Machine Learning Introduction

18.09.2019

Course Information

442173, Lectures and exercises, 5.0 ECTS

Lectures
Location: ExWi, Hörsaal B006
Time: Wednesdays, 13:15 to 15:00

Exercises
Location: ExWi, Hörsaal B006
Time: Wednesdays, 15:15 to 16:00

Courses webpage:
http://cvg.unibe.ch/teaching

Acknowledgements: Prof. Andrew Ng
Stanford University

Course Information

Lecturer:

Prof. Dr. Paolo Favaro paolo.favaro@inf.unibe.ch

Office: Room 213, Neubrückstrasse 10

Teaching assistants:

Mr. Abdelhak Lemkhenter abdelhak.lemkhenter@inf.unibe.ch

Mr. Simon Jenni simon.jenni@inf.unibe.ch

Mr. Givi Meishvili givi.meishvili@inf.unibe.ch

Office: Room 215, Neubrückstrasse 10

Review Assignments

- Topic: Basic Math and Probability
- 3 weekly review assignments
  - All 3 **mandatory** to be admitted to the exam
  - 2 on linear algebra and 1 on probability theory
  - starting today 18.09.2019; due on 09.10.2019 (in 3 weeks)
  - **pass/fail marking for each assignment (min 60 points out of 100)**
  - marks **not included** in the final course mark
  - **Each student must do and submit her work independently**
- Late submissions
  - Only in case of valid justifications (sickness, military, VISA issues etc); a certificate must be produced

Course Assignments

- 3 assignments (~3 weeks time)
- Maximum mark 100 points + bonus points
- **Admission to exam with minimum 50 points in each assignment**
- Grades contribute to the final course grade
- Purpose
  - Each assignment allows you to test and understand machine learning methods
- Late submissions
  - Max 5 days late: each late day incurs a 10 points penalty
  - No penalty if valid justifications are provided (e.g., sickness with doctor certificate, military service etc)

Course Assignments

- Requirements
  - Each assignment requires programming in Python+numpy-scikit+matplot (e.g., design a function) and answering questions
- Timing
  - Provided every three weeks and solutions must be submitted every three weeks
  - First assignment provided on 09.10.2019 and must be submitted on 30.10.2019

Exam

- Admission
  - **At least 60/100 points in each Revision Assignment**
  - **At least 50/100 points in each Course Assignment**
- Final grade
  - **75% based on the Final Exam and 25% based on the Course Assignment average mark**
- Date and place
  - **Final Exam: 8 January 2020 at 16.00-18.00 ExWi A006**

- **Notes, books and cheat sheets are not allowed**

Assignments Calendar

Mock Exams & Review Session

- Mock exams on 06.11.2019 and 18.12.2019
- 2 Mock exams (30 mins each) that cover one half of the course each

- Review session on 18.12.2019:
  - A chance for you to ask questions on the topics/problems/assignments taught during the whole course

Course Information

Resources

There is no required textbook for this course. The following books are recommended as additional reading:

- Richard Duda, Peter Hart and David Stork, *Pattern Classification*, 2nd ed. John Wiley & Sons
- Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer
- Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning, an Introduction*, MIT press
- Kevin P. Murphy, *Machine Learning, A Probabilistic Perspective*, MIT press

Course handouts and other materials can be found on Ilias.

Related activities

Computer vision course (Fall)

It is a MSc course to understand how to process images to detect, classify and segment objects and to obtain 3D geometry, 3D motion and illumination. Machine learning techniques are often used to design these algorithms.

Location: Engehaldenstrasse 8, room 002
Time: Tuesdays, 14:15 to 17:00

Check the website:
http://www.cvg.unibe.ch/computer vision.html

Related activities

Seminar in Machine Learning and Artificial Intelligence (Fall)

This course aims at providing an overview and understanding of the state of the art research in computer vision and machine learning. Recent publications are discussed in detail so that students can build a deep understanding of the methodologies and concepts in these fields.

Location: (Institute of Inf.) Neubrückstrasse 10, room 302

Time: Tuesdays, 10:15 to 12:00

Check the website:
http://www.cvg.unibe.ch/computer vision.html

Related activities

Advanced topics in machine learning course (Spring)

It is a MSc course to learn the latest techniques in machine learning. This year we will focus on **neural networks**. It is the natural continuation of the machine learning course.

Location: Engehaldenstrasse 8, room 002

Time: Tuesdays, 09:15 to 12:00

Check the website:

http://www.cvg.unibe.ch/computer vision.html

What is machine learning?

