COMP 551 ASSIGNMENT 2

NAME LASTNAME, MARC-ANDRE ROUSSEAU, AND NAME LASTNAME

ABSTRACT. Summary of task and findings

1. INTRODUCTION

(5+sentences) Context, Project Task, Data set, findings (abstract 2.0)

2. RELATED WORK

(4+sentences) Summarize Literature

3. DATASET AND SETUP

(3+sentences) Very briefly describe the dataset and any basic data pre-processing methods that are common to all your approaches (e.g., tokenizing). Note: You do not need to explicitly verify that the data satisfies the i.i.d. assumption (or any of the other formal assumptions for linear classification).

4. PROPOSED APPROACH

(7+sentences) Briefly describe the different models you implemented/compared and the features you designed, providing citations as necessary. If you use or build upon an existing model based on previously published work, it is essential that you properly cite and acknowledge this previous work. Discuss algorithm selection and implementation. Include any decisions about training/validation split, regularization strategies, any optimization tricks, setting hyper-parameters, etc. It is not necessary to provide detailed derivations for the models you use, but you should provide at least few sentences of background (and motivation) for each model.

5. RESULTS

(7+sentences)
Let us test some math formula, "X^3"
Cite something [1]
Let us test some figure here:
Let us test some table here
Provide results on the different models you implemented (e.g., accuracy on the validation set, runtimes). You should report your leaderboard test set accuracy in this section, but most of your results should be on your validation set (or from cross validation).

6. DISCUSSION AND CONCLUSION

(3+ Sentences)

7. DIVISION OF WORK

- **Name Last Name**:
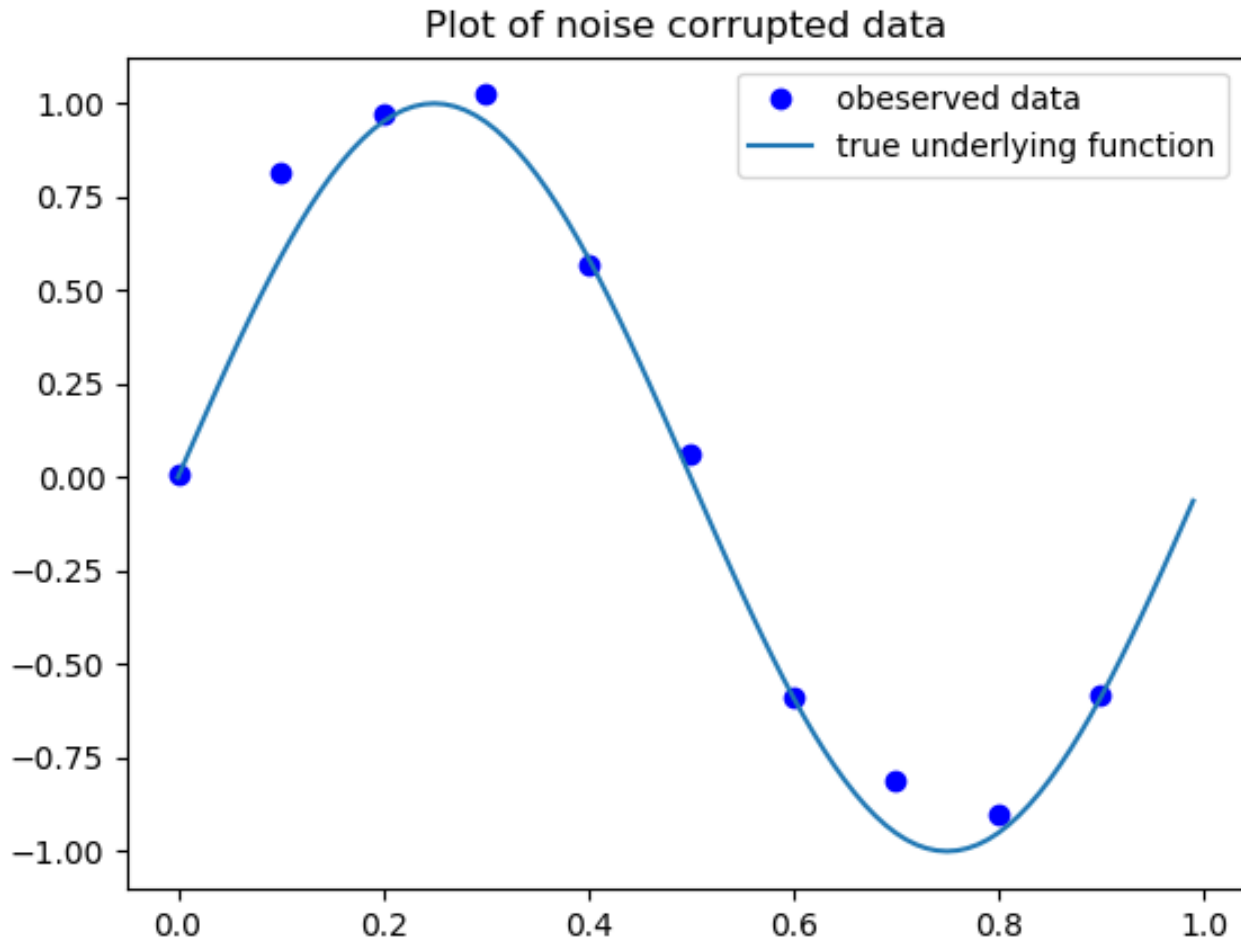- **Marc-Andre Rousseau**:
- **Name Last Name**:

FIGURE 1. run time, accuracy performance comparsion ...

REFERENCES

[1] P. Domingos, "A few useful things to know about machine learning," *Communications of the ACM*, vol. 55, no. 10, pp. 78–87, 2012.

Feature(s) w.r.t Performance	MSE of Training Data	MSE of Validation Data	MSE of Testing Data	run time (in S)	pre-run time (in S) time for computing text/numeric features and generate related matrix
Basic Model (3 numeric features)	1.084683	1.020326	1.297531	0.027982	3.207755
Model With:					
"60 high frequency words"	1.059316	0.969286	1.298629	0.026985	88.564367

TABLE 1. Performance Evaluation For Varies Implemented Features and Their Combinations