ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

93000945 - Signal processing for big data

DEGREE PROGRAMME

09AT - Master Universitario En Teoria De La Señal Y Comunicaciones

ACADEMIC YEAR & SEMESTER

2018/19 - Semester 2

Index

Learning guide

1. Description

1.1. Subject details

Name of the subject	93000945 - Signal processing for big data
No of credits	4 ECTS
Type	Optional
Academic year ot the programme	First year
Semester of tuition	Semester 2
Tuition period	February-June
Tuition languages	English
Degree programme	09AT - Master universitario en teoria de la señal y comunicaciones
Centre	09 - Escuela Tecnica Superior de Ingenieros de Telecomunicacion
Academic year	2018-19

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Jesus Gustavo Cuevas Del Rio (Subject coordinator)	B-409	gustavo.cuevas@upm.es	Sin horario. Appointment arranged by email

* The tutoring schedule is indicative and subject to possible changes. Please check tutoring times with the faculty member in charge.

3. Prior knowledge recommended to take the subject

3.1. Recommended (passed) subjects

El plan de estudios Master Universitario en Teoria de la Señal y Comunicaciones no tiene definidas asignaturas previas recomendadas para esta asignatura.

3.2. Other recommended learning outcomes

- Numerical Analysis

- Digital Signal Processing

- Matrix Algebra

- Linear Systems

4. Skills and learning outcomes *

4.1. Skills to be learned

CB06 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación

CB07 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio

CB08 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

CB09 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo

CE02 - Evaluar y sintetizar los resultados de un trabajo en equipo en proyectos relacionados con la teoría de la señal y las comunicaciones, en un entorno internacional.

CE03 - Valorar y contrastar la utilización de las diferentes técnicas disponibles para la resolución de problemas reales dentro del área de teoría de la señal y comunicaciones.

CT01 - Capacidad para comprender los contenidos de clases magistrales, conferencias y seminarios en lengua inglesa

CT03 - Capacidad para adoptar soluciones creativas que satisfagan adecuadamente las diferentes necesidades planteadas

CT04 - Capacidad para trabajar de forma efectiva como individuo, organizando y planificando su propio trabajo, de forma independiente o como miembro de un equipo

CT05 - Capacidad para gestionar la información, identificando las fuentes necesarias, los principales tipos de documentos técnicos y científicos, de una manera adecuada y eficiente

4.2. Learning outcomes

RA50 - Capability of dealing and characterizing graphs

RA51 - Capability of analysing and operating with signals defined on graphs

RA52 - - Capability for planning, design and implement applications, incorporating signal processing of signals defined on graphs

RA49 - Capability of visualize data by Tensor Networks

RA48 - Capability to understand, store data, manage, and reduce size of Tensors

RA26 - Ability of oral and written communication

* The Learning Guides should reflect the Skills and Learning Outcomes in the same way as indicated in the Degree Verification Memory. For this reason, they have not been translated into English and appear in Spanish.

5. Brief description of the subject and syllabus

5.1. Brief description of the subject

This course addresses the challenges of signal processing techniques when applied to large-scale data. The course presents processing algorithms suitable for large-scale data tasks involving sparse signals as the Sparse Fourier transform. Other introductory topics in the course are the extension of classical signal processing on data indexed by graphs (discrete signal processing on graphs, DSPG). DSPG deals with the analysis and processing of data sets in which data elements are related by dependency, similarity, physical proximity, or other properties.

Another topic is the generalization of data matrices into Tensors. Tensors are data structures indexed by more than two indices. Tensors here refer to what was previously known as multiway arrays. Tensors are becoming increasingly important in analysing Big Data. As a result of their many indices, tensors easily become very big, causing memory and storage problems. The idea is to break the data in smaller boxes, and one of the strategies consists on compressing tensors by a low rank tensor decomposition applied to both: dense or sparse tensors. At the end of all chapters, some illustrative application scenarios are given.

