

Вінницький національний технічний університет  
Факультет інформаційних технологій та комп'ютерної інженерії  
Кафедра комп'ютерних наук

## **Пояснювальна записка**

до магістерської кваліфікаційної роботи  
на тему:

### **Інформаційна технологія прогнозування результатів матчів футбольної молодшої ліги на основі нечітких правил**

Виконав: студент 2 курсу, групи 2КН-18м  
Спеціальності 122  
«Комп'ютерні науки»

Когут В.В.  
(прізвище та ініціали)

Керівник к.т.н., доц. Сілагін О. В.  
(прізвище та ініціали)

Рецензент к.т.н., доц. каф. ПЗ Войтко В.В.  
(прізвище та ініціали)

Вінниця - 2019 року

ЗАТВЕРДЖУЮ

Завідувач кафедри \_\_КН

д. т. н., проф. Яровий А.А.

(наук. ст., вч. зв., ініц. та прізви.)

(підпис)

" \_\_\_\_ " \_\_\_\_\_ 2019 р.

## ЗАВДАННЯ

на магістерську кваліфікаційну роботу на здобуття кваліфікації магістра зі спеціальності: 122 – «Комп'ютерні науки»

---

08-22.МКР.030.18.000.ПЗ

Магістранта групи 1КН-18м Когута Владислава Валентиновича

Тема магістерської кваліфікаційної роботи: «Інформаційна технологія прогнозування результатів матчів футбольної молодшої ліги на основі нечітких правил»

Вхідні дані: джерело формування вектора вхідних змінних - Веб портал із результатами футбольних матчів молодшої ліги; кількість термів вхідної лінгвістичної змінної (2-7); кількість термів вихідної лінгвістичної змінної (3) ; кількість команд навчальної тестової вибірки – не менше 50; відповідність стандарту ODBS; Інструментарій для моделювання - FUZZY TOOLBOX середовища MATLAB;

Короткий зміст частин магістерської кваліфікаційної роботи:

1. Графічна: загальний вигляд аналогів розроблюваної системи; узагальнена структура системи нечіткого логічного висновку; блок-схема системи нечіткого логічного висновку з кількісними змінними; терм-множина вхідних та вихідних змінних; нечітка база знань; схема алгоритму прогнозування результатів футбольного матчу молодшої ліги; результуюча турнірна таблиця.

2. Текстова (пояснювальна записка): Вступ. Обґрунтування доцільності розробки системи прогнозування результатів футбольних матчів молодшої ліги на основі нечітких правил. Аналіз предметної області. Класифікація методів прогнозування результатів діяльності на основі нечітких правил. Характеристика і аналіз аналогів. Проектування системи. Обґрунтування вибору методики прогнозування результатів футбольних матчів молодшої ліги на основі нечітких правил. Вибір стеку технологій. Реалізація системи. Розробка нечіткої бази знань. Формування нечітких правил. Тестування роботи системи. Висновки. Перелік посилань. Додатки.

## КАЛЕНДАРНИЙ ПЛАН ВИКОНАННЯ МКР

№ етапу	Назва етапу	Термін виконання		Очікувані результати
		початок	кінець	
1	Обґрунтування доцільності розробки системи прогнозування результатів футбольних матчів молодшої ліги на основі нечітких правил.			Аналітичний огляд літературних джерел, задачі досліджень, розділ 1 ПЗ
2	Аналіз методів прогнозування результатів діяльності на основі нечітких правил. Моделювання			Модель, інформаційна технологія, розділ 2
3	Проектування та реалізація програмного додатку.			Програмне забезпечення, розділ 3
4	Економічна частина			розділ 4
5	Тестування роботи системи.			тези доповідей/акт впровадження
6	Оформлення пояснювальної записки, графічного матеріалу та презентації			Пояснювальна записка, графічний матеріал, презентація

Консультанти з окремих розділів магістерської кваліфікаційної роботи

1. Науковий керівник \_\_\_\_\_ канд. техн. наук, доц., доц. кафедри КН

(підпис)

наук. ступінь, вчене звання (посада)

“ \_\_\_\_\_ ” \_\_\_\_\_ 20\_\_ р.

\_\_\_\_\_ О. В. Сілагін

ініціали та прізвище

2. Економічна частина \_\_\_\_\_ канд.ек.наук,доц.,доц. кафедри ЕПВМ

(підпис)

наук. ступінь, вчене звання (посада)

\_\_\_\_\_ М.В.Бальзан \_\_\_\_\_

ініціали та прізвище

“ \_\_\_\_\_ ” \_\_\_\_\_ 20\_\_ р.

Дата попереднього захисту роботи “ \_\_\_\_\_ ” \_\_\_\_\_ 20\_\_ р.

Рецензент \_\_\_\_\_ канд. техн. наук, доц., доц. кафедри  
ПЗ

(підпис)

наук. ступінь, вчене звання (посада)

\_\_\_\_\_ В. В. Войтко \_\_\_\_\_

ініціали та прізвище

Завдання видав

науковий керівник \_\_\_\_\_ канд. техн. наук, доц., доц. кафедри  
КН

(підпис)

наук. ступінь, вчене звання (посада)

\_\_\_\_\_ О. К. Сілагін \_\_\_\_\_

ініціали та прізвище

“ \_\_\_\_\_ ” \_\_\_\_\_ 20\_\_ р.

Завдання отримав магістрант \_\_\_\_\_ Когут В.В. \_\_\_\_\_

(підпис)

ініціали та прізвище

“ \_\_\_\_\_ ” \_\_\_\_\_ 20\_\_ р.

## АНОТАЦІЯ

В магістерській кваліфікаційній роботі “Інформаційна технологія прогнозування результатів матчів футбольної молодшої ліги на основі нечітких правил” досліджується можливість застосування апарату нечіткої логіки до задачі прогнозування результатів футбольних матчів на основі нечітких правил на основі експериментальних даних. Використання експертно-лінгвістичних закономірностей, що формалізуються засобами нечіткої логіки, та налаштування методом порівняння з навчальною вибіркою з мінімізацією середньо-квадратичного відхилення дозволяє отримати досить високу достовірність прогнозування в умовах невеликих експериментальних вибірок. Результати роботи можуть бути використані для широкого кола задач пов’язаних з прогнозуванням спортивних подій.

Для розробки модуля використанні інструменти Fuzzy Logic Toolbox програмного середовища MatLab.

## **ABSTRAKT**

The master's qualification work "Information technology of prediction of football junior league matches based on fuzzy rules" explores the possibility of applying the apparatus of fuzzy logic to the task of predicting the results of football matches based on fuzzy rules based on experimental data. The use of expert-linguistic regularities formalized by means of fuzzy logic and tuning by comparison with a training sample with minimization of the mean-square deviation allows to obtain a sufficiently high prediction accuracy in the conditions of small experimental samples. The results of the work can be used for a wide range of tasks related to the prediction of sports events.

To develop the module use the tools Fuzzy Logic Toolbox MatLab software environment.

## ЗМІСТ

ВСТУП.....	6
1. АНАЛІЗ ПРОБЛЕМ ТА МЕТОДОЛОГІЇ ПРОГНОЗУВАННЯ ФУТБОЛЬНИХ МАТЧІВ.....	9
1.1 Задачі і принципи прогнозування .....	9
1.2 Аналіз та обґрунтування вибору методу прогнозування.....	11
1.3 Огляд відомих технічних рішень задачі прогнозування результатів футбольних матчів.....	27
1.4 Постановка задачі та формулювання вимог.....	31
1.5 Висновок.....	31
2 МОДЕЛЮВАННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ПРОГНОЗУВАННЯ РЕЗУЛЬТАТІВ ФУТБОЛЬНИХ МАТЧІВ.....	32
2.1 Прогнозування результатів футбольних матчів як об'єкт моделювання.....	32
2.2 Вибір методології моделювання для задачі прогнозування результатів футбольних матчів.....	35
2.2.1 Аналіз можливостей нечіткої логіки для покращення інформаційної технології прогнозування.....	36
2.2.2 Аналіз доцільності моделювання на основі методу опорних векторів (SVM).....	42
2.2.3 Аналіз можливості створення моделі на основі нейронних мереж.....	46
2.3 Моделювання інформаційної технології прогнозування результатів футбольних матчів засобами нечіткої логіки.....	48
2.3.1 Формалізація моделі прогнозування результатів футбольних матчів..	51
2.3.2 Моделювання в середовищі MATLAB та робота з модулем Fuzzy Logic Toolbox.....	54
2.4 Висновок.....	64
3 ПРОГРАМНА РЕАЛІЗАЦІЯ ТЕХНОЛОГІЇ.....	65
3.1 Збір вхідних даних про результати футбольних матчів молодшої ліги..	65
3.2 Розробка та реалізація бази знань.....	67



3.3 Розробка алгоритму прогнозування футбольних матчів.....	70
3.4 Тестування роботи програмного забезпечення.....	74
3.5 Висновок.....	76
4. ЕКОНОМІЧНА ЧАСТИНА.....	77
4.1 Оцінювання комерційного потенціалу розробки.....	77
4.2 Прогнозування витрат на виконання науково-дослідної роботи та конструкторсько-технологічної роботи.....	78
4.3 Прогнозування комерційних ефектів від реалізації результатів розробки.....	81
4.4 Розрахунок ефективності вкладених інвестицій та період їх окупності.....	83
4.5 Висновок.....	86
ВИСНОВКИ.....	87
ПЕРЕЛІК ПОСИЛАНЬ.....	89
ДОДАТОК А.....	92
ДОДАТОК Б.....	96
ДОДАТОК В.....	111
ДОДАТОК Г.....	118

## ВСТУП

На сьогоднішній день сфера беттингу та букмекерства, що пов'язана з прогнозуванням спортивних подій, користується великою популярністю у широкого кола прихильників спорту. З розвитком методів штучного інтелекту стало можливим прогнозувати результати спортивних подій за допомогою відповідних інтелектуальних технологій, одержуючи таким чином значно більшу достовірність прогнозу результату, ніж виходячи з суб'єктивних припущень футбольних експертів.

Проблемами оптимізації моделей для побудови точних прогнозів результатів футбольних матчів займаються провідні школи науковці Сполучених Штатів Америки, Великої Британії, Німеччини, Японії та інших країн. Підходи та методи, що використовуються у цих дослідженнях мають ряд недоліків, через що постає гостра потреба у їх постійній модернізації та пошуку шляхів вирішення ключових проблем.

Дане дослідження може бути корисним у якості інструменту вибору правильної стратегії формування футбольних ставок. Також воно буде корисним для букмекерів, як методологія більш точного розрахунку ігрових коефіцієнтів на спортивні події. Все це робить подальші дослідження в області прогнозування результатів спортивних подій достатньо актуальними.

**Зв'язок роботи з науковими програмами, планами, темами.** Магістерська робота виконана відповідно до напрямку наукових досліджень кафедри комп'ютерних наук Вінницького національного технічного університету 22 К1 «Моделі, методи, технології та пристрої інтелектуальних інформаційних систем управління, економіки, навчання та комунікацій» та плану наукової та навчально-методичної роботи кафедри. Ця робота виконується також в рамках співробітництва між кафедрою КН ВНТУ та ТОВ «ІТІ», а результати плануються до впровадження на ТОВ «ІТІ»

**Мета та завдання дослідження.** Мета дослідження – збільшення точності прогнозування результатів футбольних матчів молодшої ліги на

основі нечітких правил. Для досягнення поставленої мети необхідно вирішити такі завдання:

- Проаналізувати та вибрати методологію прогнозування для задачі прогнозування результатів футбольних матчів;
- проаналізувати існуючі інтелектуальні технології ідентифікації задачі прогнозування результатів футбольних матчів;
- сформулювати вимоги до роботи технології та розробити ТЗ;
- допрацювати існуючу технологію та модель ідентифікації задачі прогнозування результатів футбольних матчів;
- провести математичне моделювання роботи етапів допрацьованої технології з використанням апарату нечіткої логіки;
- провести фазифікацію розроблених моделей;
- розробити та наповнити базу знань у вигляді матриць з правилами ЯКЩО-ТО;
- на основі розробленої технології виконати проектування модуля прогнозування результатів футбольних матчів;
- реалізувати та налаштувати роботу модуля прогнозування результатів футбольних матчів;
- протестувати роботу налаштованого модуля;
- виконати задачі економічного розділу.

**Об’єкт дослідження** – це процес прогнозування результатів футбольних матчів.

**Предмет дослідження** – інформаційні технології, математичні моделі, алгоритми та програмні засоби для прогнозування результатів футбольних матчів.

**Як методи дослідження** в роботі застосовано теорію нечітких множин (нечіткі рівняння, нечітка логіка), продукційна база знань, нечітке логічне виведення.

**Наукова новизна одержаних результатів** полягає в наступному:

- доопрацьовано інформаційну технологію прогнозування результатів футбольних матчів, яка відрізняється від існуючої інтелектуальної технології ідентифікації. Заде тим, що при створенні бази знань розширено перелік впливаючих факторів;
- розроблена та фазифікована спеціалізована математична модель прогнозування результатів футбольних матчів, що орієнтована на доопрацьовану технологію.

**Практичне значення одержаних результатів** полягає у наступному:

- розроблено алгоритм нечіткого логічного виведення прогнозованих результатів футбольних матчів;
- розроблено алгоритм нечіткої логічної апроксимації;
- реалізовано експериментальний програмний засіб.

**Достовірність теоретичних положень** магістерської кваліфікаційної роботи підтверджується строгістю постановки задач, коректним застосуванням математичних методів під час доведення наукових положень, строгим виведенням аналітичних співвідношень, порівнянням результатів, отриманих за допомогою розроблених у роботі методів, з відомими, та збіжністю результатів математичного моделювання з результатами, що отримані під час тестування розроблених програмних засобів.

**Особистий внесок магістранта.** Усі результати отримано автором самостійно. У працях, опублікованих у співавторстві, магістранту належать:

[1] – інформаційна технологія прогнозування результатів футбольних матчів;

**Апробація результатів роботи.** Результати досліджень апробовано в доповіді на НПК «МОЛОДЬ В НАУЦІ: ДОСЛІДЖЕННЯ, ПРОБЛЕМИ, ПЕРСПЕКТИВИ – 2020.

**Публікації.** За результатами магістерської кваліфікаційної роботи опубліковано: тези доповіді конференції [1].

Результати, одержані в процесі виконання магістерської кваліфікаційної роботи, плануються до впровадження в розробки науково-виробничого підприємства ТОВ «ІТІ».

# 1. АНАЛІЗ ПРОБЛЕМ ТА МЕТОДОЛОГІЇ ПРОГНОЗУВАННЯ ФУТБОЛЬНИХ МАТЧІВ

## 1.1 Задачі і принципи прогнозування

Прогноз – це уявлення про стан якого-небудь явища в майбутньому на основі спеціально наукового дослідження. Класифікація прогнозів здійснюється, як правило, по двох ознаках - тимчасовий і функціональний. По тимчасовій ознаці розрізняють прогнози: коротко, середньо, довгострокові і наддовгострокові. Функціональна класифікація прогнозів припускає їх розподіл на дослідницькі, програмні і ресурсні [8].

Прогнозування - процес розробки прогнозів. Залежно від виду прогнозу розрізняють нормативне, пошукове, оперативне.

Прогнозна модель - модель об'єкту прогнозування, дослідження якої дозволяє одержати інформацію про можливі стани об'єктах в майбутньому і (або) шляхах і термінах їх здійснення [8].

Залежно від того, яка задача розв'язується в першу чергу, розрізняють два види прогнозування: дослідницьке (або пошукове) і нормативне.

Формування прогнозу об'єктивно існуючих тенденцій розвитку на основі аналізу історичних тенденцій називається дослідницьким або пошуковим прогнозуванням. Цей вид прогнозування заснований на використуванні принципу інерційності розвитку, при якому орієнтація прогнозу в часі відбувається по схемі "від теперішнього часу, - до майбутнього" [8]. Дослідницький прогноз - це картина стану об'єкту прогнозу в певний момент майбутнього, одержана в результаті розгляду процесу розвитку як руху за інерцією від теперішнього часу до горизонту прогнозу. Прогнозування тенденцій розвитку об'єкту прогнозу, які повинні забезпечувати досягнення у встановлений момент майбутнього певної соціально-політичної, економічної і оборонної мети, називається

нормативним. В цьому випадку орієнтація прогнозу в часі відбувається по схемі "від майбутнього - до теперішнього часу" [9].

Не дивлячись на спільність задач, їх постановка при прогнозуванні і плануванні різна. При плануванні діє наступна схема: "мета - директивна, шляхи і засоби її досягнення - детерміновані, ресурси - обмежені". При прогнозуванні схема інша: "цілі - теоретично досяжні, шляхи і засоби їх досягнення - можливі, ресурси - вірогідні". Задачі прогнозування відрізняються широтою обхвату. Задачі прогнозування треба оцінювати як глобальні. До них можна віднести: аналіз ситуації, визначення рівнів достовірності інформації, визначення ступеня вірогідності, вироблення поточних, середньо і довгострокових прогнозів.

Прогнозування повинне носити системний характер. Необхідність системного підходу в прогнозуванні витікає з особливостей розвитку науки і техніки, народного господарства в період науково-технічної революції. Найважливішими вимогами системного підходу є комплексність прогнозів і планів і безперервний характер процесу планування.

Комплексний підхід передбачає складання прогнозів і планів у взаємозв'язку як в просторі (в галузевому і територіальному розрізі), так і в часі. Взаємозв'язок в просторі означає встановлення раціональних відносин між галузями народного господарства, економічними районами, встановлення оптимальних співвідношень між темпами розвитку науки, техніки і промислового виробництва, збалансованість потреб і ресурсів на всіх рівнях ієрархії. [4]

Чутливість прогнозу і планів до змін залежить від рівня ієрархії, термінів попередження і періодичності коректувань. Чим нижче рівень, тим чутливість вище, тим винні бути коротшим періоди коректування.

Різний ступінь невизначеності інформації, що виробляється, про майбутнє впливає на характер вживаних методів, способів і прийомів прогнозування і планування. Якщо при розробці планів перевага віддається детермінованим методам, то при прогнозуванні - стохастичним. При складанні

планів переважно вживання мають регулярні методи, при прогнозуванні – евристичні [9]

## **1.2 Аналіз та обґрунтування вибору методу прогнозування**

### **1.2.1 Методи науково-технічного прогнозування та їх класифікація**

Перш за все приведемо визначення методу прогнозування як способу теоретичної і практичної дії, направленої на розробку прогнозів. Це визначення є достатньо загальним і дозволяє розуміти термін "метод прогнозування" вельми широко: від найпростіших екстраполяційних розрахунків до складних процедур багатокрокових експертних опитів.

Для вивчення методичного апарату прогностики доцільно із самого початку деталізувати це широке поняття. Далі розрізнятимемо прості методи прогнозування і комплексні методи прогнозування. При цьому під простим методом прогнозування розумітимемо метод, нерозкладний на ще більш прості методи прогнозування, і відповідно під комплексним - метод, що складається з взаємозв'язаної сукупності декількох простих.

На першому рівні всі методи діляться на три класи по ознаці "інформаційна підстава методу". Фактографічні методи базуються на фактично наявному інформаційному матеріалі про об'єкт прогнозування і його минулий розвиток.

Експертні методи базуються на інформації, яку поставляють фахівці-експерти в процесі систематизованих процедур виявлення і узагальнення цієї думки.

Комбіновані методи виділені в окремий клас, щоб можна було відносити до нього методи із змішаною інформаційною основою, в яких як первинна інформація використовуються фактографічна і експертна. Наприклад, при проведенні експертного опиту учасникам представляють цифрову інформацію

про об'єкт або фактографічні прогнози, або, навпаки, при екстраполяції тенденції разом з фактичними даними використовують експертні оцінки.

Не слід відносити до комбінованих методів ті методи прогнозування, які до експертної початкової інформації застосовують математичні методи обробки або початкову фактографічну інформацію оцінюють експертним шляхом. В більшості випадків вони достатньо добре укладаються в перший або другий з перерахованих вище класів. [10]

Ці класи розділяються далі на підкласи за принципами обробки інформації.

Статистичні методи об'єднують сукупність методів обробки кількісної інформації про об'єкт прогнозування за принципом виявлення математичних закономірностей розвитку і математичних взаємозв'язків характеристик, що містяться в ній, з метою отримання прогнозних моделей.

Методи аналогій направлені на те, щоб виявляти схожість в закономірностях розвитку різних процесів і на цій підставі проводити прогнози.

Випереджаючі методи прогнозування будуються на певних принципах спеціальної обробки науково-технічної інформації, що реалізують в прогнозі її властивість випереджати розвиток науково-технічного прогресу.

