

Міністерство освіти і науки України  
Тернопільський національний технічний університет імені Івана Пулюя  
(повне найменування вищого навчального закладу)  
Факультет комп'ютерно-інформаційних систем і програмної інженерії  
(назва факультету)  
Кафедра комп'ютерних систем та мереж  
(повна назва кафедри)

# КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

**магістра**

(освітній ступінь)

на тему: **Методи і засоби інтелектуалізованого аналізу та прогнозування  
успішності студентів у системах електронного навчання**

Виконав: студент (ка) 6 курсу, групи СІМ-61  
спеціальності 123 «Комп'ютерна інженерія»  
(шифр і назва спеціальності)

	<hr/>	<b>Дармопук Д.В.</b> (прізвище та ініціали)
Керівник	<hr/>	<b>Жаровський Р.О.</b> (прізвище та ініціали)
Нормоконтроль	<hr/>	<b>Луцик Н.С.</b> (прізвище та ініціали)
Завідувач кафедри	<hr/>	<b>Осухівська Г.М.</b> (прізвище та ініціали)
Рецензент	<hr/>	<b>Мацюк О.В.</b> (прізвище та ініціали)

Тернопіль  
2021

Міністерство освіти і науки України  
Тернопільський національний технічний університет імені Івана Пулюя

(повне найменування вищого навчального закладу)

Факультет комп'ютерно-інформаційних систем і програмної інженерії

Кафедра комп'ютерних систем та мереж

**ЗАТВЕРДЖУЮ**

Завідувач кафедри Осухівська Г.М.

«\_\_\_\_\_» \_\_\_\_\_ 2021 р.

**ЗАВДАННЯ**  
**НА КВАЛІФІКАЦІЙНУ РОБОТУ**

на здобуття освітнього ступеня магістр  
(назва освітнього ступеня)

за спеціальністю 123 «Комп'ютерна інженерія»  
(шифр і назва спеціальності)

студенту Дармопуку Дмитру Володимировичу  
(прізвище, ім'я, по-батькові)

1. Тема проекту (роботи) Методи і засоби інтелектуалізованого аналізу та прогнозування успішності студентів у системах електронного навчання

Керівник проекту (роботи) Жаровський Руслан Олегович, к.т.н.  
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «28» жовтня 2021 року № 4/7-916

2. Термін подання студентом завершеної роботи \_\_\_\_\_  
3. Вихідні дані до роботи Показник успішності студентів, критерії систем електронного навчання, методи машинного навчання для прогнозування успішності студентів

4. Зміст роботи (перелік питань, які потрібно розробити)

Вступ. 1. Аналіз систем електронного навчання та факторів впливу на успішність студентів

2. Визначення та формалізація факторів впливу і методів прогнозування успішності студентів

3. Програмна реалізація моделей прогнозування успішності студентів

4. Охорона праці та безпека в надзвичайних ситуаціях. Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

1. Актуальність і мета дослідження. 2. Задачі дослідження, об'єкт і предмет, наукова новизна і практична цінність дослідження. 3. Властивості основних функціональних модулів СДН

4. Формалізована модель факторів впливу на успішність студентів

5,6 Методи сегментації з використанням машинного навчання

7. Прогнозування успішності студентів на основі дерев прийняття рішень та методу опорних векторів. 8. Прогнозування успішності студентів на основі моделі GritNet. 9. Висновки

\_\_\_\_\_

\_\_\_\_\_

## 6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
<i>Охорона праці та безпека в надзвичайних ситуаціях</i>	<i>Осухівська Г.М.</i>		
	<i>Стадник І.Я.</i>		

7. Дата видачі завдання \_\_\_\_\_

**КАЛЕНДАРНИЙ ПЛАН**

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	<i>Аналіз систем електронного навчання та факторів впливу на успішність студентів</i>	<i>28.10.2021-13.11.2021</i>	<i>виконано</i>
2.	<i>Визначення та формалізація факторів впливу і методів прогнозування успішності студентів</i>	<i>13.11.2021 – 20.11.2021</i>	<i>виконано</i>
3.	<i>Програмна реалізація моделей прогнозування успішності студентів</i>	<i>21.11.2021 – 28.11.2021</i>	<i>виконано</i>
4.	<i>Охорона праці та безпека в надзвичайних ситуаціях</i>	<i>28.11.2021 – 02.12.2021</i>	<i>виконано</i>
5.	<i>Оформлення пояснювальної записки</i>	<i>03.12.2021-06.12.2021</i>	<i>виконано</i>
6.	<i>Оформлення графічного матеріалу</i>	<i>07.12.2021-11.12.2021</i>	<i>виконано</i>
7.	<i>Попередній захист кваліфікаційної роботи магістра</i>	<i>15.12.2021</i>	<i>виконано</i>
8.	<i>Захист кваліфікаційної роботи магістра</i>		

Студент

\_\_\_\_\_  
(підпис)*Дармопук Д.В.*\_\_\_\_\_  
(прізвище та ініціали)

Керівник проекту (роботи)

\_\_\_\_\_  
(підпис)*Жаровський Р.О.*\_\_\_\_\_  
(прізвище та ініціали)

## АНОТАЦІЯ

Тема кваліфікаційної роботи: “ Методи і засоби інтелектуалізованого аналізу та прогнозування успішності студентів у системах електронного навчання ” // Кваліфікаційна робота // Дармопук Дмитро Володимирович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп’ютерно-інформаційних систем та програмної інженерії, група СІм-61 // Тернопіль, 2021 // с. – 97, рис. – 25, табл. – 12, аркушів А1 – 8, додат. – 1, бібліогр. – 26.

Ключові слова: метод, засіб, успішність, студент, прогнозування, система, електронне навчання.

Метою роботи є дослідження методів і засобів інтелектуалізованого аналізу та прогнозування успішності студентів у системах електронного навчання.

У кваліфікаційній роботі визначено та проаналізовано характеристики систем дистанційного навчання та онлайн навчання, що дало змогу встановити фактори і програмні модулі, які необхідні для забезпечення успішності студентів і навчального процесу в цілому, проаналізовано особливості застосування сучасних платформ для провадження навчального процесу

На основі аналізу факторів впливу на успішність студентів під час електронного навчання запропоновано їх формалізацію у вигляді елементів множини, що дало змогу побудувати модель, яка враховує ефективність програмно-апаратного забезпечення, якість навчального контенту та кваліфікацію інструкторів курсів. Для прогнозування успішності студентів обґрунтовано та формально представлено алгоритми, які використовують підходи дерев прийняття рішень та випадкових лісів, методу опорних векторів та глибокого навчання і виконано їх програмну реалізацію.

## ABSTRACT

The theme of the thesis: " Methods and means of intellectualized analysis and forecasting of students' success in e-learning systems " /Master thesis / Darmopuk Dmytro Volodymyrovych / Ternopil Ivan Pul'uj National Technical University, Faculty of Computer Information Systems and software engineering, group CIm -61 // Ternopil, 2021// p. - 97, fig. – 25, table. – 12, Sheets A1 – 8, Add – 1, Ref. – 26.

Keywords: method, tool, success, student, prediction, system, e-learning.

The aim of the work is to study the methods and means of intellectual analysis and prediction of student performance in e-learning systems.

The thesis identifies and analyzes the characteristics of distance learning and online learning, which allowed to establish the factors and software modules that are necessary to ensure student success and the learning process in general, analyzes the features of modern platforms for the learning process.

Based on the analysis of factors influencing student achievement in e-learning, their formalization in the form of set elements was proposed, which allowed to build a model that takes into account the effectiveness of software and hardware, quality of educational content and qualification of course instructors.

To predict the success of students, algorithms that use the approaches of decision-making trees and random forests, the method of reference vectors and deep learning are substantiated and formally presented, and their software implementation is performed.

## ЗМІСТ

ПЕРЕЛІК ОСНОВНИХ УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ І СКОРОЧЕНЬ .	8
ВСТУП .....	9
РОЗДІЛ 1 АНАЛІЗ СИСТЕМ ЕЛЕКТРОННОГО НАВЧАННЯ ТА ФАКТОРІВ ВПЛИВУ НА УСПІШНІСТЬ СТУДЕНТІВ .....	13
1.1. Аналіз особливостей організації освітнього середовища в Україні.....	13
1.2. Аналіз вимог до систем електронного навчання .....	17
1.3. Аналіз характеристик сучасних систем дистанційної освіти.....	21
1.4. Висновки до розділу .....	29
РОЗДІЛ 2 ВИЗНАЧЕННЯ ТА ФОРМАЛІЗАЦІЯ ФАКТОРІВ ВПЛИВУ І МЕТОДІВ ПРОГНОЗУВАННЯ УСПІШНОСТІ СТУДЕНТІВ.....	31
2.1. Фактори впливу на успішність онлайн навчання та їх формалізація.....	31
2.2. Методи інтелектуалізованого прогнозування успішності студентів.....	38
2.3. Формальне представлення алгоритмів прогнозування успішності студента з використанням дерев прийняття рішень та методу опорних векторів.....	39
2.4. Метод прогнозування успішності навчання з використанням глибоких нейронних мереж.....	48
2.5. Висновки до розділу .....	52
РОЗДІЛ 3 ПРОГРАМНА РЕАЛІЗАЦІЯ МОДЕЛЕЙ ПРОГНОЗУВАННЯ УСПІШНОСТІ СТУДЕНТІВ .....	54
3.1. Прогнозування успішності студентів у системах електронного навчання на основі алгоритмів лінійної залежності.....	54
3.1.1. Аналіз та візуалізація вхідних даних .....	54
3.1.2. Виявлення та аналіз кореляції факторів щодо успішності навчання .....	63
3.1.3. Реалізація алгоритмів прогнозування успішності студентів.....	68
3.2. Результати прогнозування успішності на основі глибоких нейронних мереж та архітектури GritNet .....	73
3.3. Висновки до розділу .....	78

ВИСНОВКИ.....	79
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	89
Додаток А Тези конференцій .....	92

ПЕРЕЛІК ОСНОВНИХ УМОВНИХ ПОЗНАЧЕНЬ,  
СИМВОЛІВ І СКОРОЧЕНЬ

БД	База Даних
ДН	Дистанційне навчання
СДН	Система дистанційного навчання
ЕП	Електронний підручник
ЖЦ	Життєвий Цикл
ІЕТІ	Інтерактивний електронна технічна інструкція
ІС	Інформаційні Системи
КНС	Комп'ютерні навчальні системи
НМК	Навчально методичний комплекс
ПС	Програмні Системи
ACID	Atomicity Consistency Isolation Durability
CASE	Computer Aided Software Engineering
ER	Entity Relations
UML	Unified Modeling Language
XML	Extended Markup Language



## ВСТУП

**Актуальність теми.** Потреба у навчанні та здобутті конкурентоспроможних знань і вмінь на сьогодні є проблемою глобального рівня. Для прикладу, у США, незважаючи на мільярди доларів, витрачених на покращення системи освіти, приблизно 35% студентів ніколи не закінчують середню школу. Рівень падіння серед деяких демографічних груп досягає 50-60%. На рівні коледжів, у США лише 30% студентів закінчують 2-річні коледжі за 3 роки і приблизно 50% закінчують 4-річні коледжі за 5 років. Таким чином, основним завданням у забезпеченні глобальної освіти є покращення успішності студентів. Під успіхом студентів, в глобальному сенсі мається на увазі підвищення показників навчання, успішного завершення навчальних предметів передбачених відповідною програмою, а також випуск і працевлаштування здобувачів освіти.

Сьогодні освіта – це вже не разова подія, а досвід, який здобувається на усе життя. Одна з причин важливості здобуття якісної освіти полягає в тому, що працездатний вік зараз настільки довгий, а технології настільки швидко змінюється, що людям потрібно продовжувати вчитися протягом усієї кар'єри.

Хоча класична модель освіти не масштабується, щоб задовольнити і врахувати швидко змінювані потреби, масовий ринок впроваджує інновації, щоб дозволити працівникам навчатися по-новому. Масові відкриті онлайн-курси (МООС), які пропонують такі компанії, як Udacity і Coursera, тепер зосереджені набагато більше на курсах, які роблять їхніх студентів більш конкурентоспроможними.

На онлайн-платформах студенти платять кошти за короткі програми, які надають мікроакредитацію з технологічних предметів, таких як самокеровані автомобілі та мобільні пристрої. Більше того, університети пропонують онлайн-курси, щоб полегшити професіоналам доступ до можливостей розвитку своїх навичок і передбачити результати навчання (наприклад, успішне завершення курсу, зняття курсу, підсумкові оцінки).

У світі MOOC значне збільшення кількості студентів робить недоцільним проведення індивідуальних оцінок навіть досвідченим інструкторам. У цьому випадку може допомогти автоматизована система, яка б могла точно прогнозувати, як студенти будуть працювати в режимі реального часу. Це був би цінний інструмент для прийняття розумних рішень про те, коли потрібно проводити живі освітні заходи під час курсу (і з ким), з метою підвищення залучення нових слухачів, забезпечення мотивації та розширення можливостей студентів для досягнення успіху. Проблема прогнозування успішності студентів частково вивчається у спільнотах аналітиків та аналізу освітніх даних у формі проблеми прогнозування провалу навчання (або завершення) студентом (яка є важливою проблемою підкласу проблеми прогнозування успішності учня).

Дослідженню методів і засобів прогнозування успішності навчання у широкому і вузькому розумінні присвячено праці багатьох науковців, як українських (І. Олексів, І. Ізонін, В. Харчук, Р. Ткаченко, А. Дорошенко та ін.), так і закордонних (В.-Н. Kim, Е. Vizitei, V. Ganapathi, P.Parkavi, V.Ramesh та ін.). Результати їхніх досліджень є важливими та показують високу достовірність і точність успішності завершення курсів студентами. Однак потребують подальшого дослідження технології прогнозування для систем електронного і дистанційного навчання, що обумовлено зростанням кількості додаткових даних, соціальних та економічних аспектів життя студентів, а також розвитком методів машинного навчання. Тому задача щодо дослідження методів і засобів інтелектуалізованого аналізу та прогнозування успішності студентів у системах електронного навчання є актуальною на сьогодні.

**Мета і задачі дослідження.** Метою роботи є дослідження методів і засобів інтелектуалізованого аналізу та прогнозування успішності студентів у системах електронного навчання.

Для досягнення вказаної мети в роботі поставлено наступні **задачі**:

- аналіз наукових публікацій та обґрунтування необхідності впровадження інтелектуалізованого прогнозування успішності студентів;
- дослідження технічних особливостей та принципів організації систем

електронного навчання;

- формалізація моделі успішності студентів;
- розробка методу інтелектуалізованого аналізу показників успішності студентів у системах електронного навчання;
- оцінювання ефективності моделі прогнозування успішності студентів;
- програмна реалізація та аналіз результатів прогнозування успішності студентів у системах електронного навчання.

**Об'єкт дослідження:** процес прогнозування успішності студентів.

**Предмет дослідження:** критерії і фактори впливу на успішність студентів, моделі, методи та алгоритми інтелектуалізованого прогнозування успішності студентів.

**Методи дослідження:** При розв'язанні задач дипломного проектування використано наступні методи: аналіз та обґрунтування – при визначенні критеріїв впливу на успішність студентів, а також при дослідженні методів і засобів її прогнозування; теорія множин, моделі та алгоритми машинного навчання – при побудові моделі успішності студентів та розробці алгоритмів прогнозування успішності; програмування – при проведенні експериментів щодо реалізації запропонованих алгоритмів; апробація – при оцінюванні ефективності моделі прогнозування успішності студентів.

**Наукова новизна отриманих результатів.** Наукова новизна результатів дослідження полягає в наступному:

- уперше на основі аналізу факторів впливу на успішність студентів під час електронного навчання запропоновано їх формалізацію у вигляді елементів множини, що дало змогу побудувати модель, яка враховує ефективність програмно-апаратного забезпечення, якість навчального контенту та кваліфікацію інструкторів курсів, які в подальшому будуть враховані при виборі інтелектуалізованих методів і засобів прогнозування успішності студентів.

- набув подальшого розвитку метод прогнозування успішності студентів на основі підходу глибоких нейронних мереж з архітектурою GritNet в основі якого лежать вбудовані шари двонаправленого LSTM, що дає змогу більш ефективно, у

порівнянні з базовою моделлю GritNet, забезпечити високу точність прогнозування фактів завершення або не завершення навчання студентів на курсі.

**Практичне значення одержаних результатів.** Практичне значення одержаних результатів полягає у програмній реалізації моделі на основі інтелектуалізованих алгоритмів прогнозування успішності навчання з використання мови програмування Python та відкритих бібліотек машинного навчання.

**Публікації.** Результати кваліфікаційної роботи апробовані на X міжнародній науково - технічній конференції молодих учених і студентів «Актуальні задачі сучасних технологій» (24-25 листопада 2021 р.) Тернопільського національного технічного університету імені Івана Пулюя та на IX науково-технічній конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (8-9 грудня 2021 року) як тези конференцій.

1. Жаровський Р.О., Дармопук Д.В. Характеристики систем електронного навчання. Матеріали X міжнародної науково - технічної конференції молодих учених і студентів «Актуальні задачі сучасних технологій» (24-25 листопада 2021 р.) Тернопільського національного технічного університету імені Івана Пулюя. Тернопіль: ТНТУ. 2021. С. 91.

2. Жаровський Р.О., Дармопук Д.В. Аналіз успішності студентів на основі технології GritNet. Матеріали IX науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (8-9 грудня 2021 року). Тернопіль: ТНТУ. 2021. С. 108.

**Структура роботи.** Кваліфікаційна робота містить розрахунково-пояснювальну записку та графічний матеріал. До складу записки входить вступу, 4 розділи, загальні висновки, список використаних джерел і додатки. Обсяг роботи: розрахунково-пояснювальна записка – 97 арк. формату А4, графічна частина – 9 аркушів формату А1.

## РОЗДІЛ 1

# АНАЛІЗ СИСТЕМ ЕЛЕКТРОННОГО НАВЧАННЯ ТА ФАКТОРІВ ВПЛИВУ НА УСПІШНІСТЬ СТУДЕНТІВ

### 1.1. Аналіз особливостей організації освітнього середовища в Україні

В Україні здобуття вищої освіти забезпечують заклади 3-4 рівня акредитації, тобто коледжі, інститути та університети. При цьому студенти можуть навчатись за різними формами, зокрема за:

- очною – передбачає постійне відвідування закладів освіти та використання матеріального і технічного забезпечення університету чи інституту;
- заочною та заочно-дистанційною – лівова частка програми спеціальності здобувається без відвідування освітніх закладів, а за допомогою використання засобів дистанційного навчання і засобів моделювання.
- дуальна – можливість навчання за місцем роботи з використанням технічних засобів підприємства і частково на базі закладу вищої освіти.

Програми навчання за технічними спеціальностями передбачають проведення лекційних занять, виконання лабораторних і практичних занять, а також перевірку знань за допомогою тестування з використанням засобів автоматизації, наприклад, платформ дистанційного навчання (засоби дистанційного навчання – ЗДН). Такі платформи дають можливість забезпечити уніфікацію при провадженні електронного навчання конкретним закладом освіти. Застосування ЗДН дає змогу розширити та візуалізувати навчальні матеріали, що сприяє кращому засвоєнню та сприйняттю матеріалів, розвиває нові навички і розширює знання здобувачів освіти.

Фахові компетенції за тією чи іншою спеціальністю формують навчальні плани, які зазвичай містять три основні розділи:

- фундаментальний – містить розділи і дисципліни, що формують базис загальної підготовки здобувачів освіти у певному професійному полі;

- професійно-орієнтований – формує знання і навички, характерні для фахівця за конкретною спеціальністю;

- вибіркового циклу – зазвичай формується на основі вподобань студентів та особливостей ринку праці у конкретному регіоні.

У результаті аналізу циклів підготовки здобувачів вищої освіти можна виявити потенційне ядро, що включає ряд дисциплін з різних блоків, які безпосередньо впливають на подальшу успішність студентів, як фахівців в конкретній області. Ефективність фахівців визначається успішністю засвоєння програми спеціальності. При цьому на успішність впливають декілька комплексних факторів, найбільш важливими з яких є:

- доступність та якість навчальних матеріалів;
- цілі та мотивація здобувачів вищої освіти;
- доступність та функціональність систем електронного навчання;
- зрілість та фаховість викладачів чи інструкторів курсів.

Особливостями здобуття вищої освіти при денній (очній) формі навчання є:

- провадження навчального процесу шляхом безпосередньої взаємодії та комунікації інструкторів (викладачів) і студентів (здобувачів освіти);
- застосування групових форм при проведенні занять: лекційні та семінарські заняття, лабораторні і практичні роботи тощо;
- визначений чіткий список предметів, які підлягають вивченню;
- період часу на вивчення кожної дисципліни є детермінованим і наперед визначеним у навчальному плані.

При заочному навчанні, здобуття освіти передбачає віддалену комунікацію між учасниками процесу. При цьому дана форма вищої освіти орієнтована на тих людей, які вже здобули певну кваліфікацію, не можуть щоденно очно відвідувати заняття у зв'язку з неможливістю залишення місця праці або з будь-яких інших причин. Однак, дана форма здобуття освіти все ж забезпечує:

- перспективу здобуття освітнього рівня, що не залежить від локації студентів;
- можливість паралельного навчання за декількома освітніми програмами;

- шанс одержати ступінь і кваліфікацію вищої освіти для різних категорій населення, наприклад, жінками, які перебувають у декретній відпустці чи виховують дітей, інвалідами чи іншими непрацездатними особами;
- можливість одночасного охоплення значної аудиторії здобувачів освіти;
- можливість навчатися за власноруч сформованим графіком і темпом у період між перевіркою знань (сесіями);
- можливість паралельно здобувати знання і навички з одночасною їх апробацією на практиці.

Дистанційне та електронне навчання на сьогодні набувають широкої популярності та стають необхідними в час пандемічних процесів. Це зумовлено потужним розвитком і впровадженням інформаційних технологій та необхідністю забезпечення дистанції між учасниками навчального процесу .

Електронна освіта повністю відображає і забезпечує трансляцію принципів, характерних для класичних, загальноприйнятих форм навчання. Застосування засобів автоматизації та інтернет-технологій дає змогу додати нові позитивні особливості у процес навчання, наприклад доступність навчальних матеріалів 24/7, підтримка відеоконтенту, середовищ автоматизованої перевірки і тестування знань.

При використанні засобів електронного навчання на інструкторів курсів покладаються додаткові завдання щодо формування не тільки фахових вмій і знань, а й аспектів виховної роботи, що одержав інтерпретацію: «принцип креативності характеру пізнавальної діяльності» [5].

Інтерактивність засобів електронного навчання дає змогу забезпечити проведення різних конкурсів серед здобувачів освіти, що дозволяє підвищити їхню мотивацію і творчість при прийнятті рішень.

Ще одним важливим принципом, дотримання якого є дуже важливим при організації як електронного навчання, так і традиційного – впровадження елементів наукових досліджень. Даний принцип базується на використанні наукових здобутків у галузі, з якою пов'язаний навчальний предмет. Це забезпечує відображення основного контенту навчальних матеріалів та наукових фактів,

понять і закономірностей, а також особливостей сучасного стану досліджень у визначеній галузі.

У термінології ДН цей принцип набув наступної інтерпретації: «принцип відповідності фундаментальності навчання пізнавальним потребам особи, яка навчається» [7]. Він передбачає, що фундаментальність навчання повинна підтверджуватись шляхом встановлення основних сутностей, взаємозалежностей і закономірностей між ними. Як наслідок, до здобувачів освіти висуваються наступні вимоги:

- необхідність високої мотивації;
- ціль щодо саморозвитку і самовдосконалення;
- гнучкість сприйняття фактів, коригування і самокритичність;
- здатність досягати поставленої мети.

