

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

Факультет інформатики та обчислювальної техніки
Кафедра інформатики та програмної інженерії

«На правах рукопису»
УДК 004.015.2

«До захисту допущено»
В.о. завідувача кафедри
_____ Едуард ЖАРИКОВ
«__» _____ 2021 р.

Магістерська дисертація

на здобуття ступеня магістра

**за освітньо-професійною програмою «Інженерія програмного
забезпечення інформаційних систем»**

зі спеціальності 121 «Інженерія програмного забезпечення»

**на тему: «Архітектурне рішення програмного забезпечення для
аналізу результатів спортивних змагань»**

Виконав:

студент II курсу, групи ІТ-04мп
Репалюк Олег Ігорович _____

Керівник:

доцент кафедри ІІІ, к.т.н., доц.,
Мажара Ольга Олександрівна _____

Рецензент:

доцент кафедри АПЕПС ТЕФ, к.т.н., доц.,
Шаповалова Світлана Ігорівна _____

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць
інших авторів без відповідних посилань.
Студент _____

Київ – 2021 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Факультет інформатики та обчислювальної техніки

Кафедра інформатики та програмної інженерії

Рівень вищої освіти – другий (магістерський)

Спеціальність – 121 «Інженерія програмного забезпечення»

Освітньо-професійна програма «Інженерія програмного забезпечення інформаційних систем»

ЗАТВЕРДЖУЮ

В.о. завідувача кафедри

_____ Едуард ЖАРІКОВ

«__» _____ 2021р.

ЗАВДАННЯ
на магістерську дисертацію студенту
Репалюку Олегу Ігоровичу

1. Тема дисертації «Архітектурне рішення програмного забезпечення для аналізу результатів спортивних змагань», науковий керівник дисертації Мажара Ольга Олександрівна, кандидат технічних наук, доцент, затверджені наказом по університету від «27» жовтня 2021 р. № 3587-с
2. Термін подання студентом дисертації «6» грудня 2021 р.
3. Об'єкт дослідження – програмне забезпечення для аналізу результатів спортивних змагань.
4. Предмет дослідження – архітектура програмного забезпечення для аналізу результатів спортивних змагань.
5. Перелік завдань, які потрібно розробити – провести дослідження існуючих рішень, що виконують задачу аналізу результатів спортивних змагань; розробити архітектуру програмного продукту; реалізувати програмне забезпечення аналізу результатів спортивних змагань; провести оцінку ефективності отриманих прогнозів результатів за двома тестовими алгоритмами.

6. Орієнтовний перелік графічного (ілюстративного) матеріалу – структурна схема варіантів використання; структурна схема класів; екранні форми.

7. Орієнтовний перелік публікацій – тези доповіді на науковій конференції.

8. Консультанти розділів дисертації

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

9. Дата видачі завдання «30» вересня 2020 р.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання	Примітка
1	Ознайомлення з завданням	1.10.2021	
2	Аналіз предметної області	3.10.2021	
3	Дослідження існуючих рішень	8.10.2021	
4	Формування вимог до програмного продукту	12.10.2021	
5	Проектування програмного забезпечення	15.10.2021	
6	Реалізація програмного продукту	10.11.2021	
7	Виконання експериментальних досліджень	15.11.2021	
8	Оформлення пояснювальної записки	20.11.2021	
9	Подання дисертації на попередній захист	22.11.2021	
10	Подання дисертації на захист	6.12.2021	

Студент

Олег РЕПАЛЮК

Науковий керівник

Ольга МАЖАРА

РЕФЕРАТ

Розмір пояснювальної записки – 110 аркушів, містить 15 ілюстрацій, 36 таблиць, 6 додатків.

Актуальність теми. У роботі розглянуто проблему архітектурного рішення програмного забезпечення для аналізу результатів спортивних змагань, показано основні особливості існуючих програмних продуктів, їх переваги та недоліки. Виявлено потребу в розробці програмного забезпечення з інтеграцією з системою управління взаємовідносинами з клієнтами.

Мета дослідження. Основною метою є удосконалення програмного забезпечення аналізу результатів спортивних змагань у архітектурному рішенні системи управління взаємовідносинами з клієнтами для розширення сфери його використання.

Об'єкт дослідження: програмне забезпечення для аналізу результатів спортивних змагань.

Предмет дослідження: архітектура програмного забезпечення для аналізу результатів спортивних змагань.

Для реалізації поставленої мети **сформульовані наступні завдання:**

- провести дослідження існуючих рішень, що виконують задачу аналізу результатів спортивних змагань;
- розробити архітектуру програмного продукту;
- реалізувати програмне забезпечення аналізу результатів спортивних змагань;
- провести оцінку ефективності отриманих прогнозів результатів за двома тестовими алгоритмами.

Наукова новизна результатів магістерської дисертації полягає в тому, що запропоновано архітектуру програмного забезпечення для аналізу результатів спортивних змагань, яка відрізняється від інших орієнтацією на використання

малими та середніми підприємствами. Результат досягнутий шляхом інтеграції з системою управління взаємовідносинами з клієнтами.

Практичне значення отриманих результатів полягає в тому, що реалізоване програмне забезпечення є адаптованим для застосування у корпоративному секторі.

Зв'язок з науковими програмами, планами, темами. Робота виконувалась на кафедрі інформатики та програмної інженерії Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського".

Апробація. Наукові положення дисертації пройшли апробацію на Міжнародній науковій інтернет-конференції «Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 62)» – Тернопіль.

Публікації. Наукові положення дисертації опубліковані в:

Репалюк О.І. Прогнозування результатів спортивних змагань на основі статистичних моделей/ О.І. Репалюк //Збірник тез доповідей Міжнародної наукової інтернет-конференції «Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 62)» – м. Тернопіль, 12 жовтня 2021р.

Ключові слова: АРХІТЕКТУРА, АНАЛІЗ РЕЗУЛЬТАТІВ СПОРТИВНИХ ЗМАГАНЬ, СИСТЕМА УПРАВЛІННЯ ВЗАЄМОВІДНОСИНАМИ З КЛІЄНТАМИ.

ABSTRACT

Explanatory note size – 110 pages, contains 15 illustrations, 36 tables, 6 applications.

Topicality. Examines the problem of architectural software solution for analysis of sports results. The paper shows the main features of existing software products, their advantages and disadvantages. The need for software development with integration with the customer relationship management system has been identified.

The aim of the study. The main target is to improve the software for analyzing the results of sports competitions in the architectural solution of the customer relationship management system to expand the scope of its use.

Object of research: software for analyzing the results of sports competitions.

Subject of research: architecture of software for analyzing the results of sports competitions.

To achieve this goal, the **following tasks** were formulated:

- to conduct research on existing solutions that perform the task of analyzing the results of sports competitions;
- develop the architecture of the software product;
- implement software for analyzing the results of sports competitions;
- to evaluate the effectiveness of the obtained predictions by two test algorithms.

The scientific novelty of the results of the master's dissertation is the proposed software architecture for analyzing the results of sports competitions, which differs from others in the focus on using by small and medium enterprises. The result is achieved by integration with the customer relationship management system.

The practical value of the obtained results is that the implemented software is adapted for using in the corporate sector.

Relationship with working with scientific programs, plans, topics. Work was performed at the Department of Informatics and Software Engineering of the National Technical University of Ukraine «Kyiv Polytechnic Institute. Igor Sikorsky».

Approbation. The scientific provisions of the dissertation were tested at the International Scientific Internet Conference "Information Society: Technological, Economic and Technical Aspects of Formation (Issue 62)" - Ternopil.

Publications. The scientific provisions of the dissertation published in:

Repaliuk O.I. Predicting the results of sports competitions based on statistical models / O.I. Repaliuk // Proceedings of the International Scientific Internet Conference "Information Society: Technological, Economic and Technical Aspects of Formation (Issue 62)" - Ternopil, October 12, 2021.

Keywords: ARCHITECTURE, ANALYSIS OF SPORTS COMPETITION RESULTS, CUSTOMER RELATIONSHIP MANAGEMENT SYSTEM.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ.....	10
ВСТУП	11
1 АНАЛІЗ РІШЕНЬ З ПРОГНОЗУВАННЯ РЕЗУЛЬТАТІВ СПОРТИВНИХ ЗМАГАНЬ.....	13
1.1 Огляд сфери прогнозування спортивних результатів.....	13
1.2 Існуючі рішення.....	16
1.2.1 Odds Wizard	16
1.2.2 BetClan	18
1.2.3 Prediction Machine	19
1.3 Формування вимог до розробки.....	20
Висновки до розділу	21
2 КОНСТРУЮВАННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ.....	22
2.1 Методи прогнозування.....	22
2.1.1 Поняття машинного навчання.....	22
2.1.2 Лінійна регресія	23
2.1.3 Логістична регресія	25
2.1.4 Штучні нейронні мережі.....	29
2.1.5 Наївний баєсів класифікатор	32
2.1.6 Метод опорних векторів	33
2.1.7 Дерева прийняття рішень.....	36
2.1.8 Вибір методів прогнозування.....	39
2.2 Визначення структури програмного забезпечення.....	39

	9
Висновки до розділу	43
3 ОПИС ПРОГРАМНОГО ТА ТЕХНІЧНОГО ЗАБЕЗПЕЧЕННЯ	44
3.1 Технології для розробки.....	44
3.2 Опис архітектури програмного продукту	48
3.3 Прогнозування методами машинного навчання	57
3.4 Результати дослідження ефективності системи	63
3.5 Керівництво користувача.....	66
Висновки до розділу	70
4 РОЗРОБЛЕННЯ СТАРТАП-ПРОЄКТУ	71
4.1 Інформаційна карта проєкту.....	71
4.2 Аналіз потенційних техніко-економічних переваг ідеї	72
4.3 Технологічний аудит ідеї проєкту	73
4.4 Аналіз ринкових можливостей стартап-проєкту.....	74
4.5 Ринкова стратегія проєкту	85
Висновки до розділу	94
ВИСНОВКИ	96
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	97
ДОДАТОК А.....	102
ДОДАТОК Б	106
ДОДАТОК В.....	107
ДОДАТОК Г	108
ДОДАТОК Д.....	109

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

CRM (Customer relationship management) – поняття, що охоплює концепції, котрі використовуються компаніями для управління взаємовідносинами зі споживачами, включаючи збір, зберігання й аналіз інформації про споживачів, постачальників, партнерів та інформації про взаємовідносини з ними

API (Application Programming Interface) – набір чітко визначених методів для взаємодії різних компонентів застосунку

REST (Representational State Transfer) – архітектурний стиль взаємодії компонентів розподіленого застосунку в мережі

SQL (Structured Query Language) – декларативна мова програмування, що застосовується для взаємодії користувача з базами даних

JSON (JavaScript Object Notation) – текстовий формат обміну даними

LINQ (Language Integrated Query) – компонент .NET Framework, що надає нативні можливості виконання запитів даних до мов, що входять у .NET

CLR (Common Language Runtime) – компонент пакету .NET Framework, віртуальна машина, на якій виконуються всі мови платформи .NET Framework

HTTP (HyperText Transfer Protocol) – протокол передачі даних, що використовується в комп'ютерних мережах

ВСТУП

На даний момент більшість популярних видів спорту успішно комерціалізувалися. Спортивні клуби хочуть перемагати та заробляти кошти на свій подальший розвиток, глядачі – відвідувати змагання, виробники спортивних товарів – нарощувати продажі, спонсори – просувати свої бренди. Тому, оцінка ефективності команд та спортсменів, а також прогнозування їх успішності є важливою для багатьох сфер бізнесу: формування спонсорських контрактів, букмекерська діяльність, проведення та обслуговування спортивних подій.

Програмне забезпечення, що реалізує моделі аналізу статистичних даних може використовуватись у щоденній діяльності для побудови довгострокової стратегії відносин зі спортивними організаціями. Однак на сьогодні компанії, які пропонують прогнози щодо результатів матчу, зазвичай не розкривають дані щодо використовуваних алгоритмів і моделей. Актуальною є задача створення відкритого програмного забезпечення для перевірки ефективності статистичних моделей прогнозування результатів на основі даних попередніх ігор та поточних характеристик.

Метою даної роботи є удосконалення програмного забезпечення аналізу результатів спортивних змагань у архітектурному рішенні системи управління взаємовідносинами з клієнтами для розширення сфери його використання.

Об'єктом дослідження є програмне забезпечення для аналізу результатів спортивних змагань. Предмет дослідження: архітектура програмного забезпечення для аналізу результатів спортивних змагань.

Для досягнення поставленої мети визначені наступні завдання:

- провести дослідження існуючих рішень, що виконують задачу аналізу результатів спортивних змагань;
- розробити архітектуру програмного продукту;
- реалізувати програмне забезпечення аналізу результатів спортивних змагань;

– провести оцінку ефективності отриманих прогнозів результатів за двома тестовими алгоритмами.

Програмний модуль допоможе отримувати в режимі реального часу дані про майбутні спортивні події хокею, баскетболу, бейсболу щодо величини, яка визначає ймовірність перемоги однієї з команд або величини, яка визначає ймовірність кількості забитих голів/набраних балів (більше/менше середнього значення).

1 АНАЛІЗ РІШЕНЬ З ПРОГНОЗУВАННЯ РЕЗУЛЬТАТІВ СПОРТИВНИХ ЗМАГАНЬ

1.1 Огляд сфери прогнозування спортивних результатів

Прогнозування спортивних результатів є цікавою та складною проблемою завдяки непередбачуваній природі спорту та, здавалося б, великій кількості потенційних факторів, які можуть вплинути на результати. Дійсно, саме ця непередбачувана природа є однією з основних причин того, чому люди захоплюються спортом.

Незважаючи на свою складність, спроба передбачити результати спорту – це те, що цікавить багато різних сфер бізнесу, і цей інтерес збільшився, оскільки спортивні дані стають все більш доступними в Інтернеті, а також за рахунок розвитку онлайн ставок на спорт. Відомих колишніх гравців та спортивних експертів часто запитують щодо їх прогнозів на майбутні матчі, які потім з'являються в Інтернеті чи у газетах.

Прогнозування спортивних результатів також може бути потенційно корисним для гравців, керівництва команди та аналітиків при виявленні найважливіших чинників, які допомагають досягти виграшних результатів на основі яких можна визначити відповідну тактику.

Враховуючи все вище викладене, не є дивним той факт що за останнє десятиліття у спортивній індустрії відбулося різке збільшення як обсягу, так і різноманітності внутрішньо-ігрової статистики, що генерується під час гри. Цьому сприяє активний розвиток інформаційних технологій.

Застосування інтелектуального аналізу даних у спорті є багатогранним, і дозволяє виявляти фактори, що сприяють успішній діяльності, оцінювати їх та розробляти моделі для прогнозування фінальних результатів конкретних ігор або турнірів.

Для досягнення більш точного прогнозу на спортивні події, перед грою необхідно проаналізувати головні показники команд, що впливають на результат. Проте, не всі з них відображають потенціал команди у рівних пропорціях. Для того, щоб врахувати ступінь значущості показника, необхідно ввести відповідні коефіцієнти. Розглянемо особливості видів спорту, для яких передбачається розробка модулю прогнозування.

Хокей із шайбою є командною спортивною грою із ключкою на льоду, що полягає в протиборстві двох команд на ковзанах, які, передаючи шайбу ключками, прагнуть закинути її якомога більшу кількість разів у ворота суперника та не пропустити в свої. Одночасно на полі з боку однієї команди повинні перебувати шість гравців: один воротар та п'ять польових. Прогнозування результатів хокею буде розглядатись у лізі NHL, що є спортивною організацією яка об'єднує клуби США та Канади. В турнірі грає 32 команди, і в регулярному сезоні кожна команда грає 82 гри – по 41 гри вдома та в гостях. Крім того, найкращі команди виходять у плей-офф, де визначають єдиного переможця[1].

Баскетбол – це спортивна командна гра з м'ячом, у якій необхідно м'яч закинути руками в кільце команди суперника. У баскетболі грають дві команди, кожна з яких складається з п'яти польових гравців. Обидві команди намагаються закинути м'яч у кільце з сіткою (корзиною) суперника, а також заважають іншій команді заволодіти м'ячом, і закинути його у свою корзину. Прогнозування результатів баскетболу буде розглядатись у лізі NBA, яка є чоловічою професійною баскетбольною лігою клубів Канади та США. В турнірі приймають участь 30 команд. Протягом регулярного сезону кожна команда має можливість зіграти 82 матчі, після чого найкращі команди потрапляють до плей-офф, де змагаються за олімпійською системою до 4 перемог[2].

Бейсбол – це спортивна гра з використанням м'яча та бити. Грають на полі дві команди, кожна з яких складається з дев'яти гравців, які по черзі грають в

нападі та захисті. Мета команди, що грає в нападі – заробити ран (пробіжку). Щоб досягнути цього завдання, гравці за допомогою бити відбивають м'яч, що кидає пітчер. Після цього вони послідовно пробігають через чотири бази, рухаючись в напрямку проти годинникової стрілки. У випадку якщо гравець пройшов всі чотири бази і повернувся на домашню пробіжка зараховується. Прогнозування результатів бейсболу буде розглядатись у лізі MLB, яка професійною бейсбольною лігою клубів Канади та США. Турнір включає 30 команд, протягом регулярного сезону кожна з яких грає по 162 матчі, кращі команди виходять у плей-офф, і за олімпійською системою визначають головного переможця[3].

Активна розробка моделей для прогнозування спортивних результатів розпочалася у минулому столітті. У 1984 році була розроблена модель для того, щоб передбачити результат баскетбольних ігор коледжу в Університеті Айови. Точність моделі була досить висока та лише трохи відрізнялася між чоловічими та жіночими командами, з успішністю 88% та 90% відповідно[4].

К. Хуанг і В. Чанг використовували нейронну мережу для прогнозування результатів чемпіонату світу з футболу 2006. Вони заявили про точність близько 75% [5].

М. Пурукер використав нейронні мережі для прогнозування результатів матчів з американського футболу ліги NFL. Успішність прогнозування склала приблизно 78.6% [6]. М. Парді досліджував прогнозування результатів матчів коледжної ліги NCAA з американського футболу, використовуючи також нейронні мережі. У результаті була досягнута точність 76% [7].

Крім того, в одному дослідженні 2008 року, вчені за допомогою штучних нейронних мереж, намагалися спрогнозувати результати у чотирьох різних видах спорту: NFL (американський футбол), AFL (австралійський футбол), супер-регбі (регбі-союз) та англійська прем'єр-ліга з футболу (EPL), використовуючи дані за 2002 рік. При цьому середній відсоток вгаданих результатів спортивних подій

алгоритмів штучних нейронних мереж склав близько 67,5%, що трохи перевищує результати аналогічних прогнозів експертів, точність яких складала близько 60–65% [8].

В 2009 році було опубліковане дослідження прогнозування результатів матчів баскетбольної ліги NBA з використанням нейронних мереж. Цифра точності прогнозування складає 74,33% [9].

1.2 Існуючі рішення

Програмне забезпечення для прогнозування результатів спортивних подій використовується значною мірою у букмекерському бізнесі, як букмекерськими компаніями так і професійними гравцями ринку, що грають у букмекерських продуктах. Тому більшість програм мають додаткові функції для розрахунку специфічних коефіцієнтів на перемогу, кількість забитих голів. Результати розрахунку дозволяють спланувати тактику здійснення ставок і отримати перевагу над букмекерськими стратегіями, що вимушені враховувати кількість ставок та відповідно коригувати коефіцієнти.

Також у існуючих рішеннях широко використовуються моделі для максимізації прибутку при ставках на результати спортивних подій. При цьому будується модель оптимальних ставок: розподіл обмеженого ресурсу на ставки для отримання максимального результату від правильних прогнозів.

1.2.1 Odds Wizard

Odds Wizard – це автоматизоване програмне забезпечення для статистичного аналізу спортивних подій. Даний програмний комплекс базуючись на минулих подіях генерує ймовірності настання певного результату для майбутніх ігор у багатьох видах спорту, таких як: бейсбол, теніс, баскетбол, футбол, хокей, регбі, американський футбол.

The screenshot displays the Odds Wizard software interface for the Premier League. The top menu includes File, Teams, Games, Fixtures, Odds, Internet, Options, Charts, Tools, and Help. The main window is divided into two panes. The left pane shows a table of team statistics, and the right pane shows a list of fixtures.

