



ПРОЄКТУВАННЯ СИСТЕМ МАШИННОГО НАВЧАННЯ

Робоча програма навчальної дисципліни (Силабус)

• Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Другий (магістерський)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>122 Комп'ютерні науки</i>
Освітня програма	<i>Комп'ютерні науки</i>
Статус дисципліни	<i>Вибіркова</i>
Форма навчання	<i>очна(денна)</i>
Рік підготовки, семестр	<i>1 курс, осінній/весняний семестр</i>
Обсяг дисципліни	<i>120 годин / 4 кредити ЄКТС (лекції – 36 год., лабораторні заняття – 18 год., СРС – 66 год.)</i>
Семестровий контроль/ контрольні заходи	<i>залік, МКР</i>
Розклад занять	<i>https://schedule.kpi.ua/ 1 год лекційних та 1 год практичних (семінарських) занять на тиждень</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	<i>Лекції та семінари проводять: канд. техн. наук, Кислий Роман Володимирович, kvrware@gmail.com</i>
Розміщення курсу	

• Програма навчальної дисципліни

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

Навчальна дисципліна "Проектування систем машинного навчання" спрямована на вивчення ключових принципів та практик у сфері MLOps, які об'єднують машинне навчання з дисциплінами DevOps. Курс охоплює вивчення методів розробки, тестування, впровадження та моніторингу моделей машинного навчання в продуктивному середовищі. Студенти дізнаються про автоматизацію процесів машинного навчання, управління життєвим циклом моделей, а також про важливість безперервної інтеграції та доставки в контексті машинного навчання. Дисципліна також включає вивчення інструментів та платформ для ефективного управління

процесами MLOps, а також методів забезпечення якості та надійності моделей машинного навчання. Курс призначений для поглибленого розуміння того, як інтегрувати машинне навчання в загальні бізнес-процеси та інфраструктуру, а також для розвитку навичок створення високоефективних, масштабованих та надійних систем машинного навчання.

1.2. Мета и завдання навчальної дисципліни

Мета:

Основною метою навчальної дисципліни "Проектування систем машинного навчання (MLOps)" є формування у студентів компетентностей, необхідних для ефективного проектування, розробки та управління системами машинного навчання. Це включає розуміння та застосування принципів MLOps для створення надійних, масштабованих та ефективних систем машинного навчання, які інтегровані з сучасними практиками DevOps.

Завдання:

- Ознайомлення з основами MLOps: Вивчення ключових концепцій та методологій, що лежать в основі MLOps, включаючи автоматизацію, моніторинг, тестування та розгортання моделей машинного навчання.
- Розуміння життєвого циклу моделей машинного навчання: Набуття знань про етапи розробки моделей машинного навчання, від збору даних до впровадження моделей у виробництво.
- Практичне застосування інструментів та платформ MLOps: Оволодіння навичками використання інструментів та платформ для управління процесами MLOps, включаючи контейнеризацію, оркестрацію та автоматизацію.
- Забезпечення якості та надійності моделей: Вивчення методів оцінки та підтримки якості моделей машинного навчання, включаючи тестування, валідацію та моніторинг.
- Інтеграція машинного навчання з бізнес-процесами: Розробка навичок інтеграції систем машинного навчання з загальними бізнес-процесами та інфраструктурою компанії.
- Розвиток критичного мислення та аналітичних навичок: Сприяння розвитку критичного мислення та аналітичних навичок, необхідних для вирішення складних проблем у сфері машинного навчання та MLOps.

Вивчення дисципліни вносить вклад в отримання наступних загальних та фахових компетенцій:

- Загальні компетенції:
 - (ЗК 1) Здатність до абстрактного мислення, аналізу та синтезу.
 - (ЗК 2) Здатність застосовувати знання у практичних ситуаціях.
 - (ЗК 5) Здатність вчитися й оволодівати сучасними знаннями.
- Фахові компетенції:
 - (ФК 4) Здатність збирати і аналізувати дані (включно з великими), для забезпечення якості прийняття проєктних рішень.

