

**Міністерство освіти і науки України
Державний університет інфраструктури та технологій**

**Інститут управління, технологій та права
Факультет управління і технологій
Кафедра «Інформаційні технології»**

«ЗАТВЕРДЖУЮ»

В.о. декана факультету

д.е.н., проф.  Світлана БОНЯР

30 серпня 2023 р.

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

**«Інноваційні технології Big Data та інформаційної безпеки в управлінні
транспортними системами»**

рівень вищої освіти другий (магістерський)

галузь знань 27 Транспорт

спеціальність 275.02 Транспортні технології (на залізничному транспорті)

освітньо-професійна програма Управління транспортними системами в умовах ризиків та криз



Co-funded by the
Erasmus+ Programme
of the European Union

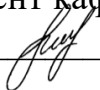
This project has been funded with support from the European Commission. This presentation reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained there in.

2023 – 2024 навчальний рік

Робоча програма дисципліни «Інноваційні технології Big Data та інформаційної безпеки в управлінні транспортними системами» для студентів рівня освіти «магістр» денної форми навчання галузі знань 27 «Транспорт» спеціальності 275 Транспортні технології (на залізничному транспорті), освітньо-професійна програма Транспортні технології (на залізничному транспорті). – Київ: ДУІТ, 2023. – 20 с.

Розробник:

доцент кафедри «Інформаційні технології»

 к.т.н., доц., Галан О.В.

Робочу програму схвалено на засіданні кафедри «Інформаційні технології»
Протокол від «29» серпня 2023 року № 1

Завідувач кафедри ІТ



(В.В. Завгородній)
(прізвище та ініціали)

© ДУІТ, 2023 рік

© Галан О.В., 2023 рік

1. Опис навчальної дисципліни

| Найменування показників | Галузь знань, напрям підготовки, освітньо-кваліфікаційний рівень | Характеристика навчальної дисципліни | |
|---|--|--------------------------------------|-----------------------|
| | | денна форма навчання | заочна форма навчання |
| Кількість кредитів – 4 | Галузь знань <u>27 Транспорт</u> (шифр і назва) | Обов'язкова | |
| | Спеціальність <u>275 Транспортні технології (залізничний транспорт)</u> (шифр і назва) | | |
| Модулів – 2 | Освітньо-професійна програма: <u>Транспортні технології (на залізничному транспорті)</u> (назва) | Рік підготовки | |
| Змістових модулів – 2 | | 1-й | |
| Індивідуальне науково-дослідне завдання: (назва) | | - | |
| Загальна кількість годин – 120 | | 1-й | |
| Тижневих годин для денної форми навчання: аудиторних – 2 самостійної роботи студента – 6 | Освітньо-кваліфікаційний рівень: <u>магістр</u> | Лекції | |
| | | 15 год. | 8 год. |
| Практичні, семінарські | | | |
| 15 год. | | 4 год. | |
| Лабораторні | | | |
| — | | | |
| Самостійна робота | | | |
| 90 год. | | 108 год. | |
| Індивідуальні завдання: | | | |
| - | | | |
| Тижневих годин для заочної форми навчання: аудиторних – 0,8 самостійної роботи студента – 7,2 | Вид контролю: | | |
| | іспит | | |

Примітка.

Співвідношення кількості годин аудиторних занять до самостійної і індивідуальної роботи становить відповідно (%):

для денної форми навчання – 25,

для заочної форми навчання - 10.

2. Мета та завдання навчальної дисципліни

Мета навчальної дисципліни – підготовка фахівців, компетентних у роботі із великими даними, їх методами і техніками аналізу, застосуванні технічних інструментів роботи з великими даними, використанні хмарних обчислень, проведенні темпорального аналізу даних, використанні технологій захисту інформації.

Основними **завданнями** вивчення дисципліни «Інноваційні технології Big Data та інформаційної безпеки в управлінні транспортними системами» є:

- ознайомити студентів з основними поняттями Big Data та IT, принципами роботи з великими даними, методами і техніками аналізу великих даних;

- надати студентам навички в застосуванні технічних інструментів роботи з Big Data технологіями, визначити тенденції розвитку та вдосконалення технологій та засобів роботи з Big Data;

- провести інтелектуальний аналіз даних, визначити область застосування машинного навчання Machine Learning;

- надати навички використання хмарних обчислень, надати розуміння інноваційних рішень технологій Map Reduce та Hadoop для роботи з Big Data;

- провести темпоральний аналіз даних, отримати навички аналізу збору даних та їх попередньої обробки;

- визначити особливості використання Python та його переваги для роботи з Big Data;

- визначити проблеми безпеки та інформаційного захисту у Big Data системах, провести аналіз технологій захисту інформації.

За підсумками вивчення дисципліни студент повинен знати:

основні положення Big Data та IT, принципами, методами і техніками роботи з великими даними та їх інформаційну базу,

інтелектуальний та темпоральний аналіз даних,

базові технології Map Reduce та Hadoop для роботи з Big Data;

машинне навчання та область його застосування;

хмарні обчислення та сферу їх використання в сучасному інформаційному середовищі;

Python та його переваги для роботи з Big Data;

проблеми безпеки та інформаційного захисту у Big Data.

