CSYE_7245_Advances_in_Data_Sciences_and_Architecture_Assignmen

March 26, 2018

0.1 Assignment 4 - Deep Learning

Due Monday, April 9, 2018

Submission: Put the data and Jupyter notebook files in a folder. Make sure all links to data are relative to the folder so the TAs can run the notebooks.

0.2 Deep Learning for analysis of your project data (if possible).

Apply a Deep Learning model to your project data (if possible). You can use another data set ONLY if it makes no sense to use a Deep Learning model for your project.

There will be NO EXTENSIONS on the assingments due in April so start early

0.2.1 Part A - Deep Learning model (40 points)

On *your project data* * Apply a Deep Learning model to your project data (if possible). Validate the accuracy. * The Deep Learning model can be a CNN, RNN, Autoencoder, Variational autoencoder (VAE), Restricted Boltzmann machine (RBM), Deep belief network (DBN) or Generative Model. It cannot be a simple multilayer perceptron (MLP).

0.2.2 Part B - Activation function (10 points)

On *your Deep Learning model data* apply at least two different activation functions.

- Change the activation function. How does it effect the accuracy?
- How does it effect how quickly the network plateaus?
- Various activation functions:

- Rectified linear unit (ReLU)

- TanH

- Leaky rectified linear unit (Leaky ReLU)

- Parameteric rectified linear unit (PReLU)

- Randomized leaky rectified linear unit (RReLU)

- Exponential linear unit (ELU)
- Scaled exponential linear unit (SELU)
- S-shaped rectified linear activation unit (SReLU)
- Identity
- Binary step
- Logistic
- ArcTan
- Softsign
- Adaptive piecewise linear (APL)
- SoftPlus
- SoftExponential
- Sinusoid
- Sinc
- Gaussian

0.2.3 Part C - Cost function (10 points)

On *your Deep Learning model data* at least two different cost functions.

- Change the cost function. How does it effect the accuracy?
- How does it effect how quickly the network plateaus?
- Various forms of cost:

- Quadratic cost (mean-square error)
- Cross-Entropy

- Hinge

- Kullback–Leibler divergence

- Cosine Proximity

- User defined

And many more, see `https://keras.io/losses/`

0.2.4 Part D - Epochs (10 points)

On *your Deep Learning model data*
* Change the number of epochs initialization. How does it effect the accuracy?
* How quickly does the network plateau?

0.2.5 Part E - Gradient estimation (10 points)

On *your Deep Learning model data* at least two qradient estimation algorithms.

- Change the gradient estimation. How does it effect the accuracy?
- How does it effect how quickly the network plateaus?
- Various forms of gradient estimation:
- Stochastic Gradient Descent
- Adagrad
- RMSProp
- ADAM
- NAG
- Adadelta
- Momentum

0.2.6 Part F - Network Architecture (10 points)

On *your Deep Learning model data* * Change the network architecture. How does it effect the accuracy? * How does it effect how quickly the network plateaus? * Various forms of network architecture: - Number of layers
- Size of each layer
- Connection type
- Pre-trained components?

0.2.7 Part G - Network initialization (10 points)

On *your Deep Learning model data* at least two network initialization techniques.

- Change the network initialization. How does it effect the accuracy?
- How does it effect how quickly the network plateaus?
- Various forms of network initialization:
- 0

- Uniform
- Gaussian
- Xavier Glorot Initialization http://andyljones.tumblr.com/post/110998971763/an-explanation-of-xavier-initialization
  - Xavier Uniform
  - Xavier Gaussian

Last update October 3, 2017