

RELACION DE PROPUESTAS DE TRABAJOS FIN DE MÁSTER EN INGENIERIA INFORMÁTICA

Los estudiantes matriculados en el primer curso del Máster en el Curso 2017/18 deberán elegir entre alguna de las propuestas que se hacen. Para ello cada alumno deberá ponerse de acuerdo con el Director/es del TFM en el que tengan interés y tras conseguir la aceptación comunicar por correo electrónico el tema elegido al Director del Máster (rmedina@uco.es)

El plazo de elección finalizará el 29 de Junio de 2018.

PROPUESTAS:

- **Título:** Aplicación de nuevas estrategias evolutivas al problema de distribución en planta

- **Director:** Laura García Hernández

Resumen: El problema de distribución en planta ha sido abordado por muchos métodos. La mayoría de ellos sólo toma en cuenta los aspectos que se pueden cuantificar. En la distribución de plantas industriales, es importante considerar aspectos tanto cuantitativos como subjetivos. Este proyecto, persigue analizar novedosas estrategias evolutivas que permitan la interacción entre el algoritmo y el experto humano. La contribución del conocimiento del DM en el enfoque deberá guiar el complejo proceso de búsqueda, ajustándolo a las preferencias del diseñador. Una vez estudiadas y seleccionas las estrategias evolutivas, lo siguiente será implementarlas y aplicarlas al problema de distribución en planta. Posteriormente habrá que realizar una experimentación en profundidad con diferentes problemas tomados de la bibliografía, así cómo un análisis exhaustivo de los resultados obtenidos.

- **Título:** Aplicación móvil para gafas de realidad aumentada
- **Directores:** Enrique Yeguas Bolivar y Rafael Muñoz Salinas

Resumen: En los últimos años la realidad virtual y aumentada ha tenido un gran auge, en gran medida al significativo aumento de la capacidad de cómputo de los dispositivos móviles. En la misma línea, han aparecido numerosos modelos de gafas que permiten convertir cualquier móvil en unas gafas de realidad aumentada y virtual.

Este proyecto pretende desarrollar una aplicación de realidad virtual para dispositivos móviles acoplados a gafas. La temática concreta de la aplicación se concretará entre el alumno y los directores de tesis. En líneas generales, el usuario se colocará las gafas con el móvil y mediante la cámara trasera del mismo, se tomarán imágenes del entorno. Para conocer la posición de la cámara en el entorno, se hará uso de la librería de posicionamiento ArUco, capaz de estimar la pose a partir de una serie de marcas artificiales previamente colocadas. La aplicación deberá mostrar algún tipo de información virtual al usuario que podrá moverse libremente por el entorno de juego.

- **Título:** Aplicación de técnicas de Deep Learning a la estimación de pose humana en imágenes
- **Directores:** Manuel Jesús Marín Jiménez y Rafael Muñoz Salinas

Resumen: El problema de estimación de pose humana a partir de imágenes consiste en obtener la localización 2D o 3D de puntos clave del cuerpo humano, como son las articulaciones (o joints), para así reconstruir la configuración de pose de las diversas partes del cuerpo (cabeza, torso, brazos, piernas, etc). La estimación de pose es algo esencial en diversos ámbitos de la vida como puede ser en rehabilitación médica, entrenamiento de deportistas de alto nivel, o en temas de ocio (videojuegos). En los últimos años, han surgido nuevas técnicas de aprendizaje automático que haciendo uso de un número suficientemente grande de muestras (imágenes en este caso) son capaces de automáticamente aprender modelos para el problema planteado. Estas técnicas se engloban dentro del ámbito conocido como Deep Learning (o Aprendizaje Profundo). Concretamente, este TFM tiene como objetivo la evaluación de técnicas basadas en Redes Neuronales Convolucionales (CNN) para el problema de la estimación de pose humana. Con este TFM, el alumno podrá aprender las técnicas más recientes en Deep Learning, en general, y sobre CNN, en particular, tan demandadas actualmente en grandes empresas tecnológicas a nivel mundial.

