**WDCC organises: Workshop Machine Learning,**

**Monday October 26, 13.00-17.30**

Are you working in Machine Learning? Do you expect Machine Learning to be part of your work in the near future? Would you like to meet colleagues working in Machine Learning, and discuss your expectations and experiences?

If you would answer any of these questions with a "yes", the workshop on Machine Learning on Monday, Oct 26, might be just the thing for you!

Four topics will be introduced by moderators, and will be followed by a discussion. In addition, details will be given about a Data Challenge, running in November and December, in which multidisciplinary teams will tackle a difficult case study. The challenge is generously sponsored by the Data-Driven & High Tech investment theme coordinated by the Wageningen Data Competence Centre, meaning that project hours will be provided for WR participants.

**Challenge:**

During the data challenge you will work in multidisciplinary teams of 5-6 people on a provided dataset and a question with coaching from Ron, Paul, Rutger and Julian (max. 25 participants). The challenge takes place between 2 Nov and 18 Dec. You can sign up for the data challenge until **Friday 30 October** by sending an email with your motivation to [data@wur.nl](mailto:data@wur.nl). Include your (1) knowledge level, (2) learning/teaching goals and (3) application and dissemination opportunities. On 2 November the applicants are informed whether they are selected. They will then recieve the group compositions and location of the dataset. The selected participants will receive financial compensation for 24 working hours to work on the challenge. Applying to the challenge implies that you have permission from your supervisor to spend this time in this time frame.

**How to participate to the first Teams Meeting:**

The workshop itself is free of charge, and there is no need to register. Simply click on the Teams link below to enter at any point in time that suits you. The schedule of the afternoon is as follows:

[Join Microsoft Teams Meeting](https://eur03.safelinks.protection.outlook.com/ap/t-59584e83/?url=https%3A%2F%2Fteams.microsoft.com%2Fl%2Fmeetup-join%2F19%253ameeting_ZDRmMDkwZGQtMWJmMy00ZTExLWI5MTctODZkYjk5MjhlZTEy%2540thread.v2%2F0%3Fcontext%3D%257b%2522Tid%2522%253a%252227d137e5-761f-4dc1-af88-d26430abb18f%2522%252c%2522Oid%2522%253a%252286ac3cff-0402-4fc6-9f27-8b75732515ef%2522%257d&data=04%7C01%7C%7C5028da17ec4549f05c4d08d870181780%7C27d137e5761f4dc1af88d26430abb18f%7C0%7C0%7C637382596062137359%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=Fny6gCD7kDmzIh8aC8U6KPRSdnLW6vqKtJKf4q1lJyM%3D&reserved=0)

**Programme:**

13.00 – 14.00: Introduction of the afternoon, and topic 1:

Overview of Machine Learning methods

(Rutger Vlek)

14.00 – 15.00: Topic 2: Validation and tuning of ML methods

(Ron Wehrens)

15.00 – 16.00: Topic 3: Preprocessing, visualisation and quality

of data (Paul Van Leeuwen)

16.00 – 17.00: Topic 4: Incorporating prior knowledge

(Paul van Leeuwen and Rutger Vlek)

17.00 – 17.30: Introduction of the Data Challenge

Feel free to contact the organisers, Paul van Leeuwen ([paul2.vanleeuwen@wur.nl](mailto:paul2.vanleeuwen@wur.nl)), Rutger Vlek ([rutger.vlek@wur.nl](mailto:rutger.vlek@wur.nl)) and Ron Wehrens ([ron.wehrens@wur.nl](mailto:ron.wehrens@wur.nl)) with your comments and suggestions.

**We hope to see you on the 26th!**

(Rutger)

Machine learning is a hot, yet complex topic. This lecture will warm you up for the more technical lectures ahead. We provide a brief overview of artificial intelligence (AI), machine learning (ML), the most important variations of ML-algorithms and the steps towards ML applications. But wait... there's more to a successful application than just the algorithm! Let's have a look at *who* uses it and *how* it is used.

(Ron)

With an almost infinite number of possibilities, it is hard to choose which Machine Learning method to use. The simple solution - use the best one! - is not so simple: all methods have their pros and cons, and it is a priori hard to say which methods will do well in a particular situation. Measuring the performance is a problem in itself, even. This talk will highlight approaches to assess prediction quality in machine learning, discuss pitfalls, and suggest some rules of thumb useful in navigating the Machine Learning seas...

(Paul)

Preprocessing your data is crucial before applying ML techniques since garbage in = garbage out. During this workshop we will discuss how to prepare your data for supervised learning algorithms. Preprocessing typically serves two goals: (1) to ensure that the data fed to your learning algorithm is representative for the learning problem, (2) to optimize the performance of your learning algorithm. This workshop touches on topics such as: dealing with missing data, feature selection, feature engineering, outlier detection.

         (Rutger en Paul)

Once you have a dataset and a collection of machine learning algorithms, you seem all set to solve a problem with the magic trick of machine learning, right? Not really!

In many cases you may be missing out on relevant prior knowledge that (often implicitly) exists about the data or the problem that you try to solve. This lecture aims to inspire you to discover relevant knowledge and it suggests several ways to make use of it in your algorithms.