

ЗАТВЕРДЖУЮ



Декан факультету інформаційних технологій

Тетяна ГОВОРУЩЕНКО
Ім'я, ПРІЗВИЩЕ

09 2025 р.

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Моделі та методи текстової аналітики

Назва дисципліни

Галузь знань – F Інформаційні технології
Спеціальність – F3 Комп'ютерні науки
Рівень вищої освіти – Другий (магістерський)
Освітньо-професійна програма – Комп'ютерні науки
Обсяг дисципліни – 5 кредитів ЄКТС, **Шифр дисципліни** – ОФП.03
Мова навчання – українська
Статус дисципліни: обов'язкова (фахової підготовки)
Факультет – Інформаційних технологій
Кафедра – Комп'ютерних наук

Форма здобуття освіти	Курс	Семестр	Загальний обсяг		Кількість годин						Курсовий проєкт	Курсова робота	Форма семестрового контролю	
					Аудиторні заняття					Самостійна робота, у т.ч. ІРС			Залік	Іспит
			Кредити ЄКТС	Години	Разом	Лекції	Лабораторні роботи	Практичні заняття	Семінарські заняття					
Д	1	1	5	150	50	16	34			100				+

Робоча програма складена на основі освітньо-професійної програми «Комп'ютерні науки» за спеціальністю F3 «Комп'ютерні науки»

Робоча програма складена


Підпис автора(ів)

канд. техн. наук, доц. Олександр МАЗУРЕЦЬ
Науковий ступінь, вчене звання, Ім'я, ПРІЗВИЩЕ автора(ів)

Схвалена на засіданні кафедри комп'ютерних наук

Протокол від 29.08.2025 № 1

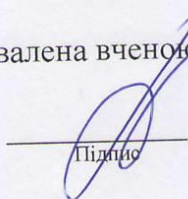
Зав. кафедри


Підпис

Олександр БАРМАК
Ім'я, ПРІЗВИЩЕ

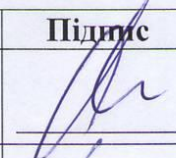
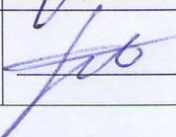
Робоча програма розглянута та схвалена вченою радою факультету інформаційних технологій

Голова вченої ради факультету


Підпис

Тетяна ГОВОРУЩЕНКО
Ім'я, ПРІЗВИЩЕ

2 ЛИСТ ПОГОДЖЕННЯ

Посада	Назва кафедри	Підпис	Ініціали, прізвище
Завідувач кафедри, д-р. техн. наук, проф.	Комп'ютерні науки		Олександр БАРМАК
Гарант освітньо-професійної програми, канд. техн. наук, доц.	Комп'ютерні науки		Руслан БАГРІЙ

3. Пояснювальна записка

Дисципліна «Моделі та методи текстової аналітики» є однією із дисциплін фахової підготовки і займає провідне місце у підготовці здобувачів другого (магістерського) рівня вищої освіти, очної (денної) (далі – денної) форми здобуття вищої освіти, які навчаються за освітньо-професійною програмою «Комп'ютерні науки» в межах спеціальності F3 «Комп'ютерні науки».

Пререквізити – Англійська мова за професійним спрямуванням (ОЗП.03), Проектування та розробка інформаційних систем (ОФП.01), Інженерія, аналітика та візуалізація великих даних (ОФП.02).

Постреквізити – Переддипломна практика (ОФП.06), Кваліфікаційна робота (ОФП.07)

Відповідно до освітньої програми дисципліна сприяє забезпеченню:

компетентностей: здатність розв'язувати задачі дослідницького та/або інноваційного характеру у сфері комп'ютерних наук (ІК); здатність до абстрактного мислення, аналізу та синтезу (ЗК01); здатність застосовувати знання у практичних ситуаціях (ЗК02); здатність спілкуватися державною мовою як усно, так і письмово (ЗК03); здатність вчитися і оволодівати сучасними знаннями (ЗК05); усвідомлення теоретичних засад комп'ютерних наук (ФК01); здатність формалізувати предметну область певного проєкту у вигляді відповідної інформаційної моделі (ФК02); здатність використовувати математичні методи для аналізу формалізованих моделей предметної області (ФК03); здатність збирати і аналізувати дані (включно з великими), для забезпечення якості прийняття проєктних рішень (ФК04); здатність застосовувати існуючі і розробляти нові алгоритми розв'язування задач у галузі комп'ютерних наук (ФК06); здатність розробляти програмне забезпечення відповідно до сформульованих вимог з урахуванням наявних ресурсів та обмежень (ФК07); здатність розробляти та застосовувати інтелектуальні моделі та методи для текстової аналітики (УК01).