- **Título:** Aprendizaje y búsqueda de vecinos con múltiples instancias y múltiples etiquetas
- **Directores:** Eva Gibaja Galindo

Resumen: En aprendizaje multi-instancia (*multi-instance learning*, MIL), en lugar de representar un ejemplo del espacio de entrada con un único vector de características (instancia), se utiliza un conjunto de estos vectores denominado *bolsa de instancias*. Por otro lado, en aprendizaje multi-etiqueta (*multi-label*, MLL), el espacio de salida no está formado por una única clase, sino que un objeto puede pertenecer a varias clases (o etiquetas) simultáneamente. Estos dos marcos de trabajo, estudiados en la asignatura NTAA, se pueden combinar permitiendo así que un patrón esté descrito por una bolsa de instancias y tenga a su vez asociadas múltiples etiquetas. En este caso hablamos de aprendizaje multi-instancia multi-etiqueta (*multi-instance multi-label*, MIML).

En la actualidad existen paquetes software tanto para MIL (Weka) como para MLL (Mulan), pero, salvo algunas implementaciones muy puntuales de algunos algoritmos, no se conoce ningún paquete software específico para MIML. El objetivo del TFM es la implementación del algoritmo kNN para MIML, uno de los algoritmos base en clasificación. Al estar cada patrón formado por una bolsa de instancias, la distancia entre bolsas y la búsqueda de vecinos son tareas no triviales y su solución requiere la implementación de un módulo específico de

búsqueda. Una vez desarrollado este módulo, la implementación de kNN es trivial. La implementación se realizará en Java, utilizando para ello las librerías Weka y MULAN.

- **Título:** Enseñanza Automática Educativa para Ingeniería
- **Director:** Antonio Araúzo Azofra

Resumen: El campo de Machine Teaching (o Enseñanza Automática) formula matemáticamente el problema de optimización de un conjunto de ejemplos de entrenamiento para una actividad formativa. Se basa en suponer el modelo de aprendizaje que va a aplicar un “estudiante”, ya sea máquina o humano. Aunque ésta es una suposición ambiciosa en el ámbito de la formación para humanos, las ventajas de aplicar este modelo pueden ser muy grandes ya que nos acerca a encontrar la formación ideal.

En este TFM, nos planteamos buscar modelos de conocimiento adecuados para una situación formativa concreta en ingeniería, dentro de las que requieren razonamiento y resolución de problemas. Por ejemplo, podemos aplicarla sobre la generación de problemas con grafos PERT con el objetivo de mejorar la experiencia formativa sobre su construcción o su cálculo.

- **Título:** Construcción eficiente de grafos AOA
- **Director:** Antonio Araúzo Azofra

Resumen: La construcción de un grafo tipo AOA (en los que se basa el método PERT) con el menor número de actividades ficticias es un problema NP-completo. Hay diversos algoritmos propuestos con diversos enfoques. La mayoría los tenemos implementados en PPCproject. En este proyecto, se pretende investigar sobre la mejora los existentes, buscando desarrollar un algoritmo que aproveche que los problemas grandes de planificación suelen tener una forma más sencilla que la definición general. Este desarrollo es útil para permitir una representación gráfica más clara y reducir los cálculos de tiempos.

- **Título:** Aplicación web la comparación de frentes de Pareto
- **Directores:** Juan Carlos Fernández y Pedro Antonio Gutiérrez Peña

Resumen: Las soluciones no dominadas aportadas por un frente de Pareto en un espacio de objetivos no son mejores o peores si se comparan unas con otras, pero si puede ser de utilidad el poder comparar un conjunto de soluciones aportadas por diferentes algoritmos o diferentes configuraciones de una misma metodología. El TFM consistiría en crear una aplicación web que implemente diferentes métricas de comparación de frentes de Pareto existentes en la literatura, de tal manera que permita comparar el rendimiento a nivel global de conjuntos de soluciones, por ejemplo que provengan de varios algoritmos multiobjetivo sobre un mismo problema.

