



Навчання з підкріпленням

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Перший (освітній)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>122 Комп'ютерні науки, 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>4 курс, весняний семестр</i>
Обсяг дисципліни	<i>3,5 кредити ЕКТС</i>
Семестровий контроль/ контрольні заходи	<i>залік</i>
Розклад занять	Асинхронно Zoom Meeting ID: 918 3662 5010 Passcode: 651168
Мова викладання	<i>Українська/Англійська</i>
Інформація про керівника курсу / викладачів	Лектор: <i>доктор фізико-математичних наук, професор, член-кореспондент НАН України Касьянов Павло Олегович, kasyanov.pavlo@iit.kpi.ua https://www.facebook.com/pkasyanov https://www.linkedin.com/in/pavlokasyanov/ https://www.researchgate.net/profile/Pavlo_Kasyanov</i> Практичні / Семінарські: <i>аспірант, Senior Research Engineer у SQUAD Андрій Миколайович Титаренко, доктор фізико-математичних наук, професор, член-кореспондент НАН України Касьянов Павло Олегович, kasyanov.pavlo@iit.kpi.ua https://www.facebook.com/pkasyanov https://www.linkedin.com/in/pavlokasyanov/ https://www.researchgate.net/profile/Pavlo_Kasyanov</i>
Розміщення курсу	https://piazza.com/national_technical_university_of_ukraine_igor_sikorsky_kyiv_polytechnic_institute/spring2024/ka0x/home

Програма навчальної дисципліни

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

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). Студенти мають оволодіти наступними **компетентностями**:

загальними - ЗК 1 Здатність застосовувати знання в практичних ситуаціях; ЗК 3 Здатність абстрактно мислити, застосовувати методи аналізу і синтезу; ЗК 7 Здатність до пошуку, оброблення та аналізу інформації з різних джерел; ЗК 11 Здатність генерувати нові ідеї (креативність); ЗК 12 Здатність працювати в команді та автономно виконувати командні рішення;

фаховими – ФК 1 Здатність використовувати системний аналіз в якості сучасної міждисциплінарної методології, заснованої на прикладах математичних методів та сучасних інформаційних технологіях, і орієнтована на вирішення задач аналізу і синтезу технічних, економічних, соціальних, екологічних та інших складних систем; ФК 6 Здатність до комп'ютерної реалізації математичних моделей реальних систем і процесів; проектувати, застосовувати і супроводжувати програмні засоби моделювання, прийняття рішень, оптимізації, обробки інформації, інтелектуального аналізу даних; ФК 7 Здатність використовувати сучасні інформаційні технології для комп'ютерної реалізації математичних моделей та прогнозування поведінки конкретних систем, а саме: об'єктно-орієнтований підхід при проектуванні складних систем різної природи, прикладні математичні пакети, застосування баз даних і знань; ФК 10 Здатність розробляти експериментальні та спостережувальні дослідження і аналізувати дані, отримані в них.

По завершенню курсу студенти мають **набути такі програмні результати навчання**: ПРН 9 Вміти створювати ефективні алгоритми для обчислювальних задач системного аналізу та систем підтримки прийняття рішень; ПРН 12 Застосовувати методи і засоби роботи з даними і знаннями, методи математичного, логіко-семантичного, об'єктного та імітаційного моделювання, технології системного і статичного аналізу; ПРН 14 Розуміти і застосовувати на практиці методи статичного моделювання і прогнозування, оцінювати вихідні дані; ПРН 17 Зберігати та примножувати досягнення і цінності суспільства на основі розуміння місця предметної області у загальній системі знань.

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

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, ФП, ООП.

