



Вступ до інтелектуального аналізу даних

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Перший (бакалаврський)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>122 Комп'ютерні науки</i>
Освітня програма	<i>Системи і методи штучного інтелекту</i>
Статус дисципліни	<i>Нормативна</i>
Форма навчання	<i>очна(денна)</i>
Рік підготовки, семестр	<i>2 курс, весняний семестр</i>
Обсяг дисципліни	<i>120 годин (4 кредита ЄКТС)</i>
Семестровий контроль/ контрольні заходи	<i>Залік</i>
Розклад занять	<i>Лекції та практики кожного тижня</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	<i>д.т.н., професор, Терент'єв Олександр Миколайович, пошта: o.terentiev@gmail.com</i>
Розміщення курсу	<i>Moodle: https://do.ipu.kpi.ua/course/view.php?id=2708</i>

Програма навчальної дисципліни

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

Предмет вивчення дисципліни становлять, головним чином, методи, підходи та технології щодо видобування даних, їхньої попередньої обробки та подальшого аналізу різноманітними методами інтелектуального аналізу даних.

*Вивчення навчальної дисципліни націлено на формування, розвиток та закріплення у здобувачів таких загальних та фахових **компетентностей**: ЗК 1 Здатність до абстрактного мислення, аналізу та синтезу, ЗК 7 Здатність до пошуку, оброблення та аналізу інформації з різних джерел, ЗК 11 Здатність приймати обґрунтовані рішення, ФК 2 Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо, ФК 11 Здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач, ФК 22 Здатність використовувати мови штучного інтелекту при розробці програмного забезпечення інтелектуальних інформаційних систем, здатність орієнтуватися в різних типах інтелектуальних систем і технологій; ставити завдання побудови інтелектуальних систем для вирішення завдання вибору варіантів в проблемній області, що погано формалізується.*

*По завершенню курсу студент має набути наступні **програмні результати навчання**: ПР 4 Використовувати методи обчислювального інтелекту, машинного навчання, нейромережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв'язання задач розпізнавання, прогнозування, класифікації, ідентифікації об'єктів керування тощо, ПР 12 Застосовувати методи та алгоритми обчислювального інтелекту та*

інтелектуального аналізу даних в задачах класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірної аналізу даних на основі технологій DataMining, TextMining, WebMining, ПР 19 Розробляти системи розпізнавання образів та класифікації в різних предметних областях, обґрунтовано вибирати та використовувати алгоритми розпізнавання образів та проводити навчання систем розпізнавання образів, ПР 24 Розуміти різні типи інтелектуальних систем і технологій; ставити завдання побудови інтелектуальних систем для вирішення завдання вибору варіантів в проблемній області, що погано формалізується.

Після засвоєння кредитного модуля студенти мають продемонструвати такі результати навчання:

знання:

- основних понять, задач, стадій інтелектуального аналізу даних;
- підходів до збереження, представлення даних у сучасних інформаційних системах;
- аналізу залежностей у даних;
- методи та інструменти інтелектуального аналізу даних;
- сучасні програмні засоби, застосовувані для інтелектуального аналізу даних;
- концепції сховищ даних, підходи до їх аналітичної обробки.

уміння:

- обґрунтовувати вибір конкретного типу моделі та методу інтелектуального аналізу даних при вирішенні поставленої практичної задачі;
- проводити необхідну попередню обробку даних, визначати тип задачі аналізу, вирішувати її адекватно обраним методом з оптимально визначеними параметрами, оцінювати результати, робити змістовні висновки та інтерпретацію;
- використовувати сучасні програмні засоби для проектування та дослідження систем інтелектуального аналізу даних;
- застосовувати технології роботи зі сховищами даних, здійснювати їх аналітичну обробку та інтелектуальний аналіз для забезпечення надійної роботи інформаційних систем;

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

Міждисциплінарні зв'язки: вивчення даної дисципліни забезпечується дисциплінами обов'язкового компоненту циклів загальної та професійної підготовки: «теорія ймовірностей», «математична логіка і теорія алгоритмів» та «математичний аналіз». Вивчення дисципліни «Вступ до інтелектуального аналізу даних» забезпечує засвоєння дисциплін нормативного блоку таких, як «Інтелектуальний аналіз даних» та «Розпізнавання образів», «Методи та системи штучного інтелекту», тощо.

