



# АНАЛІЗ ФІНАНСОВО-ЕКОНОМІЧНИХ ДАНИХ

## Робоча програма навчальної дисципліни (Силабус)

### Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Перший (бакалаврський)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>124 Системний аналіз</i>
Освітня програма	<i>Системний аналіз і управління</i>
Статус дисципліни	<i>Вибіркова</i>
Форма навчання	<i>очна (денна)</i>
Рік підготовки, семестр	<i>4 курс, осінній семестр</i>
Обсяг дисципліни	<i>4 кредити ЄКТС, 120 год. (лекції – 36 годин, практичні заняття – 18 годин, самостійна робота – 66 годин)</i>
Семестровий контроль/ контрольні заходи	<i>Залік /МКР</i>
Розклад занять	<i>Rozklad.kpi.ua</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	Лектор: <i>д.т.н., доцент, Кузнєцова Наталія Володимирівна, n.kuznietsova@kpi.ua</i> Практичні: <i>д.т.н., доцент, Кузнєцова Наталія Володимирівна</i>
Розміщення курсу	<i>Googleclassroom</i>

### Програма навчальної дисципліни

#### 1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

Дисципліна «Аналіз фінансово-економічних даних» є необхідним елементом практичної та професійної підготовки студентів для працевлаштування у фінансовому секторі. Предмет навчальної дисципліни – вивчення особливостей отримання, обробки, оцінювання фінансово-економічних даних сучасними засобами інтелектуального аналізу даних. Метою навчальної дисципліни є формування у студентів вміння збирати та коректно обробляти статистичні фінансово-економічні дані; будувати математичні і статистичні моделі різних видів; формувати та перевіряти сформовані гіпотези щодо розподілу фінансово-економічних даних; оцінювати параметри математичних моделей; застосовувати інформаційні технології ІАД до реальних фінансово-економічних задач; моделювати та розв'язувати практичні задачі

фінансово-економічного характеру. Студенти після засвоєння навчальної дисципліни мають продемонструвати такі результати навчання: знання методів збору та коректної обробки статистичних фінансово-економічних даних, уміння побудови математичних і статистичних моделей, перевірки гіпотез і оцінювання параметрів у таких моделях, набуття практичного досвіду застосування сучасних засобів інтелектуального аналізу даних до реальних фінансово-економічних задач.

У процесі навчання студент має оволодіти такими **компетентностями**: здатність застосовувати знання в практичних ситуаціях; здатність абстрактно мислити, застосовувати методи аналізу і синтезу; здатність працювати в команді та автономно виконувати командні рішення; здатність формалізувати проблеми, описані природною мовою, у тому числі за допомогою математичних методів, застосовувати загальні підходи до математичного моделювання конкретних процесів та аналізу даних; здатність будувати математично коректні моделі статичних та динамічних процесів і систем із зосередженими та розподіленими параметрами із врахуванням невизначеності зовнішніх та внутрішніх факторів; здатність визначати основні чинники, які впливають на розвиток фізичних, економічних, соціальних процесів, відокремлювати в них стохастичні та невизначені показники, формулювати їх у вигляді випадкових або нечітких величин, векторів, процесів та досліджувати залежність між ними; 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. По завершенню курсу студент має набути наступні **програмні результати навчання**: застосовувати методи і засоби роботи з даними і знаннями, методи математичного, логіко-семантичного, об'єктного та імітаційного моделювання, технології системного і статистичного аналізу; розуміти і застосовувати на практиці методи статистичного моделювання і прогнозування, оцінювати та аналізувати вихідні дані.

## **2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)**

Дисципліна «Аналіз фінансово-економічних даних» базується на результатах навчання з дисциплін: «Теорія ймовірностей» (ПОБ), «Математична статистика» (ПО23), і своєю чергою матеріал даної дисципліни тісно пов'язаний та використовується у курсах «Основи системного аналізу», «Статистичний аналіз економічних процесів», «Аналіз часових рядів». Отримані в процесі освоєння курсу практичні навички допоможуть бакалаврам при проведенні досліджень, проведення експериментів при підготовці бакалаврської роботи та при проведенні подальших наукових досліджень і роботі студентів у фінансовому секторі.

