

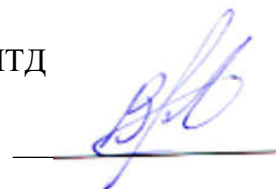
ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФРАСТРУКТУРИ ТА ТЕХНОЛОГІЙ

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

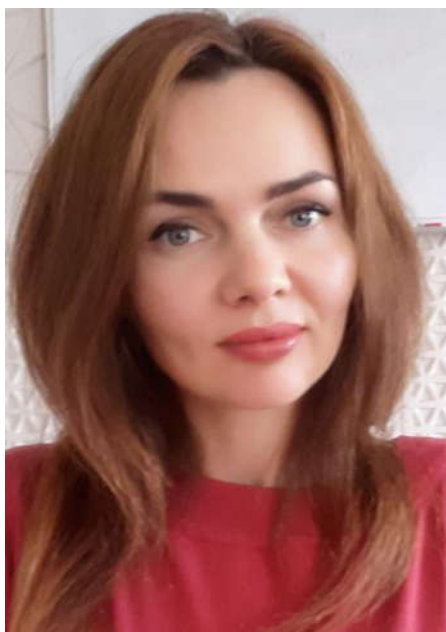
Затверджую

Завідувач кафедри ІТД

В.В. Завгородній



Протокол № 15 від 16 червня 2021 р.



ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Викладач	Кандидат технічних наук, доцент ГАЛАН Ольга Вікторівна
E-mail	Helga.halan@gmail.com
Навчальна дисципліна	Інноваційні технології Big Data та інформаційної безпеки в управлінні транспортними системами
Офіційна назва освітньої програми	Управління транспортними системами в умовах ризиків та криз
Рівень вищої освіти	другий (магістерський)
Галузь знань	27«Транспорт»
Спеціальність	275 «Транспортні технології (на залізничному транспорті)»
Обсяг дисципліни в кредитах ECTS	4
Статус дисципліни (обов'язкова, вибіркова)	Цикл дисциплін професійної підготовки, обов'язкова

Мета вивчення дисципліни	Метою навчальної дисципліни є 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.
Інтегральна компетентність	Здатність особи розв'язувати складні задачі і проблеми транспортної галузі у сфері транспортних систем і технологій та у процесі навчання, що передбачає проведення досліджень і здійснення інновацій та характеризується невизначеністю умов і вимог
Загальні компетентності	ЗК 01. Здатність працювати в міжнародному середовищі. ЗК 03. Здатність до пошуку та аналізу інформації з різних джерел.
Спеціальні (фахові) компетентності	ФК 07. Здатність до управління транспортними потоками. ФК 11. Здатність використовувати сучасні комп'ютерні програмні продукти у сфері транспортних систем та технологій.

ЗМІСТ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ ЗА ТЕМАМИ

Модуль 1

Основні принципи та методи роботи з великими даними

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.

План теми:

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

Визначаються об'єкти і завдання курсу. Розглядаються Поняття «Big Data» та уявлення про світовий обсяг даних. Приводиться основна термінологія курсу, історія винайдення технології. Проводиться аналіз методів та технік аналізу великих даних. Проводиться аналіз великих даних в промисловості та практичне застосування Big Data в експлуатації і технічному обслуговуванні залізничного транспорту. Визначаються завдання аналітики великих даних у залізничній галузі та потреба в технологіях великих даних та організаційні ефекти великих даних та технічне обслуговування рухомого складу.

Література: 1-7, 26-28.

Ключові слова: завдання курсу, Big Data, принципи роботи з великими даними, Методи і техніки аналізу великих даних.

Тема 2 Технічні інструменти та тенденції роботи з Big Data технологіями

План теми:

1. Методи аналізу великих даних
2. NoSQL
3. Hadoop та MapReduce
4. Hive, Pig та Hbase
5. Мова R
6. Python та його переваги при роботі з Big Data
7. Апаратні рішення

Розглядається 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).

Література: 8-12, 31-33.

Ключові слова: NoSQL, Hive, Pig, Hbase, R, Python, MapReduce і Hadoop.

Тема 3 Збір даних. Форми попередньої обробки даних Data Mining

План теми:

1. Збір даних (data extraction).
2. Збір даних, обробка, і завантаження (ETL процес).
3. Збір великих даних (Big Data Extraction).
4. Обробка великих даних (Big Data transformation).
5. Форми попередньої обробки даних.
6. Описова статистика та міри центральної тенденції.
7. Нормальний розподіл.
8. Дисперсія і стандартне відхилення.
9. Бімодальні розподіли.
10. Порівняння змінних і кореляція.
11. Регресія.
12. Завантаження великих даних (Big Data Loading).

