National University of Engineering (UNI)
School of Computer Science
Syllabus 2027-I

1. COURSE

AI364. AI for Cyber Security I (Elective)

2. GENERAL INFORMATION

2.1 Course	:	AI364. AI for Cyber Security I
2.2 Cycle	:	CS300 Cycle.
2.3 Credits	:	4
2.4 Hours	:	2 HT; 4 HP;
2.5 Duration of the period	:	16 weeks
2.6 Type of course	:	Elective
2.7 Learning modality	:	Face to face
2.8 Prerrequisites	:	• AI264. Deep Learning. (7^{th} Sem) • AI265. Reinforcement Learning.

3. PROFESSORS

Meetings after coordination with the professor

4. INTRODUCTION TO THE COURSE

This course covers advanced AI techniques applied to cybersecurity, including ML-based threat detection, adversarial generation, and autonomous defenses against modern attacks.

5. GOALS

- Implement AI-powered intrusion detection systems
- Design stress tests for security models
- Analyze ethical implications of offensive AI use

6. COMPETENCES

AG-C02) Ethics: Applies ethical principles and commits to professional ethics and the standards of professional computing practice. (Usage)

AG-C09) Design and Development of Solutions: Designs, implements, and evaluates solutions for complex computing problems. (Assessment)

AG-C12) Uses systemic approaches to select, develop, apply, integrate, and manage secure computing technologies to achieve user objectives. (Assessment)

7. TOPICS

Unit 1: AI/Cybersecurity Fundamentals (12 hours)	
Competences Expected: AG-C02	
Topics	**Learning Outcomes**
• AI-assisted threat taxonomy • Model architectures for malware analysis • Security datasets (CIC-IDS, NSL-KDD)	• Classify threats using neural networks [Evaluar] • Preprocess security logs for ML models [Usar]
Readings : [AL23]	

Unit 2: Attacks & Defenses (24 hours)	
Competences Expected: AG-C02	
Topics	**Learning Outcomes**
• Adversarial examples in security models • Automated exploit generation (RL-based fuzzing) • Self-adaptive detection systems	• Implement adversarial attacks on ML models [Usar] • Design metamorphic defenses [Evaluar]
Readings : [GSS14], [AR20]	

Unit 3: Cybersecurity Ethics (8 hours)	
Competences Expected: AG-C02	
Topics	**Learning Outcomes**
• Autonomous cyber weapons • Liability for AI-generated vulnerabilities • International regulations (AI Cyber Initiative)	• Evaluate risks of autonomous systems [Familiarizarse] • Design accountability protocols [Usar]
Readings : [AL23]	

8. WORKPLAN

8.1 Methodology

Individual and team participation is encouraged to present their ideas, motivating them with additional points in the different stages of the course evaluation.

8.2 Theory Sessions

The theory sessions are held in master classes with activities including active learning and roleplay to allow students to internalize the concepts.

8.3 Practical Sessions

The practical sessions are held in class where a series of exercises and/or practical concepts are developed through problem solving, problem solving, specific exercises and/or in application contexts.

9. EVALUATION SYSTEM

Theory Sessions:
The theory sessions are held in master classes with activities including active learning and roleplay to allow students to internalize the concepts.

Lab Sessions:
In order to verify their competences, several activities including active learning and roleplay will be developed during lab sessions.

Oral Presentations:
Individual and team participation is encouraged to present their ideas, motivating them with additional points in the different stages of the course evaluation.

Reading:
Throughout the course different readings are provided, which are evaluated. The average of the notes in the readings is considered as the mark of a qualified practice. The use of our virtual campus allows each student to access the course information, and interact outside the classroom with the teacher and with the other students.

The final grade NF depends on the following grades:

- PE: Partial Exam.
- FE: Final Exam.
- P: Permanent Evaluation (Average of practical work and graded assignments).

The final grade NF is calculated as follows:

$$NF = \frac{PE + FE + P}{3}$$

10. BASIC BIBLIOGRAPHY

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

[AR20] Hyrum Anderson and Phil Roth. "Metamorphic Testing for Adversarial Defenses". In: *ACM CCS* (2020).

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