МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

**МЕТОДИЧНІ РЕКОМЕНДАЦІЇ**

для виконання курсової роботи

з дисципліни

**ВСТУП ДО МАШИННОГО НАВЧАННЯ**

для студентів 123 Комп’ютерні науки

Галузь знань: 12 «Інформаційні технології»

Напрями підготовки: 123 Комп’ютерні науки

Укладач:

Полуектова Н.Р.. , д.е.н., доцент

|  |  |
| --- | --- |
|  | Розглянуто на засіданні кафедри КН  Протокол № від . . 2024 р. |
|  |  |

Запоріжжя, 2024

ЗМІСТ

|  |  |
| --- | --- |
| 1. Мета й значення курсової роботи в навчальному процесі | 3 |
| 2. Вибір теми курсової роботи | 3 |
| 3. Основні етапи виконаний курсової роботи | 4 |
| 4. Визначення набору даних для аналізу | 4 |
| 5. Структура, зміст і обсяг курсової роботи | 5 |
| 6. Порядок написання тексту курсової роботи | 7 |
| 7. Правила оформлення курсової роботи | 8 |
| 8. Захист курсової роботи | 16 |
| 9. Основні умови одержання оцінки | 17 |
| Додаток А | 20 |
| Додаток Б | 22 |
| Додаток В | 23 |
| Додаток D | 24 |

1. **Мета й значення курсової роботи в навчальному процесі**

Курсова робота є важливою формою самостійного вивчення студентами дисципліни „ ВСТУП ДО МАШИННОГО НАВЧАННЯ”. Її виконання сприяє поглибленому вивченню курсу на основі наукових праць вітчизняних та зарубіжних ІТ-фахівців, з використанням сучасного інструментарію машинного навчання на основі використання мови Python та її спеціалізованих бібліотек. При цьому студенти набувають навичок пошуку необхідних літературних джерел та матеріалів, їх аналізу та узагальнення, досвіду самостійного дослідження та письмового викладання теоретичних питань, поєднують методи творчого застосування знань з курсу та вміння пов'язувати їх з практикою програмування коду за допомогою фреймворків.

Крім того, у процесі виконання курсової роботи студенти набувають навичок, необхідних для написання кваліфікаційної бакалавської роботи, а також вмінь, які можуть бути використані в майбутній практичній роботі та наукових дослідженнях.

**2. Вибір теми курсової роботи**

Курсова робота виконується державною мовою на одному з обраних датасетів, які погоджені з науковим керівником, якого кафедра визначає кожному студенту.

Кожний студент вибирає тему за власним бажанням в межах своїх інтересів та залежно від можливостей знайти статистичний та фактичний матеріал за обраною темою. Тема курсової може бути відповідною темі науково-дослідної роботи студента.

Однак, слід звернути увагу на те, що ***в межах однієї групи теми курсових робіт не повинні повторюватися*.**

Науковий керівник курсової роботи остаточно погоджує тему, проект плану (який може бути удосконалений в процесі виконання роботи) та термін виконання. Роботи, подані на кафедру пізніше встановленого строку оцінюються (за умов виконання всіх вимог та поставлених в роботі задач) сумарною кількістю балів не вище необхідного мінімуму.

Тематика курсових робіт наводиться у додатку А.

1. **Основні етапи виконаний курсової роботи**
   * + 1. Вивчення літературних джерел.
       2. Складання плану та структури курсової роботи.
       3. Підбір та обробка статистичних даних та інших фактичних матеріалів.
       4. Написання тексту роботи.
       5. Попередня перевірка роботи викладачем.
       6. Оформлення курсової роботи.
       7. Захист курсової роботи.
2. **Визначення проблематики предметної галузі**

Мета роботи — отримати навички застосування одного або декількох сучасних методів машинного навчання для розв’язування прогнозної або класифікаційної задачі Внаслідок вивчення проблематики предметної галузі розширюються та поглиблюються теоретичні знання, розвивається логічне мислення, виробляється критичне відношення до різних точок зору на питання, що вивчаються.

У загальному вигляді здійснити представлення проблемної предметної галузі, визначаючи інструментарій, що описують предметну галузь, загальний характер взаємозв‘язків між ними, основні джерела вихідної аналітичної інформації, переваги використання обраного фреймворку, можливість практичної складової обраної теми.

**5. Структура, зміст і обсяг курсової роботи**

Курсова робота з дисципліни „ВСТУП ДО МАШИННОГО НАВЧАННЯ” як індивідуальне теоретично-прикладне дослідження повинна мати певну логічну побудову, послідовність і завершеність.

Обсяг основної частини курсової роботи має буди близько 20-40 сторінок, на яких розкривається зміст теми, яку досліджує студент. Мова курсової роботи – державна, стиль – науковий, чіткий, без орфографічних і синтаксичних помилок.

Курсова робота повинна містити:

1. титульний аркуш;
2. завдання на курсову роботу
3. зміст;
4. вступ;
5. 2-3 розділи основної частини (можливо з підрозділами);
6. висновки;
7. список використаної літератури;
8. додатки (при необхідності).

Зміст або план роботи може займати 1 сторінку. В ньому записуються назви всіх розділів та підрозділів із зазначенням початкових сторінок. Зміст роботи має відображати сутність проблеми, її складність і логіку відображення. Назви розділів та підрозділів мають бути стислими і зрозумілими, літературно грамотними, тісно пов’язаними з назвою роботи, але не повторювати її.

У вступі (1-3 с.) дається наукове обґрунтування актуальності і значення вибраної теми; формулюється мета, завдання, об’єкт та предмет дослідження; наводиться перелік застосованих методів дослідження.