N	Teams	Rating	In-total	Attack	Defence	Place	Points
1	Manchester City FC	0.631	0.856	0.712	0.455	2	15
2	Chelsea FC	0.507	0.714	0.579	0.464	14	8
3	Arsenal FC	0.493	0.747	0.588	0.441	4	13
4	Manchester United FC	0.389	0.717	0.521	0.404	1	16
5	Liverpool FC	0.316	0.787	0.520	0.332	9	11
6	Southampton FC	0.264	0.647	0.424	0.376	10	9
7	Tottenham Hotspur FC	0.219	0.744	0.450	0.305	6	12
8	Everton FC	0.214	0.688	0.419	0.330	5	12
9	Swansea City AFC	0.178	0.713	0.414	0.300	11	9
10	West Ham United FC	0.139	0.696	0.386	0.289	3	13
11	AFC Bournemouth	0.115	0.776	0.414	0.237	16	7
12	Stoke City FC	0.099	0.709	0.372	0.263	17	6
13	Leicester City FC	0.096	0.786	0.409	0.223	8	12

N	Date	Time	Home team	Away team	G1	G2	HT	Rem
5843		17:00	Leicester City FC	Arsenal FC	2	5	1:2	
5844		19:30	Newcastle United FC	Chelsea FC	2	2	1:0	
5845	27.09.15	18:00	Watford FC	Crystal Palace FC	0	1	0:0	
5846	28.09.15	22:00	West Bromwich Albion	Everton FC	2	3	1:0	
5847	03.10.15	14:45	Crystal Palace FC	West Bromwich Albion I				
5848		17:00	Sunderland AFC	West Ham United FC				
5849		17:00	AFC Bournemouth	Watford FC				
5850		17:00	Manchester City FC	Newcastle United FC				
5851		17:00	Aston Villa FC	Stoke City FC				
5852		17:00	Norwich City FC	Leicester City FC				
5853		19:30	Chelsea FC	Southampton FC				
5854	04.10.15	15:30	Everton FC	Liverpool FC				
5855		18:00	Swansea City AFC	Tottenham Hotspur FC				

N	Home team	Away team	Kickoff	Status	Min'	Score	1	X	2	1X	12	2X	AHO	1	2	Total	<	>	Both	Not	Score odds	Score odds		
	Stoke City FC	AFC Bournemouth	26.09.17.00	Finished		2:1	2,33	3,55	3,46	1,41	1,39	1,75	-0,25	2	2	2,5	2,12	1,9	1,61	2,65	1:1	6,78	2:1	8,06
	Liverpool FC	Aston Villa FC	26.09.17.00	Finished		3:2	1,38	5,38	11,2	1,098	1,23	3,63	-1,25	1,91	2,1	2,75	2,03	1,97	1,9	2,11	2:1	7,36	2:0	7,45
	Leicester City FC	Arsenal FC	26.09.17.00	Finished		2:5	4,9	3,95	1,84	2,19	1,34	1,26	+0,75	1,89	2,12	2,75	2,02	1,98	1,59	2,68	1:2	7,34	1:1	7,65
	Newcastle United F	Chelsea FC	26.09.19.30	Finished		2:2	8,69	4,63	1,5	3,02	1,26	1,13	+1	2,16	1,86	2,5	2,09	1,92	1,9	2,11	1:2	7,16	0:2	7,54
	Watford FC	Crystal Palace FC	27.09.18.00	Finished		0:1	2,74	3,25	3,06	1,49	1,44	1,57	0	1,9	2,12	2,25	1,94	2,07	1,84	2,2	1:1	5,8	1:0	8,94
	West Bromwich Alt	Everton FC	28.09.22.00	Finished		2:3	3,15	3,27	2,66	1,6	1,44	1,47	0	2,18	1,84	2,25	1,97	2,03	1,82	2,22	1:1	5,86	1:2	8,96
	Crystal Palace FC	West Bromwich Alt	03.10.14.45	Fixture			2,19	3,35	4,07	1,33	1,43	1,84	-0,25	1,87	2,15	2,25	1,93	2,07	1,88	2,14	1:1	5,95	1:0	7,9
	Sunderland AFC	West Ham United F	03.10.17.00	Fixture			3,4	3,28	2,49	1,67	1,44	1,42	+0,25	1,9	2,11	2,25	1,96	2,04	1,83	2,2	1:1	5,87	0:1	8,61
	AFC Bournemouth	Watford FC	03.10.17.00	Fixture			1,92	3,71	4,79	1,26	1,37	2,09	-0,5	1,92	2,09	2,5	1,99	2,01	1,73	2,38	1:1	6,78	2:1	7,37
	Manchester City FC	Newcastle United F	03.10.17.00	Fixture			1,15	10,1	30,8	1,034	1,11	7,6	-2,25	2,11	1,9	3,25	2,03	1,97	2,15	1,67	3:0	7,22	2:0	7,54
	Aston Villa FC	Stoke City FC	03.10.17.00	Fixture			3,44	3,39	2,41	1,71	1,42	1,41	+0,25	1,95	2,06	2,25	2,17	1,86	1,73	2,36	1:1	6,17	1:2	8,29
	Norwich City FC	Leicester City FC	03.10.17.00	Fixture			2,97	3,58	2,61	1,62	1,39	1,51	0	2,14	1,88	2,75	1,96	2,04	1,54	2,84	1:1	7,06	1:2	8,64
	Chelsea FC	Southampton FC	03.10.19.30	Fixture			1,69	3,96	6,42	1,18	1,34	2,45	-0,75	1,91	2,09	2,5	1,88	2,14	1,9	2,11	1:1	6,99	2:1	7,22
	Everton FC	Liverpool FC	04.10.15.30	Fixture			2,69	3,48	2,94	1,52	1,4	1,59	0	1,91	2,09	2,5	2,07	1,93	1,61	2,64	1:1	6,62	2:1	8,76
	Swansea City AFC	Tottenham Hotspur	04.10.18.00	Fixture			2,43	3,49	3,31	1,43	1,4	1,7	-0,25	2,08	1,92	2,5	2,03	1,97	1,64	2,57	1:1	6,55	2:1	8,26

Teams: 27 Games: 6156 Average total: 2,69 Home advantage: 0,302 Profit margin: 0% Last updated: 28.09.15

Рисунок 1.1 – Інтерфейс системи Odds Wizard

Програма була розроблена після чемпіонату світу з футболу 1998. Перші алгоритми прогнозування були досить прості, та з часом вони поступово ускладнювались. Спочатку дане програмне забезпечення підтримувало прогнозування тільки з одного виду спорту – футболу, а зараз реалізовано прогнозування більше п'яти видів спорту.

Розробники не розкривають алгоритми прогнозування, що використовуються у даному програмному забезпеченні, обмежуючись лише твердженням, що вони унікальні.

Основними можливостями програмного забезпечення Odds Wizard є:

- завантаження даних щодо наступних спортивних подій через мережу Інтернет;
- прогнозування результатів матчів у вигляді відсоткової ймовірності настання певного результату;
- побудова таблиць, що порівнюють команди з однієї ліги за різними показниками;

- букмекерський модуль, який генерує рекомендації на які саме події треба робити ставки;
- прогнозування результатів під час гри в режимі реального часу.

Варто зазначити, що для користування даним програмним забезпеченням необхідно встановити його у себе на комп'ютері. Ця програмна система є платною. Користувачу пропонується заплатити єдиноразову реєстраційну суму, а потім кожен рік продовжувати свою підписку.

Програма орієнтована на використання у ставках на спорт і має додаткові алгоритми для генерації оптимальної стратегії ставок на результати спортивних подій. Програма є аналогом букмекерської компанії і відображає ті ж самі коефіцієнти, що і букмекерські компанії, що надає можливість порівняти згенеровані прогнози з іншими джерелами.

Перевагами системи є підтримка багатьох видів спорту, та великої кількості ліг. Недоліки полягають у тому, що система не має веб-клієнту для її зручного використання, і реалізує підтримку тільки однієї операційної системи. Інтерфейс програми є застарілим. Інтеграція з іншими системами без розроблення додаткових модулів не передбачена. Це буде ускладнювати використання даної системи на підприємстві.

1.2.2 BetClan

BetClan – це ще одне рішення програмного забезпечення для аналізу та прогнозування результатів спортивних змагань. Прогнози генеруються для таких видів спорту як: футбол, крикет, баскетбол, теніс, хокей з шайбою, американський футбол, гандбол, волейбол та баскетбол.

Дане програмне забезпечення безкоштовне. Для отримання даних щодо прогнозів користувачу необхідно зайти на відповідний веб-портал, що має зручний інтуїтивний інтерфейс. Алгоритми прогнозування результатів

спортивних подій, що використовує система BetClan є невідомими. Приклад користувацького інтерфейсу системи зображений на ілюстрації нижче (рис. 1.2).

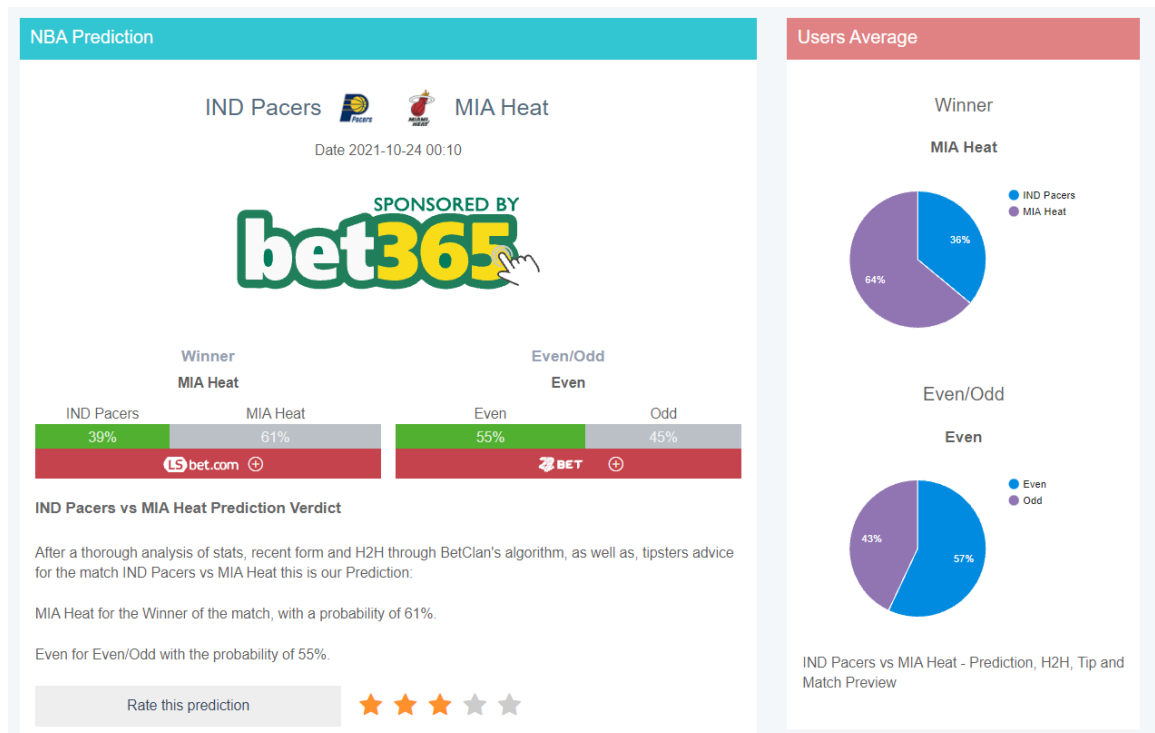


Рисунок 1.2 – Інтерфейс системи BetClan

Серед переваг цього програмного забезпечення можна виділити зручний веб-портал та розроблений мобільний додаток. Проте, існують також недоліки. По-перше, при взаємодії з даним програмним продуктом користувачу може заважати велика кількість реклами. По-друге, в програмному забезпеченні не передбачена можливість інтеграції з системою управління взаємовідносинами з клієнтами, що буде ускладнювати використання продукту у корпоративних цілях.

1.2.3 Prediction Machine

Prediction Machine – це система прогнозування спортивних подій, яка в режимі реального часу розраховує ймовірності по матчам з американського футболу, баскетболу, бейсболу. Продукт спеціалізується на прогнозуванні у спортивних професійних лігах Північної Америки.

Дане програмне забезпечення є платним. Користувачу може надаватись безкоштовний тестовий період 5 днів, щоб випробувати систему. Дані щодо алгоритмів, які використовує продукт не розкриваються. Доступна інформація, що кожна гра моделюється багато разів для того, щоб забезпечити найбільш точне прогнозування. При цьому крім результату також забезпечується прогнозування матчевої статистики кожної з команд.

Перевагою даної системи є розрахунок прогнозів на велику кількість статистичних показників гри в режимі реального часу. Основними недоліками є досить висока ціна за користування даним програмним забезпеченням, а також відсутність відкритого API для зручної інтеграції програмного забезпечення з іншими системами.

1.3 Формування вимог до розробки

Проаналізувавши існуючі аналоги, варто зазначити, що існує значна кількість систем аналізу та прогнозування результатів спортивних змагань. Проте, наявні аналоги зазвичай орієнтовані на окремих користувачів. Взаємодія з програмним забезпеченням відбувається через веб-сайт або через власний додаток. Використання деяких продуктів супроводжується великою кількістю реклами. Більшість систем не мають відкритого API.

Загалом, фактори перелічені вище ускладнюють можливу інтеграцію програмного забезпечення з іншими корпоративними системами за умови використання програмного забезпечення на підприємстві. Тому, виникає завдання створити систему, що буде більш придатною для корпоративних цілей. Це завдання може бути розв'язано шляхом інтеграції програмного продукту з системою управління взаємовідносинами з клієнтами.

Крім того, розробники аналогів не розкривають алгоритмів прогнозування, які використовуються в їх системах. Це може бути значним недоліком у випадку коли виникає сумнів у достатній ефективності працюючого алгоритму. Постає

задача створити програмне забезпечення, де будуть використовуватись декілька алгоритмів прогнозування, і можна буде оцінити їх ефективність для певного спорту, ліги тощо, що потім дасть можливість вибрати той, який покаже кращі результати на певних визначених умовах.

Отже, розроблюване програмне забезпечення повинно реалізовувати наступні можливості:

- прогнозування результатів змагань, використовуючи ефективні моделі;
- інтеграція програмного модулю з системою управління взаємовідносинами з клієнтами;
- прогнозування результатів з використанням декількох статистичних алгоритмів.

Висновки до розділу

У цьому розділі був проведений огляд сфери діяльності з прогнозування результатів спортивних змагань, описані минулі дослідження з даної теми. Розглянуто існуюче програмне забезпечення, що виконує задачу аналізу та прогнозування результатів спортивних змагань. Визначено загальні вимоги до програмного продукту.

2 КОНСТРУЮВАННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

2.1 Методи прогнозування

2.1.1 Поняття машинного навчання

Проблема прогнозування спортивних результатів деякий час розглядалася в літературі з питань статистики та операційних досліджень, однак застосування методів машинного навчання (ML) до проблеми є нещодавнім напрямом досліджень. Перше дослідження в цій галузі було опубліковане в 1996 році, і з тих пір ця тема викликала підвищений науковий інтерес[10].

Машинне навчання – розділ штучного інтелекту, що вивчає алгоритми, здатні навчатися або адаптувати свою структуру на підставі обробленої вибірки даних[11]. Це одна з інтелектуальних методологій, що показала багатообіцяючі результати в галузях класифікації та прогнозування.

Підходи машинного навчання традиційно поділяють на дві категорії в залежності від характеру «сигналу» або «зворотного зв'язку», доступних системі навчання:

– навчання з учителем, коли комп'ютерові надають множину вхідних даних та їх бажаних виходів, наданими учителем, і кінцевим завданням є вивчення загального правила, яке позначає вхідні дані на вихідні дані;

– навчання без учителя. Навчання без нагляду може бути самоціллю (виявлення прихованих закономірностей у даних) або засобом досягнення мети (навчання за ознаками).

Щоб прогнозувати спортивні змагання для формування вибірки навчальних прикладів може використовуватись історична статистика матчів. Для окремого матчу вхідний вектор може містити різні ознаки матчу, а вихідним значенням буде результат матчу[12].

В такому випадку, маючи статистику минулих результатів, їх можна використати як вхідні дані при навчанні. Тому при застосуванні машинного

навчання для прогнозування результатів спортивних змагань, варто розглядати підхід навчання з учителем.

З точки зору існуючих алгоритмів, прогнозування спортивних змагань можна розглянути з двох сторін:

- задача регресії, в якій вихідне значення представляє собою ймовірність виграшу матчу однією з команд;
- задача бінарної класифікації, в якій можна спробувати класифікувати матчі за категоріями перемога або поразка.

2.1.2 Лінійна регресія

Лінійна регресія характеризується спробою моделювання зв'язку між змінними шляхом обрання найбільш підходящого лінійного рівняння до спостережуваних даних[13]. Математичне рівняння, що розкриває поняття лінійної регресії має наступний вид:

$$y = \beta_1 x_1 + \beta_2 x_2 \dots + \beta_n x_n + \varepsilon, \quad (2.1)$$

де y – залежна змінна;

$x_1 \dots x_n$ – незалежні змінні;

$\beta_1 \dots \beta_n$ – коефіцієнти рівняння;

ε – випадкова величина.

Якщо розглядати дане рівняння з точки зору прогнозування результатів спортивних змагань, то залежна змінна являє собою результат прогнозування; незалежні змінні – фактори, що впливають на результат матчу; коефіцієнти є вагами і визначають значимість кожного фактору на кінцевий результат.

Одним із способів підбору ваги для лінійної моделі є метод найменших квадратів. Він полягає у мінімізації залишкової суми квадратів між фактичними значенням результатів y та прогнозованих значень $X\beta$, де X є матрицею вхідних значень.

При цьому для визначення коефіцієнтів фактичне значення величини можна брати з результатів спортивних подій, що вже відбулися, симулюючи їх

прогнозування. Після того як була створена модель з підібраними найкращими можливими коефіцієнтами, можна прогнозувати подальші результати. На рис. 2.1 зображено приклад побудови графіку лінійної регресії у двомірному просторі.

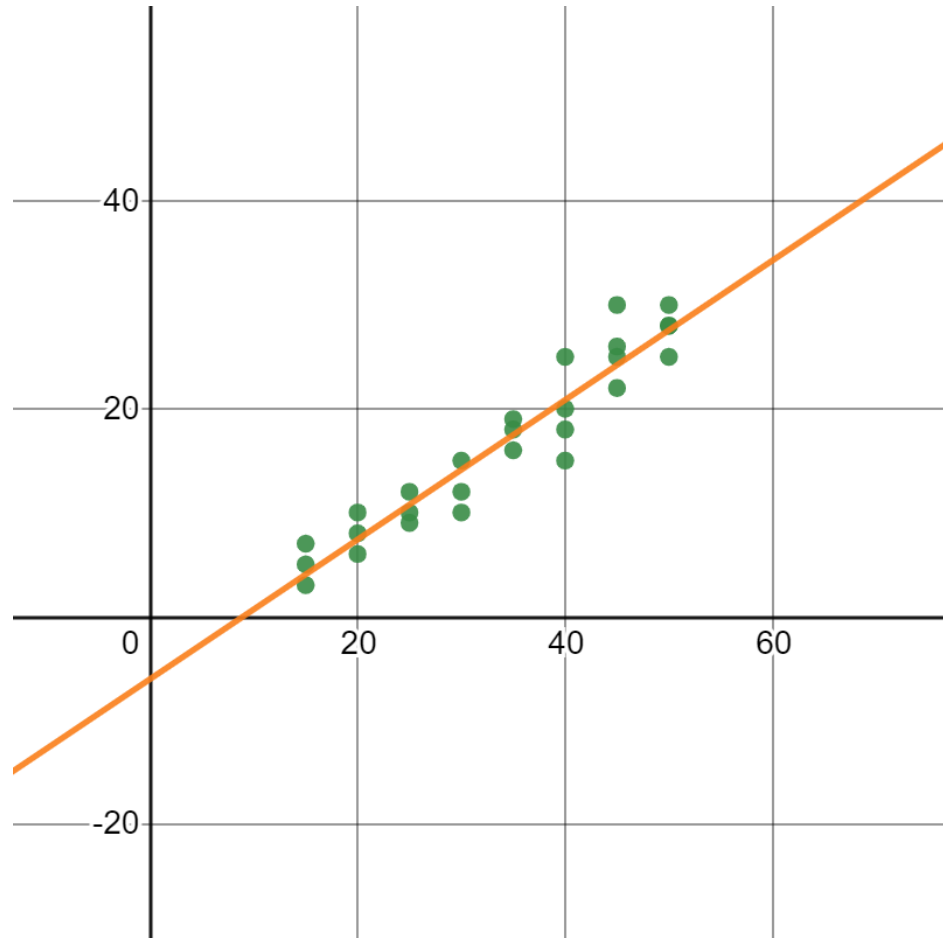


Рисунок 2.1 – Приклад графіку лінійної регресії

У даному випадку це звичайна пряма виду: $f(x) = a \cdot x + b$. З додаванням ще одного параметру, результуюча функція набуде вигляду площини. Варто зазначити, що функція не повинна обов'язково проходити через усі точки, так як дані можуть містити шум. Тому, функція повинна передавати загальну тенденцію.

Моделі на основі лінійної регресії достатньо легко будуються. Однак, якщо змінні сильно корелюють між собою, то результати, що надаватиме модель будуть низької точності.

2.1.3 Логістична регресія

Незважаючи на свою назву, логістична регресія є фактично алгоритмом класифікації для проблем, що у якості вихідних даних мають бінарні значення. Головними в алгоритмі є властивості логістичної функції. Логістична функція $\sigma(z)$ визначається за наступною формулою:

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (2.2)$$

Вхідними даними логістичної функції є дійсні значення у діапазоні від $-\infty$ до $+\infty$. При цьому вихідним значенням є значення ймовірності, яке відображається у діапазоні від 0 до 1. На рис. 2.2 зображено графік логістичної функції.

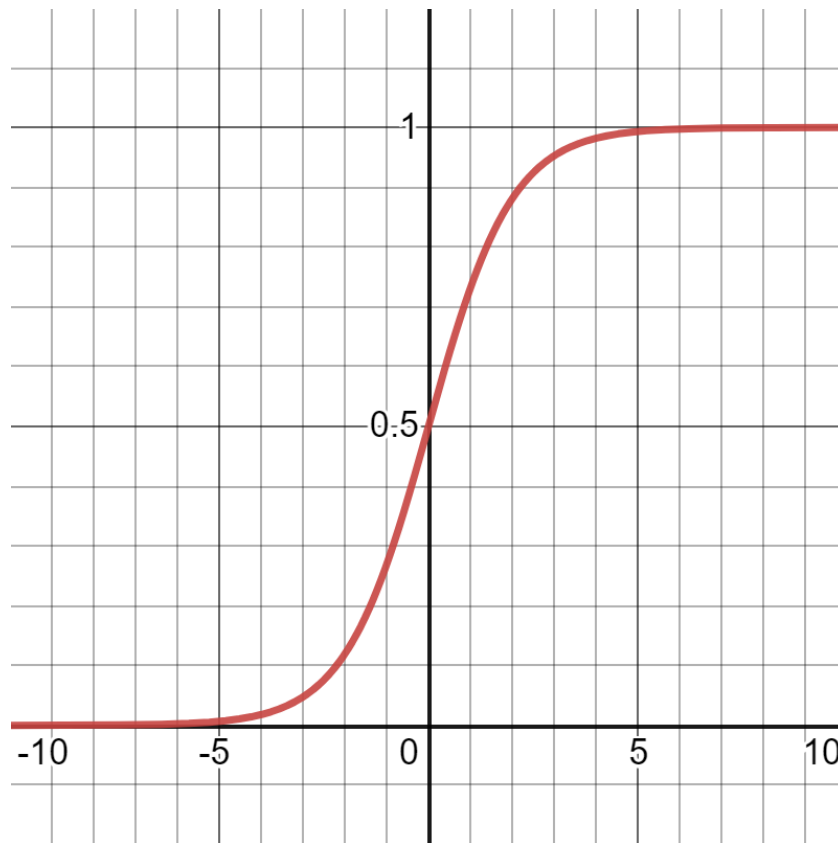


Рисунок 2.2 – Графік логістичної функції

Для прогнозування спортивних змагань модель, що базується на логістичній регресії буде складатися з вектора n показників матчу $x = (x_1 \dots x_n)$

і вектора $n+1$ параметрів моделі $\beta = (\beta_0, \beta_1 \dots \beta_n)$. Таким чином значення залежної змінної z рівняння логістичної регресії буде наступним:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \dots + \beta_n x_n \quad (2.3)$$

Навчання моделі відбувається шляхом оптимізації параметрів β так, щоб модель давала якнайкраще відтворення реальних результатів матчів для навчальної вибірки. Для навчання моделі, що базується на логістичній регресії, може застосовуватись метод максимальної правдоподібності.

