The Ministry of Education and Science of Ukraine
State Higher Education Institution
«National Mining University»

The Department of System Analysis and Management

Zheldak T.A.

WORKING PROGRAM OF EDUCATIONAL DISCIPLINE
«Intellectual Data Analysis»
For masters of specialty 124 «System Analysis»

Dnipro
NMU
2017

Zheldak T.A. Program of educational discipline «Intellectual data analysis» for masters of specialty 124 «System analysis» / T.A. Zheldak; National Mining University, Dep. System Analysis and Management. – D. : NMU, 2017. – 12 p.

Viewed and approved by the methodical commission of the specialty 124 "System Analysis" on the submission of the Department of System Analysis and Management

CONTENTS

INTRODUCTION

The program training results of the master's degree in system analysis are defined in the Standard of Higher Education by specialty 124 "System Analysis".

In the educational and professional program of masters training in this specialty of the State University "NMU" [2.1] the distribution of program education results was carried out according to the organizational forms of the educational process. The discipline "Intellectual data analysis" includes the following **professional competencies**:

PC6 Ability to design the architecture of intelligent information systems.

PC7 Ability to apply intelligent data analysis during the construction of DSS, expert and advisory systems.

PC8 Ability to develop the functions to forecast the dynamics of the processes development of different nature in a deterministic and stochastic environment and to evaluate the quality of the forecast.

PC10 Ability to apply modern information technology for solving problems of system analysis.

During the study of the discipline, the Master must learn the following general and professional **education results**:

Під час вивчення дисципліни магістр має опанувати такі загальні та професійні **результати навчання**:

GER5 To show curiosity, predisposition to risk, the ability to think, to get inspired by new ideas, implement them, to inspire others by them, to combine and to experiment.

GER6 To select and prepare information and tasks for the project team, set goals and formulate tasks for the implementation of projects and programs.

PER3 To know methods of forecasting the dynamics of different processes, be able to develop prediction functions.

PER7 To be able to develop expert and advisory systems in conditions of poorly structured data of various types.

ERS1 To use technologies and tools of search engines, methods of intellectual data and texts analysis, processing, interpreting and summarizing data.

ERS2 To be able to build typical mathematical models of system analysis objects and processes, use mathematical methods and algorithms of data processing (statistical, algebraic, combinatorial, theoretical-informational, etc.).

The purpose of the discipline is "Intellectual Data Analysis" - to form masters' skills for researching and identifying algorithms, with the help of artificial intelligence in a hidden structures "raw data", patterns or relationships that were previously unknown, non-trivial, practically useful and accessible for human interpretation and necessary for decision making in various spheres of activity.

Realization of the goal requires the transformation of the program education results in discipline, and the selection of the content of the discipline according to this criterion.

Requirements for work program structure of disciplines are given in [2.5].

1. THE SCOPE

The work program is created for

- implementation of a competent approach for shaping the structure and content of the discipline;
- internal and external quality control of specialists' skills;
- accreditation of educational program in specialty.

The work program sets:

- scope and terms of discipline teaching;
- designation of physical quantities;
- disciplinary education results and their level of difficulty;
- thematic plan and amount distribution according to the organizational forms of the educational process;
- requirements to the structure and the content of the individual task;
- tasks for independent work;
- generalized diagnostic tools, criterias and procedures for assessing applicants' achievements;
- the Complex composition of teaching and methodological provision of discipline.

2. NORMATIVE REFERENCES

The work program of the discipline is developed on the basis of the following normative documents:

2.1 Educational program of masters' training by specialty 124 "System analysis" / Ministry of Education and Science of Ukraine, NTU "Dniprovska Polytechnic". - D.: NTU "DP", 2018 - 31 p.

2.2 Licensing terms of the proceedings for the educational activities of educational institutions. Approved by the Cabinet of Ministers of Ukraine Decree, dated December 30, 2015, No. 1187. http://zakon5.rada.gov.ua/laws/show/1187-2015-p/page.

