

Силабус
освітнього компоненту
Автоматична обробка мови та мовлення

Назва дисципліни:	Автоматична обробка мови та мовлення
Рівень вищої освіти:	перший (бакалаврський)
Сторінка курсу в Moodle:	https://dl.khadi.kharkov.ua/course/view.php?id=4836
Обсяг освітнього компоненту	4 кредити (120 годин)
Форма підсумкового контролю	Залік
Консультації:	за графіком
Назва кафедри:	кафедра автоматизації та комп'ютерно-інтегрованих технологій
Мова викладання:	українська
Керівник курсу:	Гурко Олександр Геннадійович, д.т.н., професор
Контактний телефон:	(057) 738-77-92
E-mail:	gurko@khadi.kharkov.ua

Короткий зміст освітнього компоненту:

Метою вивчення дисципліни є оволодіння з технологіями автоматичної обробки природної мови.

Основними завданнями вивчення навчальної дисципліни є:

- отримання знань про лінгвістичне забезпечення комп'ютерних систем і баз даних;
- отримання знань про принципи роботи сучасних інформаційних технологій під час розв'язання завдань професійної діяльності;
- формування навичок використання технології автоматичного опрацювання природної мови та штучного інтелекту в різних сферах діяльності для розв'язання пізнавальних завдань.

Передумови для вивчення освітнього компоненту:

Основи інформаційних технологій; Комп'ютерні системи обробки текстової та графічної інформації.

Компетентності, яких набуває здобувач:

Загальні компетентності:

- Здатність спілкуватися державною мовою як усно, так і письмово.
- Здатність учитися й оволодівати сучасними знаннями.
- Здатність до пошуку, опрацювання та аналізу інформації з різних джерел.
- Здатність застосовувати знання у практичних ситуаціях.
- Навички використання інформаційних і комунікаційних технологій.

Спеціальні (фахові) компетентності:

- Здатність використовувати в професійній діяльності знання про мову як особливу знакову систему, її природу, функції, рівні.
- Здатність вільно оперувати спеціальною термінологією для розв'язання професійних завдань.

Результати навчання відповідно до освітньої програми:

1. Ефективно 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.
2. Організовувати процес свого навчання й самоосвіти.
3. Використовувати інформаційні й комунікаційні технології для вирішення складних спеціалізованих задач і проблем професійної діяльності.

Тематичний план

№ теми	Назва тем (ЛК, ЛР, ПР, СЗ, СР)	Кількість годин
		очна
1	ЛК Вступ. Штучний інтелект у лінгвістиці. Регулярні вирази, нормалізація тексту, відстань редагування.	2
	ПР (ЛР) Написання регулярних виразів. Алгоритм мінімальної відстані редагування.	2
	СР. Класифікація текстів. Визначення настрою тексту.	14
2	ЛК N-грамові мовні моделі	4
	ПР (ЛР) Оцінка біграм та триграм.	2
	СР. Згладжування. Великі мовні моделі.	14
3	ЛК Нейронні мережі та нейромовні моделі	4
	ПР (ЛР) Прямий висновок у нейромовній моделі	4
	СР. Навчання нейронних мереж.	14
4	ЛК Машинний переклад	2
	ПР (ЛР) Оцінка ефективності машинного перекладу. Порівняння систем машинного перекладу.	4
	СР. Упередженість та етичні питання при машинному перекладі.	14
5	ЛК Відповіді на запитання та пошук інформації	2
	ПР (ЛР) Оцінка відповідей на фактологічні питання.	2
	СР. Чат-боти та діалогові системи	16
6	ЛК Автоматичне розпізнавання мови та перетворення тексту на мову	2
	ПР (ЛР) Оцінка автоматичного розпізнавання мови.	2
	СР. Перетворення тексту на мову	16
Разом	ЛК	16
	ПР	16
	СР	88

Індивідуальне навчально-дослідне завдання: не передбачене

Методи навчання:

- словесний метод (пояснення, дискусія, бесіда тощо);
- практичний метод (практичні заняття / лабораторні заняття);
- наочний метод (метод ілюстрацій, метод демонстрацій);
- робота з науково-методичною літературою / робота з довідниковою літературою;
- самостійна робота.