Правдоподібність вибірки – це можливість отримання саме такої вибірки та саме таких спостережень або результатів, тобто добуток ймовірностей отримання кожного з результатів вибірки[14]. Метод максимальної правдоподібності – це метод оцінки невідомого параметра шляхом максимізації функції правдоподібності. Функція правдоподібності вибірки має вигляд:

$$P(\vec{y}|p) = \prod_{i=1}^N p^{y_i} (1-p)^{(1-y_i)} \quad (2.4)$$

При цьому p_i – це значення функції логістичного відгуку. Дана функція має вигляд:

$$p_+ = \frac{1}{1+e^{-w^T \vec{x}}} \quad (2.5)$$

Зробивши необхідні підстановки та провівши логарифмування, отримаємо що максимізація правдоподібності досягається шляхом мінімізації функції логістичних втрат. Вона дає міру похибки моделі при прогнозуванні результатів матчів, які були використані для навчання. Формула функції логістичних втрат має наступний вигляд:

$$L(p) = -\frac{1}{N} \sum_{i=1}^N p_i \log(y_i) + (1-p_i) \log(1-y_i) \quad (2.6)$$

де p_i – вірогідність прогнозу перемоги команди для матчу i ;

N – кількість матчів у вибірці;

y_i – реальний результат матчу i (при цьому 1 – перемога, 0 – поразка).

На рис. 2.3 зображено графік логістичних втрат, які виникають щодо одного матчу для різних прогнозованих ймовірностей за умови якщо матч

завершилася перемогою команди, яка і була переможцем у прогнозі. Функція досягає значення нулю при ймовірності $p = 1$.

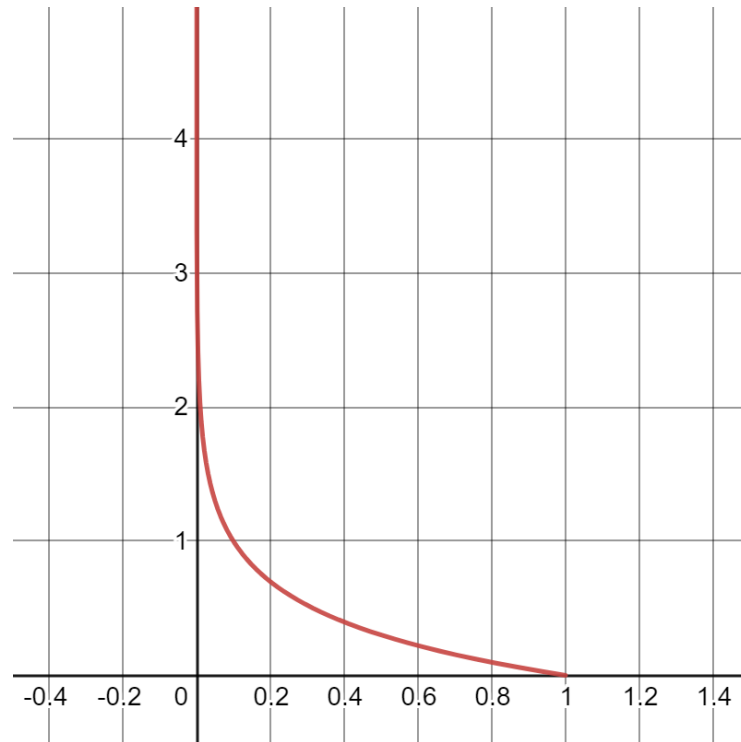


Рисунок 2.3 – Графік логістичних втрат

Після цього для знаходження оптимальних коефіцієнтів необхідно знайти похідну функції логістичної помилки, а далі використати один з методів оптимізації, наприклад метод градієнтного спуску або стохастичного градієнтного спуску.

Градієнтний спуск є методом знаходження локального мінімуму або максимуму функції за допомогою руху вздовж градієнту[15]. Він може використовуватись для мінімізації цільової функції $Q(w)$, де необхідно оцінити параметр w , який мінімізує $Q(w)$.

Стандартний метод градієнтного спуску на кожній ітерації присвоює параметру w значення: $w := w - \mu \sum_{i=1}^n \nabla Q_i(w)/n$. При цьому μ є значенням кроку або швидкістю навчання.

В деяких випадках обчислення градієнта стає значно розрахунково затратним коли функції, що сумуються мають складну формулу обчислення сум

градієнтів, тому що для того, щоб обчислити градієнт суми необхідно обчислити градієнти окремих членів суми. Для економії обчислювальних ресурсів в такому випадку доцільно використовувати метод стохастичного градієнтного спуску. Відмінність від стандартного методу полягає у тому, що істинний градієнт $Q(w)$ апроксимується градієнтом одного тренувального прикладу, тобто $w := w - \mu \nabla Q_i(w)$. Пробігаючи через навчальну множину, алгоритм виконує такий перерахунок для кожного тренувального прикладу [16].

Крім того, на кожному кроці можна обчислювати градієнт за більш ніж за одним тренувальним прикладом. Цей метод є своєрідним компромісом між обчисленням істинного градієнту та градієнту за одним тренувальним прикладом.

Також, при використанні методів градієнтного спуску можуть виникнути проблеми з пошуком абсолютного мінімуму. Існує ймовірність, що під час градієнтного спуску ваги можуть потрапити в локальний мінімум. Крім того, можливий варіант, що ваги перестрибнуть мінімум. Одним із способів для вирішення даної проблеми є використання різних значень для кроків при підборі значень для коефіцієнтів. При занадто малій швидкості навчання метод градієнтного спуску потребує багато часу щоб зійтись, при занадто великій – пропускає глобальні мінімуми та може розходитись.

Моделі, що базуються на алгоритмі логістичної регресії користуються значною популярністю при розв'язанні задач класифікації, зокрема прогнозуванні результатів спортивних змагань. Однією з переваг даних моделей є легка та швидка їх побудова. Крім того, після підбору коефіцієнтів можна інтуїтивно зрозуміти який напрямок та величина зв'язку між кожною залежною змінною показників та змінною результату. У якості вихідного результату така модель даватиме значення ймовірності настання події від 0 до 1. Це значення вже готовий результат прогнозування, оскільки воно не потребує подальших перетворень, і за необхідністю його можна перевести у відсотки.

Однак, використання моделей такого типу може мати і недоліки. У випадку сильної кореляції вхідних змінних між собою, модель на основі логістичної регресії буде видавати на виході результат низької точності.

2.1.4 Штучні нейронні мережі

Штучна нейронна мережа є системою взаємозв'язаних нейронів, що створена за моделлю біологічних нейронів.

Існує декілька способів графічного зображення нейронних мереж та нейронів. Штучна нейронна мережа може бути представлена у вигляді графу. Як правило, нейронні мережі мають декілька шарів. При цьому кожний нейрон у невхідному шарі пов'язаний з усіма нейронами минулого шару.

На рис. 2.4 зображена тришарова нейронна мережа. Усі зображені стрілки напрямлені зліва направо, тобто сигнал йде тільки від вхідного шару до вихідного. В оберненому напрямку сигнал не розповсюджується. Така модель нейронної мережі називається мережею прямого розповсюдження. Мережі даного типу часто застосовуються у задачах прогнозування, кластеризації та розпізнавання[17].

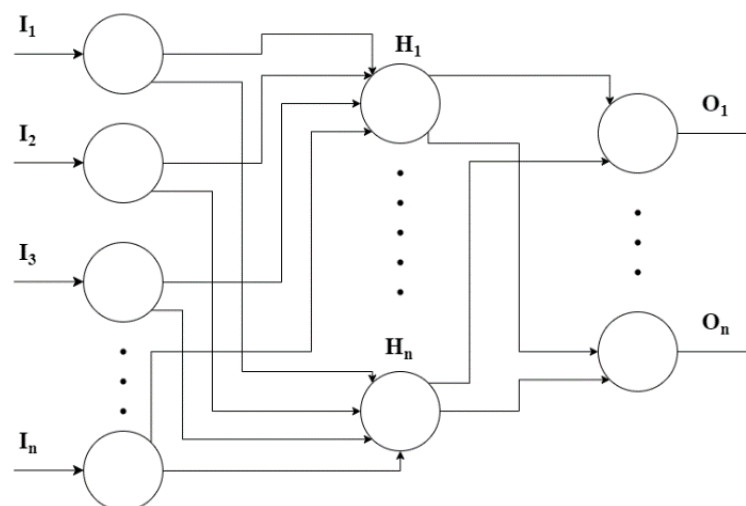


Рисунок 2.4 – Тришарова нейронна мережа

Шар нейронів, що розташований зліва, називається вхідним шаром, а його нейрони – вхідними нейронами. Шар нейронів, який знаходиться справа

називається вихідним шаром, а його нейрони – вихідними відповідно. Середній шар називається прихованим, оскільки не є вхідним чи вихідним. У нейронних мережах може бути багато прихованих шарів, проте сукупність прихованих шарів завжди можна представити у вигляді одного прихованого шару.

Кожному зв'язку в штучній нейронній мережі присвоюється своя вага. Сигнали на вході нейронів множаться на свої ваги. Добутки вхідних сигналів і коефіцієнтів сумуються, формуючи вихідний сигнал, який описується формулою:

$$f(x) = K(\sum_i w_i x_i) \quad (2.7)$$

де w_i – значення ваги вхідних даних;

K – функція активації.

Функція активації використовується для того, щоб обробити суму та сформувати адекватний вихідний сигнал. Зазвичай для даного завдання застосовуються сигмоїдальні функції, наприклад, логістична функція, яка вже була описана вище. Сигмоїдальні функції доцільні для використання у нейронних мережах, оскільки вони забезпечують гладкість результату. Це означає, що невеликі зміни до коефіцієнтів мережі зумовлять невелику зміну вихідного результату. Також, сигмоїдальні функції значно спрощують обчислення похідних завдяки властивостям експоненти при диференціюванні. Крім того, може використовуватись функція гіперболічного тангенсу. Вона дозволяє отримати на виході значення різних знаків від -1 до 1, що може бути корисним при вирішенні деяких задач з використанням нейронних мереж у випадку якщо вихідний результат може бути менше нуля.

Результати спортивних подій за допомогою штучних нейронних мереж можна прогнозувати шляхом подачі на вхідний шар статистичні показники перед матчем, проводячи значення через мережу. У випадку використання логістичної функції в якості функції активації, вихідним значенням мережі може бути вірогідність перемоги команди у матчі.

Для оптимізації коефіцієнтів штучної нейронної мережі існує досить багато алгоритмів. Одним із найбільш популярних вважається алгоритм зворотнього розповсюдження. Послідовність дій, що у ньому виконуються є наступною.

У ході першого етапу вхідний нейрон отримує сигнал і транслює його кожному із прихованих нейронів. Кожен прихований нейрон обчислює функцію активації та розсилає свій сигнал усім вихідним нейронам. Аналогічно, кожен вихідний нейрон також обчислює свій вихідний сигнал за допомогою функції активації. На виході обчислене значення порівнюється із значенням, що надане вчителем, і на основі цього вираховується відповідне значення помилки для даного вхідного шаблону.

Далі базуючись на помилці, вираховується складова коректування коефіцієнтів, що використовується при розповсюдженні помилки до усіх елементів мережі минулого шару. За таким же принципом складова коректування коефіцієнтів обчислюється для кожного прихованого нейрону. Після того як усі складові коректування були визначені, відбувається одночасне коректування коефіцієнтів усіх зв'язків [18].

Перевагами використання штучних нейронних мереж є те, що вони можуть давати результати досить високої точності. Однак, для оптимізації моделі на основі нейронної мережі необхідно досить велика кількість даних. Крім того, нейронні мережі важко піддаються інтерпретації. Зазвичай існує занадто багато шарів та нейронів, щоб зрозуміти напрямок та величину зв'язку кожної вхідної змінної з вихідною за допомогою різних коефіцієнтів. Також у нейронних мереж присутня ймовірність потрапляння у локальний мінімум при пошуку оптимальних значень для коефіцієнтів.

2.1.5 Наївний баєсів класифікатор

Ймовірнісні класифікатори є методами машинного навчання, що для заданих вхідних даних здатні передбачити розподіл ймовірності над множиною класів. Одним із найбільш відомих ймовірнісних класифікаторів є наївний баєсів класифікатор. Він базується на так званому наївному припущенні, яке полягає у тому, що усі характеристики є незалежними одна від одної[19]. Даний класифікатор використовує теорему Баєса, яка обчислює умовні ймовірності:

$$P(A|B) = \frac{P(B|A)*P(A)}{P(B)} \quad (2.8)$$

де $P(A|B)$ – як часто відбувається подія A при настанні події B ;

$P(B|A)$ – як часто відбувається подія B при настанні події A ;

$P(A)$ – ймовірність, що A не залежить від інших подій;

$P(B)$ – ймовірність, що B не залежить від інших подій.

Нехай y є цільовим класом. У випадку прогнозування результатів спортивних змагань це буде клас, що визначає перемогу певної команди у матчі, а x_i – фактори, які визначають результат. Тоді, теорему Баєса можна представити у наступному вигляді:

$$P(y|x_1, \dots, x_n) = \frac{P(x_1, \dots, x_n|y)*P(y)}{P(x_1, \dots, x_n)} \quad (2.9)$$

В формулі, що зазначена вище, для розрахунку класифікації, використовується тільки чисельник, оскільки знаменник є константою та є незалежним від мітки класів. Оскільки наївний баєсів класифікатор використовує припущення, що усі ознаки є незалежними одна від одної, то $P(x_1, \dots, x_n|y)$ можна обчислити як добуток ймовірностей зустріти ознаку x_i серед об'єктів класу y . Тоді формула набуває наступного вигляду:

$$P(X|y) \propto P(y) \prod_{i=1}^n P(x_i|y) \quad (2.10)$$

Апостеріорна ймовірність $P(y|X)$ обчислюється для кожного класу. Клас з найбільшою апостеріорною ймовірністю і буде прогнозованим результатом.

Тому треба знайти змінну класу y з максимальною ймовірністю, що можна виразити через формулу:

$$y = \operatorname{argmax}_y P(y) \prod_{i=1}^n P(x_i|y) \quad (2.11)$$

Наївний Баєсів класифікатор поділяється на три типи: мультиноміальний, Гауса та Бернуллі[20]. В мультиноміальному класифікаторі вектори ознак є значеннями частотності, тобто вони визначають частоту з якою генеруються певні події шляхом мультиноміального розподілу. В моделі Бернуллі характеристики є незалежними бінарними змінними, які описують вхідні дані. У класифікаторі типу Гауса значення всіх характеристик мають розподіл Гауса, тобто нормальний розподіл. Графіком такого розподілу є крива, яка схожа на пагорб, і є симетричною відносно середнього значення характеристик.

Перевагою використання наївних баєсівських класифікаторів є те, що вони можуть працювати з великими обсягами даних, оскільки час навчання зростає приблизно лінійно. Крім того, алгоритм має помірні вимоги до пам'яті. Однак, незважаючи на те, що класифікатор може бути достатньо надійним для того, щоб ігнорувати наївне припущення, передбачені ймовірності бувають низької точності. Ще одним з недоліків наївного баєсового класифікатора є так звана проблема нульової частоти. Вона виникає коли значення параметру та клас не зустрічаються разом. Тоді оцінка ймовірності буде дорівнювати нулю. Одним із можливих підходів в такому випадку буде додавання до кожної комбінації класу та значення параметру одиниці, якщо значення параметру є нульовим. Проте, такий прийом знижує точність отриманого результату.

2.1.6 Метод опорних векторів

Машини опорних векторів є потужними та гнучкими алгоритмами машинного навчання з вчителем. Вони використовуються для задач як класифікації, так і регресії. Вперше були представлені у 1960-х роках, пізніше у 1990-х були удосконалені[21].

Основним завданням, що вирішує метод опорних векторів, є знаходження найбільш оптимальної гіперплощини, що правильно класифікує приклади шляхом максимально можливого розмежування точок двох класів на категорії.

Спочатку алгоритм тренується на об'єктах з навчальної вибірки, для якої вже відомі мітки класів для кожного об'єкту. Після цього вже навчений алгоритм прогнозує мітку класу для об'єктів цільової вибірки.

У задачах бінарної класифікації мітки класів можуть приймати значення $Y = \{-1, +1\}$. Об'єктом є вектор з N ознаками $x = (x_1 \dots x_n)$ у N розмірному просторі R^n . У ході навчання алгоритм має побудувати функції $F(x) = y$, аргументом x якої є об'єкт з простору R^n , і видати мітку класу y .

Головна ціль методу опорних векторів як класифікатора – знайти рівняння гіперплощини $w_1x_1 + w_2x_2 + \dots + w_nx_n + w_0 = 0$ у просторі R^n , яке розділило б два класи деяким оптимальним способом. Після налаштування коефіцієнтів алгоритму w , всі об'єкти, що знаходяться по одну сторону від побудованої гіперплощини, будуть прогнозуватися як об'єкти першого класу, а об'єкти, що знаходяться на іншій стороні – як другий клас.

Результуюча гіперплощина може будуватись різноманітними способами. Площини генеруються ітеративно, далі вибирається площина, що розмежовує класи найбільш правильно. В методі опорних векторів ваги w підбираються таким чином, щоб об'єкти класів знаходились якомога далі від розмежовуючої гіперплощини. Отже, алгоритм максимізує проміжок між гіперплощиною та об'єктами класів, які розташовані найближче до гіперплощини. Такі об'єкти називають опорними векторами, звідки походить і назва для алгоритму[22]. Розмежувальна площина буде визначатись за допомогою цих точок даних. Відстань від опорних векторів до гіперплощини буде дорівнювати одиниці поділений на норму w . Отже, в процесі підбору ваги відбувається максимізація значення величини $\frac{1}{\|w\|}$. На рис. 2.5 зображений приклад вибору оптимальної гіперплощини в R^2 .

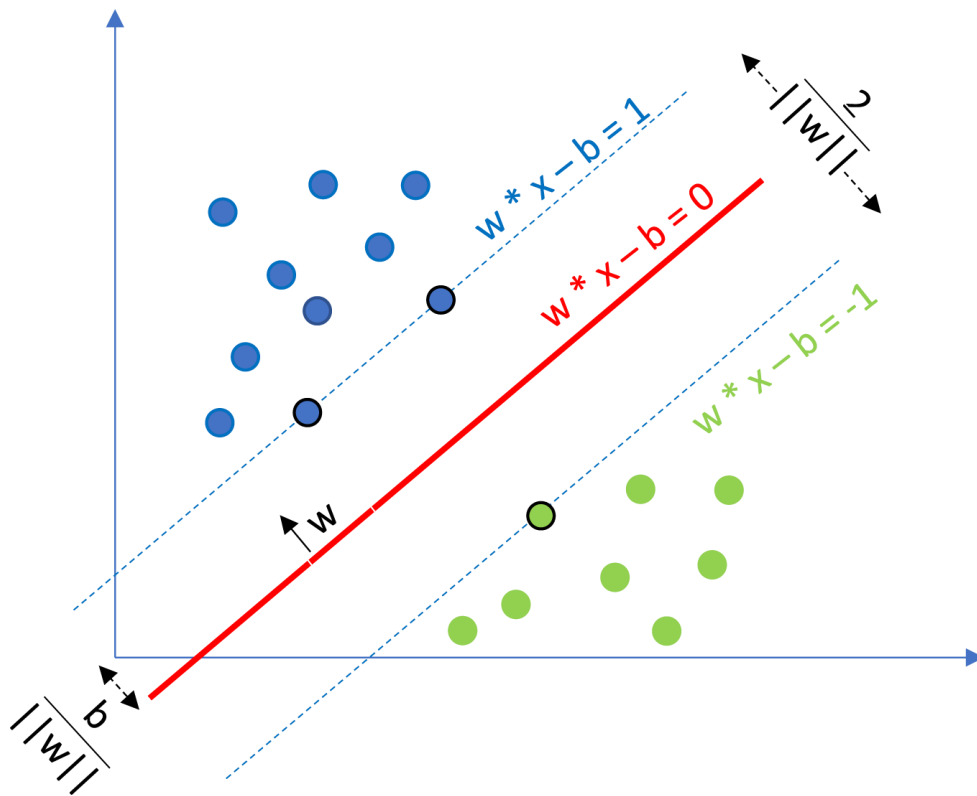


Рисунок 2.5 – Принцип роботи методу опорних векторів

При використанні алгоритмів на основі методів опорних векторів для прогнозування результатів спортивних змагань, як і в інших алгоритмах категорії класів можуть бути перемога або поразка у матчі. Новий приклад, тобто матч, що незабаром відбудеться, буде проектуватись на простір точок і класифікуватись на основі того, з якої сторони гіперплощини він знаходиться.

Перевагами алгоритмів на основі методу опорних векторів є те, що вони забезпечують досить високу точність, часто даючи прогнозування більш точні ніж нейронні мережі. Вони ефективні у багаторозмірних просторах, потребують невеликих затрат пам'яті.

