



Хмарні технології обробки даних

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни	
Рівень вищої освіти	Перший (бакалаврський)
Галузь знань	11 Математика та статистика
Спеціальність	113 Прикладна математика
Освітня програма	Математичні методи криптографічного захисту інформації
Статус дисципліни	Вибіркова
Форма навчання	очна (денна)
Рік підготовки, семестр	3 курс, весняний семестр
Обсяг дисципліни	Загальна кількість: 120 годин / 4 кредити Лекційних занять: 36 годин Лабораторних занять: 36 годин Самостійна робота студентів: 48 годин
Семестровий контроль/ контрольні заходи	Залік, МКР, поточний контроль
Розклад занять	http://ipt.kpi.ua/navchalnij-protses
Мова викладання	Українська
Інформація про керівника курсу / викладачів	Лектор: професор, доктор технічних наук, Шелестов Андрій Юрійович, andrii.shelestov@gmail.com ¹ Лабораторні:
Розміщення курсу	https://classroom.google.com/c/Mjc4MzYwMzE1NjQ1

Програма навчальної дисципліни

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

Дисципліна "Хмарні технології обробки даних" присвячена вивченню загальних понять та інфраструктури хмарних обчислень, ознайомлення і оволодіння сучасними хмарними платформами (зокрема Amazon Web Services – AWS, Google Earth Engine – GEE, та ін.), отримання навичок реалізації методів Data Science в цих хмарних платформах.

Метою навчальної ознайомлення студентів з основними поняттями хмарних сервісів, методами і принципами їх побудови, реалізації вивчених раніше алгоритмів в хмарному середовищі.

За результатами навчання студент в межах курсу буде **знати**:

- знати основи хмарних технологій, властивості та функціональні можливості сучасних хмарних платформ щодо обробки гетерогенних даних;
- знати сучасні інформаційні технології для побудови цифрових двійників;
- знати методи статистичного оцінювання параметрів багатовимірних процесів, методи оптимізації в умовах невизначеності із використанням хмарних технологій;
- знати сучасні інформаційні технології зберігання та обробки великих об'ємів гетерогенних даних, технології Data Mining.

За результатами навчання студент в межах курсу буде **уміти**:

- виконувати гармонізацію, профайлинг, трансформацію, попередню обробку, очистку та збагачення багатовимірних різнорідних даних із використанням хмарних технологій;

¹ Електронна пошта викладача або інші контакти для зворотного зв'язку, можливо зазначити прийомні години або години для комунікації у разі зазначення контактних телефонів. Для силабусу дисципліни, яку викладає багато викладачів (наприклад, історія, філософія тощо) можна зазначити сторінку сайту де представлено контактну інформацію викладачів для відповідних груп, факультетів, інститутів.

- використовувати та налаштовувати хмарне середовище та прикладне програмне забезпечення для роботи із різнорідними даними;
- використовувати методи машинного та глибокого навчання для розв'язання прикладних задач на основі гетерогенних даних;
- реалізовувати методи Data Science на хмарних технологіях (AWS, GEE).

Силабус навчальної дисципліни «Хмарні технології обробки даних» розроблено на основі гнучкого підходу до формування навчальних завдань та видів діяльності, які потрібні студентам для досягнення запланованих результатів навчання з подальшим проектуванням отриманого досвіду таким чином, щоб максимально підвищити ефективність навчання студентів в подальшому.

Загальні компетентності (ЗК)

ЗК 1 – Здатність учитися і оволодівати сучасними знаннями.

ЗК 2 – Здатність застосовувати знання у практичних ситуаціях.

ЗК 3 – Здатність генерувати нові ідеї (креативність).

Фахові компетентності (ФК)

ФК6 – Здатність розв'язувати професійні задачі за допомогою комп'ютерної техніки, комп'ютерних мереж та Інтернету, в середовищі сучасних операційних систем, з використанням стандартних офісних додатків.

ФК 9 – Здатність до проведення математичного і комп'ютерного моделювання, аналізу та обробки даних, обчислювального експерименту, розв'язання формалізованих задач за допомогою спеціалізованих програмних засобів.

