Syllabus

VK Machine learning in big data forecasting tasks

Subjects:	**Machine learning in big data forecasting tasks**
Level of higher education:	**first (undergraduate)**
Course page in Moodle:	https://dl2022.khadi-kh.com/course/index.php?categoryid=39
The scope of the educational component	**3 credits (90 hours)**
Final control form	**Test**
Consultations:	**on schedule**
Name of the department:	**department of computer technologies and mechatronics**
Teaching language:	**English**
Course leader:	**Shulyakov Vladyslav Mykolayovych, Ph.D., associate professor**
Contact phone number:	+38 (067) 737-68-69
Email:	*jason07@ukr.net*

Brief content of the educational component:

The goal isto form fundamental theoretical knowledge and practical skills about modern methods of building and analyzing various models of machine learning and data analysis, as well as their practical application for solving applied problems.

Subject: formation of knowledge and skills regarding the basics and basic principles of machine learning in big data forecasting tasks.

The main tasks of studying the academic disciplineis to teach how to work with data, build machine learning models; visualize various data structures; build intelligent models for the analysis of various processes.

Prerequisites for studying the educational component:

- Design of information intelligent systems;
- Intelligent data analysis.

Competencies acquired by the acquirer:

General competences:

- Knowledge of information data models, ability to create software for data storage, extraction and processing;
- Ability to apply fundamental and interdisciplinary knowledge to successfully solve software engineering tasks;
- Ability to algorithmic and logical thinking.

Special competences:

- Ability to develop and implement data warehouses, use methods of intelligent analysis of large data sets to support decision-making in the organization.
- The ability to accumulate, process and systematize professional knowledge regarding the creation and maintenance of software in the field of machine learning;
- Ability to search, process and analyze information from various sources.

Learning outcomes according to the educational program:

- Know and apply relevant mathematical concepts, methods of domain, system

and object-oriented analyzes and mathematical modeling for software development;

– Know and apply in practice the fundamental concepts, paradigms and basic principles of the functioning of linguistic, instrumental and computing tools of software engineering;

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

Thematic plan

Topic No	Name of topics (LK, LR, PR, SZ, SR)	Number of hours	
		ocular	extramural
1	**Lecture #1:**. Primary data analysis.	**2**	
	Practical lesson 1	**2**	
	Tasks for independent work 1	**7**	
2	**Lecture #2:**Visual data analysis.	**2**	
	Practical lesson 2	**2**	
	Assignment for independent work 2	**7**	
3	**Lecture #3:**Classification, decision trees and the method of nearest neighbors.	**2**	
	Practical lesson 3	**2**	
	Assignment for independent work 3	**7**	
4	**Lecture #4:**Linear models of classification and regression.	**2**	
	Practical lesson 4	**2**	
	Assignment for independent work 4	**7**	
5	**Lecture #5:**Construction and selection of features. Applications in image and geodata text processing tasks.	**2**	
	Practical lesson 5	**2**	
	Assignment for independent work 5	**7**	
6	**Lecture #6:**Learning without a teacher.	**2**	
	Practical lesson 6	**2**	
	Assignment for independent work 6	**7**	
7	**Lecture #7:**Time series analysis Part 1	**2**	
	Practical lesson 7	**2**	
	Assignment for independent work 7	**7**	
8	**Lecture #8:**Time series analysis Part 2.	**2**	
	Practical lesson 8	**2**	
	Assignment for independent work 8	**9**	
Together	**Lectures**	**16**	
	Practices	**16**	
	Independent work	**58**	

Individual educational and research task: not provided.

Teaching methods:

1) verbal: 1.1 traditional: lectures, explanations, stories, etc.;

1.2 interactive (non-traditional): problem lectures, discussions, etc.;
2) visual: the method of illustrations, the method of demonstrations
3) practical: 3.1 traditional: practical classes, seminars;
3.2 interactive (non-traditional): trainings, "round table", brainstorming method.

Evaluation system and requirements:

The concretization and detailing of the criteria and evaluation system, taking into account the specifics of the educational component, is carried out on the basis of the general criteria specified in STVNZ 7.1-02:2018 "Regulations on the Organization of the Educational Process at the Khnadu" (https://www.khadi.kharkov.ua/fileadmin/P_Ychebotdel/norm_dok/stvnz_7_1_02.pdf) and STVNZ 90.1-01:2021 "Evaluation of learning outcomes of higher education applicants" (https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz-90.1-01_2021.pdf).

Current performance

1 The current success of applicants for the performance of educational types of work in training sessions and for the performance of independent work tasks is evaluated using a four-point rating scale with subsequent transfer to a 100-point scale. During the evaluation of the current performance inare countedall types of work provided by the curriculumprogram

1.1 Lectures occupationare evaluated by definition quality implementation of specified tasks

1.2 Practical classes are evaluated by the quality of performance of a control or individual task, performance and design of practical work.