## **3. Зміст навчальної дисципліни**

Навчальна дисципліна містить 1 кредитний модуль: «Аналіз фінансово-економічних даних»

### РОЗДІЛ 1 Основні поняття статистичного аналізу фінансово-економічних даних

Тема 1.1. Системи управління знаннями та інструменти інтелектуального аналізу даних.

Тема 1.2. Класифікація фінансових процесів та види фінансово-економічних даних.

Тема 1.3. Основні етапи статистичного аналізу даних. Проблеми аналізу фінансово-економічних даних.

### РОЗДІЛ 2 Мережі Байєса як інструмент інтелектуального аналізу даних

Тема 2.1. Основні визначення та поняття мереж Байєса. Методи побудови та навчання МБ.

Тема 2.2. Точні та наближені методи формування висновку у мережах Байєса.

Тема 2.3. Гібридні МБ в задачах аналізу фінансово-економічних даних.

### РОЗДІЛ 3 Древа рішень в задачах аналізу фінансових даних

Тема 3.1. Основні поняття та критерії дерев рішень. Древа рішень у задачах класифікації та прогнозування. Методи побудови дерев рішень.

Тема 3.2. Методи CHAID, CART, QUEST, C4.5 для дерев рішень.

### РОЗДІЛ 4 Регресійні моделі для аналізу фінансово-економічних даних

Тема 4.1. Множинна та нелінійна регресія для аналізу фінансових даних.

Тема 4.2. Бінарна логістична регресія для аналізу кредитних ризиків.

### РОЗДІЛ 5 Застосування кластерного аналізу для аналізу фінансово-економічних даних.

Тема 5.1. Ієрархічний кластерний аналіз. Методи визначення відстані між об'єктами. Метрики для оцінки відстані.

Тема 5.2. Швидкий кластерний аналіз. Метод *k*-середніх. Модель найближчого сусіду для аналізу дефолту позичальника.

### РОЗДІЛ 6 Сучасні засоби інтелектуального аналізу даних.

Тема 6.1. Технології *data-mining*. Системи СППР. Класи задач аналізу даних.

Тема 6.2. Сучасні лідери-інструменти ІАД для аналізу фінансових даних: системи SPSS та SAS Enterprise Miner.

## **4. Навчальні матеріали та ресурси**

Базова:

1. Згуровський М.З. Байєсівські мережі в системах підтримки прийняття рішень /Згуровський М.З., Бідюк П.І., Терент'єв О.М., Просянкіна-Жарова Т.І. — Київ, ТОВ «Едельвейс», 2015. — 300 с.
2. Кузнєцова Н.В., Бідюк П.І. Теорія і практика аналізу фінансових ризиків: системний підхід: монографія / Київ: Видавництво Ліра-К, 2020. 400 с.
3. Бідюк П.І. Моделі і методи прикладної статистики/ Бідюк П.І., Коршевнік Л.О., Кузнєцова Н.В.// — Київ: НТУУ «КПІ», 2014. — 722 с.
4. Леоненко М.М., Мішура Ю.С., Пархоменко В.М., Ядренко М.Й. Теоретико-ймовірнісні та статистичні методи в економетриці та фінансовій математиці. — К.: Інформтехніка, 1995. — 380 с.
5. Андрєєв М. В. Лекції з байєсової статистики. Послідовні методи у статистиці та оптимізації стохастичних моделей: навч. посіб. / Андрєєв М.В. — К.: Акад. праці і соц. відносин Федер. проф. спілок України, 2009. — 376 с.

