



## Навчання з підкріпленням

### Робоча програма навчальної дисципліни (Силабус)

#### Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Третій (освітньо-науковий)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>124 Системний аналіз</i>
Освітня програма	<i>Системний аналіз</i>
Статус дисципліни	<i>Вибіркова</i>
Форма навчання	<i>очна(денна)/очна(вечірня)/заочна/дистанційна/змішана</i>
Рік підготовки, семестр	<i>2 курс, осінній / весняний семестр</i>
Обсяг дисципліни	<i>5 кредитів ЄКТС</i>
Семестровий контроль/ контрольні заходи	<i>залік</i>
Розклад занять	
Мова викладання	<i>Українська/Англійська</i>
Інформація про керівника курсу / викладачів	Лектор: <i>Лектор: доктор фізико-математичних наук, професор, Касьянов Павло</i> <i>Олегович, kasyanov.pavlo@lil.kpi.ua <a href="https://www.facebook.com/pkasyanov">https://www.facebook.com/pkasyanov</a></i> <i><a href="https://www.linkedin.com/in/pavlokasyanov/">https://www.linkedin.com/in/pavlokasyanov/</a></i> <i><a href="https://www.researchgate.net/profile/Pavlo_Kasyanov">https://www.researchgate.net/profile/Pavlo_Kasyanov</a></i> Практичні / Семінарські: <i>доктор фізико-математичних наук, професор, Касьянов Павло</i> <i>Олегович, kasyanov.pavlo@lil.kpi.ua <a href="https://www.facebook.com/pkasyanov">https://www.facebook.com/pkasyanov</a></i> <i><a href="https://www.linkedin.com/in/pavlokasyanov/">https://www.linkedin.com/in/pavlokasyanov/</a></i> <i><a href="https://www.researchgate.net/profile/Pavlo_Kasyanov">https://www.researchgate.net/profile/Pavlo_Kasyanov</a></i>
Розміщення курсу	<a href="#">КА 1X</a>   <a href="#">Class Profile</a>   <a href="#">Piazza</a>

#### Програма навчальної дисципліни

##### 1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

- а. **Метою** кредитного модуля є формування у студентів системного наукового світогляду, загальнокультурного кругозору та компетентностей виявляти, ставити та вирішувати проблеми дослідницького характеру в сфері комп'ютерних наук, оцінювати та забезпечувати якість виконуваних досліджень. Зокрема, засвоїти як фундаментальні засади теорії покрокового прийняття оптимальних рішень (теорія марковських процесів прийняття рішень) та динамічного програмування, так і вміти застосовувати одержані теоретичні знання для розв'язання прикладних, зокрема, задач прийняття оптимальних рішень в промисловості (технічна підтримка промислових систем, система експертизи промислової безпеки); робототехніці (автоматизоване прогнозування); бізнесі (маркетинг, керування запасами); інформатиці (усунення несправностей мереж, оптимізація запитів до серверів розподілених баз даних);

*державній безпеці та військових науках (пошук рухомих цілей, ідентифікація цілей, розподіл зброї); охороні здоров'я (медична діагностика, розробка протоколів лікування).*

### **b. Предмет вивчення.**

*Задачі та класи методів навчання з підкріпленням рівно як і область знань, що їх включає задач покрокового прийняття оптимальних рішень із частковими спостереженнями*

### **c. Основні завдання кредитного модуля.**

*Згідно з вимогами програми навчальної дисципліни аспіранти після засвоєння кредитного модуля мають продемонструвати такі результати навчання:*

#### **знання:**

*методів та засобів навчання з підкріпленням.*

#### **уміння:**

*розв'язувати реальні проблеми за допомогою методів та алгоритмів навчання з підкріпленням. Зокрема, формалізувати задачу покрокового прийняття оптимальних рішень як частково спостережуваного марковського процесу прийняття рішень з можливо невідомими перехідними ймовірностями та винагородами, застосовувати сучасні алгоритми наближеного розв'язання таких задач, уміння користуватись актуальними інформаційними технологіями та створювати власні програмні продукти для розв'язання реальних задач прийняття оптимальних рішень в промисловості (технічна підтримка промислових систем, система експертизи промислової безпеки); робототехніці (автоматизоване прогнозування); бізнесі (маркетинг, керування запасами); інформатиці (усунення несправностей мереж, оптимізація запитів до серверів розподілених баз даних); державній безпеці та військових науках (пошук рухомих цілей, ідентифікація цілей, розподіл зброї); охороні здоров'я (медична діагностика, розробка протоколів лікування).*

#### **досвід:**

*створення дослідницької лабораторії навчання з підкріпленням (парадигма організованої співпраці за досвідом провідних національних лабораторій США), де роль кожного члена команди полягає в тому, щоб спеціалізуватися на певній задачі, щоб стати найкращим у ній, маючи при цьому цілісний погляд на весь процес.*