Експертні методи розділяються на два підкласи. Прямі експертні оцінки будуються за принципом отримання і обробки незалежної узагальненої думки колективу експертів (або одного з них) за відсутності дій на думку кожного експерта думки іншого експерта і думки колективу.

Експертні оцінки із зворотним зв'язком в тому або іншому вигляді утілюють принцип зворотного зв'язку шляхом дії на оцінку експертної групи (одного експерта) думкою, одержаною раніше від цієї групи або від одного з її експертів.

Третій рівень класифікації розділяє методи прогнозування на види по класифікаційній ознаці "апарат методів". Кожний вигляд об'єднує в своєму складі методи, що мають як основа однаковий апарат їх реалізації. Так,



статистичні методи по видах діляться на методи екстраполяції і інтерполяції; методи, що використовують апарат регресійного і кореляційного аналізу; методи, що використовують аналіз чинника [9].

Клас методів аналогій підрозділяється на методи математичних і історичних аналогій. Перші як аналог для об'єкту прогнозування використовують об'єкти іншої фізичної природи, іншої області науки, галузі техніки, проте мають математичний опис процесу розвитку, співпадаючий з об'єктом прогнозування. Другі як аналог використовують процеси однакової фізичної природи, що випереджають в часі розвиток об'єкту прогнозування.

Випереджаючі методи прогнозування можна розділити на методи дослідження динаміки науково-технічної інформації; методи дослідження і оцінки рівня техніки. В першому випадку в основному використовується побудова кількісно-якісних динамічних рядів на базі різних видів НТІ і аналізу і прогнозування на їх основі відповідного об'єкту. Другий вид методів використовує спеціальний апарат аналізу кількісної і якісної інформації, що міститься в НТІ, для визначення характеристик рівня, якості існуючої і проектованої техніки [8].

Прямі експертні оцінки по ознаці апарату реалізації діляться на види експертного опиту і експертного аналізу. В першому випадку використовуються спеціальні процедури формування питань, організації отримання на них відповідей, обробки одержаних відповідей і формування остаточного результату. В другому - основним апаратом дослідження є цілеспрямований аналіз об'єкту прогнозування з боку експерта або колективу експертів, які самі ставлять і вирішують питання, що ведуть до поставленої мети.

Експертні оцінки із зворотним зв'язком в своєму апараті мають три види методів: експертне опитування; генерацію ідей; ігрове моделювання. Перший вид характеризується процедурами регламентованого неконтактного опиту експертів переміжними зворотними зв'язками в розглянутому вище значенні. Другий - побудований на процедурах безпосереднього спілкування експертів

в процесі обміну думками по поставленій проблемі. Він характеризується відсутністю питань і відповідей і направлений на взаємне стимулювання творчої діяльності експертів. Третій вид використовує апарат теорії ігор і її прикладних розділів. Як правило, реалізується на поєднанні динамічної взаємодії колективів експертів і обчислювальної машини, що імітує об'єкт прогнозування в можливих майбутніх ситуаціях [8].

Нарешті, останній, четвертий, рівень класифікації підрозділяє види методів третього рівня на окремі методи і групи методів по деяких локальних для кожного виду поєднанням класифікаційних ознак, з яких вказати один загальний для всього рівня в цілому неможливо. На рис.1.1 представлено класифікацію методів науково-технічного прогнозування.

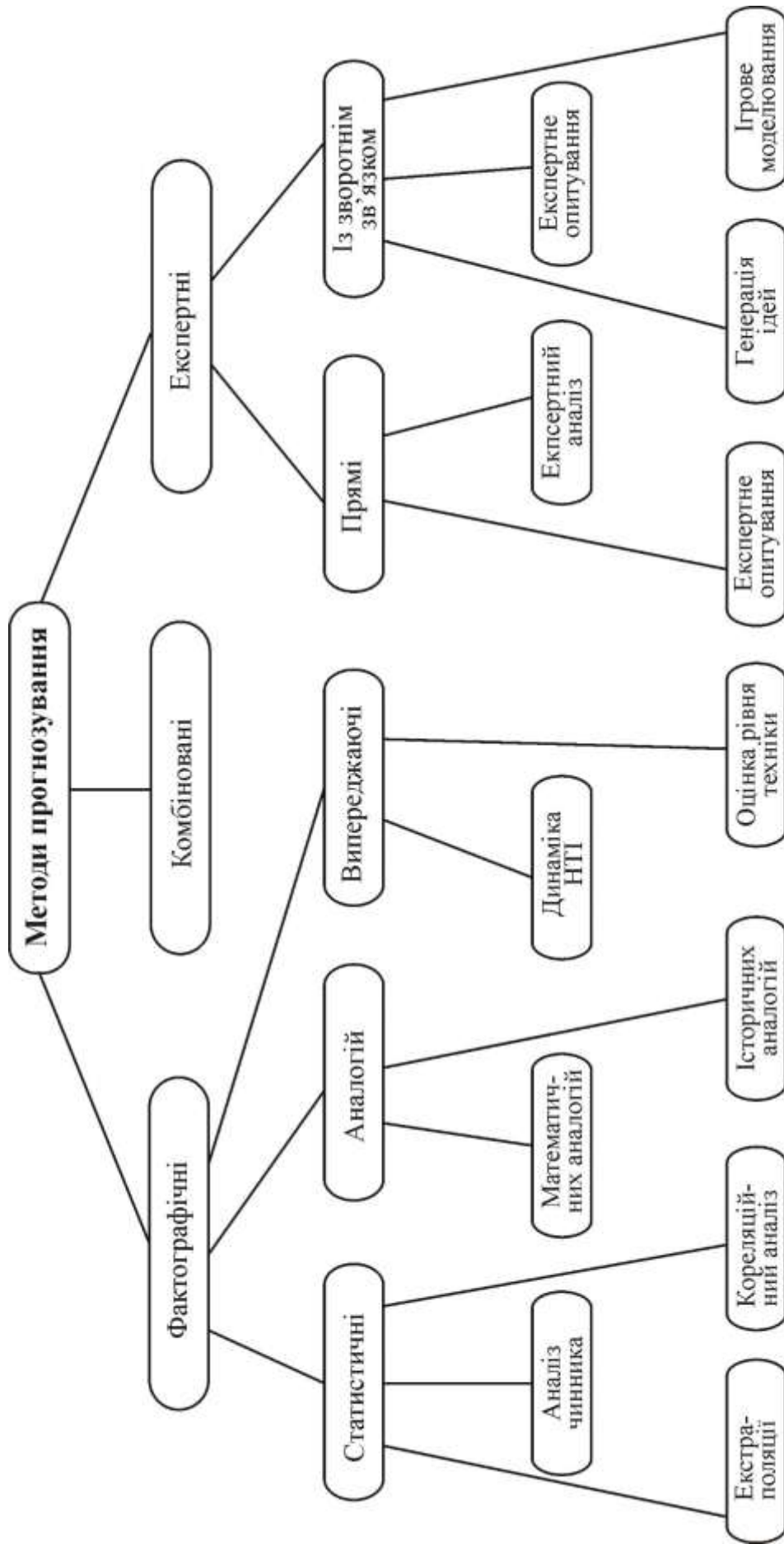


Рисунок 1.1.1. - Методи прогнозування

### 1.2.2 Екстраполяційні методи прогнозування

Методи екстраполяції тенденцій є, мабуть, найпоширенішими і самими розробленими серед всієї сукупності методів прогнозування. Використовування екстраполяції в прогнозуванні має в своїй основі припущення про те, що даний процес зміни змінної є поєднанням двох складових-регулярної і випадкової.

Вважається, що регулярна складова  $f(a, x)$  є гладкою функцією від аргументу (в більшості випадків - часу), описаною кінцевим вектором параметрів  $a$ , які зберігають свої значення на періоді попередження прогнозу. Ця складова називається також трендом, рівнем, детермінованою основою процесу, тенденцією. Під всіма цими термінами лежить інтуїтивне уявлення про якесь очищене від перешкод єство аналізованого процесу. Інтуїтивне, тому що для більшості економічних, технічних, природних процесів не можна однозначно відділити тренд від випадкової складової. Все залежить від того, яку мету переслідує це розділення і з якою точністю його здійснювати [7].

Випадкова складова  $n(x)$  звичайно вважається некорельованим випадковим процесом з нульовим математичним очікуванням. Її оцінки необхідні для подальшого визначення точнісних характеристик прогнозу.

Екстраполяційні методи прогнозування основний вибір роблять на виділення якнайкращого в деякому розумінні опису тренда і на визначення прогнозних значень шляхом його екстраполяції. Методи екстраполяції багато в чому перетинаються з методами прогнозування по регресійних моделях. Іноді їх відмінності зводяться лише до відмінностей в термінології, позначеннях або написанні формул. Проте сама по собі прогнозна екстраполяція має ряд специфічних рис і прийомів, що дозволяють зараховувати її до деякого самостійного виду методів прогнозування [].

Специфічними рисами прогновної екстраполяції можна назвати методи попередньої обробки числового ряду з метою перетворення його до вигляду,

зручного для прогнозування, а також аналіз логіки і фізики прогнозованого процесу, що робить істотний вплив як на вибір виду екстраполюючої функції, так і на визначення меж зміни її параметрів [8].

Попередня обробка початкової інформації в задачах прогнозу екстраполяції

Попередня обробка початкового числового ряду направлена на рішення наступних задач (всіх або частини з них): понизити вплив випадкової складової в початковому числовому ряду, тобто наблизити його до тренда; представити інформацію, що міститься в числовому ряду, у такому вигляді, щоб істотно понизити трудність математичного опису тренда. Основними методами рішення цих задач є процедури згладжування і вирівнювання статистичного ряду [9].

Процедура згладжування направлена на мінімізацію випадкових відхилень точок ряду від деякої гладкої кривої передбачуваного тренда процесу. Найбільш поширений спосіб усереднювання рівня по деякій сукупності навколишніх крапок, причому ця операція переміщається уздовж ряду крапок, у зв'язку з чим звичайно називається ковзаюча середня. В найпростішому варіанті згладжуюча функція лінійна і згладжуюча група складається з попередньої і подальшої крапок, в складніших - функція нелінійна і використовує групу довільного числа крапок.

Згладжування проводиться за допомогою многочленів, що наближають по методу якнайменших квадратів групи досвідчених крапок. Якнайкраще згладжування виходить для середніх точок групи, тому бажано вибирати непарну кількість крапок в згладжуваній групі.

Згладжування навіть в простому лінійному варіанті є у багатьох випадках вельми ефективним засобом виявлення тренда при накладенні на емпіричний числовий ряд випадкових перешкод і помилок вимірювання. Для

рядів із значною амплітудою перешкоди є можливість проводити багатократне згладжування початкового числового ряду. Число послідовних циклів згладжування повинне вибиратися залежно від виду початкового ряду, від ступеня передбачуваною його зашумленості перешкодою, від мети, яку переслідує згладжування. Треба мати при цьому на увазі, що ефективність цієї процедури швидко зменшується (в більшості випадків), так що доцільно повторювати її від одного до трьох разів.

Лінійне згладжування є достатньо грубою процедурою, що виявляє загальний приблизний вид тренда. Для більш точного визначення форми згладженої кривої може застосовуватися операція нелінійного згладжування або зважені ковзаючі середні. В цьому випадку ординатам крапок, що входять до ковзаючої групи, приписується різна вага залежно від їх відстані від середини інтервалу згладжування.

Якщо згладжування направлено на первинну обробку числового ряду для виключення випадкових коливань і виявлення тренда, то вирівнювання служить цілям більш зручного представлення початкового ряду, залишаючи колишнім його значення.

Самими загальними прийомами вирівнювання є логарифмування і заміна змінних.

У випадку якщо емпірична формула передбачається тією, що містить три параметри або відомо, що функція трьох параметрична, іноді вдається шляхом деяких перетворень виключити один з параметрів, а що залишилися два привести до однієї з формул вирівнювання.

Можна розглядати вирівнювання не тільки як метод представлення початкових даних, але і як метод безпосереднього наближеного визначення параметрів функції, що апроксимує початковий числовий ряд. Часто саме так і використовується цей метод в деяких екстраполяційних прогнозах. Відзначимо, що можливість безпосереднього його використання для

визначення параметрів апроксимуючої функції визначається головним чином видом початкового числового ряду і ступенем наших знань, нашої упевненості щодо виду функції, що описує досліджуваний процес.

В тому випадку, якщо вид функції нам невідомий, вирівнювання слід розглядати як попередню процедуру, в процесі якої шляхом вживання різних формул і прийомів з'ясовується самий відповідний вид функції, що описує емпіричний ряд.

Одним з різновидів методу вирівнювання є дослідження емпіричного ряду з метою з'ясування деяких властивостей функції, що описує його. При цьому не обов'язково перетворення приводять до лінійних форм. Проте результати їх готують і полегшують процес вибору апроксимуючої функції в задачах прогностичної екстраполяції. В найпростішому випадку пропонується використовувати наступні три типи диференціальних функцій зростання [8]:

- 1) Перша похідна, або абсолютна диференціальна функція зростання.
- 2) Відносний диференціальний коефіцієнт, або логарифмічна похідна
- 3) Еластичність функції

### **1.2.3 Статистичні методи**

Перш ніж приступити до аналізу статистичних методів прогнозування, розглянемо деякі загальні поняття і визначення, що відносяться до кореляційних і регресійних моделей. Дві випадкові величини є кореляційно зв'язаними, якщо математичне очікування однієї з них міняється залежно від зміни іншої [10].

Вживання кореляційного аналізу припускає виконання наступних передумов:

а) Випадкові величини  $y$  ( $y_1, y_2, \dots, y_n$ ) і  $x$  ( $x_1, x_2, \dots, x_n$ ) можуть розглядатися як вибірка з двовимірної генеральної сукупності з нормальним законом розподілу.

б) Очікувана величина погрішності і рівна нулю

в) Окремі стахостично незалежні спостереження, тобто значення даного нагляду не повинне залежати від значення попереднього і подальшого наглядів.

г) Коваріація між помилкою, пов'язаною з одним значенням залежної змінної  $y$ , і помилкою, пов'язаною з будь-яким іншим значенням  $y$ , рівна нулю.

д) Дисперсія помилки, пов'язана з одним значенням  $y$ , рівна дисперсії помилки, пов'язаній з будь-яким іншим значенням.

е) Коваріація між погрішністю і кожній з незалежних змінних рівна нулю.

ж) Безпосередня застосовність цього методу обмежується випадками, коли рівняння кривої є лінійним щодо своїх параметрів  $b_0, b_1, \dots, b_k$ . Це, проте, не означає, що саме рівняння кривої щодо змінних повинне бути лінійним. Якщо емпіричні рівняння наглядів не є лінійними, то у багатьох випадках виявляється можливим привести їх до лінійної форми і вже після цього застосовувати метод якнайменших квадратів.

з) Спостереження незалежних змінних проводяться без погрішності.

Перед початком кореляційного аналізу необхідно перевірити виконання цих передумов.

Зв'язок між випадковою і не випадковою величинами називається регресійним, а метод аналізу таких зв'язків - регресійним аналізом. Вживання регресійного аналізу припускає обов'язкове виконання передумов (б-г) кореляційного аналізу. Тільки при виконанні приведених передумов оцінки



коефіцієнтів кореляції і регресії, одержувані за допомогою способу якнайменших квадратів, будуть незміщеними і мати мінімальну дисперсію.

Регресійний аналіз тісно пов'язаний з кореляційним. При виконанні передумов кореляційного аналізу виконуються передумови регресійного аналізу. В той же час регресійний аналіз пред'являє менш жорсткі вимоги до початкової інформації." Так, наприклад, проведення регресійного аналізу можливе навіть у разі відмінності розподілу випадкової величини від нормальної, як це часто буває для техніко-економічних величин. Як залежна змінна в регресійному аналізі використовується випадкова змінна, а як незалежна - не випадкова змінна.

По ступеню комплексності статистичні дослідження можна розділити на двовимірні і багатовимірні. Перші торкаються розгляду парних взаємозв'язків між змінними (парні кореляції і регресії) і направлені в прогнозних дослідженнях на рішення таких задач, як встановлення кількісної міри тісноти зв'язку між двома випадковими величинами, встановлення близькості цього зв'язку до лінійної, оцінки достовірності і точність прогнозів, одержаних екстраполяцією регресійної залежності. Багатовимірні методи статистичного - аналізу направлені в основному на рішення задачі системного аналізу багатовимірних стохастичних об'єктів прогнозування. Метою такого аналізу є, як правило, з'ясування внутрішніх взаємозв'язків між змінними комплексу, побудова багатовимірних функцій зв'язку змінних, виділення мінімального числа характеристик, що описують об'єкт з достатнім ступенем точності. Однією з основних задач тут є скорочення розмірності опису об'єкту прогнозування.

Таким чином, статистичні методи використовуються в основному для підготовки даних, приведення їх до вигляду, придатного для виробництва прогнозу. Як правило, після їх вживання використовується один з методів екстраполяції або інтерполяції для отримання безпосередньо прогнозного результату [9].

### 1.2.4 Експертні методи

Область використання експертних методів

Методи експертних оцінок в прогнозуванні і перспективному плануванні науково-технічного прогресу застосовуються в наступних випадках:

а) в умовах відсутності достатньо представницької і достовірної статистики характеристики об'єкту (наприклад, лазери, що голографічні запам'ятовують пристрої, раціональне використання водних ресурсів на підприємствах);

б) в умовах великої невизначеності середовища функціонування об'єкту (наприклад, прогнозів людино-машинної системи в космосі або облік взаємовпливу областей науки і техніки);

в) при середньо і довгостроковому прогнозуванні об'єктів нових галузей промисловості, схильних сильному впливу нових відкриттів у фундаментальних науках (наприклад, мікробіологічна промисловість, квантова електроніка, атомне машинобудування);

г) в умовах дефіциту часу або екстремальних ситуаціях.

Експертна оцінка необхідна, коли немає належної теоретичної основи розвитку об'єкту. Ступінь достовірності експертизи встановлюється по абсолютній частоті, з якою оцінка експерта зрештою підтверджується наступними подіями. Існує дві категорії експертів - це вузькі фахівці і фахівці широкого профілю, що забезпечують формулювання крупних проблем і побудову моделей. Вибір експертів для прогнозу проводиться на основі їх репутації серед певної категорії фахівців. Проте не слід забувати і тієї обставини, що першокласний фахівець не завжди може достатньо кваліфіковано розглянути і зрозуміти загальні, глобальні, питання. Для цієї

мети потрібно привертати експертів хоча і недостатньо вузько інформованих, але володіючих здібністю до прогнозування і уяви.

"Експерт" в дослівному перекладі з латинської мови означає "досвідчений". Тому і у формалізованому, і в неформалізованому способах визначення експерта значне місце займають професійний досвід і розвинена на його основі інтуїція. Умови необхідності і достатності віднесення фахівця до категорії експертів вводяться таким чином.

Важливо встановити не абсолютний ступінь надійності експертної оцінки, а ступінь надійності в порівнянні з оцінкою середнього фахівця, а також кореляцію між вірогідністю його прогнозної оцінки і надійністю класу тих гіпотез, якими оперує експерт. Загалом, потрібно визначити, що таке експерт. Перерахуємо деякі вимоги, яким винен задовольняти експерт:

- 1) оцінки експерта повинні бути стабільні в часі і транзитивні;
- 2) наявність додаткової інформації про прогнозовані ознаки лише покращує оцінку експерта;
- 3) експерт повинен бути визнаним фахівцем в даній області знань;
- 4) експерт повинен володіти деяким досвідом успішних прогнозів в даній області знань.

Характеризуючи експертів, слід мати у вигляді, що в результаті вироблення оцінок можуть мати місце помилки двох видів. Помилки першого виду відомі в техніці вимірювань як систематичні, помилки другого виду - як випадкові. Експерт, схильний до помилок першого вигляду, видає значення, які стійко відрізняються від істинного у бік збільшення або зменшення. Вважають, що помилки цього вигляду пов'язані з складом розуму експертів. Для корекції систематичних помилок можна застосовувати поправочні коефіцієнти або ж використовувати спеціально розроблені тренувальні ігри. Помилки другого виду характеризуються величиною дисперсії. Виходячи з

аналізу основних видів помилок при винесенні експертних думок, можна додати до розглянутого раніше переліку вимог до експертів ще одне. Значення його полягає в тому, що слід віддати перевагу експерту, оцінки якого мають малу дисперсію і систематичне відхилення середньої помилки від нуля, експерту з середньою помилкою, рівною нулю, але з більшою дисперсією. На жаль, апріорі визначити здатність людини робити правильні експертні оцінки неможливо. Важливим засобом підготовки експертів є спеціальні тренувальні ігри [10].

