COURSE DESCRIPTION CARD - SYLLABUS

Course name
Fundamentals of Information Analysis [S1Cybez1>PAI]

Course

Field of study
Cybersecurity

Year/Semester
3/6

Area of study (specialization)
–

Profile of study
general academic

Level of study
first-cycle

Course offered in
Polish

Form of study
full-time

Requirements
elective

Number of hours

Lecture
16

Laboratory classes
16

Other
0

Tutorials
0

Projects/seminars
16

Number of credit points

3,00

Coordinators

dr hab. inż. Mariusz Żal
mariusz.zal@put.poznan.pl

dr hab. inż. Sławomir Hanczewski
slawomir.hanczewski@put.poznan.pl

Lecturers

Prerequisites

Basic knowledge of the Python programming language Understanding of the fundamentals of applied mathematics, including statistics and linear algebra Ability to use data analysis tools (e.g., Pandas, NumPy, scikit-learn) Basic understanding of natural language processing (NLP) techniques

Course objective

The aim of the course is to introduce students to information analysis with an emphasis on natural language processing (NLP) in the context of cybersecurity. The course focuses on the preparation and processing of textual data, understanding the mechanisms behind popular NLP models and tools, and leveraging their potential in threat analysis, detecting social engineering attacks, and other aspects related to the security of information and telecommunication systems and networks.

Course-related learning outcomes

Knowledge:

• The student understands the role of information analysis in attack detection, threat monitoring, and risk assessment. [K1_W05]
• Understands key concepts (tokenization, stemming, lemmatization, embeddings) and their importance in text processing. [K1_W16]
• Is familiar with the most widely used libraries and knows their appropriate applications. [K1_W09]
• Understands models such as BERT, GPT, and word2vec, along with their capabilities and limitations. [K1_W06]
• Knows data preprocessing techniques. [K1_W05]

Skills:
• Is able to clean and format textual data. [K1_U02]
• Can select the appropriate tool for specific tasks (text classification, sentiment analysis, named entity recognition). [K1_U02]
• Can implement language models (e.g., word2vec, BERT) and interpret the obtained results. [K1_U04]
• Is capable of processing and analyzing content related to attacks, as well as identifying patterns. [K1_U04]

Social competences:
• Is aware of regulations concerning personal data, such as GDPR, and the ethical implications of text analysis. [K1_K05]
• Is able to collaborate within a team. [K1_K05]
• Demonstrates a willingness for continuous learning and openness to others' opinions. [K1_K01][K1_K02]

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

The knowledge acquired during the lecture is assessed through a written or oral exam.
• Written Exam: Students must answer 7-10 questions (both multiple-choice and open-ended), each assigned different point values. There are three or four point groups.
• Oral Exam: Students draw one question from each point group. For each drawn question, an additional related question may be asked. The evaluation considers both the range of the response and the depth of understanding.
Each exam consists of 50-60 questions in total. A minimum of 50% of the possible points is required to pass.
Project Evaluation
Skills acquired through projects are assessed based on project presentations. The evaluation criteria include:
• Engagement in project preparation
• Tools used
• Additional knowledge acquired beyond the core material
Projects can be completed individually or in pairs. Grading scale: 2.0 - 5.0.
Laboratory Evaluation
Skills acquired in laboratory sessions are continuously assessed.
• Each laboratory session includes exercises graded on a 0 to 10-point scale.
• A minimum of 50% of the total points is required to pass the lab exercises.
Grading Scale
Score Percentage Grade
≤ 50% 2.0 (Fail)
51% - 60% 3.0
61% - 70% 3.5
71% - 80% 4.0
81% - 90% 4.5
91% - 100% 5.0

Programme content

The program includes practical preparation of textual data for analysis, starting from data cleaning and standardization to selected techniques such as lemmatization and representation (e.g., word embeddings). Participants will become familiar with popular NLP libraries, enabling them to implement

solutions for text classification, named entity recognition, and phishing detection.
A key aspect of the course is its cybersecurity context-students will explore how NLP can be applied to analyze logs, emails, and social media posts for potential threats.
Additionally, the course covers ethical considerations, including privacy and data anonymization, as well as the ability to present results in user-friendly reports and visualizations.

Course topics

• Introduction to Information Analysis - with a special focus on applications in cybersecurity
• Text Data Preparation
• Fundamental NLP Tools and Libraries
• Application of Popular NLP Models
• Text Analysis and IT Security
• Visualization and Reporting
• Latest Trends and Research Directions

Teaching methods

• Problem-based lectures, combined with case study analysis.
• Laboratory exercises, focused on implementing and testing NLP tools, working with real or near-real data.
• Team projects, where students develop comprehensive solutions for threat detection

Bibliography

Basic:
1. Jurafsky, D., Martin, J. H. Speech and Language Processing, Pearson, 2020.
3. Goldberg, Y. Neural Network Methods for Natural Language Processing, Morgan & Claypool Publishers, 2017.

Additional:
1. Clark, K., Luong, M.-T., Le, Q. V., Manning, C. D. - ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators (2020).
2. Blogs, documentation, and tutorials, including official documentation of NLP tools and libraries.
3. Cyber threat reports (e.g., CERT Polska, ENISA, Cisco Talos) - for observing real-world applications of NLP in cybersecurity.

Breakdown of average student's workload

	Hours	ECTS
Total workload	88	3,00
Classes requiring direct contact with the teacher	48	1,50
Student's own work (literature studies, preparation for laboratory classes/ tutorials, preparation for tests/exam, project preparation)	40	1,50