**Основна частина** курсової роботи умовно повинна ділитися на 2 складові: загальнотеоретичну та практичну. Перша повинна містити теоретичний досвід стосовно тематики роботи, розкривати зміст основних методів та технологій на яких буде проводитися програмування у другій частині. Студент аналізує існуючі технології для вирішення поставленої мети, розкриває переваги та недоліки їх застосування.

Друга складова основної частини обов'язково має носити практичний характер. Вона повинна складатися покрокового викладення процесу побудови моделі машинного навчання. Кожен крок містить постановку завдання, скрин-шот реалізації та пояснення.

Обсяг розділів та підрозділів повинен бути пропорційним. Розділи повинні мати завершений характер та містити в кінці логічний перехід до наступного розділу.

Перший розділ (8-15 с.) структурно складається з 2-3 підрозділів. У ньому слід розкрити стан теоретичних напрацювань з вибраної теми, подати критичний огляд літературних джерел, викласти основні теоретичні та методичні положення теми, проблемні питання, дискусійні та невирішені аспекти теми.

Студент має продемонструвати свою обізнаність з методами аналізу і на основі їх критичного огляду обґрунтувати вибір технології проектування веб-додатку.

Теоретичні положення першого розділу мають бути покладені в основу другого і третього розділу.

У другому розділі (10-20 с.) студент повинен здійснити постановку задачі та представити розробку моделі машинного навчання.

У процесі постановки задачі необхідно:

- описати сутність задачі;

- описати набір даних.

- обгрунтувати вибір методів машинного навчання, які будуть використані;

- описати бібліотеки, які будуть використані.

Слід чітко розмежувати джерела походження використаної під час аналізу інформації: що запозичена з літератури, що одержано з документів організації, а які є результатом власних спостережень, експериментів, розрахунків.

Розділ “Висновки” є завершальною частиною курсової роботи. Він містить стислий виклад актуальності теми, зроблених студентом оцінок та узагальнень під час аналізу та написання курсової роботи.

**6. Порядок написання тексту курсової роботи**

Після ретельного вивчення літературних джерел та фактичних даних, а також складання плану, необхідно приступити до написання тексту курсової роботи. Текст курсової роботи повинен бути чітко розподілений за питаннями плану. Кожному питанню плану повинен відповідати розділ тексту курсової роботи під такою ж назвою. Кожний розділ повинен починатися з нової сторінки. Назва розділу чи підрозділу не може бути аналогічною з темою курсової роботи.

Текст курсової роботи має бути відредагований, без граматичних помилок. В курсовій роботі треба уникати багатослів'я, емоційних виразів, використання зайвих слів та переказ від першої особи. Весь текст курсової роботи треба висловлювати своїми словами. Повне копіювання не тільки цілих блоків тексту, а й окремих фраз з наявних економічних публікацій не дозволяється. ***При використанні цитат окремих авторів, а також статистичних та фактичних матеріалів, треба робити посилання на літературні джерела***, зазначаючи у квадратних дужках номер, під яким знаходиться це джерело у списку використаних літературних джерел, і після коми вказати сторінку.

**7. Правила оформлення курсової роботи**

Остаточний текст курсової роботи має бути надрукований на комп'ютері шрифт Times New Roman, розмір 14, інтервал – 1,5. Текст курсової роботи необхідно друкувати, залишаючи береги таких розмірів: лівий - не менше 30 мм, правий - не менше 10 мм, верхній - не менше 25 мм, нижній – не менше 20 мм. Абзац в тексті починається відступом, який дорівнює 1.25-1.5cм. Шрифт друку повинен бути чітким, стрічка - чорного кольору середньої жирності.

Закреслення, вставки, довільні скорочення слів (крім загальноприйнятих) і будь-які позначення **не дозволяються**. До написання остаточного тексту рекомендується надати чорновий варіант роботи на перевірку викладачеві.

Для написання остаточного тексту використовувати стандартні аркуші білого паперу, як правило, формату А4 (210x297 мм), складені або підшиті до папки. Можна також використовувати папір форматів у межах від 203x288 до 210x297 мм і подати таблиці та ілюстрації на аркушах формату АЗ. ***Усі сторінки роботи (крім титулу, завдання і змісту) мають бути пронумеровані.***

ЗАГОЛОВКИ

Заголовки структурних частин курсової роботи „ЗМІСТ”, „ВСТУП”, „РОЗДІЛ”, „ВИСНОВКИ”, „СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ” „ДОДАТКИ” друкують великими літерами симетрично до тексту.

Кожну структурну частину курсової роботи (крім підрозділів) треба починати з нової сторінки.

ТИТУЛ

Титульна сторінка курсової роботи оформляється за формою, яка наведена у додатку Б.

ЗАВДАННЯ НА КУРСОВУ РОБОТУ

Другою сторінкою (після титулу) є завдання на курсову роботу (не більше 1 стор.), що містить ціль, задачі, об‘єкт, предмет дослідження, ключові слова, приклад наведений у додатку В.

ЗМІСТ

Наступною сторінкою наводиться зміст курсової роботи. Сторінки, де він наводиться, не нумеруються. Зміст містить найменування та номери початкових сторінок всіх структурних частин курсової роботи, розділів та підрозділів.

ВСТУП

Слово „ВСТУП” пишеться великими літерами симетрично до тексту по центру сторінки. Далі через 2 інтервали - текст.

У вступі (1-3 сторінки) розглядається актуальність теми, її теоретичне та практичне значення, зв'язок з іншими проблемами моделювання та практики, мета наукового дослідження. У вступі бажано вказувати прізвища науковців (вітчизняних та іноземних), які внесли значний внесок у розвиток даної теми (принаймні 5 вчених). Також вкачують задачі та методи дослідження.

