**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**ТАВРІЙСЬКИЙ ДЕРЖАВНИЙ АГРОТЕХНОЛОГІЧНИЙ УНІВЕРСИТЕТ**

**ІМЕНІ ДМИТРА МОТОРНОГО**

**Кафедра «Фінанси, облік і оподаткування»**

|  |  |
| --- | --- |

***Big Data Analytics in finance***

**«Аналіз великих даних у фінансах»**

**МЕТОДИЧНІ ВКАЗІВКИ**

**ДЛЯ ВИКОНАННЯ САМОСТІЙНОЇ РОБОТИ**

для здобувачів ступеня вищої освіти «Магістр» зі спеціальності

*071 «Облік і оподаткування»* факультет економіки та бізнесу

The Big Data Analytics in Finance individual assignments are developed in the framework of ERASMUS+ CBHE project «Digitalization of economic as an element of sustainable development of Ukraine and Tajikistan» / DigEco618270-EPP-1-2020-1-LT-EPPKA2-CBHE-JP

2022 рік

Аналіз великих даних у фінансах [Електронний ресурс]: методичні вказівки для виконання самостійної роботи за освітнім компонентом «Аналіз великих даних у фінансах» для здобувачів ступеня вищої освіти «Магістр» зі спеціальності *071 «Облік і оподаткування» за ОПП Облік і оподаткування* – Мелітополь, ТДАТУ, - 12с.

**Розробник:**

**Кюрчев В.М.,** д.т.н., професор кафедри експлуатації та технічного сервісу машин, ТДАТУ

**Мінц О. Ю.** - д.е.н., професор, завідуючий кафедрою «Фінанси і банківська справа»

**Демчук О.М.** - к.е.н., доцент, доцент кафедри фінанси, облік і оподаткування

**Рецензенти**:

**Martin Schultz** - д.е.н, професор, Університет прикладних наук Гамбургу (Німеччина)

**Легеза Д.Г.**  - д.е.н., професор, завідувач кафедри маркетингу ТДАТУ

Розглянуто та рекомендовано до друку на засіданні кафедри фінансів, обліку і оподаткування

Протокол № *11* від « *07*» червня 2022 року

Завідувач кафедри фінансів, обліку і оподаткування

д.е.н, професор \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Олег СОКІЛ

Схвалено методичною комісією факультету економіки та бізнесу для здобувачів освітнього рівня «Магістр» за спеціальності 075 «Маркетинг»

Протокол № *7* від «*20*»  *червня* 2022 року

Голова, доц. \_\_\_\_\_\_\_\_\_\_\_\_\_ Анна КОСТЯКОВА

Схвалено Навчально-методичною радою Таврійського державного агротехнологічного університету імені Дмитра Моторного для здобувачів освітнього рівня «Магістр» за спеціальності 075 «Маркетинг»

Протокол № *1* від «*26*»  *серпня* 2022 року

Голова, доц. \_\_\_\_\_\_\_\_\_\_\_\_\_ Олександр ЛОМЕЙКО

© Кюрчев В.М., 2022 рік

© Мінц О.Ю., 2022 рік

©Демчук О.М., 2022 рік

**ЗМІСТ**

[ТЕМА 1. ВСТУП. ПОНЯТТЯ ВЕЛИКИХ ДАНИХ ТА ЇХ КЛАСИФІКАЦІЯ. КОНЦЕПЦІЇ ВЕЛИКИХ ДАНИХ 4](#_heading=h.gjdgxs)

[ТЕМА 2. ПОНЯТТЯ РИНКУ ВЕЛИКИХ ДАНИХ. ЖИТТЄВИЙ ЦИКЛ АНАЛІТИКИ ДАНИХ. ЗБІР ТА ПІДГОТОВКА ДАНИХ 4](#_heading=h.30j0zll)

[ТЕМА 3. ТЕХНОЛОГІЇ ЗБЕРІГАННЯ ВЕЛИКИХ ДАНИХ 5](#_heading=h.1fob9te)

[ТЕМА 4. ТЕХНОЛОГІЇ АНАЛІЗУ ДАНИХ 5](#_heading=h.3znysh7)

[ТЕМА 5. ВІЗУАЛІЗАЦІЯ ВЕЛИКИХ ДАНИХ 6](#_heading=h.2et92p0)

