

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

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

Магістр

(назва освітнього ступеня)

на тему: Використання нейронних мереж для дослідження процесу
газоспоживання міста Тернопіль з врахуванням метеофакторів

Виконав: студент _____ 6 курсу, групи СНІМ-61
спеціальності _____

122 Комп'ютерні науки

(шифр і назва спеціальності)

	_____	Скуржанський Т.А.
	(підпис)	(прізвище та ініціали)
Керівник	_____	доц. Мацюк О.В.
	(підпис)	(прізвище та ініціали)
Нормоконтроль	_____	доц. Мацюк О.В.
	(підпис)	(прізвище та ініціали)
Завідувач кафедри	_____	доц. Боднарчук І.О.
	(підпис)	(прізвище та ініціали)
Рецензент	_____	доц. Тиш Є.В.
	(підпис)	(прізвище та ініціали)

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

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра комп'ютерних наук

(повна назва кафедри)

ЗАТВЕРДЖУЮ

Завідувач кафедри

доц. Боднарчук І.О.

(підпис)

(прізвище та ініціали)

« »

20__ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня

Магістр

(назва освітнього ступеня)

за спеціальністю

122 Комп'ютерні науки

(шифр і назва спеціальності)

студенту

Скуржанському Тарасу Андрійовичу

1. Тема роботи

Використання нейронних мереж для дослідження процесу газоспоживання міста Тернопіль з врахуванням метеофакторів

Керівник роботи

Мацюк Олександр Васильович, к.т.н., доцент

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом по університету від «28» жовтня 2021 року № 4/7-909

2. Термін подання студентом роботи 20.05.2022

3. Вихідні дані до роботи

наукові літературні джерела

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

1 Аналіз досліджень процесу газоспоживання. 2 Аналіз існуючих нейронних мереж, методи машинного навчання для процесу газоспоживання. 3 Попереднє опрацювання статистичних даних газоспоживання та метеофакторів для міста Тернопіль. 4 Побудова моделі нейронної мережі процесу газоспоживання міста з врахуванням метеофакторів. 5 Охорона праці та безпека в надзвичайних ситуаціях.

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

1. Тема роботи. 2. Актуальність. 3. Мета, задачі дослідження. 4. Об'єкт, предмет дослідження наукова новизна, практичне значення роботи. 5. Поняття нейронної мережі..

6. Принципи роботи нейронних мереж. 7. Аналіз існуючих моделей нейронних мереж.

8. Мова програмування Python та бібліотеки, що використані в дослідженні. 9. Попередній аналіз даних газоспоживання. 10. Попередній аналіз даних метеофакторів. 11. Модель нейронної мережі одношаровий перцептрон. 12. Модель нейронної мережі багатошаровий перцептрон.

13. Модель нейронної мережі LSTM. 14. Значення прогнозувань в порівнянні з реальними даними. 14.. Основні результати дослідження

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	Дмитроца Л.П, к.т.н. доцент	01.05.22	17.05.22
Безпека в НС	Клепчик В.М.	01.05.22	19.05.22

7. Дата видачі завдання _____ 2021 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Термін виконання етапів роботи	Примітка
1	Затвердження теми кваліфікаційної роботи	28.10.21	Виконано
2	Аналіз літературних джерел	29.10–18.12.21	Виконано
3	Обґрунтування актуальності дослідження	18.12–29.12.21	Виконано
4	Аналіз предмету дослідження та предметної області	02.01–26.01.22	Виконано
5	Проведення дослідження методів та засобів аналітичного опрацювання даних	27.01–22.02.22	Виконано
6	Оформлення розділу «Аналіз досліджень процесу газоспоживання»	23.03–19.03.22	Виконано
7	Оформлення розділу «Аналіз існуючих нейронних мереж, методи машинного навчання для процесу газоспоживання»	20.03–08.04.22	Виконано
8	Оформлення розділу «Попереднє опрацювання статистичних даних газоспоживання та метеофакторів для міста Тернопіль»	09.04–18.04.22	Виконано
9	Оформлення розділу «Побудова моделі нейронної мережі процесу газоспоживання міста з врахуванням метеофакторів»	19.04–03.05.22	Виконано
10	Оформлення розділу «Охорона праці та безпека в надзвичайних ситуаціях»	04.05–15.05.22	Виконано
11	Нормоконтроль	11.05–15.05.22	Виконано
12	Попередній захист роботи	16.05.22	Виконано
13	Захист кваліфікаційної роботи	26.05.22	

Студент _____
(підпис)Скуржанський Т.А.

(прізвище та ініціали)Керівник роботи _____
(підпис)Мацюк О.В.

(прізвище та ініціали)

АНОТАЦІЯ

Використання нейронних мереж для дослідження процесу газоспоживання міста Тернопіль з врахуванням метеофакторів // Кваліфікаційна робота освітнього рівня «Магістр» // Скуржанський Тарас Андрійович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем та програмної інженерії, кафедра комп'ютерних наук, група СНм-61 // Тернопіль, 2022 // С. – 105, рис. – 52, слайдів – 20, додат. – 2, бібліогр. – 62.

Ключові слова: НЕЙРОННІ МЕРЕЖІ, ОДНОШАРОВИЙ І БАГАТОШАРОВИЙ ПЕРЦЕПТРОН, БІБЛІОТЕКИ ПРОГРАМУВАННЯ PYTHON, ПРОГНОЗ ГАЗОСПОЖИВАННЯ, МЕТЕОФАКТОРИ.

Кваліфікаційна робота присвячена розробці системи прогнозування газоспоживання міста Тернопіль за допомогою нейронних мереж на основі даних про газоспоживання за період з 2006 по 2019 рік з врахуванням метеофакторів. Дослідження спрямоване на оптимізацію процесу планування газоспоживання в м.Тернополі, як складової енергетичної безпеки держави та ефективного інструменту для прийняття важливих рішень в газоспоживанні за допомогою нейронних мереж, таких як одношаровий та багатошаровий перцептрон, LSTM.

На основі аналізу наукових джерел та проведених практичних досліджень з'ясовано, що застосування нейронних мереж є одним з найпоширеніших способів прогнозування газоспоживання, описано алгоритм аналізу та оптимального відбору метеофакторів, які використані в дослідженні; здійснено порівняння ефективності тренування нейронних мереж.

Результатом роботи стало напрацювання алгоритмів застосування та принципів відбору нейронних мереж і вхідних даних для прогнозування газоспоживання.

ANNOTATION

The use of neural networks to study the process of gas consumption of Ternopil taking into account meteorological factors // Qualification work of educational level "Master" // Skurzhanskyi Taras Andriyovych // Ternopil National Technical University named after Ivan Pulyuy, Faculty of Computer Information Systems and Software Engineering, Department of Computer Science of Yuter Sciences, SNNM-61 group // Ternopil, 2022 // C. - 105, fig. - 52, slides - 20, appendix. - 2, bibliogr. - 62.

Keywords: NEURAL NETWORKS, SINGLE-LAYER AND MULTILAYER PERCEPTRON, PYTHON PROGRAMMING LIBRARIES, GAS CONSUMPTION FORECAST, METEOFATORS

Qualification work is devoted to the development of a system for forecasting gas consumption in the city of Ternopil using neural networks based on data on gas consumption for the period from 2006 to 2019, taking into account meteorological factors. The research is aimed at optimizing the process of gas consumption planning in Ternopil, as a component of energy security of the state and an effective tool for making important decisions in gas consumption using neural networks such as single-layer and multilayer perceptron, LSTM.

Based on the analysis of scientific sources and practical research, it was found that the use of neural networks is one of the most common ways to predict gas consumption, described the algorithm of analysis and optimal selection of meteorological factors used in the study.

The result was the development of algorithms for application and principles of selection of neural networks and input data for forecasting gas consumption.

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ СКОРОЧЕНЬ І ТЕРМІНІВ

ARIMA – авторегресивна ковзна середня модель

НН – нейронні мережі.

РНН – рекурентна нейронна мережа

ПГ – прогноз газоспоживання.

МАРЕ – середня абсолютна помилка

ПР – прийняття рішень.

T – Температура повітря на висоті 2м над поверхнею землі

P_0 – Атмосферний тиск на рівні станції

P_a – Тенденція зміни атмосферного тиску за останні 3 години

U – Відносна вологість повітря на висоті 2м над землею в %

F_f – Швидкість вітру на висоті 10–12м над поверхністю землі , середня за 10хв період, одиниці вимірювання м/с

F_{f10} – максимальне значення пориву вітру на висоті 10–12м над поверхнею землі за 10хв період

F_{f3} – максимальне значення пориву вітру на висоті 10–12метрів над поверхнею землі в м/с

T_n – мінімальна температура повітря за попередній період (не більше 12 годин)

T_x – максимальна температура повітря за попередній період (не більше 12 годин)

V_v – Горизонтальна дальність оглядовості в км

T_d – Температура точки роси на висоті 2метра

R_{RR} – кількість опадів в мм

t_{RR} – кількість часу за який випала та кількість опадів

T_g – мінімальна температура поверхності землі за ніч

S_{SS} – висота снігового покриву в см

ЗМІСТ

ВСТУП.....	9
1 АНАЛІЗ ДОСЛІДЖЕНЬ ПРОЦЕСУ ГАЗОСПОЖИВАННЯ.....	11
1.1 Взаємозв'язок економічного розвитку держави та споживання енергії.....	11
1.2 Прогнозування в енергетиці, як засіб оптимізації планування, підвищення ефективності газоспоживання.....	14
1.3 Огляд наукових публікацій по тематиці дослідження з метою вивчення існуючих математичних підходів до аналізу процесу газоспоживання.	16
1.4 Висновки до першого розділу.....	27
2 АНАЛІЗ ІСНУЮЧИХ НЕЙРОННИХ МЕРЕЖ, МЕТОДИ МАШИННОГО НАВЧАННЯ ДЛЯ ПРОЦЕСУ ГАЗОСПОЖИВАННЯ.....	30
2.1 Поняття нейронної мережі.....	30
2.2 Принципи роботи нейронних мереж.....	30
2.3 Аналіз існуючих моделей нейронних мереж.....	32
2.3.1 Одношаровий та багатшаровий перцептрон.....	32
2.3.2 РНН.....	38
2.3.3 ЛСТМ.....	40
2.4 Висновки до другого розділу.....	47
3 ПОПЕРЕДНЄ ОПРАЦЮВАННЯ СТАТИСТИЧНИХ ДАНИХ ГАЗОСПОЖИВАННЯ ТА МЕТЕОФАКТОРІВ ДЛЯ МІСТА ТЕРНОПІЛЬ.....	48
3.1 Мова програмування Python та бібліотеки що будуть використані в дослідженні.....	48
3.2 Перелік метеофакторів, які будуть використані в дослідженні..	50
3.3 Попередній аналіз даних газоспоживання.....	51
3.4 Попередній аналіз даних метеофакторів.....	54
3.5 Висновки до третього розділу.....	58

4	ПОБУДОВА МОДЕЛІ НЕЙРОННОЇ МЕРЕЖІ ПРОЦЕСУ ГАЗОСПОЖИВАННЯ МІСТА З ВРАХУВАННЯМ МЕТЕОФАКТОРІВ.....	59
4.1	Одношаровий та багатошаровий перцептрон.....	59
4.2	ЛСТМ.....	61
4.3	Висновки до четвертого розділу.....	65
5	ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ.....	67
5.1	Охорона праці під час експлуатації газового обладнання.....	67
5.2	Підвищення стійкості роботи об'єктів господарської діяльності у воєнний час.....	72
5.3	Висновки до п'ятого розділу.....	76
	ВИСНОВКИ.....	77
	ПЕРЕЛІК ДЖЕРЕЛ.....	78
	ДОДАТКИ.....	85
	Додаток А.....	86
	Додаток Б.....	90

ВСТУП

Актуальність теми. Системне зростання споживання природного газу у світі, ЄС та в Україні, як одного з найважливіших джерел енергії, зробили ринок природного газу комерційно надзвичайно важливим, особливо в регіонах зі зниженим рівнем видобутку газу, або його відсутністю. Не менш вагомо на проблематику прогнозування газопостачання та раціонального використання природного газу вплинули і події останніх місяців – російсько – українська війна, внаслідок чого, як на геополітичному рівні, так і на місцевому рівні проблема прогнозування газоспоживання стала ще більш актуальною.

Для вирішення питання прогнозування газоспоживання в м.Тернополі, в якому відсутні власні запаси газу, перед нами постало ряд запитань – які методики обрати, які вхідні параметри використати в дослідженні, на який період здійснювати прогнозування, щоб отримати бажаний результат.

Проведений аналіз наукових джерел засвідчив – для прогнозування газопостачання дослідниками застосовувалися різноманітні методики з найрізноманітнішими вхідними параметрами. Проте нейронні мережі є найпоширенішими методами для здійснення прогнозування газопостачання. Більшість використовуваних вхідних змінних – дані про споживання за минулі періоди та погодні дані. Багато авторів намагалися розробити моделі для прогнозування природного газу – споживання на погодинній, щоденній, щотижневій, місячній або річній основі.

З огляду на це нами було здійснено глибокий аналіз існуючих нейронних мереж та побудовано 3 моделі нейронних мереж процесу газоспоживання міста Тернополя з врахуванням метеофакторів – одношаровий перцептрон, багатошаровий перцептрон, LSTM, що дозволило здійснити аналіз швидкості тренування нейронних мереж по епохах, визначити перелік оптимальних вхідних параметрів для дослідження та напрацювати дієві методи прогнозування газоспоживання.

Мета дослідження: побудова моделі нейронної мережі процесу газоспоживання міста з врахуванням метеофакторів.

Для досягнення вказаної мети, в роботі поставлено та розв'язано **наступні задачі:**

- опрацювати наукові дослідження існуючих математичних підходів до аналізу процесу газоспоживання;
- вивчити особливості застосування нейронних мереж в прогнозуванні газоспоживання;
- провести аналіз основних параметрів дослідження – статистичних даних газоспоживання та метеофакторів для міста Тернопіль;
- визначити критерії відбору даних для дослідження;
- обрати моделі нейронних мереж для здійснення прогнозування газоспоживання у місті Тернопіль;
- побудувати моделі нейронних мереж – одношаровий перцептрон, багатошаровий перцептрон, LSTM для дослідження;
- здійснити порівняльний аналіз апробованих моделей нейронних мереж;
- розробити способи представлення, оцінки ефективності та рекомендації щодо застосування нейронних мереж для прогнозування газоспоживання.

Об'єкт дослідження: процес застосування нейронних мереж для прогнозування газоспоживання в м.Тернопіль.

Предмет дослідження: технологія побудови нейронних мереж з врахуванням метеофакторів для прогнозування газоспоживання в м.Тернопіль.

Наукова новизна роботи:

- запропоновано алгоритм імпорту та попереднього аналізу вхідних даних для м.Тернополя;
- створено математичну модель побудови нейронних мереж – одношаровий перцептрон, багатошаровий перцептрон, LSTM для прогнозування газоспоживання у м.Тернопіль;

—здійснено порівняльний аналіз швидкості тренування апробованих нейронних мереж по епохах та результатів досліджень.

Практичне значення одержаних результатів. Результати роботи планується використовувати для оптимізації прогнозування газоспоживання у м.Тернопіль, що позитивно впливатиме на планування газоспоживання та економію бюджетних ресурсів.

Апробація. Окремі результати роботи були представлені на ІХ науково–технічній конференції «Інформаційні моделі, системи та технології» Тернопільського національного технічного університету імені Івана Пулюя (08–09 грудня 2021 р.) у вигляді опублікованих тез [58].

1. АНАЛІЗ ДОСЛІДЖЕНЬ ПРОЦЕСУ ГАЗОСПОЖИВАННЯ

1.1. Взаємозв'язок економічного розвитку держави та споживання енергії.

Розвиток енергетики має вирішальний вплив на стан економіки в державі та рівень життя населення. Оскільки найвищою цінністю є людина, її добробут, безпека, захищеність, влада зобов'язана дбати як про економічний розвиток держави, так і про добробут населення. Однією з найважливіших складових добробуту в цивілізованих державах є забезпечення громадян і компаній необхідними енергоресурсами. Запорукою реалізації цієї мети має стати надійне, економічно обґрунтоване й екологічно безпечне задоволення потреб населення й економіки в енергетичних продуктах.

Природний газ є одним із найважливіших енергоресурсів, яких стає все більше, і найбільш популярний завдяки своїм екологічним перевагам, адже має нижчий вплив на забруднення навколишнього середовища. В останні роки за даними [50] споживання газу зросло повсюдно, особливо в Китаї та Близькому Сході. Головним винятком є Японія, де найбільше зростає атомна енергетика. Прогнозується, що споживання природного газу збільшиться зі 120 трильйонів кубічних футів у 2012 році до 203 у трильйонів кубічних футів у 2040 році [48]. На природній газ припадає 23,8 % споживання первинної енергії у світі [55]. Природний газ залишається основним паливом у виробництві електроенергії і промисловості.

Історично склалося, що на території України зосереджена неправомірно велика частка підприємств енергомістких галузей (гірничовидобувна, чорна та кольорова металургія, важке машинобудування) промисловості колишнього СРСР. Після відновлення своєї незалежності Україна була забезпечена природним газом за рахунок власних ресурсів лише на 20%. Незважаючи на подальшу стагнацію промислового виробництва, невиправдано високі ціни на

газ та залежність від єдиного джерела імпорту, з середини 90-х років минулого століття за рівнем споживання газу Україна увійшла в десятку найбільших країн світу. В енергетичному балансі нашої держави природний газ посідає провідне становище – частка газу в балансі енергоносіїв за останніми даними становить 45% (в Європі – 21,3%) [50]. В той же час у балансі запасів України частка газу становить лише 1% (частка вугілля – 95%). Дослідження українських науковців свідчать, що за умови таких же темпів розвитку української економіки, які були в попередні роки, та формуванні соціально-економічних запитів громадян України обсяги газоспоживання зростали б і надалі [48]. Цілком очевидно, що разом із обсягами споживання газу зростатимуть і проблеми в газовій промисловості, яких і на сьогодні є чимало.

