

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

МЕХАНІКО-МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ

Кафедра алгебри і комп'ютерної математики

«ЗАТВЕРДЖУЮ»

Заступник декана
з навчальної роботи

_____ Харитонов О.М

« ____ » _____ 2020 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

**Аналіз великих даних
для студентів**

галузь знань	01 «Освіта/Педагогіка»
спеціальність	«014 Середня освіта (за предметними спеціальностями)»
предметна спеціальність	014.04 «Середня освіта (Математика)»
освітній рівень	перший (бакалавр)
освітня програма	«Математика»
вид дисципліни	вибіркова

Форма навчання	денна
Навчальний рік	2020/2021
Семестр	8
Кількість кредитів ECTS	3
Мова викладання, навчання та оцінювання	українська
Форма заключного контролю	залік

Викладачі: Головинський Андрій Леонідович, к.ф.-м.н., доцент

Пролонговано: на 20 /20 н.р. () « » 20 р.
на 20 /20 н.р. () « » 20 р.

КИЇВ – 2020

Розробник Головинський А.Л. к.ф.-м.н., доцент

ЗАТВЕДЖЕНО

Зав. кафедри алгебри і комп'ютерної математики

(підпис)

Петравчук А.П.

Протокол № від 2020 р.

Схвалено науково-методичною комісією механіко-математичного факультету

Протокол від “ ____ ” _____ 2020 року № ____

Голова науково-методичної комісії _____ професор, д.ф.-м.н. Олійник А.С.
(підпис)

1. Мета дисципліни – ознайомлення з основними методами структурування, кодування та обробки інформації, з методами побудови математичних моделей процесів, які описуються великими даними, пошук та виправлення вад даних, оволодіння основними алгоритмами обробки даних основних типів: зображення, відео, аудіо, тексти. Оволодіння навичками застосування даних алгоритмів у навчальному шкільному процесі.

2. Попередні вимоги до опанування навчальної дисципліни:

1. Знати основні поняття, факти і теореми лінійної алгебри, алгебри і теорії чисел, дискретної математики, теорії ймовірностей, математичного аналізу, основні навички з програмування.

2. Вміти активно використовувати та творчо застосовувати зазначені вище знання в процесі опрацювання матеріалу курсу «Аналіз великих даних».

3. Володіти елементарними навичками роботи з множинами, функціями, знаходити ймовірності подій, обчислювати основні характеристики випадкових величин, знати матричне числення, вміти будувати і реалізовувати алгоритми обробки даних різної структури, працювати з бібліотеками чисельних методів.

3. Анотація навчальної дисципліни.

Навчальна дисципліна «Аналіз великих даних» є складовою освітньої програми підготовки фахівців за освітнім рівнем «бакалавр» Дана дисципліна є вибірковою. В курсі «Аналіз великих даних» висвітлюються базові відомості, поняття, факти теорії інформації і теорії кодування, теорії наближення. Зокрема, розглядаються: класифікація даних за структурою, значеннями, методи представлення даних у векторних просторах, методи зниження розмірності, алгоритми кластеризації, алгоритми візуалізації даних, методи роботи з часовими рядами, методи використання нейронних мереж в обробці даних.

Викладається у 8 семестрі 4 курсу в обсязі **90 год.** (3 кредити ECTS¹) зокрема: лекції – всього 14 год., лабораторних 12 год., консультації 4 год., самостійна робота – 60 год. У курсі передбачено 2 змістових модулі та 2 модульні контрольні роботи. Завершується дисципліна заліком у восьмому семестрі.