Розглядається збір даних (data extraction), обробка, і завантаження (ETL процес), збір великих даних (Big Data Extraction), обробка великих даних (Big Data transformation). Визначаються форми попередньої обробки даних. Приводиться описова статистика та міри центральної тенденції, нормальний розподіл, дисперсія і стандартне відхилення, бімодальні розподіли, порівняння змінних і кореляція, регресія. Розглядається процес завантаження великих даних (Big Data Loading).

Література: 11, 12, 13, 29, 33.

Ключові слова: 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.

Тема 4 Інтелектуальний аналіз даних. Машинне навчання Machine Learning

План теми:

1. Методи класу Data Mining. Інтелектуальний аналіз даних.
2. Аналіз підходів і методів вирішення задачі.
3. Властивості методів Data Mining.
4. Підхід до навчання математичних моделей.
5. Статистичні методи Data Mining.

6. Кібернетичні методи Data Mining.
7. Мета машинного навчання і сфери його застосування.
8. Що потрібно для якісного машинного навчання?
9. Основні види машинного навчання.
10. Головна користь від машинного навчання.

Розглядається інтелектуальний аналіз даних та методи класу Data Mining, аналіз підходів і методів вирішення задачі. Визначаються властивості методів Data Mining та підхід до навчання математичних моделей. Наводяться статистичні методи Data Mining та кібернетичні методи Data Mining. Визначається мета машинного навчання і сфери його застосування, основні види машинного навчання та користь від його застосування.

Література: 2, 3, 11, 12, 13, 15, 34.

Ключові слова: інтелектуальний аналіз даних, методи класу Data Mining, Machine Learning.

Тема 5 Системи аналітичної обробки даних OLAP

План теми:

1. Оперативна OLTP та аналітична OLAP обробка даних
2. Концепція сховища даних
3. Властивості інформаційних сховищ
4. Типи OLAP
5. Основні компоненти інформаційного сховища
6. Робота з гіперкубом
7. Реалізація сховищ і вітрин даних
8. Проблеми інтеграції даних
9. Наявні програмні реалізації

Розглядається оперативна OLTP та аналітична OLAP обробка даних. Визначаються концепція сховища даних, властивість інформаційних середовищ та типи OLAP. Наводяться основні компоненти інформаційного сховища. Приводиться робота з гіперкубом та реалізація сховищ і вітрини даних. Визначаються проблеми інтеграції даних та наводяться наявні програмні реалізації.

Література: 16-19, 23, 25.

Ключові слова: OLTP, OLAP обробка даних, концепція сховища даних, гіперкуб, вітрина даних.

Тема 6 Хмарні обчислення. Технологія Apache Spark для Big Data

План теми:

1. Історія хмарних технологій
2. Піраміда хмарних технологій.
3. Характеристики, послуги хмарних технологій.
4. Хмарні моделі використання сервісів.
5. Послуги, що надаються хмарними системами.
6. Можливості та недоліки хмарних обчислень.
7. Хмарні технології і віртуалізація.
8. Концепція хмарного опрацювання даних.

Наводяться історія та піраміда хмарних технологій, характеристики, послуги хмарних технологій, хмарні моделі використання сервісів. Визначаються послуги, що надаються хмарними системами. Приводяться можливості та недоліки хмарних обчислень, хмарні технології і віртуалізація. Знайомство з технологією Apache Spark для Big Data та Machine Learning.

Література: 3, 5, 7, 31, 32, 39.

Ключові слова: хмарні технології, хмарні сервіси, Apache Spark.

Модуль 2

Технології, алгоритми та мови роботи з великими даними. Безпека великих даних

Тема 7 Технологія Map Reduce для роботи з Big Data

План теми:

1. Стадія Map. Стадія Shuffle. Стадія Reduce.
2. Прийоми і стратегії розробки MapReduce-додатків. Map only job. Combine.
3. Ланцюжки MapReduce-задач
4. Distributed cach
5. Reduce Join
6. MapJoin

Розглядаються прийоми та стратегії розробки Map Reduce-додатків для роботи з Big Data. Вивчаються стадія Map, стадія Shuffle, стадія Reduce та ланцюги Map Reduce-задач.

Література: 8, 9, 19, 31.

Ключові слова: стадіяMap, стадія Shuffle, стадія Reduce, Map Reduce-задачі.

