409

github.com/enamcse • linkedin.com/in/enamcse

RESEARCH INTERESTS

Foundation models for sequential data • life-course and health (EHR) trajectories • representation learning • computational social science • natural language processing

EDUCATION

Stony Brook University, New York, USA — Ph.D., Computer Science Aug 2024 – May 2029 (expected)

Advisor: Steven Skiena • advanced to candidacy Aug 2026 (Research Proficiency Exam) • GPA 3.82/4.00

University of North Texas, Texas, USA — Ph.D., Information Science (Data Science) Aug 2023 – May 2024

Transferred to Stony Brook University • GPA 4.00/4.00

Shahjalal University of Science and Technology, Sylhet, Bangladesh — B.Sc. (Engineering), Computer Science and Engineering 2012 – 2017

CGPA 3.79/4.00 • 1st position in a class of 63 • requirements completed Oct 2016, result published Mar 2017

PUBLICATIONS

Manuscripts under review

- E. Hassan, T. Pial, F. Hafner, A. Macanovic, L. Sage, T. Emery, A. van de Rijt, and S. Skiena. “Foundation Models for Life-Course Prediction on Dutch Social Registry Data”. Under review. 2026.
- T. Pial, A. Macanovic, L. Sage, F. Hafner, E. Hassan, T. Emery, A. van de Rijt, and S. Skiena. “Fertility Prediction with Foundation Models Trained on Population Registry Data”. Under review. 2026.
- E. Sivak et al. “Births are difficult to predict even with rich survey and full-population register data”. Under submission (PreFer collaboration; E. Hassan among the authors). 2026.

Preprints

- G. Hossain, M. Shin, B. Whaley, M. Afrose, and E. Hassan. “Assessing the Social Validity of Canvas LMS-Based Cybersecurity Online Modules in High School STEM Classes”. In: *Research Square* (2026). Preprint, posted 27 July 2026. DOI: [10.21203/rs.3.rs-10107954/v1](https://doi.org/10.21203/rs.3.rs-10107954/v1).

Peer-reviewed conference proceedings

- T. Pial, F. Hafner, D. Handzlik, E. Hassan, L. Sage, A. Macanovic, T. Emery, A. van de Rijt, and S. Skiena. “Embeddings of Nation-Level Social Networks”. In: *Complex Networks & Their Applications XIV (Complex Networks 2025)*. Vol. 1266. Studies in Computational Intelligence. Springer, 2025. DOI: [10.1007/978-3-032-16719-4_27](https://doi.org/10.1007/978-3-032-16719-4_27). arXiv: [2603.29059](https://arxiv.org/abs/2603.29059).
- E. Hassan, G. Hossain, and P. Vajpayee. “Reducing Cyber Risks in IoT Device Resource Management: A Probabilistic Banker’s Algorithm Approach”. In: *2024 IEEE World AI IoT Congress (AIIoT)*. 2024.
- H. Noreen, M. Deb, and E. Hassan. “Sahacharza: Fostering Women’s Wellness via AI-Enhanced Bengali Conversations”. In: *26th Int. Conf. on Computer and Information Technology (ICIT)*. Cox’s Bazar, Bangladesh, 2023.
- R. Raihan, M. E. Sazin, and E. Hassan. “Exposition of the Issues, Difficulties, and Proposed Solutions in Bangla Text Summarization”. In: *7th International Conference on Engineering Research, Innovation and Education (ICERIE)*. Sylhet, Bangladesh, 2023.
- R. H. Rinkon, U. Habiba, and E. Hassan. “Recognizing Basic, Compound and Numerical Bangla Handwritten Characters with Convolutional Neural Network”. In: *7th International Conference on Engineering Research, Innovation and Education (ICERIE)*. Sylhet, Bangladesh, 2023.
- K. I. Islam, T. Yuvraz, M. S. Islam, and E. Hassan. “EmoNoBa: A Dataset for Analyzing Fine-Grained Emotions on Noisy Bangla Texts”. In: *Proceedings of ACL-IJCNLP*. 2022, pp. 128–134.
- N. M. Jakilim, S. Mahamudul Hasan, and E. Hassan. “A Benchmark of Machine Learning and Deep Learning Algorithms for Detecting Fake News in Bangla Language”. In: *2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI)*. 2022, pp. 1–6. DOI: [10.1109/STI56238.2022.10103235](https://doi.org/10.1109/STI56238.2022.10103235).