1.3 Laboratory occupation are evaluated quality implementation reports about implementationlaboratory work.

1.4 Seminary occupation are evaluated quality implementation individualassignment/abstract.

2 Evaluation of the current success rate of higher education applicants is carried out at each practical session(laboratoryor seminary) on a four-point scale("5", "4", "Z", "2")and are enteredinaccounting journalacademicsuccess

– "excellent": acquirerflawlessly mastered the theoretical material, demonstrates deepknowledge ofrelevant topic or academic discipline, main provisions;

– "good": the applicant has mastered the theoretical material well, possesses the main aspects from primary sources and recommended literature, presents it in a reasoned way; has practical skills, expresses his thoughts on certain problems, but certain inaccuracies and errors are assumed in the logic of the presentation of theoretical content or in the analysis of practical ones;

– "satisfactory": the applicant has basically mastered the theoretical knowledge of the educational topic or discipline, orients himself in primary sources and recommended literature, but answers unconvincingly, confuses concepts, answers additional questions uncertainly, does not have stable knowledge; when answering questions of a practical nature, reveals inaccuracy in knowledge, does not know how to evaluate facts and phenomena, connect them with the future profession;

– "unsatisfactory": the applicant has not mastered the educational material of the topic (discipline), does not know scientific facts, definitions, hardly orients himself in primary sources and recommended literature, lacks scientific thinking, practical skills are not formed.

3 Final scorebycurrent activity is recognized as an arithmetic averagesumpoints for each lesson, for individual work, current control works according to the formula:

$$K^{current} = \frac{K1+K\,2+...+Kn}{n},$$

where $K^{current}$ – final assessment of success based on the results of current control;

$K1,K\,2,...,Kn$ - evaluation of success n-th measure of current control;

n - the number of measures of current control.

Grades are convertedinpoints according to the calculation scale (table 1).

Table 1– Recalculation of the average grade for the current activity into a multi-point scale

4-point scale	100-point scale	4-point scale	100-point scale	4-point scale	100-point scale	4-point scale	100-point scale
5	100	4.45	89	3.90	78	3.35	67
4.95	99	4.4	88	3.85	77	3.3	66
4.9	98	4.35	87	3.80	76	3.25	65
4.85	97	4.3	86	3.75	75	3.2	64
4.8	96	4.25	85	3.7	74	3.15	63
4.75	95	4.20	84	3.65	73	3.1	62
4.7	94	4.15	83	3.60	72	3.05	61
4.65	93	4.10	82	3.55	71	3	60
4.6	92	4.05	81	3.5	70	from 1.78 to 2.99	from 35 to 59
						reassembly	
4.55	91	4.00	80	3.45	69	from 0 to 1.77	from 0 to 34
4.5	90	3.95	79	3,4	68	repeated study	

Final assessment

1 The exam is held after studyingeveryonetopics of the discipline and is composed of students of higher education during the examination session after the end of all classroom classes

2 Applicants of higher education who have completed all types of work prescribed by the curriculum in the discipline are admitted to the exam:

- were present ateveryoneclassroom classes (lectures, seminars,practical);
- completed all missed classes on time;
- scored the minimum number of points for the current academic performance (at least 60 points, whichrespondson a national scale"3");

If the current performance in the discipline is lower than 60 points, the higher education applicant has the opportunity to increase his current point to the minimum before the beginning of the examination session.

3 Assessment of the knowledge of applicants when taking the exam is carried out on a 100-point scale.

Assessment of the knowledge of applicants through testing is carried out according to the following scale:

– "Excellent": at least 90% of correct answers;
– "Very good": 82% to 89% correct answers;
– "Good": from 74% to 81% of correct answers;
– "Satisfactory": from 67% to 73% of correct answers;
– "Fair enough": 60% to 66% correct answers;
– "Unsatisfactory": less than 60% of correct answers.

4 The final grade for the academic discipline is defined as a weighted average grade

that takes into account the overall grade for the current academic performance and the grade for passing the exam.

5 The calculation of the overall final grade for the study of an academic discipline is carried out according to the formula:

$$PKekz=0.6\cdot Kpotoch+0.4\cdot IS,$$

where PC^{ex} - final assessment of success in disciplines, in the form of a final report controls for which there is an exam;

$K^{current}$ – final assessment of success based on the results of current control (on a 100-point scale);

IS - assessment based on the results of the exam (on a 100-point scale).

0.6 and 0.4 – coefficients of the ratio of points for current success and taking the exam.

6 For performing individual independent work and participating in scientific events, winners are awarded additional points.

6.1 Additional points are added to the total points scoredcakeof higher education for the current educational activity (for disciplines for which the final form of control is a credit), or to the final assessmentwithdisciplines, the final form of control for which is an exam.

6.2 The number of additional points awarded for different types of individual tasks depends on their volume and importance:

– prizes in the discipline at the international / all-Ukrainian competition of scientific student works - 20 points;

– prize places in the discipline at the All-Ukrainian Olympiads - 20 points;

– participation in the international / all-Ukrainian competition of scientific student works

– 15 points

– participation in international / all-Ukrainian scientific conferences of students and young scientists - 12 points;

– participation in all-Ukrainian Olympiads in the discipline - 10 points

– participation in olympiads and scientific conferences of the Khnadu in the discipline - 5 points;

– performance of individual scientific and research (educational and research) tasks of increased complexity - 5 points.