- **Título:** Implementación de algoritmos de clasificación ordinal como parte del *framework* de aprendizaje automático *scikit-learn*

- **Directores:** Pedro Antonio Gutiérrez Peña y Juan Carlos Fernández Caballero

Resumen: *scikit-learn* es una de las herramientas más completas y populares en Python para aprendizaje automático y ciencia de datos. Su estructura modular permite incluir algoritmos con bastante facilidad y, una vez integrados, existen una gran cantidad de herramientas que permiten examinar sus resultados y organizar su ejecución. La clasificación ordinal es un tipo especial de clasificación en el que las etiquetas a predecir guardan una relación de orden. *scikit-learn* no dispone de algoritmos de clasificación ordinal, por lo que el objetivo de este TFM sería la implementación de uno o varios algoritmos de clasificación ordinal, incluyendo también las métricas de evaluación correspondientes a este tipo de problemas. De esta forma, se pondría a disposición de la comunidad un conjunto de algoritmos que permitirían resolver mejor los problemas de clasificación ordinal, problemas muy frecuentes en el mundo real.

- **Título:** Diseño e implementación de un algoritmo de arrecifes de coral para el diseño de modelos de redes neuronales artificiales
- **Directores:** Juan Carlos Fernández Caballero, Antonio Manuel Durán Rosal

Resumen: Recientemente, se han desarrollado nuevos algoritmos bio-inspirados llamados *Coral Reefs Optimization (CRO)*. Este tipo de algoritmo se basa en simular el comportamiento sobre la competición entre corales por un espacio dentro de un arrecife. Cada coral representaría una solución a un problema, de forma similar a los individuos de algoritmo genético. Por tanto, el TFM consistiría en adaptar algunos de los algoritmos *CRO* más populares existentes en la literatura al diseño de modelos de Redes Neuronales Artificiales para tareas de clasificación.

- **Título:** Aprendizaje profundo ordinal
- **Director/es:** César Hervás Martínez y Pedro Antonio Gutiérrez Peña

Resumen: Este TFM consistiría en aplicar diversas técnicas de clasificación ordinal a los algoritmos de aprendizaje profundo para así resolver problemas complejos de clasificación (clasificación de imágenes, predicción de series temporales, etc.). Algunos ejemplos de los posibles problemas a abordar incluyen la clasificación de imágenes de retina para detección de retinopatía diabética, estimación de edad a partir de imágenes, análisis de emociones en imágenes, etc...

- **Título:** Técnicas de agregación de evaluaciones propias y de pares para la mejora de la validez y confianza de las mismas
- **Directores:** Carlos García Martínez

Resumen: La evaluación y calificación del alumnado es un proceso complejo sujeto a innumerables críticas. El examen final es uno de los métodos mayormente criticados por no reflejar una situación real en el que un trabajador aplique las competencias que adquirió en su

época de aprendizaje. La autoevaluación y la evaluación por pares son procesos en los que es el propio estudiante el que juzga los logros conseguidos, propios o de sus compañeros, respecto a una tarea determinada. En la actualidad se aplican con bastante aceptación, en cursos abiertos online, para su uso en el cálculo de la calificación de los estudiantes inscritos. Sin embargo, también se estudia su aplicación en cursos presenciales y más reducidos.

Actualmente, la mayoría de estudios aplican esquemas de agregación simples, como el cálculo de la media o la mediana, o aplican metodologías que incrementan considerablemente la carga de trabajo del alumnado, como la revisión por pares calibrada. El objetivo de este TFM sería analizar otras técnicas de agregación que pudiesen mejorar la confianza y validez de las calificaciones finales generadas.

