Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)
Кафедра Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Дослідження методів штучного інтелекту для формування перегонової стратегії в умовах стохастичної невизначеності
(тема)

Виконав:
здобувач другого року навчання,
групи СШІм-24-2
Ірина ГІТІС
(власне ім'я, прізвище)

Спеціальність 122 Комп'ютерні науки
(код і повна назва спеціальності)
Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)
Освітня програма Системи штучного інтелекту
(повна назва освітньої програми)
Керівник проф. Валентин ФІЛАТОВ
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ _____ Лариса ЧАЛА
(підпис) (власне ім'я, прізвище)

2026 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук

Кафедра Штучного інтелекту

Рівень вищої освіти другий (магістерський)

Спеціальність 122 Комп'ютерні науки
(код і повна назва)

Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма Системи штучного інтелекту
(повна назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
«_____» _____ 20 ___ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Гітіс Ірині Веніамінівні
(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження методів штучного інтелекту для формування перегонової стратегії в умовах стохастичної невизначеності

затверджена наказом університету від 6 квітня 20 26 р. № 269Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії 20 травня 20 26 р.

3. Вихідні дані до роботи Науково-технічні публікації щодо досліджень результатів спортивних змагань, рекурентних нейронних мереж, методів навчання з підкріпленням, алгоритмів пошуку, дані Інтернет-джерел особливостей перебігу та організації автомобільних перегонів серії «Формули-1», офіційні документи FIA з приводу регулювання «Формули-1», інші Інтернет-джерела, дані офіційного сайту компанії Pirelli щодо характеристик трас, офіційні телеметричні дані «Формули-1»

4. Перелік питань, що потрібно опрацювати в роботі
1) Аналіз предметної галузі
2) Теоретичні дослідження
3) Опис експериментальних досліджень та програмної реалізації

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / термін виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	06.04.2026	виконано
2	Аналіз особливостей та проблем предметної	09.04.2026	виконано
	галузі. Огляд досліджень спортивних змагань		
3	Постановка задачі	10.04.2026	виконано
4	Дослідження особливостей архітектур рекурент-	15.04.2026	виконано
	них нейронних мереж		
5	Дослідження алгоритмів прийняття рішень	19.04.2026	виконано
6	Формування вихідного простору ознак. Збір та	24.04.2026	виконано
	аналіз вихідних даних		
7	Побудова моделей штучних нейронних мереж.	01.05.2026	виконано
	Проведення експериментальних досліджень		
8	Програмна реалізація алгоритмів прийняття	08.05.2026	виконано
	рішень. Експериментальні дослідження		
9	Написання пояснювальної записки	16.05.2026	виконано
10	Нормоконтроль	17.05.2026	виконано
11	Попередній захист	18.05.2026	виконано
12	Рецензування	19.05.2026	виконано
13	Захист перед ЕК	20.05.2026	

Дата видачі завдання 6 квітня 2026 р.

Здобувач _____
(підпис)

Керівник роботи _____ проф. Валентин ФІЛАТОВ
(підпис) (посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка: 102 с., 14 рис., 4 табл., 3 дод., 44 джерела.

АВТОМОБІЛЬНІ ПЕРЕГОНИ, ДЕРЕВО ПОШУКУ МОНТЕ-КАРЛО, НАВЧАННЯ З ПІДКРІПЛЕННЯМ, ПРОКСИМАЛЬНА ОПТИМІЗАЦІЯ СТРАТЕГІЇ, РЕКУРЕНТНІ НЕЙРОННІ МЕРЕЖІ, СТРАТЕГІЯ, «ФОРМУЛА-1».

Об’єкт дослідження – процес формування перегонової стратегії.

Предмет дослідження – методи штучного інтелекту.

Мета роботи – підвищення ефективності перегонової стратегії в умовах стохастичної невизначеності за рахунок дослідження та застосування методів штучного інтелекту.

Методи дослідження – математичного моделювання, штучного інтелекту (рекурентні мережі, навчання з підкріпленням), прийняття рішень та оптимізації, статистичного та емпіричного аналізу.

Наукова новизна полягає в удосконаленні процесу підтримки прийняття рішень у стохастичних середовищах завдяки застосуванню алгоритму проксимальної оптимізації стратегії. Набув подальшого розвитку підхід до узагальненого нейромережевого моделювання за рахунок інтеграції попередньо навченої прогностичної моделі безпосередньо в симуляційну архітектуру.

У результаті виконання роботи було проведено комплексне дослідження методів нейромережевого моделювання та евристичного пошуку для задач підтримки прийняття рішень в умовах стохастичної невизначеності. У межах дослідження було здійснено два незалежних етапи: праведно порівняльний аналіз архітектур рекурентних нейронних мереж для задачі прогнозування деградації шин та досліджено алгоритми послідовного прийняття рішень для формування стратегії перегонів.

ABSTRACT

Master’s thesis contains: 102 pp., 14 fig., 4 tabl., 3 ann., 44 references.

AUTO RACING, FORMULA 1, MONTE CARLO TREE SEARCH, PROXIMAL POLICY OPTIMIZATION, RECURRENT NEURAL NETWORKS, REINFORCEMENT LEARNING, STRATEGY.

Research object is the process of developing a racing strategy.

Research subject is artificial intelligence methods.

The aim of the work is to improve the effectiveness of race strategy under conditions of stochastic uncertainty through the study and application of artificial intelligence methods.

Research methods: mathematical modelling, artificial intelligence (recurrent networks, reinforcement learning), decision-making and optimisation, statistical and empirical analysis.

The scientific novelty lies in the improvement of the decision-making support process in stochastic environments through the application of a proximal strategy optimisation algorithm. The approach to generalised neural network modelling has been further developed by integrating a pre-trained predictive model directly into the simulation architecture.

As a result of the work, a comprehensive study was conducted on neural network modelling methods and heuristic search for decision-making support tasks under stochastic uncertainty. The study comprised two independent stages: a comparative analysis of recurrent neural network architectures for the task of tyre degradation prediction, and an investigation of sequential decision-making algorithms for race strategy formation.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів 7
Вступ 8
1 Аналіз предметної галузі 10
1.1 Дослідження специфіки нормативного регулювання автомобільних перегонів 10
1.2 Аналіз особливостей поведінки гоночних шин 14
1.3 Дослідження стохастичних факторів та їхнього впливу на стратегію.18
1.4 Огляд існуючих досліджень автоспортивних змагань 23
1.5 Постановка задач дослідження 27
2 Теоретичні дослідження 28
2.1 Дослідження особливостей архітектур рекурентних нейронних мереж 28
2.2 Аналіз алгоритмів послідовного прийняття рішень 39
3 Опис експериментальних досліджень та програмної реалізації 50
3.1 Формування та аналіз вихідного простору ознак 50
3.2 Математична формалізація процесу деградації гоночних шин 56
3.3 Порівняльне тестування нейромережевих моделей прогнозування.. 59
3.4 Оцінка ефективності алгоритмів послідовного прийняття рішень ... 68
Висновки 78
Перелік джерел посилання 80
Додаток А Коефіцієнти впливу маси пального на час 86
Додаток Б Вихідний код програми 87
Додаток В Відомість кваліфікаційної роботи 102

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

РНМ – рекурентні нейронні мережі;

ШІ – штучний інтелект;

ШНМ – штучні нейронні мережі;

GRU – Gated Recurrent Unit – вентильний рекурентний вузол;

LSTM – Long Short-Term Memory – довга короткочасна пам’ять;

PPO – Proximal Policy Optimization – проксимальна оптимізація стратегії;

MCTS – Monte Carlo Tree Search – дерево пошуку Монте-Карло;

TRPO – Trust Region Policy Optimization – оптимізація стратегії довірчої області.

ВСТУП

На ранньому етапі розвитку автомобільного спорту змагання проводилися виключно серед гонщиків. Їхньою метою було визначити найбільш професійного, швидкого та рішучого пілота, здатного першим перетнути фінішну лінію. Саме до думки гонщиків прислухалися, адже вони були єдині, хто мав реальний досвід та розуміння того, що відбувалося безпосередньо на трасі, та єдині, хто приймав рішення.

З плином часу все змінилося: фокус змагань частково змістився з пілотів на команди та конструкторів. Проектування та будування гоночних автомобілів – трудомісткий процес, який вимагає глибокого розуміння широкого спектру фундаментальних дисциплін. Тепер переможцем ставав не тільки той, чий пілот був найбільш швидким і рішучім, а той, хто зміг створити болід, який зміг перевершити інші за продуктивністю, стабільністю та маневреністю.

Однак розвиток технологій та доступність ресурсів зрівняли ці аспекти. У сучасному автоспорті, навіть при наявності найбільш професійного гонщика та найшвидшої машини, перемога команди не є гарантованою, якщо вона приймає неправильні рішення.

У гонках прийняття рішень лежить в основі формування стратегії, що полягає у визначенні оптимальних технічних параметрів автомобілю. Однак цей процес ускладнюється необхідністю аналізу великого об’єму інформації та виникненням подій, таких як аварії та зміни погодних умов, які можуть радикально змінити хід гонки. У таких умовах вкрай актуальним є впровадження механізму, що дозволить підвищити ефективність прийняття рішень. Для цього можливе застосування методів штучного інтелекту.

На основі теми кваліфікаційної роботи та викладеної актуальності можна сформулювати об’єкт, предмет та мету дослідження.

Об’єкт дослідження: процес формування перегонової стратегії.

Предмет дослідження: методи штучного інтелекту.

Мета: підвищення ефективності перегонової стратегії в умовах стохастичної невизначеності за рахунок дослідження та застосування методів штучного інтелекту.

Для досягнення поставленої мети виконати наступні задачі:

– провести дослідження особливостей автомобільних перегонів, дослідити фактори, що мають вплив на їх перебіг; проаналізувати шляхи використання методів машинного навчання для вивчення автомобільного спорту в наукових публікаціях;

– виконати теоретичний аналіз предмету дослідження, визначити особливості реалізації та застосування методів;

– провести експериментальні дослідження методів, що вивчаються, виконати порівняльний аналіз та зробити висновки щодо їхньої ефективності.

Наукова новизна полягає в удосконаленні процесу підтримки прийняття рішень у стохастичних середовищах завдяки застосуванню алгоритму проксимальної оптимізації стратегії. Набув подальшого розвитку підхід до узагальненого нейромережевого моделювання за рахунок інтеграції попередньо навченої прогностичної моделі безпосередньо в симуляційну архітектуру.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

1.1 Дослідження специфіки нормативного регулювання автомобільних перегонів

Найвищим класом міжнародних перегонів для одномісних гоночних автомобілів є «Чемпіонат світу FIA “Формула-1”», санкціонована некомерційною асоціацією Fédération Internationale de l’Automobile (FIA), що керує правилами та положеннями для всього міжнародного чотириколісного автоспорту.

Кожного року FIA спільно з командами та комерційними структурами «Формули-1» розробляє та формує набір документів, які регулюють різні аспекти чемпіонату: технічні характеристики перегонових автомобілів, регламент проведення змагань, обмеження фінансових ресурсів, тестування двигунів та інших елементів машини.

Фінансовий регламент [1] обмежує бюджети команд та спрямований на забезпечення стійкості та зменшення розриву між учасниками. Він регулює допустимі витрати та визначає механізми контролю за їх дотриманням.

Спортивний регламент [2] регулює безпосередньо процес проведення гоночного уїк-енду та включає широкий спектр правил, що визначають як організаційні, так і змагальні аспекти. Він встановлює структуру уїк-енду, яка включає вільні заїзди, кваліфікацію та гонку, а також проведення короткого формату гонки – спринту. У регламенті детально описані процедури кваліфікації (розділ на сегменти, час, відведений на кожний сегмент, правила вибуття учасників), порядок формування стартової решітки та умови допуску автомобілів до старту.

Значна увага приділяється процедурам старту та рестарту гонки. Регламент визначає вимоги до розташування болідів перед стартом, дії пілотів на колі розігріву, правила у випадку відміни або відкладення старту.

Також регламентуються дії після введення режиму «автомобіля безпеки», включаючи обмеження на обгони та вимоги до підтримки певного темпу.

Окремий блок регламенту стосується використання шин та проведення піт-стопів. У сухих умовах команди мають використати два різних типи шин. Правила стосуються роботи на піт-лейні: обмеження швидкості проїзду автомобілю по піт-лейну, вимоги до безпеки та порядок виконання піт-стопів.

Також спортивний регламент містить положення, що стосуються поведінки гонщиків на трасі, включаючи правила обгонів, захисту позиції та дотримання границь траси. За порушення цих правил передбачена система штрафів, таких як додавання часу, проїзд по піт-лейну або штрафні позиції на старті наступної гонки. Контроль за дотриманням правил та винесення рішень у разі їх порушення здійснює спеціальна комісія, яка назначається від FIA на кожний Гран-прі.

Окрім цього, регламент визначає систему нарахування очок, умови класифікації в чемпіонаті та вимоги до фінішу. Також регулює радіопереговори та обмін інформацією між пілотами та командою.

Технічний регламент [3] визначає конструкцію автомобілів, включаючи вимоги до силової установки, аеродинаміки, маси, паливної системи та інших компонентів. Його основною задачею є обмеження надмірної технічної переваги команд з більшими фінансовими можливостями над іншими, забезпечення безпеки гонщиків та сприяння підвищенню видовищності перегонів. Особливістю технічного регламенту є його періодичний перегляд: як правило, значні зміни вносяться в рамках «регламентних циклів», які можуть тривати декілька років. Так, реформи, напрямлені на зміни в аеродинамічних елементів або силових установок, не вводяться щорічно, а раз на декілька років (приблизно 4 роки), що дозволяє командам адаптуватися та ефективно розподіляти свої ресурси.

Однак між крупними змінами можливе внесення точкових коригувань, пов’язаних з уточненням формулювань правил, усуненням

неоднозначностей та лазійок, які команди можуть використовувати для отримання переваги. Такі зміни переважно приймаються до початку сезону, проте можуть вноситися під час перебігу чемпіонату, якщо це пов’язано з питаннями безпеки або явним порушенням регламенту. Наприклад, за ліквідацію лазійки нещодавно проголосували виробники силових установок, відповідні зміни будуть внесені в текст технічного регламенту та набудуть чинності 1 червня 2026 року [4].

Важливим елементом сучасної «Формули-1» є так звана «заморожування» технічного регламенту. Під цим розуміється обмеження або повна заборона на доопрацювання певних компонентів автомобілю протягом визначеного періоду часу. Найбільш відомим прикладом є заморожування силових установок у період з 2022 по 2025 роки, коли робота над двигунами дозволялася лише в обмежених рамках.

Окрім цього, в рамках технічного регламенту діє система омологації компонентів [3]. Це означає, що певні елементи конструкції затверджуються на певний період, а їхня зміна можлива лише при дотриманні строгих процедур погодження з FIA. Така практика особливо важлива для складних та дорогих елементів, розробка яких потребує значних ресурсів.

Також технічний регламент регулює процедури збирання даних, встановлення датчиків та використання телеметрії.

Сучасні машини «Формули-1» оснащені великою кількістю датчиків, які фіксують параметри роботи різних систем: температуру та тиск в шинах, швидкість, прискорення, параметри силової установки (тиск в циліндрах, інтенсивність детонації в камері згоряння, швидкість обертання валу двигуна, температура та тиск різних рідин двигуна та ін.), положення аеродинамічних елементів (переднього та заднього антикрил) [3]. Ці дані передаються в режимі реального часу на командний місток за допомогою телеметрії та використовуються для аналізу стану автомобілю, оцінки стратегії та прийняття оперативних рішень під час гонки.

Однак регламент накладає низку обмежень. Зокрема, дозволена лише одностороння телеметрія – передача даних від автомобіля до команди [3]. Передача керуючих сигналів у зворотному напрямку, яка дозволяє змінювати параметри роботи машини дистанційно, заборонена. Це означає, що будь-які зміни мають або виконуватися самим пілотом (команда може давати інструкції по радіозв'язку), або бути запрограмовані заздалегідь. Таке обмеження напрямлене на максимальному збереженні автентичності перегонів – єдиною особою, що керує машиною, розуміє її поведінку і, внаслідок чого, приймає рішення безпосередньо на трасі, має залишатися пілот.

Окрім цього, правила обмежують призначення та кількість датчиків. Команди не можуть встановлювати довільне число сенсорів (наприклад, по декілька на окремий елемент, якщо це окремо не прописано в регламенті) або використовувати пристрої, які напряму вимірюють параметри, що вважаються критично важливими з точки зору конкурентної переваги або складності контролю. Наприклад, на машинах «Формули-1» відсутні датчики, які напряму вимірюють ступінь зносу шин, а оцінка їхнього стану виконується опосередковано, на основі інших факторів.

Також важливу роль грає стандартизація електроніки. Командам «Формули-1» забороняється розробка власних інструментів керування, натомість вони мають використовувати уніфікований електронний блок управління (ECU) [3]. Цей пристрій контролює збір та передачу даних, а також забезпечує централізований контроль над системами машини.

Додатково регламент обмежує програмне забезпечення та алгоритми, які використовуються для обробки даних. Заборонено системи, які можуть автоматично компенсувати дії гонщика або керувати автомобілем без його участі.

Таким чином, регламент «Формули-1» формує суворі обмеження як на конструкцію машин, так і на процеси збору та передачу даних. Вимоги технічного регламенту FIA обмежують використання телеметрії та датчиків,

що виключає можливість прямого внесення змін до низки характеристик, у тому числі стану шин. Через це, аналіз таких параметрів вимагає застосування опосередкованих методів та моделей, які ґрунтуються на обробці наявних телеметричних даних. Це, в свою чергу, обумовлює актуальність використання математичних методів для вилучення прихованих залежностей та прогнозування характеристик, недоступних для прямого втручання.

1.2 Аналіз особливостей поведінки гоночних шин

Головною особливістю гоночної шини – це те, що вона не є жорстким тілом. Основний матеріал протектору – гума, яка являє собою в’язкопружний еластомер, що поєднує в собі як пружні, так і в’язкі властивості. Це впливає на те, що її механічний відгук залежить не лише від величини прикладеного навантаження, а й від швидкості деформації та температури. Зокрема, зміни температурного режиму можуть суттєво впливати на пружність та коефіцієнт тертя гуми, що призводить до змін механічних властивостей шин [5] – це напряму визначає поведінку машини на трасі.

Одним з ключових механізмів, що визначає роботу гоночної шини, є гістерезис. При коченні шина постійно зазнає циклічних деформацій під дією ваги автомобілю та притискної сили. Через в’язкопружність матеріалу, з якого виготовлено шину, деформація та відновлення форми відбуваються з фазовим запізненням відносно прикладеної напруги. Це призводить до того, що частина енергії не повертається назад, а розсіюється всередині матеріалу у вигляді тепла [6]. Така властивість з одного боку дозволяє забезпечити належне зчеплення шини з асфальтованим покриттям, а з іншого боку призвести до значного погіршення її стану.

Оптимальне зчеплення гоночних шин досягається при нагріві гумових сумішей, з яких їх виготовляють, до температури, що знаходиться у

визначеному робочому діапазоні. Для різних складів гуми цей діапазон відрізняється: для найжорсткішого він дорівнює 110–140°C, а для найм'якішого – 85–115°C [7]. Температура шин поза межами цих інтервалів призводить до суттєвого зниження ефективності шини: при занизькій температурі гума стає більш жорсткою, що значно впливає на керованість автомобіля, а при зависоких – стає занадто м'якою, через що вона швидше зношується.