[ТЕМА 6. ВЕЛИКІ ДАНІ ФІНАНСОВОЇ СФЕРИ 6](#_heading=h.tyjcwt)

[ТЕМА 7. БЕЗПЕКА ТА КОНФІДЕНЦІЙНІСТЬ ВЕЛИКИХ ДАНИХ У ФІНАНСАХ 7](#_heading=h.3dy6vkm)

[ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ОПРАЦЮВАННЯ НА ОНЛАЙН РЕСУРСАХ 7](#_heading=h.1t3h5sf)

[РЕКОМЕНДОВАНА ЛІТЕРАТУРА 9](#_heading=h.4d34og8)

# ТЕМА 1. ВСТУП. ПОНЯТТЯ ВЕЛИКИХ ДАНИХ ТА ЇХ КЛАСИФІКАЦІЯ. КОНЦЕПЦІЇ ВЕЛИКИХ ДАНИХ

**Питання для самостійного опрацювання**

1. Стан поширення технологій BigData в Україні

2. Особливості застосування технологій BigData в різних сферах діяльності

3. Характеристика філософського підходу до поняття даних

4. Юридичне трактування даних та великих даних

**Пропоновані теми рефератів:**

1. Цифрові технології в умовах становлення цифрової економіки.

2. Цифрова держава в Дії – перспективи розвитку

3.Передумови виникнення та розвитку технологій BigData

# ТЕМА 2. ПОНЯТТЯ РИНКУ ВЕЛИКИХ ДАНИХ. ЖИТТЄВИЙ ЦИКЛ АНАЛІТИКИ ДАНИХ. ЗБІР ТА ПІДГОТОВКА ДАНИХ

**Питання для самостійного опрацювання**

1. Технологія збору великих даних з відкритих джерел

2. Технологія збору великих даних в соціальних мережах

3. Загрози які несуть великі дані людству

4. Алгоритми перевірки та очищення даних

**Пропоновані теми рефератів:**

1. Порівняльний аналіз джерел відкритих даних України та США

2. Порівняльний аналіз джерел відкритих даних України та країн Європи

3. Соціальні мережі – знахідка для аналітика

4. Відкриті дані – користь та приховані загрози

# ТЕМА 3. ТЕХНОЛОГІЇ ЗБЕРІГАННЯ ВЕЛИКИХ ДАНИХ

**Питання для самостійного опрацювання**

1. Де монтується обладнання для обробки великих даних?

2. З яких компонентів складається обладнання для обробки великихданих?

3. Які параметри істотні для обчислювальних вузлів кластера?

4. Які параметри істотні для системи зберігання даних?

5. Які параметри істотні для мережевої інфраструктури?

6. Скільки максимально вузлів може мати кластер, який використовуєархітектуру підключення вузлів гиперкуб, при наявності чотирьохшвидкодіючих мережевих портів в кожному вузлі?

**Пропоновані теми рефератів:**

1. Центри обробки даних

2. Порівняльний аналіз систем зберігання даних

3. Новітні технології зберігання інформації

4. Еволюція технологій зберігання інформації

# ТЕМА 4. ТЕХНОЛОГІЇ АНАЛІЗУ ДАНИХ

**Питання для самостійного опрацювання**

1. Нейронні мережі

2. Характеристика функціональних зав’язків аналітики великих даних

3. Еволюційне програмування

4. Генетичні алгоритми

5. Алгоритми обмеженого перебору

**Пропоновані теми рефератів:**

1. Еволюція технологій аналізу даних

2. Чи несе загрозу людству штучний інтелект

3. Еволюція машинного навчання

4. Новітні методи аналізу великих даних

# ТЕМА 5. ВІЗУАЛІЗАЦІЯ ВЕЛИКИХ ДАНИХ

**Питання для самостійного опрацювання**

1. Вибір техніки візуалізації даних

2. Інтерактивна візуалізація

3. Візуалізація бізнес-процесів

4. Використання інструментів візуалізації для фінансових звітів

**Пропоновані теми рефератів:**

1. Еволюція засобів візуалізації даних

2. Перспективи розвитку засобів візуалізації даних

3. Візуалізація як спосіб розуміння даних

# ТЕМА 6. ВЕЛИКІ ДАНІ ФІНАНСОВОЇ СФЕРИ

**Питання для самостійного опрацювання**

1. Виклики великим даним у фінансовій сфері

2. Застосування великих даних для аналізу біржових трендів

3. Застосування великих даних для передбачень в страховій діяльності

4. Великі дані та макро прогнозування

**Пропоновані теми рефератів:**

1. Технологічні еволюції для підтримки обробки великої кількості складних та різноманітних даних у режимі реального часу