4. Завдання (навчальні цілі): формування здатності розв'язувати складні задачі та практичні проблеми у математиці або у процесі навчання, що передбачає застосування теорій та методів математики, статистики й комп'ютерних технологій і характеризується комплексністю та невизначеністю умов; набуття знань, умінь та навичок (компетентностей) на рівні новітніх досягнень у математиці, відповідно до освітнього рівня «Бакалавр». Зокрема, професійне оволодіння компетентностями:

- 1) Здатність до абстрактного мислення, аналізу та синтезу (ЗК-1);
- 2) Здатність застосовувати знання у практичних ситуаціях (ЗК-2)
- 3) Знання й розуміння предметної області та професійної діяльності (ЗК-3)
- 4) Здатність спілкуватися державною мовою як усно, так і письмово (ЗК-4);
- 5) Навички використання інформаційних і комунікаційних технологій (ЗК-6);
- 6) Здатність учитися і оволодівати сучасними знаннями (ЗК-7)
- 7) Здатність до пошуку, обробки та аналізу інформації з різних джерел (ЗК-8);
- 8) Здатність приймати обґрунтовані рішення (ЗК-9);
- 9) Здатність працювати в команді (ЗК-10);
- 10) Здатність працювати автономно (ЗК-11);

¹ кредитів ECTS – кредит кратний 30 годинам.

- 11) Визначеність і наполегливість щодо поставлених завдань і взятих обов'язків (ЗК-12);
- 12) Здатність реалізувати свої права і обов'язки як члена суспільства, усвідомлювати цінності громадянського (вільного демократичного) суспільства та необхідність його сталого розвитку, верховенства права, прав і свобод людини і громадянина в Україні (ЗК-13);
- 13) Здатність зберігати та примножувати моральні, культурні, наукові цінності і досягнення суспільства на основі розуміння історії та закономірностей розвитку предметної області, її місця у загальній системі знань про природу і суспільство та у розвитку суспільства, техніки і технологій, використовувати різні види та форми рухової активності для активного відпочинку та ведення здорового способу життя (ЗК-14);
- 14) Здатність діяти на основі етичних міркувань (мотивів) (ЗК-15);
- 15) Здатність до адаптації та дії в новій ситуації (ЗК-16).
- 16) Здатність формулювати проблеми математично та в символній формі з метою спрощення їхнього аналізу й розв'язання (СК-1);
- 17) Здатність подавати математичні міркування та висновки з них у формі, придатній для цільової аудиторії, а також аналізувати та обговорювати математичні міркування інших осіб, залучених до розв'язання тієї самої задачі (СК -2);
- 18) Здатність до кількісного мислення (СК-3);
- 19) Здатність розробляти і досліджувати математичні моделі явищ, процесів та систем (СК-4)
- 20) Здатність застосовувати спеціалізовані мови програмування та пакети прикладних програм (СК-5);
- 21) Здатність до комунікації з фаховими спільнотами державною (українською) мовою (СК-6).

5. Результати навчання за дисципліною:

Результат навчання (1. знати; 2. вміти; 3. комунікація; 4. автономність та відповідальність)		Форми (та/або методи і технології) викладання і навчання	Методи оцінювання та пороговий критерій оцінювання за необхідності	Відсоток у підсум- ковій оцінці з дисциплі- ни
Код	Результат навчання			
PH 1.1	Знати: поняття видів даних за структурою та значеннями, основні методи представлення текстових, аудіо, відео даних, зображень у векторних просторах	лекція, самостійне опрацювання	Екзамен, контрольна робота №1, опитування під час практичних занять	10%
PH 1.2	Знати: базові методи роботи з даними: візуалізація багатомірних даних, методи зниження розмірності векторних просторів, побудову статистичних оцінок, ведення поняття відстані.	лекція, самостійне опрацювання	Екзамен, контрольна робота №1 опитування під час практичних занять	10%
PH 1.3	Знати методи кластеризації для дослідження структури у даних,	лекція	Екзамен, контро- льна робота №2	