Зносом називають поступове погіршення стану протектора шини у ході її взаємодії з дорожнім покриттям. Під час руху контакт між шиною та асфальтом супроводжується тертям, при якому відбувається поступове відшаровування частинок гуми з поверхні. Накопичення таких часток гуми на трасі призводить до так званого «прогумовування», що надає додаткове зчеплення шин з асфальтом.

На відміну від зносу, деградація являє собою процес погіршення фізичних та хімічних властивостей гуми без обов'язкового відшаровування значної кількості матеріалу.

Довготривала робота шини поза межами робочого температурного діапазону призводить до накопичення незворотних змін в структурі матеріалу, з якого виготовлено шину. Використання комплекту шин навіть на декілька перегонових кіл призводить до значних змін: доволі часто комплект, що вже було використано для однієї швидкої спроби у ході сегменту класифікації, під час другої швидкої спроби показує значне падіння темпу.

Темп зносу та деградації шин визначає низка як зовнішніх, так і внутрішніх факторів.

Як вже було описано раніше, робота шин за температур, що знаходяться поза межами робочих діапазонів, а також, загалом, високе навантаження на шуму спричиняють процеси, які призводять до деградації. Проте важливо зазначити, що ті гумові суміші, які використовуються для виробництва різних типів шин (м'якого, середнього та жорсткого) для

автомобілів «Формули-1» мають власні характеристики щодо зчеплення, терміну служби, а також, як було зазначено вище, робочих температурних інтервалів. Більш м’які суміші забезпечують високе зчеплення, але швидше зношуються, тоді як більш тверді – довговічність шин, проте демонструють зниження зчеплення [8].

Температура навколишнього середовища та асфальтного покриття також відіграє свою роль у підвищенні темпів зносу та деградації. М’які гумові суміші для шин краще працюють у прохолодних умовах, тоді як жорсткі підходять для високих температур [8].

Немаловажними та відносно складними для подолання факторами є стиль водіння та помилки пілота. Агресивне водіння, зокрема різке гальмування, швидке прискорення, вхід у поворот на надто високій швидкості, блокування колес, виїзд поза межі траси в гравійні пастки можуть прискорити знос та деградацію шин [8].

Також на знос та деградацію мають вплив умови на трасі. Кожна траса має відмінні якості, такі як тип покриття, рівень шорсткості та кількість прямих та поворотів, їх тип. Більш шорсткі поверхні призводить до швидшої деградації, більш гладкі можуть потенційно подовжити термін служби шин [8].

Умови на трасі є одним із найважливіших факторів того, наскільки швидко зношуються шини та наступає їх деградація. Офіційна компанія-постачальник шин для команд «Формули-1» Pirelli публікує низку характеристик, що визначають ці умови. Кожна властивість оцінюється за силою впливу по 5-бальній шкалі, де 1 відноситься до незначного, а 5 – до надмірного.

Однією з найбільш важливих характеристик для стану шин є абразивність асфальту. Вона вказує на те, наскільки абразивною є асфальтова поверхня, рівень її шорсткості: високий рівень абразивності покриття може призвести до швидшої деградації шин та підвищенню їх температури [9].

Іншим немаловажним фактором є адгезія асфальту, яка відображає рівень зчеплення шин із дорожнім покриттям, а також вплив зовнішніх факторів, які можуть погіршити стан асфальту і підвищити рівень адгезії [9].

Також для роботи шин немаловажним є оцінка рівня зчеплення. Вона вказує на те, наскільки вимогливим є конфігурація траси з точки зору фаз зчеплення, тобто коли на шину має діяти велика рушійна сила, особливо під час виходу з повільних поворотів [9].

Наступною характеристикою, яку оцінює Pirelli, є складність гальмування. Вона визначає, наскільки вибагливою є траса з точки зору гальмування [9].

Оцінка бічних сил відображає те, наскільки траса вимоглива до бічних сил, які впливають на автомобіль під час поворотів. Задача шин полягає в тому, щоб утримувати автомобіль на трасі та протистояти силам, що штовхають його назовні на високих швидкість [9].

Іншою характеристикою є навантаження на шини – ступінь навантаження та швидкості, що зазнають шини на трасі [9]. Наявність великої кількості поворотів, які йдуть один за одним, може підвищити вплив цього фактору на шини.

Наступна характеристика – притискна сила, вказує на рівень впливу аеродинамічних сил, що штовхають автомобіль вниз [9].

Останньою характеристикою, що виокремлює Pirelli, є еволюція асфальту. Вона відображає, наскільки умови траси можуть змінитися протягом гоночного уїк-енду через кількість гуми, що залишилася на трасі. Даний аспект безпосередньо пов’язаний з рівнем абразивності асфальту [9].

Шинам, що використовуються у «Формулі-1», притаманні характерні типи зносу та деградації [10]:

– стирання – зазвичай причиною є надто агресивний стиль водіння або надмірне навантаження шин, температура яких на той момент знаходилася поза межами робочого діапазону. Умовою для утворення таких слідів є ковзання в одному напрямку, внаслідок чого утворюється

характерний малюнок – низка майже паралельних гребнів, розташованих під прямим кутом до напрямку стирання;

– гранулювання – причиною може бути незбалансованість налаштувань автомобілю або недостатньо розігріті шини. Такий тип зносу характеризується появою зернистої структури на поверхні протектору;

– утворення пухирів – являє собою наслідок процесу вулканізації, що є основою деградації та полягає в розриві зв'язків між молекулами матеріалу, з якого виготовлено шини. Причинами можуть бути: занадто високий тиск у шині в холодному стані, надмірне підвищення тиску під час використання в поєднанні з надзвичайно високою температурою траси. Даний патерн зносу характеризується наявністю глибоких поглиблень, в наслідку розриву пухирів, у протекторі шини;

– розшаровування – як правило, причина появи є висока абразивність асфальтного покриття. Характеризується віддиранням смуг з верхнього шару шин.

Таким чином, поведінка гоночної шини визначається складною взаємодією фізичних та хімічних процесів всередині неї. Явище гістерезису відіграє ключову роль в формування температурного режиму, який визначає ефективність роботи шин в певному робочому діапазоні. Вихід із цього діапазону може призвести до зниження зчеплення та прискореної деградації. Додаткову складність вносить вплив зовнішніх факторів, таких як температура траси, шорсткість асфальтного покриття, стиль водіння та інші, які визначають як інтенсивність зносу, так і характер деградації. У результаті формуються різноманітні патерни пошкоджень протектора.

1.3 Дослідження стохастичних факторів та їхнього впливу на стратегію

Як зазначалося в п. 1.1, спортивний регламент зобов'язує команди використовувати два різних типи комплектів шин під час гонки. Це, а також

падіння темпу внаслідок зносу та деградації шин, вимагає від команд планування зупинок, піт-стопів, для заміни старого комплекту шин. Визначення правильного кола, на якому потрібно провести піт-стоп, та необхідного типу шин, якій дозволить максимально ефективно проїхати кожний відрізок гонки, уособлює собою стратегію.

При плануванні базової стратегії команда зазвичай спирається на детерміновані очікування щодо перебігу гонки: очікуваний темп деградації шин, прогнозовані витрати палива та стабільні погодні умови.

Однак реальна гонка являє собою стохастичний процес, в якому суттєву роль відіграють випадкові та важко передбачувані події, обумовлені як зовнішніми факторами (наприклад, зміна погодних умов), так і внутрішніми (наприклад, помилки пілота або команди механіків, що проводять піт-стоп).

Також важливо відрізняти два типи змін, які впливають на перебіг гонки. З одного боку, існують неперервні процеси, такі як зменшення маси машини внаслідок витрати палива або поступова деградація шин. Ці зміни можуть бути відносно точно прогнозовані, проте, можливі й винятки: темпи деградації шин значно вищі за прогнозовані, що повністю змінює визначений заздалегідь алгоритм дій. З іншого боку, існують дискретні випадкові події: аварії або раптові зміни погодних умов, момент настання яких не піддається точному прогнозуванню.

Виїзд машини безпеки (SC) на трасу та увімкнення режиму «віртуальної машини безпеки» (VSC) є одними із найбільш значущих стохастичних факторів, які можуть кардинально змінити базову стратегію гонки. Поява SC або VSC носить дискретний та важко прогнозований характер, адже це обумовлено аваріями, технічними несправностями автомобілів або станом траси.

Ключовим аспектом є відносна втрата часу на піт-стопі. При стандартних гоночних умовах виконання піт-стопа призводить до втрати приблизно 19–25 секунд [11], [12], [13] порівняно з суперниками, які в цей

час продовжують їхати в бойовому темпі. Проте при введені режимів SC або VSC швидкість всього пелотона значно зменшується, а у випадку VSC додатково вводиться вимога дотримання дельти часу [2]. Це призводить до зменшенню різниці в швидкості між болідом, що знаходиться на трасі, та машиною, яка здійснює піт-стоп, внаслідок чого відносна втрата часу скорочується приблизно до 10–15 секунд [14]. Непроведений піт-стопу при дії даних режимів вважається стратегічною помилкою, адже значне скорочення втрат часу дозволяє зберегти позицію гонщику відносно суперників.

Окрім часового фактору, SC та VSC істотно впливає на температуру шин. Як було зазначено в п. 1.2, оптимальне зчеплення досягається лише в межах робочих температурних діапазонів. В умовах руху за машиною безпеки, швидкість та навантаження на шини значно знижується, що призводить до зниження тепловиділення за рахунок гістерезису і, як наслідок, до швидкого падіння температури шин. І хоча гонщики намагаються додатково розігрівати шини під час SC та VSC (шляхом виляння на прямих), це все одного може призвести до виходу з робочого діапазону. Після рестарту гонки це може впливати на стійкість та керованість автомобілю, а також, у подальшому, схильність до гранулювання шин.

Зміна погодних умов є одним з найбільш складних стохастичних факторів в автоспорті. На відміну від інших аспектів, погодні умови можуть змінюватися як в часі, так і в просторі, неоднорідно впливаючи на різні ділянки траси. Усі команди мають метеорадари, які в режимі реального часу показують короткостроковий прогноз погоди, але це не завжди дозволяє точно спрогнозувати зміни погоди.

Ключовим поняттям в умовах мінливої погоди є так званий «кросс-поінт» – момент під час дощу, коли використання сліків (шин для сухої траси, без рисунку протектору) стає менш ефективним, ніж з проміжними – спеціальними шинами для мокрих умов, які мають рисунок на протекторі,

що забезпечує краще зчеплення, проте знижує швидкість. У цих випадках проводять незапланований піт-стоп для переходу на інший комплект.

Однак існує складність у визначенні «крос-поінту», яка обумовлена неоднорідність погодних умов: на протяжних трасах дощ може зачіпати лише окремі сектори (або невеликі сегменти), що призводить до існування і мокрих, і сухих ділянок траси. У таких умовах, зазвичай, рішення залишається за пілотом, який може оцінити як керованість автомобілю, так і власні можливості.

У результаті вибір моменту переходу на інший тип шин (і взагалі вирішення того, чи потрібен він) стає складною задачею в умовах невизначеності, адже команди мають враховувати не лише поточний стан траси, а й ймовірні прогнози подальшої зміни погодної ситуації.

У деяких випадках, коли наслідки аварії або погодні умови можуть ставити під загрозу життя гонщиків, дирекцією гонки може прийнятися рішення про зупинку гонки завдяки червоних прапорів – це є ще одним значним стохастичним фактором. На відміну від режимів SC та VSC, при яких рух автомобілів сповільнюється, червоний прапор повністю перериває гонку.

Відповідно до спортивного регламенту, під час дії червоних прапорів команди мають можливість обслуговувати машину без втрати гоночного часу. це включає заміну шин, усунення деяких типів ушкоджень, а також зміна деяких налаштувань, наприклад, заміна переднього антикрила [2].

Наслідком цього є отримання «безплатного піт-стопу», адже всі машини зупинені, а втрати часу, які пов’язані із заміною шин, зникають. Це призводить до нівелювання стратегічної переваги тих учасників, які вже зробили зупинку під час гонки. До того ж команди, що змінили комплект шин на інший під час червоних прапорів, виконують вимогу щодо використання різних комплектів за гонку і можуть більше не планувати піт-стопи.

Іншими важливими стохастичними факторами, що суттєво впливають на реалізацію гоночної стратегії, є трафік та безпосередня боротьба на трасі. Незважаючи на те, що загальне положення суперників можна спрогнозувати, їхня поведінка (темп, помилки, оборонні дії) залишаються ймовірнісними. Оборона, різниця по шинах або просто швидша машина суперника можуть стати значною перепоною для гонщика при спробах здійснити обгін на трасі.

Ключовим аспектом є ефект «брудного повітря», який виникає під час руху позаду іншого автомобіля. «Брудне повітря» – це турбулентний повітряний потік, який залишає позаду себе машина, яка їде попереду, що призводить до погіршення аеродинамічної ефективності. Зокрема, знижується притискна сила, що генерується аеродинамічними поверхнями, в особливості переднім антикрилом [15].

Зниження притискної сили призводить до недостатньої та надмірної керованості, через що передня та задня частини машини починають поводитися нестабільно в поворотах. У свою чергу, це призводить до збільшення ковзання шин, що викликає ріст температур та передчасну деградацію [15].

Через негативні наслідки «брудного повітря» на аеродинамічні елементи боліда та, у свою чергу, на його керованість, гонщикам, що потрапили в цей потік, достатньо складно отримати перевагу над машиною попереду безпосередньо на трасі. У таких випадках може бути прийнято одне зі стратегічних рішень: проведення «андеркату» (undercut) або «оверкату» (overcut).

«Андеркатом» називається обгін суперника за рахунок проведення більш раннього піт-стопу: використання нових шин дозволяє скоротити загальний час кола та проїхати швидше за суперника, який все ще їде на старому комплекті. «Оверкат» – це обгін суперника за рахунок пізнішого піт-стопу: старий комплект шин дозволяє проїхати перегонове коло за менший час, ніж суперник зі свіжими шинами, які ще не увійшли в робочий діапазон.

На більшості трас «андеркат» є основним стратегічним механізмом обгону [16].

Прийняття рішення щодо потреби в проведенні «андеркату» або «оверкату» може бути сплановано командою в рамках однієї з варіацій базового плану стратегії, проте, точно визначити необхідність цього та правильний час дуже складно. Крім того, необхідно розраховувати момент виїзду з піт-стопу. Неправильно визначений момент може призвести до того, що фактичні умови, в яких опиниться болід після піт-стопу, можуть суттєво відрізнятися від очікуваних, особливо при виїзді в щільний трафік, де пілоту буде потрібно втрачати ресурс нового комплекту шин на зайву боротьбу з тими гонщиками, які не є його прямими суперниками.

Таким чином, стохастичні фактори відіграють визначальну роль у формуванні фінальної стратегії гонки. Виїзд машини безпеки, зміни погодних умов, червоні прапори, трафік та інші важко передбачувані фактори здатні в будь-який момент змінити оптимальність базової стратегії. Окрім цього, кожний з цих факторів впливає на різні аспекти гоночного процесу. Через це, з метою забезпечення ефективності стратегічного планування необхідно спиратися не лише на середньо очікувані події, а й на аналіз ймовірнісних розподілів можливих сценаріїв.

1.4 Огляд існуючих досліджень автоспортивних змагань

На ранніх етапах розвитку автоспорту планування перегонової стратегії здебільшого ґрунтувалося на детермінованих та спрощених моделях, в яких не було закладено форс-мажорні обставини, припускався стабільний темп машини та лінійний характер зносу та деградації. Такі моделі спиралися на досвід інженерів та гонщиків, які робили припущення та приймали рішення, виходячи із суб’єктивних відчуттів у зв’язку з відсутністю точних замірів.

Із розвитком технологій ситуація суттєво змінилася. Сучасні перегонові автомобілі оснащені великою кількістю датчиків, що забезпечують високочастотну телеметрію: інженери можуть в режимі реального часу відслідковувати стан різних компонентів машини, починаючи від його швидкості, закінчуючи тиском в циліндрах двигуна. Аналіз такої кількості даних є дуже ресурсозатратним та майже неможливим для людини. Для спрощення цього процесу використовуються методи інтелектуального аналізу даних та машинного навчання.

Окрім цього, розвиток технологій та можливість збору великих масивів даних дозволили перейти до ймовірнісним та стохастичним методам моделювання. Зокрема, дуже популярним підходом в «Формулі-1» є використання методів Монте-Карло: до них вдаються як дійсні команди [17], так і дослідники в науковому середовищі [18], [19].

У роботі [19] реалізується інтеграція методу Монте-Карло в поколову симуляцію гонки для оцінки не тільки оптимальної, а й робастної стратегії. На відміну від детермінованого підходу, в якому обирається стратегія з найшвидшим часом, ймовірнісна модель дозволяє оцінити розподіл можливих результатів та обирати рішення, що є стійкими до несприятливих сценаріїв.

Для моделювання зміни часу перегонового кола та варіативності піт-стопів до базового часу додаються випадкові відхилення, які моделюються за допомогою різних розподілів ймовірності. Додатково застосовується баєсове висновування для врахування індивідуальних характеристик гонщиків та ймовірність виникнення рідкісних подій під час гонки.

Для коректного моделювання умов режиму «машини безпеки» в дослідженні вводиться концепція Safety Car Ghost («привід машини безпеки»). Під час симуляцій цей віртуальний об’єкт вводиться перед кожним гонщиком та обмежує його швидкість, імітуючи ефект збору пелотона.

Результати серії з 10 тис. симуляцій показали, що врахування стохастичних факторів суттєво змінює оцінку ефективності стратегії. Дослідники визначили: стратегії, які є оптимальними в детермінованих умовах, можуть демонструвати значний розкид результатів у ймовірнісній моделі.

Незважаючи на низку переваг, запропонований підхід має власні обмеження. Зокрема, базова стратегія в моделі задається заздалегідь та не змінюється в процесі симуляції, що обмежує можливість її використання для динамічного прийняття рішень. Визначення нової стратегії безпосередньо під час гонки також може бути неефективним через особливість самого методу Монте-Карло – вона полягає в моделювання безлічі сценаріїв одночасно. В роботі авторам, за рахунок розпаралелювання методу на всіх чотирьох ядрах процесору, вдалося провести 10 тис. симуляцій за 250–300 с. (4–5 хв.). Однак такий час на прийняття рішень є занадто великим в реаліях «Формули-1»: у середньому довжина кола не перевищує 1,5–2 хв. В той момент, коли буде розраховано оптимальну стратегію, вона вже буде неактуальною, тому що пройде декілька перегонових кіл і ситуація на трасі може кардинально змінитися.

Існують інші підходи до визначення стратегії гонки. У роботі [20] пропонується модель, в якій задача формулюється як бінарна класифікація: чи буде зроблено зупинку на даному колі, чи ні. Таким чином, замість пошуку правильного моменту для піт-стопу, алгоритм навчається відтворювати рішення команд на основі історичних даних.

У рамках дослідження було порівняно декілька алгоритмів машинного навчання: Random Forest, Support Vector Machine та Extreme Gradient Boosting – останній показав себе як найбільш ефективний при вирішенні поставленої задачі.

