



Аналіз та обробка природної мови (NLP)

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Перший (бакалаврський)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>123 Комп'ютерна інженерія, 121 Інженерія програмного забезпечення</i>
Освітня програма	<i>Комп'ютерна інженерія, Інженерія програмного забезпечення</i>
Статус дисципліни	<i>Вибіркова компонента ОП, циклу професійної підготовки</i>
Форма навчання	<i>очна (денна)</i>
Рік підготовки, семестр	<i>Зкурс, 6 семестр</i>
Обсяг дисципліни	<i>4 кредити /120 год.: лекцій 36 годин, лаб. роб. 18 год., СРС 66 год.</i>
Семестровий контроль/ контрольні заходи	<i>Залік</i>
Розклад занять	<i>http://rozklad.kpi.ua/, http://roz.kpi.ua/</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	Лектор: <i>доктор технічних наук, професор Писарчук Олексій Олександрович, pysarchuk.oleksii@ill.kpi.ua</i> Лабораторні: <i>доктор технічних наук, професор Писарчук Олексій Олександрович, pysarchuk.oleksii@ill.kpi.ua</i> <i>Асистент Баран Данило Романович, agd015979@gmail.com</i>
Розміщення курсу	<i>https://classroom.google.com/u/5/c/NzcuNDQ2NTU0OTQ4</i>

Програма навчальної дисципліни

1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

Навчальна дисципліна «Аналіз та обробка природної мови (NLP)» має на меті – набуття студентами здатності синтезувати, верифікувати математичні моделі, розробляти спеціалізоване програмне забезпечення з аналізу та обробки природної мови (Natural language processing, (NLP)). Це досягається вивченням теоретичних основ та прикладних технологій: отримання та попередньої фільтрації повідомлень природною мовою, синтаксичного аналізу повідомлень (токенізація та видалення стоп-слів; нормалізація - стемінг, лемізація; векторизація; аналіз із виділенням частин мови); семантичного (частотний та ймовірнісний) аналізу повідомлень; класифікації; аналізу подібності текстових повідомлень із методами глибинного навчання (у т.ч. з підтримки генеративного штучного інтелекту (GPT)); аналізу та обробки аудіо-повідомлень та інше.

Завдання дисципліни – надання поглиблених компетенцій з методологій та технологій процесів обробки природної мови – Natural Language Processing (NLP). Дисципліна є продовженням технологій Data Science.

Курсу включає розгляд класичних методологій NLP та результати реалізації практичних R&D проектів.

Теоретичні основи Natural language processing надаються у формі лекційних занять з обов'язковою демонстрацією розглянутих алгоритмів у формі прикладів програмного коду.

Практичні навички застосування технологій NLP набуваються на лабораторних заняттях. Практична частина орієнтована на застосування мови програмування високого рівня Python з

вивченням функціоналу бібліотек: Numpy, Pandas, Requests, NLTK, Spacy, Gensim, fastText, Scipy, Scikit-learn, Tensorflow, Keras, PyTorch та інші.

Завдання лабораторних робіт розділені за рівнями складності, побудовані за принципами нарощування функціональності в межах тем.

Зміст курсу збалансовано розкривається суть головних технологічних процесів Natural Language Processing і завершується проектним практикумом.

Курс розраховано на фахівців, що прагнуть опанувати знання, уміння та навички, які потребують посади: NLP Engineer, NLP Researcher, Data Engineer, Backend Software Engineer, Machine Learning Engineer, Data Science, Data Scientist (NLP) тощо.

Матеріали курсу може використовуватися для реалізації практичних проектів:

- Розуміння природної мови та / або генерація природної мови;
- Голосові помічники для обробки тексту та мовлення – типу Alexa, Siri тощо;
- Класифікація тексту, наприклад як Grammarly, Microsoft Word і Google Docs;
- Вилучення інформації, наприклад, як пошукові системи DuckDuckGo, Google;
- Чат-бот і відповіді на запитання, наприклад боти веб-сайтів;
- Мовний переклад, наприклад: Google Translate;
- Конспектування та реферертивна обробка тексту.

Вибіркова навчальна дисципліна «Аналіз та обробка природної мови (NLP)» підсилює фахові/загальні компетенції освітньо-професійних програм (ОПП).

Здатність до абстрактного мислення, аналізу та синтезу.

Здатність використовувати сучасні методи і мови програмування для розроблення алгоритмічного та програмного забезпечення.

