Government of Russian Federation

Federal State Autonomous Educational Institution of Higher Education

“National Research University Higher School of Economics”

Syllabus for the course
“Generative 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 "Generative methods in Machine Learning" is to provide students advanced techniques and deeper theoretical and practical knowledge in modern probabilistic learning techniques, such as:

- Basic principles, Generative Models,
- Bayesian Network, Random Markov Fields, Boltzmann Machines, Variational Auto Encoders
- Sampling and Inference, Variational inference, variational methods,
- Neural Networks,
- Deep Learning techniques.

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 probabilistic models.
- Knowledge about modern methods such as deep learning techniques.
- 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 a programming language such as R or 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 semesters (2 modules)

№	Topic	Total hours	Contact hours			Self-study
			Lec-tures	Semi-nars	Practice lessons	
1.	Introduction probabilistic modelling	12	2		1	9
2.	Bayesian Learning, Exponential Families	12	2		1	9
3.	Graphical Models, Generative Learning	24	4		2	18
4.	Sampling and Inference	12	2		1	9
5.	Variational Learning	36	6		3	27
6.	Generative models	32	4		2	26
7.	Deep learning techniques	24	4		2	18
	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 2 – bad 3 – no pass	Unsatisfactory - 2	**FAIL**
4 – pass 5 – highly pass	Satisfactory – 3	**PASS**
6 – good 7 – very good	Good – 4	
8 – almost excellent 9 – excellent 10 – perfect	Excellent – 5	

9. Course description.

Topic 1. Introduction to machine learning, Bayesian Decision Theory, Maximum Likelihood Estimation, and EM.

Basic definitions, principles and types of machine learning. Classifiers, Discriminant Functions, and Decision Surfaces, Minimum-Error-Rate Classification, Neyman-Pearson lemma, Distributions, Relation to Logistic Regression, Naïve Bayes classification, basics of MLE, learning parameters of distributions. Gaussian Mixture Models, Latent Variables, Examples, Expectation-Maximization, Latent Dirichlet Allocation.

Topic 2. Exponential Family, Sufficient Statistics.

Generalized Linear Models,

Topic 3. Graphical Models and Generative Learning

Bayesian Networks, Random Markov Fields, Conditional Random Fields, Boltzmann Machines, Energy-based methods. Hidden Markov Models. During the practice session the class implements a speech recognition system.

Topic 4. Sampling and Inference

Exact and Inexact Inference, Gibbs sampling, Bridge Sampling, Simple and Annealed Importance Sampling, Monte-Carlo EM, Junction Tree algorithm

Topic 5. Variational Learning

Mean-Field, Bethe Approximation, Variational methods, Variational Message Passing, Free-Energy, Variational Free Energy. Variational Bayes, Variational Bayes Expectation-Maximization. Mean field methods. Variational AutoEncoders.

Topic 6. Generative Models

Restricted Boltzmann Machines, Helmoltz Machines and Wake-Sleep algorithms, Energy-based methods. Generative Adversarial Networks, Generative Auto-Encoders, Belief networks, connectionist learning.

Topic 7. Deep learning techniques

Neural Networks, Shallow networks, Multilayer Neural networks, back-propagation, deep learning, Universal Approximation. Auto Encoders, Stacked Auto-Encoders, Stacked Boltzmann machines, supervised and unsupervised pre-training, Deep Belief Networks. Deep Universality theorems.

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.