Deep Learning DM873
Exam notes

Mark Jervelund

Mark@jervelund.com

Doommius.com/notes.php

SDU

UNIVERSITY OF
SOUTHERN DENMARK

IMADA

June 24, 2019

Contents

Part I

Formal course description

Aim

Machine learning has become a part in our everydays life, from simple product recommendations to personal electronic assistant to self-driving cars. More recently, through the advent of potent hardware and cheap computational power, "Deep Learning" has become a popular and powerful tool for learning from complex, large-scale data. In this course, we will discuss the fundamentals of deep learning and its application to various different fields. We will learn about the power but also the limitations of these deep neural networks. At the end of the course, the students will have significant familiarity with the subject and will be able to apply the learned techniques to a broad range of different fields.

The course builds partly on the knowledge acquired in the course DM555 but can be taken by any Computer Science or Computational BioMedicine Master student.

In relation to the competence profile of the degree it is the explicit focus of the course to:

- giving the competence to plan and execute a deep learning task by means of deep neural networks.

- providing knowledge on the different types of deep learning approaches including their advantages and disadvantages.

- transfer learned methods to new fields of applications.

- challenges the student with real-life datasets and problem-solving skills

Statement of aims

- The learning objectives of the course is that the student demonstrates the ability to:

- Describe the principles of deep neural networks in a scientific and precise language and notation

- Analyze the various types of neural networks, the different layers and their interplay

- Describe the feasibility of deep learning approaches to concrete problems

- Understand the theoretical mathematical foundations of the field

- Apply deep learning frameworks for solving concrete problems

Pensum

- All lecture slides are relevant for the exams.

- All readings noted in the lecture list are relevant for the exam.

- Ian Goodfellow, Yoshua Bengio, Aaron Courville - The Deep Learning Book

- Gareth James, Daniela Witten, Trevor Hastie - Robert Tibshirani An Introduction to Statistical Learning (ISL)

Part II

Exam topics

Exam Form

The exam will last about 15-20 minutes. At the beginning, one topic from the list below will be drawn randomly. For each topic the examinee should be prepared to make a short presentation of 5 minutes. It is allowed to bring one page of hand-written notes (DIN A4 or US-Letter, one-sided) for each of the topics. The examinee will have 2 minutes to briefly study the notes for the drawn topic before the presentation. The notes may be consulted during the presentation if needed but it will negatively influence the evaluation of the examinee's performance. During the presentation, only the blackboard can be used (you cannot use overhead transparencies, for instance).

After the short presentation, additional question about the presentation's topic but also about other topics in the curriculum will be asked.

Below is the list of possible topics and some suggested content. The listed content are only suggestions and is not necessarily complete nor must everything be covered in the short presentation. It is the responsibility of the examinee to gather and select among all relevant information for each topic from the course material. On the course website you can find suggested readings for each of these topics.

Feed-Forward Networks
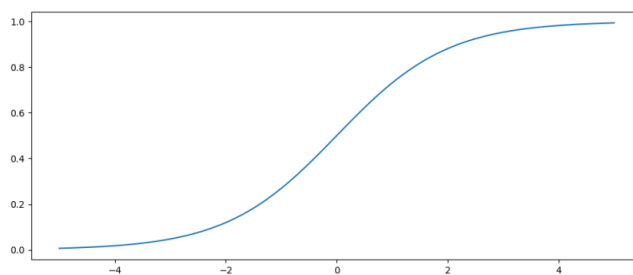
introduction

Function Principle

Output Units

Sigmoid

The sigmoid function is a binary type of output unit, which also means it can only product of a binary variable,and can only predict the Bernoulli distribution. it is of the form

$$\sigma(x) = \frac{1}{1 + exp(-x)} \tag{1}$$

The activation function for the sigmoid looks like the following.

Softmax

The softmax function is a more powerful output that is able to handle Multinoulli distribution/categorical results.
The softmax consists of two layers, first in linear layer followed by the softmax function.

$$Softmax(z)_i = \frac{e^z i}{\Sigma_j e^z j} \tag{2}$$

Hidden Units

Relu

Sigmoid

cosine

Radial Basis

Softplus

Hard tanh

Architecture design

Size

Depth

Backpropagation

introduction

Function Principle

Computational Graphs

Backpropagation through time

...

Regularization

introduction

Over/Underfitting & Model Capacity

Parameter Penalties

Bagging

Dropout

...

Convolutional Neural Networks

introduction

Function Principle

Pooling

Initialization of the kernels

...

Recurrent Neural Networks

introduction

Function Principle

Problems with long term memory

Long Short Term Memory

...

Optimization for Neural Networks

introduction

Parameter Initialization

Adaptive Learning

Batch Normalization

Pre-training

...

Autoencoders and GANs

introduction

Autoencoders

Variational Autoencoders

GANs

...