

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ЗАПОРІЗЬКА ПОЛІТЕХНІКА»**

Кафедра Системного аналізу та обчислювальної математики

(найменування кафедри)

**СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**

Методи Data Mining

(назва навчальної дисципліни)

Освітня програма: Інтелектуальні технології та прийняття рішень в складних системах

(назва освітньої програми)

Спеціальність: 124 – Системний аналіз

(найменування спеціальності)

Галузь знань: 12 – Інформаційні технології

(найменування галузі знань)

Ступінь вищої освіти: бакалаврський

(назва ступеня вищої освіти)

Затверджено на засіданні кафедри

СА та ОМ

(найменування кафедри)

Протокол № \_\_\_ від \_\_\_\_\_ р.

м. Запоріжжя 2020

<b>1. Загальна інформація</b>	
<b>Назва дисципліни</b>	Методи Data Mining, вибіркова
<b>Рівень вищої освіти</b>	Перший (бакалаврський) рівень
<b>Викладач</b>	<b>Бакурова Анна Володимирівна, д.е.н., професор</b>
<b>Контактна інформація викладача</b>	Телефон кафедри 769-8247, Телефон викладача +380509023328, abaka111060@gmail.com
<b>Час і місце проведення навчальної дисципліни</b>	Предметна аудиторія кафедри, комп'ютерні класи згідно розкладу занять
<b>Обсяг дисципліни</b> 4.5 кредитів, загальний обсяг 135 годин	розподіл годин: денне навчання: лекції-30, лабораторні заняття-14, самостійна робота 90 заочне навчання: лекції-4, лабораторні заняття-2, вид контролю- екзамен 7 семестр
<b>Консультації</b>	Згідно з графіком консультацій
<b>2. Пререквізити і постреквізити навчальної дисципліни</b>	
<p><i>Пререквізити: математичні основи системного аналізу, дискретна математика, математична логіка і алгоритми, теорія ймовірностей й математична статистика, алгоритмічне програмування, чисельні методи, методи оптимізації та дослідження операцій, математичні методи системного аналізу, теорія прийняття рішень.</i></p> <p><i>Постреквізити: інтелектуальні системи підтримки прийняття рішень, системний аналіз соціально-економічних процесів, аналіз часових рядів</i></p>	
<b>3. Характеристика навчальної дисципліни</b>	
<b>Загальні компетентності</b>	
<p>K01. Здатність до абстрактного мислення, аналізу та синтезу</p> <p>K02. Здатність застосовувати знання у практичних ситуаціях</p> <p>K03. Здатність планувати і управляти часом</p> <p>K04. Знання та розуміння предметної області та розуміння професійної діяльності</p> <p>K05. Здатність спілкуватися державною мовою усно і письмово</p> <p>K07. Здатність до пошуку, оброблення та аналізу інформації з різних джерел</p> <p>K08. Здатність бути критичним і самокритичним</p> <p>K09. Здатність до адаптації та дії в новій ситуації</p> <p>K10. Здатність працювати автономно</p> <p>K11. Здатність генерувати нові ідеї (креативність)</p> <p>K12. Здатність працювати в команді</p> <p>K14. Здатність оцінювати та забезпечувати якість виконуваних робіт</p>	
<b>Спеціальні (фахові, предметні) компетентності</b>	
<p>K17. Здатність використовувати системний аналіз як сучасну міждисциплінарну методологію, що базується на прикладних математичних методах та сучасних інформаційних технологіях і орієнтована на вирішення задач аналізу і синтезу технічних, економічних, соціальних, екологічних та інших складних систем.</p> <p>K18. Здатність формалізувати проблеми, описані природною мовою, у тому числі за допомогою математичних методів, застосовувати загальні підходи до математичного моделювання конкретних процесів.</p> <p>K20. Здатність визначати основні чинники, які впливають на розвиток фізичних, економічних, соціальних процесів, виокремлювати в них стохастичні та невизначені показники, формулювати їх у вигляді випадкових або нечітких величин, векторів, процесів та досліджувати залежності між ними.</p> <p>K22. Здатність до комп'ютерної реалізації математичних моделей реальних систем і процесів; проектувати, застосовувати і супроводжувати програмні засоби моделювання, прийняття рішень, оптимізації, обробки інформації, інтелектуального аналізу даних.</p> <p>K23. Здатність використовувати сучасні інформаційні технології для комп'ютерної реалізації математичних моделей та прогнозування поведінки конкретних систем а саме: об'єктно-орієнтований підхід при проектуванні складних систем різної природи, прикладні математичні пакети, застосування баз даних і знань.</p> <p>K25. Здатність представляти математичні аргументи і висновки з них з ясністю і</p>	

точністю і в таких формах, які підходять для аудиторії як усно так і в письмовій формі.  
 K26. Здатність розробляти експериментальні та спостережувальні дослідження і аналізувати дані, отримані в них.  
 K27. Здатність розробляти і застосовувати моделі інтелектуальних систем прийняття рішень.

