Syllabus
Course Program

Date of mining

Specialty
125 – Cybersecurity and information protection

Institute
Educational and Scientific Institute of Computer Science and Information Technology

Educational program
Cybersecurity

Department
Cybersecurity (328)

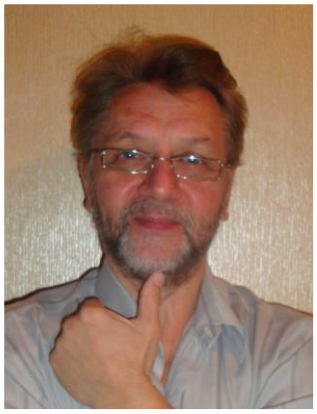
Level of education
Bachelor's level

Course type
Profile training, Selective

Semester
4

Language of instruction
English

Lecturers and course developers



Oleksandr MILOV

oleksandr.milov@khpi.edu.ua

Doctor of technical sciences, professor of the cyber security department of National Technical University "Kharkiv Polytechnic Institute".

Author of more than 200 scientific and educational and methodological works. Academic supervisor for protected candidate theses, guarantor of the educational and professional program of the second (master's) level of higher education. Leading lecturer in the disciplines: "Mathematical foundations of cryptology and cryptanalysis", "Data structures", "Industrial and office espionage", "Digital forensics", for undergraduate and graduate students, Section "Methodology of scientific and pedagogical activity in the sciences of cyber protection" for postgraduate students.
More about the lecturer on the department's website

Stanislav MILEVSKYI

Stanislav.Milevskyi@khpi.edu.ua

Candidate of economic sciences, associate professor of the department of cybersecurity of National Technical University "Kharkiv Polytechnic Institute".

Author of more than 100 scientific and educational and methodological works. Scientific Guarantor of the educational and scientific program of the second (master's) level of higher education. Leading lecturer in the disciplines: "Fundamentals of Mathematical Modeling of Security Systems", "English in Academic Applications", "Modeling of Cyber-Physical Actions" for undergraduate and graduate students.

More about the lecturer on the department's website

General information

Summary

The academic discipline "Data mining" is an optional academic discipline. The study of the discipline helps to form knowledge and skills in the field of data analysis in order to develop a complete model of functioning and development of a real business.

Course objectives and goals

Formation of students' system of theoretical knowledge and practical skills on the basics, principles and methods of intellectual data analysis. As a result of mastering the discipline, students will receive a fundamental base of knowledge on the basics, modern methodology and features of the application of intelligent data processing; study and master Data Mining standards; will acquire the ability to work with Data Mining systems of various purposes and different problem orientations; will receive practical skills for intelligent data analysis based on computational intelligence methods, including large and poorly structured data, their operational processing and visualization of analysis results in the process of solving applied problems of information protection systems and processes for forming the security contour of business processes in computer systems based on data-mining technologies and prevention of external cyber threats.

Format of classes

Lectures, laboratory classes, consultations, self-study. Final control - exam.

Competencies

GC-2. Knowledge and understanding of the domain and understanding of the profession.
GC-5. Ability to search, process and analyze information.
PC-2. Ability to use information and communication technologies, modern methods and models of information security and/or cyber security.
PC-5. The ability to ensure the protection of information processed in information and telecommunication (automated) systems for the purpose of implementing the established information and/or cyber security policy.
PC-8. Ability to carry out incident management procedures, conduct investigations, provide them with an assessment.
PC-10. Ability to apply methods and means of cryptographic and technical protection of information at objects of information activity.
PC-12. Ability to analyze, identify and evaluate possible threats, vulnerabilities and destabilizing factors to the information space and information resources in accordance with the established policy of information and/or cyber security.

Learning outcomes

LO-1. Apply knowledge of state and foreign languages in order to ensure the effectiveness of professional communication;
LO-2. Organize own professional activity, choose optimal methods and ways of solving complex specialized tasks and practical problems in professional activity, evaluate their effectiveness;
LO-3. Use the results of independent search, analysis and synthesis of information from various sources for the effective solution of specialized tasks of professional activity.
LO-4. Analyze, argue, make decisions when solving complex specialized tasks and practical problems in professional activity, which are characterized by complexity and incomplete determination of conditions, be responsible for the decisions made.
LO-5. Adapt under the conditions of frequent changes in the technologies of professional activity, to predict the final result.
LO-6. Critically understand the main theories, principles, methods and concepts in education and professional activity.
LO-7. Act on the basis of the legislative and regulatory framework of Ukraine and the requirements of relevant standards, including international ones in the field of information and/or cyber security.
LO-8. Prepare proposals for regulatory acts on ensuring information and/or cyber security.