Організація форм роботи експерта може бути програмованою або непрограмованою, а діяльність експерта може здійснюватися в усній (інтерв'ю) або у письмовій формі (відповідь на питання спеціальних таблиць експертних оцінок або вільний виклад по заданій темі).

Організація стимуляції роботи експерта полягає в розробці:

- евристичних прийомів і способів, що полегшують пошук прогнозної експертної оцінки; правових норм, що гарантують експерту оформлення пріоритету і авторства, а також нерозголошення всіх науково-технічних ідей, що висуваються їм в процесі експертизи;
- форм моральної, професійної і матеріальної зацікавленості експерта в експертних оцінках; організаційних форм роботи експерта (включення в план роботи і т.п.).

Виходячи з одержаної в результаті аналізу моделі об'єкту прогнозування, визначаються наукові і технічні напрями, по яких необхідно повернути експерта, виділяються групи експертів по приналежності питання до області фундаментальних, прикладних наук або до стикових наукових напрямів. [9]

При рішенні задачі формування експертної групи необхідно виявити і стабілізувати працездатну мережу експертів. Спосіб стабілізації експертної мережі полягає в наступному. На основі аналізу літератури по прогнозованій

проблемі вибирається будь-який фахівець, що має декілька публікацій в даній області. До нього поводяться з проханням назвати 10 самих компетентних, на його думку, фахівців по даній проблемі. Потім звертаються одночасно до кожного з десяти названих фахівців з проханням вказати 10 найкрупніших їх колег-учених. З одержаного списку фахівців викреслюються 10 первинних, а іншим розсилаються листи, що містять вказане вище прохання. Дану процедуру повторюють до тих пір, поки жоден із знов названих фахівців не додасть нових прізвищ до списку експертів, тобто поки не стабілізується мережа експертів. Одержану мережу експертів можна вважати генеральною сукупністю фахівців, компетентних в області прогнозованої проблеми. Проте через ряд практичних обмежень виявляється недоцільним привертати всіх фахівців до експертизи. Тому необхідно сформувати репрезентативну вибірку з генеральної сукупності експертів.

Визначення специфіки процедур для методів класу ПЕО (персональних експертних оцінок) здійснюється на основі аналізу вимог до експертів і їх оцінок, витікаючих з суті методів:

а) аналітичні записки пред'являють вимоги структуризації експериментальної проблеми, експлікації і ранжирування цілей, аналізу альтернативних шляхів досягнення мети, оцінки витрат на кожен альтернативу і рекомендацій за найефективнішими способами рішення проблем;

б) парні порівняння, нормування і ранжування вимагають однорідності оцінюваних ознак, наявності логічно обґрунтованих критеріїв і еталонів,

наявність однозначно певних процедур операції з критеріями, еталонами і ознаками;

в) інтерв'ю пред'являють специфічні вимоги як до експерта, так і до інтерв'юера;

г) морфологічна структуризація вимагає чіткого визначення функціональних характеристик об'єкту або проблеми, які необхідно поліпшити, класифікації наукових принципів, на основі яких можливе поліпшення характеристики; аналізу всіляких комбінацій цих принципів і відсіву явно абсурдних; оцінки комбінацій по ступеню здійсненності і витрат на їх реалізацію; порівняння комбінацій по комплексному критерію "витрати - ефективність - час".

### **1.2.5 Обґрунтування вибору методу**

Прогнозування або передбачення результатів футбольних матчів є необхідним елементом організації букмекерської діяльності, планування підготовчо-тренувального процесу спортивних команд і т.д. З формальної точки зору ця задача відноситься до широкого класу задач прогнозування дискретних послідовностей (сукупності значень в фіксовані моменти), які виникають не тільки в спорті, але і в медицині, фізиці, техніці, економіці, соціології, та інших галузях.

Нетривіальність прогнозування дискретних послідовностей спричинена тим, що на відміну від добре алгоритмізованих процедур інтерполяції [13], прогнозування потребує екстраполяції даних про минуле на майбутнє. При цьому необхідно враховувати невідому закономірність про появу, закладеному в основі процесу, який генерує дискретні послідовності. Розробці математичних моделей прогнозування присвячена велика кількість досліджень [14]. Найбільш розповсюдженими є методи, побудовані на базі ймовірно-статистичного апарату. Але їх використання потребує значної кількості експериментальних даних, які не завжди є можливість зібрати в умовах подій, які відбулися відносно недавно, наприклад, аварія на Чорнобилі.

Останнім часом у задачах прогнозування відроджується інтерес до використання штучних нейронних мереж [15]. Вони розцінюються як близькі

до людського мозку універсальні моделі, які навчаються розпізнанню невідомих закономірностей. Але, як і у випадку ймовірно-статистичних методів, для навчання нейронних мереж потрібно велика вибірка експериментальних даних. Крім того навчання нейронних мереж не допускає чіткої інтерпретації ваг дуг.

У цьому розділі пропонується перехід до прогнозуванню, який об'єднує експериментальні дані про кількість захворювань з експертно-лінгвістичної інформації про закономірностей, які вдається побачити в уже існуючих даних. Використання експертно-лінгвістичних закономірностей, які формалізуються за допомогою нечіткої логіки, дозволяє побудувати модель прогнозування в умовах малих експериментальних вибірок. Підхід, що пропонують тут ідеологічно досить близький до того що має назву «нейро - нечіткому» підходу [15], який об'єднує здатності нейронних мереж до навчання і легку здатність до інтерпретації нечітких правил ЯКЩО-ТО. Проте, на відміну від [14], ми не використовуємо нейронну мережу для навчання моделей прогнозу, а безпосередньо налаштовуємо нечіткі правила ЯКЩО-ТО за допомогою існуючих експериментальних даних [15].

### **1.3 Огляд відомих технічних рішень задачі прогнозування результатів футбольних матчів**

Під час пошуку аналогів даної системи, виявилось, що подібних систем мало. В основному, велика їх кількість пов'язана із роботою брокерських систем, де можна оцінити ймовірність ставки та перемога улюбленої команди. Також було знайдено сайти, де можна на основі попередніх статистичних даних отримати з певною ймовірністю прогноз результату футбольного матчу. Для більш точної та достовірної інформації такі системи пропонуються послуги абонентської підписки за певну оплату.

Розглянемо детальніше подібні аналоги.

### 1.3.1 Footbe.net

Дана система розроблена для прогнозування результатів футбольних матчів на основі статистичного аналізу і машинного навчання. Розробники системи стверджують що використовують стратегію Value betting, яка дозволяє своїм користувачам робити ідеальні прогнози та ставки (рис. 1.5).

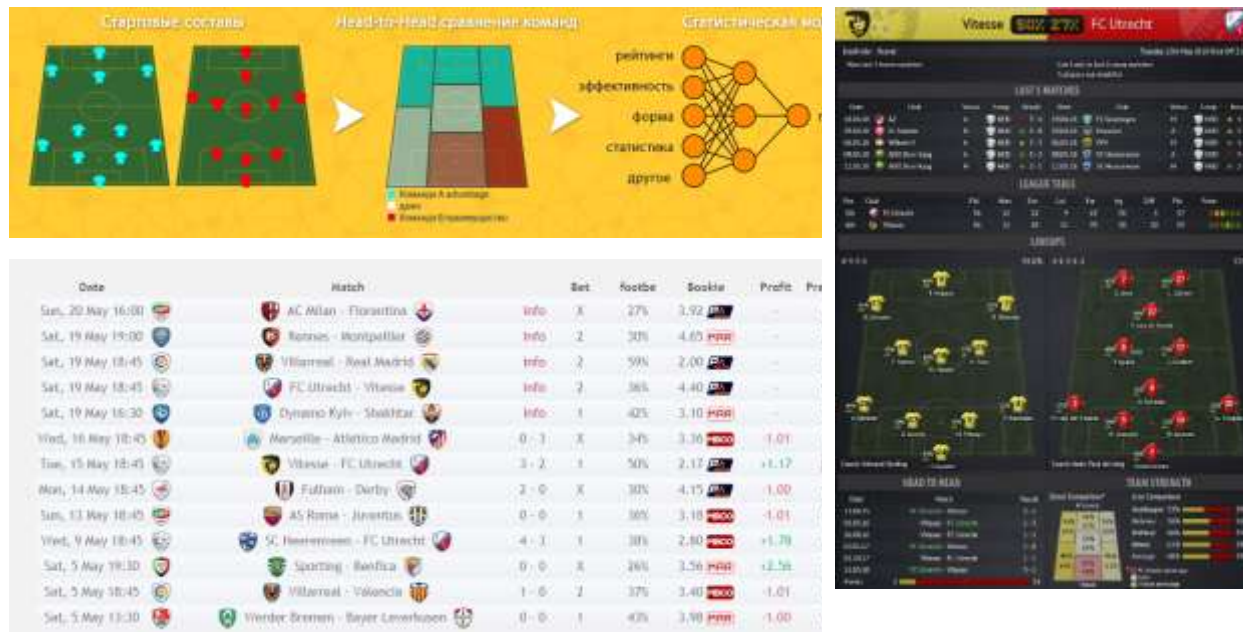


Рисунок 1.5 – Система Footbe.net

Прогнози ґрунтуються на реальних стартових складах, які оновлюються автоматично перед початком матчу. Як тільки склади команд стають відомі, пропозиція змінюється відповідно до нового прогнозу.

Екран матчу має 5 закладок, кожна з яких надає вам цінну інформацію щоб допомогти зробити ідеальну ставку. Доступна інформація про травми, дискваліфікацію, останні результати, орієнтовних складах, сильних і слабких сторонах команд, кращих гравців та т.п.

Головними недоліками даної системи є обмеження у використанні, система нова і ще проходить бета тестування та висока вартість преміум підписки – 25 у.о. на місяць або 75 у.о. за 6 міс (це із 70% знижкою при умові бета-тестування).



### 1.3.2 Betgarant.net

Betgarant.net – розробник декількох продуктів. BETGARANT - це повноцінний, автоматичний аналізатор і статист. Він позиціонує себе як персональний помічник у виборі, аналізі, і прогнозі футбольного матчу. Програма заснована на математичному аналізі і розрахунках статистики, більше 130 чемпіонатів з різних країн світу.

Ось перелік основних продуктів від BETGARANT (рис. 1.6).

BETGARANT PLUS 2017 - це автоматичний фахівець з аналізу матчів і вибірці статистики. Математичні алгоритми аналізують і оновлюють статистику понад 130 країн, професійно грають у футбол!

BETGARANT LIVE – 2017 - програма визначає, в режимі он-лайн матчі, в яких повинні бути голи. І не важливо яка це хвилина матчу. 5-я або 90-а. Betgarant Live допомагає зробити правильний вибір і в режимі он-лайн підкаже на який матч варто звернути увагу.

BETGARANT "Card & Corner" - багатофункціональна програма, що сигналізує кількість жовтих карток і кутових в майбутніх матчах, провідних чемпіонатів, індивідуальний тотал кутових для кожної команди, загальний тотал кутових, фора по угловим, індивідуальний тотал карток для кожної команди; загальний тотал карток; фора за картками.



Рисунок 1.6 – Продукти BETGARANT

Головними недоліками даної системи є те, що усі програми, які працюють на основі математичних алгоритмах працюють на основі MS Excel.

### 1.3.3 Robobet

Сайт, на якому можна безкоштовно зробити прогноз результату футбольних матчів. ROBOBET - це алгоритм прогнозування з штучним інтелектом, де робот аналізує безліч факторів, переробляючи різну інформацію і складаючи її в єдине ціле (нейроні мережі). Робот показує ймовірність того чи іншого результату у відсотках.

Приклад: Імовірність перемоги команди господарів 55%, ймовірність нічий 25% і перемоги гостей - 20%, робот виставить П1 на господарів.

ROBOBET не аналізує коефіцієнти букмекерських контор, а ґрунтується лише на основі статистичного аналізу даних з відкритих джерел. Алгоритм буде постійно вдосконалюватися для поліпшення якості прогнозів, тому розробники пропонують додати сервіс в закладки (рис. 1.7).

06-06   05-06   04-06   03-06   02-06   01-06											
Время	Футбольные матчи			Прогнозы				Коеф			Итог
	Хозяева - Гости			1	X	2	Ставка	1	X	2	
02:00	Парана - Флуминенсе			23%	37%	39%	X2	2.56	3.08	2.71	2:1
02:40	3 Февраля - Спортиво Лукеньо			21%	36%	43%	X2	2.85	3.30	2.24	0:2
03:15	Ботафого ПБ - Глобо ФК			73%	17%	11%	П1	1.52	3.60	6.10	1:1
03:15	Дельфин - Универсидад Католика			49%	35%	17%	1X	2.20	3.10	3.10	5:2
04:00	Реал Монархс - Рио Гранде Вэлли			56%	24%	19%	П1	1.31	4.90	7.60	2:0
12:00	Спарта Прага (19) - Богемиянс 1905 (19)			79%	7%	14%	П1	0.00	0.00	0.00	2:0
12:30	Логан Лайтнинг - Брисбен Сити			49%	12%	39%	12	3.19	3.92	1.77	1:4
16:30	Рэйон Спорт - Амагаджу			75%	16%	8%	П1	1.41	4.20	6.70	1:2
17:00	ПСМС Медан - Персиб Бандунг			54%	28%	17%	1X	1.93	3.30	3.60	0:3
17:00	ПКНП - Теренггану			23%	51%	27%	X2	2.40	3.40	2.55	2:1
17:00	Казахстан - Азербайджан			45%	39%	17%	1X	2.44	2.86	3.48	3:0
17:00	Кедах - Келантан			78%	13%	9%	П1	1.22	2.68	10.50	1:2
17:00	Юж. Корея (19) - Катар (19)			49%	17%	34%	12	1.95	3.65	3.50	2:1

Рисунок 1.7 – Результат роботи сервісу Robobet

## **1.4 Постановка задачі та формулювання вимог**

При розробці програми, необхідно звернути увагу на функції, які повинна вона виконувати.

Система для прогнозування результатів футбольних матчів повинна отримати вхідні дані та провести прогнозування результату футбольного матчу. Система повинна вміти відбирати фактори, які впливають на результат прогнозу. Потрібно створити та налаштувати нечітку модель прогнозування (створити та розробити лінгвістичні змінні, нечітку базу знань, розробити алгоритм прогнозування, провести налаштування нечіткої моделі). Після реалізації системи її варто протестувати на реальних даних чемпіонату молодшої ліги України.

Система є інструментом, який забезпечить зручне та ефективне прогнозування результатів футбольних матчів молодшої ліги України.

## **1.5 Висновок**

В першому розділі магістерської кваліфікаційної роботи проведено обґрунтування доцільності створення нової технології прогнозування результатів футбольних матчів де розглянуто проблеми застосування ІТ-технологій в спорті. У даному розділі було доведено актуальність і доцільність розробки системи, а також проаналізовано предметну область. Результати аналізу зведені в таблицю 1.1.

Детально розглянуто також задачу науково-технічного прогнозування, проведено порівняльний аналіз методів прогнозування, приведено їх класифікацію. Можна зробити висновок, що найбільш перспективними методами для задач медичного прогнозування є екстраполяційні, ймовірно статистичні або експертні методи прогнозування а також їх поєднання. Проведено обґрунтування вибору методу.

## **2 МОДЕЛЮВАННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ПРОГНОЗУВАННЯ РЕЗУЛЬТАТІВ ФУТБОЛЬНИХ МАТЧІВ**

### **2.1 Прогнозування результатів футбольних матчів як об'єкт моделювання**

Прогнозування футбольних матчів викликає цікавість з двох причин: продемонструвавши можливості математичних методів та бажання отримати попередній результат випадкової події.

Важливе місце моделювання у футболі відводиться математичним методам, що дають змогу виявляти особливості, закономірності, тенденції, перевіряти надійність суджень і припущень. Проте практичні приклади моделювання в дитячо-юнацькому футболі на сьогодні розроблені не

достатньо. Тому дослідження в даній предметній області є, як ніколи актуальним на сьогодні.

Щодо проблеми прогнозування й моделювання у футболі, то у цій галузі розроблено питання моделювання тактичних дій у процесі підготовки юнацьких команд із футболу. Зроблено спробу прогнозування результатів футбольних матчів на основі нечітких правил та нечіткого багатofакторного аналізу. Досліджено моделювання спеціалізованих стандартних вправ у навчально-тренувальному процесі юних футболістів. Вивчали також прогнозування фізичної підготовленості юних футболістів. Розроблено модельні варіанти структури міжігрових мікроциклів для змагального періоду при різних варіантах побудови річного макроциклу. Водночас аналіз літератури засвідчив, що практичні приклади прогнозування й моделювання в дитячо-юнацькому футболі на сьогодні розроблені недостатньо.

Будь-яка модель – це наближена подібність реальної дійсності. Але в цьому й суть великих методичних і пізнавальних можливостей моделювання. Якщо ж числові характеристики використовуються без урахування специфіки описуваних явищ, вони не лише себе не виправдовують, але й можуть призвести дослідника до серйозних помилок. Не можна забувати, що в стислості та точності математичних характеристик, у зручності відобразити складних та непередбачуваних процесів невеликою кількістю показників є не лише великі пізнавальні й методичні можливості, але й небезпека втрати зв'язку з конкретними речами, що може призвести до хибних висновків, створити видимість істини там, де її немає [7]. Якісні методи моделювання без точного урахування кількісної сторони малоефективні, оскільки не дають можливості глибокого проникнення в суть явищ, які вивчаються. Зрілість науки визначається мірою використання математики [2]. Наука лише тоді досягає досконалості, коли їй удається користуватися математикою.

Евристична роль математики полягає, по-перше, у дедуктивному характері математичних теорій, що дає можливість обчислювати й передбачати нові факти, а по-друге – у використанні певних математичних

схем (своєрідних формальних моделей), змістовне тлумачення яких часто веде до нових відкриттів у галузі конкретних наук. Не треба піддавати складній математичній обробці те, що очевидне саме собою. Здебільшого результати спостережень, зведені в прості статистичні таблиці, є настільки переконливими, що відпадає будь-яка інша форма їх інтерпретації. При математичному моделюванні виконання одного з основних етапів – побудови математичних моделей за експериментальними даними – нині просто немислиме без комп'ютера.

В останні роки, завдяки розвитку графічного інтерфейсу й графічних пакетів, широкий розвиток отримало комп'ютерне, структурно-функціональне моделювання. Під комп'ютерною моделлю найчастіше розуміють умовний образ об'єкта чи деякої системи об'єктів (або процесів), описаний за допомогою взаємозалежних комп'ютерних таблиць, блок-схем, діаграм, графіків, малюнків, анімаційних фрагментів, гіпертекстів і т. ін., що відображає структуру й взаємозв'язки між елементами об'єкта. Комп'ютерні моделі такого виду називають структурно-функціональними. Окрему програму, сукупність програм, програмний комплекс, який дає змогу за допомогою послідовності обчислень і графічного відображення їхніх результатів відтворювати (імітувати) процеси функціонування об'єкта, системи об'єктів за умови впливу на об'єкт різних, зазвичай випадкових, факторів. Такі моделі називають імітаційними. Комп'ютерне моделювання – метод виконання завдання аналізу та синтезу складної системи на основі використання її комп'ютерної моделі. Суть комп'ютерного моделювання полягає в одержанні кількісних і якісних результатів за наявною моделлю. Якісні висновки, одержувані за результатами аналізу, дають змогу виявити невідомі раніше властивості складної системи: її структуру, динаміку розвитку, стійкість, цілісність та ін. Кількісні висновки здебільшого мають характер прогнозу деяких майбутніх або пояснення минулих значень змінних, що характеризують систему. Комп'ютерне моделювання для створення нової

інформації використовує будь-яку інформацію, яку можна актуалізувати за допомогою ЕОМ.

Основні функції комп'ютера при моделюванні:

- виконувати роль допоміжного засобу для розв'язання завдань звичайними обчислювальними засобами, алгоритмами, технологіями;
- відігравати роль засобу постановки й покращення нових завдань, що не розв'язуються традиційними засобами, алгоритмами, технологіями;
- виконувати роль засобу конструювання комп'ютерних навчально-моделювальних середовищ;
- відігравати роль засобу моделювання для отримання нових знань; – виконувати роль «навчання» нових моделей (самонавчальні моделі).

Комп'ютерне моделювання стає новим інструментом, методом наукового пізнання, новою технологією також через зростаючу потребу переходу від дослідження лінійних математичних моделей систем. Предметом комп'ютерного моделювання можуть бути спортивна діяльність окремого спортсмена або команди, функціональні системи, технічна та тактична підготовка; будь-який реальний об'єкт або процес, наприклад процес розвитку, і взагалі будь-яка складна система.

