

RELACION DE PROPUESTAS DE TRABAJOS FIN DE MÁSTER EN INGENIERIA INFORMÁTICA

Los estudiantes matriculados en el primer curso del Máster en el Curso 2016/17 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 14 de Julio de 2017.

PROPUESTAS:

- **Título: Realidad aumentada para diseño de interiores**

Director: Enrique Yeguas

Resumen: El objetivo del presente Trabajo Fin de Máster es el diseño e implementación de un sistema de Realidad Aumentada para el diseño de interiores. El sistema desarrollado permitiría al usuario añadir, mediante gestos, elementos virtuales (ej. mobiliario) a una estancia en la que se encuentre presente. Ayudando de este modo, por ejemplo, a profesionales (ej. interioristas) y/o a particulares a simular cómo quedaría su vivienda o local una vez que dichos objetos se integrasen en dicho entorno.

- **Título: Sistema para la estimación subpixel de marcadores artificiales**

Director: Rafael Muñoz Salinas

Resumen: La estimación de la posición de la cámara es una tarea importante en muchas aplicaciones de visión por computador. A pesar del gran interés de las marcas naturales (como puntos de interés o texturas), los marcadores artificiales son de gran uso en la actualidad puesto que permiten realizar una rápida estimación de la pose solo siendo necesario su colocación en el entorno de trabajo.

Uno de los problemas básicos en este tipo de sistemas es la localización precisa de las esquinas, que son empleadas para el proceso de estimación de la pose. A mayor la

precisión en la estimación de las mismas, mejor la estimación final de la pose. En este sentido, los mecanismos clásicos de estimación de esquinas necesitan el uso de técnicas adicionales para obtener una precisión subpixel.

Este proyecto pretende abordar el problema de la estimación subpixel de esquinas utilizando técnicas de aprendizaje automático. Para ello, crearemos una base de datos sintética de entrenamiento mediante la aplicación de múltiples transformadas a la imagen original de una esquina. Después, probaremos el uso de Redes Neuronales, Redes Neuronales Convolucionales y Máquinas de Soporte Vectorial para realizar el proceso de entrenamiento. Finalmente, para validar los clasificadores creados, utilizamos escenas reales.

En el transcurso del proyecto el alumno aprenderá en temas relacionados con estimación de puntos de interés, realidad aumentada y técnicas de aprendizaje automático.

- **Título: Evaluación de modelos computacionales para la predicción de la actividad biológica**
Director: Gonzalo Cerruela García

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 (in silico drug design). 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.

- **Título: Soluciones IoT basadas en tecnologías sin contacto y geolocalización para el desarrollo de Smart Cities y su aplicación en Turismo y Marketing**
Directores: Irene Luque Ruiz, Miguel Ángel Gómez Nieto

RESUMEN: Las tecnologías sin contacto, como Near Field Communication (NFC), Bluetooth Low Energy (BLE), así como de geolocalización han demostrado su impacto en el desarrollo de los nuevos modelos de ciudades y negocios en el Internet of Things. Las necesidades de instituciones y empresas de una comunicación directa con el usuario, personalizada, ofreciéndole sus productos y servicios en cualquier lugar y momento, para captar y fidelizar al usuario con nuevas experiencias pueden ser resueltas con aplicaciones móviles en las que estas tecnologías posibilitan esa comunicación.

Los proyectos que se llevarán a cabo estarán destinados a participar en las investigaciones del grupo de investigación ISCBD que desarrolla en estas tecnologías y su aplicación a sectores económicos como son el Turismo, Marketing y la Educación, llevándose a cabo la propuesta de nuevos modelos de interacción, comercio y comunicación, así como el desarrollo de soluciones específicas.

- **Título: Implementación de operadores optimizados para Big Data.**
Directores: Juan Gómez Luna, José Manuel Soto Hidalgo.

Resumen: La actual acumulación de grandes cantidades de datos hace necesario el desarrollo de operadores que permitan un manejo eficiente de los mismos. Los operadores de álgebra relacional (join, filtering, cross product, unique...) son imprescindibles para acceder a los datos deseados mediante técnicas de análisis y minería de datos.

El gran ancho de banda que ofrecen las unidades de procesamiento gráfico (GPU) se puede aprovechar para optimizar esos operadores, pudiendo incrementar su rendimiento decenas de veces sobre el empleo de procesadores secuenciales (CPU).

El objetivo de este TFM es implementar operadores optimizados de álgebra relacional sobre GPU y CPUs many-core.