Розвитком наведеного вище принципу є інший, суть якого полягає у свободі вибору та інтерпретації інформації, яка одержана у результаті провадження певного виду діяльності. Дослівно він звучить наступним чином: «не існує єдиного ідеального інформаційного джерела, тому спрямованість навчання стосується безпосередньо не інформації, а шляхів її перетворення та опрацювання, за допомогою участі в дискусіях, телеконференціях, роботи з пошуковими машинами тощо» [8].

При використанні класичного навчання, ще однією важливою дидактичною особливістю є систематичність і послідовність. Вона передбачає необхідність логічної впорядкованості знань і матеріалів, чіткої їхньої класифікації, залежності з та від інших матеріалів. Це дає змогу покращити засвоєння матеріалів, сприяє підвищенню успішності та ефективності навчального процесу.

При електронному форматі навчання забезпечується самостійний вибір цілей здобуття освіти, форми і швидкості навчання. Даний принцип часто називають «індивідуальною навчальною траєкторією студента» [8].

Принципом наочності передбачено, що матеріал краще сприймається, а відповідно і відбувається краще його засвоєння, у випадку забезпечення високої концентрації уваги та відповідної реакції і дії на органи сприйняття людини.



При організації навчального процесу за допомогою електронних засобів комунікації контакт безпосередньої взаємодії слухачів та викладача відсутній, тому визначено принцип «віртуалізації навчання». Він забезпечується шляхом застосування засобів мультимедіа, матеріалів у вигляді відеоконтенту, відеоконференції і т.п..

Внаслідок активного застосування технологій інформаційного забезпечення, крім проаналізованих вище особливостей, виникла ціла низка груп принципів, що підтримуються засобами електронного навчання. Одним з них є «принцип ідентифікації», суть якого полягає у здатності системи дистанційного навчання адекватно реагувати на загрози фальсифікації у навчальному процесі. Сучасні технологічні заходи і їхня реалізація дають змогу контролювати міру самостійності виконання індивідуальних завдань, а також перевірки справжності користувача за допомогою засобів відеоконференцій та відеозв'язку.

## 1.2. Аналіз вимог до систем електронного навчання

Основними вимогами до систем електронного навчання є забезпечення можливостей:

- контролю та управління процесом навчання інструктором курсу;
- внесення змін і виправлень у навчальні матеріали;
- прозорості взаємодії між викладачем і слухачем курсу;
- комунікації між студентами записаними на курс.

Цей принцип забезпечує інтерактивність навчального процесу .

Для того, щоб забезпечити ефективність дистанційного навчання, необхідно дотримуватися ще одного важливого принципу – принцип початкових знань і вмінь. Суть даного принципу полягає у тому, що слухачі курсу електронного навчання повинні володіти базовими навиками використання комп'ютерної техніки, повинен бути забезпечений доступ до мережі Інтернет, а також інші цифрові пристрої, що вимагаються програмою курсу для ефективного його засвоєння.

Однак використання педагогічних принципів при провадженні навчального процесу повинні враховувати доцільність використання електронних засобів та інформаційних технологій, що формує ще одну вимогу, яка впливає на якість та успішність навчання здобувачів освіти.

При цьому, будь-яка дія щодо впровадження або організації дистанційної освіти потребує залучення фахівців з педагогіки щодо оцінювання відповідності визначеним критеріям та обґрунтування доцільності і впливу на складові навчання, наприклад, зміст, ціль, інструменти і т.п.

Варто відмітити, що множина педагогічних принципів і їхня структура не є строго детермінованими, а протягом певного періоду можуть зазнавати змін та адаптації. Це пов'язано з розвитком технологічної та апаратної складової засобів електронного навчання.

Проаналізовані вище принципи формують базис при проектуванні і створенні систем електронного навчання. Відповідно структура таких систем є складною конструкцією з великою кількістю взаємопов'язаних між собою програмних модулів. Для більшості сучасних систем підтримки процесів онлайн навчання характерні подібні функціональні можливості, найбільш важливими з яких є ті, аналіз яких наведено нижче.

Першою і найбільш необхідною компонентою системи дистанційного навчання є складова, що забезпечує доступ до навчальних матеріалів. З практичної точки зору, такий програмний модуль надає можливість реєстрації, ідентифікації та аутентифікації користувача, керування ролями та групами, а також забезпечує контрольований доступ до наявного контенту курсу.

Поряд із функціональністю щодо ідентифікації користувача необхідним є наявність зручних у використанні інструментів адміністрування. Вони підтримують здатність до управління користувачами під час навчання, створення та керування онлайн курсів і їх наповненості, а також проведення відповідних заходів щодо перевірки знань студентів.

Для забезпечення інтерактивності навчального процесу доцільним є використання різноманітних інструментів комунікації між учасниками процесу.

Найбільш широко використовуваними є текстові чати та інформаційні форуми, блоги, засоби електронної пошти студентів курсу та інструкторів.

Платформи дистанційного та електронного навчання повинні забезпечувати можливість підтримки моделей керування знаннями і вміннями, які формують відповідні компетенції у слухачів. А це означає, що повинна існувати підсистема створення програми курсу. При декомпозиції навчальних програм формуються моделі, що визначають певну конкретну роль користувача та множину його компетенцій. Надалі для таких компетенцій створюється сукупність електронних дисциплін для вивчення та заходів перевірки, які повинен виконати і підтвердити студент. За допомогою таких моделей, засіб електронного навчання формує відповідні програми освітнього процесу для кожного зареєстрованого слухача на платформі.

Для самоаналізу і визначення ефективності та успішності проходження курсів в системах дистанційного навчання повинен бути наявним модуль створення звітів на вимогу користувача.

Важливим аспектом при використанні систем дистанційного навчання є наявність програмних інтерфейсів для забезпечення можливості інтеграції з іншими зовнішніми системами. Це означає, що платформа електронного навчання не повинна перебувати у стані автономності та ізоляції відносно загального інформаційного середовища.

Архітектура платформ електронного навчання повинна підтримувати програмну та апаратну розподіленість, що пов'язано із забезпеченням доступу до навчальних матеріалів навіть за умови збою окремих компонентів системи. Це передбачає застосування принципів реплікації та резервного копіювання за різними вузлами інформаційної інфраструктури на якій функціонує система.

До складу базової архітектури традиційних систем електронного навчання та систем дистанційного навчання належать такі складові як:

- компонент реєстрації користувача та управління правами доступу до системи;
- компонент для провадження навчального процесу;

- модуль створення та оцінювання тестових завдань;
- компонент, що надає можливості формування та використання різних видів навчальних матеріалів;
- модуль для експорту та імпорту навчальних матеріалів з підтримкою різних форматів;
- компонент, що забезпечує інтерактивну складову взаємодії користувачів курсів: «лектор-студенти», «студенти-студенти», «студенти-лектор»;
- компонент фіксації активності користувачів.

Впровадження і застосування систем дистанційного навчання володіє рядом переваг і недоліків. Основними перевагами систем дистанційного навчання є:

- гнучкість – передбачає можливість одночасного або паралельного навчання у різних локаціях;
- свобода вибору – можливість здобувача освіти самостійно обирати швидкість навчання, формувати мітки матеріалу, наприклад, матеріал, який варто повторити, засвоєний матеріал, не цікавий і т.д.;
- формування власного розкладу навчання – передбачає навчання у зручний для студента час та у комфортних умовах;
- навчання під псевдонімом – забезпечення режиму інкогніто, що зумовлені певними обставинами;
- доступ до знань та освіти людьми з обмеженими можливостями;
- зростання особистих якостей студентів – полягає у виникненні почуття відповідальності, дисциплінованості та самоорганізованості;
- зростання навчальної аудиторії – системи дистанційного навчання дозволяють одночасно проводити навчання для більшого кола студентів, з різним рівнем знань та поглядів на життя;
- простота створення віртуального середовища – можливість викладачів, студентів тощо завдяки використанню сучасних інтернет технологій, за допомогою яких стає можливим обговорення між викладачами певних проблем, вирішення спільних завдань, обмін досвідом чи інформацією тощо.

До недоліків електронного навчання входять:

- відсутність прямої безпосередньої взаємодії між учасниками навчального процесу;
- методичні матеріали позбавлені емоційного забарвлення;
- складність створення творчої атмосфери;
- необхідність наявності відповідного апаратного і програмного забезпечення та ін.
- успішність впровадження електронного навчання напряму пов'язана з психологічними аспектами здобувачів і вимагає обов'язкової самодисципліни, і мотивації студентів;
- неможливість забезпечити постійний моніторинг здобувачів освіти та недостатність практичної підготовки.
- створення матеріалів для дистанційного курсу вимагає значних витрат трудових ресурсів, зокрема створення якісного мультимедійного контенту на час до однієї години займає приблизно тисячу годин фахівця з конкретної дисципліни.

### 1.3. Аналіз характеристик сучасних систем дистанційної освіти

Сучасні досягнення у проектуванні програмно-апаратних комплексів, застосування хмарних технологій стимулювали розвиток широкого спектру систем і платформ підтримки освітнього середовища. Такі системи розповсюджуються під різними ліцензіями, починаючи від умовно безкоштовних платформ з відкритим кодом до систем з різними тарифними платами, а також орієнтованих на широке коло користувачів або вузькоспеціалізованих.

Для прикладу, у ТНТУ використовується система дистанційного навчання Atutor, яка представляє собою модульний засіб автоматизації електронного навчання з вільною ліцензією («GNU-General Public License»).

Вимогами до програмного середовища для ефективного функціонування цієї системи є:

- сервер з інстальованим web-сервером Apache версії не нижче, ніж 1.3.x;
- мова програмування PHP згідно визначеної специфікації;

– система керування базами даних MySQL згідно технічної специфікації.

СДН Atutor створена із забезпеченням вимог доступності та гнучкості, що дозволяє виконувати адаптацію функціональності та інтерфейсу користувача у відповідності до потреб надавача освітніх послуг або користувача. Дана система є кросплатформною, тобто ефективно може працювати під будь-якою операційною системою.

Ще однією платформою, що поширюється під вільною ліцензією з відкритим кодом є Claroline LMS. Ця система може функціонувати під основними широко використовуваними ОС, зокрема, Windows, Linux та Mac. Особливістю даної системи дистанційного навчання є наявність зручного та інтуїтивно зрозумілого інтерфейсу для виконання задач адміністрування навчального процесу, а також реалізація ідеї розподілених просторів даних та матеріалів, які відображають різні аспекти електронного курсу та ведення педагогічної діяльності інструкторами. У табл. 1.1 наведено характеристику модуля для розроблення та відображення матеріалу навчального курсу у системах ATutor та Claroline.

*Таблиця 1.1*

**Модуль розроблення та представлення навчальної інформації в системі**

Система	ATutor	Claroline
Реалізація модулів	Створення курсів (вказується опис, доступ, дата публікації) Модуль відновлення курсів Редагування вмісту(ключові слова, схожі теми, попередній перегляд і перевірка відтворення в браузері) Словник Посилання на інші джерела Список літератури	Створення курсів (вказується опис, доступ) Публікація документів і посилань на сайти інструктора Завантаження файлів

Основними функціональними можливостями, що передбачені у кожному просторі, є наявність засобів для формування, моніторингу та керування навчальними матеріалами. Також особлива увага приділяється інструментам інтерактивної комунікації між користувачами з однаковими і різними ролями тощо [13].

У табл. 1.2 наведено можливості модулів інтерактивної взаємодії користувачів систем ATutor та Claroline.

Таблиця 1.2

### Інтерактивна взаємодія у системах ATutor та Claroline

Система	ATutor	Claroline
Реалізація модулів	+	+
Взаємодія студент-студент	Персональні повідомлення Модуль обміну файлами Форум	Чат Форум Оголошення Створення подій у календарі Wiki
Взаємодія викладач-студенти	Оголошення FAQ Чат Розміщення новин на баннері Персональні повідомлення Стрічка новин RSS	Оголошення Коментарі Форум Wiki Чат Створення подій у календарі
Взаємодія студенти-викладач	Форум Чат Персональні повідомлення Опитування	Форум Чат Оголошення Створення подій у календарі

На основі технології ASP («Active Server Pages») від Microsoft розроблено систему дистанційного навчання Live@EDU. Оскільки, ця система розроблена із застосуванням технологій MS, то відповідно вона функціонує на класі операційних систем Windows, використовує MS SQL Server та IIS.

Вимогами щодо забезпечення коректності функціонування цієї системи на стороні клієнта є:

- наявність доступу до мережі Інтернет.
- браузер з підтримкою протоколів HTTP/HTTPS;
- клієнтське програмне забезпечення від Live@EDU, що дає змогу переглядати і створювати навчальні матеріали.

До відносно нових систем підтримки онлайн навчання належить система eFront, що дозволяє реалізувати функціональність щодо управління освітнім процесом і забезпечує можливість формування та керування відповідними матеріалами. Система eFront ефективно застосовується при провадженні навчального процесу у закладах освіти різного рівня та у деяких підприємствах для атестації, оцінювання та підвищення кваліфікації своїх працівників. У системі визначено ролі трьох типів користувачів: «Адміністратор», «Студент», «Інструктор» [7]. У табл. 1.3 наведено характеристики модулів інтерактивної взаємодії у системах Live@EDU та eFront.

Таблиця 1.3

### Інтерактивна взаємодія у системах Live@EDU та eFront

Система	Live@EDU	eFront
Реалізація модулів	+	+
Взаємодія студент-студент	Форум Чат WWW	Форум Чат Персональні повідомлення Можливість установки модулів блогів та wiki



Система	Live@EDU	eFront
Взаємодія викладач-студенти	Оголошення Календар Chat Форум FAQ	Форум Чат Персональні повідомлення Блог Дошка оголошень FAQ Цитата дня Спільні коментарі
Взаємодія студенти-викладач	Чат Форум	Форум Чат Спільні коментарі Персональні повідомлення

Moodle представляє собою СДН у вигляді сукупності програмних модулів, що поширюється під відкритою ліцензією. Пакети цієї системи дозволяють формувати середовище для електронного навчання та публікації web-сайтів.

Характерною особливістю Moodle є орієнтація на керування навчальним процесом, що передбачає віртуальну комунікацію інструкторів курсів зі студентами і дозволяє підтримувати процеси традиційного навчання.

Вимогами до програмно-апаратної інфраструктури для використання Moodle є наявність сервера з підтримкою мови програмування PHP та однієї з реляційних СКБД: MS SQL Server, MySQL, PostgreSQL.

Moodle є кросплатформною системою дистанційного навчання, що підтримує 5 основних принципів, орієнтованих на так званий «соціальний конструктивізм»:

- середовище онлайн навчання передбачає одночасність застосування ролей як студента, так і викладача;
- орієнтація на успішність в процесі навчального процесу, що передбачає

створення чогось нового або надання конструктивних порад;

- моніторинг за поведінкою учасників онлайн навчання дає змогу підвищити його ефективність;

- врахування індивідуальних особливостей слухачів курсів для формування акцентованих рекомендацій для забезпечення успішності;

- система управління ДН повинна бути гнучкою, надавати простий та зручний у використанні інструментарій для досягнення потреб і цілей користувачів.

Яскравим представником СДН з платною підпискою є SharePoint LMS, що створена за допомогою мультифункціональної платформи «MS Office SharePoint Server» і дозволяє об'єднати в єдине інформаційне середовище усіх учасників навчального процесу та забезпечує їх засобами для спільної роботи. Дана платформа ефективно використовується не тільки закладами освіти, а й іншими організаціями. У табл. 1.4 наведено характеристики модулів інтерактивної взаємодії у системах Moodle та SharePoint.

Таблиця 1.4

#### Засоби інтерактивної взаємодії у системах Moodle та SharePoint

Система	Moodle	SharePoint
Реалізація модулів	+	+
Взаємодія студент-студент	Форум Чат Обмін повідомленнями	Персональні повідомлення Форум Чат
Взаємодія викладач-студенти	Форум Чат Обмін повідомленнями	Конференція Персональні повідомлення Рядок новин Форум
Взаємодія студенти-викладач	Чат Обмін повідомленнями	Форум Персональні повідомлення Конференція

У табл. 1.5 наведено характеристики модулів для самостійної роботи для всіх розглянутих вище платформ.

Таблиця 1.5

## Самостійна робота

Система	ATutor	Claroline
Реалізація модулів	Завдання(інструктор задає назву, суть і виконавця) Пошук в інтернеті на початковій сторінці або вкладці меню	Виконання завдань відповідно до вибраного сценарію (назва, опис, тип, дата, видимість, дозвіл на завантаження)
Система	Moodle	SharePoint
Реалізація модулів	Модуль “Завдання”. Вчитель створює опис завдання, установку на його виконання та вказує місце, куди студент зобов’язаний завантажити результати. Студент може завантажувати результати у вигляді рефератів, відеоматеріалів, презентацій, таблиць тощо. Модуль “Робочий зошит” відрізняється від модуля “Завдання” тим, що завдання складаються із Відповідей у вигляді тексту, які може редагувати студент	Модуль “Мої файли” призначений для завантаження та збереження файлів користувачів у межах курсу Модуль “Завдання” призначений для створення різних домашніх (додаткових) завдань та перегляду результатів їх виконання Модуль “Плагіат” забезпечує Викладачу можливість контролювати ступінь самостійності роботи Студента, запобігає списуванню матеріалу. Всі документи з “Мої файли” проходять перевірку.

Продовження табл. 1.5

Система	Live@EDU	eFront
Реалізація модулів	Модуль “Робоча область” забезпечує спільний простір на сервері, що робить можливим обмін файлами між студентами.	Вкладка “Проекти” модулю “Звіти” забезпечує інформацію про виконання студентами проектів.

У табл. 1.6 наведено характеристики модулів для забезпечення контролю лекційного матеріалу для всіх розглянутих вище платформ.

Таблиця 1.6

### Контроль лекційного матеріалу

Система	ATutor	Claroline
Реалізація модулів	Опитування учасників курсу (при цьому оцінки не виставляються) Статистика	Онлайн вправи зі списком питань Статистика Вибір сценарію навчання
Система	Live@EDU	eFront
Реалізація модулів	Завдання Папки завдань	Звіти по користувачу ( вкладки “Уроки”, “Курси”, “Докладніше”) Звіти по уроках вкладка “Запитання” дає інформацію про відповіді учнів на тести цього уроку “Активність” є звітом про активність студентів за вказаний період часу

Система	Moodle	SharePoint
Реалізація модулів	Журнал реєстрації активності користувачів (студентів) в блоці “Управління”. Можливими параметрами фільтрації журналу є день, назва курсу, група, учасник, виконане завдання.	Модуль “Навчальна програма” призначений для створення впорядкованої структури представлення навчальних матеріалів, а також створення системи тестування та контролю успішності студентів курсу  Модуль “Відвідування” призначений для реєстрації відвідувань студентів Курсу “Щоденник”. Модуль зберігає всю інформацію про успішність

Таким чином, проведено аналіз основних характеристик і функціональності найбільш поширених та використовуваних систем дистанційного навчання, що дало змогу виявити спільні риси, які дозволяють в подальшому врахувати їх при побудові моделі успішності навчання студентів.

#### 1.4. Висновки до розділу

У даному розділі одержано основні наукові і практичні результати:

1. Проведено аналіз особливостей організації навчального процесу в Україні та закордоном, на основі результатів якого встановлено, що набуває широкої популярності, поряд з класичними, такий вид здобуття освіти як онлайн навчання, елементи якого важливо імплементувати у традиційні форми навчального процесу.

2. Визначено та проаналізовано характеристики систем дистанційного навчання та онлайн навчання, що дало змогу встановити визначити фактори і

програмні модулі, які необхідні для забезпечення успішності студентів і навчального процесу в цілому.

3. Проаналізовано особливості застосування сучасних платформ для провадження навчального процесу, що дозволило визначити їхні переваги і недоліки, а також встановити необхідність наявності компонентів і даних для прогнозування успішності студентів із застосування інтелектуалізованих методів і засобів формування прогнозів.

## РОЗДІЛ 2

### ВИЗНАЧЕННЯ ТА ФОРМАЛІЗАЦІЯ ФАКТОРІВ ВПЛИВУ І МЕТОДІВ ПРОГНОЗУВАННЯ УСПІШНОСТІ СТУДЕНТІВ

#### 2.1. Фактори впливу на успішність онлайн навчання та їх формалізація

Класичні підходи до провадження навчального процесу та сформовані на основі них системи освіти базуються на комунікації двох груп учасників: студентів та викладачів. Вони, зазвичай, взаємодіють офлайн, що передбачає надання освітніх послуг на базі матеріально-технічної бази закладів освіти, а застосування інформаційних технологій використовується у недостатній мірі, що негативно впливає на якість одержаних знань та знижує конкурентоспроможність випускників на ринку праці. Це призводить до того, що особи після завершення навчання потребують постійної самоосвіти, особливо в галузі інформаційних технологій, витрачають додаткові кошти на самовдосконалення та підвищення кваліфікації. Можливим шляхом подолання цих проблем є застосування СДН або інших платформ електронного навчання разом з класичними формами здобуття освіти, які дають змогу значно зменшити або й взагалі уникнути появи таких ситуацій.

Дивлячись на стрімкий розвиток у галузях комп'ютерної та програмної інженерії, опрацювання великих даних, методів та систем штучного інтелекту, а також тенденції наростаючої глобалізації, доступності геолокації і мобільності серед здобувачів освіти та викладачів, системи дистанційного навчання відіграють велику роль і вже є невід'ємними компонентами процесу підготовки висококваліфікованих кадрів у будь-якій сфері.

Варто відмітити, що на відміну від послуг, які надають ЗВО при класичній організації процесу навчання, дуже популярними стають вузькоспеціалізовані курси. На освітньому ринку можна знайти багато курсів, які розповсюджуються як вільні і безоплатні, так і курси з платною підпискою.

Незалежно від того, який спеціалізований онлайн курс пройдений, спостерігається тенденція щодо значного зростання рівня знань і кваліфікації здобувача, на відміну від вивченого предмету чи комплексних курсів у традиційному освітньому закладі. Цей недолік традиційної освіти пов'язаний з великою інертністю та масштабністю університетів, які дуже повільно реагують на зміну трендів ринку праці. Окрім цього, у закладах освіти спостерігається не завжди достатня мотивованість викладацького складу і недостатність рівня кваліфікації та відповідних компетентностей, складність впровадження та не завжди задовільна якість СДН.

У даній кваліфікаційній роботі доцільно виявити та формалізувати фактори, які впливають на успішність навчання студентів при використанні платформ електронного навчання (рис. 2.1).

