



**ЧЕРКАСЬКИЙ
НАЦІОНАЛЬНИЙ
УНІВЕРСИТЕТ**
імені Богдана Хмельницького

**Силабус навчальної дисципліни
«ВЕЛИКІ МОВНІ МОДЕЛІ ТА
ОБЧИСЛЮВАЛЬНА ЛІНГВІСТИКА»**

	Статус дисципліни: навчальна дисципліна вибіркового компонента			
Галузь знань	Інформаційні технології			
Спеціальність	F1 «Прикладна математика» F6 «Інформаційні системи та технології»			
Освітня програма	Прикладна математика Веб-орієнтовані інформаційні системи			
Ступінь вищої освіти	Магістр			
Форма навчання	Денна			
Курс	1-2			
Семестр	1-3			
Обсяг дисципліни	Кредити	4	Години	120
Семестровий контроль	Залік			
Викладач	Красношлик Н.О., к.т.н.			
Контактна інформація	krasnoshlyk@vu.cdu.edu.ua			
Кафедра	кафедра прикладної математики та інформатики			
Навчально-науковий інститут / Факультет	ННІ інформаційних та освітніх технологій			
Мова викладання	українська			
Предмет навчання (Що буде вивчатися)	– основи обчислювальної лінгвістики та NLP; – токенізація, лематизація та морфологічний аналіз; – векторні представлення слів (Word2Vec, GloVe, FastText); – архітектури нейронних мереж для NLP (RNN, LSTM, GRU); – механізм уваги (Attention) та трансформери; – архітектура моделей BERT, GPT, T5; – великі мовні моделі (LLM): GPT-4, Claude, LLaMA, Gemini; – методи навчання LLM (pre-training, fine-tuning, RLHF); – prompt engineering та in-context learning; – методи адаптації моделей (LoRA, QLoRA, prefix tuning); – оцінка якості мовних моделей та метрики; – етичні аспекти та обмеження LLM; – практичні застосування: чат-боти, суммаризація, переклад, генерація коду.			
Мета (Чому це цікаво/потрібно вивчати)	Великі мовні моделі здійснили революцію в обробці природної мови та стали одним із найважливіших проривів у штучному інтелекті останніх років. Від			

	ChatGPT до автоматизації бізнес-процесів – LLM трансформують способи взаємодії людини з комп'ютерами та відкривають нові можливості для автоматизації інтелектуальної праці. Вивчення цієї дисципліни дозволяє зрозуміти, як працюють сучасні AI-асистенти, як створювати власні NLP-додатки, як адаптувати великі моделі під специфічні задачі та як інтегрувати їх у реальні продукти. Знання LLM та обчислювальної лінгвістики відкриває можливості для роботи як NLP Engineer, ML Engineer, AI Researcher, а також є критично важливим для розробників, які хочуть інтегрувати AI у свої застосунки. Це одна з найбільш динамічних та перспективних сфер сучасного IT з величезним попитом на фахівців.
Програмні результати (Чому можна навчитися)	– відшуковувати необхідну інформацію в науковій і технічній літературі, базах даних, інших джерелах, аналізувати та оцінювати цю інформацію щодо сучасних досягнень у галузі NLP та LLM; – управляти процесами розробки, впровадження та експлуатації AI-систем на основі великих мовних моделей, які є складними, непередбачуваними і потребують нових стратегічних та командних підходів; – обґрунтовувати вибір архітектури мовних моделей, методів їх навчання та адаптації з урахуванням специфіки задач та обмежень, організовувати їх впровадження та використання; – розробляти інтелектуальні системи обробки природної мови різного класу, використовувати методи моделювання, формалізації та реалізації NLP-рішень з використанням сучасних фреймворків та бібліотек; – розв'язувати задачі автоматизації інтелектуальної праці у нових або невідомих середовищах на основі спеціалізованих концептуальних знань, що включають сучасні наукові здобутки у сфері обчислювальної лінгвістики та великих мовних моделей; – демонструвати знання сучасного рівня технологій NLP з метою їх запровадження у професійній діяльності; знати принципи функціонування великих мовних моделей, методи їх навчання та адаптації; вміти розробляти програмне забезпечення на основі LLM при розв'язанні