За підсумками вивчення дисципліни студент повинен вміти:

аналізувати інформаційні джерела, вміти робити висновки у галузі інноваційних технологій;

представляти результати програмного продукту, використовуючи англійську термінологію;

застосовувати інноваційні технології BigData та методи інформаційної безпеки при вирішенні прикладних задач інших сфер діяльності;

надавати свої знання та отримані рішення в ясній та однозначній формі;

розробляти запити(програмні коди) у середовищі обробки BigData для контролю та моніторингу процесу перевезень;

обґрунтовувати доцільність застосування сучасних інноваційних технологій Big Data та інформаційної безпеки у транспортній галузі;

управляти моніторинговою системою BigData, застосовувати знання про інформаційну безпеку в управлінні транспортними системами;

використовувати сучасні комп'ютерні програмні засоби в управлінні транспортними системами;

проводити дослідницьку роботу у сфері інноваційних технологій.

На вивчення навчальної дисципліни відводиться 120 години, 4 кредитів ЄКТС.

3. Програма навчальної дисципліни

Змістовний модуль 1

Тема 1. Введення до інноваційних технологій у кризовому менеджменті. Основні поняття Big Data та IT. Принципи роботи з великими даними. Методи і техніки аналізу великих даних

Поняття «Big Data» та уявлення про світовий обсяг даних. Три "V" та принципи роботи з великими даними. Історія винайдення технології Big Data. Методи та техніки аналізу великих даних. Великі дані в промисловості та практичне застосування Big Data в експлуатації і технічному обслуговуванні залізничного транспорту. Завдання аналітики великих даних у залізничній галузі та потреба в технологіях великих даних. Організаційні ефекти великих даних та технічне обслуговування рухомого складу.

Тема 2. Технічні інструменти та тенденції роботи з Big Data технологіями

Сукупність підходів і технологій масово-паралельної обробки невизначено-структурованих даних, такі як NoSQL, алгоритми MapReduce і Hadoop. Інші рішення, що забезпечують подібні за характеристиками можливості по обробці надвеликих масивів даних, такі як NoSQL, Hive, Pig та Hbase, R, Python. А також деякі апаратно-програмні комплекси, що надають представлені рішення для обробки великих даних: Aster MapReduce appliance (корпорації Teradata), Oracle Big Data appliance, Greenplum appliance (корпорації EMC, на основі рішень поглиненої компанії Greenplum).

Тема 3. Збір даних. Форми попередньої обробки даних Data Mining

Збір даних (data extraction), обробка, і завантаження (ETL процес). Збір великих даних (Big Data Extraction), обробка великих даних (Big Data transformation). Форми попередньої обробки даних. Описова статистика та міри центральної тенденції. Нормальний розподіл, дисперсія і стандартне

відхилення, бімодальні розподіли, порівняння змінних і кореляція. Регресія. Завантаження великих даних (Big Data Loading).

Тема 4. Інтелектуальний аналіз даних. Машинне навчання Machine Learning

Інтелектуальний аналіз даних та методи класу Data Mining. Аналіз підходів і методів вирішення задачі. Властивості методів Data Mining. Підхід до навчання математичних моделей. Статистичні методи Data Mining. Кібернетичні методи Data Mining. Мета машинного навчання і сфери його застосування. Що потрібно для якісного машинного навчання. Основні види машинного навчання. Головна користь від машинного навчання.

Тема 5. Системи аналітичної обробки даних OLAP

Оперативна OLTP та аналітична OLAP обробка даних. Концепція сховища даних. Властивість інформаційних середовищ. Типи OLAP. Основні компоненти інформаційного сховища. Робота з гіперкубом. Реалізація сховищ і вітрини даних. Проблеми інтеграції даних. Наявні програмні реалізації.

Тема 6. Хмарні обчислення. Технологія Apache Spark для Big Data

Історія та піраміда хмарних технологій. Характеристики, послуги хмарних технологій. Хмарні моделі використання сервісів. Послуги, що надаються хмарними системами. Можливості та недоліки хмарних обчислень. Хмарні технології і віртуалізація. Концепція хмарного опрацювання даних. Знайомство з технологією Apache Spark для Big Data та Machine Learning.

Тема 7. Технологія Map Reduce для роботи з Big Data

Прийоми та стратегії розробки Map Reduce-додатків для роботи з Big Data. Стадія Map. Стадія Shuffle. Стадія Reduce. Ланцюги Map Reduce-задач.

Тема 8. Технологія Hadoop та її інноваційні рішення для роботи з Big Data

Для чого потрібен Hadoop та його основні принципи. Основні компоненти Hadoop. Встановлення Hadoop на кластер за допомогою дистрибутиву Cloudera Manager. Запуск MapReduce програм на Hadoop. Процес написання mapreduce-програми з використанням streaming-інтерфейсу та API Hadoop.