2.3 The Draft higher education standard for masters of specialty 124 "System analysis".

2.4 Law of Ukraine "On Higher Education" http://zakon5.rada.gov.ua/laws/show/1556-18

2.5 Standard of Higher Education of the State Higher Educational Institution "NMU" Design of the educational process. Dnipropetrovsk: NMU, 2016. - 74 p.

3. AMOUNT AND TERMS OF STUDYING THE DISCIPLINE

Total amount - 6 ECTS credits (180 academic hours).

It is taught at the 1-st year, in the 1-st semester, in the 1-st (7-th week) and 2-nd (8-th weeks) quarters.

4. DESIGNATIOM OF PHYSICAL QANTITIES

Commonly accepted notations are used.

5. EXPECTED DISCIPLINARY RESULTS OF EDUCATION

The code and the content of education results for an educational-professional program	Code and content of disciplinary education results (DES)
1	2
PER3 To know the methods of forecasting the dynamics of the different processes development, to be able to develop prediction functions	DES3-1 Know the basic methods of information primary processing, increase the data informativeness of their normalization and purification
	DES3-2 Apply checking incoming data algorithms for multicollinearity, heteroscedasticity and perform factor analysis of input data.
	DES3-3 To know the mathematical basis, idea and algorithms for classifying objects in real, categorical and mixed spaces.
	DES3-4 To know the mathematical basis and algorithms of clustering methods based on the idea of compactness, using graphs and neural networks
PER7 To be able to develop expert and advisory systems in conditions of poorly structured data.	DES7-1 To have the skills to form models and knowledge templates
	DES7-2 To use methods of obtaining knowledge, clustering and classification
	DES7-3 To be able to apply knowledge in intellectual systems, decision support systems and everyday practice.
ERS1 To use search engine technology and tools, methods for intellectual data and texts analysis, processing, interpreting and summarizing data.	ERS1-1 To use algorithms for constructing associative and classifying rules
	ERS1-2 To use neural network data and the method of grouping arguments for the approximation
ERS2 To be able to build typical mathematical models of system analysis objects and processes, to use mathematical methods and algorithms for data processing (statistical, algebraic, combinatorial, theoretical-informational, etc.).	ERS2-1 To build neural networks of different architectures with different activation functions
	ERS2-2 To apply to the neural networks with different architecture the best methods of education, testing and validation, compare them with each other.
	ERS2-3 To build evolutionary, genetic, memetic and other optimization algorithms for solving problems of modeling, optimization and control in real and binary space.

6. THEMATIC PLAN AND DISTRIBUTION OF DISCIPLINES BY TYPES OF EDUCATIONAL STUDIES

Code DES	Type and theme of educational classes	The amount of hours		
		aud	ISW	total
1	2	3	4	5
	Lectures	**30**	**60**	**90**
DES3-1	Theme 1. Entropy and informativeness of factors. Ways to increase informativeness.	2	4	6
DES3-2	Theme 2. Farrar-Glauber Test for checking input data.	2	4	6
DES3-1	Theme 3. Plural formation of significant factors.	2	4	6
ERS2-1	Theme 4. Approaches for studying neural networks with and without a teacher.	2	4	6
DES3-4 ERS2-3	Theme 5. Methods of teaching direct distribution neural networks.	4	8	12
ERS1-1	Theme 6. Search for sequential templates	2	4	6
DES3-3 DES7-3	Theme 7. Classical, iterative and heuristic classification methods	2	4	6
DES3-4 ERS2-3	Theme 8. Clustering. Hypothesis of compactness. An evolutionary search.	2	4	6
ERS2-2	Theme 9. Bayesian network.	4	8	12
DES3-3	Theme 10. Algorithms for constructing decision trees.	2	4	6
DES7-3	Theme 11. Use of fuzziness in decision trees	2	4	6
ERS1-2	Theme 12. Restoration of the information in the data sets.	4	8	12
	Laboratory work	**30**	**60**	**90**
DES3-1 DES3-2	LW1. Preprocessing information and constructing plural linear regression	4	8	12
ERS1-2 ERS2-2	LW2. Set determination of significant factors and construction of plural nonlinear regression	4	8	12
ERS1-1	LW3. Search for associative rules and sequential templates	4	12	16
DES7-2	LW4. Classification of objects in a mixed space	6	12	18
DES7-2 ERS2-2	LW5. Clustering on an objects plural of real space.	4	8	12
DES7-3 DES7-2	LW6. Obtaining knowledge from data, comparing algorithms.	6	12	18
	Control of the laboratory module	2	-	2
	Total	**60**	**120**	**180**
	Lections (auditorium - 4 hours per week)	**30**	**60**	**90**
	Laboratory classes (auditorium - 4 hours per week)	**30**	**60**	**90**
	Final (semester) control - exam: I semester, 2 quarters			

