



БАГАТОВИМІРНИЙ СТАТИСТИЧНИЙ АНАЛІЗ

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Другий (магістерський)</i>
Галузь знань	12 Інформаційні технології
Спеціальність	122 Комп'ютерні науки 124 Системний аналіз
Освітня програма	Системи і методи штучного інтелекту Системний аналіз і управління
Статус дисципліни	Вибіркова
Форма навчання	очна(денна)
Рік підготовки, семестр	III курс, весняний семестр
Обсяг дисципліни	120 годин / 4 кредити ЄКТС (лекції – 36 год., комп'ютерний практикум – 18 год., СРС – 66 год.)
Семестровий контроль/ контрольні заходи	залік, МКР
Розклад занять	https://schedule.kpi.ua/
Мова викладання	Українська
Інформація про керівника курсу / викладачів	Лекції та семінари проводять: канд. техн .наук, доцент, , доцент кафедри ММСА Джигирей Ірина Миколаївна, lab.mes@kpi.ua
Розміщення курсу	https://do.ipu.kpi.ua/course/view.php?id=5890

Програма навчальної дисципліни

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

Багатовимірний статистичний аналіз вивчає оброблення багатомірних статистичних даних з метою виявлення взаємозв'язків, їх характеру та структури. Суть багатофакторного аналізу полягає в одночасному розгляданні кількох пов'язаних змінних, причому кожна з них вважається однаково важливою. Так **аналіз головних компонент** призначено для зменшення кількості змінних, які необхідно враховувати, до невеликої кількості індексів (так званих головних компонент), які є лінійними комбінаціями початкових змінних. Таким чином, аналіз головних компонент є засобом спрощення даних шляхом зменшення кількості змінних. **Факторний аналіз** – це також спроба врахувати зміни низки початкових змінних, використовуючи меншу кількість індексних змінних або факторів. Передбачається, що кожен початкову змінну може бути виражено як лінійну комбінацію таких факторів. **Дискримінантний аналіз** пов'язано з проблемою визначення можливості виокремлення різних груп на основі наявних даних. Як і аналіз головних компонент, дискримінантний аналіз базується на ідеї пошуку відповідних лінійних комбінацій початкових змінних. **Кластерний аналіз** пов'язано з ідентифікацією груп подібних об'єктів. **Багатовимірне шкалювання** починається з даних про певну міру відстаней між кількома об'єктами. На основі цих відстаней будується карта, яка показує, як об'єкти пов'язано.

Метою опанування дисципліни є формування відповідного рівня знань і досвіду в оперуванні основними принципами, поняттями і методами багатовимірного статистичного аналізу в

контексті оцінювання стану різномасштабних соціо-економічних систем, прогнозування їх розвитку.

Предметом навчальної дисципліни є методичні та інструментальні рішення у сфері прикладної статистики для аналізу структури і тенденцій розвитку багатоозначових явищ, процесів і систем для підтримки ухвалення обґрунтованих рішень, що є необхідним складником формування майбутнього фахівця.

Дисципліна сприяє формуванню у студентів таких компетентностей:

- здатність до абстрактного мислення, аналізу та синтезу;
- здатність застосовувати знання у практичних ситуаціях;
- здатність приймати обґрунтовані рішення;
- здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач;
- 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.

Після засвоєння навчальної дисципліни студенти мають продемонструвати такі результати навчання.

ЗНАННЯ:

- теоретичних основ прикладної статистики;
- сучасних задач прикладної математичної статистики та методів їх розв'язування;
- основні принципи багатовимірного аналізу даних;
- методів багатовимірної класифікації та зниження розмірності, зокрема, кластерного аналізу, дискримінантного аналізу та факторного аналізу.

