



Propuesta de Trabajo Fin de Grado en Matemáticas (curso 2021-2022)

Responsable de tutorización: Eugenio Martínez Cámara
Departamento: Ciencias de la Computación e Inteligencia Artificial
Correo electrónico: emcamara@decsai.ugr.es

Responsable de cotutorización: Francisco Herrera Triguero
Departamento: Ciencias de la Computación e Inteligencia Artificial
Correo electrónico: herrera@decsai.ugr.es

(Rellenar sólo en caso de que la propuesta esté realizada a través de un estudiante)
Estudiante que propone el trabajo:

Título del trabajo: Modelización matemática y extracción de información en documentos policiales mediante la aplicación de Inteligencia Artificial.

Tipología del trabajo (marcar una o varias de las siguientes casillas):

- Complementario de profundización
- Divulgación de las Matemáticas
- Docencia e innovación
- Herramientas informáticas
- Iniciación a la investigación

Materias del grado relacionadas con el trabajo:

- Informática I y II (1º Curso).
- Álgebra I (1º Curso).
- Probabilidad (2º Curso).
- Análisis Matemático I y II (2º Curso).
- Álgebras, grupos y representaciones (3º Curso).
- Estadística Computacional (4º curso).

Descripción y resumen de contenidos:

Los atestados policiales son documentos en los que los agentes de policía redactan la descripción de una denuncia. Estos documentos se convierten en esenciales, dado que a partir de ellos se construye la denuncia de un ciudadano, y es punto de partida de la investigación policial. En estos documentos subyace gran cantidad de información, y de su modelización y análisis se puede obtener conocimiento de alto nivel.

En este TFG se propone la definición de una representación de alto nivel de documentos policiales, basado en la identificación y extracción de las entidades de carácter personal, geográfico y de interés para el dominio del problema. El fin de esta representación de alto nivel, es facilitar el tratamiento computacional de la información policial con el objetivo de poder comparar, relacionar y extraer conclusiones relacionadas con el dominio policial.

Actividades a desarrollar:

1. Estudio bibliográfico sobre extracción de información.

2. *Análisis de los datos.*
3. *Diseño de la representación y modelización matemática de la información textual.*
4. *Diseño del modelo de inteligencia artificial para la extracción de información.*
5. *Evaluación del modelo.*
6. *Redacción de memoria.*

Objetivos matemáticos planteados

Representación matemática de lenguaje natural

Modelización matemática de características lingüísticas

Modelización matemática del proceso de extracción automático de entidades de interés.

Bibliografía para el desarrollo matemático de la propuesta:

- Manning, C., & Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT press.
- Smith, N. A. (2011). Linguistic structure prediction. *Synthesis lectures on human language technologies*, 4(2), 1-274.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2). Cambridge: MIT press.
- Goldberg, Y. (2017). Neural network methods for natural language processing. *Synthesis lectures on human language technologies*, 10(1), 1-309.

Otras referencias (si procede):

- Quijano-Sánchez, L., Liberatore, F., Camacho-Collados, J., & Camacho-Collados, M. (2018). Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the police. *Knowledge-Based Systems*, 149, 155-168.

Firma del estudiante
(solo para trabajos propuestos por alumnos)

Firma del responsable de tutorización
(solo para trabajos propuestos por estudiantes)

Firma del responsable de cotutorización
(solo para trabajos propuestos por estudiantes)

En, Granada, a de de 2021