Lecture 13:
The world of machine learning

Meelis Kull
meelis.kull@ut.ee
Fall 2018
Lecture 01 Basics of linear classification
Lecture 02 K-nearest neighbours and Naive Bayes
Lecture 03 Linear regression and regularisation
Lecture 04 Linear classification
Lecture 05 Distance-based and kernel methods
Lecture 06 Decision trees
Lecture 07 Evaluation and scoring classifiers
Lecture 08 Class probability estimation and logistic regression
Lecture 09 Neural networks and deep learning
Lecture 10 Ensemble methods
Lecture 11 Probabilistic graphical models
Lecture 12 Bayesian machine learning
Lecture 13 The world of machine learning
Lecture 13 – Overview of machine learning

• Kaggle competition results
• What we have covered in the course
• What we have not covered in the course
• ICML sessions
• You and machine learning
## Kaggle competition results

<table>
<thead>
<tr>
<th>#</th>
<th>Team Name</th>
<th>Kernel</th>
<th>Team Members</th>
<th>Score</th>
<th>Entries</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rain Vagel &amp; Priit Paluoja</td>
<td></td>
<td></td>
<td>4.97809</td>
<td>75</td>
</tr>
<tr>
<td>2</td>
<td>Michael Nitsenko</td>
<td></td>
<td></td>
<td>4.99692</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>Ina_Boescke_Katharina_Foge...</td>
<td></td>
<td></td>
<td>5.01047</td>
<td>48</td>
</tr>
<tr>
<td>4</td>
<td>Jaagup Kirme</td>
<td></td>
<td></td>
<td>5.05619</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>MohamedAbdelrahman-Mah...</td>
<td></td>
<td></td>
<td>5.05691</td>
<td>28</td>
</tr>
<tr>
<td>6</td>
<td>Scarlet Crusade</td>
<td></td>
<td></td>
<td>5.06283</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>Andre Tätta Markus Kängsepp</td>
<td></td>
<td></td>
<td>5.06476</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>Stat</td>
<td></td>
<td></td>
<td>5.06633</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>Mahdi Mohebbian</td>
<td></td>
<td></td>
<td>5.13999</td>
<td>27</td>
</tr>
<tr>
<td>10</td>
<td>Friends</td>
<td></td>
<td></td>
<td>5.25699</td>
<td>52</td>
</tr>
</tbody>
</table>
Lecture 13 – Overview of machine learning

- Kaggle competition results
- **What we have covered in the course**
- What we have not covered in the course
- ICML sessions
- You and machine learning
Our definition of machine learning
[from Lecture 05]

Machine learning is the systematic study of algorithms and systems that improve their knowledge or performance with experience.
Machine learning is concerned with using the right features to build the right models that achieve the right tasks.

Getting all these main ingredients right is its main challenge.
Tasks in supervised learning

[from Lecture 05]

• **Supervised learning** is a general term about machine learning tasks where:
  – Training data are labelled objects
  – The task is to predict labels on test data

• Tasks in supervised learning:
  – **Classification** (when labels are categorical)
  – **Regression** (when labels are real-valued)
  – **Structured prediction** (when labels are something more complicated, for example)

• Terminology: target variable = label
Lecture 13 – Overview of machine learning

✓ Kaggle competition results
✓ What we have covered in the course
  • **What we have not covered in the course**
  • ICML sessions
  • You and machine learning
Unsupervised learning

• **Unsupervised learning** is learning from unlabeled data
• The goal is to learn structure from the data
Tasks in unsupervised learning

- **Clustering** – finding groups from the data
- **Anomaly detection** – finding unusual instances
- **Representation learning** – learning to represent the data in a better way (usually more compactly)
- **Density estimation** – how dense are the data in different parts of the instance space
- **Topic discovery** – find a way to describe each instance as covering one or several ‘topics’
More machine learning tasks

• **Semi-supervised learning** — many or most of the labels are missing in the training data

• **Active learning** — labels for training data can be ‘purchased’ one by one

• **One-shot learning** — learning from a single training instance

• **Reinforcement learning** — unlabeled training data, feedback given as rewards and punishments instead of labels

• These could all be considered as special (unusual) cases of supervised learning
Kevin Murphy’s ML textbook

• How much did we cover from the ML textbook by Kevin Murphy?
• Published in 2012
• 1070 pages
Murphy’s ML textbook

• 1 Introduction (+++), that means mostly
• 2 Probability (+, that means a bit)
• 3 Generative models for discrete data (+)
• 4 Gaussian models (+)
• 5 Bayesian statistics (+)
• 6 Frequentist statistics (+)
• 7 Linear regression (+)
• 8 Logistic regression (+)
• 9 Generalized linear models and the exponential family (-, that means almost not)
Murphy’s ML textbook

- 10 Directed graphical models (+)
- 11 Mixture models and the EM algorithm (-)
- 12 Latent linear models (-)
- 13 Sparse linear models (+)
- 14 Kernels (+)
- 15 Gaussian processes (-)
- 16 Adaptive basis function models (+)
- 17 Markov and hidden Markov models (+)
- 18 State space models (-)
Murphy’s ML textbook

