**Government of Russian Federation**

**Federal State Autonomous Educational Institution of Higher Education**

**"National Research University Higher School of Economics"**

# Syllabus for the course
# "Discriminative methods in Machine Learning"

for postgraduate program in 09.06.01 Computer Science and Computer Engineering / 05.13.01 "Systems Analysis, Control Theory, and Information Processing", 05.13.11 "Mathematical Theory and Software for Computing Machinery, Systems, and Networks", 05.13.17 "Theoretical Foundations of Computer Science", 05.13.18 "Mathematical Modeling, Numerical Methods, and Software Systems"

Author:
Attila Kertesz-Farkas, assistant professor, akerteszfarkas@hse.ru

Approved by the Academic Council of the School for Postgraduate Studies in Computer Science on October 26, 2017

Moscow - 2017

## 1. Scope of Use

This program establishes the minimal requirements to postgraduate students’ knowledge and skills for 09.06.01 Computer Science and Computer Engineering / 05.13.01 “Systems Analysis, Control Theory, and Information Processing”, 05.13.11 “Mathematical Theory and Software for Computing Machinery, Systems, and Networks”, 05.13.17 “Theoretical Foundations of Computer Science”, “05.13.18 Mathematical Modeling, Numerical Methods, and Software Systems” and determines the content of the course and educational techniques used in teaching the course.

The present syllabus is aimed at faculty teaching the course and postgraduate students studying 09.06.01 Computer Science and Computer Engineering / 05.13.01 “Systems Analysis, Control Theory, and Information Processing”, 05.13.11 “Mathematical Theory and Software for Computing Machinery, Systems, and Networks”, 05.13.17 “Theoretical Foundations of Computer Science”, 05.13.18 “Mathematical Modeling, Numerical Methods, and Software Systems”.

This syllabus meets the standards required by:

- Educational standards of National Research University Higher School of Economics;
- Postgraduate educational program for 09.06.01 Computer Science and Computer Engineering.
- University curriculum of the postgraduate program for 09.06.01 Computer Science and Computer Engineering / 05.13.01 “Systems Analysis, Control Theory, and Information Processing”, 05.13.11 “Mathematical Theory and Software for Computing Machinery, Systems, and Networks”, 05.13.17 “Theoretical Foundations of Computer Science”, 05.13.18 Mathematical Modeling, Numerical Methods, and Software Systems”, approved in 2014.

## 2. Learning Objectives

The learning objective of the course “Discriminative methods in Machine Learning” is to provide students advanced techniques and deeper theoretical and practical knowledge in modern discriminative learning techniques, such as:

• Logistic regression, Support Vector Machines, regularization, Neural Networks, Deep Neural Networks, Limits on learning, Deep Learning techniques, Neural Turing Machines, Performance evaluation techniques, optimization algorithms.

## 3. Main Competencies Developed after Completing the Study of This Discipline

After completing the study of the discipline the PhD student should have:

• Knowledge about modern discriminative methods such as deep convolutional learning techniques, kernel machines, limitations of learning methods and standard definitions such as overfitting, regularization, etc.

• Knowledge about ongoing developments in Machine Learning

• Hands-on experience with large scale machine learning problems.

• Knowledge about how to design and develop machine learning programs using programming language Python.

• Think critically with real data.

After completing the study of the discipline the student should have developed the following competencies:

| **Competence** | **Code** | **Descriptors (indicators of achievement of the result)** | **Educative forms and methods aimed at generation and development of the competence** |
|---|---|---|---|
| the ability to carry out theoretical and experimental research in the field of professional activity | ОПК-1 | PhD students obtain necessary knowledge in probabilistic generative models | Assignments, additional material/reading provided |
| the ability to develop new research methods and apply them in research in one’s professional field | ОПК-2 | The PhD student is able to choose an appropriate model for real-life problems and to calibrate the hyperparameters. | Examples covered during the lectures and tutorials. Assignments. |
| the ability to objectively evaluate the outcomes of research and development carried out by other specialists in other scientific institutions | ОПК-4 | The PhD student is able to carry out comparative testing of competing models or methods. | Examples covered during the lectures and tutorials. Assignments. |
| the ability to do research in transformation of information into data and knowledge, models of data and knowledge representation, methods for knowledge processing, machine learning and knowledge discovery methods, principles of building and operating software for automation of these processes | ПК-4 | The PhD student is able to develop and analyze machine learning models, implement them in a programming language in large scale, and select the best model using validation techniques. | Lectures, tutorials, and assignments. |

## 4. Place of the Discipline in the Postgraduate Program Structure

This is an elective course for 05.13.01 “Systems Analysis, Control Theory, and Information Processing”, 05.13.11 “Mathematical Theory and Software for Computing Machinery, Systems, and Networks”, 05.13.17 “Theoretical Foundations of Computer Science”, 05.13.18 “Mathematical Modeling, Numerical Methods, and Software Systems”.

Postgraduate students are expected to be already familiar with some statistical learning techniques, and have skills in analysis, linear algebra, optimization, computational complexity, and probability theory.

