

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

1. Загальна інформація про навчальну дисципліну

Повна назва навчальної дисципліни	Машинне навчання
Повна офіційна назва закладу вищої освіти	Сумський державний університет
Повна назва структурного підрозділу	Факультет електроніки та інформаційних технологій. Кафедра прикладної математики та моделювання складних систем
Розробник	Князь Ігор Олександрович, к.ф.-м.н., доцент
Рівень вищої освіти	Другий рівень вищої освіти, НРК – 7 рівень, QF-LLL – 7 рівень, FQ-EHEA – другий цикл
Семестр вивчення навчальної дисципліни	16 тижнів протягом 2-го семестру
Обсяг навчальної дисципліни	Обсяг навчальної дисципліни становить 5 кредитів ЄКТС, 150 годин, з яких 64 години становить контактна робота з викладачем (32 години лекцій, 32 години практичних занять), 86 годин становить самостійна робота
Мова викладання	українська

2. Місце навчальної дисципліни в освітній програмі

Статус дисципліни	Обов'язкова навчальна дисципліна для освітньої програми "Наука про дані та моделювання складних систем"
Передумови для вивчення дисципліни	Необхідні знання з: теорії ймовірності та математичної статистики, алгебри, математичного аналізу, програмування
Додаткові умови	Додаткові умови відсутні
Обмеження	Обмеження відсутні

3. Мета навчальної дисципліни

Метою дисципліни є формування у студентів сучасного наукового світогляду в області методів машинного навчання; знайомство студентів з сучасними технологіями автоматизованого інтелектуального аналізу даних та тенденціями розробки і застосування інформаційних інтелектуальних систем; подальше становлення і вдосконалення інформаційної та програмної культури майбутніх фахівців.

4. Зміст навчальної дисципліни

Модуль 1. Алгоритми машинного навчання
Тема 1 Основні концепції машинного навчання. Навчання з вчителем. Задачі машинного навчання. Історія розвитку теорії машинного навчання. Основні етапи роботи системи машинного навчання. Збір даних та їх первинна обробка. Візуалізація даних. Огляд методів класифікації. Математичні основи методів NaiveBayes, DecisionTrees, SVM, K-NN. Регресійний аналіз.

<p>Тема 2 Навчання без вчителя Математичне підгрунтя методів кластеризації. Алгоритми DBSCAN, K-Means, Mean-Shift. Пошук правил (асоціація). Алгоритми Euclat, Apriori, FP-Growth. Зменшення розмірності (узагальнення). Алгоритми PCA, SVD, LDA. Алгоритми латентно-семантичного аналізу.</p>
<p>Тема 3 Навчання з підкріпленням Математичні основи методу Q-Learning. Алгоритми SARSA, DQN, A3C. Генетичні алгоритми.</p>
<p>Тема 4 Ансамблеві методи Стекінг. Бегінг (RandomForest). Бустінг. Алгоритми XGBoost, AdaBoost, CatBoost.</p>
<p>Модуль 2. Основи проектування нейронних мереж</p>
<p>Тема 5 Перцептрони Біологічні нейрони та їх фізичні моделі. Математичні моделі нейроелементів. Поняття: синапс, ваговий коефіцієнт, поріг, дискримінантна функція, функція активації, одношаровий перцептрон. Метод найменших квадратів як основа алгоритму Уїдроу-Хоффа. Можливості і властивості одношарових перцептронів. Лінійна роздільність і лінійна нероздільність класів.</p>
<p>Тема 6 Нейронні мережі прямого поширення. Градієнтні методи навчання Багатошаровий перцептрон: модель і принципи побудови архітектури. Алгоритм зворотного поширення помилки. Градієнтні алгоритми навчання багатошарових нейромереж. Порівняння моделей та алгоритмів навчання нейромереж прямого поширення. Моделі та принципи синтезу архітектури радіально-базисних нейромереж. Методи навчання радіально-базисних нейромереж. Згорткові мережі.</p>
<p>Тема 7 Повнозв'язні нейронні мережі. Нейронні мережі Кохонена. Бінарні повнозв'язні нейромережі Хопфілда. Нейромережа Ельмана. Карти ознак самоорганізації Кохонена: нейронна мережа SOM.</p>

5. Очікувані результати навчання навчальної дисципліни

Після успішного вивчення навчальної дисципліни здобувач вищої освіти зможе:

РН1	Оцінювати можливість застосування алгоритмів машинного навчання для розв'язання поставленої задачі.
РН2	Сформулювати математичну постановку задачі, спираючись на постановку мовою предметної галузі, та обирати метод її розв'язання, що забезпечує потрібні точність і надійність результату.
РН3	Здійснювати підготовку та первинну обробку даних для їх подальшого використанні при машинному навчанні.
РН4	Аналізувати результати роботи алгоритму з метою вибору найкращого методу машинного навчання, що підходить для конкретної задачі.