ФК 12 – Здатність до пошуку, систематичного вивчення та аналізу науково-технічної інформації, вітчизняного й закордонного досвіду, пов'язаного із застосуванням математичних методів для дослідження різноманітних процесів, явищ та систем.

ФК 14 – Здатність сформулювати математичну постановку задачі, спираючись на постановку мовою предметної галузі, та обирати метод її розв'язання, що забезпечує потрібні точність і надійність результату.

Програмні результати навчання

РН 2 – Володіти основними положеннями та методами математичного, комплексного та функціонального аналізу, лінійної алгебри та теорії чисел, аналітичної геометрії, теорії диференціальних рівнянь, зокрема рівнянь у частинних похідних, теорії ймовірностей, математичної статистики та випадкових процесів, чисельними методами.

РН 3 – Формалізувати задачі, сформульовані мовою певної предметної галузі; формулювати їх математичну постановку та обирати раціональний метод вирішення; розв'язувати отримані задачі аналітичними та чисельними методами, оцінювати точність та достовірність отриманих результатів.

РН 8 – Поєднувати методи математичного та комп'ютерного моделювання з неформальними процедурами експертного аналізу для пошуку оптимальних рішень.

РН 10 – Володіти методиками вибору раціональних методів та алгоритмів розв'язання математичних задач оптимізації, дослідження операцій, оптимального керування і прийняття рішень, аналізу даних.

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

Для вивчення дисципліни «Хмарні технології обробки даних» студент має бути знайомий з класичними методами для роботи з масивами даних. Знати основні інструменти програмного забезпечення для обробки даних, основи Python для Data Science. Використовувати теорію математичної статистики та теорії ймовірності для розв'язання математичних задач.

3. Зміст навчальної дисципліни

Розділ 1. Хмарні технології. Загальні відомості. Основні характеристики

Тема 1.1 Огляд Amazon Web Services (загальні поняття, інфраструктура)

Тема 1.2 Обчислювальні сервіси (Elastic Compute, LightSail) та мережева організація

Тема 1.3 Зберігання даних (Elastic Block Store – EBS, Simple Storage Service – S3)

Розділ 2. Amazon Web Services та бази даних

Тема 2.1 Amazon Web Services та бази даних (Amazon RDS, Redshift, Athena, DynamoDB)

Тема 2.2 Сервіси Amazon, як складова сучасного Data Warehouse. Хмарні бази даних

Розділ 3. Моніторинг та масштабування ресурсів (AWS)

Тема 3.1 Моніторинг та масштабування ресурсів (Amazon CloudWatch, Elastic Load Balancing and AutoScaling)

Тема 3.2 Масштабування додатків та збільшення потужності у разі потреби, моніторингові інструменти. Оцінка вартості послуг

Розділ 4. Аспекти безпеки в AWS

Тема 4.1 Аспекти безпеки в AWS (оцінка ризиків, AWS Shared Responsibility Model, Identity and Access Management)

Тема 4.2 Ролі та акаунти у AWS (Directory Services, ізоляція доступу - Hive and VPC, endpoints, аудит на моніторинг) та моніторинг вартості використання сервісів

Тема 4.3 Моделі безпеки, розмежування прав та ізоляція доступу у AWS.

Рекомендований перелік лабораторних робіт (комп'ютерний практикум)

Лабораторна робота № 1. Основи роботи по використанню AWS Management Console.

Лабораторна робота № 2. Використанням AWS Simple Storage Service (S3).

Лабораторна робота № 3. Робота із AWS DynamoDB (serverless database).

Лабораторна робота № 4. Основи керування ресурсами AWS засобами Python SDK.

Лабораторна робота № 5 (додаткова). Встановлення та налаштування Open Data Cube.

