ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ

ЗАТВЕРДЖУЮ

Декан математичного факультету ЗНУ
_____ С.І.Гоменюк
(підпис) (ініціали та прізвище)
«_____» _____ 2026р.

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

ТЕХНОЛОГІЇ ОБРОБКИ ТЕКСТОВИХ ДАНИХ

підготовки бакалаврів

денної та заочної форми здобуття освіти

освітньо-професійна програма Прикладна лінгвістика

спеціальності 035 Філологія

галузі знань 03 - гуманітарні науки

ВИКЛАДАЧ: Добровольський Геннадій Анатолійович, к.т.н, доцент кафедри комп'ютерних наук

Обговорено та ухвалено
на засіданні кафедри комп'ютерних наук

Протокол №_1_ від "_29__" _серпня__2026 р.
Завідувач кафедри комп'ютерних наук

О. С. Пшенична
(підпис) (ініціали, прізвище)

Погоджено
Гарант освітньо-професійної програми

_____ С. В. Сабліна
(підпис) (ініціали, прізвище)

2026 рік

Зв`язок з викладачем: Добровольський Геннадій Анатолійович
E-mail: gen.dobr@gmail.com
Сезн ЗНУ повідомлення: https://moodle.znu.edu.ua/course/view.php?id=16843
Телефон (кафедра): +38-061-289-12-57
Інші засоби зв'язку: https://t.me/gen_dobr
Кафедра комп'ютерних наук, ауд. №39, 1 корпус ЗНУ

1. Опис навчальної дисципліни

Метою викладання навчальної дисципліни «Технології обробки текстових даних» є знайомство студентів з основними концепціями, технологіями та методами обробки природних мов (Natural Language Processing, NLP), їх застосуванням у реальних сценаріях.

Для вивчення дисципліни «Технології обробки текстових даних» студентам потрібно засвоїти матеріал дисциплін «Математичні основи спеціальності», «Основи програмування», «Мова програмування Python», «Об'єктно-орієнтоване програмування», «Автоматизована обробка одиниць природної мови».

Компетентності, отримані під час вивчення дисципліни «Технології обробки текстових даних», необхідні для виконання завдань виробничої практики, курсових робіт з фаху, дисципліни «Сучасні інформаційні технології перекладу».

У результаті вивчення навчальної дисципліни «Технології обробки текстових даних» студент зможе:

- готувати тексти для комп'ютерного моделювання;
- створювати комп'ютерні моделі природної мови;
- використовувати інформаційні технології для аналізу, перетворення та генерації текстів.

Паспорт навчальної дисципліни

Нормативні показники	**денна форма здобуття освіти**	**заочна форма здобуття освіти**
Статус дисципліни	**Обов'язкова**	
Семестр	5-й	5-й
Кількість кредитів ECTS	3	3
Кількість годин	90	90
Лекційні заняття	8 год.	4 год.
Практичні заняття	14 год.	4 год.
Самостійна робота	68 год.	92 год.
Консультації	Згідно розкладу: дистанційно: Zoom: Ідентифікатор: 771 5037 0584 Пароль: 9K8yqV	
Вид підсумкового семестрового контролю:	**екзамен**	
Посилання на електронний курс у СЕЗН ЗНУ (платформа Moodle)	https://moodle.znu.edu.ua/course/view.php?id=16843	

2. Методи досягнення запланованих освітньою програмою компетентностей і результатів навчання

Компетентності/ результати навчання	**Методи навчання**	**Форми і методи оцінювання**
Компетентності		
ЗК 4. Здатність бути критичним і самокритичним. ЗК 5. Здатність учитися й оволодівати сучасними знаннями. ЗК 6. Здатність до пошуку, опрацювання та аналізу інформації з різних джерел. ЗК 7. Уміння виявляти, ставити та вирішувати проблеми. ЗК 10. Здатність до абстрактного мислення, аналізу та синтезу. СК 7. Здатність до збирання й аналізу, систематизації та інтерпретації мовних фактів, інтерпретації та перекладу тексту. СК 8. Здатність вільно оперувати спеціальною термінологією для	лекція-візуалізація, пояснення, демонстрування, дискусія, аналіз, виконання завдань практичних робіт	**Поточний контроль:** захист практичних та самостійних робіт, опитування, тестування **Підсумковий контроль:** тестування

