

Національний університет водного господарства та
природокористування
Навчально-науковий інститут автоматики, кібернетики та
обчислювальної техніки

ЗАТВЕРДЖУЮ

Голова науково-методичної ради НУВГП
_____ Олег ЛАГОДНЮК

«___» _____ 20__

04-01-16S

СИЛАБУС

навчальної дисципліни

SYLLABUS

Інтелектуальний аналіз даних		Data mining
Шифр за ОП	<u>BB 43.1</u>	Code in Educational Program
Освітній рівень: бакалаврський (перший)		Educational level: Bachelor's (first)
Галузь знань: Інформаційні технології	<u>12</u>	Fields of knowledge: Information technologies
Спеціальність: Інженерія програмного забезпечення	<u>121</u>	Field of study: Software engineering
Спеціалізація: Інтернет речей		Specialization: Internet of things
Освітня програма: Інтернет речей		Educational Program: Internet of things

Силабус навчальної дисципліни **«Інтелектуальний аналіз даних»** для здобувачів вищої освіти ступеня «бакалавр», які навчаються за освітньо-професійною програмою **«Інтернет речей», 121 Інженерія програмного забезпечення**. Рівне. НУВГП. 2020. 12 стор.

ОПП «Інтернет речей» на сайті університету:
http://ep3.nuwm.edu.ua/18444/1/opp_internet_rechey_2019.pdf

Розробник силабусу: *Демчук Олена Станіславівна, к.т.н., доцент кафедри комп'ютерних наук та прикладної математики*

Силабус схвалений на засіданні кафедри комп'ютерних наук та прикладної математики
Протокол № 2 від "6" жовтня 2020 року

В.о. завідувача кафедри: *Турбал Юрій Васильович, д.т.н., професор*
Керівник освітньої програми: *Жуковський Віктор Володимирович, к.т.н., доцент*

Схвалено науково-методичною радою з якості ННІ АКOT
Протокол № 3 від "29" грудня 2020 року

Голова науково-методичної ради з якості ННІ АКOT: *Мартинюк Петро Миколайович, д.т.н., професор*

СЗ №-249 в ЕДО від 25 січня 2021 року (70-75442569).

© Демчук О.С., 2020
© НУВГП, 2020

ЗАГАЛЬНА ІНФОРМАЦІЯ*

Ступінь вищої освіти	бакалавр
Освітня програма	Інтернет речей
Спеціальність	121 Інженерія програмного забезпечення
Рік навчання, семестр	3 рік; 5 семестр
Кількість кредитів	4
Лекції:	20
Лабораторні заняття:	28
Самостійна робота:	72
Курсова робота:	ні
Форма навчання	денна
Форма підсумкового контролю	залік
Мова викладання	українська

ІНФОРМАЦІЯ ПРО ВИКЛАДАЧА*

Лектор



Демчук Олена Станіславівна, к.т.н., доцент кафедри комп'ютерних наук та прикладної математики

Вікіситет

http://wiki.nuwm.edu.ua/index.php/Демчук_Олена_Станіславівна

ORCID

<http://orcid.org/0000-0002-8318-5009>

Як комунікувати

o.s.demchuk@nuwm.edu.ua

ПРО ДИСЦИПЛІНУ

Анотація навчальної дисципліни, в т.ч. мета та цілі

Програма навчальної дисципліни «Інтелектуальний аналіз даних» складена відповідно до освітньо-професійної програми «Інтернет речей» підготовки бакалавра за спеціальністю «Інженерія програмного забезпечення».

Дисципліна «Інтелектуальний аналіз даних» належить до дисциплін вільного вибору студентів.