УМІННЯ:

- проводити багатовимірну класифікацію об'єктів та використовувати методи зниження розмірності простору ознак об'єктів і систем, що аналізують;
- підтримувати ухвалення обґрунтованих рішень на основі багатовимірного аналізу даних для розв'язання багаторівневих прикладних задач в економіці, соціології, сфері сталого розвитку і міждисциплінарних досліджень;
- використовувати знання закономірностей випадкових явищ, їх властивостей та операцій над ними, моделей випадкових процесів та сучасних програмних середовищ для розв'язування задач статистичної обробки даних і побудови прогнозних моделей;
- застосовувати методи і засоби роботи з даними і знаннями, методи математичного, логіко-семантичного, об'єктного та імітаційного моделювання, технології системного і статичного аналізу;
- збирати необхідну інформацію, використовуючи науково-технічну літературу, бази даних та інші джерела, аналізувати і оцінювати її.

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

Вивчення дисципліни базується на знанні студентами основних понять математики, дисциплін «Математичний аналіз»; «Теорія ймовірностей»; «Математична статистика».

Дисципліна забезпечує такі навчальні дисципліни у програмі підготовки фахівця як «Теорія прийняття рішень» і «Основи системного аналізу» та її тісно пов'язано з таким компонентом освітньої програми як дипломне проектування.

3. Зміст навчальної дисципліни

Тема 1 Основні поняття багатовимірного статистичного аналізу

Тема 2 Багатовимірні класифікації

Тема 3 Методи зниження розмірності простору ознак

Тема 4 Теорія індексів

Тема 5 Багатовимірний статистичний аналіз в FL-дослідженнях

Тема 6 Багатовимірний статистичний аналіз в оцінюванні вимірів та компонент розвитку суспільства

4. Навчальні матеріали та ресурси

Базова література

1. Бізнес-аналітика багатовимірних процесів : навчальний посібник [Електронний ресурс] / Т. С. Клебанова, Л. С. Гур'янова, Л. О. Чаговець та ін. – Харків : ХНЕУ ім. С. Кузнеця, 2018. – 272 с. <http://repository.hneu.edu.ua/handle/123456789/22020>
2. Григорук, Павло Михайлович. Багатомірне економіко-статистичне моделювання : навчальний посібник / Григорук П.М. - Львів : Видавництво "Новий Світ-2000", 2018. - 147 с. (НТБ ім. ім. Г. І. Денисенка)
3. Спеціальні розділи математики. Статистичний аналіз даних у середовищі STATISTICA [Електронний ресурс] : навч. посіб. для студ. спеціальності 151 «Автоматизація та комп'ютерно-інтегровані технології», спеціалізації «Комп'ютерно-інтегровані технології сталих хімічних виробництв» / КПІ ім. Ігоря Сікорського ; уклад.: І. М. Джигирей, Д. М. Складанний. – Електронні текстові данні (1 файл: 1,41 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2019. – 74 с. <https://ela.kpi.ua/bitstream/123456789/28228/1/2019srm2.pdf>
4. А. Т. Яровий, Є. М. Страхов. Багатовимірний статистичний аналіз : навчально-методичний посібник для студентів математичних та економічних фахів. – Одеса: Астропринт, 2015. – 132 с. http://ouek.onu.edu.ua/uploads/courses/msa/Yarovoy_Strakhov_MSA.pdf

Додаткова література

(факультативно / ознайомлення)