Однак на навчання моделі на основі методу опорних векторів необхідно набагато більше пам'яті, ніж, наприклад, для нейронних мереж. Крім того, моделі досить важко налаштовуються. Метод опорних векторів є нестійким до шуму. Викиди у вихідних даних мають вплив при створенні результуючої гіперплощини.

2.1.7 Древа прийняття рішень

Дерево прийняття рішень – це метод аналізу даних у машинному навчанні, який є засобом підтримки прийняття рішень. Структура даного методу є деревовидною і являє собою множину листів та гілок. В листах дерева знаходяться значення цільової функції, а в інших вузлах розташовані умови переходу виду «Якщо, то...», які визначають ребро, по якому відбувається перехід.

Умови переходу або правила генеруються завдяки множині навчальних прикладів на основі яких будується модель. Моделі на основі дерев прийняття рішень будуються за допомогою навчання з учителем. Тому, в навчальній множині для прикладів повинно бути задане цільове значення. Розрізняють два основні типи дерев:

- дерева класифікації;
- дерева регресії.

Як вже було зазначено вище, структура дерев прийняття рішень передбачає два види елементів: вузли та листя. Вузли містять правила прийняття рішення та виконують перевірку чи відповідає об'єкт певному критерію. Після проходження об'єктами перевірки в вузлі, вони розбиваються на дві підмножини: ті, що задовольняють встановлений критерій (правило); ті, що не задовольняють його[23]. Далі до кожної підмножини знову застосовується певне правило. Така операція відбувається поки результат не буде задовольняти умови зупинки алгоритму. Лист містить підмножину усіх об'єктів, які задовольняють усім критеріям гілки, яка закінчується даним листом. Об'єкт може потрапити тільки в один листок, забезпечуючи принцип, що рішення є тільки одним. На рис. 2.6 зображено приклад дерева прийняття рішень.

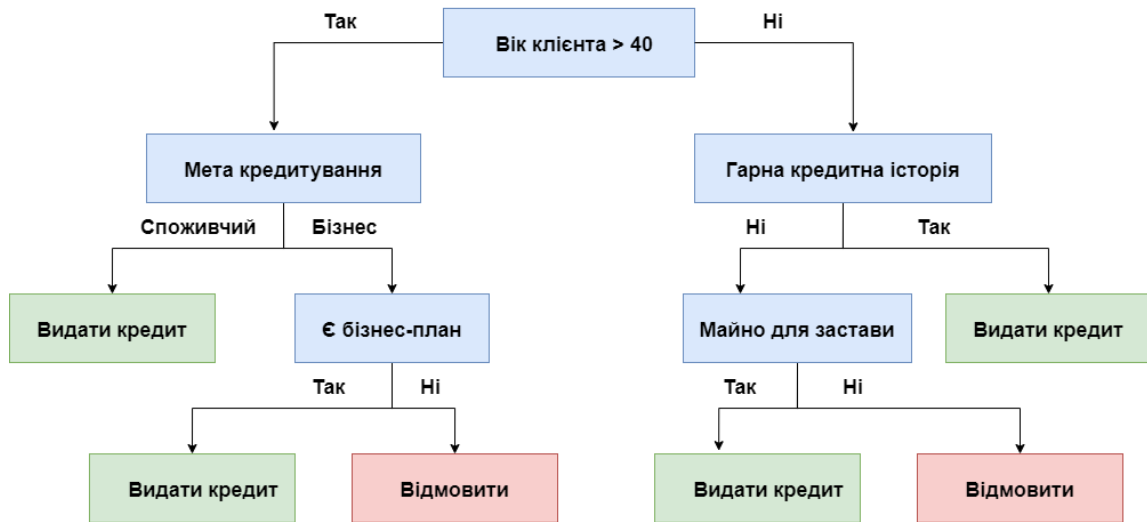


Рисунок 2.6 – Приклад дерева прийняття рішень

При вирішенні задачі класифікації листя дерева містить результуючі класи. Якщо розглядається задача регресії, то листя містить певне значення цільової функції. Для досягнення хороших результатів регресії зазвичай потрібно дерево великого розміру.

При побудові дерева рішень необхідно розділити навчальну вибірку на підмножини з використанням правил для прийняття рішення у вузлах. Це відбувається поки усі вузли в кінці гілок не перетворяться на листи. Вузол стає листом у наступних випадках: коли він містить єдиний об'єкт або об'єкт тільки одного класу; якщо було задоволено критерій зупинки алгоритму, наприклад, досягнута максимальна глибина дерева.

Якщо алгоритм буде працювати до отримання чистих підмножин з прикладами з одного класу, то існує висока ймовірність отримати дерево, в якому для кожного прикладу буде створений свій окремий лист, а отже свій шлях у дереві. Це призведе до перенавчання дерева, і дерево буде точно працювати із навчальними даними, проте не зможе давати задовільні результати на нових прикладах. Для уникнення перенавчання застосовують різноманітні підходи примусової зупинки дерева. Використовуються такі методи, як: рання зупинка – переривання роботи алгоритму при досягненні певного критерію, наприклад,

кількості правильно розпізнаних прикладів; обмеження глибини дерева – встановлення максимального числа розбиття у гілках; задання мінімально допустимого числа прикладів у вузлі – заборона алгоритму створювати вузли, де міститься кількість прикладів, що менша від заданого числа.

Побудова дерева часто базується на так званих жадібних алгоритмах, які приймають локально-оптимальні рішення, що в результаті дає оптимальне кінцеве рішення.

Однією із популярних модифікацій класичного дерева прийняття рішень є випадковий ліс. Випадкові ліси працюють шляхом створення великої кількості дерев рішень під час навчання, беручи різну частину набору даних як навчальний набір для кожного дерева. У звичайному дереві при розділенні вузла розглядається кожна можлива ознака, і обирається та, що сильніше ділить значення у вузлах. У випадковому лісі кожне дерево може обирати тільки з випадкової підмножини об'єктів. Це призводить до більшого різноманіття та до більш слабкої кореляції дерев між собою. При використанні випадкового лісу досягається досить висока точність результатів, особливо для даних, що мають велику кількість змінних.

Загалом, однією з головних переваг дерев рішень є те, що вони без труднощів інтерпретуються. Дерево легко візуалізується, і можна відслідкувати результат для певного окремого об'єкта. Дерева зручні у використанні як для задач класифікації, так і для задач регресії. Крім того, навчання моделі не займає багато часу, і вона не потребує багато параметрів.

Серед недоліків моделей на основі дерев прийняття рішень є досить значна чутливість до шуму у вхідних даних. Дерева рішень схильні до перенавчання, тому існує необхідність застосовувати методи відсічення гілок та встановлення максимальної глибини дерева. Пошук оптимального дерева часто буває складним, що зумовлює застосування різних евристик, які в результаті не гарантують знаходження оптимального дерева.

2.1.8 Вибір методів прогнозування

Розроблюване програмне забезпечення з аналізу та прогнозування результатів спортивних змагань у якості вихідних даних надаватиме ймовірність настання певного результату матчу, що буде виражена в якості відсоткового числа такої ймовірності. Тому, доцільно при розробці програмної системи використати алгоритми, що надаватимуть числове значення ймовірності настання певної події.

Для реалізації програмного забезпечення прогнозування результатів спортивних змагань вирішено обрати алгоритми лінійної та логістичної регресії. Моделі на основі обох цих алгоритмів досить легко будуються та інтерпретуються, що може бути важливим для розуміння значення впливу кожного показника. Модель, що базується на логістичній регресії є популярною для вирішення завдань прогнозування спортивних змагань, зокрема такі моделі були успішно побудовані для тенісу та американського футболу, і видавали результати задовільної точності. Рівняння алгоритму логістичної регресії є сигмоїдою, область значень якої від 0 до 1, що вже є результатом прогнозування і не потребує подальших перетворень.

2.2 Визначення структури програмного забезпечення

Архітектура програмного забезпечення є сукупністю рішень, які визначають побудову програмної системи. При створенні архітектури ПЗ відбувається вибір її структурних елементів, погоджується їх поведінка та комунікація між собою. Наразі існує багато типів архітектури програмного забезпечення, кожен з яких має свої переваги та недоліки. Розглянемо найбільш популярні підходи.

Багаторівнева архітектура є однією з найбільш розповсюджених архітектур. Вона побудована на принципі розмежування відповідальності. У такому випадку програмна система поділяється на декілька шарів, кожен з яких

має своє певне коло обов'язків, а також не має даних про структуру інших рівнів. Архітектура не передбачає строго визначену кількість рівнів, проте найчастіше використовується архітектура з трьома наступними рівнями:

- рівень представлення, який містить користувацький інтерфейс, і відповідно відповідає за забезпечення взаємодії користувача з системою;
- рівень логіки, що забезпечує бізнес-логіку додатку, і відповідальний за її зміну без впливу на інші рівні;
- рівень передачі даних, який взаємодіє зі сховищами даних.

Головною перевагою багаторівневої архітектури є те, що кожен рівень виконує певну обмежену множину функцій, а отже зміст шарів можна змінювати практично без ризику конфліктів між шарами. З недоліків – низька швидкість роботи, складність системи для знаходження помилок[24].

Подійно-орієнтована архітектура – архітектурний підхід, що полягає у визначенні поведінки системи при настанні якої-небудь події, яка впливає на стан система. Подійно-орієнтована архітектура зазвичай містить два компоненти: споживачі подій та джерела подій. При цьому події поділяються на два види: ініціалізуючі події, та події на які реагують споживачі. Перевагою такого типу архітектури є легке масштабування і забезпечення асинхронної моделі виконання. Недоліком є складність усунення помилок, оскільки одна подія може спричинити запуск цілих ланцюгів різноманітних дій.

Сервісно-орієнтована архітектура – забезпечує модульний підхід до розробки програмного забезпечення, який базується на використанні слабо зв'язаних, розподілених компонентів, які мають стандартизовані інтерфейси для взаємодії за стандартизованими протоколами[25]. Зазвичай програмні системи, які реалізують сервісно-орієнтований підхід, являють собою сукупність веб-служб, які взаємодіють за протоколом SOAP, REST або іншими. В сервісно-орієнтованій архітектурі інтерфейси компонентів інкапсулюють деталі реалізації від інших компонентів, забезпечуючи комбінування та багатократне

використання компонентів при створенні складних розподілених програмних комплексів. Така архітектура не визначається однією певною технологією, і забезпечує незалежність від платформ та засобів розробки, що сприяє покращенню масштабованості системи. Тобто сервісно-орієнтовані системи можуть мати, наприклад, сервіси розроблені на мовах програмування C# та Java, які будуть вільно взаємодіяти між собою. Отже, перевагами сервісно-орієнтованого підходу є слабка зв'язність системи, висока гнучкість та масштабованість. Недоліки полягають у підвищеному ризику збою при обміні даними між сервісами. Крім того, реалізація такого підходу є досить тривалою у часі, а також існує потреба в комплексному тестуванні програмної системи в розподіленому середовищі.

Для розробки програмної системи прогнозування результатів спортивних змагань вирішено обрати сервісно-орієнтовану архітектуру. Передбачається, що система буде використовуватись у корпоративних цілях, а сервісно-орієнтована архітектура підходить для складних систем масштабу підприємства, для яких зазвичай необхідна інтеграція з багатьма різноманітними додатками та сервісами. У ході проєктування системи прийнято рішення, що система складатиметься з наступних модулів:

- сервіс управління, що періодично запускатиме збір нової спортивної статистики, розрахунок показників та прогнозів;
- сервіс збору статистики, що підключатиметься до веб-порталу спортивної статистики та зберігатиме дані про результати та майбутні матчі до сховища;
- сервіс прогнозування, який використовуючи обрані алгоритми, бути безпосередньо розраховувати прогноз;
- сервіс аналізу даних, що буде будувати статистичні показники, які необхідні для роботи алгоритмів прогнозування.

Крім того, ще одним елементом системи буде система управління взаємовідносинами з клієнтами, через яку користувач і буде взаємодіяти з усім програмним модулем. Також, для CRM планується розробити спеціальне інтерфейсне рішення, що міститиме створені нові сутності, форми для вводу даних, представлення для перегляду записів сутності.

Система управління взаємовідносинами з клієнтами або CRM – це програмне забезпечення, за допомогою якого бізнес взаємодіє з клієнтами або споживачами. Ринок CRM систем кожен рік зростає, і такі системи впроваджують найрізноманітніші компанії: від маленьких компаній, де працюють до десяти співробітників до великих фірм з сотнями працівників. Основними завданнями, які виконують CRM системи є:

- управління продажами. Такі системи дають можливість автоматизувати діяльність виконання замовлень. CRM система відслідковує нові заявки, створює супроводжуючі документи, показує менеджерам компанії подальші кроки щодо роботи з клієнтами;

- управління взаємодіями. Збереження інформації про клієнтів, їх контактів, дані про товари які були придбані;

- керування маркетингом. Можливість автоматизувати дослідження ринку, аналізувати отримані дані, наповнювати базу даних потенційних покупців, виявляти їх потреби та формувати маркетингову стратегію;

- аналіз даних та звітність. Дозволяє, наприклад, оцінювати ефективність відділу продажів завдяки побудові інформативних графіків та діаграм;

- управління обслуговуванням. Автоматизація оброблень звернень клієнтів;

- завдання інтеграції. Об'єднання CRM системи з телефонією, електронною поштою та соцмережами, що дає можливість підвищити ефективність роботи з клієнтами та постачальниками[26].

Висновки до розділу

У даному розділі було описано сучасні існуючі алгоритми та методи для прогнозування результатів спортивних змагань, для реалізації обрано алгоритми лінійної та логістичної регресії. Розглянуто популярні архітектури, що використовуються для розробки програмних систем. Визначено, що система буде створюватись на основі сервісно-орієнтованої архітектури з інтеграцією з системою управління взаємовідносинами з клієнтами.

3 ОПИС ПРОГРАМНОГО ТА ТЕХНІЧНОГО ЗАБЕЗПЕЧЕННЯ

3.1 Технології для розробки

При розробці веб-сервісів програмного модулю було застосовано мову програмування C# та фреймворк ASP.NET. В якості системи управління взаємовідносинами з клієнтами використовувалась Microsoft Dynamics.

C# – одна з найбільш популярних мов програмування, яка є об'єктно-орієнтованою мовою, що має строгу типізацію. Ця мова є частиною сімейства мов C, тому її синтаксис схожий на мови C, C++, JavaScript, Java. Для забезпечення об'єктно-орієнтованої концепції, мова C# надає відповідні мовні конструкції для застосування при створенні програмного забезпечення. Використання мови C# дає змогу створювати стійкі та надійні програмні системи. Мова C# практично універсальна. Її можна використовувати для розроблення найрізноманітнішого програмного забезпечення: десктопних додатків для операційної системи Windows, веб-додатків, мобільних програм для операційних систем Android та iOS. Синтаксис цієї мови програмування інтуїтивно зрозумілий, що дає можливість за короткий термін часу створювати сучасні продуктивні програми. Програми C# виконуються в .NET, яка є віртуальною системою виконання, що викликає загальномовне середовище виконання CLR та набір бібліотек класів[27].

ASP.NET є безкоштовною інтернет платформою для того, щоб створювати багатофункціональні веб-сайти та веб застосунки, використовуючи багатомовне середовище .NET[28]. На даний момент одним із найбільш популярних типів програмних моделей ASP.NET додатків є Web API. Web API застосунки являють собою HTTP служби для широкого діапазону клієнтів: мобільні додатки, браузері, інші веб-служби. Такі служби працюють за принципом REST (Representation State Transfer), при цьому забезпечується обробка HTTP запитів GET, POST, PUT, DELETE. ASP.NET забезпечує продуктивну роботу веб-

застосунків на сервері, підтримує декілька способів автентифікації та використовує модель безпеки .NET.

Microsoft Dynamics 365 – сукупність продуктів компанії Microsoft для управління взаємовідносинами з клієнтами, а також для планування ресурсів підприємства. Вона складається з багатьох підмодулів, з яких задачі управління взаємовідносинами з клієнтами вирішують Dynamics 365 for Customer Engagement та Dynamics 365 for Sales. Часто ці продукти і називають відповідно Microsoft Dynamics CRM. Microsoft Dynamics входить до числа топ 5 CRM систем у світі, кількість клієнтів щороку зростає. Існує два види версій цієї CRM: online та on-premise. On-premise версія системи розгортається на власному існуючому сервері. Це дає можливість більш індивідуального налаштування системи, а також повний контроль над адмініструванням системи. Online версія працює у хмарному середовищі. Вона дозволяє користуватися системою управління взаємовідносинами з клієнтами без використання власної ІТ інфраструктури. Хмарна версія системи постійно отримує найновіші оновлення, її легше налаштувати для початку роботи і в подальшому підтримувати працездатність, вона підтримує інтеграцію з хмарними з іншими продуктами Microsoft, наприклад, Office 365, Power Automate. Microsoft Dynamics для збереження даних використовує СУБД Microsoft SQL Server. В організаційній структурі Microsoft Dynamics CRM є три головні одиниці: підрозділи, користувачі, команди. Підрозділ може містити декілька команд, команда може містити декілька користувачів. Крім того, кожен підрозділ може мати батьківський підрозділ. На основі структури підрозділів в CRM системи визначаються принципи базового, локального та глобального доступу до даних та внесення змін до системи. Microsoft Dynamics забезпечує автоматизацію трьох основних напрямів взаємовідносин з клієнтами: продажі, маркетинг та обслуговування клієнтів. Автоматизація продажів відбувається завдяки, створенню профілів з інформацією про замовника, списками зустрічей та

переговорів по кожному клієнту, оброблення замовлень. Модуль маркетингу дає можливість створення груп клієнтів за певною тематикою, розроблювати маркетингові стратегії, відправляти новини, індивідуальні пропозиції. Модуль управління обслуговування клієнтів дає змогу автоматизувати процес вирішення певної проблеми, яка виникла у клієнта. Він може аналізувати запитувану інформацію, розподіляти співробітників, що мають відповідну компетенцію, для вирішення проблем у зверненнях, при цьому враховуючи часові ресурси[29].

Microsoft Dynamics CRM є популярною зокрема завдяки тому, що її базовий функціонал можна розширювати для досягнення власних потреб. Кореневим поняттям розширення функціоналу в Microsoft Dynamics CRM є рішення або Solution. Рішення являє собою пакет, який містить повний або частковий стан певного об'єкту, що був доданий у рішення. Рішення може бути експортовано, а потім імпортовано в іншу систему. У такий спосіб можна переносити розширення функціоналу з одного екземпляру CRM в інший.

Одним із видів елементів, що може містити рішення є сутності. Сутності визначають тип запису, і є аналогами таблиць у базі даних. Екземплярами сутності є записи, тобто в базі даних це горизонтальний рядок у таблиці. Для кожної сутності в системі можна створювати нові поля, форми, представлення. Поля уособлюють атрибути таблиці в базі даних і можуть бути різних типів: числовий, текстовий, дата і час, булевий тощо. Деякі поля, наприклад, поле дати створення запису, створюються автоматично при створенні сутності. Форми використовуються для перегляду даних запису, для створення нового запису або для відображення даних запису на формі іншого запису. На формах відповідно розміщуються поля сутності. Представлення використовуються для відображення списків записів. При створенні представлення можна налаштувати поля сутності, які будуть відображатися, а також налаштувати фільтр – критерій по певному полю або декількома полями для того щоб відображались тільки потрібні записи.

Щоб змінювати вид та поведінку користувацького інтерфейсу в Microsoft Dynamics CRM використовується мова програмування JavaScript. JavaScript підтримує декілька парадигм програмування, використовується для розробки додатків та забезпечення інтерактивності веб-сторінок. Ця мова програмування є такою, що інтерпретується. Код запускається зверху вниз. В JavaScript використовується динамічна типізація, тому змінна може зберігати різні типи даних. Як вже було зазначено вище, JavaScript дозволяє писати як клієнтський, так і серверний код, проте в Microsoft Dynamics CRM код JavaScript є клієнтським. В налаштуваннях форм сутності можна підключати скрипти, які потім будуть виконуватись при завантаженні запису сутності у браузері[31]. Це дозволяє, наприклад, в залежності від бізнес-логіки, приховати певні поля на формі або зробити їх недоступними для редагування.

Для використання алгоритмів машинного навчання при реалізації програмного забезпечення прогнозування результатів спортивних змагань була застосована бібліотека ML.NET. Вона дозволяє додавати в додатки розроблені на платформі .NET можливості машинного навчання. Основою ML.NET є модель машинного навчання. Ця модель визначає кроки, які необхідно виконати, щоб отримати прогнози, базуючись на вхідних даних. За допомогою ML.NET можна навчити користувацьку модель, вказавши відповідний алгоритм. Існує можливість імпортувати вже навчені моделі[30]. Бібліотека ML.NET дозволяє отримувати прогнози наступних типів:

- класифікація;
- регресія;
- вияв аномалій;
- класифікація зображень;
- рекомендації.

Також, ML.NET дає можливість оцінити модель за новими тестовими даними. Кожен тип задач машинного навчання має власні метрики, за якими

можна оцінити точність моделі. Навчені моделі зберігаються на комп'ютері, і їх можна завантажити у потрібний момент для використання в інших застосунках.

3.2 Опис архітектури програмного продукту

Програмний модуль складається з чотирьох ASP.NET веб-сервісів:

- сервіс управління SportsPredictions.Control;
- сервіс збору статистики SportsPredictions.StatisticsCollector;
- сервіс прогнозування SportsPredictions.Predictor;
- сервіс аналізу даних SportsPredictions.DataAnalyzer.

SportsPredictions.Control є сервісом для періодичного оновлення даних для спортивного прогнозування. При розгортанні та запуску цього сервісу, починає працювати внутрішній таймер, що має інтервал в одну добу. Головна функція, що працює періодично, виконує по чергово запити на сервіс збору статистики, сервіс прогнозування та сервіс аналізу даних.

SportsPredictions.StatisticsCollector завантажує веб-сторінки веб-сайту спортивної статистики. Виконуючи парсинг розмітки, сервіс отримує дані про результати матчів, що відбулися за останній час та отримує дані про нові матчі, що відбудуться у майбутньому. Нові дані зберігаються до сховища.