програмних результатів навчання: мати спеціалізовані концептуальні знання, що включають сучасні наукові здобутки у сфері комп'ютерних наук і є основою для оригінального мислення та проведення досліджень, критичне осмислення проблем у сфері комп'ютерних наук та на межі галузей знань (ПРН01); 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 (ПРН02); розробляти концептуальну модель інформаційної або комп'ютерної системи (ПРН06); розробляти та застосовувати математичні методи для аналізу інформаційних моделей (ПРН07); розробляти математичні моделі та методи аналізу даних (включно з великим) (ПРН08); розробляти алгоритмічне та програмне забезпечення для аналізу даних (включно з великими) (ПРН09); створювати нові алгоритми розв'язування задач у сфері комп'ютерних наук, оцінювати їх ефективність та обмеження на їх застосування (ПРН11); тестувати програмне забезпечення (ПРН14); виконувати дослідження у сфері комп'ютерних наук (ПРН16); збирати, формалізувати, систематизувати і аналізувати потреби та вимоги до інформаційної або комп'ютерної системи, що розробляється, експлуатується чи супроводжується (ПРН18); аналізувати сучасний стан і світові тенденції розвитку комп'ютерних наук та інформаційних технологій (ПРН19); розробляти, застосовувати та вдосконалювати інтелектуальні моделі та методи текстової аналітики для вирішення спеціалізованих задач у сфері комп'ютерних наук (ПРН20).

Мета дисципліни. Формування у здобувачів вищої освіти системних теоретичних знань та практичних навичок із моделювання, аналізу та оброблення текстових даних з використанням сучасних методів штучного інтелекту, статистичного та нейромережевого підходів. У процесі навчання вивчаються принципи побудови інформаційних моделей текстів, методи формалізації та векторизації текстових даних, архітектури трансформерів, тематичне моделювання, семантичний пошук, пояснюваність моделей та інтеграція інтелектуальних пайплайнів. Це забезпечує здатність здобувачів застосовувати сучасні алгоритми текстової аналітики для дослідницьких, прикладних і проєктних завдань у галузі комп'ютерних наук.

Предмет дисципліни. Моделі, методи, алгоритми та інтелектуальні програмні засоби для аналізу, оброблення, моделювання та інтерпретації текстових даних.

Завдання дисципліни. Формування компетентностей та забезпечення програмних результатів навчання у сфері текстової аналітики, зокрема: оволодіння методами попередньої обробки текстів, їх формалізації та корпусного представлення; застосування статистичних, математичних та глибинних моделей для побудови векторних подань текстових даних; використання сучасних трансформерних архітектур і методів zero-shot, few-shot та fine-tuning під конкретні задачі; реалізація тематичного моделювання, кластеризації та семантичного пошуку; розроблення та тестування алгоритмічного та програмного забезпечення для аналізу текстів; використання інструментів пояснюваного ШІ (XAI) для інтерпретації моделей текстової аналітики; проєктування інтегрованих інтелектуальних пайплайнів обробки текстових даних та їх застосування в дослідницькій і прикладній діяльності; виконання інноваційних та науково-дослідницьких завдань у сфері текстових технологій і комп'ютерних наук.

Результати навчання. Після вивчення дисципліни студент повинен володіти сучасною професійною термінологією у сфері текстової аналітики, розуміти теоретичні засади інформаційного моделювання текстів і принципи їх формалізації. Він має знати методи попередньої обробки, векторизації, тематичного моделювання, семантичного пошуку та глибинного аналізу текстових даних, а також підходи zero-shot, few-shot і fine-tuning для адаптації трансформерних моделей. Студент повинен уміти проєктувати та реалізовувати інтелектуальні пайплайни для класифікації, кластеризації, розпізнавання сутностей та сумаризації текстів, здійснювати семантичне зіставлення, інтерпретувати результати з використанням методів пояснюваного ШІ та оцінювати обмеження й ефективність моделей. Він має бути здатним виконувати дослідження, аналізувати потреби та вимоги до текстових систем, розробляти алгоритмічні рішення і програмне забезпечення, інтегрувати моделі у прикладні та наукові задачі, забезпечуючи етичність, надійність і практичну застосовність отриманих результатів.

4. Структура залікових кредитів дисципліни

Назва теми	Кількість годин, відведених на:		
	лекції	лабораторні заняття	СРС
Тема 1. Теоретичні основи текстової аналітики та інформаційного моделювання	2	4	12
Тема 2. Лінгвістичні та статистичні моделі попередньої обробки текстових даних	2	4	12
Тема 3. Математичні та нейромережеві моделі векторизації текстів.	2	4	12
Тема 4. Математичні та нейромережеві моделі векторизації текстів.	2	4	12
Тема 5. Архітектури трансформерів і їх адаптація для задач NLP.	2	4	12
Тема 6. Семантичний пошук і векторні інформаційно-пошукові системи.	2	4	12
Тема 7. Пояснюваність, етичність і верифікація моделей текстової аналітики.	2	4	12
Тема 8. Проектування комплексних інтелектуальних систем текстової аналітики.	2	6	16
Разом за семестр:	16	34	100