LO-9. Implement processes based on national and international standards for detection, identification, analysis and response to information and/or cyber security incidents.
LO-10. Perform analysis and decomposition of information and telecommunication systems.
LO-11. Perform analysis of connections between information processes on remote computer systems.
LO-12. Develop threat and intruder models.
LO-13. Analyze projects of information and telecommunication systems based on standardized technologies and data transmission protocols.
LO-14. Solve the task of protecting programs and information processed in information and telecommunication systems by hardware and software tools and evaluate the effectiveness of the quality of the decisions made.
LO-15. Use modern hardware and software of information and communication technologies.
LO-16. Implement complex information security systems in the automated systems (AS) of the organization (enterprise) in accordance with the requirements of regulatory and legal documents.
LO-17. Ensure the processes of security and functioning of information and telecommunication (automated) systems based on practices, skills and knowledge, regarding structural (structural-logical) schemes, network topology, modern architectures and models of security of electronic information resources with a reflection of relationships and information flows, processes for internal and remote components.
LO-18. Use software and software-hardware complexes for the security of information resources.
LO-19. Apply theories and methods of protection to ensure information security in information and telecommunication systems.
LO-20. Ensure the functioning of special software to protect information from destructive software influences, destructive codes in information and telecommunication systems.
LO-21. Solve tasks of provision and support (including: review, testing, accountability) of the access control system according to the stated security policy in information and telecommunication (automated) systems.
LO-22. Solve the management procedures of identification, authentication, authorization of processes and users in information and telecommunication systems according to the established policy of information and/or cyber security.
LO-23. Implement measures to prevent unauthorized access to information resources and processes in information and telecommunication (automated) systems.
LO-24. Solve the problems of managing access to information resources and processes in information and telecommunication (automated) systems based on access management models (mandatory, discretionary, role-based).
LO-25. Ensure the introduction of accountability of the access management system to electronic information resources and processes in information and telecommunication (automated) systems using event registration logs, their analysis and stated protection procedures.
LO-26. Implement measures and ensure the implementation of processes of prevention of unauthorized access and protection of information, information and telecommunication (automated) systems based on the reference model of interaction of open systems.
LO-27. Solve problems of data flow protection in information and telecommunication (automated) systems.
LO-28. Analyze and evaluate the effectiveness and level of security of resources of various classes in information and telecommunication (automated) systems during tests in accordance with the established policy of information and/or cyber security.
LO-29. Evaluate the possibility of realization of potential threats of information processed in information and telecommunication systems and the effectiveness of the use of complexes of protection means under the conditions of realization of threats of various classes.
LO-30. Assess the possibility of unauthorized access to elements of information and telecommunication systems.
LO-31. Apply protection theories and methods to ensure the security of elements of information and telecommunication systems.
LO-32. Solve the tasks of managing the processes of restoring the regular functioning of information and telecommunication systems using backup procedures in accordance with the stated security policy.
LO-33. Solve the problems of ensuring the continuity of business processes of the organization on the basis of risk management theory.