7. REQUIREMENTS FOR INDIVIDUAL TASK

During the study of discipline, the implementation of an individual task is not provided.

8. TASKS FOR STUDENT INDEPENDENT WORK

The main tasks for independent work include:

- preliminary processing of information provision for each topic;
- preparation for ongoing control - solving tasks of self-control on each topic;
- preparation for the final (semester) control.

9. FORM OF FINAL CONTROL, DIAGNOSTICS, CRITERIA AND PROCEDURE OF AN EVALUATION

9.1 Form of final control

The final examination form is a written examination.

Assessment of the disciplinary competencies formation level in the form of an examination can not be carried out without student participation based on the results of current control. The latter are taken into account in the final examination at the level of the results of the written examination.

9.2 Forms of a current control

Formation level determination of the disciplinary education results during the current control is carried out for:

- testing achievements in a specific section of the discipline work program;
- laboratory work (checking and defense);

9.3 Methods of diagnosis

9.3.1 Generalized diagnostic methods

Diagnostic methods are presented in the form of theoretical questions and concretized tasks with numerical input data and are designed to assess the student's ability:

- to differentiate, integrate and unify knowledge;
- to apply rules, methods, principles, laws in a specific situation;
- to interpret circuits, graphs, diagrams;
- to analyze and evaluate the facts, events and predict the expected results from taken decisions;
- to present the material on paper logically, consistently, in compliance with the requirements of the current standards.

9.3.2 Specified diagnostic methods

Specific diagnostic methods used directly for control activities during lectures, are formed on the basis of generalized methods, by numerical or other specification of generalized methods in the form of closed and open tests.

9.4 Criteria and evaluation procedures

9.4.1 Lecture material

The results evaluation of the completed tasks is carried out by comparing them with the standards - samples of correct and completed answers through the identification of the competence formation level based on the analysis of the student's response, using the coefficient of assimilation in percent, adapting the value of the mark to the ECTS scale:

$$P_i = a / m \ (\%),$$

where a is a number of correct answers or done essential decision-making operations; m is the total number of questions or essential operations.

The results of the students' achievements (as a percentage) obtained from the described scheme are presented in the grades of the ECTS and the national scale:

Marks, %	Grade
National Differentiated Scale	
90-100	Excellent
74-89	Good
60–73	Satisfactory
1-59	Fail
Scale/ECTS	
90-100	A
82-89	B
74-81	C
64–73	D
60–63	E
35-59	Fx
1-34	F

If the level of the student’s achievements lower 60% or if the student did not appear on a control event, then he is rated with "Fx" and "Failed". In such cases, the student is obliged to learn this topic additionally and undergo a re-evaluation of his education results.

9.4.2 Laboratory work

Each laboratory work is estimated by the quality of the report with help of the assimilation coefficient or by the expert method, when the maximum grade is made by including following rules:

– the report was completed in accordance with the guidelines;
– correctness of execution;
– possession of theoretical information, the base of the subject;
– possession of experimental research methods;
– general and professional literacy, conciseness and logical sequence of material presentation;
– compliance of the report with the current standards;
– availability of references to information sources;
– independence of execution.

The level of achievements based on the results of a laboratory works set is defined as the average value of the results of the current monitoring of each.