Змістовний модуль 2

Тема 9. Hbase як засіб збереження та обробки даних в екосистемі Hadoop

Цілі використання Hbase. Модель даних та операції що підтримуються Hbase. Архітектура. Способи роботи з Hbase. Деякі особливості роботи з Hbase. Альтернативи Hbase.

Тема 10. Темпоральні зміни даних

Темпоральні інформаційні системи та часовий домен. Вимоги до моделювання темпоральних залежностей. Часова повнота, щільність та ізоморфізм. Математичне визначення часу. Часові елементи. Структурне подання часу. Класифікація інформації по відношення до часу. Часовий куб даних. Об'єкти предметної області.

Тема 11. Особливості використання Python та його переваги для роботи з Big Data

Історія та особливості мови Python. Переваги Python для роботи з Big Data та так званий «Дзен Python». Програма на Python. Основні алгоритмічні конструкції.

Тема 12. Основи роботи у Python (частина 1)

Типи даних у програмуванні. Визначення змінної. Логічні вирази. Умовний оператор. Інструкція if. Множинне розгалуження. Цикл while. Введення даних з клавіатури.

Тема 13. Основи роботи у Python (частина 2)

Рядки як послідовності символів у Python. Списки як змінні послідовності. Введення до словників. Цикл for у мові програмування Python. Функції у програмуванні. Параметри і аргументи функцій. Локальні і глобальні змінні.

Тема 14. Технології ІТ захисту інформації

Основні поняття захисту інформації. Види комп'ютерних злочинів. Причини поширення комп'ютерної злочинності. Поняття і класифікація комп'ютерних вірусів. Засоби захисту інформації.

Тема 15. Проблеми захисту та безпеки BigData

Безпека даних. Реєстрація користувачів. Керування правами доступу та специфікація повноважень. Обов'язкові методи захисту BigData. Ведення журналів доступу. Обхід системи захисту. Захист при статистичній обробці даних. Шифрування даних.

(для денної форми навчання)
4. Структура навчальної дисципліни

| 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 |
| Модуль 1. | | | | | | |
| 1. Введення до інноваційних технологій у кризовому менеджменті. Основні поняття Big Data та IT. Принципи роботи з великими даними. Методи і техніки аналізу великих даних | 7 | 1 | | | | 6 |
| 2. Технічні інструменти та тенденції роботи з Big Data технологіями | 9 | 1 | 2 | | | 6 |
| 3. Збір даних. Форми попередньої обробки даних Data Mining | 7 | 1 | | | | 6 |
| 4. Інтелектуальний аналіз даних. Машинне навчання Machine Learning | 9 | 1 | 2 | | | 6 |
| 5. Системи аналітичної обробки даних OLAP | 7 | 1 | | | | 6 |
| 6. Хмарні обчислення. Технологія Apache Spark для Big Data | 9 | 1 | 2 | | | 6 |
| 7. Технологія Map Reduce для роботи з Big Data | 7 | 1 | | | | 6 |
| 8. Технологія Hadoop та її інноваційні рішення для роботи з Big Data | 9 | 1 | 2 | | | 6 |
| Разом за змістовим модулем 1 | 64 | 8 | 8 | - | - | 48 |
| Модуль 2 | | | | | | |
| 9. Hbase як засіб збереження та обробки даних в екосистемі Hadoop | 9 | 1 | 2 | | | 6 |
| 10. Темпоральні зміни даних | 7 | 1 | | | | 6 |
| Особливості використання Python та його переваги для роботи з Big Data | 7 | 1 | | | | 6 |
| 11. Основи роботи у Python (частина 1) | 10 | 1 | 3 | | | 6 |
| 12. Основи роботи у Python (частина 2) | 9 | 1 | 2 | | | 6 |
| 13. Технології IT захисту інформації | 7 | 1 | | | | 6 |
| 14. Проблеми захисту та безпеки BigData | 7 | 1 | | | | 6 |
| 15. Разом за змістовим модулем 2 | 56 | 7 | 7 | - | - | 42 |
| Разом за семестр | 120 | 15 | 15 | - | - | 90 |
| | | | | | | |

5. Темі семінарських занять

Не передбачено.

6. Теми практичних занять

| № з/п | Назва теми | Кількість годин |
|--------------------------|--|-----------------|
| 1 | Типи даних у програмуванні. Визначення змінної. | 2 |
| 2 | Логічні вирази. Умовний оператор. Інструкція if. | 2 |
| 3 | Множинне розгалуження. Цикл while. | 2 |
| 4 | Введення даних з клавіатури. Рядки як послідовності символів у Python. | 2 |
| 5 | Списки як змінні послідовності. Введення до словників. | 2 |
| 6 | Цикл for у мові програмування Python. Функції у програмуванні. | 3 |
| 7 | Параметри і аргументи функцій. Локальні і глобальні змінні. | 2 |
| Всього за семестр | | 15 |

7. Теми лабораторних занять

Не передбачено.

