Методичні вказівки до виконання дослідницького наукового проєкту

Юлія Хлевна

Ігор Мірошниченко

2024-12-25

Зміст

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА**

**Кафедра технологій управління**

**Ю.Л. Хлевна, І.В. Мірошниченко** **МЕТОДИЧНІ ВКАЗІВКИ**

**до виконання дослідницького наукового проєкту для студентів спеціальності 122 «Комп’ютерні науки»**

освітня програма «Інформаційна аналітика та впливи» **КИЇВ – 2024**

Методичні вказівки до виконання дослідницького наукового проєкту роботи «Методологія DATA SCIENCE» для студентів спеціальності 122 «Компютерні науки», освітня програма «Інформаційна аналітика та впливи». / Ю.Л. Хлевна, І.В. Мірошниченко. К : 2024. – 62 с.

Укладачі: Ю.Л. Хлевна, І.В. Мірошниченко

Затверджено на засіданні вченої ради факультету інформаційних технологій Київського національного університету імені Тараса Шевченка, протокол № \_\_ від \_\_ \_\_\_\_\_\_\_\_\_\_ 2024 року

# 1. ВСТУП

На сьогодні компанії використовують, у своїй практиці, моделі, методи, інструменти та засоби інформаційної аналітики та впливів для прийняття раціональних управлінських рішень та конкурентоздатності компаній. В англомовній ітерації це розглядається як Data science and influences. Нерідко, це реалізовано у інтелектуальних інформаційних системах, експертних системах тощо, системах, ядром яких є база знань (БЗ) або модель предметної області, описана на мовою надвисокого рівня. Найчастіше ці системи застосовуються для вирішення складних завдань, де основна складність вирішення пов’язана з використанням слабоформалізованих знань фахівців-практиків, і де логічна (або смислова) обробка інформації превалює над обчислювальною, наприклад, розуміння природної мови, підтримка прийняття рішення в складних ситуаціях, постановка діагнозу і рекомендації щодо методів лікування, аналіз візуальної інформації, управління диспетчерськими пультами тощо.

В основі Data science and influences – методи обробки даних в умовах великих обсягів і високого рівня паралелізму, статистичні методи, методи інтелектуального аналізу даних і застосування штучного інтелекту для роботи з даними, а також методи проєктування та розробки баз даних.

Багато сучасних вебсайтів і пристроїв використовують алгоритми машинного навчання, починаючи з автоматичних рекомендацій по перегляду фільмів, замовлення їжі або покупки продуктів, і закінчуючи персоналізованими онлайн-трансляціями та розпізнаванням друзів на фотографіях. Вийшовши за межі комерційних додатків, продукти даних справили величезний вплив на наукові дослідження, керовані даними. Інструменти використовувалися для вирішення різних наукових завдань (дослідження зірок, пошук далеких планет, відкриття нових частинок, аналіз послідовностей ДНК, а також розробка персоналізованих методів лікування).

*Постановка завдання* дослідницького наукового проєкту: розробити модель аналізу даних, концепцію алгоритму та програмного забезпечення для застосування конкретній практичній сфері – застосування, у певному проєкті. Для цього потрібно аргументувати актуальність проєкту, здійснити методологічний опис проєкту аналізу даних та інтерпретацію даних, з якими Ви працюєте, і проаналізувати застосовність цих даних до завдання, яке Ви хочете вирішити. При роботі над дослідницьким науковим проєктом потрібно застосовувати методи наукового пізнання, а не вибір алгоритму навмання.

Перш ніж розпочати побудову моделі аналізу даних, необхідно виявляти потреби потенційних замовників щодо автоматизації обробки інформації, провести літературний аналіз щодо того, які методи, моделі та алгоритми застосовуються до обраної предметної галузі. Описати методологію за якою буде розроблятись продукт даних. Формалізувати, систематизувати потреби та вимоги до інформаційної або комп’ютерної системи, що буде розроблятись, експлуатується. Консолідувати дані з різних джерел інформації. Формалізувати набір даних, формалізувати обрані при літературному аналізі моделі для подальшої роботи, оцінити ваш набір даних.

Розробка моделі є найбільш ресурсозатратним етапом. Також на цьому етапі потрібно провести інженерні роботи із функціями моделі, зокрема визначення та вдосконалення функцій, які можна використовувати в моделі, провести навчання моделі за обраною в літературному аналізі технологією, оцінити модель та виберіть найефективнішу модель (моделі). Після розробленої моделі варто здійснити опис алгоритмічного та програмного забезпечення для аналізу даних (включно з великими). Представити архітектурні рішення інформаційних та комп’ютерних систем для аналізу даних. Представити який вплив розроблена модель буде мати на управлінські рішення.

Дослідницький науковий проєкт забезпечує ПРН 9, 10, 15, 16, 18, 22, 23, 24 ОНП «Інформаційна аналітика та впливи».

# 2. ЗАГАЛЬНІ ПОЛОЖЕННЯ ТА ВИМОГИ ДО ДНП

Дослідницький науковий проєкт (ДНП) є обов’язковим елементом навчального процесу другого семестру магістерської підготовки за спеціальністю 122 «Комп’ютерні науки», освітньої програми «Інформаційна аналітика та впливи». Виконання дослідницького наукового проєкту є одним з основних видів самостійної роботи студентів та першим етапом підготовки до виконання дипломної роботи. Дослідницький проєкт охоплює дисципліни, які здобувачі освіти вивчають у І та ІІ семестрах.

