


**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 28.08.2020 р.)

В.о. завідувача кафедри:

 \_\_\_\_\_ Олег БУГРИЙ

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

Львів 2020 р.

<b>Назва дисципліни</b>	Навчання з підкріпленням
<b>Адреса викладання дисципліни</b>	Головний корпус ЛНУ ім. І. Франка м. Львів, вул. Університетська 1, 79000
<b>Факультет та кафедра, за якою закріплена дисципліна</b>	Механіко-математичний факультет Кафедра математичної статистики і диференціальних рівнянь
<b>Галузь знань, шифр та назва спеціальності</b>	11 - Математика та статистика 112 – Статистика
<b>Викладачі дисципліни</b>	Доманська О.В., кандидат фізико-математичних наук, доцент кафедри математичної статистики і диференціальних рівнянь
<b>Контактна інформація викладачів</b>	<a href="mailto:olena.domanska@lnu.edu.ua">olena.domanska@lnu.edu.ua</a> , <a href="https://new.mmf.lnu.edu.ua/employee/domanska-o-v">https://new.mmf.lnu.edu.ua/employee/domanska-o-v</a> ;
<b>Консультації з питань навчання по дисципліні відбуваються</b>	Консультації в день проведення лекцій/практичних занять (за попередньою домовленістю). Головний корпус ЛНУ ім. І. Франка, каб. 267. м. Львів, вул. Університетська, 1
<b>Інформація про дисципліну</b>	Дисципліна “Навчання з підкріпленням” є вибірковою дисципліною зі спеціальності 112 – Статистика для освітньої програми “Статистичний аналіз даних”, яка викладається в 8-му семестрі в обсязі 4,5 кредитів (за Європейською Кредитно-Трансферною Системою ECTS).
<b>Коротка анотація дисципліни</b>	Курс розроблено для ознайомлення студентів з основами методів та підходів навчання з підкріпленням.
<b>Мета та цілі дисципліни</b>	<i>Мета:</i> ознайомити з основними поняттями та методами навчання з підкріпленням. <i>Цілі:</i> викласти основні алгоритми навчання з підкріпленням, навчити порівнювати та обирати найбільш ефективні методи навчання з підкріпленням для різних типів задач.
<b>Література для вивчення дисципліни</b>	1) Enes Bilgin. <i>Mastering Reinforcement Learning with Python</i> . Packt Publishing, 2020. 2) Richard S. Sutton, Andrew G. Barto. <i>Reinforcement Learning: An Introduction, 2<sup>nd</sup> edition</i> . Bradford Books, 2018. 3) Alexander Zai, Brandon Brown. <i>Deep Reinforcement Learning in Action</i> . Manning, 2020. 4) Laura Graesser, Wah Loon Keng. <i>Foundations of Deep Reinforcement Learning: Theory and Practice in Python</i> . Addison-Wesley Professional, 2019. 5) Phil Winder Ph.D. <i>Reinforcement Learning: Industrial Applications of Intelligent Agents</i> . O'Reilly Media, 2020.
<b>Обсяг курсу</b>	Загальний обсяг: 135 годин. Аудиторних занять: 78 год., з них 39 год. лекційних та 39 годин лабораторних занять. Самостійної роботи: 57 год.
<b>Очікувані</b>	У результаті вивчення даного курсу студент буде:

<b>результати навчання</b>	<b>знати:</b> формулювання основних задач та методів навчання з підкріпленням; <b>вміти:</b> використовувати алгоритми навчання з підкріпленням, такі як бандити, Марковські процеси вирішування, та різні методи глибокого навчання з підкріпленням із використанням нейронних мереж.
<b>Ключові слова</b>	Навчання з підкріпленням, глибоке навчання.
<b>Формат курсу</b>	Очний
<b>Теми</b>	Див. <b>Схема курсу</b>
<b>Підсумковий контроль, форма</b>	Залік
<b>Пререквізити</b>	Для вивчення даного курсу студенти потрібні базові знання з: - Програмування; - Машинного навчання.
<b>Навчальні методи та техніки, які будуть використовуватися під час викладання курсу</b>	Інформаційні методи (лекція, бесіда, ілюстрація, демонстрація); дедуктивні методи на основі узагальнень; евристичні методи (проблемна лекція); інтерактивні методи (дискусія)
<b>Необхідне обладнання</b>	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).
<b>Критерії оцінювання (окремо для кожного виду навчальної діяльності)</b>	Оцінювання проводиться за 100-бальною шкалою. Бали нараховуються за наступним співвідношенням: • Змістовий модуль 1: 8% семестрової оцінки за активну роботу на заняттях, 32% за виконання домашніх завдань, максимальна кількість балів 40. • Змістовий модуль 2: 8% семестрової оцінки за активну роботу на заняттях, 32% за виконання домашніх завдань, максимальна кількість балів 40. • Контрольне оцінювання: 20% семестрової оцінки, максимальна кількість балів 20. Підсумкова максимальна кількість балів 100. <b>Академічна доброчесність:</b> 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. <b>Відвідання занять</b> 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** (2 змістових модулі містять по 8 практичних роботи кожен, загалом 16 практичних робіт, максимальна кількість балів: 80) відбувається шляхом оцінки роботи студента під час проведення практичної роботи в аудиторії (0-2 балів за одну роботу) та захисту написаної студентом вдома практичної роботи (0-2 балів за одну роботу). До 1 бала студенти можуть отримати за активну роботу на заняттях.

