

Силабус
ВК Big Data (Технології обробки великих
даних)

Назва дисципліни:	Big Data (Технології обробки великих даних)
Рівень вищої освіти:	перший (бакалаврський)
Сторінка курсу в Moodle:	https://dl2022.khadi-kh.com/course/index.php?categoryid=39
Обсяг освітнього компоненту	3 кредити (90 годин)
Форма підсумкового контролю	Залік
Консультації:	за графіком
Назва кафедри:	кафедра комп'ютерних технологій та мехатроніки
Мова викладання:	українська
Керівник курсу:	Шуляков Владислав Миколайович, к.т.н., доцент
Контактний телефон:	+38 (067) 737-68-69
E-mail:	jason07@ukr.net

Короткий зміст освітнього компоненту:

Метою є ознайомлення з предметною областю великих даних (big data), наукою про дані (data science) та аналізом даних (data analytics); оволодіння навичками оперування великими даними шляхом використання спеціалізованих програмно-апаратних засобів, зокрема хмарних сервісів, спеціалізованих систем зберігання даних, розподілених файлових систем.

Предмет: розвиток уміння працювати з великими даними з урахуванням їх ключових характеристик: обсягу, різноманітності, мінливості й забезпечуючи відповідний рівень швидкості їх опрацювання; розвиток навичок ефективного використання математичного, алгоритмічного та програмного забезпечення для розв'язання основних задач предметної області великих даних.

Основними завданнями вивчення навчальної дисципліни є ефективно використовувати парадигми паралельного опрацювання даних, зокрема MapReduce та системи Apache Hadoop, Apache Spark, відповідні хмарні служби Amazon Web Services та IBM Bluemix; розгортати надійні та швидкі сховища для надвеликих обсягів даних; використовувати програмні бібліотеки та фреймворки з ефективними алгоритмами опрацювання надвеликих обсягів даних. Вміти аналізувати й ефективно застосовувати хмарні системи опрацювання великих даних.

Передумови для вивчення освітнього компоненту:

- Алгоритмізація та програмування;
- Вища математика;
- Теорія алгоритмів;
- Організація баз даних і знань.

Компетентності, яких набуває здобувач:

Загальні компетентності:

- Здатність до професійної роботи з даними: видобуток (у тому числі повторний), фільтрація, інтеграція, зберігання, перевірка актуальності, валідація і репрезентивність;
- Вибирати вихідні дані для проектування, керуючись формальними методами опису вимог та моделювання;
- Знати та вміти застосовувати інформаційні технології обробки, зберігання та передачі даних.

Спеціальні компетентності:

- Здатність розв'язувати складні спеціалізовані завдання або практичні проблеми, що характеризуються комплексністю та невизначеністю умов, із застосуванням теорій та методів інформаційних технологій;
- Здатність накопичувати, обробляти та систематизувати професійні знання щодо створення і супроводження програмного забезпечення та визнання важливості навчання протягом всього життя;
- Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

Результати навчання відповідно до освітньої програми:

- Знати і розуміти наукові положення, що лежать в основі функціонування комп'ютерних засобів, систем та мереж;
- 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 розробляти програмне забезпечення для вбудованих і розподілених застосувань, мобільних і гібридних систем, розраховувати, експлуатувати, типове для спеціальності обладнання;
- Вміти ефективно працювати як індивідуально, так і у складі команди;
- Вміти оцінювати отримані результати та аргументовано захищати прийняті рішення.

Тематичний план

№ теми	Назва тем (ЛК, ЛР, ПР, СЗ, СР)	Кількість годин	
		очна	заочна
1	Лекція №1: Аналітика Великих Даних. Вступ в аналітику Великих Даних (ВД). Основні поняття і визначення.	2	
	Практичне заняття 1	2	
	Задання для самостійної роботи 1	7	
2	Лекція №2: Управління даними, життєвий цикл ВД. Фази життєвого циклу даних	2	
	Практичне заняття 2	2	
	Задання для самостійної роботи 2	7	
3	Лекція №3: Проблеми уявлення і моделювання знань і їхній зв'язок з проблемою ВД.	2	
	Практичне заняття 3	2	
	Задання для самостійної роботи 3	7	
4	Лекція №4: Багатовимірні масиви, тензорні моделі. Традиційна DataScience - Big Data - загальні риси та відмінності.	2	
	Практичне заняття 4	2	
	Задання для самостійної роботи 4	7	

5	Лекція №5: Тензорна алгебра, застосування тензорів. Коротка історія сучасного тензорного аналізу.	2	
	Практичне заняття 5	2	
	Задання для самостійної роботи 5	7	
6	Лекція №6: Інструментальні та програмні засоби роботи з великими даними.	2	
	Практичне заняття 6	2	
	Задання для самостійної роботи 6	7	
7	Лекція №7: Принципи роботи з великими даними, парадигма MapReduce, Hadoop і корпоративні системи, Hadoop і СУРБД.	2	
	Практичне заняття 7	2	
	Задання для самостійної роботи 7	7	
8	Лекція №8: Оператори векторизації та реорганізації. Перетворення тензорів у вектори. Розгортання тензорів засобами МатЛаб.	2	
	Практичне заняття 8	2	
	Задання для самостійної роботи 8	9	
Разом	Лекції	16	
	Практики	16	
	Самостійна робота	58	

Індивідуальне навчально-дослідне завдання: не передбачене.