Дослідницький проєкт для здобувачів освіти за ОНП ІАВ - це комплексне наукове дослідження для формування наукового світогляду, розвитку дослідницьких умінь, навичок творчого вирішення практичних завдань процесів інформаційної аналітики та впливів.

*Метою* дослідницького проєкту є отримання практик дослідницького характеру під керівництвом НПП Університету, за необхідності залученням консультантів для закріплення, поглиблення та узагальнення теоретичних й практичних за ОНП ІАВ. Дослідницький проєкт включає в себе пошук, аналіз, систематизацію та узагальнення літературних джерел за темою дослідження, а також обов’язкове проведення власного дослідження, а саме:

* використання набутих у процесі навчання теоретичних знань, практичних навичок для організації та проведення методології проекту аналізу даних;
* пропонування моделей та алгоритмів при вирішенні задачі з області Data Science;
* критичний аналіз літературних та наукових джерел;
* застосування сучасних методик наукових досліджень;
* узагальнення статистичних та інших фактичних матеріалів;
* розробка й оцінювання альтернативних підходів до вирішення визначених проблем на підприємстві;
* обґрунтування запропонованих рішень з використанням сучасних методів;
* висновки та внесення пропозицій, що мають теоретичне і практичне значення;
* вміння узагальнювати переваги та недоліки обраної моделі, мови програмування та алгоритмів, що використовуються у проєкті Data Science.

*Основними завданнями ДНП є:*

* узагальнення теоретико-методологічних засад методів Data Science;
* розробка методу Data Science орієнтованого на конкретну сферу застосування та формалізація його застосування в комп’ютерних науках.

Повнота й особливості реалізації цих завдань залежать від предметної області інформаційної аналітики та його специфічних рис.

Працюючи над дослідницьким проєктом, студент має засвоїти навички правильної постановки проблеми та обґрунтування її актуальності, формулювання мети і завдань дослідження, побудови логічного плану й оптимальної структури роботи, роботи з літературними джерелами та статистичною інформацією, аналізу та оцінки різних аспектів обраної сфери інформаційної аналітики, пропозицій щодо розгортання розробленої моделі, вибір обґрунтування власних узагальнень, висновків і пропозицій.

При підготовці ДНП студент повинен дотримуватися *певних вимог*:

* ДНП має бути написана державною мовою;
* ДНП кожен студент виконує індивідуально;
* запропоновані нові рішення мають бути аргументовані й порівнюватися з уже відомими рішеннями;
* оформлення ДНП має відповідати вимогам, що ставляться до робіт, поданих до друку;
* при написанні ДНП студент повинен посилатися на автора та джерело, звідки запозичений матеріал або окремі результати.

Підготовка ДНП умовно поділяється на такі *етапи*:

* вибір і затвердження теми;
* підбір та вивчення літературних та інших джерел з теми дослідження;
* розробка концепції та плану ДНП, погодження їх з керівником;
* написання тексту пояснювальної записки ДНП та її оформлення;
* оформлення графічного (ілюстративного) матеріалу для захисту роботи;
* захист ДНП.

# 3. ПОРЯДОК ОРГАНІЗАЦІЇ ВИКОНАННЯ ДНП

ДНП виконується студентами відповідно до навчального плану.

Тематика (перелік можливих тем) ДНП затверджується рішенням відповідальної за їх проведення кафедри або спільним рішенням відповідальних кафедр. Одночасно визначається розподіл тем між відповідальними кафедрами. Для затвердження теми студенти оформлюють заяву (Додаток В), у якій формулюють назву запропонованої теми роботи та пропонують керівника цієї роботи. Під час розгляду на засіданні кафедри назви тем можуть зазнати незначних змін, а керівники тем також можуть бути змінені, враховуючи вимоги, що кожен керівник може здійснювати керівництво ДНП не більше десяти студентів.

Затверджена кафедрою тематика ДНП завчасно оприлюднюється через стенди кафедри та факультетів, студенти яких повинні їх виконувати. Відповідальність за своєчасність оприлюднення тематики ДНП згідно зі встановленими вимогами несе завідувач відповідальної кафедри.

Тематика ДНП повинна відповідати завданням ОНП «Інформаційна аналітика та впливи».

В окремих випадках, за рішенням відповідальної кафедри, допускається виконання комплексної теми кількома студентами (п. 3.9.1. Положення про організацію навчального процесу у ЗВО).

Студенти під керівництвом призначеного завідувачем кафедри викладача впродовж ***тижня*** визначають бажану для себе тему ДНП і подають заяву на ім’я завідувача відповідальної кафедри із зазначенням обраної теми.

На підставі вибраних студентами тем на засіданні кафедри затверджується розподіл тем ДНП та призначення їх керівників. Студентам, які не подали заяви про вибір теми, кафедра самостійно призначає теми ДНП, а також розв’язує конфліктні питання.

Керівники ДНП призначаються кафедрою з-поміж науково-педагогічних (педагогічних) працівників університету відповідно до розподілу педагогічного навантаження з урахуванням достатності їх кваліфікації та відповідності наукових інтересів.

Відповідальність за своєчасність розподілу тем ДНП та його оприлюднення несе завідувач кафедри.

Для викладачів, які здійснюють керівництво ДНП, методистом кафедри (впродовж тижня після затвердження розподілу тем) готуються інформаційні пам’ятки із зазначенням прізвищ та ініціалів студентів, факультету, курсу, спеціальності (групи) та затверджених тем робіт.

