

**Чернівецький національний університет імені Юрія Федьковича**

(повне найменування закладу вищої освіти)

**Факультет математики та інформатики**

(назва факультету/навчально-наукового інституту)

**Кафедра математичного моделювання**

(назва кафедри)

**“ЗАТВЕРДЖУЮ”**

Декан

**Мартинюк О.В.**

“25”

06

2024 року

**РОБОЧА ПРОГРАМА**

**навчальної дисципліни**

**Обробка структурованих та неструктурованих даних BigData**

(назва навчальної дисципліни)

**обов’язкова**

(вказати: обов’язкова / вибіркова)

**Освітньо-професійна програма «Системний аналіз»**

(назва програми)

**Спеціальність 124 Системний аналіз**

(вказати: код, назва)

**Галузь знань 12 Інформаційні технології**

(вказати: шифр, назва)

**Рівень вищої освіти другий (магістерський)**

(вказати: перший бакалаврський/другий магістерський)

**факультет математики та інформатики**

(назва факультету/ навчально-наукового інституту, на якому здійснюється підготовка фахівців за вказаною освітньо-професійною програмою)

**Мова навчання українська**

(вказати: на якій мові читається дисципліна)

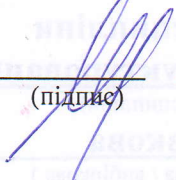
**Чернівці 2024 рік**

Робоча програма навчальної дисципліни «Обробка структурованих та неструктурованих даних BigData» складена відповідно до освітньо-професійної програми «Системний аналіз», затвердженої Вченою радою Чернівецького національного університету імені Юрія Федьковича « 29 » травня 2023 року протокол № 5.

Розробники: к. ф.-м. н., доц. Горбатенко М.Ю.  
(П.І.Б. авторів, посада, науковий ступінь, вчене звання)

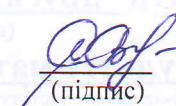
Погоджено з гарантом ОП і затверджено на засіданні кафедри математичного моделювання

Протокол № 18 від “ 25 ” червня 2024 року

Завідувач кафедри  Черевко І.М.  
(підпис) (прізвище та ініціали)

Схвалено методичною радою факультету математики та інформатики

Протокол № 11 від “25” червня 2024 року

Голова методичної ради факультету математики та інформатики  Сікора В.С.  
(підпис) (прізвище та ініціали)

**1. Мета навчальної дисципліни:** формування навичок проектування, розробки, встановлення додатків, побудова моделей з їх подальшим використанням в реальному житті. Знання, які студент повинен одержати в результаті вивчення курсу обробка структурованих та неструктурованих даних BigData, відіграватимуть важливу роль у побудові додатків, які використовуються на підприємствах.

**2. Результати навчання:** у результаті вивчення навчальної дисципліни студент повинен

**знати:** теоретичні положення про основи машинного навчання, теоретичні положення про ML.NET, теоретичні положення про вибір та побудову моделей, теоретичні положення про клієнтську обробку, теоретичні положення про взаємодію з користувачем.

**вміти:** створювати моделі машинного навчання, за допомогою ML.Net, створювати та супроводжувати проекти, коректно обробляти нештатні ситуації, вміти користуватися основними можливостями IDE Visual Studio, Visual Studio Code.

Студент повинен оволодіти програмним матеріалом, виконати лабораторні роботи, засвоїти теоретичний матеріал.

Під час вивчення дисципліни, відповідно до освітньо-професійної програми, формуються наступні

**загальні компетентності:**

ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.

ЗК4. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ЗК8. Здатність розробляти проекти та управляти ними.

**фахові компетентності:**

СК2. Здатність проектувати архітектуру інформаційних систем.

СК6. Здатність застосовувати теорію і методи Data Science для здійснення інтелектуального аналізу даних з метою виявлення нових властивостей та генерації нових знань про складні системи.

СК8. Здатність розробляти і реалізовувати наукові та прикладні проекти в галузі інформаційних технологій та дотичні до неї міждисциплінарні проекти.

СК10. Здатність до самоосвіти та професійного розвитку.

та отримуються наступні **програмні результати навчання:**

РН6. Застосовувати методи машинного навчання та інтелектуального аналізу даних, математичний апарат нечіткої логіки, теорії ігор та розподіленого штучного інтелекту для розв'язання складних задач системного аналізу.

РН7. Розробляти інтелектуальні системи в умовах слабо структурованих даних різної природи.

РН10. Зрозуміло і недвозначно доносити власні знання, висновки та аргументацію до фахівців і нефахівців, зокрема до осіб, які навчаються.

