Module Description, available in: EN

Analysis of Sequential Data

General Information

Number of ECTS Credits

3

Module code

TSM_AnSeqDa

Valid for academic year

2019-2020

Last modification

2018-10-17

Responsible of module

Giorgio Corani (SUPSI, giorgio@idsia.ch)

Explanations regarding the language definitions for each location:

- Instruction is given in the language defined below for each location/each time the module is held.
- Documentation is available in the languages defined below. Where documents are in several languages, the percentage distribution is shown (100% = all the documentation).
- The examination is available 100% in the languages shown for each location/each time it is held.

	Berne	Lausanne			Lugano	Zurich		
Instruction						**X** E 100%		
Documentation						**X** E 100%		
Examination						**X** E 100%		

Module Category

TSM Technical/scientific specialization module

Lessons

2 lecture periods and 1 tutorial period per week

Entry level competences

Prerequisites, previous knowledge

Basic knowledge in statistics and signal processing.
Programming skills (preferably R and Python).

Brief course description of module objectives and content

Many data sets are temporal by nature, i.e. time series. Typical tasks faced by data scientists are: analyzing time series in different domains and developing statistical models based on the data, in order to forecast future values or classify the time series into predefined categories.

This course provides a comprehensive introduction to analysis, forecasting and classification of sequential data. The course adopts a practical approach: theoretical concepts are illustrated and applied in specific case studies. Students will also learn to identify the tools best suited for a given task.

The first part of the course presents techniques for analysis of time series. It starts from visualization techniques; then it shows techniques for characterizing the time series; eventually it present more sophisticated modeling approaches that account for trend components and seasonal components. Several examples referring to real data sets are shown.

In the second part of the course students learn how to analyze digital signals in different domains, i.e. time and spectral domain; they learn how to extract meaningful features from digital signals suitable for classification. Finally, they learn how to set up and learn statistical models, such as HMMs or DNNs, for recognizing and classifying time series.

Aims, content, methods

Learning objectives and acquired competencies

- Students know how to visualize time series and how to characterize their main features.
- Students know how to evaluate forecast accuracy.
- Students know how to model trends, seasonalities and non-stationarities adopting exponential smoothing and ARIMA models.
- Students know how to perform model estimation, model selection and probabilistic prediction with these models.
- Students know different methods to analyse digital signals in different domains
- Students know how to extract important features used in speech processing
- Students learn to apply Bayes rule for classifying digital signals.
- Students can apply modern deep learning approaches to classify digital signals

Contents of module with emphasis on teaching content

Part 1: Forecasting sequential data (7 weeks)

- Time series graphics.
- Main features of time series.
- Assessment of the predictions.
- Exponential smoothing
- ARIMA models

Part 2: Analysis and classification of digital signals (7 weeks)

- Analysis of digital signals in different domains (1 week)
- Feature extraction (1 week)
- Modelling, classification & recognition of digital signals (5 weeks)
    - Classic Approaches: Dynamic Time Warping, Vector Quantization
    - Statistical modelling: Hidden Markov Models
    - Deep Learning Approaches

The contents are illustrated and applied in practical case studies using computational environments such as R or Python.

Teaching and learning methods

- Ex cathedra
- Self study of literature / publications
- Practical exercises with computer

Literature

Slides will be available covering the topics of the course.

In addition, recommended books are:
R. Hyndman and G. Athanasopoulos., Forecasting: Principles and Practice, Springer, 2018 (online free textbook at https://otexts.org/fpp2/)

X. Huang, A. Acero, H.-W. Hon: Spoken Language Processing, Prentice Hall, 2001, ISBN 0-13-22616-5

L. R. Rabiner und B.-H. Juang, Fundamentals of Speech Recognition. Prentice Hall, 1993.

D. Yu und L. Deng, Automatic Speech Recognition: A Deep Learning Approach. Springer London, 2014.

Assessment

Certification requirements

Module does not use certification requirements

Basic principle for exams

As a rule, all the standard final exams for modules and also all repetition exams are to be in written form

Standard final exam for a module and written repetition exam

Kind of exam

written

Duration of exam

120 minutes

Permissible aids

Aids permitted as specified below:

Permissible electronic aids

Scientific calculator

Other permissible aids

Handwritten summary pages

Special case: Repetition exam as oral exam

Kind of exam

oral

Duration of exam

30 minutes

Permissible aids

No aids permitted