Для студентів, які виконують курсові роботи, одночасно з оприлюдненням розподілу тем ДНП, на стенді кафедри може оприлюднюватися графік їх виконання (можливо, за розділами або етапами).

При необхідності зміни теми ДНП та/або керівника студент подає мотивовану заяву на ім’я завідувача кафедри (у випадку зміни теми завізовану керівником), який приймає відповідне рішення (у формі резолюції). Зміна теми ДНП або керівника менше, ніж за місяць до дати планового захисту роботи, ***не припускається***, крім випадку фізичної неможливості виконання керівником своїх функцій.

Обов’язковими компонентами керівництва виконанням студентами ДНП є:

* вступна настановча групова співбесіда тривалістю до однієї пари з усіма студентами, які виконують курсову роботу під керівництвом цього викладача;
* консультації студентів за графіком консультування студентів, який оприлюднюється кафедрою на початку кожного семестру.

У разі нехтування керівником обов’язковими компонентами керівництва виконанням ДНП, завідувач відповідальної кафедри зобов’язаний (за результатами обговорення на засіданні кафедри) замінити керівника.

Інші питання методики роботи викладачів з керівництва ДНП визначаються відповідальною (провідною) кафедрою з урахуванням вимог університету.

# 4. ОБСЯГ, СТРУКТУРА ТА ЗМІСТ ТА ВИМОГИ ДО ОФОРМЛЕННЯ ДНП

## 4.1 Обсяг та вимоги до оформлення ДНП

Обсяг завдань ДНП має відповідати не менше, ніж *90 кредитів* ЄКТС самостійної роботи студента.

Подання тексту ДНП державною мовою в друкованій формі та дотримання зразка титульного листа є вимогами університету.

ДНП, як оригінальне науково-прикладне дослідження, повинне мати певну логіку побудови, послідовність і завершеність. *Загальний обсяг ДНП* має бути в межах 35–40 сторінок машинопису формату А4 основного тексту (без урахування переліку використаних джерел та додатків).

ДНП виконується на одному боці аркуша білого паперу формату А4 (210 х 297 мм). Текст розташовується через півтора міжрядкових інтервали до тридцяти рядків на сторінці у текстовому редакторі Microsoft Word шрифтом Times New Roman, 14 кеглем. Допускається розміщувати таблиці та інші ілюстративні матеріали на аркушах формату А3 (не більше 40 рядків на сторінку).

Текст ДНП розміщується на аркуші з дотриманням таких розмірів полів:

* з лівого боку – 30 мм,
* з правого – 10 мм,
* зверху – 20 мм,
* знизу – 20 мм.

Текст має бути вирівняним з обох боків (ліворуч та праворуч). Абзаци в тексті потрібно починати з відступу, що дорівнює 1,25 пункту. Відступи мають бути однаковими впродовж всього тексту.

Роздруковані екранні форми мають відповідати формату А4, включатися до загальної нумерації сторінок ДНП і розміщуватися у вигляді рисунків з відповідною нумерацією. Великі таблиці (понад дві сторінки) або рисунки (понад дві сторінки), мають бути розміщені у додатках.

Текст основної частини ДНП поділяється на розділи, підрозділи, пункти та підпункти згідно з планом, затвердженим у завданні на виконання ДНП.

Заголовки структурних частин ДНП: ЗМІСТ, ВСТУП, РОЗДІЛ 1, РОЗДІЛ 2, РОЗДІЛ 3, РОЗДІЛ 4 (за необхідності), ПЕРЕЛІК ВИКОРИСТАНИХ ЛІТЕРАТУРНИХ ДЖЕРЕЛ, ДОДАТКИ, друкуються великими напівжирними літерами симетрично (посередині аркуша) до тексту. Заголовки підрозділів друкуються маленькими літерами (крім першої великої) з абзацу, крапка в кінці заголовка не ставиться.

Відстань між заголовком (за винятком заголовка пункту) та текстом має дорівнювати 3–4 інтервалам в основному тексті.

Кожний розділ ДНП слід починати з нової сторінки, а підрозділи продовжують з відступом на поточній сторінці.

### 4.1.1 Нумерація

Нумерацію сторінок, розділів, підрозділів, пунктів, підпунктів, рисунків, таблиць, формул подають арабськими цифрами без знака №.

Першою сторінкою ДНП є титульна сторінка, яка включається до загальної нумерації сторінок і на якій номер сторінки не ставиться. Титульна сторінка ДНП оформлюється на кожну роботу, згідно з формою, наведеною у додатку А. Нумерація наступної сторінки проставляється у правому нижньому куті.

ЗМІСТ, ВСТУП, ВИСНОВКИ, СПИСОК ВИКОРИСТАНИХ ЛІТЕРАТУРНИХ ДЖЕРЕЛ, ДОДАТКИ не нумеруються як розділи. Номер розділу ставиться арабськими цифрами з точкою після слів «РОЗДІЛ» (РОЗДІЛ 1.), підрозділи нумеруються у межах кожного розділу, використовуючи номер розділу і порядковий номер підрозділу, між якими ставиться крапка, наприклад, «1.4» (четвертий підрозділ першого розділу). Потім у тому ж порядку розміщують заголовок підрозділу, який починається з великої літери.