El alumno se acercará a los conceptos principales, modelos y técnicas para Big Data. Conocerá en detalle los algoritmos que describen los operadores y podrá proponer técnicas de optimización innovadoras para implementarlos eficientemente. Manejará lenguajes de programación paralela como CUDA y OpenCL.

- **Título: Análisis y desarrollo de optimizaciones en videojuegos para dispositivos móviles.**
Director: Ezequiel Herruzo Gómez

Resumen: El desarrollo de software multimedia y videojuegos en sistemas móviles (Smartphones y tablets especialmente) suele realizarse sin tener en cuenta tanto la arquitectura del sistema como las posibilidades de optimización del compilador o entorno de desarrollo. Incluso, en la mayoría de los casos, ni el compilador ni los entornos de desarrollo estándares utilizados para este tipo de programación contemplan la adecuación del código a la arquitectura y jerarquía de memoria del sistema.

En este sentido, este trabajo de fin de máster pretende analizar las posibilidades de optimización de código que incorpora un entorno estándar de desarrollo de videojuegos como es "Unity", el desarrollo de un videojuego con dicho entorno, la aplicación de técnicas de optimización al código del videojuego y un análisis de consumo de potencia, rendimiento y diferencias de temperatura entre el videojuego optimizado y el videojuego desarrollado de forma estándar.

- **Título: Sistema de comunicación en plataforma móvil.**
Director: Ezequiel Herruzo Gómez, Jose Manuel Soto Hidalgo

Resumen: El desarrollo de sistemas de computación empotrada que procesan información de sensores y comunican esta información a bases de datos centralizadas está, actualmente, llegando a un uso práctico en numerosas plataformas industriales. El trabajo fin de máster que se plantea trata de integrar en un mismo dispositivo un sistema de comunicaciones que permita la lectura por radiofrecuencia de información de sensores, procese dicha información en tiempo real y envíe los resultados a una base de datos centralizada. Esto ha de hacerse de forma autónoma e integrada en un sistema móvil (robot, dron, ...) guiado por una trayectoria preestablecida. Se cuenta con el sistema de guiado de la trayectoria, de modo que el TFM se centra en la recopilación de información de los sensores, procesamiento y empaquetado de esa información, envío a una base de datos e incorporación de la información en la base de datos.

- **Título: Desarrollo de modelos gráficos multiplataforma para la librería JFML.**
Director: José Manuel Soto Hidalgo.

Resumen: La librería JFML está basada en una API en Java que permite el diseño y uso de sistemas difusos para resolver problemas en muchos campos de aplicación. Esta librería está basada en el nuevo estándar IEEE 1855-2016 permitiendo, entre otras cosas, la creación de sistemas difusos interoperables y la representación de éstos de forma legible e independiente del hardware entre distintas plataformas. El objetivo de este TFM es desarrollar modelos gráficos basados en la librería JFML que permitan crear, simular, depurar y exportar sistemas difusos en cualquier plataforma de computación acorde al nuevo estándar. El alumno conocerá los conceptos principales de sistemas de control difusos así como la implementación software de sistemas en plataformas de computación.

- **Título: Análisis de las medidas empíricas de calidad de las aproximaciones poligonales de curvas cerradas de dos dimensiones**

Directores: Ángel Carmona Poyato y Nicolás Luis Fernández García

Resumen:

El trabajo consistirá en la aplicación de un algoritmo que genere aproximaciones poligonales óptimas para un contorno. La aproximación poligonal tendrá M vértices, donde M variará desde 3 (triángulo) hasta N (número de puntos del contorno). Para cada aproximación poligonal óptima de M vértices, se calculará el valor de las medidas de calidad. Posteriormente, se generará una gráfica para cada medida de calidad que permita analizar su comportamiento en función del número de vértices M de la aproximación poligonal. Una medida de calidad tendrá un mejor rendimiento si permite

identificar con claridad cual es la aproximación poligonal con M vértices que tiene un mejor ajuste al contorno. Se usaran varios algoritmos para aproximaciones poligonales ya existentes y varias medidas empíricas de calidad habitualmente usadas.

- **Título: Método probabilístico para obtener aproximaciones poligonales óptimas.**

Directores: Nicolás Luis Fernández García y Ángel Carmona Poyato

Resumen:

Con el presente trabajo se pretende desarrollar un método probabilístico que obtenga la aproximación poligonal óptima de M puntos, para un contorno cerrado de N puntos, donde $M < N$. Este problema se conoce como problema de mínima distorsión o $\min\text{-}\epsilon$.