розв’язання професійних завдань. СК 13. Спроможність аналізувати, оцінювати, перетворювати й створювати тексти різних жанрів та стилів. СК 15. Здатність проводити лінгвістичний супровід аналітичної діяльності. СК 17. Спроможність до текстової аналітики, зіставляння різнорівневих, різностильових і різномовних фактів. СК 18. Спроможність створювати мовою програмування інструментальні засоби для лінгвістичного аналізу. СК 19. Здатність працювати з інтелектуальними системами й здійснювати автоматизовану обробку тексту.		
Результати навчання		
ПРН 2. Знає принципи, технології і прийоми створення усних і письмових текстів різних жанрів і стилів державною мовою. ПРН 4. Здатний аналізувати мовні одиниці, визначати їхню взаємодію та характеризувати мовні явища і процеси, що їх зумовлюють. ПРН 5. Знає методи теоретичної та прикладної лінгвістики для розв’язання складних професійних завдань. ПРН 8. Спроможний організовувати процес свого навчання й самоосвіти. ПРН 9. Здатний працювати з інформацією: добирати необхідну інформацію з різних джерел, зокрема з фахової літератури та електронних баз, критично аналізувати й інтерпретувати її, впорядковувати, класифікувати й систематизувати. ПРН 13. Уміє використовувати інформаційні й комунікаційні технології для вирішення складних спеціалізованих задач і проблем професійної діяльності. ПРН 15. Здатний здійснювати лінгвістичний та спеціальний філологічний аналіз текстів різних стилів і жанрів. ПРН 19. Спроможний використовувати різне програмне забезпечення для проведення лінгвістичних досліджень, аналізу текстової інформації, що пов’язані з лінгвістикою.	лекція-візуалізація, пояснення, дискусія, аналіз, виконання завдань практичних робіт	**Поточний контроль:** захист практичних та самостійних робіт, опитування, тестування **Підсумковий контроль:** тестування

ПРН 20. Здатний розуміти й використовувати нові досягнення програмного й технічного забезпечення в процесі роботи з лінгвістичним матеріалом.		

3. Зміст навчальної дисципліни

Змістовий модуль 1. Огляд основних задач автоматичної обробки природних мов

Тема 1. Предмет, задачі та методи автоматичної обробки природних мов.

Основні визначення. Токенізація, визначення частин мови, стемінг та лематизація, стоп-слова, створення словників. Виявлення іменованих сутностей, аналіз настроїв, визначення залежностей, класифікація текстів, машиний переклад, розпізнавання усної мови, генерація текстів, скорочений переказ, виявлення кореферентності, розрізнення омонімів, відповіді на запитання.

Змістовий модуль 2. Методи числового представлення слів, речень, документів.

Тема 2. Методи числового представлення слів, речень, документів

Створення словників, векторна модель документа. TF-IDF. Представлення слів у R^n. Латентний семантичний аналіз. Приховане розміщення Діріхле. GloVe. Сучасні методи тематичного моделювання. Оцінка якості тематичного моделювання. Word2Vec, Fasttext. Комбінації методів.

Змістовий модуль 3. Прикладні задачі

Тема 3. Прикладні задачі.

Виявлення сутностей у тексті. Виявлення зв'язків між сутностями. Класифікація тексту. Аналіз настрою.

Змістовий модуль 4. Прикладні задачі

Тема 4. Прикладні задачі.

Машинний переклад. Генерація тексту. Скорочений переказ. Виявлення кореферентності. Розрізнення омонімів. Відповіді на запитання.

