



СИЛАБУС

вибіркового освітнього компонента ХМАРНІ ТЕХНОЛОГІЇ ТА BIG DATA

Обсяг освітнього компоненту (3 кредити / 90 годин)

Освітня програма: «Автоматизація, мехатроніка та робототехніка»
другого (магістерського) рівня вищої освіти
Спеціальність 174 «Автоматизація, комп'ютерно-інтегровані технології та
робототехніка»

ІНФОРМАЦІЯ ПРО ВИКЛАДАЧА



МИРОНОВА Наталя Олексіївна,
канд. техн. наук., доцент

Контактна інформація:
e-mail: natali.myronova@gmail.com

Профіль викладача на сайті університету:
<https://zpu.edu.ua/?q=node/676>

Час і місце проведення консультацій
Відповідно до розкладу, на платформі google meet

ОПИС КУРСУ

Освітня компонента «Хмарні технології та Big Data» є вибірковою освітньою компонентою у підготовці студентів освітньо-наукової програми «Автоматизація, мехатроніка та робототехніка».

Мета освітньої компоненти – розглянути сучасні методи та програмні засоби обробки структурованих та неструктурованих великих даних. Ознайомлення та отримання практичних навиків з аналізу та обробки великих даних з використанням хмарних технологій.

Освітня компонента спрямована на формування базових знань і практичних навичок використання програмних засобів та новітніх комп'ютерно-інтегрованих технологій для аналізу и обробки Big Data для систем автоматизації та робототехнічних систем.



МЕТА, КОМПЕТЕНТНОСТІ ТА РЕЗУЛЬТАТИ НАВЧАННЯ

1. *Мета:* розглянути сучасні методи та програмні засоби обробки структурованих та неструктурованих великих даних. Ознайомлення та отримання практичних навиків з аналізу та обробки великих даних з використанням хмарних технологій.

Завдання:

– ознайомити студентів з основними концепціями структурованих та неструктурованих даних, їхніми особливостями та сферами застосування в автоматизації і робототехніці;

– розглянути сучасні методи обробки великих даних, включаючи алгоритми машинного навчання, методи аналізу даних та інструменти для обробки текстових, числових і графічних даних;

– навчити використовувати спеціалізовані програмні платформи для обробки Big Data, такі як Hadoop, Spark, MongoDB, а також інструменти для візуалізації та аналізу даних;

– надати базові знання про хмарні технології, включаючи сервіси для обробки даних (наприклад, AWS, Google Cloud Platform, Microsoft Azure) та їх інтеграцію у виробничі процеси для автоматизації;

– сформувати навички практичної роботи з хмарними платформами, що підтримують обробку великих даних, та розробки моделей аналізу даних з урахуванням особливостей конкретних задач автоматизації;

– навчитися застосовувати методи обробки структурованих та неструктурованих даних для побудови аналітичних моделей та їх оптимізації в системах автоматизації і робототехнічних систем;

– надати практичні навички управління даними, їхньої обробки та інтерпретації результатів з використанням сучасних програмних засобів для комплексного підходу до управління інформацією в робототехнічних та автоматизованих системах.

2. *Компетентності та результати навчання, формування яких забезпечує вивчення дисципліни.*

У результаті вивчення дисципліни «Хмарні технології та Big Data» здобувач другого (магістерського) рівня вищої освіти повинен отримати:

інтегральну компетентність:

здатність розв'язувати складні задачі і проблеми автоматизації та комп'ютерно-інтегрованих технологій у професійній діяльності та/або у процесі навчання, що передбачає проведення досліджень та/або провадження інноваційної діяльності та характеризується комплексністю та невизначеністю умов і вимог;

загальні компетентності:

ЗК1. Здатність проведення досліджень на відповідному рівні.

ЗК2. Здатність генерувати нові ідеї (креативність).

ЗК3. Здатність до абстрактного мислення, аналізу та синтезу.