5. Програма навчальної дисципліни

5.1 Зміст лекційного курсу

Номер лекції	Перелік тем лекцій, їх анотації	Кількість годин
	Тема 1 Теоретичні основи текстової аналітики та інформаційного моделювання текстів	2
1	Текст як об'єкт формалізації. Корпуси, корпусна лінгвістика, теоретичні підходи. Моделі представлення знань у текстах. Літ.: [1]	2
	Тема 2. Лінгвістичні та статистичні моделі попередньої обробки текстових даних.	2
2	Морфологічний та синтаксичний аналіз. POS-тегування, NER та парсинг залежностей. Нормалізація й очищення текстів у наукових застосуваннях. Літ.: [1]	2
	Тема 3. Математичні та нейромережеві моделі векторизації текстів.	2
3	Семантичні простори та контекстуальні представлення. Distributed representations (Word2Vec, FastText). Векторизація документів і фраз.	2
	Тема 4. Математичні та нейромережеві моделі векторизації текстів.	2
4	LDA, pLSA, NMF як інструменти виявлення латентних тем. Кластеризація текстів: K-Means, йєрархічні методи, DBSCAN. Опрацювання великих корпусів.	2
	Тема 5. Архітектури трансформерів і їх адаптація для задач NLP.	2
5	Encoder/Decoder-моделі. BERT, RoBERTa, GPT, T5, XLM-R. Transfer learning, fine-tuning, LoRA та adapter-моделі.	2
	Тема 6. Семантичний пошук і векторні інформаційно-пошукові системи.	2
6	Semantic retrieval у порівнянні з традиційним пошуком. Embedding space і similarity search. FAISS, Milvus та гібридний пошук.	2
	Тема 7. Пояснюваність, етичність і верифікація моделей текстової аналітики.	2
7	Fairness, bias та відповідальне використання NLP-моделей. Attention visualization. Інструменти XAI: LIME, SHAP, Anchors.	2
	Тема 8. Проектування комплексних інтелектуальних систем текстової аналітики.	2
8	Архітектура інтегрованих пайплайнів NLP. MLOps/NLPOps та мультимодальні застосування. Прикладні кейси й дослідницькі проекти.	2
	Разом за семестр:	16

5.2 Зміст лабораторних занять

№ п/п	Тема лабораторного заняття	Кількість годин
1	Формування та попередня обробка текстових корпусів. Використання NLTK / Stanza для токенізації, лематизації, POS-тегування. Літ.: [1]	4
2	Побудова векторних представлень і семантичної подібності текстів. TF-IDF, Word2Vec/FastText. Обчислення семантичної подібності. Побудова простору ознак Літ.: [1]	4
3	LP3. Zero-shot та Few-shot класифікація текстів. Робота з трансформерними моделями без повного навчання. Prompt-based класифікація. Порівняння з fine-tuning Літ.: [1]	4
4	Тематичне моделювання та кластеризація текстових даних. K-Means, Agglomerative, DBSCAN. LDA або BERTopic. Альтернативи для великих корпусів Літ.: [1]	4
5	Налаштування та донавчання трансформерних моделей (HuggingFace). Завантаження готової моделі. Fine-tuning на користувацькому наборі. Інференс і валідація Літ.: [1]	4
6	Семантичний пошук і побудова embedding space. Semantic retrieval. Vector stores (FAISS). Побудова пошукового інтерфейсу Літ.: [1]	4
7	Інтерпретація та пояснюваність моделей текстової аналітики. Attention visualization. LIME/SHAP для текстів. Аналіз помилок. Літ.: [1]	4
8	Проектування та реалізація інтелектуального пайплайну текстової аналітики. Постановка задачі (класифікація, NER, summarization...). Конструювання обчислювального пайплайну. Презентація та інтерпретація результатів Літ.: [1]	6
	Разом:	34

5.3 Зміст самостійної (у т.ч. індивідуальної) роботи здобувача вищої освіти

Самостійна робота студентів усіх форм здобуття освіти полягає у систематичному опрацюванні програмного матеріалу з відповідних джерел інформації, підготовці до лабораторних занять, контрольних робіт, виконанні індивідуальних завдань. Крім цього до послуг студентів сторінка навчальної дисципліни у Модульному середовищі для навчання, де розміщені Робоча програма дисципліни та необхідні документи з її навчально-методичного забезпечення.

Зміст самостійної роботи студентів денної форми навчання

Номер тижня	Вид самостійної роботи	Кількість годин
1–3	Опрацювання лекційного матеріалу. Підготовка до лабораторних робіт. Підготовка до захисту лабораторної роботи. Виконання завдання для самостійної роботи №1	14
4-5	Опрацювання лекційного матеріалу. Підготовка до лабораторних робіт. Підготовка до захисту лабораторної роботи. Продовження виконання завдання для самостійної роботи №1	14
6-7	Опрацювання лекційного матеріалу. Підготовка до лабораторних робіт. Підготовка до захисту лабораторної роботи. Підготовка до проміжного контрольного заходу. Продовження виконання завдання для самостійної роботи №1. Підготовка до КР1.	14
8-9	Опрацювання лекційного матеріалу. Підготовка до лабораторних робіт. Підготовка до захисту лабораторної роботи. Підготовка до проміжного контрольного заходу. Продовження виконання завдання для самостійної роботи №1.	14
10-11	Опрацювання лекційного матеріалу. Підготовка до лабораторної роботи. Підготовка до захисту лабораторної роботи. Виконання завдання для самостійної роботи №2	14
12-13	Опрацювання лекційного матеріалу. Підготовка до лабораторної роботи. Підготовка до захисту лабораторної роботи. Виконання завдання для самостійної роботи №2	14
14-17	Опрацювання лекційного матеріалу. Підготовка до лабораторних робіт. Підготовка до захисту лабораторних робіт. Виконання завдання для самостійної роботи №2. Оформлення звіту з виконання ІЗ). Підготовка до підсумкового контрольного заходу. Презентація результатів виконання індивідуальних завдань.	16
Всього		100

Примітки: КР – контрольна робота, ІЗ – індивідуальне завдання.

На самостійне опрацювання студентів виносяться визначені у методичних рекомендаціях до лабораторних занять та самостійної роботи питання з відповідних тем. Керівництво самостійною роботою та контроль за виконанням індивідуального завдання здійснюється викладачем згідно з розкладом консультацій у позаурочний час.

