Georgia Tech College of Sciences
School of Mathematics

CRIDC 2025

GITHUB

FEEDBACK FUSION
NLP-DRIVEN STUDENT SURVEY ANALYSIS IN LARGE COURSES

Ayushi Chakrabarty (achakrabarty8@gatech.edu), Greg Mayer (greg.mayer@gatech.edu)
Georgia Institute of Technology

Introduction

Rationale:

- Large-scale MOOCs (**1000–2000+ students**) rely on start-of semester and mid-semester surveys to understand student needs.[1,2]
- Open-ended responses offer valuable qualitative insights but are **challenging to analyze manually** due to their volume leading to **delayed response times** and potentially **overlooked student concerns**.

Hypothesis:

- **Natural Language Processing (NLP)** can automate the analysis of open-ended survey responses by offering **context-aware insights** and **significantly reducing response time.**

Justification:

- Limitations in existing solutions:
  - Manual review is **time-intensive** and **lacks scalability.**[3]
  - Keyword-based search is prone to **missing nuanced insights.**[4]
  - Sentiment analysis is **limited in understanding context** and meaning.[5]

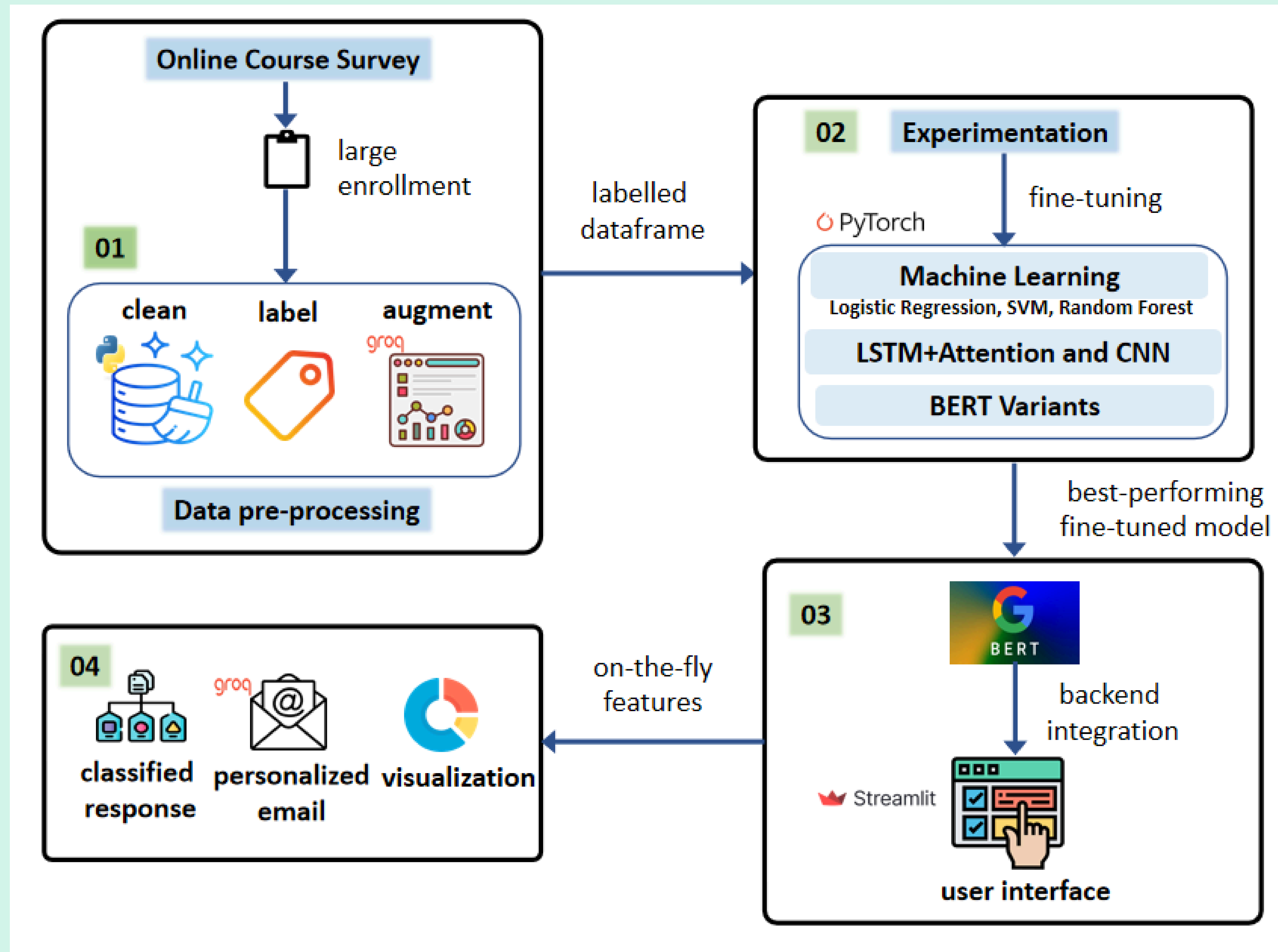
The growing demand in the **EdTech industry** for scalable and effective tools to enhance the learning experience in large-scale MOOCs makes this project highly relevant.

