



CONTENIDO DE CADA ASIGNATURA

NOMBRE DE LA ASIGNATURA	
Procesamiento de lenguaje natural	
CICLO	CLAVE DE LA ASIGNATURA
II Semestre	
OBJETIVO GENERAL	
Introducir al análisis y transformación del lenguaje mediante técnicas computacionales. Representar al texto y discurso que pueden ayudar a la predicción, la extracción y el razonamiento semántico sobre el lenguaje. Comprender el lenguaje natural para su aplicación en los hardwares y software.	
MODELO DE DISEÑO INSTRUCCIONAL	
<p>El Modelo Pedagógico que aplicará ALINNCO está fundamentado en unos referentes pedagógicos y didácticos, que ofrecen una nueva connotación al significado de enseñar y aprender; en éste se privilegia una orientación didáctica dirigida hacia la organización del proceso de aprendizaje, con el fin de explicar cómo opera la apropiación del saber teórico-práctico, cuál es la relación que mantienen los sujetos con los conocimientos teóricos, ya que los prácticos la mayoría los domina, y en qué medida la función del profesor es la de propiciar las condiciones necesarias para que la transferencia de conocimientos tenga lugar en los participantes. Este modelo permite trascender el esquema educativo tradicional creando las condiciones de una adecuada práctica pedagógica, desde la interacción con el conocimiento y el entorno social y empresarial. Además de fortalecer las competencias, el desarrollo y la investigación, lo cual contribuye en la formación de un posgraduado crítico, innovador, competitivo y autónomo.</p> <p>La metodología desarrollada por ALINNCO para el diseño y desarrollo de programas de posgrado consiste en detectar inicialmente las necesidades de formación en el sector productivo y de servicios, aspecto que surge del contacto y experiencia que tiene el grupo directivo con las empresas y a partir de esto se realiza un estudio comparativo de programas educativos similares sobre el tema del posgrado que se desea proponer, que abarque no solamente a las instituciones nacionales sino también a las internacionales.</p> <p>Como complemento para el aprendizaje, se cuenta con ALINNCO Digital, que es el campus virtual donde los profesores “colgarán” sus clases para que los alumnos las puedan consultar en sus computadoras cuando deseen. Además, tendrán acceso a la bibliografía digital a través de este sistema, pero se controlará el acceso a través de las herramientas Digital Rights Management (DRM), que son un conjunto de tecnologías electrónicas cuya función es bloquear y proteger obras.</p>	
TEMAS Y SUBTEMAS O UNIDADES DE APRENDIZAJE	

UNIDAD INTRODUCTORIA: PRESENTACIÓN Y ANÁLISIS DEL PROGRAMA ANALÍTICO.

OBJETIVO ESPECÍFICO:

Presentar a los alumnos el contenido que será impartido, haciendo un breve comentario general de cada unidad.

UNIDAD 1. Fundamentos

- 1.1 Procesamiento del lenguaje natural: problemas y perspectivas
- 1.2 Introducción / Recordar a / de cálculo de probabilidad
 - N-gramos y modelos de lenguaje
 - Modelos Markov
- 1.3 Modelos de lenguaje de red neuronal recurrente
- 1.4 Evaluación de las aplicaciones de PNL

UNIDAD 2. Corpus

- 2.1 Introducción a corpus
- 2.2 Elementos en corpus equilibrado
- 2.3 Corpus lingüístico
 - TreeBank
 - PropBank
 - WordNet
 - VerbNet
- 2.4 Gestión de recursos con XML
- 2.5 Gestión de datos lingüísticos con la ayuda de GATE, NLTK.

UNIDAD 3. Procesamiento de lenguaje natural

- 3.1 Fonética Computacional
 - Muestras de discurso: propiedades y medidas acústicas.
 - Análisis en el dominio de la frecuencia, espectrogramas
 - Aplicaciones en el campo fonético acústico.
 - Reconocimiento de voz con HMM y redes neuronales profundas
- 3.2 Tokenización y división de oraciones
- 3.3 Morfología Computacional
 - Operaciones morfológicas
 - Léxico estático, morfología de dos niveles
- 3.4 Sintaxis Computacional
 - Etiquetado de parte del discurso
 - Gramáticas para el lenguaje natural.
 - Análisis de lenguaje natural
 - Hoja de trabajo complementaria: gramáticas formales para NL
 - Lenguas formales y lenguas naturales. Complejidad del lenguaje natural.
 - Estructura de frases gramáticas, gramáticas de dependencia
 - Treebanks
 - Formalismos modernos para analizar lenguajes naturales.
- 3.5 Semántica Computacional
 - Semántica léxica: WordNet y FrameNet
 - Desambiguación del sentido de las palabras
 - Modelos de espacio de palabras
 - Enfoques lógicos para la semántica de oraciones

ACTIVIDADES DE APRENDIZAJE	
BAJO CONDUCCIÓN DE DOCENTE: Horas clase 40 h en aula Asesorías y consultas con el Profesor	
DE MANERA INDEPENDIENTE: Ejercicios independientes 40 horas	
CRITERIOS DE EVALUACIÓN Y ACREDITACIÓN	
CRITERIOS DE EVALUACIÓN	PORCENTAJE
Participación en línea	
Taller y participación en clases	30%
Producto Integrador	70%
	TOTAL 100%

Bibliografía

1. Jurafsky, D., J. H. Martin. Speech and Language Processing. USA: Prentice Hall, 2000. 975 pp.
2. Bharati, A., V. Chaintaya and R. Sangal. Natural language processing: a Paninian perspective. Prentice Hall, 1995.
3. INDURKHYA, N., F.J. Damerau. Handbook of natural language processing. 2nd. Edition. USA: Taylor and Francis Group, 2010. 676 pp. ISBN 13: 978-1-4200-8593-8
4. MANNING, C. D., Schütze, Hinric. Foundations of statistical natural language processing. USA: Massachusetts Institute of Technology, 2000. 704 pp. ISBN 0-262-13360-1