#### **Очікувані програмні результати навчання:**

ПР02. Вміти використовувати стандартні схеми для розв'язання комбінаторних та логічних задач, що сформульовані природною мовою, застосовувати класичні алгоритми для перевірки властивостей та класифікації об'єктів, множин, відношень, графів, груп, кілець, решіток, булевих функцій тощо.

ПР06. Знати та вміти застосовувати основні методи постановки та вирішення задач системного аналізу в умовах невизначеності цілей, зовнішніх умов та конфліктів.

ПР07. Знати основи теорії оптимізації, оптимального керування, теорії прийняття рішень, вміти застосовувати їх на практиці для розв'язування прикладних задач управління і проектування складних систем.

ПР09. Вміти створювати ефективні алгоритми для обчислювальних задач системного аналізу та систем підтримки прийняття рішень.

ПР12. Застосовувати методи і засоби роботи з даними і знаннями, методи математичного, логіко-семантичного, об'єктного та імітаційного моделювання, технології системного і статистичного аналізу.

ПР14. Розуміти і застосовувати на практиці методи статистичного моделювання і прогнозування, оцінювати вихідні дані.

ПР15. Розуміти українську та іноземну мови на рівні, достатньому для обробки фахових інформаційно-літературних джерел, професійного усного і письмового спілкування, написання текстів за фаховою тематикою.

ПР18. Обирати ефективні методи та здійснювати формалізоване подання складних систем і процесів з метою побудови і дослідження відповідних моделей.

ПР19. Розробляти інтелектуальні системи прийняття рішень, у тому числі за допомогою методів штучного інтелекту, моделювання систем, теорії прийняття рішень.

#### **4. Мета вивчення навчальної дисципліни**

*Метою вивчення дисципліни «Методи Data Mining» є формування у майбутніх спеціалістів знань і навичок щодо сучасних інформаційних систем і технологій взагалі та аналітичних систем і систем Data Mining, зокрема: вивчення основних моделей даних при проектуванні баз знань; вивчення механізмів вилучення знань з баз знань; вивчення математичного апарату подання і формалізації знань.*

#### **5. Завдання вивчення дисципліни**

*Завданням вивчення навчальної дисципліни є засвоєння основних методів та стандартів Data Mining та Text Mining; вміння використовувати отримані знання при розробці моделей інтелектуальних систем прийняття рішень.*

#### **6. Зміст навчальної дисципліни**

*У навчальній дисципліні «Методи Data Mining» розглядаються загальні принципи застосування методів, що використовуються для аналізу даних в системах підтримки прийняття рішень (Data Mining): регресія, метод опорних векторів та інші, їхнє використання при розв'язанні задач Data Mining; задачі аналізу текстової інформації. Частина курсу присвячена оупенсорсному програмному забезпеченню для аналізу даних методами Data Mining (мова R).*

#### **7. План вивчення навчальної дисципліни**

<b>№ тижня</b>	<b>Назва теми</b>	<b>Форми організації навчання</b>	<b>Кількість годин</b>
1.	<b>Загальні принципи методів Data Mining</b> Приклади застосування Data Mining. Доступне програмне забезпечення систем Data Mining. Нормалізація і стандартизація	Лекції, лабораторні заняття, самостійна робота	4 2 15

	вхідних та вихідних значень. Методи відбору інформативних ознак..		
2.	<b>Стандарти систем Data Mining</b> Характеристики процесів і активностей Data Mining. Архітектура, класифікація та стандарти систем Data Mining.	Лекції, самостійна робота	4 15
3.	<b>Основні задачі Data Mining</b> Логістична регресія та задача класифікації. Скоринг. Задача відновлення регресії. Проблема мультиколінерності. Лінійна регресія.	Лекції, лабораторні заняття, самостійна робота	6 4 20
4.	<b>Метод опорних векторів</b> Метод опорних векторів для задачі класифікації. Постановки задач та приклади їх розв'язання і використання. Лінійно роздільна та нероздільна вибірки. Метод опорних векторів в задачах регресії.	Лекції, лабораторні заняття, самостійна робота	6 4 20
5	<b>Методи Text Mining</b> Прикладні задачі аналізу тексту. Морфологічний аналіз та лематизація. Синтаксичний аналіз. Задача інформаційного пошуку. Прості тематичні моделі. Сингулярний розклад. Ймовірнісний латентно семантичний аналіз. Постановка задачі тематичного моделювання.	Лекції, лабораторні заняття, самостійна робота	10 4 20

