



Propuestas de Trabajo Fin de Grado

Curso 2025-2026

El grupo de investigación DSLAB ofrece una serie de propuestas de Trabajos Fin de Grado dirigidas principalmente a estudiantes de los grados en Ingeniería Informática (y similares), Matemáticas, Ciencia e Ingeniería de Datos e Inteligencia Artificial.

El DSLAB (Data Science Laboratory) es un grupo de investigación de la Universidad Rey Juan Carlos especializado en el análisis de datos y el desarrollo de modelos basados en Machine Learning e Inteligencia Artificial. Nuestro trabajo abarca aplicaciones en ámbitos muy diversos, como el deporte, la energía, el medioambiente, la salud y la sostenibilidad, siempre con un enfoque práctico y colaborativo. A través de proyectos con datos reales y socios externos, buscamos transformar la información en conocimiento útil para la toma de decisiones y la innovación tecnológica.

Si eres alumno y estás interesado en alguno de los temas, puedes ponerte en contacto con el coordinador del grupo: isaac.martin@urjc.es.

1. Sistemas de recomendación en el ámbito deportivo

Desarrollo de modelos de Machine Learning para apoyar la toma de decisiones estratégicas antes de una competición. El objetivo es explorar diferentes deportes y escenarios, aplicando técnicas de predicción y simulación.

2. Modelos de retorno a la práctica deportiva tras una lesión (Return to Play)

Construcción de modelos predictivos para estimar el tiempo óptimo de recuperación de atletas, considerando diferentes tipos de lesiones y variables médicas/deportivas. Colaboración con entidades deportivas.

3. Predicción del riesgo de lesión en deportistas

Aplicación de métodos de análisis de datos y aprendizaje automático para anticipar la probabilidad de sufrir una lesión en función de historial, carga de entrenamiento y otras métricas relevantes. Colaboración con entidades deportivas.

4. Detección temprana de talento deportivo

Uso de datos longitudinales para identificar y predecir el rendimiento futuro de jóvenes deportistas, con énfasis en la proyección de trayectorias a medio y largo plazo.

5. Evaluación y seguimiento de la carga de entrenamiento

Diseño de métricas y modelos para relacionar la intensidad y volumen de entrenamiento con el rendimiento y la prevención de lesiones. Colaboración con entidades deportivas.

6. Análisis de factores psicológicos y contextuales en el rendimiento deportivo

Exploración de cómo variables no estrictamente físicas (motivación, presión competitiva, entorno) pueden integrarse en modelos predictivos de rendimiento. Colaboración con entidades deportivas.

7. Modelización estadística y generación automática de informes en el ámbito sanitario

Diseño de un sistema de análisis exploratorio aplicado al Registro Nacional de Fracturas de Cadera (RNFC), con especial atención a la construcción de modelos estadísticos y predictivos (regresión, análisis de supervivencia, modelos de riesgo, etc.). El objetivo es desarrollar técnicas que permitan automatizar la síntesis de resultados, estimar distribuciones de variables clínicas y evaluar factores de riesgo asociados a la recuperación de los pacientes. La línea incluye también la generación de informes automáticos que integren tanto el análisis descriptivo como los modelos matemáticos implementados. Proyecto en colaboración con una entidad sin ánimo de lucro.

8. Modelos predictivos aplicados al turismo sostenible

Colaboración en el desarrollo de herramientas basadas en análisis de datos y aprendizaje automático para anticipar tendencias, optimizar recursos y fomentar prácticas sostenibles en el sector turístico. Colaboración con empresa.

9. Sistema de tutoría inteligente basado en agentes con LLMs

El objetivo del TFG es diseñar y evaluar un sistema de tutoría personalizado para estudiantes universitarios. El sistema estará compuesto por agentes especializados (planificador, evaluador, generador de recursos) que cooperan entre sí y se comunican mediante un LLM. El tutor adaptará explicaciones, ejercicios y retroalimentación en función del progreso del estudiante. Se busca demostrar cómo la orquestación de agentes con un LLM central puede mejorar la calidad de la enseñanza, permitiendo experiencias de aprendizaje dinámicas, escalables y ajustadas al perfil de cada alumno.

10. Agente autónomo para asistencia en investigación científica con LLMs

El TFG consistirá en desarrollar un agente multirol que apoye a investigadores en la búsqueda, síntesis y análisis de literatura académica. El sistema integrará un LLM con módulos-agentes especializados: uno encargado de recuperación de información, otro de resumen y otro de generación de hipótesis. El agente podrá responder preguntas complejas, sugerir conexiones entre artículos y proponer posibles líneas de investigación. Se evaluará su utilidad en distintos campos (ej. informática, biomedicina), así como su capacidad para ahorrar tiempo y mejorar la calidad del trabajo investigador.

11. Interpretación automática de dashboards con LLMs e IA Generativa

Desarrollo de sistemas que utilicen LLMs e IA generativa para proporcionar explicaciones automáticas, claras y adaptadas de los dashboards.

12. Optimización automática de prompts para LLMs

Diseño de herramientas capaces de generar y ajustar de manera autónoma prompts, optimizando el proceso de prompt y context engineering en modelos extensos de lenguaje.

13. Plataforma de análisis de datos de una liga de pádel

Diseño de sistema de recogida y almacenamiento de datos de la liga. Implementación de visualizaciones y estadísticas de datos. Sistema predictivo de ganador por trimestre.

14. Sistema de detección de patrones en video para tenis individual

Desarrollo de un sistema automático de identificación de jugadores, pista y pelota para video monocular de tenis profesional.

15. Dashboard de medidas de complejidad

Creación de un dashboard interactivo para explorar, usando distintas medidas de complejidad, la dificultad de clasificación de un conjunto de datos. El dashboard contendrá un resumen de las medidas de complejidad usadas y distintos gráficos exploratorios con los que el usuario podrá jugar. Las medidas de complejidad pretenden cuantificar aquellos factores que impactan negativamente en el proceso de aprendizaje de los clasificadores (por ejemplo, solapamiento, falta de densidad, presencia de ruido, etc.).

16. ¿Qué efectos tiene la imputación de datos? Una aproximación con medidas de complejidad

Estudiar cómo impacta la imputación de datos en la complejidad de los datos y qué métodos de imputación modifican en menor medida la distribución de los datos.

17. Integración de APIs en aplicaciones shiny para compartición de datos

Analizar las alternativas para implementar APIs de compartición de datos en aplicaciones shiny (lenguaje de programación R), incluyendo plumber, e implementar al menos una solución en un piloto.

18. Técnicas de Nowcasting para estadística pública con Ciencia de Datos

Uno de los retos de la estadística pública es la actualidad de los datos. A menudo se publican con mucho retraso por restricciones legales o metodológicas. Las técnicas de nowcasting pueden proporcionar estimaciones para los datos presentes aún no publicados. En este trabajo se explorarían las técnicas existentes y se propondrían mejoras con los métodos de ciencia de datos modernos.

19. Técnicas de agregación y desagregación de datos espaciotemporales

Estudiar las técnicas de agregación y desagregación de datos espaciotemporales. A menudo se dispone de datos con distinta resolución espacial (municipio, provincia, región, país...) y temporal (diaria, mensual, anual... que conviene agregar o desagregar para poder realizar comparaciones. En este trabajo se explorarían las técnicas existentes aplicadas a datos de estadística pública.