МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Львівський національний університет імені Івана Франка

Факультет прикладної математики та інформатики

Кафедра прикладної математики

Затверджено

На засіданні кафедри прикладної математики
факультету прикладної математики та інформатики
Львівського національного університету імені Івана Франка

Протокол № ____ від «___» _____ 20__ р.

Завідувач кафедри _____

Силабус з навчальної дисципліни

«Аналіз даних та прогнозування»

що викладається в межах освітньо-професійної програми першого (бакалаврського) рівня вищої освіти для здобувачів четвертого року навчання

Дисципліна вільного вибору студента

Львів — 2026 р.

Назва дисципліни	Аналіз даних та прогнозування
Адреса викладання дисципліни	Львівський національний університет імені Івана Франка, м. Львів
Факультет та кафедра, за якою закріплена дисципліна	Факультет прикладної математики та інформатики; кафедра прикладної математики
Галузь знань, шифр та назва спеціальності	Відповідно до освітньої програми, для якої дисципліна пропонується як вибіркова.
Рівень вищої освіти	Перший — бакалаврський
Цільова аудиторія	Здобувачі четвертого року навчання бакалаврського рівня
Статус дисципліни	Дисципліна вільного вибору студента
Семестр	8 семестр
Викладач дисципліни	Доц. Переймибіда Андрій
Контактна інформація викладача	
Консультації з питань навчання по дисципліні відбуваються	Консультації у день проведення занять або за попередньою домовленістю; очно чи онлайн.
Сторінка курсу	

Інформація про дисципліну

Курс розроблено для формування у студентів практичних навичок аналізу даних, побудови зрозумілих аналітичних висновків і створення базових прогнозів для навчальних, дослідницьких та прикладних задач. Дисципліна поєднує методологію роботи з даними, підготовку та оцінювання якості даних, описову аналітику, візуалізацію, роботу з бізнес-метриками, основи регресійного аналізу, часові ряди, перевірку якості прогнозів і представлення результатів у формі аналітичного звіту або dashboard-рішення.

Особливий акцент робиться на практичному циклі роботи з даними: постановка проблеми, розуміння джерел даних, очищення, дослідницький аналіз, побудова моделі, перевірка результату, інтерпретація обмежень і комунікація висновків для користувача або замовника. Як основний практичний інструмент може використовуватися Power BI, Excel / Google Sheets, Python або інший доступний стек — залежно від технічної підготовки групи та організації курсу.

Коротка анотація дисципліни

Дисципліна «Аналіз даних та прогнозування» є вибірковою дисципліною для здобувачів четвертого року навчання першого (бакалаврського) рівня вищої освіти. Курс спрямований на розвиток data literacy та прикладних аналітичних компетентностей: від розуміння типів даних і якості датасетів до побудови інтерактивних візуалізацій, базових прогнозних моделей і аргументованих рекомендацій. Студенти працюють із відкритими або навчальними наборами даних, формують аналітичні питання, будують моделі, оцінюють похибки та презентують результати у вигляді індивідуального проєкту.

Мета та цілі дисципліни

Мета дисципліни — сформувати у студентів здатність перетворювати дані на обґрунтовані аналітичні висновки та базові прогнози, дотримуючись коректної логіки роботи з даними, принципів валідації результатів і зрозумілої комунікації аналітичних рішень.

Цілі дисципліни: ознайомити студентів з повним циклом аналізу даних; сформувати навички постановки аналітичної задачі; навчити очищувати, структурувати й досліджувати дані; розвинути вміння будувати

візуалізації та dashboard-звіти; сформувати базове розуміння прогнозування, регресії та часових рядів; навчити оцінювати якість прогнозів і уникати типових помилок, зокрема leakage та некоректного використання тестових даних.

Обсяг курсу

Загальний обсяг: 90 годин / 3 кредити ECTS. Аудиторних занять: 42 год., з них 14 год. лекцій та 28 год. практичних / лабораторних занять. Самостійна робота: 48 год.

