Universität Duisburg-Essen   Information systems   Prof. Dr.-Ing. N. Fuhr

Information Mining - winter semester 2019

Exercise sheet 5

Exercise 1:   K-means clustering

Explain the core idea of the K-means algorithm. How important is the selection of the seeds and what are the possible causes of the seed selection? How do you determine the k value (number of clusters)?

Exercise 2:   K-means clustering – deeper understanding

- The aim of any machine learning method, including k-means clustering, is to find a global optimum. However, there are sometimes situations where the algorithm is stuck on local minimum. Give a scenario for k-means clustering where such situation can occur.

- Is it true that k-means clustering requires at least two attributes to work? Why?

- k-means clustering is a supervised approach, true or false? Why?

Exercise 3:   Gradient Descent

The file ex1data1.txt[1] contains the dataset for our linear regression problem. The first column is the population of a city and the second column is the profit of a food restaurant in that city. A negative value for profit indicates that the restaurant does not do profit but looses money. Use this data to fit a liner model: h(x)= b + mx. As shown in the lecture you can replace b and m with w0 and w1. To learn w0 and w1 you should use your own implementation of Gradient Descent. Make sure you update w0 and w1 simultaneously. Run the process with 1500 iterations. Use 0.05 for the learning rate. You can initialize w0=w1=1.

Exercise 4:   Further on Gradient Descent

What is the difference between batch gradient and stochastic gradient descent? And what is mini-batch gradient descent?

Exercise 5:   Deep Learning

What is the fundamental difference between deep learning and traditional machine learning?

---

[1]http://www.is.inf.uni-due.de/courses/im_ws19/uebung/ex1data1.txt

Exercise 6:   Continue on Deep Learning

Draw a feedforward-neural network with two inputs (x1,x2), 5 neurons on the first hidden layer, 3 neuron on the second hidden layer. The network should perform a binary classification.

Exercise 7:   Softmax

What is softmax? And what do we use it for?

Exercise 8:   Activation Functions

Name 3 activation functions and scatch their behaviour graphically. Why do we use activation functions?