

OFERTA DE TRABAJOS FIN DE MÁSTER PARA EL CURSO 2011-2012

Los alumnos que se hayan matriculado del itinerario de investigación deben escoger una línea de trabajo y un tutor para su investigación. Dicho trabajo se utilizará para evaluar tanto las prácticas del itinerario investigador (12 créditos) como el Trabajo Fin de Máster (12 créditos). A continuación se presenta la lista de líneas de investigación ofertadas por los profesores del Máster en este curso académico. Se recomienda a los alumnos del itinerario profesional que escojan su línea de investigación cuanto antes para disponer de más tiempo a lo largo del curso para realizar el trabajo escogido. Para ello deben ponerse en contacto con el director de cada trabajo (que será designado su tutor) e informar al director del Máster de la línea escogida. En caso de que varios alumnos soliciten el mismo trabajo se intentará llegar a un acuerdo o, en su defecto, la Comisión Académica del Máster decidirá la asignación en función de los currículos de los alumnos.

Por su parte, los alumnos matriculados en el itinerario profesional deberán ponerse en contacto con el Director del Máster para organizar su periodo de Prácticas en Empresa (12 créditos). En este caso, la evaluación de las prácticas del itinerario profesional se realizará en función del trabajo realizado en dichas prácticas. A estos alumnos se les asignará un tutor entre los profesores del Máster, que será el responsable de coordinar las prácticas con el tutor de la empresa. Los alumnos matriculados en el itinerario profesional deberán realizar un Trabajo Fin de Máster (12 créditos) que podrá ser una extensión del trabajo realizado en la empresa o un trabajo independiente, que podría ser alguno de los ofertados en las líneas de investigación. En cualquier caso, la dificultad del Trabajo Fin de Máster será menor que la de los alumnos del itinerario de investigación, ya que su carga crediticia es menor.

Listado de trabajos ofertados	
Director:	Dr. José Manuel Martín Ramos
Título:	Aplicaciones de la computación SIMT a los algoritmos Genéticos.
Descripción:	Se estudiará la ventaja que puedan dar los algoritmos paralelos en dispositivos hardware graficos para la reducción y/o mejora de los algoritmos Genéticos. Para ello el alumno deberá adquirir conocimientos de computación CUDA además de los conocimientos necesarios para programar los Algoritmos Genéticos todo ello en el lenguaje C++.
(ASIGNADO A Eduardo Moreno)	
Director:	Dr. José Manuel Martín Ramos
Título:	Aplicaciones de la computación SIMT a los algoritmos de Hormigas.
Descripción:	Se estudiará la ventaja que puedan dar los algoritmos paralelos endispositivos hardware graficos para la reducción y/o mejora de los algoritmos de Hormigas. Para ello el alumno deberá adquirir conocimientos de computación CUDA además de los conocimientos necesarios para programar los Algoritmos de Hormigas todo ello en el lenguaje C++.

Listado de trabajos ofertados

Director: Dr. José Luis Arjona

Título: Linked data.

Descripción: Ser capaz de estructurar la información de la web y enlazar datos de distintos dominios aportaría gran valor a la forma en la que se percibe hoy en día la web. Es por ello, que la investigación en este campo está en auge y los avances tienen una gran repercusión para las empresas. El objetivo que se persigue con la tesina es presentar una nueva propuesta que permita enlazar datos estructurados en la web.

Enlace de interés:

http://www.ted.com/talks/tim_berniers_lee_on_the_next_web.html

Director: Dr. José Luis Arjona

Título: Diseño de un crawler inteligente.

Descripción: Analizar un sitio web, clasificando sus páginas convenientemente, nos permitiría desarrollar crawlers automáticos capaces de simular a un ser humano cuando navega por ese sitio web, favoreciendo la automatización de tareas comunes. En esta tesina nos marcamos como objetivo diseñar un crawler con las características anteriores.

Enlace de interés:

<http://dollar.biz.uiowa.edu/~pant/Papers/crawling.pdf>

Director: Dr. José Luis Arjona

Título: Cloud computing.

Descripción: Desarrollo de adaptadores para conectar las soluciones de integración con aplicaciones en la nube (SaaS - Software as a Service), con el objetivo de integrar en los procesos de negocio, no sólo las aplicaciones locales, sino también aplicaciones que nos proporciona la nube, como por ejemplo Google Docs, Dropbox, ...

