

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДЕРЖАВНИЙ УНІВЕРСИТЕТ «Житомирська політехніка»

«АНАЛІЗ ВЕЛИКИХ ДАНИХ ТА БІЗНЕС-АНАЛІТИКА»

(назва навчальної дисципліни)

ПРОГРАМА
навчальної дисципліни

для здобуття освітнього ступеня магістр
за спеціальністю 051 «Економіка»
освітня програма «Економіка»

2021 рік

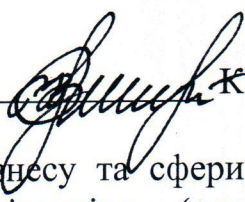
Укладачі програми: І. І. Сугоняк кандидат технічних наук, доцент,
А. В. Морозов, кандидат технічних наук, доцент.

Програма навчальної дисципліни визначає місце і значення навчальної дисципліни, її загальний зміст та вимоги до знань і вмінь.

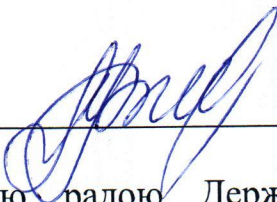
Програму розроблено в рамках проєкту Erasmus+ “Діджиталізація економіки як елемент сталого розвитку України та Таджикистану (DigEco) 618270-EPP-1-2020-1-LT-EPPKA2-CBHE-JP”/ The program is developed in the framework of ERASMUS+ CBHE project “Digitalization of economic as an element of sustainable development of Ukraine and Tajikistan” / DigEco 618270-EPP-1-2020-1-LT-EPPKA2-CBHE-JP викладачами університетів-партнерів проєкту

Цей проєкт фінансується за підтримки Європейської Комісії. Цей документ відображає лише погляди автора, і Комісія не несе відповідальності за будь-яке використання інформації, що міститься в документі/This project has been funded with support from the European Commission. This document 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.

Програма затверджена на засіданні кафедри цифрової економіки та міжнародних економічних відносин Державного університету «Житомирська політехніка» (протокол № 5 від 20.04.2021 р.)

Завідувач кафедри _____  Ж.В. Шиманська

Схвалено вченою радою факультету бізнесу та сфери обслуговування Державного університету «Житомирська політехніка» (протокол № 5 від 22.04.2021 р.)

Голова вченої ради факультету _____  Г.М. Тарасюк

Затверджено науково-методичною радою Державного університету «Житомирська політехніка» (протокол № 3 від 21.05.2021 р.)

Голова науково-методичної ради _____  А.В. Морозов



©Державний університет «Житомирська політехніка», 2021

DigEco 



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

ВСТУП

Програма вивчення навчальної дисципліни “ Аналіз великих даних та бізнес-аналітика” складена відповідно до освітньо-професійної програми підготовки здобувачів за другим (магістерським) рівнем вищої освіти за спеціальністю 051 «Економіка», освітня програма «Економіка» та спеціальністю 292 «Міжнародні економічні відносини», освітня програма «Міжнародний бізнес».

Предметом вивчення навчальної дисципліни є інструменти та аналітичні методи для використання даних для прийняття рішень, збору та впорядкування даних у масштабі та отримують розуміння того, як аналіз даних може допомогти інформувати зміни в організаціях.

Міждисциплінарні зв'язки. З дисциплінами «Стратегічне управління», «Ризик менеджмент», «Маркетинговий менеджмент».

Попередні дисципліни - менеджмент, інформатика, вища математика.

Забезпечувані дисципліни – планування і контроль на підприємстві.

Програма навчальної дисципліни складається з таких змістових модулів:

1. Сучасні аспекти цифрової трансформації бізнесу;
2. Діджиталізація бізнес-моделей.

1. Мета та завдання навчальної дисципліни

1.1. Метою викладання навчальної дисципліни “ Аналіз великих даних та бізнес-аналітика_” є підготувати фахівців зі знаннями у галузі великих даних; надання фахівцям навичок у галузі діяльності з удосконалення організації праці, виробництва та управління даними; вивчити принципи, методи та форми організації управління великими даними.

1.2. Основними завданнями вивчення дисципліни “ Аналіз великих даних та бізнес-аналітика ”

Застосовувати знання фундаментальних і природничих наук, системного аналізу та технологій моделювання, стандартних алгоритмів та дискретного аналізу при розв'язанні задач проектування і використання ІСТ.

Демонструвати знання сучасного рівня технологій інформаційних систем, практичні навички програмування та використання прикладних і спеціалізованих комп'ютерних систем та середовищ з метою їх запровадження у професійній діяльності.