спеціальні (фахові, предметні) компетентності:



СК1. Здатність здійснювати автоматизацію складних технологічних об'єктів та комплексів, створювати кіберфізичні системи на основі інтелектуальних методів управління та цифрових технологій з використанням баз даних, баз знань, методів штучного інтелекту, робототехнічних та інтелектуальних мехатронних пристроїв;

СК7. Здатність застосовувати спеціалізоване програмне забезпечення та цифрові технології для розв'язання складних задач і проблем автоматизації та комп'ютерно-інтегрованих технологій.

СК14. Здатність використовувати підходи штучного інтелекту та машинного навчання, хмарні технології, Інтернет речей та Big Data для комплексного розв'язання складних задач і проблем в галузі автоматизації, робототехніки та безпілотних систем, сприяючи підвищенню інноваційності та конкурентоспроможності технологічних рішень.

Очікувані програмні результати навчання:

РН01. Створювати системи автоматизації, кіберфізичні виробництва на основі використання інтелектуальних методів управління, баз даних та баз знань, цифрових та мережевих технологій, робототехнічних та інтелектуальних мехатронних пристроїв.

РН03. Застосовувати спеціалізовані концептуальні знання, що включають сучасні наукові здобутки, а також критичне осмислення сучасних проблем у сфері автоматизації та комп'ютерно-інтегрованих технологій для розв'язування складних задач професійної діяльності.

РН10. Розробляти і використовувати спеціалізоване програмне забезпечення та цифрові технології для створення систем автоматизації складними організаційно-технічними об'єктами, професійно володіти спеціальними програмними засобами.

РН19. Використовувати підходи штучного інтелекту та машинного навчання, хмарні технології, Інтернет речей та Big Data для комплексного розв'язання складних задач і проблем в галузі автоматизації, робототехніки та безпілотних систем, сприяючи підвищенню інноваційності та конкурентоспроможності технологічних рішень у відповідних сферах.

ПЕРЕДУМОВИ ДЛЯ ВИВЧЕННЯ ДИСЦИПЛІНИ

Вивчення освітньої компоненти «Хмарні технології та Big Data» спирається та є продовженням циклу фахових освітніх компонент наукової та професійної підготовки магістра і ґрунтується на раніше отриманих студентами знаннях та практичних навичках підготовки і захисту бакалаврської роботи.

ПЕРЕЛІК ТЕМ (ТЕМАТИЧНИЙ ПЛАН) ДИСЦИПЛІНИ

Курс освітньої компоненти «Хмарні технології та Big Data» складається з лекцій, лабораторних занять та самостійної роботи. При викладанні дисципліни лектор викладає загальні концепції, положення, теорії тощо. На



лабораторних заняттях здобувачі поглиблюють набуті теоретичні знання в обговореннях індивідуальних доповідей, у тематичних дискусіях, у співбесідах з викладачем, виконують письмові вправи тощо, а також закріплюють навчальний матеріал у самостійному розв'язанні аналогічних завдань при самостійній роботі.

Програма дисципліни «Хмарні технології та Big Data» складається з наступних змістових модулів:

1. Введення до Big Data
2. Програмне забезпечення обробки великих даних Apache Hadoop.
3. Програмне забезпечення обробки великих даних Apache Spark.
4. Хмарні сервіси
5. Вирішення прикладних практичних завдань.

Таблиця 1 – Загальний тематичний план аудиторної роботи.

Номер тижня	Теми лекцій, год.	Теми лабораторних/практичних робіт або семінарів, год.
1	2	3
Змістовий модуль 1. Введення до Big Data та Apache Hadoop		
1	Введення до Big Data. Основні поняття та визначення аналізу та обробки великих даних.	Лабораторна робота №1 Розгортання віртуальної машини з програмним забезпеченням обробки великих даних Apache Hadoop
2-3	Введення до екосистеми Apache Hadoop.	
4-5	Введення до Hadoop Distributed File System та MapReduce	Лабораторна робота №2 Вивчення компонентів ядра Apache Hadoop: Hadoop Distributed File System та MapReduce
6-8	Введення до Hadoop Streaming.	Лабораторна робота №3 Вивчення принципів роботи Hadoop Streaming
Змістовий модуль 2. Введення до Apache Spark		
9-10	Введення до екосистеми Apache Spark	Лабораторна робота №4 Основи роботи з Spark в Google Colab. Resilient Distributed Dataset та DataFrame
11-12	Поняття Resilient Distributed Dataset. Поняття DataFrame. Введення до Spark SQL. Введення до Machine learning Spark library	-
13-14	Огляд практичних прикладних завдань. Огляд хмарних сервісів для обробки	Лабораторна робота №5 Хмарний сервіс Amazon EMR Завдання



