МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ

ЗАТВЕРДЖЕНО
на засіданні кафедри
інформаційних систем.
Протокол № 1 від 22.08.2023 р.

ПОГОДЖЕНО
Проректор з навчально-методичної роботи

_____ Каріна НЕМАШКАЛО

ВСТУП ДО МАШИННОГО НАВЧАННЯ
робоча програма навчальної дисципліни (РПНД)

Галузь знань	**12 "Інформаційні технології"**
Спеціальність	**121 "Інженерія програмного забезпечення"**
Освітній рівень	**перший (бакалаврський)**
Освітня програма	**"Інженерія програмного забезпечення"**

Статус дисципліни	**вибіркова**
Мова викладання, навчання та оцінювання	**англійська**

Розробник:
к.ф.-м.н., доцент підписано КЕП Віктор ЗАДАЧИН

Завідувач кафедри
інформаційних систем _____ Дмитро БОНДАРЕНКО

Гарант програми _____ Олег ФРОЛОВ

Харків
2024

MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
SIMON KUZNETS KHARKIV NATIONAL UNIVERSITY OF ECONOMICS

APPROVED
at the meeting of the department
information systems
Protocol № 1 of 22.08.2023

AGREED
Vice-rector for educational and methodical work

Karina NEMASHKALO

INTRODUCTION TO MACHINE LEARNING

Program of the course

Field of knowledge	**12 "Information technology"**
Specialty	**121 "Software Engineering"**
Study cycle	**first (bachelor)**
Study programme	**"Software Engineering"**
Course status	**elective**
Language	**English**

Developers:
Ph.D. (Mathematical sciences),
associate professor — digital signature — Viktor ZADACHYN

Head of Information Systems Department:
Ph.D. (Technical sciences),
associate professor — Dmytro BONDARENKO

Head of Study Programme:
Ph.D. (Technical sciences),
associate professor — Oleg FROLOV

Kharkiv
2024

INTRODUCTION

The modern development of information technologies is characterized by an increasing level of use of intelligent information systems. The course "Introduction to Machine Learning" is a key stage in students mastering the basics of the modern field of information technology, which studies the theories and methods of machine learning.

The purpose of teaching this course is to provide students with higher education with special knowledge about basic concepts, methods, and algorithms of Machine Learning, as well as the development of their ability to apply this knowledge in practice, for example, for designing intelligent information systems.

Tasks in academic the course are: studying basic machine learning algorithms, acquiring data analysis skills, and working with machine learning tools.

The subject of study of the course is theoretical and applied aspects of machine learning, including classical and modern methods, optimization techniques, analysis of results, and ethical aspects of using machine learning algorithms.

The object of study of the course is the processes and methods associated with the use of computer systems for data-based learning. The principles of data selection and preparation, model development, training, and performance evaluation are considered.

In the process of learning students acquire the necessary knowledge during classroom classes: lecture and laboratory. Self-studies of students is also of great importance in the process of studying and consolidating knowledge.

The learning outcomes and competence formed by the course are defined in the table. 1.

Table 1

Learning outcomes and competencies formed by the course

Learning outcomes	Competencies
LO 01	GK 01, GK 02, GK 05, GK 06, SK 08
LO 05	GK 01, GK 02, SK 07, SK 08, SK 14
LO 18	GK01, GK02, GK05, SK07, SK14

where LO 01. Analyze, purposefully search for, and select information and reference resources and knowledge necessary for solving professional tasks, taking into account modern achievements of science and technology.

LO 05. Know and apply relevant mathematical concepts, methods of domain, system and object-oriented analysis, and mathematical modeling for software development.

LO 18. Know and be able to apply information technologies for data processing, storage and transmission.

GK 01. Ability to abstract thinking, analysis, and synthesis.

GK 02. Ability to apply knowledge in practical situations.

GK 05. Ability to learn and master modern knowledge.

GK 06. Ability to search, process, and analysis of information from various sources.

SK 07. Knowledge of information data models, and ability to create software for data storage, extraction, and processing.

SK 08. Ability to apply fundamental and interdisciplinary knowledge to successfully solve software engineering tasks.

SK 14. Ability to algorithmic and logical thinking.

COURSE CONTENT

Content module 1. Basic concepts and definitions of Machine Learning.

Topic 1. Basics of Machine Learning. Machine learning and Artificial Intelligence.

1.1. Data Mining, Big Data, and mathematical statistics

1.2. The concept of artificial intelligence, Machine Learning, Deep Learning, and Data science.

Topic 2.Classes of problems that are solved by methods of Machine Learning.

2.1. Regression.

2.2. Classification and clustering.

2.3. Ptime series forecasting.

2.4. Anomaly search.

2.5. Rimage recognition.

Topic 3. Basic types of Machine Learning.

