NYP NANYANG
Official (Open)NOVATIVE POLYTECHNIC

Applied Deep Learning
Assignment (40%)

Introduction

Computer vision and natural language processing are two areas that have seen great advances due to deep learning. In this learning unit, we have seen how deep learning models are used for the above tasks. You will apply what you have learnt for this assignment. This is an **individual** assignment.

Tasks

There are two parts to this assignment.

For Part 1, you will have to find a suitable computer vision dataset to train a deep learning model. You will (i) build and train your own deep learning model (ii) train another deep learning model (of your choice) using transfer learning. You are also expected to compare and comment on the performance of both models.

For Part 2, you will use the drug review data to build a text classifier. The dataset for Part 2 is obtained from https://www.kaggle.com/datasets/mohamedabdelwahabali/drugreview. Do remember to cite this reference in your report. A training and test dataset is provided. The outcome variable for this dataset is a rating from 1 to 10. You will have to convert this to a qualitative label (positive, negative) and train and tune your own deep learning model with the training dataset. Additionally, you should look for another sentiment dataset and use this dataset to train and tune a second deep learning model **with the same architecture/layers** as the first model. You are expected to comment on the performance of both models on test data from the **drug review dataset**.

For both of the parts, you will have to:

a. Provide an introduction and background to the problem that you are trying to solve using your deep learning models. Explain how deep learning is applied to your use case
b. Apply appropriate data pre-processing methods
c. Train and evaluate the performance of the models. For part 1, you should compare the performance of your own model with the one using transfer learning.

For part 2, you should comment on the performance of both models on test data from the drug review dataset.

d. Produce a maximum 10 minute video (5 minute for each part) summarizing and explaining what you have done.

Do remember to cite all resources used in this assignment.

Submission

Submit a report of your findings. The report should include all the necessary exploratory work you have done, description of data, important code snippets, comparisons, discussions on performance, and any insights you have gained from the study. You can use the Markdown in Jupyter notebook for this instead of submitting a separate report if you prefer.

You should also include a Jupyter notebook (ipynb file) in the submission. Additionally, do also include the dataset (or link to dataset). All the deliverables should be added to a zip archive and the zip file is to be uploaded to POLITEMall by the date specified in the delivery plan. Late submission without valid reason may result in deduction of marks.

You should also submit a video (maximum length: 10 minute) summarizing and explaining what you have done.

Assessment Rubrics

Marks will be allocated based on the completeness, thoroughness and correctness of the solution. The rubrics for both part 1 and part 2 are as below.

	Developing (0-5%)	Functional (5-8%)	Advanced (8-10%)
Introduction and Background (5%)	Confusing description on the subject introduction and background.	Describes the introduction and background of the subject.	Clearly explains the project and the process of using analytics to solve the problem.
Data Pre-processing (5%)	Little or incorrect pre-processing on the data	Some pre-processing done on the data.	Comprehensive pre-processing of data,

NYP NANYANG
Official (Open) INNOVATIVE POLYTECHNIC

	and hence resulting in incorrect or poor results.		concise and detailed explanations in report.
Training, Evaluation and Discussion of models (10%)	Model has flaws and poor performance that can be easily improved. Incorrect evaluations of model performance and no meaningful comparison of models	Model was correctly trained, but evaluation of models was incomplete. Some discussion and comparison of model performance	Model was of good design and considerations. Detailed evaluation, discussion and comparison of model performance

---- The End ----