



UNIVERSIDAD DE JAÉN

**Anexo II**

**TITULACIÓN: Grado en Biología**

**MEMORIA INICIAL DEL TRABAJO FIN DE GRADO**

**CENTRO: Facultad de Ciencias Experimentales**

**CURSO ACADÉMICO: 2013-14**



UNIVERSIDAD DE JAÉN

*Facultad de Ciencias Experimentales*

**Título del Trabajo Fin de Grado:**

Estudiar y analizar el funcionamiento de técnicas de inteligencia artificial para resolver problemas de biología.

**1. DATOS BÁSICOS DE LA ASIGNATURA**

**NOMBRE:** Trabajo Fin de Grado

**CÓDIGO:** 10216001

**CARÁCTER:** Obligatorio

**Créditos ECTS:** 12

**CURSO:** Cuarto

**CUATRIMESTRE:** Segundo

**2. TUTOR/COTUTOR(en su caso)**

**3. VARIANTE Y TIPO DE TRABAJO FIN DE GRADO (Artículo 8 del Reglamento de los Trabajos Fin de Grado)**

Especifico



UNIVERSIDAD DE JAÉN

#### 4. COMPETENCIAS (\*) Y RESULTADOS DE APRENDIZAJE

##### Competencias generales:

- CG2. Aprender el manejo de las bases de datos y de programas informáticos que pueden emplearse en el ámbito de Ciencias de la Vida
- CG6. Realizar análisis crítico de trabajos científicos y familiarizarse con su estructura.
- CG7. Utilizar las fuentes de información dentro del ámbito de las Ciencias de la Vida.
- CG9. Aplicar los principios básicos del pensamiento y del método científico.

##### Competencias transversales:

- CT1. Adquirir capacidad de gestión de la información, análisis y síntesis
- CT2. Adquirir capacidad de organización planificación y trabajo en grupo
- CT3. Ser capaz de comunicarse correctamente de forma oral y escrita en la lengua materna
- CT4. Conocer una lengua extranjera
- CT5. Ser capaz de resolver problemas y aplicar conocimientos teóricos a la práctica
- CT6. Desarrollar actitudes críticas basadas en el conocimiento
- CT7. Ser capaz de realizar aprendizaje autónomo para el desarrollo continuo profesional
- CT8. Ser capaz de adaptarse a nuevas situaciones y de tomar decisiones
- CT9. Tener sensibilidad hacia temas de índole social y medioambiental
- CT10. Formar profesionales con sólidos valores éticos relacionados con los derechos fundamentales del ser humano, y de modo destacado los relacionados con la igualdad y no discriminación entre seres humanos.

##### Competencias Específicas:

- CE10. Ser capaz de utilizar aplicaciones informáticas para el estudio de biomoléculas
- CE41. Ser capaz de utilizar programas informáticos de análisis de secuencias de ácidos nucleicos y proteínas
- CE53. Realizar análisis filogenéticos

#### Resultados de aprendizaje

<b>Resultado 216001A</b>	Capacidad de integrar creativamente sus conocimientos para resolver un problema biológico real.
<b>Resultado 216001B</b>	Capacidad para estructurar una defensa sólida de los puntos de vista personales apoyándose en conocimientos científicos bien fundados.
<b>Resultado 216001C</b>	Destreza en la elaboración de informes científicos complejos, bien estructurados y bien redactados.
<b>Resultado 216001D</b>	Destreza en la presentación oral de un trabajo, utilizando los medios audiovisuales más habituales.

#### 5. ANTECEDENTES

El trabajo consiste en estudiar el funcionamiento de técnicas de inteligencia artificial para resolver problemas de clasificación con datos biológicos.

Los problemas de clasificación permiten pronosticar el comportamiento del modelo basado en los datos disponibles, lo que puede ser muy útil para resolver problemas biológicos. Entendemos por clasificación el proceso en el que, sabiendo la existencia de ciertas clases o categorías, establecemos una regla para ubicar nuevas observaciones



UNIVERSIDAD DE JAÉN

en alguna de las clases existentes. Las clases resultan de un problema de predicción, donde cada clase corresponde a la salida posible de una función a predecir a partir de los atributos con que describimos los elementos de la base de datos. La necesidad de un clasificador surge por requerimientos de disponer de un procedimiento mecánico mucho más rápido que un supervisor humano y que a la vez pueda evitar sesgos y prejuicios adoptados por un experto. Así mismo, también nos permite evitar acciones costosas y servir de ayuda a los supervisores humanos, sobretodo en casos particularmente difíciles.

