Machine Learning for Physics and Astronomy

Juan Rojo
VU Amsterdam & Theory group, Nikhef

Natuur- en Sterrenkunde BSc (Joint Degree), Honours Track
Course overview, 03/09/2020

Course schedule

Week	Monday	Tuesday	Wednesday	Thursday
36 31/8 4/9				9-11 Hoorcollege dr. Juan Rojo Chacon Online
37 7/9 11/9	17-19 Hoorcollege dr. Juan Rojo Chacon Online			9-11 Werkcollege dr. Juan Rojo Chacon Online
38 14/9 18/9	17-19 Hoorcollege dr. Juan Rojo Chacon Online			9-11 Werkcollege dr. Juan Rojo Chacon Online
39 21/9 25/9	17-19 Hoorcollege dr. Juan Rojo Chacon Online			9-11 Werkcollege dr. Juan Rojo Chacon Online
40 28/9 2/10	17-19 Hoorcollege dr. Juan Rojo Chacon Online			9-11 Werkcollege dr. Juan Rojo Chacon Online
41 5/10 9/10	17-19 Hoorcollege dr. Juan Rojo Chacon Online			9-11 Werkcollege dr. Juan Rojo Chacon Online
42 12/10 16/10	17-19 Hoorcollege dr. Juan Rojo Chacon Online			9-11 Werkcollege dr. Juan Rojo Chacon Online
43 19/10 23/10		15-18 Tentamen Online		

- **7 lectures** (today and then Mondays 5pm-7pm) and **6 tutorials** (9am-11am on Thursdays)
- All course activities will be **online,** zoom meeting links available via Canvas. Lectures will be **recorded** and made available afterwards
- The **final presentations** used for the course assessment will take place on Tuesday 20/10 between 3pm and 6pm
- Additional discussion sessions can be scheduled upon request

Course logistics

- All the **course material** (slides, recordings, notebooks for the tutorials, other resources) will be made available via the **Canvas page of the course**, which will also be used for the **main announcements** as well as for the submission of **homework assignments**

https://canvas.uva.nl/courses/17140

- This material will also be available via the course **GitHub repository,** which will always contain the most updated versions of the tutorial's codes

https://github.com/LHCfitNikhef/ML4PA

note that this repo can be directly linked to e.g. **Google's Collab** to run the codes on the cloud rather than locally

- Further we have created a **Discord server** to streamline the communications between the instructors and the students, and to facilitate discussions during the lectures. You should all have received an invitation to join the course channel

Course evaluation

- Students should write a short report (around 4 pages) about a specific **application of Machine Learning algorithms** to a physics or astronomy problem that they find interesting. Including possible **code examples** is encouraged but not required. This report and the subsequent presentation count up to 70% of the course grade.

> You need to submit this report via Canvas by **Monday 19th** October

- **Presentation** of the contents of this report on **Tuesday 20th: 5** min + discussion

- In addition, after each tutorial we will propose some short **homework assignments** that need to be completed and submitted via Canvas. These homework assignments count up to 30% of the final course grade

> Homework assignments must be handed in **by the end of the Friday** following the tutorial (see Canvas deadlines)

References

the literature on **Machine Learning and their applications to physics** is vast.

When preparing these lectures the following resources have been used:

- *Pattern Recognition and Machine Learning*, Christopher M Bishop, Springer (2016).
- *Hands-On Machine Learning with Scikit-Learn and Tensorflow*, Geron Aurelien, O'Reilly Media (2017).
- *Introduction to Machine Learning with Python*, Andreas Muller, O'Reilly Media (2016).
- *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Trevor Hastie, Robert Tibshirani, Jerome Friedman, Springer, 2nd edition (2018), available online at https://web.stanford.edu/~hastie/Papers/ESLII.pdf.
- *An Introduction to Statistical Learning* , Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Springer, 7th edition (2018), available online at http://faculty.marshall.usc.edu/gareth-james/ISL/ISLR%20Seventh%20Printing.pdf
- *A high-bias, low-variance introduction to Machine Learning for physicists*, P. Mehta, M. Bukov, C. H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher and D. J. Schwab, Phys. Rept. **810**, 1 (2019) doi:10.1016/j.physrep.2019.03.001 [arXiv:1803.08823 [physics.comp-

Tutorials

the hands-on tutorials will allow you to familiarise with the machine learning concepts presented in the lectures by means of **practical examples**

- The **Python notebooks** with the course tutorials can be either run locally (if you have an updated python installation) or on the cloud by uploading the notebook in **Google Collab**
- Please note that some tutorials require the download of heavy datasets, make sure this is done before the tutorial
- Due to time constraints we can only offer limited assistance with **software installation** - please make sure you can install and execute the notebooks beforehand, and get back to us via Discord if there is any trouble
- The tutorial sessions will be coordinated by Maarten Both, maartenboth@outlook.com, who will be assisted by extra TAs in each session. Please contact him for any **questions related to the course tutorials**

The tutorials will include both **basic, intermediate,** and **advanced** topics concerning ML algorithms - it is up to you to decide how far you want to push!

Guest lectures

Lectures 4 to 7 will be shared between JR and a guest lecturer, with the aim to present state-of-the-art ML applications for physics and astronomy problems within our community

- Lecture 4 (and tutorial 3): **Dr. Antonia Rowlinson** (applications to Astronomy)
- Lecture 5 (and tutorial 4): **Dr. Christoph Weniger** (applications to Dark Matter)
- Lecture 6 (and tutorial 5): **Dr. Sascha Caron** (applications for High-Energy Physics)
- Lecture 7 (and tutorial 6): **Dr. Tristan Bereau** (applications to Condensed Matter)

the **corresponding tutorials** will then be based on material from the specific application presented in the guest lecture