- **Título:** Aplicación de técnicas de generación e ingeniería de características para obtener mejores modelos de predicción del fracaso de los estudiantes en LMS y/o MOOCs
- **Directores:** Cristóbal Romero Morales

Resumen: Aplicación de diferentes técnicas de selección y generación de características sobre los datos de interacción de los estudiantes con los sistemas LMS Moodle y/o MOOCs. El objetivo es mejorar el rendimiento (exactitud) de los modelos de predicción del rendimiento o fracaso (aprobar/suspender la asignatura) de los alumnos en base a la información de los logs proporcionados por el sistema LMS Moodle o MOOCs. A partir de estos ficheros crudos (raw) se pueden obtener una gran cantidad de variables a partir de las cuales se puede predecir el estado final del alumno en el curso (aprueba o suspende). Para mejorar la predicción se han utilizado clásicamente algoritmos de selección de características para seleccionar un subconjunto reducido de las características que más aportan para predecir el estado final. En este TFM se propone utilizar y comparar técnicas de generación e ingeniería de característica para generar automáticamente nuevas características en base a las características originales, que permitan mejoren la predicción de los modelos de clasificación.

- **Título:** Identificación de matrículas mediante smartphones
- **Directores:** Joaquín Olivares y José Manuel Palomares

Resumen: Implementar mecanismos de determinación de la matrícula de vehículos y la identificación de la misma utilizando los sistemas integrados en los smartphones con el envío de los datos para su posterior tratamiento a sistemas centralizados.

- **Título:** Procesamiento distribuido de imágenes mediante la GPU de sistemas empujados
- **Directores:** Joaquín Olivares y José Manuel Palomares

Resumen: Establecer los mecanismos para implementar diversos algoritmos de procesamiento de imágenes en sistemas distribuidos haciendo uso de las GPU integradas en los procesadores de dichos sistemas empotrados. Análisis del rendimiento obtenido ante diferentes grados de optimización de altas prestaciones y uso de la red.

- **Título:** Detección de bordes en imágenes en color
- **Directores:** Nicolás Luis Fernández García

Resumen: Se pretende desarrollar una interfaz gráfica que permita la ejecución de los siguientes algoritmos relacionados con la detección de bordes en imágenes en color:

- Algoritmos de realce de bordes en imágenes en color: generan imágenes monocromáticas donde se realzan los puntos de los bordes
- Algoritmos de umbralización de imágenes monocromáticas: generan los mapas de bordes
- Algoritmos de evaluación de calidad de los mapas de bordes

- **Título:** Mecanismos de interoperabilidad de herramientas basadas en flujos de trabajo en ciencia de datos
- **Directores:** Jose Raul Romero

Resumen: La ciencia de datos consiste en un conjunto de técnicas y teorías de muchos campos dentro de amplias áreas como las matemáticas, la estadística y las tecnologías de la Información. Entre estas técnicas se incluyen el procesamiento de señales, modelos probabilísticos, machine learning, aprendizaje estadístico, ingeniería de datos, el reconocimiento de patrones y la computación de altas prestaciones, entre otras.

Sin embargo, el manejo de estas técnicas requiere de expertos (científicos de datos) que sean capaces de ordenar y secuenciar todas las acciones necesarias como para poder obtener resultados eficaces de forma eficiente. Los flujos de trabajo son una forma de ordenar estas secuencias. En general, un flujo de trabajo es una técnica que permite capturar y modelar el conocimiento del experto como un conjunto de actividades que deben trabajar en coordinación para lograr un determinado objetivo. En el ámbito científico es utilizada para que los científicos puedan modelar los experimentos computacionales a un mayor nivel de abstracción, que les permite utilizar los recursos propios de su dominio de trabajo, sin necesidad de conocer los aspectos de bajo nivel de ejecución. Este tipo de técnicas facilita la definición de procesos por parte de cualquier tipo de usuario y favorece la automatización de su ejecución mediante sistemas de gestión de flujos de trabajo.

- **Título:** Validación del descriptor 'Two-Point Gait' en el dataset AVAMVG
- **Directores:** Francisco Jose Madrid Cuevas

Resumen: El TFM propuesto pretende validar los resultados obtenidos por el descriptor "Two-Point Gait" [1] en el dataset AVAMVG. Para ello el alumno deberá implementar dicho descriptor y hacer una propuesta de fusión multivista del mismo ya que la propuesta original es para una única vista.