The course also present some innovative topics in a chapter devoted to them. The sampling acquisition theory is reviewed in the context of new Big Data scenarios, as those based on the discovery that sparsity or compressibility can be exploited when acquiring large-scale signals (Compressive Sensing) such as the Sparse Fourier transform that only uses a subset of the input data at a time.

At the end of each chapter illustrative examples with their respective application scenarios, either PYTHON language or in MATLAB, are provided

5.2. Syllabus

1. Numerical techniques for small and moderate data sets

    1.1. Matrix storage versus tensor or graph storage.

    1.2. Principal Component Analysis (PCA).

    1.3. Singular Value Decomposition (SVD).

    1.4. Applications.

2. Tensors. Tensors Decompositions.

    2.1. Introduction to Tensors.

        2.1.1. Covariant, contravariant and mixed tensors. Multiway arrays and tensors.

        2.1.2. Tensor Algebra: operations on tensors. Tensors product

    2.2. Tensor decompositions and applications

        2.2.1. 2.2.1. Relation between Rank of a tensor and its decomposition.

        2.2.2. Tucker Decomposition (HOPCA).

        2.2.3. HOSVD Decomposition.

        2.2.4. Canonical Polyadic Decomposition (CP).

3. Signal Processing on Graphs.

    3.1. Introduction to Graphs and Signal on Graphs.

        3.1.1. Rudiments of graph theory and algorithms on graphs

    3.2. Discrete Signal Processing on Graphs (DSPG).

        3.2.1. Extension of basic DSP operations to Graphs: shifting, filtering, subsampling, Z transform and Fourier Transform, etc.

        3.2.2. Compresssion of signals on graphs.

        3.2.3. Applications.

    3.3. Digital Signal Processing on Graphs: Frequency Analysis.

        3.3.1. Fourier transform on Graphs: Spectrum of a Graph.

        3.3.2. Alternative basis.

        3.3.3. Frequency response of a filter on a graph. Ordering of frequencies.

        3.3.4. Filters on graphs design.

3.3.5. Applications.

3.4. Advanced Topics.

3.4.1. Strategies to fastening the Fourier transform: Sparse Fourier Transform on Graphs and Fast Fourier Transform on Graphs.

6. Schedule

6.1. Subject schedule*

Week	Face-to-face classroom activities	Face-to-face laboratory activities	Other face-to-face activities	Assessment activities
1	**Introduction to Signal Processing for Big Data. Ways of storing Big Data.** Duration: 02:00 Lecture	**Case study related to chapter 1** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 1** Problem-solving test Continuous assessment Duration: 01:00
2	**Introduction to Signal Processing for Big Data. Examples of dealing with small and moderate amounts of data. Principal Component Analysis (PCA).** Duration: 02:00 Lecture	**Case study related to chapter 1** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 1** Problem-solving test Continuous assessment Duration: 01:00
3	**Introduction to Signal Processing for Big Data. Examples of dealing with small and moderate amounts of data. Singular Value Decomposition (SVD).** Duration: 02:00 Lecture	**Case study related to chapter 1** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 1** Problem-solving test Continuous assessment Duration: 01:00
4	**Tensor Analysis. Introduction to Tensors. Definitions. Covariant and contravariant tensors.** Duration: 04:00 Lecture	**Case study related to chapter 2** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 2** Problem-solving test Continuous assessment Duration: 01:00
5	**Tensor Analysis. Introduction to Tensors. Tensor Algebra. Other Tensor operations. Applications.** Duration: 02:00 Lecture	**Case study related to chapter 2** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 2** Problem-solving test Continuous assessment Duration: 01:00
6	**Tensor Analysis. Tensor decomposition. Rank of a Tensor as related to its decomposition. Tucker Decomposition of a Tensor.** Duration: 02:00 Lecture	**Case study related to chapter 2** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 2** Problem-solving test Continuous assessment Duration: 01:00
7	**Tensor Analysis. Tensor decomposition. Rank of a Tensor as related to its decomposition. HOSVD and Polyadic decomposition.** Duration: 02:00 Lecture	**Case study related to chapter 2** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 2** Problem-solving test Continuous assessment Duration: 01:00
8	**Tensor analysis. Modelling a problem by a tensor decomposition. Analysing the element of the decomposition.** Duration: 02:00 Lecture	**Case study related to chapter 2** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 2** Online test Continuous assessment Duration: 01:00