### 4.1.2 Переліки

У тексті розділів або підрозділів ДНП можуть бути переліки ([приклад 4.1](#exm-list)). Перед кожною позицією переліку слід ставити спеціальний символ (крапка, мінус, зірочка, плюс, номер, літера, символ, залежно від вибору автора) або (за необхідності посилатися в тексті на один із переліків) малу літеру. Для подальшої деталізації переліку необхідно використовувати арабські цифри. Кожен рядок переліку має закінчувалися комою, крапкою або точкою з крапкою. Якщо після номеру або в кінці рядку переліку ставиться крапка, то подальший текст пишеться з великої літери. В кінці переліку ставиться крапка.

**Приклад 4.1 (Оформлення переліків у ДНП)** *Варіант 1:*

* невпорядкований список;
	+ підпункт 1;
	+ підпункт 2;
		- підпункт 1;
* невпорядкований список.

*Варіант 2:*

1. Впорядкований список.
2. Пункт 2.
	* підпункт 1;
	* підпункт 2.
3. Пункт 3.

### 4.1.3 Ілюстрації

Ілюстрації (схеми, графіки тощо) і таблиці слід подавати в роботі безпосередньо після тексту, де їх згадано вперше, або на наступній сторінці. Якщо вони містяться на окремих сторінках ДНП, їх включають до загальної нумерації сторінок. Ілюстративні або табличні матеріали, розміри яких є більшими за формат А4, враховують як одну сторінку і розміщують у відповідних місцях після згадування у тексті або в додатках. На всі ілюстрації мають бути посилання у тексті.

Ілюстрації позначають словом «Рисунок» та нумерують послідовно в межах розділу, за винятком ілюстрацій у додатках. Номер ілюстрації має складатися з номера розділу і порядкового номера ілюстрації, між якими ставиться крапка, наприклад, «Рисунок 1.2 − » (другий рисунок першого розділу). Номер ілюстрації, її назва і пояснювальні підписи розміщуються послідовно під ілюстрацією.

Графічний матеріал кожного додатку позначають окремою нумерацією арабськими цифрами з додаванням перед цифрою позначення додатка (наприклад, Рисунок А.1 − ). Нижче наведено приклад оформлення ілюстрації − [рисунок 4.1](#fig-figcrossref).

|  |
| --- |
| Рисунок 4.1 – Приклад оформлення рисунку |

### 4.1.4 Таблиці

За наявності у тексті таблиць, які доповнюють графічний матеріал, таблиці слід розміщувати після графічного матеріалу.

Таблиці застосовуються для уточнення та зручності порівняння показників. Назва таблиці має точно і стисло відображати її зміст.

Таблиці розміщують відразу після закінчення абзацу, у тексті якого є перше посилання на неї (наприклад, таблиця 1.1), або на наступній сторінці (за необхідності – у додатку). Таблиці нумерують послідовно у межах розділу. У правому верхньому куті над таблицею розміщують напис «Таблиця» із зазначенням її номера, який складається з номера розділу і порядкового номера таблиці, між якими ставиться крапка, наприклад: «Таблиця 2.3» (третя таблиця другого розділу). Нижче, симетрично до тексту (посередині аркуша), вказується назва таблиці. Назву і слово «Таблиця» починають з великої літери.

Якщо рядки або стовпці таблиці виходять за формат сторінки, то таблицю ділять на частини, які розміщують одна під одною, при цьому в кожній частині таблиці повторюють її заголовок («шапку»). Над першою частиною таблиці пишуть слово «таблиця» із зазначенням відповідного номера, а на наступних – «Продовження табл. 2.3» або «Закінчення табл. 2.3», використовуючи створену на початку таблиці нумерацію колонок, яку зазначають відразу після заголовків колонок таблиці. Під назвою подається зображення самої таблиці. Допускається розміщення таблиці вздовж довгого боку аркуша. Таблиці кожного додатку позначають окремою нумерацією арабськими цифрами з додаванням перед цифрою позначення додатка (наприклад, Таблиця А.1).

Заголовки стовпців і рядків таблиці слід друкувати з великої літери, під заголовки стовпців – з малої, якщо вони є продовженням заголовка, або з великої, якщо вони мають самостійне значення. У кінці заголовків та підзаголовків таблиць крапки не ставлять, заголовки і підзаголовки стовпців друкують поодинці. Заголовки стовпців і рядків таблиці, як правило, розміщують горизонтально. Приклад оформлення таблиці наведено нижче у [таблиці 4.1](#tbl-tabcrossref).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблиця 4.1 – Приклад оформлення таблиць

| Товар | В наявності | Ціна |
| --- | --- | --- |
| Python Hat | Так | 23.99 |
| SQL Hat | Так | 23.99 |
| Codecademy Tee | Ні | 19.99 |
| Codecademy Hoodie | Ні | 42.99 |

 |

### 4.1.5 Посилання

Посилання у тексті ДНП на джерела інформації слід зазначати порядковим номером посилань за переліком у списку використаних джерел та сторінки, які виділені двома квадратними дужками. На цих номерах сторінок у першоджерелі має бути розміщена пояснювальна інформація у вигляді визначень, рисунків, таблиць, висновків тощо.

Наприклад: “…….” [8, c. 12–14].

При необхідності посилання на нормативні акти слід вказати відповідні розділи або статті із зазначенням повної назви акта. Якщо в тексті ДНП наведено всі реквізити нормативного акта, а саме: назву, вид, номер та дату прийняття, то виноски не наводяться, але є обов’язковим наведення даного джерела у переліку використаних джерел. При виконанні ДНП **забороняється дослівне списування тексту з літературного джерела**, нормативних матеріалів. Якщо за текстом необхідно навести цитату, вона береться в лапки з посиланням на відповідне джерело.