Цілі комп'ютерного моделювання можуть бути різними, однак найбільш часто моделювання є, як зазначалося раніше, центральною процедурою системного аналізу, причому під ним розуміють сукупність методологічних засобів, які використовують для підготовки та прийняття рішень тактичного, організаційного або технічного характеру.

Комп'ютерна модель складної системи повинна, по можливості, відображати всі основні чинники й взаємозв'язки, що характеризують реальні ситуації, критерії та обмеження. Модель має бути достатньо універсальною, щоб по можливості описувати близькі за призначенням об'єкти, і водночас достатньо простою, аби дати змогу виконати необхідні дослідження з раціональними наслідками. Усе це вказує на те, що моделювання, яке розглядається загалом, є, передусім, мистецтвом, а не просто сформованою

наукою із самостійним набором засобів відображення явищ і процесів реального світу. Відображаючи фізичну систему (об'єкт) на математичну систему (наприклад математичний апарат рівнянь), отримують фізико-математичну модель системи, або математичну модель фізичної системи.

Розробка системи прогнозування результатів футбольних матчів молодшої ліги на основі нечітких правил - це не просто створення системи. Це система із набором нечітких правил та аналіз великої кількості фактичних даних результатів футбольних матчів молодшої ліги.

Сучасні системи можуть бути доступні в будь-якому місці і в будь-який час на комп'ютері, планшеті або мобільному пристрої. Ці особливості роблять системи дуже привабливими для вирішення широкого спектру завдань

## **2.2 Вибір методології моделювання для задачі прогнозування результатів футбольних матчів**

При всьому різноманітті різних алгоритмів і методів прогнозування результатів футбольних матчів виділяють наступні основні методи:

- нечітка логіка;
- метод опорних векторів (SVM);
- нейроні мережі.

Крім цього, в літературі можна знайти ще метод прогнозування результатів футбольних матчів на основі статистичного аналізу та кваліметрії та метод зваженої суми показників, але дані методи на розглядалися під час виконання даної роботи.

### **2.2.1 Аналіз можливостей нечіткої логіки для покращення інформаційної технології прогнозування**

Основи теорії нечітких множин і нечіткої логіки були закладені наприкінці 1960-х років у працях відомого американського математика Лотфі



Заде. Його праця “Fuzzy Sets”, опублікована у 1965 р. в журналі “Information and Control”, стала поштовхом до розвитку нової математичної теорії. Він дав назву і новій галузі науки – “fuzzy sets” (fuzzy – нечіткий, розмитий, м'який). Основною причиною появи нової теорії стали нечіткі і наближені міркування, що використовувались для опису людиною процесів, систем, об'єктів. Математична теорія нечітких множин (fuzzy sets) і нечітка логіка (fuzzy logic) є узагальненнями класичної теорії множин і класичної формальної логіки.

Перш ніж нечіткий підхід до моделювання складних систем отримав визнання в усьому світі, минуло не одне десятиліття. Що ж запропонував Л. Заде? По-перше, він розширив класичне канторовське поняття множини, припустивши, що характеристична функція (функція належності елемента множині) може набувати будь-яких значень в інтервалі  $[0, 1]$ , а не тільки значень 0 або 1. Такі множини він назвав нечіткими [21]. Л. Заде визначив також низку операцій із нечіткими множинами і запропонував узагальнення методів логічного висновку.

Ввівши згодом поняття лінгвістичної змінної і припустивши, що її значеннями (термами) є нечіткі множини, Л. Заде створив апарат для опису процесів інтелектуальної діяльності, включаючи нечіткість і невизначеність виразів (наприклад, високий, середній, незначний ризику).

Завданням нечітких множин є визначення належності деякого об'єкта чи елемента до заданої множини. Нехай  $E$  – деяка множина, а  $A$  – підмножина  $E$ , тобто  $A \subseteq E$ . Той факт, що елемент  $x$  множини  $E$  належить і множині  $A$  у теорії множин позначають так:  $x \in A$ . Щоб виразити цю належність, можна скористатися й іншим поняттям – характеристичною функцією  $\mu_A(x)$ , значення якої вказують, чи є (так або ні)  $x$  елементом  $A$ :

$$\mu_A(x) = \begin{cases} 1, & \text{якщо } x \in A, \\ 0, & \text{якщо } x \notin A. \end{cases}$$

Згідно з теорією нечітких множин, характеристична функція належності може набувати будь-якого значення в інтервалі  $[0, 1]$ , а не тільки два – 0 і 1. Відповідно до цього, елемент  $x$  множини  $E$  може не належати  $A$  ( $\mu_A(x) = 0$ ),

бути елементом  $A$  невеликою мірою (значення  $\mu_A(x)$  близьке до нуля), бути елементом  $A$  значною мірою ( $\mu_A(x)$  близьке до 1) або бути елементом  $A$  ( $\mu_A(x) = 1$ ). Отже, поняття належності узагальнюється. Нечітку під множину  $A$  універсальної множини  $E$  позначають  $A_n$  і визначають упорядкованими парами [22]:

$$A_n = \{(x, \mu_A(x)) \mid x \in E\}.$$

Характеристична функція належності (або просто функція належності)  $\mu_A(x)$  набуває значень у деякій упорядкованій множині  $M$  (наприклад,  $M = [0, 1]$ ). Ця функція належності вказує ступінь (або рівень) належності елемента  $x$  до підмножини  $A$ . Множину  $M$  називають множиною надженностей. Якщо  $M = \{0, 1\}$ , то нечітку підмножину  $A$  можна розглядати як звичайну або чітку множину.

Для нечітких множин, як і для звичайних, визначено основні логічні операції.

Рівність. Дві нечіткі множини  $A$  і  $B$  називають рівними, якщо для всіх  $x \in E$  має місце рівність їх характеристичних функцій:  $\mu_A(x) = \mu_B(x)$ . Позначення:  $A = B$ .

Домінування. Вважають, що нечітка множина  $A$  належить нечіткій множині  $B$ , якщо для всіх  $x \in E$  справедливе співвідношення:  $\mu_A(x) \leq \mu_B(x)$ ; позначають:  $A \dot{\leq} B$ . Іноді використовують термін “домінування”, тобто коли  $A \dot{\leq} B$ , кажуть, що  $B$  домінує над  $A$ .

Доповнення. Нехай  $M = [0, 1]$ ,  $A$  і  $B$  – нечіткі множини, задані на  $E$ .  $A$  і  $B$  доповнюють один одного, якщо  $\forall x \in E \mu_A(x) + \mu_B(x) = 1$ . Позначення:  $A = \bar{B}$ .

Перетин двох нечітких множин (нечітке “І”), що позначають  $A \cap B$  – найбільша нечітка підмножина, яка знаходиться одночасно в  $A$  і  $B$ . Визначають так:

$$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)).$$

Об'єднання двох нечітких множин (нечітке “АБО”), що позначають  $A \cup B$  – найменша нечітка підмножина, яка включає як  $A$ , так і  $B$ , з функцією належності

$$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x)).$$

Різниця двох нечітких множин  $A - B = A \cap \bar{B}$  з функцією належності

$$\mu_{A-B}(x) = \mu_{A \cap \bar{B}}(x) = \min(\mu_A(x), 1 - \mu_B(x)).$$

Нехай  $M = [0,1]$  і  $A$  – нечітка множина з елементами  $x$  з універсальної множини  $E$  і множиною значень функцій належності  $M$ . Величину  $\sup_{x \in E} \mu_A(x)$  називають висотою нечіткої множини  $A$ . Нечітка множина  $A$  є нормальною, якщо її висота дорівнює 1, тобто верхня межа її функції належності дорівнює 1 ( $\sup_{x \in E} \mu_A(x) = 1$ ). За  $\sup_{x \in E} \mu_A(x) < 1$  нечітку множину називають субнормальною.

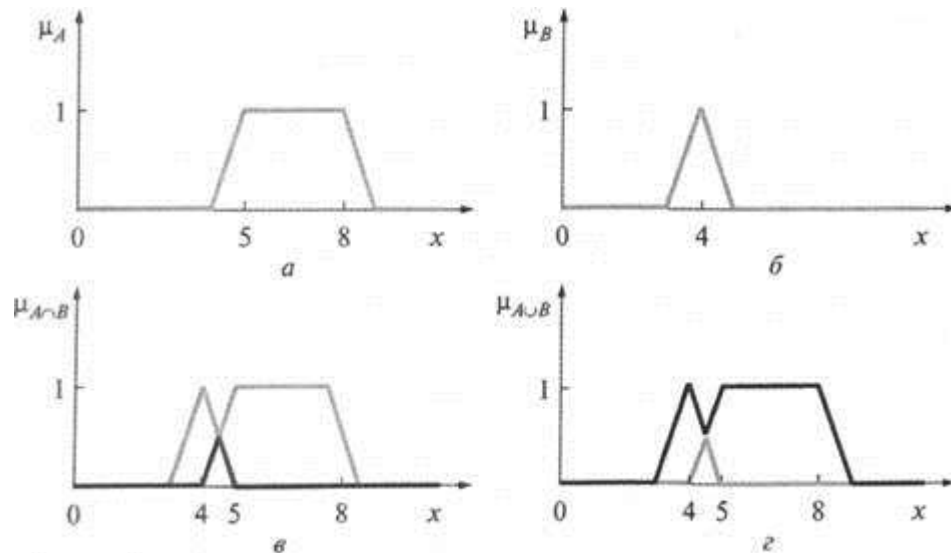
Нечітка множина є порожньою, якщо  $\forall x \in E \mu_A(x) = 0$ . Непорожню субнормальну множину можна нормалізувати за формулою

$$\mu_A(x) := \frac{\mu_A(x)}{\sup_{x \in E} \mu_A(x)}.$$

Наочне подання операцій над нечіткими множинами. Розглянемо прямокутну систему координат, на осі ординат якої відкладено значення  $\mu_A(x)$ , на осі абсцис – у довільному порядку розміщені елементи  $E$ . Якщо множина  $E$  за своєю природою впорядкована, то цей порядок бажано зберегти в розміщенні елементів на осі абсцис. Таке подання унаочнює прості операції над нечіткими множинами.

Нехай  $A$  – нечіткий інтервал між 5 і 8, а  $B$  – нечітке число, близьке до 4 (рис. 4.4, а, б) [23].

Нечітку множину між 5 і 8 І (AND) близько 4 (темна лінія) ілюструє рис. 2.1, в, нечітку множину між 5 і 8 АБО (OR) близько 4 – рис. 2.1, г (темна лінія).



Рисунку 2.1 - Приклади нечітких множин (а, б), їх перетину (в) та об'єднання (г)

Для опису нечітких множин вводять поняття нечіткої і лінгвістичної змінних. Нечітку змінну описує набір  $\langle \beta, X, A \rangle$ , де  $\beta$  – назва змінної,  $X$  – універсальна множина (область визначення  $\beta$ ),  $A$  – нечітка множина на  $X$ , що описує обмеження на значення нечіткої змінної  $\beta$ .

Значеннями лінгвістичної змінної можуть бути нечіткі змінні, тобто лінгвістична змінна знаходиться на вищому рівні, ніж нечітка змінна. Кожна лінгвістична змінна складається з: назви; множини своїх значень, що також називається базовою терм-множиною  $T$ . Елементи базової терм-множини є назвами нечітких змінних; універсальної множини  $X$  синтаксичного правила  $G$ , за яким генеруються нові терми із застосуванням слів природної або формальної мови; семантичного правила  $P$ , яке кожному значенню лінгвістичної змінної ставить у відповідність нечітка підмножина множини  $X$ .

Лінгвістичну змінну описує набір  $\langle \beta, T, X, G, M \rangle$ , де

- $\beta$  – найменування лінгвістичної змінної;
- $T$  – множина її значень (терм-множина), що є назвами нечітких змінних, областю визначення кожної з яких є множина  $X$ ; множину  $T$  називають базовою терм-множиною лінгвістичної змінної;

- $G$  – синтаксична процедура, що дає змогу оперувати елементами термножини  $T$ , зокрема генерувати нові терми (значення);
- $M$  – семантична процедура, що дає змогу перетворити кожне нове значення лінгвістичної змінної, утвореної процедурою  $G$ , на нечітку змінну, тобто сформувати відповідну нечітку множину

Найважливішим застосуванням теорії нечітких множин є контролери нечіткої логіки. Їх функціонування дещо відрізняється від роботи звичайних контролерів: для опису системи замість диференціальних рівнянь використовують знання експертів. Ці знання можуть бути виражені за допомогою лінгвістичних змінних, які описані нечіткими множинами.

Нечітке управління особливо корисне, коли технологічні процеси занадто складні для аналізу за допомогою загальноприйнятих кількісних методів, або коли доступні джерела інформації інтерпретуються на якісному рівні неточно чи невиразно. Експериментально показано, що нечітке управління у певних випадках дає кращі результати порівняно з отримуваними в разі загальноприйнятих алгоритмів управління.

Загальну структуру системи, що використовує нечітку логіку, ілюструє рис. 2.1.

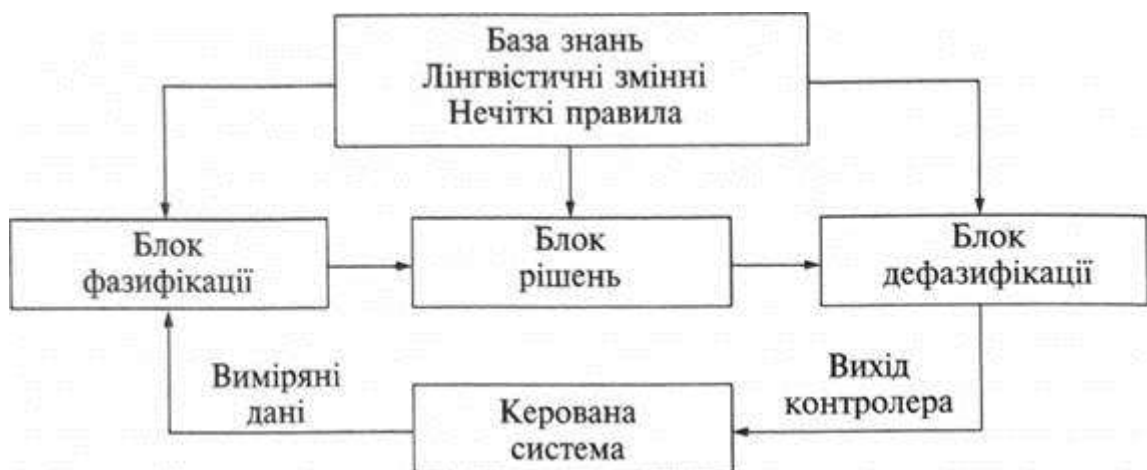


Рисунок 2.2 - Загальна структура системи з нечіткою логікою

Вона містить у своєму складі такі складові: блок фазифікації; базу знань; блок рішень; блок дефазифікації.

Блок фазифікації перетворює чіткі величини, виміряні на виході об'єкта керування, на нечіткі величини, що описані лінгвістичними змінними в базі знань.

Блок рішень використовує нечіткі умовні (if – then) правила, закладені в базі знань, для перетворення нечітких вхідних даних на необхідні керівні впливи, що також носять нечіткий характер.

Блок дефазифікації перетворює нечіткі дані з виходу блоку рішень на чітку величину, яку застосовують для керування об'єктом.

Нечітка логіка, на якій засноване нечітке управління, ближча за духом до людського мислення і вживаних мов, ніж традиційні логічні системи. Нечітка логіка в основному забезпечує ефективні засоби відображення невизначеностей і неточностей реального світу. Наявність математичних засобів відображення нечіткості вихідної інформації дає змогу побудувати модель, адекватну реальності.

### **2.2.2 Аналіз доцільності моделювання на основі методу опорних векторів (SVM)**

В машинному навчанні метод опорних векторів — це метод аналізу даних для класифікації та регресійного аналізу за допомогою моделей з керованим навчанням з пов'язаними алгоритмами навчання, які називаються опорно-векторними машинами (ОВМ, англ. support vector machines, SVM, також опорно-векторними мережами, англ. support vector networks[1]). Для заданого набору тренувальних зразків, кожен із яких відмічено як належний до однієї чи іншої з двох категорій, алгоритм тренування ОВМ будує модель, яка відносить нові зразки до однієї чи іншої категорії, роблячи це неймовірнісним бінарним лінійним класифікатором. Модель ОВМ є представленням зразків як точок у просторі, відображених таким чином, що зразки з окремих категорій

розділено чистою прогалиною, яка є щонайширшою. Нові зразки тоді відображаються до цього ж простору, й робиться передбачення про їхню належність до категорії на основі того, на який бік прогалини вони потрапляють.

На додачу до виконання лінійної класифікації, ОВМ можуть ефективно виконувати нелінійну класифікацію при застосуванні так званого ядрового трюку, неявно відображуючи свої входи до просторів ознак високої вимірності.

Коли дані не є міченими, кероване навчання є неможливим, і виникає необхідність у спонтанному навчанні, яке намагається знайти природне кластерування даних на групи, а потім відображувати нові дані на ці сформовані групи. Алгоритм кластерування, який забезпечує вдосконалення опорно-векторним машинам, називається опорно-векторним кластеруванням (англ. support vector clustering)[2], і часто використовується в промислових застосуваннях, коли дані або не є міченими, або коли лише деякі дані є міченими як попередня обробка перед проходом класифікації.

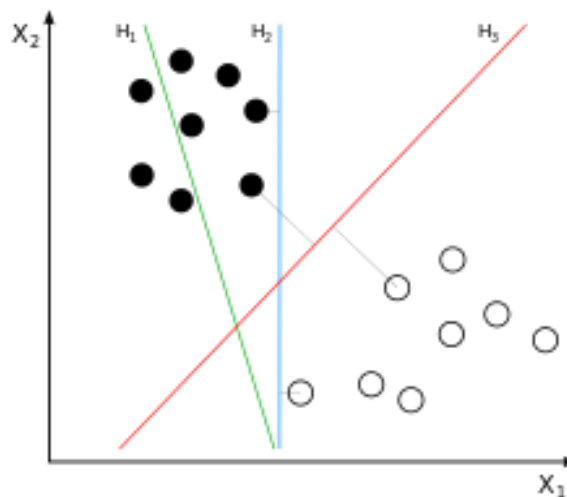


Рисунок 2.3 -  $H_1$  не розділяє класи.  $H_2$  розділяє, але лише з невеликим розділенням.  $H_3$  розділяє їх із максимальним розділенням.

В машинному навчанні поширеною задачею є класифікація даних. Розгляньмо деякі задані точки даних, кожна з яких належить до одного з двох

класів, а метою є вирішувати, в якому класі буде нова точка даних. У випадку опорно-векторних машин точку даних розглядають як вимірний вектор чисел, і хочуть дізнатися, чи можливо розділити такі точки вимірною гіперплощиною. Це називається лінійним класифікатором.

Існує багато гіперплощин, які могли би розділяти одні й ті ж дані. Одним із варіантів розумного вибору найкращої гіперплощини є такий, який пропонує найбільший проміжок, або розділення між двома класами. Тож ми обираємо гіперплощину таким чином, щоби відстань від неї до найближчих точок даних з кожного боку була максимальною. Така гіперплощина, якщо вона існує, відома як максимально розділова гіперплощина (англ. *maximum-margin hyperplane*), а лінійний класифікатор, що вона його визначає, — як максимально розділовий класифікатор (англ. *maximum margin classifier*); або, рівнозначно, як перцептрон оптимальної стабільності (англ. *perceptron of optimal stability*).

Первинний алгоритм ОВМ було винайдено Володимиром Вапником та Олексієм Червоненкісом[en] 1963 року. 1992 року Бернхард Босер, Ізабель Гійон та Володимир Вапник запропонували спосіб створювати нелінійні класифікатори шляхом застосування до максимально розділових гіперплощин ядрового трюку.[9] Поточне стандартне втілення (м'яке розділення) було запропоновано Корінною Кортес та Володимиром Вапником 1993 року й опубліковано 1995 року.

Формальніше, опорно-векторна машина будує гіперплощину, або набір гіперплощин у просторі високої або нескінченної вимірності, які можна використовувати для класифікації, регресії та інших задач. Інтуїтивно, добре розділення досягається гіперплощиною, яка має найбільшу відстань до найближчих точок тренувальних даних будь-якого з класів (так зване функційне розділення), оскільки в загальному випадку що більшим є розділення, то нижчою є похибка узагальнення класифікатора.