Видобувати знання шляхом інтеграції та аналізу великих даних, отриманих з різноманітних та різнорідних джерела інформації. Вміти обґрунтовувати вибір абстрактних типів даних та структур даних при проектуванні програмного забезпечення ІСТ.

1.3. Перелік компетенцій:

- Здатність до абстрактного мислення, аналізу та синтезу.
- Здатність до розуміння предметної області та професійної діяльності.
- Здатність до пошуку, оброблення та узагальнення інформації з різних джерел.

- Здатність аналізувати об'єкт проектування або функціонування та його предметну область.
- Здатність використовувати сучасні ІСТ (виробничі, підтримки прийняття рішень, інтелектуального аналізу даних та інші), методики й техніки кібербезпеки під час виконання функціональних завдань та обов'язків.
- Здатність вибору, проектування, розгортання інтегрування, управління, адміністрування та супроводжування ІСТ та інфокомунікацій, сервісів та інфраструктури організації.
- Здатність оволодіти сучасними технологіями програмування та тестування програмного забезпечення.
- Здатність використовувати знання з у системах штучного інтелекту (СШІ), принципи побудови СШІ, зокрема, експертних систем, технологій побудови інтелектуальних систем, представлення їх в загальній структурі ІТ.

Згідно з вимогами освітньо-професійної програми студенти повинні:

знати:

- різні типи структур даних, формати файлів, джерела даних та мови, які спеціалісти використовують у своїх повсякденних завданнях;
- різні типи сховищ даних, такі як Базы даних, Сховища даних, Мапи даних, Озера даних та Конвеєри даних;
- процес вилучення, перетворення та завантаження (ETL), який використовується для вилучення, перетворення та завантаження даних у сховища даних;
- обробка великих даних;
- різні компоненти сучасної екосистеми даних та ролі, яку в цій екосистемі відіграють аналітики даних, науковці даних та бізнес-аналітики;
- навички, необхідні для того, щоб бути аналітиком даних, і як виглядає типовий проект аналізу великих даних.

вміти:

- розробляти алгоритми великих даних;
- розуміти основні концепції програмування, включаючи абстракцію даних, їх зберігання та структури;
- розуміти обчислювальне мислення, яке включає декомпозицію, розпізнавання шаблонів та абстрагування, керовані даними проблема та дизайн алгоритму для великих даних Інтерпретації подання та аналізу даних;
- розуміти ключові математичні поняття, включаючи зменшення розмірів та парадокси моделей;
- використовувати аналітичні інструменти, такі як Java, про які студенти дізнаються;
- аналізувати різні типи аналізу даних та ключові етапи процесу аналізу даних;
- розуміти різні компоненти сучасної екосистеми даних та роль аналітиків даних, науковців даних та бізнес-аналітиків у цій екосистемі.

На вивчення навчальної дисципліни відводиться 120 годин 4 кредитів ЄКТС.

2. Інформаційний обсяг навчальної дисципліни

Змістовий модуль 1. Сучасні аспекти цифрової трансформації бізнесу

Тема 1. ІТ-фактор впливу в сучасних управлінських завданнях

Використання ІТ визначається як міра, в якій організація використовує ІТ для підтримки оперативних та стратегічних завдань. Клімат ІТ-менеджменту, управлінські ІТ-знання, ефективність процесу ІТ-управління та використання ІТ. Модель структурного рівняння розроблена для оцінки відносних ефектів та взаємозв'язків між цими конструкціями. Управлінські ІТ-знання є домінуючим фактором у поясненні високого рівня використання ІТ, а також те, що як управлінські ІТ-знання, так і ефективність процесу ІТ-управління впливають на ІТ-клімат.

Тема 2. Моделювання та аналіз даних

Модель даних - це опис того, як дані повинні використовуватися для задоволення вимог, заданих кінцевим користувачем. Моделювання даних допомагає зрозуміти інформаційні вимоги. Моделювання даних відрізняється залежно від типу бізнесу, оскільки бізнес-процеси або кожен сектор є різним, і його потрібно визначити на етапі моделювання. Початковим кроком є аналіз ситуації, збір даних. Процес моделювання даних починається зі збору вимог. При розробці належної моделі даних важливо спілкуватися із зацікавленими сторонами щодо вимог. Моделювання даних - це акт дослідження структур, орієнтованих на дані. Це можна використовувати для різних цілей. Однією з важливих функцій моделювання даних є те, що вона допомагає зрозуміти інформаційні вимоги. Особливо це полегшує життя як розробникам, так і кінцевим користувачам. Як зазначалося вище, моделювання даних допомагає кінцевим користувачам визначити свої вимоги, а розробники можуть розробити систему, яка відповідає зазначеним вимогам.