Enlace de interés:

http://en.wikipedia.org/wiki/Cloud_computing

Director: Dr. Francisco A. Márquez Hernández

Título: Modelado de Problemas de Regresión de Alta Dimensionalidad utilizando Sistemas Basados en Reglas Difusas.

Descripción: Los problemas de Alta Dimensionalidad (Large Scale Problem) son aquellos que se caracterizan por disponer de un gran número de variables o instancias. Dichos problemas son inabordables en tiempo de computación por un gran número de algoritmos. Para abordar dichos problemas se han utilizado diferentes técnicas en problemas de Clasificación. En el presente trabajo se pretende realizar un estado del arte de qué técnicas serían útiles para problemas de regresión y la aplicación de alguna de ellas.

Listado de trabajos ofertados

Director: Dr. Francisco A. Márquez Hernández
Título: Técnicas de Aprendizaje o Inteligencia Artificial Aplicadas a Sistemas de Comunicaciones WiFi.

Descripción: Existen muchas herramientas que permiten la simulación del comportamiento de los dispositivos en entornos WiFi. Muchas de ellas son educativas y libres o con licencia GNU. En el presente trabajo se pretende realizar un estado del arte de qué técnicas pueden o son utilizadas en algún entorno particular e introducir algún tipo de técnicas de Aprendizaje o Inteligencia Artificial con el fin de mejorar el comportamiento de estos entornos de manera automática. El alumno primeramente hará un estado del arte de qué técnicas son utilizadas en algunos entornos.

Director: Dr. Javier Aroba Paez
Título: Metodologías Híbridas de Soft Computing y Lógica Borrosa

Descripción: En este trabajo se pretende realizar un estudio del estado del arte de los principales algoritmos híbridos de Soft Computing y Lógica Borrosa existentes, analizando con más detalle los últimos estudios realizados en este campo, detallando sus aplicaciones y resultados obtenidos. A partir de este estudio, se seleccionarán algunos de los métodos híbridos que se consideren más relevantes y se realizarán pruebas de funcionamiento y eficiencia, con objeto de poder realizar un estudio comparativo. Como parte final del trabajo, se propondrá, y en su caso implementará, algún tipo de mejora significativa sobre alguno de los métodos híbridos estudiados, haciendo hincapié en algún aspecto del mismo (eficiencia, interpretabilidad de los resultados, interface de usuario, etc.)

Director: Dr. Javier Aroba Paez
Título: Metodologías de Obtención de Sistemas de Reglas Difusas Descriptivas (SRDD) y Aproximativas (SRDA)

Descripción: En este trabajo se pretende realizar un estudio del estado del arte de las principales metodologías de obtención de SRDD y SRDA a partir de conjuntos de datos dados, analizando en detalle los últimos estudios realizados en este campo. A partir de este estudio, se seleccionarán las metodologías que se consideren más relevantes y se realizarán pruebas sobre distintos conjuntos de datos con objeto de poder realizar un detallado estudio comparativo de las mismas. Como parte final del trabajo, se propondrá, y en su caso implementará, algún tipo de mejora significativa sobre alguna de las metodologías estudiadas, incidiendo en algún aspecto de la misma (interface de usuario, tipo de resultados que ofrece, etc.)

Listado de trabajos ofertados

Director: Dr. Francisco José Moreno Velo
Título: Desarrollo en C# de una biblioteca para el desarrollo de sistemas difusos basados en el lenguaje XFSML.
Descripción: XFSML es un nuevo lenguaje de modelado de sistemas difusos complejos basado en XML. El trabajo consiste en desarrollar una biblioteca (.dll) en C# que permita describir los sistemas definidos en XFSML.

Director: Dr. Francisco José Moreno Velo
Título: Algoritmo de extracción de listas ordenadas de reglas difusas
Descripción: Las listas ordenadas de reglas difusas han sido propuestas recientemente como una forma de representación del conocimiento más compacta que las tradicionales listas (desordenadas) de reglas difusas. Se plantea la implementación y prueba de un nuevo algoritmo de extracción de listas ordenadas basado en la combinación de la estrategia "separa y vencerás" con el algoritmo de simplificación de reglas "Quine-McCluskey generalizado".