3. Зміст навчальної дисципліни

Перелік розділів і тем дисципліни
Тема 1. Основи інтелектуального аналізу даних. Визначення Data Mining і область застосування. Задачі, моделі та методи Data Mining. Методи, стадії, задачі Data Mining. Поняття Business Intelligence.
Тема 2. Інтелектуальний аналіз даних у бізнесі Приклади застосування у різних сферах. Застосування інтелектуального аналізу для вирішення складних задач.
Тема 3. Етапи інтелектуального аналізу даних. Загальні закономірності при аналізі даних. Групи задач аналізу даних. Класифікація методів аналізу.
Тема 4. Методи аналізу часових рядів Поняття нечітких часових рядів. Методи

<i>моделювання часових рядів. Методи аналізу та прогнозування поведінки часових рядів.</i>
<i>Тема 5. Вирішення задачі кластеризації Постановка задачі кластеризації та представлення результатів. Види кластерів. Міри близькості, засновані на відстанях. Базові алгоритми кластеризації. Адаптивні методи кластеризації.</i>
<i>Тема 6. Класифікація. Постановка задачі класифікації. Контрольована непараметрична класифікація. Контрольована непараметрична нейромережева класифікація. Класифікація за методом опорних векторів. Деревя рішень. Неконтрольована класифікація.</i>
<i>Тема 7. Регресія. Поняття регресії. Основні етапи регресійного аналізу. Методи відновлення регресії.</i>
<i>Тема 8. Вирішення задачі пошуку асоціативних правил Постановка задачі пошуку асоціативних правил та представлення результатів. Секвенціальний аналіз. Різновиди задач пошуку асоціативних правил. Методи подання результатів. Алгоритми пошуку асоціативних правил. Методи пошуку асоціативних правил: метод Apriori, побудова FP-дерев пошуку шаблонів даних. Min-tax асоціації у базах даних. Побудова hash-дерев.</i>
<i>Тема 9. Застосування нейронних мереж. Самоорганізаційна карта Кохонена. Початкова ініціалізація карти. Навчання. Застосування алгоритму.</i>
<i>Тема 10. Застосування генетичних алгоритмів. Основні визначення. Схема генетичного алгоритму. Генетичні операції. Представлення даних. Відбір.</i>
<i>Тема 11. Способи візуального представлення даних. Методи візуалізації. Представлення даних в одному, двох та трьох вимірах. Паралельні координати. Якість візуалізації. Основні тенденції в області візуалізації.</i>
<i>Тема 12. Очистка та предобробка даних. Оцінка якості даних. Технології та методи оцінки якості даних. Очистка та предобробка. Фільтрація даних. Відновлення пропусків даних. Скорочення розмірності. Скорочення кількості ознак та записів. Семпінг.</i>
<i>Тема 13. Ансамблі моделей. Вступ. Беггінг. Бустінг. Альтернативні методи побудови ансамблів.</i>
<i>Тема 14. Порівняння моделей. Оцінка ефективності та порівняння моделей. Оцінка помилки моделі. ROC-аналіз. Навчання в умовах незбалансованості класів.</i>
<i>Тема 15. Оперативний аналіз даних Розгортання OLAP-кубів. Операції над OLAP-кубами (зріз, обертання, консолідація, деталізація). Архітектура OLAP-систем: MOLAP, ROLAP, HOLAP.</i>
<i>Тема 16. Сховища даних Визначення сховища даних, порівняння з базами даних, використання.. Архітектура сховища даних. ETL-процеси (добування, перетворення й завантаження даних). Вітрини даних, куби даних, багатомірна модель даних</i>
<i>Тема 17. Застосування SAS Enterprise Miner. Спеціалізоване сховище даних. Підхід SAS до створення інформаційно-аналітичних систем. Технічні вимоги пакета SASR Enterprise Miner</i>
<i>Тема 18. Інструменти Oracle Data Mining та Deductor. Oracle Data Mining. Прогнозуючі моделі. Дескрипторні моделі. Аналітична платформа Deductor.</i>

4. Навчальні матеріали та ресурси

Базова література:

1. Бідюк П.І. Прикладна статистика (навчальний посібник) / П.І. Бідюк, О.М. Терентьєв, Т.І. Просянкін-Жарова. – Вінниця : ПП “ТД Едельвейс і К”, 2013. – 288 с. – ISBN 978-966-2462-21-0. – Режим доступу: <https://ela.kpi.ua/handle/123456789/19563>
2. Згуровський М.З. Байєсівські мережі в системах підтримки прийняття рішень (навчальний посібник) / Згуровський М.З., Бідюк П.І., Терентьєв О.М., Просянкін-

Жарова Т.І. – К: ТОВ “Видавниче підприємство Едельвейс”, 2015. – 300 с. – ISBN 978-966-2748-73-4. – Режим доступу: <https://ela.kpi.ua/handle/123456789/19582>

3. Mohammed J. Zaki, Wagner Meira, Jr., *Data Mining and Machine Learning: Fundamental Concepts and Algorithms*, 2nd Edition, Cambridge University Press, March 2020. ISBN: 978-1108473989. – Режим доступу: https://dataminingbook.info/book_html

Додаткова література:

1. Сергеев-Горчинський О. О., Іщенко Г. В. *Інтелектуальний аналіз даних (комп'ютерний практикум) [Електронний ресурс] : навч. посіб. для студ. спеціальності 122 «Комп'ютерні науки та інформаційні технології», спеціалізацій «Інформаційні системи та технології проектування», «Системне проектування сервісів» / О. О. Сергеев-Горчинський, Г. В. Іщенко ; КПІ ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 1,72 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2018. – 73 с. – Режим доступу: https://ela.kpi.ua/bitstream/123456789/24971/1/Komp_prakt.pdf*
2. O. M. Trofymchuk, P. I. Bidyuk, T. I. Prosyankina-Zharova, O. M. Terentiev. *Decision Support Systems for Modeling, Forecasting and Risk Estimation (Monograph)*. – LAP LAMBERT Academic Publishing, 2019. – 176 p. – ISBN-13: 978-3-659-61214-5. – ISBN-10: 3659612146. – EAN: 9783659612145. – Режим доступу: <https://www.morebooks.shop/store/gb/book/decision-support-systems-for-modeling,-forecasting-and-risk-estimation/isbn/978-3-659-61214-5>
3. Pang-Ning Tan Michael Steinbach Vipin Kumar. *Introduction to Data Mining*. Pearson Education Limited 2014. – 719 p. – ISBN 10: 1-292-02615-4. – ISBN 13: 978-1-292-02615-2. – Режим доступу: https://www.ceom.ou.edu/media/docs/upload/Pang-Ning_Tan_Michael_Steinbach_Vipin_Kumar_-_Introduction_to_Data_Mining-Pe_NRDk4fi.pdf

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

Лекції, практичні заняття, виконання аналітично-розрахункових завдань, дискусії, аналіз конкретних ситуацій, командна робота, модульна контрольна робота.

Структура кредитного модуля

Назви розділів і тем	Кількість годин				
	Всього	у тому числі			
		Лекції	Практичні (семінарські)	Лабораторні	СРС
1	2	3	4	5	6
«Вступ до інтелектуального аналізу даних»					
Тема 1. Основи інтелектуального аналізу даних. Визначення Data Mining і область застосування. Задачі, моделі та методи Data Mining. Методи, стадії, задачі Data Mining. Поняття Business Intelligence.	5	2	1		2
Тема 2. Інтелектуальний аналіз даних у бізнесі Приклади застосування у різних сферах. Застосування інтелектуального аналізу для вирішення складних задач.	5	2	1		2
Тема 3. Етапи інтелектуального аналізу даних. Загальні	5	2	1		2