Вона спрямована на підготовку майбутніх інженерів-програмістів до ефективного використання класичних і сучасних методів, а також програмних засобів інтелектуального аналізу даних та їх застосування у

	<p>подальшій професійній діяльності.</p> <p>Мета: оволодіння методами, алгоритмами та програмними засобами Data Mining.</p> <p>Завдання: сформувати у студентів теоретичні знання та практичні вміння у сфері видобутку та інтелектуального аналізу даних, що виникають при дослідженні різноманітних інформаційних, технічних, економічних, екологічних, соціальних та інших систем.</p> <p>У результаті вивчення навчальної дисципліни студенти будуть:</p> <p>знати:</p> <ul style="list-style-type: none"> – основні поняття інтелектуального аналізу даних; – класичні статистичні та сучасні кібернетичні методи і моделі інтелектуального аналізу даних; – програмні засоби для інтелектуального аналізу даних. <p>вміти:</p> <ul style="list-style-type: none"> – вибирати методи і моделі інтелектуального аналізу даних при вирішенні конкретних практичних задач; – використовувати сучасні програмні засоби для інтелектуального аналізу даних; – аналізувати та інтерпретувати отримані результати при вирішенні прикладних задач.
<p>Посилання на розміщення навчальної дисципліни на навчальній платформі Moodle</p>	<p>https://exam.nuwm.edu.ua/course/view.php?id=4780</p>
<p>Компетентності</p>	<p>ЗК01. Здатність до абстрактного мислення, аналізу та синтезу.</p> <p>ЗК02. Здатність застосовувати знання у практичних ситуаціях.</p> <p>ЗК06. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.</p> <p>ФК08. Здатність застосовувати фундаментальні міждисциплінарні знання для успішного розв'язання завдань інженерії програмного забезпечення.</p> <p>ФК17. Здатність проводити дослідження різноманітних процесів, явищ та систем з використанням створеного спеціалізованого програмного забезпечення та проводити інтерпретацію отриманих результатів.</p> <p>ФК18. Здатність використовувати прикладні наукоємні інформаційні технології, в тому числі в рамках «інтернету речей», для вирішення спеціалізованих задач водного господарства, природокористування, екології, агросектору, охорони навколишнього середовища, геосистем.</p>
<p>Програмні результати навчання</p>	<p>ПРН01. Аналізувати, цілеспрямовано шукати і вибирати необхідні для вирішення професійних завдань інформаційно-довідникові ресурси і знання з урахуванням сучасних досягнень науки і техніки.</p> <p>ПРН05. Знати і застосовувати відповідні математичні</p>

	<p>поняття, методи доменного, системного об'єктно-орієнтованого аналізу та математичного моделювання для розробки програмного забезпечення.</p> <p>ПРН18. Знати та вміти застосовувати інформаційні технології обробки, зберігання та передачі даних.</p>
<p>Перелік соціальних, «м'яких» навичок (soft skills)</p>	<p><i>Відкритість, взаємодія з людьми, вміння працювати в команді, здатність до навчання, здатність логічно обґрунтовувати позицію, знаходити вихід з складних ситуацій, знаходити час на відпочинок, комунікаційні якості, навички міжособистісних відносин, навички усного спілкування, саморозвиток, творчі здібності, чесність.</i></p>
<p>Структура навчальної дисципліни</p>	<p>Лекції – 20 год., лабораторні – 28 год., Самостійна робота – 72 год.</p> <p style="text-align: center;">МОДУЛЬ 1 ЗМІСТОВИЙ МОДУЛЬ 1 Основні поняття ІАД. Класифікація методів та задач ІАД. Елементи факторного аналізу</p> <p>Вступ. Предмет, історія виникнення та задачі інтелектуального аналізу даних.</p> <p>Тема 1. Основні поняття, сфери застосування, методи і стадії інтелектуального аналізу даних. Поняття Data Mining. Особливості технології Data Mining та її відмінності від інших методів аналізу даних. Застосування Data Mining у бізнесі, торгівлі, банківській справі, страхуванні, медицині, науці тощо. Статистичні та кібернетичні методи Data Mining. Стадії Data Mining.</p> <p>Тема 2. Дані та їх типи. Бази даних. СУБД. Класифікація задач ІАД. Дані та їх атрибути. Набори даних. Типи даних. Шкали. Якісний аналіз даних з використанням Data Mining. Бази даних. СУБД. Класифікація задач ІАД.</p> <p>Тема 3. Статистичні методи ІАД. Факторний аналіз. Метод головних компонент. Класичні статистичні методи інтелектуального аналізу даних. Ідея факторного аналізу. Модель компонентного аналізу. Метод головних компонент та сфери його застосування.</p> <p style="text-align: center;">ЗМІСТОВИЙ МОДУЛЬ 2 Задачі інтелектуального аналізу даних. Методи і підходи до їх розв'язання</p> <p>Тема 4. Задачі класифікації та кластеризації, прогнозування та візуалізації. Задачі та види класифікації. Методи, що застосовуються для розв'язання задач класифікації. Задача кластеризації. Основні підходи до кластеризації. Задачі прогнозування. Приклади.</p>