Здатність розробляти, адаптувати, використати програмне забезпечення для покращення ефективності застосування високопродуктивних комп'ютерних систем.

Мати навички проведення експериментів, збирання даних та моделювання в комп'ютерних системах.

Вміти застосовувати знання для ідентифікації, формулювання і розв'язування технічних задач спеціальності, використовуючи методи, що є найбільш придатними для досягнення поставлених цілей.

Вміти розв'язувати задачі аналізу та синтезу засобів, характерних для спеціальності.

Вміти розробляти програмне забезпечення для вбудованих і розподілених застосувань, мобільних і гібридних систем, розраховувати, експлуатувати, типове для спеціальності обладнання.

Вміти оцінювати отримані результати та аргументовано захищати прийняті рішення.

Здатність до абстрактного мислення, аналізу та синтезу;

Володіння знаннями про інформаційні моделі даних, здатність створювати програмне забезпечення для зберігання, видобування та опрацювання даних;

Здатність до алгоритмічного та логічного мислення;

Знати і застосовувати відповідні математичні поняття, методи доменного, системного і об'єктно-орієнтованого аналізу та математичного моделювання для розробки програмного забезпечення

Уміння вибирати та використовувати відповідну задачі методологію створення програмного забезпечення.

Знати і застосовувати на практиці фундаментальні концепції, парадигми і основні принципи функціонування мовних, інструментальних і обчислювальних засобів інженерії програмного забезпечення.

Як додатковий результат навчання вибіркова навчальна дисципліна «Аналіз та обробка природної мови (NLP)» спрямована на формування здатності синтезувати, верифікувати математичні моделі, розробляти спеціалізоване програмне забезпечення з обробки і аналізу природної мови.

2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)

Перереквізити:

Успішне опанування дисципліни «Аналіз та обробка природної мови (NLP)» потребує від студента: базових знань з програмування: принципи програмування, алгоритмізація та базові алгоритми; базових знань мови програмування Python: синтаксис, типи та структури даних, базові оператори розгалужених обчислень, функціональне та ООП програмування, робота з IDE, створення оточення; базових знань з математики: елементи теорії ймовірностей, дискретна математика, теорія матриць, дослідження функцій, аналітична геометрія, тригонометрія.

Це підтримується навчальними дисциплінами ОПП:

123 Комп'ютерна інженерія:

Вища математика. Частина 1,2,3; Аналітична геометрія та лінійна алгебра; Програмування. Частина 1,2; Дискретна математика; Фізика; Теорія ймовірності та математична статистика; Структури даних та алгоритми; Організація баз даних; Інженерія програмного забезпечення; Алгоритми та методи обчислень.

121 Інженерія програмного забезпечення:

Математичний аналіз. Частина 1,2; Лінійна алгебра та аналітична геометрія; Теорія ймовірностей; Комп'ютерна дискретна математика; Алгоритми та структури даних. Частина 1,2; Основи програмування. Частина 1,2; Базы даних; Об'єктно-орієнтоване програмування; Методології і технології розроблення програмного забезпечення; Технології штучного інтелекту

Постреквізити:

Дисципліна відноситься до вибіркової компоненти ОПП, циклу професійної підготовки. Курс орієнтовано на прикладні аспекти проектування, синтезу та розробки математичних моделей, алгоритмів та програмного забезпечення складних розподілених інформаційних систем із властивостями інтелектуальності, що є міждисциплінарним зв'язком компонент освітньої програми.

Вибіркова навчальна дисципліна «Аналіз та обробка природної мови (NLP)» є прикладною та спрямована на опанувати знання, уміння та навички - компетенції, які потребують практичні посади: NLP Engineer, NLP Researcher, Data Engineer, Backend Software Engineer, Machine Learning Engineer, Data Science, Data Scientist (NLP) тощо.

Компетенції, отримані з курсу можуть також використовуватись для реалізації практичних питань із дослідження природної мови у реалізації завдань курсового проектування, розробки кваліфікаційних робіт тощо.

3. Зміст навчальної дисципліни

МОДУЛЬ 1. ВСТУП В NLP-TECHNOLOGY

Тема 1. Отримання та попередній аналіз даних.

Лекція 1. Вступ до аналізу та обробки природної мови (NLP): дослідження даних, етапи, природна мова, методи обробки.

Лекція 2. Отримання і попередня фільтрація природної мови: парсинг, очищення, трансформація мовного простору, фільтрація, властивості, форми представлення та зберігання.

Лабораторна робота 1. Отримання і попередня фільтрація природної мови.

