

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Львівський національний університет імені Івана Франка
Механіко-математичний факультет
Кафедра математичної статистики і диференціальних рівнянь



Затверджено

На засіданні
кафедри математичної статистики і
диференціальних рівнянь
механіко-математичного факультету
Львівського національного університету
імені Івана Франка
(протокол № 1 від 22.06.2023 р.)

Завідувач кафедри: Олег БУГРІЙ

A handwritten signature in black ink, appearing to be "Bugriy", written over a horizontal line.

Силабус з навчальної дисципліни
“ Предиктивна аналітика ”,
що викладається в межах ОПШ “ Статистичний аналіз даних ”
першого (бакалаврського) рівня вищої освіти для здобувачів з
спеціальності 112 - Статистика

Львів 2023 р.

Назва дисципліни	Предиктивна аналітика
Адреса викладання дисципліни	Головний корпус ЛНУ ім. І. Франка м. Львів, вул. Університетська 1, 79000
Факультет та кафедра, за якою закріплена дисципліна	Механіко-математичний факультет Кафедра математичної статистики і диференціальних рівнянь
Галузь знань, шифр та назва спеціальності	11 - Математика та статистика 112 - Статистика
Викладачі дисципліни	Доманська О.В., кандидат фізико-математичних наук, доцент кафедри математичної статистики і диференціальних рівнянь
Контактна інформація викладачів	olena.domanska@lnu.edu.ua , https://new.mmf.lnu.edu.ua/employee/domanska-o-v ;
Консультації з питань навчання по дисципліні відбуваються	Консультації в день проведення лекцій/практичних занять (за попередньою домовленістю). Головний корпус ЛНУ ім. І. Франка, каб. 267. м. Львів, вул. Університетська, 1
Сторінка курсу	https://new.mmf.lnu.edu.ua/course/preda-112-bak23
Інформація про дисципліну	Дисципліна “Предиктивна аналітика” є вибірковою дисципліною зі спеціальності 112 – Статистика для освітньої програми “Статистичний аналіз даних”, яка викладається в 6-му семестрі в обсязі 4-х кредитів (за Європейською Кредитно-Трансферною Системою ECTS).
Коротка анотація дисципліни	Курс розроблено для ознайомлення студентів з основами методів та підходів предиктивної аналітики.
Мета та цілі дисципліни	<i>Мета:</i> ознайомити з основними поняттями та методами побудови моделей предиктивної аналітики. <i>Цілі:</i> викласти основи предиктивної аналітики.
Література для вивчення дисципліни	1) Alvaro Fuentes. <i>Hands-On Predictive Analytics with Python</i> . Packt Publishing, 2018. 2) John D. Kelleher, Brian Mac Namee, Aoife D'Arcy. <i>Fundamentals of Machine Learning for Predictive Data Analytics</i> , 2 nd ed. MIT Press, 2020. 3) Eugene Charniak. <i>Introduction to Deep Learning</i> . MIT Press, 2019. 4) François Chollet. <i>Deep Learning with Python, Second Edition</i> . Manning, 2021. 5) Ovidiu Calin. <i>Deep Learning Architectures: A Mathematical Approach</i> . Springer, 2020. 6) Aston Zhang, Zachary C. Lipton, Mu Li, Alexander J. Smola. <i>Dive into Deep Learning</i> , 1 st ed. Cambridge University Press, 2023. d2l.ai
Обсяг курсу	Загальний обсяг: 120 годин. Аудиторних занять: 64 год., з них 32 год. лекційних та 32 годин лабораторних занять. Самостійної роботи: 56 год.
Очікувані	У результаті вивчення даного курсу студент буде:

