



Інтелектуальна обробка голосової інформації

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	Другий (магістерський)
Галузь знань	12 Інформаційні технології
Спеціальність	122 Комп'ютерні науки
Освітня програма	Системи і методи штучного інтелекту
Статус дисципліни	Вибіркова
Форма навчання	очна(денна)/дистанційна/змішана
Рік підготовки, семестр	1 курс, весняний семестр
Обсяг дисципліни	4 кредитів ЄКТС, 120 годин, 36 г. лекцій, 18 г. лабораторних робіт, 66 г. СРС
Семестровий контроль/ контрольні заходи	Залік
Розклад занять	Лекції один раз на тиждень. Практичні заняття один раз на 2 тижні.
Мова викладання	Українська
Інформація про керівника курсу / викладачів	Лектор: Сидорський Володимир Сергійович, volodymyr.sydorskyi@gmail.com Лабораторні: Сидорський Володимир Сергійович, volodymyr.sydorskyi@gmail.com
Розміщення курсу	Кампус КПІ ecampus.kpi.ua

Програма навчальної дисципліни

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

Програма дисципліни "Інтелектуальна обробка голосової інформації" передбачає вивчення основ обробки цифрових сигналів, перетворень цифрових сигналів у спектральну форму, а також їх обробка та трансформації за допомогою класичних алгоритмів обробки цифрових сигналів. Надалі студенти пройдуть базову загальну підготовку для роботи з фреймворком Pytorch. Після чого студенти будуть вивчати методи й алгоритми машинного й глибокого навчання для

класифікації голосових сигналів, а також сигналів іншої природи, автоматичного розпізнавання мовлення й синтезу голосової інформації з текстової. В процесі вивчення студенти використовують мову програмування Python.

Дисципліна "Інтелектуальна обробка голосової інформації" вивчається в другому семестрі на 1 курсі другого (магістерського) рівня вищої освіти і їй передують дисципліни базової і професійної та практичної підготовки, які вивчаються в період навчання на першому рівні (бакалаврська підготовка) напряму підготовки (спеціальності) 122 «Комп'ютерні науки» і забезпечують можливість її вивчення.

Серед них можна виділити наступні дисципліни: «Математичний аналіз», «Дискретна математика», «Алгоритмізація та програмування», «Алгоритми та структури даних», «Об'єктно-орієнтоване програмування», «Проектування та аналіз обчислювальних алгоритмів».

Кредитний модуль доповнює формування у студентів таких загальних та фахових компетентностей:

ЗК 1: Здатність до абстрактного мислення, аналізу та синтезу.

ЗК 2: Здатність застосовувати знання у практичних ситуаціях.

ЗК 5: Здатність вчитися й оволодівати сучасними знаннями.

ФК 5: Здатність розробляти, описувати, аналізувати та оптимізувати архітектурні рішення інформаційних та комп'ютерних систем різного призначення.

ФК 6: Здатність застосовувати існуючі і розробляти нові алгоритми розв'язування задач у галузі машинного навчання.

та програмних результатів навчання:

ПРН 1: Мати спеціалізовані концептуальні знання, що включають сучасні наукові здобутки у сфері комп'ютерних наук і є основою для оригінального мислення та проведення досліджень, критичне осмислення проблем у сфері комп'ютерних наук та на межі галузей знань.

ПРН 2: Мати спеціалізовані уміння/навички розв'язання проблем комп'ютерних наук, необхідні для проведення досліджень та/або провадження інноваційної діяльності з метою розвитку нових знань та процедур.

ПРН 11: Створювати нові алгоритми розв'язування задач у сфері машинного навчання.

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

Кредитні модулі, які передують вивченню дисципліни і забезпечують її вивчення, відносяться до дисциплін циклу професійної підготовки:

- Математичний аналіз;
- Дискретна математика;
- Алгоритмізація та програмування;
- Алгоритми і структури даних;
- Об'єктно-орієнтоване програмування;
- Проектування та аналіз обчислювальних алгоритмів.

3. Зміст навчальної дисципліни

№ з/п	Назва теми лекції та перелік основних питань (перелік дидактичних засобів, завдання на СРС з посиланням на літературу)
1	Лекція 1. Огляд основних завдань та методів обробки голосових і інших сигналів Завдання на СРС: Повторення лекційного матеріалу.
2	Лекція 2. Техніки попередньої обробки, фільтрації сигналів. Кластеризація сигналів Завдання на СРС: Повторення лекційного матеріалу.

