COMILLAS
UNIVERSIDAD PONTIFICIA
ICAI ICADE CIHS

GUÍA DOCENTE
2025 - 2026

FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura

Nombre completo	Inteligencia Artificial Aplicada a Ciberseguridad
Código	DOI-MCS-522
Impartido en	Máster Universitario en Ingeniería de Telecomunicación y Máster en Ciberseguridad [Segundo Curso] Máster en Ciberseguridad [Primer Curso]
Nivel	Master
Cuatrimestre	Semestral
Créditos	4,5 ECTS
Carácter	Optativa
Departamento / Área	Departamento de Organización Industrial
Responsable	Juan Pablo Fuentes Brea
Descriptor	El objetivo de la asignatura es conocer cómo a partir de la aplicación de técnicas de Inteligencia Artificial (IA) sobre casos de uso de ciberseguridad, se consigue un salto cualitativo y cuantitativo a nivel de detección y protección de los activos de información de una compañía gracias a la cualidad predictiva que éstas aportan. Aunque la IA puede ayudar a la defensa de dichos activos de información, también puede ser utilizada como mecanismo de ataque sobre los mismos, siendo una herramienta cada vez más empleada por los ciberatacantes. De esta forma, se presentarán los mecanismos tanto de defensa como de ataque mediante IA cercanos al estado del arte, realizando un repaso a los principales escenarios de aplicación en ciberseguridad, como son la detección de malware, análisis de fraude y análisis de comportamientos, deepfakes y casos de uso empleando IA generativa.

Datos del profesorado

Profesor

Nombre	Juan Pablo Fuentes Brea
Departamento / Área	Departamento de Organización Industrial
Correo electrónico	jpfuentes@icai.comillas.edu

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura

Prerrequisitos

Conocimientos básicos de machine learning, ciberseguridad y lenguaje Python.

Competencias - Objetivos

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

Temario

Presentación de la asignatura

- Objetivos
- Temario
- Prácticas
- Evaluación
- Frameworks
- Contacto
- Bibliografía

Detección de fraude

- Presentación de escenarios de fraude
- Detección de Fraude mediante Scoring estático
- Detección de Fraude mediante Scoring dinámico
- Introducción a Deep Learning

UEBA

- Presentación de escenarios UEBA
- Análisis temporal de actividades
- Modelos basados en autoencoders
- Modelos recurrentes
- Introducción a Pytorch

IA Responsable

- Ley de IA
- Aspectos generales
- Niveles de Riesgos
- Plan de acción
- Ciberseguridad

IA en el SDLC

- Seguridad en el ciclo de vida de IA
- Ataques de Poisoning & Evasion
- Controles en la capa de datos
- Controles en la capa de desarrollo
- Controles en la capa de despliegue

Adversarial ML

- Presentación de escenarios Adversarial ML
- Métodos de ataque
- Métodos de defensa
- Seguridad en el ciclo de vida de IA

Despliegue de Modelos IA

- Entornos de despliegue
- Tecnologías de despliegue de modelos IA
- Securización del despliegue
- Otros entornos de despliegue IA

Differential Privacy

- Problemática sobre privacidad de datos
- Membership Inference Attack
- Differential Privacy

Deepfakes

- Presentación de escenarios Deepfakes
- GANs
- Fake faces
- Fake news
- Fake speech

LLMs Security

- Introducción a LLMs
- Transformers
- RAG
- Aplicaciones en Ciberseguridad
- Ciberseguridad en IA agéntica

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

El uso de IA para crear trabajos completos o partes relevantes, sin citar la fuente o la herramienta o sin estar permitido expresamente en la descripción del trabajo, será considerado plagio y regulado conforme al Reglamento General de la Universidad.

Actividades de evaluación:

Prácticas por cada uno de los bloques (35%):

- Aplicación de IA en Ciberseguridad
- Ataque y Defensa de modelos IA
- Ciberseguridad con IA generativa

Trabajo Final de la asignatura (35%): elaboración de un proyecto final de la asignatura, cuya defensa será presencial en clase.

Examen final (30%)

Calificaciones

El mecanismo de calificación de la asignatura estará compuesto por una parte teoría y otra práctica, cuya suma deberá ser como mínimo de 5.0 para superar la asignatura. Los porcentajes de ambas partes serán las siguientes:

Nota final = 0.30 * (nota teoría) + 0.35 (nota prácticas) + 0.35 (nota proyecto final)

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

Freeman and Chio. Machine Learning and Security. O'Reilly Media 2018

Duda, Hart and Stork. Pattern Classification. Wiley & Sons 2001

Shawe-Taylor & Cristianini. Kernel Methods for Pattern Analysis. Cambridge 2004

Gollmann. Computer Security. Wiley & Sons, 2011

Szor. The Art of Computer Virus Research and Defense. Addison-Wesley, 2005

Rieck. Machine Learning for Application-Layer Intrusion Detection, Lulu 2009

Probabilistic Machine Learning - Murphy 2023