Очікувані результати навчання

Після завершення цього курсу студент буде:

Знати:

- основні етапи життєвого циклу data-driven задачі: від постановки питання до презентації результатів;
- типи даних, типові проблеми якості даних, пропуски, дублікати, викиди, неузгоджені формати та ризики зміщення;
- принципи дослідницького аналізу даних, описових статистик і базових візуалізацій;
- логіку побудови аналітичного звіту або dashboard-рішення для різних аудиторій;
- основні ідеї кореляційного та регресійного аналізу;
- базові підходи до прогнозування часових рядів: naive, seasonal naive, moving average, трендові та сезонні компоненти;
- поняття тренувальної, валідаційної і тестової вибірки, rolling-origin backtesting, MAE, RMSE, MAPE та інтерпретації похибок;
- типові помилки прогнозування: leakage, підгонка під тестовий період, некоректні агрегування, неузгоджені часові зрізи;
- принципи побудови прикладних сценаріїв best / base / worst для управлінського аналізу.

Вміти:

- формулювати аналітичну задачу, визначати KPI, одиницю аналізу, періодичність і очікуваний формат результату;
- готувати дані до аналізу: очищувати, структурувати, агрегувати, перевіряти повноту та консистентність;
- виконувати EDA та робити обґрунтовані висновки на основі описових статистик і графіків;
- створювати зрозумілі аналітичні візуалізації та базові dashboard-звіти;
- будувати прості регресійні та прогнозні моделі, пояснювати їхні припущення й обмеження;
- проводити перевірку якості прогнозу на історичних даних;
- показувати невизначеність прогнозу через операційні інтервали або сценарії;
- формувати рекомендації на основі даних і презентувати їх у стислому аналітичному форматі;
- документувати джерела даних, трансформації, припущення та обмеження аналізу.

Комунікаційні та soft skills:

- критичне мислення та інформаційна грамотність;
- вміння пояснювати числові результати не технічній аудиторії;
- уважність до якості даних і коректності інтерпретацій;
- презентація висновків і рекомендацій;
- робота з невизначеністю та обмеженнями моделей;
- відповідальність за прозорість припущень і джерел даних.

Ключові слова

Аналіз даних, data-driven рішення, прогнозування, часові ряди, регресія, EDA, візуалізація даних, dashboard, Power BI, KPI, якість даних, rolling-origin backtesting, MAE, RMSE, MAPE, сценарний аналіз, відкриті дані.

Формат курсу

Формат курсу: очний / змішаний / дистанційний — відповідно до організації навчального процесу.

Передбачено:

- лекційні заняття з демонстраціями;
- практичні / лабораторні роботи з наборами даних;
- короткі індивідуальні аналітичні завдання;
- роботу з відкритими або навчальними датасетами;
- створення візуалізацій і dashboard-сторінок;
- побудову базових прогнозів і перевірку якості моделей;
- обговорення типових помилок аналізу та прогнозування;
- підсумковий індивідуальний проєкт.