МОДУЛЬ 2. ПОПЕРЕДНІЙ АНАЛІЗ ТА ОБРОБКА ПРИРОДНОЇ МОВИ.

Тема 2. Попередня обробка та синтаксичний аналіз природної мови

Лекція 3. Токенізація та видалення стоп-слів в текстових повідомленнях.

Лекція 4. Нормалізація: стемінг, лемізація.

Лабораторна робота 2. Токенізація та нормалізація в NLP.

Лекція 5. Векторизація текстових повідомлень.

Лекція 6. Аналіз текстових повідомлень із виділенням частин мови.

Лабораторна робота 3. Векторизація та аналіз в NLP.

Тема 3. Семантичний аналіз повідомлень природною мовою

Лекція 7. Частотний та ймовірнісний аналіз текстових повідомлень.

Лекція 8. Класифікація текстових повідомлень.

Лабораторна робота 4. Частотний аналіз та кластеризація в NLP.

Лекція 9. Аналіз подібності текстових повідомлень.

Лекція 10. Визначення тональності текстових повідомлень.

Лабораторна робота 5. Аналіз подібності та визначення тональності текстових повідомлень.

МОДУЛЬ 3. МЕТОДИ І ТЕХНОЛОГІЇ ГЛИБИННОГО НАВЧАННЯ ДЛЯ ОБРОБКА ПРИРОДНОЇ МОВИ.

Тема 4. Аналіз текстових повідомлень методами глибокого навчання

Лекція 11. Основи штучних нейронних мереж.

Лекція 12. Штучні нейронні мережі для аналізу текстових повідомлень з TensorFlow.

Лекція 13. Штучні нейронні мережі для аналізу текстових повідомлень з PyTorch.

Лабораторна робота 6. Методи і технології глибокого навчання для NLP.

МОДУЛЬ 4. ПРАКТИКА ОБРОБКА ПРИРОДНОЇ МОВИ.

Тема 5. Проектний практикум з NLP-технологій.

Лекція 14. Аналіз та обробка аудіо-повідомлень.

Лабораторна робота 7. Практика аналізу аудіо-повідомлень

Лекція 15. Обробка природної мови з підтримки генеративного штучного інтелекту (GPT).

Лабораторна робота 8. Практика обробка природної мови з підтримки генеративного штучного інтелекту (GPT).

Лекція 16. Аналіз глобального інформаційного простору.

Лекція 17. Спеціальні можливості NLP-технологій.

Лабораторна робота 9. Практика аналізу глобального інформаційного простору.

Модульна контрольна робота 1 Практика обробка природної мови.

4. Навчальні матеріали та ресурси

4.1. Базова література:

1. Навчально-методичний комплекс з дисципліни: [<https://comsys.kpi.ua/napryami-pidgotovki>].

2. Ланде Д.В., Субач І.Ю., Бояринова Ю.Є. Основи теорії і практики інтелектуального аналізу даних у сфері кібербезпеки: навчальний посібник. – К.: ІСЗІ КПІ ім. Ігоря Сікорського», 2018. – 297 с.

3. Методи обробки природної мови: Феномен ChatGPT [Текст]: навч. посіб. для здобувачів ступеня бакалавра за освітньою програмою «Інженерія програмного забезпечення інтелектуальних кібер-фізичних систем в енергетиці» / автор: Ю. І. Стативка; КПІ ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 621 кбайт). – Київ: КПІ ім. Ігоря Сікорського, 2023. – 67 с.

4. К. Оніщенко та ін. Аналіз Методів Обробки Природної Мови [<https://openarchive.nure.ua/server/api/core/bitstreams/a26452b9-d866-4aaa-aa04-524a9cae5f55/content>].

5. D. Jurafsky, J.H. Martin: Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, 3rd ed. https://web.stanford.edu/~jurafsky/slp3/ed3book_jan72023.pdf

6. Hobson Lane. Natural Language Processing in Action. Understanding, analyzing, and generating text with Python/ Hobson Lane, Cole Howard, Hannes Hapke. – Manning Publications Co., 2019. – 544 p.

