

**Чернівецький національний університет імені Юрія Федьковича**

(повне найменування закладу вищої освіти)

**Навчально-науковий інститут фізико-технічних і комп'ютерних наук**

(назва інституту/факультету)

**Кафедра комп'ютерних систем та мереж**

(назва кафедри)

## **СИЛАБУС**

**навчальної дисципліни**

### ***Технології IoT & IoE для Big Data Analysis***

(вказати назву навчальної дисципліни (іноземною, якщо дисципліна викладається іноземною мовою))

#### **обов'язкова**

(обов'язкова чи вибіркова)

**Освітньо-наукова програма – Комп'ютерна інженерія технологій**

***інтернету речей та кіберфізичних систем***

**Спеціальність 123 – Комп'ютерна інженерія**

(шифр і назва спеціальності)

**Галузь знань 12 – Інформаційні технології**

(вказати: перший (бакалаврський)/другий (магістерський)/третій (освітньо-науковий))

***Навчально-науковий інститут фізико-технічних та комп'ютерних наук***

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

**Мова навчання – українська, англійська**

(мова, на якій читається дисципліна)

**Розробник: Танасюк Юлія Володимирівна, доцент кафедри КСМ, кандидат фіз.-мат. наук,**

(вказати авторів (викладач (ів)), їхні посади, наукові ступені, вчені звання)

**Профайл викладача (-ів)** <https://csn.chnu.edu.ua/employees/tanasyuk-yuliya-volodymyrivna/>

**Контактний тел.** +(38) 0372 50 94 32.

**E-mail:** [y.tanasyuk@chnu.edu.ua](mailto:y.tanasyuk@chnu.edu.ua)

**Сторінка курсу в Moodle** <https://moodle.chnu.edu.ua/course/view.php?id=5098>

**Консультації** *on-line: понеділок з 14.40 до 16.00*

## **1. Анотація дисципліни**

Вивчення дисципліни «Технології IoT & IoE для Big Data Analysis» дає змогу сформуванню у здобувачів освіти компетенції, необхідні для розв'язання практичних задач професійної та наукової діяльності, пов'язаної з обробленням та аналізом великих даних, згенерованих пристроями у середовищах Інтернету речей.

**2. Мета дисципліни** «Технології IoT & IoE для Big Data Analysis»: формування у студентів здатностей застосовувати програмні методи та засоби оброблення даних для аналізу великих даних та прийняття управлінських рішень у різних галузях науки та бізнесу.

### **2.1. Завдання дисципліни:**

- Пояснити, як організації можуть вилучати інформацію та генерувати нові ідеї з даних IoT.
- Розуміти та реалізовувати етапи життєвого циклу аналізу великих даних
- Пояснити різні типи аналітики даних: описову, прогнозувальну та перспективну.
- Використовувати Python для створення процесу одержання, маніпулювання та візуалізації даних з IoT-джерел.
- Застосовувати розвідувальний аналіз даних для виявлення новітніх тенденцій.
- Пояснити призначення машинного навчання.
- Вміти подавати та передавати різні типи даних.
- Описувати еволюцію технологій маніпулювання від SQL до NoSQL.
- Розуміти та пояснити розвиток сучасних дата-центрів та розподілених масштабованих платформ Big Data типу Apache Hadoop.

**3. Пререквізити.** Для успішного опанування матеріалу слухачам необхідно попередньо вивчити такі дисципліни, як: Основи алгоритмізації, Організація баз даних, Теорія ймовірності та матстатистика, Алгоритми та методи обчислень, Комп'ютерні системи, Основи ООП, Пристрої зв'язку з об'єктом, Мікроконтролери, Основи IoT & IoE, Комп'ютерні системи III, Технології хмарних обчислень.

## **4. Результати навчання**

У результаті вивчення дисципліни студенти повинні:

### **Знати:**

- архітектурні моделі Big Data;
- технології віртуалізації та контейнерної технології виконання програмного коду на сервері;
- технології Hadoop Big Data, MapReduce, HDFS, Kafka, Cassandra, Spark;
- можливості, що надає мова програмування Python для аналізу великих даних;
- технології імпортування даних з Інтернету та засоби для кореляційного аналізу в Pandas;
- засоби керування та оброблення даних, доступні у базаж даних SQL і NoSQL;

- технології регресійного аналізу великих даних засобами Python;
- технології візуалізації даних Tableau, Pyplot, Plotly, Scikit-Learn, Seaborn, Folium, Leaflet.js Pandas та Matplotlib;
- методи та засоби проведення аналізу великих даних з використанням хмарних, туманних і крайових обчислень;
- заходи для організації оброблення та візуалізації великих даних на базі Raspberry Pi.