### 4.1.6 Формули

Формули та рівняння розташовують безпосередньо після тексту, в якому вони згадуються, посередині рядка і нумерують арабськими цифрами. Вище і нижче кожної формули повинно бути залишено не менше одного вільного рядка.

Розміри літер та знаків повинні бути надруковані таким шрифтом: великі літери та цифри – 16, малі – 14, показники степенів та індексів над літерами та під літерами – 8. Номер формули відповідного розділу і підрозділу наводять у круглих дужках, навпроти них праворуч. Номер пишуть на рівні нижнього рядка формули, до якої він належить. При посиланні в тексті на формулу необхідно вказати її повний номер в дужках, наприклад, «у формулі (2.1)». Після формули пишуть слово «де» і розшифровують позначення словами в такій послідовності, в якій вони подані у формулі. Після слова «де» двокрапка не ставиться. Приклад оформлення наведено у [пр. 4.2](#exm-formula).

**Приклад 4.2 (Оформлення формули у тексті)** “…Метод Gradient Boosting приймає реальні значення y та шукає апроксимацію $F\left(x\right)$ у вигляді зваженої суми функцій $h\_{i}\left(x\right)$ з певного класу $H$, які називають базовими (слабкими) навчальними елементами (base/weak learners):

$$F\left(x\right)=\sum\_{i=1}^{M}γ\_{i}h\_{i}\left(x\right)+const  \left(4.1\right)$$

### 4.1.7 Перелік використаних джерел

Перелік використаних джерел повинен містити усі джерела інформації, які використані при написанні ДНП, та розміщуватись за їх значущістю в алфавітному порядку. Відомості про включені до списку джерела слід подавати відповідно до державного стандарту ДСТУ 8302:2015 − «Інформація та документація. Бібліографічне посилання. Загальні положення та правила складання». Стандарт введений в дію 01. 06. 2016 та встановлює види бібліографічних посилань, правила та особливості їхнього складання й розміщування у документах ([Додаток Б](#sec-appendix-b)).

### 4.1.8 Додатки

Додатки оформлюються як продовження ДНП і розміщуються у порядку появи посилань у тексті ДНП. При цьому розуміється, що додатки доповнюють або роз’яснюють інформацію, наведену в основному тексті ДНП. До додатків необхідно включати допоміжні матеріали, таблиці, графічний матеріал, розрахунки тощо (матеріали які стосуються ДНП, але носять довідковий або додатковий характер). Посилання на додатки в текстовій частині ДНП є обов’язковим.

Кожний додаток починають із нової сторінки. На першому аркуші додатків посередині великими літерами пишуть «Додатки». У правому верхньому куті наступної сторінки пишуть слово «Додаток» та його номер, наприклад «Додаток А». Кожний додаток повинен мати заголовок, надрукований малими літерами з першої великої літери симетрично відносно тексту сторінки. Всі додатки послідовно нумеруються великими літерами української абетки, за винятком літер Ґ, Є, З, І, Ї, Й, О, Ч, Ь. Наприклад «Додаток Д».

## 4.2 Структура ДНП

Рекомендується така структура ДНП:

* титульна сторінка;
* зміст;
* вступ;
* перший розділ;
* другий розділ;
* третій розділ;
* висновки;
* перелік використаних літературних джерел;
* додатки.

*Зміст* ДНП може займати 1–2 сторінки. У ньому вказуються назви всіх розділів і підрозділів (параграфів) із зазначенням початкових сторінок. Зміст ДНП має відображати суть проблеми, її складність та логіку дослідження. Назви розділів і підрозділів повинні бути стислими і зрозумілими, літературно грамотними, тісно пов’язаними з назвою ДНП, але не повторювати її.

Рекомендується така структура ДНП (структура та зміст пояснювальної записки):

*Вступ* *Розділ 1.* Аналітичний проєкт Data Science  1.1 Аналіз практичної сфери застосування методології Data Science  1.2. Вибір та наповнення методології проєкту\_\_\_\_\_ Data Science  1.3 Аналіз літературних джерел щодо застосування моделей аналізу даних у проєкті\_\_\_\_\_\_\_\_\_\_\_\_\_ *Результат є вибір 2-3 методів, які будуть реалізуватись у 2 розділі.*  1.4.Постановка задачі Висновки *Розділ 2.* Формалізація методів аналізу даних у проєкті\_\_\_\_\_\_\_\_\_\_\_  2.1. Математичний опис обраних методів  2.3 Переваги та недоліки обраних методів  2.4. Мова програмування, основні бібліотеки та інструменти  2.5.Алгоритм вирішення задачі Висновки *Розділ 3.* Розробка методу Data Science  3.1 Аналіз даних для моделі;  3.2 Побудова моделі  3.3. Отримання прогнозів  3.4.Оцінка якості моделі  3.5. Концепція розгортання моделі у інформаційній технології  3.5. Перспективи подальших досліджень Висновки *Загальні висновки* *Перелік використаних літературних джерел* *Додатки.*

*Об’єктом дослідження* в ДНП є процеси в проєктах розробки продуктів даних, а також методи Data Science, які об’єднують у собі методи по обробці даних в умовах великих обсягів, статистичні методи, методи інтелектуального аналізу даних, машинного навчання і додатки штучного інтелекту для роботи з даними.