7. *Natural Language Processing with Python*: <https://tjzhifei.github.io/resources/NLTK.pdf>
<https://github.com/yuvaraj119/Python-Pdf-Books/blob/master/O.Reilly.-.Natural.Language.Processing.with.Python.pdf>
8. J. Thanaki: *Python Natural Language Processing*. // Packt Publishing, 2018.
9. K. Bhavsar, N. Kumar, P. Dangeti: *Natural Language Processing with Python Cookbook*. // Packt Publishing, 2018.
10. Lewis Tunstall. *Natural Language Processing with Transformers, Revised Edition* / Lewis Tunstall, Leandro von Werra, Thomas Wolf. – Beijing-Boston-Farnham-Sebastopol-Tokyo : O'Reilly Media, Inc., 2022. – 383 p.
11. Sowmya Vajjala. *Practical Natural Language Processing* / Sowmya Vajjala, Bodhisattwa Majumder, Anuj Gupta, Harshit Surana / - O'Reilly Media, 2020 — 456 p.

4.2. Додаткова література:

1. *Natural Language Processing with Python Updated Edition: From Basics to Advanced Projects: Mastering Text Analysis, Machine Learning Models, and Chatbot ...* (Mastering the AI Revolution Book by Quantum Technologies (Author). – 2024
2. *Natural Language Processing with Transformers, Revised Edition 1st Edition, Kindle Edition* by Lewis Tunstall (Author), Leandro von Werra (Author), Thomas Wolf (Author). – 2022.
3. *Transformers for Natural Language Processing: Build, train, and fine-tune deep neural network architectures for NLP with Python, Hugging Face, and OpenAI's GPT-3, ChatGPT, and GPT-4 2nd Edition, Kindle Edition* by Denis Rothman (Author), Antonio Gulli (Foreword). – 2022.
4. *Natural Language Processing with PyTorch: Build Intelligent Language Applications Using Deep Learning 1st Edition, Kindle Edition* by Delip Rao (Author), Brian McMahan (Author). – 2019.

4.3. Інформаційні ресурси:

<https://github.com/LinguisticAndInformationSystems/mphdict>
<https://lang.org.ua/uk/>
https://github.com/UniversalDependencies/UD_Ukrainian-IU
<https://mova.institute/>
<https://spacy.io/models/uk>
<https://realpython.com/nltk-nlp-python/>
<https://www.nltk.org/>
<https://sunscrapers.com/blog/9-best-python-natural-language-processing-nlp/>
<https://www.educative.io/blog/natural-language-processing-with-python-guide>
<https://dou.ua/lenta/articles/first-steps-in-nlp-nltk/>
<https://medium.com/stinopys/%D1%81%D0%BB%D0%BE%D0%B2%D0%BD%D0%B8%D0%BA-nlp-b0fab1027551#ca94>
<https://medium.com/@felixvidalgu/nlp-for-text-classification-43a83deca302>
<https://www.nltk.org/>
<https://www.kaggle.com/>
<https://github.com/PacktPublishing/Artificial-Intelligence-with-Python>
<https://pandas.pydata.org/>
<https://scapy.net/>
<https://developers.google.com/optimization>
<https://www.tensorflow.org/>
<https://keras.io/>

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

Навчальна дисципліна «Аналіз та обробка природної мови (NLP)» включає 4 кредитів/120 год.: лекцій 36 годин, лаб. роб. 18 год., CPC 66 год.

За дисципліною передбачено проведення таких видів аудиторних занять: лекційних занять; лабораторні роботи; модульна контрольна робота; залік.

Лекційні заняття розкривають теоретичні основи основних галузей NLP, їх практичну реалізацію у програмному коді та доповнюються практикою виконання завдань лабораторних робіт.

Практична частина курсу (лабораторні роботи) орієнтована на застосування мови програмування високого рівня Python з опануванням функціоналу бібліотек: Numpy, Pandas, Requests, NLTK, Spacy, Gensim, fastText, Scipy, Scikit-learn, Tensorflow, Keras, PyTorch та інші.

Завдання лабораторних робіт розділені за рівнями складності, побудовані за принципами нарощування функціональності в межах тем.

Тематика лекційних занять.

Лекція 1. Вступ до аналізу та обробки природної мови (NLP): дослідження даних, етапи, природна мова, методи обробки.

Лекція 2. Отримання і попередня фільтрація природної мови: парсинг, очищення, трансформація мовного простору, фільтрація, властивості, форми представлення та зберігання.

Лекція 3. Токенізація та видалення стоп-слів в текстових повідомленнях.

Лекція 4. Нормалізація: стемінг, лемізація.

Лекція 5. Векторизація текстових повідомлень.

Лекція 6. Аналіз текстових повідомлень із виділенням частин мови.

Лекція 7. Частотний та ймовірнісний аналіз текстових повідомлень.