SportsPredictions.DataAnalyzer використовується для побудови параметрів по кожній спортивній команді на певну дату. Для цього аналізуються минулі матчі, у яких команда прийняла участь, та створюється і підраховується необхідна кількість параметрів, задіяних у моделях аналізу. Основні параметри, які підраховує сервіс: відсоток перемог, поразок, кількість забитих і пропущених голів, рейтинг атаки та захисту команди, місце в турнірній таблиці, показник домашньої гри. Окремо підраховуються параметри для сезону та останніх 10 ігор та останні 3 ігри. Крім того, сервіс аналізує з яким наступним суперником буде грати команда, і на основі цієї інформації будується параметр особистих зустрічей.

SportsPredictions.Predictor є сервісом, що безпосередньо виконує прогнозування результатів. Для цього він завантажує попередньо навчені моделі лінійної та логістичної регресії, отримує інформацію зі сховища про параметри обох команд, що беруть участь у матчі. Використовуючи навчені моделі машинного навчання, для кожного нового матчу вираховується відсоткова ймовірність перемоги для кожної команди. Для того, щоб отримувати прогнози інших результатів, наприклад, кількість забитих голів у матчі, або перевищення середньої кількості голів (Over/Under) у матчі, потрібно створити окрему модель і виконати процес навчання з відповідними даними. Отримані прогнози зберігаються у сховище. На рис. 3.1 зображена узагальнена схема архітектурного рішення.



Рисунок 3.1 – Узагальнена схема архітектурного рішення

На рис. 3.2 представлена узагальнена архітектура програмного забезпечення.

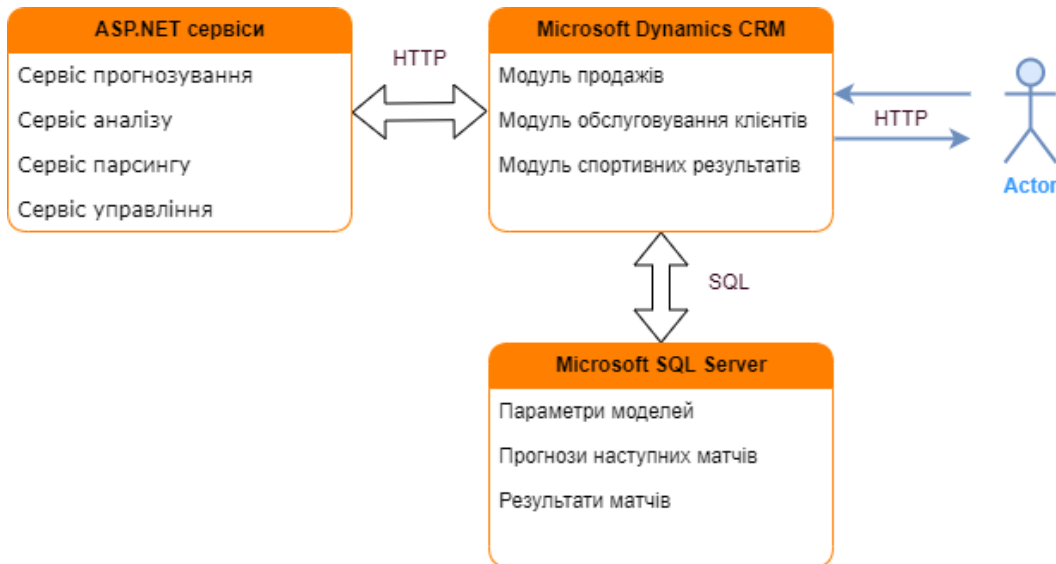


Рисунок 3.2 – Узагальнена архітектура

Для збереження усіх модифікацій, що були виконані у Microsoft Dynamics, створено рішення SportsPredictions. Воно являє собою архів та може бути перенесено на інше середовище. У рішенні містяться усі нові створені сутності, скрипти, форми, представлення, що необхідні для забезпечення роботи користувацького інтерфейсу та збереження даних у базі Microsoft SQL Server. При створенні нової сутності в програмній системі Microsoft Dynamics також створюється відповідна таблиця у базі даних з атрибутами, які є полями сутності. Розглянемо основні створені сутності для відображення та збереження даних. Крім нових створених полів, кожна створена сутність має стандартні поля, що позначаються статус запису, дату та час його створення, дату та час його зміни.

Competition являє собою сутність, що позначає змагання у якому приймають участь команди. Кожне змагання має власне ім'я та вид спорту, з якого це змагання проводиться. У таблиці 3.1 описана детальна структура сутності Competition.

Таблиця 3.1 – Структура сутності Competition

Найменування властивості	Тип даних	Характеристика
new_competitionid	Guid	Унікальний ідентифікатор змагання

Кінець таблиці 3.1

new_sport	Enum	Вид спорту
new_name	String	Ім'я змагання
statecode	Enum	Статус запису (активний/неактивний)
createdon	DateTime	Дата та час створення запису
modifiedon	DateTime	Дата та час зміни запису

Match є сутністю, яка уособлює поєдинок між двома командами. Має властивості, що позначають дату початку матчу, змагання, арену, де проводяться змагання, число глядачів, голи або очки, що здобула кожна з команд.

Таблиця 3.2 – Структура сутності Match

Найменування властивості	Тип даних	Характеристика
new_matchid	Guid	Унікальний ідентифікатор матчу
new_name	String	Назва матчу
new_startdateandtime	DateTime	Дата та час початку матчу
new_competitionid	Guid	Ідентифікатор змагання
new_seasonid	Guid	Ідентифікатор сезону
new_arenaid	Guid	Ідентифікатор арени, де відбувся поєдинок
new_attendance	Int	Кількість глядачів
new_winnerid	Guid	Ідентифікатор команди-переможця
new_wininvertime	Bool	Вказує чи перемога відбулась у додатковий час
new_hometeamid	Guid	Ідентифікатор домашньої команди

Кінець таблиці 3.2

new_homescore	Int	Рахунок домашньої команди
new_awayteamid	Guid	Ідентифікатор гостьової команди
new_awayscore	Int	Рахунок гостьової команди
statecode	Enum	Статус запису (активний/неактивний)
createddon	DateTime	Дата та час створення запису
modifiedon	DateTime	Дата та час зміни запису

Season реалізовує сутність сезону змагання, тобто конкретний екземпляр чемпіонату певного змагання, що має дату початку проведення та дату кінця. У таблиці 3.3 описано структуру сутності Season.

Таблиця 3.3 – Структура сутності Season

Найменування властивості	Тип даних	Характеристика
new_seasonid	Guid	Унікальний ідентифікатор сезону
new_name	String	Назва сезону
new_competitionid	Guid	Унікальний ідентифікатор змагання
new_startdate	DateTime	Дата початку сезону
new_enddate	DateTime	Дата кінця сезону
statecode	Enum	Статус запису (активний/неактивний)
createddon	DateTime	Дата та час створення запису
modifiedon	DateTime	Дата та час зміни запису

Сутність Arena являє собою майданчик для проведення певного матчу. Містить властивості імені та числа глядачів, яке вміщає арена. У таблиці 3.4 наведено структуру сутності Arena.

Таблиця 3.4 – Структура сутності Arena

Найменування властивості	Тип даних	Характеристика
new_arenaid	Guid	Унікальний ідентифікатор арени
new_capacity	Int	Кількість глядачів, що вміщає арена
new_name	String	Назва арени
statecode	Enum	Статус запису (активний/неактивний)
createddon	DateTime	Дата та час створення запису
modifiedon	DateTime	Дата та час зміни запису

Sports Team є сутністю, що уособлює команду, яка бере участь у змаганнях.

У таблиці 3.5 наведено повну структуру сутності Sports Team.

Таблиця 3.5 – Структура сутності Sports Team

Найменування властивості	Тип даних	Характеристика
new_sportsteamid	Guid	Унікальний ідентифікатор спортивної команди
new_name	String	Назва команди
new_sport	Enum	Вид спорту
new_division	String	Назва дивізіону, в якому виступає команда
new_founded	Int	Рік заснування команди
new_city	String	Місто, яке представляє команда
new_countryid	Guid	Країна, яку представляє команда

Кінець таблиці 3.5

new_arenaid	Guid	Унікальний ідентифікатор арени, на якій виступає команда
new_teamcode	String	Код команди
statecode	Enum	Статус запису (активний/неактивний)
createddon	DateTime	Дата та час створення запису
modifiedon	DateTime	Дата та час зміни запису

Крім того, у сутності команди існує зв'язок виду «багато до багатьох» з сутністю спортивного змагання. Такі зв'язки в Microsoft Dynamics CRM створюються у налаштуваннях сутності, після чого в базі даних з'являється нова поєднувальна таблиця з властивостями-ідентифікаторами записів, що мають між собою зв'язок. Тому для команд та змагань була створена таблиця SportsTeam_Competition (табл. 3.6).

Таблиця 3.6 – Структура зв'язку SportsTeam_Competition

Найменування властивості	Тип даних	Характеристика
new_sportsteamid	Guid	Унікальний ідентифікатор спортивної команди
new_competitionid	Guid	Унікальний ідентифікатор спортивного змагання

Прогнозування результатів відбувається відповідно до статистичних показників, що має команда. Тому, було створено сутність Team Parameters, структура якої описана у таблиця 3.7.

Таблиця 3.7 – Структура сутності Team Parameters

Найменування властивості	Тип даних	Характеристика
new_teamparametersid	Guid	Унікальний ідентифікатор запису

Кінець таблиці 3.7

new_sportsteamid	Guid	Унікальний ідентифікатор команди
new_parametersrelevancedate	DateTime	Дата, на яку параметри є актуальними
new_totalseasonwins	Int	Загальна кількість перемог у чемпіонаті
parameter_id	String	ID параметру
parameter_value	Double	Значення параметру
statecode	Enum	Статус запису (активний/неактивний)
createddon	DateTime	Дата та час створення запису
modifiedon	DateTime	Дата та час зміни запису

Основні параметри, що використовуються:

- опис параметру (Id);
- кількість перемог за сезон (WinSeason);
- кількість поразок за сезон (LostSeason);
- рейтинг атаки за сезон (OffenceSeason);
- рейтинг захисту за сезон (DefenseSeason);
- кількість перемог в останніх 10 матчах (WinsLast10);
- рейтинг атаки за останні 10 матчів (OffenceLast10);
- рейтинг захисту за останні 10 матчів (DefenseLast10);
- кількість перемог за останні 10 ігор (WinLast10).

Сутність Prediction Model визначає модель, по якій відбувається прогнозування результату. Її структура представлена у таблиці 3.8.

Таблиця 3.8 – Структура сутності Prediction Model

Найменування властивості	Тип даних	Характеристика
new_predictionmodelid	Guid	Унікальний ідентифікатор моделі прогнозування
new_name	String	Назва моделі прогнозування
statecode	Enum	Статус запису (активний/неактивний)
createddon	DateTime	Дата та час створення запису
modifiedon	DateTime	Дата та час зміни запису

Сутність Result Prediction є відповідно результатом прогнозування, який надають реалізовані в програмному забезпеченні моделі прогнозування. Структура даної сутності представлена у таблиці 3.9.

Таблиця 3.9 – Структура сутності Result Prediction

Найменування властивості	Тип даних	Характеристика
new_resultpredictionid	Guid	Унікальний ідентифікатор результату прогнозування
new_predictionmodelid	Guid	Унікальний ідентифікатор моделі прогнозування
new_matchid	Guid	Унікальний ідентифікатор матчу
new_hometeamwin-	Int	Відсоткова ймовірність, що переможе команда-господар
new_awayteamwin	Int	Відсоткова ймовірність, що переможе гостьова команда
new_score_over	Int	Відсоткова ймовірність, що загальний рахунок зустрічі перевищить середню кількість голів

Кінець таблиці 3.9

Statecode	Enum	Статус запису (активний/неактивний)
createddon	DateTime	Дата та час створення запису
modifiedon	DateTime	Дата та час зміни запису

3.3 Прогнозування методами машинного навчання

Для прогнозування результатів спортивних зустрічей було використано засоби машинного навчання. Мета полягає в тому, що для кожного матчу, який ще не відбувся, модель могла передбачити переможця зустрічі та середню кількість забитих голів.

Дані, що надаються для навчання моделі, називаються «набір навчання». Використовуючи ці дані, проводиться навчання моделі таким чином, щоб вона була здатна прогнозувати результати на даних, які не є частиною набору навчання. Дані, що використовуються для перевірки ефективності алгоритму, називаються тестовим набором.

Як вхідний, так і тестовий набір даних – це результати спортивних зустрічей минулих років, для яких вже відомі результати. Для оцінки ефективності вибрано ігри американської хокейної ліги NHL. У таблиці 3.10 описано складові елементи системи прогнозування результатів.

Таблиця 3.10 – Складові елементи системи прогнозування результатів

Тип елемента	Складові системи
Набір параметрів системи	<p>Win - число перемог за чемпіонат.</p> <p>Win10 - число перемог за останні 10 ігор.</p> <p>Offense – рейтинг атаки за чемпіонат.</p> <p>Defense – рейтинг оборони за чемпіонат.</p> <p>Offense10 – рейтинг атаки за останні 10 ігор.</p> <p>Defense10 – рейтинг оборони за 10 ігор.</p> <p>Score10 – середня кількість голів за останні 10 ігор</p>

Кінець таблиці 3.10

Обрані алгоритми	Лінійна регресія, логістична регресія
Набір навчання	Результати ігор NHL за 2014-2017 роки
Тестовий набір	Результати ігор NHL за 2018-2019 роки
Ваги параметрів	Для кожного параметра після навчання формується оптимальний ваговий коефіцієнт
Модель	Фінальна модель – це сукупність алгоритму, параметрів та вагових коефіцієнтів параметрів. Модель, застосована до актуальних параметрів нової гри, генерує необхідні результати: <ul style="list-style-type: none"> - ймовірність перемоги кожної з команд - ймовірність over (перевищення середньої кількості голів)

Під час створення моделі та машинного навчання варто взяти до уваги і уникати поширених проблем, що можуть значно погіршити якість прогнозів:

– надлишково навчена модель. Ці моделі ідеально описують «набір навчання», але мають невисоку ступінь узагальнення, і тому вони погано працюють на нових набору даних (тестових та реальних);

– спрощена модель. Ці моделі мають малий набір параметрів і погано прогнозують результати.

Для оцінки якості моделі після навчання ми використаємо «тестовий набір» даних матчів 2018 та 2019 років. Лише тестовий набір визначає ефективність моделі, навіть якщо модель показувала 100% точності на наборі для навчання.

Для всіх параметрів під час навчання формується «ваговий коефіцієнт», який входить до моделі. Після закінчення навчання, модель складається з набору вагових коефіцієнтів та функції.

Таблиця результатів змагань ліги NHL за сезон 2017-2018 для команди Anaheim має наступні дані: дата, вдома/в гостях, опонент, результат, забито шайб, пропущено шайб.

Таблиця 3.11 – Результати змагань, NHL, 2017-2018, Anaheim

Дата	Гра вдома	Опонент	Результат	Забито	Пропущено
07-04-2018		Arizona	Win	3	0
06-04-2018	+	Dallas	Win	5	3
04-04-2018	+	Minnesota	Win	3	1
01-04-2018	+	Colorado	Win	4	3
30-03-2018	+	Los Angeles	Win	2	1
27-03-2018		Vancouver	Lost	1	4
25-03-2018		Edmonton	Win	5	4
23-03-2018		Winnipeg	Lost	2	3
21-03-2018		Calgary	Win	4	0
18-03-2018	+	New Jersey	Win	4	2
16-03-2018	+	Detroit	Win	4	2
14-03-2018	+	Vancouver	Win	3	0
12-03-2018	+	St. Louis	Lost	2	4
09-03-2018		Dallas	Lost	1	2
08-03-2018		Nashville	Lost	2	4
06-03-2018	+	Washington	Win	4	0
04-03-2018	+	Chicago	Win	6	3
02-03-2018	+	Columbus	Win	4	2
25-02-2018	+	Edmonton	Lost	5	6
...					
24-09-2017	+	Vegas	Lost	2	4

Кінець таблиці 3.11

22-09-2017	+	Los Angeles	Win	4	2
20-09-2017	+	Arizona	Lost	1	5
19-09-2017		San Jose	Lost	0	5

Сезон 2017-2018 рр. (для простоти називатимемо сезон за першим роком 2017) складався з 30 команди, кожна команда зіграла в середньому 85 ігор. Всього за сезон зіграно 2550 ігор.

Для навчання моделі використано 4 сезони 2014-2017 рр., всього 10 200 ігор. Між сезонами відбувається ротація команд, з'являються нові з нижнього дивізіону, натомість декілька найгірших за результатами переходять до нижчого дивізіону. Цей момент враховується в моделі і для нових команд прогнози починають розраховуватись після 4 стартових ігор.

Модуль аналізу (SportsPredictions.DataAnalyzer) буде показники параметрів для кожного дня сезону для усіх команд, зберігаючи параметри у базі даних.

Основні параметри, що були використані в реалізації даного програмного забезпечення є: Win (кількість перемог у сезоні), Win10 (кількість перемог у останніх 10 матчах), Offense (рейтинг атаки за сезон), Defense (рейтинг захисту за сезон), Score10 (середня кількість забитих шайб за останні 10 ігор).

Таблиця 3.12 – Параметри команд NHL, сезон 2017, станом на 30.11.2017

Team	Offense	Defense	Win	Win10	Score10
Anaheim	29	17	0,42	0,50	2,50
Arizona	28	30	0,25	0,40	2,40
Boston	27	12	0,48	0,50	2,40
Buffalo	31	29	0,21	0,10	1,70
Calgary	20	25	0,52	0,50	3,50
Carolina	14	9	0,50	0,50	3,30
Chicago	8	3	0,53	0,50	3,30

Кінець таблиці 3.12

Colorado	5	14	0,55	0,40	3,20
Columbus	15	4	0,61	0,70	2,10
Dallas	18	13	0,53	0,50	2,90
Detroit	22	20	0,39	0,30	2,80
Edmonton	17	22	0,45	0,40	3,20
Florida	13	27	0,46	0,60	2,40
Los Angeles	21	1	0,57	0,30	2,10
Minnesota	23	10	0,52	0,60	3,00
Montreal	30	26	0,39	0,40	1,80
Nashville	11	11	0,57	0,80	3,90
New Jersey	6	8	0,62	0,50	2,80
NY Islanders	2	16	0,63	0,70	3,50
NY Rangers	10	15	0,53	0,70	3,50
Ottawa	9	28	0,39	0,20	2,10
Philadelphia	24	19	0,33	0,10	2,30
Pittsburgh	16	31	0,52	0,50	3,20
San Jose	25	2	0,59	0,50	2,50
St. Louis	7	7	0,59	0,60	3,60
Tampa Bay	4	5	0,63	0,60	3,10
Toronto	3	18	0,61	0,70	3,10
Vancouver	19	23	0,47	0,30	2,60
Vegas	1	24	0,62	0,60	3,40
Washington	26	21	0,50	0,60	2,80
Winnipeg	12	6	0,53	0,70	3,40

Параметри всіх команд змінюються з кожною проведеною грою, тому вони перераховуються щодня.

Таблица 3.13 – Параметры команд NHL, 2017, станом на 07.11.2017

Team	Offense	Defense	Win	Win10	Score10
Anaheim	30	17	0,40	0,30	2,30
Arizona	29	29	0,26	0,50	2,70
Boston	28	10	0,48	0,60	2,60
Buffalo	31	30	0,22	0,20	1,60
Calgary	21	24	0,49	0,40	2,80
Carolina	25	12	0,49	0,40	2,60
Chicago	12	5	0,47	0,40	3,10
Colorado	9	20	0,50	0,40	2,70
Columbus	18	3	0,58	0,70	2,50
Dallas	14	14	0,56	0,70	3,50
Detroit	20	27	0,39	0,30	3,00
Edmonton	15	25	0,44	0,40	3,40
Florida	17	26	0,42	0,40	2,20
Los Angeles	13	1	0,63	0,70	3,20
Minnesota	23	9	0,53	0,50	3,10
Montreal	26	21	0,42	0,50	3,40
Nashville	5	13	0,59	0,80	3,70
New Jersey	8	6	0,63	0,50	2,30
NY Islanders	1	22	0,60	0,70	4,10
NY Rangers	6	11	0,56	0,70	3,50
Ottawa	11	31	0,39	0,10	1,50
Philadelphia	24	18	0,36	0,20	2,90
Pittsburgh	16	28	0,53	0,60	3,60
San Jose	27	2	0,56	0,40	2,20
St. Louis	10	7	0,57	0,50	3,40

Кінець таблиці 3.13

Tampa Bay	3	4	0,65	0,60	3,20
Toronto	4	15	0,61	0,60	2,90
Vancouver	19	16	0,52	0,50	3,20
Vegas	2	23	0,61	0,70	3,50
Washington	22	19	0,53	0,70	3,40
Winnipeg	7	8	0,54	0,60	3,60

3.4 Результати дослідження ефективності системи

Для роботи було виконане навчання двох моделей: лінійної регресії та логістичної регресії. Навчання виконане на тестових даних чемпіонату з хокею NHL, сезони 2014-2017. Для кожної моделі після навчання виконана перевірка на «тестовому наборі» NHL, сезони 2018-2019.

Для прикладу, наведемо прогноз результату матчу сезону 2018 року, які надаються створеними моделями.

Матч: 23.11.2018, Anaheim (вдома) – Emdonton.

На даний матч моделі підраховали наступні ймовірності перемоги команд.

Таблиця 3.14 – Прогноз моделей «ймовірність перемоги» у зустрічі 23.11.2018

Команда	Лінійна регресія	Логістична регресія
Anaheim	62%	67%
Emdonton	38%	33%

Фактичний результат: 2:1 (перемога Anaheim). Таким чином переможець спрогнозований вірно.

Одним з часто вживаних прогнозів є Over/Under – перевищення середньої кількості голів у матчі. Референтним значенням Over/Under для хокею є 3.5. Референтне значення різне для кожного виду спорту (Наприклад, для NBA дорівнює 112).