9	**Introduction to graphs. Graphs characterization. Main parameter of a graph.** Duration: 02:00 Lecture	**Case study related to chapter 2** Duration: 01:00 Laboratory assignments		**Mid term exam corresponding to chapters 1-2. The student will have to solve a problem equivalent to the laboratory activities** Problem-solving test Continuous assessment Duration: 01:00
10	**Signals on Graphs. Time-domain basic operations on signals on graphs. Shifting of a signal, Filtering, subsampling, Strategies to define a Fourier transform and a Z-transform of a signal on a graph** Duration: 02:00 Lecture	**Case study related to chapter 3** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 3** Problem-solving test Continuous assessment Duration: 01:00
11	**Signals on graphs. Frequency Analysis. Spectrum of a Graph. Alternative basis. Frequency response of a Graph.** Duration: 02:00 Lecture	**Case study related to chapter 3** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 3** Problem-solving test Continuous assessment Duration: 01:00
12	**Signals on graphs. Frequency Analysis. Frequency-domain design of filters on graphs.** Duration: 02:00 Lecture	**Case study related to chapter 3** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 3** Problem-solving test Continuous assessment Duration: 01:00
13	**Signals on graphs. Improving the Fourier Transform on a Graph. Sparse Fourier transform and other strategies. Chapter Applications.** Duration: 02:00 Lecture	**Case study related to chapter 3** Duration: 01:00 Laboratory assignments		**Programming work report related to chapter 3** Problem-solving test Continuous assessment Duration: 01:00
14				
15				
16				
17				**Second exam, corresponding to Chapter 3 contents. The student will have to solve a problem equivalent to the laboratory activities** Problem-solving test Continuous assessment Duration: 01:00 **Final exam. The student will have to solve a problem equivalent to the laboratory activities. All the chapters will be covered.** Problem-solving test Final examination Duration: 02:00 **10% of the final mark will be assigned taking into account criteria such as: assisting to lessons, attention, attitude, ability to answer questions in class, preparation in advance of lessons, etc.** Other assessment Continuous assessment Duration: 00:00

The independent study hours are training activities during which students should spend time on individual study or individual assignments.

Depending on the programme study plan, total values will be calculated according to the ECTS credit unit as 26/27 hours of student face-to-face contact and independent study time.

* The subject schedule is based on a previous theorical planning of the subject plan and might go to through experience some unexpected changes along throughout the academic year.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Continuous assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
1	Programming work report related to chapter 1	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10
2	Programming work report related to chapter 1	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10
3	Programming work report related to chapter 1	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10

4	Programming work report related to chapter 2	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10
5	Programming work report related to chapter 2	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10
6	Programming work report related to chapter 2	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10
7	Programming work report related to chapter 2	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10
8	Programming work report related to chapter 2	Online test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CE02 CT04 CE03 CT05

							CB10
9	Mid term exam corresponding to chapters 1-2. The student will have to solve a problem equivalent to the laboratory activities	Problem-solving test	Face-to-face	01:00	30%	4 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10
10	Programming work report related to chapter 3	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10
11	Programming work report related to chapter 3	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CE02 CE03 CT05 CB10
12	Programming work report related to chapter 3	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10
13	Programming work report related to chapter 3	Problem-solving test	Face-to-face	01:00	2.5%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CT05 CB10

17	Second exam, corresponding to Chapter 3 contents. The student will have to solve a problem equivalent to the laboratory activities	Problem-solving test	Face-to-face	01:00	30%	4 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10
17	10% of the final mark will be assigned taking into account criteria such as: assisting to lessons, attention, attitude, ability to answer questions in class, preparation in advance of lessons, etc.	Other assessment	Face-to-face	00:00	10%	/ 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10