Починаючи зі вступу, здійснюється нумерація сторінок. Номери сторінок проставляються арабськими цифрами без крапки в верхньому правому кутку аркуша на відстані 5 мм від його країв. Перша сторінка вступу нумерується - 3, далі - по порядку до кінця роботи, включаючи додатки.

ОСНОВНА ЧАСТИНА

Основна частина складається з розділів та підрозділів. ***Кожний розділ починається з нової сторінки.*** Підрозділи з нової сторінки не починаються. Відстань між останнім рядком попереднього підрозділу і заголовком наступного підрозділу – 2-3 інтервали.

Номер розділу ставлять після слова “РОЗДІЛ” арабськими цифрами, після номера крапку не ставлять, потім з нового рядка друкують заголовок розділу великими літерами. Крапку в кінці заголовка не ставлять. Якщо заголовок складається з двох або більше речень, їх розділяють крапкою.

Підрозділи нумерують у межах кожного розділу. Номер підрозділу складається з номеру розділу і порядкового номера підрозділу, між якими ставлять крапку. В кінці номера підрозділу крапка не ставиться, наприклад: "2.3" (третій підрозділ другою розділу). Потім у тому ж рядку йде заголовок підрозділу. Заголовки підрозділів друкують маленькими літерами (крім першої великої) з абзацного відступу. Крапку в кінці заголовка не ставлять. Відстань між заголовком підрозділу та текстом повинна дорівнювати 2 інтервали. Відстань між заголовками розділу та підрозділу - один інтервал.

ІЛЮСТРАЦІЇ

Основними видами ілюстративного матеріалу курсової роботи є схема, діаграма і графік. Ілюстрації і таблиці необхідно подавати в курсовій роботі безпосередньо після тексту, де вони згадані вперше, або на наступній сторінці. Ілюстрації і таблиці, які розміщені на окремих сторінках курсової роботи, включають до загальної нумерації сторінок. Таблицю, малюнок або графіки, розміри яких більше формату А4, враховують як одну сторінку і розміщують у відповідних місцях після згадування в тексті або у додатках. Ілюстрації позначають словом "Рис." і нумерують послідовно в межах розділу, за виключенням ілюстрацій, поданих у додатках. ***Номер ілюстрації повинен складатися з номера розділу і порядкового номера ілюстрації, між якими ставиться крапка.*** Наприклад: Рис. 1.2 (другий рисунок першого розділу). Після номера рисунку ставиться тире і назву починають з великої літери. Номер ілюстрації, її назва і пояснювальні підписи розміщують під ілюстрацією. Якість ілюстрації повинна забезпечувати їх чітке відтворення. Вони повинні бути виконані чорнилом, тушшю або пастою чорного кольору на білому непрозорому папері. Ілюстрації повинні мати назву, яку розміщують після номера ілюстрації. При необхідності ілюстрації доповнюють пояснювальними даними (підрисунковий текст).

Ілюстрації обов'язково мають бути проаналізовані в тексті, при чому необхідно робити посилання у вигляді виразу в круглих дужках "(рис. 1.2)" або зворот типу: "...як це видно з рис. 1.2". У повторних посиланнях використовують вираз "див. рис. 1.2". До і після рисунку потрібно залишати один вільний рядок.

ТАБЛИЦІ

Таблиці нумерують послідовно (за винятком таблиць, поданих у додатках) в межах розділу. В правому верхньому куті таблиці розміщують напис "Таблиця" із зазначенням її номера. Номер таблиці повинен складатися з номера розділу і порядкового номера таблиці, між якими ставиться крапка, наприклад: "Таблиця 1.2" (друга таблиця першого розділу). Назва таблиці починається з великої літери з нової строчки. Таблицю розміщують після першого згадування про неї в тексті.

***Приклад побудови таблиці***

Таблиця (номер)

Назва таблиці (з абзацу)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | |  | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Заголовки

Підзаголовки  
 Заголовки рядків

Позначення одиниць величини, загальна для всіх даних в рядку, необхідно наводити після її назви.

При переносі частини таблиці на іншу сторінку пишуть слова ''Продовження табл." і вказують номер таблиці, наприклад: "Продовження табл. 1.2". При цьому легенду таблиці слід повторити. Цифровий матеріал, як правило, повинен оформлятися у вигляді таблиць.

Кожна таблиця повинна мати назву, яку розміщують над таблицею і друкують симетрично до тексту. Назву і слово "Таблиця" починають з великої літери. Назву не підкреслюють. Перед і після таблиці потрібно залишати один вільний рядок.

Примітки до тексту і таблиць, в яких вказуються довідкові і пояснювальні дані, нумерують послідовно в межах однієї сторінки.

ФОРМУЛИ

Формули в курсовій роботі (якщо їх більше однієї) нумерують у межах розділу. Номер формули складається з номера розділу і порядкового номера формули в розділі, між якими ставлять крапку. ***Номери формул пишуть по правому краю на рівні відповідної формули в круглих дужках***, наприклад: (3.1) (перша формула третього розділу). Пояснення значень символів і числових коефіцієнтів треба подавати безпосередньо під формулою в тій послідовності, в якій вони дані у формулі. Значення кожного символу і числового коефіцієнта треба подавати з нового рядка. Перший рядок пояснення починають зі слова "де" без двокрапки з абзацу. Рівняння і формули треба виділяти з тексту вільними рядками. Вище і нижче кожної формули потрібно залишити не менше одного вільного рядка. ***Формула друкується по центру сторінки***. Якщо рівняння не вміщується в один рядок, його слід перенести після знака рівності (=) або після знаків плюс (+), мінус (-), множення (\*) і ділення (/).