En la literatura especializada se han propuesto numerosas estrategias para abordar el problema de la clasificación; desde estrategias puramente estadísticas, hasta redes neuronales, árboles de decisión y reglas difusas. Entre las técnicas de Inteligencia Artificial existentes para resolver los problemas de clasificación el alumno utilizara técnicas básicas como árbol de decisión C4.5 [Qui93], etc..

Para el desarrollo de este trabajo de fin de grado el alumno hará uso de la herramienta Keel (<http://www.keel.es>), la cual contiene una colección de herramientas y algoritmos implementados en Java.

## **6. HIPÓTESIS DE TRABAJO**

Análisis el comportamiento de métodos de clasificación para problemas biológicos.

## **7. BREVE DESCRIPCIÓN DE LAS ACTIVIDADES A REALIZAR**

El trabajo a realizar consiste en estudiar y analizar el comportamiento de distintas técnicas de inteligencia artificial para resolver los problemas de clasificación en Biología.

El alumno utilizara datos de problemas de biología proporcionados por el tutor, obtenidos del repositorio de la UCI (Machine Learning Repository), entre los cuales podemos destacar problemas con datos de secuencias de ADN.

En primer lugar el alumno recopilara y preparara los datos de los problemas proporcionados por el tutor y utilizara métodos de preprocesamiento para tratar valores perdidos, transformar datos y seleccionar características interesantes de los datos.

En segundo lugar el alumno estudiara, aplicara y analizara el comportamiento de varios algoritmos de clasificación seleccionados por el tutor.

Por último creara una memoria justificativa del trabajo desarrollado y de las conclusiones obtenidas.

## **8. DOCUMENTACIÓN/BIBLIOGRAFÍA**

[TSK05] Tan P.-N., Steinbach M., y Kumar V. (2005) Introduction to Data Mining. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition.

[Qui93] J.R. Quinlan. C4.5: Programs for Machine Learning. Morgan Kauffman Publishers, 1993.

[Alc09] J. Alcalá-Fdez, L. Sánchez, S. García, M.J. del Jesus, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Bacardit, V.M. Rivas, J.C. Fernández, F. Herrera. KEEL: A Software Tool to Assess Evolutionary Algorithms to Data Mining Problems. Soft Computing 13:3 (2009) 307-318.



UNIVERSIDAD DE JAÉN

[Alc11] J. Alcalá-Fdez, A. Fernandez, J. Luengo, J. Derrac, S. García, L. Sánchez, F. Herrera. KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework. Journal of Multiple-Valued Logic and Soft Computing 17:2-3 (2011) 255-287.

## 9. CRONOGRAMA PROVISIONAL

- Familiarizarse con la temática del trabajo de fin de grado.
- Recopilación y preparación de datos con los que se va a trabajar.
- Aprendizaje del uso de la herramienta Keel.
- Obtención de resultados con los distintos algoritmo seleccionados.
- Descripción y análisis el comportamiento de los resultados obtenidos.
- Desarrollo de una memoria justificativa del trabajo incluyendo las conclusiones obtenidas.

**Nota informativa:** Para completar este Anexo II se recomienda consultar la guía docente de la asignatura del Trabajo Fin de Grado que está disponible en el siguiente enlace:

[https://uvirtual.ujaen.es/srv/es/informacionacademica/catalogoguiasdocentes/p/2012-13/2/102A/10216001/es/2012-13-10216001\\_es.html](https://uvirtual.ujaen.es/srv/es/informacionacademica/catalogoguiasdocentes/p/2012-13/2/102A/10216001/es/2012-13-10216001_es.html)

**Más información:**

<http://www10.ujaen.es/conocenos/centros/facexp/trabajofingrado>