Assignment 2: Practical Machine Learning Project

31005|32513 Machine Learning Spring 2019

TASK

The goal of this assignment is to develop your skills in building data models. You can do a machine learning project using real-world data and addressing a practical problem.

For students in 31005, you can alternatively implement a simple machine learning algorithm from scratch. The algorithm should be of a similar level of sophistication as the ID3 decision tree building algorithm.

We recommend using Python in the cloud computing environment provided by Google Colab (colab.research.google.com) for the implementation of the project. We recommend maintaining a Jupyter Notebook on Github to contain all your source code and the report. We will provide examples in class. If you choose to use other tools, you need to make sure your implementation can be tested in a cloud computing environment (see the replicability requirement below).

Requirements

The report will be marked with the criteria elaborated in the table below:

	Practical Project	**Algorithm Implementation***
Introduction	Define Problem; Justify Significance W-200**	Discuss the Algorithm; Define Input/Output W-100
Exploration	Identify Challenges; Design Data Structures/Plan Data Models and Tests W-300	
Methodology	Implement Algorithms; Build and Train Data Models W-300 for Practical Project W-100 for Algorithm (comments not included)	
Evaluation	Report Execution on Data; Perform and Report Testing; Perform Efficiency Analysis; Do Possible Comparative Study W-200	
Conclusion	Discuss Reflections; Propose Possible Improvements W-100	
Ethical	See below W-200	
Video Pitch	Highlight Challenges and Effort	
TOTAL	*: 31005 Only;	
REMARK	** Word counts are suggestive only. Volume is not one of the marking criteria.	

Criteria

	Practical Project
100	The practical problem is well-motivated with clear practical significance. The design/planning of the research/development is clear and logical, including considerations about data acquisition, quality control, modelling techniques, evaluation method and criteria, etc. The design of data models and experiments is logical. The implementation is correct, efficient and practically complete (comparing, or at least, considering alternatives). The evaluation report is convincing; the implementation of the experiments can be easily accessed from a cloud-based service such as Github and replicated for assessment. The report includes discussions about the social/ethical aspect of the proposed technique/project. You can adopt one or more ethical models, such as the utilitarian approach or the Kantian duty-based approach. You should anticipate the potential misuses of your technique. The report is clearly structured and well written. (31005 Only) Video pitch highlights challenges and how you or your team addressed them.
50	The practical problem is clearly defined. The design/planning of the research/development is mostly logical, including considerations about data, modelling techniques, evaluation criteria. The implementation is correct and can be easily accessed from a cloud-based service such as Github and replicated for assessment. The implementation may contain minor issues that can be addressed by a tutor without major revision. The report includes discussions about the social/ethical aspect of the proposed technique/project. You should anticipate the potential misuses of your technique. The report is written in clear English with a basic academic report structure. (31005 Only) In the video pitch you talk about challenges and mention your solution.
0	The practical problem is unclear. The design/planning of the research/development does not address the problem. The implementation cannot be replicated. Social/ethical discussion is missing. The report is not in English or not structured at all. (31005 Only) Video pitch is missing or non-informative.

	Algorithm Implementation (31005 Only)
100	The report contains a clear brief introduction to the algorithm. The input/output data formats are clearly stated. The implementation is correct. The program is well commented, explaining how the computer program codes correspond to algorithm steps. The implementation of the experiments can be easily accessed from a cloud-based service such as Github and replicated for assessment. The report includes discussions about the social/ethical aspect of the potential application of the algorithm. You can adopt one or more ethical models, such as the utilitarian approach or the Kantian duty-based approach. You should anticipate the potential misuses of your technique. The report is clearly structured and well written. Video pitch highlights challenges in the algorithm implementation, such as memory management, time efficiency or advanced functions such as parallelism, and how you or your team addressed them.

	Algorithm Implementation (31005 Only)
50	The report contains an introduction to the algorithm. The input/output data formats are listed. The implementation is mostly correct. The program is commented to be readable. The implementation of the experiments can be easily accessed from a cloud-based service such as Github. The implementation may contain minor issues that can be addressed by a tutor without major revision. The report includes discussions about the social/ethical aspect of the proposed technique/project. You should anticipate the potential misuses of your technique. The report is written in clear English and is structured. In the video pitch you talk about challenges and mention your solution.
0	The report misses an introduction to the algorithm, or the input/output data formats are not mentioned. The implementation cannot be replicated. Social/ethical discussion is missing. The report is not in English or not structured at all. The video pitch is missing or non-informative.

Marks

This assignment contributes 40% to your final mark for 31005 and 30% for 32513.

SUBMISSION

Due date 11:59pm 25 Sep 2019.

You need to submit an exported PDF of your Jupyter Notebook, including a PLAIN TEXT link to the Github file to UTSOnline.

Late Penalty 10 marks per day (round-up) past deadline.

Extension Extensions may be granted if arranged with the Subject Coordinator before the deadline and if decent progress/effort has been made at the time of application. We use github commission history of your draft report as evidence for progress/effort.

If your performance in an assessment item or items has been affected by extenuating or special circumstances beyond your control you may apply for Special Consideration. Information on how to apply can be found at http://www.uts.edu.au/current-students/managing-your-course/classes-and-assessment/special-circumstances/special.

Due to the size of our class, extensions will lead to delay in marking your assignment and **all future assignments and the final grade**.

GROUPWORK

This assignment may be done individually or in pairs. Conditions for group work are described in the subject outline. Except for exceptional circumstances (i.e. where problems occur in the group), each member will receive the same mark. If there are problems in your group, please see the Subject Coordinator.