6. Роль навчальної дисципліни у досягненні програмних результатів

Програмні результати навчання, досягнення яких забезпечує навчальна дисципліна.

ПР9	Застосовувати сучасні технології програмування та розроблення програмного забезпечення, програмної реалізації чисельних і символьних алгоритмів в галузях аналізу великих даних та побудови комп'ютерного експерименту.
ПР12	Виявляти здатність до самонавчання та продовження професійного розвитку.

PR17	Володіти математичними методами первинної обробки великих наборів даних. Вміти обирати до застосування оптимальні методи для конкретної задачі побудови моделі поведінки складної системи за існуючим набором даних .
PR18	Уміти розробляти та використовувати на практиці алгоритми, пов'язані з спрощенням даних, що описують поведінку системи, класифікацією даних за певними ознаками без навчання та за попередньої наявності класів даних.
PR19	Уміти оцінити на адекватність результат обробки великих масивів даних.

7. Види навчальних занять та навчальної діяльності

7.1 Види навчальних занять

Тема 1. Основні концепції машинного навчання. Навчання з вчителем.	
Лк1 "Задачі машинного навчання" (денна)	Задачі машинного навчання. Історія розвитку теорії машинного навчання. Основні етапи роботи системи машинного навчання. Збір даних та їх первинна обробка. Візуалізація даних.
Лк2 "Навчання з вчителем" (денна)	Огляд методів класифікації. Алгоритми NaiveBayes та DecisionTrees. Алгоритми SVM, K-NN. Регресійний аналіз.
Пр1 "Методи класифікації" (денна)	Алгоритм NaiveBayes та його програмна реалізація
Пр2 "Класифікація за допомогою DecisionTrees." (денна)	Навчання за алгоритмом DecisionTrees та його програмна реалізація.
Пр3 "Алгоритми SVM та K-NN." (денна)	Алгоритми SVM та K-NN. Програмна реалізація.
Тема 2. Навчання без вчителя	
Лк3 "Огляд методів кластеризації." (денна)	Огляд алгоритмів DBSCAN, K-Means, Mean-Shift
Лк4 "Пошук правил" (денна)	Алгоритми Euclat, Apriori, FP-Growth.
Лк5 "Зменшення розмірності (узагальнення)." (денна)	Огляд алгоритмів PCA, SVD, LDA. Алгоритми латентно-семантичного аналізу.
Пр4 "Розв'язання задач кластеризації" (денна)	Кластеризація даних за алгоритмом DBSCAN
Пр5 "Пошук правил" (денна)	Алгоритм Apriori та його програмна реалізація.
Пр6 "Компонентний аналіз" (денна)	Аналіз даних за допомогою алгоритму PCA
Тема 3. Навчання з підкріпленням	
Лк6 "Навчання з підкріпленням" (денна)	Алгоритми Q-Learning, SARSA, DQN, A3C та особливості їх програмної реалізації
Лк7 "Генетичні алгоритми" (денна)	Генетичні алгоритми
Пр7 "Навчання з підкріпленням" (денна)	Програмна реалізація алгоритму Q-Learning