	регресійні методи для визначення залежностей у даних, метод головних компонент.	самостійне опрацювання	опитування під час практичних занять	10%
PH 1.4	Знати основні поняття пов'язані з часовими рядами: кореляція та автокореляція	лекція, самостійне опрацювання	Екзамен, контрольна робота №2, опитування під час практичних занять	5%
PH 2.1	Уміти будувати нейронну мережу-класифікатор для кластеризації даних	Практичне заняття самостійне опрацювання	Екзамен, перевірка індивідуальних завдань, самостійна аудиторна робота, модульна контрольна робота №1	15%
PH 2.2	Уміти кодувати текстові дані, зображення та звуки у векторному просторі, залежно від постановки задачі. Вводити поняття відстані між елементами даних.	практичне заняття самостійне опрацювання	Екзамен, перевірка індивідуальних завдань, самостійна аудиторна робота, модульна контрольна робота №1	20%
PH 2.3	Уміти визначити вади в даних, виправити їх, візуалізувати дані та повести їх базовий аналіз.	Практичне заняття самостійне опрацювання	Екзамен, перевірка індивідуальних завдань, самостійна аудиторна робота, модульна контрольна робота №2	15%
PH 2.4	Уміти будувати алгоритми пошуку залежностей у даних та реалізувати їх у програмному забезпеченні, працювати з розміченими та нерозміченими даними.	практичне заняття самостійне опрацювання	Екзамен, перевірка індивідуальних завдань, самостійна аудиторна робота, контрольна робота №2	10%
PH 3.1	Здатність обґрунтовувати власний погляд на задачу та формулювати робочі гіпотези, спілкуватися з колегами з питань застосування математичних методів та теорій	Лекція, практичне заняття, самостійна робота	активна робота на лекції, практичних заняттях, усні відповіді	2.5%
PH 3.2	Вироблення навиків командної роботи	Лекція, практичне заняття, самостійна робота	активна робота на лекції, практичних заняттях, усні відповіді	2.5%

6. Співвідношення результатів навчання дисципліни з програмними результатами

Результати навчання дисципліни	Програмні результати навчання									
	PH 1.1	PH 1.2	PH 1.3	PH 1.4	PH 2.1	PH 2.2	PH 2.3	PH 2.4	PH 3.1	PH 3.2
PH-1. Знає основні етапи історичного розвитку математичних знань і парадигм, розуміє сучасні тенденції в математиці.	+	+	+	+	+			+	+	+
PH-2. Розуміє фундаментальну і прикладну математику на рівні, необхідному для досягнення інших вимог освітньої програми.	+	+	+	+	+	+	+	+	+	+
PH-3. Має навички використання спеціалізованих програмних засобів комп'ютерної та прикладної математики і використовувати інтернет-ресурси.	+	+	+	+	+	+	+	+	+	+
PH-4. Використовує усно і письмово професійну українську мову.	+	+	+	+	+	+	+	+	+	+
PH-8. Оперує базовими категоріями та поняттями математики.	+	+	+	+	+	+	+	+	+	+
PH-10. Застосовує міжнародні й національні стандарти та досвід у професійній діяльності.	+	+	+	+	+	+	+	+	+	+
PH-23. Володіє методикою позакласної роботи з математики, зокрема методикою підготовки учнів до математичних олімпіад та турнірів.	+	+	+	+	+	+	+	+	+	+
PH-26. Здатний оцінювати та розвивати власні математичні й методичні компетентності, усвідомлювати відповідальність за їх рівень.	+	+	+	+	+	+	+	+	+	+

7. Схема формування оцінки.

7.1. Форми оцінювання студентів:

- оцінювання впродовж навчального періоду:

1. Виконання завдань, винесених на самостійну роботу: PH2.1, PH2.2, PH2.3, PH2.4 – 8 балів/4 бали;
2. Модульна контрольна робота 1: PH1.1, PH1.2, PH2.1, PH2.2 – 20 балів/12 балів;
3. Модульна контрольна робота 2: PH1.3, PH1.4 PH2.3 – 20 балів/12 балів;
4. Розв'язання задач на лабораторних заняттях: PH2.1, PH2.2, PH2.3, PH2.4, PH3.1, PH3.2, – 12 балів/7 балів;

- підсумкове оцінювання: екзамен.

- максимальна кількість балів, які можуть бути отримані: 40 балів;
- результати навчання, які будуть оцінюватись: PH1.1, PH1.2, PH1.3, PH1.4, PH2.1, PH2.2, PH2.3, PH2.4;
- форма проведення і види завдань: письмова робота.