ПРИМІТКИ

Примітки до тексту і таблиць, в яких вказують довідкові і пояснювальні дані, нумерують послідовно в межах однієї сторінки. Якщо приміток на одному аркуші декілька, то після слова "примітка" ставлять двокрапку, наприклад:

Примітки:

1. …

2. …

Якщо є одна примітка, то її не нумерують і після слона "Примітка" ставлять тире і друкують її зміст. Примітки необхідно розміщувати після текстового, графічного матеріалу або після таблиць.

ПОСИЛАННЯ

При написанні курсової роботи студент повинен давати посилання на джерела та матеріали, окремі результати яких наводяться в роботі, або на ідеях і висновках яких розроблюються проблеми та питання, вивченню яких присвячена курсова робота. Такі посилання дають змогу відшукати документи і перевірити достовірність відомостей про цитування документа, дають необхідну інформацію щодо нього, допомагають з'ясувати його зміст.

Якщо використовують відомості, матеріали з монографій, оглядових статей, інших джерел з великою кількістю сторінок, тоді в посиланні необхідно точно вказати номери сторінок, ілюстрацій, таблиць, формул з джерела, на яке дано посилання в курсовій роботі.

Посилання в тексті курсової роботи на джерела слід позначити порядковим номером за списком використаної літератури, виділеним двома квадратними дужками, наприклад, “...у працях [1-7]...”. Якщо наводиться цитата, то поряд з номером джерела ставиться сторінка, з якої береться цитата [7, с.15].

Посилання на формули курсової роботи вказують порядковим номером формули в дужках, наприклад "... у формулі (2.1)". На всі таблиці курсової роботи повинні бути посилання в тексті, наприклад: “...в таблиці 1.2”. У повторних посиланнях на таблиці та ілюстрації треба вказувати скорочено слово "дивись", наприклад: "див. табл. 1.3".

При наведенні цифрового матеріалу в таблицях основного тексту або в додатках необхідно робити посилання на джерело (літературне або фінансовий документ підприємства), з якого бралися цифри.

висновки

У висновках (1-3 сторінки) викладаються короткі основні підсумки, що випливають із змісту роботи у відповідності до логіки роботи. Слово “ВИСНОВКИ” записують великими літерами симетрично до тексту по центру сторінки. Далі через два інтервали - текст.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

***Список використаної літератури*** ***повинен складатися з не менш, ніж 10-20 джерел***. Літературні джерела зазначаються у списку використаної літератури в алфавітному порядку прізвищ перших авторів чи заголовків книги (якщо книгу складено колективом авторів).

Перелік літературних джерел, що використовуються в курсовій роботі, необхідно робити за загальноприйнятою формою. Для книжок - це прізвище та ініціали автора, повна назва книги, том, частина, місце видання, видавництво, рік видання та кількість сторінок у книзі. Для журналів та газетних статей - це прізвище та ініціали автора, назва статі, назва журналу (газети), рік видання, місце видання або дата газети, сторінки.

Наприклад:

1. Веб-застосунок. Вікіпедія. [Електронний ресурс]. – Режим доступу: www. URL: https://uk.wikipedia.org/wiki/Вебзастосунок (дата звернення: 29.03.22)
2. Дохід українських фрілансерів на міжнародних ринках зріс на 20% [Електронний ресурс]. – Режим доступу: www. URL: https://biz.nv.ua/ukr/markets/hto-z-ukrajinciv-naybilshe-zaroblyaye-na-mizhnarodnih-rinkah-novini-ukrajini-50202014.html (дата звернення: 19.05.22)
3. Жизненный цикл компонента. [Електронний ресурс]. – Режим доступу: www. URL: https://metanit.com/web/react/2.6.php (Дата звернення: 11.05.22)
4. Основні бібліотеки та фреймворки JavaScript, які вам потрібно знати. [Електронний ресурс]. – Режим доступу: www. URL:https://code.tutsplus.com/ru/articles/essential-javascript-libraries-andframeworks-you-should-know-about--cms-29540 (дата звернення: 31.03.22)
5. Просторова організація бізнесу в регіонах України: форми та механізми регулювання. [Електронний ресурс]. – Режим доступу: www. URL: http://ird.gov.ua/irdp/p20190004.pdf (дата звернення:21.04.22) 6. Рейтинг мов програмування 2022. C# обійшов Java, TypeScript зрівнявся з PHP, а Dart — найбільш комфортна мова [Електронний ресурс]. – Режим доступу: www. URL: https://dou.ua/lenta/articles/language-rating-2022/ (дата звернення:13.05.22)
6. Рівень безробіття в Україні. Статистика трудових ресурсів України. [Електронний ресурс]. – Режим доступу: www. URL: https://index.minfin.com.ua/labour/unemploy/ (дата звернення: 03.05.22)
7. Карп’як, А.О., Карпінський, Б.А. та Федорчук, Н.В., Ціннісний підхід у системі функціонування ринку інформаційних технологій. Сучасні напрями розвитку суспільства: монографія. Харків: СГ НТМ «Новий курс», 2022. С. 169–179.

8. Гафіяк, А.М., Бородіна, О.О. та Альошин, С.П. Система формування професійної компетентності майбутніх ІТ-фахівців у закладах вищої освіти. *Інфраструктура ринку*, 2019. с.359-363.