Додаткова:

6. N. R. Draper, H Smith, *Applied Regression Analysis, Third Edition*. Wiley-Interscience, 1998, 736 p. DOI:10.1002/9781118625590.
7. Bühl A., Zöfel P. *Erweiterte Datenanalyse mit SPSS*. Springer Fachmedien Wiesbaden 2002, 396p. <https://doi.org/10.1007/978-3-322-89603-2>.
8. Sen, S. Yildirim, I. A Tutorial on How to Conduct Meta-Analysis with IBM SPSS Statistics. *Psych* 2022, 4, 640-667. <https://doi.org/10.3390/psych4040049>

9. Бідюк П. І. Основні етапи побудови і приклади застосування мереж Байєса / П. І. Бідюк, Н. В. Кузнєцова// Системні дослідження та інформаційні технології. – 2007. – № 4. – С. 26–39.
10. Tuffery S. Data mining and statistics for decision making. – New York: John Wiley & Sons, Inc., 2011. – 704 p.

## Навчальний контент

### 5. Методика опанування навчальної дисципліни(освітнього компонента)

#### Лекційні заняття

№	Назва теми лекції та перелік основних питань
1	Поняття data mining, knowledge management. Системи управління знаннями та інструменти інтелектуального аналізу даних. Магічний квадрант інтелектуальних платформ. Рекомендована література: [1,2,7, 8, 10,11].
2	Типи випадкових процесів. Приклади. Класифікація видів даних. Статистичні дані. Проблеми аналізу фінансово-економічних даних. Рекомендована література: [2 – 4].
3	Основні етапи статистичного аналізу даних. Планування експерименту та збір даних. Попередня обробка даних, формування гіпотез. Оцінювання параметрів моделей. Перевірка сформованих гіпотез та прийняття рішень. Література: [3,4].
4	Основні визначення та поняття мереж Байєса. Однозв'язні та двозв'язні мережі. Дерева та полідерева. Теорема Байєса, передача повідомлень в МБ. Ланцюгове правило та маржиналізація в МБ. Взаємна інформація та функція опису мінімальної довжини. Література: [1-5].
5	Лекція 5. Методи побудови МБ на основі оціночних функцій. Методи із застосуванням тесту на умовну незалежність. Евристичний метод побудови МБ. Навчання мереж Байєса. Література: [3, 5, 9].
6	Точні та наближені методи формування висновку в мережах Байєса. Метод кластеризації клік-дерева. Метод виключення змінної. Алгоритм Перла розповсюдження повідомлення для однозв'язних мереж (полідерев). Метод ймовірного висновку на основі навчальних даних. Наближені методи формування висновку – методи спроб. Література: [3, 5, 9].
7	Гібридні МБ в задачах аналізу фінансово-економічних даних. Умовні лінійні гаусіани. Методи формування висновку. Алгоритм калібрування сильно кореневого дерева. Література: [3, 9].
8	Методи дискретизації. Динамічно-рівномірна дискретизація. Застосування дискретизації до фінансово-економічних даних. Інструментарій для побудови та застосування МБ – GENIE 2.0 та NETICA. Література: [3, 9].
9	Основні поняття дерев рішень. Критерій щеплення, механізм відсікання гілок. Дерева рішень у задачах класифікації та прогнозування. Методи побудови дерев рішень. Зупинка побудови дерева рішень. Література: [2, 7].
10	Метод CART для побудови бінарного дерева. Індекс GINI. Методи CHAID та Exhaustive CHAID. Особливості алгоритмів QUEST, C4.5 для дерев рішень Література: [2, 7].

11	МКР-1.
12	Застосування множинної та нелінійної регресії для аналізу фінансових даних. Приклади. Література: [3, 6, 7].
13	Бінарна логістична регресія. Мультиноміальна логістична регресія. Логіт- та пробіт-моделі для аналізу кредитних ризиків. Методи Enter, прогресивної селекції та оберненої селекції для визначення параметрів моделі. Література: [2,6, 7].
14	Ієрархічний кластерний аналіз. Методи визначення відстані між об'єктами. Метрики для оцінки відстані. Література: [2, 7].
15	Швидкий кластерний аналіз. Метод k-середніх. Модель найближчого сусіду для аналізу дефолту позичальника. Література: [2, 8].
16	Технології data-mining. Системи СППР. Класи задач аналізу даних. Література: [1, 2, 3,7,8].
17	Сучасні лідери-інструменти ІАД для аналізу фінансових даних: системи SPSS та SAS Enterprise Miner [2].
18	Залік

### Практичні заняття

Основні завдання циклу практичних занять – напрацювання практичних навичок збору та обробки фінансових даних, побудови фінансово-економічних моделей, освоєння сучасного інструментарію інтелектуального аналізу даних.