### **2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)**

*Базовий рівень володіння англійською мовою, вища математика, програмування.*

### **3. Зміст навчальної дисципліни**

*Кредитний модуль 1.*

- 1. Марковські процеси прийняття рішень*
- 2. Q-навчання для табулярних задач*
- 3. Q-навчання на основі апроксимацій для задач глибокого навчання з підкріпленням*
- 4. Наближене динамічне програмування*
- 5. Policy gradient methods*
- 6. Методи актор-критик*
- 7. Наближене глибоке навчання з підкріпленням*

**Рекомендована тематика практичних (семінарських) занять**

Метою проведення практичних занять є закріплення знань, надбаних на лекційних заняттях, здобуття умінь розв'язувати реальні проблеми покрокового прийняття оптимальних рішень за допомогою методів та засобів навчання з підкріпленням.

1. Вступне заняття. Завантаження корисних ресурсів.
2. Задача багаторукового бандита
3. Марковські процеси прийняття рішень. Методи динамічного програмування. Рівняння оптимальності Белмана.
4. Методи Монте-Карло
5. Метод часових різниць
6. Алгоритми Sarsa, Expected Sarsa, Dyna-Q, Q-навчання,
7. Tile coding, бібліотеки Keras та TensorFlow для задач глибокого навчання з підкріпленням,
8. Градієнтні та напівградієнтні методи,
9. Гаусів метод актор-критик

#### 4. Навчальні матеріали та ресурси

Всі необхідні матеріали містяться на платформі Piazza

[КА 1X | Class Profile | Piazza](#)

Базова література:

1. [Reinforcement Learning, second edition: An Introduction \(Adaptive Computation and Machine Learning series\): Sutton, Richard S., Barto, Andrew G.: 9780262039246: Amazon.com: Books](#)
2. [\(PDF\) Algorithms for reinforcement learning | Csaba Szepesvari - Academia.edu](#)
3. [Markov Decision Processes | Wiley Series in Probability and Statistics](#)
4. [ELAKPI: Системний аналіз стохастично розподілених систем](#)
5. <https://www.coursera.org/specializations/reinforcement-learning>
6. [CSCPHD | Class Profile | Piazza](#)

Додаткова література:

<https://drive.google.com/drive/folders/1V9jAShWpccLvByv5S1DuOzo6GVvzd4LV>

5. Методика опанування навчальної дисципліни (освітнього компонента)

5.1. Лекційні заняття

№ з/п	Назва теми лекції та перелік основних питань (перелік дидактичних засобів, посилання на літературу та завдання на СРС)
1	8. Марковські процеси прийняття рішень. [1-6] (6 год)
2	9. Q-навчання для табулярних задач. [1-6] (2 год)
3	10. Q-навчання на основі апроксимацій для задач глибокого навчання з підкріпленням. [1-6] (2 год)
4	11. Наближене динамічне програмування. [1-6] (2 год)
5	12. Policy gradient methods. [1-6] (2 год)
6	13. Методи актор-критик. [1-6] (2 год)
7	14. Наближене глибоке навчання з підкріпленням. [1-6] (2 год)

5.2. Практичні заняття

Метою проведення практичних занять є закріплення знань, надбаних на лекційних заняттях, здобуття умінь розв'язувати реальні проблеми за допомогою фінансових аналітичних симуляцій

№ з/п	Назва теми заняття (перелік дидактичного забезпечення, посилання на літературу та завдання на СРС)
1	Вступне заняття. Завантаження корисних ресурсів. [4-6] (2 год)
2	Задача багаторукого бандита. [4-6] (2 год)
3	Марковські процеси прийняття рішень. Методи динамічного програмування. Рівняння оптимальності Белмана. [4-6] (2 год)
4	Методи Монте-Карло [4-6] (2 год)
5	Метод часових різниць. [4-6] (2 год)
6	Алгоритми Sarsa, Expected Sarsa, Dyna-Q, Q-навчання. [4-6] (2 год)
7	Tile coding, бібліотеки Keras та TensorFlow для задач глибокого навчання з підкріпленням. [4-6] (2 год)
8	Градiєнтні та напівградiєнтні методи. [4-6] (2 год)
9	Гаусів метод актор-критик [4-6] (2 год)

## 6. Самостійна робота студента/аспіранта

Самостійна робота студентів полягає в опрацюванні матеріалів та виконанні завдань на платформі дистанційного навчання Piazza

[КА 1X | Class Profile | Piazza](#); підготовці до залікової роботи.

### Політика та контроль

#### 7. Політика навчальної дисципліни (освітнього компонента)

Вимагається належне виконання усіх завдань на платформі дистанційного навчання Piazza

[КА 1X | Class Profile | Piazza](#) згідно вимог та індивідуальної стратегії, яку визначає аспірант самостійно або, за необхідності, під науковим керівництвом викладача / наукового керівника.