4. Структура навчальної дисципліни

Вид заняття /роботи	**Назва теми**	**Кількість годин о/д. ф.**	**Згідно з розкладом**
1	***2***	***3***	***4***
Лекція 1	Предмет, задачі та методи автоматичної обробки природних мов.	2	тиждень 1
Практичне заняття 1	Токенізація.	2	тиждень 2
Практичне заняття 2	Розмітка частин мови.	2	тиждень 3
Самостійна робота	Постановка задач: виявлення іменованих сутностей, аналіз настроїв, визначення залежностей, класифікація текстів, машиний переклад, розпізнавання усної мови, генерація текстів, скорочений переказ, виявлення кореферентності, розрізнення омонімів, відповіді на запитання.	17	Тиждень 1 Тиждень 2 Тиждень 3
Лекція 2	Методи числового представлення слів, речень, документів	2	тиждень 4
Практичне заняття 3	Стемінг та лематизація	2	тиждень 5
Практичне заняття 4	Виявлення та видалення стоп-слів	2	тиждень 6
Самостійна робота	Приховане розміщення Діріхле. GloVe. Сучасні методи тематичного моделювання. Оцінка якості тематичного моделювання. Word2Vec, Fasttext.	17	тиждень 4 тиждень 5 тиждень 6
Лекція 3	Прикладні задачі: Виявлення сутностей у тексті. Виявлення зв'язків між сутностями. Класифікація тексту. Аналіз настрою.	2	тиждень 7
Практичне заняття 5	Виявлення сутностей у тексті	2	тиждень 8
Практичне заняття 6	Класифікація тексту	2	тиждень 9
Самостійна робота	Виявлення зв'язків між сутностями. Класифікація тексту. Аналіз настрою.	17	тиждень 7 тиждень 8 тиждень 9
Лекція 4	Прикладні задачі: Машинний переклад. Генерація тексту. Скорочений переказ. Виявлення кореферентності. Розрізнення омонімів. Відповіді на запитання.	2	тиждень 10
Практичне заняття 7	Аналіз настрою	2	тиждень 11
Самостійна робота	Скорочений переказ. Виявлення кореферентності. Розрізнення омонімів. Відповіді на запитання.	17	тиждень 10 тиждень 11

Методичні рекомендації до практичних та самостійних занять розміщено СЕЗН ЗНУ Moodle на сторінці дисципліни.

5. Види і зміст контрольних заходів

Вид заняття/ роботи	Вид контрольного заходу	Зміст контрольного заходу*	Критерії оцінювання та термін виконання*	Усього балів
Поточний контроль				
Практичне заняття №1	Практична робота 1	Розміщено в СЕЗН ЗНУ	Повний захист виконаного завдання оцінюється в 8 балів, тільки зданий звіт - 6 балів	8
Практичне заняття №2	Практична робота 2	Розміщено в СЕЗН ЗНУ	Повний захист виконаного завдання оцінюється в 8 балів, тільки зданий звіт - 6 балів	8
Практичне заняття №3	Практична робота 3 Самостійна робота	Розміщено в СЕЗН ЗНУ	Повний захист виконаного завдання оцінюється в 8 балів, тільки зданий звіт - 6 балів	8
	Тест 1	Розміщено в СЕЗН ЗНУ	12 питань ⟩ по 0,5 балів	6
Практичне заняття №4	Практична робота 4	Розміщено в СЕЗН ЗНУ	Повний захист виконаного завдання оцінюється в 8 балів, тільки зданий звіт - 6 балів	8
Практичне заняття №5	Практична робота 5	Розміщено в СЕЗН ЗНУ	Повний захист виконаного завдання оцінюється в 8 балів, тільки зданий звіт - 6 балів	8
Практичне заняття №6	Практична робота 6	Розміщено в СЕЗН ЗНУ	Повний захист виконаного завдання оцінюється в 8 балів, тільки зданий звіт - 6 балів	8
	Тест 2	Розміщено в СЕЗН ЗНУ	12 питань ⟩ по 0,5 балів	6
Усього за поточний контроль				**60**
Підсумковий контроль				
Залік	Теоретичне завдання	Розміщено в СЕЗН ЗНУ	40 тестових питань (вибір правильної відповіді з декількох можливих) ⟩ по 1 балу	**40**
Усього за підсумковий контроль				**40**

Шкала оцінювання ЗНУ: національна та ECTS

За шкалою ECTS	За шкалою університету	За національною шкалою	
		Екзамен	Залік
A	90 – 100 (відмінно)	5 (відмінно)	Зараховано
B	85 – 89 (дуже добре)	4 (добре)	
C	75 – 84 (добре)		
D	70 – 74 (задовільно)	3 (задовільно)	
E	60 – 69 (достатньо)		

FX	35 – 59 (незадовільно – з можливістю повторного складання)	2 (незадовільно)	Не зараховано
F	1 – 34 (незадовільно – з обов'язковим повторним курсом)		