Теми

1. Вступ до курсу. Data-driven мислення та постановка аналітичної задачі — дані як основа рішень, KPI, одиниця аналізу, питання до даних, очікуваний результат.
2. Життєвий цикл аналізу даних і CRISP-DM — розуміння бізнес / навчальної задачі, розуміння даних, підготовка даних, моделювання, оцінювання, впровадження результатів.
3. Типи даних і якість даних — категоріальні та числові ознаки, часові дані, пропуски, дублікати, викиди, помилки форматів, data validation.
4. Описова статистика та EDA — розподіли, центральні тенденції, варіативність, групування, сегменти, перші аналітичні висновки.
5. Візуалізація даних і комунікація інсайтів — вибір графіків, когнітивне навантаження, сторітелінг, типові помилки візуалізації.
6. Моделювання даних для BI-аналітики — факти й довідники, гранулярність, зв’язки, календарна таблиця, підготовка моделі до звітування.
7. Побудова dashboard-звіту — структура сторінок, executive overview, деталізація, drill-through, фільтри, дата останнього оновлення моделі / даних.
8. Регресійне мислення в аналізі даних — залежна та незалежні змінні, кореляція, лінійна регресія, інтерпретація коефіцієнтів, обмеження причинно-наслідкових висновків.
9. Вступ до часових рядів — тренд, сезонність, шум, агрегування за періодами, останній повний період, ризики неповних місяців / тижнів.
10. Базові методи прогнозування — naive, seasonal naive, moving average, weighted moving average, трендові припущення, практична інтерпретація.
11. Оцінювання якості прогнозу — train / validation / test, rolling-origin backtesting, MAE, RMSE, MAPE, порівняння моделей, baseline як обов’язкова точка відліку.
12. Невизначеність прогнозу та сценарний аналіз — операційні інтервали через MAE або квантилі помилок, best / base / worst сценарії, обмеження прогнозів у dashboards.
13. Типові помилки та leakage у прогнозуванні — використання майбутньої інформації, підгонка параметрів, некоректні moving averages, порушення cutoff-логіки.
14. Підсумковий проєкт — презентація студентських робіт на відкритих або навчальних даних, обговорення якості даних, моделей, прогнозів, обмежень і практичних рекомендацій.

Підсумковий контроль, форма

Форма підсумкового контролю: залік. Залік формується за результатами виконання практичних індивідуальних завдань, активності на заняттях, мініаналізів / рефлексій і підсумкового індивідуального проєкту.

Пререквізити

Для вивчення курсу студентам бажано мати базові навички роботи з комп'ютером, електронними таблицями та онлайн-сервісами, базове розуміння математичної статистики, досвід роботи з навчальними або відкритими наборами даних, навички пошуку інформації в інтернеті та готовність критично аналізувати числові результати. Спеціальні знання з машинного навчання не є обов'язковими.

Навчальні методи та техніки, які будуть використовуватися під час викладання курсу

- лекції з демонстраціями аналітичного циклу;
- практичні завдання з підготовки та аналізу даних;
- робота з відкритими наборами даних;
- порівняння альтернативних візуалізацій і моделей;
- мініпроєкти та peer-review результатів;
- обговорення помилок інтерпретації та прогнозування;
- захист індивідуального підсумкового проєкту.

Необхідне обладнання

- комп'ютер або ноутбук;
- доступ до інтернету;
- браузер;
- Power BI Desktop або інший погоджений інструмент аналізу даних;
- Excel / Google Sheets або інші засоби роботи з табличними даними;
- за потреби — Python / Google Colab / BigQuery / Looker Studio;
- проєктор / екран для демонстрацій;
- навчальна платформа Moodle / Teams / Google Classroom або інша.

Критерії оцінювання

Оцінювання проводиться за 100-бальною шкалою. Бали нараховуються за наступним співвідношенням:

Вид діяльності	**Максимальна кількість балів**
Практичні індивідуальні завдання	50
Мініаналізи / рефлексії / порівняння підходів та інструментів	15
Активність на заняттях і участь в обговореннях	10
Підсумковий індивідуальний проєкт	25
Разом	**100**

Практичні індивідуальні завдання можуть включати:

1. Постановку аналітичної задачі та опис набору даних.
2. Очищення та первинну перевірку якості даних.
3. Побудову EDA-звіту з описовими статистиками та візуалізаціями.
4. Створення dashboard-сторінки з KPI, фільтрами та датою останнього оновлення моделі / даних.
5. Побудову базового прогнозу та порівняння з baseline.
6. Оцінювання похибки прогнозу та формування операційного інтервалу або сценаріїв best / base / worst.

Підсумковий проєкт має показати, як студент може використати дані для реальної аналітичної або прогнозної задачі. Можливі формати: dashboard на відкритих даних; прогноз попиту, відвідуваності, продажів або іншого KPI; аналіз освітніх, соціальних, економічних або міських відкритих даних; порівняння кількох підходів до прогнозування; аналітичний звіт з рекомендаціями та описом обмежень.

