Machine Learning for Physics and Astronomy

Juan Rojo
VU Amsterdam & Theory group, Nikhef

Natuur- en Sterrenkunde BSc (Joint Degree), Honours Track
Course overview, 08/09/2021

Course schedule

Week	Monday	Tuesday	Wednesday	Thursday
36 6/9 10/9		11-13 HC: Lecture dr. Juan Rojo Chacon SP G2.13	9-11 LC: Practical session dr. Juan Rojo Chacon SP B0.202	
37 13/9 17/9		11-13 HC: Lecture dr. Juan Rojo Chacon SP G2.13	9-11 LC: Practical session dr. Juan Rojo Chacon SP B0.202	
38 20/9 24/9		11-13 HC: Lecture dr. Juan Rojo Chacon SP G2.13	9-11 LC: Practical session dr. Juan Rojo Chacon SP B0.202	
39 27/9 1/10		11-13 HC: Lecture dr. Juan Rojo Chacon SP G2.13	9-11 LC: Practical session dr. Juan Rojo Chacon SP B0.202	
40 4/10 8/10		11-13 HC: Lecture dr. Juan Rojo Chacon SP G2.13	9-11 LC: Practical session dr. Juan Rojo Chacon SP B0.202	
41 11/10 15/10		11-13 HC: Lecture dr. Juan Rojo Chacon SP G2.13	9-11 LC: Practical session dr. Juan Rojo Chacon SP B0.202	
42 18/10 22/10		11-13 HC: Lecture dr. Juan Rojo Chacon SP G2.13	9-11 LC: Practical session dr. Juan Rojo Chacon SP B0.202	

- **7 lectures** (Tuesdays 11am) and **7 tutorials** (Thursdays 9am)
- All course activities will be **on campus**. Recordings of the zoom lectures of the **2020-2021 academic year** are available on Canvas.
- 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**

- 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

- Furthermore, 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, else get in touch

Course evaluation

- Students should write a short report (around 6 pages) about a specific **application of Machine Learning algorithms** to a physics or astronomy problem that they find interesting. This report and the subsequent presentation count up to **60%** of the course grade.

You need to submit this report via Canvas by **Monday 1st November**

the recorded presentations of the 2020/2021 cohort are available on Canvas for inspiration

- 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 **40%** of the final course grade

Homework assignments must be handed in **within one week** following the tutorial (see Canvas for specific 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 Tommaso Giani and Ryan van Mastrigt, the course TAs. Please contact then 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 2020/2021

In the previous year we had several guest lecturers presenting state-of-the-art ML applications for physics and astronomy problems within our community

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

Slides and recordings of these guest lectures are also available on Canvas: students are encouraged to watch them as extra study resources