[1] Lombardi, S., Nishino, K., Makihara, Y., & Yagi, Y. (2013). Two-point gait: Decoupling gait from body shape. Proceedings of the IEEE International Conference on Computer Vision, 1041–1048. <http://doi.org/10.1109/ICCV.2013.133>.

- **Título:** Escalado de algoritmos de reconocimiento de genes
- **Directores:** Aida de Haro García, Domingo Ortiz Boyer

Resumen:

Actualmente la Bioinformática es uno de los principales campos dentro del Machine Learning. La enorme cantidad de datos que genera unida a la complejidad de los problemas que trata supone uno de los mayores retos a los que se enfrenta la comunidad científica. La predicción de la estructura de los genes es un tema de enorme importancia en Bioinformática. El reconocimiento de genes es una tarea ardua debido a la ausencia de señales claras que permitan determinar un gen, puesto que las medidas a realizar para llevar a cabo esta labor son muy sutiles. El riesgo de obtener un nivel de falsos positivos por encima de lo aceptable es muy elevado en el reconocimiento de aquellos genes con una estructura poco común.

Además de la complejidad del problema, hemos de lidiar con el hecho de que las bases de datos de reconocimiento de genes poseen del orden de billones de nucleótidos. Para la mayoría de los algoritmos de Machine Learning es inmanejable y la escalabilidad del problema supone un reto en sí. Es por ello que proponemos al alumno realizar una revisión de aquellos algoritmos de Machine Learning que escalando el problema posibiliten la reconstrucción precisa del gen completo, desde su codón de inicio al codón de fin.

Como paso final, y para la puesta a prueba de los algoritmos revisados e implementados durante el desarrollo del proyecto, el alumno aplicará una selección de los mismos a diversas bases de datos genómicas reales disponibles.

- **Título:** Selección de Instancias en Clasificación Multietiqueta
- **Directores:** Juan A. Romero del Castillo, Aida de Haro García

Resumen:

En la clasificación multietiqueta cada instancia del conjunto de aprendizaje está asociada a un conjunto de etiquetas L , $|L| > 2$ lo cual permite la categorización de la instancia y la aplicación de dicha técnica a multitud de problemas en multitud de campos, como etiquetado de textos y publicaciones en cualquier soporte, y en Internet y RRSS, clasificación de imágenes, diagnóstico médico, bioinformática, big data, etc.

El lenguaje de programación Python(<http://www.python.org>) se ha convertido en una excelente herramienta para la comunidad científica en diversos campos de investigación y en concreto dentro de lo que denominamos Machine Learning. En este proyecto se trata de poner a punto un laboratorio de experimentación en Machine Learning y concretamente en clasificación multietiqueta mediante el uso de herramientas y librerías Python (Scikit-Learn, NumPy, SciPy, matplotlib, etc.)

Como paso final, y para la prueba del entorno de experimentación creado, el alumno, durante el desarrollo del proyecto, abordará el problema de la selección de instancias en problemas de clasificación multietiqueta aplicando diversas técnicas de investigación y desarrollando alguna aportación fruto de su investigación durante el proyecto.

- **Título:** Sistema para la clasificación de celdas en panales de abejas
- **Directores:** María Luque Rodríguez

Resumen:

Para la investigación de las causas de la reducción de la población mundial de abejas es necesario medir gran cantidad de valores relacionados con estos insectos, entre ellos, el recuento de los tipos de celdas en los panales de abejas. Estas medidas se realizan de forma manual a partir de fotografías digitales y será necesario centrarse en las más importantes. Se pretende diseñar e implementar un sistema basado en conocimiento capaz de aplicar selección de características y clasificar con una precisión adecuada los distintos tipos de celdas presentes en un panal de abejas.