#### **Уміти:**

- використовувати технології IoT та програмні засоби для обробки та аналізу великих даних.
- засобами мови програмування Python створювати повний цикл видобування, маніпулювання та подання даних із різних джерел;
- застосовувати пошуковий аналіз даних для генерування рішень на основі одержаних даних;
- подавати та візуалізувати великі обсяги даних;
- опрацьовувати слабоструктуровані та різнотипні дані, що надходять з різних джерел, засобами мови Python та баз даних NoSQL .

#### **Набути компетентностей:**

##### *Загальні компетенції (ЗК):*

- ЗК1. Здатність до адаптації та дій в новій ситуації.
- ЗК2. Здатність до абстрактного мислення, аналізу і синтезу.
- ЗК4. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.
- ЗК6. Здатність виявляти, ставити та вирішувати проблеми.
- ЗК7. Здатність приймати обґрунтовані рішення.
- ЗК8. Здатність спілкуватися іноземною мовою.

##### *Фахові компетентності спеціальності (СК):*

- СК2. Здатність розробляти алгоритмічне та програмне забезпечення, компоненти комп'ютерних систем та мереж, Інтернет додатків, кіберфізичних систем з використанням сучасних методів і мов програмування, а також засобів і систем автоматизації проектування.
- СК4. Здатність будувати та досліджувати моделі комп'ютерних систем та мереж.
- СК6. Здатність використовувати та впроваджувати нові технології, включаючи технології розумних, мобільних, зелених і безпечних обчислень, брати участь в модернізації та реконструкції комп'ютерних систем та мереж, різноманітних вбудованих і розподілених додатків, зокрема з метою підвищення їх ефективності.
- СК9. Здатність представляти результати власних досліджень та/або розробок у вигляді презентацій, науково-технічних звітів, статей і доповідей на науково-технічних конференціях.
- СК10. Здатність ідентифікувати, класифікувати та описувати роботу програмно-технічних засобів, комп'ютерних систем, мереж та їхніх компонентів;
- СК11. Здатність обирати ефективні методи розв'язування складних задач комп'ютерної інженерії, критично оцінювати отримані результати та аргументувати прийняті рішення.
- СК12. Здатність використовувати методи аналізу, ідентифікації й синтезу комп'ютерних систем та мереж, кіберфізичних систем, засобів Інтернету речей та IT-інфраструктур.

*Програмні результати навчання (РН):*

РН1. Застосовувати загальні підходи пізнання, методи математики, природничих та інженерних наук до розв'язання складних задач комп'ютерної інженерії.

РН2. Знаходити необхідні дані, аналізувати та оцінювати їх.

РН6. Аналізувати проблематику, ідентифікувати та формулювати конкретні проблеми, що потребують вирішення, обирати ефективні методи їх вирішення.

РН8. Застосовувати знання технічних характеристик, конструктивних особливостей, призначення і правил експлуатації програмотехнічних засобів комп'ютерних систем та мереж для вирішення складних задач комп'ютерної інженерії та дотичних проблем.

РН9. Розробляти програмне забезпечення для вбудованих і розподілених застосувань, мобільних і гібридних систем.

РН10. Здійснювати пошук інформації в різних джерелах для розв'язання задач комп'ютерної інженерії, аналізувати та оцінювати цю інформацію.

РН12. Вільно спілкуватись усно і письмово українською мовою та однією з іноземних мов (англійською, німецькою, італійською, французькою, іспанською) при обговоренні професійних питань, досліджень та інновацій в галузі інформаційних технологій.

РН14. Планувати і виконувати наукові дослідження в сфері комп'ютерної інженерії, формулювати і перевіряти гіпотези, обирати методики та інструменти, аналізувати результати, обґрунтовувати висновки.

РН15. Проводити науково-дослідну і дослідно-конструкторську роботу у сфері спеціалізованих, проблемно-орієнтованих вискоефективних комп'ютерних системи з використанням сучасних технологій IoT і КФС, мобільних і гібридних обчислень, Big Data аналізу, новітніх апаратно-програмних рішень на сучасній елементній базі, зокрема, програмованих логічних інтегральних схем, мікроконтролерів, мікрокомп'ютерів, багатоядерних процесорів.

