Course Introduction

Applied ML - 4DV117

Dr. Johan Hagelbäck

johan.hagelback@lnu.se

http://aiguy.org
Outline

• Mostly self-studies
• Pre-recorded lectures to watch plus some reading
• Three physical meetings – introduction and 2 workshops
• The workshops are 4 hours each (10-15 with lunch 12-13)
• In the workshops we will work with practical examples, so you need to bring your laptop to these
• It is not mandatory, but recommended, to attend the physical meetings
Course webpage

- Coursepress: https://coursepress.lnu.se/kurs/applied-machine-learning/
- All course contents, lectures, instructions etc.
- MyMoodle: https://mymoodle.lnu.se/course/view.php?id=32931
- Only for submissions
Welcome to the course Applied Machine Learning (Tillämpad Maskininlärning) 4DV117. This is mainly an online course where you watch pre-recorded lectures and deepen your knowledge by reading the course literature. The course includes three physical meetings: course introduction, workshop 1 and workshop 2. At the workshops we discuss the subject and work with practical problems, so you need to bring your laptop to them. You can find the schedule here: https://coursepress.linu.se/kurs/applied-machine-learning/schedule/

All meetings will be given in both Växjö and Kalmar.

All information about the course, reading instructions, preparations before workshops and pre-recorded lectures can be found here: https://coursepress.linu.se/kurs/applied-machine-learning/

Communication is mainly held through Slack. Create an account and join the channel https://coursepress.slack.com/messages/4DV117/
Applied Machine Learning, 3 credits

In the modern IT world, businesses often have access to large amounts of data collected from customer management systems, web interaction, etc. The data in itself does not bring value to the business; we must bring meaning to the data to create value. Data mining and machine learning is an area within computer science with the goal of bringing meaning to and learning from data. This course will focus on machine learning, where we learn what algorithms and approaches to apply on different types of data.

Who should take this course?

This course is for experienced developers working in the industry.

Content

The course includes the following:

- Supervised learning, different types of data and data processing
- Algorithms for handling text documents
- Algorithms for handling data with numerical and categorical attributes
- Neural Networks
Reading

• Available for free online
• Explains complex stuff in an easy way with lots of examples
• Covers much more than Machine Learning, so only a few chapters are relevant
• It is highly recommended to read those chapters
Reading

• Available for free online
• Good book, but a bit mathematical
• More comprehensive than the other book
• Read the parts you find interesting
Other good books

[Images of book covers]
Workshops

- Work on practical tasks
- Learn to use different APIs, libraries and tools on different types of data
- Work through the tasks in your own pace
- Opportunity to discuss lecture contents and practical tasks

**Before the workshop**
- Watch lectures 1 to 4 and read the corresponding chapters in the course
- Install Java, Python, Weka, R and Scikit on your laptop
- Prepare questions (if you have any) on the contents in lectures 1 to 4

**Aim of the workshop**
The aim of this workshop is to discuss the contents in lectures 1 to 4, and do

**Assignments**

**A1-1: Text classification**
- Download and install Weka if you haven’t done it already
- Download and unzip the Wikipedia_70 dataset from the Datasets page
- Classify the dataset in Weka using the algorithms NaiveBayes and Naïve differences between the two classifiers, and why do you think one is be

**A2-2: Text classification**
- Write Java code for classifying the Wikipedia_70 dataset using the Naïve
- Make sure you apply the StringToWordVector filter and set correct inde
Project

• The examination task is a project you define
• Select some data to work on, preferably from your daily work
• Gather, pre-process and learn from the data
• Evaluate different algorithms
• Describe your findings in a report
Submissions

• Short (1-2 pages) description of your project
• Deadline: coursepress.lnu.se/kurs/applied-machine-learning/deadlines/

• Project files and a technical report describing the project, how you approached the problem and your findings.
• Deadline: coursepress.lnu.se/kurs/applied-machine-learning/deadlines/

• Submit at course page on MyMoodle
Communication

• The best way to get in contact with the course manager is by Slack
• Sign up at coursepress.slack.com using your lnu.se student email address
• Join the channel #4DV117
• All course information will be posted on Slack
Course Introduction

Applied ML - 4DV117

Dr. Johan Hagelbäck

johan.hagelback@lnu.se

http://aiguy.org