### 3. Опис навчальної дисципліни

#### 3.1. Загальна інформація

Форма навчання	Рік підготовки	Семестр	Кількість		Кількість годин						Вид підсумкового контролю
			кредитів	годин	лекції	практичні	семінарські	лабораторні	самостійна робота	індивідуальні завдання	
Денна	1	1	4	120	15	-	-	30	75	-	Екзамен

#### 3.2. Структура змісту навчальної дисципліни

Назви змістових модулів і тем навчальних занять	Кількість годин											
	денна форма						Заочна форма					
	усього	у тому числі					усього	у тому числі				
		лекц.	прак./семін.	лаб.	інд.	с.р.		лекц.	прак./семін.	лаб.	інд.	с.р.
1	2	3	4	5	6	7	8	9	10	11	12	13
<b>Теми навчальних занять</b>	<b>Змістовий модуль 1. Вступ та базові моделі ML.NET</b>											
Тема 1. Основи машинного навчання в термінах ML.NET	11	2	-	-	-	9	-	-	-	-	-	-
Тема 2. Побудова регресійної моделі засобами ML.NET.	16	2	-	5	-	9	-	-	-	-	-	-
Тема 3. Побудова класифікаційної моделі засобами ML.NET.	16	2	-	5	-	9	-	-	-	-	-	-
Тема 4. Побудова моделі кластеризації засобами ML.NET	17	2	-	5	-	10	-	-	-	-	-	-
Разом за ЗМ1	60	8	-	15	-	37	-	-	-	-	-	-
<b>Теми навчальних занять</b>	<b>Змістовий модуль 2. Просунуті моделі та побудова додатків</b>											
Тема 5. Побудова моделі виявлення	17	2	-	5	-	10	-	-	-	-	-	-

аномалій засобами ML.NET												
Тема 6. Побудова моделі матричної факторизації засобами ML.NET	17	2	-	5	-	10	-	-	-	-	-	-
Тема 7. Робота з ONNX	17	2	-	5	-	10	-	-	-	-	-	-
Тема 8. Загальноприйняті принципи побудови додатків з використанням моделей	9	1	-	-	-	8	-	-	-	-	-	-
Разом за ЗМ 2	60	7	-	15	-	38	-	-	-	-	-	-
<b>Усього годин</b>	<b>120</b>	<b>15</b>	<b>-</b>	<b>30</b>	<b>-</b>	<b>75</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>

### 3.3. Тематика лабораторних занять

№	Назва теми	Кількість годин
<b>Змістовий модуль 1. Вступ та базові моделі ML.NET</b>		
1	Побудова регресійної моделі на основі обраного тренера засобами ML.NET	5
2	Побудова класифікаційної моделі на основі обраного тренера засобами ML.NET	5
3	Побудова моделі кластеризації на основі обраного тренера засобами ML.NET	5
<b>Змістовий модуль 2. Просунуті моделі та побудова додатків</b>		
4	Побудова моделі виявлення аномалій на основі обраного тренера засобами ML.NET	5
5	Робота з ONNX засобами ML.NET	5
6	Побудова додатку з використанням моделі на основі побудованої моделі ML.NET	5

### 3.4. Самостійна робота студента

№	Назва теми (форма контролю)	Кількість балів
<b>Змістовий модуль 1. Вступ та базові моделі ML.NET</b>		
1	Основні етапи побудови ML.Net додатку (звіт)	5
2	Аналіз тренерів для регресійної моделі та вибір тренера для поставленої задачі в лабораторній роботі. (звіт)	5
3	Аналіз тренерів для класифікаційної моделі та вибір тренера для поставленої задачі в лабораторній роботі. (звіт)	5
4	Аналіз тренерів для моделі кластеризації та вибір тренера для поставленої задачі в лабораторній роботі. (звіт)	5

<b>Змістовий модуль 2. Просунуті моделі та побудова додатків</b>		
5	Аналіз тренерів для моделі виявлення аномалій та вибір тренера для поставленої задачі в лабораторній роботі або інтеграція з іншими стеками/мовами. (звіт)	5
6	Інтеграція ONNX моделей в ML.NET побудованих в інших середовищах. (звіт)	5
7	Аналіз існуючих підходів та архітектур для побудови додатків з елементами машинного навчання. (звіт)	5
8	Підходи до побудови додатків з використанням моделей (звіт)	5

#### **4. Освітні технології, методи навчання і викладання навчальної дисципліни**

Використовуються як класичні методи навчання, такі як лекційні та лабораторні заняття, так і Google Classroom, де розміщуються матеріали для ознайомлення та завдання для лабораторних робіт з можливістю завантаження отриманих результатів під час самостійної роботи та роботи над лабораторними роботами, а також перевагу надається використанню Google Forms для проведення підсумкового оцінювання за наявності такої можливості. Google Meet використовується за умови змішаної або дистанційної форми навчального процесу.

#### **5. Критерії та засоби оцінювання результатів навчання з навчальної дисципліни**

##### **5.1. Критерії оцінювання результатів навчання з навчальної дисципліни**

Екзаменаційний білет містить чотири теоретичні питання, які охоплюють весь матеріал дисципліни. Повна відповідь на перше та друге питання оцінюється у 7 балів, на 3 та 4 – у 8 балів. Екзаменаційна робота оцінюється у 30 балів.

