



МЕТОДИ МАШИННОГО НАВЧАННЯ

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Перший (бакалаврський)</i>
Галузь знань	<i>11 Математика та статистика</i>
Спеціальність	<i>113 Прикладна математика</i>
Освітня програма	<i>Математичні методи криптографічного захисту інформації</i>
Статус дисципліни	<i>Вибіркова</i>
Форма навчання	<i>очна (денна)</i>
Рік підготовки, семестр	<i>IV курс, осінній семестр</i>
Обсяг дисципліни	<i>Загальна кількість: 120 години / 4 кредити Лекційних занять: – 36 годин Лабораторних занять: 36 годин Самостійна робота студентів – 48 годин</i>
Семестровий контроль/ контрольні заходи	<i>Залік, модульна контрольна робота</i>
Розклад занять	<i>http://ipt.kpi.ua/navchalnij-protses</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	<i>Лектор: д.т.н., професор, Куссуль Наталія Миколаївна, natalia.kussul.kpi@gmail.com Лабораторні: Охріменко Антон</i>
Розміщення курсу	

Програма навчальної дисципліни

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

Програму навчальної дисципліни «Методи машинного навчання» складено відповідно до освітньої програми підготовки бакалаврів спеціальності 113 Прикладна математика.

Предмет навчальної дисципліни – це методи і технології машинного навчання.

Мета вивчення дисципліни “Методи машинного навчання” — є оволодіння методами та технологіями машинного навчання, а також їх програмування з урахуванням сучасних тенденцій розвитку цієї галузі в епоху цифровізації та Industry 4.0, прищепити їм вміння розв'язувати реальні інженерні та науково-технічні задачі різної складності з використанням інтелектуальних інформаційних технологій.

Основні завдання навчальної дисципліни

Згідно з вимогами освітньо-професійної програми студенти після засвоєння навчальної дисципліни мають продемонструвати такі результати навчання:

знання:

- основних складових частин інтелектуальних технологій;
- архітектури нейронних мереж;
- алгоритмів навчання нейронних мереж;
- засоби побудови нейромережових систем;
- алгоритм навчання Support Vector Machine;
- алгоритми навчання дерев рішень та Random forest;
- основи кластеризації;
- регресійні методи.

уміння:

- визначити структуру моделі машинного навчання для розв'язання конкретної задачі;
- програмно реалізувати алгоритм її функціонування;

- визначити загальний метод навчання моделі та запрограмувати його;
- використовувати методи навчання з учителем та без нього;
- програмно реалізувати методи виявлення аномалій в даних;
- розв'язувати задачі машинного навчання на основі геопросторових даних.

досвід:

- застосування інтелектуальних систем для аналізу поставленої задачі, розробки алгоритму та програми, та, інтерпретації отриманих результатів.

У результаті вивчення даної дисципліни будуть отримані наступні загальні компетентності, фахові компетентності та програмні результати навчання у відповідності до стандарту вищої освіти:

Загальні компетентності

- ЗК01. Здатність учитися і оволодівати сучасними знаннями.
- ЗК02. Здатність застосовувати знання у практичних ситуаціях.
- ЗК06. Здатність до абстрактного мислення, аналізу та синтезу.
- ЗК07. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

Фахові компетентності

- ФК03. Здатність обирати та застосовувати математичні методи для розв'язання прикладних задач, моделювання, аналізу, проектування, керування, прогнозування, прийняття рішень.
- ФК08. Здатність використовувати сучасні технології програмування та тестування програмного забезпечення.

Програмні результати навчання

- РН14. Виявляти здатність до самонавчання та продовження професійного розвитку.
- РН15. Уміти організувати власну діяльність та одержувати результат у рамках обмеженого часу.
- РН17. Уміти здійснювати збір, опрацювання, аналіз, систематизацію науково-технічної інформації, уникаючи при цьому академічної недоброчесності.

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

Міждисциплінарні зв'язки: дисципліна «Методи машинного навчання» базується на таких курсах, як «Дискретна математика», «Математичний аналіз», «Програмне забезпечення обчислювальних систем», «Алгебра та геометрія», «Програмування» тощо.

Отримані практичні навички та засвоєні під час вивчення теоретичні знання в подальшому можна використовувати у професійній діяльності.

3. Зміст навчальної дисципліни

Розділ 1. Вступ. Біологічні основи та історія розвитку інтелектуальних обчислень.

Розділ 2. Регресія.

Тема 2.1. Лінійна регресія.

Тема 2.2. Логістична регресія.

Тема 2.3. Регуляризація.

Розділ 3. Нейронні мережі.

Тема 3.1. Мережі прямого розповсюдження. Їх архітектура.

Тема 3.2. Метод градієнтного спуску.

Тема 3.3. Метод зворотного поширення похибок.

Тема 3.4. Методи глибинного навчання.

Розділ 4. Інші традиційні класифікатори

Тема 4.1. Random forest.

Тема 4.2. Support Vector Machine.

Розділ 4. Кластеризація.