Система оцінювання та вимоги:

Поточна успішність

1 Поточна успішність здобувачів за виконання навчальних видів робіт на навчальних заняттях і за виконання завдань самостійної роботи оцінюється за допомогою чотирибальної шкали оцінок з наступним перерахуванням у 100-бальною шкалу. Під час оцінювання поточної успішності враховуються всі види робіт, передбачені навчальною програмою.

1.1 Лекційні заняття оцінюються шляхом визначення якості виконання конкретизованих завдань.

1.2 Практичні заняття оцінюються якістю виконання контрольного або індивідуального завдання, виконання та оформлення практичної роботи.

1.3 Лабораторні заняття оцінюються якістю виконання звітів про виконання лабораторних робіт.

1.4 Семінарські заняття оцінюються якістю виконання індивідуального завдання/реферату.

2 Оцінювання поточної успішності здобувачів вищої освіти здійснюється на кожному практичному занятті (лабораторному чи семінарському) за чотирибальною шкалою («5», «4», «3», «2») і заносяться у журнал обліку академічної успішності.

– «відмінно»: здобувач бездоганно засвоїв теоретичний матеріал, демонструє глибокі знання з відповідної теми або навчальної дисципліни, основні положення;

– «добре»: здобувач добре засвоїв теоретичний матеріал, володіє основними аспектами з першоджерел та рекомендованої літератури, аргументовано викладає його; має практичні навички, висловлює свої міркування з приводу тих чи інших проблем, але припускається певних неточностей і похибок у логіці викладу теоретичного змісту або при аналізі практичного;

– «задовільно»: здобувач в основному опанував теоретичні знання навчальної теми, або дисципліни, орієнтується у першоджерелах та рекомендованій літературі, але непереконливо відповідає, плутає поняття, невпевнено відповідає на додаткові питання, не має стабільних знань; відповідаючи на питання практичного характеру, виявляє неточність у знаннях, не вміє оцінювати факти та явища, пов'язувати їх із майбутньою професією;

– «незадовільно»: здобувач не опанував навчальний матеріал теми (дисципліни), не знає наукових фактів, визначень, майже не орієнтується в першоджерелах та рекомендованій літературі, відсутнє наукове мислення, практичні навички не сформовані.

3 Підсумковий бал за поточну діяльність визнається як середньоарифметична сума балів за кожне заняття, за індивідуальну роботу, поточні контрольні роботи за формулою:

$$K^{поточ} = \frac{K1 + K2 + \dots + Kn}{n},$$

де $K^{поточ}$ – підсумкова оцінка успішності за результатами поточного контролю;

$K1, K2, \dots, Kn$ – оцінка успішності n -го заходу поточного контролю;

n – кількість заходів поточного контролю.

Оцінки конвертуються у бали згідно шкали перерахунку (таблиця 1).

Таблиця 1 – Перерахунок середньої оцінки за поточну діяльність у багатобальну шкалу

4-бальна шкала	100-бальна шкала	4-бальна шкала	100-бальна шкала	4-бальна шкала	100-бальна шкала	4-бальна шкала	100-бальна шкала
5	100	4,45	89	3,90	78	3,35	67
4,95	99	4,4	88	3,85	77	3,3	66
4,9	98	4,35	87	3,80	76	3,25	65
4,85	97	4,3	86	3,75	75	3,2	64
4,8	96	4,25	85	3,7	74	3,15	63
4,75	95	4,20	84	3,65	73	3,1	62
4,7	94	4,15	83	3,60	72	3,05	61
4,65	93	4,10	82	3,55	71	3	60
4,6	92	4,05	81	3,5	70	від 1,78 до 2,99	
						повторне складання	
4,55	91	4,00	80	3,45	69	від 0 до 1,77	від 0 до 34
4,5	90	3,95	79	3,4	68	повторне вивчення	

Підсумкове оцінювання

1 Екзамен проводиться після вивчення всіх тем дисципліни і складається здобувачами вищої освіти в період екзаменаційної сесії після закінчення всіх аудиторних занять

2 До екзамену допускаються здобувачі вищої освіти, які виконали всі види робіт передбачені навчальним планом з дисципліни:

- були присутні на всіх аудиторних заняттях (лекції, семінари, практичні);
- своєчасно відпрацювали всі пропущені заняття;
- набрали мінімальну кількість балів за поточну успішність (не менше 36 балів, що відповідає за національною шкалою «3»);

Якщо поточна успішність з дисципліни нижче ніж 36 балів, здобувач вищої освіти має можливість підвищити свій поточний бал до мінімального до початку екзаменаційної сесії.