Journal articles

- T. R. Ornob, G. Roy, and E. Hassan. “CovidExpert: A Triplet Siamese Neural Network Framework for the Detection of COVID-19”. In: *Informatics in Medicine Unlocked* 37 (2023), p. 101156.
- M. Billah and E. Hassan. “Depression Detection from Bangla Facebook Status Using Machine Learning Approach”. In: *International Journal of Computer Applications* 178.43 (2019).

RESEARCH EXPERIENCE

Graduate Assistant, Department of Computer Science, Stony Brook University Aug 2026 – present
Advisor: Steven Skiena

- Extending life-sequence foundation models to longitudinal electronic health records, modelling patient trajectories over diagnoses, laboratory results and medications.

Research Project Assistant, Department of Computer Science, Stony Brook University May 2025 – Aug 2026
Advisor: Steven Skiena • on site at Statistics Netherlands (Summer 2025) and in the United Kingdom (Aug 2025 – Jan 2026)

- Pretrained transformers from 8M to 540M parameters on the full Dutch population register — roughly 23M people and 7.5B records — inside a secure government environment under output disclosure control.
- Extended the life2vec input scheme for records in which many events share one date: Time2Vec encodings of age and calendar time, same-date segment embeddings, and ReZero residuals.
- Ran controlled ablations over data class, model scale, pretraining objective, pooling and usage regime. Masked pre-training beat autoregressive on 17 of 18 tasks; richer data mattered more than model size, moving from 3 to 13 task wins at a fixed 80M parameters.
- Evaluated 24 prediction targets across seven outcome families; the models beat tabular baselines on 20 of 21 comparable tasks (fertility AUC 0.88, partnership 0.80, taxable-income R^2 0.49 $\rightarrow$ 0.63 with fine-tuning).
- Contributed the cluster-token method to the group’s population-network paper: a whitened Fibonacci-sphere quantizer that makes cluster identifiers comparable across years, which feeds a 1.4B-edge, five-layer network of 17M people into the life sequences.
- First author of the registry-scale manuscript; co-author of the network and fertility papers.

Research Assistant, SCORE Lab, Department of Information Science, University of North Texas Aug 2023 – Jul 2024

- Coordinated PERFECT, an Office of Naval Research funded cybersecurity education project run by UNT and West Texas A&M University, and developed Canvas modules for its Basic and Intermediate courses.
- First-authored a probabilistic adaptation of the Banker’s algorithm for IoT resource management under cyber risk (IEEE AIIoT 2024).

ACADEMIC APPOINTMENTS

Assistant Professor, Department of CSE, Shahjalal University of Science and Technology Sep 2019 – present
On study leave since Aug 2023

Lecturer, Department of CSE, Shahjalal University of Science and Technology Sep 2017 – Sep 2019

Adjunct Faculty (part-time), Department of CSE, Sylhet Engineering College 2017 – 2022

IT Consultant (part-time), Center for Research, Testing and Consultancy, SUST 2019 – 2023

Requirements analysis, architecture and database design, and procurement planning for the automation of Osmani Medical College and a provident-fund system for tea-estate workers; validator on a government emotion-annotation project with Gigatech Ltd.

Instructor, Cisco Networking Academy, SUST — CCNA Routing and Switching v6.0, v7.0 2017 – 2023

Instructor, AWS Academy, SUST — Cloud Computing Fundamentals 2022 – 2023

RESEARCH GRANTS

2022 – 23	Development of a Large Corpus of Spontaneous Speech for Improved Speech Recognition in Bangladeshi Bangla. Co-Investigator • ~\$4,300 • Project AS/2022/1/24
2021 – 22	BATS: A Dataset for Bangla Automatic Text Summarization. Principal Investigator • ~\$4,000 • Project AS/2021/1/19 • ~100k document–summary pairs
2020 – 21	BED: A Bangla Emotion Dataset. Principal Investigator • ~\$2,300 • Project AS/2020/1/27 • ~50k sentences tagged with six basic emotions
2019 – 20	Automated Assessment System for MCQ and Subjective Tests. Principal Investigator • ~\$2,000 • Project AS/2019/1/20
2018 – 19	SUST Online Judge: A Programming Platform for All Students. Co-Investigator • ~\$2,000 • Project AS/2018/1/11

