

УНІВЕРСИТЕТ ІМЕНІ АЛЬФРЕДА НОБЕЛЯ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Методичні вказівки
щодо виконання курсової роботи

з дисципліни **«Інтелектуальний аналіз даних»**

для спеціальності
122 «Комп'ютерні науки»

бакалаврського рівня вищої освіти

Затверджено на засіданні кафедри
Протокол №1 від 30.08.22 р.

м. Дніпро
2022

Методичні вказівки щодо виконання курсової роботи з дисципліни «Інтелектуальний аналіз даних» /Ю.М. Барташевська. – Дніпро: Університет ім. Альфреда Нобеля, 2022. – 14 с.

Укладач: Ю.М. Барташевська, к.е.н., доц.

Відповідальний за випуск: Ю.М. Барташевська, к.е.н., доцент, зав. кафедри інформаційних технологій.

ЗМІСТ

1. Мета та основні завдання курсової роботи	3
2. Обсяг та структура курсової роботи	4
3. Оформлення курсової роботи	5
4. Основні етапи підготовки курсової роботи, відповідні рекомендації для студентів	7
5. Строки виконання курсової роботи	8
6. Захист та порядок оцінювання курсової роботи	8
7. Тематика курсових робіт	9
Додатки	12

1. Мета та основні завдання курсової роботи

Курсове проектування є завершальним етапом учбового процесу з дисципліни «Інтелектуальний аналіз даних».

Головною **метою** курсового проектування є поглиблення теоретичних та практичних навичок щодо методів сучасної обробки даних (Data Mining), а саме: ознайомлення та практичного застосування методів та алгоритмів Data Mining, Text Mining, WebMining; пошуку у необроблених масивах даних знань та закономірностей, необхідних для прийняття обґрунтованих рішень.

Основними **завданнями** курсового проектування є:

- поглиблення і систематизація отриманих знань з методів Data Mining та Text Mining;
- отримання навичок збору і узагальнення теоретичного і практичного матеріалу відповідно до обраної теми;
- вміння виконувати обробку практичної інформації за допомогою R, R Studio, Tableau.

Відповідно до освітньо-професійної програми за спеціальністю 122 «Комп'ютерні науки», курсове проектування з дисципліни «Інтелектуальний аналіз даних» має забезпечувати наступні результати навчання:

Шифр	Результати навчання
PH1	Застосовувати знання основних форм і законів абстрактно-логічного мислення, основ методології наукового пізнання, форм і методів вилучення, аналізу, обробки та синтезу інформації в предметній області комп'ютерних наук
PH3	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.
PH4	Використовувати методи обчислювального інтелекту, машинного навчання, нейромережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв'язання задач розпізнавання, прогнозування, класифікації, ідентифікації об'єктів керування тощо.
PH12	Застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах

	класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірного аналізу даних на основі технологій DataMining, TextMining, WebMining.
--	--

2. Обсяг та структура курсової роботи

Загальний обсяг курсової роботи повинен складати 25 – 30 сторінок (формат А4) комп'ютерного набору (через 1,5 інтервал, розмір шрифту 14). Абзацний відступ 1,25.

Текст повинен мати поля з 4-х : верхнє, ліве і нижнє - 20 мм, а праве 10 мм.

Перша сторінка роботи оформляється як титульна (Додаток 1).

На другій сторінці подається план (зміст) роботи із зазначенням номера сторінки – початку кожного структурного елемента роботи.

Усі сторінки тексту повинні бути пронумеровані, починаючи з другої сторінки (першою вважається титульна сторінка).

Курсова робота з дисципліни "Інтелектуальний аналіз даних (Data Analysis and Data mining)" складається з наступних розділів:

ВСТУП

РОЗДІЛ 1 Теоретична частина

РОЗДІЛ 2 Практична частина

ЗАГАЛЬНІ ВИСНОВКИ

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

ДОДАТКИ

Примірна кількість розділів – 2 - 4.

1,5-2 сторінки покликаний познайомити з суттю досліджуваної теми. У актуальність теми, формулюються мету роботи і її предмет, характеризуються використані автором матеріали.