Методи навчання:

- 1) словесні: 1.1 традиційні: лекції, пояснення, розповідь тощо;
- 1.2 інтерактивні (нетрадиційні): проблемні лекції, дискусії тощо;
- 2) наочні: метод ілюстрацій, метод демонстрацій
- 3) практичні: 3.1 традиційні: практичні заняття, семінари;
- 3.2 інтерактивні (нетрадиційні): тренінги, «круглий стіл», метод мозкової атаки.

Система оцінювання та вимоги:

Конкретизація, деталізація критеріїв та системи оцінювання з урахуванням специфіки освітнього компоненту здійснюється на основі загальних критеріїв, наведених у СТБНЗ 7.1-02:2018 «Положення про організацію освітнього процесу в ХНАДУ» (https://www.khadi.kharkov.ua/fileadmin/P_Ychebotdel/norm dok/stvnz 7 1 02.pdf) та СТБНЗ 90.1-01:2021 «Оцінювання результатів навчання здобувачів вищої освіти» (https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz-90.1-01 2021.pdf).

Поточна успішність

1 Поточна успішність здобувачів за виконання навчальних видів робіт на навчальних заняттях і за виконання завдань самостійної роботи оцінюється за допомогою чотирибальної шкали оцінок з наступним перерахуванням у 100-бальною шкалу. Під час оцінювання поточної успішності враховуються всі види робіт, передбачені навчальною програмою.

1.1 Лекційні заняття оцінюються шляхом визначення якості виконання конкретизованих завдань.

1.2 Практичні заняття оцінюються якістю виконання контрольного або

індивідуального завдання, виконання та оформлення практичної роботи.

1.3 Лабораторні заняття оцінюються якістю виконання звітів про виконання лабораторних робіт.

1.4 Семінарські заняття оцінюються якістю виконання індивідуального завдання/реферату.

2 Оцінювання поточної успішності здобувачів вищої освіти здійснюється на кожному практичному занятті (лабораторному чи семінарському) за чотирибальною шкалою («5», «4», «3», «2») і заносяться у журнал обліку академічної успішності.

– «відмінно»: здобувач бездоганно засвоїв теоретичний матеріал, демонструє глибокі знання з відповідної теми або навчальної дисципліни, основні положення;

– «добре»: здобувач добре засвоїв теоретичний матеріал, володіє основними аспектами з першоджерел та рекомендованої літератури, аргументовано викладає його; має практичні навички, висловлює свої міркування з приводу тих чи інших проблем, але припускається певних неточностей і похибок у логіці викладу теоретичного змісту або при аналізі практичного;

– «задовільно»: здобувач в основному опанував теоретичні знання навчальної теми, або дисципліни, орієнтується у першоджерелах та рекомендованій літературі, але непереконливо відповідає, плутає поняття, невпевнено відповідає на додаткові питання, не має стабільних знань; відповідаючи на питання практичного характеру, виявляє неточність у знаннях, не вміє оцінювати факти та явища, пов'язувати їх із майбутньою професією;

– «незадовільно»: здобувач не опанував навчальний матеріал теми (дисципліни), не знає наукових фактів, визначень, майже не орієнтується в першоджерелах та рекомендованій літературі, відсутнє наукове мислення, практичні навички не сформовані.

3 Підсумковий бал за поточну діяльність визнається як середньоарифметична сума балів за кожне заняття, за індивідуальну роботу, поточні контрольні роботи за формулою:

$$K^{\text{поточ}} = \frac{K1 + K2 + \dots + Kn}{n}$$

де $K^{\text{поточ}}$ – підсумкова оцінка успішності за результатами поточного контролю;

$K1, K2, \dots, Kn$ – оцінка успішності n -го заходу поточного контролю;

n – кількість заходів поточного контролю.

Оцінки конвертуються у бали згідно шкали перерахунку (таблиця 1).