Аналіз важливості ознак у дослідженні виявив, що найбільш значущими є кількість кіл, які було проведено на тому чи іншому комплекті шин, та час перегонового кола. Додатково на результати прогнозування

впливають параметри, які пов’язані з особливістю керування автомобілю та умови на трасі.

Однак даний підхід також має низку обмежень. По-перше, модель навчається на історичних даних і, відтак, відтворює вже реалізовані колись стратегії, не гарантуючи їхню оптимальність. По-друге, вона не враховує довгострокові наслідки рішень, адже кожне коло розглядається окремо. По-третє, автори повністю прибрали форс-мажорні обставини із вихідних даних, що, у подальшому, значно обмежує можливості використання такої системи в реальних умовах.

Незважаючи на розвиток та різноманіття методів машинного навчання для аналізу та передбачення стратегій, такі підходи, здебільшого, обмежені рамками статистичного прогнозування. Для подолання цих проблем та створення систем, які здатні приймати рішення з урахуванням поточного стану гонки, застосовуються штучні нейронні мережі.

У роботі [21] пропонується підхід на основі використання двох різних нейромережевих моделей для побудови віртуального інженера по стратегії. Перша модель, являє собою гібрид повнозв’язної мережі та рекурентної, вирішує задачу бінарної класифікації: визначає, чи потрібно робити зупинку на поточному колі. Друга модель відповідає за прийняття рішення щодо піт-стопу та вибір типу шин. Система враховує можливі стохастичні фактори, а також вимоги регламенту.

Результати дослідження показали, що розбиття задачі на менші та делегування їх вирішення різним підсистемам, є ефективним в контексті прийняття рішень щодо стратегії перегонів. Такий підхід, у свою чергу, імітує реальну діяльність команди під час гонки.

Загалом, на основі аналізу наявних досліджень в автоспорті, можна зробити висновок про актуальність задачі визначення оптимальної стратегії перегонів. Огляд робіт показав успішність застосування методів математичного моделювання, машинного навчання, а також штучного інтелекту для вирішення задачі.

1.5 Постановка задач дослідження

Ґрунтуючись на проведеному аналізі предметної галузі, можна виділити наступні задачі, які необхідно вирішити у подальшому дослідженні:

– детально дослідити існуючі архітектури рекурентних нейронних мереж, особливості їх побудови, вивчити складнощі, які виникають під час їхнього навчання;

– розглянути методи підтримки прийняття рішень та підходи, на яких вони базуються. Вивчити особливості їхньої реалізації, а також виокремити переваги та недоліки;

– використовуючи визначені за результатами аналізу предметної галузі, фактори, що впливають на темпи деградації шин, сформувати простір ознак та провести статистичний аналіз вихідних даних. Також розглянути методи оцінки темпу деградації гоночних шин та математично формалізувати даний процес для подальшого використання методами ШІ;

– зважаючи на проведене дослідження нейронних мереж, побудувати та навчити моделі з різною архітектурою РНМ. Оцінити якість прогнозування кожної моделі та їхню робастність. Провести порівняльний аналіз та зробити висновки про ефективність досліджуваних архітектур;

– спираючись на дослідження стохастичних факторів перегонів та аналіз методів прийняття рішень, реалізувати симуляційне середовище та досліджувані алгоритми. Провести порівняльний аналіз як методів між собою, так і з історичними стратегіями переможців.

2 ТЕОРЕТИЧНІ ДОСЛІДЖЕННЯ

2.1 Дослідження особливостей архітектур рекурентних нейронних мереж

На поточний момент штучні нейронні мережі є найбільш популярною формою штучного інтелекту. Вони використовуються для вирішення широкого спектру задач у різних сферах як на промисловому, так і на побутовому рівні.

Базовою одиницею штучної нейронної мережі є штучний нейрон, який складається з вхідного оператора, оператора активації та вихідного оператора (зазвичай не виділяють окремо, мається на увазі результат, отриманий після оператора активації) [22].

Принцип роботи операторів штучного нейрону виглядає наступним чином: вхідний оператор перетворює зважені входи та подає їх на оператор активації. Вихідний сигнал нейрона являє собою перетворений вихідним оператором вихідний сигнал оператора активації [22].

Вхідний оператор нейрона задає вигляд використовуваного в нейроні перетворення зважених входів. Може бути використано функції суми, добутку, максимального або мінімального значень зважених входів [22].

Функція активації (оператор активації) описує провило переходу нейрона, що перебуває в момент часу k у стані $z(k)$, у новий стан $z(k+1)$ при знаходженні вхідних сигналів x [22].

Найпростішими функціями активації є лінійна функція та її різновиди, проте, вони мають низку недоліків, які обмежують можливості їхнього використання: лінійні активаційні функції застосовуються здебільшого у ШНМ, що припускають лінійну залежність між вхідними та вихідними змінними [22].

Іншим типом активаційних функції є сигмоїдальні функції, до яких відносяться наступні:

– експоненціальна сигмоїда [23]:

$$f(x) = \frac{1}{1 + e^{-c(x+b)}}, \tag{2.1}$$

де c – параметр, що визначає нахил функції активації;

b – значення нейронного зміщення;

– симетрована експоненціальна сигмоїда [23]:

$$f(x) = \frac{2}{1 + e^{-c(x+b)}} - 1; \tag{2.2}$$

– функція гіперболічного тангенсу [23]:

$$f(x) = \frac{e^{x+b} - e^{-(x+b)}}{e^{x+b} + e^{-(x+b)}}. \tag{2.3}$$

Їхньою перевагою є те, що вони є диференційованими на всій області визначення. Це є вкрай важливим, адже при побудові штучних мереж необхідно працювати як із самою активаційною функцією, так і з її першою похідною [22].

У загальному випадку штучна нейронна мережа являє собою сукупність штучних нейронів, згрупованих у шари та пов’язаних між собою певним зв’язком. Виокремлюють наступні типи зв’язків [23]: послідовні, перехресні, зворотні та латеральні.

Залежно від того, який тип зв’язків наявний у конкретній нейронній мережі, розрізняють їхні топології [22]: ШНМ без зворотних зв’язків (прямого поширення) та ШНМ зі зворотними зв’язками (зворотного поширення, рекурентні). Топологія мережі значно впливає на її властивості та тип задач, які вона здатна вирішити.

Штучні нейронні мережі прямого поширення є універсальними апроксиматорами [24], що робить їх придатними для вирішення широкого класу задач. Однак вони мають низку недоліків, які зумовлені їхньою топологію: через те, що ШНМ прямого поширення розглядають вхідні дані як незалежні екземпляри, вони не мають можливості оперувати інформацією про попередні стани системи. Це робить їх неефективними для задач, в яких поточний стан об'єкта критично залежить від попередньої історії, як, наприклад, задачі з часовими або контекстними залежностями.

Для вирішення таких задач можуть використовуватися рекурентні нейронні мережі. Вони долають проблему збереження попередньої інформації за допомогою особливості будови нейронів (блоків), з яких РНМ складаються, зокрема, завдяки наявності зворотних зв'язків, які дозволяють передавати інформацію з попередніх часових кроків на наступні. За рахунок цього формується прихований стан, що являє собою своєрідну пам'ять про стани системи.

Розрахування поточного прихованого стану відбувається шляхом подання результату конкатенації вектора вхідних даних на поточному часовому кроці та прихованого стану на попередньому на вхід повнозв'язного шару та застосуванням оператору активації (певної активаційної функції). Математично даний процес можна описати наступним чином [25]:

$$\mathrm{H}_t = \omega(\mathrm{X}_t \mathrm{W}_{\mathrm{xh}} + \mathrm{H}_{t-1} \mathrm{W}_{\mathrm{hh}} + \mathrm{b}_{\mathrm{h}}), \tag{2.4}$$

де ω – оператор активації;

X_t – вхідний сигнал мережі;

W_{xh} – матриця вагових коефіцієнтів на поточному часовому кроці;

H_{t-1} – прихований стан на попередньому часовому кроці;

W_{hh} – матриця вагових коефіцієнтів для стану на попередньому часовому кроці;

$\mathrm{b_h}$ – вектор зміщення.

Результати розрахування поточного прихованого стану H_t використовуються для визначення прихованого стану на наступному часовому кроці H_{t+1} та для обчислення вихідного сигналу мережі на поточному кроці [25]:

$$\mathrm{o}_t = \mathrm{H}_t \mathrm{W}_{\mathrm{hq}} + \mathrm{b}_{\mathrm{q}}, \tag{2.5}$$

де H_t – прихований стан на поточному часовому кроці;

W_{hq} – матриця вагових коефіцієнтів вихідного шару;

b_{q} – вектор зміщення вихідного шару.

Через те, що поточний стан в рекурентних нейронних мережах залежить не лише від вхідних даних, а й від попередніх станів, використання стандартного алгоритму навчання ШНМ є неможливим. Для роботи з РНМ використовується його розширення – зворотне поширення помилки в часі.

Першим етапом процесу навчання є виконання прямого проходження, в результаті якого обчислюються приховані стани h_t та виходи o_t для кожного часового кроку в послідовності ($t = 1 \dots T$). Отримані вихідні значення порівнюються з очікуваними результатами та визначається помилка роботи мережі за допомогою обраної функції втрат [25]:

$$L = \frac{1}{T} \sum_{t=1}^{T} l(\mathrm{o}_t, y_t), \tag{2.6}$$

де y_t – очікуваний вихід мережі.

Наступним етапом алгоритму є обчислення градієнтів матриць вагових коефіцієнтів прихованих та вихідного шарів. Спочатку знаходиться градієнт функції втрат по обчислених виходах мережі [25]:

$$\frac{\partial L}{\partial \mathrm{o}_t} = \frac{\partial l(\mathrm{o}_t, y_t)}{T \cdot \partial \mathrm{o}_t}. \tag{2.7}$$

Застосовуючи правило диференціювання складеної функції градієнт ваг вихідного шару $\mathrm{W_{qh}}$ дорівнює сумі градієнтів функції втрат по обчислених виходах мережі (L залежить від $\mathrm{W_{qh}}$ через $\mathrm{o}_1 \ldots \mathrm{o}_T$) [25]:

$$\frac{\partial L}{\partial \mathrm{W_{qh}}} = \sum_{t=1}^{T} \frac{\partial L}{\partial \mathrm{o}_t} \mathrm{h}_t^{\top}. \tag{2.8}$$

На останньому часовому кроці T цільова функція (функція втрат) залежить від прихованого стану h_T лише через o_T [25]:

$$\frac{\partial L}{\partial \mathrm{h}_T} = \mathrm{W_{qh}^{\top}} \frac{\partial L}{\partial \mathrm{o}_T}. \tag{2.9}$$

Для попередніх часових кроків ($t < T$) прихований стан впливає як на поточний вихід, так і на наступний прихований стан. Тому градієнт обчислюється рекурентно [25]:

$$\frac{\partial L}{\partial \mathrm{h}_t} = \mathrm{W_{hh}^{\top}} \frac{\partial L}{\partial \mathrm{h}_{t+1}} + \mathrm{W_{qh}^{\top}} \frac{\partial L}{\partial \mathrm{o}_T}. \tag{2.10}$$

Розкладання формули для всій послідовності для $t \in [1;\ T]$ дає [25]:

$$\frac{\partial L}{\partial \mathrm{h}_t} = \sum_{i=1}^{T} \left(\mathrm{W_{hh}^{\top}}\right)^{T-i} \mathrm{W_{qh}^{\top}} \frac{\partial L}{\partial \mathrm{o}_{T+t-i}}. \tag{2.11}$$

Градієнт матриці вагових коефіцієнтів попереднього стану обчислюється за формулою [25]:

$$\frac{\partial L}{\partial \mathrm{W_{hh}}} = \sum_{t=1}^{T} \frac{\partial L}{\partial \mathrm{h}_t} \mathrm{h}_{t-1}^{\top}. \tag{2.12}$$

Градієнт матриці вагових коефіцієнтів поточного стану визначається як [25]:

$$\frac{\partial L}{\partial \mathrm{W_{hx}}} = \sum_{t=1}^{T} \frac{\partial L}{\partial \mathrm{h}_t} \mathrm{x}_t^{\top}. \tag{2.13}$$

Останнім етапом є оновлення значень вагових коефіцієнтів за допомогою отриманих в результаті розрахунків градієнтів:

$$\mathrm{W} \leftarrow \mathrm{W} - \eta \frac{\partial L}{\partial \mathrm{W}}, \tag{2.14}$$

де η – швидкість навчання.

Описані вище етапи повторюються доти, поки не буде досягнуто умови виходу з циклу навчання або помилка роботи мережі не досягне прийнятної величини.

Як можна побачити з формул, обчислення градієнтів передбачає зведення матриць вагових коефіцієнтів у великі ступені, через що власні значення стають або занадто малими або занадто великими. Це призводить до проблем зникання або вибуху градієнту: у першому випадку з'являється великий ризик зупинки процесу навчання мережі, у другому – потрапляння в локальні мінімуми, тобто неможливості знаходження оптимального рішення. Ці проблеми роблять класичні РНМ непридатними для задач, які потребують вивчення довгострокових залежностей [26].

З метою подолання цих проблем у 1997 році З. Хохрайтером та Ю. Шмідгубером було запропоновано нову архітектуру рекурентних нейронних мереж – довгу короткочасну пам'ять [27].

Ключовим механізмом LSTM є Constant Error Carousel (CEC). Він має лінійну функцію активації та вагу самозв'язку w_{jj}=1, що дозволяє уникнути зменшення помилки роботи мережі та подальшого зникання градієнтів, залишивши помилку константною [27]:

$$y_j(t+1) = f_j\left(net_j(t+1)\right) = f_j\left(w_{jj}y^j(t)\right) = y^j(t), \tag{2.15}$$

де f_j – експоненціальна сигмоїда, відповідає формулі 2.1;

$y^j(t)$ – активація нейрона на поточному часовому кроці.

CEC розміщено всередині спеціального блоку – комірки пам'яті. Для захисту обчислень CEC від можливих збурень через вплив нерелевантної інформації, до комірки було додано мультиплікативні логічні елементи – вентилі входу та виходу [27].

Вентиль входу контролює, яку інформацію буде записано в пам'ять – його було введено для захисту вмісту пам'яті від збурень нерелевантними вхідними даними [26]. Активація вентилю входу на певному часовому кроці відбувається відповідно до [27]:

$$y^{in_j}(t) = f_{in_j}\left(\sum_u w_{in_j u}y^u(t-1)\right), \tag{2.16}$$

де f_{in_j} – експоненціальна сигмоїда, визначається за формулою 2.1;

$w_{in_j u}$ – матриця вагових коефіцієнтів з'єднання вхідного вентилю, що входить в блок u;

$y^u(t-1)$ – активація блоку u на попередньому часовому кроці.

Вентиль виходу вирішує, яку інформацію можна передавати далі до інших блоків – його було представлено з метою захисту інші одиниць від збурення нерелевантним вмістом пам'яті [26]. Активація вентилю виходу відбувається відповідно до [27]:

$$y^{out_j}(t) = f_{out_j}\left(\sum_u w_{out_j u} y^u(t-1)\right), \quad (2.17)$$

де f_{out_j} – експоненціальна сигмоїда, визначається за формулою 2.1;

$w_{out_j u}$ – матриця вагових коефіцієнтів з'єднання вихідного вентилю, що виходить з блоку u;

$y^u(t-1)$ – активація блоку u на попередньому часовому кроці.

Внутрішній стан комірки на певному часовому кроці t визначається шляхом додавання до її попереднього внутрішнього стану на кроці t – 1 нового вхідного сигналу мережі [27]:

$$s_{c_j}(t) = s_{c_j}(t-1) + y^{in_j}(t) g\left(\sum_u w_{c_j u} y^u(t-1)\right), \quad (2.18)$$

де t – часові кроки, $t > 1$;

$s_{c_j}(t-1)$ – попередній внутрішній стан комірки;

g – подвоєна симетрована експоненціальна сигмоїда, яка визначається за формулою 2.2;

$w_{c_j u}$ – матриця вагових коефіцієнтів, що подається на вхід блоку u.

Вихід комірки пам'яті визначається на основі її внутрішнього стану та значення активації вентиля виходу [27]:

$$y^{c_j}(t) = y^{out_j}(t) h\left(s_{c_j}(t)\right), \quad (2.19)$$

де h – функція симетрованої експоненціальної сигмоїди, визначається за формулою 2.2.

При подальших дослідженнях було виявлено, що запропонована архітектура LSTM має значний недолік при обробці безперервних потоків.

Якщо вектор вхідних даних не сегментований на конкретні фіксовані по довжині послідовності, внутрішній стан комірки пам’яті починає необмежено зростати лінійно. Це зростання веде до насичення функції стиснення виходу h, що, у свою чергу, призводить до вимивання похідної функції стиснення та втрати здатності до функціонування комірки як пам’яті. Для подолання цієї проблеми у 1999 році було запропоновано введення нового елементу до комірки пам’яті – вентилю забування [28].

Вентиль забування розроблений для скидання комірок пам’яті, якщо їхній вміст застарів і є непридатним для подальшого використання. Активація вентилю забування подібна до формул активацій інших вентилів [28]:

$$y^{\varphi_j}(t) = f_{\varphi_j}\left(\sum_u w_{\varphi_j u} y^u(t-1)\right), \tag{2.20}$$

де f_{φ_j} – експоненціальна сигмоїда, визначається за формулою 2.1;

$w_{\varphi_j u}$ – матриця вагових коефіцієнтів з’єднання вентилю забування, що входить в блок u.

З урахування вентилю забування формула 2.18 приймає вигляд [28]:

$$s_{c_j}(t) = y^{\varphi_j}(t) s_{c_j}(t-1) + y^{in_j}(t) g\left(\sum_u w_{c_j u} y^u(t-1)\right). \tag{2.21}$$

Ваги зміщення вхідного та вихідного вентилів ініціалізуються негативними значеннями, а ваги зміщення вентилю забування – позитивними. На початку навчання значення активації вентилю забування є близьким до 1, що змушує модифіковану комірку пам’яті працювати як стандартний LSTM-блок. З часом комірка навчається знижувати активацію, скидаючи себе для обробки нової інформації [28].

На рисунку 2.1 представлено схему внутрішньої будови LSTM-блоку.

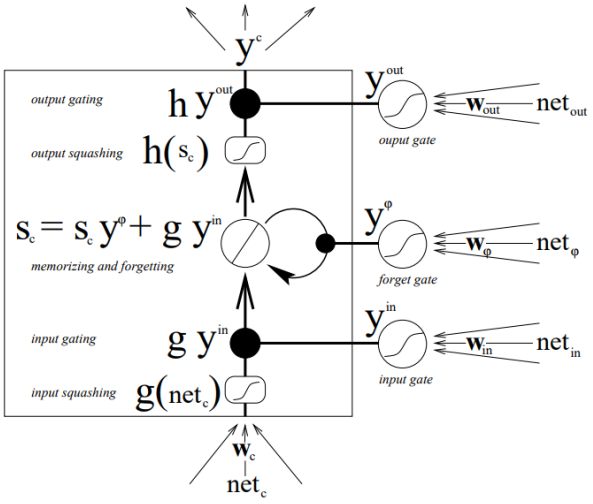


Рисунок 2.1 – Внутрішня будова LSTM-блоку

Наявність трьох вентилів, входу, виходу та забування, призводить до збільшення загальної кількості вагових коефіцієнтів мережі не більш ніж в 3^2 рази у випадку повної зв'язності [27]. Це може впливати на обчислювальні витрати, збільшуючи їх. З метою пришвидшення обчислювань було запропоновано іншу архітектуру РНМ – вентильний рекурентний вузол.

