MAIS 202 - PROJECT DELIVERABLE 2

Due: Tuesday, October 14th, 2019

Over the course of MAIS202, you will be completing a machine learning based project of your choice for the final project. At the end, you will have the option to either:

1. *Demo your project by integrating it into a webapp (or something more advanced)* ***OR***
2. *Present your work as an academic project through a poster*

McGill AI Society will be hosting a project fair where you will be showcasing your work! It will be awesome!
For both cases, you will end the project by writing a blog post about it.

Submission

This is an individual deliverable. All deliverables should be electronically submitted on Github and completed with the same academic integrity and standards expected at McGill University. Include appropriate citations.

Submit both your *well-documented* code and report.

Deliverable Description

In this deliverable, you will discuss your progress and report your preliminary results. Be precise in your explanation and report. If you discussed your approach with other students, you should honour them in your report.

1. **Problem statement**

Restate the initial project that you proposed in deliverable one in 2 - 3 sentences. Be sure to refer back to this problem statement in the following questions.

2. **Data Preprocessing**

Confirm the dataset you are working with. State any changes from the initial dataset you chose. Discuss the content of the dataset (number of samples, labels, etc). Describe and justify your data preprocessing methods.

3. **Machine learning model**

In the first deliverable, you proposed a model for your project. If you decided to change your model, explain why. Restate your chosen model and elaborate on the design decisions.
Report the following:

- Discuss the framework and tools that you used for your model. Explain your choice. Provide architecture graphs as appropriate.

- Justify any decision about training/validation/test splits, regularization techniques, optimization tricks, setting hyper-parameters, etc.
- Description of validation methods How did you test your model? Is your model overfitting or underfitting?)

Did you face any challenges implementing the model? If so, how did you solve it?

At this point, don't forget to save your trained weights! You will need them for the integration and/or testing your model!

4. **Preliminary results**

In this section, you will focus on the performance of your model. Present a detailed analysis of your results, providing graphs as appropriate. Analysis requirements differ in every field, but some things to consider reporting include:

- Confusion matrix and accuracy/precision-recall/logistic loss (classification problems).
- Mean squared error (regression problems)
- Rand index (unsupervised models)
- BLEU score with brevity penalty (text generation)
- variance of the dimension reduced set vs variance of the initial dataset (dimensionality reduction/PCA)

In addition to an evaluation metric, discuss the overall performance of the model and the feasibility of the project with these results.

Remember, graphs are beautiful and we love them!

5. **Next steps**

Discuss your next steps. Describe the pros/cons of your approach and future work. Will you be altering your model? For example, will you be fine-tuning it? At this point, if you think that your model is not performing well and/or does not work, please reach out to an exec to see what you can do to improve it.