4. Навчальні матеріали та ресурси

Основні:

1. Nayyar, Anand. *Handbook of Cloud Computing: Basic to Advance research on the concepts and design of Cloud Computing*. BPB Publications, 2019.
2. Hurwitz, Judith S., and Daniel Kirsch. *Cloud computing for dummies*. John Wiley & Sons, 2020.
3. Shelestov, Andrii, et al. "Exploring Google earth engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping." *frontiers in Earth Science* 5 (2017): 17.
4. Shelestov, Andrii, et al. "Cloud Approach to Automated Crop Classification Using Sentinel-1 Imagery." *IEEE Transactions on Big Data* (2019).
5. <https://docs.aws.amazon.com/>

Додаткова література

6. <https://developers.google.com/earth-engine/>.
7. Rolland, J. F., Haugomard, A., Castel, F., Aubrun, M., Yao, W., Datcu, M., ... & Bylicki, M. (2020, September). CANDELA: A cloud platform for Copernicus Earth observation data analytics. In *IEEE International Geoscience & Remote Sensing Symposium-IGARSS 2020*. IEEEExplore digital library.
8. Cheshire, J. (2019). *Exam Ref AZ-900 Microsoft Azure Fundamentals*. Microsoft Press.
9. <https://www.opendatacube.org/>.

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

Навчання здійснюється на основі студентоцентрованого підходу та стратегії взаємодії викладача та студента для засвоєння студентами матеріалу та розвитку у них практичних навичок. Для проведення занять застосовується практичний метод. Для лекційних занять використовуються пояснювально-ілюстративний метод та метод проблемного виконання, для проведення лабораторних робіт використовується частково-пошуковий та дослідницький методи навчання, при яких викладач ставить перед студентами проблему, і ті вирішують її самостійно або під керівництвом викладача, висувачи ідеї, перевіряючи їх, підбираючи для цього необхідні джерела інформації, методи, підходи тощо.

Загальна структура курсу

Назви змістових модулів і тем	Кількість годин				
	Всього	у тому числі			
		Лекції	Практ. (семін.)	Лаборант. (комп.пр.)	СРС
1	2	3			6
Розділ 1. Хмарні технології. Загальні відомості. Основні характеристики					
Тема 1.1 Огляд Amazon Web Services (загальні поняття, інфраструктура)	14	4			10
Тема 1.2 Обчислювальні сервіси (Elastic Compute, LightSail) та мережева організація	20	4		9	7
Тема 1.3 Зберігання даних (Elastic Block Store – EBS, Simple Storage Service – S3)	4	4			
Розділ 2. Amazon Web Services та бази даних					
Тема 2.1 Amazon Web Services та бази даних (Amazon RDS, Redshift, Athena, DynamoDB)	12	4			8
Тема 2.2 Сервіси Amazon, як складова сучасного Data Warehouse. Хмарні бази даних	21	4		9	8
Розділ 3. Моніторинг та масштабування ресурсів (AWS)					
Тема 3.1 Моніторинг та масштабування ресурсів (Amazon CloudWatch, Elastic Load Balancing and AutoScaling)	13	4		9	
Тема 3.2 Масштабування додатків та збільшення потужності у разі потреби, моніторингові інструменти. Оцінка вартості послуг	7	2			5
Розділ 4. Аспекти безпеки в AWS					
Тема 4.1 Аспекти безпеки в AWS (оцінка ризиків, AWS Shared Responsibility Model, Identity and Access Management)	9	4			5
Тема 4.2 Ролі та акаунти у AWS (Directory Services, ізоляція доступу - Hive and VPC, endpoints, аудит на моніторинг) та моніторинг вартості використання сервісів	16	2		9	5
Тема 4.3 Моделі безпеки, розмежування прав та ізоляція доступу у AWS.	4	4			
Всього годин	120	36		36	48

Лекційні заняття

№ з/п	Назва теми лекції та перелік основних питань (перелік дидактичних засобів, посилання на літературу, кількість годин)	Кількість годин
1	Лекція 1-2. Огляд Amazon Web Services (загальні поняття, інфраструктура) Література: 1, 2, 3, 4, 5.	4
2	Лекція 3-4. Обчислювальні сервіси (Elastic Compute, LightSail) та мережева організація Література: 1, 2, 3, 4, 5.	4
3	Лекція 5-6. Зберігання даних (Elastic Block Store – EBS, Simple Storage Service – S3) Література: 1, 2, 3, 4, 5.	4
4	Лекція 7-8. Amazon Web Services та бази даних (Amazon RDS, Redshift, Athena, DynamoDB) Література: 1, 2, 3, 4, 5.	4
5	Лекція 9-10. Сервіси Amazon, як складова сучасного Data Warehouse. Хмарні бази даних Література: 1, 2, 3, 4, 5.	4
6	Лекція 11-12. Моніторинг та масштабування ресурсів (Amazon CloudWatch, Elastic Load Balancing and AutoScaling) Література: 1, 2, 3, 4, 5.	4
7	Лекція 13. Масштабування додатків та збільшення потужності у разі потреби, моніторингові інструменти. Оцінка вартості послуг	2