## 5. Опис навчальної дисципліни

### 5.1. Загальна інформація

Назва навчальної дисципліни <i>Технології IoT &amp; IoE для Big Data Analysis</i>												
Форма навчання	Рік підготовки	Семестр	Кількість			Кількість годин						Вид підсумкового контролю
			кредитів	годин	змістових модулів	лекції	практичні	семінарські	лабораторні	самостійна робота	індивідуальні завдання	
Денна	2(6)	3 (11)	4	120	2	15	-	-	30	75	-	Іспит

## 5.2. Дидактична карта навчальної дисципліни

### Теоретичний зміст програми навчальної дисципліни

#### Змістовий модуль 1. Джерела та типи великих даних

Тема 1. Джерела великих даних у концепції Інтернету Речей. Визначення Big Data.

Тема 2. Основи аналізу даних та машинного навчання.

#### Змістовий модуль 2. Архітектурні моделі Big Data

Тема 3. Класифікація даних.

Тема 4. Технології віртуалізації

Тема 5. Розподілені платформи для обробки великих даних

Назви змістових модулів і тем	Кількість годин					
	Денна форма					
	усього	у тому числі				
л		п	лаб	інд	с.р.	
1	2	3	4	5	6	7
Змістовий модуль 1. Джерела та типи великих даних						
Тема 1.	20	4	-	2	-	14
Тема 2.	28	2	-	10	-	16
Разом за ЗМ1	48	6	-	12	-	30
Змістовий модуль 2. Архітектурні моделі Big Data						
Тема 3.	19	3	-	4	-	12
Тема 4.	27	3	-	6	-	18
Тема 5.	26	3	-	8	-	15
Разом за ЗМ 2	72	9	-	18	-	45
<b>Усього годин</b>	<b>120</b>	<b>15</b>	<b>-</b>	<b>30</b>	<b>-</b>	<b>75</b>

### 5.3. Тематика лабораторних занять

№	Назва теми (завдання)	Кількість годин
1.	Дослідження вимог до профілю бізнес-аналітика	2
2.	Основи підготовки та аналізу даних	6
3.	Візуалізація даних та описова статистика у Python	4
4.	Лінійна регресія та прогнозування	4
5.	Візуалізація даних за допомогою Tableau	6
6.	Розширена візуалізація. NoSQL бази даних для розподіленої обробки великих даних	8
<b>Разом:</b>		<b>30</b>

**Примітка.** Методичні рекомендації та завдання до лабораторних робіт доступні на інтернет-ресурсах: <https://moodle.chnu.edu.ua/course/view.php?id=5098>

[https://skillsforall.com/course/data-analytics-essentials?courseLang=en-US&instance\\_id=bd4725ae-c054-4bba-8553-e7be9408f90e](https://skillsforall.com/course/data-analytics-essentials?courseLang=en-US&instance_id=bd4725ae-c054-4bba-8553-e7be9408f90e)

**Програмне забезпечення** для виконання лабораторних робіт: хмарна платформа Google Colab, мова програмування Python, веб-оболонка Jupyter Notebook, онлайн-засіб візуалізації Tableau: <https://www.tableau.com/products/public/download>, хмарне середовище IBM Cloudant для роботи з базами даних NoSQL: [cloud.ibm.com](http://cloud.ibm.com).

#### 5.4. Завдання для самостійної роботи

№	Назва теми	Кількість годин
1	Дослідження джерел відкритих даних. Завантаження датасету та збереження даних в форматі csv.	6
2	Програмування Python та SQLite. Призначення утиліти csvsql.	4
3	Статистичні підходи до аналітики великих даних. Використання Pandas. Імпорт даних з мережі Інтернет. Описова статистика в Pandas. Засоби для кореляційного аналізу в Pandas.	12
4	Регресійний аналіз даних в Python	5
5	Модуль Rpyplot. Інструмент Plotly. Типи візуалізації даних.	5
6	Технології Hadoop Big Data. Розподілена обробка MapReduce. HDFS.	4
7	Kafka: розподілена потокова платформа обробки великих даних. Переваги Cassandra.	8
8	Проблема обчислювальної функції. Технологія Spark. Порівняння Spark та MapReduce.	12
9	Lambda та Карра архітектури оброблення великих даних.	8
10	Машинне навчання.	4
11	Захист великих даних.	4
12	Етика використання даних	3
<b>Разом:</b>		<b>75</b>

#### 6. Форми і методи навчання

**Форми навчання** – це проблемні й оглядові лекції, лабораторні заняття, заняття із застосуванням комп'ютерної та телекомунікаційної техніки, інтерактивні заняття з навчанням одних студентів іншими, відеолекції, відеоконференції засобами Google Meet, Cisco Webex, заняття з використанням системи електронного навчання Moodle та освітнього порталу Програми мережних академій Cisco - netacad.com, skillsforall.com.