В той час, як первинну задачу може бути сформульовано у скінченновимірному просторі, часто трапляється так, що множини, які треба



розрізняти, не є лінійно роздільними в ньому. З цієї причини було запропоновано відображувати первинний скінченновимірний простір до простору набагато вищої вимірності, згодом роблячи розділення простішим у тому просторі. Для збереження помірних обчислювальних навантажень, відображення, які використовуються методом опорних векторів, розробляють такими, щоби забезпечувати можливість простого обчислення скалярних добутків у термінах змінних первинного простору, визначаючи їх у термінах ядрових функцій, що їх обирають відповідно до задачі.[3] Гіперплощини в просторі вищої вимірності визначаються як геометричне місце точок, чий скалярний добуток з вектором у цьому просторі є сталими. Вектори, які визначають гіперплощини, можуть обиратися як лінійні комбінації з параметрами відображень векторів ознак, які трапляються в базі даних. За такого вибору гіперплощини, точки простору ознак, які відображаються на гіперплощину, визначаються відношенням.

Параметри максимально розділеної гіперплощини виводяться шляхом розв'язання задачі оптимізації. Існує кілька спеціалізованих алгоритмів для швидкого розв'язання задач КП, що виникають в ОВМ, вони здебільшого покладаються на евристичні методи для розбиття задачі на менші кусені, з якими легше впоруватися.

Іншим підходом є застосування методу внутрішньої точки[en], який використовує ньютоні-подібні ітерації для пошуку розв'язку умов Каруша — Куна — Такера прямої та двоїстої задач.[29] Замість розв'язання послідовності розбитих задач, цей підхід безпосередньо розв'язує задачу в цілому. Для уникнення розв'язання лінійної системи з великою ядровою матрицею в ядровому трюку часто використовується низькорангове наближення матриці.

Іншим поширеним методом є алгоритм послідовної мінімальної оптимізації Платта, який розбиває задачу на 2-вимірні підзадачі, які розв'язуються аналітично, усуваючи потребу в алгоритмі числової оптимізації та в зберіганні матриці. Цей алгоритм є простим концептуально, простим у

реалізації, зазвичай швидшим, і має кращі властивості масштабування для складних задач ОБМ[30].

Окремий випадок лінійних опорно-векторних машин може розв'язуватися ефективніше алгоритмами того ж роду, що й використовуються для оптимізації їхнього близького родича, логістичної регресії; цей клас алгоритмів включає субградієнтний спуск (наприклад, PEGASOS[31]) та координатний спуск(наприклад, LIBLINEAR[32]). LIBLINEAR володіє деякими привабливими властивостями часу тренування. Кожна ітерація збіжності займає час, лінійний по відношенню до часу, витраченого на читання тренувальних даних, й ітерації також володіють властивістю Q-лінійної збіжності, що робить цей алгоритм надзвичайно швидким.

Звичайні ядрові ОБМ також можуть розв'язуватися ефективніше при застосуванні субградієнтного спуску (наприклад, P-rackSVM[33]), особливо якщо дозволено розпаралелювання.

Ядрові ОБМ доступні в багатьох інструментаріях машинного навчання, включно з LIBSVM, MATLAB, SAS, SVMlight, kernlab, scikit-learn, Shogun, Weka, Shark, JKernelMachines, OpenCV та іншими.

### **2.2.3 Аналіз можливості створення моделі на основі нейронних мереж**

Основою нейромережі є прості, та у більшій мірі однотипні, елементи, які імітують роботу нейронів мозку (штучні нейрони).

Кожен нейрон має таку характеристику як поточний стан, по аналогії з нервовими клітинами головного мозку, які можуть бути збуджені або заторможені. Він володіє групою синапсів – однонаправлених вхідних зв'язків, сполучених з виходами інших нейронів, а також має аксон – вихідний зв'язок даного нейрона, з яким сигнал поступає на синапси наступних нейронів [9].

Кожен синапс характеризується величиною синаптичного зв'язку або його вагою, який за фізичним сенсом еквівалентний електричній провідності.

Загальна риса нейромереж: реалізується принцип паралельного оброблення сигналів (шляхом об'єднання великої кількості нейронів в шари и з'єднання). При чому оброблення взаємодій всіх нейронів ведеться пошарово. Теоретично кількість шарів і кількість нейронів в кожному шарі може бути довільним, однак фактично воно обмежено ресурсами комп'ютера [9].

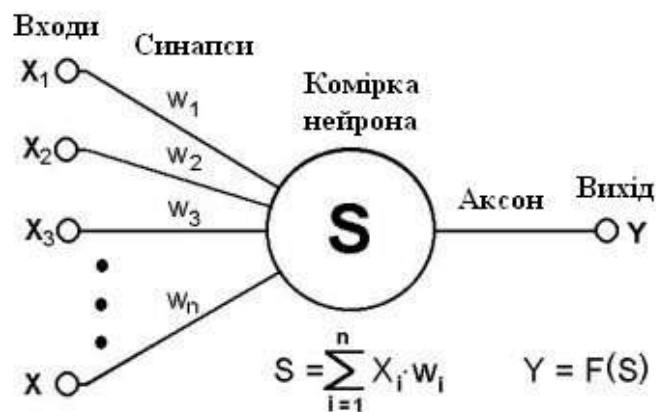


Рисунок 2.4 – Схема штучного нейрону

Архітектура багатозарової нейронної мережі складається з послідовно з'єднаних шарів, де нейрон кожного шару своїми входами пов'язаний з усіма нейронами попереднього шару, а виходами - наступного. НМ з двома вирішальними шарами може з будь-якою точністю апроксимувати будь-яку багатовимірну функцію. Навчаються багатозарові НМ за допомогою алгоритму зворотного поширення помилки, що є методом градієнтного спуску в просторі ваг з метою мінімізації сумарної помилки мережі. При цьому помилки (точніше, величини корекції ваг) поширюються в зворотному напрямку від входів до виходів, крізь ваги, що з'єднують нейрони.

Незважаючи на принцип роботи нейронної мережі, існує велика кількість підходів до її реалізації, варіантів архітектури та алгоритмів навчання. Одні нейронні мережі простіші для реалізації, інші є не тільки складними у реалізації, але й передбачають наявність значних ресурсів для навчання. Ці

твердження також можна застосувати і до алгоритмів навчання нейронних мереж.

Проаналізувавши літературу, можна зробити висновок про високу ефективність нейронної мережі прямого поширення, а також те, що складність її реалізації є задовільною, порівняно з іншими архітектурами нейронних мереж.

Нейронна мережа прямого поширення є однією з найпопулярніших архітектур нейронних мереж не тільки через складність своєї реалізації, але й продуктивність та ефективність своєї роботи.

Найбільшою відмінністю цього типу нейронних мереж є те, що її складники не формують циклу, на відміну від рекурентних нейронних мереж, для прикладу.

У нейронній мережі прямого поширення нейрон отримує дані від інших нейронів, які відносно нього знаходяться на рівні вище, і передає результат на рівень нижче. Отже, згідно з назвою нейронної мережі, інформація в ній має тільки один шлях поширення, а саме від входу до виходу, проходячи при цьому через всі приховані шари нейронної мережі.

Формування результату в цій нейронній мережі є стандартним підходом, коли інформація передається через синапси з певними вагами в нейрони, які сумують отримані результати і застосовують до отриманих даних функцію активації (передавальна функція)

Такий процес є незмінним та відбувається крізь усі приховані шари нейронної мережі.

Як зазначено вище, нейронна мережа містить приховані шари. Прихованими вважають шари нейронної мережі, які розміщені між вхідним та вихідним шарами.

Приховані шари, а саме їхня кількість, є одним з вирішальних факторів ефективності нейронної мережі, які мають безпосередній вплив на точність результату.

Чим більше прихованих шарів, тим вимогливіша до ресурсів нейронна мережа. Тому ефективною нейронною мережею можна назвати таку, в якій правильно підібрана низка факторів, враховуючи кількість прихованих шарів. Значний вплив на точність результату нейронної мережі має також і обраний алгоритм її тренування. Тренування є ключовим етапом у створенні нейронної мережі, який відбувається безпосередньо після її реалізації. Зважаючи на кількість типів нейронних мереж, алгоритмів їх навчання є також кілька. Базовий принцип навчання нейронної мережі є по суті корегування ваг на синапсах для мінімізації похибки, якою є різниця між результатом, отриманим на виході нейронної мережі, та еталонними даними.

### **2.3 Моделювання інформаційної технології прогнозування результатів футбольних матчів засобами нечіткої логіки**

Прийняття рішень у проблемно-орієнтованих інформаційних системах та системах керування здійснюється в умовах апіорної невизначеності, обумовленої неточністю або неповнотою вхідних даних, стохастичною природою зовнішніх впливів, відсутністю адекватної математичної моделі функціонування, нечіткістю мети, людським фактором [1–3] та ін. Невизначеність системи призводить до зростання ризиків від прийняття неефективних рішень, результатом чого можуть бути негативні економічні, технічні та соціальні наслідки.

Невизначеності у системах прийняття рішень компенсують за допомогою різноманітних методів штучного інтелекту. Для ефективного прийняття рішень при невизначеності умов функціонування системи застосовують методи на основі правил нечіткої логіки. Такі методи ґрунтуються на нечітких множинах і використовують лінгвістичні величини і висловлювання для опису стратегій прийняття рішень [4–6].

Методи нечітких множин особливо корисні за відсутності точної математичної моделі функціонування системи. Теорія нечітких множин дає

можливість застосувати для прийняття рішень неточні та суб'єктивні експертні знання про предметну область без формалізації їх у вигляді традиційних математичних моделей.

З використанням теорії нечітких множин вирішуються питання узгодження суперечливих критеріїв прийняття рішень, створення логічних регуляторів систем. Нечіткі множини дають змогу застосовувати лінгвістичний опис складних процесів, встановлювати нечіткі відношення між поняттями, прогнозувати поведінку системи, формувати множину альтернативних дій, виконувати формальний опис нечітких правил прийняття рішень.

Методи теорії нечітких множин є зручним засобом проектування інтерфейсів у людино-машинних системах. На основі нечіткого логічного виведення будуються системи керування, подання знань, підтримки прийняття рішень, апроксимації, структурної та параметричної ідентифікації, розпізнавання образів, оптимізації. Нечітка логіка знаходить застосування у побутовій електроніці, діагностиці, різноманітних експертних системах. Нечіткі експертні системи для підтримки прийняття рішень знаходять широке застосування у військовій справі, медицині та економіці. З їх допомогою здійснюють бізнес-прогнозування, оцінювання ризиків та прибутковості інвестиційних проектів. На основі нечіткої логіки досліджують глобальні політичні рішення та моделюють кризові ситуації [7, 8].

Важливим застосуванням теорії нечітких множин є контролери нечіткої логіки, які використовуються у різноманітних системах керування, зокрема у побутових приладах. Замість математичної моделі для опису системи такі контролери використовують інтегровані знання експертів, які за структурою подання наближаються до розмовної мови і описуються за допомогою лінгвістичних змінних та нечітких множин [9–12].

Загальна структура fuzzy-контролера містить у своєму складі такі складові: блок фазифікації; база знань; блок рішень; блок дефазифікації.

Блок фазифікації перетворює чіткі величини, виміряні на виході об'єкта керування, на нечіткі величини, описані лінгвістичними змінними у базі знань.

Блок рішень використовує нечіткі умовні (if – then) правила, закладені у базі знань, для перетворення нечітких вхідних даних на необхідні керуючі впливи, що мають також нечіткий характер.

Блок дефазифікації перетворює нечіткі дані з виходу блоку рішень на чітку величину, яка подається на виконавчий пристрій для керування об'єктом.

З огляду на широке поширення систем штучного інтелекту з інтегрованою нечіткою логікою, розроблення ефективних систем прийняття рішень на їх основі є актуальною науково-практичною проблемою.

Перспектива застосування нечіткої логіки полягає у розробленні гібридних методів штучного інтелекту, до яких можна віднести нечіткі штучні нейронні мережі, адаптивне поповнення баз нечітких правил, підтримка нечітких запитів до баз даних, побудова нечітких когнітивних карт, нечіткі графи, нечіткі мережі Петрі, нечіткі дерева прийняття рішень, нечітка кластеризація та ін. [13–15].

### 2.3.1 Формалізація моделі прогнозування результатів футбольних матчів

У даній роботі запропонована обрати підхід та модель прогнозування результатів футбольних матчів молодшої ліги України, в основу якого покладено формалізовані у вигляді нечіткої бази знань лінгвістичні висловлювання-правила типу: ЯКЩО - ТО. Для налаштування і тестування моделі використовувалися дані про матчі чемпіонатів молодшої ліги України по футболу 2017/2018 рр.

Опираючись на матеріали статті [3] можна стверджувати, що Український футбольний чемпіонат являє собою турнір чотирнадцяти команд, переможець яких визначається кількістю набраних балів в очних зустрічах з усіма суперниками.

Чемпіонат проходить в два кола, тобто кожна команда грає 26 поєдинків.

З кібернетичної точки зору завдання побудови моделі прогнозування результату футбольного матчу зводиться до пошуку функціонального відображення вигляду:

$$X = \{x_1, x_2, \dots, x_n\} \rightarrow D \in \{d_1, d_2, d_3\}, \quad (2.1)$$

де  $X$ – вектор факторів, що впливають на результат, якими можуть бути рівень майстерності команд, погодні умови, результати минулих зустрічей команд, місце проведення гри ( вдома або в гостях) і т. д. ;

$D$ – результат зустрічі, який оцінюється термами:

$d_1$ – виграш першої команди (господаря поля),

$d_2$ – нічия,

$d_3$ – виграш другої(гостьової) команди.



Для використання запропонованого методу прогнозування результатів футбольних матчів в [3] потрібно відібрати такі чинники, які мають найбільший вплив на результат поєдинку:

$x_1$ – різниця втрат головних гравців (різниця між кількістю травмованих і дискваліфікованих футболістів в першій команді (господаря поля) і кількістю травмованих і дискваліфікованих футболістів в гостьовій команді);

$x_2$ – різниця ігрових динамік (різниця балів, набраних командою-господарем поля і гостьовий командою в останніх п'яти турах);

$x_3$  - різниця в класі команд (різниця місць, які займають команда-господар і гостьова команда в поточному чемпіонаті);

$x_4$ – «фактор поля» (розраховуватися як  $HP/HG - GP/GG$ , де  $HP$  – загальна кількість балів, набраною командою-господарем поля в домашніх іграх поточного чемпіонату;  $HG$  – загальна кількість домашніх матчів, проведених командою-господарем у поточному чемпіонаті;  $GP$  – загальна кількість балів, набраних гостьовою командою в поточному чемпіонаті на виїзді;  $GG$  – загальна кількість виїзних ігор, проведених гостьовий командою в поточному чемпіонаті);

$x_5$ – «зустрічі команд» (різниця забитих і пропущених м'ячів двох команд у всіх чемпіонатах України).

Зауважимо, що значення цих факторів не є конфіденційною інформацією - їх можна легко визначити до початку футбольного матчу.

Нечітка модель являє собою апроксимацію залежності «входи – вихід» на основі лінгвістичних висловлювань типу «ЯКЩО – ТО» і операцій нечіткого логічного висновку [1, 2, 5]. Основні положення теорії нечітких множин, що використовуються в цьому розділі, наведені в Додатку Г. Типова структура нечіткої моделі показана на рис. 2.1.



Рисунок 2.1 – Узагальнена структура системи нечіткого логічного висновку

Нечітка модель містить наступні блоки:

- фаззифікатор перетворює фіксований вектор факторів, що впливають ( $X$ ) в вектор нечітких множин, необхідних для виконання нечіткого логічного висновку  $\tilde{X}$ ;
- нечітка база знань містить інформацію про залежності  $Y = f(X)$  у вигляді лінгвістичних правил типу «ЯКЩО – ТО»;
- машина нечіткого логічного висновку на основі правил бази знань визначає значення вихідної змінної у вигляді нечіткої множини  $\tilde{Y}$ , відповідного нечітким значенням вхідних змінних ( $\tilde{X}$ );
- дефаззифікатор перетворює вихідну нечітку сукупність  $\tilde{Y}$  в чітке число.

На рис. 2.2 зображено блок-схему системи нечіткого логічного висновку з  $n$  входами та одним виходом. Пакет MatLab дозволяє проектувати системи нечіткого логічного висновку двох типів [7]: Мамдані та Сугено.

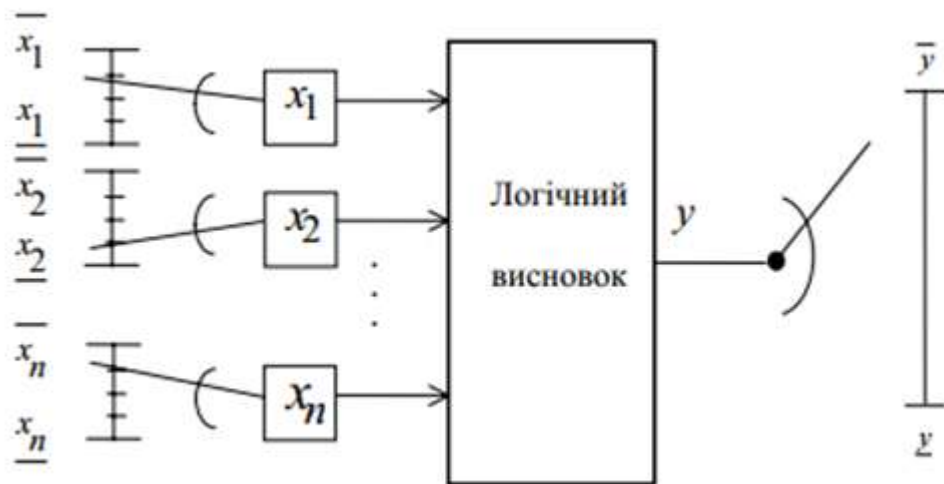


Рисунок 2.2 – Блок-схема системи нечіткого логічного висновку з кількісними змінними

Для нечіткого моделювання залежності (2.1) необхідно:

- представити вхідні ( $x_1 - x_5$ ) і вихідну ( $y$ ) змінні у вигляді лінгвістичних змінних;
- формалізувати у вигляді нечіткої бази знань експертні лінгвістичні висловлювання про взаємозв'язок входів і виходу;
- провести навчання нечіткої моделі шляхом налаштування функцій функцій приналежності і ваг правил з метою мінімізації відхилення між результатами моделювання та експериментальними даними.

Відповідно до [3, 4] синтез нечіткої бази знань може трактуватися як етап структурної ідентифікації залежності (2.1) по експертно-експериментальним даним. В цьому випадку база знань являє собою грубу модель залежності (2.1), яку потім необхідно налаштувати на етапі параметричної ідентифікації.

### 2.3.2 Моделювання в середовищі MATLAB та робота з модулем Fuzzy Logic Toolbox

MATLAB (англ. «Matrix Laboratory») термін, що відноситься до пакету прикладних програм для вирішення завдань технічних обчислень, а також до

використовуваної в цьому пакеті мови програмування. MATLAB працює на більшості сучасних операційних систем, включаючи Linux, Mac OS і Microsoft Windows. MATLAB, як мова програмування, була розроблена в кінці 1970-х років з метою полегшення процесів програмування для студентів (мова розділу Simulink отримала назву візуального проектування). Нова мова була з великим інтересом зустрінута вченими, що працюють в області прикладної математики. Вдосконалений варіант MATLAB на мові C з'явився в 1984 р. Спочатку MATLAB призначався для проектування систем управління, але швидко завоював популярність в багатьох інших наукових і інженерних областях. Він також широко використовувався і в освіті, зокрема, для викладання лінійної алгебри і чисельних методів. Мова MATLAB є високорівневою інтерпретуємою мовою програмування, що включає засновані на матрицях структури даних, широкий спектр функцій, інтегроване середовище розробки, об'єктно-орієнтовані можливості і інтерфейси до програм, написаних на інших мовах програмування.

MATLAB надає зручні засоби для розробки алгоритмів, включаючи високорівневі з використанням концепцій об'єктно-орієнтованого програмування. У ньому є всі необхідні засоби інтегрованого середовища розробки, включаючи налаштування і профайлер. Функції для роботи з цілими типами даних полегшують створення алгоритмів для мікро контролерів і інших застосувань, де це необхідно.

У складі пакету MATLAB є велика кількість функцій для побудови графіків, зокрема тривимірних, візуального аналізу даних і створення анімованих роликів. Вбудоване середовище розробки дозволяє створювати графічні інтерфейси користувача з різними елементами управління, такими як кнопки, поля введення і іншими. За допомогою компоненту MATLAB Compiler ці графічні інтерфейси можуть бути перетворені в самостійні застосування, для запуску яких на інших комп'ютерах необхідна бути встановлена бібліотека MATLAB Component Runtime.

Операції з нечіткою логікою у пакеті Matlab дає змогу виконувати модуль Fuzzy Logic Toolbox. Він дозволяє створювати системи нечіткого логічного виведення і нечіткої класифікації в рамках середовища MatLab, з можливістю їхньої інтеграції в засіб Simulink пакета Matlab.