- (ФК 5) Здатність розробляти, описувати, аналізувати та оптимізувати архітектурні рішення інформаційних та комп'ютерних систем різного призначення.
- (ФК 11) Здатність ініціювати, планувати та реалізовувати процеси розробки інформаційних та комп'ютерних систем та програмного забезпечення, включно з його розробкою, аналізом, тестуванням, системною інтеграцією, впровадженням і супроводом.
- (ФК 14) Здатність проводити планування, аналіз та моніторинг ІТ проєктів, у тому числі стартап-проєктів, на всіх етапах життєвого циклу на основі міжнародних стандартів та відповідно до концепцій та підходів сталого розвитку і захисту інтелектуальної власності.
- (ФК 15) Здатність до проєктування та програмної реалізації методів комп'ютерної обробки надвеликих за обсягом даних в інформаційних середовищах різноманітного призначення, систем управління бізнес-процесами, сервіс-орієнтованих середовищ та систем високопродуктивних обчислень.
- (ФК 17) Здатність вибрати адекватні методи і технології обчислювального інтелекту та машинного навчання, включаючи методи глибокого навчання, еволюційного моделювання, генетичні алгоритми, та використовувати їх для вирішення задач прогнозування, керування, прийняття рішень, класифікації та інтелектуального аналізу даних в умовах невизначеності та неповної інформації.

Вивчення дисципліни допомагає отримати такі програмні результати навчання:

- (ПРН 2) Мати спеціалізовані уміння/навички розв'язання проблем комп'ютерних наук, необхідні для проведення досліджень та/або провадження інноваційної діяльності з метою розвитку нових знань та процедур.
- (ПРН 4) Управляти робочими процесами у сфері інформаційних технологій, які є складними, непередбачуваними та потребують нових стратегічних підходів.
- (ПРН 8) Розробляти математичні моделі та методи аналізу даних (включно з великими).
- (ПРН 26) Застосовувати технології обчислювального інтелекту та інтелектуального аналізу даних, зокрема, нейронні мережі, нечіткі нейронні мережі, нейронні мережі глибокого навчання, методи машинного навчання для проєктування та адаптації інтелектуальних систем прийняття рішень в різних предметних сферах.

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

Міждисциплінарні зв'язки: Навчальна дисципліна відноситься до циклу професійної підготовки. Дисципліні передують наступні курси бакалаврського рівня підготовки : „Основи сервіс-орієнтованих обчислень і архітектур”, „Методи та засоби штучного інтелекту”, “Чисельні методи.

Результати вивчення дисципліни можуть бути застосовані: в курсовій роботі з дисципліни «Методи та технології обчислювального інтелекту», а також на курсі «Нечітке моделювання та управління» та при виконанні магістерської дисертації.

Дисципліною забезпечуються наступні компоненти: Науково-дослідна практика

3. Зміст навчальної дисципліни

- Тема1. Огляд ML систем: відмінності продакшн систем від дослідницьких
 - Тема1.1.Вимоги до ML систем: надійність, масштабованість, адаптивність
 - Тема1.2.Ітеративний процес розробки: CRISP-DM, ключові метрики оптимізації
- Тема2.Огляд основних систем MLOps
 - Тема2.1.Життєвий цикл моделі
 - Тема2.2.Формати даних
 - Тема2.3.Управління даними та їх версіонування (dvc, git lfs etc.)
 - Тема2.4.Training Data: підготовка та очищення (основи ETL, Airflow)
- Тема3. Збір та зберігання даних
 - Тема3.1.Feature engineering (показати на прикладах)
 - Тема3.2.Feature stores (концепція, на прикладі Feast)
 - Тема3.3.Вибір моделі (baseline, обговорення чому SOTA - це не завжди добре)
- Тема4. Minikube, платформи для MLOps
 - Тема4.1.Основні концепції (локальне встановлення, запуск контейнерів з моделлю)
 - Тема4.2.ML pipelines та експерименти
 - Тема4.3.MLFlow (використання локально з minikube)
 - Тема4.4.SageMaker, Databricks
- Тема5. Моніторинг моделей та даних
 - Тема5.1.Моніторинг метрик моделей (time series, незалежних від часу)
 - Тема5.2.Моніторинг змін розподілу даних (Deequ, Seldon)
 - Тема5.3.Адаптація моделей
- Тема6. LLMOps
 - Тема6.1.Інференс LLM
 - Тема6.2.Квантизація великих моделей
 - Тема6.3. RLHF: основи та застосування
 - Тема6.4.Побудова сервісів з LLM за допомогою langchain (і векторної бази)
 - Тема6.5.Метрики для LLM
- Тема7. MLOps on edge (on device model serving)
 - Тема7.1.Запуск моделей в середовищах з обмеженими ресурсами
 - Тема7.2.Конвертація моделей
 - Тема7.3.Фреймворки: TF light, ONNX, llama.cpp