Зважаючи на це, враховуючи комплексний підхід до розвитку економіки нашої держави Кабінетом Міністрів України було розроблено Національну економічну стратегію на період до 2030 року та Стратегію енергетичної безпеки України до 2025 року.

Енергетична безпека країни – це впевненість у наявності, доступності та можливості стабільно отримувати паливо та енергію належної якості як щодня за звичайних умов, так і у випадку за надзвичайних обставин. Іншими словами, енергетична безпека – це захищеність держави, її громадян та економіки від дефіциту енергії. Енергетична безпека України є одним з найважливіших елементів національної безпеки і невід’ємною складовою енергетичної складової Європи.

Питання енергетичної безпеки набуло особливої актуальності з моменту проголошення Незалежності України та ще більш загострилося після 2014 року. Часті випадки погроз зриву енергетичних ресурсів з Росії змусили Україну по-новому поглянути на питання енергетичної безпеки, котра є необхідною умовою існування держави.

Реформи в енергетичному секторі, які розпочала здійснювати Україна, мали на меті дати енергетичному комплексу поштовх до інтеграції з

європейськими комплексами, привести його у відповідність до норм та стандартів Європейського Союзу, сформувати повноцінні конкурентні ринки природного газу і енергії з прозорою ціновою політикою та належним захистом споживачів.

Не можливо оминати тему взаємозв'язку економічного розвитку держави та споживання енергії, в руслі російсько-української війни. Повномасштабне вторгнення Росії в Україну є не тільки фактом військової агресії, але й негативним чинником, який вплинув на послаблення енергетичної міцності, як України, так і цілого світу. Міжнародна спільнота, в тому числі і Україна, ще більш активно розпочали пошуки альтернативних російським джерел забезпечення енергією, а прогнозування газоспоживання в свою чергу стає надважливим інструментом у формуванні енергетичної безпеки держав.

1.2. Прогнозування в енергетиці, як засіб оптимізації планування, підвищення ефективності газоспоживання.

Зважаючи на ряд переваг, попит на природний газ як у світі, так і в Україні щорічно зростає.

Існуюча система розрахунку споживання природного газу населенням у випадку відсутності технічної можливості його обліку проводиться за усередненими нормами на одного мешканця чи на один квадратний метр опалювальної площі, проте вона не відображає фактичних витрат газу конкретним споживачем, які можуть бути значно меншими або значно більшими від розрахованих за нормами витрат [55]. Високоюмовірно, схожі проблеми існують і на промислових підприємствах, де кількість спожитого газу знаходять на основі теплотехнічних характеристик встановленого обладнання, яке забезпечує технологічні процеси та опалювально-вентиляційні потреби [50].

З огляду на це, при складанні місячного балансу розраховуються так звані „відтоки” чи „притоки” газу. Відтоки газу – перевитрачання газу, яке може виникати внаслідок перевищення фактично спожитих населенням обсягів газу (у випадку відсутності приладів обліку) в осінньо-зимовий період порівняно з діючими середньорічними нормами споживання газу. Притоки газу – це недобір газу, який може виникати через те, що у весняно -літній період фактично спожиті населенням (за відсутності приладів обліку газу) обсяги газу менші за діючі середньорічні норми споживання газу.

Для підприємств з газопостачання та газифікації, основним видом діяльності яких є постачання споживачам газу, вищезазначені обсяги газу є предметом купівлі з метою подальшого продажу, тобто товаром. Відповідно якість прогнозування газоспоживання суттєво впливає на фінансову складову.

Зрозуміло, що існуюча система обліку природного газу є застарілою і непридатною для контролю витрат газу, тим більше для завдань його планування.

З огляду на це застосування прогнозування в газоспоживанні стає необхідним засобом оптимізації планування та підвищення ефективності газоспоживання. Точність моделі прогнозування процесу газоспоживання, з врахуванням метеофакторів, топології споживачів є важливою для прийняття рішень щодо номінації газу.

Газові номінації зазвичай створюються газорозподільником і використовуються газотранспортером і виробником газу для оптимізації його потоку та запасів. Більше того, кожен дистриб'ютор природного газу зобов'язаний здійснити номінацію природного газу своїм постачальником, що становить обсяг газу, необхідний на наступний день (або інший майбутній період) – підтвердити точні обсяги в енергетичних одиницях і показати прогнозні обсяги в кубічних метрах. Є певні регламентовані допустимі норми відхилення від планових обсягів. У разі, якщо фактичне споживання перевищує номінальні суми, дистриб'ютор повинен сплатити певний штраф. З іншого боку, якщо

вказана сума перевищує фактичне споживання, також буде нараховано інший вид штрафу (хоча і менший). Оскільки неправильні номінації призводять до великих витрат, точні прогнози споживання природного газу на наступну добу дуже важливі для дистриб'ютора природного газу з фінансових причин.

Автори [61] визнають, що планова подача газу значно нижча фактичних потреб, тому витратами газу намагаються управляти з допомогою політики «лімітів». Особливо це стосується бюджетних установ. Однак споживачі прагнуть максимального відбору, обмежити який технічними засобами майже неможливо. Досвід показує, що практика контролю за дотриманням лімітної дисципліни не приносить бажаних результатів та негативно впливає на розвиток окремих галузей.

В міру наведених фактів можна сміливо стверджувати, що планування витрат газу є одним із найважливіших та найболючіших питань не лише газової промисловості, а й економіки загалом. А моніторинг в системі контролю, обліку, управління і прийняття рішень процесу газоспоживання виступає невід'ємною складовою ефективного функціонування галузі.

1.3. Огляд наукових публікацій по тематиці дослідження з метою вивчення існуючих математичних підходів до аналізу процесу газоспоживання.

Питання моделей прогнозування споживання природного газу було і є предметом досліджень багатьох науковців. Спробуємо представити огляд і систематичний аналіз наукових робіт, в яких висвітлено тему прогнозування споживання природного газу житловими та комерційними підприємствами, починаючи з 2002 року.

Аналіз огляду літератури проводився з використанням двох найбільш актуальних наукових баз даних Web of Science Core Collection і Scopus. Для

пошуку статей в обох базах даних були використані ключові слова «прогноз природного газу (споживання або попит)».

Пошук здійснено у базі даних у трьох індексах WoSCC: Science Citation Index Expanded (SCI–Expanded), Індекс цитування соціальних наук (SSCI) і Arts and. Індекс цитування гуманітарних наук (A&HCI).

Результатом цього пошуку стало 197 робіт із WoS та 431 стаття від Scopus. Після перегляду назви, тез і ключових слів усіх знайдених статей з WoS перші 100 статей від Scopus відсортовані за релевантністю, статті, які не стосуються моделі для прогнозування споживання природного газу для житлових або комерційних потреб були ліквідовані. Після цього 29 статей із WoS та 24 статті з Scopus залишаються, що відповідають опублікованим критеріям. Проте в обох базах даних було 14 однакових статей, тобто 39 статей були проаналізовані за кількома критеріями: методами, що використовуються для прогнозування споживання природного газу, вхідні змінні, що використовуються для моделювання, область прогнозування та горизонт прогнозування. Подібний огляд літератури проводив Солдо [36], яким проаналізовано споживання природного газу з 1949 по 2010 рік.

Аналіз даних 39 публікацій засвідчив, що, за останні три роки дослідження в 16 роботах були опубліковані прогнози споживання природного газу, що становить понад 41% усіх проаналізованих робіт, тобто з року в рік питання прогнозування газоспоживання стає більш актуальним та більш досліджуваним.

Прогноз споживання ПГ можна проводити на різних ділянках, наприклад на країну, регіон, місто або окремий район замовника [36]. Аналіз робіт показує, що автори переважно прогнозують споживання природного газу на рівні країни. Причина в тому, що даних про споживання природного газу на рівні країни більше, оскільки вони містяться в ряді оприлюднених статистичних (енергетичних) звітів.

Існує кілька горизонтів прогнозування, які можна використовувати для прогнозування природного газу споживання. Велика кількість авторів

прогнозували споживання природного газу на річний рівень. Гіл і Дефєррарі [20] представили модель, призначену для прогнозування природного газу споживання на короткому та середньому діапазоні часу для регіону в Аргентині. У проміжному діапазоні (1–5 років) модель може оцінити річний пік споживання. Се і Лі [43] прогнозували загальний попит на 2003–2005 роки, за даними з 1996 по 2002 рр. Гомперц використав інноваційну дифузійна модель для прогнозування річного споживання природного газу в Іспанії. Автори [19] використовували логістичний підхід для прогнозування річного споживання природного газу за 2006–2008 роки, тоді як Ма і Лі [28] передбачили споживання газу Китаєм у 2009–2020 рр.

Для свого дослідження Кумар і Джайн [26] використали дані, отримані від енергетичної статистики в Індії з 1970–1971 по 2005–2006 роки. Азаде та ін. [11] прогнозує споживання природного газу в чотирьох країнах Близького Сходу в період з 2008 до 2015, Ольгун та ін. [32] представили модель для оцінки річного попиту на ПГ для Туреччини до 2030 року, Фенг та ін. [18] представив прогноз значення попиту на ПГ протягом 2012–2015 років з використанням моделі ARIMA, Чжан і Ян [46] – прогнозне споживання природного газу за різними сценаріями з 2015 по 2020 р., тоді як Боран [14], Ву та ін. [42], Цзен і Лі [45] та Ма і Лю [27] використовували сірий колір моделі прогнозування для прогнозування річного споживання ПГ у Туреччині та Китаї. Форузанфар та ін. [19] змодельювали сезонне споживання ПГ (за 2006–2008 р.).

Прогнозування споживання природного газу на місячному рівні описують Вієт і Мандзюк [41], які розглядали три види передбачення – одноденне, тижневе та на чотири тижні (місяць), Арас [7], який спрогнозував побутове споживання за допомогою ген – алгоритмів на 2009–2013 роки, Акпінар і Юмусак [3], [5] передбачили споживання домогосподарств з 2009 по 2012 роки, Азаде та ін. [9] запропонували гібридний метод на основі комп'ютерного моделювання та ANFIS за 2009–2013 роки, Ізаяр та ін. [22] спрогнозували попит на опалення в Ірані та Акпінар і Юмусак [2], які спрогнозували споживання ПГ,

використовуючи різні часові ряди. Поточник та ін. [34], запропонували модель прогнозування для прогнозування оцінки ризику, автори [25] використовували нейронну мережу та нейронечітку систему для прогнозування споживання ПГ щотижня.

Найбільша кількість авторів прогнозували споживання природного газу на добовому рівні. Віет і Мандзюк [41], серед інших передбачень, провели прогноз на один день споживання природного газу, тоді як Гіл і Деферрарі представили модель, призначену для прогнозування споживання природного газу на коротких і проміжних діапазонах часу для регіону в Аргентині. На короткій дистанції модель може передбачити споживання на 1–5 днів наперед. Ельрагаль [17] запропонував модель для прогнозування добового споживання природного газу, необхідного газовим підприємствам, Поточник та співавтори [34] представили модель ризику, яка застосована для оцінки щоденного ризику прогнозування. Для цього їм довелося створити модель прогнозу споживання природного газу на день. Брабець та ін. прогнозували щоденно споживання ПГ на рівні індивідуальних споживачів. Азаде та ін. [8] спрогнозували короткостроковий попит на ПГ. Демірель та ін. [16] спрогнозували щоденне споживання газу у 2009 році на основі історичних даних, Таспінар та ін. [39] моделював щоденник споживання ПГ за допомогою трьох різних моделей, а Солдо та ін. [37] використовували сонячну радіацію як вхідну змінну для прогнозування щоденного споживання ПГ.

Чжу та ін. [47] спробували передбачити попит на природний газ у Сполученому Королівстві, Бай і Лі [12] запропонували підхід регресії опорного вектора, відкалібрований за структурою для прогнозування добового споживання, Акпінар і Юмусак [1] спрогнозували споживання ПГ на добу вперед за 2012 рік, а також Ранапакідіс і Дагумас [33], Акпінар та ін. [32] і Акпінар і Юмусак [6].

Прогнозування споживання природного газу в погодинному масштабі повідомили [40], які створили прогнозну модель споживання природного газу

на регіональному рівні за допомогою нейронних мереж, Сабо та ін. [35] запропонували математичні моделі споживання природного газу. Шоплік [38] спрогнозував споживання природного газу у Польщі за допомогою нейронних мереж.

Провівши стислий огляд горизонтів прогнозування, що використовуються для прогнозування споживання природного газу у даних роботах, ми отримали наступний результат – $\frac{2}{3}$ робіт присвячені прогнозам на річному або денному рівні (щоденно – 15, щорічно – 14, щомісячно – 7, щотижня – 3, щогодини – 3, сезонні – 2). Це цілком зрозуміло, оскільки більшість статистичних енергетичних публікацій, а також дистриб'ютори природного газу, забезпечують річне споживання природного газу.

Причина найбільшої кількості прогнозів, проведених на повсякденний рівень полягає в тому, що цей тип передбачень дуже важливий для дистриб'юторів природного газу з фінансових причин. Точні прогнози на день вперед зменшують можливість сплати штрафних санкцій.

Споживання природного газу прогнозується за допомогою різних методів прогнозування або навіть поєднання кількох методів (так звані гібридні моделі). Сольдо та ін. [36] виявили, що серед перших інструментів для прогнозування споживання природного газу була модель кривої Хабберта, яка використовувалася в 1950–х роках. З 1960–х років, коли було розроблено статистичні моделі, для прогнозів використовувалися різні статистичні моделі споживання. З кінця 1970–х і 1980–х років штучні нейронні мережі стали дуже популярним інструментом прогнозування. Останнім часом у прогнозах використовуються нові методи споживання природного газу, наприклад сірі моделі або генетичні алгоритми.

Нейронні мережі є методами, які, найчастіше шляхом ітераційного процесу з минулих даних, намагаються знайти зв'язок між вхідними та вихідними змінними моделі з метою отримання вихідного значення для нових вхідних змінних [44]. Адаптивний мережевий нечіткий висновок система (ANFIS) була

представлена Джангом [23], яка є «системою нечіткого висновку, реалізована в рамках адаптивних мереж». За даними Азаде та ін. [8], «ANFIS володіє як здатністю до навчання нейронних мереж, так і структурованими представленнями знань, що використовується в системі нечітких висновків, що є відповідним для нелінійного моделювання та прогнозування часових рядів».

Авторами, які використовували нейронні мережі, були [41], [40], [24], [39], [16], [32], [37], [18], [38], [22] та [6]. Вієт і Мандзюк [41] представили кілька нейронних і нечітких нейронних підходів. Порівняння середньої абсолютної відсоткової помилки (MAPE) значення, найкращу продуктивність для щоденних (4,04%), а також для тижневих (7,04%) передбачень, було досягнуто за допомогою трьох нейронних модулів, кожному з яких був присвячений певний температурний діапазон. Тонкович та ін. [40], Таспінар та ін. [39] описали та протестували два алгоритми нейронної мережі – багатошаровий перцептрон і мережу радіальних базисних функцій з різними функціями активації. Перший згаданий алгоритм створив найменший MAPE у всіх трьох аналізованих документах (9,36%, 0,814%, 5,477% відповідно). Таспінар та ін. [39] також порівняли нейронні мережі алгоритми та модель часових рядів. У своїх дослідженнях Демірель та співавтори [16] використали багатошаровий алгоритм перцептрона для нейронної мережі та порівняли цю модель з 2 моделями часових рядів. Нейронна мережа показала найкращий результат з MAPE 0,1833. Ольгун та ін. [32] порівняли нейронні мережі з машинами опорних векторів і вони дійшли висновку, що SVM має меншу статистичну похибку. Фенг та ін. [18] розробили три різні види моделі – модель ARIMA, модель нейронної мережі та комбіновану модель. Модель нейронної мережі (радіальна базисна функція) досягла MAPE 5,78%. Шоплік [38] також використовував багатошаровий алгоритм перцептрона для прогнозування природного споживання газу. Середній MAPE становив 8%, а нижчі значення MAPE є записані в зимові місяці. Солдо та ін. [37] досліджували вплив радіації сонця на споживання ПГ населенням. Вони використовували кілька різних методів – нейронні мережі на

двох різних наборах даних. Перший набір даних з моделі будинку (MAPE 5,06%), а другий – від місцевої дистриб'юторської компанії (MAPE 2,52%). Машину екстремального навчання як алгоритм навчання для прямого зв'язку нейронної мережі, використовували Ізадьяр та ін. [22]. Вони порівняли метод ELM з штучною нейронною мережею (зворотне поширення як правило навчання) і генетичним програмуванням. Результати точності, виміряні в термінах RMSE (середньоквадратичне значення помилки), перейшли на користь моделі ELM (0,18068). Акпінар та ін. [6] представили ABC Алгоритм (Artificial Bee Colony) як альтернативу використання нейронної мережі. Нейронна мережа ABC генерувала нижчі значення помилок (14,9% MAPE) порівняно з нейронною мережею BP (33,9% MAPE). У кількох роботах були представлені гібридні моделі. Панапакіді і Дагумас [33] запропонували цікаву модель, що поєднує вейвлет–перетворення, генетичний алгоритм, адаптивну систему нейро–нечіткого висновку та нейронну мережу з прямим зв'язком. Вони спробували перевірити надійність нової гібридної моделі обчислювального інтелекту у прогнозах попиту на природний газ на добу вперед. Їх модель була застосована до всіх газорозподільних пунктів. Найменші похибки зустрічалися у великих міських центрах. Були ще праці, які поєднували дві різні моделі. Ельрагаль [17] запропонував гібрид – модель складалася з нейронної мережі та нечітко–генетичної моделі. На першому етапі два адаптивні нейромережі–провісники працюють паралельно і виробляють незалежне передбачення. Ці передбачення потім вводяться на другий етап, який включає нечітко–генетичний модуль. Хоча нейромережі–провісники показали хороші результати (4,12% MAPE), стратегія комбінацій покращила результат (3,97% MAPE).