Перша категорія програмних інструментів пакета Fuzzy Logic Toolbox містить функції, які можуть бути викликані безпосередньо шляхом набору імені функції в командному вікні або із власних призначених для користувача додатків. Більшість із цих функцій є функціями Matlab у вигляді m-файлів. У даному випадку користувач може подивитися запрограмовані в цих функціях алгоритми, а також редагувати і корегувати ці файли.

Друга категорія програмних інструментів пакета Fuzzy Logic Toolbox містить діалогові модулі. Блоки для пакета Simulink містять модулі, які забезпечують інтеграцію систем нечіткого логічного висновку з пакетом Simulink: Fuzzy Logic Toolbox — це пакет прикладних програм, що входять до складу середовища MatLab. Він дозволяє створювати системи нечіткого логічного висновку і нечіткої класифікації в рамках середовища MatLab, з можливістю їх інтеграції в Simulink(рис. 2.3–2.4). Базовим поняттям Fuzzy Logic Toolbox є FIS-структура — система нечіткого висновку (Fuzzy Inference System). FIS-структура містить усі необхідні дані для реалізації функціонального відображення «входи-виходи» на основі нечіткого логічного висновку відповідно до схеми.

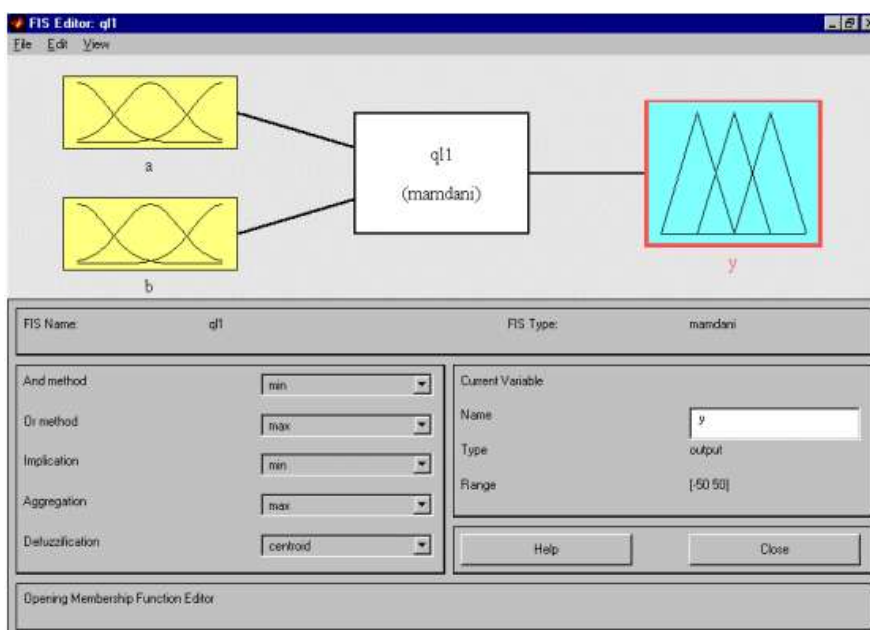


Рисунок 2.3 – Структура системи нечіткого логічного висновку

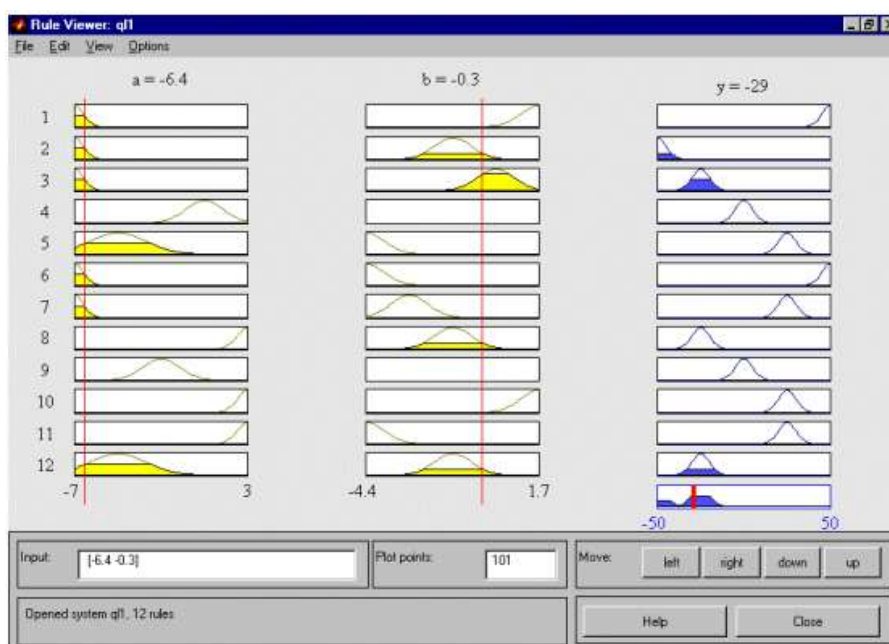


Рисунок 2.4 – Приклад візуалізації нечіткого висновку

- Fuzzy Logic Toolbox містить такі категорії програмних інструментів:
- функції;
  - інтерактивні модулі з графічним, призначеним для користувача, інтерфейсом (GUI);

- більшості функцій за допомогою графічного інтерфейсу. Крім того, ці модулі забезпечують зручне середовище для проектування, дослідження і впровадження систем на основі нечіткого логічного висновку. Для запуску інтерактивних модулів досить надрукувати ім'я модуля в командному рядку.

— блоки для пакета Simulink;

— демонстраційні приклади.

Модуль Fuzzy Logic Toolbox містить такі категорії програмних інструментів: функції; інтерактивні модулі з графічним користувальницьким інтерфейсом (GUI). Блоки для пакета Simulink можуть бути викликані з командного рядка. Для отримання переліку функцій слід ввести команду: `help fuzzy`.

У Fuzzy Logic Toolbox передбачена можливість для користувача створення власної функції приналежності. Для цього необхідно створити m-функцію, що містить два вхідних аргументи: вектор, для координат якого необхідно розрахувати значення функції приналежності, і вектор параметрів функції приналежності (рис. 2.5). Вихідним аргументом функції має бути вектор ступенів приналежності.

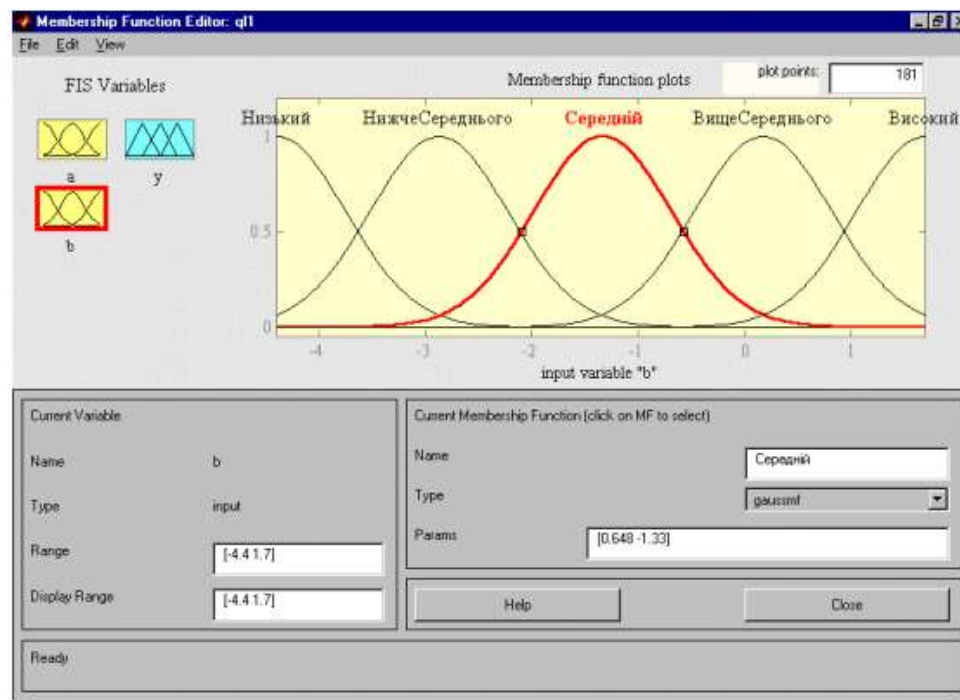


Рисунок 2.5 – Приклад функції належності змінної b

Функції FIS: `addmf` — додає функцію приналежності до FIS; `addrule` — додає правило до FIS; `addvar` — додає змінну до FIS; `defuzz` — дефазифікує функцію приналежності; `evalfis` — здійснює обчислення нечіткого виведення; `evalmf` — обчислює функцію приналежності; `gensurf`— генерує поверхню виходу FIS; `getfis` — повертає властивості нечіткої системи; `mf2mf` — транслює параметри між функціями приналежності; `newfis`— створює нову FIS; `parsrule` — аналізує нечіткі правила; `plotfis` — показує діаграму «вхід-вихід» для FIS; `plotmf`— показує усі функції приналежності для однієї змінної; `readfis` — завантажує FIS із диска; `rmmf` — видаляє функцію приналежності з FIS; `rmvar` — видаляє змінну з FIS; `setfis` — установлює властивості нечіткої системи; `showfis` — показує анотовану FIS; `showrule`— відображує правила FIS; `writefis` — зберігає FIS на диску.

Функції кластер-аналізу: `fcm`, `genfis1`, `genfis2`, `subclust`.

Різні функції: `convertfis` — перетворює нечітку матрицю структури версії 1.0 на матрицю структури версії 2.0; `discfis`— дискретизує FIS; `evalmmf` — використовується для обчислення множинних функцій приналежності; `fstrvcats` — поєднує матриці різного розміру; `fuzarith` — функція нечіткої арифметики; `findrow`— шукає рядки матриці, що відповідають вхідному рядку; `probor`— імовірнісне «АБО»; `sugmax` — максимальний вихідний діапазон для системи Сугено.

Допоміжні функції графічного користувальницького інтерфейсу: `cmfdlg` — додає діалог вибору функцій приналежності; `cmthdlg`— додає діалог вибору методу виведення; `fisgui` — дискрипторне посилання на інтерфейсні засоби модуля Fuzzy Logic Toolbox; `gfmfdlg` — генерує FIS з використанням діалогу методу ґрат; `mfdlg`— додає діалог функції приналежності; `mfdrag`— перетягування функції приналежності за допомогою миші; `popundo` — відновлює зі стеку останні зміни (відмінює останні дії); `pushundo`— передає поточні дані у стек відновлення; `savedlg` — діалог запису перед закриттям; `statmsg` — зображує повідомлення у полі статусу; `updtfis`— оновлює засоби графічного інтерфейсу Fuzzy Logic Toolbox; `wsdlg` — діалог «відкриття з» /



«збереження до» робочої зони. FIS-структура (Fuzzy Inference System — система нечіткого виведення) містить усі необхідні дані для реалізації функціонального відображення «входи-виходи» на основі нечіткого логічного виведення.

Для доступу до властивостей системи нечіткого логічного виведення треба вказати ім'я відповідного поля. Система нечіткого виведення зберігається на диску у вигляді fis-файла — текстового файлу спеціального формату. Функції readfis та writefis викорис-

Меню File — це загальне меню для всіх GUI-модулів, використовуваних із системами нечіткого виведення. За допомогою команди New-FIS користувач має можливість створити нову систему нечіткого виведення. У разі вибору цієї команди з'являться дві альтернативи: Mamdani та Sugeno, що визначають тип створюваної системи. За допомогою команди Import користувач має можливість завантажити раніше створену систему нечіткого логічного виведення. Команда Export дозволяє зберегти систему нечіткого виведення в робочу зону або на диск. Команда Print дозволяє вивести на принтер копію графічного вікна. Команда Close закриває графічне вікно.

Меню Edit. Команда Undo скасовує раніше зроблену дію. Команда Add Variable дозволяє додати в систему ще одну змінну: вхідну (input) або вихідну (output). Команда Remove Selected Variable видаляє поточну змінну із системи. Команда Membership Function відкриває редактор функцій приналежностей. Команда Rules відкриває редактор БЗ.

Меню View — це загальне меню для всіх GUI-модулів, використовуваних із системами нечіткого виведення. Це меню дозволяє відкрити вікно візуалізації нечіткого виведення (команда Rules) та вікно виведення поверхні «вхід-вихід», що відповідає системі нечіткого виведення (команда Surface).

Меню And Method, Or Method, Implication та Aggregation дозволяють установити реалізації логічних операцій «ТА», «АБО», імплікація, об'єднання, відповідно. Меню Defuzzification дозволяє вибрати метод дефазифікації. Користувач також має можливість установити власну реалізацію операцій.

Для цього необхідно вибрати команду Custom та у графічному вікні, що з'явилося, надрукувати ім'я функції, яка реалізує цю операцію.

Панель Current Variable дозволяє для поточної змінної, ім'я котрої задається у полі Name, задати у полі Type тип (вхідна —товуються для завантаження у робочу зону та збереження на диску таких файлів. FIS-редактор (FIS Editor) призначений для створення, збереження, завантаження і виведення на друк систем нечіткого логічного виведення, а також для редагування таких властивостей: тип системи; найменування системи; кількість вхідних і вихідних змінних; найменування вхідних і вихідних змінних; параметри нечіткого логічного виведення. FIS-редактор містить три загальносистемних меню: File, Edit, View та п'ять меню для вибору параметрів нечіткого виведення: And Method, Or Method, Implication, Aggregation та Defuzzification.

Редактор функцій приналежності (Membership Function Editor) призначений для задавання такої інформації про термимножини вхідних і вихідних змінних: кількість термів; найменування термів; тип і параметри функцій приналежності, котрі необхідні для подання лінгвістичних термів у вигляді нечітких множин. Редактор функцій приналежності містить меню File, Edit, View та панелі введення інформації.

Меню Edit. Команда Add MFs дозволяє додати терми у термножину, використовувану для лінгвістичної оцінки поточної змінної. При виборі цієї команди з'явиться діалогове вікно, в якому необхідно вибрати тип функції приналежності та кількість термів. Значення параметрів функцій приналежності будуть установлені автоматично таким чином, щоб рівномірно покрити область визначення змінної, заданої у полі Range. При зміні області визначення у полі Range параметри функцій приналежності будуть промаштабовані. Команда Add Custom MF дозволяє додати лінгвістичний терм, функція приналежності якого відрізняється від убудованих. Після вибору цієї команди з'явиться графічне вікно, в якому необхідно надрукувати лінгвістичний терм (поле MF-name), ім'я функції приналежності (поле M-File

function name) і параметри функції приналежності (поле Parameter list). Команда Remove Selected MF видаляє поточний терм із термножини поточної змінної. Ознакою поточної змінної є червона окантовка її прямокутника. Ознакою поточного терму є червоний колір його функції приналежності. Для вибору поточного терму необхідно позиціонувати курсор миші на графіку функції приналежності і клацнути лівою кнопкою миші. Команда Remove All MFs видаляє всі терми з термножини поточної змінної.

Панель Current Variable дозволяє для поточної змінної, ім'я котрої вказане у полі Name, а тип у полі Type, задати у полі Range діапазон значень, а у полі Display Range — діапазон відображення. Панель Current Membership Function дозволяє для поточної функції приналежності, ім'я терму котрої вказують у полі Name, задати тип (назву Matlab-функції) у полі Type та параметри у полі Params. Редактор також має панель із зображенням графіка поточної функції приналежності.

Редактор БЗ (Rule Editor) призначений для формування і модифікації нечітких правил. Редактор функцій приналежності містить чотири системних меню: File, Edit, View, Options, меню вибору термів вхідних і вихідних змінних, поля установки логічних операцій «ТА», «АБО», «НЕ» і ваг правил, а також кнопки редагування та перегляду правил.

Для введення нового правила в БЗ необхідно за допомогою миші вибрати відповідну комбінацію лінгвістичних термів вхідних і вихідних змінних, установити тип логічного зв'язування («ТА» або «АБО») між змінними всередині правила, установити наявність чи відсутність логічної операції «НЕ» для кожної лінгвістичної змінної, увести значення вагового коефіцієнта правила і натиснути кнопку Add Rule. За замовчуванням установлені параметри: логічне зв'язування змінних усередині правила — «ТА»; логічна операція «НЕ» — відсутня; значення вагового коефіцієнта правила — 1.

Можливі випадки, коли істинність правила не змінюється за довільного значення деякої вхідної змінної, тобто ця змінна не впливає на результат

нечіткого логічного виведення в даній області факторного простору. Натомість лінгвістичне значення цієї змінної необхідно установити none.

Для видалення правила з БЗ потрібно клацнути один раз лівою кнопкою миші на цьому правилі та натиснути кнопку Delete Rule.

Для модифікації правила необхідно клацнути один раз лівою кнопкою миші на цьому правилі, потім установити необхідні параметри правила і натиснути кнопку Edit Rule.

Меню Options дозволяє установити мову і формат правил БЗ. При виборі команди Language з'явиться список мов English (Англійська), Deutsch (Німецька), Francais (Французька), з якого необхідно вибрати одну. При виборі команди Format з'явиться список можливих форматів правил БЗ: Verbose — лінгвістичний; Symbolic — логічний; Indexed — індексований.

Модуль візуалізації нечіткого виведення (Rule Viewer) дає змогу проілюструвати хід логічного виведення за кожним правилом, одержати результативну нечітку множину й виконати процедуру дефазифікації. Rule Viewer містить чотири меню: File, Edit, View, Options, два поля введення інформації — Input і Plot points та кнопки прокручування зображення вліво — вправо (left-right), догори — вниз (up-down).

Кожне правило БЗ подається у вигляді послідовності горизонтально розташованих прямокутників. При цьому перші два прямокутника. Порожній прямокутник у візуалізації правила означає, що в цьому правилі посилки за змінною немає. Жовте заливання графіків функцій приналежностей вхідних змінних указує, наскільки значення входів відповідають термам даного правила. Для виведення правила у форматі Rule Editor необхідно клацнути один раз лівою кнопкою миші по номеру відповідного правила. У цьому випадку зазначене правило буде виведено в нижній частині графічного вікна. Блакитне заливання графіка функції приналежності вихідної змінної являє собою результат логічного виведення у вигляді нечіткої множини за даним правилом. Результативну нечітку множину, що відповідає логічному виведенню за всіма правилами, показано в нижньому прямокутнику

останнього стовпця графічного вікна. У цьому самому прямокутнику червона вертикальна лінія відповідає чіткому значенню логічного виведення, отриманого в результаті дефазифікації.

## 2.4 Висновок

У даному розділі обґрунтовано вибір інтелектуальної технології для ідентифікації та моделювання прогнозування результатів футбольних матчів молодшої ліги України. Для цього розглянуто три існуючі технології, такі як

- нечітка логіка;
- метод опорних векторів (SVM);
- нейроні мережі.

В результаті аналізу вибрано нечітку логіку. Ця технологія заснована на створенні нечіткої продукційної бази знань, що представляє собою матрицю нечітких правил «ЯКЩО – ТО» та механізму нечіткого логічного виводу, заснованому на дереві рішень.

Вибрану технологію доопрацьовано за рахунок введення додаткового фактору - кількість запасних гравців, що підвищує адекватність моделі прогнозування.

Вибрано та описано процеси моделювання нечітких систем в середовищі MATLAB та роботу з модулем ь Fuzzy Logic Toolbox.

Було наведено структурну організацію експертної системи та представлення результатів роботи системи.

### 3 ПРОГРАМНА РЕАЛІЗАЦІЯ ТЕХНОЛОГІЇ

#### 3.1 Збір вхідних даних про результати футбольних матчів молодшої ліги

Як зазначалось в п. 2.1 для побудови нечітких правил основними вхідними даними варто відібрати наступні параметри:

- різниця втрат головних гравців;
- різниця ігрових динамік;
- різниця в класі команд;
- «фактор поля»;
- «зустрічі команд».

Основним джерелом такого роду інформації по матчам молодшої ліги є сайт ДЮФЛУ (рис. 3.1).

