



Навчання нейронних мереж з підкріпленням

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	Другий (магістерський)
Галузь знань	12 Інформаційні технології
Спеціальність	123 Комп'ютерна інженерія
Освітня програма	Комп'ютерні системи та мережі
Статус дисципліни	Нормативна
Форма навчання	Очна (денна)
Рік підготовки, семестр	2 курс, осінній семестр
Обсяг дисципліни	6 кредитів ЄКТС
Семестровий контроль/ контрольні заходи	Екзамен
Розклад занять	Лекцій 36 годин, Лабораторні роботи -18 годин, Самостійна робота студентів – 126 годин
Мова викладання	Українська
Інформація про керівника курсу / викладачів	Лектор: асистент кафедри обчислювальної техніки, Кочура Юрій Петрович, iuriy.kochura@gmail.com Лабораторні: асистент кафедри обчислювальної техніки, Кочура Юрій Петрович, iuriy.kochura@gmail.com
Розміщення курсу	https://github.com/YKochura/rl-kpi

Програма навчальної дисципліни

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

Метою навчальної дисципліни є формування у студентів здатностей (компетенцій) розробки та використання технологій нейронних мереж з підкріпленням. За результатами вивчення дисципліни студент має бути здатним вирішувати професійні завдання та володіти такими компетенціями згідно освітньої програми: ФК12, ФК16, ПРН14, ПРН16.

Навчання з підкріпленням (англ. *reinforcement learning, RL*) — це галузь машинного навчання, а також формалізм для автоматизованого прийняття рішень на основі взаємодій. За останні 5 років глибоке навчання з підкріпленням (*deep RL*) стало одним з найінтенсивніших напрямків досліджень у сфері штучного інтелекту. Сьогодні *deep RL* дозволяє досягати надлюдської продуктивності в ряді завдань: відео ігри, покер, а також у настільних іграх, включаючи го та шахи.

Цей курс познайомить з сімейством статистичних алгоритмів, які вивчають оптимальну стратегію, метою якої є максимізація загальної винагороди, отриманої агентом при взаємодії з навколишнім середовищем.

Для проходження цього курсу потрібно володіти наступними навичками:

1. Базовий рівень володіння англійською мовою не нижче А2.
2. Базові знання з лінійної алгебри та теорії ймовірностей.

3. Досвід тренування глибоких мереж (ініціалізація, оптимізація, регуляризація, вибір методу та метрик для оцінки).

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

Дисципліни, що передують: Системи штучного інтелекту, Хмарні обчислення, Програмне забезпечення комп'ютерних систем, Мережні технології, Технології Dig Data/ Дана дисципліна є підготовкою до виконання дисертаційних магістерських робіт.

3. Зміст навчальної дисципліни

Розділ 1. Основні теоретичні відомості

Огляд основ машинного навчання (МН). Класичне програмування vs машинне навчання. Типи МН за характером навчальних даних. Моделі біологічного та штучного нейронів. Функції активації. Важливість вибору правильних ознак. Перенавчання vs недонавчання моделі. Інтерпретація та компоненти моделі. Джерела помилок. Баланс зсуву та розкиду. Важливість прийняття правильних рішень в умовах невизначеності. Визначення навчання з підкріпленням (RL). Напрямки застосування RL. Причини використання RL. Глибоке навчання з підкріпленням vs навчання з підкріпленням.

Розділ 2. Вступ до навчання з підкріпленням

Цикл взаємодії агента з середовищем. Математичний запис загальної винагороди. Гіпотеза винагороди. Функція цінності. Цінність дій агента: Q-функція. Історія агента. Поняття повністю оглядового та частково оглядового середовища. Поняття стратегії. Види стратегій. Рівняння Беллмана. Класифікація агентів. Підзадачі RL. Зв'язок передбачення та контролю. Навчання та планування.