У своїй роботі Се і Лі [43] запровадив метод моделювання сірого, оптимізований за допомогою генетичного алгоритму. традиційна сіра модель прогнозування дала відносну похибку 10,5%, тоді як така сама помилка на основі оптимізованої моделі була меншою (8,16%). Фенг та ін. [18] розробили три різні види моделі – модель ARIMA, модель нейронної мережі і комбінована модель.

MAPE комбінованої (гібридної) моделі становив 4,52%. У цьому випадку, гібридна модель не досягла кращого результату, оскільки модель ARIMA мала помилку – 3,06%. Азаде та ін. [9] створили гібридну модель адаптивної нейрон – системи нечіткого висновку та комп'ютерне моделювання, які мали значення MAPE 13,5%.

Ті ж автори [10] представили адаптивну мережеву систему нечіткого висновку (ANFIS) у своїх попередніх дослідженнях із значенням MAPE 2%. Подібне дослідження було проведено Азаде та ін. [11] з метою прогнозування споживання природного газу в країні Близького Сходу. Результати MAPE для Бахрейну, Саудівської Аравії, Сирії та Об'єднаних Арабських Еміратів становили 1,8%, 1,4%, 7,5% і 1,6% відповідно. ANFIS був використаний дослідниками [24], які представили три різні моделі для прогнозування споживання природного газу – модель нейронної мережі, модель часового ряду ARIMA та ANFIS. Останній виробив значення MAPE 5,468%, що майже таке ж, як модель нейронної мережі (5,477%). Декілька авторів використовували різні моделі сірого прогнозування для прогнозів природного газу споживання.

Іншими широко використовуваними методами для прогнозування споживання природного газу є машина опорних векторів (SVM) і опорна векторна регресія (SVR). Ольгун та ін. [32] порівняли нейронні мережі з машинами опорних векторів і прийшли до такого висновку – машини опорного вектора мали меншу статистичну похибку для оцінки попиту споживання на природний газ.

Солдо та ін. [37] використовували декілька лінійних і нелінійних моделей для передбачення. Похибки тестування, отримані нелінійними нейронними мережами (модель будинку: 5,06%, LDC: 2,52%) та моделі SVR (модель будинку: 5,28%, LDC: 2,25%) трохи вищі в порівнянні з лінійними моделями.

Чжу та ін. [47] представили метод, який інтегрував алгоритм SVR з властивостями реконструкції часового ряду і оптимізує вихідний локальний предиктор, видаляючи помилкових сусідів. MAPE цієї моделі становила 3,8%.

Бай і Лі [12] використали підхід SVR, відкалібрований за структурою. Результати показали, що середня абсолютна відсоткова помилка становила 2,36%. Арас [7] намагався спрогнозувати короткостроковий попит на природний газ у житлових будинках шляхом використання генетичних алгоритмів. За його словами, генетичні алгоритми – це стохастичний пошук методик, заснованих на механізмі природного відбору та природної генетики.

Ізадяр та ін. [22] порівняли модель генетичного програмування з моделі ELM та нейронної мережі. При генетичному програмуванні похибка передбачення була дещо вищою (0,18504), порівняно з моделлю ELM (0,18068).

Деякі автори використовували математичні або статистичні моделі для прогнозування природних споживання газу. Гіл і Деферрарі [20] розробили модель, яка здатна передбачити побутове та комерційне споживання ПГ на короткий (від 1 до 5 днів) і проміжний діапазон (від 1 до 5 років). На короткі відстані похибка прогнозу становила близько 10%, а при проміжному діапазоні така ж похибка склала 12%. Сабо та ін. [35] зробили висновок, що лінійна апроксимація дає прийнятний прогноз для практичних потреб.

Солдо та ін. [37] порівняв декілька лінійних моделей, таких як авторегресивна модель, екзогенні вхідні дані (ARX) і ступінчаста регресія. Модель ARX виробляє MAPE 3,25% для моделі будинку і 1,17% для НРС, тоді як поетапна регресія мала помилку екзогенні вхідні дані (ARX) і ступінчаста регресія. Значення 3,64% для моделі будинку і 1,33% для НРС. Акпінар і Юмусак [1] використали множинну лінійну регресію (MLR) для прогнозування споживання ПГ. Вони запропонували кількох моделей і найнижчий MAPE становив 6,8%, а найвищий був навіть 54,5%. Акпінар і Юмусак [5] також розробили модель з використанням декількох лінійних регресій. MAPE їх моделі становив 14,38%.

Проаналізувавши методи прогнозування, що використовуються для прогнозування споживання природного газу та середню помилку прогнозування для кожної групи методів прогнозування на основі вищезазначених наукових

джерел видно, що найточніше передбачення були дані генетичними алгоритмами, а потім опорним вектором машини/регресія та ANFIS. Однак до цих висновків слід ставитися обережно, оскільки існує невелика кількість статей про ці методи. Нейронні мережі також можна вважати досить точним методом, оскільки середня похибка менше 6%. Методи часових рядів і гібриди моделі мають найбільшу похибку передбачення.

Опрацьовані наукові джерела свідчать, що найбільш часто використовувані методи для прогнозу споживання природного газу ґрунтуються на нейромережах та подібних методах (ANFIS), а також використовуються різні математичні та статистичні моделі. Також можна побачити, що деякі методи пов'язані з певною областю передбачення або горизонтом. Наприклад, усі автори, які використовували сірі моделі, проводили дослідження на рівні країни для щорічного прогнозування. Нейронні мережі, а також математичні і статистичні моделі, можуть використовуватися для всіх типів областей прогнозування для всіх горизонтів передбачень.

Для створення моделей прогнозування споживання ПГ були використані різні набори даних. Відповідно, для створення моделі автори використовували різні входні змінні. Можна помітити, що найбільш часто використовуваною змінною був природний газ споживання. Ця входна змінна була використана Се і Лі [43], Кумар і Джайн [26], Кайнар та ін. [25], Акпінар і Юмусак [3], [2], [4], Фенг та ін. [18], Боран [14], Ву та ін. [42], Бай і Лі [12], Цзен і Лі [45], Ма і Лю [27] та Акпінар та ін [6]. Ці автори використовували переважно сірі моделі і різні моделі часових рядів.

Деякі автори разом із минулим споживанням газу використовували й інші змінні також. Віет і Мандзюк [41], Гіл і Деферрарі [20], Сабо та ін. [35] використовували зовнішню температуру, Ельрагаль [17] температуру, швидкість вітру, втрати тепла, день тижня, Поточником та ін. [34] використано попередні погодні дані, прогноз погоди, сезонна інформація, дні тижня, свята, промислові номінації, Тонковичем та ін. [40] використано подібні входні змінні, такі як

місяць, визначення сезону, тип дня (робочий день, свято, день за святом), день тижня, температура, напрямок вітру, швидкість вітру в різний час, Акпінар і Юмусак [5], [1] використовували температуру, вологість, швидкість вітру, вихідні дні, кількість споживачів газу, Ізадяр [22] використовував температуру і місяці, Чжу та ін. [37] температуру, швидкість вітру, ефективну температуру та псевдо сезонну нормальну ефективну температуру, в той час як Панапакідіс і Дагумас[33] – температуру, місяці та дні тижня. Арас [7] використовував для прогнозів індекс споживчих цін і середньодобові температури. Брабецом та ін. використано температуру та день тижня. Подібні змінні можна знайти в статті Шопліка [38]. Він використовував температуру, годину доби, день, тиждень, місяць і день місяця. Для прогнозування попиту на природний газ, Азаде та ін. [8] використовували день тижня, запит на той самий день попереднього року, запит на день раніше, запит за 2 дні до, як вхідні змінні. Демірель та ін. [16] використовували щоденну середню температуру, температуру в квадраті, ціну природного газу, кількість газу, споживачів, тоді як Таспінаром та ін.[39] використана вологість, атмосферний тиск, швидкість вітру, температура навколишнього повітря та середня хмарність.

Деякі автори включили валовий внутрішній продукт (ВВП) до своїх моделей. Ма і Лі [28] використаний ВВП і минуле споживання, Ольгун та ін. [32] використовували ВВП і населення, тоді як Азаде та ін. [9] використаний ВВП, населення, ціна НГ, рівень інфляції, рівень безробіття, Індекс IT/IS, індекс людського розвитку та викиди CO₂. Солдо та ін. [37] провели цікаві дослідження, в яких досліджували вплив сонячної радіації на споживання побутового природного газу. Вхідними змінними, які вони використовували, були дані про погоду (погодинна) і сонячна радіація.

З цього аналізу видно, що споживання газу в минулому є найважливішою вхідною змінною для прогнозування споживання природного газу. Її використовували в найбільшій кількості статей, особливо в статтях, де були

математичні, сірі моделі та моделі часових рядів. Важливими змінними були також різні метеорологічні дані.

Можна зробити висновок, що в дослідженнях на основі нейронних мереж метод прогнозування споживання ПГ використовується найбільша кількість вхідних змінних. Трохи менше змінні використовуються в групі різних математичних і статистичних моделей. Порівнюючи результати аналізу, можна побачити, що на точність моделі не впливає кількість вхідних змінних, оскільки найбільш точні моделі також мають дуже малу кількість змінних (генетичні алгоритми та ANFIS) на відміну від найменш точних моделей (наприклад, моделі часових рядів).

Коли справа доходить до різних областей прогнозування в прогнозуванні споживання ПГ, немає очевидної різниці в кількості використовуваних вхідних змінних. З іншого боку, коли мова йде про різний горизонт прогнозування, більшість вхідних змінних використовується в дослідженні щоденно та щомісячно.

Сучасні українські автори Назаревич О., Загородна Н., Литвиненко Я., Мацюк О., Фриз М. продовжили вивчення цієї надважливої теми, проводили ряд досліджень у сфері прогнозування газоспоживання та застосовували різноманітні методи з різними вхідними параметрами для більш точного прогнозування споживання природного газу [30, 53, 50, 52]. Було зроблено прогнози для наступних сфер застосування – установа, місто, область та часових періодів – годинний, щоденний, річний.

1.4 Висновки до першого розділу

У першому розділі нами проаналізовано взаємозв'язок між економічним розвитком держави та споживанням енергії, зокрема газоспоживанням, звернено увагу на економічну безпеку держави, яка серед іншого забезпечується

прогнозуванням в газоспоживанні, проведено аналіз досліджень процесу газоспоживання.

Результати огляду наукових статей показують, що найбільш поширений метод, який використовується для прогнозування споживання природного газу – метод нейронних мереж або методи, засновані на тих самих принципах (наприклад, ANFIS), за якими слідують методи часових рядів і різні математико–статистичні методи. Іншими популярними методами є генетичні алгоритми, підтримка векторних машин/регресії, сірих системних моделей або гібридних моделей на основі кількох методів.

Вибір відповідного методу буде залежати від обраного горизонту прогнозування або використання відповідних вхідних змінних буде залежати від використаного методу. Найнижчими середніми похибками передбачення є генетичні алгоритми, допоміжні векторні машини і ANFIS, в той час як найменш точними методами є різні методи часових рядів і гібриди моделей. Але, очевидно, немає статистично значущої різниці між точністю методів, і це може бути однією з причин використання такої величезної кількості різних методів прогнозування споживання ПГ (враховуються в першу чергу деякі інші ознаки споживання газу, наприклад кількість і тип доступних змінних, а не точність).

Аналіз показав, що для моделювання автори часто використовують природний газ минулого року, дані про споживання та дані про погоду (переважно температура) як вхідні змінні. Інший змінні включають місяць, дні тижня, швидкість вітру, кількість природного газу, споживачів, ВВП, населення тощо. Говорячи про області прогнозування, видно, що більшість статей присвячені прогнозам на рівні країни. Можна робити прогнози також на регіональному, міському чи навіть житловому рівні.

Багато авторів намагалися розробити моделі для прогнозування природного газу – споживання на погодинній, щоденній, щотижневій, місячній або річній основі. Проте аналіз виявив, що науковці переважно прогнозують споживання природного газу на добовому та річному рівні. Є небагато

документів, які прогнозують споживання природного газу за сезонними, місячними, тижневими або погодинними показниками.

2 АНАЛІЗ ІСНУЮЧИХ НЕЙРОННИХ МЕРЕЖ, МЕТОДИКИ МАШИННОГО НАВЧАННЯ ДЛЯ ПРОЦЕСУ ГАЗОСПОЖИВАННЯ

2.1 Поняття нейронної мережі.

Штучна нейронна мережа – це математична модель, а також її програмна та апаратна реалізація, побудовані за принципом функціонування біологічних нейронних мереж – мереж нервових клітин живого організму. Процеси, які відбуваються в мозку людини, лягли в основу нейронних мереж.

Нейронна мережа визначається як масивно–паралельний розподілений процесор, який складається з простих процесорів, які мають природну схильність зберігати досвідчені знання та робити їх доступними для використання.

Нейронні мережі набули широкого застосування в аналізі та класифікації даних за обраними параметрами, формуванні аналітичних прогнозів на основі вхідної інформації, порівнянні та розпізнаванні ідентичних даних.

Завдяки цьому в наш час нейронні мережі використовуються в різних галузях – медицині (діагностика захворювань), економічній сфері (прогнозування ситуації на фондовому ринку, оцінка вартості нерухомості), зв'язку (знаходження оптимального шляху трафіку між вузлами), безпеці та охороні (ідентифікація особи, розпізнаванні образів, розпізнаванні по голосу, розпізнаванні номерних знаків), кодуванні та декодуванні інформації, прогнозуванні тощо.

2.2 Принципи роботи нейронних мереж.

Обробка інформації в нейронній мережі здійснюється за допомогою нейронів. Штучні нейронні мережі є моделями нейронної структури мозку, який здатен сприймати, обробляти, зберігати та продукувати інформацію. Штучний нейрон – це спрощена модель біологічного нейрона.

Перцептрон, як кібернетична модель мозку, одна з перших моделей нейронних мереж, був розроблений в 1950–х роках вченим Франком Розенблатом, натхненним попередніми роботами Уоррена Мак–Калока та Уолтера Піттса. А в 1960–х роках Ф.Розенблат створив перший у світі нейрокомп'ютер «Марк 1».

Перцептрон отримує кілька двійкових входів, x_1, x_2, \dots, x_i створює один двійковий вихід:

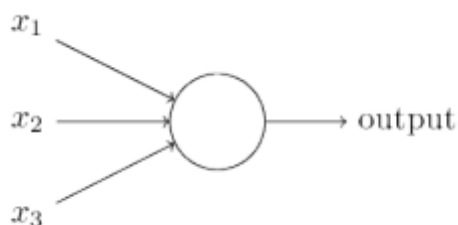


Рисунок 2.1– Перцептрон

У наведеному прикладі перцептрон має три входи: x_1, x_2, x_3 . Загалом він може мати більше або менше вхідних даних. Розенблат запропонував просте правило для обчислення результату. Він ввів ваги w_1, w_2, \dots, w_i , дійсні числа, що виражають важливість відповідних входів для виходу. Вихід нейрона, 0 або 1, визначається тим, чи є зважена сума меншою або більшою за деяке порогове значення. Так само, як і ваги, поріг — це дійсне число, яке є параметром нейрона.

Якщо сказати більш точні алгебраїчні терміни:

$$output = \begin{cases} 0 & \text{if } w_j x_j \leq threshold \\ 1 & \text{if } w_j x_j \geq threshold \end{cases} \quad (1)$$

Це основна математична модель роботи перцептрона. Перцептрон є своєрідним пристроєм, блоком, який на основі заданої інформації, приймає рішення.

2.3 Аналіз існуючих моделей нейронних мереж.

2.3.1 Одношаровий та багатошаровий перцептрон.

Очевидно, що перцептрон не є повною моделлю прийняття людських рішень, однак він може зважувати різні види доказів, щоб приймати рішення. А складна мережа перцептронів може приймати досить складні рішення:

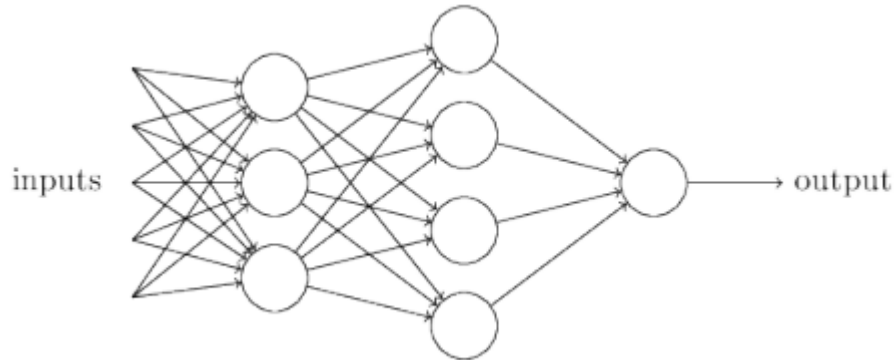


Рисунок 2.3– Архітектура багатошаровий перцептрон

У цій мережі перший стовпець перцептронів – те, що ми назвемо першим шаром перцептронів – приймає три дуже прості рішення, зважуючи вхідні докази. Кожен з перцептронів другого шару приймає рішення, зважуючи результати першого рівня прийняття рішень. Таким чином, перцептрон у другому шарі може приймати рішення на більш складному та більш абстрактнішому рівні, ніж перцептрони в першому шарі. А ще більш складні рішення може приймати перцептрон третього шару. Таким чином, багатошарова мережа перцептронів може брати участь у прийнятті складних рішень.