Таблиця 3.15 – Прогноз «Ймовірність Over» у зустрічі 23.11.2018

Команда	Лінійна регресія	Логістична регресія
Anaheim - Emdonton	45%	42%

Фактичний результат: Under (нижче середнього 3.5). Таким чином результат спрогнозований вірно.

Було проведено дослідження ефективності і точності моделей лінійної регресії та логістичної регресії для прогнозування результатів ігор NHL в сезонах 2018-2019 років.

Оцінка ефективності виконувалась наступним чином: для кожного матчу в тестовому наборі даних виконується прогноз результату із використанням двох моделей. Надалі результати прогнозу порівнюються з справжнім результатом, який відомий. Таким чином можна підрахувати кількість правильних прогнозів і оцінити ефективність моделі у відсотках для кожної з команд.

Таблиця 3.16 – Оцінка ефективності моделей для результатів ігор NHL сезону 2018-2019 за критерієм переможець матчу

Команда	Кількість правильних прогнозів, %	
	Лінійна регресія	Логістична регресія
Anaheim	65	71
Arizona	76	79
Boston	58	61
Buffalo	57	59
Calgary	58	57
Carolina	69	69
Chicago	66	70
Colorado	64	68
Columbus	60	61
Dallas	63	61

Кінець таблиці 3.16

Detroit	65	70
Edmonton	63	69
Florida	69	67
Los Angeles	59	58
Minnesota	62	66
Montreal	51	57
Nashville	56	56
New Jersey	81	80
NY Islanders	71	70
NY Rangers	61	64
Ottawa	64	67
Philadelphia	71	70
Pittsburgh	64	71
San Jose	67	67
St. Louis	66	74
Tampa Bay	70	77
Toronto	67	68
Vancouver	53	52
Vegas	69	78
Washington	76	74
Winnipeg	66	64
Середня ефективність, %	64,7	66,9

Ефективність прогнозування вище 65% вважається хорошою. Оцінка результатів дослідження двох моделей показала хорошу ефективність, яка є достатньою для застосування в реальних сферах використання. В той же час виявлено декілька важливих елементів, що можуть значно покращити

результати, і їх можливо і необхідно впровадити при подальшому розвитку програмного забезпечення:

- можливо застосувати більше різних моделей. Моделі можуть підбиратись під кожен команду окремо;
- глибший аналіз показує, що достатньо аналізувати останні 2 роки чемпіонату;
- збільшення кількості незалежних параметрів може значно покращити моделі. Наприклад, застосування рейтингів гравців команд, що безпосередньо приймають участь у зустрічі.

3.5 Керівництво користувача

Для використання запропонованого програмного забезпечення необхідно встановити рішення для Microsoft Dynamics на цільовому середовищі. Далі у меню додатків Microsoft Dynamics, користувач повинен обрати застосунок SportsPrediction. При цьому зліва буде розташовуватись зручне меню для вибору сутності, записи якої необхідно переглянути. У випадку кліку на назву сутності, відкриється одне з представлень її записів. Користувач може змінити представлення, здійснивши клік лівою кнопкою миші біля назви представлення. На рис. 3.3 зображено перегляд представлення сутності змагань. Аналогічно, можна переглянути представлення спортивних команда та матчів (рис. 3.4 – 3.5).

The screenshot displays the Dynamics 365 SportsPrediction interface. The top navigation bar includes the Dynamics 365 logo, the page title 'SportsPrediction', and utility icons for search, add, help, and user profile. The left sidebar contains a navigation menu with options like Home, Recent, Pinned, Stats, Competitions (selected), Seasons, Matches, Sports Teams, Sports Arenas, Teams Parameters, Predictions, Result Predictions, and Prediction Models. The main content area is titled 'Active Competitions' and features a table with columns for Name, Sport, and Modified On. The table lists three active competitions: MLB (Baseball), NBA (Basketball), and NHL (Ice Hockey), all with a modified date of 19.11.2021. The interface also includes a search bar, a toolbar with 'Show Chart', 'New', 'Delete', and 'Refresh' buttons, and a pagination indicator at the bottom showing '1 - 3 of 3' and 'Page 1'.

Name	Sport	Modified On
MLB	Baseball	19.11.2021 1...
NBA	Basketball	19.11.2021 1...
NHL	Ice Hockey	19.11.2021 1...

Рисунок 3.3 – Перегляд представлення змагань

The screenshot displays the Dynamics 365 SportsPrediction interface for 'MLB Teams'. The top navigation bar is similar to the previous screenshot, but includes an 'Email a Link' button in the toolbar. The left sidebar is the same, with 'Sports Teams' selected. The main content area is titled 'MLB Teams' and features a table with columns for Name, Sport, Founded, City, Arena, and Division. The table lists ten MLB teams, including Arizona Diamondbacks, Atlanta Braves, Baltimore Orioles, Boston Red Sox, Chicago Cubs, Chicago White Sox, Cincinnati Reds, and Cleveland Indians. The interface includes a search bar, a toolbar with 'Show Chart', 'New', 'Delete', 'Refresh', 'Email a Link', and 'Flow' buttons, and a pagination indicator at the bottom showing '1 - 30 of 30' and 'Page 1'.

Name	Sport	Founded	City	Arena	Division
Arizona Diamondbacks	Baseball	1998	Phoenix, ...	Chase Field	West
Atlanta Braves	Baseball	1871	Cumberla...	Truist Park	East
Baltimore Orioles	Baseball	1901	Baltimore...	Oriole Pa...	East
Boston Red Sox	Baseball	1901	Boston, ...	Fenway P...	East
Chicago Cubs	Baseball	1874	Chicago, ...	Wrigley F...	Central
Chicago White Sox	Baseball	1901	Chicago, ...	Guarante...	Central
Cincinnati Reds	Baseball	1882	Cincinnati...	Great Am...	Central
Cleveland Indians	Baseball	1901	Cleveland...	Progressi...	Central

Рисунок 3.4 – Перегляд записів спортивних команд

The screenshot displays the Dynamics 365 SportsPrediction interface. The left sidebar contains navigation options: Home, Recent, Pinned, Stats, Competitions, Seasons, Matches (highlighted), Sports Teams, Sports Arenas, Teams Parameters, Predictions, Result Predictions, and Prediction Models. The main area shows a table of 'Active Matches' with columns for Start Date, Competition, Arena, Home Team, Away Team, Home Score, and Away Score. The table lists 10 matches from 21.11.202... to 22.11.202... involving NHL teams. The bottom of the interface shows '1 - 38 of 38' and 'Page 1'.

Start Da...	Competition	Arena	Home Team	Away Team	Home Score	Away Score
22.11.202...	NHL	Staples Ce...	Los Angel...	Arizona C...	1	2
22.11.202...	NHL	Climate Pl...	Seattle Kr...	Washingt...	5	2
22.11.202...	NHL	Rogers Ar...	Vancouve...	Chicago B...	0	1
22.11.202...	NHL	UBS Arena	New York ...	Toronto ...	0	3
22.11.202...	NHL	TD Garden	Boston Br...	Calgary Fl...	0	4
22.11.202...	NHL	Madison ...	New York ...	Buffalo Sa...	5	4
22.11.202...	NHL	Amalie Ar...	Tampa Ba...	Minnesot...	5	4
21.11.202...	NHL	SAP Center	San Jose ...	Washingt...	0	4
21.11.202...	NHL	Rogers Pl...	Edmonto...	Chicago B...	5	2

Рисунок 3.5 – Перегляд записів матчів

При подвійному кліку на одному з рядків представлення, відкриється форма даного запису, що дозволяє вносити зміни у дані та зберігати їх (рис. 3.6).

Dynamics 365 | SportsPrediction

Los Angeles Kings - Saved
Sports Team

General Related

Name	Los Angeles Kings
Sport	Ice Hockey
Founded	1 967
Division	Pacific
Country	United States of America
City	Los Angeles, California
Arena	Staples Center
Team Code	LAK

Competitions

Name	Sport	Modified On
NHL	Ice Hock...	19.11.2021 ...

Рисунок 3.6 – Форма запису команди

Результати прогнозування матчів виконані у вигляді сутності Result Predictions. Відкривши представлення даної сутності, користувач може побачити відсоткові ймовірності перемоги кожної з команд (рис. 3.7).

The screenshot shows the Dynamics 365 SportsPrediction interface. The left sidebar contains navigation options: Home, Recent, Pinned, Stats, Competitions, Seasons, Matches, Sports Teams, Sports Arenas, Teams Parameters, Predictions, Result Predictions (selected), and Prediction Models. The main area displays 'Active Result Predictions' with a table of two rows. The table has columns for Match, Prediction Model, Home Team Win %, Away Team Win %, and Predicted Over, %.

Match	Prediction Model	Home Team Win %	Away Team Win %	Predicted Over, %
Anaheim-E...	Logistic Regress...	67	33	42
Anaheim-E...	Linear Regression	62	38	45

At the bottom of the table, it shows '1 - 2 of 2' and 'Page 1'.

Рисунок 3.7 – Перегляд записів результатів прогнозування

Висновки до розділу

У даному розділі було описано розроблене програмне забезпечення для аналізу та прогнозування результатів спортивних змагань. Зроблено короткий огляд технологій, що використовувались при розробці. Описано архітектуру програмного забезпечення та досліджено ефективність прогнозів, що надаються системою.

4 РОЗРОБЛЕННЯ СТАРТАП-ПРОЄКТУ

4.1 Інформаційна карта проекту

Назва проекту: «Програмне забезпечення для аналізу результатів спортивних змагань».

Зміст ідеї

Ідея проекту полягає у створенні системи для аналізу результатів спортивних змагань і прогнозування результатів на основі даних попередніх ігор і поточних характеристик.

Програмний модуль допоможе отримувати в режимі реального часу дані про майбутні спортивні події хокею, баскетболу, бейсболу щодо величини, яка визначає ймовірність перемоги однієї з команд або величини, яка визначає кількість забитих голів/набраних очків.

Опис проблеми, яку вирішує проєкт

Існує значна кількість систем аналізу та прогнозування результатів спортивних змагань. Проте, наявні аналоги зазвичай орієнтовані на окремих користувачів і використовують закриті алгоритми, а також надають результати лише у власних сайтах та програмах.

Головні цілі та завдання проєкту

Ціль: створення програмного забезпечення прогнозування та аналізу результатів спортивних змагань.

Завданням є створити проєкт с наступним функціоналом:

- прогнозування результатів змагань, використовуючи ефективні моделі;
- інтеграція програмного модулю з системою управління взаємовідносинами з клієнтами;
- прогнозування результатів з використанням декількох статистичних алгоритмів.

Очікувані результати

Використання даної системи підприємствами дозволить формувати аналіз спортивних подій. Прогнозування спортивних результатів корисне для гравців, керівництва команди та аналітиків при виявленні найважливіших чинників, які допомагають досягти виграшних результатів на основі яких можна визначити відповідну тактику. Це рішення буде привабливим для малих та середніх підприємств.

4.2 Аналіз потенційних техніко-економічних переваг ідеї

У таблиці 4.1 наведено аналіз потенційних техніко-економічних переваг ідеї у порівнянні із продуктами конкурентів. Основними перевагами системи є відкрита архітектура та реалізація елементів у формі незалежних REST-сервісів, що дозволяє інтеграцію з іншими сервісами, заміну моделі прогнозування та джерела даних про минулі результати.

Таблиця 4.1 – Аналіз потенційних техніко-економічних переваг ідеї

№	Характеристик и ідеї	Оцінка концепцій конкурентів			Слабка (W)	Нейтрал ьна (N)	Сильна (S)
		Odds Wizard	BetClan	Predicti on Machin e			
1	Зручність в користуванні	-	+	+			+
2	Експорт даних	-	-	-			+
3	Додавання нових видів спорту	-	-	-			+
4	Декілька видів спорту	+	+	+			+

Кінець таблиці 4.1

5	Можливість інтеграції із зовнішніми сервісами	-	-	-			+
6	Зміна джерела даних	-	-	-			+
7	Можливість інтеграції в CRM Dynamics	-	-	-			+
8	Швидкість аналізу	-	+	+		+	
9	Зміна моделі прогнозування	-	-	-			+
10	Веб-клієнт	-	+	+			+
11	Кросплатформність	-	+	+		+	

4.3 Технологічний аудит ідеї проекту

Аудит технологій, які необхідні для виконання проекту, наведений у таблиці 4.2.

Таблиця 4.2 – Технології, які необхідні для виконання проекту

№	Ідея проекту	Технологія та реалізація	Наявність технологій	Доступність технологій
1	Мова програмування	C#, Javascript	Так	Доступна
2	Система розробки	Microsoft Dynamics	Так	Доступна

Кінець таблиці 4.2

3	Середовище розгортання	Microsoft Dynamics	Так	Доступна
4	Збереження даних	Microsoft SQL Server	Так	Доступна
5	Джерело даних спортивних результатів минулого	Парсинг веб-сайту спортивної статистики	Необхідно розробити парсер.	Доступна, парсер розроблюється на С#
6	Розміщення модулів	Хмарні сервіси	Так	Доступна
7	Взаємодія між модулями	REST, JSON	Так	Доступна

Технологічна реалізація проєкту можлива. Для зберігання даних та розробки форм взаємодії з клієнтом обрано CRM систему Microsoft Dynamics. Для розробки REST API сервісів обрано фреймворк ASP.NET, мова С#.

4.4 Аналіз ринкових можливостей стартап-проєкту

Визначимо ринкові перспективи та загрози, що можуть перешкодити запуску проєкту.

Таблиця 4.3 – Аналіз ринкових можливостей проєкту

№	Ринковий параметр	Значення
1	Кількість конкурентів зі схожим продуктом	4
2	Розвиток ринку	Зростає.

Кінець таблиці 4.3

3	Обмеження для запуску проекту	Необхідний доступ до результатів змагань та всіх параметрів у вигляді відкритого або платного API. Не всі види спорту мають такий доступ. Можливий парсинг сайтів для завантаження необхідних результатів.
4	Особливі вимоги	Продукт потребує доступ до результатів спортивних змагань. Більшість даних є у відкритому доступі.
5	Середня рентабельність у галузі, %	20%
6	Загальний обсяг продажу, грн.	-

Аналіз показує, що ринок є достатньо привабливим для входження.

Кількість основних гравців обмежена і жоден гравець не пропонує таку гнучку систему, що буде реалізовано у стартап-проекті.

Визначимо групи потенційних клієнтів, їх характеристики та вимоги до продукту.

Таблиця 4.4 – Потенційні користувачі продукту

№	Потреба, яку формує ринок	Цільова аудиторія	Відмінності у поведінці потенційних цільових груп	Вимоги споживачів до товару
1	Оцінка результату команд у порівнянні з прогнозованим (краще/гірше).	Спортивні організації та клуби.	Клієнти потребують зручну систему з великою кількістю параметрів.	Система аналізу має бути гнучкою.

Кінець таблиці 4.4

2	Прогноз результату спортивної події для ігрових ставок у букмекерів.	Приватні гравці – клієнти букмекерів.	Клієнти потребують системи аналізу для прогнозування ставок. Можуть використовувати декілька продуктів одночасно. Використання як правило періодичне.	Максимально наближений прогноз результатів спортивних подій до коефіцієнтів букмекерів.
3	Побудова рейтингів спортивних команд: атака, захист і т.д.	Компанії, що працюють у сфері спорту, спортивні організації, букмекерські компанії.	Клієнти потребують зручну систему аналізу для щоденного використання. Продукт має легко інтегруватись з існуючими CRM та дозволяти експортувати дані для звітів.	Система аналізу має працювати надійно та в автоматичному режимі. Клієнти вимагають індивідуального підходу та повного супроводу системи (налаштування, підтримка).

Потреби, які формують ринок – це необхідність статистичної оцінки ефективності команд та прогнозування її майбутніх результатів у порівнянні з

суперниками. Клієнти вимагають безперервної роботи, гнучкого налаштування та точності прогнозів, що може конкурувати з гравцями ринку та букмекерськими компаніями.

Після аналізу груп потенційних клієнтів виконаємо аналіз ринкового середовища сформуємо таблицю факторів, що можуть перешкодити впровадженню проекту на ринку, та факторів, що сприяють його запуску. Такий аналіз допомагає заощадити час та кошти і сформувати задачі для подальшого розвитку проекту.

Таблиця 4.5 – Фактори загроз розробки системи аналізу результатів спортивних змагань

№	Фактор	Зміст загрози	Реакція компанії
1	Поява нових конкурентів	Зменшення продажу продукту	Запуск рекламних заходів серед потенційних клієнтів в Україні та у світі.
2	Відсутність зацікавленості у груп потенційних користувачів	Продукт не виконує завдання, що потребують користувачі	Розширення функціоналу, підвищення якості роботи аналізу
3	Нестабільна робота систем продукту	Продукт розміщується у хмарного провайдера (Microsoft Azure) і необхідні кваліфіковані спеціалісти для правильного налаштування служб.	Використання послуг кваліфікованих інженерів для розміщення та підтримки служб сервісу у Microsoft Azure. Розширення потужності служб за рахунок придбання більшої кількості ресурсів у провайдера.

Кінець таблиці 4.5

4	Некваліфіковані користувачі продукту	Приватні користувачі не мають бажання та часу вивчати функціонал та не можуть ефективно використовувати продукт.	Створити навчальним матеріал з відео, що спрощує навчання. Безкоштовно надавати консультації приватним користувачам. Моніторинг відгуків та скарг на продукт.
5	Неточний прогноз результатів	Не всі види спорту піддаються ефективному прогнозу за минулим результатами.	Покращувати алгоритми та моделі аналізу. Вводити додаткові показники та параметри для підвищення прогнозу результатів спортивних подій.

Таблиця 4.6 – Фактори можливостей розробки системи аналізу результатів спортивних змагань

№	Фактор	Зміст можливості	Реакція компанії
1	Можливість інтеграції з CRM	Система розроблена як модуль у CRM Microsoft Dynamics, але за потреби легко інтегрується і іншими зовнішніми системами через REST API.	Впровадження додаткових модулів інтеграції для поширених на ринку CRM

Кінець таблиці 4.6

2	Можливість налаштування додаткових параметрів аналізу	Модуль аналізу є окремим сервісом, що може розширюватись та доповнюватись новими параметрами аналізу	Розробка альтернативних модулів аналізу для використання у системі. Різні моделі аналізу дають більш точний прогноз у окремих видах спорту.
3	Можливість зміни джерела даних	Завдяки модульній архітектурі є можливість заміни модулю завантаження параметрів та результатів спортивних зустрічей.	Розробка додаткових модулів завантаження результатів спортивних зустрічей з інших веб-порталів спортивної статистики.
4	Автоматизована робота	Програма працює у автономному режимі, оновлюючи всі дані щодня у необхідний час, щоб зробити вчасно прогноз для кожного виду спорту.	Розміщення компанії на надійних Cloud серверах компанії Microsoft дозволяє гарантувати 24/7 безперебійну роботу, масштабування навантаження та моніторинг справності систем.

З часом, завдяки розширенню або зниженню ринку, відбувається зміна можливостей і загроз, тому потрібно реагувати та нові фактори. Далі проведемо ступеневий аналіз конкуренції на ринку.

Таблиця 4.7 – Ступеневий аналіз конкуренції на ринку

№	Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність компанії (дії компанії для підвищення конкурентоспроможності)
1	Тип: чиста конкуренція	На ринку присутні продукти зі схожим функціоналом. Жоден продукт не пропонує повного функціоналу пропонованої системи аналізу.	Розвивати запропоновану систему аналізу для забезпечення всіх важливих функціональних можливостей, що є у конкурентів.
2	Рівень: світова конкурентна боротьба	Система аналізу націлена на світовий ринок, тому що головні спортивні змагання відбуваються у різних країнах світу, головним чином США та Європи.	Збільшувати кількість та географію користувачів, виходити на ринки різних країн, де спорт є популярним і відповідні системи матимуть гарантований попит.

Продовження таблиці 4.7

3	Галузь: внутрішньогалузева	Гравці ринку пропонують продукти в одній галузі: програмне забезпечення для прогнозування результатів спортивних змагань	Покращувати функціонал продукту, підтримки та супроводження користувачів.
4	Види товарів: товарно-видова	Конкуренція між схожими продуктами відбувається у алгоритмах та моделях, що застосовуються для аналізу.	Покращувати якість та кількість моделей аналізу.
5	Конкурентні переваги: нецінова	Продукти на ринку відрізняються за функціоналом: окрім прогнозування є додаткові функції (коефіцієнти букмекерів, вибір найкращих ставок з точки зору ефективності), цільовою аудиторією (спортивні організації, приватні гравці)	Пропонувати готові до використання пакетні продукти для різних видів користувачів, що забезпечує швидкий старт початку ефективного використання продукту. Додаткова перевага для

Кінець таблиці 4.7

6	За інтенсивністю: марочна	Унікальна назва продукту запам'ятовується користувачами та потенційними клієнтами та викликає більшу довіру. Користувач схильний придбати продукт з більш відомою назвою.	Зареєструвати назву продукту, що легко запам'ятовується, і логотип. Захистити авторські права на торговельну марку.
---	------------------------------	--	--

Проведемо детальний аналіз конкуренції в галузі за моделлю 5 сил М.Портера.

Таблиця 4.8 – Аналіз конкуренції в галузі за М. Портером

	Складові аналізу	Висновки
Прямі конкуренти в галузі	Odds Wizard, BetClan, Prediction Machine	Конкурентів небагато, боротьба на ринку незначна і є багато нереалізованих ніш.
Потенційні конкуренти	Розробники програмного забезпечення, сайти прогнозування спортивних результатів, які можуть випустити власний продукт, такі як: pickswise.com, olbg.com/betting-tips, sportytrader.com/en/betting-tips/	Вихід на ринок можливий, потенційні конкуренти періодично з'являються, але всі є досить невеликими

Кінець таблиці 4.8

Постачальники	Електронні сервіси/сайти з результатами спортивних зустрічей у минулому.	Постачальників є багато, за потреби можна змінити постачальника результатів спортивних подій.
Клієнти	Спортивні організації та клуби, приватні клієнти та гравці.	Різні групи клієнтів висувають різні вимоги до продукту – потрібно пропонувати декілька пакетів продукту з різним функціоналом та ціною.
Товари замітники	Продукти замітники не надають повного функціоналу, але деякі системи мають частину функціоналу.	Необхідно реагувати ціною та змістом пакетних пропозицій на товари, що складають конкуренцію.