GRU було створено на основі LSTM, однак, на відміну від неї, використовує лише два вентилі, скидання та уточнення, що робить дану архітектуру більш простою в обчисленнях та реалізації [29].

Вентиль скидання вирішує ігнорувати попередній стан чи ні у подальших обчисленнях. Визначається відповідно до [29]:

$$r_j = \sigma\left(\left[\mathrm{W}_r \mathrm{x}\right]_j + \left[\mathrm{U}_r \mathrm{h}_{\langle t-1\rangle}\right]_j\right), \tag{2.22}$$

де σ –функція експоненціальної сигмоїди, відповідає формулі 2.1;

W_r, U_r – матриці вагових коефіцієнтів;

x – вектор вхідних даних;

$\mathrm{h}_{\langle t-1\rangle}$ – попередній прихований стан;

$[.]_j$ – відповідає *j*-му елементу вектору.

Вентиль уточнення контролює, яка частина інформації з попереднього прихованого стану передається до поточного. Цей вентиль працює подібно до комірки пам'яті в LSTM, допомагаючи запам'ятовувати довгострокову інформацію. Вентиль уточнення обчисляється як [29]:

$$z_j = \sigma\left([\mathrm{W}_z \mathrm{x}]_j + \left[\mathrm{U}_z \mathrm{h}_{\langle t-1\rangle}\right]_j\right). \tag{2.23}$$

Новий прихований стан обчислюється за формулою [29]:

$$\tilde{h}_j^{\langle t\rangle} = \phi\left([\mathrm{Wx}]_j + \left[\mathrm{U}\left(\mathrm{r} \odot \mathrm{h}_{\langle t-1\rangle}\right)\right]_j\right), \tag{2.24}$$

де ϕ – функція гіперболічного тангенсу, наведено у формулі 2.3;

$\odot$ – операція поелементного множення.

Якщо значення вентилю скидання є близьким до 0, прихований стан на попередньому часовому кроці ігнорується, а скидання здійснюється виключно на основі поточного вхідного сигналу. Це дозволяє прихованому стану відкинути ту інформацію, яка є нерелевантною для подальших обчислень [29].

Новий поточний стан визначається відповідно до [29]:

$$h_j^{\langle t\rangle} = z_j h_j^{\langle t-1\rangle} + \left(1 - z_j\right)\tilde{h}_j^{\langle t\rangle}. \tag{2.25}$$

Оскільки кожний GRU-блок має окремі вентилі скидання та уточнення, різні блоки можуть навчатися фіксувати залежності в різних часових масштабах. Вентиль скидання блоків, які фіксують короткострокові залежності, частіше активується, ніж вентиль уточнення. І навпаки, вентиль уточнення блоків, які фіксують довгострокові залежності, знаходиться переважно в активному стані, на відміну від вентилю скидання [29].

На рисунку 2.2 представлено схему внутрішньої будови GRU-блоку.

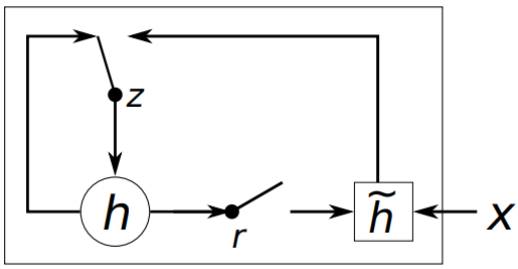


Рисунок 2.2 – Внутрішня будова GRU-блоку

Однак, попри наявність переваг та недоліків у кожної з вищеописаних архітектур рекурентних нейронних мереж, їхня ефективність та обчислювальна складність критично залежить від складності вирішуваної задачі. Таким чином, висновки про ефективність та доцільність використання кожної окремої архітектури мають ґрунтуватися на результатах досліджень, проведених в контексті однієї конкретної задачі.

2.2 Аналіз алгоритмів послідовного прийняття рішень

Моделювання складних систем та прийняття оптимальних рішень в умовах невизначеності є однією з основних задач сучасного машинного навчання. Існують різноманітні методи для розв’язання таких задач, традиційно використовуються алгоритми пошуку по дереву, які добре працюють з динамікою середовища.

Дерево пошуку Монте-Карло – це алгоритм прийняття рішень, який полягає у пошуку в комбінаторних просторах, представлених у вигляді дерев. Вузли дерев уособлюють собою стани (конфігураційні задачі), а ребра визначають переходи (дії) з одного стану в інший [30].

Алгоритм MCTS було вперше запропоновано у 2006 році в роботах Коксіса, Шепешварі та Колума як алгоритм для створення комп'ютерних гравців у грі Go [30].

На відміну від багатьох методів, які використовувалися раніше в комбінаторних іграх, MCTS не є суто евристичним. Це означає, що додаткові знання, окрім правил гри (або задачі, загалом) не потрібні. Однак евристики можуть бути використані та включені в MCTS для підвищення його ефективності та покращити збіжність [30].

Формально, MCTS застосовується до проблем, змодельованих як марковські процеси вирішування, що визначаються як кортеж (S, A_s, P_a, R_a), де [30]:

– S – це множина станів, можливих у середовищі (просторі станів). Конкретний стан $s_0 \in S$ визначається як початковий стан;

– A_s – множина дій, доступних для виконання у стані s;

– $P_a(s, s')$ – це функція переходу, змодельована ймовірністю того, що дія a, виконана у стані s, призведе до стану s';

– $R_a(s)$ – безпосередня винагорода за досягнення стану s дією a.

MCTS є алгоритмом з можливістю переривання в будь-який момент, тобто його можна зупинити в будь-який час і отримати найкращу на даний момент дію (рішення) відповідно до [30]:

$$a^* = \arg\max_{a \in A(s)} Q(s, a)\,, \tag{2.26}$$

де $A(s)$ – набір дій, доступних у стані s, в якому має бути прийнято рішення;

$Q(s, a)$ – емпіричний середній результат виконання дії a у стані s.

Алгоритм MCTS працює ітеративно, на кожній ітерації виконується наступні чотири кроки [30]:

– вибір: алгоритм здійснює пошук у тій частині дерева, яка вже представлена в пам'яті. Вибір завжди починається з кореневого вузла і на

кожному рівні обирається наступний вузол відповідно до стратегії вибору, яку також називають стратегією дерева. Цей етап завершується, коли відвідується кінцевий вузол, а наступний вузол або ще не представлений у пам’яті, або досягнуто кінцевого стану гри (задачі);

– розширення: якщо вибір не досяг кінцевого стану, розширення додає принаймні один новий дочірній вузол до дерева, представленого в пам’яті. Новий вузол відповідає стану, якого було досягнуто при виконанні останньої дії на етапі вибору. Коли розширення досягає кінцевого стану, що трапляється вкрай рідко, поточна ітерація переходить безпосередньо до зворотного поширення;

– симуляція: виконує повне випадкове моделювання гри/задачі, тобто досягнення кінцевого стану та отримання виграшів – частина алгоритму «Монте-Карло»;

– зворотне поширення: поширює виграші для всіх змодельованих агентів у грі назад до всіх вузлів уздовж шляху від останнього відвіданого вузла в дереві (листя) до кореня. Статистика оновлюється.

На рисунку 2.3 представлено схему ітерації алгоритму MCTS [30].

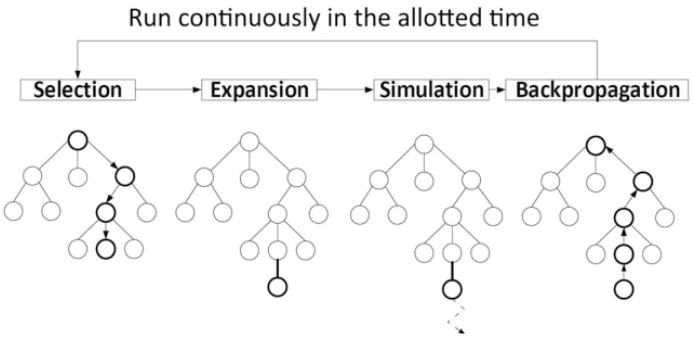


Рисунок 2.3 – Ітерація алгоритму MCTS

Збільшення кількості ітерацій призводить до більш впевненої статистики і більшої ймовірність того, що рекомендована найкраща дія дійсно є оптимальною [30].

Метою стратегії вибору є підтримка належного балансу між дослідженням (тестуванням дій, які ще не були оцінені) та використанням (наданням переваги перевіреним діям, які було визначено як найкращі). Найпоширеніший метод, який фактично є основою MCTS, називається «Upper Confidence Bounds applied for Trees» (UCT). Спочатку потрібно перевірити кожну дію один раз, а потім за наступною формулою [30]:

$$a^* = \arg\max_{a \in A(s)} \left\{ Q(s,a) + C\sqrt{\frac{ln[N(s)]}{N(s,a)}} \right\}, \qquad (2.27)$$

де $N(s)$ – кількість разів, коли стан s було відвідано на попередніх ітераціях;

$N(s,a)$ – кількість разів, коли дія a була вибрана у стані s;

C – константа, яка контролює баланс між дослідженням та використанням; рекомендоване значення для тестування становить $\sqrt{2}$, однак, сильно залежить від задачі, тому потребує додаткового налаштування.

Завдяки формулі UCT та методу випадковій відбору, MCTS здійснює асиметричний пошук у дереві гри, на відміну від традиційних методів пошуку по деревах. Найбільш перспективні лінії гри досліджуються більш ретельно [30].

Однак, через свою архітектуру, алгоритм MCTS може мати обмеження при вирішенні задач, в яких середовище має високий рівень стохастичності. На етапі симуляції випадкові події можуть викликати шум, що призводить до неправильних оцінок гілок дерева рішень і призводить до консервативної поведінки агента при виборі наступних дій. Для вирішення цих проблем можливе використання підходів, здатних узагальнювати глобальний досвід взаємодії із середовищем, зокрема, методи навчання з підкріпленням.

Навчання з підкріпленням – це клас методів для вирішення різних видів задач послідовного прийняття рішень [31]. В рамках цих методів розглядається взаємодія агента та середовища: агент навчається шляхом прийняття певних рішень, а середовище реагує на них, надає нові ситуації агенту та винагороджує його у разі виконання правильних дій, або карає при неправильних [32]. Фактично, середовище формалізовано як марковській процес вирішування, які було описано раніше, та визначається кортежом (S, A_s, P_a, R_a). На рисунку 2.4 представлено схему взаємодії агента та середовища в марковському процесі вирішування.

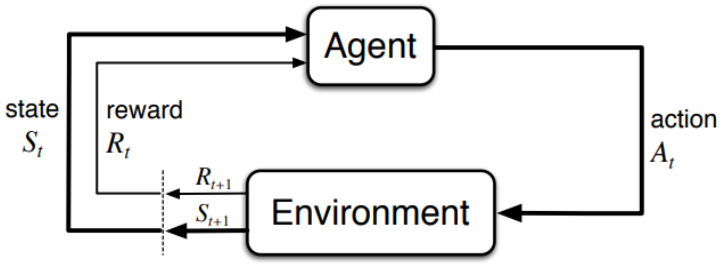


Рисунок 2.4 – Взаємодія агента та середовища в марковському процесі вирішування

У кожний дискретний момент часу t $(t = 0, 1, 2\ldots)$ агент отримує певне представлення стану середовища $S_t \in S$ і, на його основі, обирає дію $A_t \in A(s)$. Через один часовий крок, частково як наслідок своєї дії, агент отримує числову винагороду, $R_{t+1} \in R \subset \mathbb{R}$, і опиняється в новому стані – S_{t+1}. Таким чином, створюється послідовність або траєкторія, яка починається як: $S_0, A_0, R_1, S_1, A_1, R_2, \ldots$ [32].

Загалом, метою агента є максимізація загальної суми винагороди, яку він отримує, тобто максимізацію не негайної винагороди (R_t), а кумулятивної винагороди в довгостроковій перспективі. Ця кумулятивна винагорода називається очікуваною віддачою і визначається як певна

специфічна функція послідовності винагород. У найпростішому випадку, вона відповідає сумі винагород [32]:

$$G_t \doteq R_{t+1} + R_{t+2} + R_{t+3} + \cdots + R_T, \tag{2.28}$$

де T – кінцевий часовий крок.

Однак визначення віддачі таким чином (як у формулі 2.28) можливо лише при наявності певного термінального стану задачі (тобто кінцевий час T відомий заздалегідь). При $T = \infty$ кумулятивна винагорода буде прагнути до нескінченості, через що збіжність методу не може бути гарантованою.

Для вирішення цієї проблеми використовується концепція знецінювання, згідно з якою агент намагається обирати дії таким чином, щоб сума знецінюваних винагород, які він отримує в майбутньому, була максимізована. Зокрема, він вибирає A_t, щоб максимізувати очікувану знецінену віддачу [32]:

$$G_t \doteq R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+1} + \cdots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}\,, \tag{2.29}$$

де γ – коефіцієнт знецінювання ($\gamma \in [0;1]$).

Прагнучи максимізувати віддачу, агент приймає рішення щодо виконання тих чи інших дій, спираючись на певну внутрішню логіку – стратегію.

Формально, стратегія – це відображення станів у ймовірності вибору кожної можливої дії. Якщо агент дотримується стратегії π у момент часу t, то $\pi(a|s)$ – це ймовірність того, що $A_t = a$, якщо $S_t = t$. Методи навчання з підкріпленням визначають, як змінюється політика агента в результаті його досвіду [32].

Функція цінності стану s за стратегії π, що позначається $v_\pi(s)$, це очікувана віддача, починаючи з s і наступаючи за π, визначається як [32]:

$$v_\pi(s) \doteq \mathbb{E}_\pi[G_t \mid S_t = s], \tag{2.30}$$

де $\mathbb{E}_\pi[\cdot]$ – очікуване значення випадкової величини за умови, що агент дотримується стратегії π;

t – будь-який часовий крок.

Функція цінності дії a за стратегії π позначається як $q_\pi(s, a)$ і визначає очікувану віддачу, починаючи з s, виконання дії a та подальшого дотримання стратегії π [32]:

$$q(s) \doteq \mathbb{E}_\pi[G_t \mid S_t = s, A_t = a]. \tag{2.31}$$

Розв'язання задачі навчання з підкріпленням означає знаходження оптимальної стратегії (π_*), тобто такої, для якої очікувана віддача більша або дорівнює віддачі іншої стратегії для всіх станів – $\pi \geq \pi'$ тоді і тільки тоді, коли $v_\pi(s) \geq v_{\pi'}(s)$ для всіх $s \in S$. Оптимальних стратегії може бути декілька, і вони всі мають однакову оптимальну функцію цінності стану [32]:

$$v_*(s) \doteq \max_\pi v_\pi(s)\,. \tag{2.32}$$

Оптимальні стратегії також мають однакову оптимальну функцію цінності дії [32]:

$$q_*(s, a) \doteq \max_\pi q_\pi(s, a)\,. \tag{2.33}$$

Для пари стан-дія (s, a) ця функція визначає очікувану віддачу від виконання дії a у стані s та подальшого дотримання оптимальної стратегії. Отже, q_* може бути виражена через v_* наступним чином [32]:

$$q_*(s,a) = \mathbb{E}[R_{t+1} + \gamma v_*(S_{t+1}) \mid S_t = s,\ A_t = t]. \qquad (2.34)$$

Загалом, класичні методи навчання з підкріпленням мають високу ефективність під час вирішення задач в середовищах із дискретними та скінченними просторами станів. Однак вони можуть страдати від «прокляття розмірності» при переході до задач з безперервними або багатовимірними просторами. Для вирішення цих проблем було запропоновано гібридні нейромережеві підходи, наприклад, актор-критик.

Методи актор-критик – це клас методів навчання з підкріпленням, які поєднують переваги алгоритмів, заснованих на оптимізації градієнту стратегії та функції цінності. У таких методах навчаються два окремих компонента:

– актор – параметризація стратегії $\pi(a|s,\theta)$, яка відповідає за вибір дій;

– критик – параметризація функції цінності стану $\tilde{v}(s,\mathrm{w})$, яка служить базовою лінією та оцінює переходи, здійснені актором.

На відміну від багатьох методів градієнту стратегії, в яких функція цінності оцінює лише початковий стан переходу, алгоритми актор-критик застосовують функцію цінності також до наступного стану. Це дозволяє реалізовувати бутстрепінг, тобто, коли оцінка другого стану (знецінена та додана до винагороди) дає однокрокову оцінку віддачу ($G_{t:t+1}$) для оцінки обраної дії [32].

Однокрокові методи актор-критик є повністю онлайн-методами та працюють інкрементально, що дозволяє уникнути складнощів, пов’язаних з розрахунком повної віддачі. Основою механізму оновлення є помилка TD, яка визначається як [32]:

$$\delta_t = R_{t+1} + \gamma \hat{v}(S_{t+1}, \mathrm{w}_t) - \hat{v}(S_t, \mathrm{w}_t), \qquad (2.35)$$

де w – ваги функції цінності стану.

З урахуванням цієї помилки параметри стратегії оновлюються з метою підвищенням ймовірність дій відповідно до [32]:

$$\theta_{t+1} \doteq \theta_t + \alpha\delta_t \frac{\nabla\pi(A_t|S_t,\theta_t)}{\pi(A_t|S_t,\theta_t)}. \quad (2.36)$$

Загалом, велика кількість моделей навчання з підкріпленням, що використовують апроксиматори функцій на основі нейронних мереж, спиралися на такі методи, як глибоке Q-навчання, класичні методи градієнту стратегії або методи довірчої області. Класичні методи градієнту стратегії працюють шляхом обчислення оцінки градієнта стратегії та використання стохастичного градієнтного підйому, диференціюючи цільову функцію, визначену як [33]:

$$L^{PG}(\theta) = \widehat{\mathbb{E}}_t\left[\log \pi_\theta(a_t|s_t)\hat{A}_t\right], \quad (2.37)$$

де $\widehat{\mathbb{E}}_t[\cdot]$ – емпіричне середнє значення для скінченної сукупності вибіркових даних;

π_θ – стохастична стратегія;

$\hat{A}_t$ – оцінка функції переваги на кроці часу t.

Однак, виконання декількох кроків оптимізації за цією стандартної функцією втрат із використанням однієї й тієї ж траєкторії часто призводить до надмірно великих оновлень стратегії та низької ефективності використання даних [33].

Метод оптимізації стратегії довірчої області дозволяє вирішити цю проблему шляхом максимізації сурогатної цільової функції з урахуванням жорсткого обмеження на розмір оновлення стратегії, використовуючи дивергенцію Кульбака-Лейблера (KL). Незважаючи на свої ефективність, TRPO є складною в реалізації та несумісна з архітектурами, що передбачають спільне використання параметрів або шум. Метод

проксимальної оптимізації стратегії було запропоновано для досягнення ефективності використання даних та надійної роботи TRPO, але з використанням набагато простішої оптимізації першого порядку [33].

З метою запобігання надмірно великих оновлень стратегії без використання складних обмежень KL, PPO містить нову сурогатну цільову функцію, яка штрафує за зміни, що занадто віддаляють стратегію від її попереднього стану.

