Machine Learning 2
Ma Data Science
Winter Semester 2019/20

BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences

Prof. Tim Downie

Course Information

Course Overview: The principles and methods learnt in Machine Learning 1 will be continued in this course to more sophisticated supervised learning methods. The first half of the subjects covered are *Coping with missing values*, *non-linear regression* methods such as *spline smoothing* and *GAM*s, the *naive Bayes classifier* and *support vector machines*. The rest of the course will focus on artificial neural networks.

The workshops focus on implementing the methods learnt in that week lecture using the statistical software R.

Moodle-Site: All course material and news will be distributed via Moodle:
http://lms.beuth-hochschule.de/moodle/course/view.php?id=17296
Short name: ml2-wise1920

Lecturer: Prof. Tim Downie (TD) (FB II, Room A 126)

Email: tim.downie@beuth-hochschule.de

Office Hour: Please email TD to arrange a time.

Texts: As in Machine Learning I, you will be using James et al. as the main text book, which is available to download for free. For the more advanced subjects, you will use Haste et al. which also covers more technical details of the subjects covered in James.

James, Witten, Hastie and Tibshirani (2013) *An Introduction to Statistical Learning with Applications in R*.
Springer-Verlag.
`http://www-bcf.usc.edu/~gareth/ISL/index.html`

Hastie, Tibshirani and Friedman (2009)
The Elements of Statistical Learning (2nd edition). Springer-Verlag.
`https://web.stanford.edu/~hastie/ElemStatLearn/`

Baumer, Kaplan and Horton (2017) *Modern Data Science with R*. Chapman and Hall .

W.N. Venables and B.D. Ripley (2002). *Modern Applied Statistics with* S (fourth edition, Springer).

Timetable:
Lectures (SU): Wednesdays 14:15 in room A 324
Workshops (Ü): Wednesdays 16:00 in room A 328

Semester Outline:

Please see the PDF file in Moodle

Assessment for examination grading: The course will be evaluated by a project and a 90 minute exam. The project contributes 30 marks and the exam 70 marks. To pass the course you need to gain at least 45 out of 100 marks (exam plus project).

Project: You will work in small groups analysing a medium sized data set using supervised learning. Methods learnt in *Machine Learning 1* and the first part of *Machine Learning 2* will be relevant. Finding a suitable data set is a part of the project. You should check with the lecturer that the data are appropriate before analysing the data. The provisional time plan for the project is:

- 27th November – Project details given out.
- 11th December or before – details of your proposed data set and the students in your group to be submitted to the lecturer.
- 10th January 2019 – Deadline for submitting your work.

The exam in the 1st exam period will be on 22 January 2020 at 14:15 (provisional).

The exam in the 2nd exam period is on 29 March 2019 at 10:00 (provisional).

Information on all exam times and locations can be found at:
https://pruefungen.beuth-hochschule.de/M-DS

The following grading scheme will be used

Percentage	<45	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-100
Grade (Note)	5.0	4.0	3.7	3.3	3.0	2.7	2.3	2.0	1.7	1.3	1.0