	Література: 1, 2, 3, 4, 5.	
8	Лекція 14-15. Аспекти безпеки в AWS (оцінка ризиків, AWS Shared Responsibility Model, Identity and Access Management) Література: 1, 2, 3, 4, 5.	4
9	Лекція 16. Полі та акаунти у AWS (Directory Services, ізоляція доступу - Hive and VPC, endpoints, аудит на моніторинг) та моніторинг вартості використання сервісів Література: 1, 2, 3, 4, 5.	2
10	Лекція 17-18. Моделі безпеки, розмежування прав та ізоляція доступу у AWS. Література: 1, 2, 3, 4, 5.	4
	Всього годин	36

Лабораторні заняття (комп'ютерний практикум)

Основні завдання циклу лабораторних занять (комп'ютерного практикуму) присвячені закріпленню знань, отриманих на лекційних заняттях та практичному оволодінню методами інтелектуальних обчислень.

№ з/п	Назва лабораторної роботи (комп'ютерного практикуму)	Кількість ауд. годин
1	Основи роботи по використанню AWS Management Console	9
2	Використанням AWS Simple Storage Service (S3)	9
3	Робота із AWS DynamoDB (serverless database)	9
4	Основи керування ресурсами AWS засобами Python SDK	9
5	Встановлення та налаштування Open Data Cube	
	Всього	36

Контрольні роботи

МКР поділяється на 2 контрольні роботи. Метою проведення контрольних робіт є перевірка рівня засвоєння теоретичного і практичного матеріалу отриманого на лекційних та практичних заняттях.

6. Самостійна робота студента/аспіранта

До СРС відносяться такі види робіт: вивчення теоретичного матеріалу, виконання лабораторних робіт (ЛР). При здачі лабораторних робіт оформлюються протоколи, які підписуються викладачем. Результати навчальної діяльності студентів оцінюються за 100-бальною шкалою. Система рейтингових (вагових) балів та критерії оцінювання наведені у Положенні про рейтингову систему оцінювання з даного кредитного модуля. Семестровим контролем є залік.

Самостійна робота

№ з/п	Назва розділу, теми (окремого питання), що виноситься на самостійне опрацювання	Кі-сть годин СРС
1	Історія розвитку хмарних технологій. Google Earth Engine (GEE), Amazon Web Service (AWS), Microsoft Azure та ін.	10
2	Знайомство з Google Earth Engine (GEE)	7
3	Хмарний сервіс GEE та основні функціональні можливості для роботи з супутниковими даними	8
4	Налаштування програмного забезпечення для глибокого навчання на вузлах EC2	8
5	Методи Data Science на хмарних технологіях	5
6	Класифікація та кластеризація засобами GEE	5
7	Розв'язання прикладних задач на створених вузлах EC2 та S3	5
	Всього	48

Студент допускається до заліку за наявності зданих та підписаних всіх лабораторних робіт. Залік з дисципліни «Хмарні технології обробки даних» є підсумком всієї роботи студента

в семестрі і враховує всі види робіт, які студенти зобов'язані виконати протягом семестру згідно Рейтингової системи оцінювання.

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Відвідування занять

Відвідування лекцій та лабораторних занять, а також відсутність на них, не оцінюється. Однак студентам рекомендується відвідувати заняття, оскільки на них викладається теоретичний матеріал та розвиваються навички, необхідні для успішного виконання лабораторних робіт та отримання заліку. Система оцінювання орієнтована на отримання балів за активність студента, а також виконання завдань, які здатні розвинути його практичні уміння та навички.