Метод базується на коефіцієнті ймовірності, який порівнює ймовірності дій за поточною та попередньою стратегією [33]:

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}. \tag{2.38}$$

Однак максимізація цільової функції (як це виконується в TPRO) без обмежень зазвичай призводить до значного відхилення $r_t(\theta)$ від 1 та спричиняє надмірно великі оновлення стратегії. Для вирішення цих проблем, у PPO використовується усічена сурогатна цільова функція [33]:

$$L^{CLIP}(\theta) = \hat{\mathbb{E}}_t\left[\min\left(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t\right)\right], \tag{2.39}$$

де ϵ – гіперпараметр відсікання ($\epsilon = 0{,}2$).

Беручи мінімум усіченого та неусіченого членів, функція створює песимістичну оцінку, формуючи нижню межу неусіченої цільової функції. Зрештою, це гарантує, що зміни коефіцієнта ймовірності ігноруються, якщо вони призводять покращення цільової функції, але вони враховуються, якщо вони її погіршують [33].

Більшість методів обчислення оцінки функції переваги зі зниженою дисперсією потребують вивченої цінності стану $V(s)$. При використанні архітектури нейронної мережі, в якій параметри є спільними для стратегії та цієї функції цінності, функція втрат має охоплювати обидва компоненти.

Для забезпечення достатнього рівня дослідження під час навчання, також додається бонус ентропії S. Поєднуючи ці члени, на кожній ітерації PPO оптимізує наступну цільову функцію [33]:

$$L_t^{CLIP+VF+S}(\theta) = \hat{\mathbb{E}}_t[L_t^{CLIP}(\theta) - c_1 L_t^{VF}(\theta) + c_2 S[\pi_\theta](s_t)], \quad (2.40)$$

де c_1, c_2 – коефіцієнти;

S – бонус ентропії;

L_t^{VF} – це втрата за квадратичною помилкою, яка обчисляється відповідно до:

$$L_t^{VF}(\theta) = \left(V_\theta(s_t) - V_t^{\text{targ}}\right)^2. \quad (2.41)$$

Завдяки чергуванню збору вибіркових даних із стратегії та виконанню декількох ітерацій оптимізації на цих вибіркових даних із використанням цієї цільової функції, PPO забезпечує баланс між складністю вибірки, часом виконання та простотою реалізації [33].

Таким чином, розглянуті алгоритми мають різні підходи при визначенні дій. MCTS покладається на симуляції майбутніх станів безпосередньо під час прийняття рішень, що робить його вразливим до виникнення збурень в реальному часі. Методи навчання з підкріпленням, зокрема PPO, формує параметризовану стратегію, якою керується при визначенні наступних дій, на етапі навчання. Це дозволяє агенту приймати робастні рішення на основі глобального узагальненого досвіду.

3 ОПИС ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ ТА ПРОГРАМНОЇ РЕАЛІЗАЦІЇ

3.1 Формування та аналіз вихідного простору ознак

Вихідний набір даних, який було використано у ході досліджень, являє собою гібридний масив, сформований шляхом злиття інформації з двох офіційних джерел.

Першим джерелом даних була відкрита бібліотека FastF1 [34], яка забезпечує доступ до офіційного історичного хронометражу та телеметрії етапів «Формули-1». Завдяки даному API було отримано дані про погодні умови, статуси траси під час гонки та поколову статистику (типи шин, терміни їхньої експлуатації, час перегонового кола та інші).

Другим джерелом даних був офіційний сайт компанії-постачальника шин для «Формули-1» – Pirelli [35]. Завдяки їхнім аналітичним публікаціям було отримано дані про характеристики трас: рівень абразивності та адгезії асфальту, ступінь впливу бічних сил та інші.

Вихідний набір містив дані, які було взято у період з 2022 по 2025 рр. Це було зумовлено заморожування технічного регламенту саме в цей період (описано в п. 1.1). Використання даних, отриманих під час заморожування, гарантує сталість роботи машини з точки зору силових елементів. Проте воно не стосується фіксації процесів розвитку аеродинамічних компонент, що, потенційно, може призвести до змін в ефективності роботи машини та, як наслідок, до наявності певного зсуву в даних.

Незважаючи на те, що аналітичні публікації Pirelli містять багато інформації про специфіку трас з боку як особливостей асфальтного покриття, так і впливу фізичних сил, деякі оцінки, важливі для розуміння роботи шин, не є представленими.

Зокрема, Pirelli надає оцінку рівню впливу бічних сил у поворотах, проте, оцінка повздовжніх сил відсутня, хоча є не менш важливою для моделювання деградації шин. Її дія пов'язана з наявністю повздовжнього ковзання в місці контакту шини з дорожнім покриттям та характеризується коефіцієнтом ковзання. Зі збільшенням повздовжньої сили зростає рівень ковзання, що призводить до збільшення енергії в шині, її нагрівання (виводить її з робочих температурних діапазонів, описаних в п. 1.2) і, як наслідок, прискорює темп зносу протектора.

Відповідно до класичних понять динаміки автомобілю [36], повздовжня сила може бути як рушійною, так і гальмівної, залежно від режиму руху: розгону або гальмуванню. На основі наданих Pirelli оцінок рушійних та гальмівних сил було обчислено індекс поздовжньої сили, характерної для кожної траси:

$$L = \frac{T + B}{2}, \tag{3.1}$$

де T – оцінка рушійної сили;

B – оцінка гальмівної сили.

З метою репрезентації особливостей гумових сумішей, з яких виготовлюють шини для болідів «Формули-1», було введено опосередковану змінну, що відображає, наскільки певний комплект шин є схильним до деградації:

$$TDR = AA \cdot CSS, \tag{3.2}$$

де AA – оцінка рівня абразивності асфальтного покриття (надається Pirelli);

CSS – оцінка «м'якості» гумової суміші.

Для «Формули-1» Pirelli виготовляє шини з різних гумових сумішей C1–C6 від найжорсткішого складу до найм'якішого. CSS було визначено на

основі даних сумішей та закодовано від 1 до 6, де 1 відповідає найбільш «жорсткому» комплекту (найменш «м'якої» суміші), а 6 – найменш «жорсткому» (найбільш «м'якої» суміші). Відповідно, чим більші оцінки рівня абразивності асфальту траси та «м'якості» шини, тим вона більш схильна до надмірної деградації, і навпаки.

У вихідному наборі телеметрії шини були представлені номінальними типами: Hard, Medium та Soft. У рамках регламенту «Формули-1» ці типи є контекстно-залежними і призначаються окремо для кожного Гран-прі з набору гумових сумішей C1–C6, які було описано вище.

Використання номінальних типів шин у якості вхідних ознак вводить приховану неоднорідність в навчальну вибірку. Наприклад, шини Soft (C3) на Гран-прі Бахрейну та Soft (C5) на Гран-прі Саудівської Аравії мають принципово різні властивості з точки зору зчеплення та темпів деградації. Для забезпечення структурної однорідності простору ознак типи шин було замінено на індекси гумових сумішей відповідно до офіційного розподілу шин для кожного конкретного етапу.

У таблиці 3.1 представлено умовні позначення ознак та їх значення.

Таблиця 3.1 – Назви ознак

№ ознаки	Назва ознаки	Значення
1	Tyre Life	Кількість гоночних кіл, проведених на шинах
2	Track Temp	Температура траси у °C
3	Grip Evolution State	Еволюція зчеплення траси
4	Longitudinal Energy	Повздовжні сили
5	Lateral	Бічні сили
6	Thermal Degradation Risk	Схильність комплекту шин до деградації
7	Slip Ratio	Коефіцієнти ковзання
8	Asphalt Grip	Адгезія асфальту
9	Comp C1	Гумова суміш C1
10	Comp C2	Гумова суміш C2
11	Comp C3	Гумова суміш C3
12	Comp C4	Гумова суміш C4
13	Comp C5	Гумова суміш C5
14	Comp C6	Гумова суміш C6

У таблиці 3.2 представлено типи вхідних ознак.

Таблиця 3.2 – Типи ознак

№ ознаки	Назва ознаки	Тип ознаки
1	Tyre Life	Дискретна
2	Track Temp	Безперервна
3	Grip EvolutionState	Безперервна
4	Longitudinal Energy	Безперервна
5	Lateral	Порядкова
6	Thermal Degradation Risk	Порядкова
7	Slip Ratio	Безперервна
8	Asphalt Grip	Порядкова
9	Comp C1	Бінарна
10	Comp C2	Бінарна
11	Comp C3	Бінарна
12	Comp C4	Бінарна
13	Comp C5	Бінарна
14	Comp C6	Бінарна

Через помилки синхронізації телеметрії, вихідний набір даних містив пропущені значення. Їх було оброблено в залежності від типу ознаки. Категоріальні змінні (наприклад, тип шин) було видалено повністю через потенційний ризик порушення цілісності даних під час заповнення пропусків. Для безперервних змінних (наприклад, час кола) було застосовано методи заповнення попереднім або наступним значенням в залежності від локалізації пропуску в межах гоночного відрізку (на початку або в кінці), проведеного на конкретному комплекті шин.

Перед використанням набору було проведено масштабування ознак. Цей крок був обов’язковим, адже значення ознак мали великий розбіг, що могло вплинути на результати навчання штучних нейронних мереж.

У якості методу масштабування було використано нормалізацію, яка визначається за формулою [23]:

$$x_i^{\text{norm}} = \frac{(x_i - x_{min})(b - a)}{x_{max} - x_{min}} + a, \tag{3.3}$$

де i – індекс спостереження у вихідному наборі даних;

x_i^{norm} – нормалізоване i-е значення вхідного сигналу;

x_i – вихідне i-е значення сигналу;

x_{min} та x_{max} – мінімальне та максимальне значення сигналу у вихідному наборі даних;

a та b – нижня та верхня границі нормалізованого діапазону, було обрано -1 та 1 відповідно.

Для дослідження сформованого простору ознак було побудовано теплову матрицю рангових кореляцій Спірмена. Даний метод дозволяє провести оцінку не лише строго лінійних, а й монотонних нелінійних залежностей в даних. На рисунку 3.1 представлено отримані результати.

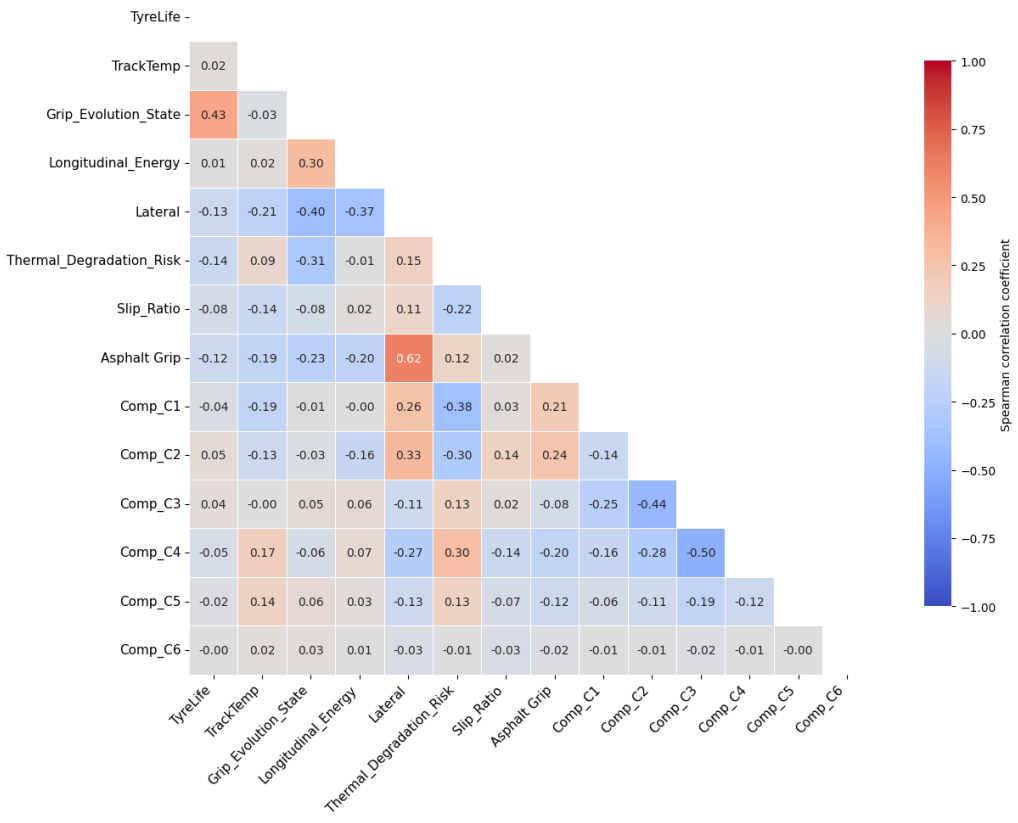


Рисунок 3.1 – Візуалізація результатів обчислення коефіцієнтів рангової кореляції Спірмена

Як можна побачити, між ознаками, що відображають дії бічних (Lateral) та повздовжніх сил на автомобіль (Longitudinal_Energy), присутній зворотній зв’язок (-0,37). Цей результат відповідає фізичним обмеженням зчеплення шини з дорожнім покриттям. Відповідно до концепції «кола зчеплення» [36], результуюча сила, що реалізується шиною, обмежена та розподіляється між поздовжньою та поперечною компонентами. Збільшення бічного навантаження в повороті призводить до зниження поздовжньої сили.

Найсильніший позитивний зв’язок (0,62) наявний між адгезією асфальтного покриття (Asphalt Grip) та дією бічних сил на болід (Lateral). На основі даного результату можна підтвердити, що при вищому рівні зчеплення шин з трасою машина спроможна на більшій швидкості проходити повороти.

Також можна побачити позитивну рангову кореляцію (0,43) між тривалістю відрізка (TyreLife) та еволюцією зчеплення траси (Grip_Evolution_State). Ця залежність відображає та підтверджує наявність процесу «прогумовування» траси під час перебігу гонки, який було описано в п. 1.2.

Окрім цього, отримані результати демонструють залежність між складом гумової суміші та її схильністю до деградації. Більш жорсткі склади (Comp_C1, Comp_C2) демонструють зворотну кореляцію зі схильністю до деградації (Thermal_Degradation_Risk): -0,38 та -0,3 відповідно. Більш м’які склади (Comp_C3, Comp_C4, Comp_C5) мають пряму кореляцію (0,13 та 0,3). Варто зазначити, що низькі значення коефіцієнтів кореляції, які відповідають складу гуми C6 (Comp_C6), зумовлені невеликою кількістю наявних даних, адже дана гумова суміш була представлена в 2025 р. та показала низьку ефективність в гоночних умовах, через що використовувалася командами рідко.

Таким чином, вихідний простір ознак складається з 14 змінних, які описують різні аспекти, що впливають на автомобільні шини під час

перегонів. Проведений аналіз визначених ознак показав наявність зв’язків між ними, які необхідно враховувати під час моделювання процесу деградації шин.

3.2 Математична формалізація процесу деградації гоночних шин

Як було зазначено у п. 1.1, на шинах болідів «Формули-1» відсутні датчики вимірювання рівня зносу протектора, що унеможливлює кількісне визначення та розуміння реальних темпів деградації. Проте оцінка стану шин може здійснюватися завдяки опосередкованим методам, наприклад, через дослідження зміни поведінки автомобіля при проходженні поворотів, візуальний аналіз протектору (наявність певного патерну деградації, описано у п. 1.2), суб’єктивні відчуття пілота, тривалість експлуатації поточного комплекту шин, а також моніторинг часу проходження перегонового кола.

Першим індикатором деградації шин є збільшення часу, за який пілот проїжджає коло. Як правило, це зумовлено втратою зчеплення шин з трасою: машина втрачає стійкість та керованість, через що гонщик стає більш обережним при проходженні траси, зокрема, проїжджає повороти на нижчій швидкості, або починає скидати швидкість набагато раніше потрібної точки гальмування.

Однак виміряний час проходження гоночного кола не є об’єктивною оцінкою темпів зносу: на початку гонки баки автомобілів заправлені повністю. У ході гонки пальне згорає, що робить болід легшим та, як наслідок, швидшим. Для визначення реального часу проходження кола необхідно робити поправку на пальне:

$$T = T_l - \gamma m_f, \tag{3.4}$$

де T_l – виміряний час перегонового кола;

m_f – маса пального;

γ – коефіцієнт впливу маси пального на час (~0,03 с/коло, проте може відрізнятися в залежності від конфігурації траси).

Так як «Формула-1» не надає офіційну телеметрію маси пального, її було визначено опосередковано шляхом віднімання поточного номера кола від загальної дистанції гонки:

$$m_f = N - n, \tag{3.5}$$

де N – загальна дистанція гонки (загальна кількість перегонових кіл);

n – номер поточного кола.

Таким чином, змінна m_f являла собою спадну послідовність чисел (наприклад, від 57 до 0), імітуючи поступове згорання пального та зменшення його маси від максимально потрібної для гоночної дистанції до мінімального.

Для визначення коефіцієнту впливу маси пального на час, γ, притаманний для кожної окремої траси, було використано лінійну регресію.

Набір даних було відфільтровано з метою виокремлення найбільш тривалих відрізів, при яких було використано найжорсткіші комплекти шин. Це зумовлено тим, що склад гуми в таких комплектах є найменш чутливим до термодинамічних змін, менш схильні до передчасної та надмірної деградації, що дозволяє отримати валідні гоночні кола.

Як зазначалося раніше, на час гоночного кола впливає, за умови відсутності виникнення ситуаційних факторів, два основних чинника: зменшення загальної ваги машини в результаті згорання пального (що робить болід швидшим) та зношування протектору шини (що сповільнює автомобіль). Спираючись на це, було визначено наступне рівняння множинної лінійної регресії:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon, \tag{3.6}$$

де Y – час кола;

X_1 – рівень зношування протектору шини;

X_2 – маса пального;

$\beta_0, \beta_1, \beta_2$ – коефіцієнти регресії;

ε – похибка моделі.

Так як рівень зношування протектору шини (X_1) окремо не вимірюється на машинах «Формули-1», визначення даного фактору в моделі було здійснено опосередковано, шляхом використання ознаки з вихідного набору даних, що відповідає тривалості експлуатації поточного комплекту шин (TyreLife).

У результаті вирішення задачі лінійної регресії методом найменших квадратів було отримано значення коефіцієнтів впливу маси пального на час для трас «Формули-1», представлені в додатку А.

Однак через те, що траси, санкціоновані для проведення гонок «Формули-1» мають різну довжину, використання абсолютного часу кола для прогнозування деградації шин може призвести до низької узагальнюючої здатності моделі. Крім того, незважаючи на наявність технічного регламенту, машини команд відрізняються за технічними характеристиками та роботою з шинами. Для коректного розуміння та прогнозування темпів деградації необхідно використовувати різницю з кола:

$$\Delta T = T - T_b, \tag{3.7}$$

де T – час кола з поправкою на витрати пального;

T_b – базовий час кола для шин конкретного відрізку.

У якості базового часу кола на конкретному відрізку T_b кола було обрано найшвидше з 3 перших кіл у цьому відрізку. Як було зазначено у п. 1.2, кожній гумовій суміші, з якої виготовляють комплекти шин, властивий певний робочий температурний діапазон, при якому шина набуває оптимального зчеплення з трасою. Після виїзду з піт-лейну

температура щойно встановлених шин є занадто низькою та знаходиться поза межами цього інтервалу, що негативно впливає на швидкість боліда та, як наслідок, збільшує час кола. Зазвичай, гонщикам потрібно декілька гоночних кіл, щоб ввести шини в робочий температурний діапазон, через що час першого кола є нерепрезентативний для оцінки реального темпу машини та ефективності роботи шин.

