

Postgraduate Seminar Series

Venue: Graduate Seminar Room

Date & Time: September 5, 2026 at 2:30 PM

Speaker Information

Anika Bintee Aftab (Std No. 0422052054) is a part-time M.Sc. student in the department of CSE, BUET. She completed her undergraduate studies from Ahsanullah University of Science and Technology (AUST) in 2022. She serves as a faculty member in the department of CSE at Ahsanullah University of Science and Technology (AUST), Bangladesh. Her research interest lies in the field(s) of Health Informatics, Machine Learning, Deep Learning and Large Language Models. She is currently doing her postgraduate thesis under the supervision of Dr. Tanzima Hashem. She will be speaking about her ongoing research in this talk.



An Ontology-Driven LLM-Based Framework for Detecting Prenatal Mental Distress Triggering Events from Social Media Posts

Prenatal mental distress critically impacts maternal and fetal well-being. While expectant mothers frequently share psychosocial struggles on platforms like Reddit, existing computational approaches primarily treat distress analysis as a basic classification task, overlooking underlying root triggers. Moreover, relying on Large Language Models (LLMs) for end-to-end extraction and reasoning over *triggering events* introduces analytical hallucinations—manifesting as inaccurate frequency counts, arbitrary rankings, and missing occurrences during dataset-level aggregation, while lacking verifiable traceability. Additionally, existing medical ontologies remain predominantly clinical, failing to capture informal social media narratives and anticipatory maternal fears.

To address these limitations, we propose an ontology-driven LLM-based framework for extracting explainable prenatal mental distress triggering events from social media. Our multi-stage pipeline extends beyond simple classification to generate structured epidemiological insights. First, we construct a custom ontology formalizing pregnancy-specific socio-environmental, interpersonal, and anticipatory stressors.

Using advanced LLMs guided by this ontology, the framework extracts (Subject → Relation → Object) factual triplets from raw posts, compressing the raw token payload by 29.07% and abstracting them into high-level canonical triggering events.

A dual-evaluation strategy, comprising human experts and an LLM-as-a-Judge, rigorously validates our framework. Triplet extraction achieves over 97.0% Adjacent Agreement and a Quadratic Weighted Kappa (QWK) of 0.66 against human ground truth, confirming substantial reliability. Furthermore, against a human-validated benchmark with 97.3% consensus Semantic F1-Score, our triggering event extraction achieves high semantic coverage (0.956) and superior abstraction fidelity (0.682 Smart Coverage), significantly outperforming direct LLM baselines that suffer from abstraction failures.

Ultimately, by decoupling semantic extraction from analytical computation, our framework transforms the social media corpus into a reusable, queryable knowledge base. Using deterministic Python querying with zero additional LLM inference, the system generates hallucination-free epidemiological insights, such as trimester-specific anxiety shifts and symptom-trigger correlations. To demonstrate generalizability, we apply the pipeline to distinct demographic subgroups (e.g., pregnancy over 35/40, working mothers) to explore how stressors shift across societal contexts. This *extract once, analyze infinitely* paradigm ensures traceability, establishing a scalable foundation for maternal mental health surveillance.