- **Título:** Simulador de optimizaciones del compilador para dispositivos móviles
- **Directores:** Jose Manuel Soto Hidalgo, Ezequiel Herruzo Gómez

Resumen:

Los compiladores y entornos de desarrollo realizan o deberían realizar la adecuación del código o del resultado de la compilación a la arquitectura del sistema donde se ejecuta el programa, o al menos a la arquitectura donde el programa ha sido compilado. Actualmente esto no sucede siempre así, llegando a tener incrementos de rendimiento en la ejecución de programas muy superiores con optimizaciones muy simples. Estas optimizaciones pueden ser realizadas tanto en el propio código fuente desarrollado por el programador, como internamente por parte del compilador en alguna de las fases de compilación. El trabajo fin de máster que se propone pretende desarrollar una herramienta que muestre de forma sencilla qué modificaciones en el código permiten mejorar el rendimiento en la ejecución del programa, contemplando distintos tipos de arquitectura y entornos de programación. La idea final es que esta herramienta permita asimilar y afianzar hábitos de programación eficiente.
Asignatura: Optimizaciones para Computación de Alto Rendimiento

- **Título:** Integración de servicios distribuidos IoT en arquitecturas Cloud
- **Directores:** Jose Manuel Soto Hidalgo, Ezequiel Herruzo Gómez

Resumen:

El Internet de las cosas (IoT) es un término que se refiere a la interconexión digital de objetos cotidianos o cosas con Internet. Aunque las cosas, Internet y la conectividad son los tres componentes principales del IoT, el valor se logra al cerrar la distancia entre el mundo físico y el mundo digital en sistemas que se refuerzan y mejoran automáticamente. Si se pudiera conocer el estado de cualquier cosa del mundo, y se pudiera razonar a partir de los datos extraídos a partir de una serie de dispositivos conectados, podríamos resolver gran cantidad de problemas o retos. Para ello AWS ha incorporado el IoT a servicios específicos, como AWS Greengrass y AWS IoT. Estos le ayudan a recopilar y enviar datos a la nube, facilitando la carga y el análisis de esa información, y le permiten administrar sus

dispositivos, para que pueda concentrarse en desarrollar aplicaciones que satisfacen sus necesidades. El objetivo de este TFM consiste en definir, desplegar e integrar una plataforma de comunicación para IoT en arquitecturas Cloud como base para implementaciones de problemas aplicados a Smart Cities.

- **Título:** SMART LEGO CITIES: Sistema de diseño de soluciones de Inteligencia Ambiental con sensores/actuadores y piezas LEGO
- **Directores:** Jose Manuel Soto Hidalgo, Ezequiel Herruzo Gómez

Resumen:

El objetivo de este TFM consiste en definir un sistema que permitiría a un usuario sin mayores conocimientos técnicos implementar una idea para Inteligencia Ambiental con piezas Lego comercial. Se hará uso de un conjunto de micros (Raspberry o Arduino) y un set de sensores/actuadores que se integrarán con las piezas Lego. Se definirán archivos xml que permitan describir las características del micro, como por ejemplo: un listado de los puertos e interfaces que el micro tiene disponible, conexiones GPIO, etc. También se definirán archivos xml para describir los sensores/actuadores en función de qué pines necesita, qué interfaz utiliza, si es analógico o digital, rango de funcionamiento, etc. También, en los archivos xml, se incorporará el código necesario para la integración de micro/sensor y su funcionamiento. Finalmente, se elaborará un repositorio con los sensores/actuadores y micros más comunes junto con sus archivos xml asociados y, se desarrollará una interfaz gráfica que facilite el uso de éstos a modo que un usuario pueda diseñarse y configurarse su idea para Inteligencia Ambiental y ésta funcione de manera sencilla y transparente al usuario. Por ejemplo, llevar a cabo ideas como: cuando no hay nadie en una habitación, apagar la luz o si la temperatura es superior a 35 grados y no hay luz, encender riego por goteo media hora, todo ello representado con piezas lego e implementado de forma sencilla con la interfaz gráfica.