En la actualidad, los únicos métodos que garantizan el óptimo global se basan en la aplicación de un método que garantiza el óptimo local siempre y cuando se fije un punto inicial de la aproximación poligonal.

Los métodos más habituales que garantizan el óptimo local son el método de Pérez y Vidal y el método de Salotti. Para obtener el óptimo global, se aplica cualquiera de ellos N veces, utilizando en cada una de ellas un punto diferente del contorno. De esta forma se obtienen N aproximaciones poligonales que serían óptimas localmente para el punto de inicio utilizado. Seleccionando la mejor de ellas obtendríamos el óptimo global.

Evidentemente esta estrategia es de un alto coste computacional ya que el valor de N suele ser muy elevado en contornos reales.

Existen varias estrategias para encontrar cual sería el punto inicial óptimo (como mínimo habría M puntos iniciales óptimos) que garantice el óptimo global, aunque dichas estrategias tampoco garantizan que se encuentre un punto inicial óptimo.

En este trabajo se pretende implementar un método probabilístico que combine dichas estrategias para encontrar el óptimo global con una alta probabilidad.

- **Título: Aproximaciones a la extracción de reglas de clasificación para aprendizaje multivista con algoritmos evolutivos**

Directores: Carlos García y Sebastián Ventura

Resumen:

Dentro del contexto del aprendizaje automático, el aprendizaje multivista consiste en el entrenamiento de más de un modelo con el objetivo de obtener mejores resultados que entrenando únicamente uno de ellos. Esta metodología es particularmente apropiada cuando la información proviene de varias fuentes, por ejemplo, un mismo estudiante puede venir descrito por sus patrones de comportamiento en una plataforma educativa, como por sus calificaciones en las materias cursadas. Uno de los contextos en los que se usa más frecuentemente el aprendizaje multivista es el semi-supervisado. En este contexto, sólo un pequeño porcentaje de los patrones de entrenamiento fueron etiquetados por el experto, y se desea que los modelos aprovechen las relaciones entre los patrones de entrenamiento, tanto etiquetados como sin etiquetar, para obtener mejores resultados que aprendiendo sólo sobre aquellos etiquetados. Hasta el momento,

en la literatura no se encuentran estudios de extracción de modelos basados en reglas para aprendizaje multivista, lo que constituye el objetivo de esta propuesta de TFM.

- **Título:** Generador de señales sintéticas con comportamiento realístico

Directores: José Manuel Palomares Muñoz y Joaquín Olivares Bueno

Resumen:

En la actualidad, la simulación de sistemas ubicuos requiere la incorporación de multitud de señales de sensores que analizan y muestrean el entorno físico real. Estos flujos de datos permiten la toma de decisiones distribuidas en tiempo de ejecución bajo unos límites temporales ligados a la realidad. A pesar de que existen multitud de repositorios de libre acceso en Internet con datos de sensores en entornos reales, el número de señales disponible no es suficiente para obtener un volumen de datos que permita una variabilidad en el entorno de funcionamiento que habilite su validación. Por ello, se plantea el desarrollo de un generador de señales que sea capaz de construir señales aleatorias con comportamientos similares a los que muestran las señales de los sensores reales. Se analizarán las señales reales y los parámetros que permiten modelar dichas señales. Posteriormente, se ajustarán esos parámetros mediante alguna técnica de Inteligencia Artificial y se comprobará la consistencia de las señales generadas sintéticamente con las señales reales obtenidas de los sensores. Se almacenarán diferentes conjuntos de parámetros que permitirán la generación de diferentes familias de señales, en función del sensor y la variable física que se desee simular. Finalmente, se validarán las señales obtenidas mediante técnicas cuantitativas y cualitativas.

- **Título:** Aplicación de técnicas de Minería de Datos a Moodle y MOOCs para la predicción del rendimiento académico de los estudiantes

Director: Cristóbal Romero Morales

Resumen:

Aplicación de diferentes técnicas de minería de datos y análisis de datos sobre los datos de interacción de los estudiantes con los sistemas Moodle y MOOCs, para la predicción del rendimiento académico. Se estudiarán y evaluarán diferentes aproximaciones como son la generación de modelos de predicción basados en atributos genéricos frente a la utilización de atributos específicos, predicción utilizando diferentes fuentes de datos frente a una única fuente, selección del subconjunto de datos que mejor predice para un determinado datasets, y la detección lo antes posible o detección temprana de alumnos en riesgo de suspender o abandonar el curso. El objetivo del TFM sería:

- Preparación/Preprocesado los datos de Moodle y MOOC.
- Estudio de la aplicación de diferentes técnicas de minería de datos a dichos datos.
- Realización de la experimentación aplicando diferentes algoritmos y técnicas, para compararlos y analizar los resultados obtenidos.

