ML take-home challenge

Context

Automated short answer grading (ASAG) using computational techniques is a key component of online tutoring platforms, enabling instant, granular feedback to students within interactive assessments.

A core challenge for building robust ASAG systems is the lack of available domain-specific data and the difficulty of grading open-ended questions. State-of-the-art solutions are based on recent progress in Natural Language Processing and Deep Learning.

Task

Build, train and evaluate a ML model for ASAG to achieve:

- High grading accuracy
- Robustness to linguistic challenges common to NLP tasks

Data

Use Atomi's public version of the open-source SemEval 2013 Task 7 dataset to train and evaluate ML models for ASAG:

🤗 Atomi/semeval_2013_task_7_beetle_5way · Datasets at Hugging Face

The dataset contains ~12,000 examples of question, reference answer and student answer triples graded by domain experts.

Requirements

- You should work individually
- You should cap development time to ~4 hours and aim for a reasonable baseline solution that can be delivered in that time
- You should use Python and any ML tools/models that you think is relevant for the task
- You should build a low cost solution and avoid/minimise any experimentation cost

Deliverables

A Github repository with:

- A `README.md` file summarising your findings with the following presentation structure:
  - Objectives
  - Methodology, including:
    - ML approaches/tasks considered
    - Evaluation metrics chosen
    - Etc.
  - Results, including:
    - Data preprocessing details
    - Training details
    - High level results
    - Error analysis
    - Etc.
  - Concluding remarks, including:

    - Strengths and weaknesses of chosen solution
    - Opportunities with more time/resources/budget
- Experimentation source code as Jupyter notebooks or Python scripts

Evaluation criteria

- **Problem Solving:** Ability to select an appropriate ML approach/task for solving the problem
- **Experimentation Practices**: Ability to design and execute experiments, analyse results and iterate, following best ML practices
- **Technical Competency:** Quality of the machine learning models and code
- **Agility**: Ability to deliver a reasonably good baseline solution with the suggested time constraint
- **Presentation Skills:** Clarity and thoroughness in communicating your methodology and insights.

Submission

Please reach out if you need any clarification about the task.

You have ~1 week to complete the challenge, although this can be extended to adapt to your schedules.

Once done, notify the recruiter with the link to your Github repository.