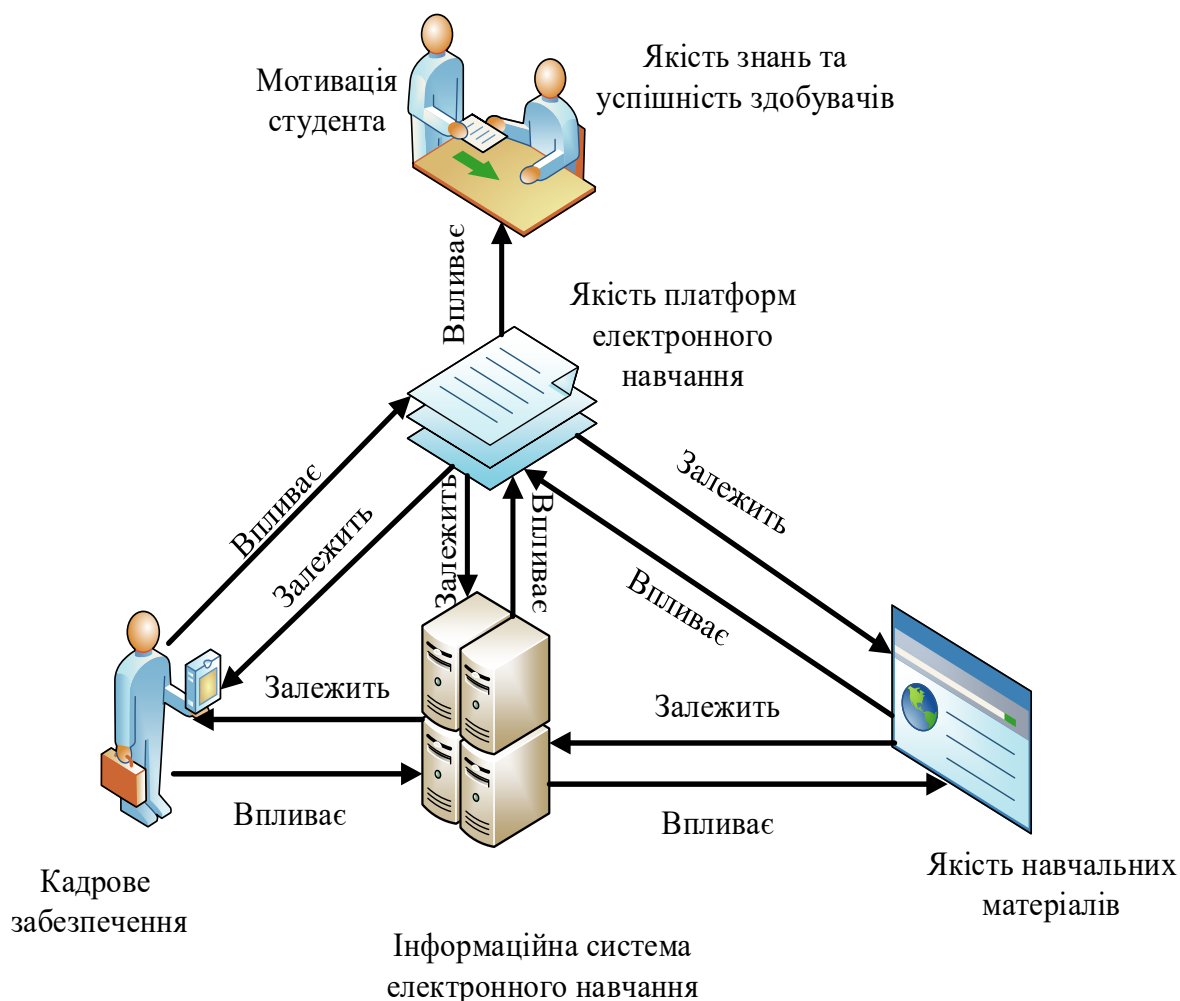


Рис. 2.1. Фактори, що визначають успішність студентів



Успішність студентів є частиною комплексної характеристики «якість знань і вмінь», що безпосередньо залежить від програмно-технічного забезпечення онлайн навчання, навчальних матеріалів і його форм, наявності інтерактивних інструментів та засобів контролю успішності. На рис. 2.1 показано залежності між основними факторами на успішність та якість знань студентів при дистанційному навчанні.

Впровадженню систем електронного навчання зазвичай передують аналіз наявних на ринку існуючих рішень та обрання оптимальної, яка найбільше відповідає вимогам надавача освітніх послуг. Розв'язання цієї проблеми є доволі складною задачею, оскільки формальні процедури та інструменти щодо оцінювання функціональності, надійності, продуктивності та зручності використання практично відсутні і не дають достовірної інформації про переваги однієї системи над іншою.

Для того, щоб вибрати оптимальну платформу провадження освітнього процесу на рівні реалізації програмної та апаратної складових доцільно скористатись серією стандартів міжнародних стандартів ISO 250xxx. Вони у більш-менш адекватній формі надають інформацію про міру реалізації властивостей і враховують потреби тих, хто надає освітні послуги. Математично моделі ISO 250xxx можна представити у вигляді сукупності елементів відповідних множин або у термінах теорії категорій.

Оскільки на успішність студентів впливає якість навчального матеріалу, то для формального їх опису доцільно застосовувати статистичні підходи при аналізі використання матеріалу, перевірки знань та відгуків користувачів, а також методів експертних оцінок і штучного інтелекту.

Окрім цього, для визначення професійності кадрового складу, який безпосередньо бере участь у навчальному процесі, при формалізації варто використати підхід теорії множин з врахуванням статистичних показників відносно результатів методичної і наукової роботи викладача.

Слід розуміти, що описані вище фактори, які впливають на успішність навчання є комплексними характеристиками, які потребують деталізації шляхом визначення елементарних атрибутів і засобів їх кількісного оцінювання.

Базуючись на фідбеках учасників освітнього процесу найбільш широко використовуваними та популярними платформами електронного і дистанційного навчання у ЗВО є наступні: MS Learning Gateway, Moodle, ATutor, Прометей, , IBM Lotus LMS, WebCT, SharePoint LMS .

Для опису факторів, які мають значний вплив на успішність студентів щодо завершення навчання та на якість здобутих знань, одержаних при використанні засобів онлайн навчання потрібно врахувати особливості характеристик наведених на рис. 2.1.

Зважаючи на той факт, що критерії впливу на успішність завершення навчання є комплексними, то доцільно використати ієрархічну модель, яка на верхньому рівні описує комплексну сутність, на середньому – атрибути, які формують характеристику вищого рівня, а на нижньому – метрики, які дозволяють кількісно виражати значення атрибутів.

Для загального випадку, успішність студентів можна записати у вигляді функції, що залежить від кортежа факторів впливу, які також є функціями

$$Success = \langle Q_{PI}, Q_M, Q_T \rangle, \quad (2.1)$$

де *Success* – успішність студентів;

$Q_{PI}$  – якість та ефективність платформи онлайн навчання;

$Q_M$  – якість і зручність використання опублікованих у системі матеріалів;

$Q_T$  – професійність інструктора курсу.

Для формального представлення якості та ефективності платформи дистанційного навчання можна скористатись залежністю, що описує вплив технічних факторів (ПЗ і АЗ) на успішність та якість сприйняття навчальних матеріалів.

Модель  $Q_{Pl}$  описується таким кортежем моделей:

$$Q_{Pl} = \langle Q_S(Q_{use}, Q_{ext}, Q_{in}), Q_A \rangle, \quad (2.2)$$

де  $Q_S$  – якість програмної реалізації платформи, що формується на основі моделей  $Q_{use}, Q_{ext}, Q_{in}$ , які відповідно описують моделі стандарту ISO 25010 і формують враження на рівні якості використання, якості програмних модулів і програмного коду;

$Q_A$  – якість технічного забезпечення та використання апаратних ресурсів.

Формальне представлення опублікованих у системі онлайн навчання матеріалів можна відобразити набором відповідних атрибутів чи властивостей. Однак, при їх визначенні, потрібно також обґрунтувати метрику для оцінювання міри реалізації того чи іншого атрибуту. У зв'язку з цим формальне представлення якості та ефективності навчальних матеріалів пропонується представити так, як у формулі (2.3)

$$Q_M = \{H_i^m, A_{ij}^m, M_{ij}^m\} \quad (2.3)$$

$H_i^m$  – деяка комплексна властивість, що описує ефективність і якість наявного матеріалу,  $i = 1..n$  - кількість комплексних властивостей;

$A_{ij}^m$  – простий атрибут, який відображає якість опублікованого навчального матеріалу,  $j = 1..k$  - кількість простих властивостей;

$M_{ij}^m$  – кількісна міра для представлення значення властивості якості навчального матеріалу.

Структурно модель представлення навчальних матеріалів щодо успішності навчання пропонується зображати у вигляді структури з трьома рівнями ієрархії таким чином, як наведено на рис. 2.2.

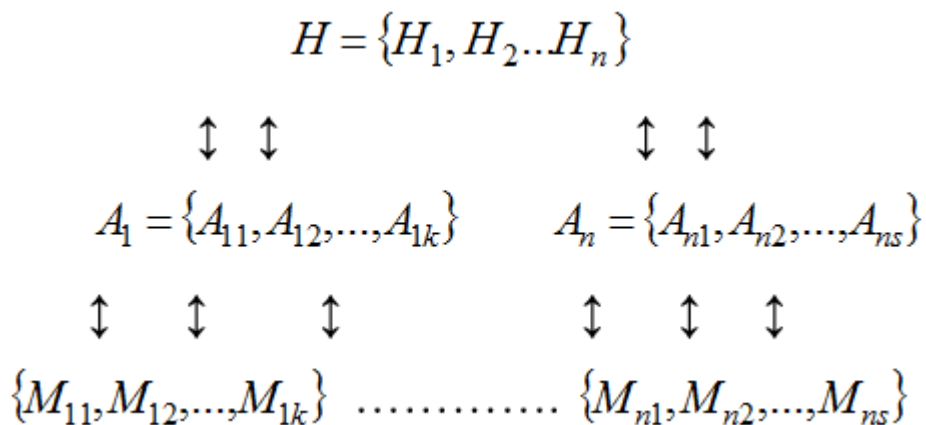


Рис 2.2. Структура моделі з трьома рівнями ієрархії

Комплексні властивості  $H_1, \dots, H_n$  навчального матеріалу формують верхній рівень ієрархії. На наступному рівні представлені прості вимірні властивості  $\{A_{im}\}$ , які стосуються специфіку навчального предмету, а найнижчим рівнем ієрархії є метрики  $\{M_{im}\}$ , які слугують засобом представлення оцінки властивостей середнього рівня.

Аналогічно до (2.3) можна описати професійність інструкторів курсів при провадженні дистанційної освіти. Формула (2.4) дуже подібна до попередньої, єдине, що відрізняється складом та інтерпретацією властивостей.

$$Q_T = \{H_i^T, A_{ij}^T, M_{ij}^T\} \quad (2.4)$$

$H_i^T$  – комплексна властивість показника професійності інструктора курсу,  
 $i = 1..n'$  - кількість комплексних властивостей;

$A_{ij}^T$  – елементарна властивість, що описує ефективність інструкторів курсу,  
 $j = 1..k'$  - кількість елементарних властивостей;

$M_{ij}^T$  – метрика для кількісного представлення оцінки професійності викладача.

Програмно-технічне забезпечення платформ віддаленого навчання та їх ефективність можна представити у вигляді моделей якості, наведених у [10].

Важливим показником і критерієм, який необхідно врахувати, щоб забезпечити та визначити успішність навчального процесу і наведених вище факторів впливу є відгук про курс тих, хто навчається. Для цього у платформах онлайн і дистанційної освіти передбачено модулі зворотного зв'язку, або функціональність, що дає змогу під час навчання або після проходження певних курсів формувати відгук про його актуальність, якість, простоту роботи та ін. Зазвичай, інформація у вигляді відгуків може бути прокласифікована за такими значеннями: «нейтральний», «позитивний», «негативний». Для формального представлення відгуків з позиції прогнозування успішності студентів можна представити в наступному вигляді:

$$UF = \{Q(A), Feedback\} \quad (2.5)$$

де  $Q(A)$ , – сукупність атрибутів представлених у (2.1);

*Feedback* – одне із значень множини, що належать категорійними змінним відносно типу відгуку і дає відповідь на питання щодо якості курсу та його впливу на успішність засвоєння і подачі матеріалу.

*Feedback* можна представити у вигляді трьохкомпонентної множини

$$Feedback = \{P, N, Nt\} \quad (2.6)$$

де  $P$  – значення, що позитивно визначає вплив на успішність;

$N$  – значення, що відповідає негативному відгуку і впливу на успішність;

$Nt$  – значення, що відповідає за те, що студент нічого нового для себе не дізнався і курс жодним чином не впливає на його успішність.

Таким чином побудовано та математично представлено модель, яка дозволяє описати та визначити фактори впливу на успішність студентів при використанні СДН та платформ електронного навчання.

## 2.2. Методи інтелектуалізованого прогнозування успішності студентів

Проблема прогнозування успішності студентів частково досліджувалась спільнотою аналітиків та фахівцями з аналізу освітніх даних у формі прогнозування імовірності відрахування (або завершення навчання) студента. Ця задача є важливою проблемою підкласу прогнозування успішності здобувачів освіти. Більшість таких робіт можна поділити на дві групи, в залежності від використовуваних підходів.

Перший класичний підхід в основному ґрунтується на узагальнених лінійних моделях, які включають логістичну регресію, лінійний SVM та аналіз «виживання». Кожна модель враховує різні типи поведінкових і прогнозних особливостей, добутих з різних неопрацьованих записів активності студентів (наприклад, потоку кліків, оцінок, форуму, оцінки).

Другий, новий підхід, включає дослідження успішності з використанням нейронних мереж. Дослідники використовували моделі глибоких нейронних мереж (DNN), рекурентні та згорткові нейронні мережі. Проте всі ці нові моделі, поки що, показали примітивну продуктивність. Це пов'язано головним чином із тим, що моделі все ще покладаються на розробку функцій для зменшення вхідних розмірів даних, що, здається, обмежує розробку більших (тобто кращих) моделей нейронних мереж. Записи діяльності студентів, зібрані з різних курсів, часто мають різну довжину, формат і зміст, тому функції, ефективні в одному курсі, можуть бути не такими в іншому. Навіть ретельно розроблені розміри об'єкта зазвичай мають бути малими

Обидва підходи формують вхідні дані, які поки що занадто обмежені, щоб використати всі переваги моделей послідовного глибокого навчання. Щоб уникнути недоліків попередніх робіт, GritNet сприймає навчальну діяльність здобувачів освіти у часі як вихідні дані і (неявно) шукає частини послідовності вбудовування подій, які є найбільш дискримінаційними для прогнозування успішності студента без необхідності розробляти ці частини як (явні) вхідні ознаки.

У кваліфікаційній роботі на основі алгоритмів машинного навчання, зокрема дерев прийняття рішень та методу опорних векторів, пропонується визначити найбільш вагомі фактори впливу на успішність студента, а за допомогою нейронних мереж спрогнозувати майбутні оцінки здобувача освіти.

### 2.3. Формальне представлення алгоритмів прогнозування успішності студента з використанням дерев прийняття рішень та методу опорних векторів

У машинному навчанні досить потужним та ефективним інструментом для розв'язання задач прогнозування, класифікації і регресії є підхід на основі дерев прийняття рішень та побудованих на ньому підході «випадкових лісів». Дерева можна зобразити у вигляді моделі, яка у циклі перебирає варіанти реалізації різних функцій, що виконуються над даними послідовно. Якщо говорити про випадковий ліс, то він представляє собою сукупність чи по-іншому ансамбль таких дерев, що застосовують довільне, випадкове впорядкування даних у дата фреймі.

Можна розглянути приклад використання decision tree для прогнозування рішення друга спортсмена щодо гри у футбол. Професійний футболіст займається спортом кожного понеділка і на тренування запрошує свого товариша, який інколи приймає запрошення, а деколи відмовляється.

Друг приймає рішення щодо відвідування тренування, аналізуючи і враховуючи множину факторів:

- чи хороша погода;
- яка температура повітря;
- чи сильний вітер;
- чи він не надто втомлений у той час.

Професійний футболіст звертає увагу на такі чинники впливу і спостерігає, як товариш приймає рішення щодо відвідування чи не відвідування тренування. Такий набір факторів доцільно використати при формуванні прогнозів з використанням підходу «дерев прийняття рішень», як показано на рис. 2.1.

Кожне дерево містить два види елементів:

- «вершини» – є сегментами дерева і у яких відбувається розгалуження з врахуванням значення певного критерію чи фактора впливу;
- «ребра» – зв'язки, які утворюються у результаті декомпозиції вершини і які зв'язують інші вузли дерева.

Аналізуючи рис. 2.1, можна сказати, що дерево складається з вершин, які містять прогнозовані значення, показники вологості повітря і силу та швидкість вітру. Крім цього, на даному рисунку можна побачити також і крайні значення інтервалу для кожного з факторів впливу відносно прийняття рішення щодо відвідування тренування.

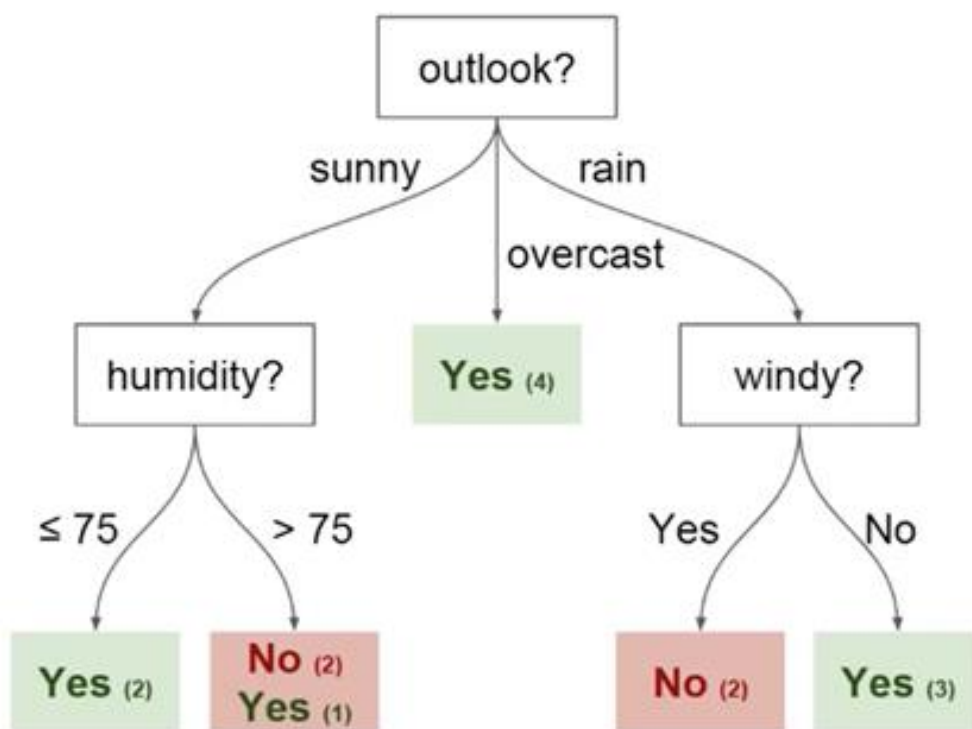


Рис. 2.1. Структура «дерева прийняття рішень»

Термінологічно, при використанні «дерев прийняття рішень» керуються означеннями характерними для ієрархічних структур даних, зокрема:

- кореневий вузол – основна початкова вершина дерева з якої стартує декомпозиція;
- листки – кінцеві вершини, які не містять розгалужень, а призначені для формування прогнозного значення.



Процедура побудови і використання підходу decision trees є досить складною і трудомісткою, ніж може здаватися спочатку. Фахівці з машинного навчання будують не одне, а множину дерев з використанням рандомних факторів й ознак наявних в наборі даних і призначених для реалізації декомпозиції. По-іншому кожен раз створюються нові випадкові сукупності ознак і подаються на кожну ієрархічну структуру дерева. Таким чином, утворюється не одне дерево прийняття рішення, а цілі «випадкові ліси».

Загальний принцип за яким визначають об'єм випадкових ознак ( $m$ ) повинен відповідати значенню, що рівне квадратному кореню від загальної кількості факторів або ознак ( $p$ ) у наявному фреймі даних. Тобто  $m = \sqrt{p}$ , а конкретна ознака випадково обирається з утвореного набору  $m$ .

Приклад бінарної класифікації при розв'язанні задачі прогнозування показано на рис. 2.2.

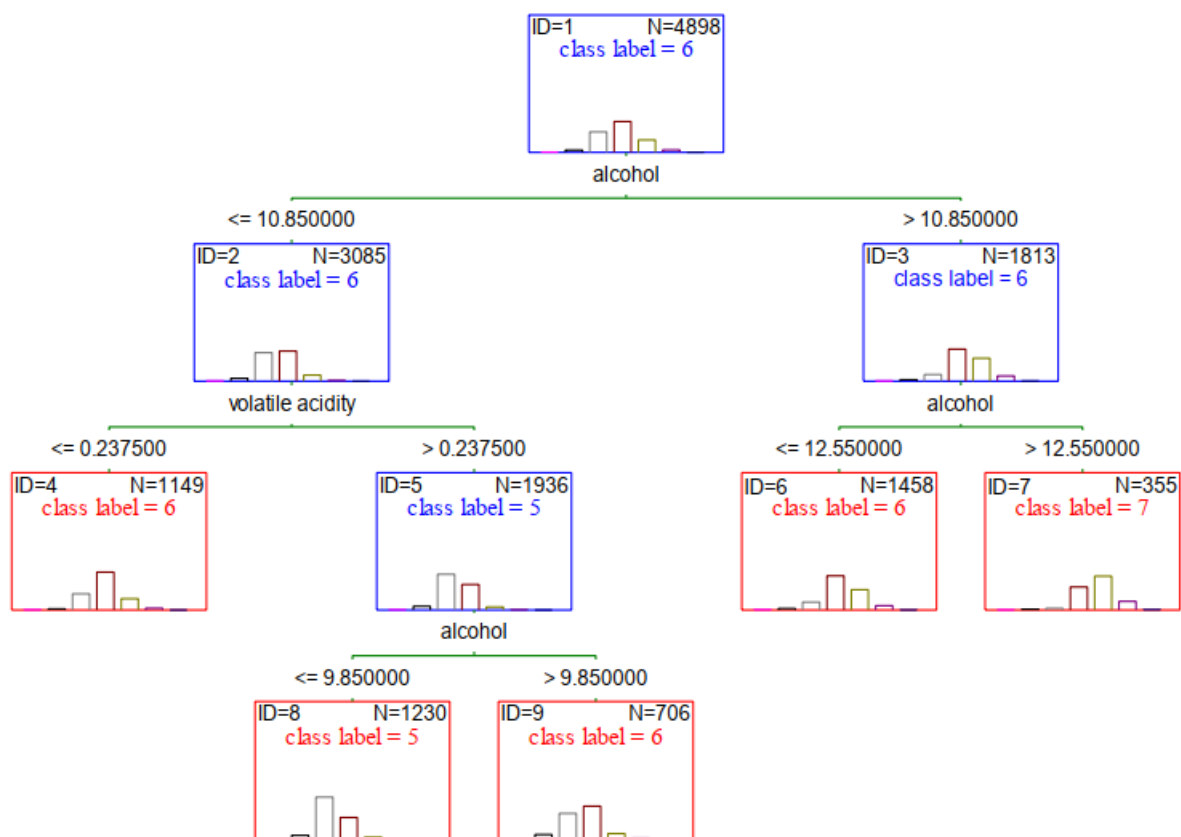


Рис. 2.2. Приклад бінарної класифікації із застосуванням дерева прийняття рішень

Основною перевагою застосування підходу на основі «випадкового лісу» є здатність опрацювання значного за розміром фрейму даних, який містить ознаку, що дуже сильно впливає на результат прогнозування. По-іншому, у сукупності вхідного набору ознак наявна така ознака, що за її значенням прогнозований результат є достовірним і практично не залежить від інших факторів впливу, що містяться у наборі даних.

При ручному формуванні decision tree доцільно саме цю ознаку чи фактор поміщати у найвищу вершину дерева для виконання подальшої декомпозиції. Це призводить до появи декількох ієрархічних структур, значення прогнозу яких є сильно корельованими. Цього потрібно уникати, оскільки математичне сподівання сильно залежних між собою змінних не забезпечує значного зниження дисперсії.

Для виконання процесу декореляції та зменшення дисперсії результатів прогнозування потрібно також використовувати випадкові множини ознак для кожної конкретної структури дерева, що належить до «випадкового лісу».

Застосування методу декореляції забезпечує зниження рівня залежності між ознаками і є основною перевагою «випадкових лісів» над простими «деревами прийняття рішень».

До переваг застосування алгоритмів «дерев прийняття рішень» та «випадкових лісів» належать:

- висока ефективність результатів аналізу нелінійних залежностей між ознаками у фреймі даних, що характеризується значною продуктивністю у порівнянні з різними видами регресії, та практично на одному рівні з підходом на основі нейронних мереж;

- базові алгоритми decision trees є доволі простими для розуміння і практичної імплементації, а ліміти, визначені при тренуванні моделі є зрозумілими та обґрунтованими.

До недоліків такого підходу належать:

- схильність до перенавчання і складна ієрархічна структуру, яка містить надлишкові дані;

– практична реалізація «випадкових лісів» передбачає використання значних програмно-апаратних ресурсів, що відображається на часовій продуктивності алгоритму прогнозування та використанні значного об'єму оперативної пам'яті.