Прогнозування і часові ряди. Тренд, сезонність і цикл. Види помилок та прогнозів. Візуалізація інструментів Data Mining. Методи візуалізації. Принципи компонування візуальних засобів.

Тема 5. Методи дерев рішень, класифікації та прогнозування.

Метод дерев рішень та його переваги. Алгоритми дерев рішень. Метод опорних векторів. Лінійний SVM. Метод «найближчого сусіда». Байєсівська класифікація.

Тема 6. Методи кластерного аналізу. Методи пошуку асоціативних правил.

Методи кластерного аналізу. Ієрархічний кластерний аналіз. Алгоритми неієрархічної кластеризації. Задача пошуку асоціативних правил. Характеристики і методи пошуку асоціативних правил. Приклади розв'язання задач.

Тема 7. Штучні нейронні мережі для розв'язання задач ІАД.

Поняття штучної нейронної мережі. Модель нейрону. Архітектура НМ. Основні парадигми навчання НМ. Нейронні мережі із зворотним розповсюдженням помилки. Закон навчання Хебба. Конкурентне навчання. Карти Кохонена. Нейронні мережі Хопфільда та Хемінга.

Тема 8. Методи прийняття рішень на основі нечітких множин та нечіткої логіки.

Основи теорії нечітких множин. Методи нечіткої логіки. Нечіткі експертні системи. Етапи побудови нечіткого логічного виводу. Нечіткі алгоритми Мамдані, Цукамото, Такагі-Сугено та ін. Нейро-нечіткі технології.

Тема 9. Генетичні алгоритми.

Основні поняття генетичних алгоритмів. Класичний генетичний алгоритм та його модифікації. Види еволюційних алгоритмів: еволюційні стратегії, еволюційне програмування, генетичні алгоритми, генетичне програмування. Гібридні системи.

Теми лабораторних робіт

Розв'язання задачі факторного аналізу даних методом головних компонент.

Розв'язання задачі класифікації за допомогою нейронних мереж.

Застосування НМ для побудови регресійної моделі.

Побудова регресійної моделі і прогнозування часового ряду за допомогою НМ.

Застосування НМ для розв'язання задачі класифікації даних часового ряду.

Кластерний аналіз за допомогою НМ Кохонена.

Розв'язання задачі класифікації за допомогою дерева рішень.

Проектування нечіткої експертної системи Мамдані для побудови регресійної залежності вигляду $y=f(x_1, x_2)$.
Застосування генетичних алгоритмів для розв'язання задачі оптимізації функції однієї змінної.
Застосування генетичних алгоритмів для розв'язання задачі оптимізації функції кількох змінних.

Завдання для самостійної роботи

Історія виникнення та причини розвитку Data Mining.
Сфери застосування технології Data Mining.
Класифікація систем Data Mining.
Задачі інтелектуального аналізу даних.
Типи даних. Формати збереження даних.
Рівні аналізу даних.
Інформація. Властивості інформації.
Задачі класифікації. Методи і підходи до класифікації.
Методи пошуку асоціативних правил.
Дерева рішень. Метод опорних векторів.
Метод "найближчого сусіда".
Байєсовська класифікація.
Кластеризація даних з використанням методів Data Mining.
Ієрархічна кластеризація.
Неієрархічна кластеризація.
Порівняння ієрархічних і неієрархічних методів кластеризації.
Аналіз часових рядів.
Методи прогнозування даних часового ряду.
Використання нейронних мереж для розв'язання задач інтелектуального аналізу даних.
Карти Кохонена.
Методи прийняття рішень на основі нечітких множин та нечіткої логіки.
Класичний генетичний алгоритм і його модифікації.
Еволюційні алгоритми.
Процес Data Mining. Дублювання даних. Очищення даних.