Вимоги до виконання індивідуального домашнього завдання викладені в Модульному середовищі для навчання на сторінці навчальної дисципліни.

6. Технології та методи навчання

Процес навчання з дисципліни ґрунтується на використанні традиційних та сучасних технологій та методів навчання, зокрема: лекції (з використанням методів візуалізації, проблемного й інтерактивного навчання, мотиваційних прийомів, інформаційно-комунікаційних технологій); лабораторні заняття (з використанням інструктування, демонстрування, розв'язування типових і прикладних задач, аналізу кейсів, ситуаційних завдань, елементів дискусії тощо); самостійна робота (опрацювання теоретичного матеріалу, підготовка до виконання лабораторних робіт, поточного та підсумкового контролю, виконання індивідуальних завдань), з використанням інформаційно-комп'ютерних технологій та технологій дистанційного навчання.

7. Методи контролю

Поточний контроль здійснюється під час аудиторних лабораторних занять, а також у дні проведення контрольних заходів, встановлених робочою програмою і графіком освітнього процесу, в т.ч. з використанням Модульного середовища для навчання. При цьому використовуються такі методи поточного контролю:

- оцінювання результатів виконання індивідуального завдання;
- оцінювання контрольних робіт;
- оцінювання результатів захисту лабораторних робіт.

При виведенні підсумкової семестрової оцінки враховуються результати як поточного контролю, так і підсумкового контролю, який проводиться з усього матеріалу дисципліни за білетами, попередньо розробленими і затвердженими на засіданні кафедри. Здобувач вищої освіти, який набрав з будь-якого виду навчальної роботи, суму балів нижчу за 60 відсотків від максимального балу, **не допускається** до семестрового контролю, поки не виконає обсяг роботи, передбачений Робочою програмою. Здобувач вищої освіти, який набрав позитивний середньозважений бал (60 відсотків і більше від максимального балу) з усіх видів поточного контролю і не склав іспит, вважається таким, який **має** академічну заборгованість. Ліквідація академічної заборгованості із семестрового контролю здійснюється у період екзаменаційної сесії або за графіком, встановленим деканатом відповідно до «Положення про контроль і оцінювання результатів навчання здобувачів вищої освіти у ХНУ».

8. Політика дисципліни

Політика навчальної дисципліни загалом визначається системою вимог до здобувача вищої освіти, що передбачені чинними положеннями Університету про організацію і навчально-методичне забезпечення освітнього процесу. Зокрема, проходження інструктажу з техніки безпеки; відвідування занять з дисципліни є обов'язковим. За об'єктивних причин (підтверджених документально) теоретичне навчання за погодженням із лектором може відбутись в он-лайн режимі. Успішне опанування дисципліни і формування фахових компетентностей і програмних результатів навчання передбачає необхідність підготовки до лабораторного заняття (вивчення теоретичного матеріалу з теми роботи, попередню підготовку протоколу роботи (наведені у Методичних рекомендаціях до лабораторних занять)), активно працювати на занятті, якісно підготувати звіт, захистити результати виконаної роботи, брати участь у дискусіях щодо прийнятих конструктивних рішень при виконанні здобувачами лабораторних робіт тощо.

Здобувачі вищої освіти мають дотримуватися встановлених термінів виконання всіх видів навчальної роботи відповідно до робочої програми навчальної дисципліни. Термін захисту лабораторної роботи вважається своєчасним, якщо студент захистив її на наступному після виконання роботи занятті. Пропущене лабораторне заняття студент зобов'язаний відпрацювати у встановлений викладачем термін, але не пізніше, ніж за два тижні до кінця теоретичних занять у семестрі.

Засвоєння студентом теоретичного матеріалу з дисципліни оцінюється за результатами опитування під час лабораторних занять, виконання індивідуального завдання та контрольних робіт. Виконання індивідуального завдання завершується його здачею на перевірку у терміни, встановлені графіком самостійної роботи.

Здобувач вищої освіти, виконуючи самостійну роботу або індивідуальну роботу з дисципліни, має дотримуватися політики доброчесності (заборонені списування, плагіат (в т.ч. із використанням мобільних девайсів)). У разі виявлення порушення політики академічної доброчесності в будь-яких видах навчальної роботи здобувач вищої освіти отримує незадовільну оцінку і має повторно виконати завдання з відповідної теми (виду роботи), що передбачені робочою програмою. Будь-які форми порушення академічної доброчесності **не допускаються**.

У межах вивчення навчальної дисципліни допускається визнання та зарахування результатів навчання, здобутих шляхом неформальної освіти на відкритих освітніх платформах. Такі результати можуть враховуватися за умови їх відповідності компетентностям і програмним результатам навчання, визначеним робочою програмою дисципліни, або у разі забезпечення вивчення окремих тем і видів робіт, що передбачені програмою. (детальніше у Положенні про порядок визнання та зарахування результатів навчання здобувачів вищої освіти у ХНУ).

9. Оцінювання результатів навчання студентів у семестрі

Оцінювання академічних досягнень здобувача вищої освіти здійснюється відповідно до «Положення про контроль і оцінювання результатів навчання здобувачів вищої освіти у ХНУ». При поточному оцінюванні виконаної здобувачем роботи з кожної структурної одиниці і отриманих ним результатів викладач виставляє йому певну кількість балів із встановлених Робочою програмою для цього виду роботи. При цьому кожна структурна одиниця навчальної роботи може бути зарахована, якщо здобувач набрав не менше 60 відсотків (мінімальний рівень для позитивної оцінки) від максимально можливої суми балів, призначеної структурній одиниці.

