



CD-2986-17



Universidad de Buenos Aires  
Facultad de Ciencias Exactas y Naturales

Planilla a completar para presentación de Cursos de Posgrado

1.- DEPARTAMENTO de COMPUTACIÓN

2.- NOMBRE DEL CURSO: Procesamiento de Lenguaje Natural

3.- DOCENTES:

RESPONSABLE/S: Franco M. Luque

COLABORADORES: .....

AUXILIARES: .....

4.- CARRERA de DOCTORADO

5.- AÑO: 2018

CUATRIMESTRE/S: Verano

6.- PUNTAJE PROPUESTO PARA CARRERA DE DOCTORADO: 1 (un) punto

7.- DURACIÓN (anual, cuatrimestral, bimestral u otra): mensual

8.- CARGA HORARIA SEMANAL:

Teóricas: .....

Problemas: .....

Laboratorio: .....

Seminarios: .....

Teórico – Práctico: 8hs

Salida a Campo: .....

9.- CARGA HORARIA TOTAL: 32 horas (8 horas semanales, 4 semanas).

10.- FORMA DE EVALUACIÓN: El curso será evaluado a través de trabajos prácticos en los que se deben realizar diversos ejercicios de programación de modelos y experimentación con datos. El último trabajo práctico será de temática libre. Para la aprobación del curso, deben ser aprobados todos los trabajos prácticos.

## 11.- PROGRAMA ANALÍTICO:

- Procesamiento básico de texto: Expresiones regulares, tokenización, segmentación, normalización, lematización y stemming.
- Modelado de lenguaje: N-gramas, suavizado add-one y por interpolación, back-off. Evaluación con perplexity. Aplicaciones: Generación de lenguaje y atribución de autoría.
- Etiquetado de secuencias: Etiquetado morfosintáctico ( PoS tagging ) y Reconocimiento de Entidades Nombradas (NER). Modelos Ocultos de Markov (HMMs). Modelos de Máxima Entropía (MEMMs). Algoritmo de Viterbi y beam search . Ingeniería de features, evaluación y análisis de error.
- - Análisis sintáctico: Estructuras de constituyentes y dependencias. Gramáticas Libres de Contexto Probabilísticas (PCFGs). Algoritmo CKY. PCFGs lexicalizadas. PCFGs con variables latentes. Métricas de evaluación.

## 12.- BIBLIOGRAFÍA:

[JM08] Daniel Jurafsky and James H. Martin. Speech and Language Processing, 2nd Edition . Prentice Hall, 2nd edition, May 2008.

[MS99] Christopher D. Manning and Hinrich Schütze. Foundations of statistical



Universidad de Buenos Aires  
Facultad de Ciencias Exactas y Naturales

Ref. Expte. N° 508397 V. 01

Buenos Aires, 18 DIC 2017

**VISTO**

la nota a fs. 3 del Comité Académico del Departamento de Ciencias de la Computación, mediante la cual eleva la información y el programa del curso de posgrado Procesamiento de Lenguaje Natural, que será dictado durante el verano de 2018 por el Dr. Franco M. Luque,

**CONSIDERANDO**

lo actuado por la Comisión de Doctorado,

lo actuado por la Comisión de Posgrado,

lo actuado por este Cuerpo en Sesión Ordinaria realizada en el día de la fecha,

en uso de las atribuciones que le confiere el artículo 113º del Estatuto Universitario,

**EL CONSEJO DIRECTIVO DE LA FACULTAD  
DE CIENCIAS EXACTAS Y NATURALES  
RESUELVE:**

Artículo 1º: Aprobar el dictado del NUEVO curso de posgrado Procesamiento de Lenguaje Natural de 32 hs. de duración.

Artículo 2º: Aprobar el programa del curso de posgrado Procesamiento de Lenguaje Natural obrante a fs. 2 del expediente de la referencia.

Artículo 3º: Aprobar un puntaje máximo de un (1) punto para la Carrera del Doctorado.

Artículo 4º: Comuníquese a la Dirección del Departamento de Ciencias de la Computación, a la Dirección de Alumnos, a la Secretaría de Posgrado y a la Biblioteca de la FCEyN (con fotocopia del programa incluida). Cumplido archívese.

**RESOLUCION CD N° 2986**

Dr. PABLO J. PAZOS  
Secretario Adjunto de Posgrado  
FCEyN - UBA

Dr. JUAN CARLOS REBOREDA  
DECANO