6.3 The number of additional points cannot exceed 20 points.

7 General final rating by study educational disciplines not maybeexceed 100 points.

The overall final grade for the study of the academic discipline is determined according to the scale given in Table 2.

Table 2– The scale for evaluating the knowledge of the students based on the results of the final control of the academic discipline

Score in points	Assessment by national scale		Evaluation according to the ECTS scale	
	examination	test	Rating	Criteria
90-100	**Perfectly**	**Enrolled**	**A**	The theoretical content of the course has been mastered in its entirety, without gaps, the necessary practical skills for working with the mastered material have been formed, all the educational tasks provided for in the training program have been completed, the quality of their performance has been assessed by the number of points close to to the maximum
80–89	**Okay**	**Enrolled**	**B**	The theoretical content of the course has been mastered in its entirety, without gaps, the necessary practical skills for working with the mastered material have mainly been formed, all educational tasks provided for in the training program have been completed, the quality of most of them has been evaluated with a number of points close to the maximum
75-79			**WITH**	The theoretical content of the course has been mastered completely, without gaps, some practical skills of working with the mastered material have not been sufficiently developed, all the educational tasks provided for by the training program have been completed, the quality of none of them has been evaluated with a minimum number of points, some types tasks were completed with errors
67-74	**Satisfactorily**	**Enrolled**	**D**	The theoretical content of the course is partially mastered, but the gaps are not of a significant nature, the necessary practical skills for working with the mastered material are basically formed, most of the educational tasks provided by the training program have been completed, some of the completed tasks, may contain errors
60–66			**E**	The theoretical content of the course has been partially mastered, some practical work skills have not been formed, many of the educational tasks provided by the training program have not been completed, or the quality of some of them has been assessed with a number of points close to minimal

35–59	**Unsatisfactorily**	**Not counted**	**FX**	The theoretical content of the course has been partially mastered, the necessary practical work skills have not been formed, most of the prescribed training programs of educational tasks have not been completed, or the quality of their implementation has been assessed with a number of points close to the minimum; with additional independent work on the course material, it is possible to improve the quality of the performance of educational tasks (with the possibility of repeated drafting)
0–34	**Unacceptable**		**F**	The theoretical content of the course has not been mastered, the necessary practical work skills have not been formed, all completed educational tasks contain gross errors, additional independent work on the course material will not lead to any significant improvement in the quality of the performance of educational tasks (with a mandatory repeat course)

Course policy:

– the course involves working in a team, the environment in the classroom is friendly, creative, open to constructive criticism;

– mastering the discipline involves mandatory attendance at lectures and practical classes, as well as independent work;

– independent work involves the study of individual topics of the academic discipline, which are presented in accordance with the program for independent study, or were considered briefly;

– all tasks provided by the program must be completed within the set time;

– if the student of higher education is absent from classes for a valid reason, he presents the completed tasks during independent preparation and consultation of the teacher;

– course work must be protected no later than a week before the beginning of the examination session;

– while studying the course, students of higher education must adhere to the rules of academic integrity set forth in the following documents: "Rules of Academic Integrity participants educational process LOOKING FOR"(https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz_67_01_dobroch_1.p _df), "Academic Integrity. Verification of academic, scientific and qualification texts works on plagiarism" (https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz_85_1_01.pdf), "Moral and ethical code of participants in the educational process of the National Academy of Sciences (https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz_67_01_MEK_1.pdf).

– in case of detection of plagiarism, the applicant receives 0 points for the task and must repeat the tasks provided for in the syllabus;

– writing off during tests and exams is prohibited (including using mobile devices). Mobile devices are allowed to be used only during online testing.

Recommended Books:

1. J Manyika, M Chui, B Brown, J Bughin, R Dobbs, C Roxburgh, AH Byers, Big data: the next frontier for innovation, competition, and productivity (McKinsey Global Institute, USA, 2011).

2. CW Tsai, CF Lai, MC Chiang, LT Yang, Data mining for internet of things: a survey. IEEE Commun Surv Tut 16(1), 77–97 (2014).

3. XW Chen, X Lin, Big data deep learning: challenges and perspectives. IEEE Access 2, 514–525 (2014).

4. C Rudin, KL Wagstaff, Machine learning for science and society. Mach Learn 95(1), 1–9 (2014).

Additional sources:

1. A Sandryhaila, JMF Moura, Big data analysis with signal processing on graphs: representation and processing of massive data sets with irregular structure. IEEE Signal Proc Mag 31(5), 80–90 (2014).

2. Che, M Safran, Z Peng, From big data to big data mining: challenges, issues, and opportunities, in Proceedings of the 18th International Conference on DASFAA (Wuhan, 2013), pp. 1–15.

Developer(s)
the syllabus of the academic discipline
Ph.D. technical of Science, Assoc. _____ Vladyslav Shulyakov
signature

Head of the KTM department _____ Oleg NIKONOV
dr. technical Sciences, Prof. signature