При оцінюванні результатів навчання здобувачів вищої освіти з будь-якого виду навчальної роботи (структурної одиниці) рекомендується використовувати наведені нижче узагальнені критерії:

Таблиця – Критерії оцінювання навчальних досягнень здобувача вищої освіти

Оцінка та рівень досягнення здобувачем запланованих ПРН та сформованих компетентностей	Узагальнений зміст критерія оцінювання
Відмінно (високий)	Здобувач вищої освіти глибоко і у повному обсязі опанував зміст навчального матеріалу, легко в ньому орієнтується і вміло використовує понятійний апарат; уміє пов'язувати теорію з практикою, вирішувати практичні завдання, впевнено висловлювати і обґрунтовувати свої судження. Відмінна оцінка передбачає логічний виклад відповіді мовою викладання (в усній або у письмовій формі), демонструє якісне оформлення роботи і володіння спеціальними приладами та інструментами, прикладними програмами. Здобувач не вагається при видозміні запитання, вміє робити детальні та узагальнюючі висновки, демонструє практичні навички з вирішення фахових завдань. При відповіді допустив дві–три несуттєві помилки .
Добре (середній)	Здобувач вищої освіти виявив повне засвоєння навчального матеріалу, володіє понятійним апаратом, орієнтується у вивченому матеріалі; свідомо використовує теоретичні знання для вирішення практичних задач; виклад відповіді грамотний, але у змісті і формі відповіді можуть мати місце окремі неточності, нечіткі формулювання правил, закономірностей тощо. Відповідь здобувача вищої освіти будується на основі самостійного мислення. Здобувач вищої освіти у відповіді допустив дві–три несуттєві помилки .
Задовільно (достатній)	Здобувач вищої освіти виявив знання основного програмного матеріалу в обсязі, необхідному для подальшого навчання та практичної діяльності за професією, справляється з виконанням практичних завдань, передбачених програмою. Як правило, відповідь здобувача вищої освіти будується на рівні репродуктивного мислення, здобувач

	вищої освіти має слабкі знання структури навчальної дисципліни, допускає неточності і суттєві помилки у відповіді, вагається при відповіді на видозмінене запитання. Разом з тим, набув навичок, необхідних для виконання нескладних практичних завдань, які відповідають мінімальним критеріям оцінювання і володіє знаннями, що дозволяють йому під керівництвом викладача усунути неточності у відповіді.
Незадовільно (недостатній)	Здобувач вищої освіти виявив розрізнені, безсистемні знання, не вміє виділяти головне і другорядне, допускається помилок у визначенні понять, перекручує їх зміст, хаотично і невпевнено викладає матеріал, не може використовувати знання при вирішенні практичних завдань. Як правило, оцінка «незадовільно» виставляється здобувачеві вищої освіти, який не може продовжити навчання без додаткової роботи з вивчення навчальної дисципліни.

Структурування дисципліни за видами навчальної роботи і оцінювання результатів навчання студентів денної форми здобуття освіти у семестрі

Аудиторна робота								Контрольні заходи		Самостійна робота		Семестровий контроль	Разом
Сьомий семестр													
Лабораторні роботи №:								Контрольні роботи №:		ІЗ №:		Іспит	Сума балів
1	2	3	4	5	6	7	8	1	2	1	2		
Кількість балів за вид навчальної роботи (мінімум-максимум)													
3-5	3-5	3-5	3-5	3-5	3-5	3-5	3-5	3-5	3-5	3-5	3-5	24-40	60-100
24-40								6-10		6-10		24-40	

Примітки: *За набрану з будь-якого виду навчальної роботи з дисципліни кількість балів, нижче встановленого мінімуму, здобувач отримує незадовільну оцінку і має її перездати у встановлений викладачем (деканом) термін. Інституційна оцінка встановлюється відповідно до таблиці «Співвідношення інституційної шкали оцінювання і шкали оцінювання ЄКТС».

Оцінювання результатів захисту лабораторної роботи

Виконана й оформлена відповідно до встановлених Методичними рекомендаціями вимог лабораторна робота комплексно оцінюється викладачем при її захисті з урахуванням таких критеріїв: самостійність та правильність виконання; повнота відповіді та знання.

Результат виконання і захисту здобувачем вищої освіти кожної лабораторної роботи оцінюється відповідно до таблиці Критеріїв оцінювання навчальних досягнень здобувача вищої освіти.

У випадку виявлення здобувачем рівня знань, нижчого ніж 60 відсотків від максимального балу, встановленого Робочою програмою для кожної структурної одиниці, лабораторна робота йому **не зараховується** і для її захисту він має детальніше опрацювати матеріал з теми роботи, методику її виконання, виправити грубі помилки та повторно вийти на її захист у призначений для цього викладачем час.

Оцінювання контрольної роботи

Контрольна робота передбачає виконання п'яти завдань. При оцінюванні контрольної роботи враховуються: повнота відповіді та якість виконання. Кожне завдання оцінюється 1 балом, загальна сума балів на позитивну оцінку становить від 3 до 5.