закономірності при аналізі даних. Групи задач аналізу даних. Класифікація методів аналізу.					
Тема 4. Методи аналізу часових рядів Поняття нечітких часових рядів. Методи моделювання часових рядів. Методи аналізу та прогнозування поведінки часових рядів.	5	2	1		2
Тема 5. Вирішення задачі кластеризації Постановка задачі кластеризації та представлення результатів. Види кластерів. Міри близькості, засновані на відстанях. Базові алгоритми кластеризації. Адаптивні методи кластеризації.	5	2	1		2
Тема 6. Класифікація. Постановка задачі класифікації. Контрольована непараметрична класифікація. Контрольована непараметрична нейромережева класифікація. Класифікація за методом опорних векторів. Дерева рішень. Неконтрольована класифікація.	5	2	1		2
Тема 7. Регресія. Поняття регресії. Основні етапи регресійного аналізу. Методи відновлення регресії.	5	2	1		2
Тема 8. Вирішення задачі пошуку асоціативних правил Постановка задачі пошуку асоціативних правил та представлення результатів. Секвенціальний аналіз. Різновиди задач пошуку асоціативних правил. Методи подання результатів. Алгоритми пошуку асоціативних правил. Методи пошуку асоціативних правил: метод Apriori, побудова FP-дерев пошуку шаблонів даних. Min-тах асоціації у базах даних. Побудова hash-дерев.	5	2	1		2
Тема 9. Застосування нейронних мереж. Самоорганізаційна карта Кохонена. Початкова ініціалізація карти. Навчання. Застосування алгоритму.	5	2	1		2
Тема 10. Застосування генетичних алгоритмів. Основні визначення. Схема генетичного алгоритму. Генетичні операції. Представлення даних. Відбір.	5	2	1		2
Тема 11. Способи візуального представлення даних. Методи візуалізації. Представлення даних в одному, двох та трьох вимірах. Паралельні координати. Якість візуалізації. Основні тенденції в області візуалізації.	5	2	1		2
Тема 12. Очистка та предобробка даних. Оцінка якості даних. Технології та методи оцінки якості	5	2	1		2

даних. Очистка та предобробка. Фільтрація даних. Відновлення пропусків даних. Скорочення розмірності. Скорочення кількості ознак та записів. Семпінг.					
Тема 13. Ансамблі моделей. Вступ. Бегінг. Бустінг. Альтернативні методи побудови ансамблів.	5	2	1		2
Тема 14. Порівняння моделей. Оцінка ефективності та порівняння моделей. Оцінка помилки моделі. ROC-аналіз. Навчання в умовах незбалансованості класів.	5	2	1		2
Тема 15. Оперативний аналіз даних Розгортання OLAP-кубів. Операції над OLAP-кубами (зріз, обертання, консолідація, деталізація). Архітектура OLAP-систем: MOLAP, ROLAP, HOLAP.	5	2	1		2
Тема 16. Сховища даних Визначення сховища даних, порівняння з базами даних, використання. Архітектура сховища даних. ETL-процеси (добування, перетворення й завантаження даних). Вітрини даних, куби даних, багатомірна модель даних	5	2	1		2
Тема 17. Застосування SAS Enterprise Miner. Спеціалізоване сховище даних. Підхід SAS до створення інформаційно-аналітичних систем. Технічні вимоги пакета SASR Enterprise Miner	5	2	1		2
Тема 18. Інструменти Oracle Data Mining та Deductor. Oracle Data Mining. Прогнозуючі моделі. Дескрипторні моделі. Аналітична платформа Deductor.	5	2	1		2
Разом за розділом	90	36	18		36
<i>Залік</i>					
Всього годин	90	36	18		36

Лекційні заняття

№ з/п	Назва теми лекції та перелік основних питань
1.	Основи інтелектуального аналізу даних. Визначення Data Mining і область застосування. Задачі, моделі та методи Data Mining. Методи, стадії, задачі Data Mining. Поняття Business Intelligence.
2.	Інтелектуальний аналіз даних у бізнесі Приклади застосування у різних сферах. Застосування інтелектуального аналізу для вирішення складних задач.
3.	Етапи інтелектуального аналізу даних. Загальні закономірності при аналізі даних. Групи задач аналізу даних. Класифікація методів аналізу.
4.	Методи аналізу часових рядів Поняття нечітких часових рядів. Методи моделювання часових рядів. Методи аналізу та прогнозування поведінки часових рядів.