	прикладних науково-виробничих задач і задач бізнесу.
Компетентності (Як можна користуватися набутими знаннями і вміннями)	– Здатність до абстрактного мислення, аналізу та синтезу при роботі з мовними моделями та NLP задачами. – Здатність оцінювати та забезпечувати якість розроблюваних NLP-систем та AI-додатків. – Здатність застосовувати знання великих мовних моделей у практичних ситуаціях для вирішення реальних бізнес-задач. – Здатність формувати вимоги до етапів життєвого циклу інтелектуальних систем на основі LLM. – Здатність проектувати NLP-системи з урахуванням особливостей їх призначення, обмежень моделей та етичних аспектів використання AI. – Розробляти і реалізовувати інноваційні проекти у сфері обчислювальної лінгвістики та застосування великих мовних моделей. – Здатність управляти та користуватися сучасними фреймворками та API для роботи з LLM, у першу чергу, орієнтованими на інтеграцію AI у додатки. – Обґрунтовувати вибір методів навчання та адаптації мовних моделей з урахуванням їх взаємодії та потенційного впливу на вирішення організаційних проблем, організовувати їх впровадження та використання. – Розробляти архітектури інтелектуальних систем обробки природної мови різного класу, використовувати методи prompt engineering, fine-tuning та RAG для реалізації моделей з використанням сучасних комп'ютерних засобів. – Розв'язувати задачі впровадження AI-технологій у нових або невідомих середовищах на основі спеціалізованих концептуальних знань, що включають сучасні наукові здобутки у сфері NLP, досліджень та інтеграції знань з різних галузей. – Демонструвати знання сучасного рівня технологій великих мовних моделей з метою їх запровадження у професійній діяльності; знати принципи функціонування трансформерів, архітектури BERT та GPT, методи RLHF; вміти розробляти програмне забезпечення на основі LLM при розв'язанні прикладних науково-виробничих задач і задач бізнесу.

Зміст дисципліни	Змістовий модуль 1 Тема 1. Основи обчислювальної лінгвістики та NLP. Токенізація, лематизація та морфологічний аналіз. Тема 2. Векторні представлення слів (Word2Vec, GloVe, FastText). Методи embeddings. Тема 3. Архітектури нейронних мереж для NLP (RNN, LSTM, GRU). Механізм уваги (Attention). Тема 4. Трансформери та їх архітектура. Моделі BERT, GPT, T5. Змістовий модуль 2 Тема 5. Великі мовні моделі (LLM): GPT-4, Claude, LLaMA, Gemini. Методи навчання LLM (pre-training, fine-tuning, RLHF). Тема 6. Prompt engineering та in-context learning. Методи адаптації моделей (LoRA, QLoRA, prefix tuning). Тема 7. Побудова RAG систем. Використання LangChain для створення LLM-додатків. Тема 8. Практичні застосування: чат-боти, суммаризація, переклад, генерація коду. Оцінка якості та етичні аспекти використання LLM.	
Розподіл годин	Лекційні	12
	Практичні/семінарські	-
	Лабораторні	24
	Самостійна робота	84
Критерії оцінювання роботи студентів	Завданням поточного контролю є систематична перевірка розуміння та засвоєння програмного матеріалу шляхом усного та письмового опитування, аналіз виконання завдань практичних занять, індивідуальної та самостійної роботи, умінь самостійно опрацьовувати навчальний матеріал, здатності публічно, письмово чи в електронному форматі представляти певний матеріал. Критеріями оцінювання у ході поточного контролю є: а) під час поточної аудиторної роботи на лекційних та практичних заняттях: – активна участь у дискусіях та пропонуваннях форм роботи на лекційних та практичних заняттях; – доповнення та запитання на лекційних та практичних заняттях. б) при усних відповідях: – повнота розкриття питання; – логіка викладення, культура мовлення;	

	– впевненість, емоційність та аргументованість; – використання основної та додаткової літератури (підручників, навчальних посібників, журналів, інших періодичних видань, інтернет-ресурсів тощо); – аналітичні міркування, уміння робити порівняння, висновки. в) при виконанні письмових завдань: – повнота розкриття питання; – цілісність, систематичність, логічна послідовність; – підготовка матеріалу за допомогою комп'ютерної техніки, різних технічних засобів. г) при виконанні завдань для самостійної та індивідуальної роботи: – повнота виконання завдання; – творчість та самостійність виконання. Критерієм успішного проходження здобувачем освіти підсумкового оцінювання може бути досягнення ним мінімальних порогових рівнів оцінок за кожним запланованим результатом навчання навчальної дисципліни, який визначається до кожного завдання через якісні критерії і трансформується у мінімальну позитивну оцінку обраної для даної дисципліни шкали. Після завершення курсу використана шкала перенормовується у накопичувальну 100-бальну і ЄКТС (A, B, C, D, E, FX, F) шкали.
Інформаційне забезпечення (лінк на е-НМЗНД)	
Матеріально-технічне забезпечення	Аудиторія теоретичного навчання, комп'ютерний клас для виконання лабораторних робіт з потужними GPU або доступом до cloud платформ, ноутбук, проектор, навчальна та наукова література, презентаційні матеріали, встановлене програмне забезпечення (Python, Hugging Face Transformers, LangChain, векторні бази даних).