6. Основні навчальні ресурси

Рекомендована література

1. Jurafsky D., Martin J. H. Speech and Language Processing : An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. 3rd ed. Stanford : Stanford University, 2024. 591 p. URL: http://files.znu.edu.ua/files/Bibliobooks/Inshi80/0060166.pdf .
2. Lane H., Howard C., Hapke H. M. Natural Language Processing in Action.Understanding, analyzing, and generating text with Python. Manning, 2019. 545 p.
3. Rao D., McMahan B. Natural Language Processing with PyTorch. Build Intelligent Language Applications Using Deep Learning. O'Reilly. 2019. 210 p.
4. Gad A. F., Jarmouni F. E. Introduction to Deep Learning and Neural Networks with Python™ : A Practical Guide. Amsterdam : Elsevier, 2021. 204 p. URL: http://ebooks.znu.edu.ua/files/Bibliobooks/Inshi66/0048321.zip.
5. Monte-Serrat D. M., Cattani C. The Natural Language for Artificial Intelligence. London : Academic Press, 2021. 234 p. URL: http://files.znu.edu.ua/files/Bibliobooks/Inshi74/0054454/.
6. Rao D., McMahan B. Natural Language Processing with PyTorch : Build Intelligent Language Applications Using Deep Learning. Sebastopol : O'Reilly, 2019. 210 p. URL: http://files.znu.edu.ua/files/Bibliobooks/Inshi80/0060167.pdf.
7. Witten I. H., Frank E., Hall M. A., Pal C. J. Data Mining : Practical Machine Learning Tools and Techniques. 4th ed. Amsterdam : Elsevier, 2017. 654 p. URL: http://ebooks.znu.edu.ua/files/Bibliobooks/ScienceDirect/0046080.zip

7. Регуляції і політики курсу

Відвідування занять. Регуляція пропусків.
Відвідування усіх занять є обов'язковим. Студенти зобов'язані дотримуватися усіх строків, визначених для виконання усіх видів робіт, передбачених даною дисципліною. Пропуски та запізнення на заняття є недопустимими.

Політика академічної доброчесності
Кожний студент зобов'язаний дотримуватися принципів академічної доброчесності. Письмові завдання з використанням часткових або повнотекстових запозичень з інших робіт без зазначення авторства – це *плагіат*. Використання будь-якої інформації (текст, фото, ілюстрації тощо) мають бути правильно процитовані з посиланням на автора! Якщо ви не впевнені, що таке плагіат, фабрикація, фальсифікація, порадьтеся з викладачем. До студентів, у роботах яких буде виявлено списування, плагіат чи інші прояви недоброчесної поведінки можуть бути застосовані різні дисциплінарні заходи (див. посилання на Кодекс академічної доброчесності ЗНУ в додатку до силабусу). Неприпустиме складання роботи, виконаної іншою особою.

Використання комп'ютерів/телефонів на занятті
Використання мобільних телефонів, ноутбуків та інших гаджетів під час лекційних та лабораторних занять дозволяється виключно у навчальних цілях (з активованим режимом «без звуку»).

Комунікація
Комунікація викладача зі студентами здійснюється безпосередньо на заняттях та додатково за допомогою месенджерів (наприклад, Telegram), електронної пошти і в СЕЗН Moodle (форум курсу, приватні повідомлення)

ДОДАТКОВА ІНФОРМАЦІЯ

ГРАФІК ОСВІТНЬОГО ПРОЦЕСУ 2026-2027 н. р. доступний за адресою: https://tinyurl.com/yckze4jd.

НАВЧАЛЬНИЙ ПРОЦЕС ТА ЗАБЕЗПЕЧЕННЯ ЯКОСТІ ОСВІТИ. Перевірка набутих студентами знань, навичок та вмінь (атестації, заліки, іспити та інші форми контролю) є невід'ємною складовою системи забезпечення якості освіти і проводиться відповідно до Положення про організацію та методику проведення поточного та підсумкового семестрового контролю навчання студентів ЗНУ: https://tinyurl.com/y9tve4lk.