5. Гороховатський, Володимир Олексійович. Методи інтелектуального аналізу та оброблення даних : навчальний посібник / В.О. Гороховатський, І.С. Творошенко ; Міністерство освіти і науки України, Харківський національний університет радіоелектроніки. - Харків : ХНУРЕ, 2021. - 90 с. (НТБ ім. ім. Г. І. Денисенка)
6. Мармоза, Анатолій Тимофійович. Практикум з математичної статистики : навчальний посібник для студентів вищих навчальних закладів / А.Т. Мармоза. - Київ : Кондор, 2021. - 257 с. (НТБ ім. ім. Г. І. Денисенка)
7. Основи статистичного моделювання: навч. посібник / за загальною редакцією С.В. Чугаєвської, Н.В. Ковтун. Житомир: Видавництво ПП "Рута", 2022. – 604 с. (бібліотека дистанційного курсу дисципліни)
8. Аналіз сталого розвитку — глобальний і регіональний контексти / Міжнар. рада з науки (ISC) та ін.; наук. кер. проекту М. З. Згуровський. — К. : КПІ ім. Ігоря Сікорського, 2019. — Ч. 1. Глобальний аналіз якості і безпеки життя (2019). — 216 с. — URL: <http://wcd.org.ua/sites/default/files/SD2019-P1-FULL-UA.pdf>
9. Wolfgang Härdle and Léopold Simar, Applied multivariate statistical analysis. 5th edition. Berlin: Springer Nature Switzerland AG, 2019. (за запитом викладачу)
10. Krolovie Peter James. Factor Analysis: Excellent Guide With SPSS. Independently published, 2021. 404 p. (за запитом викладачу)
11. C. M. Hennig, Meilă M., Fionn Murtagh, and R. Rocci, Handbook of cluster analysis. Boca Raton: CRC Press, Taylor and Francis Group, 2017. (за запитом викладачу)
12. B. F. J. Manly and J. A. Navarro, Multivariate statistical methods: a primer. 4th Edition. Boca Raton: Chapman and Hall/CRC, 2017. (за запитом викладачу)
13. S. R. Terrell, Statistics translated: a step-by-step guide to analyzing and interpreting data. 2nd edition. New York: The Guilford Press, 2021. (за запитом викладачу)

5. Методика опанування навчальної дисципліни (освітнього компонента)

Навчальна дисципліна охоплює 36 годин лекцій та 18 годин комп'ютерного практикуму, а також виконання модульної контрольної роботи, яка охоплює три частини за темами навчальної дисципліни тривалістю 0,67 акад. год. кожна.

Комп'ютерний практикум з дисципліни проводять з метою закріплення теоретичних положень навчальної дисципліни «Багатовимірний статистичний аналіз» і набуття студентами умінь і досвіду їх практичного застосування під керівництвом викладача шляхом виконання відповідно сформульованих завдань. Виходячи з розподілу часу на вивчення дисципліни, рекомендується дев'ять семінарських занять (з врахуванням часу на модульну контрольну роботу та залік).

Термін виконання (тиждень)	Назви розділів і тем
Тема 1. Основні поняття багатовимірного статистичного аналізу	
1	Лекція 1. Основні поняття багатовимірного статистичного аналізу
1	Комп'ютерний практикум 1. Кореляційний аналіз у середовищі STATISTICA
2	Лекція 2. Графічний аналіз багатовимірних даних
3	Комп'ютерний практикум 2. Графічний аналіз даних у середовищі STATISTICA
Тема 2. Багатовимірна класифікація	
3	Лекція 3. Міри подібності
4	Лекція 4. Ієрархічні методи кластерного аналізу
5	Лекція 5. Ітераційні методи кластерного аналізу. Оцінка якості класифікації
5	Комп'ютерний практикум 3. Розрахунок мір подібності вживаних у кластерному аналізі
7	Комп'ютерний практикум 4. Кластерний аналіз у STATISTICA
6	Лекція 6. Основні поняття дискримінантного аналізу
7	Лекція 7. Задачі дискримінантного аналізу
9	Комп'ютерний практикум 5. Дискримінантний аналіз у STATISTICA
Модульна контрольна робота (частина I)	
Тема 3. Методи зниження розмірності простору ознак	
8	Лекція 8 Метод головних компонент
9	Лекція 9 Факторний аналіз
10	Лекція 10 Багатовимірне шкалювання
11	Комп'ютерний практикум 6. Факторний аналіз у STATISTICA
Тема 4. Теорія індексів	
11	Лекція 11 Методи нормалізування та агрегування даних в оцінюванні складних систем на основі індексів
12	Лекція 12 Агрегатні індекси у прикладах
13	Комп'ютерний практикум 7. Індексний метод
Модульна контрольна робота (частина II)	
Тема 5. Багатовимірний статистичний аналіз в FL-дослідженнях	
13	Лекція 13 Багатофакторне прогнозування
14	Лекція 14 Методи багатовимірного аналізу в форсайт-дослідженнях
15	Комп'ютерний практикум 8. FL-проект. Частина I
15	Лекція 15 Застосування методів штучного інтелекту для вирішення задач класифікування, зниження розмірності та прогнозування
Тема 6. Багатовимірний статистичний аналіз в оцінюванні вимірів та компонент розвитку суспільства	
16	Лекція 16 Застосування багатовимірного аналізу в соціології та економіці
17	Лекція 17 Багатовимірний аналіз в оцінюванні сталого розвитку

