

Guía docente

270709 - AHLT - Tecnologías Avanzadas del Lenguaje Humano

Última modificación: 12/02/2020

Unidad responsable: Facultad de Informática de Barcelona
Unidad que imparte: 723 - CS - Departamento de Ciencias de la Computación.

Titulación: MÁSTER UNIVERSITARIO EN INTELIGENCIA ARTIFICIAL (Plan 2017). (Asignatura optativa).
MÁSTER UNIVERSITARIO EN INNOVACIÓN E INVESTIGACIÓN EN INFORMÁTICA (Plan 2012). (Asignatura optativa).
MÁSTER UNIVERSITARIO EN TECNOLOGÍAS AVANZADAS DE TELECOMUNICACIÓN (Plan 2019). (Asignatura optativa).
MÁSTER UNIVERSITARIO EN INGENIERÍA DE TELECOMUNICACIÓN (Plan 2013). (Asignatura optativa).

Curso: 2019 **Créditos ECTS:** 5.0 **Idiomas:**

PROFESORADO

Profesorado responsable:

Otros:

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:

CEA3. Capacidad de comprender los principios básicos de funcionamiento de las técnicas principales de Aprendizaje Automático, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CEA5. Capacidad de comprender los principios básicos de funcionamiento de las técnicas de Procesamiento del Lenguaje Natural, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

Genéricas:

CG3. Capacidad para la modelización, cálculo, simulación, desarrollo e implantación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Inteligencia Artificial.

Transversales:

CT3. TRABAJO EN EQUIPO: Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CT6. RAZONAMIENTO: Capacidad de evaluar y analizar de manera razonada y crítica sobre situaciones, proyectos, propuestas, informes y estudios de carácter científico-técnico. Capacidad de argumentar las razones que explican o justifican tales situaciones, propuestas, etc.

CT7. ANALISIS Y SINTESIS: Capacidad de análisis y resolución de problemas técnicos complejos.

Básicas:

CB6. 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.

CB8. 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.

CB9. 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.

METODOLOGÍAS DOCENTES

La metodología docente se describe en las Actividades del curso



OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTADO

Tipo	Horas	Porcentaje
Horas grupo mediano	15,0	25.00
Horas grupo grande	30,0	50.00
Horas grupo pequeño	15,0	25.00

Dedicación total: 60 h

CONTENIDOS

(CAST) Statistical Models for NLP

Descripción:

(CAST) Introduction to statistical modelling for language. Maximum Likelihood models and smooting. Maximum entropy estimation. Log-Linear models

Objetivos específicos:

(CAST)

Actividades vinculadas:

(CAST)

(CAST) Distances and Similarities

Descripción:

(CAST) Distances (and similarities) between linguistic units. Textual, Semantic, and Distributional distances. Semantic spaces (WN, Wikipedia, Freebase, Dbpedia).

Objetivos específicos:

(CAST)

Actividades vinculadas:

(CAST)

(CAST) Sequence Predicion

Descripción:

(CAST) Prediction in word sequences: PoS tagging, NERC. Local classifiers, HMM, global predictors, Log-linear models.

Objetivos específicos:

(CAST)

Actividades vinculadas:

(CAST)



(CAST) Syntactic Parsing

Descripción:

(CAST) Parsing constituent trees: PCFG, CKY vs Inside/outside
Parsing dependency trees: CRFs for parsing. Earley algorithm

Objetivos específicos:

(CAST)

Actividades vinculadas:

(CAST)

(CAST) Document-level modelling

Descripción:

(CAST) Document representation: from BoW to NLU.
Document similarities.
Document classification.

