CSCI433/933
Machine learning algorithms and applications
Assignment 2 - Intelligent Shakespeare Chatbot

Professor Philip O. Ogunbona

Due dates and mark distribution
Part 1 (30%) - Week 9
Part 2 (70%) - Week 12

Assignment Overview

This assignment challenges students to explore and implement a lightweight, domain-specific intelligent chatbot capable of engaging in dialogue about Shakespearean literature. The chatbot should be able to answer factual and interpretive questions about Shakespeare's plays, generate relevant quotes, and provide summaries of specific scenes. The project is divided into two major components: research-driven prompt engineering and full system design and implementation.

Learning Objectives

- Develop an understanding of how domain-specific chatbots can be created using compact language models.
- Gain experience in prompt engineering and guided exploration using large language models (LLMs) like ChatGPT.
- Build practical skills in the design, training, and deployment of a task-oriented conversational AI system.
- Integrate theoretical insights from natural language processing and dialogue systems with implementation.
- Practice referencing and cross-validating AI-generated knowledge with academic and technical literature.

Part One: Prompt Engineering and Technical Exploration (30%)

Students are required to engage with ChatGPT (or equivalent LLMs) to explore the technical and design considerations of developing a Shakespeare-specific chatbot. Students must:

1. Formulate and refine intelligent prompts to investigate:
    - How a domain-specific chatbot can be built using a small language model.
    - Suitable model architectures (e.g., distilled transformers, RNNs, quantized LLMs).
    - Data preparation and pretraining/finetuning on Shakespearean text.
    - How to generate contextual quotes and perform scene-based summarization.
    - Tools, frameworks, and constraints for deployment (e.g., HuggingFace, LangChain, RAG, etc.).
2. Cross-validate insights from ChatGPT with peer-reviewed papers and tool documentation (Please do not reference blogs or Medium-type of articles.).
3. Document their findings in a structured report of 1500–2000 words, clearly outlining:
    - Prompts used and corresponding responses.
    - Critical assessment of the validity of ChatGPT outputs.
    - Bibliographic references (peer-reviewed) supporting or contesting ChatGPT's responses.
    - A proposed system design architecture for Part Two.

Deliverable

A well-organized report in PDF format. Proper citations using style in accompanying template are required.

Part Two: Chatbot Implementation and Demonstration (70%)

Using insights gathered in Part One, students are to implement a functional prototype of an intelligent Shakespeare chatbot with the following capabilities:

- Engage in basic multi-turn dialogue about Shakespearean plays.
- Answer factual and thematic questions about characters, events, and scenes.
- Generate short quotes or passages from Shakespeare's corpus.
- Summarize a selected scene from a specified play in prose.

Minimum Requirements

- Use of a compact/custom-trained or distilled model (not a full LLM).
- Inclusion of a retrieval-based component or prompt templating if training is infeasible.
- A user interface (CLI, web, or chatbot framework such as Streamlit or Gradio).
- Clear documentation, including setup instructions and usage examples.

Deliverables

1. Fully documented codebase submitted via subject Moodle site.
2. Demonstration video (5–7 minutes) explaining and showcasing functionality.
3. Brief technical report (max 1000 words) explaining design choices, challenges encountered, and future improvements.

Assessment Criteria

- **Part One:** Depth and relevance of prompt exploration; clarity and coherence of report; use of references.
- **Part Two:** System functionality; innovation and creativity; usability; technical accuracy; documentation quality.