LO-34. Participate in the development and implementation of an information security and/or cyber security strategy in accordance with the goals and objectives of the organization.
LO-35. Solve the tasks of providing and supporting complex information security systems, as well as countering unauthorized access to information resources and processes in information and information-telecommunication (automated) systems in accordance with the stated policy of information and/or cyber security.
LO-36. Detect dangerous signals of technical means.
LO-37. Measure the parameters of dangerous and interfering signals during the instrumental control of information security processes and determine the effectiveness of information security against leakage through technical channels in accordance with the requirements of regulatory documents of the technical information security system.
LO-38. Interpret the results of special measurements using technical means, monitoring the characteristics of information and telecommunication systems in accordance with the requirements of regulatory documents of the technical information security system.
LO-39. Carry out attestation (based on accounting and survey) of regime territories (zones), premises, etc. under the conditions of compliance with the secrecy regime, recording the results in the relevant documents.
LO–40. Interpret the results of special measurements using technical means, control of ITS characteristics in accordance with the requirements of regulatory documents of the technical information security system.
LO-41. Ensure the continuity of the event and incident logging process based on automated procedures.
LO-42. Implement processes of detection, identification, analysis and response to information and/or cyber security incidents.
LO-43. Apply national and international regulatory acts in the field of information security and/or cyber security to investigate incidents.
LO-44. Solve the problems of ensuring the continuity of the organization's business processes on the basis of risk management theory and the stated information security management system, in accordance with national and international requirements and standards.
LO-45. Apply early classes of information security and/or cyber security policies based on risk-based access control to information assets.
LO-46. Analyze and minimize the risks of information processing in information and telecommunication systems.
LO-47. Solve the problems of protection of information processed in information and telecommunication systems using modern methods and means of cryptographic protection of information.
LO-48. Implement and maintain intrusion detection systems and use cryptographic protection components to ensure the required level of information security in information and telecommunications systems.
LO-49. Ensure the proper functioning of the monitoring system of information resources and processes in information and telecommunication systems.
LO-50. Ensure the functioning of software and software-hardware complexes for detecting intrusions of various levels and classes (statistical, signature, statistical-signature).
LO-51. Maintain operational efficiency and ensure configuration of intrusion detection systems in information and telecommunication systems.
LO-52. Use tools for monitoring processes in information and telecommunication systems.
LO-53. Solve problems of software code analysis for the presence of possible threats.
LO-54. Be aware of the values of a civil (free democratic) society and the need for its sustainable development, the rule of law, the rights and freedoms of a person and a citizen in Ukraine.

Student workload

The total volume of the course is 90 hours (3 ECTS credits): lectures - 16 hours, laboratory classes - 16 hours, self-study - 58 hours.

Course prerequisites

Mathematical foundations of cryptology, Information security of the state, Basics of programming, Basics of cryptographic protection.

Features of the course, teaching and learning methods, and technologies

In the course of teaching the discipline, the teacher uses explanatory-illustrative (informational-receptive) and reproductive teaching methods. Presentations, conversations, and master classes are used as teaching methods aimed at activating and stimulating the educational and cognitive activities of applicants.

Program of the course

Topics of the lectures

Topic 1. Intelligent data analysis - DataMining.

The concepts of DataMining are considered in detail. The emergence, prospects, problems of DataMining are described. A look at DataMining technology as a part of the information technology market is given. The concept of data is considered in detail. The meaning of the concepts object and attribute, sample, dependent and independent variable is explained. The types of scales are discussed in detail. Different types of data sets are given. The concepts of database and DBMS are briefly considered. Describes the stages of DataMining and the actions performed within these stages. Known classifications of DataMining methods are considered. The comparative characteristics of some methods based on their properties are given. The main essence of DataMining tasks and their classification is characterized. The concepts of "information", "knowledge", as well as comparison and comparison of these concepts are considered in detail.

Topic 2. Visual data analysis — VisualMining.

Issues of visual data analysis are considered. The characteristics of data visualization tools, visualization methods and methods of geometric transformations are given. Pixel-oriented methods, as well as hierarchical image analysis and icon display methods, are compared. Methods and means of visual presentation of information are considered, in particular, ways of presenting information in one-, two-, and three-dimensional dimensions, as well as ways of displaying information in more than three dimensions. The principles of qualitative visualization are described. The main trends in visualization are outlined.

Topic 3. Analysis of text information — TextMining.

Text analysis tasks are formulated (stages of text analysis, text pre-processing, TextMining tasks).The stages of text analysis are considered, such as the extraction of key concepts from the text (a general description of the process of extracting concepts from the text, the stage of local analysis, the stage of integration and conclusion of concepts), classification of text documents (description of tasks of text classification, methods of classification of text documents), methods of text clustering documents (guidance of text documents, hierarchical methods of text clustering, binary methods of text clustering), text annotation (execution of text annotation, methods of extracting fragments for annotation). Various tools for analyzing text information are compared (Oracle tools - Oracle Text, IBM tools - Intelligent Miner for Text, SAS Institute tools - Text Miner, Mega-computer Intelligence - Text Analyst tools).

Topic 4. Extracting knowledge from the Web - WebMining.

