PROPUESTA DE TRABAJO DE FIN DE GRADO

Curso 2026-27

1.- Titulaciónes para la que se ofrece el TFG: Grado en Matemáticas y Estadística / Grado en Matemáticas y Ciencia de Datos / Grado en Ingeniería Matemática / Grado en Matemáticas / Doble Grado en Economía – Matemáticas y Estadística / Doble Grado en Matemáticas – Física / Doble Grado en Ingeniería Informática - Matemáticas

2.- Título del TFG propuesto: Aprendizaje por refuerzo distribucional: de la teoría a la práctica. Aplicaciones en finanzas y control de videojuegos.

3.- Descripción del TFG (objetivos, metodología): Estudio teórico, computacional y aplicado del marco del aprendizaje por refuerzo distribucional (DRL). El DRL desplaza el foco habitual del aprendizaje por refuerzo desde la aproximación del valor esperado $Q(s,a)$ de las recompensas o retornos de la acción a en el estado s, a la estimación de la distribución completa de estos retornos, con base en la correspondiente variable aleatoria $Z(s,a)$. A nivel teórico, se revisarán primero los fundamentos clásicos de procesos de decisión de Markov, ecuación y operador de Bellman y análisis de su convergencia mediante el teorema de punto fijo, para luego realizar la transición de estas nociones al contexto distribucional, formulando la ecuación de Bellman distribucional y llevando a cabo el análisis teórico de la convergencia del operador correspondiente bajo la métrica de Wasserstein.

La base anterior se puede desarrollar en múltiples direcciones. En primer lugar, hacia la revisión de los fundamentos de las *Deep Q-Networks* (DQN), a partir de la cuales se estudiará el modelo *Quantile Regression* DQN, que plantea una extensión distribucional de las DQN basado en el empleo de cuantiles de la distribución de retornos, y en la que tendrá un papel relevante el análisis de la pérdida de regresión de Huber y sus propiedades matemáticas para aproximar la métrica de Wasserstein. El contexto de aplicación que se plantea es la resolución y control automático de videojuegos.

En segundo lugar, hacia el estudio de las medidas de riesgo coherentes (axiomas de Artzner) y el análisis de métricas como el *Value at Risk* (VaR) y su versión condicional, CVaR, profundizando en el uso de la transformada inversa para modelar geométricamente la curva completa de la distribución de $Z(s,a)$ y extraer de forma natural las métricas de riesgo sin una discretización previa. Computacionalmente, se estudiará e implementará el algoritmo *Implicit Quantile Networks* (IQN) para el diseño de políticas aversas al riesgo en un contexto de toma de decisiones bajo incertidumbre, como la optimización de carteras financieras.

4.- Profesor/profesores tutores: Juan Tinguaro Rodríguez González

5.- Requisitos (si procede):

6.- Número de plazas: 2

7.- Nombre del alumno (en su caso):