



## ОБРОБКА ПРИРОДНИХ МОВ

**Викладач:** к. т. н., ст. викл. кафедри комп'ютерних наук, Добровольський Геннадій Анатолійович

**Кафедра:** кафедра комп'ютерних наук, I корпус, ауд. 39

**E-mail:** gen.dobr@gmail.com

**Телефон:** (061) 289-12-57

**Інші засоби зв'язку:** Moodle (форум курсу, приватні повідомлення), Telegram:

[https://t.me/gen\\_dobr](https://t.me/gen_dobr)

<b>Освітня програма, рівень вищої освіти:</b>		Інформаційні системи та технології Магістр					
<b>Статус дисципліни:</b>		Вибіркова					
<b>Кредити ECTS</b>	3	<b>Навч. рік:</b>	2023-24	<b>Рік навчання</b>	2	<b>Тижні</b>	12
<b>Кількість годин</b>	90	<b>Кількість змістових модулів</b>	4	<b>Лекційні заняття – 24</b> <b>Лабораторні заняття – 24</b> <b>Самостійна робота – 42</b>			
<b>Вид контролю:</b>		Залік					
<b>Посилання на курс в Moodle</b>			<a href="https://moodle.znu.edu.ua/course/view.php?id=15093">https://moodle.znu.edu.ua/course/view.php?id=15093</a>				
<b>Консультації:</b>							

### ОПИС КУРСУ

В межах дисципліни «Обробка природних мов» вивчаються методи дослідження, попередньої обробки та перетворення текстових даних.

**Метою** вивчення дисципліни «Обробка природних мов» є оволодіння знаннями про основні методи комп'ютерного моделювання природних мов та застосування моделей до вирішення практичних задач.

**Основними завданнями** вивчення дисципліни «Обробка природних мов» є:

- оволодіння базовими знаннями щодо особливостей обробки природних мов;
- набуття вмінь та навичок створення числових представлень елементів природної мови — слів, словосполучень, речень, довгох текстів;
- ознайомлення з принципами застосування нейронних мереж до обробки текстів;
- набуття вмінь та навичок класифікації, кластеризації, перетворення текстів.

### ОЧІКУВАНІ РЕЗУЛЬТАТИ НАВЧАННЯ

У результаті вивчення дисципліни студенти повинні:

**знати** основні методи моделювання природних мов, методи класифікації, кластеризації, перетворення текстів.

**вміти:**

- створювати словники;
- створювати векторні представлення слів та речень;
- обчислювати семантичну схожість слів та речень;
- групувати тексти за змістом;
- аналізувати структуру тексту;
- моделювати текст як послідовність слів;
- створювати моделі послідовність-в-послідовність;
- застосовувати модель уваги;



- створювати чат-боти;
- знаходити в текстах іменовані сутності та зв'язки між ними.

Згідно з вимогами освітньо-професійної програми студенти повинні досягти таких **компетенцій**:

- ЗК 1 Здатність до абстрактного мислення, аналізу та синтезу.
- ЗК 5 Здатність оцінювати та забезпечувати якість виконуваних робіт.
- СК 4 Здатність розробляти математичні, інформаційні та комп'ютерні моделі об'єктів і процесів інформатизації.
- СК 8 Здатність розробляти та використовувати методи штучного інтелекту для аналізу даних і прогнозування в інформаційних системах.

та наступних **програмних результатів навчання**:

- РН 8 Розробляти моделі інформаційних процесів та систем різного класу, використовувати методи моделювання, формалізації, алгоритмізації та реалізації моделей з використанням сучасних комп'ютерних засобів.
- РН 12 Розробляти та використовувати методи штучного інтелекту.
- РН 14 Розробляти мережі штучних нейронів.

## ОСНОВНІ НАВЧАЛЬНІ РЕСУРСИ

Презентації лекцій, методичні рекомендації до виконання лабораторних робіт та індивідуального завдання, тести у системі Moodle

## КОНТРОЛЬНІ ЗАХОДИ

### Поточні контрольні заходи

Поточний контроль передбачає такі **теоретичні** завдання:

- усне опитування з теоретичного матеріалу за темою на початку кожного лабораторного заняття. Перелік питань з кожної лабораторної роботи розміщено у файлі з завданням до лабораторної роботи у системі Moodle.
- поточний тест за пройденим матеріалом.

Поточний контроль передбачає таке **практичне** завдання:

- виконання лабораторних робіт.