- 19 Undirected graphical models (Markov random fields) (-)
- 20 Exact inference for graphical models (-)
- 21 Variational inference (-)
- 22 More variational inference (-)
- 23 Monte Carlo inference (-)
- 24 Markov chain Monte Carlo (MCMC) inference (-)
- 25 Clustering (-)
- 26 Graphical model structure learning (-)
- 27 Latent variable models for discrete data (-)
- 28 Deep learning (+)
Lecture 13 – Overview of machine learning

✓ Kaggle competition results
✓ What we have covered in the course
✓ What we have not covered in the course
  • ICML sessions
  • You and machine learning
Key approaches in modern ML

- Deep learning
- Probabilistic graphical modelling
- All other methods
Session titles in ICML 2017

- **ICML** –
  International Conference on Machine Learning
  (ICML and NIPS are the 2 best ML conferences)
- At ICML a session is a set of 4-5 talks
  - Each talk presents one paper
  - Usually gathered based on similarity
- **ICML 2017** – 91 sessions (over 400 papers)
  - 21 on deep learning
  - 12 on probabilistic modelling
  - 25 on various tasks in machine learning
  - 27 on theory and methods
  - 6 on applications
### ICML 2018 vs ICML 2017

- Topics of sessions on ICML 2017 and 2018
  - Note: somewhat subjective categorisation

<table>
<thead>
<tr>
<th>Topic</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep learning</td>
<td>21</td>
<td>30</td>
</tr>
<tr>
<td>Probabilistic modelling</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>Various tasks in machine learning</td>
<td>25</td>
<td>51</td>
</tr>
<tr>
<td>Theory and methods</td>
<td>27</td>
<td>33</td>
</tr>
<tr>
<td>Applications</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>91</strong></td>
<td><strong>140</strong></td>
</tr>
</tbody>
</table>
ICML 2017 sessions

- Deep learning (21 sessions):
  - 9x Deep learning
  - 4x Recurrent neural networks
  - 3x Deep generative models
  - 3x Deep learning theory
  - 2x Deep reinforcement learning
ICML 2017 sessions

• Probabilistic modelling (12 sessions):
  – 3x Probabilistic inference
  – 3x Probabilistic learning
  – 2x Causal Inference
  – 1x Bayesian Nonparametrics
  – 1x Bayesian Optimization
  – 1x Gaussian processes
  – 1x Latent feature models
ICML 2017 sessions

- Tasks (25 sessions):
  - 5x Reinforcement learning
  - 4x Online learning
  - 3x Supervised learning
  - 2x Clustering
  - 1x Active learning
  - 1x High dimensional estimation
  - 1x Learning To Learn
  - 1x Large scale learning
  - 1x Metalearning
  - 1x Metric learning
  - 1x Networks & relational learning
  - 1x Ranking and preferences
  - 1x Robust Estimation
  - 1x Structured prediction
  - 1x Transfer and multitask learning
ICML 2017 sessions

• Theory and methods (27 sessions):
  – 7x Continuous optimization
  – 3x Matrix factorization
  – 2x Combinatorial optimization
  – 2x Learning theory
  – 2x Monte Carlo methods
  – 2x Privacy and security
  – 2x Sparsity
  – 1x Continuous control
  – 1x Distributed optimization
  – 1x Ensemble methods
  – 1x Game theory and multiagents
  – 1x Information theory
  – 1x Kernel methods
  – 1x Spectral methods
ICML 2017 sessions

• Applications (6 sessions):
  – 3x Language
  – 1x Healthcare
  – 1x ML and programming
  – 1x Time series
Lecture 13 – Overview of machine learning

✓ Kaggle competition results
✓ What we have covered in the course
✓ What we have not covered in the course
✓ ICML sessions

• You and machine learning
12 related courses

- Fundamentals:
  - LTAT.02.XXX Machine Learning II (spring)
  - LTAT.02.001 Neural Networks (spring)
  - LTAT.02.003 Big Data Management (spring)
  - LTMS.00.025 Statistical Data Science (spring)
  - MTMS.01.100 Statistical Data Science and Visualization (spring)

- Seminars:
  - MTAT.03.317 Special Course in Machine Learning (autumn & spring)
  - MTAT.03.277 Research Seminar in Data Mining (autumn & spring)

- Applications:
  - LTAT.01.001 Natural language processing
  - MTAT.03.291 Introduction to Computational Neuroscience
  - MTAT.03.239 Bioinformatics
  - MTAT.08.043 Transportation Theory and Applications
  - MTAT.03.319 Business Data Analytics
Your thesis work on ML?

• Are you interested in doing your thesis in machine learning?
• Please write an e-mail to me
  – The earlier the better
  – Preferably already in spring / summer
• See also https://ml.cs.ut.ee/
My research interests

• Deep learning
• Loss measures
• Uncertainty and calibration
• Probabilistic modelling
• Dataset shift and transfer learning
• Ensemble methods and boosting
• Applications
Lecture 13 – Overview of machine learning

✓ Kaggle competition results
✓ What we have covered in the course
✓ What we have not covered in the course
✓ ICML sessions
✓ You and machine learning
Congratulations!
The final slide of the ML course...

Thank you and good luck!