Лекція 8. Класифікація текстових повідомлень.

Лекція 9. Аналіз подібності текстових повідомлень.

Лекція 10. Визначення тональності текстових повідомлень.

Лекція 11. Основи штучних нейронних мереж.

Лекція 12. Штучні нейронні мережі для аналізу текстових повідомлень з TensorFlow.

Лекція 13. Штучні нейронні мережі для аналізу текстових повідомлень з PyTorch.

Лекція 14. Аналіз та обробка аудіо-повідомлень.

Лекція 15. Обробка природної мови з підтримки генеративного штучного інтелекту (GPT).

Лекція 16. Аналіз глобального інформаційного простору.

Лекція 17. Спеціальні можливості NLP-технологій.

Тематика лабораторних занять.

Лабораторна робота 1. Отримання і попередня фільтрація природної мови.

Лабораторна робота 2. Токенізація та нормалізація в NLP.

Лабораторна робота 3. Векторизація та аналіз в NLP.

Лабораторна робота 4. Частотний аналіз та кластеризація в NLP.

Лабораторна робота 5. Аналіз подібності та визначення тональності текстових повідомлень.

Лабораторна робота 6. Методи і технології глибинного навчання для NLP.

Лабораторна робота 8. Практика обробки природної мови з підтримки генеративного штучного інтелекту (GPT).

Лабораторна робота 9. Практика аналізу глобального інформаційного простору.

6. Самостійна робота здобувача вищої освіти денної форми навчання

Види самостійної роботи (66 годин):

самостійне опрацювання матеріалів лекційних занять (1 година x 17 лекцій = 17 годин);

підготовка та оброблення проведення розрахунків за первинними даними, отриманими на лабораторних заняттях, виконання лабораторних робіт, розв'язок задач надання на перевірку (рекомендовано 2 години x 9 лабораторних робіт = 18 годин);

виконання модульної контрольної роботи (МКР = 4 години);

підготовка до заліку (8 години);

налаштування віртуального середовища для виконання лабораторних робіт (3 години);

опрацювання тем на самостійну роботу (16 годин).

Теми на самостійне опрацювання.

МОДУЛЬ 1. ВСТУП В NLP-TECHNOLOGY

Тема 1. Отримання та попередній аналіз даних.

Лекція 1. Вступ до аналізу та обробки природної мови (NLP): дослідження даних, етапи, природна мова, методи обробки.

Лекція 2. Отримання і попередня фільтрація природної мови: парсинг, очищення, трансформація мовного простору, фільтрація, властивості, форми представлення та зберігання.

Лабораторна робота 1. Отримання і попередня фільтрація природної мови.

МОДУЛЬ 2. ПОПЕРЕДНІЙ АНАЛІЗ ТА ОБРОБКА ПРИРОДНОЇ МОВИ.

Тема 2. Попередня обробка та синтаксичний аналіз природної мови

Лекція 3. Токенізація та видалення стоп-слів в текстових повідомленнях.

Лекція 4. Нормалізація: стемінг, лемізація.

Лабораторна робота 2. Токенізація та нормалізація в NLP.

Лекція 5. Векторизація текстових повідомлень.

Лекція 6. Аналіз текстових повідомлень із виділенням частин мови.

Лабораторна робота 3. Векторизація та аналіз в NLP.

Тема 3. Семантичний аналіз повідомлень природною мовою

Лекція 7. Частотний та ймовірнісний аналіз текстових повідомлень.

Лекція 8. Класифікація текстових повідомлень.

Лабораторна робота 4. Частотний аналіз та кластеризація в NLP.

Лекція 9. Аналіз подібності текстових повідомлень.

Лекція 10. Визначення тональності текстових повідомлень.

Лабораторна робота 5. Аналіз подібності та визначення тональності текстових повідомлень.

МОДУЛЬ 3. МЕТОДИ І ТЕХНОЛОГІЇ ГЛИБИННОГО НАВЧАННЯ ДЛЯ ОБРОБКА ПРИРОДНОЇ МОВИ.

Тема 4. Аналіз текстових повідомлень методами глибокого навчання

Лекція 11. Основи штучних нейронних мереж.

Лекція 12. Штучні нейронні мережі для аналізу текстових повідомлень з TensorFlow.

Лекція 13. Штучні нейронні мережі для аналізу текстових повідомлень з PyTorch.

Лабораторна робота 6. Методи і технології глибокого навчання для NLP.