4. Навчальні матеріали та ресурси

1. Mills T.C. Applied Time Series Analysis A Practical Guide to Modeling and Forecasting, 2019. – 356 p.
2. Blitzstein J.K. et al. Time Series for Data Science, 2022. – 529 p.
3. Foster D. Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play, 2019. – 330 p..

Електронні ресурси:

4. База даних <https://www.kaggle.com/datasets>

5. Відео-курс «Machine Learning — Andrew Ng, Stanford University»
https://www.youtube.com/playlist?list=PLLsT5z_DsK-h9vYZkQkYNWcltqhIRJLN
6. Відео-курси з Random Forest:
 - ✓ <https://www.youtube.com/watch?v=psvL-kpc9uw>
 - ✓ <https://www.youtube.com/watch?v=NsuQRe-9tb4>
 - ✓ <https://www.youtube.com/watch?v=3kYujfDgmNk>
7. Приклади з використанням геопросторових даних:
 - ✓ simple: https://github.com/acgeospatial/Satellite_Imagery_Python/blob/master/Clustering_KMeans-Sentinel2.ipynb
 - ✓ middle: https://colab.research.google.com/drive/1tGgypw1qM6TJD9A2NLWI_X1JMoyiuC_b?usp=sharing#scrollTo=cppWnp9SBgnE
 - ✓ advanced:
 1. <https://towardsdatascience.com/whats-growing-there-a5618a2e6933>
 2. <https://github.com/jakevdp/PythonDataScienceHandbook/>
 3. <https://www.kaggle.com/learn/intro-to-machine-learning>
 4. <https://realpython.com/linear-regression-in-python/>

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

Навчання здійснюється на основі студентоцентрованого підходу та стратегії взаємодії викладача та студента для засвоєння студентами матеріалу та розвитку у них практичних навичок. Для проведення занять застосовується практичний метод. Для лекційних занять використовуються пояснювально-ілюстративний метод та метод проблемного виконання, для проведення лабораторних робіт використовується частково-пошуковий та дослідницький методи навчання, при яких викладач ставить перед студентами проблему.

Загальна структура курсу

Назви змістових модулів і тем	Кількість годин				
	Всього	у тому числі			
		Лекції	Практ. (семін.)	Лаборат. (комп. пр.)	СРС
1	2	3	4	5	6
Розділ 1. Вступ. Біологічні основи та історія розвитку інтелектуальних обчислень.					
Розділ 2. Регресія					
Тема 2.1 Лінійна регресія.		4		4	
Тема 2.2 Логістична регресія		4			
Тема 2.3 Регуляризація		4			
Розділ 3. Нейронні мережі					
Тема 3.1 Мережі прямого розповсюдження. Їх архітектура.		4		4	
Тема 3.2 Метод градієнтного спуску		4		4	
Тема 3.3. Метод зворотного поширення похибок.		4		4	
Тема 3.4. Методи глибокого навчання.		4			
Розділ 4. Інші традиційні класифікатори					
Тема 4.1 Random Forest		4		8	
Тема 4.2 Support Vector Machine.		4		12	
Розділ 5. Кластеризація					
Всього годин					
	120	36		36	48

Практичні заняття в межах навчальної дисципліни не передбачені.

Рекомендований перелік лабораторних робіт:

1. Лабораторна робота № 1. Регресія

Ознайомитись з основами машинного навчання та аналізу даних для розв'язання задачі регресії, реалізувати методи, що базуються на принципі лінійної регресії (Lasso, Ridge, ElasticNet) та ансамблювати їх за допомогою бібліотеки scikit-learn.

2. **Лабораторна робота № 2.** Перцептрон Розенблатта. Багатошаровий перцептрон
Реалізувати одношаровий перцептрон, використовуючи такі мови програмування як C++, Java тощо. За допомогою реалізованого перцептрона розв'язати задачу згідно з номером варіанту. Навчити нейронну мережу на навчальній вибірці, використовуючи алгоритм Розенблатта. Перевірити роботу перцептрона на тестових даних.

3. **Лабораторна робота № 3.** Древа прийняття рішень
Ознайомитись з основами машинного навчання та аналізу даних, реалізувати процедуру побудови дерева прийняття рішень (Decision Tree) та ансамблю дерев прийняття рішень (Random Forest) із використанням бібліотеки scikit-learn.

4. **Лабораторна робота № 4.** Support Vector Machine
Ознайомитись з основами машинного навчання та аналізу даних, реалізувати процедуру класифікації із використанням методу опорних векторів (Support Vector Machine, SVM) бібліотеки scikit-learn.

5. **Лабораторна робота № 5.** Кластеризація
Ознайомитись з основами машинного навчання та аналізу даних без учителя для розв'язання задачі кластеризації даних, реалізувати метод, що базується на принципі навчання без учителя (K-середніх) за допомогою бібліотеки scikit-learn.