Екзамен проводиться з використанням Google Forms. Студент отримує посилання на Google Form, яке буде містити 4 питання, студент повинен дати відповідь на запитання у вигляді відкритої відповіді.

Нижче наведена шкала оцінювання. Кожне з питань оцінюється так:

- 1) робота виконана повністю без помилок або з незначними помилками 7-8 балів;
- 2) робота виконана повністю з помилками, які не впливають на кінцевий результат 5-6 балів;
- 3) робота виконана повністю з суттєвими помилками, але витримано алгоритм викладання матеріалу 4 бали;
- 4) робота виконана не повністю з суттєвими помилками, але витримано загальний алгоритм викладання матеріалу 3 бали;
- 5) робота виконана не повністю з суттєвими помилками 2 бали;
- 6) робота не виконана або виконана не повністю з суттєвими помилками 1-0 балів;

Підсумкова оцінка за екзаменаційну роботу відповідає загальній сумі балів, отриманих під час поточного контролю (максимально 70 балів) та під час іспиту (максимально 30 балів).

### Шкала оцінювання: національна та ECTS

Оцінка за національною шкалою	Оцінка за шкалою ECTS	
	Оцінка (бали)	Пояснення за розширеною шкалою
<b>Відмінно</b>	A (90-100)	відмінно
<b>Добре</b>	B (80-89)	дуже добре
	C (70-79)	добре
<b>Задовільно</b>	D (60-69)	задовільно
	E (50-59)	достатньо
<b>Незадовільно</b>	FX (35-49)	(незадовільно) з можливістю повторного складання
	F (1-34)	(незадовільно) з обов'язковим повторним курсом

### 5.2. Засоби оцінювання

Засобами оцінювання та демонстрування навчання є аналітичні звіти та вихідний код проєктів виконаний самостійно або під час командної роботи.

### 6. Форми поточного контролю

Формами поточного контролю є захист лабораторних робіт, звіти.  
Формою підсумкового контролю є екзамен.

### 7. Рекомендована література

#### 7.1. Фахова (основна)

1. Machine Learning with Microsoft ML.Net: Develop Machine Learning Models with ML.NET from the Developer's perspective // Marco Zamana, 2023, Packt Publishing,
2. Programming ML.NET (Developer Reference) 1st Edition // Dino Esposito, Francesco Esposito, 2022 Microsoft Press.- 256 pp.
3. ML.NET Revealed: Simple Tools for Applying Machine Learning to Your Applications 1st Edition // Sudipta Mukherjee, 2020 Apress.- 192 pp.
4. Microsoft ML.NET Machine Learning for .NET Developers using C#.NET // Dr. A. F. Salam, Jakia Salam, 2019, Independently published.- 175 pp.
5. Automated Machine Learning with Microsoft Azure: Build highly accurate and scalable end-to-end AI solutions with Azure AutoML 1st Edition // Dennis Michael Sawyers, 2021, Packt Publishing.- 340pp.

## 7.2. Допоміжна

1. Introducing Machine Learning (Developer Reference) // Dino Esposito, Francesco Esposito, 2020, Microsoft Press.- 400 pp.
2. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems 3rd Edition // Aurélien Géron, 2022, O'Reilly Media.- 861 pp
3. Programming Machine Learning: From Coding to Deep Learning // Paolo Perrotta, 2020, Pragmatic Bookshelf.- 342 pp.

## 8. Інформаційні ресурси

1. <https://dotnet.microsoft.com/en-us/apps/machinelearning-ai/ml-dotnet>
2. <https://learn.microsoft.com/en-us/dotnet/machine-learning/how-does-ml-dotnet-work>
3. <https://github.com/dotnet/machinelearning>

Додатково

(для контролю та самоконтролю роботи студента)

### Розподіл балів, які отримують студенти

Поточне оцінювання (аудиторна та самостійна робота)								Кількість балів (екзамен)	Сумарна к-ть балів
Змістовий модуль 1 (35 балів)				Змістовий модуль 2 (35 балів)				<b>30</b>	<b>100</b>
T1	T2	T3	T4	T5	T6	T7	T8		
<b>5</b>	<b>10</b>	<b>10</b>	<b>10</b>	<b>10</b>	<b>10</b>	<b>10</b>	<b>5</b>		

T1, T2 ... T8 – теми змістових модулів.

### Політика освітнього процесу

Студенти зобов'язані своєчасно та якісно виконувати всі отримані завдання і акуратно їх оформлювати у вигляді звіту з наданням програмного коду у випадку чисельних розрахунків. За необхідністю з метою з'ясування всіх незрозумілих під час самостійної та індивідуальної роботи питань, відвідувати консультації викладача. Кожний студент зобов'язаний дотримуватися принципів академічної доброчесності. Виконаний студентом не свій варіант завдання не оцінюється. Складання (перескладання) екзамену проводиться за встановленим деканатом розкладом.