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

План теми:

1. Навіщо потрібен Hadoop?
2. Основні принципи Hadoop.
3. Основні компоненти Hadoop.
4. Встановлення Hadoop на кластер за допомогою дистрибутива Cloudera Manager.
5. Запуск MapReduce програм на Hadoop.
6. Процес написання mapreduce-програми з використанням streaming-інтерфейса і API Hadoop.

Розглядається для чого потрібен Hadoop та його основні принципи. Визначаються основні компоненти Hadoop. Розглядається встановлення Hadoop на кластер за допомогою дистрибутиву Cloudera Manager. Розглядається процес написання mapreduce-програми з використанням streaming-інтерфейсу та API Hadoop.

Література: 3, 11, 21, 41.

Ключові слова: Hadoop, основні компоненти Hadoop.

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

План теми:

1. Для чого потрібен Hbase
2. Модель даних
3. Операції, що підтримуються
4. Архітектура
5. Способи роботи з Hbase
6. Деякі особливості роботи з HBase
7. Альтернативи

Розглядаються цілі використання Hbase, модель даних та операції що підтримуються Hbase. Визначаються архітектура та способи роботи з Hbase та деякі особливості роботи з Hbase.

Література: 6, 16, 31, 32.

Ключові слова: модель даних Hbase, способи роботи з Hbase.

Тема 10 Темпоральні зміни даних

План теми:

1. Часовий домен
2. Вимоги до моделювання часових залежностей. Часова повнота. Часова щільність. Часовий ізоморфізм.
3. Математичне визначення часу. Часові елементи.
4. Класифікація інформації за відношенням до часу. Часовий куб даних.

Розглядаються темпоральні інформаційні системи та часовий домен. Визначаються вимоги до моделювання темпоральних залежностей. Наводиться часова повнота, щільність та ізоморфізм, математичне визначення часу, часові елементи, структурне подання часу. Розглядаються класифікація інформації по відношення до часу, часовий куб даних та об'єкти предметної області.

Література: 4, 5, 6, 10, 11, 18.

Ключові слова: часовий домен, часовий ізоморфізм, часові елементи, часовий куб даних.

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

План теми:

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

Розглядається історія та особливості мови Python. Визначаються переваги Python для роботи з Big Data та так званий «Дзен Python». Наводяться основні алгоритмічні конструкції. Типи даних у програмуванні. Визначення змінної. Логічні вирази. Умовний оператор. Інструкція if. Множинне розгалуження. Цикл while. Введення даних з клавіатури. 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.

Література: 3, 9-12, 15, 34, 37.

Ключові слова: Python, програма на Python, алгоритмічні конструкції Python.

Тема 12 Технології ІТ захисту інформації

План теми:

1. Основні поняття захисту інформації.
2. Види комп'ютерних злочинів.
3. Причини поширення комп'ютерної злочинності.
4. Поняття і класифікація комп'ютерних вірусів.
5. Засоби захисту інформації.
6. Комп'ютерні віруси як загроза інформаційним системам
7. Класифікація засобів захисту інформації

8. Основні заходи з антивірусного захисту
9. Методи криптографічного захисту
10. Біометричний захист інформації
11. Організація захисту комп'ютерних інформаційних систем

Розглядаються основні поняття захисту інформації, види комп'ютерних злочинів, причини поширення комп'ютерної злочинності. Визначаються поняття і класифікація комп'ютерних вірусів та засоби захисту інформації.

Література: 3, 16-21, 27, 30, 31.

Ключові слова: засоби захисту інформації, заходи з антивірусного захисту, криптографічний захист, біометричний захист.

Тема 13 Проблеми захисту та безпеки BigData

План теми:

1. Безпека даних.
2. Реєстрація користувачів.
3. Керування правами доступу та специфікація повноважень.
4. Обов'язкові методи захисту BigData.
5. Ведення журналів доступу.
6. Обхід системи захисту.
7. Захист при статистичній обробці даних.
8. Шифрування даних.
9. SQL ін'єкція

Розглядаються безпека даних, реєстрація користувачів, керування правами доступу та специфікація повноважень. Визначаються обов'язкові методи захисту BigData. Розглядаються варіанти обходу системи захисту, захист при статистичній обробці даних та шифрування даних.

Література: 3, 27, 30, 31.

Ключові слова: методи захисту BigData, обхід системи захисту, шифрування даних, SQL ін'єкція.