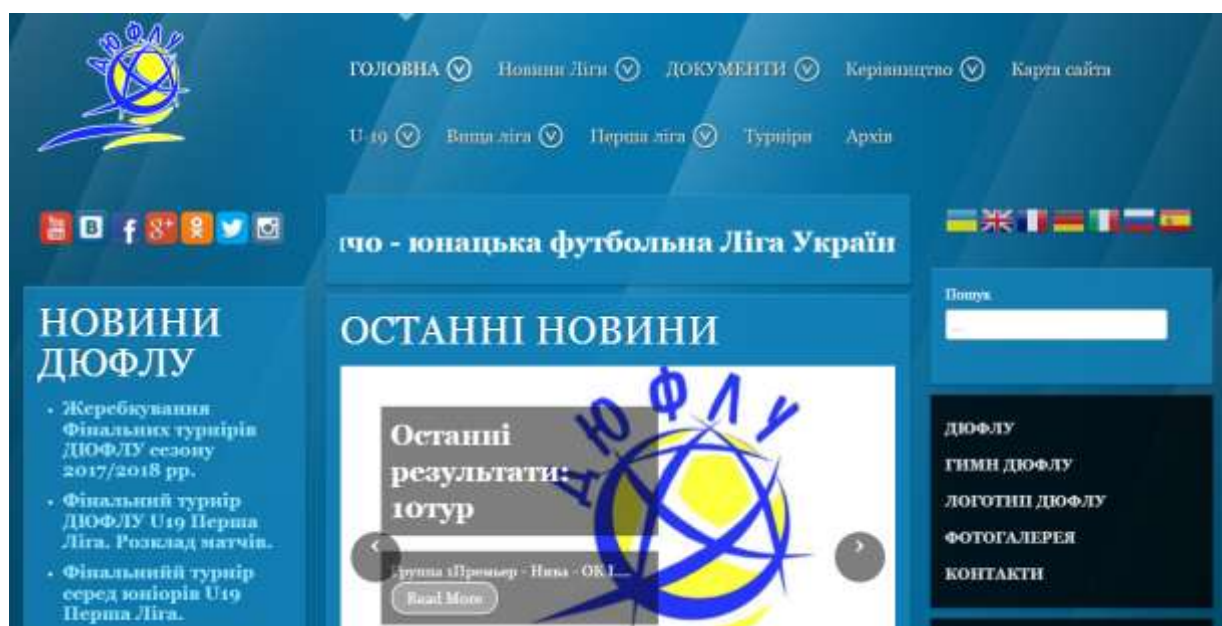


Рисунок 3.1 – Загальний вигляд сайту ДЮФЛУ

На сайті представлена інформація як про матчі в турнірі (рис. 3.2.)

Також можна переглянути результати матчів, які вже пройшли (рис. 3.3).



Кісткові матчі Всеукраїнських змагань з футболу  
серед команд дитячо-юнацьких спортивних закладів  
сезону 2017/2018 років.



### Вища ліга.

#### U-17 перші матчі

№	місце	Команда	-	місце	команда	дата
1	4-е м. гр.1	Атлет м. Київ	-	1-е м. гр. 4	УФК-Карпати м. Львів	02.06.18
2	3-е м. гр.1	КОДЮСШ Арсенал сест. Шахове	-	2-е м. гр. 4	Карпати м. Львів	02.06.18
3	4-е м. гр.4	ВРВ-ВІК м. Володимир-Вол.	-	1-е м. гр. 1	Зіна-Оболонь м. Київ	02.06.18
4	3-е м. гр.4	Прикарпаття м. Івано-Франківськ	-	2-е м. гр. 1	Арсенал-Київ м. Київ	02.06.18
5	4-е м. гр.2	ДЮСШ Металург м. Запоріжжя	-	1-е м. гр. 3	МФК Кремень м. Кременчук	02.06.18
6	3-е м. гр.2	Дніпро м. Дніпро	-	2-е м. гр. 3	ДЮФШ ФК Ворскла м. Полтава	02.06.18
7	4-е м. гр.3	Арсенал Харків	-	1-е м. гр. 2	Динамо м. Київ	02.06.18
8	3-е м. гр.3	ДЮСШ-11 Чорноморець м. Одеса	-	2-е м. гр. 2	Шахтар м. Донецьк	03.06.18

#### U-17 матчі відповіді

№	місце	Команда	-	місце	команда	дата
9	1-е м. гр.4	УФК-Карпати м. Львів	-	4-е м. гр. 1	Атлет м. Київ	10.06.18
10	2-е м. гр.4	Карпати м. Львів	-	3-е м. гр. 1	КОДЮСШ Арсенал сест. Шахове	09.06.18
11	1-е м. гр.1	Зіна-Оболонь м. Київ	-	4-е м. гр. 4	ВРВ-ВІК м. Володимир-Вол.	09.06.18
12	2-е м. гр.1	Арсенал-Київ м. Київ	-	3-е м. гр. 4	Прикарпаття м. Івано-Франківськ	10.06.18
13	1-е м. гр.3	МФК Кремень м. Кременчук	-	4-е м. гр. 2	ДЮСШ Металург м. Запоріжжя	09.06.18
14	2-е м. гр.3	ДЮФШ ФК Ворскла м. Полтава	-	3-е м. гр. 2	Дніпро м. Дніпро	09.06.18
15	1-е м. гр.2	Динамо м. Київ	-	4-е м. гр. 3	Арсенал Харків	09.06.18
16	2-е м. гр.2	Шахтар м. Донецьк	-	3-е м. гр. 3	ДЮСШ-11 Чорноморець м. Одеса	10.06.18

#### U-16 перші матчі

№	місце	Команда	-	місце	команда	дата
1	4-е м. гр.1	Юність м. Чернігів	-	1-е м. гр. 4	УФК-Карпати м. Львів	02.06.18
2	3-е м. гр.1	Барса м. Суми	-	2-е м. гр. 4	Карпати м. Львів	02.06.18
3	4-е м. гр.4	ВРВ-ВІК м. Володимир-Вол.	-	1-е м. гр. 1	Зіна-Оболонь м. Київ	02.06.18
4	3-е м. гр.4	ФК Скала м. Моршин	-	2-е м. гр. 1	Арсенал-Київ м. Київ	03.06.18
5	4-е м. гр.2	ДЮСШ-15 м. Київ	-	1-е м. гр. 3	Торпедо-ВУФК м. Миколаїв	02.06.18
6	3-е м. гр.2	Дніпро м. Дніпро	-	2-е м. гр. 3	Зірка м. Кропивницький	02.06.18
7	4-е м. гр.3	ДЮФШ ФК Ворскла м. Полтава	-	1-е м. гр. 2	Динамо м. Київ	02.06.18
8	3-е м. гр.3	ФК Маріуполь м. Маріуполь	-	2-е м. гр. 2	Шахтар м. Донецьк	02.06.18

Рисунок 3.2 – Інформація про матчі в турнірі на сайті ДЮФЛУ

Команда	Ігор	Пе	Ні	По	Очки	М'ячі
Колос Ковалівка	1	1	0	0	3	1-0
Волинь Луцьк	1	1	0	0	3	3-2
УФК-Олімпік Харків	2	0	0	2	0	2-4
Коло	Тур	Дата	Господарі	М'ячі	Гості	
Коло	1					
	Тур	1				
		27-05-2018				
			УФК-Олімпік Харків	2:3	Волинь Луцьк	
	Тур	2				
		31-05-2018				

Рисунок 3.3 – Результати матчів, [першість України 2017/2018, перша ліга](#)

U19

Тому проблеми із отриманням вхідних даних – немає. Далі варто продумати і прописати етапи розробки нечіткої моделі прогнозування.

### 3.2 Розробка бази знань

Як зазначалось в п.2.1 після визначення вхідних перемінних в першу чергу варто визначитися із лінгвістичними змінними. Для підвищення високої точності прогнозування при оцінці вихідної змінної варто використовувати п'ять лінгвістичних терм:

- «Велика поразка»;
- «Поразка»
- «Нічия»;
- «Виграш»;
- «Великий виграш».

Ці лінгвістичні змінні варто задати для універсальної множини різниці голів, забитих командою-господарем і гостьовою командою. В даному випадку:

$d_1$  – Виграш та Великий виграш;

$d_2$  – Нічия;

$d_3$  – Поразка та Велика поразка.

Для лінгвістичної оцінки вхідних і вихідних змінних сформуємо множину термів. Усю можливу множину термів наведено в табл. 3.1.



Таблиця 3.1 – Терм-множина вхідних та вихідних змінних

Змінна	Терм-множини	Параметри функції приналежності	
		c	b
$x_1$	Велика лава (BC)	2,55	-6
	Однакова лава (OC)	2,55	0
	Коротка лава (KC)	2,55	6
$x_2$	Суттєва поразка (СП)	4,25	-15
	Поразка (П)	4,25	-5
	Виграш (В)	4,25	5
	Суттєвий виграш (СВ)	4,25	15
$x_3$	Лідер (Л)	2,76	-13
	Верхня половина (ВП)	2,76	-6,5
	Середина (С)	2,76	0
	Нижня половина (НП)	2,76	6,5
	Аутсайдер (А)	2,76	13
$x_4$	Абсолютна навдача (АНд)	0,7	-2
	Невдача (Нд)	0,7	-0,33
	Перевага (Пр)	0,7	1,33
	Абсолютна перевага (АПр)	0,7	3
$x_5$	Позорна зустріч (Пз)	8,5	-20
	Рівна зустріч (Р)	8,5	0
	Розгромна зустріч (Рз)	8,5	20
$x_6$	Велика поразка (ВП)	0,64	-3
	Поразка (П)	0,44	-0,9
	Нічия (Н)	0,44	0
	Виграш (В)	0,44	0,9
	Великий виграш (ВВ)	0,64	3

Формалізацію лінгвістичних термів проводиться за допомогою гауссівської функції приналежності [6]:

$$\mu^t(x) = e^{\frac{-(x-b)^2}{2c^2}}, \quad (3.1)$$

де,  $\mu^t(x)$  – функція належності змінної  $x$  до терма  $t$ ;

$b$  – параметр функції належності, який відповідає координаті факсимуму ( $\mu^t(x) = 1$ );

$c$  - параметр стиснення-розтягування функції належності.

Параметри функцій належності для кожного лінгвістичного терма наведені в табл. 3.1. Нацьому етапі параметри функцій приналежності видібрано експериментально, згодом на етапі налаштування вони будуть оптимізовані.

Наступним кроком є формування нечіткої бази знань.

Експертні лінгвістичні висловлювання, які відображають взаємозв'язок між факторами  $x_1 - x_5$  і результатом футбольного матчу ( $y$ ), наведено в табл. 3.2.

Таблиця 3.2 – Нечітка база знань

№	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$y$
1	BC	CB	Л	Апр	Рз	ВВ
2	OC	В	ВП	Пр	Рз	ВВ
3	OC	П	Л	Пр	Рз	ВВ
4	BC	В	ВП	Пр	Р	ВВ
5	OC	В	С	Нд	Рз	В
6	КС	П	ВП	Пр	Р	В
7	OC	В	С	Нд	Рз	В
8	BC	CB	НП	Пр	Р	В
9	OC	В	С	Нд	Р	Н
10	КС	СП	С	Нд	Р	Н
11	OC	П	НП	Пр	Пз	Н
12	BC	СП	ВП	Нд	Р	Н
13	BC	П	С	АНд	Р	П

14	ОС	В	НП	Нд	Пз	П
15	КС	СП	С	Пр	Пз	П
16	ОС	П	А	Нд	Р	П
17	КС	СП	А	АНд	Р	ВП
18	ОС	СП	НП	Нд	Пз	ВП
19	КС	П	НП	АНд	Р	ВП
20	ВС	СП	НП	Нд	Пз	ВП

Кожна стрічка таблиці відповідає одному правилу, наприклад, для першої стрічки це наступне правилоа:

ЯКЩО  $x_1 = \text{«Велика лава»}$  І

$x_2 = \text{«Суттєвий виграш»}$  І

$x_3 = \text{«Лідер»}$  І

$x_4 = \text{«Абсолютна перевага»}$  І

$x_5 = \text{«Розгромна зустріч»}$

ТО  $y = \text{«Великий виграш»}$ .

Отримана нечітка база знань дозволяє сформувати нечіткий логічний висновок результату футбольного матчу. Саме ці залежності і використовувались в для написання правил в Matlab. На онові отриманих результатів маємо підстави, щоб сформулювати алгоритм прогнозування футбольних матчів молодшої ліги.

### 3.3 Розробка алгоритму прогнозування футбольних матчів

Прогнозування результатів конкретного футбольного матчу можна проводити на наступним алгоритмом:

Крок 1. Визначити значення факторів, які характерні для конкретного матчу ( $X^* = x_1, x_2, x_3, x_4, x_5$ ).

Крок 2. Розрахувати ступені приналежності значень факторів до нечітких термів з бази знань, які наведені в табл 3.2.

Крок 3. Підставити знайдені в кроці 2 значення в систему нечітких логічних рівнянь, які сформовані на основі нечіткої бази знань и визначити степінь приналежності рішення до термів «Велика поразка», «Поразка», «Нічия», «Виграш», «Великий виграш».

Крок 4. Визначити результируючу нечітку множину як об'єднання відповідних нечітких множин:

$$\bar{y} = \bigcup_{q \in (\text{ВП, П, Н, В, ВВ})} \int_{-3}^3 \min(\mu_q(X^*), \mu_q(y)) / y \quad (3.2)$$

Крок 5. Провести дефазифікацію отриманої на попередньому кроці множини з використанням методу центра тяжіння [7]:

$$y = \frac{\int_{-3}^3 y \mu_{\bar{y}}(y) dy}{\int_{-3}^3 \mu_{\bar{y}}(y) dy}, \quad (3.3)$$

Крок 6. По отриманій різниці голів визначити результат футбольного матчу:

$$D = \begin{cases} d_1, & \text{if } y \in (0.5, 3] \\ d_2, & \text{if } y \in [-0.5, 0.5] \\ d_3, & \text{if } y \in [-3, -0.5) \end{cases} \quad (3.4)$$

Схема алгоритму прогнозування результату футбольного матчу молодшої ліги на основі нечітких правил наведено на 3.4.



Рисунок 3.4 – Схема алгоритму прогнозування результату футбольного матчу молодшої ліги на основі нечітких правил

Для перевірки правильності роботи алгоритму було розглянуто матч «Прем'єр-Нива» Вінниця та «Атлет» Київ, який закінчився із рахунком 3:1.

Матчу відповідають наступні значення впливаючих факторів:

$$x_1 = -2;$$

$$x_2 = 7;$$

$$x_3 = -8;$$

$$x_4 = 1,75;$$

$$x_5 = 17.$$

Запропонована нечітка модель прогнозує різницю голів  $y = 1,72$ , що відповідає рішенням  $d_1$  – домашня перемога. На рис. 3.5 проілюстровано процедуру нечіткого логічного висновку для цього матчу з використанням інструментів Fuzzy Logic Toolbox програмної середовища MatLab.

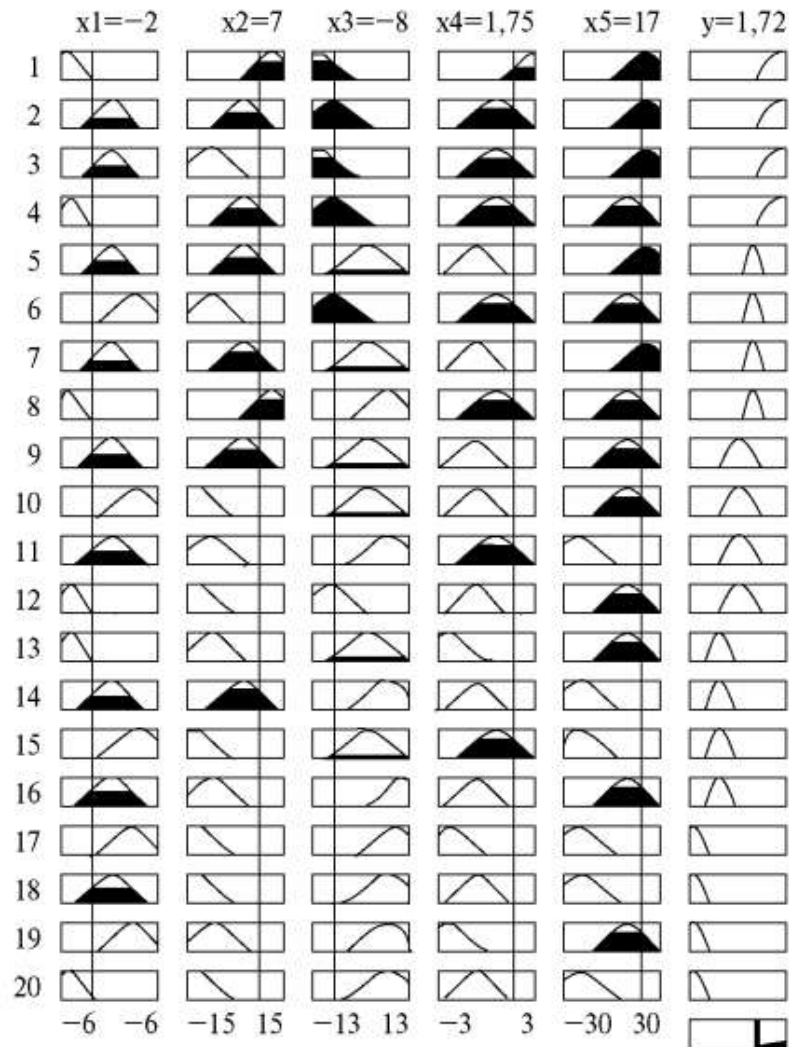


Рисунок 3.5 – Процес прогнозування результату матчу «Прем'єр-Нива»  
Вінниця та «Атлет» Київ.

Відомо, що налаштування нечіткої моделі полягає в знаходженні таких її параметрів, які мінімізують середньоквадратичне відхилення між експериментальними даними з навчальної вибірки і результатами моделювання. Параметри моделі, які можна налаштувати є параметри функції приналежності для кожного лінгвістичного терма з бази знань і ваги правил.

Загальна кількість параметрів, що настраюються становить 64 (20 правил і 22 терма по два параметра функції приналежності для кожного).

Налаштування нечіткої моделі було проведено на вибірці з 65 матчів, зіграних у двох чемпіонатах молодшої Першої ліги U16 та U17.

Оптимальні параметри нечіткої моделі - функції приналежності і ваги правил наведені в табл. 3.3

Таблиця 3.3 – Оптимальні ваги правил

№	1	2	3	4	5	6	7	8	9	10
Вага	0,5	0,95	0,5	0,63	0,5	0,75	0,5	1	0	0,22
№	11	12	13	14	15	16	17	18	19	20
Вага	0,1	0,08	0,01	0,29	0,3	1	1	1	1	1

### 3.4 Тестування роботи програмного забезпечення

Тестування моделі на вибірці з 105 матчів показує, що середньоквадратичне відхилення становить 1.60, при цьому статистична оцінка ймовірності правильного передбачення результату поединку (виграш, нічия, програш) складає 0,64, що вдва рази краще, ніж при випадковому вгадуванні.

Проілюструємо можливості запропонованих моделей на прикладі прогнозування результатів Молодшої Першої ліги Першості України 2017/2018 рр. U17. Вихідною інформацією, необхідною для прогнозування, є результати ігор за попередні роки та результати перших трьох турів Першості України 2017/2018 рр.

На основі цієї інформації необхідно спрогнозувати результати решти 10 турів. У кожному турі відбувається 6 матчів, тому необхідно спрогнозувати результати  $10 \times 6 = 60$  ігор.

Використовуючи підхід, який описано в [3] прогнозування будемо здійснювати за аналогічною методикою.

Спочатку спрогнозуємо результати 4-го туру. Потім з урахуванням отриманих даних перерахуємо значення факторів  $x_2 - x_5$  і спрогнозуємо результати 5-го туру, потім 6-го і т. д. по 13-йтур.

Результати прогнозування у вигляді залежності точності прогнозування від кількості турів у вигляді підсумкової таблиці чемпіонату показані в табл. 3.4.

Таблиця 3.4 – Результуюча турнірна таблиця

Команда	Результати Першості		Прогноз нечіткої моделі	
	бали	місце	бали	місце
«ДВУОР» ім. С.Бубки	24	1	25	1-2
«МДЮСШ», Дніпро	22	2	25	1-2
«Поділля», Хмельницький	17	3	19	3
«Арена», Харків	16	4	17	4
«Геліос», Харків	14	5	15	5
«СДЮСШОР», Металург Запоріжжя	13	6	14	6-7
«Ковель-Волинь», Ковель	12	7	14	6-7
«ДЮСШ-1», Хмельницький	11	8	11	9
ФК «Мрія-ЮФК», Куп'янськ	9	9	12	8
«ДЮСШ-2», Дніпро	8	10	9	10

Як видно з таблиці 3.4, нечітка модель достатньо точно визначає лідерів і аутсайдерів, тобто найбільш важливі результати першості - команди, що займають призові місця команди, які вилітають із Першої ліги.