Таким чином, визначення реального темпу деградації є складним процесом через неможливість можливості безпосереднього вимірювання, спричиненого відсутністю відповідних датчиків. Для цього командами використовуються опосередковані методи, зокрема, на основі часу, за який пілот проїжджає гоночне кола. Однак на виміряний час кола може впливати безліч факторів, що може вливати на об’єктивність оцінки. З метою отримання якомога чистого темпу використовується різниця часу з гоночного кола.

3.3 Порівняльне тестування нейромережевих моделей прогнозування

Для програмної реалізації експериментальної частини дослідження було обрано мову програмування Python. Розробка та тестування моделей здійснювалися у середовищі Google Colab з використанням хмарних графічних прискорювачів (GPU).

У якості інструменту проектування, навчання та оцінки архітектур рекурентних нейронних мереж було обрано TensorFlow з API Keras. Вибір даної бібліотеки було зроблено, спираючись на наявність модулів реалізації шарів РНМ, алгоритмів оптимізації та інших гіперпараметрів навчання нейронних мереж.

Додатково було використано бібліотеки NumPy та Matplotlib для математичних обчислень та візуалізації відповідно.

Для проведення чистого експерименту та забезпечення об’єктивного порівняльного аналізу, топологію усіх трьох досліджуваних мереж було

уніфіковано. Єдиною архітектурною компонентною, що відрізняється в моделях, є тип рекурентного шару: класична РНМ, LSTM та GRU. Даний підхід було обрано з метою дослідження впливу наявності вентилів та їх типів на ефективність прогнозування, виключивши потенційний вплив кількості шарів та параметрів оптимізації на різницю в якості роботи моделей.

Уніфікована базова архітектура тестових моделей складається з двох прихованих рекурентних шарів, по 32 нейрони на першому та 16 нейрони на другому шарі, одного прихованого повнозв'язного шара, який містить 8 нейронів з функцією активації ReLU, та вихідного повнозв'язного шара з 1 нейроном і лінійною активаційною функцією. На рисунку 3.2 представлено уніфіковану архітектуру досліджуваних моделей.

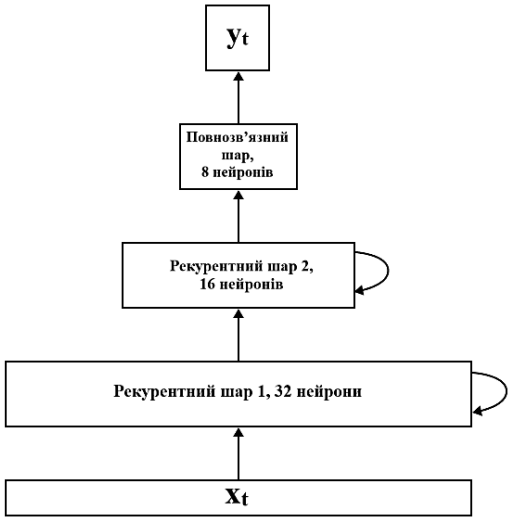


Рисунок 3.2 – Архітектура досліджуваних моделей

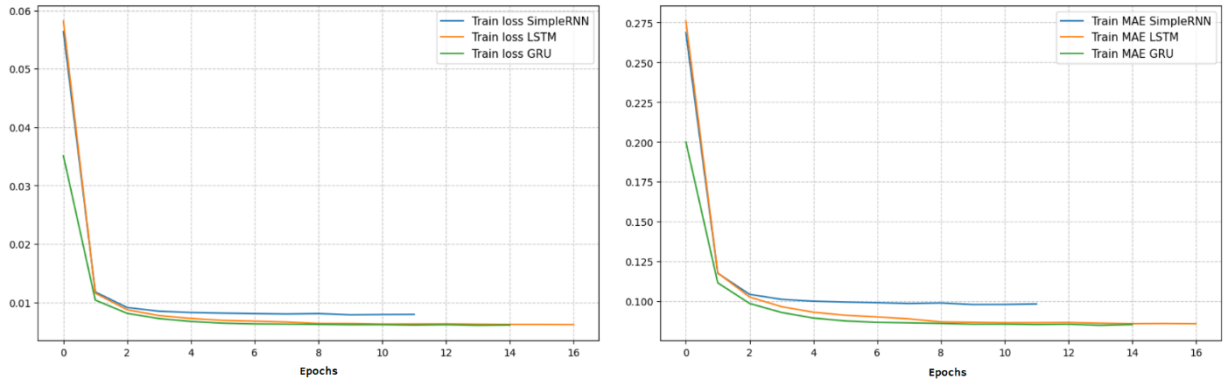
У ході навчання процес оптимізації вагових коефіцієнтів мереж здійснювався з використанням алгоритму адаптивної оптимізації Adam, який, згідно дослідженням емпірично більш ефективний за інші методи стохастичної оптимізації для різних архітектур ШНМ, включаючи глибокі [37]. Початкова швидкість навчання була встановлена рівною 0,001.

Так як прогнозування темпу деградації шин фактично є задачею регресії, для оцінки помилки роботи мережі було використано функцію, відповідну даному класу – функцію втрат Хьюбера [38]. Особливістю даної функції є те, що вона змінює свою форму в залежності від величини помилки: якщо помилка завелика – функція приймає форму середньої абсолютною похибки, а якщо її значення замалі – форму середньоквадратичної. Оцінка величини помилки та, внаслідок, зміна поведінки функції Хьюбера здійснюється відповідно до значення параметру k. У попередніх дослідженнях [39] було зроблено висновок про доцільність визначення значення параметру k шляхом медіанної оцінки сукупності. Так як вихідні дані містять декілька змінних, для знаходження k було застосовано медіану стандартних відхилень. Для вихідного набору даних $k = 0{,}2259$.

Для максимізації узагальнюючої здатності моделей та попередженню перенавчанню було використано додаткові техніки. Зокрема, після кожного рекурентного шару було інтегровано механізм регуляризації Dropout із коефіцієнтом рівним 0,3. Dropout вирішує проблему перенавчання шляхом випадкового видалення нейронів та зв’язків між ними в мережі, що запобігає коадаптації нейронів [40]. Додатково було використано механізм адаптивного зниження швидкості навчання, який автоматично знижував крок оптимізації у разі виходу функції втрат на плато або її зростання. Також було застосовано механізм ранньої зупинки – переривання процесу навчання, якщо метрики не демонструють покращень протягом 12 епох поспіль. З метою стабілізації процесу навчання та підвищенню стійкості мереж було додано шари пакетної нормалізації. Це дозволило мінімізувати ефект внутрішнього коваріативного зсуву та прискорити збіжність.

Для оцінки похибки навчання досліджуваних мереж під час навчання було використано середню абсолютну похибку (MAE).

На рисунку 3.3 представлено порівняльний графік функцій втрат та похибки навчання протягом навчання моделей.



а) б)

Рисунок 3.3 – Порівняльний графік а) функції втрат та б) похибки навчання

Виходячи з отриманих графіків, можна зробити висновок, що процес навчання всіх трьох досліджуваних мереж був стабільним – коливання функцій втрат та помилки відсутні.

Також можна побачити, що час навчання моделей був різним – найшвидше за всіх було навчено класичну РНМ (SimpleRNN), найдовше – LSTM. Дане явище зумовлено наявністю вентилів в LSTM та GRU, які, як зазначалося в п. 2.1, призводять до збільшення кількості вагових коефіцієнтів. У випадку LSTM кількість вентилів дорівнює 3 – це збільшує кількість параметрів мережі у 3^2 разів у найгіршому випадку, що призвело до найдовшого часу навчання. GRU має 2 вентилі – це зумовлює середній час навчання. Нейрони класичної РНМ не містять додаткових елементів, що призводить до найменшої кількості вагових коефіцієнтів у мережах даного типу та зумовлює найнижчий час процесу навчання.

Для оцінки точності прогнозування досліджуваних моделей, окрім MAE, було використано квадратний корінь із середньоквадратичної похибки (RMSE).

У таблиці 3.3 наведено результати оцінки якості прогнозування побудованих моделей.

Таблиця 3.3 – Результати оцінки роботи досліджуваних моделей

Модель / Оцінка	SimpleRNN	LSTM	GRU
MAE	0,932	0,745	0,749
RMSE	1,263	1,004	1,016

Можна побачити, що найнижче значення помилки прогнозування темпу деградації на тестовій вибірці показала LSTM-модель: MAE склала ~0,7 с. Отримані результати можна вважати цілком прийнятними в контексті перегонів, адже на час кола, а відповідно й на темп деградації, може впливати безліч ситуаційних факторів, зокрема, виникнення поривів вітру, потрапляння машини в трафік, дію брудного повітря, активацію системи DRS, вибір неоптимальної гоночної траєкторії та помилки пілота.

На рисунку 3.4 представлено порівняння реального темпу деградації та прогнозованого досліджуваними моделями.

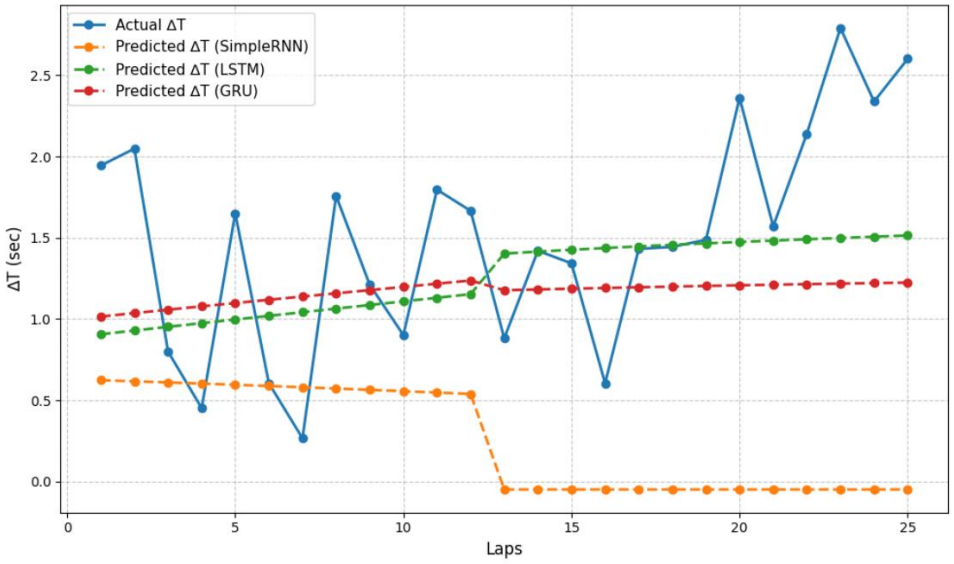


Рисунок 3.4 – Порівняльний графік реального темпу деградації та прогнозованого досліджуваними моделями

На отриманих графіках можна побачити, що прогноз, якій надала LSTM-модель, є найкращим: він формує плавну криву, яка відображає

кумулятивний тренд деградації шин. GRU є найбільш близькою за результатами, проте, занадто сильно занижує прогноз, починаючи з 12 кола.

На графіку, який відповідає прогнозу класичної РНМ (SimpleRNN), можна побачити, що прогнозоване значення падає нижче нульової позначки на 13 часовому кроці тестової послідовності. Можна припустити, що такий результат є наслідком проблеми зникання градієнту, характерної для класичних РНМ (описано у п. 2.1). Внаслідок цього нейронна мережа не здатна ефективно вивчати довгострокові залежності, що проявляється у зниженні чутливості моделі до змін у вхідних сигналах на пізніших кроках послідовності.

Для визначення якості роботи розроблених моделей в умовах стохастичної невизначеності, притаманної «Формули-1», було проведено експеримент з оцінки їхньої робастності до аномальних збурень у телеметричних даних. Задачею даного тесту було дослідження інерційності прогнозу та здатності розроблених моделей адаптуватися до раптових змін даних, викликаних зовнішніми факторами.

Експеримент полягав у введені штучного збурення, що імітував увімкнення режиму «віртуальної машини безпеки», до чистого відрізку гонки у 25 кіл. Збурення охоплювало 3 кола, під час яких було понижено температуру (для імітації охолодження шин), знижено телеметричні показники, що визначають вплив розгону, гальмування та поворотів на шини (для імітації зниження темпу). Після завершення відрізку зі збуреннями, дані було повернено до граничних гоночних значень.

Особливістю режиму VSC є те, що під час нього силове навантаження та термодинамічний ефект на шину знижується, що призводить до зниження темпів її зносу та деградації, однак, не зупиняють ці процеси повністю. Очікуваний ефект даного режиму полягає у незначному зниженню деградації безпосередньо під час VSC та поступовому поверненню її темпу до початкового тренду.

У ході експерименту кожною досліджуваною моделлю було зроблено два прогнози темпу деградації: перший – на чистому 25-коловому відрізку, другий – на 25-коловому відрізку зі збуреннями на 10–12 колах. На рисунку 3.5 представлено результати проведення експерименту на моделі, що відповідає класичній архітектурі PHM.

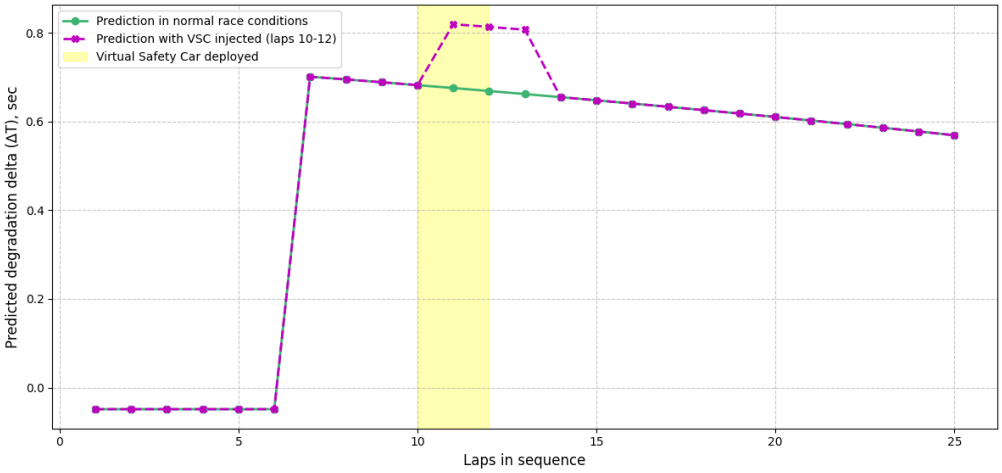


Рисунок 3.5 – Результати експерименту на моделі з класичній PHM

На рисунку 3.6 представлено результати проведення експерименту на моделі, що відповідає архітектурі LSTM.

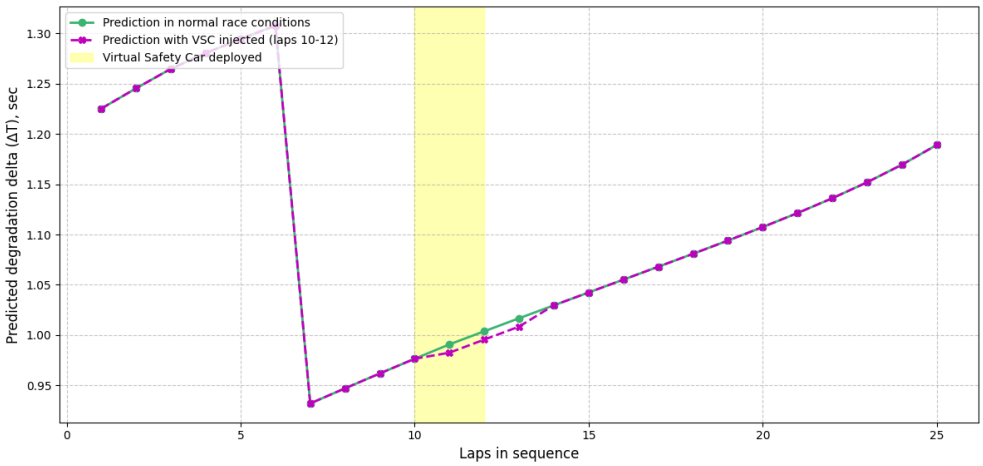


Рисунок 3.6 – Результати експерименту на моделі з LSTM-блоками

На рисунку 3.7 представлено результати проведення експерименту на моделі, що відповідає архітектурі GRU.

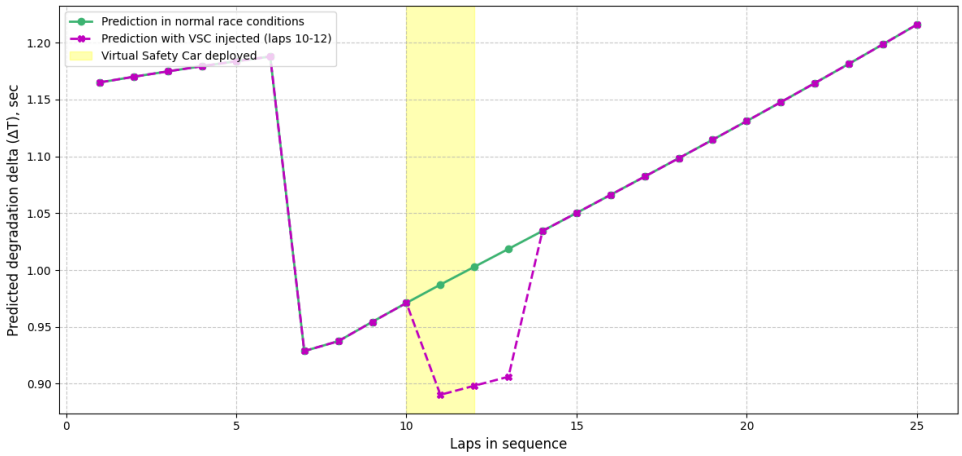


Рисунок 3.7 – Результати експерименту на моделі з GRU-блоками

Аналізуючи отримані результати експерименту, можна зробити висновок про те, що класична PHM (SimpleRNN) показала найгірші результати. Під час режиму VSC прогноз показав певне збільшення темпу деградації, що повністю суперечить очікуваному ефекту та фізиці процесу сповільнення машини. Після закінчення режиму VSC прогноз різко повернувся до початкової лінії тренду. Проте варто зазначити, що ця лінія тренду є спадною. На основі цього можна стверджувати про низьку узагальнюючу здатність даної моделі, а також її схильність до накопичення та розповсюдження по всій послідовності помилок, які виникли на більш ранніх етапах, що призводить до викривлення загальної лінії прогнозу.

Окрім цього, можна побачити, що на початкових часових кроках (гоночних колах) тестової послідовності, прогнозовані значення темпу деградації є близькими до 0. Такі результати зумовлені ініціалізацією початкового стану мережі (для рекурентних мереж початковий стан h_0 дорівнює 0) та відсутністю механізму керування інформаційними потоками. Це призвело до періоду розігріву моделі, який тривав приблизно 6 часових

кроків, поки вікно спостереження частково не заповнилося реальними телеметричними даними.