Календарний рубіжний контроль

Календарний контроль проводиться двічі на семестр як моніторинг поточного стану виконання вимог силабусу. Календарний контроль базується на поточній рейтинговій оцінці. Умовою позитивної атестації є значення поточного рейтингу студента не менше 50% від максимально можливого на час атестації. Бал, необхідний для отримання позитивного календарного контролю доноситься до студентів викладачем не пізніше ніж за 2 тижні до початку календарного контролю.

Академічна доброчесність

Політика та принципи академічної доброчесності визначені у розділі 3 Кодексу честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Детальніше: <https://kpi.ua/code>.

Норми етичної поведінки

Норми етичної поведінки студентів і працівників визначені у розділі 2 Кодексу честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Детальніше: <https://kpi.ua/code>.

Процедура оскарження результатів контрольних заходів

Студенти мають можливість підняти будь-яке питання, яке стосується процедури контрольних заходів та очікувати, що воно буде розглянуто згідно із наперед визначеними процедурами.

Студенти мають право оскаржити результати контрольних заходів, але обов'язково аргументовано, пояснивши з яким критерієм не погоджуються відповідно до оціночного листа та/або зауважень.

Дистанційне навчання (необов'язковий пункт)

Дистанційне навчання через проходження онлайн-курсів за певною тематикою допускається за умови погодження зі студентами. У разі, якщо невелика кількість студентів має бажання пройти онлайн-курс за певною тематикою, вивчення матеріалу за допомогою таких курсів допускається, але студенти повинні виконати всі завдання, які передбачені у навчальній дисципліні.

Інклюзивне навчання

Навчальна дисципліна «Хмарні технології обробки даних» може викладатися для більшості студентів з особливими освітніми потребами, крім студентів з серйозними вадами зору, які не дозволяють виконувати завдання за допомогою персональних комп'ютерів, ноутбуків та/або інших технічних засобів.

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

Рейтинг студента з дисципліни складається з балів, що отримуються:
за 2 комп'ютерні практикуми, 2 контрольні роботи (1 МКР поділяється на 2 контрольні роботи).

Семестровим контролем є **залік**.

Система рейтингових (вагових) балів та критерії оцінювання:

1. Виконання та захист лабораторних робіт (комп'ютерного практикуму):

- повне виконання, повна відповідь при захисті – 20 балів;
- працююча програма, але неповна відповідь при захисті – 10–15 балів;
- непрацююча програма, роботу не зараховано – 0 балів;
- **максимальна** оцінка за практикум номер 1–10 балів

2. Модульний контроль:

1 МКР поділяється на 2 контрольні роботи, кожна з яких оцінюється так:

- повне виконання завдання 13...15;
- неповне виконання завдання (є неprincipові помилки, неточності) 9...12;
- неповне виконання завдання (є принципіві помилки) 5...8;
- незадовільне виконання завдання 0...4.

4. Заохочувальні бали:

- виконання додаткового завдання з комп'ютерного практикуму №4 1...3 бали;
- участь в олімпіадах з програмування, модернізація методичних матеріалів з дисципліни до 10 заохочувальних балів.

Розрахунок шкали рейтингу:

$$R = \text{ЛРН}\#1 (10 \times 1) + \text{ЛРН}\#2-4 (20 \times 3) + \text{МКР} (15 \times 2) = 100 \text{ балів}$$

Для отримання студентом відповідної семестрової оцінки його рейтинг має бути:

Рейтингові бали	Оцінка ECTS	Визначення	Традиційна оцінка
95...100	A	Відмінно	Відмінно
85...94	B	Дуже добре	Добре
75...84	C	Добре	
65...74	D	Задовільно	Задовільно
60...64	E	Достатньо	
50...59	FX	Незадовільно	

9. Додаткова інформація з дисципліни (освітнього компонента)

- Проходження додаткових курсів за тематикою навчальної дисципліни та наявність сертифікатів за їх результатом дає можливість підвищити свій загальний рейтинг після узгодження із викладачем.
- Виконання 5 додаткової лабораторної роботи дає можливість отримати додаткові бали до свого загального рейтингу.

Робочу програму навчальної дисципліни (силабус):

Складено д.т.н., проф., Шелестов А.Ю.

Ухвалено кафедрою ММАД (протокол № 14 від 22.06.2022),

Погоджено Методичною комісією НН ФТІ (протокол № 6 від 30.06.2022)