Director: Dr. Pedro J. Abad Herrera
Título: Estudio de la influencia de las características categóricas en los problemas de clasificación One-Class
Descripción: Los problemas de clasificación one-class son aquellos donde sólo hay disponible información sobre una de las clases. El objetivo de los clasificadores one-class es reconocer la pertenencia o no de un nuevo caso como perteneciente a dicha clase. El fin es, por tanto, definir la frontera alrededor de la clase objetivo, con objeto de reconocer los casos correspondientes a dicha clase. Existen diversos métodos que abordan dicho problema desde distintas perspectivas.

Existen diversas técnicas de clasificación One-Class, pero la mayoría tratan con características numéricas, y se basan en distancias o densidades para calcular la pertenencia de un nuevo caso al conjunto. Cuando en el conjunto de datos a clasificar existen características categóricas (no numéricas), dichas medidas quedan alteradas, falseando, por tanto, la precisión del modelo.

El objetivo del presente trabajo es realizar un estudio estadístico de cómo afecta la presencia de características no numéricas a la precisión de la clasificación, proponiendo, en su caso, tratamientos específicos para este tipo de características.

Listado de trabajos ofertados

Director: Dr. Pedro J. Abad Herrera
Título: Aplicación de técnicas de machine learning a la detección de microaneurismas en imágenes de fondo de ojo.

Descripción: La retinopatía diabética es una complicación ocular de la diabetes que está causada por el deterioro de los vasos sanguíneos que irrigan la retina. El daño de los vasos sanguíneos de la retina puede tener como resultado que estos sufran una fuga de fluido o sangre. Si la enfermedad avanza se forman nuevos vasos sanguíneos y prolifera el tejido fibroso en la retina, lo que tiene como consecuencia que la visión se deteriore, pues la imagen enviada al cerebro se hace borrosa. [Tomado de Wikipedia]

La evaluación cuantitativa del número de microaneurismas es un buen indicador de la progresión de la retinopatía diabética. Sin embargo, una comparación manual de las diferentes imágenes temporales del paciente consume mucho tiempo del especialista.

El objetivo del presente trabajo es aplicar técnicas de machine learning a la detección automática de microaneurismas en imágenes de fondo de ojo. Para ello, se dispondrá de un conjunto de imágenes donde los microaneurismas están identificados y se pretende extraer un modelo que permita caracterizar dicha lesión de forma automática.

(ASIGNADO A Paki Rodríguez)

Directores: Dr. Ana M. Roldán y Dr. Antonio Peregrín Rubio
Título: Data Mining en entornos educacionales.

(ASIGNADO A Marco Antonio Callejo)

Directores: Dr. Ana M. Roldán y Dr. Antonio Peregrín Rubio
Título: Web mining.

Director: Dr. Antonio Peregrín Rubio
Título: Aprendizaje con instancias múltiples.
Descripción: El aprendizaje con instancias múltiples (Multiple Instance Learning, MIL) es una variación sobre el paradigma clásico de aprendizaje supervisado para problemas en los que se tiene información incompleta sobre clases o instancias. A diferencia del modelo clásico, en MIL se clasifican conjuntos de instancias, en lugar de instancias individuales, pero no se sabe a cual de las instancias presentes en el conjunto se debe la clasificación. Nosotros nos centraríamos en el uso de técnicas basadas en sistemas difusos evolutivos y aprendizaje evolutivo de instancias para este problema.

Listado de trabajos ofertados

Director: Dr. Antonio Peregrín Rubio

Título: Multiclasificadores.

Descripción: El diseño de multiclasificadores es un área que ha adquirido especial importancia en el aprendizaje automático en los últimos años. Un multiclasificador (o ensemble) consiste en la combinación de diferentes clasificadores de forma homogénea (todos del mismo tipo) o heterogénea (de tipos distintos) para realizar la clasificación de forma conjunta. Las propuestas más importantes dentro de éste área se conocen como Bagging y Boosting. Bagging consiste en la aplicación iterativa de subconjuntos del conjunto de entrenamiento para ajustar cada uno de los componentes del multiclasificador. Por otro lado, Boosting se centra en el ajuste de los propios componentes, generalmente siguiendo esquemas de pesos, para obtener un multiclasificador final que modele fielmente al conjunto de entrenamiento. Nosotros estamos interesados en el uso de técnicas basadas aprendizaje evolutivo de instancias para el diseño de algoritmos coevolutivos para clasificación y su aplicación en problemas de DM.