МОДУЛЬ 4. ПРАКТИКА ОБРОБКА ПРИРОДНОЇ МОВИ.

Тема 5. Проектний практикум з NLP-технологій.

Лекція 14. Аналіз та обробка аудіо-повідомлень.

Лабораторна робота 7. Практика аналізу аудіо-повідомлень

Лекція 15. Обробка природної мови з підтримки генеративного штучного інтелекту (GPT).

Лабораторна робота 8. Практика обробки природної мови з підтримки генеративного штучного інтелекту (GPT).

Лекція 16. Аналіз глобального інформаційного простору.

Лекція 17. Спеціальні можливості NLP-технологій.

Лабораторна робота 9. Практика аналізу глобального інформаційного простору.

Модульна контрольна робота 1 Практика обробки природної мови.

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Для виконання лабораторних робіт та модульних контрольних робіт встановлюються дедлайни.

Виконання лабораторних робіт поза встановлених термінів супроводжуються штрафними балами, які вираховуються із оцінки за протокол (п. 2.9. ПОЛОЖЕННЯ про систему оцінювання результатів навчання). МКР не приймається поза встановлені терміни.

Штрафні бали виставляються за: невчасну здачу лабораторної роботи. Кількість штрафних балів не більше 10 (9 лабораторних робіт + МКР). Штрафні бали та жорсткі дедлайни не запроваджуються у період військового стану.

Заохочувальні бали виставляються за: R&D результати виконання лабораторних робіт; активну участь на лекціях; виконання поточних домашніх завдань, ведення конспекту, підготовка повідомлення з презентацією по одній із тем СРС дисципліни тощо. Кількість заохочуваних балів не більше 10.

Виконанню кожної лабораторної роботи передую виконання індивідуального завдання і оформлення його у вигляді протоколу. Студент, який прийшов на заняття без оформленого протоколу до лабораторної роботи не допускається. Першим етапом студент захищає результати отримані під час виконання індивідуального завдання до лабораторної роботи, на другому етапі – захищає теорію шляхом усного опитування або тестування (за необхідності). Захист лабораторних робіт може бути проведено за результатами аналізу повноти і якості виконання протоколу. Бали отримані за виконання лабораторної роботи, за теоретичною частиною та за протокол входять в оцінку за лабораторну роботу.

Перездача лабораторної роботи у разі її позитивної оцінки з метою підвищення оцінки – не передбачено.

Виконання лабораторних робіт є обов'язковими для допуску до семестрового контролю. Умовою допуску до семестрового контролю є зарахування всіх лабораторних робіт та стартовий рейтинг не менше 60 балів.

Модульна контрольна робота пишеться на лекційному занятті без застосування допоміжних засобів (мобільні телефони, планшети та ін.); результат в електронному вигляді надаються викладачеві.

Модульна контрольна робота не переписується за умови негативної оцінки. Негативна оцінка за МКР прирівнюється до 0 балів, в цьому випадку МКР не зараховується.

Оцінка, яку студент може отримати за виконання кожної лабораторної роботи та за модульну контрольну роботу наведені в таблиці 1 оцінювання семестрових робіт, розділ 8 силабусу.

Таким чином мінімальна оцінка, яку повинен отримати студент для допуску до семестрового контролю дорівнює 60 балів, максимальна – 100 балів за виконання всіх поточних робіт за семестр.

Здобувачі, які виконали всі умови допуску до заліку (виконали всі лабораторні роботи) та мають рейтингову оцінку менше 60 балів, а також ті здобувачі, хто бажає підвищити свою рейтингову оцінку, на останньому за розкладом занятті мають можливість пройти семестровий контроль у вигляді залікової контрольної роботи.

У разі виконання залікової контрольної роботи рейтингова оцінка визначається як сума балів за залікову контрольну роботу та балів за індивідуальні семестрові завдання.

8. Види контролю та рейтингова система оцінювання результатів навчання (PCO)

Для навчальної дисципліни «Аналіз та обробка природної мови (NLP)», як прикладної дисципліни, що спрямована на отримання комплексного ґрунтового теоретичного базису та потужних практичних навичок програмної реалізації методів, математичних моделей і алгоритмів технологічних процесів NLP застосовується PCO-1 (п. 2.1. ПОЛОЖЕННЯ про систему оцінювання результатів навчання в КПІ ім. Ігоря Сікорського, https://osvita.kpi.ua/sites/default/files/downloads/Pologennia_RSO_2022.pdf).