Розділ 3. Марковські процеси прийняття рішень

Марковські процеси прийняття рішень (Markov Decision Processes, MDPs). Дослідження Марковських процесів, включаючи процеси винагороди та процеси прийняття рішень.

Розділ 4. Динамічне програмування

Планування за допомогою динамічного програмування. Знайомство з оцінкою стратегії та функції цінності.

Розділ 5. Безмодельне передбачення

Безмодельне передбачення. Знайомство з методами навчання Монте-Карло та навчанням за часовими відмінностями.

Розділ 6. Методи апроксимації

Методи апроксимації функції цінності. Інкрементні методи та пакетні методи апроксимації функції цінності.

Розділ 7. Градієнтні методи стратегії

Градієнтні методи стратегії. Інтеграція навчання та планування.

4. Навчальні матеріали та ресурси

Базова:

- 1 Haykin Simon Neural Networks: A Comprehensive Foundation/Prentice Hall 2005- 826 p.
https://cdn.preterhuman.net/texts/science_and_technology/artificial_intelligence/Neural%20Networks%20-%20A%20Comprehensive%20Foundation%20-%20Simon%20Haykin.pdf
- 2 Тимошук П.В. Штучні нейронні мережі. Львівської політехніки, 2017 р.- 444 с.
- 3 Навчання нейронних мереж з підкріпленням. Лабораторній практикум Навчальний посібник для здобувачів ступеня магістра за спеціальністю 123 «Комп'ютерні системи та мережі» / О.Стіренко., Ю.Кочура - К.: КПІ ім. Ігоря Сікорського, 2022. - 26 с. Електронний ресурс. Гриф надано Методичною радою ФІОТ КПІ ім. Ігоря Сікорського (протокол № 8 від 18.04.2022 р.)

Додаткова

- 4 Основи нейрокомп'ютерингу : навчально-методичний посібник до практичних занять / В. Д. Дмитрієнко, О. Ю. Заковоротний, В. І. Носков, М. В. Мезенцев. - Х.: НТМТ, 2014. - 140 с.
- 5 Троцько В.В. Методи штучного інтелекту: навчально-методичний і практичний посібник. - Київ: Університет економіки та права «КРОК», 2020 - 86 с.
- 6 Дорогий Я. Ю. Розробка архітектури системи для моделювання згорточних нейронних мереж / Дорогий Я. Ю., Цуркан В. В., Хренов О. І. // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка : збірник наукових праць. – 2013. – Вип. 59. – С. 73–77. – Бібліогр.: 16 назв Електронний ресурс . <https://ela.kpi.ua/handle/123456789/15562>
- 7 Глибовець М.М., Гулаєва Н.М. Еволюційні алгоритми: підручник. – Київ: НаУКМА, 2013. - 828 с.
- 8 Субботін С.О. Нейронні мережі: навч. посібник / С.О. Субботін, А.О. Олійник; за ред. С.О. Субботіна. - Запоріжжя : ЗНТУ, 2014. - 132 с.
- 9 Субботін С. О. Нейронні мережі : теорія та практика: навч. посіб. / С. О. Субботін. - Житомир : Вид. О. О. Євенок, 2020. - 184 с.
- 10 Russell, S., & Norvig, P. (3d or 4th Edition). Artificial intelligence: a modern approach.
- 11 Sutton, R. S., & Barto, A. G. (2020). Reinforcement learning: An introduction. MIT press.