Director: Dr. Manuel Maña López

Título: Asistencia a la consulta en sistemas de búsqueda en el dominio médico.

Descripción: Cuando el dominio al que pertenece una colección de documentos es muy especializado, como el caso de la medicina, los usuarios de sistemas de recuperación de información se encuentran con graves problemas para definir su consulta. En general, el conocimiento que tiene el usuario sobre la terminología utilizada en este dominio es muy escaso. El resultado son consultas inadecuadas o poco precisas.

El objetivo de este trabajo es crear un sistema de ayuda al usuario que, haciendo uso de recursos ontológicos y terminológicos, le preste asistencia en la definición de la consulta. Esta asistencia vendrá dada en forma de definiciones, sugerencia de términos y grafos que muestren términos relacionados con el propuesto por el usuario.

Listado de trabajos ofertados

Director: Dr. Manuel Maña López

Título: Detección de la negación en documentos del dominio biomédico en español.

Descripción: Uno de los problemas que aparece en cualquier sistema de búsqueda en texto libre y, particularmente, en el dominio biomédico es el tratamiento correcto de la negación. Con el simple uso de palabras clave en la búsqueda, los sistemas recuperan textos que contienen dichas palabras aunque éstas aparezcan de forma afirmativa o negativa.

En este momento no existe una colección de documentos en español en la que estén identificados los conceptos biomédicos negados (similar al BioScope corpus para el inglés). Por tanto, resultaría inviable aplicar técnicas de aprendizaje automático supervisado. El objetivo de este trabajo es crear un sistema basado en expresiones regulares para la detección de la negación en español.

Director: Dr. Jacinto Mata

Título: Desarrollo de técnicas para integrar recursos léxicos en español en el acceso a la información en el dominio biomédico

Descripción: En los sistemas de recuperación de información es habitual la utilización de recursos léxicos externos para mejorar la eficacia de la consulta. En este sentido, y en el ámbito biomédico, es frecuente hacer uso de tesauros y ontologías como UMLS, SNOMED, Gene Ontology, WordNet, etc. Sin embargo, debido a la escasez de recursos léxicos en español no resulta tan usual incluir este tipo de recursos en los sistemas de recuperación de información cuando se realizan búsquedas en colecciones de documentos en español. Se pretende desarrollar técnicas que mejoren la eficacia en las búsquedas bien mediante una indexación conceptual, o bien mediante la expansión de consultas basadas en recursos externos. Para poder llevar a cabo esta tarea será necesario contar con algún recurso léxico-semántico en dicho ámbito. Se hará uso de SNOMED y de otras ontologías disponibles en español. El objetivo es desarrollar un conjunto de técnicas capaces de mejorar la tradicional búsqueda por palabras clave.

Listado de trabajos ofertados

Director:	Dr. Jacinto Mata
Título:	Estudio de diferentes técnicas de agrupación de genes en rutas metabólicas para mejorar las técnicas de minería de datos en la investigación sobre el cáncer de mama
Descripción:	El objetivo de esta propuesta es el estudio y análisis de diferentes técnicas de agrupación de genes en rutas metabólicas. Se realizará una comparativa entre el conocimiento que se extrae con los datos en bruto (todos los genes) con el obtenido a partir de la reducción de genes basada en las rutas metabólicas. El estudio se focalizará en una base de datos que recoge información sobre pacientes con cáncer de mama. Esta base de datos, como la mayoría de las genéticas, contiene un número elevado de atributos (genes). Mediante esta agrupación se pretende mejorar las técnicas de minería de datos para obtener un conocimiento más interpretable, útil y novedoso para los profesionales de la biomedicina.
Director:	Dr. José Luis Álvarez
Título:	Implementación de un algoritmo de extracción de registros en páginas Web basado en la representación visual de la información en el navegador.
Director:	Dr. José Luis Álvarez
Título:	Polyphone: Extractor de información en Redes Sociales. Implementación y su uso en Sistemas de Recomendación.
Director:	Dr. José Luis Álvarez
Título:	Implementación y análisis de varios operadores de cruce con codificación real en un problema sobre rendimiento de placas solares.