Методи оцінювання та структура оцінки

Оцінювання навчальних досягнень студентів за усіма видами навчальних робіт проводиться за поточним та підсумковим контролем. Поточний контроль знань та контроль самостійної роботи студентів з навчальної дисципліни проводиться:

- за допомогою тестів: контрольні завдання включають тестові питання трьох рівнів складності;
- з лабораторних робіт – за допомогою перевірки виконаних завдань та теоретичної підготовки до занять.

Усі контрольні заходи включено до 100-бальної шкали оцінювання.

Підсумковий семестровий контроль знань відбувається у вигляді заліку, є обов'язковим і проводиться у формі тестування лише для тих студентів, які, з тих чи інших причин, не пройшли модульний контроль. Основними критеріями, що характеризують рівень компетентності студента при оцінюванні результатів поточного та підсумкового контролю з навчальної дисципліни

	<p>«Інтелектуальний аналіз даних» є:</p> <ul style="list-style-type: none"> – виконання всіх видів навчальної роботи, що передбачені робочою програмою навчальної дисципліни; – глибина і характер знань навчального матеріалу за змістом навчальної дисципліни. – характер відповідей на поставлені питання (чіткість, лаконічність, логічність, послідовність тощо); – обґрунтування вибору методу для розв'язання задачі; – рівень вміння аналізувати одержані результати. <p>Оцінювання результатів усіх форм контролю передбачено у 100-бальній шкалі.</p> <p>Критерії оцінювання результатів поточної роботи (завдань, що виконуються на практичних заняттях, результати самостійної роботи студентів) проводиться у % від кількості балів, виділених на завдання, із заокругленням до цілого числа:</p> <ul style="list-style-type: none"> 0% – завдання не виконано; 40% – завдання виконано частково та містить суттєві помилки методичного або розрахункового характеру; 60% – завдання виконано повністю, але містить суттєві помилки у розрахунках або в методиці; 80% – завдання виконано повністю і вчасно, проте містить окремі несуттєві недоліки (розмірності, висновки, оформлення тощо); 100% – завдання виконано правильно, вчасно і без зауважень.
<p>Місце навчальної дисципліни в освітній траєкторії здобувача вищої освіти</p>	<p>Вивчення курсу передбачає наявність систематичних і ґрунтовних знань із курсів «Математичний аналіз», «Комп'ютерна дискретна математика», «Програмування», «Методи обчислень», «Теорія ймовірностей та математична статистика», «Математична логіка, теорія алгоритмів і програмування».</p> <p>Дисципліни, що вивчаються супутньо із зазначеною дисципліною: «Теорія систем, системний аналіз та теорія прийняття рішень», «Математичне та комп'ютерне моделювання» та ін.</p>
<p>Інформаційні ресурси</p>	<ol style="list-style-type: none"> 1. Чубукова І.А. <i>Data Mining</i> – М. : Бином ЛБЗ, 2008. – 384 с. 2. <i>Методы и модели анализа данных OLAP и Data Mining</i> / [А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод]. – СПб : БВХ-Петербург, 2004. – 336 с. 3. Дранишников Л.В. <i>Інтелектуальні методи в управлінні: навчальний посібник</i> – Кам'янське: ДДТУ, 2018. – 416 с. 4. Грицюк П. М., Остапчук О.П. <i>Аналіз даних: навчальний посібник</i> – Рівне: НУВГП, 2008. – 218 с. 5. Зайченко Ю.П. <i>Нечеткие модели и методы в интеллектуальных системах. Учебное пособие для студентов ВУЗов.</i> – К.: Издательский дом "Слово", 2008. – 344 с. 6. Ульянов С., Литвинцева Л., Добрынин В, Мишин А. <i>Интеллектуальное робастное управление:</i>

технологии мягких вычислений. — 1-е изд. — М: PronetLabs, 2011. — 406 с.

