

**Federal State Autonomous Educational Institution of Higher Education "Moscow
Institute of Physics and Technology
(National Research University)"**

APPROVED

**Head of the Phystech School of
Applied Mathematics and
Informatics**

A.M. Raygorodskiy

Work program of the course (training module)

course: Machine Learning with Graphs/Машинное обучение на графах
major: Applied Mathematics and Informatics
specialization: Modern State of Artificial Intelligence/Современные методы искусственного интеллекта
“Pusk” Online and Supplementary Education Centre
Chair of Machine Learning and Digital Humanities
term: 1
qualification: Master

Semester, form of interim assessment: 2 (spring) - Grading test

Academic hours: 45 AH in total, including:

lectures: 15 AH.

seminars: 30 AH.

laboratory practical: 0 AH.

Independent work: 45 AH.

In total: 90 AH, credits in total: 2

Number of course papers, tasks: 2

Author of the program: R.G. Neychev, assistant

The program was discussed at the Chair of Machine Learning and Digital Humanities 04.06.2020

Annotation

Graphs allow us to represent data and processes naturally in many situations. This course extends the scope of the Machine Learning methods to data represented as graphs and focuses on novel approaches and applied techniques of graph analysis. The main goal of this course is to build a thorough theoretical background and provide practical experience to apply Graph Machine Learning methods in real projects and to use them in other domains like Computer Vision, Natural Language Processing or Reinforcement Learning as well..

1. Study objective

Purpose of the course

1. Get familiar with classical approaches to graph analysis
2. Learn the novel methods of information retrieval using Graph Neural Networks
3. Get hands on experience in working with graph representation of the data

Tasks of the course

1. Graph Machine Learning problem statement and ability to develop the general pipeline of the solution
2. Choose relevant approach to and model for particular problem
3. Essential experience with Python, PyTorch and PyTorch frameworks

2. List of the planned results of the course (training module), correlated with the planned results of the mastering the educational program

Mastering the discipline is aimed at the formation of the following competencies:

Code and the name of the competence	Competency indicators
Gen.Pro.C-4 Combine and adapt current information and communications technologies (ICTs) to meet professional challenges	Gen.Pro.C-4.2 Apply ICTs to solve the task in hand, to draw conclusions, and to evaluate the obtained results
	Gen.Pro.C-4.3 Create original algorithms and use software tools and modern smart technologies for professional problem-solving
Pro.C-2 Understands and is able to apply modern mathematical apparatus and algorithms, the basic laws of natural science, modern programming languages and software; operating systems and networking technologies in research and applied activities	Pro.C-2.3 Use fundamental knowledge in the field of information theory to carry out research tasks

3. List of the planned results of the course (training module)

As a result of studying the course the student should:

know:

fundamental concepts, laws, theories of algebraic methods in number theory;
modern problems of the relevant sections of the theory of algebraic methods in number theory;
concepts, axioms, methods of proof and proof of the main theorems in the sections included in the basic part of the cycle of the theory of algebraic methods in number theory;
basic properties of the corresponding mathematical objects;
analytical and numerical approaches and methods for solving typical applied problems of the theory of algebraic methods in number theory.

be able to:

understand the task;
use your knowledge to solve fundamental and applied problems;
evaluate the correctness of the problem statements;
strictly prove or disprove the statement;
independently find algorithms for solving problems, including non-standard ones, and conduct their analysis;
independently see the consequences of the results;
accurately represent mathematical knowledge in topology orally and in writing.

master:

skills of mastering a large amount of information and solving problems (including complex ones);
 skills of independent work and mastering new disciplines;
 the culture of the formulation, analysis and solution of mathematical and applied problems that
 require the use of mathematical approaches and methods for their solution;
 the subject language of topology and the skills of competent description of problem solving and
 presentation of the results.

4. Content of the course (training module), structured by topics (sections), indicating the number of allocated academic hours and types of training sessions

4.1. The sections of the course (training module) and the complexity of the types of training sessions

№	Topic (section) of the course	Types of training sessions, including independent work			
		Lectures	Seminars	Laboratory practical	Independent work
1	Graph properties	5	10		15
2	Node embeddings	5	10		15
3	Link prediction	5	10		15
AH in total		15	30		45
Exam preparation		0 AH.			
Total complexity		90 AH., credits in total 2			

4.2. Content of the course (training module), structured by topics (sections)

Semester: 2 (Spring)

1. Graph properties

Traditional ML methods for graphs, Graph neural networks, Graph traversing methods

2. Node embeddings

Knowledge graph embeddings, Node propagation

3. Link prediction

Geometrical priors in ML

5. Description of the material and technical facilities that are necessary for the implementation of the educational process of the course (training module)

A standard classroom.

6. List of the main and additional literature, that is necessary for the course (training module) mastering

Main literature

1. Graph theory /J. A. Bondy, U. S. R. Murty. London, Springer, 2008

Additional literature

7. List of web resources that are necessary for the course (training module) mastering

<http://dm.fizteh.ru/>

8. List of information technologies used for implementation of the educational process, including a list of software and information reference systems (if necessary)

Multimedia technologies can be used in lectures and practical exercises, including presentations.

9. Guidelines for students to master the course

1. It is recommended to successfully pass the test papers, as this simplifies the final certification in the subject.
2. To prepare for the final certification in the subject, it is best to use the lecture materials.

Assessment funds for course (training module)

major: Applied Mathematics and Informatics
specialization: Modern State of Artificial Intelligence/Современные методы искусственного интеллекта
“Pusk” Online and Supplementary Education Centre
Chair of Machine Learning and Digital Humanities
term: 1
qualification: Master
Semester, form of interim assessment: 2 (spring) - Grading test
Author: R.G. Neychev, assistant

1. Competencies formed during the process of studying the course

Code and the name of the competence	Competency indicators
Gen.Pro.C-4 Combine and adapt current information and communications technologies (ICTs) to meet professional challenges	Gen.Pro.C-4.2 Apply ICTs to solve the task in hand, to draw conclusions, and to evaluate the obtained results
	Gen.Pro.C-4.3 Create original algorithms and use software tools and modern smart technologies for professional problem-solving
Pro.C-2 Understands and is able to apply modern mathematical apparatus and algorithms, the basic laws of natural science, modern programming languages and software; operating systems and networking technologies in research and applied activities	Pro.C-2.3 Use fundamental knowledge in the field of information theory to carry out research tasks

2. Competency assessment indicators

As a result of studying the course the student should:

know:

fundamental concepts, laws, theories of algebraic methods in number theory;
modern problems of the relevant sections of the theory of algebraic methods in number theory;
concepts, axioms, methods of proof and proof of the main theorems in the sections included in the basic part of the cycle of the theory of algebraic methods in number theory;
basic properties of the corresponding mathematical objects;
analytical and numerical approaches and methods for solving typical applied problems of the theory of algebraic methods in number theory.

be able to:

understand the task;
use your knowledge to solve fundamental and applied problems;
evaluate the correctness of the problem statements;
strictly prove or disprove the statement;
independently find algorithms for solving problems, including non-standard ones, and conduct their analysis;
independently see the consequences of the results;
accurately represent mathematical knowledge in topology orally and in writing.

master:

skills of mastering a large amount of information and solving problems (including complex ones);
skills of independent work and mastering new disciplines;
the culture of the formulation, analysis and solution of mathematical and applied problems that require the use of mathematical approaches and methods for their solution;
the subject language of topology and the skills of competent description of problem solving and presentation of the results.

3. List of typical control tasks used to evaluate knowledge and skills

Example Lab assignment

Use PyTorch Geometric framework to build Graph Neural Network for one of the following tasks:

- Graph/subgraph property prediction
- Node property prediction
- Link prediction
- Graph classification

using Open Graph Benchmark datasets.

Answer the inline questions on the data properties. Measure the time and memory complexity of the graph machine learning models. Write down the report explaining the results.

4. Evaluation criteria

Questions

1. Graph representations: adjacency matrix, adjacency list, etc.
2. Node-level features, graphlets, their properties
3. Graph-level features, bag-of-graphlets, their properties
4. Node embeddings: Node2Vec, DeepWalk: main ideas, advantages
5. Link analysis. PageRank algorithm, applications
6. Graph traversal algorithms. Complexity analysis.
7. Graph Neural Networks: GCNs, GNNs.

Assessment “excellent (10)” is given to a student who has displayed comprehensive, systematic and deep knowledge of the educational program material, has independently performed all the tasks stipulated by the program, has deeply studied the basic and additional literature recommended by the program, has been actively working in the classroom, and understands the basic scientific concepts on studied discipline, who showed creativity and scientific approach in understanding and presenting educational program material, whose answer is characterized by using rich and adequate terms, and by the consistent and logical presentation of the material;

Assessment “excellent (9)” is given to a student who has displayed comprehensive, systematic knowledge of the educational program material, has independently performed all the tasks provided by the program, has deeply mastered the basic literature and is familiar with the additional literature recommended by the program, has been actively working in the classroom, has shown the systematic nature of knowledge on discipline sufficient for further study, as well as the ability to amplify it on one’s own, whose answer is distinguished by the accuracy of the terms used, and the presentation of the material in it is consistent and logical;

Assessment “excellent (8)” is given to a student who has displayed complete knowledge of the educational program material, does not allow significant inaccuracies in his answer, has independently performed all the tasks stipulated by the program, studied the basic literature recommended by the program, worked actively in the classroom, showed systematic character of his knowledge of the discipline, which is sufficient for further study, as well as the ability to amplify it on his own;

Assessment “good (7)” is given to a student who has displayed a sufficiently complete knowledge of the educational program material, does not allow significant inaccuracies in the answer, has independently performed all the tasks provided by the program, studied the basic literature recommended by the program, worked actively in the classroom, showed systematic character of his knowledge of the discipline, which is sufficient for further study, as well as the ability to amplify it on his own;

Assessment “good (6)” is given to a student who has displayed a sufficiently complete knowledge of the educational program material, does not allow significant inaccuracies in his answer, has independently carried out the main tasks stipulated by the program, studied the basic literature recommended by the program, showed systematic character of his knowledge of the discipline, which is sufficient for further study;

Assessment “good (5)” is given to a student who has displayed knowledge of the basic educational program material in the amount necessary for further study and future work in the profession, who while not being sufficiently active in the classroom, has nevertheless independently carried out the main tasks stipulated by the program, mastered the basic literature recommended by the program, made some errors in their implementation and in his answer during the test, but has the necessary knowledge for correcting these errors by himself;

Assessment “satisfactory (4)” is given to a student who has discovered knowledge of the basic educational program material in the amount necessary for further study and future work in the profession, who while not being sufficiently active in the classroom, has nevertheless independently carried out the main tasks stipulated by the program, learned the main literature but allowed some errors in their implementation and in his answer during the test, but has the necessary knowledge for correcting these errors under the guidance of a teacher;

Assessment “satisfactory (3)” is given to a student who has displayed knowledge of the basic educational program material in the amount necessary for further study and future work in the profession, not showed activity in the classroom, independently fulfilled the main tasks envisaged by the program, but allowed errors in their implementation and in the answer during the test, but possessing necessary knowledge for elimination under the guidance of the teacher of the most essential errors;

Assessment “unsatisfactory (2)” is given to a student who showed gaps in knowledge or lack of knowledge on a significant part of the basic educational program material, who has not performed independently the main tasks demanded by the program, made fundamental errors in the fulfillment of the tasks stipulated by the program, who is not able to continue his studies or start professional activities without additional training in the discipline in question;

Assessment “unsatisfactory (1)” is given to a student when there is no answer (refusal to answer), or when the submitted answer does not correspond at all to the essence of the questions contained in the task.

5. Methodological materials defining the procedures for the assessment of knowledge, skills, abilities and/or experience

During examination the student are allowed to use the program of the discipline.