- **Título:** Modelos Computacionales para la Predicción de la Actividad Biológica.
- **Directores:** Gonzalo Cerruela García, Nicolás García Pedrajas

Resumen:

Actualmente las industrias farmacéuticas, institutos de investigación privados y públicos tienen como uno de los grandes retos el desarrollo de herramientas computacionales para el descubrimiento de nuevos fármacos. En esta línea de investigación existe un especial interés en el desarrollo de metodologías que relacionen numéricamente estructuras moleculares con sus actividades biológicas. Todo esto, con el fin de hacer una predicción teórica de la actividad biológica, evitando pasar por un proceso de prueba y error de síntesis orgánica cuyo coste actual se estima en 1.8 billones de dólares.

El objetivo fundamental del trabajo fin de máster es, el estudio y evaluación en bases de datos moleculares de diferentes métodos computacionales para la predicción de la actividad biológica. Para ello, se utilizarán bases de datos moleculares de dominio público sobre las que serán evaluados, algoritmos de clustering, clasificación, selección de características y diferentes modelos para la representación molecular.

Con este proyecto el alumno se introduce en una de las principales líneas de investigación realizadas en la industria farmacéutica, universidades y centros de investigación, orientadas al descubrimiento de nuevos fármacos.

- **Título:** Estudio comparativo de métodos de clasificación multi-etiqueta
- **Directores:** Nicolás García Pedrajas y Gonzalo Cerruela García.

Resumen:

En la clasificación multietiqueta cada instancia del conjunto de aprendizaje está asociada a un conjunto de etiquetas L , $|L| > 2$, donde a diferencia de la clasificación estándar una instancia puede pertenecer a cualquier subconjunto de L . Esto permite la categorización de la instancia y la aplicación de dicha técnica a multitud de problemas en multitud de campos, como etiquetado de textos y publicaciones en cualquier soporte, y en Internet y RRSS, clasificación de imágenes, diagnóstico médico, bioinformática, big data, etc.

El presente trabajo fin de máster pretende hacer un estudio comparativo de los principales métodos de clasificación multietiqueta existentes para comprobar las fortalezas y debilidades de cada uno de ellos.

Las tareas a realizar en este trabajo fin de máster serían:

1. Implementación de los métodos necesarios en caso de que no haya herramientas que los incluyan. Aquellos métodos ya implementados en herramientas de uso común en minería de datos serán usados en ellas.
2. Selección de un conjunto suficientemente significativo de problemas multietiqueta para que la comparación se significativa.
3. Ejecución experimental de los métodos seleccionados en el apartado 1.
4. Comparación estadística de los resultados y extracción de conclusiones.

- **Título:** Desarrollo de modelos de Deep Learning para Procesamiento del Lenguaje Natural. Aplicación a problemas de inteligencia social.
- **Directores:** Sebastián Ventura Soto

Resumen: Los modelos de aprendizaje profundo (deep learning) se están utilizando para la resolución de múltiples tareas en el ámbito del procesamiento del lenguaje natural (clasificación de textos, extracción de tokens, etc). En concreto, este trabajo centrará en problemas de clasificación de texto como *sentiment analysis* y *topic detection* sobre contenido publicado en Twitter, haciendo uso de modelos basados en aprendizaje profundo como redes convolucionales y redes recurrentes.

Para poder estimar los parámetros de estos modelos, en problemas donde la entrada es una secuencia de unidades lingüísticas, se requiere una representación que retenga la información más discriminativa para una tarea determinada (información contextual, semántica, sintáctica, entre otras).

Por este motivo, en el presente trabajo se propondrán representaciones adecuadas para las tareas tratadas y se realizarán comparaciones entre las diferentes representaciones,

así como entre los diferentes modelos de redes neuronales que hacen uso de dichas representaciones.

Los modelos desarrollados se evaluarán frente a bechmarks desarrollados en talleres de congresos nacionales e internacionales como SemEval e IberEval.

- **Título:** Algoritmo multi-criterio para la recomendación de asignaturas en los Grados Universitarios
- **Directores:** Cristóbal Romero y Amelia Zafra

Resumen:

Los sistemas de recomendación se han convertido hoy día en herramientas esenciales. Debido a la gran cantidad de información disponible para poder tomar cualquier decisión, estas herramientas nos facilitan la tarea, recomendándonos aquellas opciones que nos pueden resultar más interesantes, de acuerdo a nuestros intereses.