- **Título:** Aplicación de optimización y búsqueda basados en simulación de arrecifes de coral a problemas de selección de características

Director: Antonio Araúzo Azofra

Resumen:

Los algoritmos basados en simulación de arrecifes de coral (CRO) son una meta-heurística novedosa que permite la utilización de estrategias similares a las de otras meta-heurísticas pero con más variaciones y ajustes. Por otra parte, la selección de características (SC) persigue elegir un conjunto de características adecuado para un proceso de aprendizaje. Nos centraremos en el aprendizaje supervisado de clasificadores. Este proyecto pretende investigar la aplicación de CRO a la búsqueda de conjuntos de características adecuados. Se compararan los resultados con otros algoritmos de SC disponibles en una plataforma propia sobre problemas de clasificación públicos.

- **Título:** Técnicas de gamificación en dirección de proyectos

Director: Antonio Araúzo Azofra

Resumen:

Investigar y/o desarrollar aplicaciones que hagan uso de técnicas, elementos y dinámicas propias de los juegos y el ocio con el fin de potenciar la motivación o reforzar las conductas positivas para la correcta ejecución de un proyecto.

- **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 algoritmos 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 de la minería de repositorios para el aprendizaje de buenas prácticas en el desarrollo de software
- **Director:** José Raúl Romero Salguero

Resumen:

El desarrollo de software complejo es un proceso cuyo éxito depende altamente de la experiencia y habilidad de los participantes. Por ello, aprender de proyectos anteriores resulta de especial interés para detectar aquellas buenas prácticas y acciones que lograron obtener los mejores y más efectivos resultados. Sin embargo, extraer cualquier tipo de conocimiento en el ámbito del desarrollo software constituye un reto en sí mismo, pues va a estar oculto entre miles de líneas de código y en otros artefactos software. En este contexto surge la minería de repositorios software (MRS), cuyo objetivo es descubrir información interesante y previamente desconocida a partir de los recursos y plataformas para la gestión y configuración del software, como repositorios software, documentos y artefactos software, foros de programadores o *bug trackers*, entre otros. Estas fuentes se analizan para descubrir aspectos escondidos en esos datos pero relevantes al proceso de la ingeniería del software. Técnicas como la minería de textos, clasificación, minería de reglas de asociación, regresión o *clustering* son habituales en MRS. En esta propuesta de TFM se propone que el estudiante explore estas técnicas para la inspección y mejora de código fuente en base a proyectos históricos, en concreto, el descubrimiento y propuesta de uso de patrones de diseño que mitiguen las deficiencias del programador o los problemas introducidos por los continuos cambios en el fuente.

- **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 Zhang . "A k-nearest neighbor based multi-instance multi-label learning algorithm".

- **Título:** Aplicación de algoritmos genéticos para el remuestreo de patrones en bases de datos no balanceadas.

Directores: Pedro Antonio Gutiérrez Peña y Juan Carlos Fernández Caballero.

Resumen

El TFM consistiría en la utilización de algoritmos genéticos para refinar los resultados obtenidos por mecanismos de remuestreo en bases de datos no balanceadas. Entre dichos mecanismos, figura SMOTE, que se basa en interpolar, de forma aleatoria, parejas de patrones de la clase minoritaria. La idea del proyecto sería refinar dicho algoritmo para que los patrones generados no fuesen aleatorios, generándose con vista a minimizar el error para un clasificador dado.

- **Título:** Mejora y ampliación de una herramienta de programación evolutiva con redes neuronales (NNEP)

Directores: Pedro Antonio Gutiérrez Peña y Juan Carlos Fernández Caballero.

Resumen

El TFM consistiría en refactorizar y organizar las clases y código fuente de un software de Programación Evolutiva con Redes Neuronales Artificiales llamado NNEP (Neural Network Evolutionary Programming), de forma que quede integrado en el entorno de desarrollo integrado (IDE) NetBeans.

Una vez integrado se podrían implementar y/o adaptar algoritmos evolutivos multiobjetivo para el diseño de modelos de Redes Neuronales Artificiales para tareas de clasificación y/o regresión.