3 Оцінювання знань здобувачів при складанні екзамену здійснюється за 100-бальною шкалою.

Оцінювання знань здобувачів шляхом тестування здійснюється за шкалою:

- «Відмінно»: не менше 90 % правильних відповідей;
- «Дуже добре»: від 82 % до 89 % правильних відповідей;
- «Добре»: від 74 % до 81 % правильних відповідей;
- «Задовільно»: від 67 % до 73% правильних відповідей;
- «Задовільно достатньо»: від 60 % до 66 % правильних відповідей;
- «Незадовільно»: менше 60 % правильних відповідей.

4 Підсумкова оцінка з навчальної дисципліни визначається як середньозважена оцінка, що враховує загальну оцінку за поточну успішність і оцінку за складання екзамену.

5 Розрахунок загальної підсумкової оцінки за вивчення навчальної дисципліни проводиться за формулою:

$$PK^{екз} = 0,6 \cdot K^{поточ} + 0,4 \cdot E,$$

де $PK^{екз}$ – підсумкова оцінка успішності з дисциплін, формою підсумкового контролю для яких є екзамен;

$K^{поточ}$ – підсумкова оцінка успішності за результатами поточного контролю (за 100-бальною шкалою);

E - оцінка за результатами складання екзамену (за 100-бальною шкалою).

0,6 і 0,4 – коефіцієнти співвідношення балів за поточну успішність і складання екзамену.

6 За виконання індивідуальної самостійної роботи та участь у наукових заходах здобувачам нараховуються додаткові бали.

6.1 Додаткові бали додаються до суми балів, набраних здобувачем вищої освіти за поточну навчальну діяльність (для дисциплін, підсумковою формою контролю для яких є залік), або до підсумкової оцінки з дисципліни, підсумковою формою контролю для якої є екзамен.

6.2 Кількість додаткових балів, яка нараховується за різні види індивідуальних завдань, залежить від їх об'єму та значимості:

- призові місця з дисципліни на міжнародному / всеукраїнському конкурсі наукових студентських робіт – 20 балів;
- призові місця з дисципліни на всеукраїнських олімпіадах – 20 балів;
- участь у міжнародному / всеукраїнському конкурсі наукових студентських робіт – 15 балів
- участь у міжнародних / всеукраїнських наукових конференціях студентів та молодих вчених – 12 балів;

- участь у всеукраїнських олімпіадах з дисципліни – 10 балів
- участь в олімпіадах і наукових конференціях ХНАДУ з дисципліни – 5 балів;
- виконання індивідуальних науково-дослідних (навчально-дослідних) завдань підвищеної складності – 5 балів.

6.3 Кількість додаткових балів не може перевищувати 20 балів.

7 Загальна підсумкова оцінка за вивчення навчальної дисципліни не може перевищувати 100 балів.

Загальна підсумкова оцінка за вивчення навчальної дисципліни визначається згідно зі шкалою, наведеною в таблиці 2.

Таблиця 2 – Шкала оцінювання знань здобувачів за результатами підсумкового контролю з навчальної дисципліни

Оцінка в балах	Оцінка за національною шкалою		Оцінка за шкалою ЄКТС	
	екзамен	залік	Оцінка	Критерії
90-100	Відмінно	Зараховано	A	Теоретичний зміст курсу освоєний цілком, без прогалин, необхідні практичні навички роботи з освоєним матеріалом сформовані, усі передбачені програмою навчання навчальні завдання виконані, якість їхнього виконання оцінено числом балів, близьким до максимального
80–89	Добре	Зараховано	B	Теоретичний зміст курсу освоєний цілком, без прогалин, необхідні практичні навички роботи з освоєним матеріалом в основному сформовані, усі передбачені програмою навчання навчальні завдання виконані, якість виконання більшості з них оцінено числом балів, близьким до максимального
75-79			C	Теоретичний зміст курсу освоєний цілком, без прогалин, деякі практичні навички роботи з освоєним матеріалом сформовані недостатньо, усі передбачені програмою навчання навчальні завдання виконані, якість виконання жодного з них не оцінено мінімальним числом балів, деякі види завдань виконані з помилками
67-74	Задовільно		D	Теоретичний зміст курсу освоєний частково, але прогалини не носять істотного характеру, необхідні практичні навички роботи з освоєним матеріалом в основному сформовані, більшість передбачених програмою навчання навчальних завдань виконано, деякі з виконаних завдань, можливо, містять помилки
60–66			E	Теоретичний зміст курсу освоєний частково, деякі практичні навички роботи не сформовані, багато передбачених програмою навчання навчальних завдань не виконані, або якість виконання деяких з них оцінено числом балів, близьким до мінімального.

