Universidad Nacional de Ingeniería (UNI)
Escuela Profesional de
Ciencia de la Computación
Sílabo 2027-I

1. CURSO

AI364. Inteligencia Artificial para Ciberseguridad I (Electivo)

2. INFORMACIÓN GENERAL

2.1 Curso	:	AI364. Inteligencia Artificial para Ciberseguridad I
2.2 Ciclo	:	CS300 Ciclo.
2.3 Créditos	:	4
2.4 Horas	:	2 HT; 4 HP;
2.5 Duración del periodo	:	16 semanas
2.6 Condición	:	Electivo
2.7 Modalidad de aprendizaje	:	Presencial
2.8 Prerrequisitos	:	▪ AI264. Aprendizaje Profundo. (7^{mo} Sem) ▪ AI265. Aprendizaje por Refuerzo.

3. PROFESORES

Atención previa coordinación con el profesor

4. INTRODUCCIÓN AL CURSO

Este curso cubre técnicas avanzadas de IA aplicadas a ciberseguridad, incluyendo detección de amenazas mediante aprendizaje automático, generación de adversarios y defensas autónomas contra ataques modernos.

5. OBJETIVOS

- Implementar sistemas de detección de intrusos basados en IA
- Diseñar pruebas de estrés para modelos de seguridad
- Analizar implicaciones éticas del uso ofensivo de IA

6. RESULTADOS DEL ESTUDIANTE

AG-C02) Ética: Aplica principios éticos y se compromete con la ética profesional y las normas de la práctica profesional de la computación. (Usage)

AG-C09) Diseño y Desarrollo de Soluciones: Diseña, implementa y evalúa soluciones para problemas complejos de computación. (Assessment)

AG-C12) Usa enfoques sistémicos para seleccionar, desarrollar, aplica, integrar y administrar tecnologías seguras de computación para lograr los objetivos del usuario. (Assessment)

7. TEMAS

Unidad 1: Fundamentos de IA en Cybersecurity (12 horas)	
Resultados esperados: AG-C02	
Temas	**Aprendizaje esperado (*Learning Outcomes*)**
▪ Taxonomía de amenazas asistidas por IA ▪ Arquitecturas de modelos para análisis de malware ▪ Datasets de seguridad (CIC-IDS, NSL-KDD)	▪ Clasificar amenazas usando redes neuronales [Evaluar] ▪ Preprocesar logs de seguridad para modelos ML [Usar]
Lecturas : [AL23]	

Unidad 2: Ataques y Defensas (24 horas)	
Resultados esperados: AG-C02	
Temas	**Aprendizaje esperado (*Learning Outcomes*)**
▪ Adversarial examples en modelos de seguridad ▪ Generación automática de exploits (Fuzzing con RL) ▪ Sistemas de detección autoadaptativos	▪ Implementar ataques adversarios a modelos ML [Usar] ▪ Diseñar defensas metamórficas [Evaluar]
Lecturas : [GSS14], [AR20]	

Unidad 3: Ética en Cybersecurity (8 horas)	
Resultados esperados: AG-C02	
Temas	**Aprendizaje esperado (*Learning Outcomes*)**
▪ Armas cibernéticas autónomas ▪ Responsabilidad en vulnerabilidades IA-generadas ▪ Regulaciones internacionales (AI Cyber Initiative)	▪ Evaluar riesgos de sistemas autónomos [Familiarizarse] ▪ Diseñar protocolos de rendición de cuentas [Usar]
Lecturas : [AL23]	

8. PLAN DE TRABAJO

8.1 Metodología

Se fomenta la participación individual y en equipo para exponer sus ideas, motivándolos con puntos adicionales en las diferentes etapas de la evaluación del curso.

8.2 Sesiones Teóricas

Las sesiones de teoría se llevan a cabo en clases magistrales donde se realizarán actividades que propicien un aprendizaje activo, con dinámicas que permitan a los estudiantes interiorizar los conceptos.

8.3 Sesiones Prácticas

Las sesiones prácticas se llevan en clase donde se desarrollan una serie de ejercicios y/o conceptos prácticos mediante planteamiento de problemas, la resolución de problemas, ejercicios puntuales y/o en contextos aplicativos.

9. SISTEMA DE EVALUACIÓN

Sesiones Teóricas:
Las sesiones de teoría se llevan a cabo en clases magistrales donde se realizarán actividades que propicien un aprendizaje activo, con dinámicas que permitan a los estudiantes interiorizar los conceptos.

Sesiones de Laboratorio:
Para verificar que los alumnos hayan alcanzado el logro planteado para cada una de las unidades de aprendizaje, realizarán actividades que les permita aplicar los conocimientos adquiridos durante las sesiones de teoría y se les propondrá retos que permitan evaluar el desempeño de los alumnos.

Exposiciones individuales o grupales:
Se fomenta la participación individual y en equipo para exponer sus ideas, motivándolos con puntos adicionales en las diferentes etapas de la evaluación del curso.

Lecturas:
A lo largo del curso se proporcionan diferentes lecturas, las cuales podrían ser evaluadas. El uso de herramientas Online permite a cada estudiante acceder a la información del curso, e interactuar fuera de aula con el profesor y con los otros estudiantes.

La nota final NF depende de de las siguientes notas:

- EP: Examen Parcial.
- EF: Examen Final.
- P: Evaluacion permanente (Promedio de prácticas y Trabajos calificados).

La nota final NF se calcula así:

$$NF = \frac{EP + EF + P}{3}$$

10. BIBLIOGRAFÍA BÁSICA

[GSS14] Ian Goodfellow, Jonathon Shlens y Christian Szegedy. *Explaining and Harnessing Adversarial Examples*. ICLR, 2014.

[AR20] Hyrum Anderson y Phil Roth. ■Metamorphic Testing for Adversarial Defenses■. En: *ACM CCS* (2020).

[AL23] Giovanni Apruzzese y Pavel Laskov. *AI for Cybersecurity*. Springer, 2023.