2. Хмарні рішення для передачі великих даних у фінансах

3. Аналіз фондового ринку на основі великих даних

4. Перспективи розвитку спектру банківських послуг на основі аналізу великих даних

# ТЕМА 7. БЕЗПЕКА ТА КОНФІДЕНЦІЙНІСТЬ ВЕЛИКИХ ДАНИХ У ФІНАНСАХ

**Питання для самостійного опрацювання**

1. Види шахрайств в банківській сфері

2. Методи протидії шахрайствам банківської сфери

3. Інформаційна «гігієна» клієнтів банку

**Пропоновані теми рефератів:**

1. Скорингові технології оцінювання ризиків шахрайства

2. Сучасні технології кібер безпеки банку

3. Правовий захист персональних даних клієнтів банку

4. Технології шифрування даних

# ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ОПРАЦЮВАННЯ НА ОНЛАЙН РЕСУРСАХ

Пройдіть запропоновані курси для поглибленого засвоєння матеріалу і в підтвердження проведеної роботи отримайте сертифікати:

На платформі Prometheus(попередньо авторизувавшись та знайшовши відповідний тренінг в пошуку):

https://courses.prometheus.org.ua

1. Основи фінансів та інвестицій

2. Аналіз даних та статистичне виведення на мові R

3. Візуалізація даних

На платформі Дія. Цифрова освіта (попередньо авторизувавшись та знайшовши відповідний тренінг в пошуку):

https://osvita.diia.gov.ua

1. Доступ до публічної інформації

2. Відкриті дані для державних службовців

3. Доступ до публічної інформації

3. Відкриті дані для бізнесу

# РЕКОМЕНДОВАНА ЛІТЕРАТУРА

**Основна**

1. Chapman, Arthur D. “Principles and methods of data cleaning.” 2005
2. D. Hand, H. Mannila, P. Smyth. 2001. Principles of Data Mining
3. Daniel Keim, JörnKohlhammer, Geoffrey Ellis und Florian Mansmann. „Visual Analytics“. 2010
4. Data Science from Scratch by Joel Grus (O‘REILLY) Elements of Statistical Learning: Data Mining, Inference, and Prediction, by Trevor Hastie, Robert Tibshirani and Jerome Friedman
5. Dimitri P. Bertsekas and John N. Tsitsiklis. Introduction to Probability. Charles Wheelan. Naked Statistics: Stripping the Dread from the Data. W. W. Norton and Company, 2013.
6. E. Friedmann, T. Dunning. 2015. Time Series Databases: New Ways to Store and Access Data
7. Guyon, Isabelle, Nada Matic and Vladimir Vapnik. "Discovering Informative Patterns and Data Cleaning." 1996
8. J. Han, M. Kamber. 2011. Data Mining. Concepts and Techniques
9. James Thomas und Kristin Cook. „Illuminating the Path“. 2005 vis.pnnl.gov/pdf/RD\_Agenda\_VisualAnalytics.pdf
10. Osborne, Jason W. “Best practices in data cleaning: A complete guide to everything you need to do before and after collecting your data.” 2013
11. Python for Data Analysis by Wes McKinney (O‘REILLY)
12. R. M. Müller, H.-J. Lenz. 2013. Business Intelligence
13. Steven Skiena. “The Data Science Design Manual” <http://www.data-manual.com/>
14. Гурвиц Дж., Ньюджент А., Халпер Ф., Кауфман М. Просто о больших данных: пер. с англ. М.:Эксмо, 2015. 400 с.
15. Кобзаренко Д.Н., Мустафаев А.Г. Учебное пособие дисциплины«Анализ больших данных» для направления подготовки 38.03.05«Бизнесинформатика», профиль «Электронный бизнес». – Махачкала: ДГУНХ, 2019 г.– 107 с.
16. Майер-Шенбергер В. Большие данные. Революция, которая изменит то, как мы живем, работаем имыслим. – М.: Манн, Иванов и Фербер, 2017. – 240 c
17. Радченко И.А, Николаев И.Н. Технологии и инфраструктура Big Data. СПб:Университет ИТМО, 2018. 52 с.
18. Силен Дэвид, МайсманАрно, Али МохамедОсновы Data Science иBig Data. Python и наука о даных. СПб.: Питер, 2017. 336 с.
19. Томас Єрл, ВаджидХаттак, Пол БулерОснови Big Data: Концепції, алгоритми та технології/Пер.з англ. Анатолія Гладуна;За наук.ред. Олексія Найди. Дніпро: «Баланс Бізнес Букс», 2018. 320 с.
20. Фрэнк Билл. Революция в аналитике. Как в эпоху Big Data улучшить вашбизнес с помощью операционной аналитики М.: Альпина Паблишер, 2014. 430 с
21. Фрэнкс Билл. Укрощение больших данных М.: Манн, Иванов и Фербер, 2014. 352 с.
22. ЫнАнналин, Су Кеннет Теоритический минимум по Big Data, что нужно знать о больших даных. СПб.: Питер, 2019. 208 с.