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

5.1. Лекційні заняття

№ лекції	Назва лекції та перелік основних питань	Кількість ауд. годин
1	<p>Основні теоретичні відомості.</p> <p>Огляд основ машинного навчання (МН). Класичне програмування vs машинне навчання. Типи МН за характером навчальних даних. Моделі біологічного та штучного нейронів. Функції активації. Важливість вибору правильних ознак. Перенавчання vs недонавчання моделі. Інтерпретація та компоненти моделі. Джерела помилок. Баланс зсуву та розкиду. Важливість прийняття правильних рішень в умовах невизначеності. Визначення навчання з підкріпленням (RL). Напрямки застосування RL. Причини використання RL. Глибинне навчання з підкріпленням vs навчання з підкріпленням.</p>	4
2	<p>Вступ до навчання з підкріпленням</p> <p>Цикл взаємодії агента з середовищем. Математичний запис загальної винагороди. Гіпотеза винагороди. Функція цінності. Цінність дій агента: Q-функція. Історія агента. Поняття повністю оглядового та частково оглядового середовища. Поняття стратегії. Види стратегій. Рівняння Беллмана. Класифікація агентів. Підзадачі RL. Зв'язок передбачення та контролю. Навчання та планування.</p>	4
3	<p>Марковські процеси прийняття рішень</p> <p>Марковські процеси прийняття рішень (Markov Decision Processes, MDPs). Дослідження Марковських процесів, включаючи процеси винагороди та процеси прийняття рішень.</p>	6
4	<p>Динамічне програмування</p> <p>Планування за допомогою динамічного програмування. Знайомство з оцінкою стратегії та функції цінності.</p>	4
5	<p>Безмодельне передбачення</p> <p>Безмодельне передбачення. Знайомство з методами навчання Монте-Карло та навчанням за часовими відмінностями.</p>	6
6	<p>Методи апроксимації</p> <p>Методи апроксимації функції цінності. Інкрементні методи та пакетні методи апроксимації функції цінності.</p>	6
7	<p>Градiєнтні методи стратегії</p> <p>Градiєнтні методи стратегії. Інтеграція навчання та планування.</p>	6
	Разом:	36

5.2. Лабораторні/практичні роботи

№ з/п	Назва лабораторної/практичної роботи	Кількість ауд. годин
1	Перевернутий маятник (Cartpole), понг (Pong) та MountainCar Вирішуючи це завдання, студенти познайомляться з тим як можна створити середовище різної складності та навчити агента дотримуватись в ньому певної стратегії.	2
2	N-рукий бандит Вирішення задачі про N-рукого бандита. Пошук агентом компромісу між вивчення поточної ситуації у середовищі та використанням раніше набутих знань.	4
3	Глибинне Q-навчання Передбачення найкращих станів та дій агента.	4
4	Проектування системи Закріплення та застосування на практиці отриманих у рамках цієї дисципліни теоретичних знань для вирішення прикладної проблеми.	8
	Разом:	18

6. Самостійна робота студента/аспіранта

6.1. Теми, які виносяться на самостійне опрацювання.

№ з/п	Назва теми, що виносяться на самостійне опрацювання	Кількість годин
1	Вирішення проблем методами актор-критик	10
2	Мультиагентне навчання з підкріпленням	10

3	Інтерпретоване навчання з підкріпленням: увага і реляційні моделі	18
4	Підготовку відповідно до теми лекції	72
5	Підготовку до лабораторної роботи	8
6	Підготовка до екзамену	8
	Разом:	126

Перед кожним заняттям студенти здійснюють підготовку відповідно до теми лекції або лабораторної роботи не менше двох годин. Підготовка до екзамену має складати не менше 8 годин.

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Ви можете обговорювати завдання практичних робіт у групах. Однак, кожен студент/студентка повинен/повинна підготувати розв'язки завдань самостійно. Під час проходження цього курсу Ви зобов'язані дотримуватись Кодекс честі КПІ ім. Ігоря Сікорського та усі наступні правила:

- 1. Кожен з Вас повинен відправляти на перевірку власно виконану роботу. Використання чужих розв'язків або програмного коду і представлення їх за свої напрацювання є плагіатом та серйозним порушенням основних академічних стандартів.*
- 2. Ви не повинні ділитися своїми розв'язками з іншими студентами, а також просити інших ділитися своїми розв'язками з Вами.*
- 3. Якщо Ви отримували допомогу у вирішенні певного завдання, Ви маєте зазначити це у звіті, а саме: від кого та яку допомогу отримали.*

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

Види контролю з навчальної дисципліни включають:

Лабораторні роботи:

Заплановано самостійне виконання чотирьох лабораторних робіт. Темати лабораторних робіт узгоджені у часі та за змістом з темами лекцій. Виконання лабораторних робіт у повному обсязі дозволяє набутти практичних навичок використання нейронних мереж з підкріпленням.