Оцінка в балах	Оцінка за національною шкалою		Оцінка за шкалою ЄКТС	
	екзамен	залік	Оцінка	Критерії
35–59	Незадовільно	Не зараховано	FX	Теоретичний зміст курсу освоєний частково, необхідні практичні навички роботи не сформовані, більшість передбачених програм навчання навчальних завдань не виконано, або якість їхнього виконання оцінено числом балів, близьким до мінімального; при додатковій самостійній роботі над матеріалом курсу можливе підвищення якості виконання навчальних завдань (з можливістю повторного складання)
0–34			F	Теоретичний зміст курсу не освоєно, необхідні практичні навички роботи не сформовані, усі виконані навчальні завдання містять грубі помилки, додаткова самостійна робота над матеріалом курсу не приведе до якого-небудь значущого підвищення якості виконання навчальних завдань (з обов'язковим повторним курсом)

Політика курсу:

- курс передбачає роботу в колективі, середовище в аудиторії є дружнім, творчим, відкритим до конструктивної критики;
- освоєння дисципліни передбачає обов'язкове відвідування лекцій і практичних занять, а також самостійну роботу;
- самостійна робота передбачає вивчення окремих тем навчальної дисципліни, які винесені відповідно до програми на самостійне опрацювання, або ж були розглянуті стисло;
- усі завдання, передбачені програмою, мають бути виконані у встановлений термін;
- якщо здобувач вищої освіти відсутній на заняттях з поважної причини, він презентує виконані завдання під час самостійної підготовки та консультації викладача;
- курсова робота повинна бути захищена не пізніше, ніж за тиждень до початку екзаменаційної сесії;
- під час вивчення курсу здобувачі вищої освіти повинні дотримуватись правил академічної доброчесності, викладених у таких документах: «Правила академічної доброчесності учасників освітнього процесу ХНАДУ» (https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz_67_01_dobroch_1.pdf), «Академічна доброчесність. Перевірка тексту академічних, наукових та кваліфікаційних робіт на плагіат» (https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz_85_1_01.pdf), «Морально-етичний кодекс учасників освітнього процесу ХНАДУ» (https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz_67_01_MEK_1.pdf).
- у разі виявлення факту плагіату здобувач отримує за завдання 0 балів і повинен повторно виконати завдання, які передбачені у силабусі;
- списування під час контрольних робіт та екзаменів заборонені (в т.ч. із використанням мобільних пристроїв). Мобільні пристрої дозволяється використовувати лише під час он-лайн тестування.

Рекомендована література:

1. Daniel Jurafsky. Speech and Language Processing. Pearson, 2016. 636 p.
2. Булгакова О. С. Методи та системи штучного інтелекту: теорія та практика : навч. посіб. / О. С. Булгакова, В. В. Зосімов, В. О. Поздєєв ; Миколаїв. нац. ун-т ім. В.О. Сухомлинського. Херсон : ОЛДІ-ПЛЮС, 2020. 353 с.
3. Савченко О.С. Методи та системи штучного інтелекту: навч. посібник / А.С. Савченко, О.О. Синельников. К.: НАУ, 2017. 176 с.

Додаткові джерела:

1. Субботін С.О. Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень. Навчальний посібник / С.О. Субботів. – Запоріжжя. ЗНТУ, 2008. – 341 с.
2. Introduction to Information Retrieval. URL: <https://nlp.stanford.edu/IR-book/>.
3. SAS Visual Text Analytics. URL: https://www.sas.com/ru_ua/software/viya.html.

Розробник
силабусу навчальної дисципліни


підпис

Олександр ГУРКО
ПІБ

Завідувач кафедри


підпис

Олександр ГУРКО
ПІБ