Через специфічну і високодисперсну природу регресії, регресори дерева прийняття рішень потрібно ретельно обрізати. Проте, підхід до регресії нерегулярний – замість того, щоб обчислювати значення безперервно, він приходить до заданим кінцевих вузлів. Якщо регресор обрізаний занадто сильно, у нього занадто мало кінцевих вузлів, щоб належним чином виконати своє завдання.

Отже, дерево прийняття рішень повинно бути обрізане так, щоб воно мало найбільшу свободу (можливі вихідні значення регресії – кількість кінцевих вузлів), але недостатньо, щоб воно було занадто глибоким. Якщо його не обрізати, то і без того високодисперсний алгоритм стане надмірно складним через природу регресії

Метод опорних векторів або англійською мовою SVM («Support Vector Machine») представляю собою один з видів машинного навчання під наглядом, який широко використовується в задачах розпізнавання образів і класифікації.

Алгоритм SVM виконує класифікацію шляхом побудови багатовимірної гіперплощини, яка оптимально розрізняє два класи, максимізуючи запас між двома кластерами даних.

Цей алгоритм досягає високої дискримінаційної потужності, використовуючи спеціальні нелінійні функції, які називаються ядрами, для перетворення вхідного простору ознак у багатовимірний простір [17].

Основна ідея техніки цієї техніки полягає в тому, щоб побудувати  $n-1$  вимірну гіперплощину поділу даних для виділення двох класів у  $n$ -вимірному просторі. Точка даних розглядається як  $n$ -вимірний вектор.

Наприклад, дві змінні в наборі даних утворюють двовимірний простір. Гіперплощина, яка виконує поділ, була б представлена прямою (одновимірною), що розділяє простір навпіл.

Коли існує більше, ніж два класи вимірів, метод опорних векторів шукає оптимальну розділювальну гіперплощину, яка називається гіперплощиною відокремлення з максимальним запасом.

На рис. 2.3 показано приклад нерозривного двовимірного простору, який стає роздільним після перетворення вхідного простору з низьковимірною у багатовимірний.

Підхід SVM має тенденцію до класифікації сутностей без надання оцінок ймовірності приналежності до класу в наборі даних, що є принциповою відмінністю від множинної логістичної регресії.

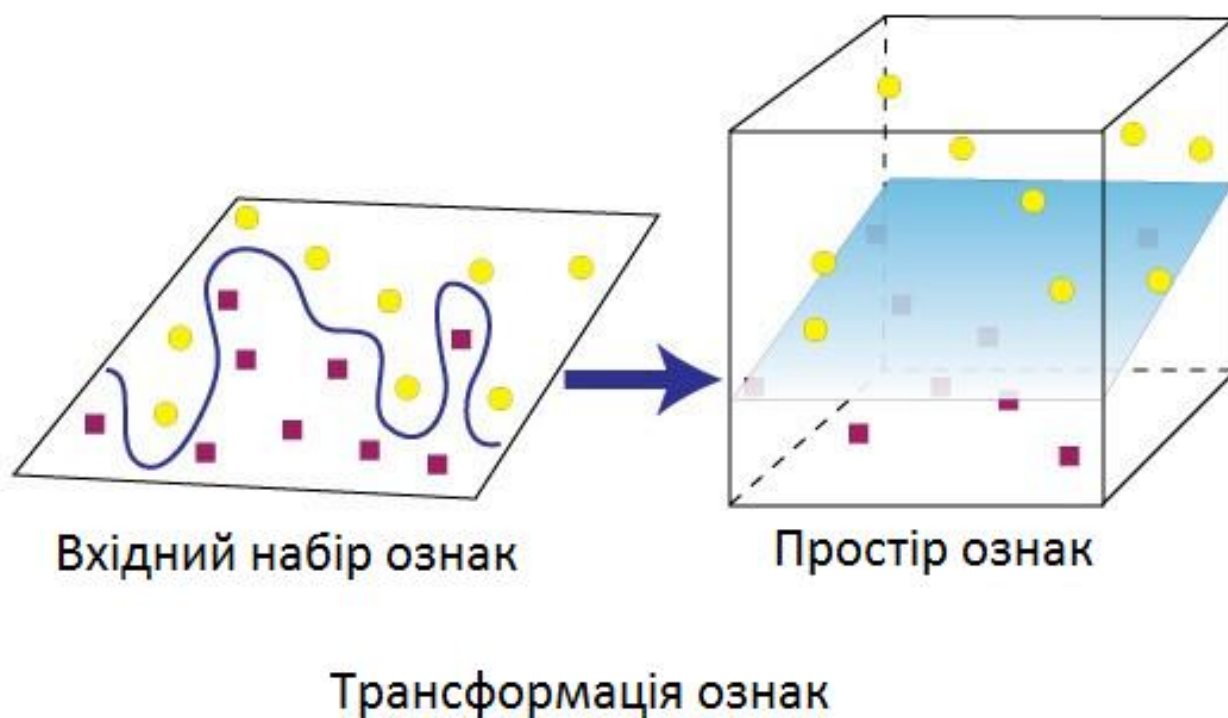


Рис. 2.3. Приклад застосування методу опорних векторів у просторі ознак

Відстань між гіперплощиною та найближчою точкою даних на кожній стороні (так звані опорні вектори) максимізується. Найкращий сценарій полягає в тому, що два класи розділені лінійною гіперплощиною. Однак реальні ситуації не завжди такі прості. Деякі точки даних у двох класах можуть потрапити у так звану «сіру» область, яку нелегко відокремити.

SVM вирішує цю проблему:

- дозволяючи деяким даним вказувати на неправильну сторону гіперплощини, вводячи заданий користувачем параметр  $C$ , який визначає компроміс між мінімізацією помилкової класифікації та максимізацією маржі;
- використання функцій ядра (зазвичай включаючи лінійні, поліноміальні, сигмоїдні та радіальні базисні функції) для додавання більшої кількості вимірів до низьковимірного простору, у результаті чого два класи можуть бути розділені у просторі високої розмірності.

Для створення оптимальної моделі SVM необхідно попередньо вибрати два ключових параметри для ядер –  $C$  і  $\gamma$ . Параметр  $C$  контролює надмірне пристосування моделі, вказуючи допуск на неправильну класифікацію. Параметр  $\gamma$  контролює ступінь нелінійності моделі. Приклад застосування і суті методу опорних векторів показано на рис. 2.5.

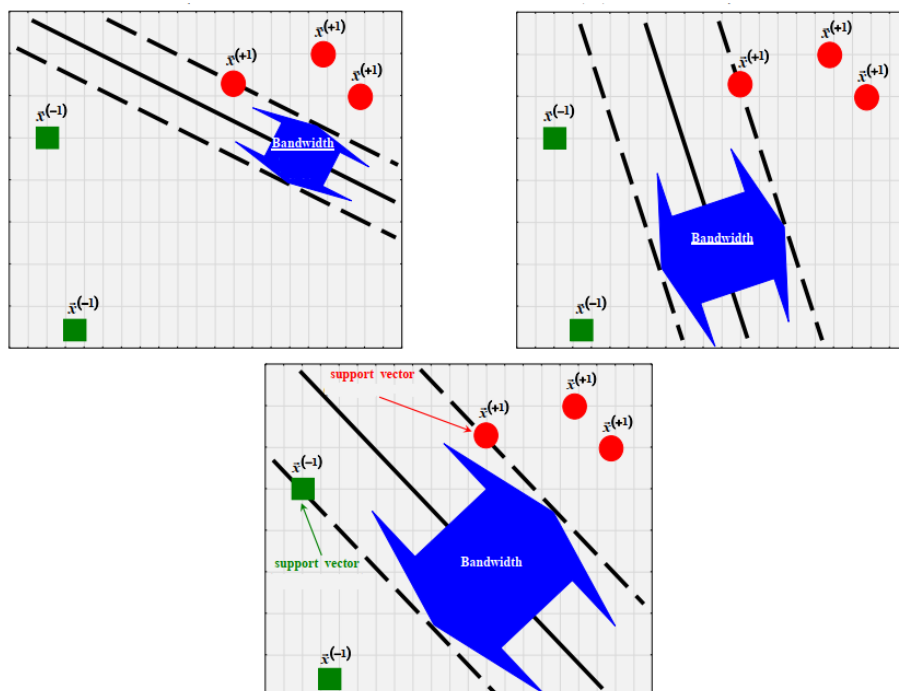


Рис. 2.5. Приклад застосування SVM

Дані, які підлягають відокремленню/класифікації при прогнозуванні, представлені як унікальна точка в просторі, де кожна з них представлена деяким вектором ознак

$$x \in R^D \quad (2.6)$$

де  $R^D$  – векторний простір розмірності  $D$ .

Далі, відображаючи точку даних у комплексному просторі ознак  $x$ , отримаємо

$$\Phi(x) \in R^M \quad (2.7)$$

Перетворений простір ознак для кожного вхідного об'єкта, відображеного на перетворений базисний вектор  $\Phi(x)$ , можна визначити як:

$$\phi(x): R^D \rightarrow R^M \quad (2.8)$$

Отже, тепер, коли точки візуалізовано, наступний крок полягає в тому, щоб розділити ці точки за допомогою лінії, і ось тут з'являється термін «межа рішення». Межа рішення є основним роздільником для поділу точок на відповідні класи.

Рівняння головної роздільної лінії називається рівнянням гіперплощини показано на рис. 2.6.

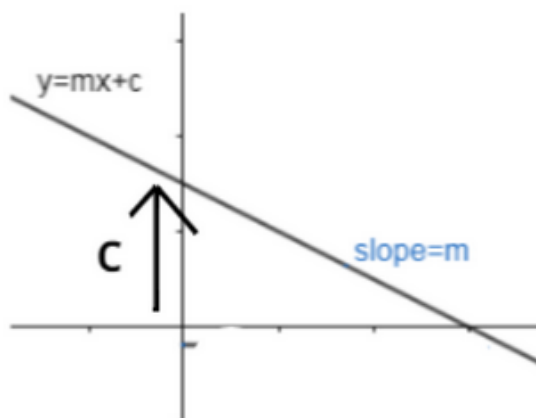


Рис. 2.6. Рівняння гіперплощини

Тепер подивимося на рівняння для прямої з нахилом  $m$  і перерізом  $c$ . Рівняння має вигляд:

$$mx + c = 0 \quad (2.9)$$

У даному випадку обрано пряму лінію, яка знаходиться в одно- або двовимірному просторі. Рівняння гіперплощини, що розділяє точки при класифікації, тепер можна легко записати як:

$$H: w^T(x) + b = 0 \quad (2.10)$$

де  $b$  – член перехоплення та зміщення рівняння гіперплощини. У багатовимірному просторі  $D$  гіперплощина завжди буде оператором  $D-1$ . Наприклад, для двовимірного простору гіперплощина є прямою.

Тепер зрозуміло, як представити точки даних і як розташувати лінію поділу точками. Але, встановлюючи таку лінію, очевидно, необхідним є одержання такої лінії, яка б могла найкращим чином відокремити точки даних, маючи найменшу кількість помилок неправильної класифікації. Отже, щоб мати найменші помилки в класифікації точок даних, ця концепція вимагатиме спочатку визначення відстані між точкою даних і розділювальною лінією (рис. 2.7).

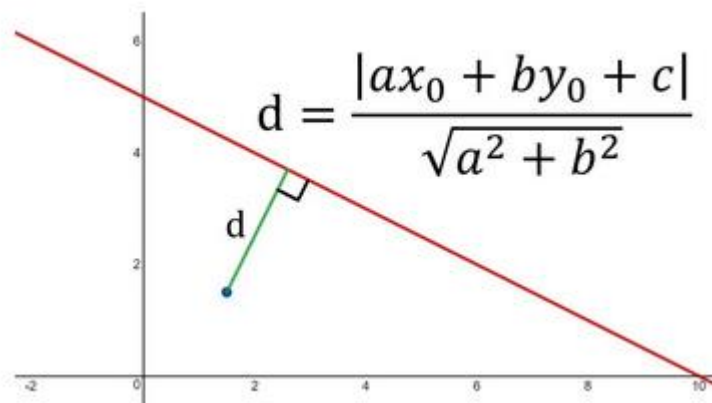


Рис. 2.7. Відстань від точки даних до лінії поділу

По аналогії, відстань від точки до рівняння гіперплощини (2.10) від заданого вектора точки  $\Phi(x_0)$  можна легко записати як формулу гіперплощини

$$d_H(\phi(x_0)) = \frac{|\omega^T(\phi(x_0)) + b|}{\|\omega\|_2} \quad (2.11)$$

де  $\|\omega\|_2$  – евклідова відстань, що обчислюється за формулою:

$$\|\omega\|_2 =: \sqrt{\omega_1^2 + \omega_2^2 + \omega_3^2 + \dots + \omega_T^2} \quad (2.12)$$

Таким чином, проаналізовано та обґрунтовано формальне представлення дерев прийняття рішень та методу опорних векторів для прогнозування успішності студентів, що в подальшому реалізовано у вигляді програмної моделі.

#### 2.4. Метод прогнозування успішності навчання з використанням глибоких нейронних мереж

Задачу прогнозування успішності студентів можна виразити у вигляді задачі послідовного прогнозування подій [8]: враховуючи попередню (історичну) послідовність  $o \triangleq (o_1, o_2, \dots, o_T)$  у яких бере участь здобувач освіти потрібно оцінити імовірність майбутньої послідовності подій  $y \triangleq (y_{T+D}, \dots, y_{T'})$ , де  $D \in Z_+$ .

У формі онлайн-занять кожна подія представляється у вигляді деякої діяльності студента, яка пов'язана з відповідною міткою часу. Іншими словами,  $o_t$  визначається як кортеж  $(a_t, d_t)$ .

Кожна дія  $a_t$  може представляти собою, наприклад, «переглянуто відео лекції», «відповідь на тестове запитання правильне/неправильне» або «проект виконано та здано/не здано», а  $d_t$  вказує відповідну (zareestrovannu) мітку часу.

Тоді логарифмічна функція правдоподібності  $p(y|o)$  може бути записана у вигляді рівняння наведеного нижче, заданої фіксованої розмірності вкладеності  $v$ .



$$\log p(y|o) \cong \sum_{i=T+D}^{T'} \log p(y_i|v) \quad (2.13)$$

Таким чином, метою GritNet є обчислення індивідуального журналу логарифмічної правдоподібності  $p(y_i|v; \theta)$ , і ці оцінені бали можна просто додати для оцінки довгострокових результатів успішності студента.

Щоб оцінити, яку додану вартість приносить GritNet, логістична регресія використовується як базова модель. Крім того, застосовується модель «мішка слів» для представлення послідовності минулих подій кожного студента  $o$ .

Після перетворення даних щодо всієї діяльності здобувачів онлайн освіти у «мішок слів» проводиться обчислення того, скільки разів кожна унікальна діяльність з'являється в  $o$ .

Нехай представлення ознаки  $v$  з фіксованою розмірністю є  $N$ -вимірним вектором ознак, де  $v_j \in \mathbb{Z}_{\geq 0}$ . Врахування  $v$  у моделі логістичної регресії  $\log p(y_i|v)$  виконується таким чином:

$$\log p(y_i = 1|v; \theta) = \frac{1}{1 + \exp(-\theta^T v)} \quad (2.14)$$

де  $\theta \in \mathbb{R}^N$  – параметри моделі логістичної регресії.

Для  $M$  навчальних екземплярів  $\{(v^{(k)}, y^{(k)})\}_{k=1}^M$ ,  $L_2$  логістична регресія з регуляризаторами знаходить параметри  $\theta$ , які вирішують наступну задачу оптимізації:

$$\arg \max_{\theta} \sum_{k=1}^m \log p(y^{(k)}|v^{(k)}; \theta) + \alpha \|\theta\|_2 \quad (2.15)$$

Однак буде зручніше розглядати регуляризовану логістичну регресію  $L_1$  замість рівняння (2.14) для опрацювання невідповідних ознак [17]. Варто відмітити, що ще простіші методи вибору ознак (наприклад, на основі оцінки  $\chi^2$ ) у поєднанні

з упорядкованою логістичною регресією  $L_2$  дають результати, які дуже схожі до результатів на основі  $L_1$ .

Архітектура нейронної мережі GritNet показано на рис. 2.8.

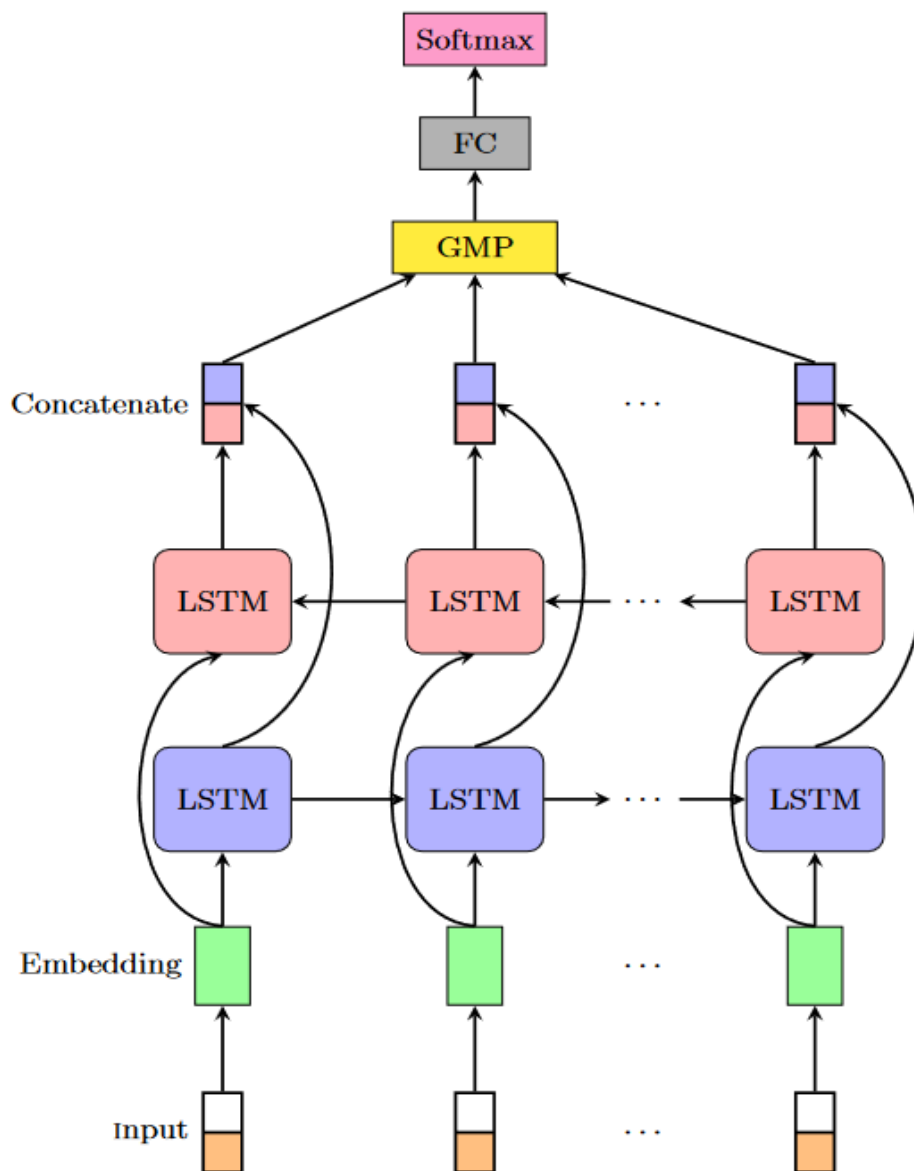


Рис. 2.8. Архітектура нейронної мережі GritNet

До складу нейронної мережі GritNet входять:

- вхідний шар ознак успішності студента;
- Приховані шарів на основі LSTM;
- GMP – забезпечує оптимізацію параметрів моделі;

- FC – повнозв'язний шар;
- Softmax – функція мультиноміальної класифікації.

Для передачі вихідних записів студентів до GritNet, необхідно закодувати часові мітки (впорядковані послідовно) у послідовність вхідних векторів фіксованої довжини. Це можна просто реалізувати за допомогою техніки «one-hot encoding». Закодований за допомогою цієї техніки вектор можна представити як  $\mathbb{1}(a_t) \in \{0,1\}^L$ , де  $L$  – кількість унікальних дій студента. Елементи вектора можна визначити як:

$$\mathbb{1}(a_t) = \begin{cases} 1, \text{ якщо } j = a_t \\ 0, \text{ в ін. випадку} \end{cases} \quad (2.16)$$

Формула (2.16) використовується, щоб відрізнити кожен діяльність  $a_t$  від будь-якої іншої. Після цього виконується конкатенація закодованих векторів одного і того ж студента в довгу векторну послідовність, щоб представити всю послідовну його діяльність в  $o$ .

Хоча цей метод кодування зберігає інформацію про порядок, на відміну від методу WoW, він накладає обмеження щодо швидкості навчання студентів. Змінна швидкість навчання є важливою частиною інформації, що залежить від часу, яка відображає прогрес студента та/або складність змісту курсу. Оскільки пряме використання кожної позначки часу  $d_t$  надто швидко збільшує вхідний простір, визначаємо дискретизовану різницю в часі між сусідніми подіями як:

$$\Delta_t \triangleq d_t - d_{t-1} \quad (2.17)$$

Після цього, таким самим чином закодовується  $\Delta_t$  у  $\mathbb{1}(\Delta_t)$  і конкатенується з відповідним  $\mathbb{1}(a_t)$ , що дозволяє представити  $\mathbb{1}(o_t)$  у вигляді:

$$\mathbb{1}(o_t) \triangleq [\mathbb{1}(a_t); \mathbb{1}(\Delta_t)] \quad (2.18)$$

Таким чином, виконано заповнення вихідної послідовності, даними, які є меншими і коротшими за максимальну довжину послідовності подій всіх  $o$  векторів.

Ядром запропонованої моделі GritNet є вбудовані шари двонаправленого LSTM та GMP, які навчені для прийому і навчання на історичних студентських подіях і прогнозування ймовірності успішності в майбутньому.

Перший вбудований шар визначає і досліджує матрицю  $E^o \in \mathbb{R}^{E \times |O|}$ , де  $E$  та  $|O|$  – розмірність інтеграції та кількість унікальних подій, що відповідає розміру вхідного вектора  $\mathbb{l}(o_t)$  з метою перетворення його перетворення у простір меншої розмірності  $v_t$ , який визначається наступним чином

$$v_t \triangleq E^o \mathbb{l}(o_t) \quad (2.19)$$

Після цього, дана вбудована подія передається в BLSTM, і вихідні вектори формуються шляхом конкатенації кожного виходу у прямому і зворотному напрямі. Далі перед вихідним шаром додається шар GMP. За допомогою шарів GMP GritNet вчиться фокусувати і «звертати увагу» на найважливішу частину послідовність вбудованих подій, ігноруючи решту. Ця операція GMP здається вирішальною для підвищення ефективності прогнозування, особливо для незбалансованих даних, які надаються без будь-якої функції інженерії даних. Вихід шару GMP, в кінцевому результаті, подається у повнозв'язний шар і шар softmax (функція активації сигмоподібна) послідовно для обчислення функції правдоподібності  $\log p(y_i; v)$ .

## 2.5. Висновки до розділу

У даному розділі одержано наступні наукові та практичні результати:

1. На основі аналізу факторів впливу на успішність студентів під час електронного навчання запропоновано їх формалізацію у вигляді елементів множини, що дало змогу побудувати модель, яка враховує ефективність

програмно-апаратного забезпечення, якість навчального контенту та кваліфікацію інструкторів курсів, які в подальшому будуть враховані при виборі інтелектуалізованих методів і засобів прогнозування успішності студентів.

2. Для прогнозування успішності студентів обґрунтовано та формально представлено алгоритми, які використовують підходи дерев прийняття рішень та випадкових лісів, методу опорних векторів, що в перспективі дасть змогу прогнозувати успішність студентів за трьома класами: «висока успішність», «достатня успішність», «відрахування студента».