Поточний контроль:

Передбачено проведення опитування

Іспит:

Проводиться для об'єктивного визначення рівня знань, умінь та практичних навичок,

отриманих за семестр

Семестровий рейтинг студента складається з балів, які він отримує за види робіт відповідно до таблиці 1.

Таблиця 1

Оцінювання окремих видів навчальної роботи студента (у балах)

Вид навчальної роботи	Всього за видом роботи
Виконання та захист лабораторної роботи № 1	15
Виконання та захист лабораторної роботи № 2	15
Виконання та захист лабораторної роботи № 3	15
Виконання та захист лабораторної роботи № 4	15
Rп	60
Іспит (Ri)	40
Усього за семестр (R = Rп+ Ri)	100

Індивідуальний поточний рейтинг студента (R_n) складається з балів, які він отримує за виконання лабораторних робіт. Протягом семестру студенти виконують 4 лабораторні роботи. Максимальна кількість балів за кожну лабораторну роботу - 15. Бали нараховуються за:

- Теоретична складова - 8 балів,
- Практична складова - 7 балів

Максимальний можливий бал за лабораторну роботу - 15 балів.

Максимальна кількість балів за всі лабораторні роботи $15 \times 4 = 60$ балів.

Розрахунок розміру шкали (R) рейтингу.

Сума вагових балів контрольних заходів протягом семестру становить:

$$R = R_n + R_i,$$

де R_n - семестровий рейтинг студента (лабораторні роботи).

R_i - іспит.

Розмір рейтингової шкали для навчальної дисципліни становить:

$$R = R_n + R_i = 60 + 40 = 100 \text{ балів.}$$

Необхідною умовою допуску студента до іспиту є його індивідуальний семестровий рейтинг (R_n), не менший, ніж 59 балів, та відсутність заборгованості з лабораторних робіт. При невиконанні згаданих вимог студент до іспиту не допускається.

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою

Табл. 1 Відповідність рейтингових балів оцінкам за університетською шкалою

Кількість балів	Оцінка ECTS
95...100	Відмінно
85...94	Дуже добре
75...84	Добре
65...74	Задовільно

60...64	Достатньо
$R_c < 60$	Незадовільно
$R_c \leq 50$ або не виконані інші умови допуску до екзамену	Не допущено

Додаткова інформація з дисципліни (освітнього компонента)

теоретичні та практичні питання, що виносяться на семестровий контроль, відповідають переліку основних тем, що входять до програми вивчення дисципліни

Умова зарахування додаткових балів.

В рамках вивчення навчальної дисципліни допускається зарахування балів, одержаних в результаті

- дистанційних курсів на платформі "Coursera", за умови попереднього погодження програми даного курсу з викладачем та за умови отримання офіційного сертифікату,
- виконання індивідуального науково-дослідного проєкту за темою дисципліни та умови попереднього погодження напряму дослідження з викладачем та подачі статті про результати дослідження на міжнародну наукову конференцію рівне не нижче IEEE/ACM.

9.Додаткова інформація з дисципліни (освітнього компонента)

- можливість зарахування сертифікатів проходження дистанційних чи онлайн курсів за відповідною тематикою;

Робочу програму навчальної дисципліни (силабус):

Склав асистент кафедри обчислювальної техніки, Кочура Ю. П.

Ухвалено кафедрою обчислювальної техніки (протокол № 10 від 25.05.2022)

Погоджено Методичною комісією факультету інформатики та обчислювальної техніки (протокол № 10 від 09.06.2022)