РЕЗУЛЬТАТИ НАВЧАННЯ

<p>Програмні результати навчання</p>	<p>ПРН1: Аналіз інформаційних джерел, вміння робити висновки у галузі інноваційних технологій.</p> <p>ПРН2: Представляти результати програмного продукту, використовуючи англійську термінологію.</p> <p>ПРН3: Застосовувати інноваційні технології BigData та методи інформаційної безпеки при вирішенні прикладних задач інших сфер діяльності.</p> <p>ПРН4: Вміти надавати свої знання та отримані рішення в ясній та однозначній формі.</p> <p>ПРН8: Розробляти запити(програмні коди) у середовищі обробки BigData для контролю та моніторингу процесу перевезень.</p> <p>ПРН10: Вміння обґрунтовувати доцільність застосування сучасних інноваційних технологій Big Data та інформаційної безпеки у транспортній галузі.</p>
--------------------------------------	--

	<p>ПРН12: Вміння управляти моніторинговою системою BigData, застосовувати знання про інформаційну безпеку в управлінні транспортними системами.</p> <p>ПРН14: Використовувати сучасні комп'ютерні програмні засоби в управлінні транспортними системами.</p> <p>ПРН15: Здатність проводити дослідницьку роботу у сфері інноваційних технологій.</p>
--	---

ОЦІНЮВАННЯ

Форми поточного та підсумкового контролю	Поточний контроль – 70 балів Проміжний контроль - 20 балів Підсумковий контроль – (екзамен) - 10 балів
КРИТЕРІЇ ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ	
Підсумкові бали з навчальної дисципліни визначаються як сума балів, отриманих здобувачем протягом семестру та балів, набраних на підсумковому контролі (екзамен, залік).	
Підсумкові бали навчальної дисципліни	$= \text{Загальна кількість балів (перед підсумковим контролем)} + \text{Кількість балів за підсумковим контролем}$

ШКАЛА ОЦІНЮВАННЯ: НАЦІОНАЛЬНА ТА ECTS			
Оцінка в балах	Оцінка за національною шкалою	Оцінка за шкалою ECTS	
		Оцінка	Пояснення
90-100	Відмінно («зараховано»)	A	«Відмінно» - теоретичний зміст курсу освоєний цілком, необхідні практичні навички роботи з освоєним матеріалом сформовані, всі навчальні завдання, які передбачені програмою навчання виконанні в повному обсязі, відмінна робота без помилок або з однією незначною помилкою.
80-89	Добре («зараховано»)	B	«Дуже добре» - теоретичний зміст курсу освоєний цілком, необхідні практичні навички роботи з освоєним матеріалом в основному сформовані, всі навчальні завдання, які передбачені програмою навчання виконанні, якість виконання більшості з них оцінено числом балів, близьким до максимального, робота з двома – трьома незначними помилками.
75-79		C	«Добре» - теоретичний зміст курсу освоєний цілком, практичні навички роботи з освоєним матеріалом в основному сформовані, всі навчальні завдання, які передбачені програмою навчання виконанні, якість виконання жодного з них не оцінено мінімальним числом балів, деякі види завдань виконані з помилками, робота з декількома незначними помилками, або з однією – двома значними помилками.
65-74	Задовільно («зараховано»)	D	«Задовільно» - теоретичний зміст курсу освоєний не повністю, але прогалини не носять істотного характеру, необхідні практичні навички роботи з освоєним матеріалом в основному сформовані, більшість передбачених програмою навчання навчальних завдань виконано, деякі з виконаних завдань, містять помилки, робота з трьома значними помилками.
60-64		E	«Достатньо» - теоретичний зміст курсу освоєний частково, деякі практичні навички роботи не сформовані, частина передбачених програмою навчання навчальних завдань не виконані, або якість

			виконання деяких з них оцінено числом балів, близьким до мінімального, робота, що задовольняє мінімум критеріїв оцінки.
21-59	Незадовільно («не зараховано»)	FX	«Умовно незадовільно» теоретичний зміст курсу освоєний частково, необхідні практичні навички роботи не сформовані, більшість передбачених програм навчання, навчальних завдань не виконано, або якість їхнього виконання оцінено числом балів, близьким до мінімального; при додатковій самостійній роботі над матеріалом курсу можливе підвищення якості виконання навчальних завдань (з можливістю повторного складання), робота, що потребує доробки
1-20		F	«Безумовно незадовільно» теоретичний зміст курсу не освоєно, необхідні практичні навички роботи не сформовані, всі виконані навчальні завдання містять грубі помилки, додаткова самостійна робота над матеріалом курсу не приведе до значимого підвищення якості виконання навчальних завдань, робота, що потребує повної переробки

СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

Базова

1. Faeze Ghofrani, QingHe, Rob M.P.Goverde, Xiang Liu. Recent applications of big data analytics in railway transportation systems: A survey.
2. П.Ю. Бойко, Е.М. Быков, Е.И. Соколов, Д.А. Яроцкий Применение машинного обучения к ранжированию инцидентов на Московской железной дороге.
3. Efreim Turban, Ramesh Sharda Business Intelligence.
4. Efreim Turban, Jay E. Aronson Decision Support and Business Intelligence Systems, 8/E.
5. 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/>.
6. 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/>.
7. Brighten your day with cognitive UEM (2017) Available at: <http://www-03.ibm.com/security/mobile/maas360.html>.
8. 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>.
9. United Nations (2016) World Economic Situation and Prospects. Available at: <http://www.unic.ru/sites/default/files/Доклад%20WESP2016.pdf>.