ДОДАТКИ

Матеріал, який доповнює текст, допускається розміщувати в додатках (таблиці великого формату, програми завдань, які вирішують на ЕОМ та ін.).

Додатки оформлюють як продовження курсової роботи на наступних її сторінках, розміщуючи їх у порядку появи посилань у тексті роботи. Кожний додаток повинен починатися з нової сторінки. Додаток повинен мати заголовок, надрукований угорі малими літерами з першої великої симетрично відносно тексту сторінки. Посередині рядка над заголовком малими літерами з першої великої друкується слово "Додаток\_\_" і велика літера, що позначає додаток. Додатки слід позначати послідовно великими літерами української абетки, за винятком літер Г, Є, І, Ї, Й, О, Ч, Ь, наприклад, додаток А, додаток Б і т.д. Під словом "Додаток..." в дужках пишуть слово "обов'язковий" (якщо в додатку наводяться великі таблиці, які недоцільно включати в текст), або "довідковий" (якщо в ньому наводиться додатково інформація).

Ілюстрації, таблиці і формули, які розміщені в додатках, нумерують в межах кожного додатка, наприклад: рисунок Д.2 - другий рисунок додатка Д); формула (А.1)- перша формула додатка А.

У такому вигляді курсова робота віддається на перевірку на кафедру ІТ не пізніше встановленого строку.

**8. Захист курсової роботи**

Курсову роботу для перевірки науковим керівником студент подає на кафедру в термін вказаний в навчальному плані. У випадку незадовільної оцінки робота має бути перероблена відповідно до зауважень наукового керівника і повторно подана у встановлений строк.

Заключним етапом є захист курсової роботи. Курсова робота, що відповідає викладеним у методичних рекомендаціях вимогам, залежно від її якості, оцінюється на "відмінно", "добре", "задовільно", "не задовільно".

Комісія на захисті курсової роботи від кафедри розглядає й оцінює роботу студента, дає відзив з рекомендацією до захисту (на титульному аркуші) перед кафедральною комісією, яка призначається завідувачем кафедри.

Під час захисту курсової роботи студент повинен підготувати презентацію своїх основних результатів та має виявити глибокі знання з дисципліни „ВСТУП ДО МАШИННОГО НАВЧАННЯ”, вміти розкривати в усній доповіді зміст розглянутих у курсовій роботі положень, відповісти на поставлені запитання.

Пояснювальна записка курсової роботи, яку захищено, зберігається на кафедрі.

**9. Основні умови одержання оцінки**

Загальна ***бальна оцінка,*** що представлена в таблиці 1, виводиться простим сумуванням одержаних студентом балів за всі види поточних завдань при виконанні курсової роботи.

Таблиця 1

Бальна оцінка

|  |  |  |
| --- | --- | --- |
| Вид роботи | Бальна оцінка | |
| максимально можлива | мінімально необхідна |
| Виконання поставлених задач у повному обсязі | 40 | 30 |
| Оформлення роботи в межах вимог | 20 | 10 |
| Прилюдний захист курсової роботи | 40 | 20 |

Максимально можлива бальна оцінка, яку може набрати студент за виконання всіх вимог до курсової роботи, дорівнює 100 балам

***Курсова робота до захисту не допускається****,* якщо вона:

* подана науковому керівникові на перевірку зі значними порушеннями строків, установлених календарним графіком виконання курсової роботи;
* виконана несамостійно;
* за структурою не відповідає вимогам;
* за змістом не розкриває теми курсової роботи;
* недбало оформлена.

Додаток А

(обов’язковий)

**Тематика курсових робіт з курсу**

**«ВСТУП ДО МАШИННОГО НАВЧАННЯ»**

**1. Прогнозування цін на нерухомість**

* **Опис:** Використання алгоритмів машинного навчання для прогнозування цін на нерухомість на основі наборів даних з Kaggle. Можна використовувати такі набори, як "House Prices: Advanced Regression Techniques".
* **Методи:** Лінійна регресія, дерева рішень.
* **Ціль:** Побудувати модель для прогнозування ціни на нерухомість на основі різних ознак (площа, кількість кімнат, розташування тощо).

**2. Прогнозування якості вина**

* **Опис:** Використання даних про якість вина для створення моделі, яка передбачає рейтинг вина за його властивостями.
* **Набір даних:** "Wine Quality" з Kaggle.
* **Методи:** Лінійна регресія, дерева рішень.
* **Ціль:** Розробити модель для прогнозування якості вина на основі хімічних характеристик.

**3. Прогнозування ймовірності захворювання (наприклад, діабет)**

* **Опис:** Створення прогнозної моделі для визначення ймовірності розвитку діабету на основі медичних показників.
* **Набір даних:** "Pima Indians Diabetes Database".
* **Методи:** Логістична регресія, k-Nearest Neighbors
* **Ціль:** Створити модель для прогнозування ризику діабету за допомогою медичних даних.

**4. Класифікація зображень: Визначення категорії тварин**

* **Опис:** Використання моделей глибокого навчання для класифікації зображень різних тварин на основі набора зображень.
* **Набір даних:** "Dogs vs. Cats" або "CIFAR-10".
* **Методи:** Конволюційні нейронні мережі (CNN), Transfer Learning.
* **Ціль:** Створити модель для класифікації зображень котів і собак (або інших тварин) за допомогою глибоких нейронних мереж.

**5. Прогнозування відтоку клієнтів**

* **Опис:** Створення моделі для прогнозування ймовірності відтоку клієнтів з певної компанії або сервісу.
* **Набір даних:** "Telco Customer Churn" з Kaggle.
* **Методи:** Логістична регресія, дерева рішень.
* **Ціль:** Побудувати модель для визначення ймовірності того, що клієнт відмовиться від послуг компанії.