№	Назва теми занять
1	Збір, очистка та формування вибірок фінансово-економічних даних.
2	Застосування мереж Байєса до аналізу фінансово-економічних даних. Робота з інструментарієм Genie 2.0.
3	Аналіз кредитоспроможності юридичних осіб за допомогою мереж Байєса. Знайомство з інструментарієм Netica.
4	Дерева рішень в задачах прогнозування фінансово-економічних даних.
5	Дерева рішень у задачах скорингу.
6	Логістична регресія в задачах аналізу фінансово-економічних даних. Бінарна логістична регресія для прогнозування дискретного атрибуту в фінансових задачах.
7	Практичне заняття № 7. Моделювання кредитних ризиків на основі моделей Кокса та пропорційних ризиків.
8	Контрольна робота за темами «Мережі Байєса», «Дерева рішень» та «Регресійні моделі».

9	Особливості застосування інструментарію інтелектуального аналізу даних SPSS та SAS Enterprise Miner для розв'язання задач аналізу фінансово-економічних даних.
---	--

Методика освоєння тем даної дисципліни така: (1) – лекційні заняття; (2) – самостійне вивчення та аналіз лекційного матеріалу; (3) – опрацювання матеріалу за підручником у відповідності до плану занять; (4) – виконання вправ та практичних робіт для самостійного вивчення, спрямованих на закріплення теоретичного матеріалу; (5) – опрацювання матеріалу за додатковою спеціальною літературою (слайди, статті, монографії).

Використання комп'ютерних технологій при вивченні дисципліни: (1) – на протязі вивчення курсу викладач надає додаткові матеріали в електронній формі – ілюстрації, презентації, статті, приклади розрахунків з використанням фактичних даних і т. ін.; (2) – студенти активно використовують існуюче програмне забезпечення: SPSS, SAS, Econometric Views, GeNIe 2.0, Netica, Matlab, Matcad, STATISTICA, що сприяє швидкому засвоєнню матеріалу та набуванню практичних навичок застосування різних методів та комп'ютерних продуктів для аналізу і оцінювання ризиків різної природи.

## 6. Самостійна робота студента/аспіранта

Студентам пропонуються індивідуальні семестрові завдання у вигляді підготовки рефератів та/або домашніх контрольних робіт. Індивідуальні завдання сприяють розвиненню у студентів творчих навичок та практичних вмінь, отриманих за кредитним модулем «Аналіз фінансово-економічних даних» для постановки і розв'язання власних практичних фінансово-економічних задач. Студентам надається проект для виконання, який містить три задачі на основі вхідних даних та і з постановкою, заданою викладачем, а також варіативна частина: на вибір напрям та задача прикладного характеру, яку необхідно розв'язати за допомогою методів та засобів інтелектуального аналізу даних.

Напрями завдань варіативної частини:

- засоби розв'язання задачі кредитного скорингу;
- прогнозування обсягів продажів компанії;
- аналіз попиту на продукцію кіноіндустрії;
- прогнозування обсягів депозитів та потенціальних клієнтів-вкладників банку;
- оцінювання попиту компанії для прогнозування логістичних витрат;
- прогнозування обсягів благодійних внесків та сегментів-благодійників;
- аналіз дзвінків мобільного оператора та напрацювання рекомендацій стосовно нових тарифних пакетів;
- прогнозування грошової бази на основі даних про грошові агрегати.