Тема 3. Архітектура та методи зберігання даних

Дизайн сховища даних використовує певну тематику. Він надає інформацію про тему, а не про діяльність підприємства. Ці теми можуть бути пов'язані з продажами, рекламою, маркетингом тощо.

Замість того, щоб зосередитись на ділових операціях або операціях, зберігання даних робиться акцент на бізнес-аналітиці (BI), тобто; відображення та аналіз даних для прийняття рішень. Він також пропонує пряму та стислу інтерпретацію певної теми, виключаючи дані, які можуть бути не корисними для тих, хто приймає рішення.

Тема 4. Видобування, перетворення та навантаження (ETL)

Зараз трудомістка пакетна операція, ETL що рекомендується частіше для створення менших цільових сховищ даних, які вимагають менш частого оновлення, тоді як використовуються інші методи інтеграції даних, включаючи

ELT (витяг, завантаження, перетворення), CDC та віртуалізація даних. інтегрувати все більші обсяги потоків даних, що постійно змінюються, або потоків даних у режимі реального часу.

Змістовий модуль 2. Діджиталізація бізнес-моделей

Тема 5. Надання інформації (звітування, інформаційні панелі)

Конфігурація як код - це практика зберігання конфігурації вашої системи як набору контрольованих версій, зручних для читання файлів конфігурації, а не в базі даних. Ці файли конфігурації можна використовувати повторно в усіх середовищах, щоб уникнути дублювання ресурсів.

У міру зростання кількості інформаційних панелей та джерел даних в організації, керування змінами вручну може стати схильним до помилок. Заохочення повторного використання стає важливим, щоб уникнути переробки декількох команд однакових інформаційних панелей.

Тема 6. Аналітичний життєвий цикл та методи: кластеризація, класифікація, машинне навчання

Доступ до даних та їх збір можуть бути досить простими, якщо дані перебувають у сховищі даних, однак якщо дані розкидані у різних файлах, базах даних та операційних системах, що може бути нудним процесом. Якщо дані знаходяться в операційній системі, їх доведеться витягти в окреме сховище, щоб інтелектуальний аналіз даних не заважав нормальній роботі. Дані також потрібно оцінювати на предмет повноти, надмірності, відсутніх значень, дублювання, а іноді може знадобитися зменшення складності, наприклад, округлення даних та зменшення кількості десяткових значень. Наступним кроком є підготовка даних, і цей етап може бути як мистецтвом, так і наукою. Це справді найбільш трудомістка частина процесу видобування даних, і це часто може зайняти від 60 до 80% загальних зусиль. Однак, як тільки буде створений сценарій коду підготовки даних, побудова наступної моделі на даних одного типу може бути набагато швидшою, оскільки все, що потрібно зробити, це запустити цей скрипт на необроблених даних та завантаженні даних, і підготовка виконується автоматично (зазвичай з від декількох хвилин до декількох годин або на ніч).

Тема 7. Моделювання великих даних

Моделювання великих даних залежить від багатьох факторів, включаючи структуру даних, які операції можуть виконуватися з даними та які обмеження накладені на моделі. Big Data створює нові вимоги, засновані на складності збору даних, їх зберігання, аналізу та візуалізації даних.

Тема 8. Архітектура та розгортання

Big Data вимагає нових можливостей для управління та управління внутрішніми та зовнішніми інформаційними потоками, щоб перетворити їх у стратегічні ресурси, здатні визначати стратегії позиціонування, які відповідають потребам споживачів, які стають дедалі більш поінформованими та вимогливими. як технічні, накладають кардинальну зміну бізнес-моделей та

людських ресурсів з точки зору інформаційної спрямованості та вдосконалення корпоративного інформаційного активу як підтримку процесу прийняття стратегічних рішень. Менеджерам часто важко приймати рішення щодо інноваційних технологій, головним чином через відсутність технічних знань. Буде надано інструмент для підтримки цих оцінок з метою зменшення ризику, пов'язаного з інвестицією у великі дані. Цей інструмент пов'язує звичайні бар'єри, що перешкоджають успішному проекту, з технічним вибором, який слід зробити під час впровадження архітектури великих даних.