Термін виконання (тиждень)	Назви розділів і тем
18	Лекція 18 Новітні методи багатовимірного статистичного аналізу
17	Комп'ютерний практикум 9. FL-проект. Частина II Модульна контрольна робота (частина III)

6. Самостійна робота студента

Самостійна робота студента охоплює такі складники як підготування до поточних опитувань, комп'ютерних практикумів, модульної контрольної роботи та підготування електронних звітів до комп'ютерних практикумів у вказаний викладачем термін.

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Відвідування занять. Відсутність на аудиторному занятті не передбачає нарахування штрафних балів, оскільки фінальний рейтинговий бал студента формується виключно на основі оцінювання результатів навчання. Разом з тим, оцінювання результатів опанування питань лекційних занять оцінюватимуться під час аудиторних занять.

Пропущені контрольні заходи оцінювання. Кожен студент має право відпрацювати пропущені з поважної причини (лікарняний, мобільність тощо) заняття за рахунок самостійної роботи. Детальніше за посиланням: <https://kpi.ua/files/n3277.pdf>.

Процедура оскарження результатів контрольних заходів оцінювання. Студент може підняти будь-яке питання, яке стосується процедури контрольних заходів та очікувати, що воно буде розглянуто згідно із наперед визначеними процедурами. Студенти мають право аргументовано оскаржити результати контрольних заходів, пояснивши з яким критерієм не погоджуються відповідно до оціночного.

Календарний контроль проводиться з метою підвищення якості навчання студентів та моніторингу виконання студентом вимог силабусу.

Критерій		Перший календарний контроль	Другий календарний контроль
Термін календарного контролю		Тиждень 8	Тиждень 14
Умови отримання позитивної оцінки	Поточний рейтинг	≥ 10 балів	≥ 30 балів

Академічна доброчесність. Політика та принципи академічної доброчесності визначені у розділі 3 Кодексу честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Детальніше: <https://kpi.ua/code>.

Норми етичної поведінки. Норми етичної поведінки студентів і працівників визначені у розділі 2 Кодексу честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Детальніше: <https://kpi.ua/code>.

Інклюзивне навчання. Засвоєння знань та умінь в ході вивчення дисципліни «Багатовимірний статистичний аналіз» може бути доступним для більшості осіб з особливими освітніми потребами, окрім здобувачів з серйозними вадами зору, які не дають змогу виконувати завдання за допомогою персональних комп'ютерів, ноутбуків та/або інших технічних засобів.

Навчання іноземною мовою. У ході виконання завдань студентам може бути рекомендовано звернутися до англomовних джерел.

Призначення заохочувальних та штрафних балів. Відповідно до Положення про систему оцінювання результатів навчання сума всіх заохочувальних балів не може перевищувати 10% рейтингової шкали оцінювання.

Заохочувальні бали		Штрафні бали	
Критерій	Ваговий бал	Критерій	Ваговий бал
Написання тез, статті, оформлення курсової роботи як наукової роботи для участі у конкурсі студентських наукових робіт (за тематикою навчальної дисципліни)	5-10 балів	Несвоєчасне здавання електронного звіту з виконання завдань комп'ютерного практикуму	-2 бали
Участь у міжнародних, всеукраїнських та/або інших заходах та/або конкурсах (за тематикою навчальної дисципліни)	5-10 балів	-	-

Підготування до семінарських занять та контрольних заходів здійснюється під час самостійної роботи студентів з можливістю консультування з викладачем у визначений час консультацій або за допомогою електронного листування (електронна пошта, месенджери).