<p>Пр8 "Модифікації методу Q-learning" (денна) Приклади практичного застосування методів SARSA (A3C)</p>
<p>Лк8 "Математичні основи методу навчання з підкріпленням" (денна) Математичні основи методів групи Reinforcement Learning та їхніх модифікацій.</p>
<p>Тема 4. Ансамблеві методи</p>
<p>Лк9 "Ансамблеві методи" (денна) Особливості методів стекінгу, бегінгу (RandomForest), бустінгу. Алгоритми XGBoost, AdaBoost, CatBoost.</p>
<p>Пр9 "Ансамблеві методи" (денна) Побудова ансамблю у рамках стекінгу. Реалізація бегінгу.</p>
<p>Тема 5. Персептрони</p>
<p>Лк10 "Загальна характеристика та основні принципи побудови нейромереж" (денна) Біологічні нейрони та їх фізичні моделі. Математичні моделі нейроелементів. Персептрон. Алгоритм Уідроу-Хоффа.</p>
<p>Лк11. "Математичні основи алгоритмів навчання персептронів" Метод градієнтного спуску. Пакетне навчання.</p>
<p>Пр10 "Навчання персептрона" (денна) Програмна реалізація алгоритму Уідроу-Хоффа..</p>
<p>Тема 6. Нейронні мережі прямого поширення. Градієнтні методи навчання</p>
<p>Лк12 "Багатошарові мережі" (денна) Багатошарові мережі. Алгоритми навчання. Моделі та принципи синтезу архітектури радіально-базисних нейромереж. Алгоритм зворотнього поширення помилки.</p>
<p>Лк13. "Згорткові мережі" (денна) Архітектура згорткових мереж.</p>
<p>Лк14. Математичні основи навчання загорткових нейромереж Математичні основи навчання загорткових нейромереж. Фільтри. Згортки. Алгоритми навчання.</p>
<p>Пр11 "Класифікація за допомогою нейромереж" (денна) Алгоритм зворотнього поширення помилки та його програмна реалізація</p>
<p>Пр12 "Згорткові мережі" (денна) Застосування згорткових мереж для класифікації масиву вхідних зображень</p>
<p>Пр13 "Навчання радіально-базисних нейромереж" (денна) Архітектура радіально-базисних нейромереж. Методи навчання та їх програмна реалізація</p>
<p>Тема 7. Повнозв'язні нейронні мережі. Нейронні мережі Кохонена.</p>
<p>Лк15 "Повнозв'язні нейронні мережі" (денна) Архітектура мереж Хопфільда та Хеммінга. Математичні основи методів навчання та приклади практичного застосування</p>
<p>Лк16 "Математичні основи алгоритму навчання мережі Хопфільда "</p>

Пр14 "Побудова мережі Хопфілда" (денна) Мережа Хопфілда. Програмна реалізація методу навчання.
Пр15 "Карті Кохонена" (денна) Застосування мереж Кохонена у задачах кластер-аналізу.
Пр16 "Використання нейромереж для розв'язання практичних задач" (денна) Нейромережі у задачах прогнозування та оцінювання.

7.2 Види навчальної діяльності

НД1	Самонавчання
НД2	Робота з підручниками та релевантними інформаційними джерелами
НД3	Підготовка та презентація доповіді
НД4	Підготовка до практичних занять
НД5	Підготовка до поточного та підсумкового контролю
НД6	Виконання індивідуальних розрахунково-аналітичних завдань

8. Методи викладання, навчання

Дисципліна передбачає навчання через:

МН1	Інтерактивні лекції (або он-лайн лекції)
МН2	Лекції-дискусії
МН3	Поблемно-пошуковий метод
МН4	Репродуктивний метод
МН5	Творчий метод
МН6	Аналіз конкретних ситуацій
МН7	Практико-орієнтоване навчання

Лекції надають студентам матеріали з фундаментальних основ теорії алгоритмів машинного навчання, що є основою для самостійного навчання здобувачів вищої освіти (РН1, РН2). Лекції доповнюються практичними заняттями, що надають студентам можливість застосовувати теоретичні знання на практичних прикладах (РН3, РН4).

Soft skills:

ЗК01. Здатність вчитися і оволодівати сучасними знаннями. ЗК02. Здатність до адаптації та дії в новій ситуації. ЗК03. Здатність бути критичним і самокритичним. ЗК05. Здатність генерувати нові ідеї (креативність).

9. Методи та критерії оцінювання

9.1. Критерії оцінювання

Шкала оцінювання ECTS	Визначення	Чотирибальна національна шкала оцінювання	Рейтингова бальна шкала оцінювання
5 (відмінно)	Відмінне виконання лише з незначною кількістю помилок	A	$90 \leq RD \leq 100$
4 (добре)	Вище середнього рівня з кількома помилками	B	$82 \leq RD < 89$
4 (добре)	Загалом правильна робота з певною кількістю помилок	C	$74 \leq RD < 81$
3 (задовільно)	Непогано, але зі значною кількістю недоліків	D	$64 \leq RD < 73$
3 (задовільно)	Виконання задовольняє мінімальні критерії	E	$60 \leq RD < 63$
2 (незадовільно)	Можливе повторне складання	FX	$35 \leq RD < 59$
2 (незадовільно)	Виконання задовольняє мінімальні критерії	F	$0 \leq RD < 34$

9.2 Методи поточного формативного оцінювання

МФО1	Захист презентацій та рефератів
МФО2	Настанови викладача в процесі виконання практичних завдань
МФО3	Обговорення та самокорекція виконаної роботи студентами
МФО4	Опитування та усні коментарі викладача за його результатами
МФО5	Самостійне виконання студентами ситуаційних вправ на практичних заняттях та їх обговорення.