Проведемо аналіз слабких та сильних сторін системи аналізу.

Таблиця 4.9 – Порівняльний аналіз сильних та слабких сторін

№	Фактор конкурентоспроможності	Бали 1-20	Рейтинг товарів-конкурентів у порівнянні з системою аналізу.						
			-3	-2	-1	0	+1	+2	+3
1	Зручність в користуванні	10				+			
2	Експорт даних	15	+						

Кінець таблиці 4.9

3	Додавання нових видів спорту	10	+						
4	Декілька видів спорту	10						+	
5	Можливість інтеграції із зовнішніми сервісами	15	+						
6	Зміна джерела даних	10			+				
7	Можливість інтеграції в CRM Dynamics	20	+						
8	Швидкість аналізу	10					+		
9	Зміна моделі прогнозування	20	+						
10	Веб-клієнт	15						+	
11	Кросплатформність	5					+		

Заключним етапом аналізу ринкових можливостей є SWOT-аналіз: таблиця сильних та слабких сторін, загроз та можливостей. На основі SWOT-аналізу виконаємо опис альтернатив ринкової поведінки та орієнтовний час їх реалізації, враховуючи проекти конкурентів.

Таблиця 4.10 – SWOT-аналіз стартап-проєкту

Сильні сторони	Слабкі сторони
Інтеграція з CRM та іншими системами	Перша версія проєкту підтримуватиме лише 3 види спорту.
Зміна моделі прогнозування	
Додавання нових видів спорту	
Експорт даних	
Можливості	Загрози

Кінець таблиці 4.10

Можливість налаштування додаткових параметрів аналізу	Нестабільна робота систем продукту
Можливість зміни джерела даних	Некваліфіковані користувачі продукту
Автоматизована робота	Неточний прогноз результатів

Розробимо альтернативи ринкової поведінки на основі SWOT-аналізу. Розглянемо оптимальний час виводу на ринок з огляду на конкурентів та стан ринку.

Таблиця 4.11 – Альтернативи ринкового впровадження стартап-проекту

№	Альтернатива ринкової поведінки (комплекс заходів)	Ймовірність отримання ресурсів	Термін реалізації
1	Додавання більшої кількості видів спорту до аналізу	Середня	6 міс.
2	Додавання нових моделей аналізу	Висока	6 міс.
3	Інтеграція з CRM Dynamics	Висока	3 міс.

Вибираємо альтернативу 3 з мінімальним терміном реалізації та високою ймовірністю отримання ресурсів.

4.5 Ринкова стратегія проекту

Розробимо стратегію розвитку та охоплення ринку. Спочатку визначимо цільові групи споживачів у таблиці 4.12.

Таблиця 4.12 – Цільові групи споживачів

№	Профіль групи потенційних споживачів	Готовність споживачів прийняти продукт	Попит в межах групи споживачів	Інтенсивність конкуренції в групі	Простота входу в групу
1	Спортивні організації та клуби	Середня	Високий	Низька	Низька
2	Професійні гравці	Висока	Високий	Середня	Висока
3	Приватні споживачі	Середня	Середній	Висока	Середня

За результатом аналізу груп споживачів виберемо спортивні організації та клуби, а також приватних користувачів. Вибір обумовлено меншою конкуренцією в цих групах та простотою входу на ринок. Серед приватних споживачів дуже висока конкуренція інших проектів з великою кількістю безкоштовних та низькоякісних сайтів, що ускладнює вихід на ринок іншим продуктам.

Для спортивних організацій та клубів, а також професійних гравців, ціна продукту не грає високої ролі, а головним чинником виступає якість продукту, точність аналізу та підтримка.

Таблиця 4.13 – Визначення базової стратегії розвитку

№	Обрана стратегія розвитку	Стратегія охоплення ринку	Ключові позиції відповідно до обраної стратегії	Базова стратегія
1	Розробка пакетних рішень, що задовольняють потреби різних цільових груп споживачів. Індивідуальне налаштування у кінцевого споживача за потреби.	Таргетована реклама серед цільової аудиторії. Кількість спортивних клубів та організацій обмежена, тому можна її охопити повністю. Масована реклама на тематичних ресурсах для професійних гравців.	Інноваційні рішення для аналізу, гарантована якість продукту, інтеграція із CRM Dynamics для організацій, експорт даних та API для інтеграції із сторонніми системами, Значна відмінність від існуючих рішень на ринку.	Стратегія спеціалізації

Стратегія спеціалізації вибрана базовою стратегією розвитку проекту. Стратегія передбачає концентрацію функціоналу та реклами на потребах кожної

групи окремо, без намагання входу на весь ринок одночасно. Метою є максимально швидко та якісно охопити вибрані групи, пропонуючи повне «коробочне» рішення з переліком переваг над конкурентами у кожній групі окремо.

Якість роботи системи аналізу, точність прогнозів та ефективне і автоматизоване використання результатів прогнозів дає змогу встановити високу ціну на продукцію, зважаючи на відсутність конкурентів за аналогічним по обсягу функціоналом та інтеграцією.

Надання підтримки та гарантія якості є важливим фактором для роботи середніх та великих підприємств у цільовій групі спортивних організацій. Надалі виберемо стратегію конкурентної поведінки.

Таблиця 4.14 – Визначення базової стратегії конкурентної поведінки

№	Проект є «першопрохідцем» на ринку?	Компанія буде шукати нових споживачів, чи забирати у конкурентів?	Компанія буде копіювати основні характеристики продукту у конкурентів?	Стратегія конкурентної поведінки
1	Ні	Шукати нових і забирати існуючих клієнтів у конкурентів	Лише деякі, головним чином це власний продукт із інноваційними розробками	Наступальна

Обрана наступальна стратегія, яка передбачає збільшення своєї частки ринку. Стратегія передбачає максимальне використання інноваційних розробок

та ідей без очікування появи подібних у конкурентів. Такий підхід забезпечує максимальний відрив від конкурентів та створення пулу лояльних користувачів.

Розробимо стратегію позиціонування на основі вимог споживачів та в залежності від базової стратегії розвитку.

Таблиця 4.15 – Визначення стратегії позиціонування

№	Вимоги групи споживачів до продукту	Базова стратегія розвитку	Ключові позиції проекту	Асоціації, що формують позицію проекту
1	Точність прогнозів	Стратегія спеціалізації	Висока точність прогнозів результатів та формування прихильності користувачів	Якість і точність, технічна підтримка
2	Розширення функціоналу	Стратегія спеціалізації	Інноваційні рішення	Постійне вдосконалення продукту
3	Інтеграція з CRM та зовнішніми сервісами	Стратегія спеціалізації	Гнучкість, широкий вибір API для інтеграції, інтеграція з професійною CRM Microsoft Dynamics	Якість, підтримка, безперебійність роботи.

Користувачі продукту потребують високу точність прогнозів, а також швидку підтримку для ефективної роботи. Продукт має постійно оновлюватись та вдосконалювати функціонал.

При створенні програми маркетингу розробимо маркетингову концепцію продукту і підведемо підсумки конкурентоспроможності продукту.

Таблиця 4.16 – Визначення ключових переваг концепції продукту

№	Потреба	Вигода	Ключові переваги над конкурентами
1	Точність прогнозу результатів	Конкурентні переваги користувачів продукту (точніші прогнози, кращі результати).	Використання декількох моделей аналізу дозволяє отримувати кращі прогнози, що більш пристосовані до кожного виду спорту.
2	Стабільність роботи	Автономна робота продукту, вчасні оновлення та безперебійна робота.	Модулі проекту розміщені на хмарних ресурсах, що мають гарантовану безперебійну роботу. Випуск оновлень та підтримка забезпечують безперервну роботу.

Кінець таблиці 4.16

3	Інтеграція з CRM та зовнішніми системами	Зручне використання результатів продукту в CRM, або інших зовнішніх сервісах, з якими звикли працювати користувачі.	Використання CRM систем для управління результатами прогнозів спортивних змагань. Інтеграція із зовнішніми системами через API.
---	--	---	---

Перевагами даного продукту є точність прогнозів, стабільність роботи, інтеграція із зовнішніми системами та CRM Dynamics, що надає перевагу у порівнянні з конкурентами. Оновлення та впровадження інноваційних рішень дозволяє вивести продукт на вищий рівень.

Далі розробимо трирівневу маркетингову модель товару, щоб деталізувати ідею продукту, його складові та особливості.

Таблиця 4.17 – Опис трьох рівнів моделі продукту

Рівні товару	Сутність та складові
1. Товар за задумом	Система прогнозу результатів спортивних змагань для різних видів спорту. Програма є модульною і інтегрується з CRM Microsoft Dynamics та іншими системами за необхідності.
2. Товар у реальному виконанні	Характеристики: модульна автоматизована система можливість використання декількох моделей аналізу інтеграція з CRM Microsoft Dynamics

Кінець таблиці 4.17

3. Товар з підкріпленням	Додавання нових моделей аналізу для різних видів спорту Додавання видів спорту Розширення параметрів аналізу
--------------------------	--

Визначимо межі при встановленні ціни на продукт. Процес передбачає аналіз цін конкурентів та можливостей оплати споживачів продукту.

Таблиця 4.18 – Визначення меж встановлення ціни, грн.

№	Ціни на продукти-замінники	Ціни на продукти-аналоги	Рівень доходів цільової групи споживачів	Нижня і верхня межа встановлення ціни на продукт
1	0-5000	100-5000	10000-100000	1000-10000

Унікальність продукту та його функціонал і підтримка дає змогу встановити високі початкові ціни і бути гнучким у наданні індивідуальної знижки покупцям. При збільшенні кількості користувачів є можливість зменшувати ціну.

Визначимо оптимальну систему збуту продукту.

Таблиця 4.19 – Система збуту

№	Особливості закупівельної поведінки клієнтів	Функції збуту постачальника товарів	Глибина каналу збуту	Оптимальна система збуту
1	Попереднє замовлення клієнтами. Оплата за продукт складається з фіксованої оплати при покупці ліцензії, та щомісячній абонентській оплаті. Після покупки клієнт скачує та самостійно встановлює програму, або може замовити підтримку для допомоги в початкових налаштуваннях.	Доставка програми відбувається при скачуванні програми за унікальним посиланням після оплати. Можливо замовити послугу налаштування програми.	Пряма, перший рівень. Продукт і підтримка продається напряму розробником.	Пряма, одного рівня.

Для обслуговування вузького спеціалізованого сегменту ринку використовуємо прямий канал збуту. Розробимо концепцію маркетингових комунікацій, спираючись на обране позиціонування та поведінку груп споживачів.

Таблиця 4.20 – Концепція маркетингових комунікацій

№	Специфіка поведінки цільових клієнтів	Канали комунікацій, якими користуються цільові клієнти	Ключові позиції, обрані для позиціонування	Завдання рекламних повідомлень	Концепція рекламних звернень
1	Перехід за рекламним посиланням, прямий контакт з потенційним клієнтом відділу маркетингу	Канали комунікацій, вказані у самому продукті на контактній сторінці. Можливість віддаленого налаштування службою підтримки.	Точні прогнози результатів спортивних змагань. Безперебійна робота та підтримка. Інтеграція з CRM Microsoft Dynamics та іншими системами клієнта.	Ознайомити потенційних клієнтів з можливостями системи та вигодами її використання саме для певної цільової групи клієнтів.	Розробка рекламних повідомлень з урахуванням специфічних вимог та вигоди цільової групи клієнтів.

Висновки до розділу

В даному розділі пропонується стартап-проект «Програмне забезпечення прогнозу результатів спортивних змагань». Система прогнозує результати на основі даних попередніх ігор і поточних характеристик, використовує декілька найбільш пристосованих моделей аналізу для різних видів спорту.

Сильними сторонами системи є гнучка інтеграція з системами користувача, безперебійна робота та надання результатів прогнозу в режимі реального часу про майбутні спортивні події хокею, баскетболу, бейсболу.

Результати аналізу представлені у поширеній CRM - Microsoft Dynamics. Також можливо інтегрувати дану систему з іншими системами користувача за рахунок API.

Основними споживачами системи є спортивні організації, спортивні клуби та приватні гравці. Прогнозування спортивних результатів корисне для гравців, керівництва команди та аналітиків при виявленні найважливіших чинників, які допомагають досягти виграшних результатів на основі яких можна визначити відповідну тактику.

Завдяки тому, що кількість спортивних організацій та клубів є обмеженою, можливо надавати персональну маркетингову пропозицію кожному потенційному клієнту в цій групі.

Впровадження стартап-проєкту є перспективним завдяки наявності унікальних функцій, що привабливі як спортивним організаціям, так і приватним клієнтам.

ВИСНОВКИ

У ході виконання магістерської дисертації було розглянуто питання пов'язані з архітектурою програмного забезпечення для аналізу результатів спортивних змагань.

На основі даних, отриманих в процесі аналізу, сформульовано задачу створити програмне забезпечення, що буде більш придатним для корпоративних цілей ніж аналоги, а також матиме можливість надавати прогнози за декількома алгоритмами.

Розроблено програмне забезпечення для прогнозування результатів у вигляді сервісів, що інтегровані з системою взаємовідносинами з клієнтами, через яку відбувається взаємодія користувача з розробленим програмним модулем.

Створено та проведено опис стартап-проєкту. Наведено інформаційну карту проєкту, здійснено аналіз потенційних техніко-економічних переваг ідеї. Виконано аналіз ринкових можливостей стартап-проєкту, та технологічний аудит ідеї проєкту. Визначено ринкову стратегію проєкту.

Наукова новизна одержаних результатів магістерської дисертації полягає в тому, що запропоновано архітектуру програмного забезпечення для аналізу результатів спортивних змагань, яка відрізняється від інших реалізацією інтеграції з системою управління взаємовідносинами з клієнтами.

Практична значимість одержаних результатів полягає у розробці програмного забезпечення для аналізу результатів спортивних змагань, яке є адаптованим для застосування у корпоративному секторі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- 1) NHL [Електронний ресурс]: (Стаття) / Wikipedia – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу:
https://en.wikipedia.org/wiki/National_Hockey_League
- 2) NBA [Електронний ресурс]: (Стаття) / Wikipedia – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу:
https://en.wikipedia.org/wiki/National_Basketball_Association
- 3) MLB [Електронний ресурс]: (Стаття) / Wikipedia – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу:
https://en.wikipedia.org/wiki/Major_League_Baseball
- 4) Shanahan, Kathleen Jean. A model for predicting the probability of a win in basketball. // University of Iowa. - 1984. - P.1-10.
- 5) Huang, K.Y., Chang, W.L. A neural network method for prediction of 2006 World Cup football game. // Neural Networks (IJCNN). - IEEE. - 2010. - P.1-8. – ISBN 978-1-4244-6918-5.
- 6) Purucker, M.C. Neural network quarterbacking. //IEEE Potentials. - 1996. - P.9-15.
- 7) Pardee, M. An artificial neural network approach to college football prediction and ranking. // University of Wisconsin. - 1999.
- 8) Alan McCabe, Jarrod Trevathan. Artificial Intelligence in Sports Prediction. // Fifth International Conference on Information Technology: New Generations (ITNG 2008). - 2008.
- 9) Loeffelholz, B., Bednar, E., Bauer, K.W. Predicting NBA games using neural networks. // Journal of Quantitative Analysis in Sports. - 2009. P.1-17.
- 10) Rory Bunkera, Teo Susnjak. The Application of Machine Learning Techniques for Predicting Results in Team Sport: A Review. // Sport Result Prediction. - 2019.

- 11) Машинне навчання [Електронний ресурс]: (Стаття) / Wikipedia – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу: https://en.wikipedia.org/wiki/Machine_learning
- 12) Машинне навчання для прогнозування тенісу [Електронний ресурс]: (Стаття) / Habr – Електрон. дан. (1 файл). – 2016. – Режим доступу до ресурсу: <https://habr.com/post/307422/>
- 13) Linear Regression [Електронний ресурс]: (Стаття) / Yale – Електрон. дан. (1 файл). – 1997. – Режим доступу до ресурсу: <http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm>
- 14) Пережовуючи логістичну регресію – [Електронний ресурс] – Режим доступу до ресурсу: [Електронний ресурс]: (Стаття) / Habr – Електрон. дан. (1 файл). – 2020. – Режим доступу до ресурсу: <https://habr.com/post/485872/>
- 15) Градієнтний спуск [Електронний ресурс]: (Стаття) / Wikipedia – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу: https://en.wikipedia.org/wiki/Gradient_descent
- 16) Стохастичний градієнтний спуск [Електронний ресурс]: (Стаття) / Wikipedia – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу: https://en.wikipedia.org/wiki/Stochastic_gradient_descent
- 17) Основи штучних нейронних мереж [Електронний ресурс]: (Стаття) / Neuralnet – Електрон. дан. (1 файл). – 2017. – Режим доступу до ресурсу: <https://neuralnet.info/chapter/основы-инс>
- 18) Метод зворотнього розповсюдження помилки [Електронний ресурс]: (Стаття) / Habr – Електрон. дан. (1 файл). – 2013. – Режим доступу до ресурсу: <https://habr.com/post/198268/>
- 19) Баєсові класифікатори [Електронний ресурс]: (Стаття) / Craftappmobile – Електрон. дан. (1 файл). – 2019. – Режим доступу до ресурсу: <https://craftappmobile.com/naive-bayes-classifier-and-bayes-classifiers/>

20) Наївний баєсів класифікатор [Електронний ресурс]: (Стаття) / Towardsdatascience – Електрон. дан. (1 файл). – 2018. – Режим доступу до ресурсу: <https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c>

21) Машина опорних векторів [Електронний ресурс]: (Стаття) / Coderlessons – Електрон. дан. (1 файл). – 2018. – Режим доступу до ресурсу: <https://coderlessons.com/tutorials/python-technologies/uznaite-mashinnoe-obuchenie-s-python/ml-mashina-opornykh-vektorov-svm>

22) SVM [Електронний ресурс]: (Стаття) / Habr – Електрон. дан. (1 файл). – 2020. – Режим доступу до ресурсу: <https://habr.com/company/ods/blog/484148/>

23) Що таке дерево рішень [Електронний ресурс]: (Стаття) / Habr – Електрон. дан. (1 файл). – 2020. – Режим доступу до ресурсу: <https://habr.com/company/productstar/blog/523044/>

24) Типи архітектур ПЗ [Електронний ресурс]: (Стаття) / Habr – Електрон. дан. (1 файл). – 2018. – Режим доступу до ресурсу: <https://habr.com/company/1cloud/blog/424911/>

25) SOA архітектура [Електронний ресурс]: (Стаття) / Arhitekturainformacionnyhsist – Електрон. дан. (1 файл). – 2017. – Режим доступу до ресурсу: <https://sites.google.com/site/arhitekturainformacionnyhsist/lekcia-7-servis-orientirovanna-arhitektura-soa>

26) Переваги CRM систем [Електронний ресурс]: (Стаття) / Envybox – Електрон. дан. (1 файл). – 2020. – Режим доступу до ресурсу: <https://envybox.io/blog/preimushhestva-crm-sistem/>

27) Мова C# [Електронний ресурс]: (Стаття) / Microsoft Docs – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу: <https://docs.microsoft.com/ru-ru/dotnet/csharp/tour-of-csharp/>

28) ASP.NET Огляд [Електронний ресурс]: (Стаття) / Microsoft Docs – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу: <https://docs.microsoft.com/ru-ru/aspnet/overview>

29) Microsoft Dynamics CRM [Електронний ресурс]: (Стаття) / Wikipedia – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу:

https://uk.wikipedia.org/wiki/Microsoft_Dynamics_CRM

30) Бібліотека ML.NET [Електронний ресурс]: (Стаття) / Microsoft Docs – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу:

<https://docs.microsoft.com/ru-ru/dotnet/machine-learning/how-does-mldotnet-work>

31) Dynamics 365. Використання мови Javascript [Електронний ресурс]: (Стаття) / Microsoft Docs – Електрон. дан. (1 файл). – 2021. – Режим доступу до ресурсу: <https://docs.microsoft.com/en-us/dynamics365/customerengagement/on-premises/developer/use-javascript?view=op-9-1>

ДОДАТКИ

ДОДАТОК А

Лістинг коду

new_actions.js

```

function getPredictionsByMatch(){
    debugger;
    var req = new XMLHttpRequest();
    req.open("POST", Xrm.Page.context.getClientUrl() + "/api/data/v9.2/new_GetPredictionsByMatch",
true);
    req.setRequestHeader("OData-MaxVersion", "4.0");
    req.setRequestHeader("OData-Version", "4.0");
    req.setRequestHeader("Accept", "application/json");
    req.setRequestHeader("Content-Type", "application/json; charset=utf-8");
    req.onreadystatechange = function() {
        if (this.readyState === 4) {
            req.onreadystatechange = null;
            if (this.status === 200 || this.status === 204) {
                Xrm.Utility.alertDialog("Predictions are generated");
                window.parent.Xrm.Page.data.refresh(true).then(function()
                {}),
function(error) {alert(error.message);});
            } else {
                Xrm.Utility.alertDialog(this.statusText);
            }
        }
    };
    req.send();
}

```

SportsPredictions.Plugins GetPredictionsByMatch.cs

```

using System;
using System.Collections.Generic;
using System.Linq;
using System.Net.Http;
using System.Net.Http.Headers;
using System.ServiceModel;
using System.Text;
using System.Threading.Tasks;
using Microsoft.Xrm.Sdk;
using Newtonsoft.Json;

namespace SportsPredictions.Plugins
{
    public class GetPredictionsByMatch : IPlugin
    {
        public void Execute(IServiceProvider serviceProvider)
        {
            IPluginExecutionContext context =
(IPluginExecutionContext)serviceProvider.GetService(typeof(IPluginExecutionContext));
            IOrganizationServiceFactory serviceFactory =
(IOrganizationServiceFactory)serviceProvider.GetService(typeof(IOrganizationServiceFactory));
            IOrganizationService service = serviceFactory.CreateOrganizationService(context.UserId);
            try
            {