Допоміжна

10. Data Science: an introduction. Wikibooks [Electronic resource]. – 2017. – Mode of access: https://en.wikibooks.org/wiki/Data_Science:An_Introduction.
11. Forecasting: principles and practice [Electronic resource]. – 2017. – Mode of access: <https://www.otexts.org/fpp>.
12. Gabriel Martos. Cluster Analysis with R [Electronic resource] / RPubS. – 2015. – Mode of access: <https://rpubs.com/gabrielmartos/ClusterAnalysis>.
13. 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/>.
14. Golemund, G. R for Data Science [Electronic resource] / Garrett Golemund, Hadley Wickham. – 2016. – Mode of access: <http://r4ds.had.co.nz/index.html>. – Date of access: 01.09.2016.

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.
15. 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>.
16. 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>.
17. 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>.
18. Ogurtsov, A. Tidy data (Перевод) [Electronic resource]. – 2015. – Mode of access: <http://biostat-r.blogspot.com.by/2016/01/tidy-data.html>.
19. 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>.
20. Simko, V. From R code to R package [Electronic resource]. – 2015. – Mode of access: <https://github.com/vsimko/course-rpackages>.
21. Victor Lavrenko. Introductory Applied Machine Learning [Electronic resource]. – 2017. – Mode of access: <https://www.youtube.com/channel/UCs7alOMRnxhzhfKAJ4JjZ7Wg>.
22. Wickham, H. R packages [Electronic resource]. – 2015. – Mode of access: <http://rpkgs.had.co.nz/package.html>.
23. 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>.
24. Wickham, H. Tidy data [Electronic resource]. – 2017. – Mode of access: <ftp://cran.r-project.org/pub/R/web/packages/tidyr/vignettes/tidy-data.html>.
25. Wickham, H. Vignettes: long-form documentation [Electronic resource] / R packages. – 2015. – Mode of access: <http://r-pkgs.had.co.nz/vignettes.html>.
26. 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
27. Improving rolling stock maintenance using big data [Електронний ресурс] – Режим доступу: <https://www.globalrailwayreview.com/article/28433/big-data-trenitalia-rolling-stock-maintenance/>
28. Big data in railway operations and maintenance [Електронний ресурс] – Режим доступу: <https://www.globalrailwayreview.com/article/61515/big-data-railway-operations-maintenance-2/>.
29. 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>.
30. 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>.
31. Великі дані [Електронний ресурс] – Режим доступу: https://en.wikipedia.org/wiki/Big_data.
32. Big Data, методи і техніки аналізу великих даних [Електронний ресурс] – Режим доступу: <https://www.it.ua/ru/knowledge-base/technology-innovation/big-data-bolshie-dannye>.
33. Data Mining — добыча данных [Електронний ресурс] – Режим доступу: <https://basegroup.ru/community/articles/data-mining>.
34. Машинное обучение — это легко [Електронний ресурс] – Режим доступу: <https://habr.com/ru/post/319288/>.

35. Основы статистики та аналізу даних [Електронний ресурс] – Режим доступу: <http://socialdata.org.ua/manual4/>.
36. Что такое базы данных NoSQL? [Електронний ресурс] – Режим доступу: <https://aws.amazon.com/ru/nosql/>.
37. Python. Высокоуровневый язык программирования [Електронний ресурс] – Режим доступу: <https://habr.com/ru/hub/python/>.
38. Работа с датами на языке R (базовые возможности) [Електронний ресурс] – Режим доступу: <https://habr.com/ru/hub/r/>.
39. Big Data от А до Я: Hadoop [Електронний ресурс] – Режим доступу: <https://habr.com/ru/company/dca/blog/268277/>.
40. Зачем менеджеру язык программирования R: 7 причин освоить аналитический инструмент Big Data и Machine Learning [Електронний ресурс] – Режим доступу: <https://www.bigdataschool.ru/bigdata/>.
41. Hadoop: что, где и зачем [Електронний ресурс] – Режим доступу: <https://habr.com/ru/post/>