### Підсумкові контрольні заходи:

**Індивідуальне завдання** – Чат-бот.

**Підсумковий тест** – підсумкове тестування з курсу у системі Moodle.

**ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**  
**МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ**  
**Силабус навчальної дисципліни**



Контрольний захід		Термін виконання	% від загальної оцінки
<b>Поточний контроль (max 60%)</b>			
Змістовий модуль 1	Лабораторна робота 1	Тиждень 1, 2, 3	10
Змістовий модуль 2	Лабораторна робота 2	Тиждень 4, 5, 6	10
	Поточний тест №1	Тиждень 6	10
Змістовий модуль 3	Лабораторна робота 3	Тиждень 7, 8, 9	10
Змістовий модуль 4	Лабораторна робота 4	Тиждень 10, 11, 12	10
	Поточний тест №2	Тиждень 12	10
<b>Підсумковий контроль (max 40%)</b>			
Підсумковий тест			20
Індивідуальне завдання			20
<b>Разом</b>			<b>100%</b>

**Шкала оцінювання: 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)		

**РОЗКЛАД КУРСУ ЗА ТЕМАМИ І КОНТРОЛЬНІ ЗАВДАННЯ**

Тиждень і вид заняття	Тема заняття	Контрольне завдання	Кількість балів
Змістовий модуль 1.			
Тиждень 1 Лекція 1	Предмет, типові задачі та методи обробки природних мов,		
Тиждень 2 Лекція 2	Створення словника, векторна модель документа.		
Тиждень 3 Лекція 3	Тематичне моделювання.		
Тиждень 1, 2, 3 Лабораторна робота 1	Класифікація новин	Усне опитування з теми. Виконання завдань лабораторної роботи	10
Змістовий модуль 2.			
Тиждень 4 Лекція 4	Прості нейронні мережі та їх навчання.		
Тиждень 5 Лекція 5	Представлення слів у $R^n$ . Word2Vec. FastText. Оперування		

**ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**  
**МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ**  
**Силабус навчальної дисципліни**



	представленнями		
Тиждень 6 Лекція 6	Згорточні нейронні мережі для обробки текстів		
Тиждень 4, 5, 6 Лабораторна робота 2	Аналіз настроїв у відгуках	Усне опитування з теми. Виконання завдань лабораторної роботи	10
Тиждень 6 Поточний тест			10
<b>Змістовий модуль 3.</b>			
Тиждень 7 Лекція 7	Моделювання тексту за допомогою рекурентних нейронних мереж. RNN		
Тиждень 8 Лекція 8	Моделювання тексту за допомогою нейронної мережі з довгою короткочасною пам'яттю. LSTM		
Тиждень 9 Лекція 9	Методи генерації текстів		
Тиждень 7, 8, 9 Лабораторна робота 4	Генерація тексту	Усне опитування з теми. Виконання завдань лабораторної роботи	10
<b>Змістовий модуль 4.</b>			
Тиждень 10 Лекція 10	Перетворення послідовності в послідовність		
Тиждень 11 Лекція 11	Модель кодування-декодування		
Тиждень 12 Лекція 12	Механізм уваги		
Тиждень 10, 11, 12 Лабораторна робота 5	Модель кодування-декодування	Усне опитування з теми. Виконання завдань лабораторної роботи	10
Тиждень 12 Поточний тест			10

## ОСНОВНІ ДЖЕРЕЛА

1. Daniel Jurafsky, James H. Martin Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Third Edition. Stanford University and University of Colorado at Boulder. 2023. 636 p.
2. Hobson Lane, Cole Howard, Hannes Max Hapke. Natural Language Processing in Action. Understanding, analyzing, and generating text with Python. Manning. 2019. 545 p.
3. Delip Rao and Brian McMahan. Natural Language Processing with PyTorch. Build Intelligent Language Applications Using Deep Learning. O'Reilly. 2019. 210 p.

## ІНФОРМАЦІЙНІ РЕСУРСИ

4. "10 Standard Datasets for Practicing Applied Machine Learning"  
<https://machinelearningmastery.com/standard-machine-learning-datasets/> Last Updated on October 20, 2021.



## РЕГУЛЯЦІЯ 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)



ДОДАТОК ДО СИЛАБУСУ ЗНУ – 2023-2024 рр.

ГРАФІК ОСВІТНЬОГО ПРОЦЕСУ 2023-2024 н. р. доступний за адресою:  
<https://tinyurl.com/yckze4jd>.

**АКАДЕМІЧНА ДОБРОЧЕСНІСТЬ.** Студенти і викладачі Запорізького національного університету несуть персональну відповідальність за дотримання принципів академічної доброчесності, затверджених **Кодексом академічної доброчесності ЗНУ**: <https://tinyurl.com/ya6yk4ad>. Декларація академічної доброчесності здобувача вищої освіти (додається в обов'язковому порядку до письмових кваліфікаційних робіт, виконаних здобувачем, та засвідчується особистим підписом): <https://tinyurl.com/y6wzzlu3>.

**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 методичку 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>.

**НЕФОРМАЛЬНА ОСВІТА.** Порядок зарахування результатів навчання, підтверджених сертифікатами, свідоцтвами, іншими документами, здобутими поза основним місцем навчання, регулюється Положенням про порядок визнання результатів навчання, отриманих у неформальній освіті: <https://tinyurl.com/y8gbt4xs>.

**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/ycyfws9v>. 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: **Борисов Костянтин Борисович**

Електронна адреса: [uv@znu.edu.ua](mailto:uv@znu.edu.ua) Гаряча лінія: Тел. (061) 228-75-50

**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-

**ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**  
**МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ**  
**Силабус навчальної дисципліни**



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.

**ЕЛЕКТРОННЕ 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](mailto: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>