9.3 Методи підсумкового сумативного оцінювання

МСО1	Звіт за результатами виконання практичних робіт
МСО2	Написання за захист курсової роботи
МСО3	Складання комплексного письмового модульного контролю

Контрольні заходи:

2-й семестр		100 балів
МСО1. Звіт за результатами виконання практичних робіт		60
	6x10	60
МСО3. Складання комплексного письмового модульного контролю		40
	2x20	40

Контрольні заходи в особливому випадку:

Курсова робота:

2-й семестр	100 балів
МСО2. Написання за захист курсової роботи	100

Оцінка студента формується таким чином: 1. звіт за результатами виконання практичної роботи (виконання, презентація, захист) – 10 балів за кожну роботу (усього 6 практичних робіт); 2. підсумковий модульний контроль (ПМК) у кожному модульному циклі – по 20 балів (форма підсумкового контролю – залік, що проводиться у письмовій формі за тестовими технологіями). Захист практичної роботи є комплексним і включає захист теоретичного матеріалу згідно теми практичної роботи, захист базового та індивідуальних завдань за даною темою. Відповідно рейтингові бали розподіляються наступним чином: 4 бали нараховуються за теоретичну частину (від 2,4 до 4 балів), 6 балів – за практичне виконання базового та індивідуальних завдань (від 3.6 до 6 балів). Курсова робота з дисципліни передбачає проведення усіх етапів машинного навчання на реальних даних: підготовку даних; вибір оптимального методу навчання; програмну реалізацію та інтерпретацію результатів. Дані та предмет дослідження для курсової роботи обираються студентами самостійно та узгоджуються з викладачем до кінця 1 модулю навчання. Шкала оцінювання курсової роботи з навчальної дисципліни: R=100 балів (при позитивному оцінюванні від 60 до 100 балів). Оцінюванню підлягають такі етапи курсової роботи:

- розуміння даних, доменної галузі, формулювання проблеми, актуальність – 20 б.
- чіткий та деталізований аналіз даних, що відповідає поставленій меті – 20 б.;
- ефективність використаних підходів та методів – 20 б.;
- наявність чітких та обґрунтованих рекомендацій та висновків – 20 б.;
- захист та презентація результатів роботи – 20 б.

10. Ресурсне забезпечення навчальної дисципліни

10.1 Засоби навчання

ЗН1	Інформаційно-комунікаційні системи
ЗН2	Бібліотечні фонди
ЗН3	Комп'ютери, комп'ютерні системи та мережи

10.2 Інформаційне та навчально-методичне забезпечення

Основна література	
1	Knyaz', I. Modelling of Neural Networks: lecture notes in two parts. Part1 / I. Knyaz'. – Електронне вид. каф. ПМ та МСС. – Sumy : SumyStateUniversity, 2017. – 62 p.
2	Knyaz', I. A. Methodological instructions for practical training in "Modelling of Neural Networks": for students of the speciality 8.04030101 "Applied Mathematics" / I. A. Knyaz'. – Електронне вид. каф ПМ та МСС – Sumy : SumyStateUniversity, 2017.
3	Хайкин С. Нейронные сети: полный курс, 2-е издание.: Пер. с англ.. – М.: Издательский дом «Диалектика-Вильямс», 2018. – 1104 с.
4	Золотых, Н. Ю. Машинное обучение и анализ данных / Н. Ю. Золотых. – 2018. – 693 с.
5	Мюллер, А. Введение в машинное обучение с помощью Python: руководство для специалистов по работе с данными / А. Мюллер, С. Гвидо. – М., 2016-2017. – 393 с.
Допоміжна література	
6	Рассел С., Питер Н. Искусственный интеллект: современный подход, 2-е изд.: Пер. С англ.. – М.: Издательский дом «Вильямс», 2015. – 1408 с.