7.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
17	Final exam. The student will have to solve a problem equivalent to the laboratory activities. All the chapters will be covered.	Problem-solving test	Face-to-face	02:00	100%	5 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
Extraordinary exam. For those students who are under continous evaluation, the student will have to solve a problem equivalent to the laboratory activities.	Problem-solving test	Face-to-face	02:00	100%	5 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CT05 CB10

7.2. Assessment criteria

Students will be qualified through continuous evaluation by default. According to the Normativa de Evaluación del Aprendizaje de la Universidad Politécnica de Madrid. Sstudents willing to renounce to continuous evaluation must complete the Moodle task entitled "Renounce to continuous evaluation" during the first ten weeks of the course (deadline will be announced in Moodle).

Evaluation will assess if students have acquired all the competences of the subject. Thus, evaluation through final assessment will be carried out considering all the evaluation techniques used in continuous evaluation (EX, ET, TG, etc.), and will be celebrated in the exam period approved by Junta de Escuela for the current academic semester and year. Evaluation activities that assess learning outcomes that cannot be evaluated through a single exam can be carried out along the semester.

Extraordinary examination will be carried out exclusively by the final assessment method.

After each chapter, the student will have to program a problem in MATLAB/Python similar to the case study described by the instructor. After completing the program and running some simulations, a short report has to be written and sent to the instructor for evaluation. Each report will be scored with 2,5% of the final mark

One mid term exam including the first 2 chapters counting 30% of the final mark has to be completed. The student will have to solve a certain problem and answer some theoretical aspects.

A second mid term exam including the chapter 3 counting 30% of the final mark has to be completed. The student will have to solve a certain problem and answer some theoretical aspects

10% of the final mark will be assigned taking into account criteria such as: assisting to lessons, attention, attitude, ability to answer questions in class, preparation in advance of lessons, etc.

Those students who have not passed either continuous evaluation of Final Exam will have still an Extraordinary Exam by the end of June with the same structure as that of the Final Exam in 17th week. All the chapters will be covered.

The Final Assessment method will consists on a exam covering all the chapters in the subject. There will be two parts (40% + 60%). The first one will be a few short questions in order to assess the theoretical knowledge. The second part will consist of two problems in order to evaluate the student's ability to solve practical situations.

8. Teaching resources

8.1. Teaching resources for the subject

Name	Type	Notes
Signal processing for Big Data	Bibliography	Notes describing some of the contents of the course.
Python or Matlab code of case studies	Others	We provide a Matlab/Python code solving some of the case studies proposed in the course.
Course slides	Bibliography	Slides to be presented by the instructor to support the explanations.

COMON, Pierre. 'Tensors: A brief Introduction', IEEE Signal Processing Magazine, n0. 44, May 2014.	Bibliography	Important reference of chapter 2
KOLDA, T. G. and BADER, B. W., 'Tensor Decompositions and Applications', SIAM Rev., vol. 51, n0. 3, pp. 455-500, 2009	Bibliography	Important reference to chapter 2
SHUMAN, D. I., NARANG, S. K. FROSSARD, P., ORTEGA, A., VANDERGHEYNST, P.,'The Emerging Field of Signal Processing on Graphs', IEEE Signal Processing Magazine, Vol. 83, May 2013.	Bibliography	Important reference chapter 3
SANDRYHAILA, A., MOURA, J.M, 'Discrete Signal Processing on Graphs', IEEE Trans. On Signal Processing, vol. 61, n0. 7, pp. 1644-1656, April 2013.	Bibliography	Important reference chapter 3
SANDRYHAILA, A., MOURA, J.M, 'Discrete Signal Processing on Graphs: Frequency Analysis', IEEE Trans. On Signal Processing, vol. 62, n0. 12, pp. 3042-3054, June 2014.	Bibliography	Important reference chapter 3