Семестровий рейтинг студента розраховується, виходячи із 100-бальної шкали. Рейтинг складається з балів, що студент отримав за виконання 9 лабораторних робіт РЛ та однієї модульної контрольної роботи РМКР.

Завдання лабораторних робіт розділені за рівнями складності. Високий рівень складності передбачає отримання максимум 10 балів, середній рівень – 8 балів.

Максимальна кількість балів за лабораторні роботи (РЛ) за високим рівнем складає 90 балів, за середнім рівнем - 72 бали.

Розподіл балів за виконання лабораторних робіт.

1.1. Якість / повнота оформлення протоколу з лабораторної роботи – 1 бал.

1.2. Своєчасний захист роботи – 1 бал.

1.3. Повнота аналізу отриманих результатів – 1 бал.

1.4. Якість та повнота виконання технічних умов завдання, функціональність розробленої технічної продукції (програмного скрипта) -5 бали.

1.5. Рівень теоретичної підготовки – 2 бали.

*** Для умов дистанційного навчання бали за теоретичну підготовленість (п.1.4) можуть нараховуватись за результатами аналізу вмісту протоколу з лабораторної роботи.

*** Для умов військового стану – своєчасність захисту лабораторної роботи (п.1.2) – не застосовується а додається до п.1.4.

Максимальна кількість балів за МКР - РМКР = 10.

Розподіл балів за виконання МКР.

2.1. Повнота розкриття суті та оригінальність відповіді на теоретичні питання – 2 бали за кожне питання (2 питання білету) – загалом – 2 балів;

2.2. Повнота, оригінальність та якість виконання практичного питання – 6 бали.

У разі **значних запозичень / порушення вимог доброчесності** в звітному протоколі з лабораторної роботи - робота може бути не зарахована, або повернення на переопрацювання. За таких умов МКР та залікова робота – не зараховуються.

Календарна атестація студентів (на 8 та 14 тижнях семестрів) з дисципліни проводиться за значенням поточного рейтингу студента на час атестації. Якщо значення цього рейтингу не менше 50 % від максимально можливого на час атестації, студент вважається атестованим. В іншому випадку в атестаційній відомості виставляється «не атестовано».

Таким чином, порядок визначення загального рейтингу пояснюється Таб.1 та визначається за виразом

$$R = RЛ + РМКР.$$

Таблиця 1

Загальний рейтинг за дисципліною

Звітність	Лр1	Лр2	Лр3	Лр4	Лр5	Лр6	Лр7	Лр8	Лр9	МКР	СУМА
Високий рівень	10	10	10	10	10	10	10	10	10	10	100
Середній рівень	8	8	8	8	8	8	8	8	8	10	82

Умовою атестації є отримання 60 балів за виконання будь-якої звітності з навчальної дисципліни.

Здобувачі, які мають рейтингову оцінку 60 і більше балів, отримують відповідну до набраного рейтингу оцінку без додаткових випробувань.

Зі здобувачами, які мають рейтингову оцінку не менше 50 балів (виконали умови допуску до заліку), а також з тими здобувачами, хто бажає підвищити свою рейтингову оцінку, на останньому за розкладом занятті з дисципліни в семестрі проводиться семестровий контроль у вигляді залікової контрольної роботи.

Максимальна кількість балів за залікову контрольну роботу складає R3 = 10 балів.

3.1. Повнота розкриття суті та оригінальність відповіді на теоретичні питання – 2 балів за кожне питання – загалом – 4 бали;

3.2. Повнота, оригінальність та якість виконання практичного питання – 5 бали.

У разі **значних запозичень / порушення вимог доброчесності** залікова контрольна робота – не зараховується.

У разі виконання залікової контрольної роботи сума підсумкової семестрової рейтингової оцінки R та залікової контрольної роботи R3 у балах становить підсумкову семестрову рейтингову оцінку. Максимальна кількість балів за семестр не перевищує RC = 100.

З урахуванням одержаної суми балів кінцева оцінка визначається за національною шкалою та шкалою ECTS відповідно до табл. Табл.2.

Таблиця 2

Визначення семестрової оцінки

Кількість балів	Оцінка
100-95	Відмінно
94-85	Дуже добре
84-75	Добре
74-65	Задовільно
64-60	Достатньо
Менше 60	Незадовільно
Не виконані умови допуску	Не допущено

9. Нормативні положення політики навчальної дисципліни (освітнього компонента)

Правила відвідування занять. Відвідування занять є вільним, бали за присутність на лекціях та практичних заняттях не додаються, відповідно штрафні бали не передбачаються. Втім, вагома частина рейтингу формується через активну участь у заходах на практичних заняттях.