- Not easy to define
- Here is a first attempt

- Arthur Samuel (1959)
  - **Machine learning**: „Field of study that gives computers the ability to learn without being explicitly programmed“
    - Samuels wrote a checkers playing program
      - He had the program play 10000 games against itself
      - and work out which board positions were good and bad depending on wins/losses

What is machine learning?

- Here is a more recent attempt

- Tom Michel (1999)
  - **Well-posed learning problem**: „A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.“
    - The checkers example
      - E = 10000s games
      - T is playing checkers
      - P if you win or not

What is machine learning?

- There are several types of learning algorithms
  - Supervised learning
    - Teach the computer how to do something, then let it use its new found knowldege to do it
  - Unsupervised learning
    - Let the computer learn how to do something, and use this to determine structure and patterns in data
  - Reinforcement learning
    - Let the computer learn from its own decisions/actions
- This course
  - Look at practical advice for applying learning algorithms
  - Learning a set of tools and how to apply them

Supervised Learning - introduction

- Probably the most common problem type in machine learning

<u>Example problem</u>: „Given this data, a friend has a house 750 square feet – What is the value of the property?“

Supervised Learning - regression

- What approaches can we use to solve it?
  - Straight line through data
    - Maybe $ 150.000
  - Second order polynomial
    - Maybe $ 200.000
  - One thing we discuss later – how to choose straight or curved lines?
  - Each of these approaches represents a way of doing supervised learning
- What does this mean?
  - We gave the algorithm a data set where a „right answer“ was provided
    - This is the **training data**
  - The algorithm should then produce more right answers based on new data where we do not already know the price, i.e., predict the price
    - We call this a **regression problem**

Supervised Learning - classification

- Example
  - Can we define breast cancer as malignant or benign based on tumor size?

  - Looking at data
    - Can you estimate prognosis based on tumor size?
    - This is an example of a **classification problem**

Supervised Learning - classification

- Classification Problem
  - Classify data into one of two discrete classes – no in between, but either malignant or not
  - In classification problems we can have a discrete number of possible values for the output
    - E.g. maybe four values
      - 0 – benign
      - 1 – type 1
      - 2 – type 2
      - 3 – type 3
  - In classification problems we can plot data in a different way

Supervised Learning - classification

- In the previous example we use only one attribute (size)
  - In other problems we may have multiple attributes
  - For example, we may know age and tumor size

Supervised Learning - classification

- Based on that data, you can try and define separate classes by
  - Drawing a straight line between the two groups
  - Using a more complex function to define the two groups (which we will discuss later)
  - Then, when you have an individual with a specific tumor size and who is a specific age, you can hopefully use that information to place them into one of your classes
- You might have many features to consider
  - Clump thickness
  - Uniformity of cell size
  - Uniformity of cell shape
- The most exciting algorithms can deal with an infinite number of features
  - How do you deal with an infinite number of features?
  - Neat mathematical trick in support vector machine (which we discuss later)
  - If you have an infinitely long list - we can develop an algorithm to deal with that

Unsupervised learning - introduction

- Second major problem type
  - In unsupervised learning, we get unlabeled data
  - Given a data set, can you structure it?

One way of doing this is to cluster data into groups
This is a **clustering algorithm**

Unsupervised learning - clustering

- Example of clustering algorithm
  - Google news
    - Groups news stories into cohesive groups
  - Used in any other problems as well
    - Genomics
    - Microarray data
      - Have a group of individuals
      - On each measure expression of a gene
      - Run algorithm to cluster individuals into types of people

Unsupervised learning - clustering

Image segmentation

Unsupervised learning - Clustering

Social Network Analysis

Astronomical data analysis

Market segmentation

Cocktail party problem

Cocktail party problem

- Record slightly different versions of the conversation depending on where your microphone is
  - But overlapping
- Have recordings of the conversation from each microphone
  - Give them to a cocktail party algorithm
  - Algorithm processes audio recordings
    - Determines that there are two audio sources
    - Separates out the two sources

Cocktail party problem

- Is this a very complicated problem?
  - Algorithm can be done with one line of code!

```
[W,s,v] = svd((repmat(sum(x.*x,1), size(x,1),1).*x)*x');
```

  - Not easy to identify
  - But, programs can be short!
  - Using octave (or MATLAB) for example
    - Often prototype algorithms in octave/MATLAB to test as it's very fast
    - Only when it works migrate it to C++
    - Gives a much faster agile development
- Understanding this algorithm
  - svd - linear algebra routine which is built into octave
    - In C++ this would be very complicated!
  - Shown that using MATLAB to prototype is a really good way to do this

Reinforcement Learning

- Problems where a sequence of decisions over time is required
  - A few wrong decisions could be acceptable
  - Basic idea: reward function