Title

Report for the CEGE0004: Assignment

**─**

Group Name

Your Name 1

Your Portico Number 1

Your Email 1

Your Name 2

Your Portico Number 2

Your Email 2

Your Name 3

Your Portico Number 3

Your Email 3

Your GitHub Repo Link

Use the text highlighted in orange as a guideline and delete it before submitting the report.

General recommendations:

* Delete the sections that you have not completed.
* You can copy and paste functions/methods/classes/bash-commands here and describe how they work. Please don’t copy and paste screenshots (penalties will occur if you do that).

# Introduction

Describe the problem and look at the big picture.

This section should answer the following questions:

* Why would anybody care about this task? Motivate your work.
* Which learning tasks are you trying to solve?
* How should performance be measured?
* How would you solve this problem manually?

# Learning Tasks

Formalize mathematically the tasks you solved. For each task, you should clearly describe what kind of features your model will deal with and what is the expected output.

# Material

Describe the data.

Do some data exploration.

Study each attribute and its characteristics:

* Name
* Type (categorical, int/float, bounded/unbounded, text, structured, etc.)
* % of missing values
* Noisiness and type of noise (stochastic, outliers, rounding errors, etc.)
* Type of distribution (Gaussian, uniform, logarithmic, etc.)

Identify the target attribute(s).

Visualize the data.

Study the correlation between attributes.

Identify any common (among the tasks) transformations you may want to apply.

Describe how you sampled the test set.

# Technology

Describe any technology you used in order to solve this assignment.

# Learning Algorithms

## Decision Trees

Give a general introduction to this family of algorithms and the specific one you picked. Describe which task you solved using this algorithm. Present any specific preprocessing step you needed to perform. Describe how you implemented this algorithm, how you tuned its hyperparameters, and validate its performance.

## Instance-Based Learning

Give a general introduction to this family of algorithms and the specific one you picked. Describe which task you solved using this algorithm. Present any specific preprocessing step you needed to perform. Describe how you implemented this algorithm, how you tuned its hyperparameters, and validate its performance.

## Bayesian Learning

Give a general introduction to this family of algorithms and the specific one you picked. Describe which task you solved using this algorithm. Present any specific preprocessing step you needed to perform. Describe how you implemented this algorithm, how you tuned its hyperparameters, and validate its performance.

## Neural Networks

Give a general introduction to this family of algorithms and the specific one you picked. Describe which task you solved using this algorithm. Present any specific preprocessing step you needed to perform. Describe how you implemented this algorithm, how you tuned its hyperparameters, and validate its performance.

## Model Ensemble

Give a general introduction to the ensemble method you chose and make an ensemble of the models above for a shared task. Describe which task you solved using this algorithm. Present any specific preprocessing step you needed to perform. Describe how you implemented this algorithm, how you tuned its hyperparameters, and validate its performance.

# Conclusion

Present a comparison of the models studied above on the test results and make some hypotheses about why you observed these behaviors.

# Git Log

Copy and paste here the output of git log.