Course subject: **Machine learning in Astronomy (PhD course for 7.5 ECTS)**
Course advisor: **Rahul Biswas,** Stockholm University
Local course evaluator: **Matthew Hayes**

Course overview:
The first part of the course is a self-study part, where the students will study the textbook and work with some practical problems. After this part, students should have some familiarity with a range of machine learning algorithms, and some basic understanding of how to implement them to solve real astronomical problems. The second part of the course is a project, where students will pick one problem and apply machine learning algorithm to it. The aim of the project part is to gain a deeper understanding for the selected algorithm, and the kind of real-world challenges one faces when implementing it.

Course material:

- Text book: **Statistics, Data Mining, and Machine Learning in Astronomy** by Zeljko Ivezic, Andrew Connolly, Jacob VanderPlas, and Alex Gray (2014)
- Online material: http://www.astroml.org/

Requirements to "pass" the course:

- Presence in hands-on sessions: half of the total hands-on sessions, 3
- Presence in lectures: ⅔ of total lectures, 6
- Give a presentation/lead a lecture session: (number will depend on number of students)
- Final project (written report + presentation)
- Peer-review evaluation

Parts of the course:

- Lectures ( 9 X 2hr) : Everyone needs to read the material before the session. 1 or 2 student(s) will research the topic more thoroughly and give a presentation and/or lead the discussion about the topic.
    - Chapter 1 (Introduction to the topic covering this chapter and his own research, by Rahul)
    - Chapter 2 & 3
    - Chapter 4
    - Chapter 5
    - Chapter 6
    - Chapter 7
    - Chapter 8
    - Chapter 9
    - Chapter 10
- Hands-on sessions (6 X 2hr): Remaking the book figures. 1 or 2 student(s) will research the topic more thoroughly (preferably including more examples/exercises). Everyone gets to explore the methods introduced in the corresponding lecture through the piece of code available in the book. (Each hands-on session has a mini-goal, such using a similar

method on a different dataset (in cases where relevant pre-processed data sets are available), or “discovering" some abilities of the method/function that was not discussed in the book. This differs a lot from one problem to another.)
    - Chapter 4 & 5
    - Chapter 6
    - Chapter 7
    - Chapter 8
    - Chapter 9
    - Chapter 10
- Projects (individually or in small groups)
- Project evaluation (Peer review + Advisor review)
- Final project presentation (2 X 2hr): Depending on the number of students and how many projects there are
- Machine learning and statistics in research (A chance to invite some researchers to give short talks about the use of machine learning and statistics in their field of research, introducing more advanced uses of ML)
    - Daniel Mortlock (lecture on Bayesian statistics in research)
    - Suggested speakers: Will be suggested by course participants

Project criteria:

- Select a subject related to Astronomy/Astrophysics, and find a suitable dataset.
- Formulate the problem in machine-learning terminology. Identify training examples, features, labels, cost function.
- Select at least one method discussed in the book, discuss its pros and cons for the problem at hand, and apply it to the data.
- Use methods discussed in the course to tune model parameters and get generalizable results.
- Evaluate results and suggest future improvements.

Timeline:

- Average one lecture/hands-on session per week, starting end of January.
- Plans for projects need to be finalized by mid-March (topic and data set)
- Deadline for project reports: Tentatively 14/5
- Deadline for peer-reviews: (Deadline for project reports + 1 Week)
- Project presentation: Last two weeks of May
- ML in research talk session, dependent on organization of speakers