8. Самостійна робота

| № з/п | Назва теми | Кількість годин |
|-------|--|-----------------|
| 1 | Введення до інноваційних технологій у кризовому менеджменті. Основні поняття Big Data та IT. Принципи роботи з великими даними. Методи і техніки аналізу великих даних | 6 |
| 2 | Технічні інструменти та тенденції роботи з Big Data технологіями | 6 |
| 3 | Збір даних. Форми попередньої обробки даних Data Mining | 6 |
| 4 | Інтелектуальний аналіз даних. Машинне навчання Machine Learning | 6 |
| 5 | Системи аналітичної обробки даних OLAP | 6 |
| 6 | Хмарні обчислення. Технологія Apache Spark для Big Data | 6 |
| 7 | Технологія Map Reduce для роботи з Big Data | 6 |
| 8 | Технологія Hadoop та її інноваційні рішення для роботи з Big Data | 6 |
| 9 | Hbase як засіб збереження та обробки даних в екосистемі Hadoop | 6 |
| 10 | Темпоральні зміни даних | 6 |
| 11 | Особливості використання Python та його переваги для роботи з Big Data | 6 |
| 12 | Основи роботи у Python (частина 1) | 6 |

| | | |
|---------------|-------------------------------------|-----------|
| 13 | Основи роботи у Python (частина 2) | 6 |
| 14 | Технології ІТ захисту інформації | 6 |
| 15 | Проблеми захисту та безпеки BigData | 6 |
| Разом: | | 90 |

9. Індивідуальні завдання

Не передбачено

10. Методи навчання

Лекції із застосуванням мультимедійної техніки; практичні заняття; розв'язування тестових завдань, робота з довідковою літературою, складання програмних скриптів, а також самостійна робота, що включає роботу з конспектом, вирішення завдань, роботу з учбовою спеціальною літературою та індивідуальні завдання.

11. Методи контролю

Контроль знань студентів здійснюється згідно п. 12.

Більшу кількість рейтингових балів студент отримує завдяки ритмічній аудиторній та самостійній роботі протягом семестру, тоді як на заходи проміжного та підсумкового контролю припадає не більше 20 – 30% балів.

Курс дисципліни «Інноваційні технології Big Data та інформаційної безпеки в управлінні транспортними системами» поділено на два модулі, що складаються впродовж семестру.

Кожний модуль складається із двох частин: теоретичної (лекційний матеріал), практичної (протягом семестру) робіт, а також постійної самостійної роботи студента. Кожний розділ є обов'язковим для опанування і оцінюється відповідною кількістю балів. Балами оцінюється також самостійна робота студентів.

Оцінка по модулю визначається кількістю балів, які отримав студент за теоретичну, практичну частини та самостійну роботу.

Критерії оцінювання модульного завдання

Модульний контроль проводиться у формі тестування. Всього в тесті 24 питання, за кожен вірну відповідь студент отримує 0,5 балів, за невірну відповідь студент отримує 0 балів. Максимальна кількість балів, яку може отримати студент за модульну контрольну роботу, становить 12 балів.

Критерії оцінювання екзаменаційного завдання

| | |
|----------|-----------------|
| Завдання | Кількість балів |
|----------|-----------------|

| | |
|--------------|-----------|
| Питання 1 | 4 |
| Питання 2 | 4 |
| Питання 3 | 5 |
| Разом | 13 |

12. Розподіл балів, які отримують студенти

| Від заняття або контрольного заходу (наприклад) | Балів за одне заняття або контрольний захід | За семестр | | До 1-й атестації | |
|---|--|--|------------|--|------------|
| | | кількість занять або контрольних заходів | сума балів | кількість занять або контрольних заходів | сума балів |
| Практичні, в тому числі: | | | | | |
| - виконання | 8 | 7 | 70 | 4 | 40 |
| - здача звіту | 2 | | | | |
| Модульна контрольна робота | 10 | 2 | 20 | 1 | 10 |
| Іспит | | | 10 | | |
| Ітого | | | 100 | | 50 |
| Наукова робота | Додатково за участь у наукових конференціях, семінарах, круглих столах, студентських олімпіадах та конкурсах за тематикою дисципліни, що вивчається, студент може отримати до 15 балів | | | | |

| Поточне тестування та самостійна робота | | | | | | | | | | | | | | | Підсумковий тест (екзамен) | Сума |
|---|----|----|----|----|----|----|----|--------------------|-----|-----|-----|-----|-----|-----|----------------------------|------|
| Змістовий модуль 1 | | | | | | | | Змістовий модуль 2 | | | | | | | | |
| T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 | T11 | T12 | T13 | T14 | T15 | 13 | 100 |
| 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 5 | 5 | 6 | 6 | | |

Шкала оцінювання: національна та ЄКТС

| Сума балів за всі види навчальної діяльності | Оцінка ЄКТС | Оцінка за національною шкалою | |
|--|-------------|--|---|
| | | для екзамену, курсової роботи, практики | для заліку |
| 90 – 100 | A | відмінно | зараховано |
| 82-89 | B | добре | |
| 74-81 | C | | |
| 64-73 | D | задовільно | |
| 60-63 | E | | |
| 35-59 | FX | незадовільно з можливістю повторного складання | не зараховано з можливістю повторного складання |
| 0-34 | F | незадовільно з обов'язковим повторним вивченням дисципліни | не зараховано з обов'язковим повторним вивченням дисципліни |

13. Методичне забезпечення

1. Опорний конспект лекцій – у вигляді презентацій (PowerPoint) надсилається студентам в електронному вигляді.
2. Інтерактивний комплекс навчально-методичного забезпечення дисципліни включає: програма навчальної дисципліни; тексти лекцій; методичні вказівки до практичних занять; модулі перевірки знань; основна та додаткова рекомендована література; тести для контролю; перелік літератури до курсу, підготовлений в електронному форматі з посиланнями до Інтернет-ресурсів; нормативні документи; ілюстративні дидактичні матеріали.