HONORS AND AWARDS

2024 – 26	Chairman's Fellowship, Department of Computer Science, Stony Brook University
2023	1st place, BUET CSE Fest Inter-University Programming Contest — as coach
2022	50th place, 45th ACM-ICPC World Finals, Dhaka — as coach
2021	4th place (Regional) and 2nd place (Regional Preliminary), ACM-ICPC Dhaka — as coach
2019	43rd ACM-ICPC World Finals, Porto — as co-coach
2014 – 15	ACM-ICPC Dhaka Regional, 12th (2014) and 14th (2015); SUBIUPC 8th (2015) — as contestant
2012 – 17	University Scholarship, SUST, awarded on annual results throughout the undergraduate degree
2017	1st position in a class of 63, B.Sc. in CSE, SUST

SELECTED PROJECTS AND SYSTEMS

2017 – 19	MBBS and BDS admission test result processing. Result-processing system for the national medical and dental college entrance tests, covering roughly 60 colleges and 100,000 candidates. Team led by Prof. Muhammed Zafar Iqbal.
2020 – 21	SUST Exam Management System. Online examination platform supporting multiple-choice, matching, coding and descriptive questions with program-based grading and basic proctoring. Supervised across three student cohorts.
2020 – 21	SUST Class Routine Management. Scheduling system that allocates scarce classrooms while respecting instructor preferences. Supervised as a course project.
2016	NanoMapper. Read-mapping tool for long, noisy Oxford Nanopore reads using gapped minimizers and FM-indexing. Undergraduate thesis, supervised by Md. Eamin Rahman.

TEACHING

Twenty-five courses across 68 offerings at SUST, Sylhet Engineering College and Sylhet International University, 2017–2023.

Programming and data structures	Structured Programming Language (6 offerings) • Structured Programming Language Lab (5) • Data Structures (4) • Data Structures Lab (4) • Introduction to Computer Language (5) • Introduction to Computer Language Lab (7) • Introduction to Programming with Python Lab • Introduction to Computing Applications (2)
Databases	Database Management System (3) • Database Management System Lab (2) • Database Management and Programming (2) • Database Management and Programming Lab (3) • Advanced Database System (2) • Advanced Database System Lab (2)
AI and machine learning	Artificial Intelligence (4) • Artificial Intelligence Lab (4) • Machine Learning • Machine Learning Lab
Web and cloud	Web Engineering (2) • Web Engineering Lab (2) • Web Technologies Lab • Cloud Computing • Cloud Computing Lab
Competitive programming	Introduction to Competitive Programming
Projects and thesis	Project Work II (2) • Project and Thesis I/II — supervisor, report evaluator and committee member
Departments served	CSE, Mathematics, Genetic Engineering and Biotechnology, Electrical and Electronic Engineering, Statistics, Business Administration, Chemistry, Economics, Social Work, Software Engineering, and the Institute of Information and Communication Technology

Teaching Assistant, CSE 220 System Fundamentals I, Stony Brook University

Fall 2024, Spring 2025

MENTORING AND SUPERVISION

2017 – 23	Supervised 45 final-year undergraduate thesis and project students (29 at SUST, 16 at Sylhet Engineering College).
2018 – 23	Evaluated approximately 80 final-year thesis and project reports across natural language processing, image processing, deep learning, blockchain and machine learning; served on final-year committees for five cohorts.
2020 – 23	Lead Coach, Competitive Programming Wing, Department of CSE, SUST.

SERVICE

Reviewing	ICCIT 2024 • ICERIE 2019, 2021 • ICBSLP 2019 • Journal of Information & Knowledge Management
Department and university	Chairman (2020–23) and member (2018–23), examination committees • Undergraduate Coordinator, CSE (2020–22) • Curriculum Committee, CSE (2020–23) • Technical Sub-committee, Admission Test (2017–23) • Examiner, technical staff recruitment (2022–23) • Member, Total SUST, a project to make administrative processing paperless (2022–23) • Member, Natural Language Processing Lab, SUST
Community	Problem setter and judge, SUST Intra-University Programming Contest (2014–present) • Advisor, SUST Science Arena (2021–24) • Advisor, RoboSUST (2023–24) • Organisational administrator for Coursera and edX programmes during the pandemic (2020–23) • Contributor, Stack Overflow (reputation 5.3k+)

TECHNICAL SKILLS

Machine learning	PyTorch, Hugging Face Transformers, scikit-learn, NumPy, pandas, NLTK
Languages	Python, C/C++, Java, JavaScript, SQL, Bash, MATLAB, $\text{\LaTeX}$
Data and systems	PostgreSQL, MySQL, SQLite, MongoDB, Spark, Hadoop, Linux, Git, Docker, AWS, GCP
Research environments	Statistics Netherlands secure microdata environment, output disclosure control

REFERENCES

Available on request.