**Допоміжна**

1. 36 кращих інструментів для візуалізації даних. URL: <https://toplead.com.ua/ua/blog/id/38-luchshih-instrumentov-dlja-vizualizacii-dannyh-160/>
2. Big data от А до Я. Часть 3: Приемы и стратегии разработки MapReduce-приложений URL: https://habrahabr.ru/company/dca/blog/270453/ 5
3. BigData от А до Я. Часть 1: Принципы работы с большими данными, парадигма MapReduce URL: <https://habrahabr.ru/company/dca/blog/267361/>
4. Большие данные в финансовой отрасли: обзор и оценка перспектив развития мирового и российского рынков. URL: [file:///C:/Users/PC/Downloads/idc-26012016.pdf](about:blank)
5. Великі дані в індустрії фінансових послуг – від даних до аналізу. URL: <https://www.finextra.com/blogposting/17847/big-data-in-the-financial-services-industry---from-data-to-insights>
6. Великі дані у фінансах – ваш посібник з аналізу фінансових даних. URL: <https://www.talend.com/resources/big-data-finance/>
7. Вуколов Э.А. Основы статистического анализа. Практикум по статистическим методами исследованию операций с использованием пакетов STATISTICA и EXCEL: учебноепособие для студентов вузов. – М.: ФОРУМ: ИНФРА-М, 2013. – 464 с.
8. Гобарева Я.Л., Городецкая О.Ю., Золотарюк А.В. Бизнес-аналитикасредствами Excel М.: Вузовский учебник, ИНФРА-М, 2013. –336 с.
9. Захист персональних даних . URL: <https://www.atcominvestbank.com/uk/about/personal-data-protection> URL:
10. ІЕ безпека і захист інформації. URL: <https://infotel.ua/ua/IT-bezopasnost-i-zacshita-informatsii-1/>
11. Ілюк А. Ризики, пов'язані із захистом персональних даних в контексті BigData. *Юридична газета online*. URL: https://yur-gazeta.com/publications/practice/inshe/riziki-povyazani-iz-zahistom-personalnih-danih-v-konteksti-big-data.html
12. Квартальний Н. Великі дані у фінансах – роль аналізу фінансових даних. URL: https://inoxoft.com/blog/big-data-in-financial-services-role-of-financial-data-analysis-inoxoft/
13. Константинівна І.О. Сучасні криптосистеми. URL: <https://sites.google.com/site/sucasnikriptosistemik/home>
14. Круковець Д. Можливості Data Science в центральних банках: огляд. URL: https://journal.bank.gov.ua/uploads/articles/249\_2\_Krukovets\_Ukr.pdf
15. Мега Шах. Як використовувати аналітику великих даних для вдосконалення фінансової індустрії URL: https://www.techfunnel.com/fintech/how-to-use-big-data-analytics-to-improve-finance-industry/
16. Основные понятия безопасности для работы с Кластеры больших данных SQL Server URL: <https://docs.microsoft.com/ru-ru/sql/big-data-cluster/concept-security?view=sql-server-ver15>
17. Охорона банківської таємниці: правові засади. URL: <http://obt.inf.ua/page10.html#top>
18. Про Систему BankID Національного банку. URL: <https://bank.gov.ua/ua/bank-id-nbu>
19. Распределенные базы и хранилища данных : Электронный учебник / Н. Аносова, О. Бородин, Е. Гаврилов и др. – НОУ "ИНТУИТ" URL: <http://www.intuit.ru/studies/courses/1145/214/info>
20. Распределенные файловые системы. Технологии хранения и обработкибольших объемов данных / Computer Science Center URL: https://compscicenter.ru/ courses/ big-data/ 2015-spring/ classes/1117/.
21. Сем Палмер. Аналіз великих даних у фінансових послугах. URL: <https://www.devteam.space/blog/big-data-analytics-in-financial-services/>
22. Ситник В. Ф., Краснюк М. Т. Інтелектуальний аналіз даних (дейтамайнінг): навч. посіб.К.: КНЕУ, 2007. 376 с.
23. Черняк Л. Безопасность больших данных. URL:<https://www.osp.ru/os/2013/02/13034551>
24. Что нужно знать о Big Data в финансах: краткое руководство. URL: <https://aboutdata.ru/2017/06/06/big-data-and-finance/>
25. Шандрівська О. Є., Кириленко А. А.Особливості ідентифікації ризиків ринку big data. [URL: http://science.lpnu.ua/sites/default/files/journal-paper/2021/jun/23774/menedzhment121-84-97.pdf](about:blank)
26. Шипунов А.Б., Балдин Е.М., Волкова П.А., Коробейников А.И., Назарова С.А., ПетровС.В., Суфиянов В.Г. Наглядная статистика. Используем R! - М.: ДМК Пресс, 2012. – 298с.
27. Шулин С. Будущее фынансового сектора за Big Data. URL: <https://bluescreen.kz/articles/budushhee-finansovogo-sektora-za-bigdata/>
28. Що таке великі дані у фінансах. URL: <https://corporatefinanceinstitute.com/resources/knowledge/other/big-data-in-finance/>
29. Як великі дані покращують банківські та фінансові системи. URL: <https://innovecs.com/blog/big-data-in-banking-and-financial-systems/>
30. Як і для чого використовують візуалізацію даних? URL: http://eidos.org.ua/novyny/yak-i-dlya-choho-vykorystovuvaty-vizualizatsiyu-danyh/