*Предметом дослідження* в ДНП є методологічне формування аналітичних проєктів, інструменти виявлення закономірностей в даних та отримання знань з даних в узагальненій формі.

У *вступі* (1–2 стор.) дається обґрунтування актуальності і практичного значення обраної теми; визначаються основні характеристики проєкту, що розглядається, формулюються мета, завдання, об’єкт і предмет дослідження; наводиться перелік застосованих методів дослідження; зазначається практичне значення отриманих результатів, **публікації за темою**.

*Перший розділ* (7-10 стор.) складається із опису проєкту дослідження, вибору методології за якою буде здійснено аналіз даних, літературного аналізу використання моделей, методів DS у аналітичному проєкті. Розробляється зміст фаз завдань за обраною методологією аналізу даних – план реалізації аналітичного проєкту. Постановка завдання.

*В другому розділі* слід зазначити (7-10 стор.) опис обраних методів та їх основних структур та підрозділів. Наводиться необхідна для впровадження модель, подається її математичний опис. Вибір доцільної мови програмування для обраного методу, наводяться особливості необхідних бібліотек, інструментів та функцій. Обґрунтування переваг та недоліків обраної мови програмування. Наводить алгоритм вирішення задачі.

*В третьому розділі* (10–20 стор.) виконується розробка обраного методу Data Science, при цьому виконується концептуалізація предметної області, побудова моделі, вилучення знань з тестового датасету, включаючи фрагменти коду, отримання прогнозів оцінки ефективності моделі, концепцію алгоритму та програмного забезпечення для застосування у конкретній практичній сфері – застосування, у певному проекті. Представити архітектурні рішення інформаційних та комп’ютерних систем для аналізу даних. Представити який вплив розроблена модель буде мати на управлінські рішення. Перспективи подальших досліджень Дані для проекту можуть мати синтетичний характер.

*Розділ «Висновки»* (1–2 стор.) є завершальною частиною ДНП. Він містить стислий виклад зроблених оцінок та узагальнень процесу виконання проєкту. Ознайомлення з текстом висновків повинно формувати у читача уявлення про ступінь реалізації автором поставленої мети і завдань ДНП.

*Розділ «Перелік використаних літературних джерел»* містить складений за чинними правилами перелік використаних літературних джерел, наукових публікацій, вебсайтів тощо.

*У розділ «Додатки»* виносяться таблиці допоміжного характеру, діаграми, графіки, зразки комп’ютерних екранів, роздруковані звіти з результатами комп’ютерного моделювання, методологічні описи проєкту.

# 5. ТЕМАТИКА ДНІ

Орієнтовний перелік тем ДНП (**в дужках рекомендований один із методів, але це не обов’язково**):

1. Розробка технології Data Science у проектах електронної комерції (метод кластеризації k-середніх).
2. Розробка технології Data Science у сфері аналітики продажів (метод агломеративної кластеризації у сфері)
3. Розробка технології Data Science у сфері маркетингових досліджень (метод DBSCAN кластеризації)
4. Розробка технології Data Science у сфері виявлення шахрайської діяльності (метод факторизації невід’ємних матриць)
5. Розробка технології Data Science у сфері обробки геоданих (метод k найближчих сусідів)
6. Розробка технології Data Science у сфері дослідження змін клімату (метод лінійної регресії)
7. Розробка технології Data Science у сфері прийняття рішень (метод бінарної класифікації)
8. Розробка технології Data Science у сфері пошуку аномальних паттернів поведінки на веб-сайті (метод логістичної регресії)
9. Розробка технології Data Science у сфері сегментування клієнтів (метод класифікації опорних векторів)
10. Розробка технології Data Science у сфері сегментування продажів (метод мультикласової класифікації)
11. Розробка технології Data Science у сфері виявлення спаму у соціальних мережах (метод наївного байєсівського класифікатора)
12. Розробка технології Data Science у сфері банківської діяльності (метод аналізу дерева рішень)
13. Розробка технології Data Science у сфері телекомунікацій (метод випадкового лісу)
14. Розробка технології Data Science у сфері оцінки нерухомості (метод градієнтного бустінгу)
15. Розробка технології Data Science у сфері прийняття рішень (метод нейронних мереж)
16. Розробка технології Data Science у сфері розпізнавання образів (метод нейронних мереж)
17. Розробка технології Data Science у сфері прогнозування фінансових котувань (метод нейронних мереж)
18. Розробка технології Data Science у сфері оптимізації (метод нейронних мереж)
19. Розробка технології Data Science у сфері синтезу мовлення (метод нейронних мереж)
20. Розробка технології Data Science у сфері скорингових систем (метод градієнтного бустінгу)

# 6. ВИКОНАННЯ ТА ЗАХИСТ ДНП

Підготовка ДНП умовно поділяється на такі етапи:

* установча лекція з методики підготовки ДНП;
* вибір і затвердження напрямку досліджень та теми ДНП;
* підбір та вивчення літературних та інших джерел з теми дослідження;
* розробка концепції та плану ДНП, узгодження їх з науковим керівником;
* виконання досліджень за темою ДНП та написання пояснювальної записки та її оформлення;
* захист ДНП.

Визначення напряму дослідження здійснюється виходячи з наукових інтересів студента, тематики наукових досліджень кафедри, замовлень підприємств тощо.