5.	Вирішення задачі кластеризації Постановка задачі кластеризації та представлення результатів. Види кластерів. Міри близькості, засновані на відстанях. Базові алгоритми кластеризації. Адаптивні методи кластеризації.
6.	Класифікація. Постановка задачі класифікації. Контрольована непараметрична класифікація. Контрольована непараметрична нейромережева класифікація. Класифікація за методом опрних векторів. Дерева рішень. Неконтрольована класифікація.
7.	Регресія. Поняття регресії. Основні етапи регресійного аналізу. Методи відновлення регресії.
8.	Вирішення задачі пошуку асоціативних правил Постановка задачі пошуку асоціативних правил та представлення результатів. Секвенціальний аналіз. Різновиди задач пошуку асоціативних правил. Методи подання результатів. Алгоритми пошуку асоціативних правил. Методи пошуку асоціативних правил: метод Apriori, побудова FP-дерев пошуку шаблонів даних. Min-max асоціації у базах даних. Побудова hash-дерев.
9.	Застосування нейронних мереж. Самоорганізація карта Кохонена. Початкова ініціалізація карти. Навчання. Застосування алгоритму.
10.	Застосування генетичних алгоритмів. Основні визначення. Схема гентичного алгоритму. Генетичні операції. Представлення даних. Відбір.
11.	Способи візуального представлення даних. Методи візуалізації. Представлення даних в одному, двох та трьох вимірах. Паралельні координати. Якість візуалізації. Основні тенденції в області візуалізації.
12.	Очистка та предобробка даних. Оцінка якості даних. Технології та методи оцінки якості даних. Очищення та предобробка. Фільтрація даних. Відновлення пропусків даних. Скорочення розмірності. Скорочення кількості ознак та записів. Семпінг.
13.	Ансамблі моделей. Вступ. Бегінг. Бустінг. Альтернативні методи побудови ансамблів.
14.	Порівняння моделей. Оцінка ефективності та порівняння моделей. Оцінка помилки моделі. ROC-аналіз. Навчання в умовах незбалансованості класів.
15.	Оперативний аналіз даних Розгортання OLAP-кубів. Операції над OLAP-кубами (зріз, обертання, консолідація, деталізація). Архітектура OLAP-систем: MOLAP, ROLAP, HOLAP.
16.	Сховища даних Визначення сховища даних, порівняння з базами даних, використання.. Архітектура сховища даних. ETL-процеси (добування, перетворення й завантаження даних). Вітрини даних, куби даних, багатомірна модель даних
17.	Застосування SAS Enterprise Miner. Спеціалізоване сховище даних. Підхід SAS до створення інформаційно-аналітичних систем. Технічні вимоги пакета SASR Enterprise Miner
18.	Інструменти Oracle Data Mining та Deductor. Oracle Data Mining. Прогнозуючі моделі. Дескрипторні моделі. Аналітична платформа Deductor.

Практичні заняття

Основні завдання циклу практичних занять набути практичних навичок застосування методів інтелектуального аналізу даних

№ з/п	Назва теми заняття
1	Методи, стадії, задачі Data Mining. Поняття Business Intelligence. Застосування

	інтелектуального аналізу для вирішення задач у різних сферах
2	Етапи інтелектуального аналізу даних. Загальні закономірності при аналізі даних. Методи аналізу та моделювання часових рядів
3	Вирішення задачі кластеризації
4	Класифікація. Дерева рішень.
5	Вирішення задачі пошуку асоціативних правил
6	Застосування нейронних мереж
7	Способи візуального представлення даних. Методи візуалізації
8	Ансамблі моделей. Беггінг. Бустінг. Альтернативні методи побудови ансамблів.
9	Сховища даних, використання. Архітектура сховища даних

6. Самостійна робота студента/аспіранта

До самостійної роботи студентів відноситься підготовка до аудиторних занять, опанування тем, що винесені на самостійне вивчення, проведення обчислень та написання програм, за первинними даними, отриманими на практичних заняттях, підготовка до модульної контрольної роботи та заліку. Відповідно до навчального плану на самостійну роботу відводиться 66 годин.