Проект сприяє поглибленому засвоєнню методів аналізу фінансово-економічних даних, і допомагає отримати студентам практичні навички збору і підготовки, а також попередньої обробки і інтелектуального аналізу даних, характерних фінансовій сфері. Методичні рекомендації до виконання проекту, варіанти завдань, термін виконання надає лектор всім групам потоку і зазначає у гугл-класі, роботи перевіряють та виставляються як рейтингові бали через гугл-клас.



## 7. Політика навчальної дисципліни (освітнього компонента)

Під час занять (як лекцій, так і практичних) студентам надаються інтерактивні вправи у вигляді задач, ігор, оцінюється як активність, так і успішність освоєння матеріалу (в тому числі під час виконання домашніх завдань і підготовки до лекцій), призначаються заохочувальні бали.

Всі завдання з описом відповідних етапів їх виконання, а також відповідних контрольних строків завантажені у Гугл-класі дисципліни. Усі роботи студенти мають завантажувати в особистому кабінеті гугл-класу. Дедлайни кожного завдання позначені в завданнях у гугл-класі. Захисти індивідуальних робіт відбуваються на практичних заняттях. Роботи мають бути виконані з дотриманням академічної доброчесності. У період роботи в дистанційному режимі лектор може запропонувати студентам пройти запропоновані ним онлайн-курси на платформі Coursera. Також сертифікати/проміжні результати цих курсів можуть бути частково зараховані згідно до [Положення](#).

## 8. Види контролю та рейтингова система оцінювання результатів навчання (PCO)

Поточний контроль: експрес-опитування, опитування за темою заняття, МКР, тест.

Календарний контроль: проводиться двічі на семестр як моніторинг поточного стану виконання вимог силабусу.

Семестровий контроль: залік

Умовою першого семестрового контролю є поточний рейтинг не менше 50% запланованих балів. Умова другого семестрового контролю є поточний рейтинг не менше 50% запланованих балів та зарахування командного проекту.

Рейтинг студента з дисципліни за семестр виставляється лектором та складається з балів, що він отримує за:

- виконання індивідуального завдання (проект);
- домашні роботи;
- виконання контрольних робіт;
- розв'язання задач і роботи на практичних і лекційних заняттях.

### 1. Критерії нарахування балів.

1) Контрольна робота (МКР-1 оцінюється у 20 балів, МКР-2 у 25 балів).

2) Розв'язання практичних задач, що оцінюється у 15 балів.

3) Проект складається з 4 етапів, кожен етап послідовно доповнює один одного. Другий- четвертий етапи є командною роботою. Вирішується задача аналізу фінансово-економічних даних реального проекту, що включає постановку задачі, розв'язання, презентацію і захист роботи, і оцінюється у 40 балів.

Етап проекту і роботу не зараховано (завдання не виконане або є грубі помилки) – 0 балів. Невчасна здача проекту оцінюється -3 бали за кожний тиждень запізнення.

Додаткові індивідуальні задачі і робота на практичних і лекційних заняттях додатково оцінюється у 10 балів.

Залікова контрольна робота оцінюється у 40 балів.

Сума рейтингових балів, отриманих студентом протягом семестру, переводиться до підсумкової оцінки згідно з таблицею. Якщо сума балів менш ніж 60, студент виконує залікову контрольну роботу. У цьому разі сума балів за виконання проекту залікову контрольну роботу переводиться до підсумкової оцінки згідно з таблицею.

2. Студент, який у семестрі отримав не менш ніж 60 балів, може прийняти участь у заліковій контрольній роботі. У цьому разі, бали отримані ним на контрольній роботі є остаточними.