№ з/п	Назва теми лекції та перелік основних питань (перелік дидактичних засобів, завдання на СРС з посиланням на літературу)
3	Лекція 3. Представлення голосових сигналів через приховані ознаки нейронних мереж Завдання на СРС: Повторення лекційного матеріалу.
4	Лекція 4. Вступ до фреймворку Pytorch Завдання на СРС: Повторення лекційного матеріалу.
5	Лекція 5. Методи розв'язування задачі класифікації та тегування послідовності за допомогою фреймворку Pytorch Завдання на СРС: Повторення лекційного матеріалу.
6	Лекція 6. Методи вирішення задачі генерації за допомогою фреймворку Pytorch Завдання на СРС: Повторення лекційного матеріалу.
7	Лекція 7. Класифікація голосових сигналів і сигналів іншої природи Завдання на СРС: Повторення лекційного матеріалу.
8	Лекція 8. Побудова різних схем валідації алгоритмів машинного навчання Завдання на СРС: Повторення лекційного матеріалу.
9	Лекція 9. Задача виокремлення мовців Завдання на СРС: Повторення лекційного матеріалу.
10	Лекція 10. Метрики й базові підходи до задачі розпізнавання мовлення Завдання на СРС: Повторення лекційного матеріалу.
11	Лекція 11. Розпізнавання мовлення за допомогою нейронних моделей часткового навчання Завдання на СРС: Повторення лекційного матеріалу.
12	Лекція 12. Розпізнавання мовлення за допомогою авторегресивних нейронних моделей Завдання на СРС: Повторення лекційного матеріалу.
13	Лекція 13. Огляд сучасних генеративних концепцій Завдання на СРС: Повторення лекційного матеріалу.
14	Лекція 14. Побудова нейронної моделі синтезу мовлення Завдання на СРС: Повторення лекційного матеріалу.
15	Лекція 15. Побудова авторегресивної нейронної моделі синтезу мовлення Завдання на СРС: Повторення лекційного матеріалу.
16	Лекція 16. Контрольна робота.

4. Навчальні матеріали та ресурси

Базова література:

1. ISCA Archive, "ISCA Archive," [Online]. Available: <https://www.isca-archive.org/>.
2. Hugging Face, "Introduction to Audio Processing," [Online]. Available: <https://huggingface.co/learn/audio-course/chapter0/introduction>.
3. PyTorch, "PyTorch Tutorials," [Online]. Available: <https://pytorch.org/tutorials/>.
4. Kaggle, "Kaggle Competitions," [Online]. Available: <https://www.kaggle.com/competitions>.

Додаткова література:

5. Глибовець М.М., Олецький О.В. Штучний інтелект – Київ: Видавничий дім “Академія”, 2002.– 366с.
6. SPANIAS, Andreas; PAINTER, Ted; ATTI, Venkatraman. *Audio signal processing and coding*. John Wiley & Sons, 2006.

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

Лекційні заняття

Назви розділів і тем	Кількість годин				
	Всього	у тому числі			
		Лекції	Практичні	Лабораторні	СРС
Розділ 1. Базова обробка й представлення голосової інформації					
Тема 1. Огляд основних завдань та методів обробки голосових і інших сигналів	6	2			4
Тема 2. Техніки попередньої обробки, фільтрації сигналів. Кластеризація сигналів	6	2			4
Тема 3. Представлення голосових сигналів через приховані ознаки нейронних мереж	8	2		4	4
Разом за розділом 1	20	6		4	12
Розділ 2. Фреймворк Pytorch					
Тема 1. Вступ до фреймворку Pytorch	6	2			4
Тема 2. Методи розв'язування задачі класифікації та тегування послідовності за допомогою фреймворку Pytorch	6	2			4
Тема 3. Методи вирішення задачі генерації за допомогою фреймворку Pytorch	8	2		2	4
Разом за розділом 2	20	6		2	12
Розділ 3. Класифікація сигналів					
Тема 1. Класифікація голосових сигналів і сигналів іншої природи	8	4			4
Тема 2. Побудова різних схем валідації алгоритмів машинного навчання	6	2			4
Тема 3. Задача виокремлення мовців	8	2		2	4
Разом за розділом 3	22	6		2	12
Розділ 4. Автоматичне розпізнавання мовлення					
Тема 1. Метрики й базові підходи до задачі	6	2			4