**Методи:** проблемний виклад матеріалу, частково-пошукові та дослідницькі лабораторні практикуми, презентації, консультації і дискусії, робота в інтернет-класі: електронні лекції, лабораторні роботи, дистанційні консультації та ін., спрямовані на активізацію і стимулювання навчально-пізнавальної діяльності студентів.

**Підходи до навчання:** використовуються студентоцентрикований, проблемно-орієнтований, діяльнісний, комунікативний, професійно-орієнтований, міждисциплінарний підходи.

**Реалізація навчального процесу** здійснюється під час лекційних, лабораторних занять, самостійної позааудиторної роботи з використанням сучасних інформаційних засобів навчання, програмних симуляторів, віртуальних середовищ колективної роботи, консультацій з викладачами.

Для **формувань умінь та навичок** застосовуються такі **методи навчання:**

- вербальні/словесні (лекція, пояснення, розповідь, бесіда, інструктаж);
- наочні (спостереження, ілюстрація, демонстрація);
- практичні (проведення експерименту, практики);

- пояснювально-ілюстративний, який передбачає пред'явлення готової інформації викладачем та її засвоєння студентами;
- репродуктивний (виконання лабораторних завдань за зразком);
- метод проблемного викладу матеріалу на лабораторних заняттях.

## 7. Система контролю та оцінювання

Засоби оцінювання та демонстрування результатів навчання:

- хмарні середовища проектування та візуалізації; - завдання на лабораторному обладнанні; - лабораторні роботи; - тести; - презентації та обґрунтування результатів виконаних завдань.

Формами поточного контролю рівня знань є усна та письмова відповідь студента при захисті виконаних лабораторних робіт, подання результатів практичних завдань у формі звіту, кількість отриманих балів при виконанні тестових завдань. Підсумковий контроль здійснюється шляхом складання фінального іспиту.

### 7.1. Критерії оцінювання результатів навчання з навчальної дисципліни

Критерієм успішного проходження здобувачем освіти підсумкового оцінювання є досягнення ним мінімальних порогових рівнів оцінок за кожним запланованим результатом навчання навчальної дисципліни.

#### Шкала та критерії оцінювання: національна та ЄКТС

Оцінка за шкалою ЄКТС	Критерії	Пояснення	Оцінка за 100-бальною шкалою	Оцінка за національною шкалою (іспит/залік)
<b>A</b>	Відмінний рівень компетентностей у межах обов'язкового матеріалу, з можливими незначними недоліками	відмінно/зараховано	<b>90 – 100</b>	<b>відмінно/зараховано</b>
<b>B</b>	Достатньо високий рівень компетентностей у межах обов'язкового матеріалу без суттєвих (грубих) помилок	дуже добре/зараховано	<b>80-89</b>	<b>добре/зараховано</b>
<b>C</b>	В цілому добрий рівень компетентностей із незначною кількістю помилок	добре/зараховано	<b>70-79</b>	
<b>D</b>	Посередній рівень компетентностей із значною кількістю недоліків, достатній для подальшого навчання або професійної діяльності	задовільно/зараховано	<b>60-69</b>	<b>задовільно/зараховано</b>
<b>E</b>	Мінімально можливий допустимий рівень компетентностей	достатньо/зараховано	<b>50-59</b>	
<b>FX</b>	Незадовільний рівень компетентностей, з можливістю повторного перескладання за умови належного самостійного доопрацювання	(незадовільно/не зараховано) з можливістю повторного складання	<b>35-49</b>	
<b>F</b>	Дуже поганий рівень компетентностей, що вимагає повторного вивчення дисципліни	(незадовільно/не зараховано) з обов'язковим повторним вивченням дисципліни	<b>1-34</b>	

## 7.2. Розподіл максимально можливої кількості балів, які отримують студенти за виконання всіх видів навчальної діяльності

Поточне оцінювання (індивідуальна робота та самостійна робота)					Основи аналізу даних	Підсумковий контроль (іспит)	Сумарна кількість балів
Змістовий модуль 1		Змістовий модуль 2					
T1	T2	T3	T4	T5	МК	40	100
6,5	13,5	6,5	6,5	7	20		

Формами підсумкового контролю рівня знань є складання тестів, виконання лабораторних робіт та контрольних робіт у системі дистанційного навчання.

### **Змістовий модуль 1. Джерела та типи великих даних**

T1. Джерела великих даних у концепції Інтернету Речей. Визначення Big Data. (виконання та захист л/р №1 – 10 балів). BDA1 – підсумковий тест за матеріалами Теми 1 (10 балів).

T2. Основи аналізу даних та машинного навчання. (виконання та захист л/р № 2 – 10 балів, виконання та захист л/р № 3 – 10 балів).

