Machine Learning for Business Applications 46-887

Carnegie Mellon University
Tepper School of Business

Homework 2

Overview

The goals of this second assignment are:

1. To train *at least* 1 ML model based on the source(s) (textbooks, blog posts, . . .) and the dataset(s) that your team identified in the previous assignment.

2. To identify appropriate statistical performance metrics to evaluate the quality of the ML model(s).

3. To quantify the statistical performance of the model(s) using the dataset(s) that your team identified in the previous assignment.

In one sentence: **by the end of this second assignment, your team needs to have at least 1 ML model that provably solves the business problem of your choice *to some degree* (as measured by the performance metrics that your team has identified).**

In the following assignments, your team will build on top of this model to create a POC ML system.

Deliverable

Your team will submit a project update of no more than 3 pages describing:

1. What model(s) the team has trained and evaluated.

2. What metrics have been used for the evaluation of the model(s).

3. What best-practices (e.g., data splitting, cross-validation, . . .) the team has applied (e.g., to tune and evaluate the model(s)).

4. How the team has organized the code that is responsible for the training and evaluation of the model(s). E.g., has the team used one or multiple Jupyter notebooks? Is the code version-controlled somewhere (e.g., on GitHub)? What assets (e.g., model files) does the code generate, and where are these assets saved? Has your team produced a "runbook" or a set of instructions that others can read to use the same code to reproduce your team's work? ...

Your team is welcome to include screenshots and other figures as needed. These do not count towards the 3 page limit.

Your team can submit the project update as a PDF document or as a Word document.

Checklist

Here is a checklist to help your team make sure that the most important items are covered in this project update.

☐ Train at least 1 ML model based on the sources that your team identified in the previous assignment.

☐ Identify appropriate metrics to quantify the statistical performance of the ML models that your team has trained, and justify why these metrics are relevant and appropriate for the business problem that your team is trying to solve.

☐ Evaluate the trained model(s) with respect to the metrics that your team has identified.

☐ Discuss the best practices that your team has applied when training and evaluating the model(s).

☐ Discuss how your team has organized the code that is responsible for the training and evaluation of the model(s), and whether your team has produced any documentation to allow others to use the same code to replicate your team's work.

Other notes

- For the purpose of this assignment, it is OK to closely replicate the training and evaluation strategies used in the sources that your team has previously identified. If your team sees opportunities to make improvements, go ahead and make those improvements. However, your team is only asked to *replicate* the training and the evaluation of models that were identified in the previous assignment.

- **If your team is using a publicly available pretrained model, please use this alternative checklist:**

☐ Read the relevant documentation/papers associated with the pre-trained model(s) that you selected and describe in detail how these models were trained by their authors.

☐ Discuss in detail what metrics were used by the authors of these models to evaluate their statistical performance, and explain why these metrics are relevant and appropriate for the business problem that your team is trying to solve.

☐ Read and report the performance that these models achieved on the datasets that the authors used for their evaluation.

☐ Discuss the best practices that the authors have applied when training and evaluating these model(s).

☐ Discuss how your team has organized the code that is responsible for downloading and using these pretrained models, and whether your team has produced any documentation to allow others to use the same code to replicate your team's work.