**Sofia University “St. Kliment Ohridski”**  
Faculty of Mathematics and Informatics  
Specialty: Information Systems  
Course: Prompt Engineering

**Coursework Proposal**  
**Student Name:** Kaloyan Mehandzhiyski  
**Faculty Number:** 5MI0700025

**Project Title:**

AI-Powered Automated Grader for Short Answers

**Assignment Description:**

The goal of this project is to implement a prompt-engineered solution for automated grading of student short-answer responses. The system will compare a student’s answer to a model answer and grade it on a numeric scale, also providing short, constructive feedback. This will support teachers in evaluating open-ended responses more efficiently and fairly.

* **Input:**
  + Student answer (free-text)
  + Model answer (reference)
  + Grading rubric (criteria-based)
* **Output:**
  + Numeric score (e.g., 0 to 10)
  + Short feedback on correctness and clarity

**Prompt Design:**

* **Few-shot prompting** will be used to provide examples of graded responses and rubric interpretation.
* **Meta prompting** will instruct the model to critique and score the answer based on explicit criteria.
* **Self-consistency** technique will be applied to generate multiple scoring attempts and average them for a more balanced evaluation.

**Example Q&As (Source Prompts):**

**Prompt:** "Grade this answer on a scale of 0–5 based on accuracy and explanation clarity. Provide a reason."

* **Question:** "Explain Newton's 2nd law."
* **Student Answer:** "The force is equal to the change in motion."
* **Output:** "Score: 3/5. Reason: The concept is mostly correct but lacks mention of mass, acceleration, and vector notation."

**Model Evaluation:**

* **Dataset:** A small corpus of 20 short questions in Programming on C++, each of it with 5 different answers.

**(Test with 5 questions, then learn from the other 15 and then test again with the same 5 to have results BEFORE and AFTER the learning to compare)**

* **Evaluation Metrics:**
  + Scores from 0 to 10 for each question where:

|  |  |
| --- | --- |
| **Grade** | **Meaning** |
| 0 | Completely wrong / nothing right |
| 1 | Very bad answer |
| 2–3 | Poor answer |
| 4 | Bad answer |
| 5 | Medium / partial understanding |
| 6-7 | Good answer |
| 8–9 | Great answer |
| 10 | Absolutely correct |

* + Final result (average grade)
* **Challenges:** Handling ambiguous or partially correct answers; avoiding over-rewarding verbose but inaccurate answers

**Ethical and Trustworthiness Evaluation:**

* The system will include transparent feedback for every grade issued.
* Confidence scores will be provided to highlight uncertain grading outcomes.
* Teachers will be encouraged to use the tool for support, not full automation.
* Bias toward writing style or length will be minimized with rubric anchoring.

**Submission Date:** [09.04.2025]  
**Instructor:** Assoc. Prof. Ioannis Patias