

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE INGENIERÍA



PROGRAMA DE ESTUDIO

PROCESAMIENTO DEL LENGUAJE NATURAL

0751

8°, 9°

06

Asignatura

Clave

Semestre

Créditos

Ingeniería Eléctrica

Ingeniería en Computación

Ingeniería en Computación

División

Departamento

Carrera en que se imparte

Asignatura:

Obligatoria de elección

Optativa

Horas:

Teóricas

Prácticas

Total (horas):

Semana

16 Semanas

Aprobado:
Consejo Técnico de la Facultad
Consejo Académico del Área de las Ciencias
Físico Matemáticas y de las Ingenierías

Fecha:
25 de febrero, 17 de marzo y 16 de junio de 2005
11 de agosto de 2005

Modalidad: Curso

Asignatura obligatoria antecedente: Ninguna.

Asignatura obligatoria consecuyente: Ninguna.

Objetivo(s) del curso:

El alumno comprenderá los métodos y problemas centrales del Procesamiento del Lenguaje Natural (PLN) y será capaz de proponer soluciones a problemas concretos en este campo.

Temario

NÚM.	NOMBRE	HORAS
1.	Introducción al procesamiento del lenguaje natural (PLN)	3.0
2.	Bases computacionales del PLN	6.0
3.	Recursos del PLN	6.0
4.	Análisis morfosintáctico	9.0
5.	Análisis sintáctico	9.0
6.	Análisis semántico	9.0
7.	Aplicaciones	6.0
		48.0
	Prácticas de laboratorio	0.0
	Total	48.0



1 Introducción al procesamiento del lenguaje natural (PLN)

Objetivo: El alumno se familiarizará con los problemas de representación del conocimiento lingüístico y con los métodos de aproximación al PLN.

Contenido:

- 1.1 Antecedentes
- 1.2 La representación del conocimiento lingüístico
- 1.3 Dos aproximaciones diferentes: métodos estadísticos vs. métodos basados en reglas
- 1.4 Objetivos: Análisis y generación

2 Bases computacionales del PLN

Objetivo: El alumno profundizará sus conocimientos sobre los esquemas de representación de lenguajes formales con respecto a su aplicación en la generación y reconocimiento de lenguas naturales.

Contenido:

- 2.1 Tipología de lenguajes
- 2.2 Autómatas de estados finitos
- 2.3 Procesadores de lenguaje: reconocedores, transductores, analizadores y generadores
- 2.4 Algoritmos de n-gramas
- 2.5 Algunos problemas interesantes
 - 2.5.1 Desambiguación léxica
 - 2.5.2 Resolución de anáforas

3 Recursos del PLN

Objetivo: El alumno conocerá los recursos del PLN para resolver problemas típicos del este campo de especialización.

Contenido:

- 3.1 Lexicones y bases de conocimiento
- 3.2 Corpus textuales
 - 3.2.1 Toquenización
 - 3.2.2 Análisis estadísticos
- 3.3 Información semántica: WordNet y otras ontologías

4 Análisis morfosintáctico

Objetivo: El alumno dominará las técnicas de etiquetamiento morfosintáctico de corpus de lenguajes naturales.

Contenido:

- 4.1 Análisis morfológico y morfosintáctico
- 4.2 Categorización gramatical (POS)



- 4.3 Construcción de etiquetadores morfosintácticos
- 4.4 Estrategias de entrenamiento
- 4.5 Etiquetadores en cascada
- 4.6 Morfología de dos niveles
- 4.7 Evaluación y modificación de las reglas

5 Análisis sintáctico

Objetivo: El alumno dominará las técnicas de segmentación del discurso en sintagmas y se familiarizará con las técnicas básicas de análisis sintáctico.

Contenido:

- 5.1 Gramáticas independientes de contexto
- 5.2 Creación de reglas para el segmentador sintagmático
- 5.3 Identificación de sintagmas nominales
 - 5.3.1 Sintagmas nominales
 - 5.3.2 Sintagmas verbales
 - 5.3.3 Sintagmas preposicionales
- 5.4 Análisis sintáctico
 - 5.4.1 Técnicas: métodos ascendente y descendente
 - 5.4.2 Estrategias de búsqueda: profundidad y amplitud
 - 5.4.3 Chart-parsing
 - 5.4.4 Gramáticas de transición aumentadas y de unificación
- 5.5 Resolución anafórica

6 Análisis semántico

Objetivo: El alumno conocerá los principios del análisis semántico típicos del PLN.

Contenido:

- 6.1 Representación del significado
- 6.2 Análisis semántico
- 6.3 Interacción con WordNet
- 6.4 Dependencia conceptual

7 Aplicaciones

Objetivo: El alumno conocerá algunas aplicaciones del PLN en las que se utilizan las técnicas y recursos aprendidos en el curso.

Contenido:

- 7.1 Recuperación y extracción de información
- 7.2 Categorización y resúmenes automáticos
- 7.3 Generación de lenguaje natural
- 7.4 Sistemas conversacionales



Bibliografía básica:

ALLEN, James

Natural Language Understanding

2nd edition

U.S.A.

Addison Wesley, 1994

JURAFSKY, Daniel, JAMES H. Martin

Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition

Prentice-Hall, 2000

MANNING, Christopher, FRIDRICH, Schütze

Foundations in Statistical Natural Language Processing

Cambridge, U.S.A.

The MIT Press, 1999

GRISHMAN, Ralph

Computational Linguistics

Cambridge University Press, 1986

CHARNIAK, Eugene

Statistical Language Learning

Cambridge, U.S.A.

The MIT Press, 1996

Bibliografía complementaria:

MITKOV, Ruslan, et al.

The Oxford Handbook of Computational Linguistics

Oxford

Oxford UP, 2005

MORENO, Lidia, et al.

Introducción al procesamiento de lenguaje natural

Murcia

Universidad de Alicante, 1999



Sugerencias didácticas:

Exposición oral	<input checked="" type="checkbox"/>
Exposición audiovisual	<input checked="" type="checkbox"/>
Ejercicios dentro de clase	<input checked="" type="checkbox"/>
Ejercicios fuera del aula	<input checked="" type="checkbox"/>
Seminarios	<input type="checkbox"/>

Lecturas obligatorias	<input type="checkbox"/>
Trabajos de investigación	<input type="checkbox"/>
Prácticas de taller o laboratorio	<input checked="" type="checkbox"/>
Prácticas de campo	<input type="checkbox"/>
Otras	<input type="checkbox"/>

Forma de evaluar:

Exámenes parciales	<input type="checkbox"/>
Exámenes finales	<input type="checkbox"/>
Trabajos y tareas fuera del aula	<input checked="" type="checkbox"/>

Participación en clase	<input checked="" type="checkbox"/>
Asistencias a prácticas	<input type="checkbox"/>
Otras	<input type="checkbox"/>

Perfil profesiográfico de quienes pueden impartir la asignatura

Ingeniero en computación familiarizado con el procesamiento del lenguaje natural.