Політика академічної доброчесності

Очікується, що всі роботи студентів будуть результатом їхнього власного аналізу, критичного осмислення та самостійного прийняття рішень. Використання зовнішніх інструментів, шаблонів, AI-асистентів або готових фрагментів коду допускається лише за умови прозорого документування, пояснення логіки використання та здатності студента захистити власний результат.

- вказувати джерела даних і описувати виконані трансформації;
- не фабрикувати дані, джерела або результати моделей;
- не подавати чужий dashboard, код, аналітичний текст або прогноз як власну роботу;
- перевіряти коректність формул, моделей і висновків;
- не використовувати інструменти автоматизації для обходу навчальної мети завдання.

Порушення академічної доброчесності, зокрема списування, фальсифікація результатів, некритичне копіювання згенерованих пояснень або подання чужої роботи як власної, є підставою для незарахування відповідного завдання.

Політика відвідування

Відвідування занять є важливою складовою навчання, оскільки значна частина курсу базується на практичних демонстраціях, обговоренні прикладів і спільному аналізі результатів. Студенти мають дотримуватися строків виконання завдань. У разі неможливості відвідати заняття або вчасно виконати завдання студент повинен повідомити викладача заздалегідь.

Політика використання AI-інструментів

У межах дисципліни AI може використовуватися як допоміжний інструмент для пояснення понять, генерації ідей, коментування коду, пошуку можливих помилок, формування чернеток тексту або структури презентації. Остаточні висновки, моделі, перевірка даних, інтерпретація похибок і рекомендації залишаються відповідальністю студента.

Студент повинен позначати використання AI-інструментів у роботах, якщо вони вплинули на зміст, структуру, код або висновки. Не допускається використання AI для фабрикації джерел, даних, результатів або для повної заміни власного аналізу.

Література для вивчення дисципліни

Основна література

1. Provost F., Fawcett T. Data Science for Business. O’Reilly Media.
2. Wickham H., Grolemund G. R for Data Science. O’Reilly Media.
3. James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning.
4. Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice.
5. Few S. Information Dashboard Design.
6. Microsoft Learn. Power BI documentation and learning paths.
7. Офіційні матеріали з відкритих даних України та набори даних data.gov.ua.

8. Навчальні матеріали курсу, демонстраційні датасети та приклади, надані викладачем.

Питання до заліку

Що означає data-driven підхід до прийняття рішень?

Які етапи має типовий життєвий цикл задачі аналізу даних?

Чим відрізняються дані, інформація, інсайт і рекомендація?

Що таке KPI і чому важливо правильно визначити одиницю аналізу?

Які типові проблеми якості даних виникають у наборах даних?

Як обробляти пропуски, дублікати та викиди?

Які задачі вирішує EDA?

Як вибрати відповідний тип візуалізації для аналітичного питання?

Які принципи побудови ефективного dashboard-звіту?

Навіщо у звіті показувати дату останнього оновлення моделі або даних?

Що таке кореляція і чому вона не доводить причинність?

Яка базова ідея лінійної регресії?

Що таке часовий ряд?

Чим тренд відрізняється від сезонності?

Чому важливо працювати з останнім повним періодом?

Що таке baseline forecast?

Як працюють naive, seasonal naive і moving average підходи?

Що таке rolling-origin backtesting?

Як інтерпретувати MAE, RMSE і MAPE?

Що таке leakage у прогнозуванні?

Як побудувати операційний інтервал прогнозу на основі історичних похибок?

У чому різниця між best / base / worst сценаріями?

Які обмеження має прогноз у dashboard-звіті?

Як презентувати аналітичний результат для не технічної аудиторії?

Опитування

Анкету-оцінку з метою оцінювання якості курсу буде надано студентам після завершення дисципліни. Метою опитування є збір зворотного зв’язку щодо змісту курсу, практичної цінності завдань, якості навчальних матеріалів, обраних інструментів і доцільності формату підсумкового проєкту.