COLLEGE OF
INNOVATION & TECHNOLOGY

CSC-530

Advanced Information Storage & Retrieval

Phase 2 and 3 Project- Winter 25

Deadline: 02/12/2025@ 23:59

[Total Mark for each phase is 100]

Student Details:

Name: TARAKA RAM DONEPUDI	ID: 80678315
Name: LOCHAN KRISHNA DASARI	ID1: 92690298
Name: AMEEN VATHIMATTOM ASHRAF	ID2: 27612818
CRN: **25869**	

Submission Instructions:

- This assignment must be submitted on canvas **pdf format only**).
- The page number should appear on each page.
- All the text and diagrams must be electronically produced.
- Email submission will not be accepted.
- You are advised to make your work clear and well-presented; marks may be reduced for poor presentation. This includes filling your information on the cover page.
- You MUST show all your work, and text must not be converted into an image, unless specified otherwise by the question.
- The project proposal should contain the following sections:

Phase 1

1. **Phase 2 Intellectual Merit**

(the value and quality of the proposed work in terms of advancing knowledge and understanding within its field or across multiple disciplines)

Foundation for Broader Applications: Highlights the project's foundational

This sentiment analysis project on Amazon reviews serves as a foundational framework for a wide range of applications beyond e-commerce. By developing a robust NLP-based sentiment classification system, the methodologies and models used in this project can be adapted to various domains, including:

- **Customer Feedback Analysis:** Applied to service-based industries such as hospitality, healthcare, and banking to analyze customer sentiment and improve service quality.
- **Brand Reputation Management:** Helps businesses monitor and respond to public perception on social media and review platforms.
- **Market Research and Trend Analysis:** Enables companies to track consumer trends and preferences for strategic decision-making.
- **Fake Review Detection:** Contributes to identifying spam or fraudulent reviews, ensuring a more reliable e-commerce ecosystem.

Key Benefit:

The primary benefit of this project is **enhanced decision-making for businesses and consumers** through accurate sentiment classification of Amazon product reviews. By automatically analyzing thousands of reviews, businesses can:

- Identify key strengths and weaknesses of their products.
- Improve product development based on customer feedback.
- Optimize marketing strategies by targeting customer sentiments.
- Detect negative sentiment spikes early and address customer concerns proactively.

For consumers, sentiment analysis provides a **data-driven approach to making informed purchasing decisions** by summarizing product sentiment trends and filtering out misleading or fake reviews.

How the project can contribute to Explainable AI:

Explainable AI (XAI) aims to make machine learning models transparent, interpretable, and trustworthy. This project contributes to XAI in the following ways:

1. **Feature Importance Analysis:** By using models like logistic regression, decision trees, or attention-based deep learning approaches, the project can highlight key words and phrases that contribute to positive, negative, or neutral sentiment classifications.
2. **Bias Detection and Mitigation:** Ensuring that sentiment analysis models do not exhibit biases toward specific product categories, brands, or user demographics.
3. **Trust and Transparency:** Enabling users (both businesses and consumers) to understand **why** a review was classified as positive, negative, or neutral, increasing trust in AI-driven decision-making.
4. **Rule-Based and Hybrid Approaches:** Combining traditional lexicon-based sentiment analysis with deep learning models to enhance interpretability while maintaining high accuracy.

2.Phase 3 Broader Impacts

(refers to the potential of a project or research to benefit society, contribute to desired societal outcomes, and positively influence fields beyond its immediate scope)

Main Contribution:

The main contribution of this project lies in its ability to enhance **consumer trust and business decision-making** through automated sentiment analysis. By leveraging AI and NLP techniques, the project contributes to:

- **Improved Consumer Experience:** Helping customers make informed purchasing decisions based on real-time sentiment insights.
- **Business Growth and Product Innovation:** Enabling businesses to refine their products and marketing strategies based on customer sentiment trends.
- **Advancements in AI and NLP:** Developing scalable, explainable, and adaptable sentiment analysis models that can be extended to other domains like healthcare, finance, and political sentiment analysis.

Dissemination Plan:

(outlines how the outcomes, findings, or results of a project will be shared with relevant audiences to maximize its impact and ensure broad accessibility)

Sharing code, datasets, and pre-trained models on platforms like GitHub, Kaggle, and Hugging Face.

Providing detailed documentation and tutorials to make sentiment analysis techniques accessible to developers and researchers.

Broader Impact on STEM Education:

Encouraging AI and NLP Learning:

- Providing datasets and tutorials to help students and researchers experiment with sentiment analysis models.
- Creating Jupyter Notebooks and Google Colab projects for practical engagement.

Promoting Ethical AI Awareness:

- Teaching students about bias mitigation and explainability in AI models.
- Discussing real-world implications of sentiment analysis, such as misinformation detection and brand reputation management.