Обираючи тему, слід враховувати її актуальність для організації, де планується реалізувати проект, можливість одержання відповідних матеріалів (планових, звітних, статистичних), проведення власних спостережень, розрахунків, експериментів, наявність літературних джерел, власні напрацювання.

Тема ДНП повинна бути актуальною, мати теоретичне і прикладне значення, відповідати сучасному стану та перспективам розвитку інформаційної аналітики. Обов’язковою вимогою до ДНП є її корисність і практична цінність.

Пропозиції щодо вибору теми ДНП студент обговорює з науковим керівником від кафедри. За поданням наукових керівників кафедра розглядає пропозиції студентів і після схвалення готує проект відповідного наказу.

Обов’язками керівників є:

* консультування з питань вибору теми ДНП, розробки її плану, добору спеціальної літератури, підготовки окремих розділів;
* видача завдання на роботу із зазначенням термінів виконання кожного розділу;
* контроль за дотриманням регламенту підготовки ДНП;
* контроль якості ДНП.

Неодмінною умовою якісного виконання ДНП є ґрунтовне ознайомлення з інформаційними джерелами за обраною темою. Вивчення літератури є важливим напрямком роботи, який дає можливість студенту врахувати досягнення попередників, що присвятили свої дослідження обраній ним науковій проблемі. Матеріали попередніх досліджень являють собою великий за обсягом матеріал, з якого можна отримувати корисну інформацію. Він може стати підставою для формування нових ідей та науково-світоглядних поглядів.

Після ґрунтовного опрацювання літературних джерел і з’ясування теоретичної бази дослідження студент приступає до ознайомлення з практичною роботою по реалізації конкретного проєкту.

Перед студентом стоїть завдання, виходячи з плану ДНП, конкретизувати *об’єкт та предмет дослідження*.

Під час виконання ДНП згідно з її планом студент звертається до свого керівника, керуючись графіком консультацій, затвердженого кафедрою. Готова та оформлена ДНП захищається студентом не пізніше кінцевого терміну захисту, встановленого кафедрою.

## 6.1 *Підготовка до захисту*

Роздруковану пояснювальну записку до ДНП разом із додатками **зшивають** за допомогою скорозшивача (на дві дірки) з прозорою обкладинкою. За додатками в кінці тексту роботи розміщують прозорий файл з компакт-диском, який має бути підписаним і містити текст самої ДНП, презентацію та розроблені автором компоненти проекту та його проміжні і кінцевий результати. Оформлена належним чином курсова робота підписується автором та керівником роботи.

Оформлена належним чином робота **за тиждень** до визначеної дати захисту подається автором на кафедру для проведення попередньої експертизи та нормативного контролю. Якщо робота пройшла ці етапи – магістр отримує допуск до захисту. Допуск надає завідувач кафедри. Студенту може бути відмовлено у захисті ДНП за невідповідність вимогам, що наведені у даних методичних вказівках або за відсутністю схвалення з боку керівника роботи.

Час та місце проведення захисту ДНП визначається деканатом за погодженням з кафедрою та навчальним відділом, вноситься до розкладу заліково-екзаменаційної сесії та оприлюднюється на стенді факультету разом із складом комісії та переліком студентів, які повинні в ній захищатись.

До захисту ДНП подаються такі матеріали:

* оформлені ДНП (з дотриманням вимог методичних вказівок, з підписом студента та візою керівника на титульному листі);
* залікові книжки студентів (подаються студентами особисто);
* відгуки керівників ДНП;
* екзаменаційні відомості на захист ДНП (надає деканат).

Недотримання умов першого або другого пункту з наведеного вище переліку тягне недопуск студента до захисту. Дозвіл на захист ДНП без залікової книжки може бути надано деканом факультету на підставі мотивованої заяви студента.

Студенти можуть запропонувати додатковий роздатковий матеріал для захисту, а також скористатись мультимедійним проектором, який має бути встановлений на час захисту. Своєчасне замовлення мультимедійного проектора покладається на методиста кафедри.

## 6.2 *Захист ДНП*

Завершену ДНП студенту необхідно подати на кафедру за тиждень до захисту в паперовому та електронному вигляді (в форматі *.doc,* .docx, *.rtf,* .rar, *.zip на карті пам’яті); презентацію − не пізніше, ніж за 2 дні до захисту в паперовому та електронному вигляді у форматі* .pрt, *.pptx,* .pdf. Електронна версія ДНП додається до паперової під час передачі роботи до архіву університету.

Студент готує до захисту ілюстративний матеріал (таблиці, графіки, діаграми, рисунки) для всіх членів комісії.

Захист ДНП студентів відбувається прилюдно згідно з розкладом. Кожному студентові комісія надає заздалегідь обумовлений час на доповідь та відповіді на запитання, який складає **до п’ятнадцяти хвилин**.

Час виступу не повинен перевищувати **10 хвилин**, має бути стислим, конкретним, з використанням ілюстративного матеріалу: плакатів, роздаткового матеріалу (набору таблиць, графіків, діаграм, схем або відповідних слайдів). Головне призначення таких додатків − детально та наочно проілюструвати ті чи інші твердження автора, тому необхідно вчасно посилатися на відповідний матеріал, акцентувати на ньому увагу членів комісії.

У разі використання слайдів або комп’ютерної презентації треба переконатися у наявності проектора, екрана, можливості затемнення, доступних джерел електричного струму та пам’ятати про специфіку підготовки та застосування цього методу презентації. Кількість слайдів презентації − 12-14 аркушів.