Самостійна робота

№ з/п	Назви тем і питань, що виносяться на самостійне опрацювання та посилання на навчальну літературу	Кількість годин СРС
1	Основи інтелектуального аналізу даних. Визначення Data Mining і область застосування. Задачі, моделі та методи Data Mining. Методи, стадії, задачі Data Mining. Поняття Business Intelligence.	2
2	Інтелектуальний аналіз даних у бізнесі Приклади застосування у різних сферах. Застосування інтелектуального аналізу для вирішення складних задач.	2
3	Етапи інтелектуального аналізу даних. Загальні закономірності при аналізі даних. Групи задач аналізу даних. Класифікація методів аналізу.	2
4	Методи аналізу часових рядів Поняття нечітких часових рядів. Методи моделювання часових рядів. Методи аналізу та прогнозування поведінки часових рядів.	4
5	Вирішення задачі кластеризації Постановка задачі кластеризації та представлення результатів. Види кластерів. Міри близькості, засновані на відстанях. Базові алгоритми кластеризації. Адаптивні методи кластеризації.	4
6	Класифікація. Постановка задачі класифікації. Контрольована непараметрична класифікація. Контрольована непараметрична нейромережева класифікація. Класифікація за методом опрних векторів. Дерева рішень. Неконтрольована класифікація.	4
7	Регресія. Поняття регресії. Основні етапи регресійного аналізу. Методи відновлення регресії.	4
8	Вирішення задачі пошуку асоціативних правил Постановка задачі пошуку асоціативних правил та представлення результатів. Секвенціальний аналіз. Різновиди задач пошуку асоціативних правил. Методи подання результатів. Алгоритми пошуку асоціативних правил. Методи пошуку асоціативних правил: метод Apriori, побудова FP-дерев пошуку шаблонів даних. Min-max асоціації у базах даних. Побудова hash-дерев.	4
9	Застосування нейронних мереж. Самоорганізація карта Кохонена. Початкова ініціалізація карти. Навчання. Застосування алгоритму.	4
10	Застосування генетичних алгоритмів. Основні визначення. Схема	4

	гентичного алгоритму. Генетичні операції. Представлення даних. Відбір.	
11	Способи візуального представлення даних. Методи візуалізації. Представлення даних в одному, двох та трьох вимірах. Паралельні координати. Якість візуалізації. Основні тенденції в області візуалізації.	4
12	Очистка та предобробка даних. Оцінка якості даних. Технології та методи оцінки якості даних. Очищення та предобробка. Фільтрація даних. Відновлення пропусків даних. Скорочення розмірності. Скорочення кількості ознак та записів. Семпінг.	4
13	Ансамблі моделей. Вступ. Беггінг. Бустінг. Альтернативні методи побудови ансамблів.	4
14	Порівняння моделей. Оцінка ефективності та порівняння моделей. Оцінка помилки моделі. ROC-аналіз. Навчання в умовах незбалансованості класів.	4
15	Оперативний аналіз даних Розгортання OLAP-кубів. Операції над OLAP-кубами (зріз, обертання, консолідація, деталізація). Архітектура OLAP-систем: MOLAP, ROLAP, HOLAP.	4
16	Сховища даних Визначення сховища даних, порівняння з базами даних, використання. Архітектура сховища даних. ETL-процеси (добування, перетворення й завантаження даних). Вітрини даних, куби даних, багатомірні моделі даних	4
17	Застосування SAS Enterprise Miner. Спеціалізоване сховище даних. Підхід SAS до створення інформаційно-аналітичних систем. Технічні вимоги пакета SASR Enterprise Miner	4
18	Інструменти Oracle Data Mining та Deductor. Oracle Data Mining. Прогнозуючі моделі. Дескрипторні моделі. Аналітична платформа Deductor.	4

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Під час занять (як лекцій, так і практичних) студентам надаються інтерактивні вправи у вигляді задач, оцінюється як активність, так і успішність освоєння матеріалу.

(в тому числі під час виконання домашніх завдань і підготовки до лекцій), призначаються заохочувальні бали.

Всі завдання з описом відповідних етапів їх виконання, а також відповідних контрольних строків завантажені у систему Moodle дисципліни. Усі роботи студенти мають завантажувати в особистому кабінеті Moodle. Дедлайни кожного завдання позначені в завданнях у Moodle. Захисти індивідуальних робіт відбуваються на практичних заняттях. Роботи мають бути виконані з дотриманням академічної доброчесності. У період роботи в дистанційному режимі лектор може запропонувати студентам пройти запропоновані ним онлайн-курси на платформі Coursera. Сертифікати/проміжні результати цих курсів зараховуються бути частково зараховані згідно до Положення розміщеного на сайті КПІ ім. Ігоря Сікорського (https://document.kpi.ua/files/2020_7-124.pdf).

Крім того, передбачається, в межах вивчення навчальної дисципліни, участь в конференціях, форумах, круглих столах тощо.

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 студентів 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)

Семестрова атестація проводиться у виді заліку. Для оцінювання результатів навчання застосовується 100-бальна рейтингова система і університетська шкала оцінювання.

Рейтинг студента з кредитного модуля розраховується зі 100 балів, з них 52 бали складає стартова шкала. Стартовий рейтинг (протягом семестру) складається з балів, що студент отримує за:

- роботу на практичних заняттях (9 занять);
- виконання завдань індивідуального навчально-дослідного завдання.