результати навчання	знати: формулювання основних задач предиктивної аналітики вміти: використовувати алгоритми предиктивної аналітики
Ключові слова	Предиктивна аналітика, глибоке навчання.
Формат курсу	Очний
Теми	Див. Схема курсу
Підсумковий контроль, форма	Залік
Пререквізити	Для вивчення даного курсу студенти потрібні базові знання з: - Програмування; - Машинного навчання.
Навчальні методи та техніки, які будуть використовуватися під час викладання курсу	Інформаційні методи (лекція, бесіда, ілюстрація, демонстрація); дедуктивні методи на основі узагальнень; евристичні методи (проблемна лекція); інтерактивні методи (дискусія)
Необхідне обладнання	Для проведення лекційних занять: комп'ютер (мінімальні характеристики: процесор Intel Core i3, 4ГБ оперативної пам'яті), доступ до мережі Internet, засоби мультимедіа (в т.ч. проектор). Для проведення практичних/лабораторних занять: комп'ютер (мінімальні характеристики: процесор Intel Core i3, 4ГБ оперативної пам'яті), доступ до мережі Internet. Необхідне програмне забезпечення включає в себе ОС Windows 10, програмні додатки (MS Teams, MS Excel, Jupyter Notebook з вбудованим компілятором мови програмування Python).
Критерії оцінювання (окремо для кожного виду навчальної діяльності)	Оцінювання проводиться за 100-бальною шкалою. Бали нараховуються за наступним співвідношенням: • Змістовий модуль 1: 10% семестрової оцінки за активну роботу на заняттях, 30% семестрової оцінки за виконання практичних аудиторних і домашніх завдань, максимальна кількість балів 40. • Змістовий модуль 2: 10% семестрової оцінки за активну роботу на заняттях, 30% семестрової оцінки за виконання практичних аудиторних і домашніх завдань, максимальна кількість балів 40. • Контрольне тестування: 20% семестрової оцінки, максимальна кількість балів 20. Підсумкова максимальна кількість балів 100. Академічна доброчесність: Очікується, що роботи студентів будуть оригінальними дослідженнями чи міркуваннями. Списування та втручання в роботу інших студентів становлять, але не обмежують, приклади можливої академічної недоброчесності. Виявлення ознак академічної недоброчесності в написанні завдань є підставою для її незарахування викладачем, незалежно від масштабів плагіату чи обману. Жодні форми порушення академічної доброчесності не толеруються. Відвідання занять є важливою складовою навчання. Очікується, що всі студенти відвідають усі лекції та практичні/лабораторні заняття курсу. Студенти повинні інформувати викладача про неможливість відвідати

заняття. У будь-якому випадку студенти зобов'язані дотримуватися термінів визначених для виконання всіх видів робіт, передбачених курсом.

Література. Уся література, яку студенти не зможуть знайти самостійно, буде надана викладачем виключно в освітніх цілях без права її передачі третім особам. Студенти заохочуються до використання також й іншої літератури та джерел, яких немає серед рекомендованих.

Політика виставлення балів. Враховуються бали, набрані при поточному контролі та бали підсумкового тестування. При цьому обов'язково враховуються присутність на заняттях та активність студента під час практичного заняття; недопустимість пропусків та запізнь на заняття; користування мобільним телефоном, планшетом чи іншими мобільними пристроями під час заняття в цілях не пов'язаних з навчанням; списування та плагіат; несвоєчасне виконання поставленого завдання і т. ін.

Оцінювання практичних робіт відбувається шляхом оцінки роботи студента під час проведення практичної роботи в аудиторії та захисту написаної студентом вдома практичної роботи.

Бали оцінювання аудиторного виконання практичних робіт нараховуються за наступним співвідношенням:

100% – студент в повному обсязі володіє навчальним матеріалом, має повне розуміння розглянутої теми, надає правильні відповіді на запитання по темі, код програми функціонує відповідно до завдання;

75% – студент достатньо розуміє розглянутий матеріал та принципи написаного ним коду програми, присутні неточності та незначні помилки у відповідях на запитання по темі, код програми функціонує відповідно до завдання;

50% – студент не досить добре розуміє розглянутий матеріал та написаний ним код програми, вагається та надає неточні/не конкретні відповіді на запитання по темі, код програми функціонує з помірними недоліками;

25% – студент погано розуміє розглянутий матеріал та написаний ним код програми, студент в більшості надає помилкові відповіді на питання по темі, код програми не функціонує належним чином;

0 - студент зовсім не засвоїв розглянутий матеріал, написаний ним код програми не відповідає темі/не функціонує взагалі.

Бали оцінювання домашнього завершення виконання практичних робіт та наданого звіту нараховуються за наступним співвідношенням:

100% – звіт цілком і повністю відображає індивідуальне завдання студента, містить правильні висновки, ілюстрований (за потреби) відповідними графіками і таблицями які правильно відображають суть виконаного завдання, студент має повне розуміння розглянутої теми, надає правильні відповіді на запитання по темі, код програми функціонує відповідно до завдання;

75% – звіт в достатній мірі відображає індивідуальне завдання студента, містить допустимі висновки, ілюстрований (за потреби) відповідними графіками і таблицями які частково відображають суть виконаного завдання, студент достатньо розуміє принципи написаного ним коду програми, присутні неточності та незначні помилки у відповідях на запитання по темі, код програми функціонує відповідно до завдання;

50% – звіт містить загальні формулювання завдання, висновки нечіткі, необхідні ілюстрації чи таблиці відсутні, студент не досить добре розуміє