14. Рекомендована література

Основна

1. Faeze Ghofrani, QingHe, Rob M.P.Goverde, Xiang Liu. Recent applications of big data analytics in railway transportation systems: A survey.
2. Efreim Turban, Ramesh Sharda Business Intelligence.
3. Efreim Turban, Jay E. Aronson Decision Support and Business Intelligence Systems, 8/E.
4. Al Powell, Harvard Staff Writer (2017). Bringing big data to the farm. Available at: <http://news.harvard.edu/gazette/story/2017/04/the-future-of-food-will-be-proactive-efficientand-digitized-or-else/>.
5. Alyson Y. (2017) Our partnership with Imperial College London: The Digital Asset Research Lab. Available at: <https://blog.blockchain.com/2017/01/19/our-partnership-with-imperial-college-london-the-digital-asset-research-lab/>.
6. Brighten your day with cognitive UEM (2017) Available at: <http://www-03.ibm.com/security/mobile/maas360.html>.
7. Polonski V. (2014) The evolution of social networking sites: The rise of content-centric platforms which favour the perpetual present. Available at: <http://www.kellogg.ox.ac.uk/blog/the-evolution-of-social-networking>.

Допоміжна

8. Data Science: an introduction. Wikibooks [Electronic resource]. – 2017. – Mode of access: https://en.wikibooks.org/wiki/Data_Science: An Introduction.
9. Forecasting: principles and practice [Electronic resource]. – 2017. – Mode of access: <https://www.otexts.org/fpp>.
10. Gabriel Martos. Cluster Analysis with R [Electronic resource] / R Pubs. – 2015. – Mode of access: <https://rpubs.com/gabrielmartos/ClusterAnalysis>.
11. Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An Introduction to Statistical Learning with Applications in R [Electronic resource]. – 2017. – Mode of access: <http://www-bcf.usc.edu/~gareth/ISL/>.
12. Grolemond, G. R for Data Science [Electronic resource] / Garrett Grolemond, Hadley Wickham. – 2016. – Mode of access: <http://r4ds.had.co.nz/index.html>. – Date of access: 01.09.2022.
12. IData Science: an introduction to. Jeffrey Stanton, Syracuse

University [Electronic resource]. – 2012. – Mode of access: https://ischool.syr.edu/media/documents/2012/3/DataScienceBook1_1.pdf.

13. Introducing Data Science: Big data, machine learning, and more, using Python tools. Davy Cielen, Arno D. B. Meysman, and Mohamed Ali [Electronic resource]. – May 2016. ISBN 9781633430037. 320 pages. – Mode of access: <https://www.manning.com/books/introducing-data-science>.

14. Intro-to-h2o.R [Electronic resource] / h2o-tutorials/h2o-open-tour-2016/chicago. – 2016. – Mode of access: <https://github.com/h2oai/h2o-tutorials/blob/master/h2oopen-tour-2016/chicago/intro-to-h2o.R>.

15. Leisch, F. Creating R Packages: a Tutorial [Electronic resource]: Department of Statistics, Ludwig-Maximilians-Universitat Munchen, and R Development Core Team, 2009 / Friedrich Leisch. – Mode of access: <ftp://cran.rproject.org/pub/R/doc/contrib/Leisch-CreatingPackages.pdf>.

16. R Markdown Reference Guide [Electronic resource] / RStudio Cheat Sheets. – 2014. – Mode of access: <https://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf>.

17. Simko, V. From R code to R package [Electronic resource]. – 2015. – Mode of access: <https://github.com/vsimko/course-rpackages>.

18. Wickham, H. R packages [Electronic resource]. – 2015. – Mode of access: <http://rpkgs.had.co.nz/package.html>.

19. Wickham, H. Tidy data [Electronic resource] / Hadley Wickham // The Journal of Statistical Software. – Vol. 59. – 2014. – Mode of access: <http://vita.had.co.nz/papers/tidy-data.html>.

20. Wickham, H. Tidy data [Electronic resource]. – 2017. – Mode of access: <ftp://cran.r-project.org/pub/R/web/packages/tidyr/vignettes/tidy-data.html>.

21. Wickham, H. Vignettes: long-form documentation [Electronic resource] / R packages. – 2015. – Mode of access: <http://r-pkgs.had.co.nz/vignettes.html>.