**Рекомендовані джерела інформації**

1. Законодавчі та нормативно-правові документи. URL: [http://zakon.rada.gov.ua/](http://zakon.rada.gov.ua/7)
2. <https://books.google.com/ngrams>
3. <https://www.netflixprize.com/assets/GrandPrize2009_BPC_BellKor.pdf>
4. [https://www.data.gov](https://www.data.gov/)
5. <http://data.europa.eu/euodp/de/about>
6. <https://www.govdata.de/>
7. <http://transparenz.hamburg.de/open-data/>
8. <https://dasl.datadescription.com/datafiles/>
9. <http://linkedscience.org/>
10. <https://www.opensciencedatacloud.org/>
11. <https://philogb.github.io/jit/>
12. <http://tagesnetzwerk.de/>
13. [https://tweetping.net](https://tweetping.net/)
14. <https://flowingdata.com/2017/08/18/catalog-of-visualization-types-to-find-the-one-that-fits-your-dataset/>
15. <https://flowingdata.com/2015/12/15/a-day-in-the-life-of-americans/>
16. <https://callingbullshit.org/tools/tools_proportional_ink.html>
17. <http://www.visualisingdata.com/2014/04/the-fine-line-between-confusion-and-deception>
18. <http://tylervigen.com/spurious-correlations>
19. www.domo.com/learn/infographic-data-never-sleeps
20. http://www.originlab.de/
21. http://support.minitab.com
22. https://flowingdata.com/tag/treemap/
23. http://mars.wiwi.hu-berlin.de/mediawiki/teachwiki/index.php/Analysis\_of\_Mortality
24. http://jn.physiology.org/content/104/4/2103
25. http://www.sigchi.org/chi96/proceedings/papers/Tweedie/lt1txt.htm
26. http://rosuda.org/software/Gauguin/gauguin.html