Перцептрон має лише один вихід. На зображенні вище перцептрони виглядають так, ніби мають кілька виходів. Насправді вони все ще є єдиним виходом. Стрілки з кількома вихідними сигналами є лише корисним способом вказати, що вихідні дані з перцептрона використовуються як вхідні дані для кількох інших перцептронів.

Використовуючи зміщення замість порогового значення функції, правило перцептрона можна переписати:

$$output = \begin{cases} 0 & \text{if } wx + b \leq 0 \\ 1 & \text{if } wx + b \geq 0 \end{cases} \quad (2)$$

Перцептрони розглядаються як метод зважування доказів для прийняття рішень. Іншим способом використання перцептронів є обчисленням елементарних логічних функцій, які ми зазвичай вважаємо основними обчисленнями, таких як логічне І, АБО та NAND.

Наприклад, припустимо, що у нас є перцептрон з двома входами, кожен з вагою -2 і загальним зміщенням 3 . Ось наш перцептрон:

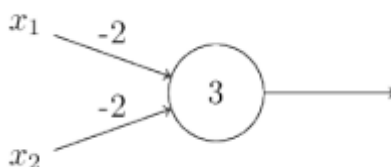


Рисунок 2.5 – Спосіб активації перцептрона

Тоді ми бачимо, що вхід 00 дає вихід 1 , оскільки $(-2) * 0 + (-2) * 0 + 3 = 3$ є додатним. Подібні розрахунки показують, що входи 01 і 10 дають вихід 1 . Але вхід 11 дає вихід 0 , оскільки $(-2) * 1 + (-2) * 1 + 3 = -1$ є негативним. І тому наш перцептрон реалізує вентиль NAND. Приклад NAND показує, що ми можемо використовувати перцептрони для обчислення простих логічних функцій. Насправді, ми можемо використовувати мережі перцептронів для обчислення будь-якої логічної функції взагалі. Причина в тому, що вентиль NAND є універсальним для обчислень, тобто ми можемо побудувати будь-які обчислення з елементів NAND. Наприклад, ми можемо використовувати вентиля NAND, щоб побудувати схему, яка додає два біти, x_1 . та x_2 . Для цього потрібно обчислити порозрядну суму, а також біт переносу, який встановлюється в 1 , коли

обидва x_1 і x_2 дорівнюють 1, тобто біт переносу є лише побітовим добутком x_1 , x_2 .

Щоб отримати еквівалентну мережу перцептронів, ми замінюємо всі вентиля NAND на перцептрони з двома входами, кожен з вагою -2 і загальним зміщенням 3 . Ось результуюча мережа. Перцептрон, що відповідає нижньому правому вентилю NAND трохи переміщено, щоб було легше малювати стрілки на схемі:

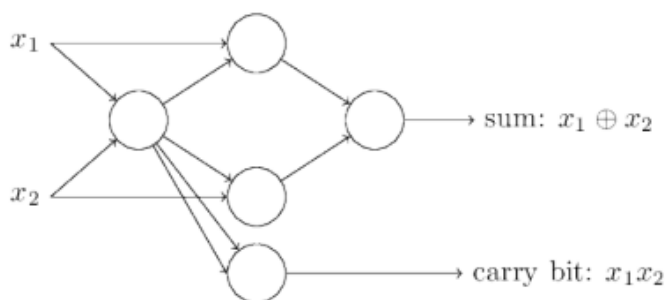


Рисунок 2.6 – Принцип роботи перцептрона

Одним з помітних аспектів цієї мережі перцептронів є те, що вихідні дані з крайнього лівого перцептрона використовуються двічі як вхідні дані для самого нижнього перцептрона. З цією зміною мережа виглядає наступним чином, з усіма непозначеними вагами, рівними -2 , усіма зміщеннями, рівними 3 , і одна вага -4 , як зазначено:

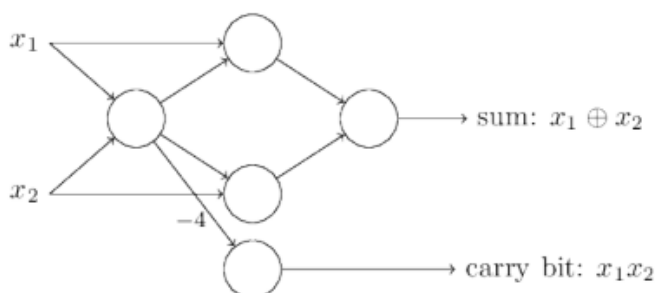


Рисунок 2.7 – Перцептрон з параметрами ваг

До цього часу я малював вхідні дані, такі як x_1 та x_2 , як змінні, що плавають ліворуч від мережі перцептронів. Насправді, звичайним є зобразити додатковий шар перцептронів – вхідний шар – для кодування вхідних даних:

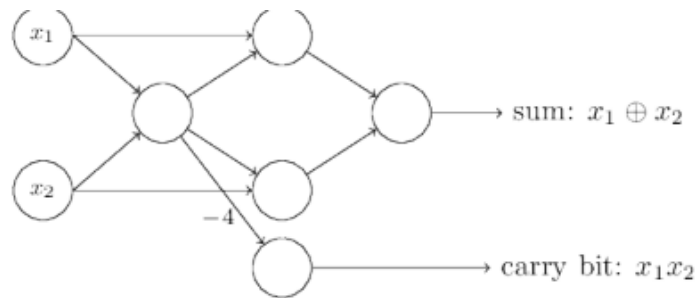


Рисунок 2.8 – Додатковий шар перцептронів з кодуванням вхідних даних

Це позначення для вхідних перцептронів, в яких у нас є вихід, але немає входу – скороченням



Рисунок 2.9 – Перцептрон без входу

Насправді це не означає перцептрон без входів. Щоб побачити це, припустимо, що у нас є перцептрон без вхідних даних. Тоді зважена сума P_j завжди дорівнюватиме нулю, і тому перцептрон виведе 1, якщо $b > 0$, і 0, якщо $b \leq 0$. Тобто перцептрон просто виведе фіксоване значення, а не потрібне значення (x_1 , у прикладі вище). Краще думати про вхідні перцептрони не як про перцептрони взагалі, а як про спеціальні одиниці, які просто визначені для виведення бажаних значень, x_1, x_2 .

Приклад суматора демонструє, як мережа перцептронів може бути використана для моделювання схеми, що містить багато елементів NAND. І оскільки вентилі NAND є універсальними для обчислень, впливає, що перцептрони також універсальні для обчислень. Обчислювальна універсальність

перцептронів одночасно заспокоює і розчаровує. Це заспокоює, оскільки говорить нам, що мережі перцептронів можуть бути настільки ж потужними, як і будь-який інший комп'ютерний пристрій. Але це також розчаровує, тому що створюється враження, ніби перцептрони — це просто новий тип воріт NAND. Однак ситуація краща, ніж припускає ця точка зору. Виявилось, що ми можемо розробити алгоритми навчання, які можуть автоматично налаштувати ваги та зміщення мережі штучних нейронів. Ця настройка відбувається у відповідь на зовнішні подразники, без прямого втручання програміста. Ці алгоритми навчання дозволяють нам використовувати штучні нейрони способом, який радикально відрізняється від звичайних логічних вентилів. Замість того, щоб чітко викладати схему NAND та інших вентилів, наші нейронні мережі можуть просто навчитися вирішувати проблеми, іноді проблеми, коли було б надзвичайно важко безпосередньо спроектувати звичайну схему.

Багатошарові перцептрони мережі широко застосовуються в інженерних додатках. Зміщена зважена сума вхідних даних для кожного обчислювального блоку (організованого у багатошарову мережу прямого зв'язку) передається через передатну функцію для обчислення відповідного результату. По суті, ANN розроблено визначення схеми взаємозв'язку, передатних функцій і процес навчання, описаний раніше. Розглядаючи повністю підключену нейронну мережу з прямим зв'язком, головне завдання з проектування взаємозв'язку шаблон вибирає кількість прихованих шарів і кількість одиниць у всіх шари. У багатьох випадках входи/виходи визначаються відповідно до проблеми.

Щодо результатів, зокрема на практичних прикладах регресійні програми, нейронні мережі часто мають лише один нейрон в їх вихідному шарі. У нейронних мережах з кількома вихідними змінними процес навчання може бути складнішим, оскільки множинні невідомі функції необхідно виявити одночасно. Це особливо важливо для обчислювальних блоків у мережі. Тому практично, найкращий підхід – це мати ансамбль менших мереж, кожен з однією вихідною змінною, і запустити їх окремо. Вибір оптимальної мережі здійснюється методом

проб і помилок. Це означає, що не існує попередньо визначеного правила для вибору кількості прихованих шарів. Проте більша кількість прихованих шарів підвищення продуктивності узагальнення MLP, збільшення часу обчислень, коли для навчання використовується алгоритм зворотного поширення. Тут проектування мережі MLP спочатку починається з розгляду лише одного прихованого шару, а потім інші шари розглядаються та порівнюються з першою мережею.

Одного прихованого шару достатньо, щоб достатньо добре наблизити основну функцію. Після того, як кількість прихованих шарів мережі вибрано, ще потрібно визначити параметри, такі як розмір прихованого шару, швидкість навчання, функція активації та мінімальний градієнт. Так само вони зазвичай визначаються експериментально, що означає – потрібно шукати великий простір параметрів. Прийняття рішення про розумну кількість прихованих одиниць досить складне і трудомістке. Залежно від типу мережі можна використовувати різні способи вибору найкращої кількості прихованих одиниць. Для цього створюються і навчаються подібні штучні нейронні мережі, але з різною кількістю прихованих нейронів.

Потім з мережі визначається оптимальна кількість прихованих блоків що дає найменшу помилку в підмножині перевірки.

Як правило, похибки оцінки та перевірки зменшуються протягом початкового етапу навчання. Однак, як зазначалося раніше, за рахунок збільшення кількості прихованих нейронів, внаслідок перенавчання, помилка в наборі перевірки збільшується. У всіх випадках навчання припиняється як тільки кількість ітерацій досягне максимальної межі, або обчислена помилка відповіді зменшується до значення, меншого за бажане значення.

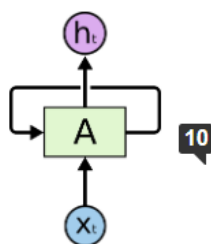
Навчання припиняється коли кількість ітерацій досягає певного числа або коли помилка відповіді зменшена до бажаного цільового значення [13].

У всіх випадках мережі ваги та зміщення вибираються як остаточні параметри моделі в точці, де досягається мінімальна помилка в підмножині

перевірки. Тому, кількість прихованих вузлів вибирається на основі того, коли була досягнута найменша помилка моделлю з такою кількістю вузлів.

2.3.2 РНН.

Люди не починають думати з нуля кожену секунду. Читаючи це речення, ми розуміємо кожне слово на основі нашого розуміння попередніх слів. Ми не кидаємо все і знову починаємо думати з нуля. Наші думки мають послідовність.



Recurrent Neural Networks have loops.

Рисунок 2.10 – Фрагмент рекурентної нейронної мережі

На наведеній вище діаграмі фрагмент нейронної мережі A переглядає вхідні дані x_t і виводить значення h_t . Цикл дозволяє передавати інформацію від одного кроку мережі до наступного.

Завдяки цим циклам повторювані нейронні мережі виглядають загадковими. Однак, якщо подумати трохи більше, то виявиться, що вони не так вже й відрізняються від звичайної нейронної мережі. Рекуррентну нейронну мережу можна розглядати як кілька копій однієї мережі, кожна з яких передає повідомлення наступнику [31].

Розглянемо, що станеться, якщо розгорнути цикл:

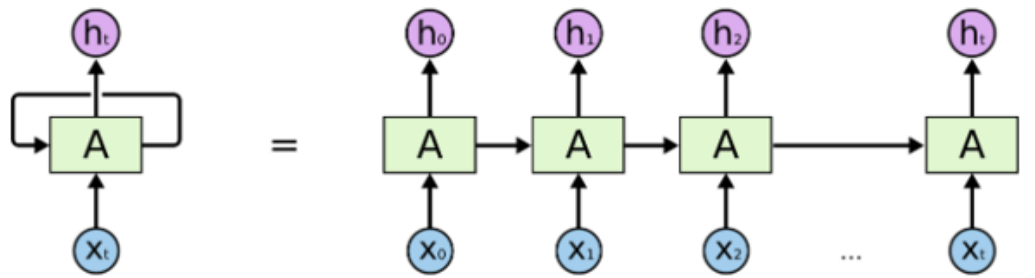


Рисунок 2.11 – Розгорнута нейронна мережа

Ця ланцюгова природа показує, що повторювані нейронні мережі тісно пов'язані з послідовностями і списками. Вони є природною архітектурою нейронної мережі для використання таких даних. За останні кілька років досягнуто неймовірного успіху в застосуванні РНН для вирішення різноманітних проблем: розпізнавання мовлення, моделювання мови, перекладу, підписів зображень.

Однією з привабливих переваг РНН є ідея, що вони можуть використати попередню інформацію до поточного завдання, наприклад, використання попередніх відеокадрів може допомогти зрозуміти поточний кадр. Іноді, щоб виконати поточне завдання, нам потрібно лише подивитися на останню інформацію. Якщо ми намагаємося передбачити останнє слово «хмари на небі», нам не потрібен додатковий контекст – цілком очевидно, що наступним словом буде небо. У таких випадках, коли розрив між відповідною інформацією та місцем, у якому вона потрібна, невеликий, РНН можуть навчитися використовувати минулу інформацію.

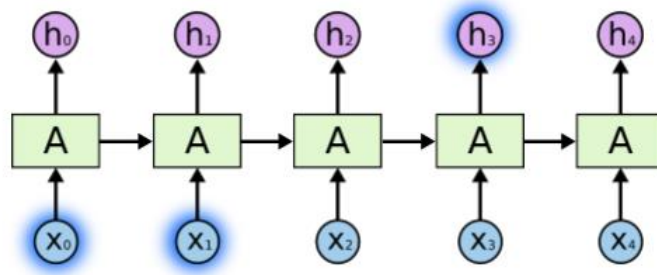


Рисунок 2.12 – Використання РНН минулої інформації

Але є також випадки, коли нам потрібно більше контексту. Спробуймо передбачити останнє слово в тексті «Я виріс у Франції... я вільно розмовляю французькою». Останні відомості свідчать про те, що наступне слово, ймовірно, є назвою мови, але якщо ми хочемо звузити мову, нам потрібен контекст Франції з дальнього краю. Цілком можливо, що розрив між відповідною інформацією та точкою, де вона потрібна, стане дуже великим.

На жаль, оскільки цей розрив зростає, РНН не можуть навчитися зв'язувати інформацію.

2.3.3 ЛСТМ.

Мережі довгострокової пам'яті, які зазвичай називають просто ЛСТМ – це особливий тип РНН, здатний вивчати довгострокові залежності. Вони були представлені Hochreiter і Schmidhuber (1997), і були вдосконалені та популяризовані багатьма людьми в наступних роботах. Вони надзвичайно добре працюють над широким спектром проблем і зараз широко використовуються. ЛСТМ явно розроблені, щоб уникнути довгострокової проблеми залежності. Запам'ятовування інформації протягом тривалого періоду часу – це практично їхня поведінка за замовчуванням, а не те, що їм важко вивчити.

Усі рекурентні нейронні мережі мають вигляд ланцюга повторюваних модулів нейронної мережі. У стандартних РНН цей повторюваний модуль матиме дуже просту структуру, таку як один шар \tanh .

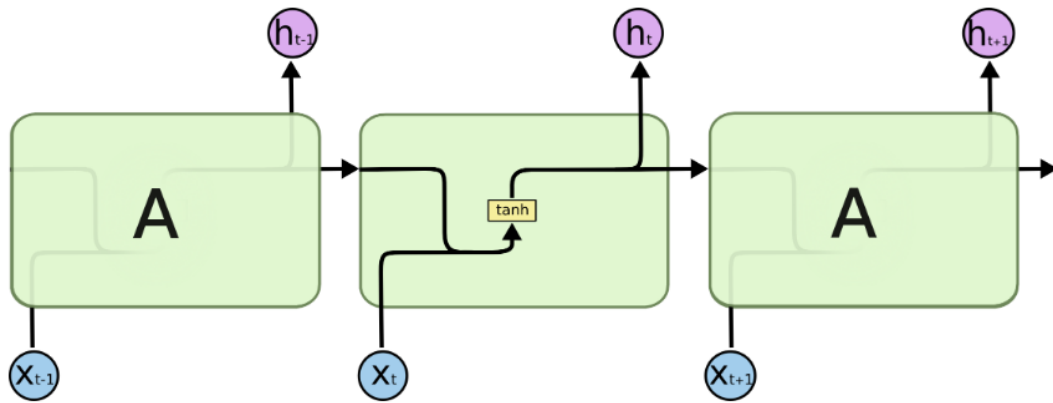
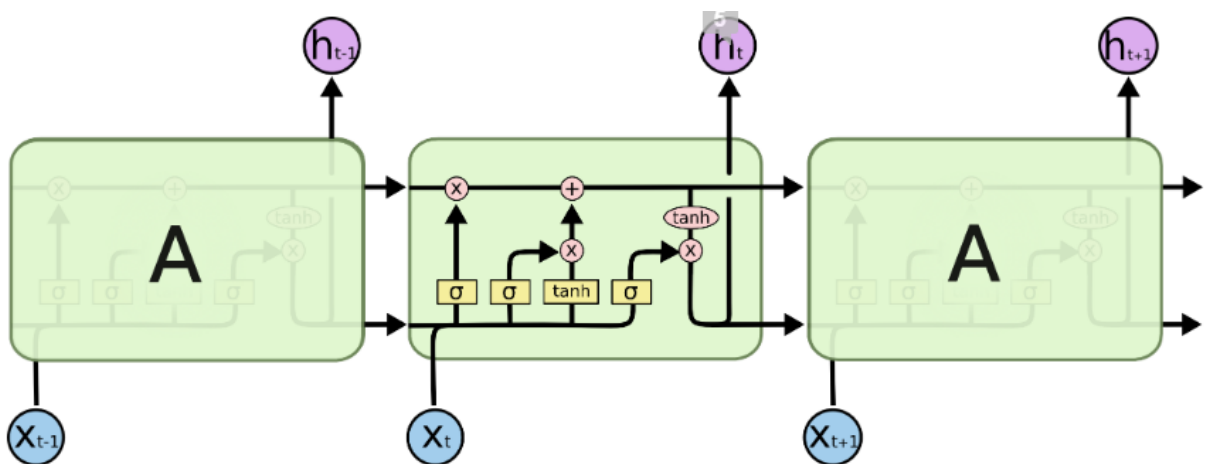


Рисунок 2.13– Повторювальний модуль

Повторюваний модуль у стандартній РНН містить один шар.