3. Запропоновано метод прогнозування успішності студентів на основі підходу глибоких нейронних мереж з архітектурою GritNet в основі якого лежать вбудовані шари двонаправленого LSTM, що дає змогу більш ефективно, у порівнянні з базовою моделлю GritNet, забезпечити високу точність прогнозування фактів завершення або не завершення навчання студентів на курсі.

## РОЗДІЛ 3

### ПРОГРАМНА РЕАЛІЗАЦІЯ МОДЕЛЕЙ ПРОГНОЗУВАННЯ УСПІШНОСТІ СТУДЕНТІВ

#### 3.1. Прогнозування успішності студентів у системах електронного навчання на основі алгоритмів лінійної залежності

Мета, яка поставлена у даному підрозділі кваліфікаційної роботи полягає у програмній реалізації і дослідженні алгоритмів машинного навчання, зокрема дерев прийняття рішень та методу опорних векторів, які були б здатними адаптуватися до категорійних даних і спрогнозувати успішність студентів.

##### 3.1.1. Аналіз та візуалізація вхідних даних

В якості вхідного фрейму даних використано інформацію зі спеціалізованого ресурсу для опрацювання даних засобами штучного інтелекту «<https://www.kaggle.com>» (рис. 3.1).

gender	Nationality	PlaceofBirth	StageID	GradeID	Section
Male or Female					
M 64%	KW 37%	KuwaIT 38%	MiddleSchool 52%	G-02 31%	A
F 36%	Jordan 36%	Jordan 37%	lowerlevel 41%	G-08 24%	B
	Other (129) 27%	Other (124) 26%	Other (33) 7%	Other (217) 45%	Other ( )
M	KW	KuwaIT	lowerlevel	G-04	A
M	KW	KuwaIT	lowerlevel	G-04	A
M	KW	KuwaIT	lowerlevel	G-04	A
M	KW	KuwaIT	lowerlevel	G-04	A
M	KW	KuwaIT	lowerlevel	G-04	A
F	KW	KuwaIT	lowerlevel	G-04	A
M	KW	KuwaIT	MiddleSchool	G-07	A
M	KW	KuwaIT	MiddleSchool	G-07	A
F	KW	KuwaIT	MiddleSchool	G-07	A
F	KW	KuwaIT	MiddleSchool	G-07	B
M	KW	KuwaIT	MiddleSchool	G-07	A
M	KW	KuwaIT	MiddleSchool	G-07	B
M	KW	KuwaIT	lowerlevel	G-04	A
M	lebanon	lebanon	MiddleSchool	G-08	A

Рис. 3.1. Вхідний набір даних для прогнозування успішності студентів

Вхідний набір даних є структурованим і описує властивості, які впливають на успішність студентів та представлений у табличному вигляді. На рис. 3.1. наведено фрагмент дата фрейму на основі якого буде проводитись прогнозування успішності студентів.

В загальному випадку, вхідний набір даних міститиме наступну інформацію за стовпцями, як представлено у табл. 3.1.

Таблиця 3.1

### Поля набору даних для прогнозування успішності студента

Поле даних	Опис
gender	поле, що відповідає за стать студента і приймає значення «male» і «female»
NationalITy	текстове поле, що містить інформації про національність студента
PlaceofBirth	регіон або країна народження студента
StageID	Освітній рівень до якого належить студент (може набувати значень «Elementary», «Middle», «High School»).
GradeID	Рік навчання студента.
SectionID	Клас або група, у якому навчається студент.
Topic	Назва курсу в системі електронного навчання
Semester	Семестр навчального року («F» - осінній, «S» - весняний)
Relation	Відповідальність батьків за студента
raisedhands	Кількість разів, коли студент проявляв активність на занятті
VisITedResources	Кількість разів відвідування студентом контенту навчального курсу

Поле даних	Опис
AnnouncementsView	Кількість разів перевірки студентом оголошень курсу
Discussion	Кількість разів участі студента у дискусіях
ParentAnsweringSurvey	Чи беруть батьки участь в опитуванні закладу освіти (бінарне поле: так або ні)
ParentschoolSatisfaction	Задоволеність батьків успішністю студента
StudentAbsenceDays	Кількість випадків, коли студент був відсутній протягом більше, ніж семи днів
Class	Мітка успішності студента: «L» – для студентів, які отримали одержували оцінки «незадовільно» (0 – 69 балів); «M» – для студентів, які отримали низький прохідний бал (від 70 до 89 балів); «H» – для студентів, які досягли високих оцінок у своїх курсах (від 90 до 100 балів)

Зчитування даних відбувається шляхом виконання коду, представленого у лістингу 3.1.

### Лістинг 3.1. Зчитування даних про успішність студентів

```
import pandas as pd
import numpy as np
import seaborn as sns import warnings
# sklearn is a big source of pre-written and mostly optimized
ML algorithms.
# Here we use their Decision trees, Support Vector Machines,
and the classic Perceptron.
from sklearn import preprocessing, svm
from sklearn.linear_model import Perceptron
from sklearn.tree import DecisionTreeClassifier
#ignore warnings
warnings.filterwarnings('ignore')
data = pd.read_csv('../input/xAPI-Edu-Data.csv')
data.head()
```



Далі проведемо попередню візуалізацію даних, мета якої полягає в отриманні уявлення про вигляд і розподіл даних, а також побачити, чи можна визначити будь-які можливі відхилення.

Оскільки, успішність представляється у вигляді категоріальних змінних, тому особливу увагу варто звернути на графіки кількості полів даних і класів. Також виконується перевірка щодо надлишковості даних або їх незрозумілості. Для цього виконується лістинг 3.2.

Лістинг 3.2. Візуалізація важливих даних щодо успішності студентів

```
ax = sns.countplot(x='Class', data=data, order=['L', 'M', 'H'])  
for p in ax.patches:  
    ax.annotate('{:.2f}%'.format((p.get_height() * 100) /  
len(data)), (p.get_x() + 0.24, p.get_height() + 2))  
plt.show()
```

Результати візуалізації щодо розподілу значень успішності студентів за класами «L», «M», «H», які відповідають низькому, достатньому і високому рівню знань показано на рис. 3.

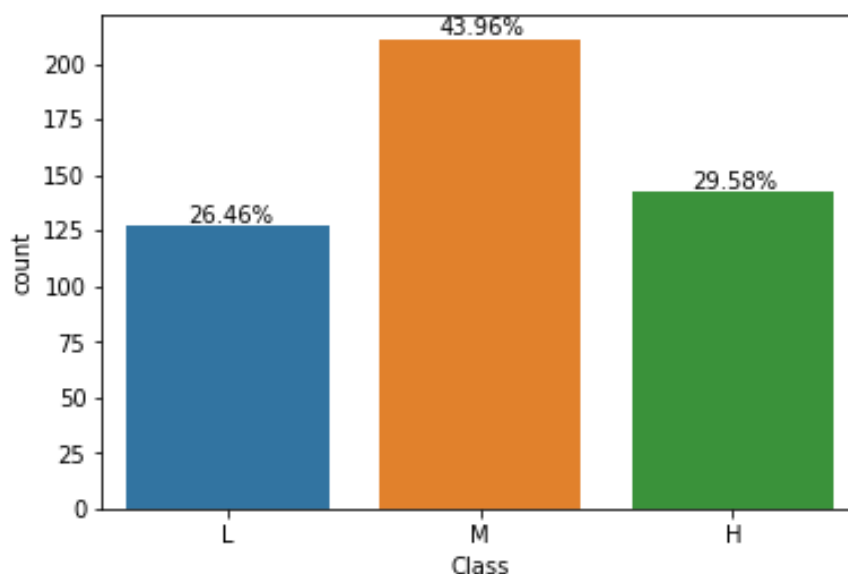


Рис. 3.2. Розподіл даних за успішністю студентів

З проведеного аналізу та візуалізації розподілу успішності, показаного на рис. 3.2., видно, що лише приблизно 30% студентів мають високі успіхи у навчанні, близько 44% – володіють достатнім рівнем знань і 26% є неуспішними.

Важливо також визначити схильність до навчання серед студентів чоловічої і жіночої статі. Для цього реалізовано скрипт мовою Python, як показано у лістингу 3.3.

### Лістинг 3.3. Розподіл успішності за статтю студентів

```
fig, axarr = plt.subplots(2, figsize=(10,10))
sns.countplot(x='gender', data=data, order=['M', 'F'],
ax=axarr[0])
sns.countplot(x='gender', hue='Class', data=data, order=['M',
'F'], hue_order = ['L', 'M', 'H'], ax=axarr[1])
plt.show()
```

Візуалізація результатів лістингу 3.3 показано на рис. 3.3

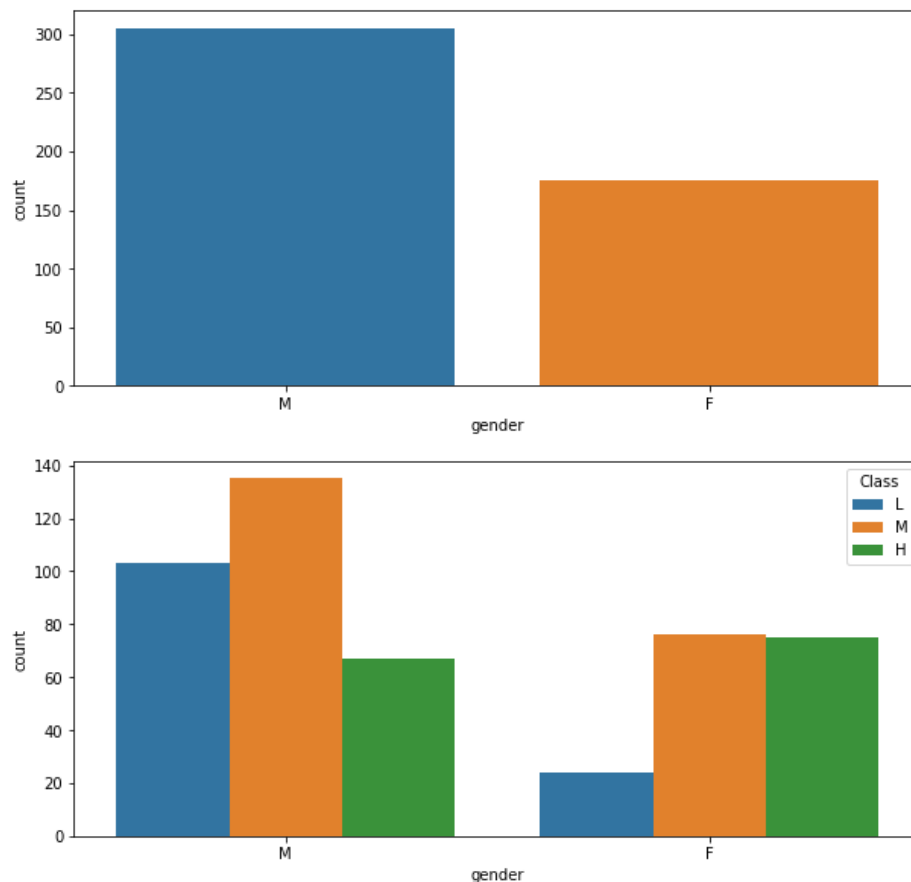


Рис. 3.3. Візуалізація розподілу успішності за статтю студентів

З результатів, показаних рис. 3.3. видно, що кількість студентів чоловічої статі, які здобувають освіту, майже на 40% відсотків більше, ніж студентів жіночої статі. Якщо говорити про успішність серед різних статей, то відсоток неуспішності серед чоловічої статі домінує. Для студентів жіночої статі характерним є приблизно однаковий розподіл за достатнім і високим рівнем успішності.

Лістинг 3.4 дозволяє визначити розподіл студентів за країнами їхнього походження, за національністю та відповідно за успішністю.

#### Лістинг 3.4. Розподіл студентів за національністю та успішністю

```
fig, axarr = plt.subplots(2, figsize=(10,10))
sns.countplot(x='NationalITY', data=data, ax=axarr[0])
sns.countplot(x='NationalITY', hue='Class',
data=data, hue_order = ['L', 'M', 'H'], ax=axarr[1])
plt.show()
```

На рис. 3.4 можна побачити результати щодо розподілу студентів за національністю і їхню успішність.

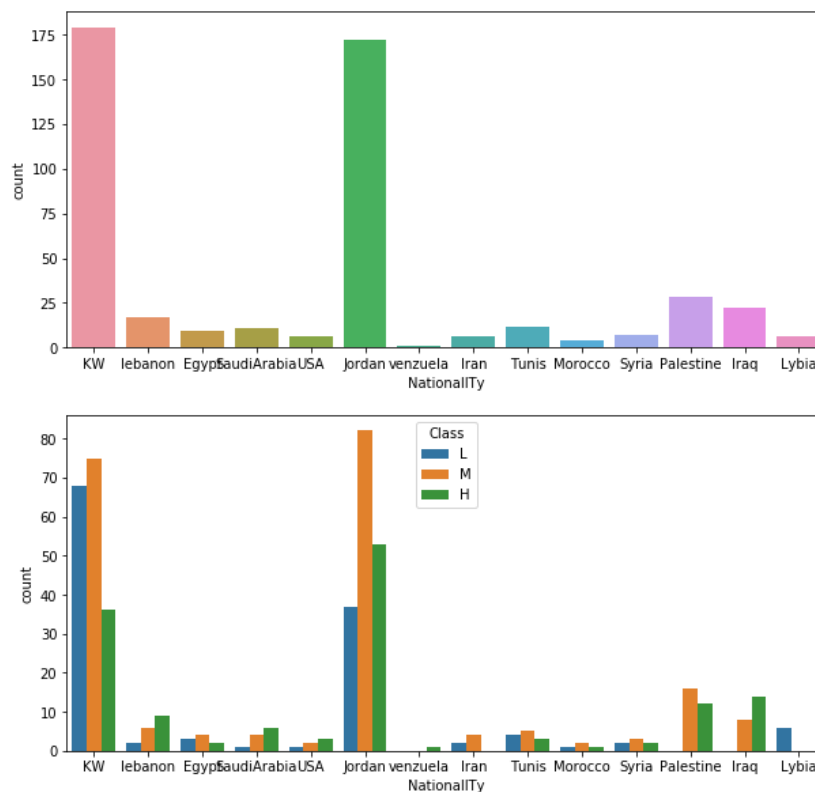


Рис. 3.4. Розподіл студентів за національністю та успішністю

Якщо говорити про національність студентів, то найбільше з тих хто навчається походять з Кувейту, Йорданії, Іраку та Палестини. Серед палестинців та іракців відсутні неуспішні студенти, а найбільший відсоток неуспішних здобувачів освіти серед тих, хто за національністю з Кувейту та Лівії.

Потенційно важливим фактором, що впливає на успішність є те, у якій школі навчається здобувач (рис. 3.5).

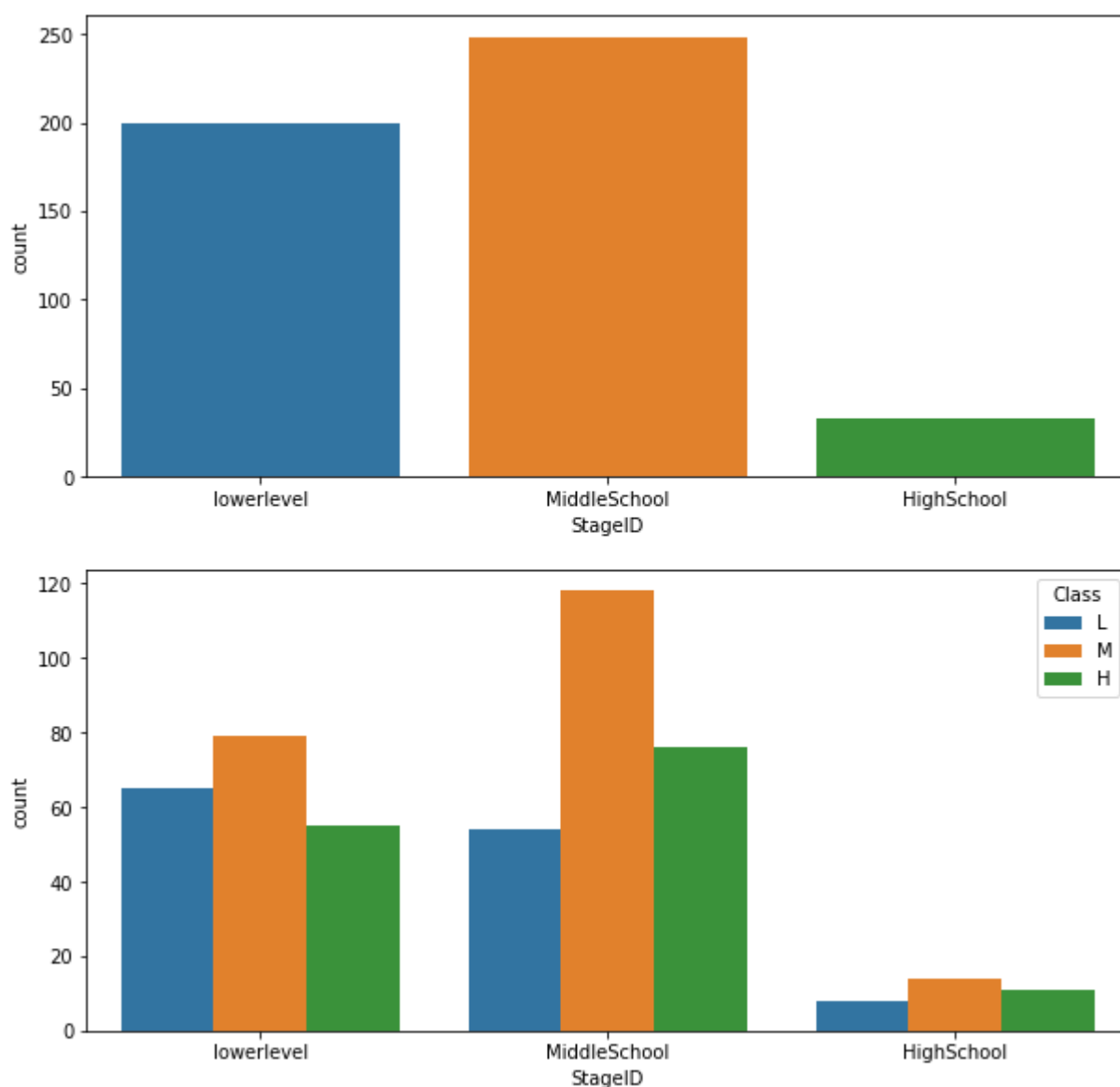


Рис. 3.5. Розподіл студентів за національністю та успішністю

Аналізуючи результати (рис. 3.5), найбільша кількість здобувачів освіти навчається у середній і початковій школі, а найменший відсоток – у вищій. Однак

спільною рисою усіх трьох рівнів освіти є те, що в них практично однаковий відсоток тих, хто відноситься до класу з достатнім рівнем успішності.

Виконавши програмний код лістингу 3.5 можна дізнатися розподіл успішності за різними рівнями освіти.

### Лістинг 3.5. Розподіл успішності за освітніми рівнями

```
fig, axarr = plt.subplots(2, figsize=(10,10))
sns.countplot(x='GradeID', data=data, order=['G-02', 'G-04',
'G-05', 'G-06', 'G-07', 'G-08', 'G-09', 'G-10', 'G-11', 'G-
12'], ax=axarr[0])
sns.countplot(x='GradeID', hue='Class', data=data, order=['G-
02', 'G-04', 'G-05', 'G-06', 'G-07', 'G-08', 'G-09', 'G-10',
'G-11', 'G-12'], hue_order = ['L', 'M', 'H'], ax=axarr[1])
plt.show()
```

Графічне представлення результатів програмного коду лістингу 3.5 представлено на рис. 3.6.

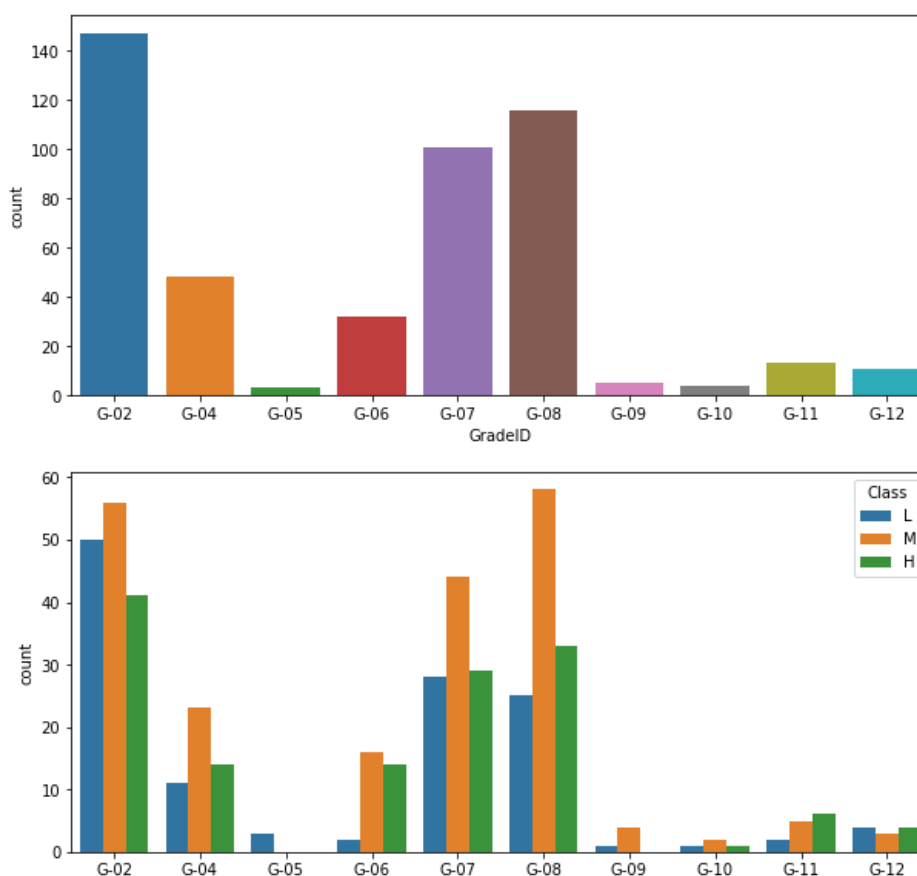


Рис. 3.6. Розподіл успішності студентів за класами (освітніми рівнями)

Дивлячись на результати (рис. 3.6) учні 5-х, 9-х і 10-х класів мають дуже мало оцінок. Крім того, жоден з учнів 5-го класу не здає завдань і жоден з учнів 9-х класів не досягає високих оцінок. Можна припустити, що це викиди і їх можна виключити з вибірки даних для прогнозування успішності студентів при навчанні за допомогою засобів електронного навчання. Однак потрібно більш детально переглянути дані за учнями та їх успішністю за зазначеними класами (лістинг 3.6).

### Лістинг 3.6. Детальна інформація щодо потенційних викидів

```
#Students in Grade 5
data.loc[data['GradeID'] == 'G-05']
#Students in Grade 9
data.loc[data['GradeID'] == 'G-09']
```

Детальна інформація щодо учнів 5 і 9 класів показано на рис. 3.7.

	gender	NationalTy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester	Relation	raisedhands	VisITedRe:
33	M	KW	KuwaIT	lowerlevel	G-05	A	English	F	Father	8	22
46	M	KW	KuwaIT	lowerlevel	G-05	A	English	F	Father	7	10
60	F	Jordan	Jordan	lowerlevel	G-05	A	English	F	Mum	21	10
< >											
	gender	NationalTy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester	Relation	raisedhands	VisITedRe:
42	M	KW	KuwaIT	HighSchool	G-09	A	IT	F	Father	10	12
43	F	KW	KuwaIT	HighSchool	G-09	A	IT	F	Father	30	35
44	F	KW	KuwaIT	HighSchool	G-09	A	IT	F	Father	33	33
76	M	Iran	Iran	HighSchool	G-09	A	IT	F	Mum	15	70
77	M	KW	KuwaIT	HighSchool	G-09	A	IT	F	Father	20	80
< >											

Рис. 3.7. Детальна інформація про учнів 5 і 9 класів

Після перегляду рядків записів, учні 5 класу мають подібні дані з усіма іншими учнями, які не здали завдань (пропущені більше 7 днів, низькі числові значення, відсутність опитування в школі тощо). І знову ж таки, після вивчення даних для учнів 9 класу це також виглядає так, як можна було б очікувати для кожної категорії.