The following knowledge and competences are needed to study the discipline:

- A good command of the English language, both oral and written.
- A sound knowledge of probability theory, complexity theory, optimization, and linear algebra

## 5. Schedule for one 1 module

| № | Topic | Total hours | Contact hours | | | Self-study |
|---|---|---|---|---|---|---|
| | | | Lectures | Seminars | Practice lessons | |
| 1. | Introduction of Machine Learning, performance evaluation | 12 | 2 | | 1 | 9 |
| 2. | Regression, Logistic regression, Support Vector Machines, Neural Networks, collaborative filtering, K-nearest Neighbor, decision trees | 48 | 8 | | 4 | 36 |
| 3. | Kernels and distance functions | 12 | 2 | | 1 | 9 |
| 4. | Auto-encoders, deep models, convolutional Neural Networks | 12 | 2 | | 1 | 9 |
| 5. | Sequential data, Recurrent Neural Networks, and long-short term memory models. | 24 | 4 | | 2 | 18 |
| 6. | Neural Turing Machines | 12 | 2 | | 1 | 9 |
| 7. | Optimization and Regularization | 12 | 2 | | 1 | 9 |
| 8. | Algorithm independent machine learning and No-free-lunch theorems | 20 | 2 | | 1 | 17 |
| | **Total** | 152 | 24 | | 12 | 116 |

## 6. Requirements and Grading

| | | |
|---|---|---|
| Mid-Term Exam | 1 | Mid-semester test. Written exam. |
| Presence | 1 | |
| Exam | 1 | Written exam. Preparation time – 180 min. |

## 7. Assessment

*Final assessments* are based on the mid-exam and the final exam. Students have to demonstrate knowledge of the material covered during the entire course.

## 8. The grade formula

***The exam*** is worth 60% of the final mark.

***Final course mark*** *is obtained from the following formula*: Final=0.2*(Mid-term exam)+ 0.2*(Presence on all lectures and seminars)+0.6*(Exam).

All grades having a fractional part greater than 0.5 are rounded up.

**Table of Grade Accordance**

| **Ten-point grading Scale** | **Five-point grading Scale** | |
|---|---|---|
| 1 - very bad<br>2 – bad<br>3 – no pass | Unsatisfactory - 2 | **FAIL** |
| 4 – pass<br>5 – highly pass | Satisfactory – 3 | **PASS** |
| 6 – good<br>7 – very good | Good – 4 | |
| 8 – almost excellent | Excellent – 5 | |

| | | |
|---|---|---|
| 9 – excellent<br>10 – perfect | | |

## 9. Course description.

### Topic 1. Introduction to machine learning, Evaluation techniques

Basic definitions or machine learning, principles and types of machine learning, performace metrics, errors and type of errors. ROC characteristics.

### Topic 2. Basic methods.

Regression, Logistic regression, Support Vector Machines, Neural Networks, Collaborative Filtering, K-nearest Neighbor, decision trees, random forests.

### Topic 3. Kernels and distance functions

Kernel functions for real-valued vectors and for discrete models. Distance functions, edit distance, and information distance. Course of dimensionality.

### Topic 4. Deep Neural Networks

Auto Encoders, deep neural networks, stacked auto encoders, convolutional layers and max-pooling. Deep data vs. wide data, universal approximators, supervised and unsupervised pre-training

### Topic 5. Methods for sequential data

Sequential data, Recurrent Neural Networks, and long-short term memory models.

### Topic 6. Neural Turing Machines

And is applications

### Topic 7. Optimization and Regularization

Error Surfaces, Optimization and Regularization methods: stochastic gradient descent, momentum methods, polyak averaging, coordinate descent, adaptive learning rates, line-search, adaGrad, RMprop, Second order methods: Levenberg–Marquardt, Newton, conjugate gradients, Broyden–Fletcher–Goldfarb–Shanno. Regularization: parameter norm penalty, early stopping, data augmentation, sparse coding, mini batches vs sharp minima, batch normalization.

### Topic 8. Algorithm independent machine learning and No-free-lunch theorems

Regularization, overfitting-underfitting, bias-variance decomposition in model selection, model capacity, minimum description length, parameters and hyper parameters, other problems: missing values and class imbalance problem. Bootstrap and Jackknife estimations. No-Free-lunch theorems. Interpretability. Bias.

During the practice sessions the class implements neural networks in python+tensorflow, and some deep convolutional models from scratch.

## 10. Educational technologies

The following educational technologies are used in the study process:

- discussion and analysis of the results during the tutorials;
- regular assignments to test the progress of the PhD student;

- consultation time on Monday afternoons.

## 11. Final exam questions

The final exam will consist of a selection of problems equally weighted. No material is allowed for the exam. Each question will focus on a particular topic presented during the lectures.

The questions consist in exercises on any topic seen during the lectures. To be prepared for the final exam, PhD students must be able to answer questions from the topics covered during the lecture.

## 12. Reading and Materials

**Literature:**

1. Kevin Murphy, Machine Leaning: A probabilistic Perspective, 2013, MIT press
2. C. Bishop: Pattern Recognition and Machine Learning,
3. I. Goodfellow, Y. Bengio and A. Courville, Deep Learning. 2016 MIT press
4. G. James, D. Witten, T. Hastie, R. Tibshirani. An introduction to Statistical Learning, 2013, Springer
5. Li Deng, Dong Yu: Deep Learning: Methods and Applications, 2014, Now publishers.
6. M. J. Wainwright, M. I. Jordan: Graphical Models, Exponential Families, and Variational Inference, 2008, Now publishers

**Literature for self-study:**

1. **Y Bengio:** Learning Deep Architectures for AI; Machine Learning, 2009, Vol. 2, No. 1,

## 13. Equipment.

The course requires a computer room, laptop and a projector.