Розподіл балів при оцінюванні завдань контрольної роботи

Кількість правильних відповідей	1	2	3	4	5
Відсоток правильних відповідей	0-59		60	80	100
Кількість отриманих балів	0		3	4	5

При отриманні негативної оцінки контрольну роботу слід перездати до терміну **наступного** контролю.

Оцінювання результатів виконання індивідуального завдання

Виконане та оформлене відповідно до вимог, визначених Методичними рекомендаціями, індивідуальне завдання (ІЗ) комплексно оцінюється викладачем з урахуванням таких критеріїв: самостійність виконання; правильність розв'язання поставлених задач; обґрунтованість вибору методів розв'язання; повнота пояснень та аргументованість відповідей; якість оформлення та дотримання вимог до структури і змісту роботи.

Результат виконання здобувачем вищої освіти кожного ІЗ оцінюється відповідно до таблиці **Критеріїв оцінювання навчальних досягнень здобувача вищої освіти** з урахуванням рівня досягнення запланованих програмних результатів навчання та сформованих компетентностей. За підсумками захисту присвоюється відповідна сума балів (мінімальний позитивний бал – 3 бали, максимальний – 5 балів).

У разі, якщо здобувач вищої освіти виявив рівень знань і виконання ІЗ, що нижчий ніж 60 відсотків від максимальної кількості балів, встановленої Робочою програмою для цієї структурної одиниці, завдання не

зараховується. У такому випадку студент має повторно опрацювати зміст завдання, усунути помилки та здати на перевірку доопрацьоване ІЗ у терміни, погоджені з викладачем.

Оцінювання результатів підсумкового семестрового контролю (іспит)

Освітня програма передбачає підсумковий семестровий контроль з дисципліни у формі іспиту, завданням якого є системне й об'єктивне оцінювання як теоретичної, так і практичної підготовки здобувача з навчальної дисципліни. Складання іспиту відбувається за попередньо розробленими і затвердженими на засіданні кафедри білетами. Відповідно до цього в екзаменаційному білеті пропонується поєднання питань як теоретичного (в т.ч. у тестовій формі), так і практичного характеру.

Таблиця – Оцінювання результатів підсумкового семестрового контролю здобувачів денної форми навчання (40 балів для підсумкового контролю)

Види завдань	Для кожного окремого виду завдань		
	Мінімальний (достатній) бал (задовільно)	Потенційні позитивні бали* (середній бал) (добре)	Максимальний (високий) бал (відмінно)
Завдання № 1	6	8	10
Завдання № 2	6	8	10
Завдання № 3	6	8	10
Завдання № 4	6	8	10
Разом:	24		40

Примітка. *Позитивний бал за іспит, відмінний від мінімального (24 бали) та максимального (40 балів), знаходиться в межах 25-39 балів та розраховується як сума балів за усі структурні елементи (завдання) іспиту.

Для кожного окремого виду завдань підсумкового семестрового контролю застосовуються критерії оцінювання навчальних досягнень здобувача вищої освіти, наведені вище (**Таблиця – Критерії оцінювання навчальних досягнень здобувача вищої освіти**).

Підсумкова семестрова оцінка за інституційною шкалою і шкалою ЄКТС визначається в автоматизованому режимі після внесення викладачем результатів оцінювання у балах з усіх видів навчальної роботи до електронного журналу. Співвідношення інституційної шкали оцінювання і шкали оцінювання ЄКТС наведені нижче у таблиці «Співвідношення».

Семестровий іспит виставляється, якщо загальна сума балів, яку набрав студент з дисципліни за результатами поточного контролю, знаходиться у межах від 60 до 100 балів. При цьому за інституційною шкалою ставиться оцінка «відмінно/добре/задовільно», а за шкалою ЄКТС – буквене позначення оцінки, що відповідає набраній студентом сумі балів відповідно до таблиці Співвідношення.

Таблиця – Співвідношення інституційної шкали оцінювання і шкали оцінювання ЄКТС

Оцінка ЄКТС	Рейтингова шкала балів	Інституційна оцінка (рівень досягнення здобувачем вищої освіти запланованих результатів навчання з навчальної дисципліни)	
		Залік	Іспит/диференційований залік
A	90-100	Зараховано	Відмінно/Excellent – високий рівень досягнення запланованих результатів навчання з навчальної дисципліни, що свідчить про безумовну готовність здобувача до подальшого навчання та/або професійної діяльності за фахом
B	83-89		Добре/Good – середній (максимально достатній) рівень досягнення запланованих результатів навчання з навчальної дисципліни та готовності до подальшого навчання та/або професійної діяльності за фахом
C	73-82		Задовільно/Satisfactory – Наявні мінімально достатні для подальшого навчання та/або професійної діяльності за фахом результати навчання з навчальної дисципліни
D	66-72		
E	60-65		
FX	40-59	Незараховано	Незадовільно/Fail – Низка запланованих результатів навчання з навчальної дисципліни відсутня. Рівень набутих результатів навчання є недостатнім для подальшого навчання та/або професійної діяльності за фахом
F	0-39		Незадовільно/Fail – Результати навчання відсутні

10. Питання для самоконтролю результатів навчання

1. Які основні завдання охоплює текстова аналітика?
2. Що таке формалізація текстових даних і для чого вона застосовується?
3. У чому полягає різниця між корпусною лінгвістикою та лінгвоінженерією?
4. Які етапи включає побудова текстового корпусу?
5. Які формати зберігання корпусів використовуються в NLP-системах?
6. Як здійснюється морфологічний аналіз тексту в автоматичному режимі?
7. У чому відмінність лематизації від стемінгу?
8. Які проблеми виникають при попередній обробці текстів природною мовою?