### 3.1.2. Виявлення та аналіз кореляції факторів щодо успішності навчання

Перед тим, як проводити кореляційний аналіз, було досліджено залежність між успішністю учнів та складом сім'ї. Так у лістингу 3.7 реалізовано функціонал візуалізації таких залежностей, а на рис. 3.8 – наведено графічне представлення.

#### Лістинг 3.7. Кореляція між успішністю і складом сім'ї

```
fig, axarr = plt.subplots(2, figsize=(10,10))
sns.countplot(x='Relation', data=data, ax = axarr[0])
sns.countplot(x='Relation', hue='Class', data=data, hue_order =
['L', 'M', 'H'], ax = axarr[1])
plt.show()
```

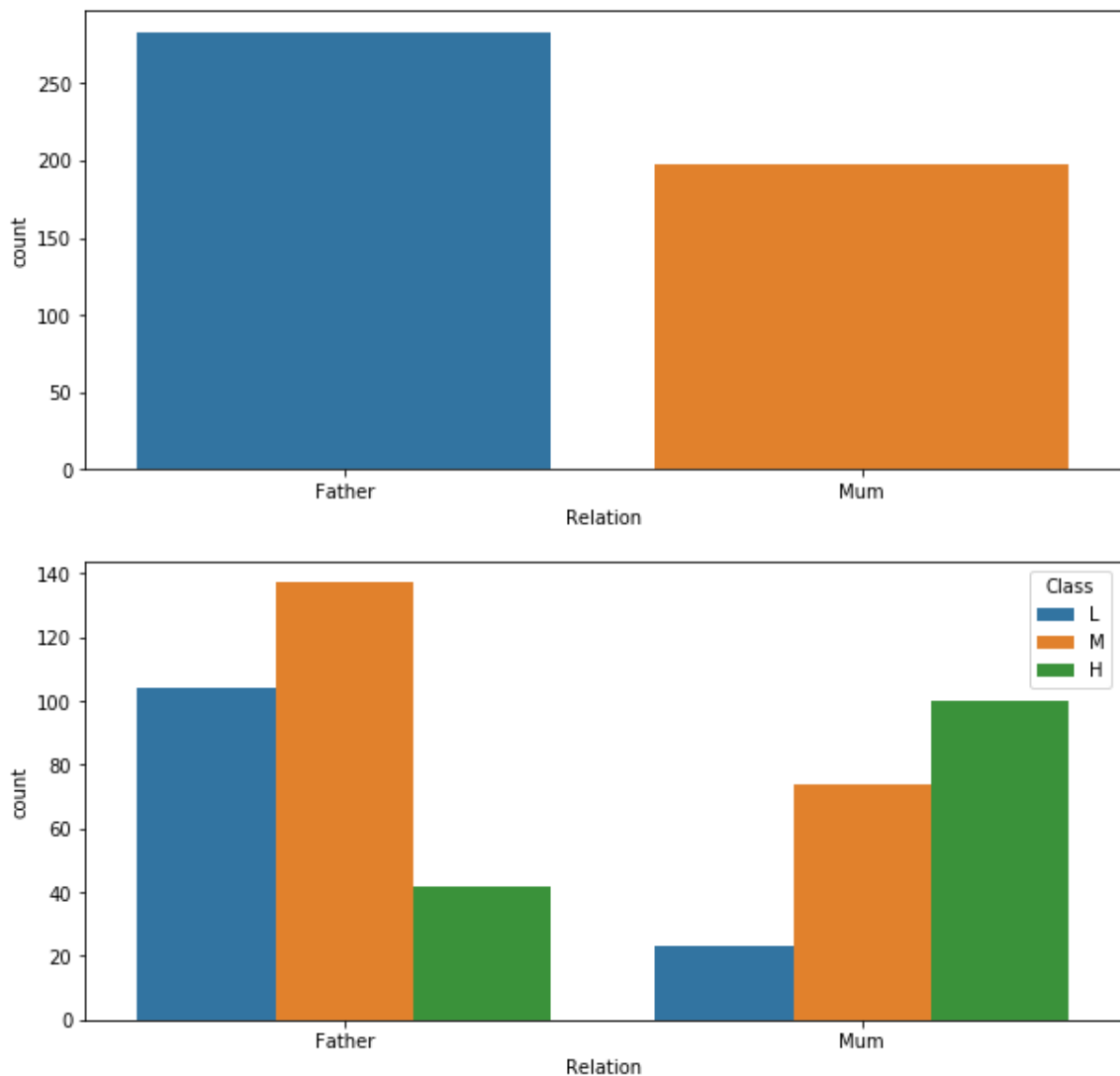


Рис. 3.8. Кореляційний зв'язок між успішністю і складом сім'ї

Дивлячись на рис. 3.8 і детальні дані попереднього аналізу, здається, існує кореляція між учнями, у яких матері є основними опікунами, та учнями, які менш схильні до невдач.

Далі можна подивитися та проаналізувати кількісні показники, тобто зафіксована кількість разів, коли студент:

- проявив активність на занятті;
- відвідав ресурси курсу;
- переглянув сторінку оголошення онлайн-курсу;
- відвідав сторінки обговорення.

Для легшого візуального порівняння ці результати зведено на один рисунок разом (рис. 3.9).

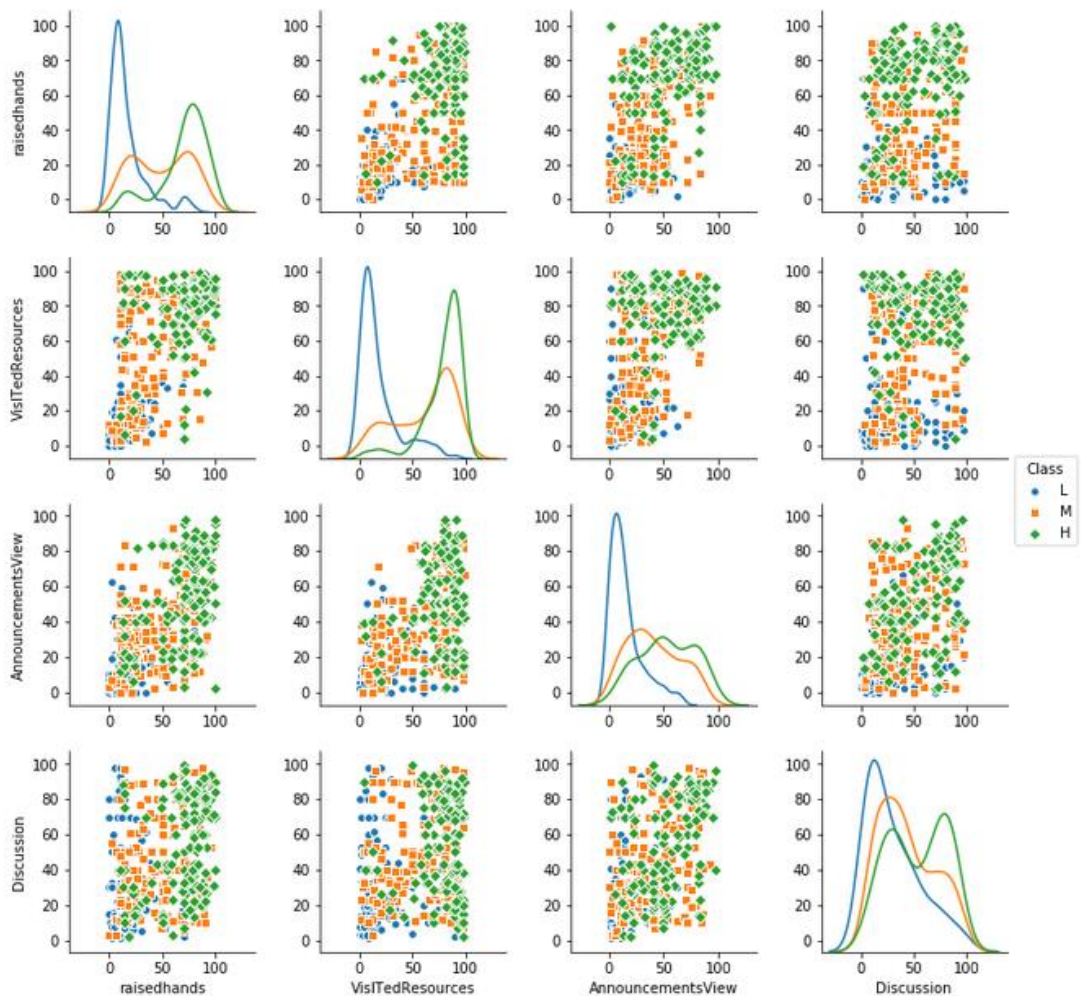


Рис. 3.9. Візуалізація залежностей за кількісними показниками



Фактор «StudentAbsenceDays» має сильну кореляцію із змінною класу. Дуже небагатьом студентам, які пропустили більше 7 днів, вдалося досягти високих оцінок, і дуже мало студентів, які пропустили менше 7 днів, провалили курс.

Попереднє опрацювання даних у наявному дата фреймі полягає у проведенні процедури стандартизації для заміни значень полів, які подібні до GradeID, на числове значення, таким чином, щоб зберегти відстань. Крім того, необхідно поставити у відповідність трьом класам числові результати із збереженою дистанцією. Існує кілька способів зробити це.

Пропонується для класів успішності встановити наступні значення «L» = -1, «M» = 0 і «H» = 1. Крім того, можна обчислити для кожного класу середнє значення за категорією на відносній відсотковій шкалі («L» = 34,5, «M» = 79,5 і «H» = 95) . Таким чином зберігається відстань між категоріальними значеннями.

Крім того, потрібно масштабувати значення числових полів, щоб вони були більш значущими для комплексного порівняння. Для цього можна використати вбудовану можливість масштабування попередньої обробки бібліотеки scikit learn (лістинг 3.8).

### Лістинг 3.8. Стандартизація числових значень

```
# Translate GradeID from categorical to numerical
gradeID_dict = {"G-01" : 1,
                "G-02" : 2,
                "G-03" : 3,
                "G-04" : 4,
                "G-05" : 5,
                "G-06" : 6,
                "G-07" : 7,
                "G-08" : 8,
                "G-09" : 9,
                "G-10" : 10,
                "G-11" : 11,
                "G-12" : 12}

data = data.replace({"GradeID" : gradeID_dict})

class_dict = {"L" : -1,
              "M" : 0,
              "H" : 1}
data = data.replace({"Class" : class_dict})
```

```

# Scale numerical fields
data["GradeID"] = preprocessing.scale(data["GradeID"])
data["raisedhands"] = preprocessing.scale(data["raisedhands"])
data["VisITedResources"] =
preprocessing.scale(data["VisITedResources"])
data["AnnouncementsView"] =
preprocessing.scale(data["AnnouncementsView"])
data["Discussion"] = preprocessing.scale(data["Discussion"])

# Use dummy variables for categorical fields
data = pd.get_dummies(data, columns=["gender",
                                     "NationalITy",
                                     "PlaceofBirth",
                                     "SectionID",
                                     "StageID",
                                     "Topic",
                                     "Semester",
                                     "Relation",
                                     "ParentAnsweringSurvey",

                                     "ParentschoolSatisfaction",
                                     "StudentAbsenceDays"])

# Show preprocessed data
data.head()

```

Результат стандартизації числових значень і кодування категоріальних змінних показано на рис. 3.10.

	GradeID	raisedhands	VisITedResources	AnnouncementsView	Discussion	Class	gender_F	gender_M	NationalITy_Eg
0	-0.563838	-1.033429	-1.174075	-1.351167	-0.843326	0	0	1	0
1	-0.563838	-0.870813	-1.053029	-1.313549	-0.662225	0	0	1	0
2	-0.563838	-1.196046	-1.446426	-1.426401	-0.481125	-1	0	1	0
3	-0.563838	-0.545579	-0.901723	-1.238315	-0.300024	-1	0	1	0
4	-0.563838	-0.220346	-0.145191	-0.974994	0.243279	0	0	1	0

Рис. 3.10. Стандартизований дата фрейм

Одним з основних методів опрацювання категоріальних даних є перетворення полів з багатьма значеннями у двійкові «фіктивні» змінні. Це гарантує, що алгоритми не інтерпретують відстань у стовпці за багатьма можливими категоріями. У даному випадку хорошим прикладом цього є значення

стовпця «Національність». Він має 16 можливих значень, і просте перетворення значень у число означало б деяку відстань між ними, що в даному випадку не має сенсу. Як видно, шляхом стандартизації виконано перехід від 17 стовпців (16 змінних і 1 клас) до 64 (63 змінні та 1 клас). Створення фіктивних змінних, як це, безумовно, може збільшити складність задачі, але більшість із них дуже рідко є розрідженими. Це важливо при застосуванні методів машинного навчання. Реалізація функції для встановлення кореляції між ознаками і прогнозованим значенням показано у лістингу 3.9, а результат – на рис. 3.11.

Лістинг 3.9. Визначення кореляції між ознаками та прогнозованим значенням

```
corr = data.corr()
corr.iloc[[5]]
```

	GradeID	raisedhands	VisiTedResources	AnnouncementsView	Discussion	Class	gender_F	gender_M	Nationality
Class	0.071654	0.646298	0.677094	0.52737	0.308183	1.0	0.26349	-0.26349	-0.02631

Рис. 3.11. Матриця кореляцій при прогнозуванні успішності студентів

Як показано на рис. 3.11, кортеж кореляційної матриці містить лише прогнозовану змінну та атрибути, які впливають на її значення. Можна помітити, що «StudentAbsenceDays» має сильну кореляцію, як і інші стовпці.

Важливо відзначити, що значення стовпців, які були перетворені в прості двійкові фіктивні пари, завжди будуть мати «рівну» кореляцію. (наприклад, стать\_F і стать\_M).

До найбільш впливових факторів на успішність учнів відносяться такі поля дата фрейму:

- відвідувані ресурси;
- дні відсутності студентів;
- активність на занятті;
- перегляди оголошень;
- відповіді на опитування;

- стосунки з батьками;
- задоволеність батьків;
- обговорення;
- семестр.

Встановивши залежність між ознаками, наступний крок полягає у безпосередньому прогнозуванні успішності студентів з використанням лінійних алгоритмів та багатошарового перцептрону.

### 3.1.3. Реалізація алгоритмів прогнозування успішності студентів

Перший алгоритм, який було реалізовано базується на простому багатошаровому перцептрону. Рядок коду, який відповідає за його конфігурацію має наступний вигляд:

```
«perc = Perceptron(n_iter=100, eta0=0.1, random_state=15).
```

Процес навчання і тестування результатів щодо прогнозування значень успішності показано у лістингу 3.12.

#### Лістинг 3.12. Навчання і тестування перцептрону

```
results = []
predMiss = []

for _ in range(1000):
    # Randomly sample our training data
    data_train = data.sample(frac=0.7)
    # train data without label
    data_train_X = data_train.loc[:, lambda x: [l for l in
data if l != "Class"]]
    # labels of train data
    data_train_Y = data_train.loc[:, lambda x: "Class"]

    # The rest is test data
    data_test = data.loc[~data.index.isin(data_train.index)]
    # Test data without label
    data_test_X = data_test.loc[:, lambda x: [l for l in data
if l != "Class"]]
    # labels of test data
    data_test_Y = data_test.loc[:, lambda x: "Class"]

    # Train svm
    perc.fit(data_train_X, data_train_Y)
```

```

    predMiss.append((data_test_Y !=
perc.predict(data_test_X)).sum())
    # Score the mean accuracy on the test data and append
    results in a list
    results.append(perc.score(data_test_X, data_test_Y))

# Convert results to an array and look at the minimum and the
average
predErr = np.hstack(predMiss)
Final = np.hstack(results)
print('Minimum Accuracy Score:    %.8f' %
Final[Final.argmin()])
print('Maximum Accuracy Score:    %.8f' %
Final[Final.argmax()])
print('Average Accuracy Score:    %.8f' % np.average(Final))
print('Minimum Prediction Misses:  %d' %
predErr[predErr.argmin()])
print('Maximum Prediction Misses:  %d' %
predErr[predErr.argmax()])
print('Average Prediction Misses:  %.2f' %
np.average(predErr))

```

У результаті прогнозування успішності на основі перцептронну одержано наступні характеристики щодо точності прогнозу (рис. 3.12).

```

Minimum Accuracy Score:    0.47916667
Maximum Accuracy Score:    0.77083333
Average Accuracy Score:    0.64804861
Minimum Prediction Misses:  33
Maximum Prediction Misses:  75
Average Prediction Misses:  50.68

```

Рис. 3.12. Результати точності прогнозування

Як видно з результатів, представлених на рис. 3.12, застосування перцептронну не формує і не забезпечує високої точності прогнозування і стабільності результатів. Найнижчий показник прогнозування на рівні майже 48%, що не дозволяє говорити про ефективність даного алгоритму.

Наступний алгоритм, який потрібно реалізувати – метод опорних векторів. У лістингу 3.13 показано програмний код для конфігурації SVM.

### Лістинг 3.13. Конфігурація SVM

```

# Create the radial basis function kernel version of a Support
Vector Machine classifier
rbf_clf = svm.SVC(C=1.0, cache_size=200, class_weight=None,
coef0=0.0,
                decision_function_shape='ovo', degree=3,
gamma='auto', kernel='rbf',
                max_iter=-1, probability=False,
random_state=None, shrinking=True,
                tol=0.001, verbose=False)
# Create the linear kernel version of a Support Vector Machine
classifier
lin_clf = svm.SVC(C=1.0, cache_size=200, class_weight=None,
coef0=0.0,
                decision_function_shape='ovo', degree=3,
gamma='auto', kernel='linear',
                max_iter=-1, probability=False,
random_state=None, shrinking=True,
                tol=0.001, verbose=False)
# Create the polynomial kernel version of a Support Vector
Machine classifier
poly_clf = svm.SVC(C=1.0, cache_size=200, class_weight=None,
coef0=0.0,
                decision_function_shape='ovo', degree=3,
gamma='auto', kernel='poly',
                max_iter=-1, probability=False,
random_state=None, shrinking=True,
                tol=0.001, verbose=False)
# Create the sigmoid kernel version of a Support Vector
Machine classifier
sig_clf = svm.SVC(C=1.0, cache_size=200, class_weight=None,
coef0=0.0,
                decision_function_shape='ovo', degree=3,
gamma='auto', kernel='sigmoid',
                max_iter=-1, probability=False,
random_state=None, shrinking=True,
                tol=0.001, verbose=False)

```

Варто відмітити, що метод опорних векторів навчали за допомогою різних ядерних функцій та лінійних функцій:

- SVM на основі радіальної базисна функція;
- поліноміальний SVM;
- лінійний SVM;
- SVM на основі сигмоїдної функції;

Результат прогнозування успішності представлено на рис. 3.13.

```

RBF Minimum Accuracy Score: 0.63194444
RBF Maximum Accuracy Score: 0.84027778
RBF Average Accuracy Score: 0.74390972
-----
Linear Minimum Accuracy Score: 0.66666667
Linear Maximum Accuracy Score: 0.84722222
Linear Average Accuracy Score: 0.75675000
-----
Polynomial Minimum Accuracy Score: 0.34722222
Polynomial Maximum Accuracy Score: 0.54861111
Polynomial Average Accuracy Score: 0.43920833
-----
Sigmoid Minimum Accuracy Score: 0.58333333
Sigmoid Maximum Accuracy Score: 0.84027778
Sigmoid Average Accuracy Score: 0.72760417
=====
RBF Average Prediction Misses: 36.88
Linear Average Prediction Misses: 35.03
Polynomial Average Prediction Misses: 80.75
Sigmoid Average Prediction Misses: 39.23

```

Рис. 3.13. Результат прогнозування успішності із застосуванням SVM

Ще один алгоритм, який застосовується при прогнозуванні успішності – дерева прийняття рішень. Конфігурація дерев наведена у лістингу 3.14.

Лістинг 3.14. Конфігурація та реалізація дерев прийняття рішень

```

tree3 = DecisionTreeClassifier(random_state=56,
criterion='gini', max_depth=3)
tree5 = DecisionTreeClassifier(random_state=56,
criterion='gini', max_depth=5)
results_3 = []
results_5 = []
predMiss_3 = []
predMiss_5 = []

for _ in range(1000):
    # Randomly sample our training data
    data_train = data.sample(frac=0.7)
    # train data without label
    data_train_X = data_train.loc[:, lambda x: [1 for 1 in
data if 1 != "Class"]]
    # labels of train data

```

```

data_train_Y = data_train.loc[:, lambda x: "Class"]

# The rest is test data
data_test = data.loc[~data.index.isin(data_train.index)]
# Test data without label
data_test_X = data_test.loc[:, lambda x: [l for l in data
if l != "Class"]]
# labels of test data
data_test_Y = data_test.loc[:, lambda x: "Class"]

# Train svm
tree3.fit(data_train_X, data_train_Y)
tree5.fit(data_train_X, data_train_Y)
#Sum the prediction misses. Since this is a smaller
dataset,
predMiss_3.append((data_test_Y !=
tree3.predict(data_test_X)).sum())
predMiss_5.append((data_test_Y !=
tree5.predict(data_test_X)).sum())
# Score the mean accuracy on the test data and append
results in a list
results_3.append(tree3.score(data_test_X, data_test_Y))
results_5.append(tree5.score(data_test_X, data_test_Y))

```

Одержані результати прогнозування на основі підходу дерев прийняття рішень показано на рис.3.14.

```

3-depth Tree Minimum Accuracy Score: 0.54166667
3-depth Tree Maximum Accuracy Score: 0.79166667
3-depth Tree Average Accuracy Score: 0.68047917
-----
5-depth Tree Minimum Accuracy Score: 0.56250000
5-depth Tree Maximum Accuracy Score: 0.83333333
5-depth Tree Average Accuracy Score: 0.71739583

```

Рис. 3.14. Результати прогнозування на основі дерев прийняття рішень

Зведені, усереднені результати прогнозування при використанні розглянутих алгоритмів представлено у табл. 3.2.



Таблиця 3.2

**Усереднені результати щодо точності прогнозування успішності студентів**

№	Алгоритм	Середнє значення точності прогнозування
1.	Багатошаровий перцептрон	64,74
2.	Метод опорних векторів (поліноміальний)	73,27
3.	Метод опорних векторів (лінійний)	75,62
4.	Метод опорних векторів з RBF	74,33
5.	Метод опорних векторів з сигмоїдною функцією	72,77
6.	Випадковий ліс (глибина 3)	68,08
7.	Випадковий ліс (глибина 5)	71,70

Як видно з узагальнених результатів, кращим при прогнозуванні успішності є метод опорних векторів з функцією лінійної залежності.

Перейдемо до результатів прогнозування на основі глибоких нейронних мереж, які реалізуються архітектурою GritNet.

3.2. Результати прогнозування успішності на основі глибоких нейронних мереж та архітектури GritNet

В якості вхідних наборів даних при застосуванні глибоких нейронних мереж, зокрема GritNet, обрано набори студентських даних за двома програми: ND-A і ND-B. Обґрунтування вибору цих двох програм полягає в тому, що вони відрізняються одна від одної за багатьма критеріями.

Наприклад, навчальна програма ND-A передбачає нижчі очікування попередніх технічних знань і відносно вищий показник випуску, ніж ND-B.

