Memory protocol of Deep Learning for Speech and Language Processing

16.03.2022

1 Supervised Learning

SVM

With relations to support vector machines (SVMs), define max-margin, support vectors and kernels.

Logistic Regression

You are given the input, weights and true label vector.
Calculate C_{MSE} when using logistic regression.

Explain in 1-2 sentences

1. Explain overfitting?

2. How can you use SVMs for multiclass classifications?

3. Name two stopping conditions for k-means.

4. Explain why decision trees are more transparent compared to neural networks?

2 Neural Networks

FFNN calculation

You are given the input, weights and true label vector.
Calculate the results of a feed-forward neural network (FFNN) that contains 2 x hidden layers, both using the ReLU activation function and an output layer using Softmax. Biases aren't given.

Calculate the cross entropy loss function with log_{10} (writing the equation was enough, since we didn't have calculators).

Backpropagation

Explain the intuition behind using backpropagation in neural networks.

Parameter initialization

Explain why parameter initialization has an effect on the final result of the model. Give two examples of how we can use the Kaiming method.

Explain in 1-2 sentences

1. What can be a problem when using the sigmoid activation function?

2. Give three hyperparameters of a FFNN.

3. What do we use the Softmax function

4. Why do we use non-linear activation functions in NN?

5. What's the reason behind using deep neural networks?

3 CNN

CNN usage

Explain why and how convolutional neural networks (CNN) can be used in a natural language processing task like sentiment analysis.

CNN calculation

Calculate the results of the CNN given a 4x4 matrix, 2x2 filter, 1x2 stride and 2x2 max pooling. The ReLU activation function is used after the convolutional step.

Explain in 1-2 sentences

1. Give three hyperparameters of a CNN.

2. What are the trainable parameters of a CNN?

3. What's something you can do when you have an input of different lengths?

4 RNN

RNN calculation

You are given the input, weights and biases.
Calculate the output of a recurrent neural network (RRN) given sentence *"happy day"* and a vocabulary $vocab = \{happy, plus, make, money\}$ as a bag-of-words.

RNN problems

Give examples of three problems that an RNN can have during training.

LSTM

Give your intuitive explanation behind the workings of a LSTM network.

Explain in 1-2 sentences

1. Explain the difference between Jordan Elman RNN network.

2. What does "time" mean when talking about the backpropagation through time algorithm?

3. Why does a LSTM have 4x the amount of paramaeters compared to a vanilla RNN?

4. Why do we need RNN's when dealing with sequential data?

5. What's the functionality of gates in LSTM.

5 Tricks

Learning curves

Draw learning curves for training and development. Give an example of how to improve the model in the end?

Adaptive learning

Explain the reason for using adaptive learning rate and explain how AdaGrad works.

Explain in 1-2 sentences

1. Explain gradient clipping.

2. What do we do with Dropout when testing on the development set?

3. Why do we use data augmentation?

4. Explain residual networks.

6 Reinforced learning

Calculate agent action

The network was the same as the one in the DL mock exam WS21.
Calculate agent action in state s3 and give an answer as to what action it would take given it had a greedy approach.

Markov property

Explain the Markov property.

Exploration-exploitation dilemma

Explain the exploration-exploitation dilemma.

Agent-Environment interaction

Draw the agent-environment interaction diagram.

Set rewards

Set rewards on a state diagram in such a way that a greedy agent is suboptimal.

7 General questions

Self-attention

Describe self-attention. What are the parameters and hyperparameters of self-attention?

Adversarial examples

Explain adversarial examples give two reasons why we should give a bigger focus on them.

Explainable AI

Explain XAI and give links to two existing ethical guidelines.