



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

FICHA TÉCNICA DE LA ACTIVIDAD FORMATIVA

Datos básicos de la Actividad Formativa	
Name	Explainable Artificial Intelligence: Use of Neural Networks in Social Sciences for Explanatory Models
Type	Elective
Length	20 hours
Term	See calendar



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

Contents

The course addresses how to use neural network models with post-hoc analysis in Social Sciences. Neural network (NN) models are widely used across multiple disciplines due to their enormous potential to capture non-linear patterns. However, in the field of Social Sciences, the use of NNs is much more limited, mainly due to their lack of transparency (interpretability issues: NNs are often referred to as "black box" models). Various techniques have been developed to mitigate this lack of transparency, such as the LIME or SHAP methods, which assess the importance of the explanatory variables in estimating the output variable. However, these tools are qualitative (the importance measures allow for comparison between variables, but it is difficult to truly interpret the impact of that importance). Nonetheless, recent research conducted by Pizarroso et al. (2022) has proposed an innovative methodology (NeuralSens) based on sensitivity calculations, which interprets NN models just as it is done in a linear regression model. That is, it allows for a direct interpretation of the variables and their importance. Therefore, NNs are no longer "black box" models. The enormous advantage of this methodology is that NNs are very powerful for identifying complex relationships in data, regardless of whether or not one is aware of their existence, as NNs do not require a prior functional specification.

This technique is very attractive for scientific journals in Social Sciences, both for its novelty and its power. One of the instructors of this course has already published 4 JCR papers in Social Sciences using this methodology, proving that NNs can be used not only for predictive purposes but also for developing explanatory models currently.

In this course, starting from zero level (no prior knowledge is required), students will be provided with the necessary knowledge to start working with NeuralSens in their research. All modules will include a practical part, where the explained theory will be applied.

1. Module 1: Introduction to Programming in R (2 hours - optional). **No prior programming knowledge is required. Students who wish will have access to a "button-based" app similar to Stata or Gretl, designed specifically for the course, which will allow them to work with NeuralSens without the need for coding.**
2. Module 2: Review of Classical Causal Models: Linear Regression and Logit Models (2 hours)
3. Module 3: Introduction to Neural Networks (4 hours)
4. Module 4: NeuralSens (6 hours)
5. Module 5: Alpha Curves (2 hours)
6. Module 6: Hypothesis Testing with NeuralSens (1 hour)
7. Module 7: Final Practical Case (3 hours)



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

Learning outcomes	<p>This course will contribute to the development of the following skills and abilities:</p> <p>CB11 Systematic understanding of a field of study and mastery of research skills and methods related to that field.</p> <p>CB12 Ability to conceive, design or create, implement and adopt a substantial investigative or creative process.</p> <p>CA03. Design, create, develop and undertake novel and innovative projects in their field of knowledge.</p> <p>The fundamental learning outcomes will be:</p> <ul style="list-style-type: none">• Understand the fundamentals of neural network models, their characteristics, and application scope.• Be capable of fitting a neural network model given a dataset.• Be able to interpret the results obtained in a neural network model for explanatory purposes.• Learn to use the results obtained from a neural network model to identify nonlinear effects.• Be capable of identifying local effects in causal neural network models using the Alpha curves methodology.
Language of instruction	Spanish



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

Assessment

To obtain the pass rating, it is necessary to have attended 75% of the sessions and deliver, within 4 weeks of the end of the last session a simple practical application exercise that demonstrates the acquisition of the basic skills of the course.

At the beginning of each course, the student will talk with his tutor the plan of activities to develop in the course in question, which will be included in the individualized plan of activities.

Likewise, each doctoral school and IQS School or Management will open a registration period for the activities planned for each program. In the event that the number of applications exceeds the number of places, the Program Committee (in the case of specific activities) or the Committee of each doctoral school (in the case of transversal activities) will decide on admission, taking into account the interest of activity in the formative progress of the researcher.

The professor will deliver the transcript with the assessment of each student to the AC of the corresponding program, who will receive the grade of "Pass" or "Fail". In the event that the evaluation is "Pass", the activity will be included in the individualized record of student activities. If the evaluation is "Fail", the doctoral student will have to agree with his tutor to take a new course depending on the offer.