Таблиця 1 – Перерахунок середньої оцінки за поточну діяльність у багатобальну шкалу

4-бальна шкала	100-бальна шкала	4-бальна шкала	100-бальна шкала	4-бальна шкала	100-бальна шкала	4-бальна шкала	100-бальна шкала
5	100	4,45	89	3,90	78	3,35	67
4,95	99	4,4	88	3,85	77	3,3	66
4,9	98	4,35	87	3,80	76	3,25	65
4,85	97	4,3	86	3,75	75	3,2	64
4,8	96	4,25	85	3,7	74	3,15	63
4,75	95	4,20	84	3,65	73	3,1	62
4,7	94	4,15	83	3,60	72	3,05	61
4,65	93	4,10	82	3,55	71	3	60

4,6	92	4,05	81	3,5	70	від 1,78 до 2,99	від 35 до 59
						повторне складання	
4,55	91	4,00	80	3,45	69	від 0 до 1,77	від 0 до 34
4,5	90	3,95	79	3,4	68	повторне вивчення	

Підсумкове оцінювання

1 Екзамен проводиться після вивчення всіх тем дисципліни і складається здобувачами вищої освіти в період екзаменаційної сесії після закінчення всіх аудиторних занять

2 До екзамену допускаються здобувачі вищої освіти, які виконали всі види робіт передбачені навчальним планом з дисципліни:

- були присутні на всіх аудиторних заняттях (лекції, семінари, практичні);
- своєчасно відпрацювали всі пропущені заняття;
- набрали мінімальну кількість балів за поточну успішність (не менше 60 балів, що відповідає за національною шкалою «3»);

Якщо поточна успішність з дисципліни нижче ніж 60 балів, здобувач вищої освіти має можливість підвищити свій поточний бал до мінімального до початку екзаменаційної сесії.

3 Оцінювання знань здобувачів при складанні екзамену здійснюється за 100-бальною шкалою.

Оцінювання знань здобувачів шляхом тестування здійснюється за шкалою:

- «Відмінно»: не менше 90 % правильних відповідей;
- «Дуже добре»: від 82 % до 89 % правильних відповідей;
- «Добре»: від 74 % до 81 % правильних відповідей;
- «Задовільно»: від 67 % до 73% правильних відповідей;
- «Задовільно достатньо»: від 60 % до 66 % правильних відповідей;
- «Незадовільно»: менше 60 % правильних відповідей.

4 Підсумкова оцінка з навчальної дисципліни визначається як середньозважена оцінка, що враховує загальну оцінку за поточну успішність і оцінку за складання екзамену.

5 Розрахунок загальної підсумкової оцінки за вивчення навчальної дисципліни проводиться за формулою:

$$PK^{екз} = 0,6 \cdot K^{поточ} + 0,4 \cdot E,$$

де $PK^{екз}$ – підсумкова оцінка успішності з дисциплін, формою підсумкового контролю для яких є екзамен;

$K^{поточ}$ – підсумкова оцінка успішності за результатами поточного контролю (за 100-бальною шкалою);

E - оцінка за результатами складання екзамену (за 100-бальною шкалою).

0,6 і 0,4

– коефіцієнти співвідношення балів за поточну успішність і складання екзамену.

6 За виконання індивідуальної самостійної роботи та участь у наукових заходах здобувачам нараховуються додаткові бали.

6.1 Додаткові бали додаються до суми балів, набраних здобувачем вищої освіти за поточну навчальну діяльність (для дисциплін, підсумковою формою контролю для яких є залік), або до підсумкової оцінки з дисципліни, підсумковою формою контролю для якої є екзамен.

6.2 Кількість додаткових балів, яка нараховується за різні види індивідуальних завдань, залежить від їх об'єму та значимості:

- призові місця з дисципліни на міжнародному / всеукраїнському конкурсі наукових студентських робіт – 20 балів;
- призові місця з дисципліни на всеукраїнських олімпіадах – 20 балів;
- участь у міжнародному / всеукраїнському конкурсі наукових студентських робіт – 15 балів
- участь у міжнародних / всеукраїнських наукових конференціях студентів та молодих вчених – 12 балів;
- участь у всеукраїнських олімпіадах з дисципліни – 10 балів
- участь в олімпіадах і наукових конференціях ХНАДУ з дисципліни – 5 балів;
- виконання індивідуальних науково-дослідних (навчально-дослідних) завдань підвищеної складності – 5 балів.

6.3 Кількість додаткових балів не може перевищувати 20 балів.

7 Загальна підсумкова оцінка за вивчення навчальної дисципліни не може перевищувати 100 балів.

Загальна підсумкова оцінка за вивчення навчальної дисципліни визначається згідно зі шкалою, наведеною в таблиці 2.