22. Big data in railways. European Union Agency for railways [Електронний ресурс] – Режим доступу: https://www.era.europa.eu/sites/default/files/activities/docs/cor_big_data_en.pdf

23. Improving rolling stock maintenance using big data [Електронний ресурс] – Режим доступу: <https://www.globalrailwayreview.com/article/28433/big-data-trenitalia-rolling-stock-maintenance/>

24. Big data in railway operations and maintenance [Електронний ресурс] – Режим доступу: <https://www.globalrailwayreview.com/article/61515/big-data-railway-operations-maintenance-2/>.

25. Big Data Solutions: Example of The Development Cost [Електронний ресурс] – Режим доступу: <https://medium.com/existek/big-data-solutions-example-of-the-development-cost-e3d173d97064>.

26. Can Machine Learning improve railway operational performance [Електронний ресурс] – Режим доступу: <https://www.rssb.co.uk/Insights-and-News/Blogs/Can-Machine-Learning-improve-railway-operational-performance>.

27. Великі дані [Електронний ресурс] – Режим доступу: https://en.wikipedia.org/wiki/Big_data.

(для заочної форми навчання)

4. Структура навчальної дисципліни

| 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 |
| Модуль 1. | | | | | | |
| 16. Введення до інноваційних технологій у кризовому менеджменті. Основні поняття Big Data та IT. Принципи роботи з великими даними. Методи і техніки аналізу великих даних | 7 | 1 | | | | 6 |
| 17. Технічні інструменти та тенденції роботи з Big Data технологіями | 9 | 1 | | | | 8 |
| 18. Збір даних. Форми попередньої обробки даних Data Mining | 7 | | | | | 7 |
| 19. Інтелектуальний аналіз даних. Машинне навчання Machine Learning | 9 | 1 | | | | 8 |
| 20. Системи аналітичної обробки даних OLAP | 7 | | | | | 7 |
| 21. Хмарні обчислення. Технологія Apache Spark для Big Data | 9 | 1 | | | | 8 |
| 22. Технологія Map Reduce для роботи з Big Data | 7 | | 1 | | | 6 |
| 23. Технологія Hadoop та її інноваційні рішення для роботи з Big Data | 9 | | 1 | | | 8 |
| Разом за змістовим модулем 1 | 64 | 4 | 2 | - | - | 58 |
| Модуль 2 | | | | | | |
| 24. Hbase як засіб збереження та обробки даних в екосистемі Hadoop | 9 | 1 | | | | 8 |
| 25. Темпоральні зміни даних | 7 | 1 | | | | 6 |
| Особливості використання Python та його переваги для роботи з Big Data | 7 | | | | | 7 |
| 26. Основи роботи у Python (частина 1) | 10 | | 1 | | | 9 |
| 27. Основи роботи у Python (частина 2) | 9 | | 1 | | | 8 |
| 28. Технології IT захисту інформації | 7 | 1 | | | | 6 |
| 29. Проблеми захисту та безпеки BigData | 7 | 1 | | | | 6 |
| 30. Разом за змістовим модулем 2 | 56 | 4 | 2 | - | - | 50 |
| Разом за семестр | 120 | 8 | 4 | - | - | 108 |
| | | | | | | |

5. Темі семінарських занять

Не передбачено.

6. Теми практичних занять

| № з/п | Назва теми | Кількість годин |
|--------------------------|--|-----------------|
| 1 | Типи даних у програмуванні. Визначення змінної. | 0,5 |
| 2 | Логічні вирази. Умовний оператор. Інструкція if. | 0,5 |
| 3 | Множинне розгалуження. Цикл while. | 0,5 |
| 4 | Введення даних з клавіатури. Рядки як послідовності символів у Python. | 0,5 |
| 5 | Списки як змінні послідовності. Введення до словників. | 0,5 |
| 6 | Цикл for у мові програмування Python. Функції у програмуванні. | 0,5 |
| 7 | Параметри і аргументи функцій. Локальні і глобальні змінні. | 1 |
| Всього за семестр | | 4 |

7. Теми лабораторних занять

Не передбачено.

9. Самостійна робота

| № з/п | Назва теми | Кількість годин |
|-------|--|-----------------|
| 1 | Введення до інноваційних технологій у кризовому менеджменті. Основні поняття Big Data та IT. Принципи роботи з великими даними. Методи і техніки аналізу великих даних | 6 |
| 2 | Технічні інструменти та тенденції роботи з Big Data технологіями | 8 |
| 3 | Збір даних. Форми попередньої обробки даних Data Mining | 7 |
| 4 | Інтелектуальний аналіз даних. Машинне навчання Machine Learning | 8 |
| 5 | Системи аналітичної обробки даних OLAP | 6 |
| 6 | Хмарні обчислення. Технологія Apache Spark для Big Data | 8 |
| 7 | Технологія Map Reduce для роботи з Big Data | 7 |
| 8 | Технологія Hadoop та її інноваційні рішення для роботи з Big Data | 8 |
| 9 | Hbase як засіб збереження та обробки даних в екосистемі Hadoop | 8 |
| 10 | Темпоральні зміни даних | 6 |
| 11 | Особливості використання Python та його переваги для роботи з Big Data | 6 |
| 12 | Основи роботи у Python (частина 1) | 9 |

| | | |
|---------------|-------------------------------------|------------|
| 13 | Основи роботи у Python (частина 2) | 9 |
| 14 | Технології ІТ захисту інформації | 6 |
| 15 | Проблеми захисту та безпеки BigData | 6 |
| Разом: | | 108 |

9. Індивідуальні завдання

Не передбачено

10. Методи навчання

Лекції із застосуванням мультимедійної техніки; практичні заняття; розв'язування тестових завдань, робота з довідковою літературою, складання програмних скриптів, а також самостійна робота, що включає роботу з конспектом, вирішення завдань, роботу з учбовою спеціальною літературою та індивідуальні завдання.