**6. Прогнозування продажів на основі історичних даних**

* **Опис:** Використання даних про продажі товарів для прогнозування майбутніх продажів.
* **Набір даних:** "Store Sales Forecasting".
* **Методи:** Часові ряди (ARIMA, LSTM), випадкові ліси.
* **Ціль:** Створити модель для прогнозування майбутніх продажів на основі історичних даних про продажі.

**7. Прогнозування результатів спортивних подій**

* **Опис:** Створення моделі для прогнозування результатів спортивних матчів на основі історичних даних.
* **Набір даних:** "Football Results" або подібні.
* **Методи:** Класифікація, регресія, дерева рішень.
* **Ціль:** Розробити модель для прогнозування результатів футбольних матчів на основі статистичних даних про команди та попередні результати.

**8. Аналіз тексту та класифікація: Визначення емоцій в текстах**

* **Опис:** Використання моделей обробки природної мови (NLP) для класифікації текстів за емоціями.
* **Набір даних:** "Sentiment140" або "IMDB Movie Reviews".
* **Методи:** Наївний баєсівський класифікатор, LSTM, трансформери.
* **Ціль:** Створити модель, яка класифікує текст за категоріями емоцій (позитивний, негативний, нейтральний).

**9. Прогнозування рейтингу фільмів за допомогою даних про перегляди**

* **Опис:** Розробка моделі для прогнозування рейтингу фільмів на основі даних про їх популярність, жанр, режисерів та акторів.
* **Набір даних:** "MovieLens" або подібні.
* **Методи:** регресія.
* **Ціль:** Створити модель для прогнозування рейтингів фільмів на основі їх характеристик та минулої популярності.

**10. Аналіз фінансових даних для прогнозування фондового ринку**

* **Опис:** Використання фінансових даних для створення моделей, які можуть прогнозувати рух цін на фондовому ринку.
* **Набір даних:** Дані про фондовий ринок з Kaggle (наприклад, "Stock Market Data").
* **Методи:** Часові ряди (ARIMA, LSTM), регресія.
* **Ціль:** Побудувати модель для прогнозування цін акцій на основі історичних даних.

**11.** **Прогнозування ціни мобільних телефонів на основі їх характеристик.**

* Опис: Використовуючи набір даних "Mobile Price Classification", створіть модель машинного навчання для прогнозування цін мобільних телефонів на основі їх технічних характеристик, таких як кількість ядер, оперативна пам'ять, роздільна здатність екрану та інші характеристики.
* Методи: Класифікація, випадкові ліси.

12. **Прогнозування академічної успішності студентів на основі їх характеристик та умов навчання***.*

* Опис: Використовуючи набір даних "Student Study Performance", створіть модель для прогнозування успішності студентів (оцінки з предметів) на основі їх особистих характеристик, соціально-економічного статусу та умов навчання.
* Методи: Регресія

### 13. Прогнозування цін на електричну енергію за допомогою моделей машинного навчання.

* Опис: Створіть прогнозну модель для передбачення майбутніх цін на електричну енергію на основі історичних даних про ціни, попит, температуру та інші фактори. (**electricity-prices)**
* Методи: Часові ряди, ARIMA, LSTM, регресія.

14 – свій варіант

Додаток Б

(довідковий)

**Титульний лист курсової роботи**

|  |
| --- |
| Міністерство освіти та науки України  Запорізький національний університет  Кафедра КН  **ПОЯСНЮВАЛЬНА ЗАПИСКА**  **до курсової роботи**  з дисципліни: «ВСТУП ДО МАШИННОГО НАВЧАННЯ»  на тему: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Виконав: Студент\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  (ПІП студента)  Групи\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Керівник: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  (ПІП викладача)  Дата перевірки\_\_\_\_\_\_\_\_\_\_\_\_\_  Дата захисту\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  (Підпис викладача)  Запоріжжя  20\_\_ |

Додаток В

(довідковий)

|  |  |  |
| --- | --- | --- |
| **МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  **ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ** | | |
| Факультет | математичний | |
| Кафедра | комп’ютерних наук | |
| Дисципліна | | Вступ до машинного навчання |
| Спеціальність | | 123 Комп’ютерні науки |
| Освітня програма | | інформаційні системи та технології |

**ЗАВДАННЯ**

**НА КУРСОВУ РОБОТУ СТУДЕНТОВІ (СТУДЕНТЦІ)**

|  |
| --- |
| Свириденко Василь Станiславович |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1. Тема роботи | створення моделі машинного навчання для класифікації товарів за їхніми назвами на основі текстових даних | | | | | |
|  |  |  | | | | | |
|  | 2. Строк здачі студентом закінченої роботи | | | |  | | |
|  | 3. Вихідні дані до роботи | | | 1. Постановка задачі. | | | |
|  | 2. Перелік літератури. | | | | | | |
|  | 3. Практична реалізація коду | | | | | | |
|  | 4. Зміст роботи(перелік питань, які потрібно розробити) | | | | |  | |
|  | 1. Постановка задачі. | | | | | | |
|  | 2. Основні теоретичні відомості. | | | | | | |
|  | 3. Модель машинного навчання | | | | | | |
|  | 5. Перелік графічного матеріалу (з точним зазначенням обов’язкових креслень) | | | | | |  |
| 6. Дата видачі завдання | | | 26.03.2025 | | | | | |