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

https://piazza.com/national_technical_university_of_ukraine_igor_sikorsky_kyiv_polytechnic_institute/spring2024/ka0x/home

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. <https://piazza.com/national technical university of ukraine igor sikorsky kyiv polytechnic institute/spring2024/ka0x/home>

Додаткова література:

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

Навчальний контент

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	Марковські процеси прийняття рішень. [1-6] (6 год)
2	Q-навчання для табулярних задач. [1-6] (2 год)
3	Q-навчання на основі апроксимацій для задач глибокого навчання з підкріпленням. [1-6] (2 год)
4	Наближене динамічне програмування. [1-6] (2 год)
5	Policy gradient methods. [1-6] (2 год)
6	Методи актор-критик. [1-6] (2 год)
7	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 багаторукового 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, Дупа-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	Градентні та напівградентні методи. [4-6] (2 год)

6. Самостійна робота студента/аспіранта

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

https://piazza.com/national_technical_university_of_ukraine_igor_sikorsky_kyiv_polytechnic_institute/spring2024/ka0x/home

Політика та контроль**7. Політика навчальної дисципліни (освітнього компонента)**

Вимагається належне виконання усіх завдань на платформі дистанційного навчання Piazza

https://piazza.com/national_technical_university_of_ukraine_igor_sikorsky_kyiv_polytechnic_institute/spring2024/ka0x/home згідно вимог та індивідуальної стратегії, яку визначає слухач 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 :

а) **модельні задачі:** 5 домашніх завдань на платформі дистанційного навчання https://piazza.com/national_technical_university_of_ukraine_igor_sikorsky_kyiv_polytechnic_institute/spring2024/ka0x/home (є рекомендовані обмеження за крайніми термінами виконання), кожне з яких може бути оцінене максимум у 20 балів;

б) **практичні проблеми:** виконання та відеопрезентація практичного проекту у складі команди (1-5 студентів). Максимум 85 балів (кількість балів за проект дорівнює сумі експертної оцінки (максимум 40 балів) та "кількості позитивних реакцій на відео проекту (напр., [YouTube](#)) та його код (напр., [GitHub: Where the world builds software](#) - [GitHub](#))".

в) **навчально-методичні проблеми:** запис відеолекції на книгу [[Reinforcement Learning, second edition: An Introduction \(Adaptive Computation and Machine Learning series\): Sutton, Richard S., Barto, Andrew G.: 9780262039246: Amazon.com: Books](#)] у складі команди (1-5 студентів), кожна із яких може бути оцінена у 75 балів шляхом голосування (позитивні реакції, коментарі тощо) інших членів колективу (напр., [YouTube](#)).

г) **Заохочувальні бали за виконання завдань із удосконалення дидактичних матеріалів з дисциплін** надається від 20 до 40 заохочувальних балів.

Календарний контроль: проводиться як моніторинг поточного стану виконання вимог syllabus (пропорційно кількості робочих тижнів за семестр).

Семестровий контроль: залік диференційований (сума балів за семестр, довиконання видів робіт пп. а,б,в,г)

Умови допуску до семестрового контролю: семестровий рейтинг бажано мати не менше 20 балів.

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

Кількість балів	Оцінка
-----------------	--------

<i>100-95</i>	<i>Відмінно</i>
<i>94-85</i>	<i>Дуже добре</i>
<i>84-75</i>	<i>Добре</i>
<i>74-65</i>	<i>Задовільно</i>
<i>64-60</i>	<i>Достатньо</i>
<i>Менше 60</i>	<i>Незадовільно</i>
<i>Не виконані умови допуску</i>	<i>Не допущено</i>

9. Додаткова інформація з дисципліни (освітнього компонента)

Всі необхідні матеріали містяться на платформі навчання Piazza

https://piazza.com/national_technical_university_of_ukraine_igor_sikorsky_kyiv_polytechnic_institute/spring2024/ka0x/home

Робочу програму навчальної дисципліни (силабус):

Складено директор ІПСА, д.ф.-м.н., професор, член-кореспондент НАН України Касьянов Павло Олегович
асп., провідний інженер-дослідник SQUAD, Андрій Миколайович Титаренко

Ухвалено кафедрою математичних методів системного аналізу (протокол № 13 від 05.06.2024 р.)

Погоджено Методичною комісією факультету (протокол № 10 від 24.06.2024 р.)