Бали оцінювання аудиторного виконання практичних робіт нараховуються за наступним співвідношенням:

2 – 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;

1.5 – 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;

1 – 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;

0.5 – 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:

2 – 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;

1.5 – звіт в достатній мірі відображає індивідуальне завдання студента, містить допустимі висновки, ілюстрований (за потреби) відповідними

	<p>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;</p> <p>1 – звіт містить загальні формулювання завдання, висновки нечіткі, необхідні ілюстрації чи таблиці відсутні, студент не досить добре розуміє розглянутий матеріал та представлений код програми, надає неточні/не конкретні відповіді на запитання по темі, код програми функціонує з помірними недоліками;</p> <p>0.5 – 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;</p> <p>0 – звіт відсутній/не відповідає темі, студент зовсім не засвоїв розглянутий матеріал, написаний ним код програми не відповідає темі/не функціонує взагалі.</p> <p><b>Оцінювання контрольного тестування</b> 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><b>Критерії оцінювання результатів неформальної освіти:</b>  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>
<b>Питання для контрольного тестування</b>	Навчання з підкріпленням
<b>Опитування</b>	Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.

**Схема курсу “ Навчання з підкріпленням ”  
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 навчання з підкріпленням.	3	Налаштування робочого середовища. Огляд застосувань.	3	4	[1]-[5], Сайт курсу
2	Багаторукі бандити. Контекстуальні бандити.	3	Практичні завдання по темі. Вправи на епсилон-жадібність, аналіз їх переваг і недоліків. Гаусові винагороди.	3	4	[1]-[5], Сайт курсу
3	Марковський процес вирішування.	3	Розв'язування Марковського процесу вирішування. Обчислення n-крокових переходів. Оцінка значень стану (state-values).	3	4	[1]-[5], Сайт курсу
4	Динамічне програмування.	3	Алгоритм ітерації значень. Написання коду із використанням policy iteration та value iteration алгоритмів.	3	4	[1]-[5], Сайт курсу
5	Методи Монте-Карло.	3	Тренування агента методами Монте-Карло. Розв'язання задач контролю методами Монте-Карло.	3	4	[1]-[5], Сайт курсу
6	Навчання методами часової різниці.	3	Розв'язання задач контролю методами часової різниці. Порівняння методів Монте-Карло та часової різниці. Практичні завдання на використання SARSA.	3	4	[1]-[5], Сайт курсу
7	N-крокові методи часової різниці. Апроксимація функцій.	3	Розв'язання задач навчання із використанням n-крокового SARSA. Практичні завдання на методи апроксимації.	3	4	[1]-[5], Сайт курсу
8	Вступ у глибоке навчання з підкріпленням. Policy gradient алгоритми.	3	Реалізація алгоритмів глибого Q-навчання. Бібліотека RLlib. Написання коду, що реалізує методи policy gradient, із використанням RLlib.	3	4	[1]-[5], Сайт курсу
9	Методи актор-критик. Trust-region методи.	3	Вправи на реалізацію алгоритмів A2C та A3C. Вправи на імплементацію алгоритмів TRPO та PPO.	3	5	[1]-[5], Сайт курсу
10	Off-policy методи.	3	Практичні завдання на	3	5	[1]-[5], Сайт курсу

			реалізацію DDPG, TD3, SAC, IMPALA.			
11	Model-based методи.	3	Реалізація model-based агента. Порівняння поведінки model-free та model-based агентів.	3	5	[1]-[5], Сайт курсу
12	Багатоагентне навчання з підкріпленням.	3	Вправи на тренування багатьох агентів. Тренування гри у хрестики-нулики.	3	5	[1]-[5], Сайт курсу
13	Підсумкове заняття.	3	Підсумкове заняття. Перевірка знань та вмінь студентів.	3	5	[1]-[5], Сайт курсу
	<b>Разом</b>	<b>39</b>		<b>39</b>	<b>57</b>	
	Викладач: Доманська О.В.		Викладач: Доманська О.В.			