4. Навчальні матеріали та ресурси

Базова література

1. Stanford University - CS329s: Machine Learning Systems Design [Електрон. ресурс]. – Режим доступу: <https://stanford-cs329s.github.io/syllabus.html>
2. Coursera - Machine Learning Engineering for Production (MLOps) Specialization [Електрон. ресурс]. – Режим доступу: <https://www.coursera.org/specializations/machine-learning-engineering-for-production-mlops>
3. Weights & Biases Courses - Data Validation for Machine Learning [Електрон. ресурс]. – Режим доступу: <https://www.wandb.courses/courses/data-validation-for-machine-learning>
4. GitHub - MLflow: An Open Source Platform for the Machine Learning Lifecycle [Електрон. ресурс]. – Режим доступу: <https://github.com/mlflow/mlflow>
5. Minikube - Start Guide [Електрон. ресурс]. – Режим доступу: <https://minikube.sigs.k8s.io/docs/start/>

Додаткова література

(факультативно / ознайомлення)

6. Databricks Glossary - Lifecycle of Machine Learning Operations (MLOps) [Електрон. ресурс]. – Режим доступу: <https://www.databricks.com/glossary/mlmlops>
7. O'Reilly Media - Designing Machine Learning Systems [Електрон. ресурс]. – Режим доступу: <https://www.oreilly.com/library/view/designing-machine-learning/9781098107956/>

• Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

Назви розділів і тем

Тема 1. Проблеми сталого розвитку	
1	Лекція 1-2. Огляд ML систем: відмінності продакшн систем від дослідницьких
2	Лабораторне заняття 1. Створення діаграми життєвого циклу моделі Завдання на СРС: п.6 №2
3	Лекція 3-4. Огляд основних систем MLOps Завдання на СРС: п.6 №3
4	Лабораторне заняття 2. Застосування принципів MLOps для моделі
5	Лекція 5-6. Збір та зберігання даних Завдання на СРС: п.6 №5
6	<i>Лекційне заняття 7: Модульна контрольна робота (частина I)</i>
7	Лекція 8. Платформи для MLOps Завдання на СРС: п.6 №8

Назви розділів і тем

8	Лабораторне заняття 3-4. Використання Kubeflow для керування поведінкою моделі
9	Лекція 9-10. Моніторинг моделей та даних Завдання на СРС: п.6 №10
10	<i>Лекційне заняття 11: Модульна контрольна робота (частина II)</i>
11	Лекція 12. LLMOps. Завдання на СРС: п.6 №13
12	Лабораторне заняття 5-6. Запуск LLM і використання на своїх даних
13	Лекція 13-14. RLHF: основи та застосування Завдання на СРС: п.6 №15
14	Лекція 15-16. Квантизація великих моделей Завдання на СРС: п.6 №18
15	Лабораторне заняття 7-8. Використання принципу RLHF для тюнінгу своїх моделей Завдання на СРС: п.6 №19
16	Лекція 17-18. MLOps on edge (on device model serving) Завдання на СРС: п.6 №20
17	Лабораторне заняття 9. Використання принципу RLHF для тюнінгу своїх моделей

6. Самостійна робота студента

Дисципліна «Проектування систем машинного навчання» охоплює такі складники самостійної роботи студента з підготовки до аудиторних занять на теоретичні і практичні теми, а саме підготування до поточних опитувань, підготування до лабораторних робіт, підготування до виконання частин модульної контрольної роботи.