На відміну від класичної PHM, LSTM та GRU на початкових часових кроках тестової послідовності надавали стабільний та відмінних від 0 прогноз. Такий ефект зумовлений наявністю вентилів в блоках обох моделей: навіть за наявності нулів при ініціалізації початкового стану, вентилі блокували цей шум, прогнозуючи завчений базовий рівень деградації шин. Після акумулювання даних телеметрії поточного відрізку гонки, моделі скорегували попередній прогноз під фактичний темп деградації, що зумовлює різке падіння лінії прогнозу на 7 кроці.

Однак, подібно до класичної PHM, модель GRU продемонструвала надмірну чутливість до введення збурень до вхідних даних. На 11 колі, після початку VSC, модель продемонструвала різке падіння прогнозу, а після завершення VSC, прогноз деградації різко повернувся до початкової лінії тренду. Можна припустити, що такі результати зумовлені внутрішньою будовою нейронів GRU, зокрема роботою вентилів уточнення та скидання. Через те, що внутрішній стан є одночасно і безпосереднім виходом моделі, вентиль уточнення безперешкодно оновлює стан і прогноз моделі на основі збурених даних, що призвело до падіння темпу деградації на 11 часовому кроці послідовності. Після повернення вхідних даних до граничних гоночних значень, вентиль скидання відкидає попередній стан, який було сформовано на основі нерелевантних показниках телеметрії, повертаючи прогноз до загального тренду на 14 колі.

Найкращі результати в ході експерименту із введення аномалій у вхідні дані показала LSTM-мережа. Під час моделювання режиму VSC модель, на відміну від класичної PHM та GRU, показала незначне пониження темпу деградації, уникнувши різкого та значного падіння. Після закінчення періоду, протягом якого діяв режим VSC, та повернення телеметричних показників до граничних гоночних значень, прогноз темпу деградації було плавно відновлено до початкової лінії тренду. Такий

результат було отримано за рахунок особливості внутрішньої будови блоку LSTM: наявність вентилів входу та виходу ізолюють внутрішній стан комірки пам’яті від вхідних збурень та від їхнього розповсюдження на остаточний прогноз, а додавання вентилю забування виконує захисну функцію, нівелюючи ризик накопичення нерелевантної інформації.

Таким чином, спираючись на проведений аналіз результатів, отриманих в ході як стандартного, так і адверсаріального тестувань, можна зробити висновок, що архітектура рекурентних нейронних мереж LSTM є найбільш ефективною для вирішення задачі прогнозування темпу деградації гоночних шин «Формули-1». Вона продемонструвала найнижче значення помилки на тестовій вибірці, стабільність прогнозування, здатність до формування кумулятивної лінії тренду темпу деградації, а також продемонструвала робастність до аномальних збурень, викликаних стохастичними факторами, притаманними даній серії перегонів.

3.4 Оцінка ефективності алгоритмів послідовного прийняття рішень

Програмну реалізацію алгоритмів прийняття рішень, а також розробку самого симуляційного середовища, було виконано за допомогою мови програмування Python у середовищі Google Colab.

У якості інструментів розробки було використано бібліотеки Pandas (для роботи з даними та їх обробки), NumPy, tf2onnx та ONNX Runtime (для реалізації алгоритму Monte Carlo Tree Search), Gymnasium та Stable Baselines3 (для проектування та навчання агента PPO).

Як було визначено в п. 2.2, як MCTS, так і будь-який метод навчання з підкріпленням працюють з задачами сформульованими у вигляді марковського процесу вирішування, згідно якого розроблена система потребує чітко визначених меж того, яку інформацію агент отримує від середовища на кожному часовому кроці, та які дії він може обирати.

Для забезпечення агента контекстом для прийняття рішень, поточний стан середовища (гонки) на кожному кроці (гоночному колі) було закодовано у вигляді багатовимірного вектора ознак, який складається з наступних ключових метрик:

– номера поточного кола: даний параметр необхідний для розуміння горизонту планування;

– поточного комплекту шин (Soft, Medium або Hard);

– переліку використаних комплектів шин: цей параметр необхідний для відслідковування виконання вимоги спортивного регламенту щодо обов'язкового використання двох різних типів шин;

– віку шин: кількості повних кіл, пройдених на поточному комплекті від моменту останнього піт-стопу;

– наявності або відсутності режиму «машини безпеки» або «віртуальної машини безпеки»: ці параметри необхідні для правильного розрахунку гоночного кола та розуміння горизонту планування;

– переліку доступних комплектів шин: так як кількість шин, які надаються командам на гонку є чітко визначеними правилами «Формули-1», даний параметр обмежує неконтрольоване та нераціональне використання шин, а також формує розуміння горизонту планування;

– очікуваного часу перегонового кола: очікуваний час кола з урахуванням прогнозу темпу деградації.

Оскільки прийняття рішень у рамках формування стратегії має категоріальний характер, простір дій було спроектовано як дискретний, і на кожному часовому кроці можливий вибір однієї з чотирьох наступних дій:

– залишитися на трасі (Stay out): агент приймає рішення залишитися на поточному комплекті шин;

– проведення піт-стопу (Soft): агент приймає рішення про проведення піт-стопу для зміни шин на м'який комплект;

– проведення піт-стопу (Medium): агент приймає рішення про проведення піт-стопу для зміни шин на комплект з середньої за жорсткістю сумішшю гуми;

– проведення піт-стопу (Hard): агент приймає рішення про проведення піт-стопу для зміни шин на найбільш жорсткий комплект.

Після вибору однієї з чотирьох дій, середовище розраховує фактичний час гоночного кола, який є головною метрикою симуляції. Цей розрахунок залежить від обраної дії та стохастичних факторів:

– залишитися на трасі (Stay out): якщо агент обирає продовжити рух на поточному комплекті шин, час гоночного кола формується як сума базового часу (оптимального темпу без зносу, розраховується на основі історичних даних для кожної траси окремо) та прогнозованого темпу деградації, який надається побудованою LSTM-моделлю. Вік поточного комплекту збільшується на 1 коло;

– проведення піт-стопу: якщо агент вирішує зробити зупинку, до загального часу додається константа витрати часу на піт-лейні (яка залежить від траси та обчислюється на основі історичних даних). Після цього вік шин скидається до 0, параметр поточного комплекту змінюється відповідно обраному, новий тип шин додається до переліку використаних;

– виїзд машини безпеки або ввімкнення режиму «віртуальної машини безпеки»: якщо на поточному колі вводить режими SC або VSC, час гоночного кола помножується на відповідну константу. Якщо агент обирає дію проведення піт-стопу під час дії даних режимів, константа витрати часу на піт-лейні змінюється відповідно до режиму (піт-стоп під час SC займає найменший час, під час VSC – дещо більший; залежать від траси та обчислюються на основі історичних даних);

– фактор трафіку: для симуляції гоночного середовища (боротьба за позицію, обгін колових) до часу кола додається стохастична змінна, яка відповідає часу, витраченому в трафіку. Ця величина генерується

випадковим чином із заданого розподілу ймовірностей на кожному часовому кроці.

Так як мета алгоритмів MCTS та PPO полягає в максимізації кумулятивної винагороди, а мета гоночної стратегії – у мінімізації витраченого часу, функція винагороди формалізується як від'ємне значення розрахованого часу кола:

$$R_t = -T. \tag{3.8}$$

Додатково до базової функції винагороди було визначено систему штрафів, яка активується у випадках, коли дії агента суперечать логіці та регламенту «Формули-1»:

– штраф за використання лише одного типу шин: відповідно правил гонщик має використати мінімум два різних типи шин за гонку;

– штраф за перевищення максимального терміну експлуатації комплекту шин: перед початком Гран-прі компанія-постачальник шин Pirelli визначає приблизний термін служби шин, його перевищення можливе, проте, не є бажаним;

– штраф за надмірну кількість піт-стопів: запобігає ситуації, коли агент проводить піт-стопи на кожному колі гонки; максимальне число зупинок, як може зробити агент дорівнює 3 (спираючись на попередні дослідження [39], в середньому гонщики не роблять більше 3 піт-стопів за гонку).

Відповідно, цільова функція винагороди 3.8 на кожному часовому кроці t набуває вигляду:

$$R_t = \begin{cases} -\left(T_b + \Delta T + T_{tr} + T_{pl} \cdot \mathbb{I}_p\right) + \sum P \text{, якщо SC та VSC відсутні} \\ -\left(m_{VSC} \cdot \left(T_b + \Delta T + T_{tr} + T_{pl} \cdot \mathbb{I}_p\right)\right) + \sum P \text{, якщо VSC} \\ -\left(m_{SC} \cdot \left(T_b + \Delta T + T_{tr} + T_{pl} \cdot \mathbb{I}_p\right)\right) + \sum P \text{, якщо SC} \end{cases}, \tag{3.9}$$

де T_b – базовий час кола;

ΔT – прогнозований темп деградації;

T_{tr} – час, витрачений в трафіку;

T_{pl} – час витрачений на піт-стоп;

$\mathbb{I}_p$ – індикаторна функція, дорівнює 1 у разі проведення піт-стопу та 0 у протилежному випадку;

P – штраф;

m_{VSC}, m_{SC} – константи-модифікатори темпу, $m_{VSC} = 1{,}3$, $m_{SC} = 1{,}45$.

Тестування алгоритмів проводилася на основі 5 різних етапів «Формули-1»: Китаю, Австрії, Угорщини, Азербайджану та Бахрейну. Даний вибір трас забезпечує репрезентативність, оскільки охоплює різні види трас:

– траса Гран-прі Китаю є стаціонарним автодромом, її конфігурація характеризується великою кількістю поворотів (зокрема повільних), має високі оцінки дії бічних сил, навантажень на шини та еволюції дорожнього покриття;

– траса етапу в Австрії також є стаціонарним автодромом, однак, вона має меншу кількість поворотів, невеликий піт-лейн, а оцінки рівня навантаження на машину та шини є помірними за своєю силою;

– траса Гран-прі Угорщини також є стаціонарним автодромом, але дуже вузьким, за рахунок чого обгони є дуже складними, має велику кількість повільних поворотів, оцінки рівня притискної сили, зчеплення та еволюції асфальту є дуже високими;

– траса Гран-прі Азербайджану є міською, що сильно позначається на якості дорожнього покриття та його еволюції, її конфігурація характеризується великою кількістю довгих прямих та крутих поворотів, через що оцінка складності гальмування є дуже високою;

– траса етапу в Бахрейні є стаціонарним автодромом, містить велику кількість прямих та поворотів, однак, асфальтне покриття має дуже високий рівень абразивності.

Для кожної з 5 трас було проведено 100 незалежних симуляцій гонки, що дозволило нівелювати вплив випадкових збурень середовища та отримати статистично значимі результати.

Оцінка якості сформованих стратегій здійснювалася індивідуально для кожної траси з використанням наступних критеріїв:

– середній час гонки та стандартне відхилення: розраховувалися окремо для реальної стратегії, стратегії MCTS та стратегії PPO за результатами 100 симуляцій;

– частка перемог: попарне порівняння сформованих стратегій за часом (реальна стратегія та MCTS, реальна стратегія та PPO, MCTS та PPO).

У таблиці 3.4 наведено зведені результати 100 симуляцій гонок для кожної тестової траси. Окрім вищеописаних критеріїв оцінювання, для підтвердження достовірності переваги агента PPO над MCTS, було включено результати розрахунку T-статистики та p-значення.

Таблиця 3.4 – Зведені результати симуляцій

Траса	Стратегія	Час гонки (середній час ± σ), с	Частка перемог (/Реальна), %	Частка перемог (/MCTS), %	T-статистика	p-значення
Австрія	Реальна	5508,58 ± 108,75	-	-	-	-
	MCTS	5497,27 ± 107,27	59	-	-	-
	PPO	5229,98 ± 102,01	100	100	-18,058	<0,001
Бахрейн	Реальна	6176,04 ± 199,47	-	-	-	-
	MCTS	6146,12 ± 202,82	64	-	-	-
	PPO	6092,54 ± 212,23	95	83	-1,8252	0,049
Азербайджан	Реальна	5993,25 ± 200,24	-	-	-	-
	MCTS	5983,76 ± 202,71	59	-	-	-
	PPO	5897,9 ± 214,62	97	95	-2,893	0,004
Китай	Реальна	6195,31 ± 164,32	-	-	-	-
	MCTS	6177,24 ± 173,71	65	-	-	-
	PPO	6064,37 ± 185,83	100	99	-4,437	<0,001
Угорщина	Реальна	6127,51 ± 125,66	-	-	-	-
	MCTS	6125,22 ± 126,82	60	-	-	-
	PPO	6098,42 ± 126,7	97	94,7	-2,589	0,01

Для аналізу часової ефективності та стабільності досліджуваних алгоритмів було побудовано діаграму розмаху, яку представлено на рисунку 3.8.

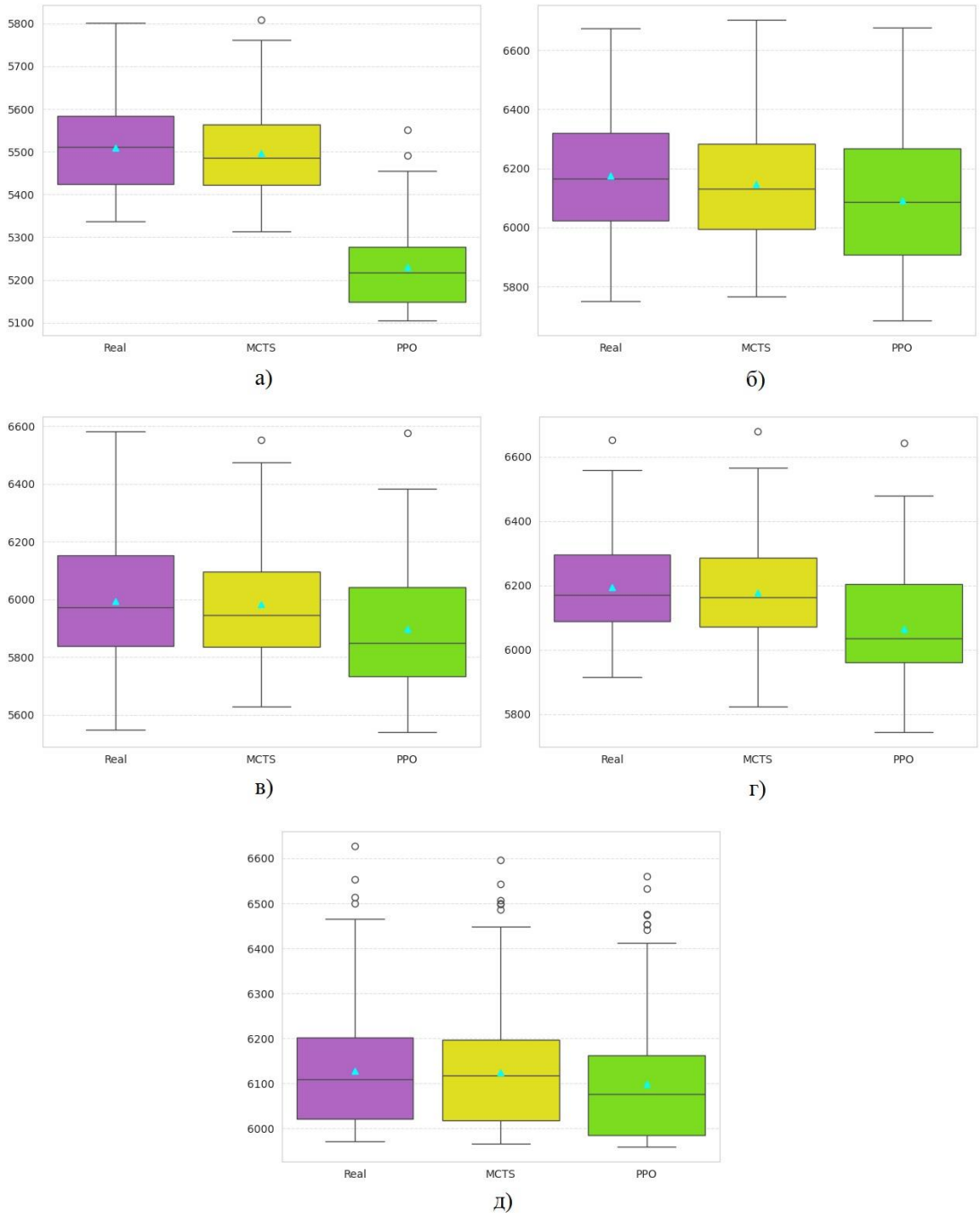


Рисунок 3.8 – Діаграми розмаху сумарного часу за кожної стратегією на тестових трасах Гран-прі а) Австрії, б) Бахрейну, г) Азербайджана, в) Китаю та д) Угорщини

Кожен блок діаграми візуалізує розподіл сумарного часу проходження дистанції гонки за умови впровадження трьох стратегій (реальної, сформованих MCTS та PPO) на основі вибірки зі 100 незалежних гоночних симуляцій. Вісь ординат відповідає загальному часу гонки. Окремим маркером синього кольору було позначено розрахований середній час.

Як можна побачити на наведених графіках, обидва агенти (MCTS та PPO) демонструють перевагу на всіх досліджуваних автодромах. І медіанний час (горизонтальна лінія всередині блоку), і середній час проходження дистанцій гонок зі сформованими стратегіями агентами MCTS та PPO є нижчими за показники реальної історичної стратегії.

Також можна спостерігати, що на окремих трасах, результати алгоритму MCTS мають менший міжквартильний розмах (меншу висоту блоків). Такий розподіл відображає архітектурну схильність MCTS приймати надмірно консервативні рішення через переоцінку ризиків під час симуляцій. Це призводить до вибору безпечних, але повільних стратегій.

Дещо більший розмах результатів PPO є наслідком його адаптивності. Замість використання певного безпечного шаблону стратегії, агент підлаштовується під конкретні умови та стохастичні фактори, які виникають піл час гонки.

Однак важливо зазначити, що розподіли результатів агента PPO знаходиться суттєво нижче, тобто в зоні кращого часу, порівняно з медіанним часом як реальної стратегії, так і стратегій, сформованих MCTS.

Для порівняльного аналізу загальної ефективності розроблених агентів було побудовано груповану стовпчикову диграму, представлену на рисунку 3.9, яка візуалізує частку перемог у результаті попарних порівнянь стратегій для кожної тестової траси.

Як можна побачити, на всіх трасах обидва агенти (MCTS та PPO) демонструють частку перемог вищу за 60% при порівнянні з реальними стратегіями переможців останніх гонок. Агент PPO має найвищі значення

частки перемог над реальною стратегією, яка досягає близько 100% для трас Гран-прі Китаю та Австрії.