ЛСТМ також мають таку ланцюжкову структуру, але повторюваний модуль має іншу структуру. Замість того, щоб мати один шар нейронної мережі, є чотири, які взаємодіють дуже особливим чином.



The repeating module in an LSTM contains four interacting layers.

Рисунок 2.14 – Повторювальний модуль в ЛСТМ з чотирма ітеративними шарами

У наведеній вище діаграмі кожен рядок несе цілий вектор від виходу одного вузла до входу інших. Рожеві кола представляють точкові операції, як-от додавання векторів, а жовті квадрати — це вивчені шари нейронної мережі. Злиття рядків позначає конкатенацію, а розгалуження рядків означає, що його вміст копіюється, а копії спрямовуються в різні місця.

Ключем до LSTM є стан комірки, горизонтальна лінія, що проходить через верхню частину діаграми.

Стан клітини нагадує конвеєрну стрічку. Він проходить прямо по всьому ланцюжку, лише з деякими незначними лінійними взаємодіями. Інформації дуже легко просто текти по ній без змін.

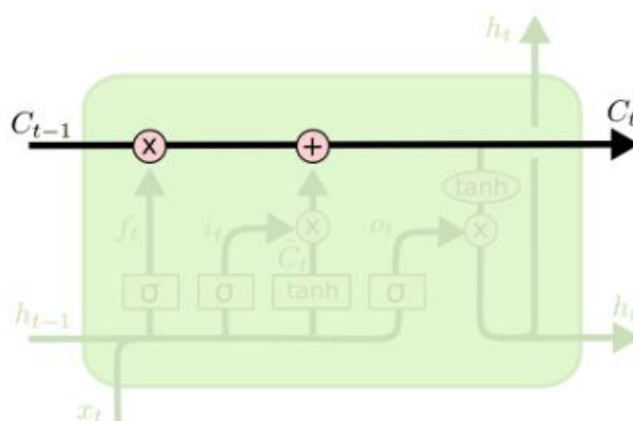


Рисунок 2.15 – Зміна стану системи

LSTM дійсно має можливість видаляти або додавати інформацію до стану вузла, ретельно регульованого структурами, які називаються воротами.

Ворота – це спосіб додаткового пропускання інформації. Вони складаються із шару сигмовидної нейронної мережі та операції поточкового множення.

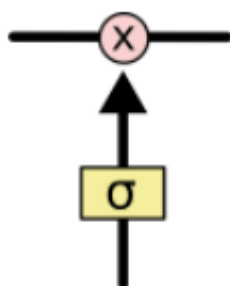


Рисунок 2.16– Схема додаткового пропускання інформації

Сигмовидний шар виводить числа від нуля до одиниці, описуючи, скільки кожного компонента слід пропустити. Нульове значення означає «нічого не пропустити», а значення одиниці означає «пропустити все!»

ЛСТМ має троє таких воріт для захисту та контролю стану клітини.

Здійснимо покроковий огляд ЛСТМ.

Перший крок у нашому ЛСТМ — вирішити, яку інформацію ми збираємось викинути зі стану клітинки. Це рішення приймається сигмовидним шаром, який називається «шар воріт забути». Він розглядає h_{t-1} і x_t і виводить число від 0 до 1 для кожного числа в стані комірki C_{t-1} . 1 означає «повністю зберегти це», а 0 означає «повністю позбутися цього».

Повернемося до нашого прикладу мовної моделі, яка намагається передбачити наступне слово на основі всіх попередніх. У такій задачі стан клітинки може включати стать теперішнього суб'єкта, щоб можна було використовувати правильні займенники.

Коли ми бачимо новий предмет, ми хочемо забути ознаки старого предмета.

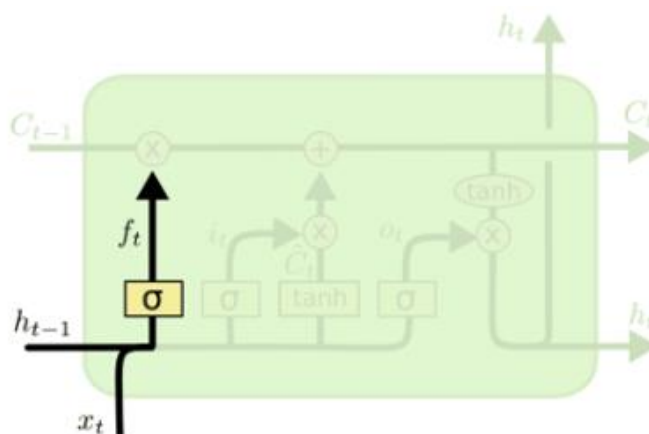


Рисунок 2.17 – Крок алгоритму

Наступний крок — вирішити, яку нову інформацію ми будемо зберігати в стані клітинки. Це має дві частини. По-перше, сигмовидний шар, який називається «вхідним шаром шлюза», вирішує, які значення ми будемо оновлювати. Далі шар \tanh створює вектор нових значень-кандидатів, C_t , які можна додати до стану. На наступному кроці ми об'єднаємо ці два, щоб створити оновлення стану.

У прикладі нашої мовної моделі ми б хотіли додати стать нового суб'єкта до стану клітинки, щоб замінити старий, який ми забуваємо.

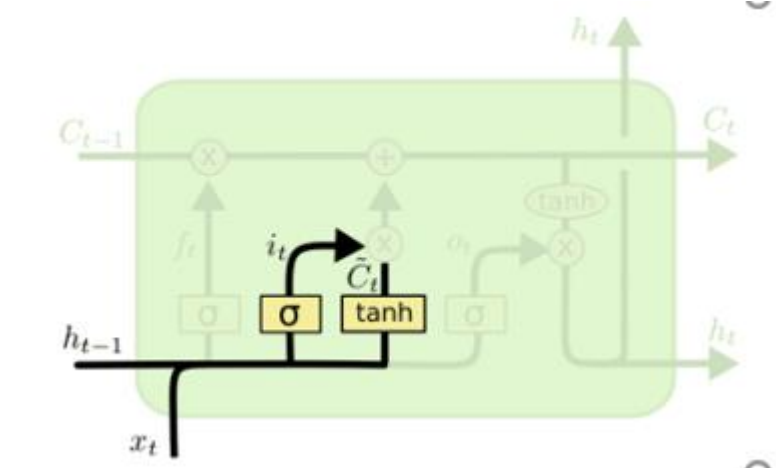


Рисунок 2.18 – Додавання нової інформації

Тепер настав час оновити старий стан комірки, C_{t-1} , на новий стан комірки C_t .

Ми множимо старий стан на f_t , забуваючи речі, які вирішили забути раніше. Потім додаємо його C_t . Це нові значення-кандидати, масштабовані відповідно до того, наскільки ми вирішили оновити значення кожного стану.

У випадку з мовною моделлю ми фактично скидаємо інформацію про стать старого суб'єкта та додаємо нову інформацію, як ми вирішили на попередніх кроках.

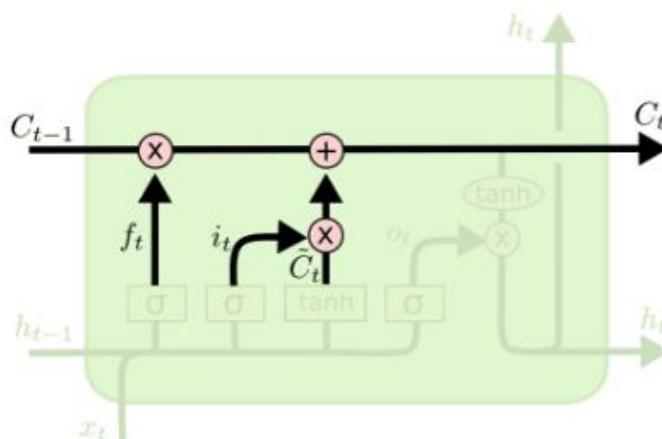


Рисунок 2.19 – Принцип передачі інформації про новий об'єкт

Нарешті, нам потрібно вирішити, що ми збираємося вивести. Цей вихід буде заснований на стані нашої клітинки, але буде відфільтрованою версією. Спочатку ми запускаємо сигмовидний шар, який вирішує, які частини стану клітинки ми збираємося вивести. Потім ми переносимо стан комірki через \tanh (щоб значення були між -1 і 1) і множимо його на вихід сигмоподібних воріт, щоб ми виводили лише ті частини, які ми вирішили.

Для прикладу мовної моделі, оскільки він щойно побачив тему, він може захотіти вивести інформацію, що стосується дієслова, якщо це буде те, що буде далі. Наприклад, він може вивести, чи є підмет однини чи множини, щоб ми знали, у яку форму дієслово має відмінюватися, якщо це те, що буде далі.

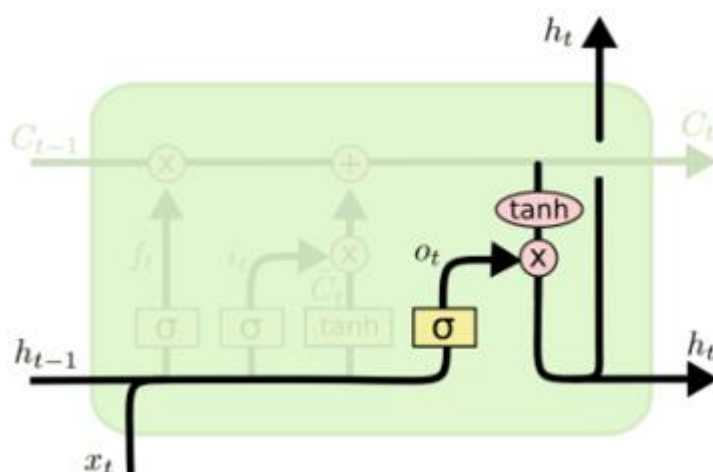


Рисунок 2.20 – Виведення інформації

Є ряд різновидів ЛСТМ.

Один з популярних варіантів ЛСТМ, представлений Gers і Schmidhuber (2000), Це означає, що ми дозволяємо шарам затвора дивитися на стан комірki.

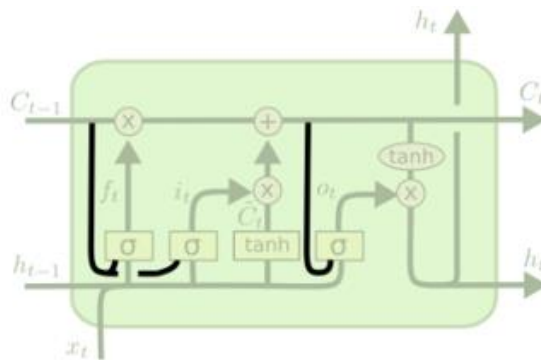


Рисунок 2.21 – Архітектура Gers і Schmidhuber

Вищенаведена діаграма додає вічки до всіх воріт, але багато документів дають одні вічки, а інші ні.

Інший варіант полягає у використанні поєднаних шлюзів забуття та введення. Замість того, щоб окремо вирішувати, що забути і до чого додати нову інформацію, ми приймаємо ці рішення разом. Ми забуваємо лише тоді, коли збираємося щось ввести на місце. Ми вводимо нові значення в стан лише тоді, коли забуваємо щось старше.

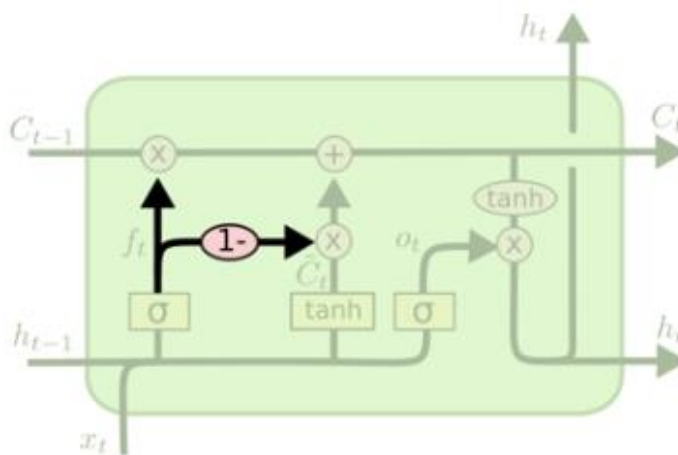


Рисунок 2.22 – Принци роботи поєднаних шлюзів з пам'яттю

Також варіацією ЛСТМ є Рекурентні вузли на виходах, представлений Чо та ін. (2014). Він поєднує ворота забуття та введення в єдиний «шлюз оновлення». Він також об'єднує стан комірки і прихований стан і вносить деякі інші зміни. Отримана модель є простішою, ніж стандартні моделі ЛСТМ, і стає все популярнішою.

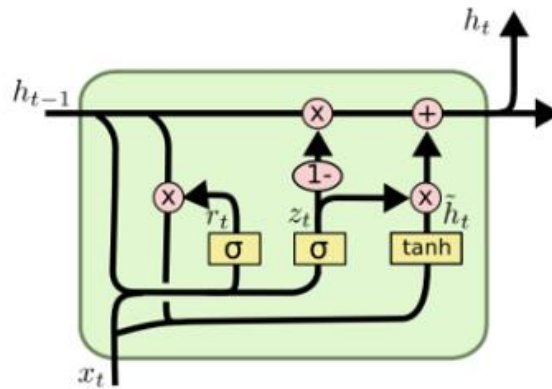


Рисунок 2.23 – Рекурентна LSTM при вузлах-воротах

Це лише деякі з найбільш помітних варіантів LSTM. Йозефович та ін. перевірили понад десять тисяч архітектур РНН, знайшовши деякі, які працювали краще, ніж LSTM для певних завдань. Проте варта зазначити, що за допомогою LSTM досягається ряд чудових результатів та є надзвичайно перспективними.

2.4 Висновки до другого розділу

У другому розділі нами висвітлено поняття нейронної мережі, проаналізовано існуючі нейронні мережі, методи машинного навчання для процесу газоспоживання. Зокрема висвітлено питання одношарового та багатошарового перцептрона, рекурентних нейронних мереж, LSTM, як різновиду РНН.

Опрацьований матеріал є основою для використання наявних наукових здобутків у сфері нейронних мереж при прогнозуванні процесу газоспоживання міста Тернополя.

3 ПОПЕРЕДНЄ ОПРАЦЮВАННЯ СТАТИСТИЧНИХ ДАНИХ ГАЗОСПОЖИВАННЯ ТА МЕТЕОФАКТОРІВ ДЛЯ МІСТА ТЕРНОПІЛЬ.

3.1 Мова програмування Python та бібліотеки що будуть використані в дослідженні.

Python – це мова комп'ютерного програмування, яка часто використовується для створення веб-сайтів і програмного забезпечення, автоматизації завдань і аналізу даних. Python — мова загального призначення. Це означає, що її можна використовувати для створення різноманітних програм і вона не спеціалізується на будь-яких конкретних проблемах. Ця універсальність разом із зручністю для початківців зробили її однією з найбільш використовуваних мов програмування сьогодні. Опитування, проведене аналітичною фірмою RedMonk, показало, що це друга за популярністю мова програмування серед розробників у 2021 році [60].

Стандартна бібліотека Python містить точний синтаксис, семантику та токени Python. Вона має вбудовані модулі, які надають доступ до базової функціональності системи, як-от вводу-виводу та деяких інших основних модулів. Більшість бібліотек Python написані мовою програмування C. Стандартна бібліотека Python складається з понад 200 основних модулів. Усе це працює разом, щоб зробити Python мовою програмування високого рівня. Стандартна бібліотека Python відіграє дуже важливу роль. Без цього програмісти не зможуть отримати доступ до функцій Python. Поряд з тим в Python є кілька інших бібліотек, які полегшують життя програміста. Проаналізуємо деякі з найбільш використовуваних бібліотек:

TensorFlow: ця бібліотека була розроблена Google у співпраці з Brain Team. Дана бібліотека з відкритим вихідним кодом, яка використовується для високорівневих обчислень, в алгоритмах машинного навчання та глибинного навчання. Вона містить велику кількість тензорних операцій. Дослідники також

використовують цю бібліотеку Python для вирішення складних обчислень з математики та фізики.

Matplotlib: ця бібліотека відповідає за побудову числових даних і тому вона використовується в аналізі даних. Це також бібліотека з відкритим вихідним кодом, яка відображає чіткі фігури, такі як кругові діаграми, гістограми, діаграми розсіювання, графіки тощо.

Pandas: Pandas є важливою бібліотекою для наукових даних. Це бібліотека машинного навчання з відкритим вихідним кодом, яка надає гнучкі високорівневі структури даних і різноманітні інструменти аналізу, що полегшує аналіз даних, маніпулювання даними та очищення даних. Pandas підтримує такі операції, як сортування, повторна індексація, ітерація, конкатенація, перетворення даних, візуалізації, агрегації тощо.