	<p>розглянутий матеріал та представлений код програми, надає неточні/не конкретні відповіді на запитання по темі, код програми функціонує з помірними недоліками; 25% – звіт не містить формулювання завдання, висновки необґрунтовані чи неповні, необхідні ілюстрації чи таблиці відсутні, студент погано розуміє розглянутий матеріал та представлений код програми, студент в більшості надає помилкові відповіді на питання по темі, код програми не функціонує належним чином; 0 – звіт відсутній/не відповідає темі, студент зовсім не засвоїв розглянутий матеріал, написаний ним код програми не відповідає темі/не функціонує взагалі.</p> <p>Оцінювання контрольного тестування (завдання з тематики кожного змістового модуля) відбувається шляхом оцінки письмових відповідей студента на поставлені запитання. Відсотки нарахування балів оцінювання відповіді на кожне запитання нараховуються за наступним співвідношенням: 75-100% – тема відтворюється в повному обсязі, правильно, обґрунтовано, логічно; 50-75% – відтворюється значна частина розглянутої теми, проте присутні неточності та/або невідповідності; 25-50% – виявлено множинні неточності та невідповідності, пояснення відсутні чи частково помилкові; 0-25% – тему майже не розкрито, кількість викладеного матеріалу не відповідає загальним нормам обраного виду роботи.</p> <p>Критерії оцінювання результатів неформальної освіти: Нарахування балів відбувається за публікацію студентом тез доповідей на конференціях, наукових статей, за участь студента у діяльності наукових гуртків, семінарів, круглих столів, конкурсів, участь у заходах неформальної освіти, за отримання сертифікатів про проходження навчання на різних освітніх платформах (Coursera, Prometheus тощо), курсах провідних ІТ компаній за тематикою навчальної дисципліни. Кількість балів визначається відсотком покриття результатів відповідної активності до вимог результатів навчання з навчальної дисципліни</p>
Питання для контрольного тестування	методи зведення тексту до числового вектора, підходи до автокорекції та автозаповнення незакінчених речень, підходи до перекладу текстів з однієї мови на іншу.
Опитування	Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.

**Схема курсу “Предиктивна аналітика”
для студентів спеціальності 112 - Статистика**

Тижні	Лекційний курс		Лабораторні заняття		К-сть год СР	Літе- ратура
	Назва теми	К-сть год	Назва теми	К-сть год		
1	2	3	4	5	6	7
1	Основні поняття предиктивної аналітики. Оцінка ризиків. Діагностика. Класифікація. Моделювання схильності.	2	Вступне заняття. Налаштування необхідного програмного середовища.	2	4	[1]-[6], Сайт курсу
2	Збір даних. Постановка аналітичної задачі. Дизайн структур даних. Визначення критеріїв оцінки аналітичної моделі.	2	Міні-проекти по збору фінансових даних, оцінці кредитних ризиків	2	3	[1] -[6], Сайт курсу
3	Аналіз набору даних. Нормалізація даних. Обрахування різноманітних статистичних показників. Візуалізація.	2	Міні-проекти по медичній діагностиці	2	4	[1] -[6], Сайт курсу
4	Практична підготовка до моделювання. Train-test split. Dimensionality reduction using PCA. Стандартизація.	2	Практичні завдання по візуалізації різноманітних характеристик набору даних	2	3	[1] -[6], Сайт курсу
5	Передбачення чисельних значень. Множинна регресія. Метод лассо. KNN метод.	2	Задачі по однофакторному і багатофакторному аналізу даних	2	4	[1] -[6], Сайт курсу
6	Класифікація. Логістична регресія. Прості і складні моделі логістичної регресії. Мультиноміальна логістична регресія.	2	Практичні завдання із використанням scikit-learn	2	3	[1] -[6], Сайт курсу
7	Класифікаційні дерева. Random forests. Gradient boosting.	2	Задачі на порівняння різних моделей	2	4	[1] -[6], Сайт курсу

	Рекурсивний бінарний поділ. Bagging та boosting.					
8	Наївні Баєсові класифікатори. Гаусовий Баєс. Баєсові мережі.	2	Практичні задачі на передбачення категорій із використанням логістичної регресії	2	3	[1] -[6], Сайт курсу
9	Регресія з нейронними мережами.	2	Практичні задачі на передбачення категорій із використанням логістичної регресії	2	4	[1] -[6], Сайт курсу
10	Класифікація з нейронними мережами.	2	Вступ до GraphViz	2	3	[1] -[6], Сайт курсу
11	Оцінка регресивних моделей.	2	Задачі на візуалізацію дерев	2	4	[1] -[6], Сайт курсу
12	Оцінка класифікаційних моделей.	2	Задачі на Баєсову класифікацію	2	3	[1] -[6], Сайт курсу
13	Оптимізація гіперпараметрів.	2	Практичні задачі на передбачення з використанням нейронних моделей	2	4	[1] -[6], Сайт курсу
14	Оптимізація моделей.	2	Практичні задачі на передбачення з використанням нейронних моделей	2	3	[1] -[6], Сайт курсу
15	Презентація аналітичних результатів	2	Порівняння моделей	2	4	[1] -[6], Сайт курсу
16	Презентація аналітичних результатів	2	Підсумкове заняття	2	3	[1] -[6], Сайт курсу
	Разом	32		32	56	
	Викладач: Доманська О.В.		Викладач: Доманська О.В.			