**COMP 551 – Machine Learning   
Project #1**

**Machine Learning 101**

**Group #  
Emmanuelle Coutu-Nadeau (260681550)    
Lynn ()    
Ilyes ()**

**McGill University   
September 30, 2021**

**Abstract (100-250 words)**

Summarize the project task and your most important findings. For example, include sentences like “In this project we investigated the performance of two classification models, namely k-nearest neighbours and decision trees, on predicting if the income of an adult exceeds $50K/yr from various factors, such as age, sex, nationality, etc...”, “We found that the k-nearest neighbour regression approach achieved worse/better accuracy than decision trees and was significantly faster/slower to train.”

**Introduction (5+ sentences)**

Summarize the project task, the two datasets, and your most important findings. This should be similar to the abstract but more detailed. You should include background information and potential citations to relevant work, if any, (e.g., other papers analyzing these datasets).

**Datasets (5+ sentences)**

Very briefly describe the datasets and how you processed them. How did you handle the missing data? If you have come up with new new features to get better results, you should explain it here. Present your efforts for better understanding of the data, e.g. through visualization plots.

**Results (7+ sentences, possibly with figures or tables)**

Describe the results of all the experiments as well as any other interesting results you find. Elements we expect to see: 1. Comparing performances between KNN and decision trees  
2. Revealing how changing hyperparameters affects performances for both models  
3. Describe how reducing the amount of data impacts results

**Discussion and Conclusion (5+ sentences)**

Summarize the key takeaways from the project and possibly directions for future investigation.

**Statement of Contributions (1-3 sentences)**

State the breakdown of the workload across the team members.