В обох програмах закінчення визначається як завершення кожного з необхідних проектів за програмою ND з прохідною оцінкою. Коли користувач

офіційно закінчує навчання, його запис про зарахування анутується міткою часу, тому можна використовувати наявність цієї мітки часу як цільової мітки.

Користувачі повинні закінчити навчання до 2017-09-30, щоб вважатися, що вони успішно завершили навчання. Навчальна програма кожної програми ND містить суміш відеоконтенту, письмового вмісту, вікторин і проектів.

Потрібно зауважити, що для отримання диплому не обов'язково взаємодіяти з кожною частиною вмісту чи виконувати кожну вікторину. Нижче наведено детальні характеристики кожного набору даних, використаного для цього дослідження.

Набір даних ND-A містить інформацію про студентів, які зареєструвалися за програмою ND-A (з 2017-03-07 по 2017-09-30). Обрано 1853 студентів, які активно співпрацювали зі своїм наставником під час навчання, вважаючи, що це студенти, які демонструють повну взаємодію з навчальною програмою в цілому. Цей набір із 1853 студентів включає 777 студентів, які закінчили навчання, що становить 41,9%.

Довжина потоків подій кожного студента коливається від 0 до 4175 подій, у середньому 536 подій. Навчальний план програми ND-A містить 9 проектів, 1025 унікальних сторінок вмісту для відвідування та 77 вікторин.

Набір даних ND-B складається з даних про 8301 студентів, які активно комунікують зі своїм наставником у класі і є зарахованими на програму ND-B (з 20.06.2016 по 30.09.2017).

Цей набір із 8301 студента включає 1005 студентів, які закінчили навчання, що становить 12,1%. Тривалість потоків подій кожного студента коливається від 1 події до 4554 подій, у середньому 242 події.

Навчальний план програми ND-B складається з 6 проектів, 668 унікальних сторінок із вмістом та 347 вікторин. Для обох наборів даних подія означає, що користувач виконує певну дію (наприклад, переглядає відео, читає текстову сторінку, намагається пройти вікторину або отримує оцінку за проект) у певний час.

Деякі невідповідні дані відфільтровуються під час попередньої обробки, наприклад, події, які відбуваються до офіційної реєстрації користувача в результаті безкоштовного пробного періоду.

Слід зазначити, що в ці дані не включено жодних особистих відомостей і визначається рівність студентів через непрозорі унікальні ідентифікатори. Аналізуючи дані за кількістю студентів в наборі ND-A та ND-B, можна побудувати графік залежності, який показано на рис. 3.15.

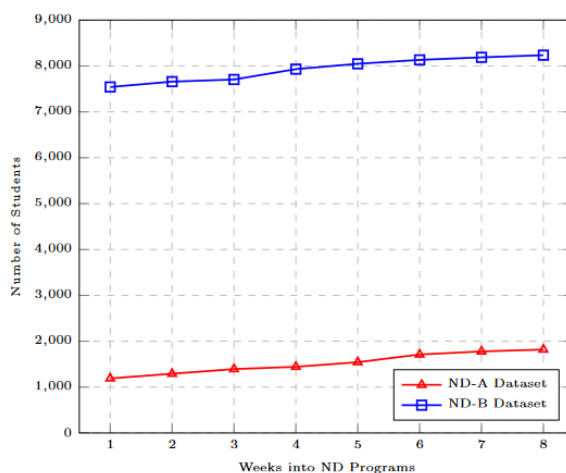


Рис. 3.15. Кількість студентів, які навчаються за програмами ND-A та ND-B

Залежність, що відображає кількість навчальних матеріалів за обома програмами, які розглядаються показано на рис. 3.16.

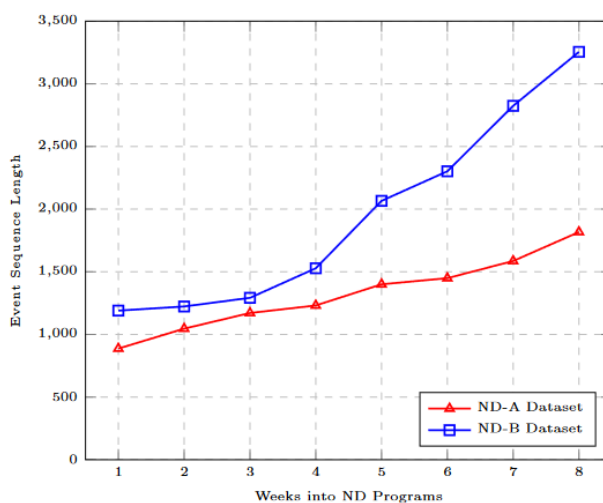


Рис. 3.16. Кількість навчальних матеріалів за програмами ND-A та ND-B

Встановлено, що моделі GritNet досить легко піддаються навчанню. Метою навчання є негативна логарифмічна функція правдоподібності щодо спостережуваної послідовності подій діяльності студента за визначеною моделлю.

Втрати бінарної перехресної ентропії мінімізуються за допомогою стохастичного градієнтного спуску на міні-партіях. У даному експерименті використовується BLSTM з прямими і зворотними шарами LSTM, що містять 128 комірок у кожному напрямку.

Розмір вбудованого шару визначають шляхом пошуку найкращих параметрів на основі набору даних: від 1024 до 3584 для набору ND-A і від 1024 до 5120 для набору ND-B. Рівень показника викидів становить від 10 до 20%, що застосовувався до виходу BLSTM і добре працює для обох наборів даних, щоб запобігти перенавчанню під час тренування моделі з розміром міні-партії 32.

Як для базової моделі, так і для моделей GritNet проведено навчання різних моделей протягом різних тижнів на основі записів студентів за тижнями, щоб спрогнозувати здатність студента успішно закінчити навчання.

Щоб продемонструвати переваги GritNet, основна увага зосереджена на прогнозуванні випуску студентів. Оскільки справжня бінарна цільова мітка (1: «завершити навчання», 0: «не завершити») є незбалансованою (тобто кількість 0 переважає кількість 1), точність не є показником ефективності моделі.

Замість цього запропоновано використати метрику ROC для оцінки якості прогнозів GritNet. Крива ROC була створена шляхом побудови графіка справжньої позитивної частоти (TPR) проти хибнопозитивної частоти (FPR).

У даному випадку TPR – це відсоток студентів, які закінчили навчання, який GritNet маркує позитивним, а FPR – це відсоток студентів, які не закінчили навчання, які GritNet неправильно позначає позитивними.

Точність прогнозування вимірювалася площею під кривою ROC (AUC), яка може приймати значення від 0 до 100% (чим вище, тим краще) – із випадковим припущенням, що завжди давало 50%. Окрім цього, використано 5-кратну перехресну перевірку на рівні студентів, переконавшись, що кожна частина містить приблизно однакові пропорції двох груп (випускників і невипускників) студентів.

Для справедливого порівняння базова продуктивність була оптимізована шляхом зміни значень  $\alpha$  щотижня. GritNet також вимагає невеликої оптимізації гіперпараметрів для оптимальної точності протягом кожного тижня навчання.

Показано, що GritNet дійсно покращує точність прогнозування успішності завершення навчання студентом протягом кількох тижнів. Результати прогнозування щодо успішності завершення навчання показано на рис. 3.17.

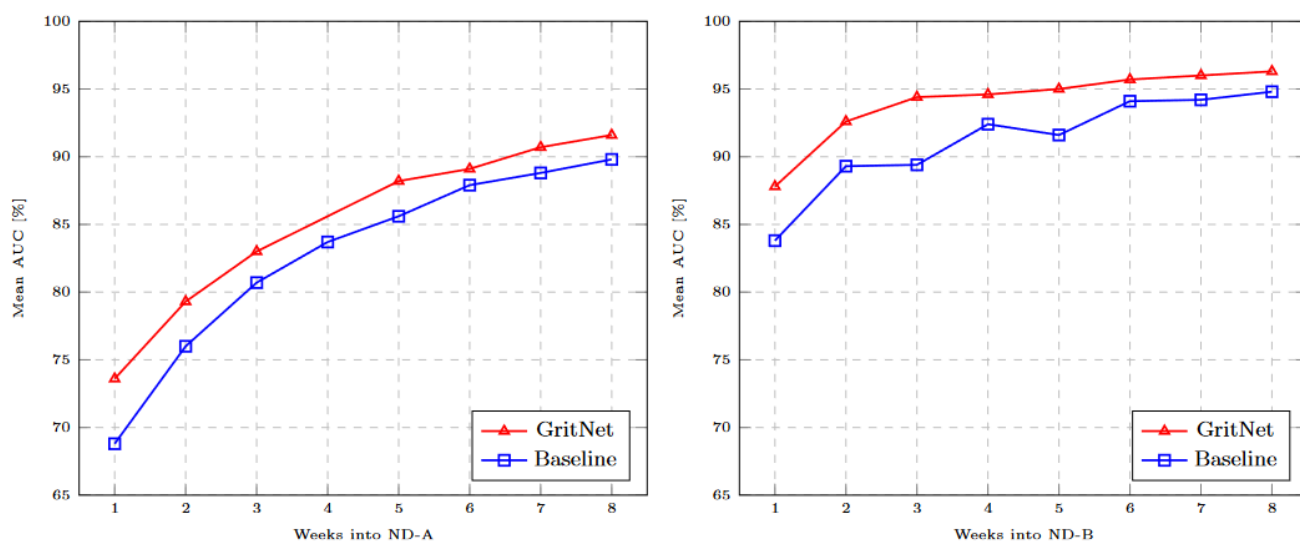


Рис. 3.17. Показники точності прогнозування успішного завершення навчання студентами за двома наборами даних

З результатів прогнозування для обох наборів даних Udacity чітко видно, що продуктивність у базовій моделі та моделях GritNet подібна після восьми тижневого отримання даних про даного студента.

Однак GritNet може досягти значних покращень якості прогнозування протягом перших кількох тижнів навчання студентів. Зокрема, GritNet зміг досягти кращої продуктивності більш ніж на 5,0% як на наборі даних ND-A (на 1-му тижні), так і на наборі даних ND-B (на 3-му тижні). Більше того, у наборі даних ND-B базовій моделі потрібно було почекати два місяці, щоб досягти точності прогнозування, яку показав GritNet, протягом трьох тижнів. Отже, важливою перевагою GritNet є створення швидко адаптованих моделей та одержання точного показника для оцінки довгострокових результатів студентів.

### 3.3. Висновки до розділу

1. Програмно реалізовано метод прогнозування успішності студентів з використанням інтелектуалізованих методів прогнозування, зокрема, підходу випадкового лісу, методу опорних векторів та багатошарового перцептронну, що дало змогу обґрунтувати кращий з них – метод опорних векторів з лінійною функцією, що забезпечила точність прогнозування на рівні 75% на заданому вхідному наборі даних.

2. Реалізовано програмну імплементацію методу прогнозування успішності студентів на основі підходу глибоких нейронних мереж, зокрема з модифікованою архітектурою GritNet, що дало змогу підвищити ефективність результатів прогнозування за значно коротший проміжок часу та з вищою точністю (до 98%).

## РОЗДІЛ 4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

### 4.1. Охорона праці

Під час проведення дослідження методів і засобів інтелектуалізованого аналізу та прогнозування успішності студентів у системах електронного навчання усі процеси повинні виконуватись із врахуванням вимог техніки безпеки на робочому місці, пожежної безпеки, відповідно з діючими нормативно-правовими актами та встановленими нормами. Дослідження передбачають пошук та аналіз інформації, розробку методу інтелектуального аналізу даних, реалізацію взаємодії між компонентами системи прогнозування успішності студентів, а також розробку моделі опису факторів впливу на успішність, що передбачає використання ЕОМ.

Перед початком усіх робіт проведено інструктаж з техніки безпеки на робочому місці відповідно до вимог Типового положення про інструктажі, спеціальне навчання та перевірку знань з питань пожежної безпеки на підприємствах, в установах та організаціях України ДБН В.1.1-7-2016 „Пожежна безпека об’єктів будівництва та НПАОП 0.00-4.36-05.

Розробка програмного забезпечення та проведення досліджень відбувались відповідно до Закону України "Про охорону праці" та "Правил охорони праці при експлуатації електронно-обчислювальних машин" та Державних санітарних правил і норм роботи з візуальними дисплейними терміналами електронно-обчислювальних машин, а також НПАОП 0.00-7.15-18 «Вимоги щодо безпеки та захисту здоров’я працівників під час роботи з екранними пристроями» [24].

Робоче місце відповідає ергономічним, організаційним вимогам, вказаних у вищеперерахованих документах.

Електробезпека на робочому місці, повинна відповідати вимогам Правил безпечної експлуатації електроустановок споживачів, затверджених наказом Держнаглядохоронпраці від 09.01.98 N 4 , зареєстрованих у Міністерстві юстиції України 10.02.98 за N 93/2533 (НПАОП 40.1-1.21-98) [20].

Для забезпечення електробезпеки при роботі з ЕОМ під час проведення досліджень електромережа зроблена відповідно до правил:

- лінія електромережі живлення виконана як окрема групова трипровідна мережа, шляхом прокладання фазового, нульового робочого та нульового захисного провідників;
- нульовий захисний провідник використовується для заземлення(занулення) електроприймачів;
- електричне устаткування, електропроводи та кабелі мають апаратуру захисту від струму короткого замикання та інших аварійних режимів;
- під час монтажу та експлуатації електромережі необхідно повністю унеможливити виникнення електричного джерела загоряння внаслідок короткого замикання та перевантаження проводів, обмежено застосування проводів з легкозаймистою ізоляцією та застосовано негорючу ізоляцію;
- електричне устаткування підключено до мережі лише за допомогою справних штепсельних з'єднань і розеток заводського виготовлення;
- у штепсельних з'єднаннях та електророзетках є спеціальні контакти для підключення нульового захисного провідника [20].

Електробезпека під час розробки прототипу для тестування методу взаємодії забезпечувалась завдяки використанню безпроводних технологій та напруги живлення в діапазоні 3.3 – 9 В, що зменшує шанс ураження струмом та його вплив при виникненні контакту з мережею чи аварійних ситуацій.

Пожежна безпека будівель та приміщень, де розміщені робочі місця, обладнані ЕОМ з ВДТ і ПП відповідає вимогам, встановленим:

- державними будівельними нормами ДСТУ Б.В.1.1-36:2016 „Визначення категорій приміщень, будинків та зовнішніх установок за вибухопожежною та пожежною небезпекою”;
- ДБН В.1.1-7-2016 „Пожежна безпека об'єктів будівництва”;



Будівлі та приміщення, де розміщені робочі місця операторів, також повинні відповідати вимогам чинних санітарних норм, санітарних норм і правил, вказаних у ДСанПіН 3.3.2-007-98[24].

На робочих місцях з ВДТ забезпечені оптимальні значення параметрів мікроклімату: температури, відносної вологості й рухливості повітря, відповідно до вимог ДСН 3.3.6.042 – 99 „Санітарні норми мікроклімату виробничих приміщень”.

Приміщення для роботи з ВДТ повинні мати природне та штучне освітлення відповідно до ДБН В.2.5-28-2006. Природне освітлення забезпечується через світлові прорізи, орієнтовані переважно на північ чи північний схід і коефіцієнт природною освітленості (КПО) не нижче ніж 1,5%. Розраховується КПО за методикою, викладеною в ДБН В.2.5-28-2006.

Штучне освітлення в приміщенні з робочими місцями, обладнаними ВДТ ЕОМ та ПЕОМ, забезпечувалось системою загального рівномірного освітлення. В якості джерела світла для штучного освітлення використані люмінесцентні лампи типу ЛБ.

Також під час дослідження методів та засобів інтелектуалізованого аналізу та прогнозування успішності студентів в системах електронного навчання не було порушено наступних вимог:

- самостійного ремонту ЕОМ, ВДТ чи ПП;
- внесення змін до конструкції ЕОМ;
- не торкатись задньої панелі системного блоку при включеному живленні;
- вживання напоїв та їжі на робочому місці;
- нагромадження біля ЕОМ, ВДТ, ПП предмети, які не використовуються

для поточної роботи.

Дослідження методів за засобів інтелектуалізованого аналізу та прогнозування успішності студентів в системах електронного навчання проведено з дотриманням вимог техніки безпеки та охорони праці, електробезпеки, пожежної безпеки та санітарних норм при роботі з ЕОМ та ВДТ.

#### 4.2. Планування та порядок проведення евакуації населення з районів наслідків впливу НС техногенного та природного характеру

В умовах неповного забезпечення захисними спорудами в містах та інших населених пунктах, що мають об'єкти підвищеної небезпеки, основним засобом захисту населення є евакуація і розміщення його у зонах, які є безпечними для проживання людей.

Евакуації підлягає населення, яке проживає в населених пунктах, що знаходяться у зонах можливого катастрофічного затоплення, можливого небезпечного радіоактивного забруднення, хімічного ураження, в районах виникнення стихійного лиха, аварій і катастроф (якщо виникає безпосередня загроза життю та здоров'ю людей). Залежно від обставин, які склалися на час надзвичайної ситуації, може бути проведено загальну або часткову евакуацію населення тимчасового або безповоротного характеру. Загальна евакуація проводиться за рішенням Кабінету Міністрів України для всіх категорій населення і планується на випадок:

- можливого небезпечного радіоактивного забруднення територій навколо атомних електростанцій (якщо виникає безпосередня загроза життю та здоров'ю людей, які проживають в зоні ураження);
- виникнення загрози катастрофічного затоплення місцевості з чотиригодинним добіганням проривної хвилі.

Часткова евакуація проводиться за рішенням Кабінету Міністрів України у разі загрози або виникнення надзвичайної ситуації техногенного та природного характеру.

Під час проведення часткової евакуації завчасно вивозиться не зайняте у сферах виробництва та обслуговування населення: діти, учні навчальних закладів, вихованці дитячих будинків, разом з викладачами та вихователями, студенти, пенсіонери та інваліди, які утримуються у будинках для осіб похилого віку, разом з обслуговуючим персоналом і членами їх сімей.

У сфері захисту населення і територій від надзвичайних ситуацій техногенного та природного характеру евакуація населення планується на випадок:

- аварії на атомній електростанції з можливим забрудненням території; усіх видів аварій з викидом сильнодіючих отруйних речовин; загрози катастрофічного забруднення місцевості :

- лісових і торф'яних пожеж, землетрусів, зсувів, інших геофізичних і гідрометеорологічних явищ з тяжкими наслідкам, що загрожують населеним пунктам.

Загальна евакуація проводиться шляхом вивезення основної частини населення з міст і небезпечних районів усіма видами наявних транспортних засобів на відповідній адміністративній території та виведення найбільш витривалої його частини пішки. Часткова евакуація проводиться з використанням транспортних засобів, що експлуатуються за діючим графіком. На органи виконавчої влади, органи місцевого самоврядування та керівників об'єктів, які проводять евакуацію населення, покладається:

- планування і проведення евакуації працівників та членів їх сімей;
- подання до відповідних транспортних органів розрахунків потреби у транспортних засобах для вивезення працівників і членів їх сімей до безпечних районів;

- контроль за плануванням, підготовкою і проведенням евакуаційних заходів підвідомчими об'єктами;

- визначення та підготовка безпечного району для розміщення евакуйованих працівників і членів їх сімей.

Інші заходи та порядок проведення евакуації викладено у постанові Кабінету Міністрів від 26 жовтня 2001р. № 1432 про затвердження Положення про порядок проведення евакуації населення у разі загрози або виникнення надзвичайних ситуацій техногенного та природного характеру.

У плані евакуації, складовою частиною якого є карта (схема), зазначаються:

- висновки з оцінки обстановки у разі виникнення надзвичайної ситуації;

- порядок оповіщення населення про початок евакуації;
- кількість населення, яке підлягає евакуації, за віковими категоріями;
- терміни проведення евакуації;
- склад евакуаційних органів і терміни приведення їх у готовність;
- кількість населення, яке вивозиться різними видами транспортних засобів окремо і виводиться пішки;
- розподілення об'єктів за збірними евакуаційними пунктами, пунктами посадки, районами (пунктами) розміщення та евакуаційними напрямками; маршрути евакуації;
- райони (пункти) розміщення евакуйованого населення; пункти посадки на транспортні засоби, пункти висадки у безпечному районі, порядок доставки населення з пунктів висадки до районів (пунктів) розміщення;
- заходи щодо організації приймання, розміщення, захисту та життєзабезпечення евакуйованого населення у безпечному районі;
- порядок організації управління і зв'язку.

Розділ плану, в якому визначаються види забезпечення евакуації, розробляється відповідними службами. До цього розділу включаються:

- основні завдання служби;
- перелік сил і засобів, які залучаються для виконання евакуаційних заходів;
- терміни виконання завдань.

Евакуаційна комісія відповідного органу виконавчої влади, на території якої планується розміщення евакуйованого населення, розробляє план його приймання і розміщення у безпечному районі з картою (схемою).

У плані зазначаються:

- кількість евакуйованого населення за віковими категоріями, яке прибуває у район, місто, район у місті, селище, село;
- кількість об'єктів і їх розподіл за районами у місті, сільськими і селищними радами, населеними пунктами;
- чисельність населення, яке проживає на відповідній території;

- будівлі і споруди для розміщення об'єктів господарювання;
- пункти висадки евакуйованого населення;
- порядок і терміни доставки евакуйованого населення з приймальних евакуаційних пунктів до районів (пунктів) розміщення;
- порядок розміщення евакуйованого населення;
- порядок забезпечення евакуйованого населення продуктами харчування, водою, предметами першої необхідності, медичним та іншими видами обслуговування;
- порядок оповіщення посадових осіб, які відповідають за приймання евакуйованого населення, про початок евакуації і терміни прибуття населення.

План приймання і розміщення евакуйованого населення включає також розділ з транспортного забезпечення евакуації, в якому зазначається:

- кількість транспортних засобів кожного виду і термін їх подачі до пунктів посадки;
- кількість населення, яке підлягає евакуації;
- терміни відправлення евакуйованого населення у безпечні райони;
- терміни прибуття евакуйованого населення до пунктів посадки;
- маршрути руху транспортних засобів;
- кількість рейсів.

На всіх громадян, які підлягають евакуації, завчасно складаються списки за об'єктами і житлово-експлуатаційними організаціями у трьох примірниках, один з яких залишається на об'єкті або в житлово-експлуатаційній організації, другий (у разі одержання рішення про проведення евакуації) після уточнення списків надсилається на збірний евакуаційний пункт, третій - до евакуаційної комісії району (пункту) розміщення.

З отриманням рішення (сигналу) про проведення евакуації евакуаційні комісії уточнюють завдання керівникам об'єктів щодо проведення евакуаційних заходів, контролюють стан оповіщення населення, його збору, формування колон (через начальників маршрутів), забезпечують переміщення їх до пунктів евакуації, а також разом з транспортними службами - готовність транспортних

засобів до перевезень, уточнюють порядок їх використання, підтримують постійний зв'язок з начальниками маршрутів та з органами виконавчої влади безпечних районів, інформують їх про хід евакуації.

У райони розміщення евакуаційних органів та населення, яке підлягає евакуації, направляються представники евакуаційних комісій для вирішення питань приймання, розміщення і життєзабезпечення евакуйованого населення.

Керівник органу виконавчої влади і евакуаційна комісія безпечного району, організують підготовку пунктів висадки, розгортають приймальний евакуаційний пункт, уточнюють кількість прибулих і порядок подачі транспортних засобів для їх вивезення з пунктів висадки, а також з проміжних пунктів евакуації до пунктів розміщення, контролюють роботу керівників об'єктів безпечних районів з прийому і розміщення евакуйованого населення.

У разі оголошення евакуації громадяни самостійно на міських транспортних засобах, які у цей період працюють цілодобово, прибувають на збірні евакуаційні пункти. Працівники цих пунктів розподіляють громадян, які підлягають евакуації, за транспортними засобами, інструктують їх і забезпечують посадку на транспортні засоби.

