ITMO University

# Machine Learning

| Course Workload | | Assessment form (examination/ graded test/ ungraded test) |
|---|---|---|
| ECTS | Hours | |
| 3 | 108 | Exam |

Students will gain an understanding of the basic tasks and methods of machine learning and data mining, study the statistical foundations of machine learning theory, learn how to solve the problem of classification, regression, clustering using machine learning methods and algorithms, and evaluate the quality of solutions, learn to apply software for machine learning for solving domain-specific problems.

**Course structure:**

## 1. Classification and regression methods

1.1. Supervised learning (tasks and algorithms)

1.2. Empirical risk minimization principle

1.3. Likelihood maximization principle. Bayes theorem

1.4. Offset error and scatter error. The scatter-displacement dilemma

1.5. Generalized metric classifier. Distance functions

1.6. Logical classification algorithm. Informativeness criteria

1.7. Types of features. Feature selection. One-hot-encoding.

1.8. Classification metrics. Accuracy, completeness, F-measure. Types of errors. Confusion matrix.

1.9. Decision tree construction algorithms. Gini index

1.10. Random forest. Algorithm for constructing an ensemble of decision trees

1.11. Decision tree reduction

1.12. Ensembles of models. Bagging

1.13. Boosting. AdaBoost Algorithm

1.14. AnyBoost algorithm. Gradient boosting

1.15. Gradient boosting over decisive trees

## 2. Learning without a teacher

2.1. Unsupervised learning (tasks and algorithms)

2.2. Nearest Neighbors Method

2.3. Parzen windows method

2.4. Multidimensional data visualization and dimensionality reduction algorithms

2.5. Clustering algorithms. Distance metrics. K-means and SVD

## 3. Neural networks

3.1. McCulloch-Pitts model. Rosenblatt's perceptron. Neural implementation of logical functions

3.2. Activation functions and error functions. Classification and regression

3.3. Gradient descent and its types

3.4. Multilayer neural network. Backpropagation method. The problem of fading gradients

3.5. Neural network training as an optimization problem. Stochastic gradient descent with moments, RMSProp and Adam algorithms

3.6. Training data: train, test, sample validation. Monitoring the network learning process

3.7. Using Parallel Computing in Deep Learning

3.8. Parallel optimization of neural network hyperparameters

3.9. Convolutional neural network architecture. Shared weights, folding, pooling, padding

3.10. Convolutional neural network architecture. Feature maps and links to computer vision algorithms

3.11. Modern convolutional networks. VGG-16 and ResNet models. Attention mechanism. Problems of convolutional networks and capsule networks

3.12. Sequence Analysis Problems. The principle of constructing recurrent neural networks

3.13. Recurrent neural networks. LSTM and GRU cells. Attention mechanism

3.14. Natural language processing tasks. A bag of words approach. TF-IDF method. Stamming. Lemmatization. Stop words

3.15. Vector word representation and embeddings. The word2vec model

3.16. Autoencoder architecture. Encoder and decoder

3.17. Autoencoders and code space. Generation of new objects. Variational autoencoders

3.18. Generative adversarial networks. GAN learning algorithm

3.19. Reinforcement learning problem. Markov decision making process

3.20. Reinforcement learning and the Q-learning algorithm

3.21. Gradients by strategy and the REINFORCE algorithm