Назви розділів і тем	Кількість годин				
	Всього	у тому числі			
		Лекції	Практичні	Лабораторні	СРС
розпізнавання мовлення					
Тема 2. Розпізнавання мовлення за допомогою нейронних моделей часткового навчання	6	2			4
Тема 2. Розпізнавання мовлення за допомогою авторегресивних нейронних моделей	10	2		4	4
Разом за розділом 4	22	6		4	12
Розділ 5. Синтез Мовлення					
Тема 1. Огляд сучасних генеративних концепцій	10	4			6
Тема 2. Побудова нейронної моделі синтезу мовлення	6	2			4
Тема 3. Побудова авторегресивної нейронної моделі синтезу мовлення	10	2		4	4
Разом за розділом 5	26	8		4	14
Модульна контрольна робота з розділів 1, 3, 4, 5	6	2			4
Залік	2			2	
Всього годин	120	36		18	66

Лабораторні заняття

Основні завдання циклу лабораторних занять полягають у закріпленні лекційного матеріалу та отриманні студентами практичних навичок.

№ з/п	Назва лабораторної роботи	Кількість ауд. годин
1	Побудова моделі машинного навчання яка буде визначати акустичні фонемі за представленнями з прихованих ознак нейронних мереж	4
2	Розробка моделі класифікації або тегування послідовності за допомогою фреймворку Pytorch	2
3	Розробка моделі класифікації аудіо сигналів і побудова схеми валідації для неї	2

4	Побудова системи автоматичного розпізнавання мовлення на базі нейронних мереж	4
5	Побудова синтезу мовлення на базі нейронних мереж	4
6	Залік	2

6. Самостійна робота студента

№ з/п	Назви тем і питань, що виносяться на самостійне опрацювання та посилання на навчальну літературу	Кількість годин СРС
1	Огляд основних завдань та методів обробки голосових і інших сигналів. Завдання на СРС: Запрограмувати віконне перетворення Фур'є.	4
2	Техніки попередньої обробки, фільтрації сигналів. Кластеризація сигналів. Завдання на СРС: Побудувати алгоритм детекції присутності звукового сигналу в аудіопотоці за допомогою кластеризації.	4
3	Представлення голосових сигналів через приховані ознаки нейронних мереж. Завдання на СРС: Використати приховані ознаки нейронних мереж для визначення фонем.	4
4	Вступ до фреймворку Pytorch. Завдання на СРС: Повторення лекційного матеріалу.	4
5	Методи розв'язування задачі класифікації та тегування послідовності за допомогою фреймворку Pytorch. Завдання на СРС: Побудувати класифікатор за допомогою фреймворку Pytorch.	4
6	Методи вирішення задачі генерації за допомогою фреймворку Pytorch. Завдання на СРС: Побудувати базовий генеративний алгоритм.	4
7	Класифікація голосових сигналів і сигналів іншої природи. Завдання на СРС: Побудувати класифікатор голосу або співу птахів або сигналів EEG.	4
8	Побудова різних схем валідації алгоритмів машинного навчання. Завдання на СРС: Створити схему валідацію для моделі з попередньої лекції	4
9	Задача виокремлення мовців. Завдання на СРС: Створити систему виокремлення мовців.	4
10	Метрики й базові підходи до задачі розпізнавання мовлення. Завдання на СРС: Повторення лекційного матеріалу.	4

№ з/п	Назви тем і питань, що виносяться на самостійне опрацювання та посилання на навчальну літературу	Кількість годин СРС
11	Розпізнавання мовлення за допомогою нейронних моделей часткового навчання. Завдання на СРС: Побудувати алгоритм автоматичного розпізнавання мовлення за допомогою нейронних моделей часткового навчання.	4
12	Розпізнавання мовлення за допомогою авторегресивних нейронних моделей. Завдання на СРС: Побудувати алгоритм автоматичного розпізнавання мовлення за допомогою авторегресивних нейронних моделей.	4
13	Огляд сучасних генеративних концепцій. Завдання на СРС: Повторення лекційного матеріалу.	6
14	Побудова нейронної моделі синтезу мовлення. Завдання на СРС: Протестувати моделі для синтезу мовлення й виокремити найкращі.	4
15	Побудова авторегресивної нейронної моделі синтезу мовлення. Завдання на СРС: Побудувати модель синтезу мовлення.	4
16	Модульна контрольна робота. Завдання на СРС: Підготовка до контрольної роботи.	4

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Вимоги, яких має дотримуватися студент в рамках даної дисципліни:

- Відвідування лекційних та лабораторних занять є бажаним, але не обов'язковим;
- Дозволяється, при необхідності, використання засобів зв'язку для пошуку потрібної інформації на платформі дистанційного навчання та/або в інтернеті;
- Лабораторні роботи мають бути розміщені на платформі GitHub, Kaggle, Colab або інших платформах за попереднім узгодженням з викладачем. Роботи можуть бути не загальнодоступними, але відкритими для викладача. Робота може бути оцінена без участі студента, але студент має бути готовий відповісти на всі питання щодо роботи, що виникнуть у викладача. Студент в свою чергу може оскаржити оцінку викладача й попросити детальне пояснення від викладача з можливістю надалі підвищити оцінку;
- Заохочувальні бали можуть призначатися за активність на лекціях та нестандартні рішення при виконанні лабораторних робіт, а також за додаткові роботи, які можуть бути запропоновані як і студентом (за згодою викладача), так і викладачем;
- штрафні бали можуть призначатися за несвоєчасне виконання лабораторних робіт;
- при виконанні лабораторних робіт потрібно дотримуватися графіка, який доводиться до відома студентів викладачем на початку семестру;
- обов'язковим є дотримання академічної доброчесності.

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

Поточний контроль: лабораторна робота № 1 (13 балів практична частина, 5 балів теоретична), лабораторна робота № 2 (13 балів практична частина, 5 балів теоретична),

лабораторна робота № 3 (13 балів практична частина, 5 балів теоретична), лабораторна робота № 4 (13 балів практична частина, 5 балів теоретична), лабораторна робота № 5 (13 балів практична частина, 5 балів теоретична), МКР (10 балів).

Поточний контроль проводиться методом контрольних питань під час приймання лабораторних робіт. За протермінування здачі лабораторної роботи нараховується 9 штрафних балів.

МКР буде запропонована у вигляді аналізу статті за темою курсу. Критерії оцінки модульної контрольної роботи:

- стаття повністю розібрана, присутній повний і тезовий опис, схеми системи, програмна реалізація системи - 10 балів;
- стаття повністю розібрана, присутній повний і тезовий опис, схеми системи, проте відсутня програмна реалізація – 7 балів;
- стаття частково розібрана, присутній повний і тезовий опис - 4 бали
- стаття не розібрана або студент не може відповісти на контрольні запитання - 0 балів.

Календарний контроль: провадиться двічі на семестр як моніторинг поточного стану виконання вимог силабусу.

Умова першого календарного контролю – зарахування першої лабораторної роботи, умова другого календарного контролю – зарахування двох лабораторних робіт.

Семестровий контроль: залік.

Умова допуску до семестрового контролю: зарахування усіх лабораторних робіт.

Сума рейтингових балів, отриманих студентом протягом семестру, переводиться до підсумкової оцінки згідно з таблицею. Якщо сума балів менша ніж 60, студент виконує додаткову залікову практичну роботу. У цьому разі сума балів за виконання залікової контрольної роботи додається до рейтингової суми балів.

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

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. Додаткова інформація з дисципліни (освітнього компонента)

При вивченні матеріалу курсу рекомендується особливу увагу приділити закріпленню базових положень, поступово ускладнюючи матеріал із наведенням відповідних прикладів задач. Під час МКР дозволяється використання конспекту, що зменшує психологічний тиск на студентів та стимулює більш активну роботу на лекціях. Постійний зворотний зв'язок зі студентами через сучасні засоби комунікації дозволяє швидко вирішувати труднощі в навчанні та є засобом індивідуального навчання.

За погодженням з викладачем, студент має можливість

- Взяти участь у змаганні, що відповідає темі однієї з лекцій розділі 1,3,4,5 й отримати до 20 балів рейтингу
- Розібрати одну з останніх статей, що відповідає темі однієї з лекцій розділі 1,3,4,5 й

отримати до 10 балів рейтингу

Приклади питань, які виносяться на модульну контрольну роботу:

- Які перетворення цифрових сигналів використовують для побудови моделей машинного навчання розпізнавання й синтезу мовлення?
- Які мають переваги й недоліки приховані ознаки нейронних мереж в порівнянні з перетворенням Фур'є в контексті обробки голосової інформації?
- Які генеративні підходи використовують для задачі синтезу мовлення?
- Як можна оцінити якість роботи моделі синтезу мовлення?
- Як і чому можна використати згорткові мережі для обробки зображень для класифікації цифрових сигналів?
- Яким чином можливо дискретизувати голосову інформацію для використання великих мовних моделей для синтезу аудіо сигналу?
- Назвіть приклади нейронних мереж з обробки природних мов або зображень, які можна застосувати для задач аналізу голосової інформації?

Робочу програму навчальної дисципліни (силабус):

Складено Сидорським Володимиром Сергійовичем

Ухвалено кафедрою штучного інтелекту (протокол № 14 від "11" червня 2024 р.)

Погоджено Методичною комісією ННІПСА (протокол № 10 від "24" червня 2024 р.)