Евакуйовані громадяни повинні мати при собі паспорт, військовий квиток, документ про освіту, трудову книжку або пенсійне посвідчення, свідоцтво про народження, гроші і цінності, продукти харчування і воду на 3 доби, постільну білизну, необхідний одяг і взуття загальною вагою не більш як 50 кілограмів на кожного члена сім'ї. Дітям дошкільного віку вкладається у кишеню або пришивається до одягу записка, де зазначається прізвище, ім'я та по батькові, домашня адреса, а також ім'я та по батькові матері і батька.

## ВИСНОВКИ

Основні наукові та практичні результати полягають в наступному.

1. Проведено аналіз особливостей організації навчального процесу в Україні та закордоном, на основі результатів якого встановлено, що набуває широкої популярності, поряд з класичними, такий вид здобуття освіти як онлайн навчання, елементи якого важливо імплементувати у традиційні форми навчального процесу.

2. Визначено та проаналізовано характеристики систем дистанційного навчання та онлайн навчання, що дало змогу встановити фактори і програмні модулі, які необхідні для забезпечення успішності студентів і навчального процесу в цілому.

3. Проаналізовано особливості застосування сучасних платформ для провадження навчального процесу, що дозволило визначити їхні переваги і недоліки, а також встановити необхідність наявності компонентів і даних для прогнозування успішності студентів із застосування інтелектуалізованих методів і засобів формування прогнозів.

4. На основі аналізу факторів впливу на успішність студентів під час електронного навчання запропоновано їх формалізацію у вигляді елементів множини, що дало змогу побудувати модель, яка враховує ефективність програмно-апаратного забезпечення, якість навчального контенту та кваліфікацію інструкторів курсів, які в подальшому будуть враховані при виборі інтелектуалізованих методів і засобів прогнозування успішності студентів.

5. Для прогнозування успішності студентів обґрунтовано та формально представлено алгоритми, які використовують підходи дерев прийняття рішень та випадкових лісів, методу опорних векторів, що в перспективі дасть змогу прогнозувати успішність студентів за трьома класами: «висока успішність», «достатня успішність», «відрахування студента».

6. Запропоновано метод прогнозування успішності студентів на основі підходу глибоких нейронних мереж з архітектурою GritNet в основі якого лежать вбудовані шари двонаправленого LSTM, що дає змогу більш ефективно, у порівнянні з базовою моделлю GritNet, забезпечити високу точність прогнозування фактів завершення або не завершення навчання студентів на курсі.

7. Програмно реалізовано метод прогнозування успішності студентів з використанням інтелектуалізованих методів прогнозування, зокрема, підходу випадкового лісу, методу опорних векторів та багат шарового перцептрон, що дало змогу обґрунтувати кращий з них – метод опорних векторів з лінійною функцією, що забезпечила точність прогнозування на рівні 75% на заданому вхідному наборі даних.

8. Реалізовано програмну імплементацію методу прогнозування успішності студентів на основі підходу глибоких нейронних мереж, зокрема з модифікованою архітектурою GritNet, що дало змогу підвищити ефективність результатів прогнозування за значно коротший проміжок часу та з вищою точністю (до 98%).



## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Y. Bengio, R. Ducharme, and P. Vincent. A neural probabilistic language model. In *Advances in Neural Information Processing Systems 13 (NIPS 2000)*. 2001. p. 932–938.
2. Graves and J. Schmidhuber. Framewise phoneme classification with bidirectional LSTM networks. In *2005 International Joint Conference on Neural Networks (ICJNN'05)*. 2005. p. 23–43.
3. J. D. Keeler, D. E. Rumelhart, and W. K. Leow. Integrated segmentation and recognition of hand-printed numerals. In *Advances in Neural Information Processing Systems 3 (NIPS 1990)*. 1991. p. 557–563.
4. F. Mi and D.-Y. Yeung. Temporal models for predicting student dropout in massive open online courses. In *Proceedings of 15th IEEE International Conference on Data Mining Workshop (ICDMW 2015)/ Atlantic City, New Jersey*. 2015. p. 256–263.
5. Y. Ng. Feature selection, L1 vs. L2 regularization, and rotational invariance. In *Proceedings of the 21st International Conference on Machine Learning (ICML 2004)*. Banff, Alberta, Canada. 2004. p. 78–85.
6. C. Rudin, B. Letham, A. Salieb-Aouissi, E. Kogan, and D. Madigan. Sequential event prediction with association rules. In *24th Annual Conference on Learning Theory (COLT 2011)*. 2011. p. 615–634.
7. Жаровський Р.О., Дармопук Д.В. Характеристики систем електронного навчання. Матеріали X міжнародної науково - технічної конференції молодих учених і студентів «Актуальні задачі сучасних технологій» (24-25 листопада 2021 р.) Тернопільського національного технічного університету імені Івана Пулюя. Тернопіль: ТНТУ. 2021. С. 91.
8. Жаровський Р.О., Дармопук Д.В. Аналіз успішності студентів на основі технології GritNet. Матеріали IX науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (8-9 грудня 2021 року). Тернопіль: ТНТУ. 2021. С. 108.

9. Python-recsys on Github. URL: <https://github.com/ocelma/python-recsys> (дата звернення 22.10.2021 р).
10. Preprocessing data. URL: <https://scikit-learn.org/stable/modules/preprocessing.html#preprocessing> (дата звернення 02.11.2021 р.).
11. API reference. URL: <https://pandas.pydata.org/docs/reference/index.html> (дата звернення 10.11.2021 р.).
12. NumPy Reference. URL: <https://numpy.org/doc/stable/reference/index.html> (дата звернення 12.11.2021 р.)
13. Барсегян А. Анализ данных и процессов. 3 изд. БХВ-Петербург. 2009. 512 с.
14. Breese J., Heckerman D., Kadie C. Empirical Analysis of Predictive Algorithms for Collaborative Filtering. Proc. 14th Conf. Uncertainty in Artificial Intelligence. 1998. pp. 223-234.
15. Install TensorFlow. URL: <https://www.tensorflow.org/install> (дата звернення 17.11.2021 р.).
16. TensorFlow Datasets: a collection of ready-to-use datasets. URL: <https://www.tensorflow.org/datasets> (дата звернення 17.11.2021 р.)
17. Cuda. URL: <https://opencv.org/platforms/cuda/> (дата звернення 21.04.2021 р.)
18. Микитишин А.Г., Митник М.М., Стухляк П.Д., Пасічник В.В. Комп'ютерні мережі. Книга 1. Львів, «Магнолія 2006». 2013. 256 с.
19. Микитишин А.Г., Митник М.М., Стухляк П.Д., Пасічник В.В. Комп'ютерні мережі. Книга 2. Львів, «Магнолія 2006», 2014. 312 с.
20. Beginner's Guide to Object Detection Algorithms. URL: <https://medium.com/analytics-vidhya/beginners-guide-to-object-detectionalgorithms-6620fb31c375> (дата звернення 01.12.2021).
21. NumPy v1.20 Manual. URL: <https://numpy.org/doc/stable/> (дата звернення 25.11.2021р.).
22. Python Tutorial. URL: <https://www.w3schools.com/python/default.asp> (дата звернення 15.10.2021 р.).

23. Pandas documentation. URL: <https://pandas.pydata.org/docs/index.html>  
(дата звернення 28.11.2021 р.).

24. НПАОП 0.00-7.15-18 «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями». Київ. 2018.

25. Катренко Л.А., Катренко А.В. Охорона праці в галузі комп'ютерингу. Львів: Магнолія-2006. 2012. 544 с.

26. Желібо Е.Н. Безпека життєдіяльності: Навчальний посібник. Київ: «Каравела», Львів: «Новий світ - 2000». 2001. 320с.

Додаток А  
Тези конференцій

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
Тернопільський національний технічний університет імені Івана Пулюя (Україна)  
Університет імені П'єра і Марії Кюрі (Франція)  
Маріборський університет (Словенія)  
Технічний університет у Кошице (Словаччина)  
Вільнюський технічний університет ім. Гедимінаса (Литва)  
Білоруський національний технічний університет (Республіка Білорусь)  
Міжнародний університет цивільної авіації (Марокко)  
Наукове товариство ім. Т.Шевченка

# АКТУАЛЬНІ ЗАДАЧІ СУЧАСНИХ ТЕХНОЛОГІЙ

**Збірник**  
тез доповідей  
**Том I**

**X Міжнародної науково-практичної  
конференції молодих учених та студентів**  
24-25 листопада 2021 року



**УКРАЇНА**  
**ТЕРНОПІЛЬ – 2021**

Матеріали X Міжнародної науково-практичної конференції молодих учених та студентів  
«АКТУАЛЬНІ ЗАДАЧІ СУЧАСНИХ ТЕХНОЛОГІЙ» – Тернопіль 24-25 листопада 2021 року

- |     |   |    |
|-----|---|----|
| 9.  | <b>Ю.І. Пиндус, В.П.Калушка, Р.Р. Заверуха, О.Ю. Пиндус, Ю.І. Пипко</b>           | 77 |
|     | ДОСЛІДЖЕННЯ ТЕПЛОВОГО БАЛАНСУ РОБОТИ ДВИГУНА НА<br>ДИЗЕЛЬНОМУ ПАЛИВІ ТА БІОПАЛИВІ |    |
| 10. | <b>Р.М. Рогатинський, В.Л. Дмитроца, М.В. Грубенюк, Р.П. Цапик</b>                | 79 |
|     | ТРАНСПОРТУВАННЯ НАСИПНОГО ПАЛИВА ГВИНТОВИМИ<br>КОНВЕСРАМИ                         |    |
| 11. | <b>Р.М. Рогатинський, Р. В. Хорошун, А.Д. Бобков, Р.Б. Шимків</b>                 | 81 |
|     | МОДЕЛЮВАННЯ РУХУ АВТОМОБІЛЯ ПО КРИВОЛІНІЙНІЙ ТРАСІ                                |    |
| 12. | <b>В.В. Ткачук, Ю.С. Шуберт</b>   | 83 |
|     | ІМІТАЦІЙНЕ МОДЕЛЮВАННЯ РОБОТИ ЛОГІСТИЧНОГО СКЛАДУ                                 |    |

#### **СЕКЦІЯ: КОМП'ЮТЕРНО-ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА СИСТЕМИ ЗВ'ЯЗКУ**

- |     |  |     |
|-----|--|-----|
| 1.  | <b>О.В. Балакунець, Є.В. Тиш</b>   | 84  |
|     | МЕТОДИ ТА ПРОГРАМНО-АПАРАТНЕ ЗАБЕЗПЕЧЕННЯ СИСТЕМИ<br>РЕЗЕРВНОГО ЖИВЛЕННЯ В КОМП'ЮТЕРНИХ СИСТЕМАХ   |     |
| 2.  | <b>О.М. Барановський, А.В. Жилін, Г.С. Голич</b>   | 85  |
|     | ЗАСТОСУВАННЯ МЕТОДУ МАШИННОГО НАВЧАННЯ ДЛЯ<br>ВИРІШЕННЯ ЗАДАЧІ ДЕТЕКТУВАННЯ ТОЧКОВИХ АНОМАЛІЙ У<br>МЕРЕЖЕВОМУ ТРАФІКУ ЗАСОБАМИ SIEM SPLUNK |     |
| 3.  | <b>В.П. Волоський, Ю.З. Лещинин, Н.Р. Романишин</b>  | 87  |
|     | КОМП'ЮТЕРНА СИСТЕМА КОНТРОЛЮ ТА БАЛАНСУВАННЯ ЛІТІЙ-<br>ІОННИХ АКУМУЛЯТОРНИХ БАТАРЕЙ  |     |
| 4.  | <b>А. В. Гайдар, В. А. Готович</b>   | 89  |
|     | ЗАСТОСУВАННЯ КОМП'ЮТЕРНО-ІНФОРМАЦІЙНИХ ЗАСОБІВ В<br>ПРОЦЕСІ НАВЧАННЯ   |     |
| 5.  | <b>О.Р. Гончаренко, Є.В. Тиш</b>   | 90  |
|     | СИСТЕМИ КЕРУВАННЯ СОНЯЧНИХ ТРЕКЕРІВ  |     |
| 6.  | <b>Р.О.Жаровський, Д.В.Дармопук</b>  | 91  |
|     | ХАРАКТЕРИСТИКИ СИСТЕМ ЕЛЕКТРОННОГО НАВЧАННЯ  |     |
| 7.  | <b>С.А.Криськова</b>   | 92  |
|     | ХАРАКТЕРИСТИКА ГЕОІНФОРМАЦІЙНИХ СИСТЕМ. ПОБУДОВА<br>ГЕОІНФОРМАЦІЙНОЇ КАРТИ «ВИДАТНІ УКРАЇНСЬКІ ВЧЕНІ»                                      |     |
| 8.  | <b>Д.В. Кунинець, Ю.З. Лещинин</b>   | 94  |
|     | ЗАСТОСУНОК ДЛЯ МОНІТОРИНГУ ДАНИХ РОЗУМНОГО БУДИНКУ   |     |
| 9.  | <b>Р. М. Кучерешко</b>   | 95  |
|     | СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ НА ОСНОВІ НЕЧІТКОЇ<br>ЛОГІКИ ДЛЯ ВИЯВЛЕННЯ ШКІДЛИВИХ ПРОГРАМ  |     |
| 10. | <b>А.Д. Лавренів, І.В. Бойко</b>   | 97  |
|     | РОЗРОБКА МЕТОДІВ ДОСЛІДЖЕННЯ НЕЙРОННИХ МЕРЕЖ З<br>ВИКОРИСТАННЯМ СЕРЕДОВИЩА WOLFRAM MATHEMATICA ТА<br>МОВИ ПРОГРАМУВАННЯ C++                |     |
| 11. | <b>Я.Р. Лапшин</b>   | 99  |
|     | АНАЛІЗ ЗАГРОЗ КІБЕРБЕЗПЕКИ ТА ДОСЛІДЖЕННЯ ЙОГО ВПЛИВУ  |     |
| 12. | <b>Р.В. Ларіоник, Н.С. Луцик</b>   | 100 |
|     | КОМП'ЮТЕРНА СИСТЕМА ДЛЯ ДИСТАНЦІЙНОГО КОНТРОЛЮ<br>ЯКОСТІ АТМОСФЕРНОГО ПОВІТРЯ  |     |
| 13. | <b>Ю.З. Лещинин, З.В. Кузик</b>  | 101 |
|     | МЕТОДИ ТА ЗАСОБИ АВТОМАТИЗОВАНОЇ РОЗРОБКИ ТЕХНІЧНОЇ<br>ДОКУМЕНТАЦІЇ МЕРЕЖЕВИХ КАБЕЛЬНИХ СИСТЕМ   |     |

УДК 004.4

Жаровський Р.О. канд. техн. наук, Дармопук Д.В.

Тернопільський національний технічний університет імені Івана Пулюя

## ХАРАКТЕРИСТИКИ СИСТЕМ ЕЛЕКТРОННОГО НАВЧАННЯ

Zharovskyi R.O. PhD, Darmopuk D. V.

### CHARACTERISTICS OF E-LEARNING SYSTEMS

До складу базової архітектури традиційних систем електронного навчання та систем дистанційного навчання належать такі складові як:

- компонент реєстрації користувача та управління правами доступу до системи;
- компонент для провадження навчального процесу;
- модуль створення та оцінювання тестових завдань;
- компонент, що надає можливості формування та використання різних видів навчальних матеріалів;
- модуль для експорту та імпорту навчальних матеріалів з підтримкою різних форматів;
- компонент, що забезпечує інтерактивну складову взаємодії користувачів курсів: «лектор-студенти», «студенти-студенти», «студенти-лектор»;
- компонент фіксації активності користувачів.

Впровадження і застосування систем дистанційного навчання володіє рядом переваг і недоліків. Основними перевагами систем дистанційного навчання є:

- гнучкість – передбачає можливість одночасного або паралельного навчання у різних локаціях;
- свобода вибору – можливість здобувача освіти самостійно обирати швидкість навчання, формувати мітки матеріалу, наприклад, матеріал, який варто повторити, засвоєний матеріал, не цікавий і т.д.;
- формування власного розкладу навчання – передбачає навчання у зручний для студента час та у комфортних умовах;
- навчання під псевдонімом – забезпечення режиму інкогніто, що зумовлені певними обставинами;
- доступ до знань та освіти людьми з обмеженими можливостями;
- зростання особистих якостей студентів – полягає у виникненні почуття відповідальності, дисциплінованості та самоорганізованості;
- зростання навчальної аудиторії – системи дистанційного навчання дозволяють одночасно проводити навчання для більшого кола студентів, з різним рівнем знань та поглядів на життя;
- простота створення віртуального середовища – можливість викладачів, студентів тощо завдяки використанню сучасних інтернет технологій, за допомогою яких стає можливим обговорення між викладачами певних проблем, вирішення спільних завдань, обмін досвідом чи інформацією тощо.

До недоліків електронного навчання входять: відсутність прямої безпосередньої взаємодії між учасниками навчального процесу, методичні матеріали позбавлені емоційного забарвлення, складність створення творчої атмосфери, необхідність наявності відповідного апаратного і програмного забезпечення та ін.

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ТЕРНОПЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ІВАНА ПУЛЮЯ**

**МАТЕРІАЛИ**

**ІХ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ**

**«ІНФОРМАЦІЙНІ МОДЕЛІ,  
СИСТЕМИ ТА ТЕХНОЛОГІЇ»**



**8–9 грудня 2021 року**

**ТЕРНОПЛЬ  
2021**

<b>О.В. Балакунець, Є.В. Тиш</b> ПРИНЦИПИ ОРГАНІЗАЦІЇ ТА РОБОТИ КОНТРОЛЕРА РЕЗЕРВНОГО ЖИВЛЕННЯ <b>O.V. Balakunets, Ye.V. Tysh</b> PRINCIPLES OF ORGANIZATION AND WORK OF THE CONTROLLER RESERVE POWER SUPPLY	105
<b>В.П. Волоський, Ю.З. Лещинин, Н.Р. Романишин</b> АЛГОРИТМ БАЛАНСУВАННЯ LI-ION АКУМУЛЯТОРНИХ БАТАРЕЙ НА ОСНОВІ ПОТОЧНОЇ НАПРУГИ ТА НАПРУГИ ПРИ РОЗІМКНЕНОМУ КОЛІ <b>V.P. Voloskyi, N.R. Romanishin</b> LI-ION BATTERY BALANCING ALGORITHM BASED ON CURRENT VOLTAGE AND OPEN CIRCUIT VOLTAGE	106
<b>В.О. Дармограй, С.А. Лупенко</b> ТЕХНОЛОГІЯ АНАЛІЗУ ІОТ-ІНФРАСТРУКТУР AZURE DIGITAL TWINS В УМОВАХ КАРАНТИНУ COVID <b>V.O. Darmohrai, S.A. Lupenko</b> AZURE DIGITAL TWINS IOT-INFRASTRUCTURE ANALYSIS TECHNOLOGY IN COVID QUARANTINE CONDITIONS	107
<b>Р.О. Жаровський, Д.В. Дармопук</b> АНАЛІЗ УСПІШНОСТІ СТУДЕНТІВ НА ОСНОВІ ТЕХНОЛОГІЇ GRITNET <b>R.O. Zharovskyy, D.V. Darmopuk</b> STUDENT PERFORMANCE ANALYSIS BASED ON GRITNET TECHNOLOGY	108
<b>Ю.О. Дорош, М.М. Митник</b> ДОСЛІДЖЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ ДЛЯ НАКОПИЧЕННЯ КРИПТОВАЛЮТИ <b>Y.O. Dorosh, M.M. Mytnyk</b> RESEARCH OF THE AUTOMATED SYSTEM OF CRYPTO CURRENCY ACCUMULATION	109
<b>Д.О. Льченко, Р.О. Жаровський</b> МЕТОДИ ФІЛЬТРАЦІЇ СПАМУ В СУЧАСНИХ ПОШТОВИХ СИСТЕМАХ <b>D. Pchenko, R. Zharovskyy</b> SPAM FILTERING METHODS IN MODERN MAIL SYSTEMS	110
<b>Д.О. Льченко, Р.О. Жаровський</b> СЕМАНТИЧНІ МЕТОДИ ФІЛЬТРАЦІЇ СПАМУ <b>D. Pchenko, R. Zharovskyy</b> SEMANTIC METHODS OF SPAM FILTRATION	111
<b>В.В. Кохан, Є.В. Тиш</b> МЕТОДИ ОЦІНЮВАННЯ ЕМОЦІЙНОГО НАХИЛУ ТЕКСТІВ ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ <b>V.V. Kokhan, Ye.V. Tysh</b> METHODS OF EVALUATION OF SENTIMENT ANALYSIS OF TEXTS BY MEANS OF ARTIFICIAL INTELLIGENCE	112



УДК 004.89

Р.О. Жаровський канд. техн. наук, Д.В. Дармопук

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

**АНАЛІЗ УСПІШНОСТІ СТУДЕНТІВ НА ОСНОВІ ТЕХНОЛОГІЇ GRITNET**

UDC 004.89

R.O. Zharovsky PhD, D.V. Darmopuk

**STUDENT PERFORMANCE ANALYSIS BASED ON GRITNET TECHNOLOGY**

Проблема прогнозування успішності студентів частково досліджувалась спільнотою аналітиків та фахівцями з аналізу освітніх даних у формі прогнозування імовірності відрахування (або завершення навчання) студента. Ця задача є важливою проблемою підкласу прогнозування успішності здобувачів освіти. Більшість таких робіт можна поділити на дві групи, в залежності від використовуваних підходів.

Перший класичний підхід в основному ґрунтується на узагальнених лінійних моделях, які включають логістичну регресію, лінійний SVM та аналіз «виживання». Кожна модель враховує різні типи поведінкових і прогнозних особливостей, добутих з різних неопрацьованих записів активності студентів (наприклад, потоку кліків, оцінок, форуму).

Другий, новий підхід, включає дослідження успішності з використанням нейронних мереж. Дослідники використовували моделі глибоких нейронних мереж (DNN), рекурентні та згорткові нейронні мережі. Проте всі ці нові моделі, поки що, показали примітивну продуктивність. Це пов'язано головним чином із тим, що моделі все ще покладаються на розробку функцій для зменшення вхідних розмірів даних, що, здається, обмежує розробку більших (тобто кращих) моделей нейронних мереж. Записи діяльності студентів, зібрані з різних курсів, часто мають різну довжину, формат і зміст, тому функції, ефективні в одному курсі, можуть бути не такими в іншому. Навіть ретельно розроблені розміри об'єкта зазвичай мають бути малими.

Обидва підходи формують вхідні дані, які поки що занадто обмежені, щоб використати всі переваги моделей послідовного глибокого навчання. Щоб уникнути недоліків попередніх робіт, GritNet сприймає навчальну діяльність здобувачів освіти у часі як вихідні дані і (неявно) шукає частини послідовності вбудовування подій, які є найбільш дискримінаційними для прогнозування успішності студента без необхідності розробляти ці частини як (явні) вхідні ознаки.

Задачу прогнозування успішності студентів можна виразити у вигляді задачі послідовного прогнозування подій: враховуючи попередню (історичну) послідовність  $o \triangleq (o_1, o_2, \dots, o_T)$  у яких бере участь здобувач освіти потрібно оцінити імовірність майбутньої послідовності подій  $y \triangleq (y_{T+D}, \dots, y_{T'})$ , де  $D \in \mathbb{Z}_+$ .

У формі онлайн-занять кожна подія представляється у вигляді деякої діяльності студента, яка пов'язана з відповідною міткою часу. Іншими словами,  $o_t$  визначається як кортеж  $(a_t, d_t)$ .

Кожна дія  $a_t$  може представляти собою, наприклад, «переглянуто відео лекції», «відповідь на тестове запитання правильне/неправильне» або «проект виконано та здано/не здано», а  $d_t$  вказує відповідну (zareєстровану) мітку часу.

Таким чином, метою GritNet є обчислення індивідуального журналу логарифмічної правдоподібності  $p(y_t|v; \theta)$ , і ці оцінені бали можна просто додати для оцінки довгострокових результатів успішності студента.