Таблиця 2 – Шкала оцінювання знань здобувачів за результатами підсумкового контролю з навчальної дисципліни

Оцінка в балах	Оцінка за національною шкалою		Оцінка за шкалою ЄКТС	
			Оцінка	Критерії
	екзамен	залік		
90-100	Відмінно	Зараховано	A	Теоретичний зміст курсу освоєний цілком, без прогалин, необхідні практичні навички роботи з освоєним матеріалом сформовані, усі передбачені програмою навчання навчальні завдання виконані, якість їхнього виконання оцінено числом балів, близьким до максимального
80–89	Добре	Зараховано	B	Теоретичний зміст курсу освоєний цілком, без прогалин, необхідні практичні навички роботи з освоєним матеріалом в основному сформовані, усі передбачені програмою навчання навчальні завдання виконані, якість виконання більшості з них оцінено числом балів, близьким до максимального

75-79			C	Теоретичний зміст курсу освоєний цілком, без прогалин, деякі практичні навички роботи з освоєним матеріалом сформовані недостатньо, усі передбачені програмою навчання навчальні завдання виконані, якість виконання жодного з них не оцінено мінімальним числом балів, деякі види завдань виконані з помилками
67-74	Задовільно	Зараховано	D	Теоретичний зміст курсу освоєний частково, але прогалини не носять істотного характеру, необхідні практичні навички роботи з освоєним матеріалом в основному сформовані, більшість передбачених програмою навчання навчальних завдань виконано, деякі з виконаних завдань, можливо, містять помилки
60-66			E	Теоретичний зміст курсу освоєний частково, деякі практичні навички роботи не сформовані, багато передбачених програмою навчання навчальних завдань не виконані, або якість виконання деяких з них оцінено числом балів, близьким до мінімального.
35-59	Незадовільно	Не зараховано	FX	Теоретичний зміст курсу освоєний частково, необхідні практичні навички роботи не сформовані, більшість передбачених програм навчання навчальних завдань не виконано, або якість їхнього виконання оцінено числом балів, близьким до мінімального; при додатковій самостійній роботі над матеріалом курсу можливе підвищення якості виконання навчальних завдань (з можливістю повторного складання)
0-34			F	Теоретичний зміст курсу не освоєно, необхідні практичні навички роботи не сформовані, усі виконані навчальні завдання містять грубі помилки, додаткова самостійна робота над матеріалом курсу не приведе до якого-небудь значущого підвищення якості виконання навчальних завдань (з обов'язковим повторним курсом)

Політика курсу:

- курс передбачає роботу в колективі, середовище в аудиторії є дружнім, творчим, відкритим до конструктивної критики;
- освоєння дисципліни передбачає обов'язкове відвідування лекцій і практичних занять, а також самостійну роботу;
- самостійна робота передбачає вивчення окремих тем навчальної дисципліни, які винесені відповідно до програми на самостійне опрацювання, або ж були розглянуті стисло;
- усі завдання, передбачені програмою, мають бути виконані у встановлений термін;

- якщо здобувач вищої освіти відсутній на заняттях з поважної причини, він презентує виконані завдання під час самостійної підготовки та консультації викладача;
- курсова робота повинна бути захищена не пізніше, ніж за тиждень до початку екзаменаційної сесії;
- під час вивчення курсу здобувачі вищої освіти повинні дотримуватись правил академічної доброчесності, викладених у таких документах: «Правила академічної доброчесності учасників освітнього процесу ХНАДУ» (https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz_67_01_dobroch_1.pdf), «Академічна доброчесність. Перевірка тексту академічних, наукових та кваліфікаційних робіт на плагіат» (https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz_85_1_01.pdf), «Морально-етичний кодекс учасників освітнього процесу ХНАДУ» (https://www.khadi.kharkov.ua/fileadmin/P_Standart/pologeniya/stvnz_67_01_MEK_1.pdf).
- у разі виявлення факту плагіату здобувач отримує за завдання 0 балів і повинен повторно виконати завдання, які передбачені у силабусі;
- списування під час контрольних робіт та екзаменів заборонені (в т.ч. із використанням мобільних пристроїв). Мобільні пристрої дозволяється використовувати лише під час он-лайн тестування.

Рекомендована література:

1. John D. Kelleher, Brendan Tierney Data Science. – The MIT Press, 2018. – 280 p.
2. Cathy O'Neil Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. – Crown, 2016. – 272 p.
3. Paul Deitel Intro to Python for Computer Science and Data Science: Learning to Program With AI, Big Data and the Cloud. – Pearson, 2019. – 831 p.

Додаткові джерела:

1. Big Data & Analytics Tutorial. Електронний ресурс: https://www.tutorialspoint.com/big_data_tutorials.htm
2. Big Data Tutorial Library. Електронний ресурс: <https://dataflair.training/blogs/big-data-tutorials-home/>

Розробник (розробники)
силабусу навчальної дисципліни
канд. техн. наук, доц.



підпис

Владислав ШУЛЯКОВ

Завідувач кафедри КТМ
докт. техн. наук, проф.



підпис

Олег НІКОНОВ