INNS Open Assessment

Y3843100

March 1, 2020

1 [20 marks] Discussion of architectures.

This section should:

- describe (briefly) the data you have, and how much there is of it.
- identify the type of problem
- identify which classes of architectures would be suitable
- give a brief discussion of the technical features of the architectures, and the advantages and disadvantages of each
- state which class of architecture you are going to use and justify your choice, relating the characteristics of the problem to the advantages/disadvantages of the architecture.

To do this you might need to:

- do some preliminary experiments with simple versions of the architecture to get a feel for what will work
- do some exploratory data analysis to see what the characteristics of the data are
- consider the principles involved and relate them to the problem.

The dataset contains 2126 fetal cardiotocograms (CTGs) from different patients each of which has 23 different recorded features. The CTGs have been annotated by three expert obstetricians creating two categories of classes.[1] One is a 10 tuple with respect to the fetal heart rate FHR patterns and the other is a three tuple regarding fetal state. This gives us two classification problems with respectively 10 and 3 distinct classes. Neural Network architectures that can handle classification problems need to have appropriate activation functions and the ability to specify the targets (outputs) as a finite set of discrete classes. We will discuss **three** different conceptual models with respect to their ability to be configured for classifying our high dimensional CTG dataset.

[1] Ayres de-Campos Bernardes Garrido Marques-de Sa Pereira-Leite. *UCI Machine Learning Repository*. 2000. URL: `https://archive.ics.uci.edu/ml/datasets/cardiotocography`.

1.0.1 Perceptron

The *perceptron* is a basic network structure in which our output class y is determined by a weighted sum of our inputs X that is evaluated against some hard limit (threshold or activation function) $y = H(\sum_i^X x_i w_i)$. Both the advantages and disadvantages of the perceptron are in its simplicity. On one hand we have intuitive behaviour in the fact that the perceptron creates a line that bipartitions our data space, but on the other we are limited only to linearly separable classes. Furthermore a perceptron can perform only binary classification or at best *one-vs-many*. Our data does not appear to be linearly separable and our task is to discriminate between all the distinct classes, therefore we rule out the perceptron as a viable architecture.

1.0.2 Multi-Layer Perceptron MLP networks

The shortcomings of the single perceptron are addressed by its orchestrated counterpart, the multi-layer perceptron *MLP*. The three main differences as highlighted by Haykin:[2]

- activation functions must be differentiable, unlike the hard limits we had before
- between our input and output nodes we construct one or more *hidden layers* containing one or more neurons with our modified activation function
- the input, output and hidden neurons are highly connected

1.0.3 Radial Basis Function RBF networks

Relating the aforementioned properties of the different architectures we see that the MLP and RBF are theoretically viable for both of our classification problems. **[[relate net property to problem i.e. what the dataset gives us and what the network provides to solve that, high dimensions, many classes, not enough data (prone to overfitting), etc.]]**. After further investigation of the semantic meaning of the we will empirically compare the two networks with different formulation of our problem.

2 [40 marks] Creation and application of neural networks.

This section should

- Describe the chosen inputs to (and outputs from) the networks.
- Describe how the data you started with have been preprocessed.
- Give sufficient detail for someone else to process a new batch of data for use with the final trained network.

[2] Simon Haykin. *Neural Networks: A Comprehensive Foundation*. 2nd. USA: Prentice Hall PTR, 1998. ISBN: 0132733501.

- State which training algorithm you selected, and explain how you selected that training algorithm. For this training algorithm, give sufficient detail to enable someone to use the same training algorithm in exactly the same way. This does NOT mean (for example) describing gradient descent in great detail. It DOES mean giving any parameters, initialisation, etc, even if they are the toolbox defaults.
- Explain the process you went through in making the selection of the final architecture, for example, the number of neurons or the number of layers to use.

To do this you might need to:

- Test of one or more networks to demonstrate the effect of different preprocessing choices on the performance of the network.
- Try different training algorithms on one or more networks to compare performance.
- Evaluate a number of networks, and record details of their structures and how

3 [20 marks] Results and evaluation

This section should

- Explain the metric or metrics you have used for comparison between networks.
- Give a synopsis of the results obtained from the final selected network.
- Evaluate the results, in relation to the problem posed in the scenario.

To do this you might need to:

- Consider different metrics for performance, appropriate to the problem. Remember that a Mean Squared Error (MSE) on its own is not always helpful in judging how well something works.
- Identify anything of interest in the results, such as areas of particularly good or poor performance, or variation between different training runs.
- Reflect on the conclusions that you may draw from the results, and whether they are showing that the neural network is useful in this case.

4 [20 marks] Further application

In the previous sections you used a neural network to convert cardiotocogram features into a diagnosis. Another tool for detection and diagnosis of fetal abnormalities is the ultrasound scan, that produces an image of the a section through the fetus. Interpreting fetal scans is a highly complex task which require years of training. Assume the availability of a large

number of fetal scan images, both normal and with some abnormality, and labelled to indicate different types of abnormalities. The task is for a neural network to process new ultrasound images, and to indicate which images needed further investigation. This section should discuss the issues you would need to consider in relation to:

- selection of an architecture
- construction of the network
- use of data for training
- evaluation of the network