Objetivos específicos:

(CAST)

Actividades vinculadas:

(CAST)

(CAST) Deep Learning approaches - Introduction

Descripción:

(CAST) Introduction to ANN for NLP
Lexical semantics. Word Embeddings

Objetivos específicos:

(CAST)

Actividades vinculadas:

(CAST)

(CAST) Deep Learning approaches - Word Sequences

Descripción:

(CAST) PoS tagging, NERC

Objetivos específicos:

(CAST)

Actividades vinculadas:

(CAST)



(CAST) Deep Learning Approaches - Sentences

Descripción:

(CAST) Sentence similarity, sentence classification. LSTM. BERT. Sentence embeddings

Objetivos específicos:

(CAST)

Actividades vinculadas:

(CAST)

(CAST) Deep Learning approaches - Document Level

Descripción:

(CAST) Document similarity, document classification, document embeddings - doc2vec

Objetivos específicos:

(CAST)

Actividades vinculadas:

(CAST)

(CAST) Deep Learning Approaches - Machine Translation

Descripción:

(CAST) Neural Machine Translation

Objetivos específicos:

(CAST)

Actividades vinculadas:

(CAST)



ACTIVIDADES

(CAST) Course Introduction

Descripción:

(CAST)

Objetivos específicos:

(CAST) 2, 4

Material:

(CAST)

Entregable:

(CAST)

Competencias relacionadas:

CG3. Capacidad para la modelización, cálculo, simulación, desarrollo e implantación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Inteligencia Artificial.

CEA5. Capacidad de comprender los principios básicos de funcionamiento de las técnicas de Procesamiento del Lenguaje Natural, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CEA3. Capacidad de comprender los principios básicos de funcionamiento de las técnicas principales de Aprendizaje Automático, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CT7. ANALISIS Y SINTESIS: Capacidad de análisis y resolución de problemas técnicos complejos.

CT6. RAZONAMIENTO: Capacidad de evaluar y analizar de manera razonada y crítica sobre situaciones, proyectos, propuestas, informes y estudios de carácter científico-técnico. Capacidad de argumentar las razones que explican o justifican tales situaciones, propuestas, etc.

CB6. 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.

Dedicación: 3 h

Grupo grande/Teoría: 2h

Grupo mediano/Prácticas: 1h



(CAST) Distances and Similarities

Descripción:

(CAST)

Objetivos específicos:

(CAST) 2, 4

Material:

(CAST)

Entregable:

(CAST)

Competencias relacionadas:

CG3. Capacidad para la modelización, cálculo, simulación, desarrollo e implantación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Inteligencia Artificial.

CEA5. Capacidad de comprender los principios básicos de funcionamiento de las técnicas de Procesamiento del Lenguaje Natural, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CEA3. Capacidad de comprender los principios básicos de funcionamiento de las técnicas principales de Aprendizaje Automático, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CT7. ANALISIS Y SINTESIS: Capacidad de analisis y resolucion de problemas tecnicos complejos.

CT6. RAZONAMIENTO: Capacidad de evaluar y analizar de manera razonada y critica sobre situaciones, proyectos, propuestas, informes y estudios de caracter cientifico-tecnico. Capacidad de argumentar las razones que explican o justifican tales situaciones, propuestas, etc.

CB6. 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.

Dedicación: 8 h

Grupo grande/Teoría: 5h

Grupo mediano/Prácticas: 3h



(CAST) Sequence Models in NLP

Descripción:

(CAST)

Objetivos específicos:

(CAST) 2, 4

Material:

(CAST)

Entregable:

(CAST)

Competencias relacionadas:

CG3. Capacidad para la modelización, cálculo, simulación, desarrollo e implantación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Inteligencia Artificial.

CEA5. Capacidad de comprender los principios básicos de funcionamiento de las técnicas de Procesamiento del Lenguaje Natural, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CEA3. Capacidad de comprender los principios básicos de funcionamiento de las técnicas principales de Aprendizaje Automático, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CT7. ANALISIS Y SINTESIS: Capacidad de análisis y resolución de problemas técnicos complejos.

CT6. RAZONAMIENTO: Capacidad de evaluar y analizar de manera razonada y crítica sobre situaciones, proyectos, propuestas, informes y estudios de carácter científico-técnico. Capacidad de argumentar las razones que explican o justifican tales situaciones, propuestas, etc.

CB6. 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.

