

PROGRAMA DE ASIGNATURA

1. NOMBRE DE LA ASIGNATURA

Seminario de Lingüística Computacional

2. NOMBRE DE LA ASIGNATURA EN INGLÉS

Seminar in Computational Linguistics

3. TIPO DE CRÉDITOS DE LA ASIGNATURA

SCT/

UD/

OTROS/

4. NÚMERO DE CRÉDITOS

8 créditos

5. HORAS DE TRABAJO PRESENCIAL DEL CURSO

3 (Lunes de 18:00 a 21:00, Segundo Semestre 2020)

6. HORAS DE TRABAJO NO PRESENCIAL DEL CURSO

6

7. OBJETIVO GENERAL DE LA ASIGNATURA

Introducir a los estudiantes a la Lingüística Computacional y del Corpus y al Procesamiento del Lenguaje Natural (NLP) como herramienta para el trabajo cuantitativo en lingüística. Se introduce al estudiante al campo interdisciplinario de la Lingüística Computacional (que utiliza procedimientos de la Inteligencia Artificial para abordar problemas de la Lingüística), sus técnicas, métodos y algunas de sus áreas de investigación. El curso se organiza por medio de clases expositivas desarrolladas en modalidad de seminario y de práctica con funcionalidades computacionales.

8. OBJETIVOS ESPECÍFICOS DE LA ASIGNATURA

1. Introducir al estudiante al campo interdisciplinario de la Lingüística Computacional.
2. Adiestrarse en el uso de herramientas computacionales para la elaboración y procesamiento de corpora.
3. Adiestrarse en el uso de herramientas computacionales para el procesamiento del lenguaje natural.

9. SABERES / CONTENIDOS

1. Introducción al curso, lectura de programa, indicaciones trabajo final, discusión sobre "El Gran Despertar de la IA".
2. Nociones Fundamentales: Ley de Zipf y de Heaps, KWIC, Corporas, TreeBanks, Precision, Recall & F, Colocaciones.
3. Los corpora electrónicos, fundamentos. Usando las redes sociales para crear nuestros propios corpora. El uso de las APIs.
4. Buscando en los corpora electrónicos. RegEx (Antconc), Notepad++, DTSearch, KFNgram.
5. Lectura y Discusión de las Introducciones de Manning & Schütze (1999).
6. Análisis cualitativo: el modelo Nvivo.
7. Análisis cuantitativo: el modelo de Machine Learning. Uso de Weka.
8. Ejemplos en Análisis Computacional de la Literatura: modelos Martínez-Gamboa (2003, 2012, 2015, 2016). Social Network Analysis. Silabizador.
9. El uso de POS-Taggers, Parsers y otros mecanismos de análisis lingüístico (niveles). El uso de phpsyntaxtree.
10. Análisis de Sentimientos. Polaridades. Novedad.
11. Avance Trabajos de Investigación
12. Presentaciones en modalidad de seminario I (Mitkov, 2002): Fonología, Morfología, Lexicología.
13. Presentaciones en modalidad de seminario II (Mitkov, 2002): Sintaxis, Semántica, Discurso.
14. Presentaciones en modalidad de seminario III (Mitkov, 2002): Segmentación, POS-Tagging, Parsing.
15. Entrega Final Trabajos de Investigación.

10. METODOLOGÍA

Clases expositivas y prácticas sobre contenidos del programa; trabajo de investigación; lectura crítica en modalidad de seminario de textos seleccionados.

11. METODOLOGÍAS DE EVALUACIÓN

Desarrollo de un trabajo de investigación (60%).
Presentaciones en modalidad de seminario (40%).

12. REQUISITOS DE APROBACIÓN

ASISTENCIA (*indique %*): 75%

NOTA DE APROBACIÓN MÍNIMA (*Escala de 1.0 a 7.0*): 4,0

REQUISITOS PARA PRESENTACIÓN A EXÁMEN: 4,0

OTROS REQUISITOS:

13. PALABRAS CLAVE

Lingüística Computacional, Lingüística de Corpus, Machine Learning, Análisis Cualitativo, Social Network Analysis.

14. BIBLIOGRAFÍA OBLIGATORIA

- Gefter, A. (2015). The man who tried to redeem the world with logic. *Nautilus*, 21.
- Lewis-Kraus, G. (2016). The great AI awakening. *The New York Times Magazine*, 14.
- Manning, C. D., & Schütze, H. (1999). Foundations of statistical natural language processing. MIT press.
- Mitkov, R. (Ed.). (2004). *The Oxford handbook of computational linguistics*. Oxford University Press.

15. BIBLIOGRAFÍA COMPLEMENTARIA

Brown, K. (ED) (2005). Encyclopedia of Language & Linguistics, 2nd. Ed. Amsterdam: Elsevier Press.

16. RECURSOS WEB

Weka
NVivo
Pajek
Twitter API