



Facultad de Ingeniería Comisión Académica de Posgrado

Formulario de Aprobación Curso de Posgrado

Asignatura: Procesamiento Semántico del Lenguaje Natural

(Si el nombre contiene siglas deberán ser aclaradas)

Profesor de la asignatura ¹: Profesora Titular Dra.Dina Wonsever

(título, nombre, grado o cargo, Instituto o Institución)

Profesor Responsable Local ¹:

(título, nombre, grado, Instituto)

Otros docentes de la Facultad:

(título, nombre, grado, Instituto)

Docentes fuera de Facultad:

(título, nombre, cargo, Institución, país)

Instituto ó Unidad: Instituto de Computación

Departamento ó Area: Procesamiento de Lenguaje Natural

¹ Agregar CV si el curso se dicta por primera vez.

(Si el profesor de la asignatura no es docente de la Facultad se deberá designar un responsable local)

Horas Presenciales: 34

Nº de Créditos: 8

Público objetivo y Cupos:

Estudiantes de posgrado en Informática con conocimientos previos en Procesamiento de Lenguaje Natural. No hay cupos.

Objetivos: Adquisición de conocimientos y experimentación práctica en distintos tipos de representaciones semánticas para el lenguaje natural y en métodos y procesos para su obtención a partir de enunciados lingüísticos. Conocimiento de recursos existentes, de metodologías de construcción de nuevos recursos y de propuestas de transferencia entre distintas lenguas.

Conocimientos previos exigidos: Lógica de predicados
Probabilidad y estadística
Aprendizaje automático
Programación
Procesamiento de Lenguaje Natural

Los conocimientos requeridos en los primeros cuatro ítems se adquieren en la carrera de Ingeniería en Computación. Los conocimientos previos requeridos en Procesamiento de Lenguaje Natural se adquieren en alguno de estos dos cursos : Introducción al Procesamiento de Lenguaje Natural o Gramáticas Formales para el Lenguaje Natural. Esto significa que al menos uno de estos dos cursos es previo.

Conocimientos previos recomendados: Además de los conocimientos exigidos es recomendable tener conocimientos y experiencia en Aprendizaje Automático

Metodología de enseñanza:

(comprende una descripción de las horas dedicadas por el estudiante a la asignatura y su distribución en horas presenciales -de clase práctica, teórico, laboratorio, consulta, etc.- y no presenciales de trabajo personal del estudiante)



Facultad de Ingeniería Comisión Académica de Posgrado

En el curso se realizarán tres tipos de actividades: clases teórico-práctica impartidas por el docente, clases de presentación y discusión de artículos por parte de los estudiantes y realización de un proyecto con escritura de un artículo científico sobre este.

- Horas clase (teórico): 30
- Horas clase (práctico):
- Horas clase (laboratorio):
- Horas consulta:
- Horas evaluación: 4
 - Subtotal horas presenciales: 34
- Horas estudio: 30
- Horas resolución ejercicios/prácticos:
- Horas proyecto final/monografía: 60
 - Total de horas de dedicación del estudiante: 124

Forma de evaluación:

El curso se evalúa en base al proyecto / artículo científico final (contribuye con 30% al puntaje total), la presentación de un artículo en clase (contribuye con 30% al puntaje total) y una prueba individual escrita (contribuye con 40% al puntaje total). Cada una de estas partes tiene un mínimo de suficiencia de un 60%; es imprescindible llegar a estos mínimos para la aprobación del curso.

Temario:

- Revisión de nociones básicas de Procesamiento de Lenguaje Natural
- Representaciones semánticas en PLN: semántica lógica/composicional, semántica léxica, etiquetas semánticas en dependencias, semántica específica orientada a acciones un dominio.
- Anotación semántica, esquema AMR, esquemas particulares para algunos fenómenos: temporalidad, relaciones espaciales, afectividad, otros. Esquemas transversales y esquemas orientados a dominio.
- Recursos para el análisis semántico:
 - Repositorios léxicos, relaciones entre unidades léxicas: diccionarios, tesauros, bases de datos léxicas (WordNet). Repositorios para el español y el inglés
 - Corpora: texto puro, texto anotado, representaciones vectoriales.
- Construcción de representaciones semánticas para expresiones lingüísticas (*parsing semántico*):
 - Aprendizaje supervisado en CCG
 - Aprendizaje semi-supervisado
 - Algoritmos basados en grafos

Bibliografía:

Consulta (disponible en la biblioteca del InCo)

Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, (2nd edition), D. Jurafsky y J. H. Martin. Prentice Hall, Upper Saddle River, N.J. ISBN-10: 0131873210. 2008.

Artículos

Liang, P., Jordan, M. I., and Klein, D. (2011). Learning dependency-based compositional semantics. In



Facultad de Ingeniería Comisión Académica de Posgrado

Proceedings of ACL-11 pages 590-599. (Disponible web,
http://www.mitpressjournals.org/doi/pdf/10.1162/COLI_a_00127)

Liang,P. and Potts,C. (2015).Bringing machine learning and compositional semantics together. Annual Reviews of Linguistics, 1(1):355- 376. (disponible web en versión casi final,
<http://web.stanford.edu/~cgpotts/manuscripts/liang-potts-semantics.pdf>)

Graph-based Algorithms for Semantic Parsing Jeffrey Flanigan, Samuel Thomson, David Bamman, Jesse Dodge, Manaal Faruqui, Brendan O'Connor, Nathan Schneider, Swabha Swayamdipta, Chris Dyer and Noah A. Smith ACL 2014 (disponible web)

Wong, Y. W. and Mooney, R. J. (2006). Learning for semantic parsing with statistical machine translation. In Proceedings of HLT-NAACL-2006, pages 439-446. (disponible web : <http://dl.acm.org/citation.cfm?id=1220891>)

(Se verán artículos adicionales durante el curso)



Facultad de Ingeniería Comisión Académica de Posgrado

Datos del curso

Fecha de inicio y finalización: 15 de setiembre al 15 de diciembre de 2016

Horario y Salón: Martes y Jueves de 8 a 9.30, salón 705