Основна частина курсової роботи послідовно, відповідно до змісту і складається з розділів/підрозділів. Кожен розділ починають з нової сторінки.

Теоретична частина курсової роботи містить методологічні аспекти рішення по

наданій темі, наводяться дані, що підтверджують певні висновки. При цьому автор повинен робити посилання на джерело тих даних, які він використовує в роботі. В цьому розділі вказуються також існуючі методи щодо вирішення даної проблеми, їх порівняння.

Практична (проектна) частина курсової роботи складається з

- постановки задачі, опису вхідних даних;
- опису алгоритму рішення задачі за допомогою R, R Studio, Tableau. Можливе наведення частин програмного коду з описом, скріншоти. Обов'язково навести пояснення отриманого результату.

Висновок. У висновку коротко відбиваються результати виконаної курсової роботи, робляться висновки і даються рекомендації про можливість використання виконаних розробок.

Література містить перелік матеріалів, використовуваних при курсовому проектуванні, дається наприкінці курсового проекту перед додатком. У перелік включаються статті, монографії, методичні вказівки.

Додатки оформляються як продовження курсової роботи і розміщуються в порядку посилань в тексті.

3. Оформлення курсової роботи

Курсова робота, як правило, виконується українською мовою.

Текст друкують за допомогою комп'ютера на одному боці аркуша білого паперу формату А-4 через 1,5 міжрядкових (комп'ютерних) інтервали, шрифт Times New Roman, 14.

Абзацний відступ – 1,25. Текст повинен мати поля з 4-х : верхнє, ліве і нижнє - 20 мм, а праве 10 мм.

Текст основної частини поділяють на розділи та підрозділи.

Заголовки структурних частин „ЗМІСТ”, „ВСТУП”, „РОЗДІЛ”, „ВИСНОВКИ”, „ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ (ЛІТЕРАТУРИ)”, „ДОДАТКИ” друкують

великими літерами симетрично до тексту.

Заголовки підрозділів друкують маленькими літерами (крім першої великої) з абзацного відступу.

Крапку в кінці заголовка не ставлять. Якщо заголовок складається з двох чи більше речень, їх розділяють крапкою. Перенесення слів в заголовок розділу не допускається. Відстань між заголовком і текстом повинна дорівнювати 1-2 інтервалам.

Кожен **розділ** курсової роботи слід починати з нової сторінки.

Додатки слід позначати послідовно великими буквами української абетки, за винятком Є, І Ї, Й, Ч, Ь, наприклад, Додаток А, Додаток Б і т.д.

Ілюстрації позначають словом «Рис. » і нумерують послідовно в межах розділу, за винятком малюнків, поданих в додатках.

Таблиці нумерують послідовно (за винятком тих, які розміщені в додатках) в межах розділу. У правому верхньому кутку розміщують напис «Таблиця» з вказівкою її номера, який складається з номера розділу і порядкового номера таблиці, між якими ставиться крапка: наприклад, «Таблиця 2.3» (третя таблиця другого розділу), нижче розміщується назва таблиці.

У разі перенесення таблиці на іншу сторінку над її частинами пишеться: наприклад, «Продовження табл. 2.3».

Формули нумеруються в межах розділу. Номер формули складається з номера розділу і порядкового номера формули в розділі, між якими ставлять крапку.

Перелік використаних джерел будують у порядку появи посилань у тексті або в алфавітному порядку.

Вимоги до оформлення посилань:

а) посилання на книгу:

1. Брюс П., Брюс Э. Практическая статистика для специалистов Data Science: 50 важнейших понятий. СПб: БХВ, 2018. 304 с.
2. Новиков Л.А. Художественный текст и его анализ / Л.А. Новиков. – М.: Изд-во ЛКИ, 2009. – 304 с.

б) посилання на статтю:

1. Борисов Л.А., Орлов Ю.Н., Осминин К.П. Идентификация автора текста по распределению частот буквосочетаний/ Л.А. Борисов, Ю.Н. Орлов, К.П. Осминин // Препринты ИПМ им. М.В. Келдыша. - 2013. - № 27. - 26 с.