9. Для чого застосовується POS-тегування?
10. Як визначаються сутності та структури тексту на синтаксичному рівні?
11. У чому перевага розріджених моделей подання текстів (TF-IDF)?
12. Які особливості моделей Word2Vec (CBOW і Skip-gram)?
13. Як FastText працює з підсловниковими одиницями?
14. Які недоліки мають статичні словникові вектори?
15. Що таке embedding space і як він використовується?
16. Як обчислюється семантична подібність текстів?
17. Які метрики застосовуються для порівняння векторних моделей?
18. Чим контекстуальні вектори відрізняються від класичних?
19. У чому суть тематичного моделювання?
20. Як працює модель LDA?
21. Що таке BERTopic та в чому її новизна?
22. Які особливості кластеризації текстових колекцій?
23. Чим відрізняється K-Means від Agglomerative Clustering?
24. Як визначають оптимальну кількість тематичних кластерів?
25. У чому полягає відмінність між кластеризацією та тематичним моделюванням?
26. Які обмеження мають RNN, LSTM та GRU в NLP?
27. Яку роль відіграє механізм attention?
28. У чому полягає архітектура Transformer?
29. Як побудовано модель BERT?
30. Чим GPT відрізняється від BERT?
31. Що таке encoder-only, decoder-only і encoder-decoder моделі?
32. Які переваги має fine-tuning над класичним навчанням моделей?
33. Що таке LoRA та adapters у донавчанні трансформерів?
34. Як працює zero-shot класифікація?
35. Чим few-shot підхід відрізняється від zero-shot?
36. У чому різниця між sparse та dense retrieval?
37. Що таке векторні бази даних (FAISS, Milvus)?
38. Як здійснюється семантичний пошук у великих корпусах?
39. Що таке embedding-based indexing?
40. Чим semantic retrieval відрізняється від ключового пошуку?
41. Які ризики пов'язані з упередженістю моделей NLP?
42. У чому полягає концепція fairness у ШІ?
43. Як працює LIME для текстів?
44. Для чого застосовується SHAP у NLP?
45. Що таке attention visualization?
46. Як оцінюється інтерпретованість результатів?
47. З яких етапів складається пайплайн текстової аналітики?
48. Як забезпечується якість моделей у MLOps/NLPOps середовищі?
49. Які особливості інтеграції моделей у прикладні системи?
50. Які напрями використання текстової аналітики є найбільш актуальними для комп'ютерних наук?

11. Навчально-методичне забезпечення

Освітній процес з дисципліни «Об'єктно-орієнтоване проектування» повністю і в достатній кількості забезпечений необхідною навчально-методичною літературою. Зокрема, викладачами кафедри підготовлені і видані такі роботи:

1. Мазурець О. В. Моделі та методи інтелектуального аналізу текстової інформації та машинного навчання : курс лекцій для студентів спеціальності «Комп'ютерні науки». Хмельницький : ХНУ, 2021. 98 с. URL: http://lib.khmnu.edu.ua/EL_LIBRARY/vidavn/metod/mtd2021_1e/70/index.pdf.

12. Матеріально-технічне та програмне забезпечення дисципліни (за потреби)

Інформаційна та комп'ютерна підтримка: ПК, планшет, смартфон або інший мобільний пристрій, проектор. Програмне забезпечення: програми Microsoft Office або аналогічні, доступ до мережі Інтернет, робота з презентаціями.

Вивчення навчальної дисципліни не потребує використання спеціального програмного прикладного забезпечення, крім загальноновживаних програм і операційних систем та такого, що є безкоштовним або ж надається студентам за академічними ліцензіями.

13. Рекомендована література:

Основна

2. Jurafsky, D., & Martin, J. (2025). Speech and Language Processing (3rd ed.) [Draft PDF]. Stanford University. URL: <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>

3. Терейковський І.А., Бушуєв Д.А., Терейковська Л.О. Штучні нейронні мережі: базові положення. Навчальний посібник. Київ, КПІ ім. Ігоря Сікорського. 2022. 123с. URL: <https://ela.kpi.ua/bitstream/123456789/50135/1/ANN.pdf>

Додаткова

4. Hugging Face Transformers: Documentation and API Reference Hugging Face, 2024. URL: <https://huggingface.co/docs/transformers>
5. TensorFlow Text API and Guides Google Research, 2024. URL: <https://www.tensorflow.org/text>

14. Інформаційні ресурси

6. Модульне середовище для навчання. URL : <https://msn.khmnu.edu.ua>
7. Електронна бібліотека університету. URL: <http://library.khmnu.edu.ua/>
8. Репозитарій ХНУ. URL : <https://elar.khmnu.edu.ua/home>

МОДЕЛІ ТА МЕТОДИ ТЕКСТОВОЇ АНАЛІТИКИ

Тип дисципліни	Обов'язкова
Рівень вищої освіти	Другий (магістерський)
Мова викладання	Українська
Семестр	Другий
Кількість призначених кредитів ЄКТС	5,0
Форми здобуття освіти, для яких викладається дисципліна	Очна (денна)