Захист ДНП в індивідуальному порядку проводиться за такою самою процедурою.

У випадку захисту ДНП, яка виконувалась кількома студентами за комплексною темою, кожен з них представляє і захищає власний доробок. Оцінка ставиться кожному студенту окремо.

Якщо своєчасний захист ДНП не відбувся з поважної причини, то деканат факультету разом з кафедрою визначає дату захисту ДНП в індивідуальному порядку з урахуванням обставин, які унеможливили її своєчасний захист. Поважність причини визначається деканом факультету на підставі поданих документів.

У випадку несвоєчасного захисту ДНП без поважної причини (отримання негативної оцінки, неявка, недопуск у зв’язку з несвоєчасним поданням роботи чи недопуск у зв’язку з неподанням належних для захисту матеріалів тощо), проводиться її повторний захист у порядку, передбаченому наказом ректора від 15 листопада 2004 року № 219-5.

Результати захисту ДНП визначаються оцінками у відповідності з робочою навчальною програмою дисципліни і оголошуються у цей же день. У Комісію можуть бути подані також інші матеріали, що характеризують наукову та практичну цінність ДНП − друковані або підготовлені до друку статті за темою роботи, тези доповідей на конференціях, документи, які підтверджують практичне застосування результатів, макети, зразки матеріалів, виробів тощо.

Захист роботи розпочинається з доповіді, в якій студент повідомляє тему роботи, яка ним захищається та має розкрити: актуальність теми, зміст роботи, об’єкт та предмет дослідження, характеристика проведених досліджень, зміст запропонованих заходів та обґрунтування їх ефективності, отримані результати, висновки.

Завершуючи доповідь, студент має відзначити: які проблеми були при реалізації проекту; які його рекомендації можуть вплинути на роботу підприємства; яка фактична чи очікувана соціально-економічна ефективність запропонованих ним заходів.

Під час доповіді слід звертатися до ілюстративного матеріалу, коротко пояснюючи його зміст. Члени Комісії, присутні на захисті, можуть ставити студентові запитання з метою визначення рівня його підготовки, ерудиції та обізнаності за темою ДНП.

Після обговорення підсумків захисту на закритому засіданні Комісії виноситься її рішення. У разі незгоди між членами Комісії думка голови є вирішальною. Рішення комісії оголошується її головою в той самий день.

У процесі визначення оцінки враховується низка важливих показників якості ДНП:

1. Змістовні аспекти ДНП:
* актуальність обраної теми дослідження;
* спрямованість роботи на розробку реальних практичних рекомендацій;
* відповідність логічної побудови ДНП поставленим цілям і завданням;
* широта й адекватність методологічного та діагностичного апарату;
* професійний рівень обґрунтування та представлення запропонованих рішень;
* ступінь самостійності проведення дослідження;
* розвиненість мови ДНП та її загальне оформлення.
1. Якість захисту ДНП:
* уміння стисло, послідовно й чітко викласти сутність і результати дослідження;
* здатність аргументовано захищати свої пропозиції, думки, погляди;
* загальний рівень підготовки студента.

# 7. ПЕРЕЛІК ВИКОРИСТАНИХ ЛІТЕРАТУРНИХ ДЖЕРЕЛ

1. Матвієнко О.В., Цивін М.Н. Інформаційний менеджмент: навч. посіб. Київ: Слово, 2007. 200 с. URL: http://iub.at.ua/\_ld/0/2\_64\_kniga\_IM\_Mat.pdf (дата звернення 17.12.2023)
2. Тринько Р. І., Тадник М. Є. Основи теоретичної і прикладної статистики: навч. посіб. Київ: Знання, 201(@) 400 с.
3. Черняк О.І., Захарченко П.В. Інтелектуальний аналіз даних: підручник. Київ: Знання, 2014. 599 с
4. Чубукова И.А. Data Mining: конспект лекций, Киев, 2018. 326 с.
5. Ланде Д.В., Субач І.Ю., Бояринова Ю.Є. Основи теорії і практики інтелектуального аналізу даних у сфері кібербезпеки: навч. посіб. Київ: ІСЗЗІ КПІ ім. Ігоря Сікорського, 2018. 297 с.
6. Data Analysis, Machine Learning and Application / C.Preisach, H.Burkhardt, L.Schmidt-Thieme, R.Decker and etc.; Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation e.V., Albert-Ludwigs-Universität Freiburg, March 7-9, 2007. – Berlin Heidelberg: Springer-Verlag, 2008 – 703 p.

**Додаткова література**

1. Інформаційні системи і технології в економіці: Посібник для студентів вищих навчальних закладів/ За ред. В.С. Пономаренка. Київ: Академія, 2002. 544 с.
2. Інформаційний менеджмент. Менеджмент : веб-сайт. URL: http://ru.osvita.ua/vnz/reports/management/13966 (дата звернення 15.12.2023)
3. Мармоза А. Т. Теорія статистики : підручник. Київ: Центр учбової літератури, 2013. 592 с.
4. Опря А. Т. Статистика. Математична статистика. Теорія статистики : навч. пос. Київ : ЦНЛ, 2005. 496 с.
5. ДСТУ8302:2015 Бібліографічне посилання. Загальні положення та правила складання. [Чинний від 2016-06-01]. Київ : УкрНДНЦ, 2016. 20 с (Інформація та документація).

# Додатки