

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСА ШЕВЧЕНКА  
ФАКУЛЬТЕТ КОМП'ЮТЕРНИХ НАУК ТА КІБЕРНЕТИКИ  
Кафедра прикладної статистики**



**«ЗАТВЕРДЖУЮ»**

Заступник декана  
з навчальної роботи

Кашпур О. Ф.

«11» листопада 2020 року

**РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ  
СТАТИСТИЧНІ МЕТОДИ АНАЛІЗУ ДАНИХ /STATISTICAL  
METHODS FOR DATA ANALYSIS  
для студентів / for students**

галузь знань  
спеціальність  
освітній рівень  
освітня програма

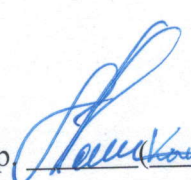
12 «Інформаційні технології»/ Information Technologies  
122 «Комп'ютерні науки»/ Computer Science  
магістр / Master  
«Математичні Методи Штучного Інтелекту»  
/ «Mathematical Methods of Artificial Intelligence»

вид дисципліни

вибіркова / free choice

|                                            |                      |
|--------------------------------------------|----------------------|
| Форма навчання                             | денна                |
| Навчальний рік                             | 2020/2021            |
| Семестр                                    | 3                    |
| Кількість кредитів ECTS                    | 6                    |
| Мова викладання, навчання<br>та оцінювання | англійська / English |
| Форма заключного контролю                  | іспит / exam         |

Викладачі: д.ф.-м.н, доц. Розора І.В.

Пролонговано: на 2021/2022 н.р.  «11» 02 2021 р.

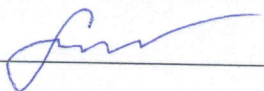
на 20\_\_/20\_\_ н.р. \_\_\_\_\_ (\_\_\_\_\_) «\_\_» \_\_\_\_ 20\_\_ р.

**КИЇВ – 2020**

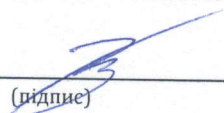
Розробник: **Розора І.В.** д. фіз.-мат. н., доцент, доцент кафедри прикладної статистики

ЗАТВЕРДЖЕНО

Завідувач кафедри прикладної статистики

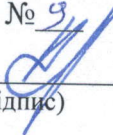
 (Лебедев Є.О.)

Протокол № 4 від «2» квітня 2020 р.

Схвалено Гарантом освітньо-наукової програми «Математичні методи штучного інтелекту»  (Завадський І.О.)

«2» 04 2020 р.

Схвалено науково-методичною комісією ф-ту комп'ютерних наук та кібернетики  
Протокол від «2» 04 2020 року № 9

Голова науково-методичної комісії  (Омельчук Л.Л.)

«2» 04 2020 року

**1. Мета дисципліни** – ознайомлення студентів з сучасними математичними методами аналізу даних та практична підготовка студентів до проведення статистичних досліджень, їх аналізу та обробки.

**Discipline aim.** The purpose of the discipline " Statistical methods for Data Analysis", acquaintance of students with modern mathematical methods of data analysis and practical preparation of students for carrying out statistical researches, their analysis and processing..

## **2 Попередні вимоги до опанування або вибору навчальної дисципліни**

*Знати:* основи дискретної математики, теорії ймовірностей та математичної статистики

*Вміти:* формалізувати умови задач та складати план розв'язку

*Володіти елементарними навичками:* розв'язувати типові задачі з теорії ймовірностей, математичної статистики та дискретної математики.

### **Preliminary demands to master or choice of the course discipline:**

*Know:* basics of discrete mathematics, probability theory and mathematical statistics

*Be able to:* formalize the conditions of problems and make a solution plan

*Have basic skills:* solve typical problems in probability theory, mathematical statistics and discrete mathematics.