Mobility

NA



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

FICHA TÉCNICA DE LA ACTIVIDAD FORMATIVA

Datos básicos de la Actividad Formativa	
Título	Inteligencia Artificial Explicable: uso de redes neuronales en Ciencias Sociales para modelos explicativos
Carácter	Transversal
Tipo	Optativa
Duración	20 horas
Planificación temporal	A realizar a lo largo de la duración de los estudios de doctorado
Contenidos	<p>El curso aborda cómo emplear modelos causales de redes neuronales con análisis post-hoc en Ciencias Sociales. Los modelos de red neuronal (NN) se utilizan ampliamente en múltiples disciplinas debido a su enorme potencial para capturar patrones no lineales. Sin embargo, en el campo de las Ciencias Sociales en general, el uso de NNs es mucho más limitado, principalmente debido a su falta de transparencia (problemas de interpretabilidad: las NNs se denominan frecuentemente modelos de "caja negra"). Se han desarrollado varias técnicas para mitigar esta falta de transparencia, como los métodos LIME o SHAP, que evalúan la importancia de las variables explicativas para estimar la variable de salida. Sin embargo, estas herramientas son cualitativas (las medidas de importancia permiten la comparación entre variables, pero es difícil interpretar realmente el impacto de esa importancia). No obstante, investigaciones recientes llevadas a cabo por Pizarroso et al. (2022) han propuesto una metodología innovadora (NeuralSens) basada en cálculos de sensibilidad, que interpreta los modelos de NN de forma idéntica a como se hace en un modelo de regresión lineal. Es decir, permite una interpretación directa de las variables y su importancia. Por lo tanto, las NNs ya no son modelos de "caja negra". La enorme ventaja que esta metodología presenta es que las NNs son muy poderosas para identificar relaciones complejas en los datos, independientemente de que se tenga o no consciencia de su existencia, pues las NNs no requieren una especificación funcional a priori.</p>



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

Resultados de aprendizaje	<p>Este curso contribuirá al desarrollo de las competencias y capacidades:</p> <p>CB11 Comprensión sistemática de un campo de estudio y dominio de las habilidades y métodos de investigación relacionados con dicho campo. CB12 Capacidad de concebir, diseñar o crear, poner en práctica y adoptar un proceso sustancial de investigación o creativo. CA03 - Diseñar, crear, desarrollar y emprender proyectos novedosos e innovadores en su ámbito de conocimiento.</p> <p>Los resultados de aprendizaje fundamentales serán:</p> <ul style="list-style-type: none">• Conocer los fundamentos de los modelos de redes neuronales, sus características y ámbito de aplicación.• Ser capaz de ajustar un modelo de red neuronal dado un conjunto de datos.• Ser capaz de interpretar los resultados obtenidos en un modelo de red neuronal con fines explicativos.• Aprender a utilizar los resultados obtenidos en un modelo de red neuronal para identificar efectos no lineales.• Ser capaz de identificar efectos locales en modelos causales de redes neuronales usando la metodología de las curvas Alpha.
Lengua	Español



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

Procedimiento de control

Para obtener la calificación de apto es preciso haber asistido al 75 % de las sesiones y entregar, en el plazo de 4 semanas a la finalización de la última sesión, y obtener la valoración de apto, un ejercicio de aplicación práctica sencillo que demuestre la adquisición de las competencias básicas del curso. Al inicio de cada curso, el estudiante consensuará con su tutor el plan de actividades a desarrollar en el curso en cuestión, lo que se recogerá en el plan de actividades individualizado.

Asimismo, en cada escuela de doctorado y en IQS School o Management se abrirá un período de inscripción en las actividades planificadas para cada programa. En caso de que el número de solicitudes excediera al de plazas, la Comisión de Programa (en el caso de actividades específicas) o la Comisión de cada escuela de doctorado (en el caso de actividades transversales) decidirán sobre la admisión, atendiendo al interés de la actividad en el progreso formativo del investigador.

La evaluación de los cursos tendrá un carácter continuo a través de la asistencia y participación en las clases y las actividades entregadas por el doctorando.

El profesor entregará el acta con las calificaciones a la comisión académica del programa con la evaluación de los estudiantes, que recibirán la calificación de "Apto" o "No Apto". En caso de que la evaluación sea "Apto", la actividad será recogida en el registro individualizado de actividades del estudiante. Si la evaluación fuera "No Apto", el doctorando tendrá que consensuar con su tutor la realización de un nuevo curso en función de la oferta.

Actuaciones de movilidad

No aplica