#### 8. Види контролю та рейтингова система оцінювання результатів навчання (PCO)

**Поточний контроль:** кожен студент визначає стратегію виконання завдань (самостійно або, за необхідності, під науковим керівництвом викладача / наукового керівника), ставлячи на меті одержати в кінці семестру 100 балів.

Види контролю :

1) дві відповіді (кожного студента в середньому) на практичних заняттях (за умови, що на одному занятті в середньому опитуються 8 студентів при максимальній чисельності групи 30 осіб);

2) виконання однієї контрольної роботи (в дистанційному режимі – тестів та завдань).

### СИСТЕМА РЕЙТИНГОВИХ БАЛІВ

#### 1. Практичне заняття

Максимальна кількість балів на всіх практичних заняттях дорівнює 20 балів X 2 =40 балів.

Критерії оцінювання:

0–8 балів – задача не розв’язана, при цьому студент володіє певними теоретичними відомостями щодо теми практичного заняття;

9–14 балів – задача розв’язана не до кінця або розв’язок містить грубі технічні недоліки, при цьому студент вільно володіє теоретичними відомостями щодо теми практичного заняття;

15–20 балів – задача в цілому розв’язана, при цьому студент вільно володіє теоретичними відомостями щодо теми практичного заняття.

#### 2. Модульний контроль.

Максимальні кількість балів за контрольну роботу (проєкт) дорівнює 60 балів.

Критерії оцінювання:

0–20 балів – задача в цілому не розв’язана або розв’язок містить грубі технічні недоліки, відповіді на теоретичне питання немає;

21–50 балів – задача в цілому розв’язана, теоретичне питання розкрито;

51–60 балів – задача розв’язана, відповідь на теоретичне питання є вичерпною.

Штрафні та заохочувальні бали за:

- виконання завдань із удосконалення дидактичних матеріалів з дисциплін надається від 15 до 30 заохочувальних балів.

За результатами навчальної роботи за перші 7 тижнів «ідеальний студент» має набрати 20 балів. На першій атестації (8-й тиждень) студент отримує «зараховано», якщо його поточний рейтинг не менш ніж 10 балів. За результатами 13 тижнів навчання «ідеальний студент» має

набрати 40 балів. На другій атестації (14-й тиждень) студент отримує «зараховано», якщо його поточний рейтинг не менш ніж 20 балів.

Максимальна сума балів складає 100. Необхідною умовою допуску до заліку є позитивна оцінка з контрольної роботи. Для отримання заліку з кредитного модуля "автоматом" потрібно мати рейтинг не менш ніж 60 балів, а також зараховану контрольну роботу (більш ніж 30 балів). Студенти, які наприкінці семестру мають рейтинг менше 60 балів, а також ті, хто хоче підвищити оцінку у системі ECTS, виконують залікову роботу. При цьому до балів з контрольної роботи додаються бали за залікову роботу і ця рейтингова оцінка є остаточною. Контрольне завдання цієї роботи складається з двох питань робочої програми з переліку, що наданий у методичних рекомендаціях до засвоєння кредитного модуля. Додаткове питання з тем практичних занять отримують студенти, які не приймали участі у роботі певного практичного заняття. Незадовільна відповідь з додаткового питання знижує загальну оцінку на 4 бали.

Кожне питання оцінюється з 20 балів відповідно до системи оцінювання:

- “відмінно”, повна відповідь (не менше 90% потрібної інформації) – 20...18 балів;
- “добре”, достатньо повна відповідь (не менше 75% потрібної інформації, або незначні неточності) – 17...14 бал;
- “задовільно”, неповна відповідь (не менше 60% потрібної інформації та деякі помилки) – 13...11 балів;
- “незадовільно”, незадовільна відповідь – 0 балів.

Сума балів: за кожне з двох запитань залікової роботи та контрольну роботу переводиться до залікової оцінки згідно з таблицею.

*Бали Оцінка*

100...95 Відмінно

94...85 Дуже добре

84...75 Добре

74...65 Задовільно

64...60 Достатньо

Менше 60 Незадовільно

ДКР не зараховано Не допущено

## **9. Додаткова інформація з дисципліни (освітнього компонента)**

*Всі необхідні матеріали містяться на платформі навчання Piazza*

[КА 1X | Class Profile | Piazza](#)

### **Робочу програму навчальної дисципліни (силабус):**

**Складено** директор ІПСА, д.ф.-м.н., професор, Касьянов Павло Олегович

акад. НАН України, д.т.н., проф. Михайло Захарович Згуровський

**Ухвалено** кафедрою математичних методів системного аналізу (протокол № 11 від 08.07.2022 р.)

**Погоджено** Методичною комісією факультету (протокол № 8 від 17.06.2022 р.)