## **3 Анотація навчальної дисципліни**

Дисципліна «*Статистичні методи аналізу даних*» є складовою частиною циклу професійної підготовки фахівців освітньо-кваліфікаційного рівня „магістр”; вона включає вивчення таких понять, як регресійні (прості, множинні, узагальнені) моделі, логістичні моделі, кореляційний аналіз, дисперсійний аналіз тощо. Особлива увага приділяється візуалізації даних та методам прогнозування. Викладається у 3-му семестрі, обсяг 180 год. (6 кредити ECTS), з них лекції – 38 год., лабораторних – 20 год. консультації – 2 год., самостійна робота – 120 год. Передбачено 1 частину та іспит.

### **Synopsis of the course:**

The discipline " Statistical methods for Data Analysis" is an integral part of the cycle of professional training of specialists of educational and qualification level "master"; it includes the study of such concepts as regression (simple, multiple, general) models, logistic models, correlation analysis, ANOVA analysis. Particular attention is paid to visualization and data forecasting. Taught in the 3th semester, volume 180 hours. (6 ECTS credits), of which lectures - 38 hours, laboratory - 20 hours. consultations - 2 hours, independent work - 120 hours. Part 1 and exam are provided.

## **4 Завдання (навчальні цілі)**

Набуття знань, умінь та навичок (компетентностей) на основі використання системного підходу та його принципів для створення інтелектуальних бізнес систем, на основі сучасних інформаційних та комп'ютерних технологій, моделей і методів штучного інтелекту. Зокрема, розвивати:

- Здатність спілкуватися іноземною мовою.
- Здатність аналізувати та використовувати статистичні методи аналізу даних.

### **Objectives of study:**

Acquisition of knowledge, skills and abilities (competencies) based on the use of a systems approach and its principles for creating intelligent business systems, based on modern information and computer technologies, models and methods of artificial intelligence. In particular, to develop:

- Ability to communicate in a foreign language.
- Ability to analyze and use statistical methods for data analysis.

## 5 Результати навчання за дисципліною

| Результат навчання (РН)<br>(1 – знати; 2 – вміти; 3. комунікація; 4. автономність та відповідальність) |                                                                                                                                                                                            | Форми викладання та навчання                                                           | Методи оцінювання                                                                        | Відсоток у підсумковій оцінці з дисципліни |
|--------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|--------------------------------------------|
| Код                                                                                                    | Результат навчання                                                                                                                                                                         |                                                                                        |                                                                                          |                                            |
| РН1.1                                                                                                  | Знати і розуміти основні методи побудови DataSet з використанням різних методів відборів /<br>Know and understand the basic methods of building a DataSet using different sampling methods | Лекції, лабораторні заняття, самостійна робота /<br>Lecture, Lab.work, Individual work | Поточне оцінювання (ПО) контрольна робота, іспит /<br>Current assessment (PO) test, exam | 40%                                        |
| РН2.1                                                                                                  | Вміти будувати DataSet з використанням різних методів відборів. /<br>Be able to build a DataSet using different sampling methods.                                                          |                                                                                        |                                                                                          |                                            |
| РН3.1                                                                                                  | Демонструвати навички взаємодії з іншими людьми, уміння працювати в командах /<br>Demonstrate skills of interaction with other people, ability to work in teams                            | Лекції, лабораторні заняття, самостійна робота /<br>Lecture, Lab.work, Individual work | Поточне оцінювання (ПО) контрольна робота /<br>Current assessment (PO) test              |                                            |
| РН.4.1                                                                                                 | Виявляти здатність до самонавчання та продовження професійного розвитку /<br>Demonstrate the ability to self-study and continue professional development                                   | Лекції, лабораторні заняття, самостійна робота /<br>Lecture, Lab.work, Individual work | Поточне оцінювання (ПО) контрольна робота, іспит /<br>Current assessment (PO) test, exam | 50%                                        |
| РН.4.2                                                                                                 | Уміти організувати власну діяльність та одержувати результат у рамках обмеженого часу /<br>Be able to organize their own activities and get results within a limited time                  | Лекції, лабораторні заняття, самостійна робота /<br>Lecture, Lab.work, Individual work | Поточне оцінювання (ПО) контрольна робота /<br>Current assessment (PO) test              | 10%                                        |

