Swedish University of Agricultural Sciences

P000137, Machine Learning for Agriculture and Natural Sciences, 7.5 Hp

Syllabus

Finalized by: Finalized by: FUN-VH, 2025-04-08, 2025-04-08

Level within study regulation:

Third cycle

Grading scale:

Pass / Failed

Course language:

Swedish

Entry requirements:

Admitted to a postgraduate program in animal science, biology, veterinary medicine, food science, nutrition, nursing, or related subjects, or to a residency program in veterinary science. Alternatively, admitted to a postgraduate program in Bioinformatics or Biological field with a significant computational aspect. VMAS course Statistics I: Basic Statistics or similar (including regression). Introduction to Python for data science, PVG0045, or equivalent Python coursework or experience. If you want clarification on what constitutes sufficient Python knowledge or advice on how to obtain it, please contact the course coordinator.

Objectives:

On completion of the course, the student should be able to:

1. *Understand basics of statistical machine learning, including variable selection, cross-validation, and supervised and unsupervised learning.*
2. *Understand basic concepts of Deep Learning and its historical development, including perceptrons, back-propagation, and under/over training. *
3. *Understand and select input/output encoding, activation function, as well as the types of layers. *
4. *Solve problems in binary and multi-class classification, especially in image analysis.*
5. *Apply the below listed concepts to important problems in Natural Sciences *

Content:

Supervised Learning Applications in Agriculture and Natural Sciences. Explore the fundamentals of supervised learning algorithms tailored for agricultural andnatural science datasets. Implement and interpret models such as logistic regression, linear regression, regularization and shrinkage techniques, non-linear regression, discriminant analysis, regression trees, random forests, Gradient Boosting, and support vector machines.

Cross Validation Techniques: Understand the importance and application of cross-validation methods to assess model performance and generalization ability. Effectively implement and interpret cross-validation techniques to optimize model parameters and mitigate overfitting.

Variable Selection in High-Dimensional Data: Explore strategies for variable selection in machine learning models, particularly in the context of high-dimensional datasets common in agriculture and natural sciences. Identifying and prioritize relevant features to enhance model accuracy and interpretability.

Unsupervised Learning Techniques: Investigate unsupervised learning methods for extracting meaningful insights from unlabelled data in agriculture and natural sciences. Understand and apply principal component analysis (PCA) for dimensionality reduction and visualization, as well as various clustering algorithms including K-means, K-Medoids, DBSCAN, and Clara

Review of Python in preparation for Deep Learning portion of course.

The XOR problem, and its relation to hidden layers. Back-propagation: how it works, and how it relates to your choice of activation function.

Regularization, Frameworks, Regression, Convolutional layers and their role in image classification.

UNet and ResNet.

In person:

Course days will begin with a theory discussion (1-2 hours or as needed). The rest of the day will consist of laboratory exercises.

Labs early in the week will cover specific concepts, while labs later in the week will be directed towards a practical problems, e.g. gene selection, image classification. For the lab exercises you will turn in code and answer a quiz. Evaluation will be based on conceptual understanding and appropriate of application of the learned techniques. There will be time scheduled on the last in-person day to kick off the post-course assignment.

As a post-course assignment, the student may complete a novel task, e.g. in image classification. A specific problem will be suggested, or alternatively come from the student's research. A report and code will be turned in.

Grading of this will be based on fulfilment of objectives, e.g. use of specific techniques, as well as appropriateness of the approach. While some degree of predictive accuracy is desired, demonstration of conceptual knowledge is more important.

Modes of assessment:

Examination will be primarily by assignments to be turned in.

In these assignments, the student will implement methods, turn in code, and critically analyze results. The assignments will be in written or quiz form. Evaluation will be based on theoretical understanding, technical performance, and ability to interpret results.

There will also be quiz questions testing theoretical knowledge. - If a student has failed an examination, the examiner has the right to issue supplementary assignments. This applies if it is possible and there are grounds to do so.

- The examiner can provide an adapted assessment to students entitled to study support for students with disabilities following a decision by the university. Examiners may also issue an adapted examination or provide an alternative way for the students to take the exam.
- If this syllabus is withdrawn, SLU may introduce transitional provisions for examining students admitted based on this syllabus and who have not yet passed the course.
- For the assessment of an independent project (degree project), the examiner may also allow a student to add supplemental information after the deadline for submission. Read more in the Education Planning and Administration Handbook.

Organisation:

Department of Animal Biosciences

Supplementary information

Other information:

- The right to participate in teaching and/or supervision only applies for the course instance the student was admitted to and registered on.
- If there are special reasons, students are entitled to participate in components with compulsory attendance when the course is given again. Read more in the Education Planning and Administration Handbook.