3.1. Learning with a teacher.

3.2. Learning without a teacher.

3.3. Semi-supervised learning.

3.4. Reinforcement learning.

Topic 4. The process of working with data. Data preprocessing and visualization.

4.1. Statistical data processing.

4.2. INdata visualization.

4.3. Missing and anomalous data.

4.4. Data Scaling.

Topic 5. Classical methods of solving problems of regression analysis and forecasting.

5.1. Methods of estimation of regression model parameters.

5.2. Time series forecasting methods.

5.3. Assessment of model quality.

Topic 6. Classical methods of solving problems classification and clustering.

6.1. Metrics.

6.2. WITHbad luck classification. decision tree, naive Bayes, method of k nearest neighbors.

6.3. WITHbad luck clustering. k-means methods, fuzzy c-means.

6.4. Assessment of model quality.

Content module 2. Neural networks

Topic 7. Apparatus of artificial neural networks

7.1. Artificial neuron. Types of transfer (activation) functions. The architecture of a neural network.

7.2. Learning artificial neural network.

7.3. Rosenblatt's perceptron. MLP network.

7.4. Classesneural networks.Direct distribution networks. Recurrent neural networks. Convolutional neural networks.

Topic 8. Solving regression problems using neural networks.

8.1. Stages of model construction. Architecture search, training, and model quality assessment. Application of the model.

8.2. Software tools for building a regression model.

Topic 9. Solving problems classification using neural networks.

9.1. Stages of model construction. Architecture search, training, and model quality assessment. Application of the model.

9.2. Software tools for building a model classification.

Topic 10. Solving problems clustering using neural networks.

10.1. Kohonen Neural Network (SOM).

10.2. Stages of model construction. Architecture search, training, and model quality assessment. Application of the model.

10.3. Software tools for building a model clustering.

Topic 11. Solving problems forecasting time series using neural networks.

11.1. Stages of model construction. Architecture search, training, and model quality assessment. Application of the model.

11.2. Software tools for building a model forecasting time series.

Topic 12. Solving problems pattern recognition using neural networks.

12.1. Hopfield and BAM networks.

12.2. Stages of model construction. Architecture search, training, and model quality assessment. Application of the model.

12.3. Software tools for building model recognition.

The list of laboratory studies in the course is given in table 2.

Table 2

The list of laboratory studies

Topic name	Content
Topic 2-5, 7, 8.	Solving regression problems using Machine Learning methods.
Topic 2-4, 6, 7, 9.	Solving problems classification by Machine Learning methods.
Topic 2-4, 6, 7, 10.	Solving problems clustering by Machine Learning methods.
Topic 2, 5, 7, 11.	Solving problems with time series forecasting by Machine Learning methods.
Topic 2, 7, 12.	Solving problems pattern recognition using neural networks.

The list of self-studies in the course is given in table 3.

Table 3

List of self-studies

Topic name	Content
Topic 1 - 12	Studying lecture material
Topic 2 - 12	Preparation for laboratory classes
Topic 1 - 12	Preparation for the exam

The number of hours of lecture and laboratory studies and hours of self-study is given in the technological card of the course.

TEACHING METHODS

In the process of teaching the course, in order to acquire certain learning outcomes, to activate the educational process, it is envisaged to use such teaching methods as:

Verbal (lecture (Topic 1, 2, 3, 4, 5, 6, 7, 8, 9), problem lecture (Topic 10, 12), lecture-visualization (Topic 11)).

In-person (demonstration (Topic 1-12)).

Laboratory work (Topic 4, 5, 6, 7, 8, 9, 10, 11, 12), case method (Topic 5).

FORMS AND METHODS OF ASSESSMENT

The University uses a 100-point cumulative system for assessing the learning outcomes of students.

Current control is carried out during lectures and laboratory classes and is aimed at checking the level of readiness of the student to perform specific work and is evaluated by the sum of points scored:

– for courses with a form of semester control an exam: the maximum amount is 60 points; the minimum amount that allows a student of higher education to pass an exam is 35 points.

Final control includes semester control and certification of the student of higher education.

Semester control is carried out in the form of a semester exam .

The maximum number of points that a student of higher education can receive during the exam is 40 points. The minimum amount for which the exam is considered passed is 25 points.

The final grade in the course is determined by summing the points received during the current control and the exam grade.

During the teaching of the course, the following control measures are used:

Current control: defense of laboratory works (50 points), written control works (10 points).

Semester control: Exam (40 points).

More detailed information on the assessment system is provided in the technological card of the course.

An example of an exam card and assessment criteria.