## 6 Співвідношення результатів навчання дисципліни з програмними результатами навчання

| Результати навчання дисципліни                                    |  | РН1.1 | РН2.1 | РН3.1 | РН4.1 | РН4.2 |
|-------------------------------------------------------------------|--|-------|-------|-------|-------|-------|
| Програмні результати навчання                                     |  |       |       |       |       |       |
| (з опису освітньої програми)                                      |  |       |       |       |       |       |
| ПРН19.2. Демонструвати знання статистичних методів аналізу даних. |  | +     | +     | +     | +     | +     |

## 7 Схема формування оцінки/ Evaluation scheme

## 7.1 Форми оцінювання студентів: / Forms of evaluation:

### - семестрове оцінювання: / semester evaluation:

1. Контрольна робота / Test (PH1.1, PH2.1, PH3.1, PH4.2): 30 балів (points) / 18 балів (points).
2. Поточне оцінювання / Current assessment (PH1.1, PH2.1, PH3.1, PH4.2): 30 балів (points) / 18 балів (points).

### Підсумкове оцінювання / final evaluation: іспит /exam:

- максимальна кількість балів які можуть бути отримані студентом: 40;
- результати навчання, які оцінюються: PH1.1, PH2.1, PH.4.1;
- форма проведення: письмова робота
- види завдань: два теоретичні питання (40%), дві задачі (60%).

Студент допускається до іспиту, якщо в семестрі набрав не менше ніж 36 балів. Для отримання загальної позитивної оцінки з дисципліни оцінка за екзамен має бути не менше 24 балів.

- the maximum number of points that can be obtained by a student: 40 points;
- learning outcomes that are evaluated: PH1.1, PH2.1, PH.4.1;
- form of conducting and types of tasks: written work;
- types of tasks: two theoretical questions (40%), two tasks (60%).

A student is admitted to the exam if he scored at least 36 points in the semester. To receive an overall positive grade in the discipline, the grade for the exam must be at least 24 points.

## 7.2 Організація оцінювання

Терміни проведення оцінювання

Контрольна № 1 – до 7 тижня.

Поточне оцінювання – протягом семестру.

За відсутності студента з поважних причин перездача екзамену здійснюється відповідно до «Положення про порядок оцінювання знань студентів при кредитно-модульній системі організації навчального процесу» від 1 жовтня 2010 року.

## 7.2 Assessment

Terms of assessment

Test № 1 - up to 7 weeks.

Formative assessment - during the semester.

In the absence of a student for valid reasons, the reexamination is carried out in accordance with the "Regulations on the procedure for assessing students' knowledge in the credit-module system of the educational process" from October 1, 2010.

## 7.3 Шкала відповідності оцінок / Rating scale

|                                  |        |
|----------------------------------|--------|
| <b>Відмінно</b> / Excellent      | 90-100 |
| <b>Добре</b> / Good              | 75-89  |
| <b>Задовільно</b> / Satisfactory | 60-74  |
| <b>Незадовільно</b> / Fail       | 0-59   |

## 8. Структура навчальної дисципліни. Тематичний план лекцій і лабораторних занять /

The structure of the discipline. Thematic plan of lectures and laboratory classes