Numpy: назва «Numpy» означає «Числовий Python» – найпоширеніша бібліотека. Популярна бібліотека машинного навчання, яка підтримує великі матриці та багатовимірні дані. Складається з вбудованих математичних функцій для легких обчислень. Навіть такі бібліотеки, як TensorFlow, використовують Numpy для виконання кількох операцій над тензорами. Інтерфейс масиву є однією з ключових функцій цієї бібліотеки.

SciPy: назва «SciPy» означає «Scientific Python». Це бібліотека з відкритим вихідним кодом, яка використовується для наукових обчислень високого рівня. Вона побудована на основі розширення Numpy та працює з Numpy для обробки складних обчислень. Хоча Numpy дозволяє сортувати та індексувати дані масиву, числовий код даних зберігається в SciPy. Дана бібліотека також широко використовується розробниками та інженерами додатків.

Scikit-learn: відома бібліотека Python для роботи зі складними даними. Scikit-learn — це бібліотека з відкритим вихідним кодом, яка підтримує машинне навчання. Вона підтримує різноманітні контрольовані та неконтрольовані алгоритми, такі як лінійна регресія, класифікація, кластеризація тощо. Ця бібліотека працює разом із Numpy та SciPy.

Seaborn — бібліотека для створення статистичної графіки на Python. Побудована на основі matplotlib і тісно інтегрується зі структурами даних pandas. Seaborn допомагає нам досліджувати та розуміти наші дані. Її функції побудови графіків працюють з фреймами даних і масивами, що містять цілі набори даних, і внутрішньо виконують необхідне семантичне відображення та статистичну агрегацію для створення інформативних графіків. Декларативний API бібліотеки, орієнтований на набір даних, дозволяє зосередитися на тому, що означають різні елементи графіків, а не на деталях того, як їх намалювати [39].

3.2 Перелік метеофакторів, які будуть використані в дослідженні

Для того щоб спрогнозувати кількість газоспоживання на річний чи денний проміжок часу в роботі використано ряд метеофакторів. Статистичні спостереження взято за період з 2006 до 2019 року в м. Тернопіль. Джерелом інформації є сайт gr5.ua.

Вхідними даними дослідження виступають: температура повітря на висоті 2 м над поверхнею землі (в градусах Цельсія) – T , атмосферний тиск на рівні станції – P_0 , атмосферний тиск до середнього рівня моря – P , тенденція зміни атмосферного тиску за останні 3 години – P_a , відносна вологість повітря на висоті 2м над землею в % – U , швидкість вітру на висоті 10 – 12 м над поверхністю землі, середня за 10 хв період, одиниці вимірювання м/с – F_f , максимальне значення пориву вітру на висоті 10–12 м над поверхнею землі за 10 хв період – F_{f10} , максимальне значення пориву вітру на висоті 10–12 метрів над поверхнею землі в м/с – F_{f3} , мінімальна температура повітря за попередній період (не більше 12 годин) – T_n , максимальна температура повітря за попередній період (не більше 12 годин) – T_x , горизонтальна дальність оглядовості в км – V_v , температура точки роси на висоті 2 метри – T_d , кількість опадів в мм – R_{RR} , кількість часу, за який випала та кількість опадів – t_R ,

мінімальна температура поверхності землі за ніч – T_g , висота снігового покриву в см – S_{SS} . Також отримано дані газоспоживання Q в мЗ.

```
1 data.head()
```

	id	T	Po	P	Pa	U	Ff	ff10	ff3	Tn	Tx	WV	Td	RRR	tR	Tg	sss	Q
0	2626	-2.8	732.7	763.5	0.3	82.0	3.0	19.0	18.0	-4.550	0.6250	10.0	-5.5	1.76	10.00000	-2.0	1.0	38551.7
1	2627	-4.5	733.3	764.4	0.3	86.0	2.0	19.0	18.0	-4.775	0.7375	10.0	-6.5	1.72	9.66667	-2.0	1.0	42794.1
2	2628	-4.1	733.3	764.4	0.3	89.0	3.0	19.0	18.0	-5.000	0.8500	10.0	-5.7	1.68	9.33333	-2.0	1.0	52405.8
3	2629	-1.8	734.4	765.3	0.3	79.0	3.0	19.0	18.0	-4.650	0.9625	10.0	-4.9	1.64	9.00000	-2.0	1.0	45367.2
4	2630	1.2	733.4	763.9	0.3	74.0	3.0	19.0	18.0	-4.300	1.0750	10.0	-3.0	1.60	8.66667	-2.0	1.0	42925.1

Рисунок 3.1 – частина таблиці з вхідними даними дослідження

3.3 Попередній аналіз даних газоспоживання

На першому етапі аналізу даних газоспоживання нами виконано імпорт всіх необхідних бібліотек для подальшої зручності роботи з даними.

На вхід дані газоспоживання були подані в форматі csv та розміщені на GoogleDrive. Для можливості в подальшому виконувати над ними операції було виконано підключення за допомогою функції mount() класу drive до GoogleDisk та імпортовано дані функцією read_csv() з бібліотеки Pandas.

```
1 from google.colab import drive
2 drive.mount("/content/drive", force_remount=True)
3
```

Рисунок 3.2– Підключення до GoogleDisk

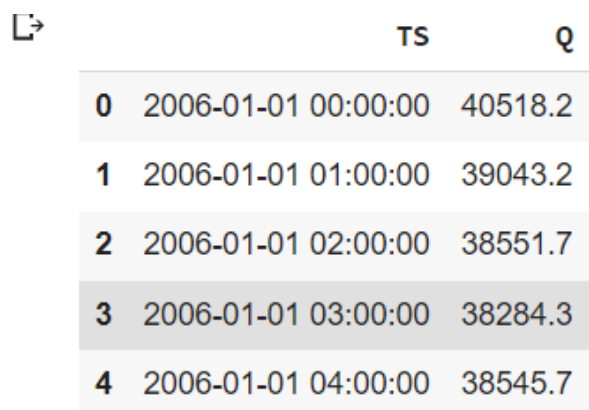
```
1 gas_file = Path("/content/drive/MyDrive/2006-2019(+day_of_week)_t_gaz_sum_hours_full.csv")
2 wether_file = Path("/content/drive/MyDrive/2006-2019(3hour)_meteo_rp5ua.csv")
```

Рисунок 3.3 – Присвоєння змінним шляху до даних

```
1 gas_df = pd.read_csv(gas_file)
```

Рисунок 3.4 – Імпортування даних газоспоживання

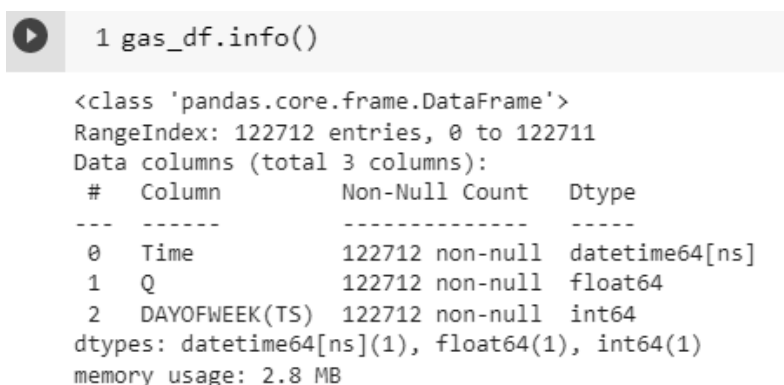
Після імпортування даних розпочато їхнє опрацювання. Зображено початок таблиці з допомогою функції `head()`.



	TS	Q
0	2006-01-01 00:00:00	40518.2
1	2006-01-01 01:00:00	39043.2
2	2006-01-01 02:00:00	38551.7
3	2006-01-01 03:00:00	38284.3
4	2006-01-01 04:00:00	38545.7

Рисунок 3.5– Дані газоспоживання погодинно у Тернополі з 2006 по 2019 рік

Користуючись функцією `.info()` отримано загальну інформацію про датасет.



```
1 gas_df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 122712 entries, 0 to 122711
Data columns (total 3 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Time                  122712 non-null  datetime64[ns]
1   Q                      122712 non-null  float64
2   DAYOFWEEK(TS)         122712 non-null  int64
dtypes: datetime64[ns](1), float64(1), int64(1)
memory usage: 2.8 MB
```

Рисунок 3.6– Дані газоспоживання погодинно

Основним показником є кількість спостережень, а саме інформація про газоспоживання протягом 122712 годин. До колонок таблиці входять час у форматі «Рік–місяць–день–година», кількість споживаного природнього газу в м3 та день тижня. Кількість використаної пам'яті становить 2.8мб.

Для того, щоб краще зрозуміти поведінку даних газоспоживання зобразимо декілька графіків.

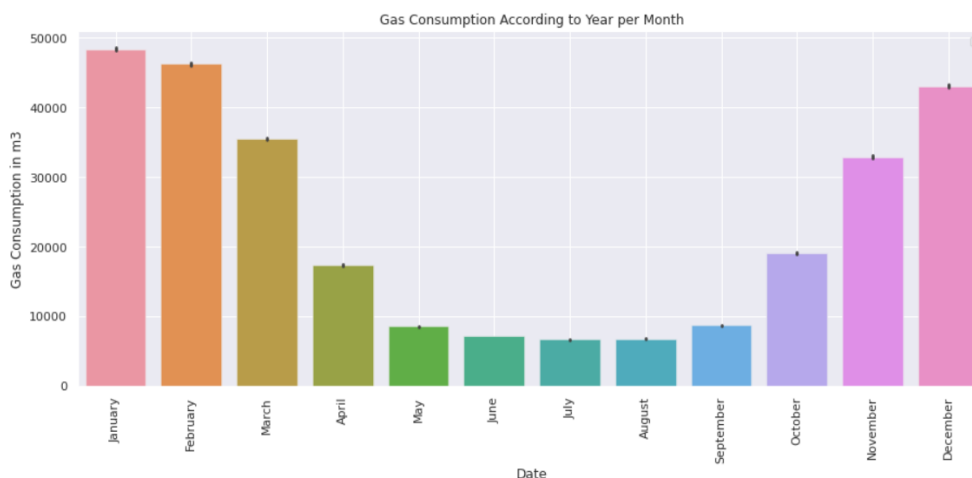


Рисунок 3.7 – Кількість газоспоживання в Тернополі з 2006 по 2019 рік по місяцях за годину

Аналіз даних підтверджує загальноприйняту тенденцію максимального газоспоживання у зимові місяці (січень, лютий, грудень) з наступним зменшенням в весняно – осінній період (березень, листопад, жовтень) та малою часткою газоспоживання у літні місяці.

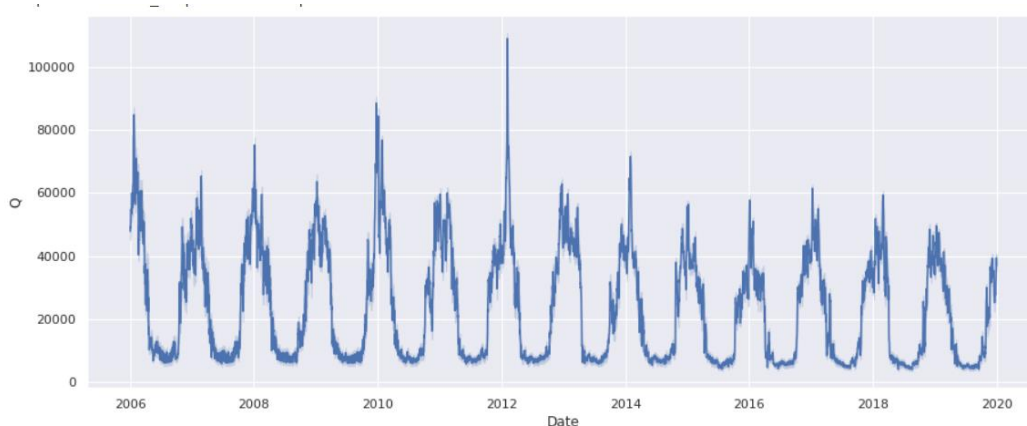


Рисунок 3.8 – Поведінка даних газоспоживання в Тернополі з 2006 по 2019 рік

Інформація, зображена на даному рисунку, свідчить про відносну стабільність газоспоживання в Тернополі протягом зазначеного періоду – приблизно на рівні 6000 м³, проте у 2006, 2008, 2010, 2012 та 2014 роках спостерігається значно більше газоспоживання, особливо у 2012 році. Натомість у 2019 році помітна тенденція до зменшення газоспоживання.

3.4 Попередній аналіз даних метеофакторів

З таблиці видно що спостереження за метеофакторами здійснювались кожні 3 години. Вхідні параметри описані в розділі 3.1

1 weather_df.head(10)

	id	Time	T	Po	P	Pa	U	Ff	ff10	ff3	Tn	Tx	WV	Td	RRR	tR	Tg	sss
0	2626	2006-01-01 02:00:00	-2.8	732.7	763.5	0.3	82.0	3.0	19.0	18.0	-4.550	0.6250	10.0	-5.5	1.760000	10.00000	-2.0	1.0
1	2627	2006-01-01 05:00:00	-4.5	733.3	764.4	0.3	86.0	2.0	19.0	18.0	-4.775	0.7375	10.0	-6.5	1.720000	9.66667	-2.0	1.0
2	2628	2006-01-01 08:00:00	-4.1	733.3	764.4	0.3	89.0	3.0	19.0	18.0	-5.000	0.8500	10.0	-5.7	1.680000	9.33333	-2.0	1.0
3	2629	2006-01-01 11:00:00	-1.8	734.4	765.3	0.3	79.0	3.0	19.0	18.0	-4.650	0.9625	10.0	-4.9	1.640000	9.00000	-2.0	1.0
4	2630	2006-01-01 14:00:00	1.2	733.4	763.9	0.3	74.0	3.0	19.0	18.0	-4.300	1.0750	10.0	-3.0	1.600000	8.66667	-2.0	1.0
5	2631	2006-01-01 17:00:00	1.0	733.6	764.0	0.3	75.0	6.0	19.0	18.0	-3.950	1.1875	10.0	-2.9	1.560000	8.33333	-2.0	1.0
6	2632	2006-01-01 20:00:00	-0.6	733.6	764.2	0.3	78.0	5.0	19.0	18.0	-3.600	1.3000	10.0	-4.0	1.520000	8.00000	-2.0	1.0
7	2633	2006-01-01 23:00:00	-1.9	733.8	764.6	0.3	81.0	5.0	19.0	18.0	-3.250	1.5500	10.0	-4.7	1.480000	7.66667	-2.0	1.0
8	2874	2006-01-02 02:00:00	-1.6	733.0	763.9	0.3	94.0	5.0	19.0	18.0	-3.050	-1.6500	4.0	-2.4	0.483333	6.00000	-2.0	3.0
9	2875	2006-01-02 05:00:00	-1.2	732.5	763.3	0.3	93.0	8.0	19.0	18.0	-2.825	-1.2750	4.0	-2.2	0.475000	9.00000	-2.0	3.0

Рисунок 3.9 – Метеодані з інтервалом 3 години у Тернополі з 2006 по 2019 рік

З допомогою функції describe() на таблиці зображено мінімальне, максимальне, середнє значення по кожному параметрі, також середнє квадратичне відхилення та значення для 25 і 75% вибірки.

	id	T	Po	P	Pa	\
count	41052.000000	41052.000000	41052.000000	41052.000000	41052.000000	
mean	23151.500000	8.434419	733.181666	762.820318	0.148344	
std	11850.835962	9.958278	5.707831	6.200644	0.571741	
min	2626.000000	-32.200000	703.800000	732.600000	-14.700000	
25%	12888.750000	1.000000	729.700000	758.800000	0.000000	
50%	23151.500000	8.500000	733.300000	762.700000	0.300000	
75%	33414.250000	16.000000	736.700000	766.600000	0.300000	
max	43677.000000	34.600000	754.300000	789.100000	15.700000	

	U	Ff	ff10	ff3	Tn	\
count	41052.000000	41052.000000	41052.000000	41052.000000	41052.000000	
mean	77.135767	3.843430	18.519012	19.338169	4.441564	
std	17.537281	2.476638	0.600688	1.387916	8.245509	
min	13.000000	0.000000	18.000000	18.000000	-33.300000	
25%	67.000000	2.000000	18.000000	18.250000	-0.765625	
50%	83.000000	4.000000	18.300200	19.000000	4.812500	
75%	91.000000	5.000000	18.915650	19.979800	11.137500	
max	100.000000	20.000000	23.000000	27.000000	22.400000	

	Tx	VV	Td	RRR	tR	\	Tg	sss
count	41052.000000	41052.000000	41052.000000	41052.000000	41052.000000		count	41052.000000
mean	12.701023	8.939914	4.071782	2.430168	10.278549		mean	-0.458309
std	10.885235	4.555583	8.192065	3.407527	2.095693		std	2.565736
min	-20.300000	0.050000	-33.300000	0.100000	3.000000		min	-11.000000
25%	3.400000	10.000000	-1.300000	0.579262	9.000000		25%	-2.000000
50%	13.625000	10.000000	4.300000	1.133575	12.000000		50%	-2.000000
75%	21.975000	10.000000	10.600000	2.897232	12.000000		75%	1.750000
max	35.200000	71.000000	22.300000	49.000000	18.000000		max	7.000000

Рисунок 3.10 – Описова статистика метеофакторів в Тернополі з 2006 по 2019 рік

Проведемо аналіз метеофакторів .

```
[594] 1 weather_df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 41052 entries, 0 to 41051
Data columns (total 18 columns):
#   Column  Non-Null Count  Dtype
---  -
0   id      41052 non-null  int64
1   Time    41052 non-null  datetime64[ns]
2   T        41052 non-null  float64
3   Po      41052 non-null  float64
4   P        41052 non-null  float64
5   Pa      41052 non-null  float64
6   U        41052 non-null  float64
7   Ff      41052 non-null  float64
8   ff10    41052 non-null  float64
9   ff3     41052 non-null  float64
10  Tn      41052 non-null  float64
11  Tx      41052 non-null  float64
12  VV      41052 non-null  float64
13  Td      41052 non-null  float64
14  RRR     41052 non-null  float64
15  tR      41052 non-null  float64
16  Tg      41052 non-null  float64
17  sss     41052 non-null  float64
dtypes: datetime64[ns](1), float64(16), int64(1)
memory usage: 5.6 MB
```

Рисунок 3. 11– Дані метеофакторів, крок вимірювань 3години

Стосовно даних газоспоживання спостерігається значно менша вибірка у розмірі 41052 вхідних даних по кожному з параметрів. Згідно з колонкою «Count» зрозуміло, що порожніх значень по кожному з параметрів немає. Усі необхідні параметри подані з типом float що відповідає їх сутності. Також дані часу подані у відповідному форматі, тому згідно з попереднім аналізом дані перетворенню не підлягають.

