Statistical Learning in Practice (L24)

Alberto J. Coca

Statistical learning is the process of using data to guide the construction and selection of models, which are then used to predict future outcomes. In this course, consisting of roughly 12 lectures and 12 practical classes, we will examine some of the most successful and widely used statistical methodologies in modern applications. The practical classes will deal with an introduction to R, exploratory data analysis and the implementation of the statistical methods discussed in the lectures. We aim to cover a selection of the following topics:

- Generalised linear models for regression and classification

- Model selection and regularisation

- Mixed effects models and quasi-likelihood methods

- Linear discriminant analysis and support vector machines

- Introduction to neural networks

- Introduction to time series

Pre-requisites

Elementary probability theory. Maximum likelihood estimation, hypothesis tests and confidence intervals. Linear models.

Previous experience with R is helpful but not essential.

Literature

1. Agresti, A. (2015) *Foundations of Linear and Generalized Linear Models.* Wiley.

2. Brockwell, P.J. and Davis, R.A. (1996) *Introduction to time series and forecasting.* Springer.

3. Dobson, A.J. and Barnett A. (2008) *An Introduction to Generalized Linear Models.* Third edition. Chapman & Hall/CRC.

4. Faraway, J. J. (2005) *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models.* CRC press.

5. Hastie, T., Tibshirani, R. and Friedman, J. (2009) *The Elements of Statistical Learning.* Second Edition. Springer.

6. Shumway, R. H., and Stoffer, D. S. (2010) *Time Series Analysis and Its Applications: with R Examples.* Springer Science & Business Media.

Additional support

This course includes practical classes, in which statistical methods are introduced in a practical context and students carry out analysis of datasets using R. In the practical classes, the students have the opportunity to discuss statistical questions with the lecturer. Four examples sheets will be provided and there will be four associated examples classes (the last one probably to be held in Easter term). There will be a revision class in the Easter Term.