```

```

Guid matchid = (Guid)context.InputParameters["matchid"];

string url = Config.PredictionBaseAddress + "/predictmatch";

var client = new HttpClient();

var match = new {MatchId = matchid};

var json = JsonConvert.SerializeObject(match);
var data = new StringContent(json, Encoding.UTF8, "application/json");

string username = Config.IntegrationUserName;
string password = Config.IntegrationPassword;

var byteArray = Encoding.ASCII.GetBytes($"{username}:{password}");
client.DefaultRequestHeaders.Authorization = new
System.Net.Http.Headers.AuthenticationHeaderValue("Basic", Convert.ToBase64String(byteArray));

    var response = client.PostAsync(url, data).Result;
    }
    catch (FaultException<OrganizationServiceFault> ex)
    {
        throw new InvalidPluginExecutionException($"An error occurred in
SportsPredictions.Plugins.GetPredictionsByMatch plugin. {ex?.Message}", ex);
    }
}
}
}
}

```

SportsPredictions.Predictor HomeController.cs

```

using Microsoft.Xrm.Sdk;
using System;
using System.Collections.Generic;
using System.Configuration;
using System.Web;
using System.Web.Http;
using System.Web.Mvc;
using Microsoft.Xrm.Sdk.Query;
using Microsoft.Xrm.Tooling.Connector;
using SportsPredictions.Predictor.Repositories;
using SportsPredictions.Predictor.Services;

namespace SportsPredictions.Predictor.Controllers
{
    public class HomeController : Controller
    {
        private readonly MatchRepository _matchRepository;
        private readonly PredictionRepository _predictionRepository;
        private readonly SportsTeamRepository _sportsTeamRepository;
        private readonly TeamParametersRepository _teamParamsRepository;
        private readonly PredictionModelRepository _predictionModelRepository;

        public HomeController()
        {
            string connectionString = ConfigurationManager.ConnectionStrings["crm"].ToString();

```

```

CrmServiceClient crmService = new CrmServiceClient(connectionString);

_matchRepository = new MatchRepository(crmService);
_predictionRepository = new PredictionRepository(crmService);
_sportsTeamRepository = new SportsTeamRepository(crmService);
_teamParamsRepository = new TeamParametersRepository(crmService);
_predictionModelRepository = new PredictionModelRepository(crmService);
}

[Authorize]
[HttpPost]
public ActionResult PredictMatch([FromBody] dynamic match)
{
    try
    {
        Guid matchId = match.MatchId;
        PredictMatch(matchId);
        var response = new { Status = 200, Message = "Predictions created" };
        return Json(response);
    }
    catch (Exception ex)
    {
        throw new Exception(ex.Message);
    }
}

[Authorize]
[HttpPost]
public ActionResult PredictMatches([FromBody] List<Guid> matches)
{
    try
    {
        foreach (var match in matches)
        {
            PredictMatch(match);
        }

        var response = new { Status = 200, Message = "Predictions created" };
        return Json(response);
    }
    catch (Exception ex)
    {
        throw new Exception(ex.Message);
    }
}

[Authorize]
[HttpPost]
public ActionResult PredictNextMatchOfTeam([FromBody] dynamic team)
{
    try
    {
        Guid teamId = team.teamId;

        EntityReference nextMatch = _matchRepository.GetNextMatchOfTeam(teamId);

        Guid nextMatchId = nextMatch.Id;

```

```

        PredictMatch(nextMatchId);

        var response = new { Status = 200, Message = "Predictions created" };
        return Json(response);
    }
    catch (Exception ex)
    {
        throw new Exception(ex.Message);
    }
}

[Authorize]
[HttpPost]
public ActionResult PredictMatchesOfDate([FromBody] dynamic data)
{
    try
    {
        DateTime matchesDate = data.DateTimeOfMatches;

        List<Entity> matches = _matchRepository.GetMatchesOfDate(matchesDate);

        foreach (var match in matches)
        {
            PredictMatch(match);
        }
        var response = new { Status = 200, Message = "Predictions created" };
        return Json(response);
    }
    catch (Exception ex)
    {
        throw new Exception(ex.Message);
    }
}

private void PredictMatch(Guid matchId)
{
    Entity matchEntity = _matchRepository.GetMatchById(matchId);

    EntityReference homeTeam = matchEntity.GetAttributeValue<EntityReference>("new_hometeamid");
    EntityReference awayTeam = matchEntity.GetAttributeValue<EntityReference>("new_awayteamid");

    Entity homeTeamParams = _teamParamsRepository.GetTeamParamsForMatch(homeTeam.Id, matchId);
    Entity awayTeamParams = _teamParamsRepository.GetTeamParamsForMatch(awayTeam.Id, matchId);

    var predictor = new PredictionWorker();

    List<Entity> availablePredictionModels = _predictionModelRepository.GetCurrentPredictionModels();

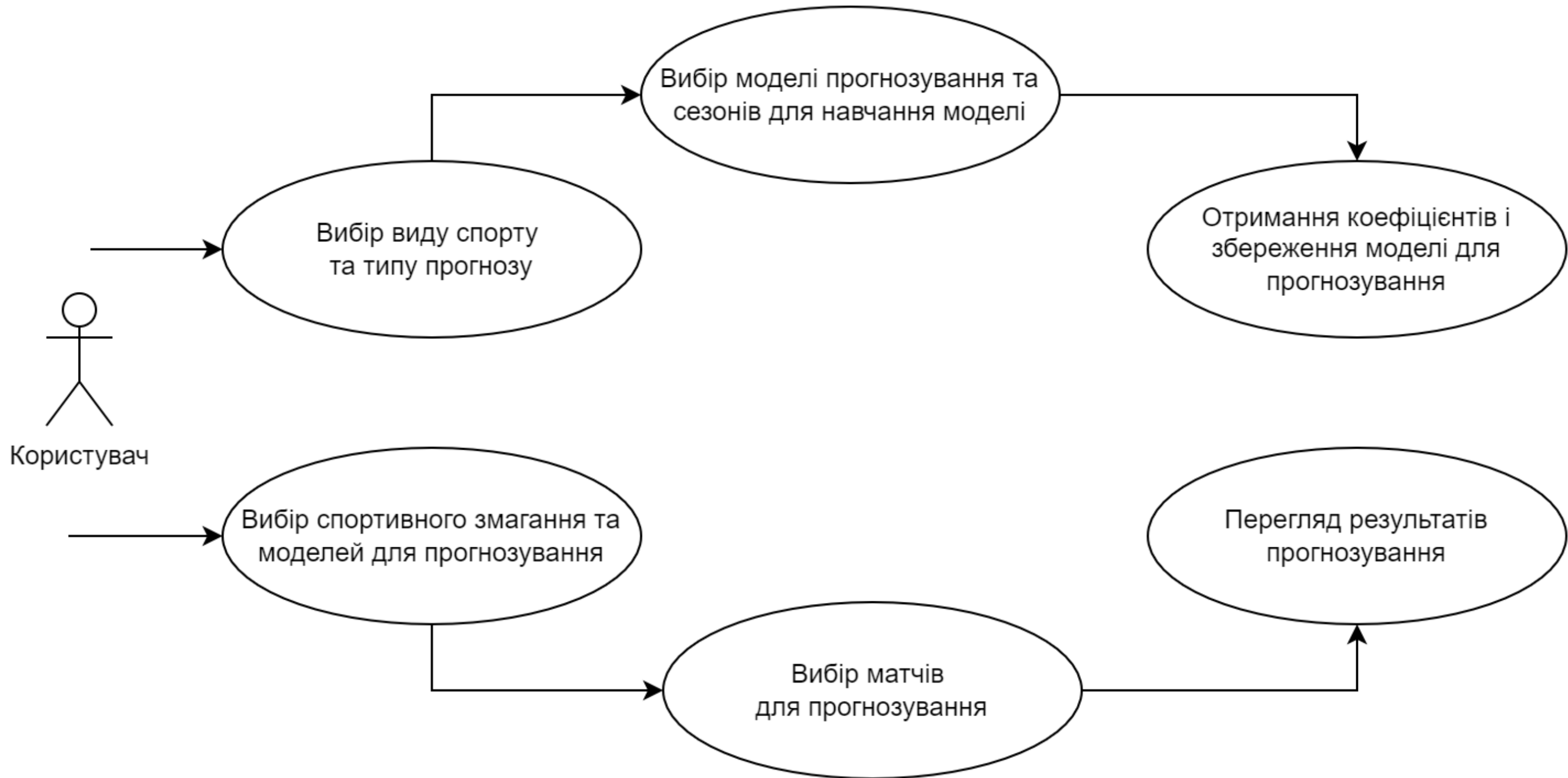
    List<Entity> predictions = new List<Entity>();

    foreach (var model in availablePredictionModels)
    {
        Entity prediction = predictor.GetPrediction(model, matchEntity, homeTeamParams, awayTeamParams);
        _predictionRepository.SaveNewPrediction(prediction);
    }
}
}
}

```

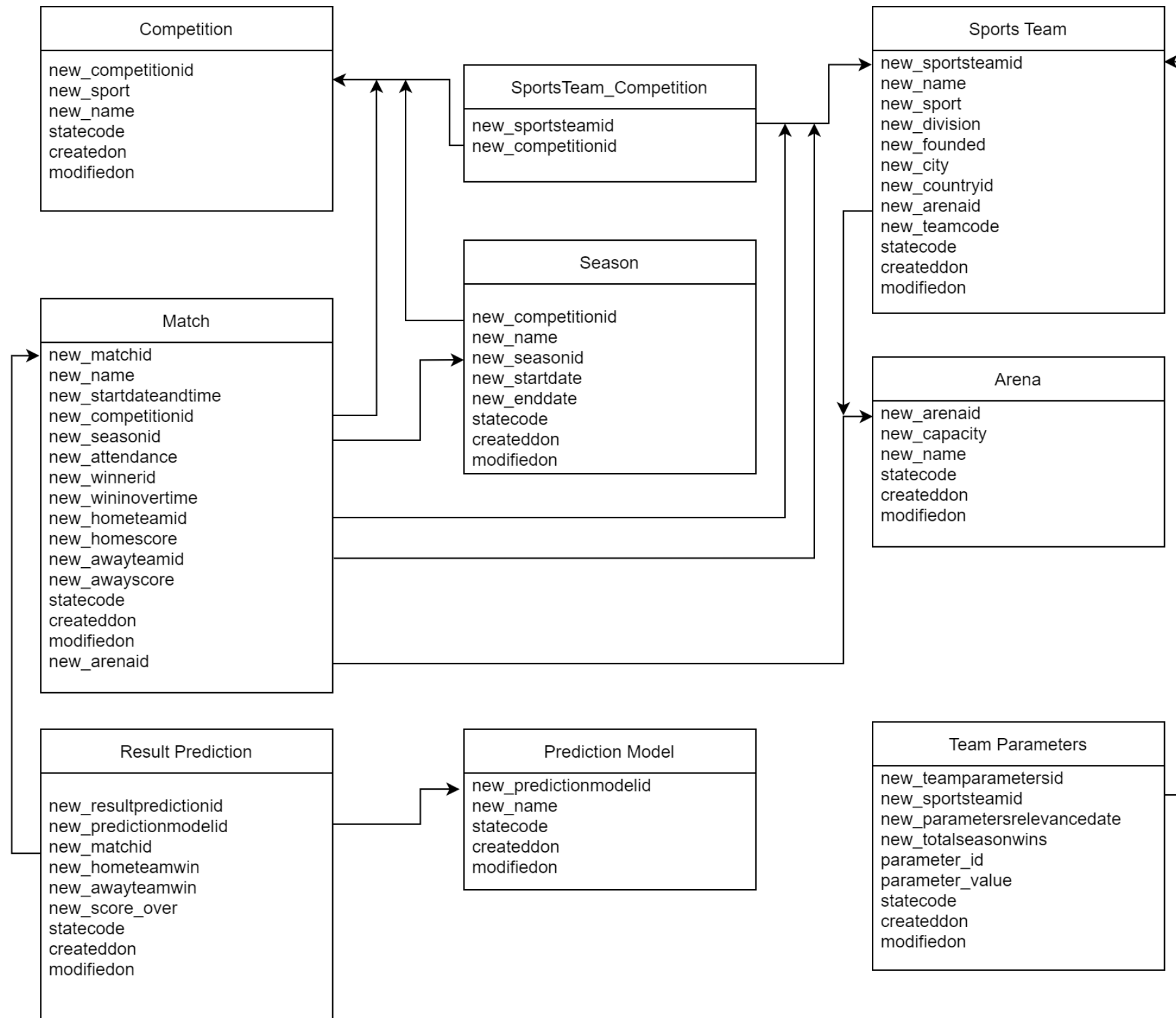
ДОДАТОК Б

Структурна схема варіантів використання



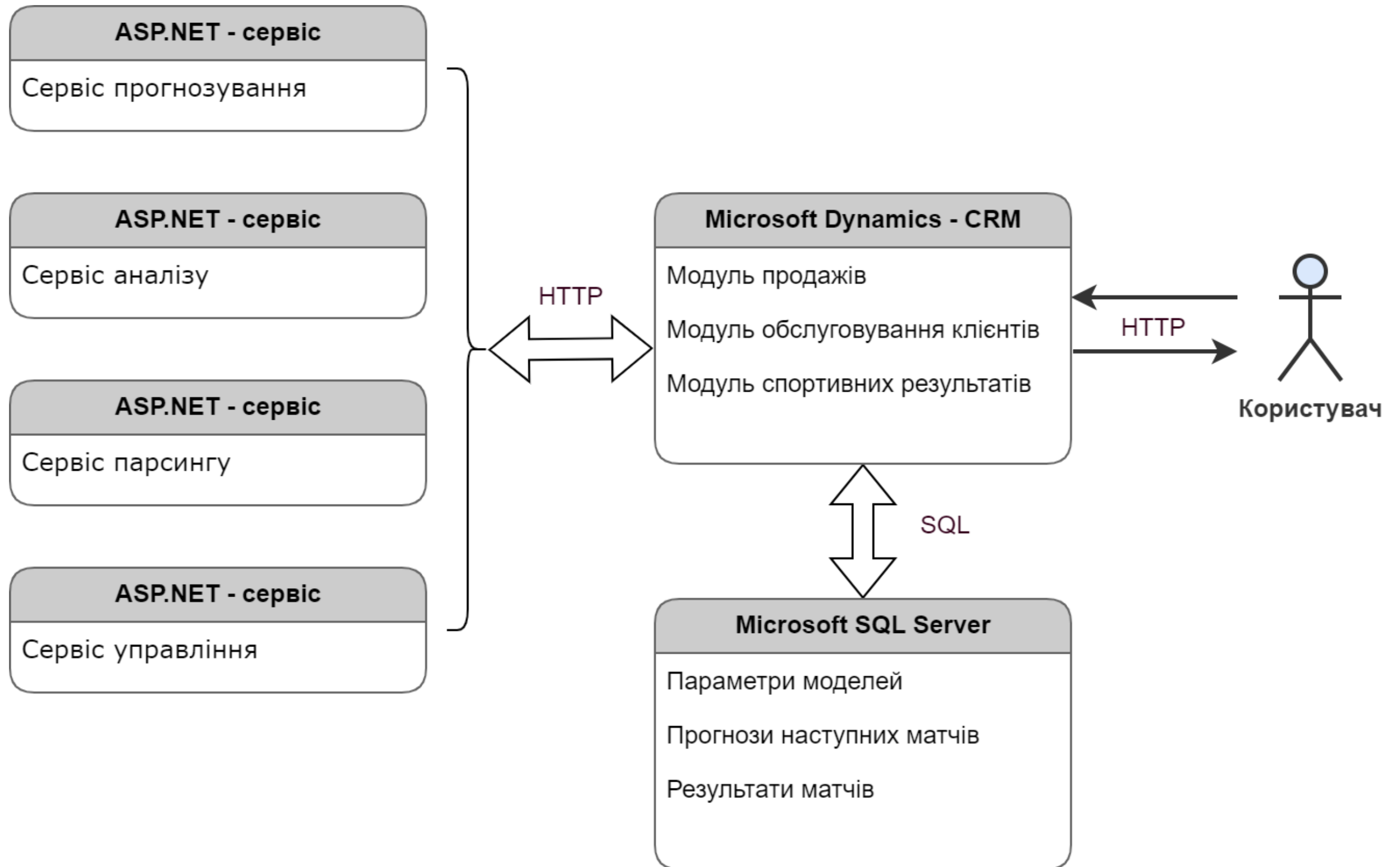
ДОДАТОК В

Структурна схема класів програмного забезпечення



ДОДАТОК Г

Узагальнена схема архітектури



ДОДАТОК Д

Екранні форми

This screenshot shows the 'MLB Teams' view in the Dynamics 365 SportsPrediction application. The left-hand navigation pane is expanded to 'Sports Teams'. The main content area displays a table of MLB teams with columns for Name, Sport, Founded, City, Arena, and Division. The table lists 10 teams, including Arizona Diamondbacks, Atlanta Braves, Baltimore Orioles, Boston Red Sox, Chicago Cubs, Chicago White Sox, Cincinnati Reds, and Cleveland Indians. The interface includes a search bar, a toolbar with options like 'Show Chart', 'New', 'Delete', 'Refresh', and 'Email a Link', and a footer showing '1 - 30 of 30' items on 'Page 1'.

Name	Sport	Founded	City	Arena	Division
Arizona Diamondbacks	Baseball	1998	Phoenix, ...	Chase Field	West
Atlanta Braves	Baseball	1871	Cumberla...	Truist Park	East
Baltimore Orioles	Baseball	1901	Baltimore...	Oriole Pa...	East
Boston Red Sox	Baseball	1901	Boston, ...	Fenway P...	East
Chicago Cubs	Baseball	1874	Chicago, ...	Wrigley F...	Central
Chicago White Sox	Baseball	1901	Chicago, ...	Guarante...	Central
Cincinnati Reds	Baseball	1882	Cincinnati...	Great Am...	Central
Cleveland Indians	Baseball	1901	Cleveland...	Progressi...	Central

This screenshot shows the 'Active Matches' view in the Dynamics 365 SportsPrediction application. The left-hand navigation pane is expanded to 'Matches'. The main content area displays a table of active matches with columns for Start Date, Competition, Arena, Home Team, Away Team, Home Score, and Away Score. The table lists 10 matches, including NHL games between Los Angeles and Arizona, Seattle and Washington, Vancouver and Chicago, New York and Toronto, Boston and Calgary, New York and Buffalo, Tampa Bay and Minnesota, San Jose and Washington, and Edmonton and Chicago. The interface includes a search bar, a toolbar with options like 'Show Chart', 'New', 'Delete', 'Refresh', and 'Email a Link', and a footer showing '1 - 38 of 38' items on 'Page 1'.

Start Date	Competition	Arena	Home Team	Away Team	Home Score	Away Score
22.11.202...	NHL	Staples Ce...	Los Angel...	Arizona C...	1	2
22.11.202...	NHL	Climate Pl...	Seattle Kr...	Washingt...	5	2
22.11.202...	NHL	Rogers Ar...	Vancouve...	Chicago B...	0	1
22.11.202...	NHL	UBS Arena	New York ...	Toronto ...	0	3
22.11.202...	NHL	TD Garden	Boston Br...	Calgary Fl...	0	4
22.11.202...	NHL	Madison ...	New York ...	Buffalo Sa...	5	4
22.11.202...	NHL	Amalie Ar...	Tampa Ba...	Minnesot...	5	4
21.11.202...	NHL	SAP Center	San Jose ...	Washingt...	0	4
21.11.202...	NHL	Rogers Pl...	Edmonto...	Chicago B...	5	2

This screenshot shows the 'Active Competitions' view in the Dynamics 365 SportsPrediction application. The left-hand navigation pane is expanded to 'Competitions'. The main content area displays a table of active competitions with columns for Name, Sport, and Modified On. The table lists 3 competitions: MLB (Baseball), NBA (Basketball), and NHL (Ice Hockey). The interface includes a search bar, a toolbar with options like 'Show Chart', 'New', 'Delete', and 'Refresh', and a footer showing '1 - 3 of 3' items on 'Page 1'.

Name	Sport	Modified On
MLB	Baseball	19.11.2021 1...
NBA	Basketball	19.11.2021 1...
NHL	Ice Hockey	19.11.2021 1...

This screenshot shows the 'Active Result Predictions' view in the Dynamics 365 SportsPrediction application. The left-hand navigation pane is expanded to 'Result Predictions'. The main content area displays a table of active result predictions with columns for Match, Prediction Model, Home Team Win %, Away Team Win %, and Predicted Over, % . The table lists 2 predictions for the Anaheim-E... match, using Logistic Regression and Linear Regression models. The interface includes a search bar, a toolbar with options like 'Show Chart', 'New', 'Delete', 'Refresh', and 'Email a Link', and a footer showing '1 - 2 of 2' items on 'Page 1'.

Match	Prediction Model	Home Team Win %	Away Team Win %	Predicted Over, %
Anaheim-E...	Logistic Regress...	67	33	42
Anaheim-E...	Linear Regression	62	38	45



Имя пользователя:
Лісовиченко Олег Іванович

ID проверки:
1009402992

Дата проверки:
29.11.2021 12:59:36 EET

Тип проверки:
Doc vs Internet + Library

Дата отчета:
29.11.2021 13:00:08 EET

ID пользователя:
76913

Название файла: IT_04мп_Репалюк_ПЗ

Количество страниц: 49 Количество слов: 10342 Количество символов: 76528 Размер файла: 107.84 KB ID файла: 1009421749

2.66% Совпадения

Наибольшее совпадение: 0.44% с источником из Библиотеки (ID файла: 1005770804)

0.64% Источники из Интернета 45 Страница 51

2.61% Источники из Библиотеки 190 Страница 51

0% Цитат

Исключение цитат выключено

Исключение списка библиографических ссылок выключено

0% Исключений

Нет исключенных источников

Модификации

Обнаружены модификации текста. Подробная информация доступна в онлайн-отчете.

Замененные символы 6