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

<p><i>Бали:</i>  <i>Практики + МКР + проект</i>  <i>Або</i>  <i>Практики+ Coursera + проект</i>  <i>або</i>  <i>Залік: проект + Залікова контрольна робота</i></p>	<p><i>Оцінка</i></p>
100...95	Відмінно
94...85	Дуже добре
84...75	Добре
74...65	Задовільно
64...60	Достатньо
Менше 60	Незадовільно
РР не зараховано або менше 30	Не допущено

## 9. Додаткова інформація з дисципліни (освітнього компонента)

Теоретичні питання:

1. Поняття «знання» та «управління знаннями».
2. Етапи та особливості статистичного аналізу даних (в загальному).
3. Програмні продукти для аналізу даних. Особливості застосування.
4. Визначення мереж Байєса та основні етапи побудови і застосування.
5. Дані: визначення, види, типи і категорії. Поняття метаданих. Приклади.
6. Перший та другий етапи статистичного аналізу даних: планування та попередня обробка даних.
7. Критерії та характеристики якості моделей.
8. Поняття умовної незалежності та маржиналізації в мережах Байєса.
9. Попередня обробка даних: статистичні параметри та етапи.
10. Показники якості моделей: специфічність, чутливість тощо.
11. Методи побудови мереж Байєса.
12. Поняття Data Mining. Приклади інтелектуальних платформ.
13. Методи оцінювання параметрів статистичних і математичних моделей.
14. Поняття мереж Байєса.
15. Поняття взаємної інформації в мережах Байєса. Навести приклад методу, що її застосовує.
16. Поняття статистичних моделей. Приклади.
17. Перевірка сформульованих гіпотез в статистичному аналізі даних.



18. Однозв'язні та багатозв'язні мережі.
19. Навчання мереж Байєсе. Особливості та відмінності в залежності від структури та спостережуваності.
20. Поняття статистичних даних. Приклади.
21. Основні проблеми аналізу фінансових даних.
22. Поняття дерев та полідерев.
23. Загальна ідея та особливості EM-алгоритму.
24. Інструменти інтелектуального аналізу даних: Statistica та Deductor.
25. Випадкові процеси. Моделювання часових рядів за допомогою некорельованих випадкових величин.
26. Статистичний аналіз розподілу логарифмів цін та волатильності.
27. Модифікації EM-алгоритму для МБ.
28. Наближені методи формування висновку у мережах Байєса.
29. Марківські ланцюги та гібридні мережі Байєса.
30. Практичні задачі класифікації клієнтів та маркетингових досліджень за допомогою дерев рішень.
31. Методи Exhaustive CHAID, C4.5 для дерев рішень.
32. Множинна та нелінійна регресія для аналізу фінансових даних. Пробіт-моделі аналізу кредитних ризиків.
33. Кластерний аналіз: метрики, методи, модифікації. Приклади задач.
34. Сховища даних, OLAP-системи та інформаційні СППР в фінансових установах. Особливості data mining в таких системах. Інтеграція та особливості роботи з фінансовими даними систем SPSS та SAS Enterprise Miner.

*У період роботи в дистанційному режимі лектор може запропонувати студентам пройти такі онлайн-курси на платформі Coursera: Probabilistic Graphical Courses -*

*<https://www.coursera.org/specializations/probabilistic-graphical-models#courses> спеціалізована серія з 3х курсів:*

*а) <https://www.coursera.org/learn/probabilistic-graphical-models?specialization=probabilistic-graphical-models>*

*б) <https://www.coursera.org/learn/probabilistic-graphical-models-2-inference>*

*в) <https://www.coursera.org/learn/probabilistic-graphical-models-3-learning#instructors>, (кожен 15 балів)*

*Machine Learning <https://www.coursera.org/learn/machine-learning?action=enroll#enroll> (10 балів) і*

*Econometrics: Methods and Applications (Erasmus University Rotterdam)*

*<https://www.coursera.org/learn/erasmus-econometrics/home/welcome> (20 балів). Сертифікати, а також*

*проміжні (контрольні тестові) результати цих курсів можуть бути частково зараховані згідно до Положення.*

### **Робочу програму навчальної дисципліни (силабус):**

**Склала** професор кафедри математичних методів системного аналізу, д.т.н., доцент, Кузнєцова Наталія Володимирівна

**Ухвалено** кафедрою ММСА (протокол № 13 від 05.06.2024)

**Погоджено** Методичною комісією НН ІПСА (протокол № 10 від 24.06.2024 )