великих даних AWS, Google Cloud Platform, Microsoft Azure

САМОСТІЙНА РОБОТА

Перелік завдань на СРС, вид завдань, контрольні заходи.

Назва теми	Графік самостійної роботи	Кількість годин	Контрольні заходи
Історія розвитку Big Data.	Виконання індивідуального завдання	5	40 балів у загальних 100
Історія розвитку програмного забезпечення обробки великих даних Apache Hadoop. Огляд дистрибутивів та хмарних сервісів Apache Hadoop.		5	
Огляд компонентів Apache Hadoop: HDFS, MapReduce, YARN, HBase, Hive, Pig, Sqoop, Flume, Kafka, ZooKeeper, Spark		5	
Парадигма програмування MapReduce		10	
Програмування для Hadoop Streaming з використанням мови Python		10	
Порівняльна характеристика Apache Hadoop та Apache Spark. Компоненти Apache Spark. Функціональне програмування з використанням мови Python. Лямбда функції, загальні функціональні функції, структури даних.		10	
Операції над Resilient Distributed Dataset. Операції над DataFrame		10	
Вивчення основних функцій Spark SQL		5	
Вивчення основних функцій Machine learning Spark library		5	
Огляд хмарних сервісів для обробки великих даних AWS, Google Cloud Platform, Microsoft Azure		5	

РЕКОМЕНДОВАНІ ІНФОРМАЦІЙНІ ТА НАВЧАЛЬНО-МЕТОДИЧНІ ДЖЕРЕЛА

Навчально-методичні розробки:

Навчально-методичні розробки в системі дистанційної навчання moodle.zp.edu.ua



Літературні джерела:

1. Erl, R. Puttini, Z. Mahmood, “Cloud Computing: Concepts, Technology & Architecture (The Pearson Service Technology Series from Thomas Erl)”, Pearson, 2013
2. Jon Hall, Joshua Andrews, “VMware Certified Professional Data Center Virtualization on vSphere 6.7 Study Guide: Exam 2V0-21.19”, Sybex, 2020
3. Elver Sena Sosa, “Mastering VMware NSX for vSphere”, Sybex, 2020
4. Brown, H. Cartwright, M. Gavanda, A. Mauro, K. Novak, P. Valsecchi, “The Complete VMware vSphere Guide: Design a virtualized data center with VMware vSphere 6.7”, Packt Publishing, 2019
5. Silverman, M. Solberg, “OpenStack for Architects: Design production-ready private cloud infrastructure”, Packt Publishing, 2018 (2nd Edition)
6. Jackson, C. Bunch, E. Sigler, J. Denton, “OpenStack Cloud Computing Cookbook: over 100 practical recipes to help you build and operate OpenStack cloud computing, storage, networking, and automation”, Packt Publishing, 2018 (4th Edition)
7. White, “Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale”, O'Reilly Media, 2015 (4th Edition)
8. Sridhar Alla, “Big Data Analytics with Hadoop 3”, Packt Publishing, 2018
9. Holmes, “Hadoop in Practice”, Manning Publications, 2015 (Second Edition)
10. Aven, “Sams Teach Yourself Hadoop™ in 24 Hours”, Pearson Education, 2017
11. Hrishikesh Karambelkar, “Apache Hadoop 3 Quick Start Guide”, Packt Publishing, 2018
12. Aven, “Sams Teach Yourself Apache Spark™ in 24 Hours”, Pearson Education, 2017
13. Bill Chambers&Matei Zaharia, “Spark: The Definitive Guide: Big Data Processing Made Simple”, O'Reilly Media, 2018
14. Jean-Georges Perrin, “Spark in Action, Second Edition: Covers Apache Spark 3 with Examples in Java, Python, and Scala”, Manning Publications, 2020
15. Jules S. Damji, Brooke Wenig, Tathagata Das, Denny Lee, “Learning Spark: Lightning-fast Data Analytics”, O'Reilly Media, 2020, 2nd Edition
16. The Hadoop Ecosystem Table <https://hadoopecosystemtable.github.io/>
17. Hadoop Streaming <https://hadoop.apache.org/docs/current/hadoop-streaming/HadoopStreaming.html>
18. Apache Spark <https://spark.apache.org/>
19. PySpark Tutorial <https://sparkbyexamples.com/pyspark-tutorial/>