11. Методи контролю

Контроль знань студентів здійснюється згідно п. 12.

Більшу кількість рейтингових балів студент отримує завдяки ритмічній аудиторній та самостійній роботі протягом семестру, тоді як на заходи проміжного та підсумкового контролю припадає не більше 20 – 30% балів.

Курс дисципліни «Інноваційні технології Big Data та інформаційної безпеки в управлінні транспортними системами» поділено на два модулі, що складаються впродовж семестру.

Кожний модуль складається із двох частин: теоретичної (лекційний матеріал), практичної (протягом семестру) робіт, а також постійної самостійної роботи студента. Кожний розділ є обов'язковим для опанування і оцінюється відповідною кількістю балів. Балами оцінюється також самостійна робота студентів.

Оцінка по модулю визначається кількістю балів, які отримав студент за теоретичну, практичну частини та самостійну роботу.

Критерії оцінювання модульного завдання

Модульний контроль проводиться у формі тестування. Всього в тесті 24 питання, за кожен вірну відповідь студент отримує 0,5 балів, за невірну відповідь студент отримує 0 балів. Максимальна кількість балів, яку може отримати студент за модульну контрольну роботу, становить 12 балів.

Критерії оцінювання екзаменаційного завдання

| Завдання | Кількість балів |
|--------------|-----------------|
| Питання 1 | 4 |
| Питання 2 | 4 |
| Питання 3 | 5 |
| Разом | 13 |

12. Розподіл балів, які отримують студенти

| Від заняття або контрольного заходу (наприклад) | Балів за одне заняття або контрольний захід | За семестр | | До 1-й атестації | |
|---|--|--|------------|--|------------|
| | | кількість занять або контрольних заходів | сума балів | кількість занять або контрольних заходів | сума балів |
| Практичні, в тому числі: | | | | | |
| - виконання | 8 | 7 | 70 | 4 | 40 |
| - здача звіту | 2 | | | | |
| Модульна контрольна робота | 10 | 2 | 20 | 1 | 10 |
| Іспит | | | 10 | | |
| Ітого | | | 100 | | 50 |
| Наукова робота | Додатково за участь у наукових конференціях, семінарах, круглих столах, студентських олімпіадах та конкурсах за тематикою дисципліни, що вивчається, студент може отримати до 15 балів | | | | |

| Поточне тестування та самостійна робота | | | | | | | | | | | | | | | Підсумковий тест (екзамен) | Сума |
|---|----|----|----|----|----|----|----|--------------------|-----|-----|-----|-----|-----|-----|----------------------------|------|
| Змістовий модуль 1 | | | | | | | | Змістовий модуль 2 | | | | | | | | |
| T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 | T10 | T11 | T12 | T13 | T14 | T15 | 13 | 100 |
| 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 5 | 5 | 6 | 6 | | |

Шкала оцінювання: національна та ЄКТС

| Сума балів за всі види навчальної діяльності | Оцінка ECTS | Оцінка за національною шкалою | |
|--|-------------|--|---|
| | | для екзамену, курсової роботи, практики | для заліку |
| 90 – 100 | A | відмінно | зараховано |
| 82-89 | B | добре | |
| 74-81 | C | | |
| 64-73 | D | задовільно | |
| 60-63 | E | | |
| 35-59 | FX | незадовільно з можливістю повторного складання | не зараховано з можливістю повторного складання |

| | | | |
|------|---|--|---|
| 0-34 | F | незадовільно з обов'язковим повторним вивченням дисципліни | не зараховано з обов'язковим повторним вивченням дисципліни |
|------|---|--|---|

13. Методичне забезпечення

1. Опорний конспект лекцій – у вигляді презентацій (PowerPoint) надсилається студентам в електронному вигляді.

2. Інтерактивний комплекс навчально-методичного забезпечення дисципліни включає: програма навчальної дисципліни; тексти лекцій; методичні вказівки до практичних занять; модулі перевірки знань; основна та додаткова рекомендована література; тести для контролю; перелік літератури до курсу, підготовлений в електронному форматі з посиланнями до Інтернет-ресурсів; нормативні документи; ілюстративні дидактичні матеріали.