The problems of analyzing information from the Web, the stages of WebMining, WebMining and other Internet technologies, as well as the categories of WebMining are considered. Methods of extracting Web content are described (extraction of Web content in the process of information search, extraction of Web content for the formation of databases), as well as methods of extracting Web structures (presentation of Web structures, assessment of the importance of Web structures, search for Web documents from taking into account hyperlinks, clustering of Web structures). The results of research on the use of Web resources are given (research information, preprocessing stage, template extraction stage, template analysis stage and their application).

Topic 5. Process analysis tools - ProcessMining.

Considered means of automating the execution of business processes (business processes, formalization of business processes, Workflow systems, service-oriented architecture, design of business processes). Process analysis was performed (ProcessMining technology, protocol analysis, MXML protocol recording standard, ProcessMining tasks, protocol analysis problems). ProcessMining methods are compared (the first probabilistic ProcessMining methods, the method of constructing a disjunctive Workflow scheme, (□-algorithm, methods based on genetic algorithms). The library of Process Mining-Pro algorithms (Pro architecture, ProImport Framework) is described.

Topic 6. Search for associative rules - RulesMining.
The problem statement is completed. Forms of presentation of results considered (classification rules, classification trees, mathematical functions), methods of construction of classification rules (1-rules construction algorithm, NaiveBayes method), as well as methods of construction of classification trees ("divide and conquer" method, coverage algorithm), construction methods of mathematical functions (general view, linear methods, method of least squares, nonlinear methods, SupportVectorMachines (SVM), regularization networks (RegularizationNetworks), discretizations and sparse meshes). The formulation of the problem of finding associative rules (formal formulation of the problem, sequential analysis, types of problems of finding associative rules), algorithms (the Apriori algorithm, a variant of the Apriori algorithm) are considered.
Topic 7. Distributed data analysis.
Mobile agent systems are considered (basic concepts, standards of multi-agent systems, mobile agent systems, JADE mobile agent system). The use of mobile agents for data analysis is demonstrated (problems of distributed data analysis, analytical agents, options for distributed data analysis). A distributed data analysis system was built (a general approach to system implementation, an agent for collecting information about the database, an agent for collecting statistical information, an agent for solving a single task of intelligent data analysis, an agent for solving an integrated task of intelligent data analysis)..

Topics of the workshops

Not provided for in the curriculum.

Topics of the laboratory classes

Topic 1. Skills for working with the Deductor program.
Topic 2. Visual data analysis — VisualMining.
Topic 3. Analysis of text information — TextMining. Programs Text Analyzer, VaalMini.
Topic 4. Extracting knowledge from the Web - WebMining. Program C5.4.
Topic 5. Means of process analysis - ProcessMining. Using ProM.
Topic 6. Search for associative rules - RulesMining. See 5 program.
Topic 7. Distributed data analysis. Agent modeling using JADE and Xelopes.

Self-study

A student's independent work is one of the forms of organization of learning, the main form of mastering educational material in free time from classroom training. During independent work, students study lecture material, do individual homework, prepare for tests, tests and exams. Students are also recommended additional materials (videos, articles) for self-study and analysis.

Non-formal education

Within the framework of non-formal education, according to the relevant Regulation (http://surl.li/pxssv), the educational component or its individual topics may be taken into account in the case of independent completion of professional courses/trainings, civic education, online education, vocational training, etc.
Subjects are not considered for this component in case of successful completion of the courses.

Course materials and recommended reading

Basic literature:

1 O. I. Chernyak Intellectual data analysis: a textbook / O. I. Chernyak, P. V. Zakharchenko. - K.: Znannia, 2014. - 599 p.
https://moodle.znu.edu.ua/pluginfile.php/593075/mod_folder/intro/%D0%91%D0%B0%D0%B7%D0%BE%D0%B2%D0%B8%D0%B9%20%D0%BF%D1%96%D0%B4%D1%80%D1%83%D1%87%D0%BD%D0%B8%D0%BA_3%20%28%D0%A7%D0%B5%D1%80%D0%BD%D1%8F%D0%BA%20%D0%9E.%20%D0%86.%20%D0%86%D0%BD%D1%82%D0%B5%D0%BB%D0%B5%D0%BA%D1%82%D1