7. Ситник В. Ф., Краснюк М. Т. Интеллектуальный анализ данных (дейтамайнінг): Навч. посіб. — К.: КНЕУ, 2007. — 376 с.
8. Зайченко Ю.П. Основы проектирования интеллектуальных систем. Навчальний посібник. — К.: Видавничий Дім “Слово”, 2004. — 352 с.
9. Ситник С. Л. Лекции. “Основы проектирования систем искусственного интеллекта”.
10. Рутковская Д., Пилинський М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И.Д. Рудинского. — М.: Горячая линия – Телеком, 2006. — 452 с.: ил.
11. Джонс М.Т. Программирование искусственного интеллекта в приложениях: Пер. с англ. — М.: ДМК Пресс, 2004. — 312 с.: ил.
12. Кононюк А.Ю. Нейронні мережі і генетичні алгоритми / А.Ю. Кононюк. — К.: ПП Корнійчук. — 2008. — 160 с.
13. Хайкин Саймон. Нейронные сети: полный курс, 2-е изд.: Пер. с англ. — М.: ООО «И. Д. Вильямс», 2006. — 1104 с.
14. Тимощук П.В. Штучні нейронні мережі: навч. посібник. — Львів: Видавництво Львівської політехніки, 2011. — 444 с.
15. Ротштейн А.П. Интеллектуальные технологии идентификации: нечеткая логика, генетические алгоритмы, нейронные сети. — Винница: УНИВЕРСУМ—Винница, 1999. — 320 с.
16. Нейронные сети. Statistica Neural Networks: Методология и технологии современного анализа данных / Под редакцией В. П. Боровикова. - 2-е изд. перераб. и доп. - М.: Горячая линия - Телеком, 2008. - 392 с.

ПРАВИЛА ТА ВИМОГИ (ПОЛІТИКА)*

Дедлайни та перескладання

Завдання до лабораторних та самостійних робіт з відповідної теми повинні бути виконані і здані на оцінювання протягом 14 днів з дати заняття. У випадку порушення термінів кількість балів знижується на 10%.

Кінцевим терміном здачі завдань є останній робочий день навчального семестру.

Порядок повторного проходження контрольних заходів у НУВГП врегульовано «Положенням про семестровий поточний та підсумковий контроль навчальних досягнень здобувачів вищої освіти»: <http://ep3.nuwm.edu.ua/5040/>.

Усі перездачі проходять за погодженням з директором ННІ. Правила ННЦНО стосовно повторного тестування наведено у документах: <http://nuwm.edu.ua/struktturni-pidrozdzili/navch-nauk-tsentr-nezaleznoho-otsiniuvannia-znan/dokumenti>.

Перша перездача проводиться через ННЦНО згідно з розкладом перездач, який розміщено в додатку Мій НУВГП та ПС-Студент WEB: <http://desk.nuwm.edu.ua/cgi-bin/shell.cgi?n=999>.

	<p>У випадку отримання незадовільної оцінки, здобувач направляє на комісію з перездачі дисципліни, яка формується деканатом ННІ. Після трьох невдалих спроб здачі семестрового підсумкового контролю з навчальної дисципліни вважається, що здобувач має академічну заборгованість. Рішення про повторне вивчення навчальної дисципліни або відрахування здобувача приймає ректор на підставі звернення директора ННІ, як це передбачено «Порядком ліквідації академічних заборгованостей у НУВГП»: http://ep3.nuwm.edu.ua/id/eprint/4273.</p> <p>У випадку нездачі підсумкового контролю через хворобу чи з інших поважних причин, здобувач має написати заяву на ім'я директора ННІ для зміни строків сесії.</p>
<p>Правила академічної доброчесності</p>	<p>Викладач та здобувачі несуть спільну відповідальність за створення сприятливого творчого навчального середовища, яке базується на взаємній повазі.</p> <p>До кожного заняття здобувачі повинні наперед ознайомитися з матеріалами та інформаційними ресурсами, наведеними у методичних вказівках і розміщеними на сторінці дисципліни в Moodle.</p> <p>Здобувачі освіти повинні дотримуватися Кодексу честі студентів: http://nuwm.edu.ua/strukturni-pidrozdili/vvrsdev/dokumenti</p> <p>Принцип студентоцентризму передбачає розуміння серйозності ставлення до академічної недоброчесності та неправомірної поведінки. Студенти мають самостійно виконувати і здавати на оцінювання лише результати власних зусиль та оригінальної праці. При виконанні лабораторних робіт з дисципліни студентам рекомендується працювати в навчальних групах, порівнювати отримані результати та обговорювати застосовувані методи. Однак виконуючи поставлені завдання, студенти повинні індивідуально здійснити кожен розрахунок. Обмін виконаними завданнями чи їх частинами у формі тексту, таблиці, програмного коду чи у будь-якій іншій формі є недопустимим. Не існує прийнятого приводу для плагіату чи обману. Здобувачі освіти не можуть копіювати виконані завдання у інших студентів, ділитися виконаними завданнями з іншими студентами і мають дотримуватися Положення про виявлення та запобігання академічного плагіату в НУВГП http://nuwm.edu.ua/sp/akademichna-dobrochesnistj</p> <p>У випадку плагіату при виконанні завдання здобувач не отримує бали і повинен виконати завдання повторно.</p> <p>Перевірка дотримання доброчесності під час модульного та підсумкового контролю може здійснюватися засобами відеонагляду.</p> <p>Здобувачі можуть робити аудіозапис аудиторного заняття для свого особистого освітнього використання тільки за погодженням з викладачем і не мають права розміщувати такий запис в соціальних мережах.</p>
<p>Вимоги до</p>	<p>Здобувачі вищої освіти зобов'язані відвідувати усі лекційні та лабораторні заняття з дисципліни згідно розкладу</p>