Dedicación: 10 h

Grupo grande/Teoría: 6h

Grupo mediano/Prácticas: 4h



(CAST) Syntax and Parsing

Descripción:

(CAST)

Objetivos específicos:

(CAST) 2, 4

Material:

(CAST)

Entregable:

(CAST)

Competencias relacionadas:

CG3. Capacidad para la modelización, cálculo, simulación, desarrollo e implantación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Inteligencia Artificial.

CEA5. Capacidad de comprender los principios básicos de funcionamiento de las técnicas de Procesamiento del Lenguaje Natural, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CEA3. Capacidad de comprender los principios básicos de funcionamiento de las técnicas principales de Aprendizaje Automático, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CT7. ANALISIS Y SINTESIS: Capacidad de análisis y resolución de problemas técnicos complejos.

CT6. RAZONAMIENTO: Capacidad de evaluar y analizar de manera razonada y crítica sobre situaciones, proyectos, propuestas, informes y estudios de carácter científico-técnico. Capacidad de argumentar las razones que explican o justifican tales situaciones, propuestas, etc.

CB6. 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.

Dedicación: 9 h

Grupo grande/Teoría: 6h

Grupo mediano/Prácticas: 3h



(CAST) Document-level modelling

Descripción:

(CAST)

Objetivos específicos:

(CAST) 2, 4

Material:

(CAST)

Entregable:

(CAST)

Competencias relacionadas:

CG3. Capacidad para la modelización, cálculo, simulación, desarrollo e implantación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Inteligencia Artificial.

CEA5. Capacidad de comprender los principios básicos de funcionamiento de las técnicas de Procesamiento del Lenguaje Natural, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CEA3. Capacidad de comprender los principios básicos de funcionamiento de las técnicas principales de Aprendizaje Automático, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CT7. ANALISIS Y SINTESIS: Capacidad de analisis y resolucion de problemas tecnicos complejos.

CT6. RAZONAMIENTO: Capacidad de evaluar y analizar de manera razonada y critica sobre situaciones, proyectos, propuestas, informes y estudios de caracter cientifico-tecnico. Capacidad de argumentar las razones que explican o justifican tales situaciones, propuestas, etc.

CB6. 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.

Dedicación: 6 h

Grupo grande/Teoría: 4h

Grupo mediano/Prácticas: 2h



(CAST) Neural Machine Translation

Descripción:

(CAST)

Objetivos específicos:

(CAST) 2, 4

Material:

(CAST)

Entregable:

(CAST)

Competencias relacionadas:

CG3. Capacidad para la modelización, cálculo, simulación, desarrollo e implantación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Inteligencia Artificial.

CEA5. Capacidad de comprender los principios básicos de funcionamiento de las técnicas de Procesamiento del Lenguaje Natural, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CEA3. Capacidad de comprender los principios básicos de funcionamiento de las técnicas principales de Aprendizaje Automático, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CT7. ANALISIS Y SINTESIS: Capacidad de análisis y resolución de problemas técnicos complejos.

CT6. RAZONAMIENTO: Capacidad de evaluar y analizar de manera razonada y crítica sobre situaciones, proyectos, propuestas, informes y estudios de carácter científico-técnico. Capacidad de argumentar las razones que explican o justifican tales situaciones, propuestas, etc.

CB6. 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.

Dedicación: 6 h

Grupo grande/Teoría: 4h

Grupo mediano/Prácticas: 2h



(CAST) Final Exam

Descripción:

(CAST)

Objetivos específicos:

(CAST) 2, 3, 4

Material:

(CAST)

Entregable:

(CAST)

Competencias relacionadas:

CG3. Capacidad para la modelización, cálculo, simulación, desarrollo e implantación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Inteligencia Artificial.

CEA5. Capacidad de comprender los principios básicos de funcionamiento de las técnicas de Procesamiento del Lenguaje Natural, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CEA3. Capacidad de comprender los principios básicos de funcionamiento de las técnicas principales de Aprendizaje Automático, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CT7. ANALISIS Y SINTESIS: Capacidad de análisis y resolución de problemas técnicos complejos.