BDA2 – підсумковий тест за матеріалами Теми 2 (10 балів).

BDA3 – підсумковий тест за матеріалами Теми 2 (10 балів).

### **Змістовий модуль 2. Архітектурні моделі Big Data**

T3. Класифікація даних. (виконання та захист л/р № 4 – 10 балів).

BDA4 – підсумковий тест за матеріалами Теми 3 (10 балів).

T4. Технології віртуалізації (виконання та захист л/р № 5 – 10 балів).

BDA5 – підсумковий тест за матеріалами Теми 4 (10 балів).

T5. Розподілені платформи для обробки великих даних (виконання та захист л/р № 6 – 10 балів).

BDA6 – підсумковий тест за матеріалами Теми 5 (10 балів).

МК – підсумковий результат з вивчення онлайн-курсу «Основи аналізу даних» (Data Analytics Essentials) від Cisco Networking Academy. (20 балів).

Підсумковий контроль (іспит) – 40 балів.

**Загальна сума балів за семестр визначається за формулою:**

$$\text{ЗСБ} = ((\text{ЗСЛР} + \text{ТТ} + \text{МК} * 0,6) / 3) + \text{ПК},$$

де ЗСБ – загальна сума балів за семестр (максимум – 100);

ЗСЛР – загальна сума балів за лабораторні роботи (максимум – 60);

ТТ – загальна сума балів за тести по темах (максимум – 60);

МК – підсумковий зважений бали за онлайн-курс «Основи аналізу даних» (максимум – 100 балів);

ПК – результат підсумкового контролю (іспиту) (максимум – 40).



### 7.3. Умови зарахування результатів неформальної освіти

Студент, згідно Положення ЧНУ «Про неформальну освіту» може отримати додаткові бали, або бути звільненим від певних видів роботи з окремих тем, якщо у нього наявні сертифікати про неформальну освіту з питань, які окреслює дисципліна “Технології IoT&IoE для Big Data Analysis”.

## 8. Рекомендована література

### 8.1. Основні (фахова)

1. Інтернет речей для індустріальних і гуманітарних застосунків. У трьох томах. Том 1. Основи і технології / За ред. В. С. Харченка. – Міністерство освіти і науки України, Національний аерокосмічний університет ХАІ, 2019. – 605 с.
2. IoT Fundamentals: Big Data & Analytics. URL: <https://www.netacad.com/courses/iot/big-data-analytics>
3. Data Analytics Essentials. URL: <https://skillsforall.com/course/data-analytics-essentials?courseLang=en-US>
4. О’Нил К. Big Data. Зброя математичного знищення. – К: Форс, 2019. – 336 с.
5. Провост Ф., Фоусет Т. Data Science для бізнесу. Як збирати, аналізувати і використовувати дані. – К.: Наш Формат, 2019. 400 с.
6. Nokeri T. C. Data Science Solutions with Python. – Apress, 2022. 119 p.

### 8.1. Допоміжна

1. Klosterman S. Data science projects with Python: a case study approach to gaining valuable insights from real data with machine learning? 2<sup>nd</sup> edition. Packt Publishing Limited. 2021. 432 p.
2. Kats P. Learn Python by building data science applications. 1<sup>st</sup> edition. Packt Publishing. 2019. 482 p.
3. Ерл Т. Основи Big Data: концепції, алгоритми і технології. – К: Баланс бізнес бус, 2018. – 320 с.
4. Bruce P. Practical statistics for data scientists: 50 essential concepts. 1<sup>st</sup> edition. O’Reilly Media. 2017. 318 p.
5. Muller A.C. Introduction to machine learning with Python: a guide for data scientists. 1<sup>st</sup> edition. O’Reilly Media. 2016. 402 p.

### Інтернет-ресурси

1. Big Data: 20 free Big Data sources everyone should know. URL: <https://www.smartdatacollective.com/big-data-20-free-big-data-sources-everyone-should-know/>
2. Big Data & Analytics Tutorial. URL: [https://www.tutorialspoint.com/big\\_data\\_tutorials.htm](https://www.tutorialspoint.com/big_data_tutorials.htm)
3. Big Data Tutorial Library. URL: <https://data-flair.training/blogs/big-data-tutorials-home/>
4. Machine learning. URL: [https://www.w3schools.com/python/python\\_ml\\_getting\\_started.asp](https://www.w3schools.com/python/python_ml_getting_started.asp)
5. Python Machine Learning Tutorials. URL: <https://realpython.com/tutorials/machine-learning/>
6. Data Science Tutorial for Beginners. URL: <https://www.simplilearn.com/tutorials/data-science-tutorial>