Критерії нарахування балів:

1. Робота на практичних заняттях:

- активна творча робота – 4 бали;
- плідна робота – 3 бали.

2. Виконання індивідуальних навчально-дослідних завдань:

- творча робота – 16-15 балів;
- роботу виконано з незначними недоліками – 14-12 балів;
- роботу виконано з певними помилками – 11-10 балів;
- роботу не зараховано (завдання не виконане або є грубі помилки) – 0 балів.

3. Умовою першої атестації є отримання не менше 8 балів та виконання завдань всіх практичних робіт (на час атестації). Умовою другої атестації – отримання не менше 22 балів, виконання завдань всіх практичних робіт (на час атестації) та зарахування виконання індивідуальної роботи.

4. Умовою допуску до заліку є зарахування всіх практичних робіт, індивідуальної роботи та стартовий рейтинг не менше 26 балів.

5. На заліку студенти виконують письмову контрольну роботу. Кожне завдання містить два теоретичних запитання (завдання) і одне практичне. Перелік запитань наведений у Рекомендаціях до засвоєння кредитного модуля. Кожне запитання (завдання) оцінюється у 16 балів за такими критеріями:

- «відмінно», повна відповідь, не менше 90% потрібної інформації (повне, безпомилкове розв'язування завдання) – 16-15 балів;
- «добре», достатньо повна відповідь, не менше 75% потрібної інформації або незначні неточності (повне розв'язування завдання з незначними неточностями) – 14-12 балів;
- «задовільно», неповна відповідь, не менше 60% потрібної інформації та деякі помилки (завдання виконане з певними недоліками) – 11-10 балів;
- «незадовільно», відповідь не відповідає умовам до «задовільно» – 0 балів.

6. Сума стартових балів та балів за залікову контрольну роботу переводиться до залікової оцінки згідно з таблицею:

Бали: практичні заняття + ІДЗ + + залікова контрольна робота	Оцінка	Бали ECTS
100...95	Відмінно	A

94...85	Дуже добре	B
84...75	Добре	C
74...65	Задовільно	D
64...60	Достатньо	E
Менше 60	Незадовільно	FX
Є не зараховані практичні роботи або не зарахована розрахункова робота або стартовий рейтинг менше 26 балів	Не допущено	F

9. Додаткова інформація з дисципліни (освітнього компонента)

У студентів є можливість отримати позитивну оцінку по дисципліні за умов отримання сертифікатів проходження дистанційних чи онлайн курсів за відповідною тематикою.

У додатку до силабусу знаходиться перелік питань до заліку.

Робочу програму навчальної дисципліни (силабус):

Складено доцентом кафедри математичних методів системного аналізу, д.т.н., професором Терентьевим О.М.

Ухвалено кафедрою штучного інтелекту (протокол № 14 від 24. 05. 2023)

Погоджено Методичною комісією НН ІПСА (протокол № 4 від 16.06.2023)

ДОДАТОК А.

Перелік питань до заліку.

1. Модель як засіб дослідження об'єкту, системи, явища.
2. Основні підходи до побудови моделей даних
3. Властивості моделі
4. Суть та мета інтелектуального аналізу даних
5. Дослідження систем за аналітичного підходу
6. Інформаційний підхід до моделювання
7. В чому полягає концепція «моделей від даних»
8. Основні групи аналітичних задач
9. Способи візуалізації даних
10. Тиражування знань: переваги та недоліки
11. Етапи моделювання
12. Процес побудови моделі
13. Дані, їх види та форми їх представлення
14. Структуровані дані як основна форма представлення відомостей в базах даних
15. Типи структурованих даних
16. Особливості підготовки даних до аналізу
17. Формалізація даних та вимоги до них
18. Особливості технологій KDD та Data Mining
19. Сучасні аналітичні платформи
20. Методи та алгоритми data mining. Загальна характеристика.
21. Навчальна вибірка. Правила формування.
22. Навчання з учителем та без.
23. Помилки, що виникають у процесі навчання.
24. Перенавчання. Ефект перенавчання
25. Консолідація даних: джерела, задачі, схема процесу
26. Типи сховищ даних.
27. Особливості та концепції сховищ даних
28. Деталізовані та агреговані дані
29. Архітектура сховищ даних
30. Багатовимірні та гібридні сховища даних