### **3.5 Висновок**

Третій розділ магістерської кваліфікаційної роботи присвячено проектуванню та програмній реалізації модуля прогнозування результатів футбольних матчів молодшої ліги на основі доопрацьованої інформаційної технології. В ході виконання даного розділу було проведено формалізацію та фазифікацію доопрацьованої спеціалізованої моделі. Було знайдено платформу для збору вхідних даних про результати футбольних матчів молодшої ліги. На основі зібраної інформації та експертних висновків було розроблено нечітку базу знань прогнозування та алгоритм прогнозування результатів футбольних матчів молодшої ліги. Також було проведено відповідне тестування роботи системи. Лістинг програми наведено в додатку В, інструкцію користувача роботи в системі – в додатку Б.

## 4 ЕКОНОМІЧНА ЧАСТИНА

### 4.1 Оцінювання комерційного потенціалу розробки

Метою проведення технологічного аудиту є оцінювання комерційного потенціалу розробки. Для проведення технологічного аудиту було залучено 2-х незалежних експертів. Такими експертами будуть доц.каф.КН Сілагін Олексій Віталійович та доц.каф.КН Арсенюк Ігор Ростиславович

Здійснюємо оцінювання комерційного потенціалу розробки за 12-ма критеріями за 5-ти бальною шкалою.

Результати оцінювання комерційного потенціалу розробки наведено в таблиці 4.1.

Таблиця 4.1 – Результати оцінювання комерційного потенціалу розробки

Критерії	Прізвище, ініціали, посада експерта	
	1. Експерт 1	2. Експерт 2
	Бали, виставлені експертами:	
1	4	4
2	3	3
3	4	4
4	4	4
5	3	3
6	3	4
7	4	3
8	3	4
9	4	4
10	4	4
11	3	4
12	3	4
Сума балів	СБ <sub>1</sub> = 43	СБ <sub>2</sub> = 45
Середньоарифметична сума балів $\overline{СБ}$	$\overline{СБ} = \frac{\sum_{i=1}^3 СБ_i}{2} = 44$	

Отже, з отриманих даних таблиці 4.1 видно, що нова розробка має високий рівень комерційного потенціалу.

#### 4.2 Прогнозування витрат на виконання науково-дослідної роботи та конструкторсько–технологічної роботи.

Для розробки нового програмного продукту необхідні такі витрати.

Основна заробітна плата для розробників визначається за формулою (4.1):

$$Z_o = \frac{M}{T_p} \cdot t, \quad (4.1)$$

де M- місячний посадовий оклад конкретного розробника;

$T_p$  - кількість робочих днів у місяці,  $T_p = 22$  дні;

t - число днів роботи розробника, t = 45 днів.

Розрахунки заробітних плат для керівника і програміста наведені в таблиці 4.2.

Таблиця 5.2 – Розрахунки основної заробітної плати

Працівник	Оклад M, грн.	Оплата за робочий день, грн.	Число днів роботи, t	Витрати на оплату праці, грн.
Науковий керівник	6000	272,72	5	1363,6
Інженер-програміст	3500	159,09	45	7159,05
Всього:				8522,65

Розрахуємо додаткову заробітну плату:

$$Z_{\text{дод}} = 0,1 \cdot 8522,65 = 852,26 \text{ (грн.)}$$

Нарахування на заробітну плату операторів НЗП розраховується як 37,5...40% від суми їхньої основної та додаткової заробітної плати:

$$H_{\text{зп}} = (Z_o + Z_p) \cdot \frac{\beta}{100}, \quad (5.2)$$

$$H_{\text{зп}} = (8522,65 + 852,26) \cdot \frac{36,3}{100} = 3403,09 \text{ (грн.)}.$$

Розрахунок амортизаційних витрат для програмного забезпечення виконується за такою формулою:

$$A = \frac{Ц \cdot H_a}{100} \cdot \frac{T}{12}, \quad (4.3)$$

де Ц – балансова вартість обладнання, грн;

$H_a$  – річна норма амортизаційних відрахувань % (для програмного забезпечення 25%);

T – Термін використання (T=3 міс.).

Таблиця 5.3 – Розрахунок амортизаційних відрахувань

Найменування програмного забезпечення	Балансова вартість, грн.	Норма амортизації, %	Термін використання, міс.	Величина амортизаційних відрахувань, грн
Персональний комп'ютер	10000	25	3	625
Всього:				625

Розрахуємо витрати на комплектуючі. Витрати на комплектуючі розрахуємо за формулою:

$$K = \sum_1^n H_i \cdot Ц_i \cdot K_i, \quad (4.4)$$

де n – кількість комплектуючих;

$H_i$  - кількість комплектуючих і-го виду;

$Ц_i$  – покупна ціна комплектуючих і-го виду, грн;

$K_i$  – коефіцієнт транспортних витрат (прийmemo  $K_i = 1,1$ ).

Таблиця 4.4 - Витрати на комплектуючі, що були використані для розробки ПЗ.

Найменування матеріалу	Одиниці виміру	Ціна, грн.	Витрачено	Вартість витрачених матеріалів, грн.
Пачка паперу	уп.	150	1	150
Ручка	шт.	10	1	10
Всього з урахуванням транспортних витрат				176

Витрати на силову електроенергію розраховуються за формулою:

$$V_e = V \cdot \Pi \cdot \Phi \cdot K_{\Pi} ; \quad (4.5)$$

де  $V$  – вартість 1кВт-години електроенергії ( $V=1,7$  грн/кВт);

$\Pi$  – установлена потужність комп'ютера ( $\Pi=0,6$ кВт);

$\Phi$  – фактична кількість годин роботи комп'ютера ( $\Phi=150$  год.);

$K_{\Pi}$  – коефіцієнт використання потужності ( $K_{\Pi} < 1$ ,  $K_{\Pi} = 0,9$ ).

$$V_e = 1,7 \cdot 0,6 \cdot 150 \cdot 0,9 = 137,7 \text{ (грн.)}$$

Розрахуємо інші витрати  $V_{ін}$ .

Інші витрати  $I_v$  можна прийняти як (100...300)% від суми основної заробітної плати розробників та робітників, які були виконували дану роботу, тобто:

$$V_{ін} = (1..3) \cdot (Z_o + Z_p). \quad (4.6)$$

Отже, розрахуємо інші витрати:

$$V_{ін} = 1 * (8522,65 + 852,26) = 9374,91 \text{ (грн.)}$$

Сума всіх попередніх статей витрат дає витрати на виконання даної частини роботи:

$$V = Z_o + Z_d + H_{зп} + A + K + V_e + I_v$$

$$V = 8522,65 + 852,25 + 3403,09 + 625 + 176 + 137,7 + 9374,91 = 23091,61 \text{ (грн.)}$$

Розрахуємо загальну вартість наукової роботи  $B_{\text{заг}}$  за формулою:

$$B_{\text{заг}} = \frac{B_{\text{ін}}}{\alpha} \quad (4.7)$$

де  $\alpha$  – частка витрат, які безпосередньо здійснює виконавець даного етапу роботи, у відн. одиницях = 1.

$$B_{\text{заг}} = \frac{23091,61}{1} = 23091,61$$

Прогнозування загальних витрат ЗВ на виконання та впровадження результатів виконаної наукової роботи здійснюється за формулою:

$$\text{ЗВ} = \frac{B_{\text{заг}}}{\beta} \quad (4.8)$$

де  $\beta$  – коефіцієнт, який характеризує етап (стадію) виконання даної роботи.

Отже, розрахуємо загальні витрати:

$$\text{ЗВ} = \frac{23091,61}{0,9} = 25657,34 \text{ (грн.)}$$

### **4.3 Прогнозування комерційних ефектів від реалізації результатів розробки.**

Спрогнозуємо отримання прибутку від реалізації результатів нашої розробки. Зростання чистого прибутку можна оцінити у теперішній вартості грошей. Це забезпечить підприємству (організації) надходження додаткових коштів, які дозволять покращити фінансові результати діяльності .

Оцінка зростання чистого прибутку підприємства від впровадження результатів наукової розробки. У цьому випадку збільшення чистого прибутку підприємства  $\Delta \Pi_i$  для кожного із років, протягом яких очікується отримання позитивних результатів від впровадження розробки, розраховується за формулою:

$$\Delta \Pi_i = \sum_1^n (\Delta \Pi_{\text{я}} \cdot N + \Pi_{\text{я}} \Delta N)_i \quad (4.9)$$

де  $\Delta \Pi_{\text{я}}$  – покращення основного якісного показника від впровадження результатів розробки у даному році;

$N$  – основний кількісний показник, який визначає діяльність підприємства у даному році до впровадження результатів наукової розробки;

$\Delta N$  – покращення основного кількісного показника діяльності підприємства від впровадження результатів розробки;

$\Pi_{\text{я}}$  – основний якісний показник, який визначає діяльність підприємства у даному році після впровадження результатів наукової розробки;

$n$  – кількість років, протягом яких очікується отримання позитивних результатів від впровадження розробки.

В результаті впровадження результатів наукової розробки витрати на виготовлення інформаційної технології зменшаться на 25 грн (що автоматично спричинить збільшення чистого прибутку підприємства на 25 грн), а кількість користувачів, які будуть користуватись збільшиться: протягом першого року – на 200 користувачів, протягом другого року – на 175 користувачів, протягом третього року – 150 користувачів. Реалізація інформаційної технології до впровадження результатів наукової розробки складала 1000 користувачів, а прибуток, що отримував розробник до впровадження результатів наукової розробки – 200 грн.

Спрогнозуємо збільшення чистого прибутку від впровадження результатів наукової розробки у кожному році відносно базового.

Отже, збільшення чистого продукту  $\Delta P_1$  протягом першого року складатиме:

$$\Delta P_1 = 25 \cdot 1000 + (200 + 25) \cdot 200 = 70000 \text{ грн.}$$

Протягом другого року:

$$\Delta P_2 = 25 \cdot 1000 + (200 + 25) \cdot (200 + 175) = 109375 \text{ грн.}$$

Протягом третього року:

$$\Delta P_3 = 25 \cdot 1000 + (200 + 25) \cdot (200 + 175 + 150) = 143125 \text{ грн.}$$

#### **4.4 Розрахунок ефективності вкладених інвестицій та період їх окупності**

Визначимо абсолютну і відносну ефективність вкладених інвестором інвестицій та розрахуємо термін окупності.

Абсолютна ефективність  $E_{\text{абс}}$  вкладених інвестицій розраховується за формулою:

$$E_{\text{абс}} = (\text{ПП} - PV), \quad (4.10)$$

де  $\Delta P_i$  – збільшення чистого прибутку у кожному із років, протягом яких виявляються результати виконаної та впровадженої НДДКР, грн;

$t$  – період часу, протягом якого виявляються результати впровадженої НДДКР, 3 роки;

$\tau$  – ставка дисконтування, за яку можна взяти щорічний прогнозований рівень інфляції в країні; для України цей показник знаходиться на рівні 0,1;

$t$  – період часу (в роках) від моменту отримання чистого прибутку до точки 2, 3, 4.

Рисунок, що характеризує рух платежів (інвестицій та додаткових прибутків) буде мати вигляд, рисунок 5.1.





Рисунок 4.1 – Вісь часу з фіксацією платежів, що мають місце під час розробки та впровадження результатів НДДКР

Розрахуємо вартість чистих прибутків за формулою:

$$\text{ПП} = \sum_1^m \frac{\Delta\Pi_i}{(1+\tau)^t} \quad (4.11)$$

де  $\Delta\Pi_i$  – збільшення чистого прибутку у кожному із років, протягом яких виявляються результати виконаної та впровадженої НДДКР, грн;

$t$  – період часу, протягом якого виявляються результати впровадженої НДДКР, роки;

$\tau$  – ставка дисконтування, за яку можна взяти щорічний прогнозований рівень інфляції в країні; для України цей показник знаходиться на рівні 0,1;

$t$  – період часу (в роках) від моменту отримання чистого прибутку до точки.

Отже, розрахуємо вартість чистого прибутку:

$$ПП = \frac{25657,34}{(1+0,1)^0} + \frac{70000}{(1+0,1)^2} + \frac{109375}{(1+0,1)^3} + \frac{143125}{(1+0,1)^4} = 263439,92 \text{ (грн.)}$$

Тоді розрахуємо  $E_{\text{абс}}$ :

$$E_{\text{абс}} = 263439,92 - 25657,34 = 237782,58 \text{ грн.}$$

Оскільки  $E_{\text{абс}} > 0$ , то вкладання коштів на виконання та впровадження результатів НДДКР буде доцільним.

Розрахуємо відносну (щорічну) ефективність вкладених в наукову розробку інвестицій  $E_{\text{в}}$  за формулою:

$$E_{\text{в}} = \sqrt[T]{1 + \frac{E_{\text{абс}}}{PV}} - 1 \quad (4.12)$$

де  $E_{\text{абс}}$  – абсолютна ефективність вкладених інвестицій, грн;

$PV$  – теперішня вартість інвестицій  $PV = 3B$ , грн;

$T_{\text{ж}}$  – життєвий цикл наукової розробки, роки.

Тоді будемо мати:

$$E_{\text{в}} = \sqrt[3]{1 + \frac{237782,58}{25657,34}} - 1 = 1,17 \text{ або } 117 \%$$

Далі, розраховану величина  $E_{\text{в}}$  порівнюємо з мінімальною (бар'єрною) ставкою дисконтування  $\tau_{\text{мін}}$ , яка визначає ту мінімальну дохідність, нижче за яку інвестиції вкладатися не будуть. У загальному вигляді мінімальна (бар'єрна) ставка дисконтування  $\tau_{\text{мін}}$  визначається за формулою:

$$\tau = d + f,$$

де  $d$  – середньозважена ставка за депозитними операціями в комерційних банках; в 2019 році в Україні  $d = 0,2$ ;

$f$  – показник, що характеризує ризикованість вкладень, величина  $f = 0,1$ .

$$\tau = 0,2 + 0,1 = 0,3$$

Оскільки  $E_B = 117\% > \tau_{\text{мін}} = 0,3 = 30\%$ , то у інвестор буде зацікавлений вкладати гроші в дану наукову розробку.

Термін окупності вкладених у реалізацію наукового проекту інвестицій. Термін окупності вкладених у реалізацію наукового проекту інвестицій  $T_{\text{ок}}$  розраховується за формулою:

$$T_{\text{ок}} = \frac{1}{E_B}$$

$$T_{\text{ок}} = \frac{1}{1,17} = 0,85 \text{ року}$$

Обрахувавши термін окупності даної наукової розробки, можна зробити висновок, що фінансування даної наукової розробки буде доцільним.

## ВИСНОВКИ

В процесі виконання магістерської кваліфікаційної роботи було доопрацьовано існуючу інформаційну технологію прогнозування результатів футбольних матчів та створено експертну систему для прогнозування результатів футбольних матчів молодшої ліги України на основі нечітких правил.

Всі задачі, поставлені в завданні на магістерську кваліфікаційну роботу, виконано в повному об'ємі, а саме:

- проаналізовано та вибрано методологію прогнозування для задачі прогнозування результатів футбольних матчів;
- проаналізовано існуючі інтелектуальні технології ідентифікації задачі прогнозування результатів футбольних матчів та вибрано базову;
- сформульовано вимоги до роботи технології та розроблено ТЗ;
- допрацьовано існуючу технологію та модель ідентифікації задачі прогнозування результатів футбольних матчів;
- проведено математичне моделювання роботи етапів допрацьованої технології з використанням апарату нечіткої логіки;
- проведено фазифікацію розроблених моделей;
- розроблено та наповнено базу знань у вигляді матриць з правилами ЯКЦО-ТО;
- на основі допрацьованої технології виконано проектування та програмну реалізацію модуля прогнозування результатів футбольних матчів;
- проведено налаштування роботи модуля прогнозування результатів футбольних матчів та бази знань;
- протестовано роботу налаштованого модуля;
- виконані задачі економічного розділу.

Мета дослідження – збільшення точності прогнозування результатів футбольних матчів досягнута за рахунок введення в модель прогнозування додаткового фактору – кількості та стану запасних гравців, що підвищує адекватність моделі прогнозування і відповідно точність прогнозу.

## ПЕРЕЛІК ПОСИЛАНЬ

1. В.В.Когут, Б.В. Поліщук, О.В.Сілагін А.В.Шалак. Інжинірингова технологія створення експертних систем нечіткої логіки в *Матеріалах конференції «Молодь в науці: дослідження, проблеми, перспективи (МН-2020)»*, Вінниця, 2020. [Електронний ресурс].  
Режимдоступу: <https://conferences.vntu.edu.ua/index.php/all-mn/index/pages/view/zbirn2020> Дата звернення: Черв. 2020
2. Кулябичев Ю.П., Пивторацкая С.В. // Современные технологии в задачах управления, автоматизации и обработки информации: труды XVII международного научно-технического семинара. Алушта, сентябрь 2008 г. – СПб.: ГУАП, 2008. – С. 207.
3. А. А. Шумейко, С. Л. Сотник, М. В. Лысак. Использование генетических алгоритмов в задачах классификации текстов. Мир, 1993. – 220 с.
4. Штовба С.Д., Вивдюк В.В. Прогнозирование результатов футбольных матчей на основе нечетких правил // Вестник молодых ученых. Серия: Экономические науки. – 2002. – №6. – С.54–60.
5. Вентцель Е.С. Исследование операций: задачи, принципы, методология / Е.С. Вентцель. – М.: Наука, Главная редакция физико-математической литературы, 1980. – 320 с.
6. Строгалева В.П. Имитационное моделирование / В.П. Строгалева, И.О. Толкачева. – М.: МГТУ им. Баумана, 2008. – С.697-737.
7. Прикладные нечеткие системы: Пер. с япон. / К. Асан, Д. Ватада, С. Иван и др. Под ред. Т. Тэрано, К. Асан, М. Сугэно. М.: Мир, 1993. – 220 с.
8. Ротштейн А.П. Интеллектуальные технологии идентификации. Винница: Вінниця-УНІВЕРСУМ, 1999. – 240 с.
9. Ротштейн А.П. Кательников Д.и. Идентификация нелинейных зависимостей нечеткими базами знаний // Кибернетика и системный анализ. 1998. № 5. С. 53–61

10. Результати матчів туру ДЮФЛУ 2017/2018 р.р. – [Режим доступу до сервісу: <http://duflu.org.ua/index.php/novini-ligi/1642-rezultati-matchiv-turu-dyuflu-2017-2018-r-r>]
11. В.Ф. Худяков, В.А. Хабuzов. Моделирование источников вторичного электропитания в среде MATLAB 7.x: учебное пособие. СПб.: ГУАП, 2008, - 332 с.
12. И.Ф. Цисарь, В.Г. Нейман. Компьютерное моделирование экономики. М.: ДИАЛОГ-МИФИ, 2008.-384с.
13. И.Ф. Цисарь. MATLAB Simulink. Компьютерное моделирование экономики. М.: Солон-Пресс, 2008. - 256 с.
14. Теория вероятностей и математическая статистика на базе MATLAB. Издательство НТУ "ХПИ", 2006, Харьков, Украина, - 612 с.
15. Adams, T. (2017). Training an artificial neural network. Retrieved from: <https://www.solver.com/training-artificial-neural-networkintro>.
16. Badde, S., Gupta, A., & Patki, K. (2009). Cascade and Feed Forward Back propagation Artificial Neural Network Models for Prediction of Compressive Strength of Ready Mix Concrete. Second International Conference on Emerging Trends in Engineering (SICETE). Retrieved from: [http://www.iosrjournals.org/iosr-jmce/papers/sicete\(civil\)-volume3/26.pdf](http://www.iosrjournals.org/iosr-jmce/papers/sicete(civil)-volume3/26.pdf)
17. Galushkin, A. I. (2000). Teoriia neuronnykh setei. Moscow: IPRZhR. 415 p. [in Russian].
18. Gorban, A. N., & Rossiev, D. A. (1996). Neironnye seti na personalnom kompiutere. Novosibirsk: Nauka. 276 p. [in Russian].
19. Hryhorkiv, V. S., Yaroshenko, O. I., & Filipchuk, N. V. (2012). Neironni merezhi ta yikh vykorystannia dlia prohnozuvannia tendentsii rynku nerukhomosti [Neural networks and their application for real estate market trends forecasting]. Scientific Bulletin of UNFU, 22(5), 324–330.
20. Kohonen, T. (2013). Essentials of the self-organizing map. Neural Networks, 37, 52–65. <https://doi.org/10.1016/j.neunet.2012.09.018>

21. Kruglov, V. V., & Borisov, V. V. (2001). *Iskusstvennye neironnye seti: teoriia i praktika*. Moscow: Goriachaia liniia-Telekom. 382 p. [in Russian].
22. Liubun, Z. M. (2006). *Osnovy teorii neiromerezh*. Lviv: Vyd. tsentr LNU im. Ivana Franka. 140 p. [in Ukrainian].