7.2. Організація оцінювання:

Самостійна робота передбачає активну самостійну роботу по розв'язанню задач і по формулюванню основних теоретичних положень під час практичних занять, при цьому кожен студент отримує індивідуальне завдання, яке він повинен виконати за невеликий проміжок часу (складність завдання пропорційно відведеному часу).

Критично-розрахунковий мінімум балів за навчання впродовж семестру становить **20** балів, рекомендований мінімум, розрахований з урахуванням специфіки дисципліни становить **35** балів.

Студенти, які протягом семестру набрали сумарно меншу кількість балів ніж рекомендований мінімум **35** балів для підвищення балів отримують можливість написати додаткову контрольну роботу та доскласти домашні завдання. Мінімальна кількість балів, які додаються до семестрових – 24 бали, тобто, якщо оцінка студента на заліку є нижчою від мінімального порогового рівня (24 бали), то бали за залік не додаються до семестрової оцінки (вважаються рівними нулю), а підсумкова оцінка із дисципліни є незадовільною, тобто залік не зараховується.

Терміни проведення форм оцінювання:

1. Модульна контрольна робота №1: на 5-му тижні 2 семестру 4-го курсу.
2. Модульна контрольна робота №2: на 9-му тижні 2 семестру 4-го курсу.
3. Оцінювання завдань самостійної роботи за РН2.1 на 3-му тижні, за РН2.2 на 6 тижні, за РН2.3 на 12 тижні.

Форма заліку – письмово-усна. Білет складається із 5 завдань, перші два з яких є теоретичними, три інших – задачі. Кожне завдання оцінюється від 0 до 7 балів. Додатково від 0 до 5 балів студент отримує за усне опитування. Всього за залік можна отримати від 0 до 40 балів.

У випадку відсутності студента з поважних причин відпрацювання та перездачі форм контролю здійснюються у відповідності до „Положення про організацію освітнього процесу в Київському національному університеті імені Тараса Шевченка” (2018), <http://www.univ.kiev.ua/pdfs/official/Organization-of-the-educational-process.pdf>.

7.3 Шкала відповідності оцінок:

Відмінно/ Excellent	90 – 100
Добре/ Good	75 – 89
Задовільно/ Satisfactory	60 – 74
Не задовільно/ Fail	0 – 59
Зараховано/ Passed	60 – 100
Не зараховано/ Fail	0 – 34

8. СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ. ТЕМАТИЧНИЙ ПЛАН ЛЕКЦІЙ І ПРАКТИЧНИХ ЗАНЯТЬ

№ п/п	Назва теми	Кількість годин				
		Лекції	лабораторні заняття	Самост. робота	Модульна контрольна	Інші форми контролю
Змістовий модуль 1 „Великі дані, базові методи роботи з даними”						
1	Основні поняття представлення даних	4	4	15		
2	Базові методи роботи з даними	4	4	15	2	
Змістовий модуль 2 „ Використання нейронних мереж для обробки даних”						
3	Розмічені та не розмічені дані, задача класифікації	4	2	15		
4	Автоенкодера, часові ряди,	2	2	15	2	

	рекурентні нейронні мережі					
	Всього годин	14	12	60	4	

Загальний обсяг 90 годин, у тому числі:
лекції – 14 годин,
лабораторні заняття – 12 годин,
консультації – 4 годин,
самостійна робота – 60 годин.

9. Рекомендовані джерела

Основні:

1. P. Simon. Too Big to Ignore: The Business Case for Big Data. Wiley, 2013, 256 p.
2. Т. Ерл. Основы Big Data: концепции, алгоритмы и технологии. 2018, 320 с.
3. The white book of Big Data. Fujitsu, 66 p.

Додаткові:

1. R. Buyya. Big Data: Principles and Paradigms, 2016, 370 p.
2. C. Borgman. Big Data, Little Data, No Data, Scholarship in the Networked World. 2017, 416 p.