14. Рекомендована література

Основна

1. Faeze Ghofrani, QingHe, Rob M.P.Goverde, Xiang Liu. Recent applications of big data analytics in railway transportation systems: A survey.
2. Turban, Ramesh Sharda Business Intelligence.
3. Efreim Turban, Jay E. Aronson Decision Support and Business Intelligence Systems, 8/E.
4. Al Powell, Harvard Staff Writer (2017). Bringing big data to the farm. Available at: <http://news.harvard.edu/gazette/story/2017/04/the-future-of-food-will-be-proactive-efficientand-digitized-or-else/>.
5. Alyson Y. (2017) Our partnership with Imperial College London: The Digital Asset Research Lab. Available at: <https://blog.blockchain.com/2017/01/19/our-partnership-withimperial-college-london-the-digital-asset-research-lab/>.
6. Brighten your day with cognitive UEM (2017) Available at: <http://www-03.ibm.com/security/mobile/maas360.html>.
7. Polonski V. (2014) The evolution of social networking sites: The rise of content-centric platforms which favour the perpetual present. Available at: <http://www.kellogg.ox.ac.uk/blog/the-evolution-of-social-networking>.

Допоміжна

8. Data Science: an introduction. Wikibooks [Electronic resource]. – 2017. – Mode of access: https://en.wikibooks.org/wiki/Data_Science:An_Introduction.
9. Forecasting: principles and practice [Electronic resource]. – 2017. – Mode of access: <https://www.otexts.org/fpp>.
10. Gabriel Martos. Cluster Analysis with R [Electronic resource] / R Pubs. – 2015. – Mode of access: <https://rpubs.com/gabrielmartos/ClusterAnalysis>.
11. Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An Introduction to Statistical Learning with Applications in R [Electronic resource]. – 2017. – Mode of access: <http://www-bcf.usc.edu/~gareth/ISL/>.

12. Grolemund, G. R for Data Science [Electronic resource] / Garrett Grolemund, Hadley Wickham. – 2016. – Mode of access: <http://r4ds.had.co.nz/index.html>. – Date of access: 01.09.2022. 12. IData Science: an introduction to. Jeffrey Stanton, Syracuse University [Electronic resource]. – 2012. – Mode of access: https://ischool.syr.edu/media/documents/2012/3/DataScienceBook1_1.pdf.

13. Introducing Data Science: Big data, machine learning, and more, using Python tools. Davy Cielen, Arno D. B. Meysman, and Mohamed Ali [Electronic resource]. – May 2016. ISBN 9781633430037. 320 pages. – Mode of access: <https://www.manning.com/books/introducing-data-science>.

14. Intro-to-h2o.R [Electronic resource] / h2o-tutorials/h2o-open-tour-2016/chicago. – 2016. – Mode of access: <https://github.com/h2oai/h2o-tutorials/blob/master/h2oopen-tour-2016/chicago/intro-to-h2o.R>.

15. Leisch, F. Creating R Packages: a Tutorial [Electronic resource]: Department of Statistics, Ludwig-Maximilians-Universitat Munchen, and R Development Core Team, 2009 / Friedrich Leisch. – Mode of access: <ftp://cran.rproject.org/pub/R/doc/contrib/Leisch-CreatingPackages.pdf>.

16. R Markdown Reference Guide [Electronic resource] / RStudio Cheat Sheets. – 2014. – Mode of access: <https://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf>.

17. Simko, V. From R code to R package [Electronic resource]. – 2015. – Mode of access: <https://github.com/vsimko/course-rpackages>.

18. Wickham, H. R packages [Electronic resource]. – 2015. – Mode of access: <http://rpkgs.had.co.nz/package.html>.

19. Wickham, H. Tidy data [Electronic resource] / Hadley Wickham // The Journal of Statistical Software. – Vol. 59. – 2014. – Mode of access: <http://vita.had.co.nz/papers/tidy-data.html>.

20. Wickham, H. Tidy data [Electronic resource]. – 2017. – Mode of access: <ftp://cran.r-project.org/pub/R/web/packages/tidyr/vignettes/tidy-data.html>.

21. Wickham, H. Vignettes: long-form documentation [Electronic resource] / R packages. – 2015. – Mode of access: <http://r-pkgs.had.co.nz/vignettes.html>.

22. Big data in railways. European Union Agency for railways [Электронный ресурс] – Режим доступа: https://www.era.europa.eu/sites/default/files/activities/docs/cor_big_data_en.pdf

23. Improving rolling stock maintenance using big data [Электронный ресурс] – Режим доступа: <https://www.globalrailwayreview.com/article/28433/big-data-trenitalia-rolling-stock-maintenance/>

24. Big data in railway operations and maintenance [Электронный ресурс] – Режим доступа: <https://www.globalrailwayreview.com/article/61515/big-data-railway-operations-maintenance-2/>.

25. Big Data Solutions: Example of The Development Cost [Электронный ресурс] – Режим доступа: <https://medium.com/existek/big-data-solutions-example-of-the-development-cost-e3d173d97064>.

26. Can Machine Learning improve railway operational performance [Електронний ресурс] – Режим доступу: <https://www.rssb.co.uk/Insights-and-News/Blogs/Can-Machine-Learning-improve-railway-operational-performance>.

27. Великі дані [Електронний ресурс] – Режим доступу: https://en.wikipedia.org/wiki/Big_data.