3. Рекомендована література

1. C. Chen, M. Lin, and X. Guo, “High-level modeling and synthesis of smart sensor networks for Industrial Internet of Things,” *Computers & Electrical Engineering*, vol. 61, pp. 48–66, 2017.
2. Daniel Keim, Jörn Kohlhammer, Geoffrey Ellis und Florian Mansmann. „Visual Analytics“. 2010
3. 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.
4. F. Liu, Y. Liu, D. Jin, X. Jia, and T. Wang, “Research on Workshop-Based Positioning Technology Based on Internet of Things in Big Data Background,” *Complexity*, vol. 2018, Article ID 875460, 11 pages, 2018.
5. H. Mora, M. Signes-Pont, D. Gil, and M. Johnsson, “Collaborative Working Architecture for IoT-Based Applications,” *Sensors*, vol. 18, no. 6, p. 1676, 2018.
6. H. Tahaei, R. Salleh, S. Khan, R. Izard, K.-K. R. Choo, and N. B. Anuar, “A multi-objective software defined network traffic measurement,” *Measurement*, vol. 95, pp. 317–327, 2017.
7. Hariri, R.H., Fredericks, E.M. & Bowers, K.M. Uncertainty in big data analytics: survey, opportunities, and challenges. *J Big Data* 6, 44 (2019). <https://doi.org/10.1186/s40537-019-0206-3>
8. INMON, W.H.; LINSTEDT, D.: *Data architecture a primer for the data scientist: big data, data warehouse and data vault*. 2014.
9. J. Han, M. Kamber. 2011. *Data Mining. Concepts and Techniques Visualize This* by Nathan Yau
10. J. Pan and J. McElhannon, “Future edge cloud and edge computing for internet of things applications,” *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 439–449, 2018.
11. L. J. M. Nieuwenhuis, M. L. Ehrenhard, and L. Prause, “The shift to Cloud Computing: The impact of disruptive technology on the enterprise software business ecosystem,” *Technological Forecasting & Social Change*, vol. 129, pp. 308–313, 2018.

12. M. Giacobbe, R. Di Pietro, A. Longo Minnolo, and A. Puliafito, "Evaluating Information Quality in Delivering IoT-as-a-Service," in Proceedings of the 2018 IEEE International Conference on Smart Computing (SMARTCOMP), pp. 405–410, June 2018.
13. M. Osman, "A novel big data analytics framework for smart cities," *Future Generation Computer Systems*, vol. 91, pp. 620–633, 2019.
14. Marrone, M. and Hazelton, J. (2019), "The disruptive and transformative potential of new technologies for accounting, accountants and accountability: A review of current literature and call for further research", *Meditari Accountancy Research*, Vol. 27 No. 5, pp. 677-694. <https://doi.org/10.1108/MEDAR-06-2019-0508>
15. Osborne, Jason W. "Best practices in data cleaning: A complete guide to everything you need to do before and after collecting your data." 2013
16. R. Lovas, A. Farkas, A. C. Marosi et al., "Orchestrated Platform for Cyber-Physical Systems," *Complexity*, vol. 2018, Article ID 8281079, 16 pages, 2018.
17. R. M. Müller, H.-J. Lenz. 2013. *Business Intelligence*
18. R. Y. Zhong, X. Xu, E. Klotz, and S. T. Newman, "Intelligent Manufacturing in the context of industry 4.0: a review," *Engineering Journal*, vol. 3, no. 5, pp. 616–630, 2017.
19. Steven Skiena. "The Data Science Design Manual" <http://www.data-manual.com/>
20. TURBAN, EFRAIM ; SHARDA, RAMESH ; DELEN, DURSUN ; KING, DAVID: *Business intelligence: a managerial approach*. Boston, Mass. : Pearson, Prentice Hall, 2011 www.vismaster.eu/wp-content/uploads/2010/11/VisMaster-book-lowres.pdf
21. X. Wang, C. Xu, G. Zhao, K. Xie, and S. Yu, "Efficient Performance Monitoring for Ubiquitous Virtual Networks Based on Matrix Completion," *IEEE Access*, vol. 6, pp. 14524–14536, 2018.
22. Y. Guo, Z. Yang, S. Feng, and J. Hu, "Complex Power System Status Monitoring and Evaluation Using Big Data Platform and Machine Learning Algorithms: A Review and a Case Study," *Complexity*, vol. 2018, Article ID 8496187, 21 pages, 2018.
23. Y. Su, X. Meng, Q. Kang, and X. Han, "Dynamic Virtual Network Reconfiguration Method for Hybrid Multiple Failures Based on Weighted Relative Entropy," *Entropy*, vol. 20, no. 9, p. 711, 2018.

4. Форма підсумкового контролю успішності навчання – іспит

5. Засоби діагностики успішності навчання – поточний контроль, модульний контроль, семестрове домашнє завдання

Примітки:

1. Програма навчальної дисципліни визначає місце і значення навчальної дисципліни, її загальний зміст та вимоги до знань і вмінь.
2. Засоби діагностики успішності навчання (розділ 5) - Усне та письмове опитування, тестування, поточна і модульна контрольна робота, захист лабораторної роботи, оцінка присутності та активності на лекціях, практичних і семінарських заняттях, захист або оцінка самостійної роботи.