відвідування	<p>http://desk.nuwm.edu.ua/cgi-bin/timetable.cgi Відвідування консультацій не обов'язкове.</p> <p>У випадку відсутності з поважних причин (індивідуальний план, лікарняний, мобільність тощо) здобувач самостійно опрацьовує теоретичний матеріал і виконує завдання з відповідної практичної роботи.</p> <p>Завдання до лабораторних робіт розміщено на сайті університету за посиланням: http://ep3.nuwm.edu.ua/2164/</p> <p>Файл (файли) із виконаними розрахунками здобувач прикріплює до відповідних завдань на платформі Moodle або надсилає викладачу на електронну пошту для перевірки. Захист роботи відбувається на наступному занятті, консультації або онлайн у відеорежимі.</p> <p>На лекціях і лабораторних заняттях студенти можуть використовувати свої ноутбуки, планшети чи смартфони для роботи.</p>
Неформальна та інформальна освіта	<p>Визнання (перезарахування) результатів навчання, здобутих у неформальній та інформальній освіті, відбувається відповідно до «Положення про неформальну та інформальну освіту в НУВГП»: http://nuwm.edu.ua/sp/neformalna-osvita</p> <p>Здобувачі можуть пройти відкриті онлайн курси, близькі за темою до даної навчальної дисципліни, таких платформ як Coursera, Prometheus, edEx, edEra, VUMOnline, FutureLearn тощо.</p>
ДОДАТКОВО	
Правила отримання зворотної інформації про дисципліну*	<p>Здобувач має право звертатися до викладача за додатковим поясненням матеріалу теми, змісту завдань лабораторних робіт та самостійної роботи протягом семестру усно (під час занять чи консультацій), корпоративною електронною поштою або через систему повідомлень Moodle.</p> <p>Консультації можуть проводитися онлайн із застосуванням сервісу Google Hangouts Meet.</p> <p>Здобувачі вищої освіти можуть подавати свої критичні зауваження, а також ідеї та рекомендації щодо наповнення навчальної дисципліни і методів викладання шляхом анонімного онлайн анкетування через Google Forms, яке проводиться наприкінці кожного семестру.</p> <p>Незалежне оцінювання якості викладання проводиться Відділом якості освіти. http://nuwm.edu.ua/struktturni-pidrozdili/vyo/dokumenti</p>
Оновлення*	<p>Викладач періодично оновлює зміст даної навчальної дисципліни на основі сучасних практик та опитування випускників кафедри</p>
Навчання осіб з інвалідністю	<p>Організація навчання людей з інвалідністю проводиться за дотриманням вимог нормативних документів, розроблених в НУВГП: http://nuwm.edu.ua/sp/dlja-osib-z-invalidnistju</p>

Лектор

Демчук Олена Стеніславівна, к.т.н.