Після злиття двох таблиць проведено інтерполяцію для зведення показників до інтервалу спостережень в 1 год . При виконанні даного завдання

застосовано функцію interpolate(), яка заповнювала втрачені дані на основі середнього по двох сусідніх показниках. По завершенні виконання функції отримано таблицю виду

	Time	Q	DAYOFWEEK(TS)	id	T	Po	P	Pa	U	Ff	ff10	ff3	Tn	Tx	W	Td	RRR	tR	Tg	sss
122702	2019-12-31 14:00:00	14000.1	4.0	43565.666667	2.133333	733.700000	763.966667	0.066667	96.333333	0.666667	18.0	22.271567	1.450000	12.129167	10.0	1.533333	1.0	12.0	-3.0	2.0
122703	2019-12-31 15:00:00	14716.5	4.0	43566.000000	1.900000	733.700000	764.000000	0.000000	97.000000	0.000000	18.0	22.276800	1.500000	12.075000	10.0	1.400000	1.0	12.0	-3.0	2.0
122704	2019-12-31 16:00:00	21679.6	4.0	43566.333333	1.966667	733.666667	763.966667	-0.033333	95.666667	1.333333	18.0	22.282000	1.550000	12.020833	10.0	1.300000	1.0	12.0	-3.0	2.0
122705	2019-12-31 17:00:00	25094.3	4.0	43566.666667	2.033333	733.633333	763.933333	-0.066667	94.333333	2.666667	18.0	22.287200	1.600000	11.966667	10.0	1.200000	1.0	12.0	-3.0	2.0
122706	2019-12-31 18:00:00	26563.3	4.0	43567.000000	2.100000	733.600000	763.900000	-0.100000	93.000000	4.000000	18.0	22.292400	1.650000	11.912500	10.0	1.100000	1.0	12.0	-3.0	2.0
122707	2019-12-31 19:00:00	28844.6	4.0	43567.333333	2.166667	733.700000	764.000000	0.033333	93.000000	3.000000	18.0	22.297633	1.700000	11.858333	10.0	1.166667	1.0	12.0	-3.0	2.0
122708	2019-12-31 20:00:00	29502.1	4.0	43567.666667	2.233333	733.800000	764.100000	0.166667	93.000000	2.000000	18.0	22.302867	1.750000	11.804167	10.0	1.233333	1.0	12.0	-3.0	2.0
122709	2019-12-31 21:00:00	30441.7	4.0	43568.000000	2.300000	733.900000	764.200000	0.300000	93.000000	1.000000	18.0	22.308100	1.800000	11.750000	10.0	1.300000	1.0	12.0	-3.0	2.0
122710	2019-12-31 22:00:00	28379.4	4.0	43568.333333	3.466667	734.166667	764.333333	0.466667	92.000000	0.666667	18.0	22.313333	1.733333	11.695833	10.0	2.266667	1.0	12.0	-3.0	2.0
122711	2019-12-31 23:00:00	21092.4	4.0	43568.666667	4.633333	734.433333	764.466667	0.633333	91.000000	0.333333	18.0	22.318567	1.666667	11.641667	10.0	3.233333	1.0	12.0	-3.0	2.0

Рисунок 3.12 – Описова статистика метеофакторів в Тернополі з 2006 по 2019 рік

Проведемо їх аналіз.

Після отримання повної таблиці з даними ключовим завданням, було обрано позбутися надлишковості. Так на основі кореляційної матриці та її зображеної на графіку проведено кореляційний аналіз, при якому дані, що несуть схожу інформацію, такі як (різного роду температури, тиск, пориви вітру) було вилучено з таблиці.

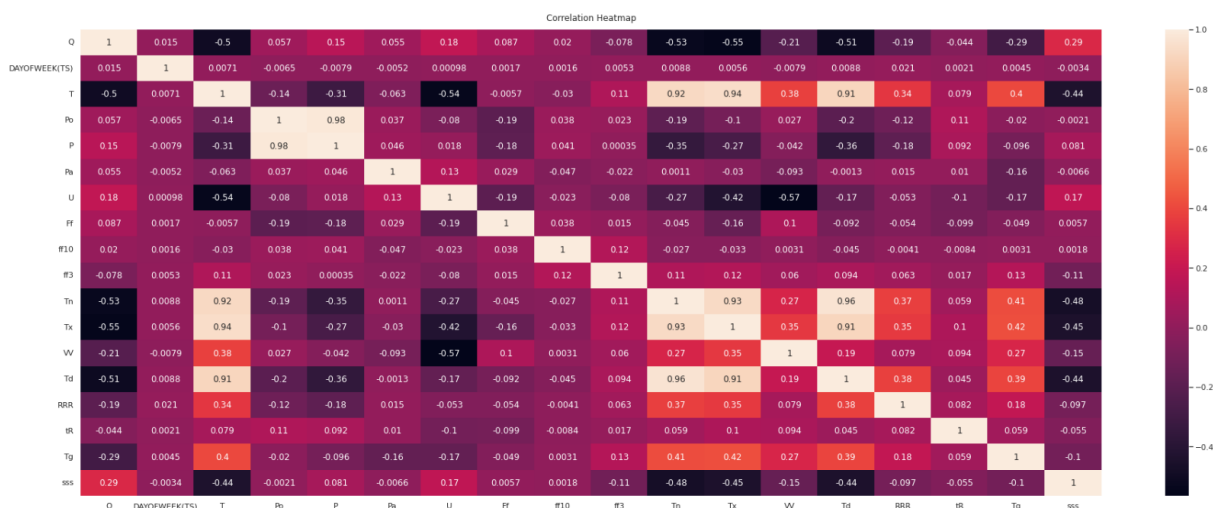


Рисунок 3.13– Матриця кореляції

Після проведеної роботи таблиця виглядає так:

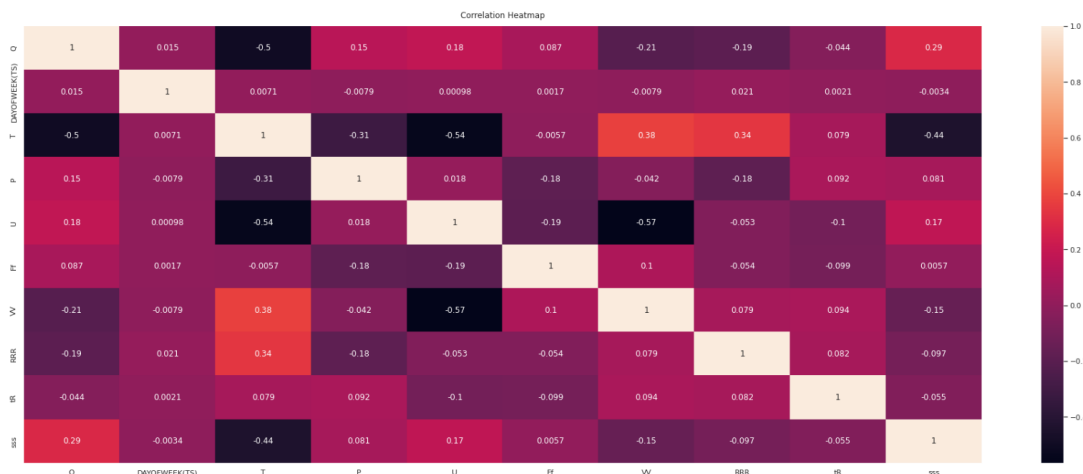


Рисунок 3.14– Матриця кореляції

На основі кореляційного аналізу відібрано метеорологічні дані для тренування нейронної мережі:

температура повітря на висоті 2 м над поверхнею землі; атмосферний тиск до середнього рівня моря; відносна вологість повітря на висоті 2 м над землею в %; швидкість вітру на висоті 10–12 м над поверхністю землі, середня за 10 хв період, одиниці вимірювання м/с; горизонтальна дальність оглядовості в км; кількість опадів в мм ; кількість часу за який випала та кількість опадів; висота снігового покриву в см.

Здійснимо кореляційний аналіз між тими метеофакторами, що залишилися та газоспоживанням. Аналіз свідчить про низьку кореляцію опрацьованих нами метеофакторів, крім температури повітря на висоті 2 м над поверхнею землі, з газоспоживанням. В наступному розділі побачимо, як це впливає на прогноз газоспоживання з допомогою моделей одношарового та багатшарового перцептрона. Як альтернативу здійснимо тренування з допомогою нейронної мережі LSTM та порівняємо отримані результати. Варта зазначити, що нейронна мережа LSTM не потребує таких вхідних даних, як метеофактори, а приймає на вхід лиш зміну газоспоживання.

Здійснивши відбір параметрів, що найбільш сприяють правильному тренуванню моделі, ми виконали нормалізацію даних з допомогою функції

MinMaxScaler() та звели дані до вигляду в діапазоні від 0 до 1 для кращого розуміння машиною.

3.5 Висновки до третього розділу.

У третьому розділі нами висвітлено питання мови програмування Python, як другої за популярністю мови програмування серед розробників у 2021 році, окреслено бібліотеки Python, що використані в дослідженні, розкрито перелік вхідних даних дослідження та джерела інформації, здійснено попереднє опрацювання статистичних даних газоспоживання та метеофакторів для міста Тернопіль. На основі попереднього аналізу даних, оцінки їх природи, провівши кореляцію даних, нам вдалося позбутися надлишкових параметрів, які суттєво не впливають на результати дослідження, а, навпаки, перевантажують його. Нами також була проведена нормалізація даних та підготовлено весь набір вхідних даних для безпосереднього проведення побудови моделі нейронних мереж процесу газоспоживання міста Тернополя з врахуванням метеофакторів.

4. ПОБУДОВА МОДЕЛЕЙ НЕЙРОННИХ МЕРЕЖ ПРОЦЕСУ ГАЗОСПОЖИВАННЯ МІСТА З ВРАХУВАННЯМ МЕТЕОФАКТОРІВ.

Попередній аналіз та опрацювання даних газоспоживання і метеофакторів м.Тернопіль дозволяє приступити до побудови моделей нейронних мереж. В дослідженні буде побудовано : одношаровий перцептрон, багатшаровий перцептрон та LSTM модель.

4.1. Одношаровий перцептрон та багатшаровий перцептрон.

На вхід одношарового перцептрона подаються скореговані параметри на основі кореляційного аналізу.

```
3 df.info()
↳ <class 'pandas.core.frame.DataFrame'>
Int64Index: 122712 entries, 0 to 122711
Data columns (total 9 columns):
#   Column  Non-Null Count  Dtype
---  -
0   Q        122712 non-null float64
1   T        122712 non-null float64
2   P        122712 non-null float64
3   U        122712 non-null float64
4   Ff       122712 non-null float64
5   VV       122712 non-null float64
6   RRR      122712 non-null float64
7   tR       122712 non-null float64
8   sss      122712 non-null float64
dtypes: float64(9)
memory usage: 9.4 MB
```

Рисунок 4.1 –Загальна інформація по параметрах моделі

Щоб добитися кращого тренування моделі виконано нормалізацію даних та розбито вибірку на тренувальну та тестову за допомогою функції `train_test_split()`, що поділила вибірку на 0.999% тренувальні дані та 0.001 тестові.

```
1 train_X, test_X, train_y, test_y = train_test_split(X, y, train_size=0.999, test_size=0.001)
```

Рисунок 4.2 – Виконання ділення вибірки на навчальні та тестові дані

Після завершення підготовки даних створюємо моделі

```
1 model_2 = Sequential()  
2 model_2.add(Dense(8, input_dim=8, activation='linear'))  
3 model_2.add(Dense(1))  
4 model_2.compile(loss='mean_squared_error', optimizer='adam')
```

Рисунок 4.3 – Модель одношарового перцептрона

```
1 # create model  
2 model_8 = Sequential()  
3 model_8.add(Dense(8, input_shape=(8,), kernel_initializer='normal', activation='linear'))  
4 model_8.add(Dense(4, activation='linear'))  
5 model_8.add(Dense(1, activation='linear'))  
6 # Compile model  
7 model_8.compile(loss='mean_absolute_error', optimizer='adam')
```

Рисунок 4.4 – Модель багатошарового перцептрона

В обох моделях на вхід подається 8 параметрів. Варта зазначити, що перший шар складається з 8 нейронів, активація нейрона здійснюється лінійно тобто в даному випадку вхідні значення множаться на 1. Оптимізація моделі відбувається на основі градієнтного спуску. В багатошаровому перцептроні також є проміжний шар з 4 нейронами. Обидві моделі на виході мають нейрон, що повертає спрогнозоване значення газоспоживання. Структуру моделі описано на рисунках 4.3 та 4.4 відповідно.

```
1 model_2.fit(train_X, train_y, epochs=20, batch_size=2, verbose=2)
```

Рисунок 4.5 – Функція навчання моделі

Навчання моделей проводиться в 20 епох, що є достатньою кількістю для доброго тренування моделі.

Після процесу тренування отримуємо порівняльну таблицю

index	REAL	MLP	SLP
0	36930.8	3152797696.0	24182.578125
1	49375.4	4108761088.0	30127.296875
2	6962.39	1845355392.0	17261.96484375
3	18958.4	2686684672.0	22248.732421875
4	11052.3	924331456.0	12689.8310546875

Рисунок 4.6 – Порівняльна таблиця

На основі порівняльної таблиці можна побачити, що обидва алгоритми погано справились з прогнозуванням, що підтверджує результати кореляційного аналізу. Також помічаємо що багатошаровий перцептрон за рахунок внутрішніх шарів показав значно гірші результати, ніж одношаровий, що вказує на недоцільність застосування багатошаровості для розв’язання лінійних задач.

4.2 ЛСТМ

На вхід ЛСТМ моделі подаємо лиш один параметр – газоспоживання. Тому починаємо працювати зі змінною `dataset_gas`, що містить в собі інформацію про газоспоживання Щоб зробити погодинне прогнозування застосовуємо функцію `resample()` з параметром ‘H’, щоб розбити вибірку на годинні відрізки та позбутись зайвих колонок, функція `mean()` для отримання середніх значень для кожної години.

```
[244] 1 NewDataSet = dataset_gas.resample('H').mean()
      2 print(NewDataSet)
```

TS	Q	DAYOFWEEK(TS)	Month	Year	Week
2006-01-01 00:00:00	40518.2	1.0	1.0	2006.0	52.0
2006-01-01 01:00:00	39043.2	1.0	1.0	2006.0	52.0
2006-01-01 02:00:00	38551.7	1.0	1.0	2006.0	52.0
2006-01-01 03:00:00	38284.3	1.0	1.0	2006.0	52.0
2006-01-01 04:00:00	38545.7	1.0	1.0	2006.0	52.0
...
2019-12-31 19:00:00	43434.8	3.0	12.0	2019.0	1.0
2019-12-31 20:00:00	41943.0	3.0	12.0	2019.0	1.0
2019-12-31 21:00:00	39574.2	3.0	12.0	2019.0	1.0
2019-12-31 22:00:00	36841.5	3.0	12.0	2019.0	1.0
2019-12-31 23:00:00	33227.7	3.0	12.0	2019.0	1.0

[122712 rows x 5 columns]

Рисунок 4.7 – Вибірка для LSTM

На наступному етапі розбиваємо вибірку на тестові та тренувальні дані, щоб першим дати моделі для навчання, а з допомогою других перевірити точність прогнозування. Для оцінки моделі залишаємо 100 спостережень. Всі інші залишаємо для навчання.

Після цього проводимо нормалізацію та створюємо об'єкт моделі з допомогою функції Sequential().

```
3 regressor = Sequential()
```

Рисунок 4.8 – Ініціалізація об'єкта моделі

На черговому етапі нами застосовано функцію add() для додавання шарів в нейронну мережу.

У дослідженні в довільному порядку було вибрано архітектуру з одним вхідним шаром з 50 нейронів, трьома прихованими, що теж мають по 50 нейронів кожен та вихідним шаром з 1 нейроном.

