

Programa de Asignatura

Ingeniería en Computación - In.Co.

Nombre de la Asignatura	Introducción al Procesamiento de Lenguaje Natural
Créditos	12 créditos
Objetivo de la Asignatura	<p>Presentar los principales problemas del área de PLN: traducción automática, resumen automático, reconocimiento del habla, recuperación de información, etc. Dar un panorama del estado de las aplicaciones existentes hoy día.</p> <p>Desarrollar los enfoques y técnicas comúnmente utilizados en el área: análisis morfológico, análisis sintáctico de superficie, análisis semántico, análisis pragmático y modelos del discurso, métodos de estado finito, aprendizaje automático, etc. Estudiar las limitaciones teóricas y materiales de los diferentes enfoques propuestos.</p> <p>Presentar algunas de las herramientas y recursos más utilizados: <i>taggers</i>, diccionarios, ontologías, gramáticas, corpus paralelos, etc. Utilizar estas herramientas en la resolución de un problema práctico.</p>
Metodología de enseñanza	Los temas se expondrán en clases teórico-prácticas. Los estudiantes realizarán un trabajo de laboratorio durante el desarrollo del curso en el que aplicarán los conocimientos teóricos adquiridos.
Temario	<ul style="list-style-type: none"> • Introducción al área. Motivación. Aplicaciones más comunes. Enfoques simbólicos, estadísticos, híbridos. • Nociones básicas de gramática del español. • Análisis léxico. Tokenización. Segmentación en oraciones. Entidades con nombre. Morfología. • Análisis sintáctico. Gramáticas. Métodos de parsing. • Creación de corpus. Realización de POS tagging usando métodos estadísticos. • Análisis semántico. Ambigüedad. Anáforas y catáforas. Representación de conocimiento del mundo. • Lingüística textual. Modelos del discurso. • Aplicaciones
Bibliografía	<p>Bibliografía básica:</p> <ul style="list-style-type: none"> • <i>Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition</i>, Prentice Hall, Upper Saddle River, N.J., 2000. D. Jurafsky, James H. Martin. ISBN 0130950696. • <i>Foundations of Statistical Natural Language Processing</i>, Manning & Schütze, Mit Press Cambridge, 1999. ISBN: 9780262133609.

Bibliografía complementaria:

- *Handbook of natural Language Processing*, R. Dale, H. Moisl, H.Somers, ed., Marcel Dekker, New York, 2000. ISBN 0824790006.
- *Natural Language Understanding*, J. Allen, Benjamin/Cummings Publishing Company, 1995. ISBN 0805303340.
- *Finite State Morphology*, Kenneth R. Beesley and Lauri Karttunen, CSLI Publications, ISBN (Paperback): 1575864347.
- Durante el curso se indicarán artículos científicos a ser presentados.

**Conocimientos
previos exigidos y
recomendados**

Se requiere conocimiento de los siguientes temas :

- Lógica de predicados.
- Probabilidad y estadística.
- Gramáticas formales.
- Programación.

Anexo:

1) Cronograma tentativo.

- Semana 1: Introducción al área. Conceptos básicos.
- Semanas 2 – 4: Análisis léxico. Morfología.
- Semanas 5 – 8: Análisis sintáctico.
- Semanas 9 – 11: Análisis semántico.
- Semana 12: Lingüística textual. Modelos del discurso.
- Semanas 13 – 14: Presentación de aplicaciones.
- Semana 15: Presentación de artículos.

Se estima un total de aproximadamente 180 horas de trabajo del estudiante, desglosadas en: 60 horas aula, 20 horas de lectura y presentación de artículos, y 100 horas de trabajo de laboratorio.

2) Modalidad del curso y procedimiento de evaluación.

El curso es participativo y está pensado para un grupo de no más de 30 personas. Se propone la siguiente metodología:

1. Presentación y cuestionamiento de artículos

Los estudiantes deberán realizar presentaciones de artículos seleccionados. Los docentes del curso ofrecerán un conjunto inicial de artículos y los estudiantes (en grupos de dos personas) deberán elegir un artículo a presentar y un artículo (que será presentado por otro grupo) a cuestionar. La dinámica será la siguiente: el grupo que presenta el artículo debe, además, entregarle al docente, el día de la presentación, un resumen de una página conteniendo los conceptos clave del artículo. El grupo que cuestiona el artículo, deberá entregarle al docente, el día de la presentación, una lista de al menos cinco preguntas a hacer al grupo que presenta el artículo. Las preguntas las formularán los estudiantes de este grupo y el docente será

el moderador de la discusión (en la que puede participar el resto de la clase), y evaluará la pertinencia e interés de las preguntas.

2. Proyectos

Los estudiantes deberán realizar un proyecto a lo largo del curso. Los docentes del curso pondrán a disposición de los estudiantes las herramientas y recursos, y una lista de proyectos.

3. Prueba final

Se prevé una prueba individual escrita final, con el objetivo de evaluar los conocimientos del curso.

4. Evaluación

El curso se evaluará a partir de:

- la presentación del artículo: 15%
- el cuestionamiento del artículo: 10%
- el proyecto: 35%
- la prueba individual escrita: 40%

Las cuatro instancias son obligatorias y se requerirá como mínimo el 60% del puntaje de la prueba individual escrita. El curso se aprueba obteniendo más del 60% de los puntos totales.

3) Materia.

Inteligencia Artificial y Robótica.

4) Previaturas.

- Plan 87
 - Lógica
 - Programación III
 - Teoría de la Programación I
 - Taller III
- Plan 97
 - Lógica (E-C)
 - Programación 3 (E-C)
 - Probabilidad y Estadística (E-C)
 - Teoría de Lenguajes (E-C)
 - Programación 4 (E-C)
 - Taller de Programación

5) Cupo

30 personas.

El criterio de selección se distribuye de la siguiente forma: 15 estudiantes serán seleccionados según escolaridad en asignaturas afines y los restantes 15 mediante sorteo. La cantidad asignada se justifica por la modalidad propuesta para el curso, que es participativo y requiere que todos los estudiantes realicen presentaciones de artículos. A su vez, como es el primer año de su dictado es necesario mantener un cupo controlado por parte de los docentes.

Aprobado por Res. del Consejo de Facultad el 29.6.06 - Exp. 060120-001492-06