Data Categories

- Utilizes **start/mid semester survey** data collected from large-scale MOOCs.
- The dataset contains **8000 samples** in total with almost equivalent data samples for each category. These were manually labelled into four categories.

Code	Description	Example
AC	Academic Concern	The exams since they are given such a large weight.
PC	Personal Concern	Could you reschedule the exam as I am sick?
TC	Technical Concern	Technological issues with the gatech instructure portal.
NC	No Concerns	There is nothing in particular!

Design Flowchart



Methods

Modeling approach:

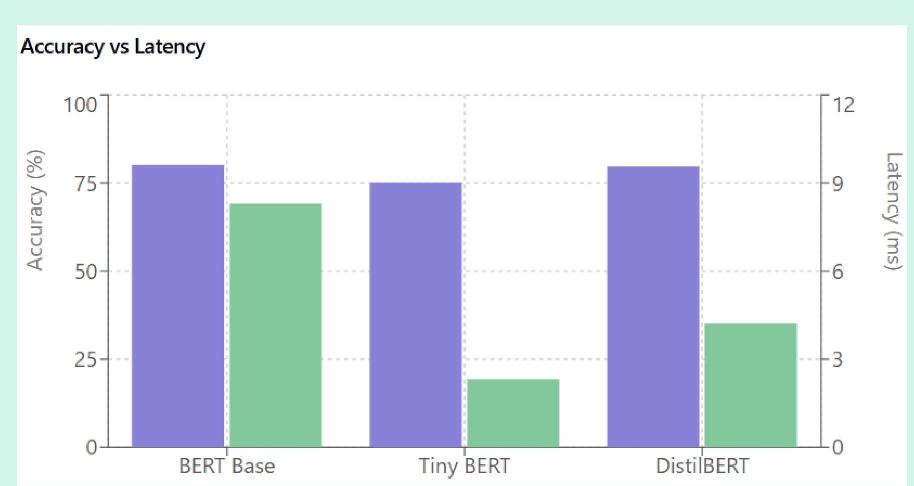
- Three **BERT-based** model variants fine-tuned on survey feedback data to allow for flexibility in choosing a model based on the trade-off between **computational efficiency** and **classification performance.**