ОЦІНЮВАННЯ

Оцінювання навчальних успіхів здобувачів реалізується шляхом проведення поточного та підсумкового контролю успішності.

Форма підсумкового контролю – залік, у формі підсумкового опитування(проходження тесту).



Формами поточного контролю є надання результатів лабораторних робіт, які оцінюються згідно табл.3.

Здобувач освіти отримує позитивну оцінку при сумі балів за всіма видами контролю, не нижче 60 балів. Розрахунок балів наведено у табл.3.

Таблиця 3 – Розрахунок балів із врахуванням контрольного опитування

Поточне оцінювання					Підсумкове опитування (залік)	Сума балів
ЛБ1	ЛБ2	ЛБ3	ЛБ4	ЛБ5	40	100
12	12	12	12	12		

Отже, сумарна кількість балів, яку отримує студент впродовж семестру, складає 100. В залежності від отриманої суми балів до залікової відомості та в залікову книжку виставляється оцінка згідно національної шкали.

Шкала оцінювання: національна та ECTS

Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою	
		для екзамену, курсового проєкту (роботи), практики	для заліку
90 – 100	A	відмінно	зараховано
85 – 89	B	добре	
75 – 84	C		
70 – 74	D	задовільно	
60 – 69	E		
35 – 59	FX	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного складання
1 – 34	F	незадовільно з обов'язковим повторним вивченням дисципліни	не зараховано з обов'язковим повторним вивченням дисципліни

ПОЛІТИКИ КУРСУ

При вивченні навчальної дисципліни від здобувача освіти очікуються – дотримання вимог до складання курсу, старанність, коректне і взаємно цікаве ділове спілкування здобувач/викладач, дотримання засад академічної доброчесності.

Політика щодо відвідування, дедлайнів та перескладання:

- усі види робіт, передбачені курсом, усі завдання, передбачені програмою, мають бути виконані у встановлений термін;
- самостійну роботу здобувач виконує відповідно до методичних вказівок та визначених викладачем завдань і термінів;
- ліквідація заборгованості відбувається під час проведення консультацій з дисципліни, за оприлюдненим графіком.



Політика щодо академічної доброчесності:

- у нашому університеті академічна доброчесність передбачається за замовчуванням; це означає, що викладач очікує, що всі здані роботи є результатом розумової праці та творчості конкретного здобувача; під час вивчення дисципліни здобувачі повинні дотримуватись основних принципів та цінностей академічної доброчесності та етики академічних взаємовідносин;
- Ви зобов'язані з повагою та толерантністю ставитися до всіх членів академічної спільноти; не допускати поведінку, яка ставить під сумнів чесність та сумлінність Вашого навчання; складати всі завдання самостійно без допомоги сторонніх осіб; надавати для оцінювання лише результати власної роботи; не вдаватися до кроків, які можуть нечесно покращити Ваші результати чи погіршити/покращити результати інших здобувачів;
- дотримання академічної доброчесності здобувачами освіти передбачає: самостійне виконання навчальних завдань, завдань поточного та підсумкового контролю результатів навчання; недопущення академічного плагіату, фальсифікації, фабрикації й посилення на джерела інформації у разі використання ідей, розробок, тверджень, відомостей;
- порушення здобувачами освіти норм чинного законодавства про академічну доброчесність та етику взаємовідносин, може спричинити застосування заходів дисциплінарного характеру, і Ви будете нести академічну (повторне вивчення дисципліни, позбавлення академічної стипендії, відрахування з Університету та ін.) та/або інші види відповідальності.