#### **8. Самостійна робота**

*Самостійна робота студентів передбачає ознайомлення з основами оупенсорсного програмного забезпечення для розв'язування основних задач Data Mining та Text Mining.*

#### **9. Система та критерії оцінювання курсу**

*В якості контрольного заходу з курсу «Методи Data Mining» передбачено екзамен. Екзамен виставляється за результатами захисту 4-х лабораторних робіт та 2-х модульних контрольних робіт*

#### **10. Політика курсу**

*Академічна доброчесність.* Студенти при використанні чужих ідей і тверджень у власних роботах обов'язково посилаються на використані джерела інформації. Під час оцінювання результатів навчання самостійно виконують навчальні завдання, завдання поточного та підсумкового контролю.

*Відвідування занять.* Очікується, що всі студенти відвідають усі лекції та практичні заняття курсу в аудиторіях або дистанційно (за умовами карантину). За згоди декана та викладача дозволяється перейти на індивідуальний графік занять. У будь-якому випадку студенти зобов'язані дотримуватися термінів виконання усіх видів робіт, передбачених робочою програмою курсу.

#### **11. Рекомендована література**

##### **Базова**

1. Гладун А. Я. Data Mining: пошук знань в даних / А. Я. Гладун, Ю. В. Рогушина. – Київ. ТОВ «ВД «АДЕФ- Україна», 2016. — 452 с. ISBN 978- 617-7393-22-0
2. Гладун А. Я. Семантичні технології: принципи та практики /А. Я. Гладун, Ю. В. Рогушина. – К.: ТОВ «ВД «АДЕФ-Україна», 2016. – 308 с. ISBN 978- 617-7156-95-5
3. Марченко О. О., Россада Т.В. Актуальні проблеми Data Mining: Навчальний посібник для студентів факультету комп'ютерних наук та кібернетики. — Київ. — 2017. — 150 с.
4. Литвин В.В. Методи та засоби інженерії даних та знань / В.В.Литвин – Львів: Магнолія-2006, 2012. – 241 с.
5. Майер-Шенбергер, В. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / Виктор Майер-Шенбергер, Кеннет Кукьер ; пер. с англ.

Инны Гайдюк. — М. : Манн, Иванов и Фербер, 2014. — 240 с.

6. Олійник А.О. Інтелектуальний аналіз даних : навчальний посібник / А. О. Олійник, С. О. Субботін, О.О.Олійник. – Запоріжжя : ЗНТУ, 2012. – 278 с.

#### **Допоміжна**

7. В. Дюк, А. Самойленко. Data Mining: Учебный курс – СПб.: Питер, 2001. – 368с.

8. Ситник В.Ф. Інтелектуальний аналіз даних. К.: КНЕУ, 2007.

9. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям: Учеб. пособие. 2-е изд., перераб. и доп. \_ СПб.: Питер, 2010. – 704 с.

10. Барсегян А.А., Куприянов М.С. и др. Методы и модели анализа данных: OLAP и Data Mining. – СПб.: БХВ-Петербург, 2004. – 336 с.

11. Чубукова И.А. Data mining: учебное пособие – М.: Интернет-университет информационных технологий: БИНОМ: Лаборатория знаний, 2006. – 382 с. – ISBN 5-9556-0064-7.

12. Снитюк В.Є. Прогнозування: Моделі, методи, алгоритми. – К.: Маклаут, 2008. – 364 с. Навчальний посібник з грифом МОН України.

13. Bakhrushin V. Risks of Data Inconsistency in Information Systems Used for Predicting the Pandemics Development / Bakhrushin V., Bakurova A., Pasichnyk M, Tereschenko E. // 1st International Workshop on Computational & Information Technologies for Risk-Informed Systems (CITRisk 2020) co-located with XX International scientific and technical conference on Information Technologies in Education and Management (ITEM 2020). Kherson, Ukraine, October 15-16, 2020. - Kherson National Technical University: 2020. – Vol-2805. – P. 1-15 .  
Режим доступу: <http://ceur-ws.org/Vol-2805/invited1.pdf>

#### **Інформаційні ресурси**

Єдиний державний веб-портал відкритих даних (<https://data.gov.ua>)

Державна служба статистики України (<http://www.ukrstat.gov.ua>)