Самостійна робота студента

До СРС відносяться такі види робіт: вивчення теоретичного матеріалу, виконання додаткових практичних завдань. При здачі лабораторних робіт оформлюється протоколи, які підписуються викладачем. Самостійна робота включає в себе підготовку до лекційних занять, а саме, підготовка презентаційного матеріалу за тематикою дисципліни. Нижче наведено орієнтовні теми, за якими можна підготувати доповідь:

1. Convolutional Neural Networks (CNN)
2. VGG, Inception
3. Long short-term memory (LSTM)
4. Інші архітектури U-Net, ResNet
5. Generative Adversarial Networks (GANs)
6. Pretrained neural networks
7. Бібліотеки для реалізації Deep Learning
8. Навчальні ресурси по Deep Learning

Політика та контроль

Політика навчальної дисципліни (освітнього компонента)

Відвідування занять

Відвідування лекцій та занять з комп'ютерного практикуму, а також відсутність на них, не оцінюється. Однак студентам рекомендується відвідувати заняття, оскільки на них викладається теоретичний матеріал та розвиваються практичні навички, необхідні для успішного виконання лабораторних робіт та МКР. Система оцінювання орієнтована на отримання балів за активність студента, а також виконання завдань, які здатні розвинути його практичні уміння та навички.

Оголошення результатів контрольних заходів

Результати модульної контрольної роботи (або тесту) вказуються на бланках для модульної контрольної роботи (завдання, які виконували студенти) з позначенням усіх помилок, коректної або некоректної відповіді, а також з коментарями, зауваженнями тощо.

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

Процедура оскарження результатів контрольних заходів

Студенти мають можливість підняти будь-яке питання, яке стосується процедури проведення та/або оцінювання контрольних заходів, та очікувати, що воно буде розглянуто згідно із наперед визначеними процедурами.

Студенти мають право оскаржити результати контрольних заходів, але обов'язково аргументовано, пояснивши, з яким критерієм не погоджуються відповідно до оціночного листа та/або зауважень.

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

Рейтинг студента з дисципліни складається з балів, що отримуються:

за **5 основних лабораторних робіт та модульну контрольну роботу** (поділяється на дві контрольні роботи).

Семестровим контролем є залік.

Система рейтингових (вагових) балів та критерії оцінювання:

- *Виконання та захист основних ЛР:*

- *ЛР №1, №2б* – до 15 балів

- повне виконання, повна відповідь при захисті 10-15;

- працююча програма, але неповна відповідь при захисті 9-5;

- непрацююча програма (в тому числі неспроможність виконати додаткове завдання), роботу не зараховано 0..3;

- за несвоєчасне виконання лабораторного практикуму нараховуються штрафні бали – від 1 до 9.

- *ЛР №2а, №3, №4, №5* – до 6 балів

- повне виконання, повна відповідь при захисті 6;

- працююча програма, але неповна відповідь при захисті 3..5;

- непрацююча програма (в тому числі неспроможність виконати додаткове завдання), роботу не зараховано 0..2;

- за несвоєчасне виконання лабораторного практикуму нараховуються штрафні бали – від 1 до 4.

- *Модульний контроль:*

- *повне виконання завдання: 40 балів;*

- *неповне виконання завдання (є неprincipові помилки, неточності): 30 балів;*

- *неповне виконання*

- *завдання (є принципіві помилки): 10-20 балів;*

- *незадовільне виконання*

- *завдання: 5 балів.*

№ з/п	Контрольний захід	Макс бал	Кількість	Всього
1	МКР	40	1	40
2	ЛР №1, №2б	15	2	30
3	ЛР №2а, №3, №4, №5	6	5	30
				100

На останньому за розкладом занятті викладач проводить семестрову атестацію у вигляді співбесіди зі студентами, які не змогли отримати за рейтингом позитивну оцінку (набрали протягом семестру менше ніж 60 балів (RD < 60)), але були допущені до семестрової атестації. Ці студенти зобов'язані проходити співбесіду. У даному випадку рейтингова оцінка студента буде складатись з результатів роботи в семестрі (RD) та результатів співбесіди, але не вище 60 балів.

Студенти, які протягом семестру отримали більш ніж 60 балів, можуть пройти співбесіду з метою підвищення оцінки. У даному випадку семестровий рейтинг студента анулюється, і студент отримує оцінку за результатами співбесіди.

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

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

Заохочувальні бали:

1. Додаткові доповіді на лекціях на задану тему (до 5 балів)
2. Доопрацювання/модифікація існуючих лабораторних робіт (до 5 балів за одну роботу).
3. Додаткові бали (+1 бал) за правильну відповідь на питання на лекції
4. Додаткові бали за проміжні тести.
5. Участь в олімпіаді KPI-Open та/або інших аналогічних заходах (участь з підтвердженням – 5 балів, призове місце – 10 балів).
6. Додаткові курси за тематикою курсу – до 10 балів при умові доповіді на лекції.

Додаткова інформація з дисципліни (освітнього компонента)

За успішне проходження дистанційних чи онлайн курсів за відповідною тематикою та отримання відповідного сертифікату можна отримати додаткові бали (за домовленістю з викладачем).

Складено д.т.н., проф., Куссуль Н.М.

Ухвалено кафедрою ММАД (протокол № 14 від 22.06.2022),

Погоджено Методичною комісією НН ФТІ (протокол № 6 від 30.06.2022)