ПОВТОРНЕ ВИВЧЕННЯ ДИСЦИПЛІН, ВІДРАХУВАННЯ. Наявність академічної заборгованості до 6 навчальних дисциплін (в тому числі проходження практики чи виконання курсової роботи) за результатами однієї екзаменаційної сесії є підставою для надання студенту права на повторне вивчення зазначених навчальних дисциплін. Порядок повторного вивчення визначається Положенням про порядок повторного вивчення навчальних дисциплін та повторного навчання у ЗНУ: https://tinyurl.com/y9pkmmp5. Підстави та процедури відрахування студентів, у тому числі за невиконання навчального плану, регламентуються Положенням про порядок переведення, відрахування та поновлення студентів у ЗНУ: https://tinyurl.com/ycds57la.

ВИРІШЕННЯ КОНФЛІКТІВ. Порядок і процедури врегулювання конфліктів, пов'язаних із корупційними діями, зіткненням інтересів, різними формами дискримінації, сексуальними домаганнями, міжособистісними стосунками та іншими ситуаціями, що можуть виникнути під час навчання, регламентуються Положенням про порядок і процедури вирішення конфліктних ситуацій у ЗНУ: https://tinyurl.com/57wha734. Конфліктні ситуації, що виникають у сфері стипендіального забезпечення здобувачів вищої освіти, вирішуються стипендіальними комісіями факультетів, коледжів та університету в межах їх повноважень, відповідно до: Положення про порядок призначення і виплати академічних стипендій у ЗНУ: https://tinyurl.com/yd6bq6p9; Положення про призначення та виплату соціальних стипендій у ЗНУ: https://tinyurl.com/y9r5dpwh.

ПСИХОЛОГІЧНА ДОПОМОГА. Телефон довіри практичного психолога **Марті Ірини Вадимівни** (061) 228-15-84, (099) 253-78-73 (щоденно з 9 до 21).

УПОВНОВАЖЕНА ОСОБА З ПИТАНЬ ЗАПОБІГАННЯ ТА ВИЯВЛЕННЯ КОРУПЦІЇ Запорізького національного університету: **Банах Віктор Аркадійович**
Електронна адреса: v_banakh@znu.edu.ua
Гаряча лінія: тел. (061) 227-12-76, факс 227-12-88

РІВНІ МОЖЛИВОСТІ ТА ІНКЛЮЗИВНЕ ОСВІТНЄ СЕРЕДОВИЩЕ. Центральні входи усіх навчальних корпусів ЗНУ обладнані пандусами для забезпечення доступу осіб з інвалідністю та інших маломобільних груп населення. Допомога для здійснення входу у разі потреби надається черговими охоронцями

навчальних корпусів. Якщо вам потрібна спеціалізована допомога, будь ласка, зателефонуйте (061) 228-75-11 (начальник охорони). Порядок супроводу (надання допомоги) осіб з інвалідністю та інших маломобільних груп населення у ЗНУ: https://tinyurl.com/ydhcsagx.

РЕСУРСИ ДЛЯ НАВЧАННЯ

НАУКОВА БІБЛІОТЕКА: http://library.znu.edu.ua. Графік роботи абонементів: понеділок-п`ятниця з 08.00 до 16.00; вихідні дні: субота і неділя.
СИСТЕМА ЕЛЕКТРОННОГО ЗАБЕЗПЕЧЕННЯ НАВЧАННЯ (MOODLE): https://moodle.znu.edu.ua
Якщо забули пароль/логін, направте листа з темою «Забув пароль/логін» за адресою: moodle.znu@znu.edu.ua.
У листі вкажіть: прізвище, ім'я, по-батькові українською мовою; шифр групи; електронну адресу.
Якщо ви вказували електронну адресу в профілі системи Moodle ЗНУ, то використовуйте посилання для відновлення паролю https://moodle.znu.edu.ua/mod/page/view.php?id=133015.
ЦЕНТР ІНТЕНСИВНОГО ВИВЧЕННЯ ІНОЗЕМНИХ МОВ: http://sites.znu.edu.ua/child-advance/
ЦЕНТР НІМЕЦЬКОЇ МОВИ, ПАРТНЕР ГЕТЕ-ІНСТИТУТУ: https://www.znu.edu.ua/ukr/edu/ocznu/nim
ШКОЛА КОНФУЦІЯ (ВИВЧЕННЯ КИТАЙСЬКОЇ МОВИ): http://sites.znu.edu.ua/confucius