Результати навчання. Після вивчення дисципліни студент повинен володіти сучасною професійною термінологією у сфері текстової аналітики, розуміти теоретичні засади інформаційного моделювання текстів і принципи їх формалізації. Він має знати методи попередньої обробки, векторизації, тематичного моделювання, семантичного пошуку та глибинного аналізу текстових даних, а також підходи zero-shot, few-shot і fine-tuning для адаптації трансформерних моделей. Студент повинен уміти проєктувати та реалізовувати інтелектуальні пайплайни для класифікації, кластеризації, розпізнавання сутностей та сумаризації текстів, здійснювати семантичне зіставлення, інтерпретувати результати з використанням методів пояснюваного ШІ та оцінювати обмеження й ефективність моделей. Він має бути здатним виконувати дослідження, аналізувати потреби та вимоги до текстових систем, розробляти алгоритмічні рішення і програмне забезпечення, інтегрувати моделі у прикладні та наукові задачі, забезпечуючи етичність, надійність і практичну застосовність отриманих результатів.

Зміст навчальної дисципліни. Дисципліна охоплює сучасні підходи до формалізації, моделювання та аналізу текстових даних у межах комп'ютерних наук. Розглядаються базові принципи корпусної лінгвістики та інформаційного моделювання текстів, включно з токенизацією, лематизацією, частиномовним тегуванням, синтаксичним аналізом і виділенням сутностей. Вивчаються статистичні та нейромережеві методи представлення текстів, зокрема TF-IDF, Word2Vec, FastText і контекстуальні вектори. Тематичне моделювання розглядається як інструмент виявлення латентної структури текстових колекцій, а кластеризація — як засіб групування неструктурованих даних. Особливе місце займають архітектури трансформерів, включно з моделями BERT, GPT, RoBERTa, T5, їх донавчанням, переносом знань, zero-shot і few-shot підходами. Окремо аналізуються методи семантичного пошуку, побудова embedding-простору, застосування векторних баз даних та розроблення інтелектуальних пошуково-аналітичних систем. Значну увагу приділено питанням пояснюваності, етичності та якості NLP-моделей, застосуванню таких інструментів, як SHAP, LIME та attention-based аналіз. Завершальним елементом є проєктування та інтеграція інтелектуальних пайплайнів текстової аналітики для прикладних, наукових та дослідницьких завдань, що забезпечує здатність студентів реалізовувати комплексні системи оброблення текстів у межах сучасних інформаційних технологій.

Пререквізити – Англійська мова за професійним спрямуванням (ОЗП.03), Проєктування та розробка інформаційних систем (ОФП.01), Інженерія, аналітика та візуалізація великих даних (ОФП.02).

Постреквізити – Переддипломна практика (ОФП.06), Кваліфікаційна робота (ОФП.07)

Запланована навчальна діяльність: Мінімальний обсяг навчальних занять в одному кредиті ЄКТС навчальної дисципліни для *першого* (бакалаврського) рівня вищої освіти за денною формою здобуття освіти становить 10 годин на 1 кредит ЄКТС.

Форми (методи) навчання: Процес навчання з дисципліни ґрунтується на використанні традиційних та сучасних технологій та методів навчання, зокрема: лекції (з використанням методів візуалізації, проблемного й інтерактивного навчання, мотиваційних прийомів, інформаційно-комунікаційних технологій); лабораторні заняття (з використанням інструктування, демонстрування, розв'язування типових і прикладних задач, аналізу кейсів, ситуаційних завдань, елементів дискусії тощо); самостійна робота (опрацювання теоретичного матеріалу, підготовка до виконання лабораторних робіт, поточного та підсумкового контролю, виконання індивідуальних завдань), з використанням інформаційно-комп'ютерних технологій та технологій дистанційного навчання.

Форми оцінювання результатів навчання: оцінювання лабораторних та контрольних робіт; виконання індивідуального завдання; підсумковий контрольний захід.

Вид семестрового контролю: іспит – 2 семестр.

Навчальні ресурси:

1. Мазурець О. В. Моделі та методи інтелектуального аналізу текстової інформації та машинного навчання : курс лекцій для студентів спеціальності «Комп'ютерні науки». Хмельницький : ХНУ, 2021. 98 с. URL: http://lib.khmnu.edu.ua/EL_LIBRARY/vidavn/metod/mtd2021_1e/70/index.pdf.

2. Jurafsky, D., & Martin, J. (2025). Speech and Language Processing (3rd ed.) [Draft PDF]. Stanford University. URL: <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>
3. Терейковський І.А., Бушуєв Д.А., Терейковська Л.О. Штучні нейронні мережі: базові положення. Навчальний посібник. Київ, КПІ ім. Ігоря Сікорського. 2022. 123с. URL: <https://ela.kpi.ua/bitstream/123456789/50135/1/ANN.pdf>
Додаткова
4. Hugging Face Transformers: Documentation and API Reference Hugging Face, 2024. URL: <https://huggingface.co/docs/transformers>
5. TensorFlow Text API and Guides Google Research, 2024. URL: <https://www.tensorflow.org/text>
6. Модульне середовище для навчання. URL : <https://msn.khmnu.edu.ua>
7. Електронна бібліотека університету. URL: <http://library.khmnu.edu.ua/>
8. Репозитарій ХНУ. URL : <https://elar.khmnu.edu.ua/home>

Викладачі: канд. техн. наук, доц. Олександр МАЗУРЕЦЬ