

Hyperparameter Study for Feedforward Neural Networks on Classifying Text Data

Marius Kleidl
06.07.2018

Motivation

- Text classification is part of our daily life
- Distinguish bills and ads in letters, e-mails, notifications on your smartphone
- Computers use it to filter spam, grouping mails etc.

Motivation

- Neural networks are commonly used for classification
- However, optimizing neural networks is not easy
 - We conduct an experiment to see how different hyperparameters affect a neural networks' accuracy

Agenda

1. Motivation
2. Experiment
3. Data Set
4. Results
5. Conclusion

Experiment

1. Setup simple feature selection for text data
2. Train feedforward neural network using different hyperparameter values
3. Evaluate accuracy using test data set

Experiment – Hyperparameters

We limit our experiment to three parameters:

1. Number of layers (1 or 2)
2. Number of perceptrons per layer (50 or 100)
3. Learning rate (1, 0.1, 0.01, 0.001, 0.0001)
 - In total: 20 different combinations

Data Set

20 Newsgroups

~ 18 000 posts from 20 different news groups with overlapping topics

comp.graphics comp.os.ms-windows.misc comp.sys.ibm.pc.hardware comp.sys.mac.hardware comp.windows.x	rec.autos rec.motorcycles rec.sport.baseball rec.sport.hockey	sci.crypt sci.electronics sci.med sci.space
misc.forsale	talk.politics.misc talk.politics.guns talk.politics.mideast	talk.religion.misc alt.atheism soc.religion.christian

Table from <http://qwone.com/~jason/20Newsgroups/>

Data Set – Example

From: JC924@uacsc2.albany.edu
Subject: Why are our desktop fonts changing?
Organization: University at Albany, Albany NY 12222
X-Newsreader: NNR/VM S_1.3.2
Lines: 17

One of our users is having an unusual problem. If she does an Alt/Tab to a full-screen DOS program, when she goes back to Windows her desktop fonts have changed. If she goes back to a full-screen DOS program and then goes back to Windows, the font has changed back to its default font. It's not a major problem (everything works and the font is legible), but it is annoying. Does anyone have any idea why this happens. By the way, she has a DEC 486D2LP machine.

```
+++++
Jeffrey M. Cohen                               Voice: 518-442-3510
Office for Research (AD 218)                   Fax: 518-442-3560
The University at Albany                       E-mail: JC924@uacsc2.albany.edu
State University of New York
1400 Washington Ave.
Albany, NY 12222
+++++
```

Example from the comp.os.ms-windows.misc newsgroup

Neural Network

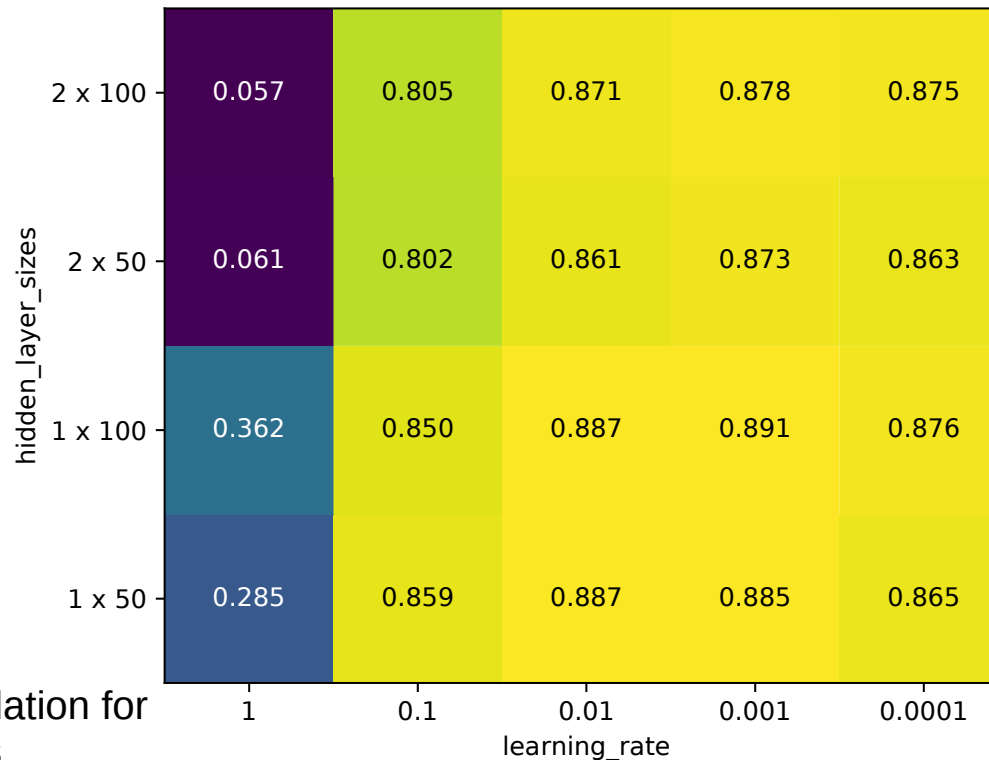
- Remove English stopwords (e.g. 'the')
- Text is represented as a Bag of Words
 - No sentence structure is preserved
- Normalize count matrix using Tf-Idf
 - Limit impact of common words
- No further text processing is done (e.g. stemming)

Training and Validation

- Training is stopped if accuracy stops improving for 0.1 p.p. for two iterations
- Each neural network is evaluated using four-folded cross-validation
- Final accuracy score is determined using previously unused test set

Results

- Best combination: 1 layer with 100 perceptrons and learning rate of 0.001
- Accuracy on unseen test set: **81.5%**



Accuracy from cross-validation for different hyperparameters

Conclusion

- Tuning hyperparameters is important for optimization
- However, other network architectures and advanced text processing may be needed for further increased accuracy