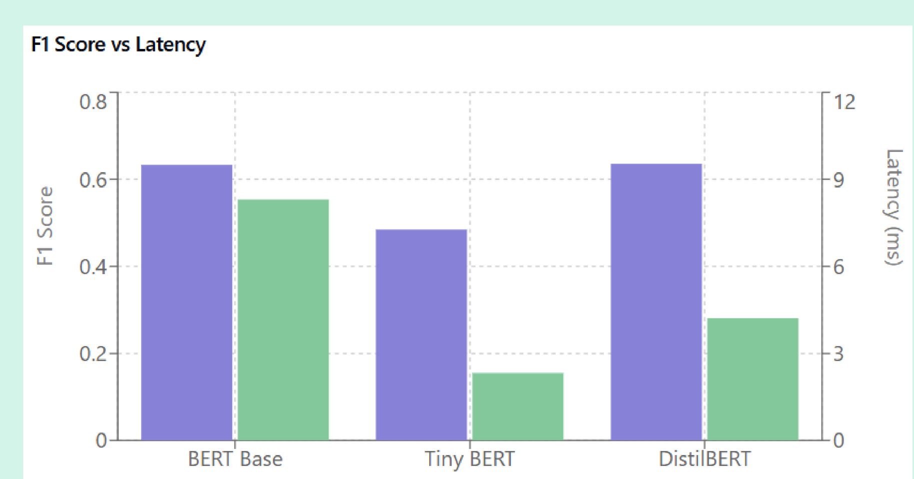
System Design:

- **Backend:** Python-based NLP pipeline using transformers & scikit-learn.
- **Frontend:** Streamlit-powered UI for easy instructor interaction to allow intructors to download categorized responses, highlight sensitive personal concerns, and visualize the distribution of concerns across categories.

Evaluation:

- Assess model performance using classification **accuracy, F1-score** and **latency**.





Results

This system demonstrates **first-of-its-kind LLM-powered classification system that complies with FERPA** empowering instructors with intelligent analysis of student feedback while advancing UN SDG 4 (Quality Education).

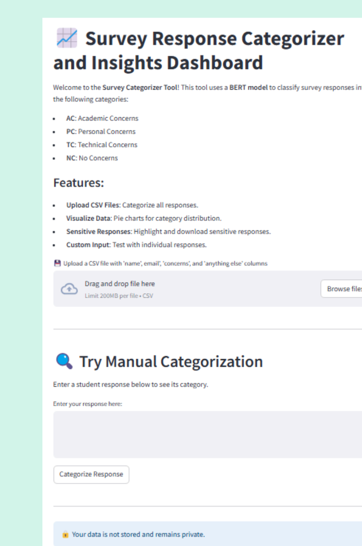
Key Findings:

- BERT Base and DistilBERT had similar **80% accuracy**, but DistilBERT was **2x faster** at 4.2ms latency.
- Tiny BERT had lower 76% accuracy but **fastest 2ms latency**.

Scientific Rigor & Congruence:

- Assessed on consistent performance metrics and validated across **multiple iterations** to confirm stability and reliability.
- Results align with expected **performance-efficiency tradeoffs** in BERT model compression.
- **Class imbalance** impacts mirror challenges in educational data mining literature.

Model	Avg. Accuracy (%)	Avg. Latency (ms)
BERT Base	80.13	8.30
DistilBERT	79.67	4.22
TinyBERT	75.12	2.32



Limitations & Future Directions

- The model struggles with minority classes as constrained by the original dataset diversity where augmentation **replicates limited patterns**.
- Expand to **diverse MOOC data** for generalization.
- **Real-world deployment** validation.
- **Multilingual** support.

References

[1] Richardson, J. T. (2005). Instruments for obtaining student feedback: A review of the literature. Assessment & evaluation in higher education, 30(4), 387-415.
[2] Angelo, T. A., & Cross, K. P. (2012). Classroom assessment techniques. Jossey Bass Wiley.
[3] Morgan, M., Nylén, A., Butler, M., Eckerdal, A., Thota, N., & Kinnunen, P. (2017, November). Examining manual and semi-automated methods of analysing MOOC data for computing education. In Proceedings of the 17th Koli Calling International Conference on Computing Education Research (pp. 153-157).
[4] Jiang, Z., Miao, C., & Li, X. (2017). Application of keyword extraction on MOOC resources. International Journal of Crowd Science, 1(1), 48-70.
[5] Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review, 55(7), 5731-5780.