| №                   | Назва лекції                                                                                                                                                                                                                                                    | Кількість годин |                  |                   |
|---------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|------------------|-------------------|
|                     |                                                                                                                                                                                                                                                                 | Лекції          | Лабора-<br>Торні | Самост.<br>Робота |
| Змістовий модуль 1. |                                                                                                                                                                                                                                                                 |                 |                  |                   |
| 1                   | <b>Тема 1.</b> Вступ. Основні поняття. Попередня обробка даних. / Introduction. Basic concepts. Data processing.<br><i>Самостійна робота/ individual work:</i> Вступ. Основні поняття. Попередня обробка даних. / Introduction. Basic concepts. Data processing | 4               | 2                | 15                |
| 2                   | <b>Тема 2.</b> Методи візуалізації даних за допомогою мови програмування R/ Data visualization with a help of R<br><i>Самостійна робота/ individual work:</i> Мова програмування R/ Program language R                                                          | 4               | 4                | 15                |
| 3                   | <b>Тема 3.</b> Regression analysis/ Regression analysis.<br><i>Самостійна робота/ individual work:</i> Regression analysis/ Regression analysis.                                                                                                                | 8               | 4                | 15                |
| 4                   | <b>Тема 4.</b> Логістичні моделі/ Logistic models.<br><i>Самостійна робота/ individual work:</i> Логістичні моделі/ Logistic models.                                                                                                                            | 4               | 2                | 15                |
| 5                   | <b>Тема 5.</b> Кореляційний аналіз/ Correlation analysis<br><i>Самостійна робота individual work:</i> Кореляційний аналіз/ Correlation analysis                                                                                                                 | 4               | 2                | 15                |
| 6                   | <b>Тема 6.</b> Дисперсійний аналіз/ Analysis of variance (ANOVA)<br><i>Самостійна робота/ individual work:</i> Дисперсійний аналіз/ Analysis of variance (ANOVA)                                                                                                | 6               | 2                | 15                |
| 7                   | <b>Тема 7.</b> Аналіз часових рядів/analysis of time series<br><i>Самостійна робота / individual work:</i> Аналіз часових рядів/analysis of time series                                                                                                         | 6               | 2                | 15                |
|                     | Контрольна робота                                                                                                                                                                                                                                               | 2               | 2                | 15                |
| Іспит               |                                                                                                                                                                                                                                                                 |                 |                  |                   |
| ВСЬОГО              |                                                                                                                                                                                                                                                                 | 38              | 20               | 120               |

Загальний обсяг 180 год.<sup>1</sup>, в тому числі:

Лекцій – **38 год.**

Лабораторні – **20 год.**

Самостійна робота - **120 год.**

<sup>1</sup> Загальна кількість годин, відведених на дану дисципліну згідно навчального плану.

## 9. Рекомендовані джерела / Literature

### Основні:

- 1) Майборода Р.Є. Комп'ютерна статистика – професійний старт, 2018, 482 ст.
- 2) Майборода Р.Є., Сугакова О.В. Аналіз даних за допомогою пакета R.  
[http://matphys.rpd.univ.kiev.ua/downloads/courses/mmatstat/Statistics\\_with\\_R.pdf](http://matphys.rpd.univ.kiev.ua/downloads/courses/mmatstat/Statistics_with_R.pdf)
- 3) J. Maindonald, W. John Braun, Data Analysis and Graphics Using R – an Example-Based Approach, Cambridge Univ. Press.(2010).
- 4) P. Dalgaard, Introductory Statistics with R, 2-nd edition, Springer (2008).
- 5) W. Hardle, L. Simar, Applied Multivariate Statistical Analysis, second edition, Springer (2007).
- 6) G. James, D. Witten, T. Hastie, R. Tibshirani, An Introduction to Statistical Learning with Applications in R, Springer (2013).
- 7) R.H. Shumway and D.S. Stoffer, Time Series Analysis and Its Applications. With R Examples. 2nd edition. Springer (2006).

### Додаткові:

- 1) Айвазян С.А., Енюков И.С., Мешалкин Л.Д., Прикладная статистика. Основы моделирования и первичная обработка данных. М.: Финансы и статистика, 1983.
- 2) Лагутин М.Б., Наглядная математическая статистика, М.: БИНОМ, 2007.
- 3) Бриллинджер, Д., Временные ряды. Обработка данных и теория. (1980)
- 4) Майборода Р.Є., Регресія: лінійні моделі. ВПЦ “Київський університет”, 296 р. – 2007.  
<http://probability.univ.kiev.ua/userfiles/mre/ora0.pdf>