270209 - TEOI - Information Theory

Coordinating unit: 270 - FIB - Barcelona School of Informatics
Teaching unit: 749 - MAT - Department of Mathematics
Academic year: 2018
Degree: BACHELOR'S DEGREE IN DATA SCIENCE AND ENGINEERING (Syllabus 2017). (Teaching unit Compulsory)
ECTS credits: 6  Teaching languages: Catalan

Prior skills
Fundamentals of probability, statistics and stochastic processes

Degree competences to which the subject contributes

Specific:
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE7. Demonstrate knowledge and ability to apply the necessary tools for the storage, processing and access to data.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.

General:
CG5. To be able to draw on fundamental knowledge and sound work methodologies acquired during the studies to adapt to the new technological scenarios of the future.

Transversal:
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.

Teaching methodology
50% of lectures in which the participation of students is stimulated, followed by 50% of practical classes based on exercises and programming of algorithms with the aim of bringing information theory to practical applications related to data science engineering.

Learning objectives of the subject
1. To acquire the knowledge necessary to understand the basic principles of treatment, compression, cryptography and analysis of data based on Shannon's theory.

Study load

<table>
<thead>
<tr>
<th>Total learning time: 150h</th>
<th>Hours large group: 60h</th>
<th>40.00%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guided activities: 6h</td>
<td></td>
<td>4.00%</td>
</tr>
<tr>
<td>Self study: 84h</td>
<td></td>
<td>56.00%</td>
</tr>
</tbody>
</table>
## Content

### Discrete random variables and processes

<table>
<thead>
<tr>
<th>Degree competences to which the content contributes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description:</td>
</tr>
<tr>
<td>Probability, ensembles of random variables, stochastic processes, Márkov processes</td>
</tr>
</tbody>
</table>

### Measures of information

<table>
<thead>
<tr>
<th>Degree competences to which the content contributes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description:</td>
</tr>
<tr>
<td>Information theory, entropy, joint entropy and mutual information, data processing inequality, Fano's inequality, applications</td>
</tr>
</tbody>
</table>

### Information of data sources

<table>
<thead>
<tr>
<th>Degree competences to which the content contributes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description:</td>
</tr>
<tr>
<td>Codes, asymptotic equipartition property, data compression, the high probability set, non-independent sources</td>
</tr>
</tbody>
</table>

### Source coding

<table>
<thead>
<tr>
<th>Degree competences to which the content contributes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description:</td>
</tr>
<tr>
<td>Properties of codes, unique decodability, minimum average length, Huffman codes, dictionary codes</td>
</tr>
</tbody>
</table>

### Capacity of discrete channels

<table>
<thead>
<tr>
<th>Degree competences to which the content contributes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description:</td>
</tr>
<tr>
<td>Joint typical sequences, channel capacity theorem, separability of source and channel coding.</td>
</tr>
</tbody>
</table>

### Channel codes

<table>
<thead>
<tr>
<th>Degree competences to which the content contributes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description:</td>
</tr>
<tr>
<td>Introduction to error correction codes, block codes</td>
</tr>
</tbody>
</table>
## Capacity of real channels

**Degree competences to which the content contributes:**

**Description:**
Differential entropy, capacity of Gaussian channels, bandlimited channels

## Statistical Inference

**Degree competences to which the content contributes:**

**Description:**
Law of large numbers and large deviation theory; hypothesis testing; estimation and sufficient statistics; parameter estimation criteria: maximum entropy, maximum mutual information, cross-entropy optimization versus classification error; Fischer information and the Cramer-Rao inequality.

## Cryptography

**Degree competences to which the content contributes:**

**Description:**
Shannon theory of secrecy systems; main theorem; one-time pad; symmetric cryptography in practice
### Planning of activities

| Development of lecture | "Discrete random variables and processes" | **Hours:** 8h  
Theory classes: 2h  
Practical classes: 2h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 4h |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific objectives:</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

| Development of lecture | "Measures of information"               | **Hours:** 16h  
Theory classes: 4h  
Practical classes: 4h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 8h |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific objectives:</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

| Development of lecture | "Information of data sources"           | **Hours:** 8h  
Theory classes: 2h  
Practical classes: 2h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 4h |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific objectives:</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

| Development of lecture | "Source coding"                         | **Hours:** 16h  
Theory classes: 4h  
Practical classes: 4h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 8h |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific objectives:</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
## Development of lecture "Capacity of discrete channels"

**Hours:** 8h  
Theory classes: 2h  
Practical classes: 2h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 4h  

**Specific objectives:**  
1

## Development of lecture "Channel codes"

**Hours:** 8h  
Theory classes: 2h  
Practical classes: 2h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 4h  

**Specific objectives:**  
1

## Development of lecture "Capacity of real channels"

**Hours:** 16h  
Theory classes: 4h  
Practical classes: 4h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 8h  

**Specific objectives:**  
1

## Development of lecture "Statistical inference"

**Hours:** 16h  
Theory classes: 4h  
Practical classes: 4h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 8h  

**Specific objectives:**  
1
Development of lecture "Cryptography"

<table>
<thead>
<tr>
<th>Hours: 8h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory classes: 2h</td>
</tr>
<tr>
<td>Practical classes: 2h</td>
</tr>
<tr>
<td>Laboratory classes: 0h</td>
</tr>
<tr>
<td>Guided activities: 0h</td>
</tr>
<tr>
<td>Self study: 4h</td>
</tr>
</tbody>
</table>

Specific objectives:
1

Qualification system

There will be a mid-term test of two hours duration at the 8th week and a final exam. The grade is calculated as the maximum of (final exam grade, 0.6 * final exam grade + 0.4 * mid-term exam grade).

The re-evaluation exam, for fails who have attended lectures and final exam, will consist of one exam to be held in July and that will be considered at 100% for the final grading.

Bibliography

Basic:

Complementary: