

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

Факультет комп'ютерних наук та кібернетики
Кафедра математичної інформатики



«ЗАТВЕРДЖУЮ»

Заступник декана
з навчальної роботи

Кашпур О.Ф.

«квітень» 2020 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ Навчання з підкріпленням/Reinforcement Learning

для студентів / for students

галузь знань 12 “Інформаційні технології”/“Information Technologies”
спеціальність 122 “Комп'ютерні науки”/“Computer Science”
освітній рівень магістр/masters
освітня програма Математичні Методи Штучного Інтелекту /
Mathematical Methods of Artificial Intelligence
вид дисципліни вільний вибір студента / free choice

Форма навчання	денна
Навчальний рік	2020/2021
Семестр	3
Кількість кредитів ECTS	6
Мова викладання, навчання та оцінювання	англійська \ English
Форма заключного контролю	іспит/exam

Викладачі: к.т.н., доцент Радивоненко О.С.
ас. Терещенко Я.В.

Пролонговано: на 2021/2022 н.р.
на 20 /20 н.р.

 «11» 02 2021 р.
() « » 20 р.

КИЇВ – 2020

Розробники: Радивоненко Ольга Сергіївна, к.т.н., доцент, керівник підрозділу
Самсунг Електронікс Україна Компані
Терещенко Ярослав Васильович, асистент кафедри математичної інформатики

ЗАТВЕРДЖЕНО

Зав. кафедри математичної інформатики

(підпис)

(Терещенко В.М.)
(прізвище та ініціали)

Протокол № 2 від «2» листопада 2020 р.

Схвалено Гарантом освітньо-наукової програми «Математичні методи штучного
інтелекту» _____ (Завадський І.О.)

Схвалено науково-методичною комісією факультету комп'ютерних наук та кібернетики

Протокол від «2» листопада 2020 року № 9

Голова науково-методичної комісії _____ (Омельчук Л.Л.)
(підпис) (прізвище та ініціали)

«2» листопада 2020 року

1. Мета дисципліни «Навчання з підкріпленням/Reinforcement Learning» - ознайомлення студентів з основами машинного навчання з підкріпленням, як однієї із основних галузей штучного інтелекту, та сучасними методами розв'язання задач побудови алгоритмів прийняття рішень. Підготувати студента до ефективного використання сучасних методів таких, як методи машинного навчання для створення систем штучного інтелекту у подальшій професійній діяльності; допомогти набуту навичок практичної роботи із сучасними програмними засобами для побудови інтелектуальних моделей.

/

Discipline aim. The aim of the discipline is to acquaint students with the basics of reinforcement learning as one of the main branches of artificial intelligence, and modern methods for solving problems of building decision-making algorithms. To prepare the student for the effective use of modern methods such as machine learning methods to create artificial intelligence systems in further professional activities; help to acquire skills of practical work with modern software for building intelligent models.

2. Попередні вимоги до опанування або вибору навчальної дисципліни:

1. *Знати:* базову дисципліну - «Машинне навчання», володіти методами математичного аналізу, лінійної алгебри та геометрії, теорії ймовірностей та математичної статистики, програмування, алгоритмів і обробки структур даних.

2. *Вміти:* розробляти, аналізувати та застосовувати алгоритми та програмне забезпечення для розв'язання завдань та прикладних задач, використовуючи сучасні методи розробки програм.

/

1. *To know:* the basic discipline - "Machine Learning", master the methods of mathematical analysis, linear algebra and geometry, probability theory and mathematical statistics, programming, algorithms and data structures processing.

2. *To be able to:* develop, analyze and apply algorithms and software to solve problems and applied tasks using modern software development methods.

3. Анотація навчальної дисципліни / Synopsis of the course:

Навчальна дисципліна «Навчання з підкріпленням/Reinforcement Learning» є складовою освітньо-професійної програми підготовки фахівців за освітньо-кваліфікаційним рівнем «магістр» галузі знань 12 «Інформаційні технології» спеціальності 122 «Комп'ютерні науки», освітньої програми «Математичні методи штучного інтелекту». Вона є базовою дисципліною вишів, які спеціалізуються у галузі інформаційних та комп'ютерних технологій, а також ефективним інструментом розв'язування наукових та інженерних задач.

/

The discipline "Reinforcement Learning" is a component of the educational-professional training program at the educational-qualification level "Master" in the field of knowledge 12 "Information Technology" specialty 122 "Computer Science", educational program "Mathematical methods of artificial intelligence". It is a basic discipline of universities specializing in information and computer technology, as well as an effective tool for solving scientific and engineering problems.

4. Завдання (навчальні цілі) / Objectives of study:

- здатність аналізувати та використовувати інтелектуальні інформаційні технології машинного навчання (СК20.1);
- здатність до проектування та реалізації систем штучного інтелекту (СК21.1);
- оволодіти нейронними мережами як одним із сучасних провідних наукових напрямків штучного інтелекту, а також оволодіти технологією розв'язання широкого класу задач науки та техніки (зокрема, задач розпізнавання зображень, обробки природної мови,

пошуку рішень, комп'ютерного зору, інтелектуального управління) за допомогою методів, підходів та алгоритмів нейронних мереж.

/

Objectives (learning objectives):

- ability to analyze and use intelligent information technologies of machine learning (SC20.1);
- ability to design and implement artificial intelligence systems (SC21.1);
- to master neural networks as one of the modern leading scientific directions of artificial intelligence, and also to master technology of the decision of a wide class of problems of science and technology (in particular, problems of image recognition, natural language, processing, search of decisions, computer sight, intellectual management) by means of methods, approaches and algorithms of neural networks.

5. Результати навчання за дисципліною/ Results of learning:

Результат навчання (РН) (1. знати; 2. вміти; 3. комунікація[□]; 4. автономність та відповідальність[□])	Форми (та/або методи і технології) викладання і навчання	Методи оцінювання та пороговий критерій оцінювання (за необхідності)		Відсоток у підсумковій оцінці з дисципліни
Код	Результат навчання			
РН 1.1	Знати постановку основних задач навчання з підкріпленням / To know the formulation of the main tasks of reinforcement learning	Лекція / Lecture	Контрольні роботи, активна робота на лекції, усні відповіді, іспити/ Tests, active work on lectures, oral answers, exam	20%
РН 1.2	Знати основні підходи до розв'язання задач навчання з підкріпленням / To know the basic approaches to solving reinforcement learning problems	Лекція / Lecture		
РН 1.3	Знати основні поняття та методи машинного навчання / To know the basic concepts and methods of machine learning	Лекція / Lecture	Контрольні роботи, активна робота на лекції, усні відповіді, іспити/ Tests, active work on lectures, oral answers, exam	20%
РН 1.4	Знати основні поняття та принципи роботи штучних нейронних мереж / To know the basic concepts and principles of artificial neural networks	Лекція / Lecture		
РН 2.1	Вміти проводити аналіз задачі для вибору найкращого методу її розв'язання / Be able to analyze the problem to choose the best method for solving it	Лекція, самостійна робота / Lecture, individual work	Контрольні роботи, виконання завдань, винесених на самостійну роботу, іспити / Tests, assignments submitted for	20%

			<i>independent work, exam</i>	
PH 2.2	Вміти формулювати професійні завдання на мові машинного навчання і основі формулювань базових задач навчання з підкріпленням/ Be able to formulate professional tasks in the language of machine learning and on the basis of formulations of basic problems of reinforcement learning	<i>Лекція, самостійна робота/ Lecture, individual work</i>	<i>Контрольні роботи, виконання завдань, винесених на самостійну роботу, ісnum / Tests, assignments submitted for independent work, exam</i>	20%
PH 2.3	Вміти застосовувати методи машинного навчання для розв'язання прикладних задач/ Be able to apply machine learning methods to solve applied problems	<i>Лекція, самостійна робота/ Lecture, individual work</i>	<i>Контрольні роботи, активна робота на лекції, усні відповіді, ісnum / Tests, active work on lectures, oral answers, exam</i>	20%

6. Співвідношення результатів навчання дисципліни із програмними результатами навчання (необов'язково для вибірових дисциплін, які не входять до блоків спеціалізації) / Correspondence between learning outcomes and program study outcomes.

Результати навчання дисципліни Програмні результати навчання	PH 1.1	PH 1.2	PH 1.3	PH 1.4	PH 2.1	PH 2.2	PH 2.3
ПРН18.1. Володіти основними методами машинного навчання. / PLO18.1. Master basic methods of machine learning.	+	+	+	+			
ПРН19.1. Використовувати теорію ігор, мультиагентні системи, навчання з підкріпленням та генетичні алгоритми. / PLO19.1. Use game theory, multi-agent systems, reinforced learning and genetic algorithms.					+	+	+

7. Схема формування оцінки.

7.1. Форми оцінювання студентів:

- семестрове оцінювання:

1. Активна робота на лекції, усні відповіді: PH1.1, PH1.2, PH1.3, PH1.4, PH2.3 – 5 балів / 3 бали;
2. Виконання завдань, винесених на самостійну роботу: PH2.1, PH2.2 – 10 балів / 6 балів;
3. Контрольна робота 1: PH1.1, PH1.2, PH2.1 – 15 балів / 9 балів;
4. Контрольна робота 2: PH1.3, PH2.2 – 15 балів / 9 балів;
5. Контрольна робота 3: PH1.4, PH2.3 – 15 балів / 9 балів;

Здобувач освіти може бути недопущений до підсумкового оцінювання, якщо під час семестру він: 1) не досяг мінімального порогового рівня (60%) оцінки тих результатів навчання, які не можуть бути оцінені під час підсумкового контролю; 2) набрав кількість балів, що є

недостатньою для отримання позитивної оцінки навіть у випадку досягнення ним на підсумковому контролі максимально можливого результату.

Рекомендований мінімум – 36 балів

/

1. *Active work on lectures, oral answers:* LO1.1, LO1.2, LO1.3, LO1.4, LOH2.3 – 5 points / 3 points;
2. *Assignments submitted for independent work:* LO2.1, LO2.2 – 10 points / 6 points;
3. *Test 1:* LO 1.1, LO 1.2, LO2.1 - 15 points / 9 points.
4. *Test 2:* LO1.3, LO2.2 - 15 points / 9 points.
5. *Test 3:* LO1.4, LO2.3 - 15 points / 9 points.

An applicant may not be admitted to the final assessment if during the semester he: 1) has not reached the minimum threshold level (60%) of the assessment of those learning outcomes that cannot be assessed during the final control; 2) scored the number of points, which is insufficient to obtain a positive assessment, even if he achieves the maximum possible result in the final control.

The recommended minimum is 36 points.

- підсумкове оцінювання: іспит.

- *максимальна кількість балів які можуть бути отримані студентом:* 40 балів;

- *результати навчання які будуть оцінюватись:* PH1.1, PH1.2, PH1.3, PH1.4, PH2.1, PH2.2, PH2.3;

- *форма проведення і види завдань:* письмова робота.

Рекомендований мінімум – 24 балів.

/

- final assessment: exam.

- the maximum number of points that can be obtained: 40 points;

- learning outcomes that will be evaluated: LO1.1, LO1.2, LO1.3, LO1.4, LO2.1, LO2.2, LO2.3;

- form and types of tasks: written work.

The recommended minimum is 24 points.

Види завдань: 4 письмових питання.

- 1 питання: PH1.1, PH1.3, PH2.1;
- 2 питання: PH1.2, PH1.3, PH2.2;
- 3 питання: PH1.3, PH1.3, PH2.3;
- 4 питання: PH1.4, PH2.1, PH2.2, PH2.3;

За розгорнуту відповідь на кожне завдання студент може отримати від 1 до 10 балів.

Критерії оцінювання відповіді студента на питання:

- повнота розкриття питання – 1-4 бали;
- логіка викладення – 1-2 бали;
- аналітичні міркування – 1-4 бали.

/

Types of tasks: 4 written questions.

- 1th question: LO1.1, LO1.3, LO2.1;
- 2nd question: LO1.2, LO1.3, LO2.2;
- 3rd question: LO1.3, LO1.3, LO2.3;
- 4th question: LO1.4, LO2.1, LO2.2, LO2.4.

For a detailed answer to each task, a student can receive from 1 to 10 points.
Criteria for evaluating the student's answer to the question:

- completeness of the question – 1-4 points;
- logic of presentation – 1-2 points;
- analytical considerations – 1-4 points.

7.2. Організація оцінювання:

Терміни проведення форм оцінювання:

1. Активна робота на лекції, усні відповіді: протягом семестру;
2. Виконання завдань, винесених на самостійну роботу: протягом семестру;
3. Контрольна робота 1: РН1.1, РН1.2, РН2.1 – до 5 тижня семестру;
4. Контрольна робота 2: РН1.3, РН2.2 – до 10 тижня семестру;
5. Контрольна робота 3: РН1.4, РН2.3 – до 14 тижня семестру;

Студент має право на одне перескладання кожної контрольної роботи із можливістю отримання максимально 80% початково визначених за цю контрольну роботу балів. Термін перескладання визначається викладачем.

У випадку відсутності студента з поважних причин відпрацювання та перездачі контрольних робіт здійснюються у відповідності до „Положення про порядок оцінювання знань студентів при кредитно-модульній системі організації навчального процесу” від 1 жовтня 2010 року.

/

Terms of evaluation forms:

1. Active work on lectures, oral answers: during the semester;
2. Assignments submitted for independent work: during the semester;
3. Test 1: up to 5 weeks of the semester..
4. Test 2: up to 10 weeks of the semester..
5. Test 5: up to 14 weeks of the semester..

he student has the right to one retake of each test with the possibility of obtaining a maximum of 80% of the points initially determined for this test. The term of reassembly is determined by the teacher.

In case of absence of a student for valid reasons working off and transfer of tests are carried out according to "Regulations on the order of an estimation of knowledge of students at the credit-modular system of the organization of educational process" from October 1, 2010.

7.3 Шкала відповідності оцінок / Rating scale:

Відмінно / Excellent	90-100
Добре / Good	75-89
Задовільно / Satisfactory	60-74
Незадовільно / Fail	0-59

8. Структура навчальної дисципліни. Тематичний план лекцій

/ The structure of the discipline. Thematic plan of lectures and independent work

№ лекції	Назва лекції	Кількість годин		
		Лекції	Семінари/ лабораторні	Самостійна робота

Частина І. Вступ до навчання з підкріпленням/ Module I. “Introduction to reinforcement learning”				
1	Тема 1. Основні поняття навчання з підкріпленням <i>Самостійна робота:</i> Провести аналіз програми і змісту курсу та рекомендованої літератури. / Theme 1 Basic concepts of reinforcement learning: To analyze the program and course content, as well as recommended literature.	2		12
2	Тема 2. Марківський процес вирішування для планування <i>Самостійна робота:</i> Застосувати алгоритми PI, VI для різних вхідних даних / Theme 2. Markov decision process (MDP) for planning . <i>Individual work:</i> Apply Policy Iteration, Value Iteration algorithms on different input data	2	2	12
3	Тема 3. Навчання з підкріпленням для визначення поведінки. Пошук оптимальною політики методами Монте Карло та TD – навчання. Застосування динамічного програмування для оптимізації. <i>Самостійна робота:</i> Провести аналіз та застосувати MC та TD для різних конфігурацій середовища. / Theme 3. Reinforced learning policy evaluation. Optimal policy search using Monte Carlo and TD-learning methods. Application of dynamic programming for optimization.. <i>Individual work:</i> Analyze and apply MC and TD for different environment configurations.	2		14
4	Тема 4. Q-навчання. Узагальнений метод PI. Керування методом Монте Карло використовуючи жадібні алгоритми для визначення політики. TD методи, SARSA та Q-навчання для табличного контролю. <i>Самостійна робота:</i> Застосовувати методи Q-навчання для різних умов. Провести оцінку складності алгоритмів. / Theme 4. Q-learning. Generalized Policy Iteration method. MC Control with e-greedy policies. TD methods, SARSA and Q-learning for tabular control. <i>Individual work:</i> Apply Q-learning methods for different conditions. Estimate the complexity of the algorithms.	2	2	14
Контрольна робота I/ Control work I		2		1

Частина II. Глибоке навчання з підкріпленням/ Module II. Deep reinforcement learning				
5	Тема 5. Основні поняття навчання з підкріпленням. Приклади існуючих систем. <i>Самостійна робота:</i> Провести аналіз існуючих систем та оцінити їхню складність. / Theme 5. Basic concepts of deep reinforcement learning. Examples of existing systems. <i>Individual work:</i> Analyze existing systems and assess their complexity.	2	2	14
6	Тема 6. Імітаційне навчання. <i>Самостійна робота:</i> Навчити власну модель використовуючи імітаційне навчання для певного класу задач. / Theme 6. Imitation learning <i>Individual work:</i> Train own model using imitation learning for a certain class of tasks.	2	2	14
7	Тема 7. Визначення градієнтів політики <i>Самостійна робота:</i> Опрацювати методи визначення градієнтів політики / Theme 7. Policy Gradients <i>Individual work:</i> To develop methods for determining policy gradients	2		12
8	Тема 8. Глибоке навчання з підкріпленням використовуючи Q-learning. <i>Самостійна робота:</i> Опрацювати алгоритми Q-learning для глибокого навчання з підкріпленням / Theme 8. Deep reinforcement learning using Q-learning. <i>Individual work:</i> To develop Q-learning algorithms for deep reinforcement learning	2	2	14
Контрольна робота 2/ Control work 2		2		1
Частина III. Просунуті алгоритми глибокого навчання з підкріпленням/ Module III. Advanced algorithms of deep reinforcement learning				
9	Тема 9. Навчання з підкріпленням та апроксимація функції. Просунуті градієнти політики. Оптимальний контроль та планування. <i>Самостійна робота:</i> Опрацювати основні методи апроксимації функцій. Провести аналіз підходів для планування та контролю. / Theme 9. Reinforced learning and function approximation. Advanced policy gradients. Optimal control and planning.	4	2	14

	<i>Individual work:</i> To develop the main approximation methods. Analyze approaches to planning and control.			
10	Тема 10. Навчання з підкріпленням на базі існуючої моделі. Визначення політики на базі існуючої моделі. <i>Самостійна робота:</i> Опрацювати глибоке навчання з підкріпленням на базі існуючої моделі / Theme 10. Model-Based Reinforcement Learning. Model-Based Policy Learning. <i>Individual work:</i> To develop main methods of Model-Based Reinforcement Learning.	2	2	14
<i>Контрольна робота 3</i>		2		
ВСЬОГО		28	14	136

Загальний обсяг 180 *годин*, в тому числі:

Лекцій – 28 год.

Лабораторних – 14 год.

Консультації – 2 год.

Самостійна робота – 136 год.

9. РЕКОМЕНДОВАНІ ДЖЕРЕЛА

Основні / Main:

1. Richard S. Sutton, Andrew G. Barto. Reinforcement Learning: An Introduction, MIT Press, Cambridge, MA, 2018
2. Ian Goodfellow. Deep Learning, MIT Press, 2017
3. Sebastian Raschka, Vahid Mirjalili. Python Machine Learning, 2019
4. Artur Sahakjan. A Review of Recent Advancements in Deep Reinforcement Learning, 2018
5. Richard Szeliski. Computer Vision: Algorithms and Applications, Springer, 2010
6. Дэвид А. Форсайт, Жан Понс. Компьютерное зрение. Современный подход, 2004
7. Линда Шапиро, Джордж Стокман. Компьютерное зрение. Лаборатория знаний. 2013
8. Stephen Marsland. Machine Learning: An Algorithmic Perspective, 452 p., 2015.
9. Christopher M Bishop. Pattern recognition. Machine Learning, 128 p., 2006.
10. Ethem Alpaydin. Introduction To Machine Learning, 584 p., 2009.
11. Tom M. Mitchell. Machine Learning [<http://www.cs.cmu.edu/~tom/mlbook.html>]
12. Yaser S. Abu-Mostafa. Learning from data, 215 p., 2017
13. Alex Smola. Introduction to Machine Learning, 234 p., 2008.
14. Люгер Дж.Ф. Искусственный интеллект: стратегии и методы решения сложных проблем. – М.: Вильямс, 2005. – 864 с.
15. Trevor Hastie, Robert Tibshirani, Jerome Friedman. The Elements of Statistical Learning: Data Mining, Inference and Prediction, 764 p., 2008.

Додаткові / Additional:

16. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И.Д. Рудинского. – М.: Горячая линия-Телеком, 2004. – 452 с.
17. Рассел С., Норвиг П. Искусственный интеллект. Современный подход . – М.: Вильямс, 2006. – 1408 с.

18. Simon J.D. Prince. Computer Vision: Models, Learning, and Inference.- Cambridge: Cambridge University Press. - 567 p.- 2017.[<http://www.cambridge.org/>].
19. Andrej Karpathy. The unreasonable effectiveness of recurrent neural networks. <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>, 2015. Online; accessed 11-December.
20. Andrew Moore. Statistical Data Mining Tutorials [<http://www.autonlab.org/tutorials/>]
Pierre Baldi and Søren Brunak. Chapter 6. Neural networks: applications. In Bioinformatics: The Machine Learning Approach. MIT press, 2001.