

Industrial Engineering (Deep Learning) 874

Formative Assessment: Exploring Deep Learning

Department of Industrial Engineering

Deadline: 25 June 2020, 23:59

Total: 70

Instructions

The focus of this assignment is to test your understanding of certain concepts covered in the lectures from days 1-2. In addition, your implementation of these concepts on real-world data will also be tested in this assignment.

- Answer all the questions below.
- Submit your typed answers as a pdf document. Please also submit all other documents required to obtain your answer. For instance, submit your Python script(s) that you used for any of the questions. In addition, also submit the video recording explaining your coding process.
- Please make sure that you do and submit your own work. Plagiarism will not be tolerated.
- Note that late submissions cannot be accepted and that no extensions to the deadline can be provided.

Building and Improving a Deep Learning Model

[75]

In this assessment you are given the choice of modelling *any* interesting problem of your choice and building a relevant deep learning model using the concepts covered in the lectures during days 1-2. Specifically, you will choose a publicly available data set (e.g., a data set from Kaggle, UCI Machine Learning Repository, etc.) and build a deep learning model (in your programming language of choice). Note that you are *not* allowed to use the IRIS/MNIST/Boston house prices datasets. The deep learning model must be a multi-layer perceptron (MLP) model or any of the autoencoders that we covered in the lectures. The problem can be any of the following:

- Classification
- Regression
- Compression (using an autoencoder)

As part of your submission, do the following:

- 1) Describe the problem you are modelling. Specifically, discuss the reason you chose the problem and provide a high-level discussion on how you are approaching the problem. What value would modelling the problem of your choice provide (to the research

community or in practice)? Also give an overview of the data set you are using. If you are using a well-known data set, discuss how other people have approached the problem. **[10]**

- 2) Create a base deep learning model and discuss your model architecture and all hyperparameter choices. **[10]**
- 3) Split your data set into training, validation, and test sets using any appropriate ratio. Implement *at least* 3 techniques on your base model to improve the performance (choose an applicable performance metric(s)). These techniques include, but are not limited to dropout, early stopping, changes to model architecture, and other regularization techniques. Plot the performance metrics you are measuring during this optimization process. **[20]**
- 4) Report your results on the test set (or plot reconstructed images if using autoencoders) and discuss the effects of implementing the measures in (3). **[15]**
- 5) Create a 10-minute video recording of you working through and explaining your code. Refer to your data pre-processing, model development, model training, model improvement, and model testing. Submit this video as part of your submission. **[20]**