Semyon Kuznets Kharkiv National University of Economics
First (bachelor) level of higher education
Specialty "Software Engineering"
Study programme "Software Engineering"
Semester VII
The course "Introduction to Machine Learning"

EXAM CARD

Task (heuristic, 40 points).
Carry out clustering of multidimensional objects for "Adult" data from the "UCI Machine Learning repository" Repository (http://archive.ics.uci.edu/ml/index.php), where the data sample corresponds to a folder with the appropriate name, which contains a file with a description of the subject area and a data file directly.
When recording the results of the task, it is necessary to comply with the following requirements: give a formulation of the general statement of the task to which the task refers; explain the meaning of the main attributes; conduct data preparation; build a model; evaluate the quality of the model based on test data; apply the model to solve the problem; give an interpretation of the obtained results; to build graphs, which are necessary for the interpretation of both the input data and the obtained results.
Protocol No. ____ dated "___"_____20___ was approved at the meeting of the Department of Information Systems.

Examiner PhD, Associate Professor Zadachynn V.

Chief Department PhD, Associate Professor Bondarenko D.

Assessment criteria

The examination ticket consists of one heuristic task. The final grade for the exam is a maximum of 40 points.

As a result of the task, there should be two files: the first is in DOC format (explanations and screenshots of the results of the program), and the second is a program script (Python or R) that implements the task. These files should be attached as an answer to the task on the PNS course.

When recording the results of the task (in the first file), the following requirements must be observed: give the formulation of the general statement of the task to which the task refers; explain the meaning of the main attributes; conduct data preparation; build a model; evaluate the quality of the model based on test data; apply the model to solve the problem; give an interpretation of the obtained results; to build graphs, which are necessary for the interpretation of both the input data and the obtained results.

The task is evaluated according to the scale presented in the table. 4.

Table 4

Assessment criteria of the heuristic task

40 points	The task is completed in full. The script of the program is presented, the results obtained, the analysis of the obtained results with reference to the essence of the problem, and the answers to the questions posed in the condition of the problem are given. Conclusions made. All requirements for recording the results of the task have

	been met.
35 points	The task is completed in full. The script of the program is presented, the results obtained, the analysis of the obtained results with reference to the essence of the problem, and the answers to the questions posed in the condition of the problem are given. Conclusions made. Not all requirements for recording performance results have been met.
30 points	The task is completed in full. The script of the program is presented, the results obtained, the analysis of the obtained results with reference to the essence of the problem is carried out, and the answers are not given to all the questions posed in the condition of the problem. Conclusions made. Not all requirements for recording performance results have been met.
25 points	Task accomplished. There is a program script, and the program works, but there are minor inconsistencies with the condition of the task. The results were obtained, but the analysis of the obtained results was not carried out, there are no answers to the questions posed in the condition of the problem. Not all requirements for recording performance results have been met.
20 points	There is a program script, but it does not work. Not all requirements for recording the results of the task are met, but the modeling logic is given.
10 points	Task not completed. There is a fragment of the program script and the simulation logic is given. The requirements for recording the results of the task have not been met.
0 points	Task not completed.

The task can be evaluated in intermediate points (for example, 24-29, 31-34, 36-39) depending on the availability and degree of detail of the explanations for the task performance process.

RECOMMENDED LITERATURE

Main

1. Yu.V. Nikolskyi Artificial intelligence systems: Study guide / Yu.V. Nikolskyi, V.V. Pasichnyk, Yu. M. Shcherbina. – Kyiv: Magnolia, 2021. – 280 p.

2. Trotsko V.V. Methods of artificial intelligence: educational, methodological and practical manual / V.V. Trotsky – Kyiv: KROK University, 2020. – 86 p.

3. Data pre-processing and analysis: laboratory practice for students. specialty 113 "Applied mathematics" /Compiled by: N. E. Kondruk. Uzhhorod: UzhNU, 2023. - 41 p.

Additional

4. Stephan S. Jones, Frank M. Groom Artificial Intelligence and Machine Learning for Business for Non-Engineers. - CRC Press, 2019. - 148 p.

5. Aurélien Géron. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. 2nd Edition – O'Reilly Media, 2019. –856 p.

6. Burkov A.The Hundred-Page Machine Learning Book–Andriy Burkov, 2019. –141 p.

7. Zadachyn V. M. System modeling and optimization methods: methodological recommendations for laboratory work for students of the field of

knowledge 12 "Information technologies" of the first (bachelor) level: [Electronic edition]- Kharkiv: Ed. HNEU named after S. Kuznetsa, 2019. – 210 p. http://repository.hneu.edu.ua/handle/123456789/22458

Information resources

8. Introduction to Mathematical Optimization: with Python– https://indrag49.github.io/Numerical-Optimization/

9. Collaborative –https://colab.research.google.com/

10. Numerical Programming with Python– https://python-course.eu/numerical-programming/