Дистанційний режим навчання. У разі запровадження обмежень на відвідування університету, пов'язаних з введенням карантину або режиму воєнного стану в державі, освітній процес здійснюється у дистанційному режимі відповідно до Положення про дистанційне навчання в КПІ ім. Ігоря Сікорського (<https://osvita.kpi.ua/index.php/node/188>), Регламенту організації освітнього процесу в дистанційному режимі (<https://profkom.kpi.ua/reglament-organizatsiyi-osvitnogo-protsesu-v-distantsiynomu-rezhimi>) та Регламенту проведення семестрового контролю в дистанційному режимі (<https://osvita.kpi.ua/node/148>). У режимі дистанційного навчання заняття відбуваються у вигляді онлайн-конференції на платформах BigBlueButton, Google Meet, Zoom. Посилання на конференцію видається на початку семестру і розміщується в АС «Електронний кампус». З метою забезпечення якісної підготовки здобувачів, дистанційний курс дисципліни розміщено на Платформі дистанційного навчання «Сікорський». Результати оцінювання висвітлюють у АС «Електронний кампус» на особистій сторінці здобувача (<https://ecampus.kpi.ua>).

Правила поведінки на заняттях. На заняттях слід дотримуватись норм етичної поведінки визначених у Кодексі честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» (<https://kpi.ua/code>). На території університету здобувачі мають поводити себе відповідно до Правил внутрішнього розпорядку (<https://kpi.ua/admin-rule>).

Визнання результатів навчання, набутих у неформальній/інформальній освіті. Положення про визнання результатів навчання, набутих у неформальній / інформальній освіті (<https://osvita.kpi.ua/index.php/node/179>) регламентує визнання результатів навчання, набутих у неформальній/інформальній освіті:

у разі навчання на неформальній освіті, яку студент обрав самостійно і хоче зарахувати результати навчання в рамках дисципліни, проходиться процедура валідації, що передбачає

подання здобувачем заяви на ім'я декана, декларації підтверджувальних документів; рішення про визнання чи не визнання приймається комісією у складі завідувача кафедри, викладача, гаранта освітньо-професійної програми.

за рішенням викладача окремі змістовні модулі або теми дисципліни можуть бути зараховані на підставі результатів проходження здобувачем онлайн-курсу або іншого елементу неформальної освіти; перелік завдань та елементів неформальної освіти визначається викладачем заздалегідь, про що студентам повідомляються на початку семестру; у цьому випадку здобувач звільняється від виконання відповідних завдань і отримує максимальний бал згідно з рейтинговою системою оцінювання.

За даним курсом можуть бути визнані результати навчання здобуті у неформальній / інформальній освіті в обсязі, що не перевищує 10% від загального обсягу навчального курсу (п.2.6 Положення).

Політика використання штучного інтелекту. Використання штучного інтелекту регламентується «Політикою використання штучного інтелекту для академічної діяльності в КПІ ім. Ігоря Сікорського» (<https://osvita.kpi.ua/node/1225>). Усі завдання, як під час виконання навчальних завдань з дисципліни, так і індивідуальні завдання, мають бути результатом власної оригінальної роботи здобувача. Використання ШІ для автоматичної генерації відповідей без подальшого їх аналізу та доопрацювання заборонено.

Порушення академічної доброчесності: використання плагіату. У разі виявлення фактів використання плагіату під час виконання навчальних завдань, студент втрачає право на подальше виконання поточних завдань у межах навчальної дисципліни до з'ясування обставин порушення.

Студент зобов'язаний аргументовано пояснити причини використання плагіату, а також засвідчити готовність дотримуватись принципів академічної доброчесності в подальшому навчанні; повторно виконати завдання або, за потреби, виконати додаткове завдання відповідно до бачення викладача з урахуванням вимог щодо самостійності.

Зазначені дії регламентуються Кодексом честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» (<https://kpi.ua/code>).

Робочу програму навчальної дисципліни (силабус):

Складено професором кафедри обчислювальної техніки, доктором технічних наук, професором Писарчуком Олексієм Олександровичем.

Ухвалено кафедрою обчислювальної техніки (протокол № 12 від 23.06.2025).

Погоджено Методичною комісією факультету інформатики та обчислювальної техніки (протокол № 11 від 27.06.2025).