**КАЛЕНДАРНИЙ ПЛАН**

|  |  |  |  |
| --- | --- | --- | --- |
| **№** | **Назва етапів курсової роботи** | **Термін виконання етапів роботи** | **Примітка** |
| 1. | Розробка плану роботи | 26.03.2025 |  |
|  |  |  |  |
| 2. | Збір вихідних даних | 30.03.2025 |  |
|  |  |  |  |
| 3. | Обробка методичних джерел | 03.04.2025 |  |
|  |  |  |  |
| 4. | Розробка першого розділу | 09.04.2025 |  |
|  |  |  |  |
| 5. | Розробка другого розділу | 18.04.2025 |  |
|  |  |  |  |
| 6. | Оформлення курсової роботи | 01.05.2025 |  |
|  |  |  |  |
| 7. | Захист курсової роботи | 10.05.2025 |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Студент | |  |  |  |
|  |  | (підпис) |  | (ініціали та прізвище) |
|  |  | | | |
| Керівник роботи | |  |  |  |
|  |  | (підпис) |  | (ініціали та прізвище) |

Додаток D

(довідковий)

**Приклад виконання 2 розділу**

РОЗДІЛ 2. РЕАЛІЗАЦІЯ МОДЕЛІ МАШИННОГО НАВЧАННЯ

2.1. Постановка задачі

2.2. Обгрунтування обраної моделі машинного навчання

2.3. Опис використаних інструментів

2.4. Побудова та аналіз моделі

2.1. Постановка задачі

Розглянемо набір даних World Happinest Index https://www.kaggle.com/datasets/unsdsn/world-happiness, який містить індекс щастя у різних країнах а також показники розвитку цих країн. Тому, ми можем досліджувати, як ці показники розвитку впливают на рівень щастя, та прогнозувати для деякої іншої країни цей індекс, якщо відомі ці показники.

Прогнозування індексу щастя для різних країн на основі різноманітних соціально-економічних та політичних факторів.

Опис:

Метою цієї роботи є побудова моделі машинного навчання для прогнозування індексу щастя для різних країн на основі наданого набору даних "Happiness Countries" з Kaggle. Індекс щастя (Happiness Index) є комплексною мірою, яка оцінює загальний рівень благополуччя в країні, враховуючи такі фактори, як дохід, соціальна підтримка, очікувана тривалість життя, свобода вибору, корупція та інші соціально-економічні параметри.

Дані: Набір даних містить різноманітні ознаки для кожної країни, включаючи:

Happiness Score (Індекс щастя) — цільова змінна.

GDP per capita (Валовий внутрішній продукт на душу населення).

Social support (Соціальна підтримка).

Healthy life expectancy (Очікувана тривалість життя).

Freedom to make life choices (Свобода приймати життєві рішення).

Generosity (Щедрість).

Perceptions of corruption (Відчуття корупції).

Інші соціальні та економічні фактори.

Мета: Збудувати регресійну модель, яка на основі доступних факторів передбачатиме індекс щастя для різних країн.

Завдання:

Попередня обробка даних: Очищення набору даних, обробка відсутніх значень, нормалізація або стандартизація числових змінних.

Аналіз даних: Аналіз кореляції між різними ознаками та індексом щастя для виявлення найважливіших факторів, що впливають на щастя.

Побудова моделі: Використати методи регресії, такі як лінійна регресія, регресія на основі випадкових лісів, градієнтний бустинг для побудови моделі прогнозування індексу щастя.

Оцінка моделі: Оцінка точності моделі за допомогою метрик регресії, таких як середня квадратична помилка (MSE), R², середня абсолютна помилка (MAE).

Інтерпретація результатів: Аналіз результатів моделі та визначення основних факторів, що найбільше впливають на індекс щастя.

2.2. Обгрунтування обраної моделі машинного навчання

Для задачі прогнозування індексу щастя можна використовувати різноманітні методи регресії. Оскільки ми маємо числові ознаки, задача є класичним прикладом для побудови регресійної моделі, яка передбачатиме значення індексу щастя на основі соціально-економічних та політичних характеристик країн.

1. **Лінійна регресія:** Це базовий підхід, який можна використовувати для прогнозування цільової змінної на основі лінійних зв'язків між ознаками та цільовою змінною.
2. **Регресія на основі випадкових лісів (Random Forest Regression):** Це більш складний метод, який застосовує ансамбль рішень для більш точного прогнозування, здатний обробляти нелінійні зв'язки між ознаками.

Оцінка моделей виконуватиметься за допомогою таких методів як

* + - Середня квадратична помилка (MSE): для вимірювання середньої величини помилки між прогнозом та реальним значенням.
    - R² (коефіцієнт детермінації): для оцінки того, наскільки добре модель пояснює варіацію цільової змінної.
    - Середня абсолютна помилка (MAE): для вимірювання середньої абсолютної різниці між прогнозованими та реальними значеннями.

2.3. Опис використаних інструментів

1. Pandas: для обробки та маніпуляції даними. Pandas дозволяє легко завантажувати дані, очищати їх та здійснювати попередню обробку.

import pandas as pd

data = pd.read\_csv('happiness\_data.csv')

1. NumPy: для математичних обчислень і роботи з масивами, що дозволяє швидко виконувати операції над даними.

import numpy as np

1. Matplotlib / Seaborn: для візуалізації даних та результатів моделювання.

import matplotlib.pyplot as plt

import seaborn as sns

sns.heatmap(data.corr(), annot=True)

1. Scikit-learn: для побудови моделей машинного навчання, підбору гіперпараметрів і оцінки точності.
   * Лінійна регресія:

from sklearn.linear\_model import LinearRegression

model = LinearRegression()

model.fit(X\_train, y\_train)

* + Випадкові ліси:

from sklearn.ensemble import RandomForestRegressor

rf\_model = RandomForestRegressor()

rf\_model.fit(X\_train, y\_train)

2.4. Побудова та аналіз моделі

Після вибору набору даних необхідно визначити ті стовбці, які є впливаючими, і той, який є цільовим.

