

WEEK 1

Web Application:

- that allows people sign using google authenticator
- that allows to create a new assessment
- allows the person to gives the links to the google spreadsheet:
 - student ID + email
 - questions, rubrics & criteria

WEEK 2

LLMs:

- LLMs installed exposed by/via a some kind of REST API
- 50 questions experimented with 20 different prompts

WEEK 3

Evaluation:

- Previous week (dummy marks) ,Now put the mark using a REST call to LLM to give you the mark
- assess numerically how well is your LLMs doing the mark
- 50 questions -> corresponding rubrics, marks, sample answers
- FIND OUT : How many times is the mark given by LLM matching the mark in your given marks.

-Table:

Questions	Rubrics	Answers	Given Marks	LLM marks	Matched (%)	Not Matched (%)

Code+integration

REPORT:

- explore ways of improving % of matches
- write down the prompt templates of CoT used for these questions
 - explanation
 - matched % record
 - table

50 questions -> average 3 answers for each question -> 150 marking -> check all of your prompts