№ з/п	<i>Назва теми, що виноситься на самостійне опрацювання</i>	<i>Кількість годин</i>	<i>Література</i>
1	<i>Підготовка до лекції 1</i>	0.5	1-5
2	<i>Підготовка до лабораторної роботи 1 (в тому числі вивчення відповідної літератури та написання звіту)</i>	9	1, 2
3	<i>Підготовка до лекції 2</i>	0.5	3, 4, 5
4	<i>Підготовка до лекції 3</i>	0.5	2, 3
5	<i>Підготовка до частини I модульної контрольної роботи (за темами 1-2)</i>	2	1, 3
7	<i>Підготовка до лекції 4</i>	0.5	2-4
8	<i>Підготовка до лабораторної роботи 2 (в тому числі вивчення відповідної літератури та написання звіту)</i>	9	3, 6
9	<i>Підготовка до лекції 5</i>	0.5	2, 3, 4
10	<i>Підготовка до лекції 6</i>	0.5	1, 3, 2

№ з/п	Назва теми, що виноситься на самостійне опрацювання	Кількість годин	Література
12	Підготовка до лабораторної роботи 3 (в тому числі вивчення відповідної літератури та написання звіту)	10	1, 4, 5
13	Підготовка до лекції 7	0.5	1, 3
14	Підготовка до лекції 8	0.5	2, 4
15	Підготовка до лекції 9	0.5	1, 4
16	Підготовка до лекції 10	0.5	2, 3
18	Підготовка до лекції 11	0.5	3, 4, 5
19	Підготовка до лабораторної роботи 4 (в тому числі вивчення відповідної літератури та написання звіту)	10	3, 4
20	Підготовка до частини II модульної контрольної роботи (за темами 2-5)	2	1-6
21	Підготовка до лекції 12	0.5	3,4
22	Підготовка до лекції 13	0.5	2,4
23	Підготовка до лекції 14	0.5	1,5,3
24	Підготовка до лекції 15	0.5	2,4
25	Підготовка до лекції 16	0.5	5,6
26	Підготовка до лабораторної роботи 5 (в тому числі вивчення відповідної літератури та написання звіту)	10	7,8
27	Підготовка до лекції 17	0.5	3,7
28	Підготовка до лекції 18	0.5	1-8
	Підготовка до заліку	5	
	Разом	66	

● Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Застосовуються наступні вимоги до здобувачів:

- відвідування занять, як лекцій, так і лабораторних є обов'язковим;
- враховується активність на лекціях;
- лабораторні роботи повинні бути захищені персонально і в чітко визначені терміни;
- застосовується політика щодо академічної доброчесності, всі лабораторні роботи повинні бути виконані персонально з можливою перевіркою на плагіат
- додатково можуть застосовуватись інші вимоги, що не суперечать законодавству України та нормативним документам Університету.

Академічна доброчесність

- Політика та принципи академічної доброчесності визначені у розділі 3 Кодексу честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Детальніше:

<https://kpi.ua/code>

Норми етичної поведінки

- Норми етичної поведінки аспірантів і працівників визначені у розділі 2 Кодексу честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Детальніше: <https://kpi.ua/code>

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

Поточний контроль: експрес-опитування, опитування за темою заняття, МКР (складається з двох частин)

Календарний контроль: провадиться двічі на семестр як моніторинг поточного стану виконання вимог Силабусу.

Семестровий контроль: Залік

Умови допуску до семестрового контролю: зарахування усіх лабораторних робіт та семестровий рейтинг не менше ніж 50 балів.

Рейтинг студента з дисципліни складається з балів, що отримуються за наступне:

1. Виконання лабораторних робіт.
2. Виконання 2-х частин МКР.

Система рейтингових (вагових) балів та критерії оцінювання

1. Виконання лабораторних робіт

Ваговий бал - 10. Максимальна кількість балів на всіх лекціях дорівнює 10 бали *5 = 50 балів.

2. Виконання контрольних робіт.

Ваговий бал – 10 (максимально можливий) за кожну.

Штрафні бали за:

- нездачу лабораторної роботи вчасно - 1 бал за кожен тиждень.

Умови позитивної проміжної атестації

Календарна атестація студентів (на 8 та 14 тижнях семестру) проводиться викладачем за значенням поточного рейтингу студента на час атестації. Для отримання "зараховано" з першої проміжної атестації (8-ий тиждень) студент повинен мати не менше ніж 20 балів. Для отримання "зараховано" з другої проміжної атестації (14-ий тиждень) повинен мати не менше ніж 50 балів.