%83%D0%B0%D0%BB%D1%8C%D0%BD%D0%B8%D0%B9%20%D0%B0%D0%BD%D0%B0%D0%BB%D1%96%D0%B7%20%D0%B4%D0%B0%D0%BD%D0%B8%D1%85%29.pdf
2. Data Mining: searching for knowledge in data / A. Ya. Gladun, Yu. V. Rogushina – K.: LLC "VD "ADEF-Ukraine", 2016. – 452 p.
https://www.researchgate.net/publication/304025285_Data_Mining_Search_for_Knowledge_in_Data
3. Horokhovatsky V.O., Tvoroshenko I.S. Methods of intellectual analysis and data processing: teaching. manual. – Kharkiv: Khnure, 2021. – 92 p.
https://openarchive.nure.ua/server/api/core/bitstreams/2e55d639-52fd-48d9-b7b7-14989f49f291/content
4. Intellectual data analysis: a study guide / A. O. Oliynyk, S. O. Subbotin, O. O. Oliynyk. – Zaporizhzhia: ZNTU, 2012. – 278 p.
https://ru.scribd.com/document/460020140/%D0%9E%D0%BB%D1%96%D0%B9%D0%BD%D0%B8%D0%BA-%D0%90-%D0 %9E-%D0%86%D0%BD%D1%82%D0%B5%D0%BB%D0%B5%D0%BA%D1%82%D1%83%D0%B0%D0%BB%D1% 8C%D0%BD%D0%B8%D0%B9-%D0%B0%D0%BD%D0%B0%D0%BB%D1%96%D0%B7-%D0%B4%D0%B0%D0% BD%D0%B8%D1%85
5. Applied tasks of intelligent data analysis (DATA MINING). - K.: KNU named after Taras Shevchenko, 2018. – 152 c.
https://www.researchgate.net/profile/Vitalii-Akimenko/publication/325474310_DATA_MINING/links/5b1024bb4585150a0a5deaf6/DATA-MINING.pdf
6. Bakhrushin V. E. Methods of data analysis: study guide for students / V. E. Bakhrushin. – Zaporizhzhia: KPU, 2011. – 268 p.
http://kist.ntu.edu.ua/textPhD/metDataManing.pdf.

Additional literature:

7. Lande D.V., Subach I.Yu., Boyarynova Yu.E. Fundamentals of the theory and practice of intelligent data analysis in the field of cyber security: a study guide. — K.: ISZZI KPI named after Igor Sikorsky", 2018. — 300 p.
http://dwl.kiev.ua/art/oiad/oiad.pdf
8. Lyubun Z. M. Basics of the theory of neural networks / Z. M. Lyubun /: Text of lectures. – Lviv: Ivan Franko LNU Publishing Center, 2007. – 142 p.
https://f.eruditor.link/file/236389/
9. Liubun Z. Hover Signal-Profile Detection / Liubun, V. Mandziy, H. Klein, O. Karpin, V. Rabyk // Proceedings of the XV International Scientific and Technical Conference “Computer Science and Information Technologies” – 2020. P 7 – 10. (Scopus).
10. Bobalo Yu.Ya., Horbaty I.V. (ed.) Information security. Study guide. — Lviv: Publishing House of Lviv Polytechnic, 2019. — 580 p. — ISBN 978-966-941-339-0.
https://pdf.lib.vntu.edu.ua/books/2021/Bobalo_2019_580.pdf.

Assessment and grading

Criteria for assessment of student performance, and the final score structure

Points are awarded according to the following ratio:
- laboratory work: 40% of the semester grade;
- independent work: 10% of the semester grade;
- control work: 10% of the semester grade;
- exam: 40% of the semester grade.

Grading scale

Total points	National	ECTS
90–100	Excellent	A
82–89	Good	B
75–81	Good	C
64–74	Satisfactory	D
60–63	Satisfactory	E
35–59	Unsatisfactory (requires additional learning)	FX
1–34	Unsatisfactory (requires repetition of the course)	F

Norms of academic integrity and course policy

The student must adhere to the Code of Ethics of Academic Relations and Integrity of NTU "KhPI": to demonstrate discipline, good manners, kindness, honesty, and responsibility. Conflict situations should be openly discussed in academic groups with a lecturer, and if it is impossible to resolve the conflict, they should be brought to the attention of the Institute's management.
Regulatory and legal documents related to the implementation of the principles of academic integrity at NTU "KhPI" are available on the website: http://blogs.kpi.kharkov.ua/v2/nv/akademichna-dobrochesnist/

Approval

Approved by

28.08.2024 — Head of the department
Serhii YEVSEIEV

28.08.2024 — Guarantor of the educational program
Serhii YEVSEIEV