в) посилання на електронні джерела інформації:

1. Соціально-економічне становище України. – [Електронний ресурс]. – Режим доступу: <http://www.ukrstat.gov.ua>

г) посилання на нормативні документи:

1. Інформаційні технології. Процеси життєвого циклу програмного забезпечення: ДСТУ 3918-1999 (ISO/IEC 12207:1995). – К.: Державний комітет стандартизації, метрології і сертифікації України, 2000. – 37 с.

4. Основні етапи підготовки курсової роботи, відповідні рекомендації для студентів

Курсова робота повинна бути виконана не пізніше як за два тижні до початку сесії, оформлена відповідно до вимог та подана на перевірку керівнику. Науково-педагогічні працівники – керівники робіт забезпечують перевірку курсових робіт здобувачів на наявність запозичень шляхом застосування ресурсів, які знаходяться у вільному доступі та рекомендовані до використання. Перевірка проводиться під час рецензування, яке здійснюється протягом 5 (п'яти) календарних днів з дня реєстрації поданої роботи на кафедрі (з дня розміщення роботи у гугл-класрум). Роботи вважаються оригінальними за наявності не більше 20% запозичень. Роботи, оригінальність яких становить менше 80%, або які містять запозичення без відповідних посилань, повертаються на доопрацювання. Роботи, у яких міститься понад 60% запозичень до розгляду не приймаються. Результати оформлюються у

вигляді протоколу перевірки на наявність текстових запозичень та зберігаються на кафедрі разом з роботами.

5. Строки виконання курсової роботи

На початку виконання роботи студенту необхідно ознайомитися з основною літературою за обраною темою і скласти план в аспекті основних розділів і параграфів. План має бути узгоджений з керівником роботи. В ході роботи перелік літератури доповнюється, а план роботи коригується і деталізується.

Виконання роботи здійснюється студентом за планом-графіком.

Керівник контролює хід виконання роботи. Виконані розділи подаються керівникові і відповідно до його зауважень уточнюються, доповнюються чи переробляються.

Після підготовки розділів курсової роботи та ознайомлення з ними керівника, студент згідно із зауваженнями та вказівками доопрацьовує розділи роботи. Завершена робота подається на підпис керівникові не пізніше ніж за 10 днів до початку сесії.

6. Захист та порядок оцінювання курсової роботи

Захист курсової роботи проводиться індивідуально до здачі заліку. Як правило, студент захищає роботу перед науковим керівником.

При підготовці до захисту студентові необхідно виконати всі вказівки, дані в рецензії, врахувати зауваження в тексті роботи і заздалегідь відповісти на поставлені питання.

На захисті курсової роботи студент повинен бути готовий до короткого викладу основного змісту роботи і її результатів, до співбесіди по окремих моментах роботи, до відповіді на будь-які питання як по даній темі, так і по всьому курсу. Головне завдання захисту – виявити дійсну ступінь самостійності студента у виконанні курсової роботи та глибину його розуміння тих питань, які викладені у курсовій роботі.

За рецензії і захисту курсової роботи виставляється оцінка. При незадовільній оцінці студент зобов'язаний повторно виконати роботу по новій темі або переробити колишню. Повторний захист робіт повинен завершитися до початку сесії. Студенти,

що не здали і не захистили в строк курсову роботу, до сесії не допускаються.

Керівник курсової роботи оцінює роботу за 100-бальною шкалою, спираючись на наступну схему:

1. *Оформлення роботи* – максимально 20 балів. Оцінюється ступінь відповідності оформлення курсової роботи встановленим вимогам (обсяг та структура роботи, оформлення титульного листа, рисунків, таблиць, діаграм, посилань, списку літератури тощо).
2. *Своєчасність подання роботи керівнику* – максимально 40 балів (при повторному представленні роботи після доопрацювання – максимум 20 балів).
3. *Оцінка захисту* – максимально 40 балів. Під час захисту керівник роботи в ході співбесіди оцінює ступінь самостійності виконання роботи студентом, розуміння ним теоретичних та практичних аспектів обраної теми.

Переведення результатів 100-бальної шкали в 4-бальну здійснюється відповідно до Положення про поточний та підсумковий контроль знань студентів відповідно до Європейської кредитно-трансферної системи (ЄКТС).