До заліку допускаються студенти, у яких зараховані всі лабораторні роботи, а також значення $R > 50$ (50% від R).

Розрахунок шкали (R) рейтингу:

Сума вагових балів контрольних заходів протягом семестру складає:

$R_s = 50 + 20 = 70$ балів.

Атестація проводиться за поточним рейтингом студента. Якщо поточний рейтинг складає не менше 50% від максимально можливого на цей момент, студент вважається атестованим.

Всі студенти повинні з'явитись на залік незалежно від набраного рейтингу. Оцінку на заліку студенти отримують згідно таблиці:

R	Оцінка а ECTS	Оцінка традиційна
95...100	A	Відмінно
85..94	B	Дуже добре
75...84	C	Добре
65...74	D	Задовільно
60...64	E	Достатньо
R < 60	FX	Незадовільно
R < 30	F	Недопущений

Якщо студент отримав за рейтингом $R < 30$ балів (менш ніж 30% від R) і по початку заліку виконав необхідну додаткову роботу (підвищив свій рейтинг), то він допускається до заліку.

Ваговий бал за залік **Re** складає **30 балів** і одержується за наступне:

1. Відповідь на 2 теоретичні питання оцінюються максимально в 10 балів по 5 балів за кожну вірну відповідь.

- 2. Відповідь на 2 практичних питання оцінюється максимально в 20 бали по 10 балів за кожну вірну відповідь.

Загальний рейтинг: **$R_r = R_s + R_e = 70 + 30 = 100$ балів**

9. Додаткова інформація з дисципліни (освітнього компонента)

Перелік основних питань, які виносяться на семестровий контроль:

1. Що таке MLOps і як він інтегрується з DevOps?
2. Які основні принципи та практики MLOps?
3. Опишіть типовий життєвий цикл моделі машинного навчання.
4. Які етапи включає процес розробки та впровадження моделей машинного навчання?
5. Які інструменти та платформи часто використовуються в MLOps?
6. Як контейнеризація та оркестрація впливають на процеси MLOps?
7. Які методи автоматизації використовуються в MLOps для підтримки моделей машинного навчання?
8. Які стратегії моніторингу ефективні для відстеження продуктивності моделей машинного навчання?
9. Як забезпечується якість та надійність моделей машинного навчання в MLOps?
10. Які методи тестування та валідації використовуються для оцінки моделей машинного навчання?
11. Як машинне навчання інтегрується з загальними бізнес-процесами в організації?
12. Які виклики можуть виникнути при інтеграції систем машинного навчання з існуючою інфраструктурою компанії?

13. Наведіть приклад успішного впровадження MLOps у реальному проекті.
14. Які заходи безпеки необхідно враховувати при розробці систем машинного навчання?
15. Як MLOps допомагає у забезпеченні конфіденційності даних?
16. Які стратегії використовуються для масштабування моделей машинного навчання в MLOps?
17. Як оптимізувати продуктивність моделей машинного навчання без втрати точності?
18. Як хмарні технології впливають на розвиток та впровадження MLOps?
19. Які переваги та виклики пов'язані з використанням хмарних платформ для MLOps?
20. Що таке континуальне навчання в контексті MLOps і чому воно важливе?
21. Як забезпечити, щоб моделі машинного навчання адаптувалися до змінних умов даних та середовища?
22. Які етичні розгляди необхідно враховувати при розробці та впровадженні моделей машинного навчання?
23. Як забезпечити відповідальне використання моделей машинного навчання в бізнесі?
24. Як організувати ефективну командну роботу між розробниками, інженерами машинного навчання та операційними командами?
25. Які навички та компетенції важливі для команди, що працює в області MLOps?

Робочу програму навчальної дисципліни (силабус):

Склав асистент кафедри системного проектування,
кандидат технічних наук, **Кислий Роман Володимирович**

Ухвалено кафедрою системного проектування (протокол № 10 від 12 червня 2023 р.)

Погоджено Методичною комісією ІПСА (протокол № 4 від 16 червня 2023 р.)

Погоджено науково-методичною комісією КПІ ім. Ігоря Сікорського зі спеціальності 122 (протокол № 6 від 27 червня 2023 р.)