En el contexto educativo, también nos encontramos con múltiples opciones. Los estudiantes tienen que enfrentarse a la selección de itinerarios durante sus estudios, la selección de carreras universitarias, y de asignaturas optativas dentro de los estudios. En muchas ocasiones, el estudiante no está seguro de la elección, y es aquí, donde estas herramientas resultan muy útiles, para ayudar a los estudiantes a tomar esta decisión.

En este TFM se propone el diseño e implementación de un algoritmo de recomendación personalizado, que utilizando múltiples criterios, relacionados con los estudiantes, y con las asignaturas, ayude a los alumnos a la selección de las asignaturas optativas dentro de un grado universitario.

- **Título:** Algoritmo basado en el vecino más cercano para trabajar con el aprendizaje con múltiples instancias y múltiples etiquetas
- **Directores:** Eva Gibaja y Amelia Zafra

Resumen:

Conforme las aplicaciones de clasificación se vuelven más complejas, encontrar una representación apropiada de la información disponible para poder realizar una clasificación eficiente, se convierte en una tarea fundamental, a la vez que difícil. La experiencia demuestra, que una representación precisa que sea capaz de reflejar todas las relaciones e interacciones existentes en los datos iniciales, influye directamente en una resolución más efectiva del

problema. En este contexto, el aprendizaje con múltiples instancias y múltiples etiquetas aparece como un marco de aprendizaje que introduce una mayor flexibilidad y ambigüedad en la representación de los objetos, proporcionando una formulación muy natural e intuitiva para representar objetos complejos.

La representación utilizando aprendizaje con múltiples instancias, es una alternativa a la representación con instancias simples y permite representar objetos complejos usando un conjunto de instancias, en lugar de una única instancia. Aunque permite representar objetos complejos, cada objeto debe pertenecer exclusivamente a una única clase. Sin embargo, existen escenarios de clasificación en los que los ejemplos pueden pertenecer a varias clases. En esta situación, el aprendizaje con múltiples etiquetas nos permitiría introducir una mayor flexibilidad en el espacio de salida, ya que cada objeto puede pertenecer a varias clases.

De este modo, el aprendizaje con múltiples instancias y múltiples etiquetas combina ambos marcos de aprendizaje para introducir una mayor flexibilidad tanto en la representación de la entrada, como en las salidas. Cada objeto es representado por una bolsa de instancias y se le permite tener múltiples etiquetas de clase. De este modo, este aprendizaje permite formalizar objetos en diferentes problemas complejos. Por ejemplo, una imagen contiene de forma natural diferentes regiones (instancias) y la imagen completa puede representar diferentes clases, tales como nubes, leones, y paisajes. En categorización de textos, cada documento normalmente consiste en diferentes secciones o párrafos (instancias), y cada documento puede ser asignado a diferentes ítems, tales como deportes, política, y ocio.

En este trabajo, se plantea profundizar en este marco de aprendizaje desarrollando un modelo basado en el vecino más cercano que sea capaz de trabajar en este contexto. La propuesta estaría basada en el trabajo de Zhang¹

¹ Zhang, M. L.: A k-nearest neighbor based multi-instance multi-label learning algorithm. In: Gregoire, E. (eds.) Proceedings of the 22nd IEEE International Conference on Tools with Artificial Intelligence (ICTAI), 2, pp. 207–212. IEEE Computer Society, Los Alamitos (2010)

- **Título:** Fabricación aditiva de piezas tridimensionales a partir de imágenes para su posterior impresión.
- **Directores:** Francisco José Jiménez Hornero y Eduardo Gutiérrez de Ravé Agüera

Resumen:

Este trabajo explorará la impresión 3D de objetos tomado como dato de partida imágenes de los mismos. Con este fin, se estudiarán dos alternativas para modelar los objetos. La primera de ellas se basa en la reconstrucción del objeto a partir distintas fotografías del mismo en las que se varía la posición del observador. La segunda opción consiste en usar imágenes de diferentes secciones del modelo a reconstruir.