В прикладі з набору даних World Happinest Index (дані за 2019 р були відкриті в Excel) зеленим кольром показані вхідні стовбці, жовтим – цільовий.

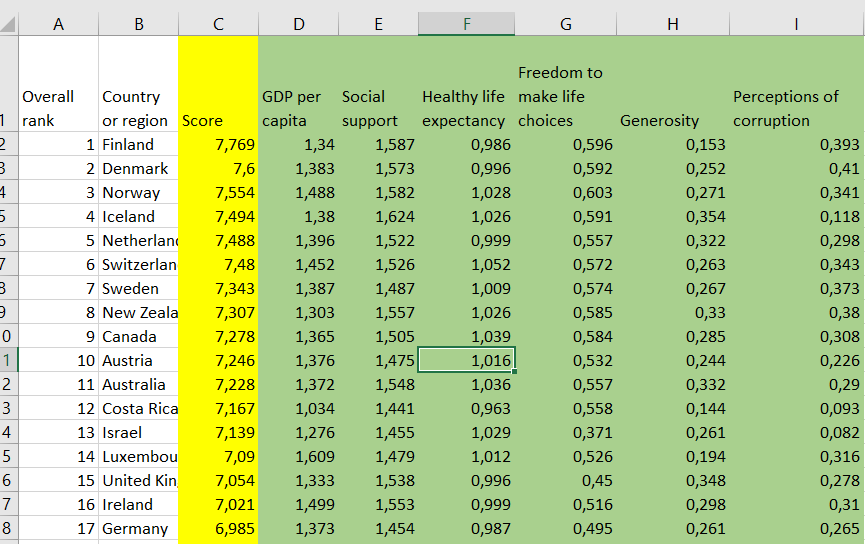


Рис. 2.1 – Фрагмент даних

Далі аналіз виконується згідно прикладам, розглянутим в курсі <https://www.kaggle.com/learn/intro-to-machine-learning>

**Опис моделі машинного навчання:**

<https://www.kaggle.com/code/dansbecker/how-models-work>

**Перевірка Ваших даних**

<https://www.kaggle.com/code/dansbecker/basic-data-exploration>

Крок 1.

Завантажуємо файл з набором даних у робочий каталог проєкту

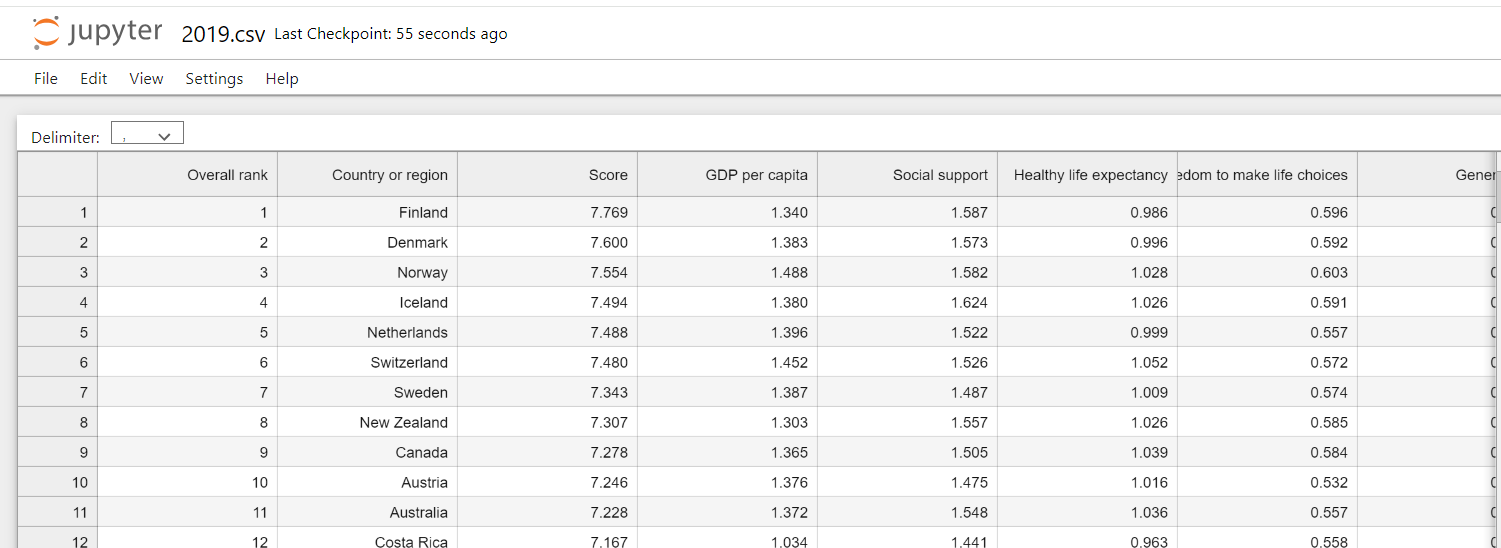


Рис. 2.2. Завантаження даних у середовище моделювання

Крок 2. Створюємо новий блокнот в якому виконуємо підключення до даних а також отримуємо основну описову статистику.



Рис. 2.3. – Описова статистика для набору даних

Описуємо отримані показники даних…

Створюємо першу модель машинного навчання

Робимо прогноз цільових показників на основі моделі

Оцінюємо прогноз порівнянням з даними вхідного набору

…