7. Тематика курсових робіт

1. Кластеризація областей України за рівнем соціально-економічного розвитку (3 роки).
2. Статистичні та нейромережеві методи дослідження ціноутворення на ринку нерухомості міста/ області.
3. Порівняльний аналіз статистичних та нейромережевих методів прогнозування (на прикладі мережевих атак).
4. Аналіз думок і настроїв в текстах.
5. Використання алгоритмів Text Analytics для аналізу популярності Youtube відео.
6. Аналіз відгуків готелів з Інтернет сайтів.
7. Застосування алгоритмів аналізу тексту для прогнозування очікувань клієнтів.
8. Лінгвістична текстова аналітика для вивчення досвіду клієнтів (аналіз книг з IT-тематики).

9. Аналіз та прогнозування продажів медичних страховок.
10. Аналіз та прогнозування популярності Інтернет-курсів.
11. Аналіз та прогнозування Kickstarter проєктів.
12. Аналіз результатів виборів в Україні та побудова прогнозу.
13. Аналіз сільськогосподарського виробництва за регіонами України.
14. Аналіз та прогнозування викидів в атмосферу.
15. Аналіз доступності дитячих садочків в Україні.
16. Аналіз рівня зайнятості в Україні (за сімейним положенням та полом).
17. Аналіз рівня освіти населення за рівнем освіти, віком і статтю.
18. Аналіз послуг з перевезення вантажів залізничним транспортом за видами транспорту.
19. Аналіз та прогнозування транспортування нафти по трубопроводах.
20. Аналіз та прогнозування тенденцій у транспорті
21. Застосування елементів аналізу текстів для визначення емоцій користувачів Twitter (<https://www.kaggle.com/parulpandey/emotion-dataset>).
22. Аналіз кібератак в Інтернеті.
23. Аналіз та прогнозування землетрусів
24. Аналіз (сегментація) даних клієнтів торговельної мережі.
25. Аналіз та прогнозування попиту на бронювання номерів в отелях.
26. Аналіз та прогнозування Інтернет-користування.
27. Аналіз популярності ретро або б/у авто
28. Аналіз та прогнозування популярності ІТ-напрямків в Україні.
29. Аналіз та прогнозування рівня з/п ІТ-фахівців в Україні.
30. Аналіз популярності фільмів на платформах Netflix, Prime Video, Hulu и Disney +.
31. Застосування алгоритмів Text Analytics для моніторингу якості послуг.
32. Аналіз відповідей відкритого опитування для кластеризації преференцій споживачів.
33. Часовий аналіз полеміки навколо теми в засобах масової інформації.

34. Часовий аналіз тематичного розвитку обраної галузі науки.
35. Аналіз та прогнозування розвитку туристичної галузі України.
36. Аналіз популярності мобільних додатків як елемент утримання майбутніх користувачів.
37. Аналіз мобільних транзакцій з метою визначення та запобігання шахрайства.
38. Аналіз та прогнозування плинності кадрів організації
39. Аналіз даних щодо визначення та запобігання корупції країн світу
40. Застосування алгоритмів аналізу тексту для прогнозування очікувань клієнтів.
41. Лінгвістична текстова аналітика для вивчення досвіду клієнтів (аналіз книг з ІТ-тематики).

ЗРАЗОК ОФОРМЛЕННЯ ТИТУЛЬНОГО ЛИСТА

УНІВЕРСИТЕТ ІМЕНІ АЛЬФРЕДА НОБЕЛЯ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

КУРСОВА РОБОТА

з дисципліни

«Інтелектуальний аналіз даних»

на тему _____

Здобувача _____ (П.І.Б.)

Групи _____ (шифр групи)

Керівник курсової роботи _____

(П.І.Б., науковий ступень, вчене звання)

Дніпро
2022

Навчальне видання

Методичні вказівки щодо виконання курсової роботи з дисципліни
«Інтелектуальний аналіз даних»

Підписано до друку . Формат 60 X 84/16. Ум. друк. арк. 0,69

Оперативна поліграфія. Зам. № . Тираж прим.

Університет імені Альфреда Нобеля.

49000, м. Дніпро, вул. Січеславська Набережна, 18.