При вивченні курсу політика дотримання академічної доброчесності визначається Кодексом академічної доброчесності НУ «Запорізька політехніка» https://zp.edu.ua/uploads/dept_nm/Nakaz_N253_vid_29.06.21.pdf

За погодженням можливе перезарахування освітньої компоненти у випадку участі студента в рамках міжнародної академічної мобільності (очно, онлайн або дистанційно) та вивчення курсу з відповідного напрямку обсягом не менше 3 кредитів ECTS.

Здобувачі, що бажають перезарахувати результатів неформальної (самостійної) освіти шляхом проходження онлайн курсів, можуть самостійно зареєструватися на курсах платформ Coursera/Udemy або інших онлайн-платформ, попередньо узгодивши тематику обраного курсу або курсів, повинні отримати відповідний сертифікат або сертифікати і показати його (їх) викладачу(в окремих випадках оформити звіт з проходження курсу у вигляді звіту з самостійної роботи). Кількість балів буде виставлена пропорційно до успіхів студента (досягнення на курсі згідно зі статистикою Coursera/Udemy або інших онлайн-платформ, сумарна мінімальна кількість годин курсу або курсів - 30 або 1 кредит ECTS).



Рекомендовані курси онлайн платформ:

1. Hadoop Starter Kit <http://www.hadoopinrealworld.com/starterkit/>
2. Hadoop Platform and Application Framework
<https://www.coursera.org/learn/hadoop>
3. Big Data Essentials: HDFS, MapReduce and Spark RDD
<https://www.coursera.org/learn/big-data-essentials>
4. Cloud Computing Applications, Part 2: Big Data and Applications in the Cloud
<https://www.coursera.org/learn/cloud-applications-part2>
5. CCA 175 Spark and Hadoop Developer Certification using Scala
<https://www.udemy.com/course/cca-175-spark-and-hadoop-developer-certification-using-scala>

Також можлива участь студентів в міжнародних школах та семінарах з тематики систем управління роботами і технологій студенти повинні отримати відповідний сертифікат та показати його викладачу (мінімальна кількість годин 30 або 1 ECTS).

Написання та публікація тез доповіді (одних з дисципліни) на науково-практичну конференцію викладачів, науковців, молодих учених, аспірантів та студентів «Тиждень науки» оцінюється в 10 додаткових балів. Написання та публікація тез доповіді (одних з дисципліни) на міжнародну конференцію або однієї спільної з викладачем наукової статі, що реферуються в SCOPUS, оцінюється в 20 додаткових балів.

Положення про порядок реалізації права на академічну мобільність учасників освітнього процесу НУ «Запорізька політехніка»
https://zp.edu.ua/uploads/dept_nm/Polozhennia_pro_akademichnu_mobilnist.pdf

ТЕХНІЧНІ ВИМОГИ ДЛЯ РОБОТИ НА КУРСІ

Служби підтримки:

- Система дистанційного навчання НУ «Запорізька політехніка» (Система Moodle) <https://moodle.zp.edu.ua/>;
- Електронний Інституційний репозиторій НУ "Запорізька політехніка" <http://eir.zp.edu.ua/>;
- Інформаційні електронні ресурси наукової бібліотеки НУ "Запорізька політехніка" <http://library.zp.edu.ua/>.

Щоб мати доступ до навчально-методичних розробок курсу необхідно мати особистий доступ до університетської навчальної платформи Moodle.

Курс розроблено на основі матеріалів, створених під час виконання міжнародного проекту Virtual Master Cooperation Data Science (ViMaCs) за підтримки Німецької служби академічних обмінів DAAD