During the examination, the assessment of laboratory work is determined by the percentage of correct algorithm steps for its implementation.

Integral assessment of achievements for all laboratory works is positive (the level of student achievement is higher than 60% or more than 60 points) only in the case when all the laboratory works, that provided by the work program of the discipline, are done and evaluated.

9.4. Integral level of student achievements in discipline

The integral level of the student's achievements in lerning the discipline material as a whole is calculated as an average value of the competence formation level on lecture, practical and laboratory classes:

$$IP = \sum_{i=1}^{n} \frac{(P_i \times T_i)}{T}, \%,$$

where n is a number of classes types;

P_i is a level of achievements for the i-th type of classes, %;

T_i is an amount of the i-th type of classes;

T is the total discipline amount.

The student's achievements in mastering a certain discipline in general can not be assessed positively, if the student has not received a positive evaluation for any planned control event in this discipline.

If the level of any current control event is higher than 60%, then the grade is "scaled" according to the national scale.

If the level of any current control event is lower than 60%, then the discipline is graded with "Fx" and, if below 35%, then "F". In this case, student gets the "Fail" mark according to the national scale.

10. THE COMPLEX COMPOSITION OF EDUCATIONAL AND METHODOLOGICAL SUPPORT OF DISCIPLINES

The complex of educational and methodological support of the discipline must be downloaded on the website of the system analysis and management department and should contain:

1) work program of the discipline;
2) educational content (information provision of lectures);
3) tasks and methodical provision of laboratory works;
4) methodical support of the student's independent work:
– preliminary study of information support that is included in lectures;
– solving tasks for self-control on each topic
6) generalized tasks for the current control of the disciplinary competencies formation level in the form of typical situational exercises with examples of solutions;
7) tasks for post-certification monitoring of disciplinary formation competencies level.

11. RECOMMENDED LITERATURE

11.1 Basic

1. Paklin N.B. Business Analytics: From Data to Knowledge: A Textbook. The allowance / Н.Б. Paklin, V.I. Oreshkov. - 2 nd ed., rem. and add. - St. Petersburg: Peter. - 2010. - 704 p.
2. Zaychenko Yu.P. Fundamentals of Designing Intelligent Systems. Tutorial. - K .: Publishing House "Word". - 2004 - 352 pp.
3. Snityuk V.E. Prediction: Models. Methods. Algorithms. Tutorial. - K: "Maclaut". - 2008 - 364 pp.
4. Khaikin S. Neural networks: full course / Simon Khaikin. - 2nd ed.: Trans. from English - M .: Williams Publishing House. - 2006. - 1104 pp.
5. Rutkovskaya D. Neural networks, genetic algorithms and fuzzy systems / D. Rutkovskaya, M. Pylinky, L. Rutkovsky: Trans. from Polsk. I.Д. Rudinsky - M .: Hot line - Telecom. - 2007. - 452 pp.

11.2 Additional

6. Lyuger J.F. Artificial Intelligence: Strategies and Methods for Solving Complex Problems. - 4th edition: Trans. from English - M .: Williams Publishing House, 2003. - 864 pp.
7. Duke V. Data mining: training course / V. Duke, A. Samoilenko. - St. Petersburg: Peter. - 2001. - 368 pp.
8. Barseghyan A.A. Data Analysis Techniques: Data Mining, Visual Mining, Text Mining, OLAP / A.A. Barseghyan, M.S. Kupriyanov, V.V. Stepanenko, II Cold. - 2nd ed., Pererab. and add - SPb .: BHV-Petersburg. - 2007. - 384 pp.

Educational publishing

Zheldak Timur Anatolievich

WORKING PROGRAM OF EDUCATIONAL DISCIPLINE
«Intellectual Data Analysis»
For masters of specialty 124 «System Analysis»

Published
In State Higher Educational Institution
«National Mining University».
Certificate of entry in the State Register of Civil Code № 1842 dated 11.06.2004
49005, Dnipro, Yavornitsky av., 19.