| | | | |
|---|---|---|---|
| **Subject code:**<br>I.7(3) | **Subject name:** Advanced Information Technology | | |
| **Study load:**<br>4 ECTS | **Load of contact hours:** 36 | **Study semester:**<br>Autumn | **Assessment:**<br>Credit / No credit |
| **Objectives:** | The aim of the course - to give an idea of the modern, advanced information technologies. | | |
| **Course outline:** | Topics covered:<br>1. Overview of approaches to solving the problems of classification and regression. Statistical methods. Machine learning.<br>2. Basics of statistical information processing. Linear regression.<br>3. The choice of dimension feature space.<br>4. Logistic regression.<br>5. Methods of time series analysis.<br>6. Fuzzy uncertainty, fuzzy inference.<br>7. Review of ensembles of classification models.<br>8. Methods of decision trees.<br>Contact lessons will be divided into two parts: lectures&practicals and information technology workshops with individual tasks. | | |
| **Learning Outcomes:** | At the end of the course, the student should have a holistic view of advanced information technologies as an instrumental environment for solving regression and classification problems.<br><br>The student should be able to choose effective models, methods and algorithms that are adequate to the given conditions and application environment. | | |
| **Assessment Methods:** | Assessment of student knowledge is carried out according to the results of solving two independent tasks according to the statistics provided by the teacher. The first is the construction of multidimensional dependence, the second is the classification problem. Software implementation tools are selected by the student on their own. | | |
| **Teacher(s):** | Mikhail Matveev | | |
| **Prerequisite subject(s):** | Preliminary knowledge of mathematics, probability theory and mathematical statistics, computational methods, and programming is required. | | |
| **Compulsory Literature:** | Statistical analysis of data in R. / A.G. Bukhovets, P.V. Moskalev, V.P. Bogatova, T.Ya. Biryuchinskaya – Voronezh, VGAU. 2010. - 124 p. | | |

<table>
<tr><td></td><td>Paklin N.B. Business Intelligence: From Data to Knowledge: A Study Guide. 2nd ed. corrected / N.B. Paklin, V.I. Oreshkov - St. Petersburg: Peter. 2013 .-- 704 p.</td></tr>
<tr><td>Replacement Literature:</td><td>Rafolovich V. Data mining, or Data mining for the employed. Practical course. /AT. Rafolovich - “I-trade”, 2014.</td></tr>
<tr><td>Participation requirements:</td><td>Lower limit of lectures attendance is 80%, each task must be presented by end of the course.</td></tr>
<tr><td>Independent work:</td><td>1. The task of building dependency<br>2. The task of classification</td></tr>
<tr><td>Grading criteria scale or the minimal level necessary for passing the subject:</td><td><table><tr><td>Failed</td><td>< 50 points</td></tr><tr><td>Passed</td><td>>= 50 points</td></tr></table><br>Points distribution:<br>Ongoing assessment:<br>The task of building dependencies: 50 points<br>Classification Problem: 50 points</td></tr>
<tr><td>Information about the course:</td><td>Room ___, on ___ at ___</td></tr>
<tr><td>1) Date 1</td><td>Lecture 1. Overview of approaches to solving classification and regression problems. Statistical methods. Machine learning.<br><br>The unity and difference of classification and regression problems. Classification of uncertainties. Problem solving in the face of uncertainty. Statistical methods: linear regression, autoregression, logistic regression, Bayesian classifier. Machine learning: neural networks, decision trees, decision rules.</td></tr>
<tr><td rowspan="2">2) Date 2</td><td>Lecture 2. Fundamentals of statistical information processing. Linear regression.<br><br>Characteristics of random events, random variables, random processes. Selective method. The concept of stochastic dependence. Linear regression model. Least square method.</td></tr>
<tr><td>Practical 2<br>Knowledge consolidation on examples.</td></tr>
<tr><td>3) Date 3</td><td>Lecture 3. Choice of dimension of feature space.<br><br>Features of the studied object, measured on quantitative and qualitative scales. Requirements for the system of features. Basic concepts of factor analysis, selection of features in regression analysis, the method of principal components.<br>Practical 3<br>Knowledge consolidation on examples.</td></tr>
<tr><td>4) Date 4</td><td>Lecture 4. Logistic regression.</td></tr>
</table>

| | |
|---|---|
| | Statement of the classification problem. The concept of logit conversion. Approaches to the assessment of logistic regression parameters. |
| | **Practical 4**<br>Knowledge consolidation on examples. |
| **5) Date 5** | **Lecture 5.** Time series analysis methods.<br><br>Time series as a tool for describing a dynamic process. Definition of stationary and non-stationary time series. Models of autoregression and moving average. Violation of the Gauss-Markov conditions. Method of moments for estimating autoregressive parameters. Models for predicting the behavior of a series. |
| | **Practical 5**<br>Knowledge consolidation on examples. |
| **6) Date 6** | **Lecture 6.** ***Fuzzy uncertainty, fuzzy inference.***<br><br>Fuzzy sets as a method of formalizing fuzzy uncertainty. The basics of fuzzy logic. Fuzzy production systems. Fuzzy neural production networks like ANFIS. |
| | **Practical 6**<br>Knowledge consolidation on examples. |
| **7) Date 7** | **Lecture 7.** Methods of decision trees.<br><br>Classification using decision trees. Algorithms for constructing decision trees, measures of partition quality. An example of a partition using the Shannon entropy. The concept of random forest. |
| | **Practical 7**<br>Knowledge consolidation on examples. |
| **8) Date 8** | **Lecture 8.** ***Browse ensembles of classification models.***<br><br>Theoretical substantiation of the capabilities of the ensemble of "weak" algorithms. Types of ensembles: homogeneous, heterogeneous; component training, ensemble training. The formation of the output of the ensemble. The methods of formation of the ensemble: running, boosting, stacking. |
| | **Practical 8**<br>Knowledge consolidation on examples. |
| **9) Date 9** | **Assessment** |