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

Семестровий контроль проводиться у вигляді заліку. Для оцінювання результатів навчання застосовується 100-бальна рейтингова система та університетська шкала.

Поточний контроль: фронтальні опитування, електронне звітування, МКР.

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

Семестровий контроль: залік

Якщо семестровий рейтинг більше 60 балів студент може не виходити на залікову контрольну роботу, а отримати оцінку «автоматом». Модульна контрольна робота. Кожна з трьох частин модульної контрольної роботи містить 15 комплексних питань тестового, розрахункового або відкритого типу, які оцінюються в один бал. За правильну відповідь на питання студент отримує 1 бал, неправильну – 0 балів. Фронтальне опитування – тестування за вмістом лекційних занять, кожне з яких оцінюється у 1 бал.

№ з/п	Контрольний захід оцінювання	%	Ваговий бал	Кількість	Разом
1.	Електронне звітування за результатами виконання завдань комп'ютерних практикумів	40%	5	8	40
2.	Фронтальне опитування	15%	1	15	15
3.	Модульна контрольна робота	45%	45	1	45
	Разом				100

Для отримання заліку з навчальної дисципліни «автоматом» потрібно мати рейтинг не менше 60 балів, а також зараховані вісім електронних звітів за результатами виконання завдань комп'ютерних практикумів. Студенти, які наприкінці семестру мають рейтинг менше 60 балів за умови зарахування електронних звітів з комп'ютерних практикумів, а також ті, хто бажає підвищити оцінку, виконують залікову контрольну роботу. Доступні дві опції складання залікової контрольної роботи за вибором студента.

Опція 1. Залікова контрольна робота виконується на платформі дистанційного навчання протягом 2 академічних годин та містить 100 закритих тестових і відкритих запитань різного рівня складності з ваговими балами від 0,5 до 2, сума яких становить 100 балів.

Опція 2. Письмова залікова контрольна робота, білети якої містять шість питань теоретичного, системного і розрахунково-аналітичного характеру за кожною з тем навчальної дисципліни, виконується протягом 2 академічних годин. Перші чотири питання оцінюються в 17 балів, останні два – 16 балів: «відмінно», творче, системне і повне розкриття питання, вільне володіння матеріалом – 16-17 (15-16)^{1*} балів; «дуже добре», розкриття питання, вільне володіння матеріалом – 14-15 (13-14) балів; «добре», достатнє розкриття питання, володіння матеріалом – 12-13 (11-12) балів; «задовільно», обґрунтоване розкриття питання, неповне володіння матеріалом – 11 (10) балів; «достатньо», часткове розкриття питання – 10 (9) балів.

¹ для двох останніх питань

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

<i>Кількість балів</i>	<i>Оцінка</i>
100-95	Відмінно
94-85	Дуже добре
84-75	Добре
74-65	Задовільно
64-60	Достатньо
Менше 60	Незадовільно

9. Додаткова інформація з дисципліни (освітнього компонента)

Перелік питань, які виносяться на семестровий контроль представлено у Додатку А.

Методи та форми навчання охоплюють не лише традиційні університетські лекції та практичні заняття, а також застосовуються стратегії активного навчання, які визначаються такими методами та технологіями: методи проблемного навчання (дослідницький метод); особистісно-орієнтовані технології, засновані на таких формах і методах навчання як кейс-технологія і проектна технологія; візуалізація та інформаційно-комунікаційні технології, зокрема електронні презентації для лекційних занять. Комунікація з викладачем будується за допомогою використання інформаційної системи «Електронний кампус», платформи дистанційного навчання «Сікорський», а також такими інструментами комунікації, як електронна пошта, месенджер Telegram.

Робочу програму навчальної дисципліни (силабус):

Складено:

*доцент кафедри штучного інтелекту, канд. техн .наук, доцент,
Джигирей Ірина Миколаївна*

Ухвалено кафедрою штучного інтелекту (протокол № 14 від "11" червня 2024 р.)

Погоджено Методичною комісією ННІПСА (протокол № 10 від "24" червня 2024 р.)