CT6. RAZONAMIENTO: Capacidad de evaluar y analizar de manera razonada y crítica sobre situaciones, proyectos, propuestas, informes y estudios de carácter científico-técnico. Capacidad de argumentar las razones que explican o justifican tales situaciones, propuestas, etc.

CB6. 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.

CB8. 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.

CB9. 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.

Dedicación: 13 h

Actividades dirigidas: 3h

Aprendizaje autónomo: 10h 30m



(CAST) Project

Descripción:

(CAST)

Objetivos específicos:

(CAST) 1, 2, 4

Material:

(CAST)

Entregable:

(CAST)

Competencias relacionadas:

CG3. Capacidad para la modelización, cálculo, simulación, desarrollo e implantación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Inteligencia Artificial.

CEA5. Capacidad de comprender los principios básicos de funcionamiento de las técnicas de Procesamiento del Lenguaje Natural, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CEA3. Capacidad de comprender los principios básicos de funcionamiento de las técnicas principales de Aprendizaje Automático, y saber utilizarlas en el entorno de un sistema o servicio inteligente.

CT7. ANALISIS Y SINTESIS: Capacidad de análisis y resolución de problemas técnicos complejos.

CT6. RAZONAMIENTO: Capacidad de evaluar y analizar de manera razonada y crítica sobre situaciones, proyectos, propuestas, informes y estudios de carácter científico-técnico. Capacidad de argumentar las razones que explican o justifican tales situaciones, propuestas, etc.

CT3. TRABAJO EN EQUIPO: Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CB6. 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.

CB8. 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.

Dedicación: 45 h

Aprendizaje autónomo: 45h

SISTEMA DE CALIFICACIÓN

Sin traducir



BIBLIOGRAFÍA

Básica:

- Clark, A.; Fox, C.; Lappin, S. (eds.). The handbook of computational linguistics and natural language processing [en línea]. Oxford: Wiley-Blackwell, 2010 [Consulta: 01/04/2020]. Disponible a: <https://onlinelibrary.wiley.com/doi/book/10.1002/9781444324044>. ISBN 9781444324044.
- Dale, R.; Moisl, H.; Somers, H. (eds.). Handbook of natural language processing. New York: Marcel Dekker, 2000. ISBN 0824790006.
- Stymne, S.; Lhoneux, Miryam de. Syntactic analysis in language technology: syntactic parsing [en línea]. Uppsala: Uppsala Universitet, 2017 [Consulta: 20/04/2020]. Disponible a: <https://cl.lingfil.uu.se/~sara/kurser/5LN455-2017/>.
- Indurkha, N.; Damerau, F.J. (eds.). Handbook of natural language processing [en línea]. Boca Raton: Chapman and Hall/CRC, 2010 [Consulta: 19/02/2020]. Disponible a: <https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=565922>. ISBN 9781420085938.
- Jurafsky, D.; Martin, J.H. Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. 2nd ed. Upper Saddle River: Prentice Hall, 2008. ISBN 9332518416.
- Mitkov, R. (ed.). The Oxford handbook of computational linguistics. Oxford: Oxford University Press, 2003. ISBN 0198238827.
- Manning, C.D.; Schütze, H. Foundations of statistical natural language processing. Cambridge, Mass.: MIT Press, 1999. ISBN 0262133601.
- Smith, N.A. Linguistic structure prediction. San Rafael, California: Morgan & Claypool, 2011. ISBN 9781608454051.
- Manning, C.; See, A. Natural language processing with deep learning [en línea]. Stanford University, [Consulta: 25/05/2020]. Disponible a: <http://web.stanford.edu/class/cs224n/>.
- Collins, M. Natural language processing [en línea]. Columbia University, [Consulta: 19/02/2020]. Disponible a: <http://www.cs.columbia.edu/~mcollins/cs4705-fall2018/>.
- Titov, I. Natural language processing [en línea]. Universiteit van Amsterdam, [Consulta: 19/02/2020]. Disponible a: <http://ivan-titov.org/teaching/nlp1-15/index.html>.

RECURSOS

Enlace web:

- <http://www.lsi.upc.edu/~ageno/anlp>