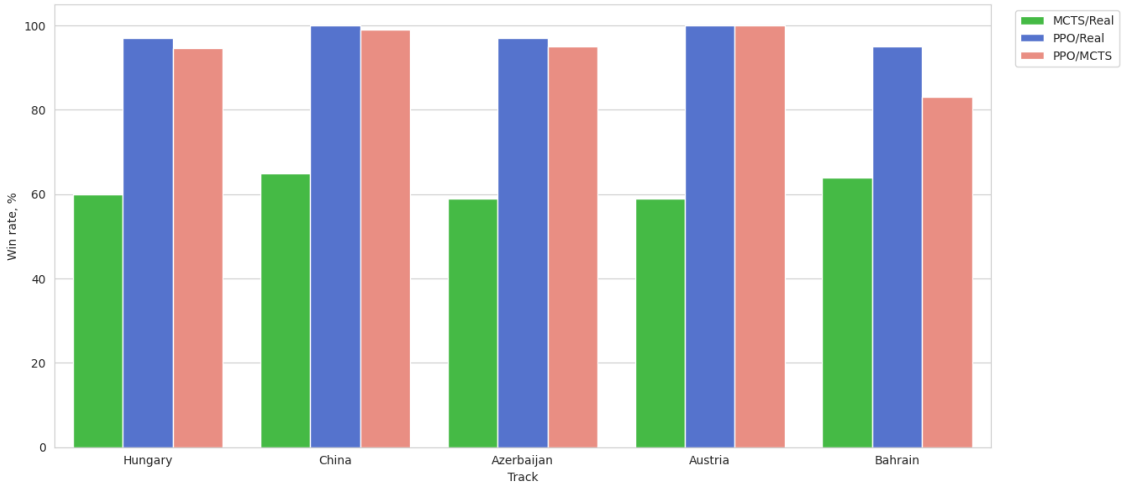


Рисунок 3.9 – Візуалізація частки перемог у результаті попарних порівнянь стратегій для кожної тестової траси

Водночас при порівнянні обох агентів, PPO демонструє перевагу над MCTS на кожній з тестових трас (частка перемог PPO над MCTS становить 80–99%). Можна припустити, що такі результати обумовлені наявністю стохастичних факторів, які впливають на перебіг гонки, та тим, як кожен з алгоритмів реагує на них. На етапі симуляції в алгоритмі MCTS стохастичні фактори створюють шум, який призводить до хибних оцінок гілок дерева рішень. Через це, агент або починає приймати неоптимальні рішення, або дотримується консервативної стратегії, яка є неефективною в контексті виникаючих факторів. Як результат, загальний час збільшується. З іншого боку, агент PPO при виборі дій спирається на оптимізовану параметризовану стратегію. На етапі навчання модель сформувала уявлення про динаміку середовища та вивчила глобальні закономірності, що дозволяє їй приймати рішення попри виникнення випадкових та нетипових подій.

Для підтвердження статистичної значущості результатів тестів було застосовано t-критерій Стьюдента (таблиця 3.4). За базову гіпотезу було

прийнято припущення, що різниця у середньому часі проведення гонки між алгоритмами є наслідком випадковості стохастичного середовища.

Отримані p-значення для всіх попарних порівнянь агентів PPO та MCTS на кожній з тестових трас є меншим за поріг (p=0,05). На основі цього базове припущення про випадковість було відхилено. Розраховані метрики є статистично значущими, наявна перевага агента PPO відображає реальну ефективність підходу.

На рисунку 3.10 представлено порівняння реальної стратегії та стратегій сформованих агентами MCTS та PPO для траси етапу в Австрії.

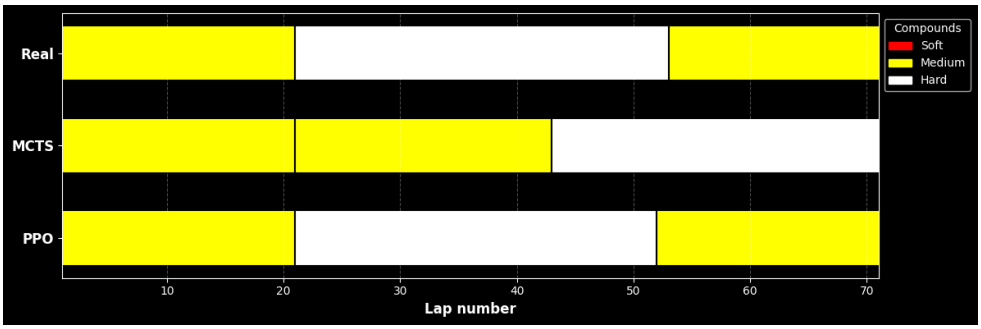


Рисунок 3.10 – Порівняльна діаграма стратегій

Таким чином, спираючись на проведений аналіз отриманих результатів, можна зробити висновок про ефективність запропонованого підходу на основі алгоритму PPO для вирішення задач підтримки прийняття рішень. Незважаючи на те, що евристичні підходи демонструють кращі результати, ніж статично сплановані рішення, вони часто страждають від шуму в середовищі, викликаному стохастичними факторами. Це призводить до вибору неоптимальних дій та консервативних стратегій, що значною мірою вливає на загальну продуктивність. Завдяки здатності нейромережевого підходу до глобального узагальнення станів, алгоритм PPO показав кращу адаптивність до випадкових збурень середовища. Це дозволило мінімізувати загальний час гонки та забезпечити найвищий відсоток успішних рішень.

ВИСНОВКИ

У результаті виконання кваліфікаційної роботи було проведено комплексне дослідження методів нейромережевого моделювання та евристичного пошуку для задач підтримки прийняття рішень в умовах стохастичної невизначеності.

У процесі виконання роботи було отримано наступні результати:

– було вивчено засади системи регулювання різних аспектів автомобільних перегонів серії «Формули-1» та визначено, що регуляторні документи містять обмеження щодо механізмів збору та передачі даних від машини до команди. Було розглянуто, яким чином фізичні та хімічні властивості матеріалу, з якого виготовляють шини, впливають на їхнє зчеплення, знос та деградацію, а також визначено фактори, які впливають на ці процеси. Було детально досліджено стохастичні чинники (виїзд машини безпеки, помилки пілота тощо) та те, яким чином вони впливають на різні аспекти гонки (часові втрати, температуру шин). Також було проаналізовано наявні дослідження в автоспорті та визначено, що вирішення задачі визначення перегонової стратегії є актуальним напрямком наукових робіт;

– використовуючи наукову літературу, було проведено дослідження особливостей реалізації та навчання рекурентних нейронних мереж. У рамках дослідження було розглянуто різні архітектури РНМ (класичні, LSTM та GRU), вивчено їхні переваги та недоліки;

– було проведено детальний аналіз методів прийняття рішень, що базуються на різних алгоритмічних парадигмах: евристичному пошуку (MCTS) та нейромережевого на основі глибинного навчання з підкріпленням (PPO). Було розглянуто особливості їхньої реалізації, а також досліджено сильні та слабкі сторони;

– на основі результатів аналізу предметної галузі та визначених факторів, що впливають на темпи зносу та деградації, було сформовано вихідний простір ознак. Із використанням коефіцієнту рангової кореляції

Спірмена було проведено дослідження набору даних та детально проаналізовано визначені зв’язки між ознаками;

– було визначено, що через відсутність датчиків на шинах болідів «Формули-1», оцінка зносу та деградації шин виконується опосередковано, через моніторинг часу гоночного кола. Однак було зроблено висновок, що виміряний час гоночного кола не є об’єктивною оцінкою деградації через поступове згорання пального під час гонки та різну довжину трас. Для знаходження дійсного темпу деградації було запропоновано та використано різницю з гоночного кола з поправкою на пальне;

– спираючись на дослідження рекурентних нейронних мереж, було побудовано та навчено три моделі з різними типами рекурентних шарів. Для забезпечення об’єктивного аналізу топологія моделей була уніфікованою. Було проведено стандартне та адверсаріальне тестування досліджуваних моделей, за результатами яких було зроблено висновок про високу ефективність та робастність архітектури LSTM для вирішення задачі прогнозування темпу деградації гоночних шин. Середня абсолютна похибка на тестовій вибірці склала 0,7;

– на основі проведеного аналізу методів прийняття рішень було реалізовано досліджувані алгоритми MCTS та PPO. Ґрунтуючись на дослідженні предметної галузі, було реалізовано симуляційне середовище: визначено простір станів та дій, а також описано функцію винагороди. На основі п’яти різних за характеристиками трас було проведено порівняльний аналіз реалізованих агентів між собою та відносно історичних стратегій переможців. За результатами аналізу було зроблено висновок про перевагу алгоритму PPO при вирішенні задач підтримки прийняття рішень в умовах стохастичної невизначеності (частка перемог PPO в порівнянні з історичною стратегією для всіх тестових трас в середньому склала 97,8%, в порівнянні з MCTS – 94,34%).

Результати досліджень було апробовано у конференціях та представлено публікаціях [39], [41], [42], [43], [44].

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Section d: financial regulations (f1 teams). *Federation Internationale de l'Automobile*. URL: https://www.fia.com/system/files/documents/fia_2026_f1_regulations_-_section_d_financial_-_f1_teams_-_iss_05_-_2026-02-27.pdf (date of access: 08.05.2026).

2. Section c: technical regulations. *Federation Internationale de l'Automobile*. URL: https://www.fia.com/system/files/documents/fia_2026_f1_regulations_-_section_c_technical_-_iss_16_-_2026-02-27.pdf (date of access: 08.05.2026).

3. Section b: sporting regulations. *Federation Internationale de l'Automobile*. URL: https://www.fia.com/system/files/documents/fia_2026_f1_regulations_-_section_b_sporting_-_iss_05_-_2026-02-27.pdf (date of access: 08.05.2026).

4. F1. FIA issue statements on commission and power unit meetings. *F1 - The Official Home of Formula 1® Racing*. URL: https://www.formula1.com/en/latest/article/fia-issue-statements-after-f1-commission-and-power-unit-advisory-committee.vs2Is3TIMy1eT4aD1ArRG (date of access: 08.05.2026).

5. Liu J., Wang M., Yin H. A study of the friction characteristics of rubber thermo-mechanical coupling. *Polymers*. 2024. Vol. 16, no. 5. P. 596. URL: https://doi.org/10.3390/polym16050596 (date of access: 08.05.2026).

6. Thermo-Mechanical coupling analysis of inserts supporting run-flat tires under zero-pressure conditions / C. Xue et al. *Machines*. 2024. Vol. 12, no. 8. P. 578. URL: https://doi.org/10.3390/machines12080578 (date of access: 08.05.2026).

7. Pirelli. What's new with pirelli's 2019 formula 1 tyres?. *What's new with Pirelli's 2019 Formula 1 tyres?*. URL: https://press.pirelli.com/whats-new-with-pirellis-2019-formula-1-tyres/ (date of access: 08.05.2026).

8. How to prevent tyre degradation in Formula 1. *Motorsport Engineer Online Courses: Home*. URL: https://motorsportengineer.net/how-to-prevent-tyre-degradation-in-formula-1/ (date of access: 08.05.2026).

9. Pirelli. Formula 1 for dummies: how tyres are chosen for a gp?. *Pirelli Global: Discover our world | Pirelli*. URL: https://www.pirelli.com/global/en-ww/race/racingspot/formula-1/formula-1-for-dummies-the-choice-of-tyres-in-formula-1-53864/ (date of access: 08.05.2026).

10. Tyre management in Formula 1 racing. *Formula 1 Dictionary*. URL: https://www.formula1-dictionary.net/f1-tyre-management-racing/ (date of access: 08.05.2026).

11. F1. Need to know: facts, stats and trivia for Monaco. *F1 - The Official Home of Formula 1® Racing*. URL: https://www.formula1.com/en/latest/article/need-to-know-the-most-important-facts-stats-and-trivia-ahead-of-the-2025-monaco-grand-prix.7FbX7r8dNJZQF6jnyLnoJC (date of access: 08.05.2026).

12. F1. The most important stats ahead of the Hungarian GP. *F1 - The Official Home of Formula 1® Racing*. URL: https://www.formula1.com/en/latest/article/need-to-know-the-most-important-facts-stats-and-trivia-ahead-of-the-2025-hungarian.3MROLEbJtsLuQUIaFvYnv0 (date of access: 08.05.2026).

13. F1. Facts, stats and trivia ahead of the 2026 Japanese GP. *F1 - The Official Home of Formula 1® Racing*. URL: https://www.formula1.com/en/latest/article/need-to-know-the-most-important-facts-stats-and-trivia-ahead-of-the-2026-japanese-grand-prix.59xoCL07qjNCWPVzBJOmK2 (date of access: 08.05.2026).

14. F1. Monday morning debrief: how Leclerc's crash caused a domino effect that shaped Sainz's French Grand Prix charge | Formula 1®. *F1 - The Official Home of Formula 1® Racing*. URL: https://www.formula1.com/en/latest/article/monday-morning-debrief-

how-leclercs-crash-caused-a-domino-effect-that-shaped.1SpXv2Tjm1YuWIiuIW8yCV (date of access: 08.05.2026).

15. What's the difference between slipstream and dirty air?. *raceteq*. URL: https://www.raceteq.com/articles/2025/07/slipstream-vs-dirty-air-explained (date of access: 08.05.2026).

16. F1. Monday morning debrief: undercut vs overcut – why tyre strategy was so finely poised in Monaco, and why Hamilton couldn't make either work | Formula 1®. *F1 - The Official Home of Formula 1® Racing*. URL: https://www.formula1.com/en/latest/article/undercut-vs-overcut-why-tyre-strategy-was-so-finely-poised-in-monaco-and-why.1YYMDkEBnFols8bDWtSXiz (date of access: 08.05.2026).

17. F1 race strategy: tyres, pit stops, and tactics. *Formula 1 Dictionary*. URL: https://www.formula1-dictionary.net/f1-race-strategy/ (date of access: 08.05.2026).

18. Bekker J., Lotz W. Planning Formula One race strategies using discrete-event simulation. *Journal of the operational research society*. 2009. Vol. 60, no. 7. P. 952–961. URL: https://doi.org/10.1057/palgrave.jors.2602626 (date of access: 08.05.2026).

19. Application of monte carlo methods to consider probabilistic effects in a race simulation for circuit motorsport / A. Heilmeier et al. *Applied sciences*. 2020. Vol. 10, no. 12. P. 4229. URL: https://doi.org/10.3390/app10124229 (date of access: 08.05.2026).

20. A gradient boosting analytics framework for predictive pit stop decision making in formula one racing / T. Ametsbichler et al. 2026. 36 p. (Preprint). URL: https://doi.org/10.2139/ssrn.6271405 (date of access: 08.05.2026).

21. Virtual strategy engineer: using artificial neural networks for making race strategy decisions in circuit motorsport / A. Heilmeier et al. *Applied sciences*. 2020. Vol. 10, no. 21. P. 7805. URL: https://doi.org/10.3390/app10217805 (date of access: 08.05.2026).

22. Руденко О. Г., Бодянський Є. В. Штучні нейронні мережі : навч. посіб. Харків : Компанія СМІТ, 2006. 404 с.

23. Гітіс В. Б. Нейромережні технології : навч. посіб. Краматорськ : ДДМА, 2021.

24. Cybenko G. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals, and systems*. 1992. Vol. 5, no. 4. P. 455. URL: https://doi.org/10.1007/bf02134016 (date of access: 08.05.2026).

25. Zhang A., Lipton Z. C. Dive into deep learning. Cambridge University Press, 2023.

26. Гітіс Т., Підгора Є., Гітіс І. Прогнозування динаміки попиту та пропозиції на ринку праці україни. *Herald of Khmelnytskyi National University. Economic sciences*. 2025. Т. 340, № 2. С. 68–72. URL: https://doi.org/10.31891/2307-5740-2025-340-10.

27. Hochreiter S., Schmidhuber J. Long Short-Term Memory. *Neural computation*. 1997. Vol. 9, no. 8. P. 1735–1780. URL: https://doi.org/10.1162/neco.1997.9.8.1735 (date of access: 08.05.2026).

28. Gers F. A., Schmidhuber J., Cummins F. Learning to forget: continual prediction with LSTM. *Neural computation*. 2000. Vol. 12, no. 10. P. 2451–2471. URL: https://doi.org/10.1162/089976600300015015 (date of access: 08.05.2026).

29. Learning representations using RNN encoder-decoder for edge security control / W. Guo et al. *Computational intelligence and neuroscience*. 2022. Vol. 2022. P. 1–8. URL: https://doi.org/10.1155/2022/4199044 (date of access: 08.05.2026).

30. Monte Carlo Tree Search: a review of recent modifications and applications / M. Świechowski et al. *Artificial intelligence review*. 2022. URL: https://doi.org/10.1007/s10462-022-10228-y (date of access: 08.05.2026).

31. Murphy K. Reinforcement learning: an overview. 2025. 253 p. URL: https://doi.org/10.48550/arXiv.2412.0526 (date of access: 08.05.2026).

32. Sutton R. S., Barto A. G., Bach F. Reinforcement learning: an introduction. MIT Press, 2018. 552 p.

33. Proximal policy optimization algorithms / J. Schulman et al. URL: https://doi.org/10.48550/arXiv.1707.06347 (date of access: 08.05.2026).

34. Introduction to FastF1 – FastF1 (3.8.3). *FastF1*. URL: https://docs.fastf1.dev/index.html (date of access: 08.05.2026).

35. Pirelli global: discover our world | pirelli. *Pirelli*. URL: https://www.pirelli.com/global/en-ww/homepage/ (date of access: 08.05.2026).

36. Milliken W. F. Race car vehicle dynamics. Warrendale, PA, U.S.A : SAE International, 1995. 890 c.

37. Kingma D. P., Ba J. L. Adam: a method for stochastic optimization. *International conference on learning representations (ICLR)*, San Diego, CA, 7–9 May 2026. URL: https://arxiv.org/pdf/1412.6980 (date of access: 08.05.2026).

38. Huber P. J. Robust estimation of a location parameter. *The annals of mathematical statistics*. 1964. Vol. 35, no. 1. P. 73–101.

39. Gitis I., Gitis V. Neural network decision support system for formulating a racing team strategy. *Computer systems and information technologies*. 2025. No. 2. P. 145–153. URL: https://doi.org/10.31891/csit-2025-2-17 (date of access: 08.05.2026).

40. Ali S. Y., Maseeh H. Dropout: an effective approach to prevent neural networks from overfitting. *Asian journal of research in computer science*. 2025. Vol. 18, np. 2. P. 163–185. URL: https://doi.org/10.9734/ajrcos/2025/v18i2569 (date of access: 08.05.2026).

41. Гітіс І. В. Застосування штучних нейронних мереж для прогнозування стану шин в умовах автомобільних перегонів / І. В. Гітіс // Радіоелектроніка та молодь у XXI столітті : матеріали 30-го Міжнар. молодіж. форуму, 22–24 квітня 2026 р. – Харків : ХНУРЕ, 2026. – Т. 6 – С. 34–36.

42. Гітіс В.Б., Гітіс І.В. Розробка нейромережевої системи формування стратегії перегонової команди // Нейромережні технології та їх застосування НМТіЗ-2025: збірник наукових праць XXIV Міжнародної наукової конференції «Нейромережні технології та їх застосування НМТіЗ-2025» / за заг. ред. д-ра техн. наук, проф. С.В. Ковалевського. – Краматорськ-Тернопіль: ДДМА, 2025. – С. 233-238.

43. Гітіс І. В. Аналіз релевантності даних для системи формування стратегії автоперегонів. Автоматизація та комп'ютерно-інтегровані технології у виробництві та освіті: стан, досягнення, перспективи розвитку : матеріали Всеукр. науково-практ. Інтернет-конф., м. Черкаси, 11–17 берез. 2024 р. С. 354–355. URL: https://conference.ikto.net/pub/akit_2024_11-17march_1.pdf (дата звернення: 08.05.2026).

44. Gitis I. Construction of a neural network decision support system for the formation of a racing strategy. Штучний інтелект: досягнення, виклики та ризики : матеріали Міжнар. наук. конф., м. Київ, 15–16 берез. 2024 р. P. 400–403. URL: https://www.ipai.net.ua/docs/ai_conf_15.03.2024.pdf (date of access: 08.05.2026).