Кожен шар повертає послідовність в наступний, а параметр dropout показує спосіб виключень повторних послідовностей, що входять в нейрон.

```

1 #MODEL
2
3 regressor = Sequential()
4
5 # Adding the first LSTM layer and some Dropout regularisation
6 regressor.add(LSTM(units = 50, return_sequences = True, input_shape = (X_Train.shape[1], 1)))
7 regressor.add(Dropout(0.2))
8
9 # Adding a second LSTM layer and some Dropout regularisation
10 regressor.add(LSTM(units = 50, return_sequences = True))
11 regressor.add(Dropout(0.2))
12
13 # Adding a third LSTM layer and some Dropout regularisation
14 regressor.add(LSTM(units = 50, return_sequences = True))
15 regressor.add(Dropout(0.2))
16
17 # Adding a fourth LSTM layer and some Dropout regularisation
18 regressor.add(LSTM(units = 50))
19 regressor.add(Dropout(0.2))
20
21 # Adding the output layer
22 regressor.add(Dense(units = 1))
23

```

Рисунок 4.9 – Архітектура LSTM моделі для прогнозування газоспоживання в місті Тернопіль

Після задання архітектури виконано компіляцію моделі з параметрами оптимізації та функцією втрат, що виражають спосіб навчання моделі.

```

# Compiling the RNN
regressor.compile(optimizer = 'adam', loss = 'mean_squared_error')

```

Рисунок 4 .10 – Компіляція RNN

Тепер приступаємо до процесу тренування, запускаємо процес тренування за допомогою функції fit() вхідними параметрами якої є навчальна вибірка, встановлюємо кількість епох тренування 50.

```

1 regressor.fit(X_Train, Y_Train, epochs = 50, batch_size = 32)

```

Рисунок 4.11– Функція тренування NN

```
1 regressor.fit(X_Train, Y_Train, epochs = 50, batch_size = 32)

Epoch 1/50
3831/3831 [=====] - 335s 87ms/step - loss: 4.1056e-04
Epoch 2/50
3831/3831 [=====] - 336s 88ms/step - loss: 4.1092e-04
Epoch 3/50
2688/3831 [=====>.....] - ETA: 1:40 - loss: 4.1030e-04
```

Рисунок 4.12 – Процес тренування

Коли навчання моделі завершено, спрогнозуємо на тестовій вибірці газоспоживання для міста на 100 годин з кроком 1 година

```
19 # Pass to Model
20 predicted_Q = regressor.predict(X_test)
21
```

Рисунок 4.13 – Виконання прогнозування

Зберігаємо дані в таблицю та візуалізуємо отримані результати:

	Date	TrueM3	PredictedM3
0	2019-12-27 20:00:00	38842.4	38259.347656
1	2019-12-27 21:00:00	38225.3	37046.214844
2	2019-12-27 22:00:00	36877.4	34316.808594
3	2019-12-27 23:00:00	32856.5	30565.976562
4	2019-12-28 00:00:00	30854.9	29490.365234
...
95	2019-12-31 19:00:00	43434.8	43133.640625
96	2019-12-31 20:00:00	41943.0	42687.921875
97	2019-12-31 21:00:00	39574.2	40895.941406
98	2019-12-31 22:00:00	36841.5	37004.902344
99	2019-12-31 23:00:00	33227.7	31284.580078

100 rows x 3 columns

Рисунок 4.14 – Таблиця значень прогнозувань в порівнянні з реальними даними

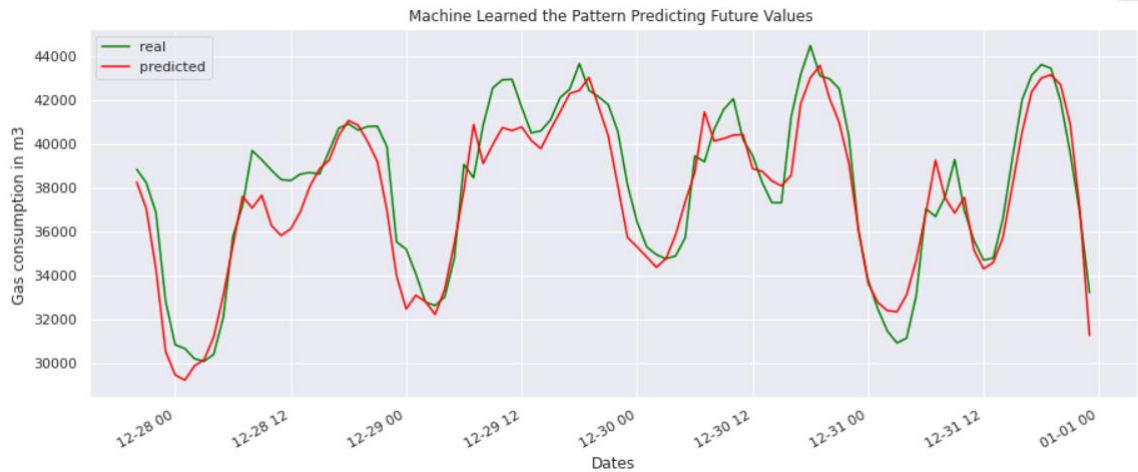


Рисунок 4.15 – Порівняльний графік прогнозування газоспоживання

Аналіз порівняльного графіка прогнозування газоспоживання у м.Тернопіль за допомогою моделі ЛСТМ дає можливість зробити наступні висновки:

- графіки реальних та прогнозованих показників мають мінімальні показники відхилення,
- дана модель оптимально підходить для прогнозування часових рядів газоспоживання,
- на її прогнозах можна побудувати модель закупівлі газу для міста та застосовувати розумне керування його споживанням

4.3 Висновки до четвертого розділу

У четвертому розділі побудовано 3 моделі нейронних мереж - одношаровий перцептрон, багатошаровий перцептрон та ЛСТМ модель для прогнозування газоспоживання у м.Тернопіль. Як вхідні дані при побудові моделей нами використані попередні дані газоспоживання у місті та 8 метеофакторів, відібраних на основі кореляційного аналізу.

Результати тренування моделей нейронних мереж засвідчили про те, що для такого роду завдань нейронна мережа багатошаровий перцептрон не

підходить, оскільки за рахунок внутрішніх шарів дана мережа показала значно гірші результати, ніж одношаровий перцептрон. Для розв'язання лінійних задач одношаровий перцептрон дає вищі результати.

В той же час оптимальною для здійснення прогнозу газоспоживання міста Тернополя виявилася LSTM модель нейронної мережі, про що свідчать графіки реальних та прогнозованих показників газоспоживання.

5 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

5.1. Охорона праці під час експлуатації газового обладнання.

Природному газу, як джерелу енергії, відводиться особлива роль. Протягом останніх років спостерігається тенденція до збільшення частки використання природнього газу у виробництві та побуті.

Попри зручність у використанні, експлуатація газового обладнання вимагає дотримання правил техніки безпеки, а охорона праці є одним із пріоритетних завдань держави та керівництва установи, організації, в якій експлуатується газове обладнання.

Найважливіші аспекти щодо реалізації конституційного права працівників на охорону їх життя і здоров'я у процесі трудової діяльності, на належні, безпечні і здорові умови праці регулює Закон України «Про охорону праці» (зі змінами) [51].

На основі статті 28 даного Закону Міністерством енергетики та вугільної промисловості України 15 травня 2015 року видано Наказ №285 «Про затвердження Правил безпеки систем газопостачання», який є основним нормативно – правовим документом з питань охорони праці при експлуатації систем газопостачання. Ці Правила поширюються на всіх суб'єктів господарювання незалежно від форм власності та організаційно–правової форми, діяльність яких пов'язана з будівництвом, ремонтом, розширенням, реконструкцією, технічним переоснащення, експлуатацією систем газопостачання, а також роздрібною торгівлею зрідженим вуглеводним газом у балонах.

Відповідальність за створення належних і безпечних умов праці при експлуатації газового обладнання покладається на роботодавця, який повинен одержати дозвіл на виконання робіт підвищеної небезпеки і на експлуатацію

(застосування) машин, механізмів, устаткування підвищеної небезпеки та забезпечити утримання систем газопостачання відповідно до вимог чинного законодавства України. Крім того роботодавець – власник (балансоутримувач та/або орендар (наймач)) відповідає за технічний стан і безпечне користування газопроводами, газовими приладами, димовими та вентиляційними каналами.

При облаштуванні робочого місця, створенні безпечних і нешкідливих умов праці роботодавець зобов'язаний керуватися «Загальними вимогами стосовно забезпечення роботодавцями охорони праці працівників», затвердженими наказом Міністерства надзвичайних ситуацій України від 25 січня 2012 року № 67 [54].

Даним нормативним документом передбачено вимоги щодо облаштування робочих зон, невиробничих приміщень, вибір безпечних засобів праці, вимоги безпечного проведення робіт із застосуванням засобів праці.

На роботодавця з метою забезпечення охорони праці поширюються наступні обов'язки:

- безпечне облаштування робочих місць, зокрема – виключення небезпек, якщо це є можливим і реальним; обмеження небезпек, яких уникнути неможливо; усунення небезпек у їх першоджерелах, виключення або максимальне обмеження впливу небезпечних і шкідливих виробничих чинників; забезпечення пріоритету колективних засобів захисту над індивідуальними; врахування людського фактора, зокрема під час вибору засобів виробництва, технології, організації праці, устаткування робочих місць тощо,
- організація проведення попередніх та періодичних медичних оглядів працівників певних категорій,

- забезпечення працівників спеціальним одягом, спеціальним взуттям та іншими засобами індивідуального захисту, відповідно до вимог «Положення про порядок забезпечення працівників спеціальним одягом, спеціальним взуттям та іншими засобами індивідуального захисту» [56],

- допуск до роботи працівників, які відповідно до «Типового положення про порядок проведення навчання і перевірки знань з питань охорони праці» пройшли навчання та мають відповідні документи,

- інформування та інструктування працівників щодо алгоритму дій у випадку аварійних ситуацій,

- вжиття термінових заходів для евакуації працівників, надання першої медичної допомоги потерпілим, локалізації і ліквідації аварії або пожежі, усунення їх наслідків при виникненні на підприємстві надзвичайних ситуацій і нещасних випадків,

- здійснення обліку нещасних випадків, професійних захворювань та аварій на виробництві,

- розроблення планів локалізації і ліквідації можливих аварій в системі газопостачання,

- не допущення експлуатації газового обладнання при несправності або при відключенні приладів контролю і захисту,

- не допущення використання приміщень, в яких прокладено газопроводи, встановлено газорегулюючі пристрої, установки, що використовують газ, і запірні пристрої під склади, майстерні тощо,

- не навантаження газопроводів та не використання їх як заземлення.

На підприємстві, де експлуатується система газопостачання власними силами, повинна бути створена газова служба, затверджено відповідне положення про газову службу та визначено особу, відповідальну за газове господарство.

Якщо роботодавець планує експлуатувати газове обладнання під тиском, то на нього покладається обов'язок утримання обладнання під тиском у справному стані, забезпечення належного технічного обслуговування, технічного огляду, експертного обстеження.

Відповідним наказом призначається відповідального працівника за справний стан і безпечну експлуатацію обладнання під тиском, який пройшов навчання та перевірку знань з охорони праці у встановленому порядку.

В необхідній кількості призначається обслуговуючий персонал, який також має відповідну кваліфікацію та пройшов навчання з охорони праці.

Права і обов'язки персоналу, задіяного в обслуговуванні газового обладнання під тиском затверджуються у Посадовій інструкції, яка підписується у двох примірниках – один зберігається у роботодавця, а другий дається працівникові.

Посадова інструкція особи, відповідальної за газове господарство, є основним документом, який регламентує її правовий статус, вимоги до кваліфікації і знань, права, взаємовідносини за посадою та відповідальність. Про ознайомлення з інструкцією свідчить підпис працівника з проставленням дати ознайомлення.

Працівник, відповідальний за газове господарство, зобов'язаний:

проходити перевірку знань з питань охорони праці, технічної експлуатації і пожежної безпеки — у встановленому порядку до початку виконання своїх функціональних обов'язків, а також періодично раз на три роки;

вступні інструктажі з охорони праці та пожежної безпеки;

первинні інструктажі — на робочому місці;

повторні інструктажі: з охорони праці — двічі на рік, з пожежної безпеки щорічно.

Роботодавець розробляє форму журналу для запису результатів огляду і перевірки газових систем, а саме ведення ретельного спостереження за дорученим устаткуванням шляхом його огляду, перевірки справності арматури,

запобіжних клапанів, засобів сигналізації і захисту, живильних пристроїв вноситься у журнал.

Перед початком опалювального періоду та під час нього мають здійснюватися технічні огляди, технічне обслуговування та при необхідності ремонт систем газопостачання. Конкретні терміни технічного огляду затверджуються відповідною посадовою особою підприємства з урахуванням строків, зазначених в паспортах або інструкціях заводів–виробників обладнання та місцевих умов експлуатації.

Поточний ремонт обладнання проводиться не рідше, ніж один раз в рік, контрольний огляд – не рідше, ніж два рази в рік, технічне обслуговування – щомісячно.

До початку опалювального сезону, не рідше, ніж один раз на рік, перевіряються газоходи і димові труби, печі, котли, інше обладнання. Результати перевірки оформляються актом. Усі вищезазначені заходи сприяють підготовці газового обладнання до опалювального періоду, створюючи безпечні та нешкідливі умови праці, як для відповідального за експлуатацію газового обладнання, так і для всієї організації.

З метою охорони праці та безпеки усіх працівників організації не допускається експлуатація газового обладнання у випадках:

- наявності витоків газу,
- несправностей димових і вентиляційних каналів, відсутності тяги в них,
- несправної автоматики безпеки.

Безпечні умови праці – один з пріоритетних напрямків державної політики в галузі охорони праці. Стаття 4 Закону України «Про охорону праці» передбачає також запобігання нещасним випадкам і професійним захворюванням. У статті 13 цього Закону серед обов'язків роботодавця визначено необхідність забезпечення усунення причин, що призводять до нещасних випадків.

Таким чином держава розглядає людину, як найвищу цінність, та зобов'язує роботодавців на основі ряду нормативно – правових актів створювати належні умови для безпеки праці працівникам, які обслуговують газове обладнання.

5.2 Підвищення стійкості роботи об'єктів господарської діяльності у воєнний час.

Відповідно до Концепції захисту населення і територій у разі загрози та виникнення надзвичайних ситуацій, схваленої Указом Президента України від 26 березня 1999 року № 284/99 «Забезпечення захисту населення і територій у разі загрози надзвичайних ситуацій, які згідно з класифікацією поділяються за характером на техногенні, природні, воєнні та соціально – політичні, а за рівнем – на загальнодержавні, регіональні, місцеві та об'єктові, є одним з найважливіших завдань держави» [59].

Приділяючи велику увагу зміцненню обороноздатності нашої країни, урядом неодноразово підкреслювалося, що оборонна міць держави складається не тільки з високої готовності й оснащення Збройних сил, а й нерозривно пов'язана з високим рівнем економічного розвитку країни, підготовкою населення й об'єктів народного господарства до захисту від зброї масового ураження. Це пов'язано з цілісним підходом до забезпечення повноцінного функціонування держави в умовах НС, в тому числі у воєнний час.

24 лютого 2022 року Україна зазнала повномасштабного військового вторгнення з боку Російської Федерації. За таких обставин питання підвищення стійкості роботи об'єктів господарської діяльності у воєнний час стало ще більш актуальним.

Під стійкістю роботи промислового об'єкта (об'єкта господарювання будь-якої форми власності) розуміють здатність його в умовах надзвичайних ситуацій мирного і воєнного часу випускати продукцію в запланованому обсязі

й номенклатурі, а при одержанні слабких і середніх руйнувань, порушенні зв'язків по кооперації і постачаннях відновлювати виробництво в мінімальний термін.

В основних напрямках Концепції захисту населення і територій у разі загрози та виникнення надзвичайних ситуацій також належне місце відводиться даному напрямку роботи – «розроблення та забезпечення цільових і науково–технічних програм, спрямованих на запобігання надзвичайним ситуаціям та забезпечення сталого функціонування підприємств, установ, організацій незалежно від форм власності та підпорядкування, а також підвідомчих їм об'єктів виробничого і соціального призначення» [59].

Для забезпечення сталого функціонування промислових та сільськогосподарських підприємств в умовах воєнного часу важливим є захист та належне функціонування чотирьох основних елементів сучасного виробництва:

- виробничий персонал – робітники та службовці, як головна продуктивна сила;
- будинки і споруди з технологічним устаткуванням;
- система постачання енергією, водою, паливом, устаткуванням і ремонтною базою;
- система виробничих і кооперативних зв'язків з іншими об'єктами.

Система захисту персоналу – працівників та їх родин – є першочерговим завданням для забезпечення сталого функціонування підприємств. Вона може здійснюватися шляхом інформування людей про правила поведінки в умовах воєнних дій (поведінка в умовах сигналу «повітряна тривога», при хімічній, бактеріологічній, радіаційній небезпеці, при обвалі будинків, наданні першої медичної допомоги тощо). В умовах військового стану у зв'язку з російсько – українською війною на офіційних сторінках державних органів влади, в засобах масової інформації систематично розробляється та оприлюднюється така інформація.

Важливими способами захисту людей від зброї масового знищення є:

- укриття в захисних спорудах, протирадіаційних сховищах,
- використання засобів індивідуального захисту,
- здійснення евакуації.

Захист засобів виробництва – будинків, споруд, устаткування полягає в підвищенні фізичної опірності будинків, споруджень і конструкцій об'єкта до впливу вражаючих факторів ядерного вибуху, захисту технологічного і верстатного устаткування, засобів зв'язку й інших засобів, що складають матеріальну основу виробничого процесу.

Розрізняють оцінку стійкості об'єкта господарювання

– до впливу ударної хвилі – встановлені сім послідовних значень величини надмірного тиску ДРф: 1,5; 10; 20; 30; 40; 50; 100 кПа (0,015; 0,1; 0,2; 0,3; 0,4; 0,5; 1,0 кгс/см²,

– до впливу світлового випромінювання – при умові прозорості атмосфери прийняті шість значень величин світлового імпульсу л/св : 12,4; 29,3; 54,4; 83,7; 113,0; 259,5 Дж/см² (2,7, 13, 20, 27 і 62 кал/см²,

– до впливу проникаючої радіації і радіоактивного зараження – прийнято 12 значень величини гамма-радіації Д7: 40, 100, 200, 300, 400, 600, 800, 1200, 2000, 3000, 5000, 7000 Р.

У зв'язку з тим, що на практиці можуть бути різні проміжні вказаних величин, прийнято групувати сусідні градації значень в інтервали або діапазони. Послідовність діапазонів утворюють шкалу діапазонів.

Більш готовими до стійкої роботи будуть ті об'єкти, які адекватно оцінять фактори, їх несприятливий вплив на виробництво і розроблять відповідні заходи. Заздалегідь проведені організаційні, агрохімічні, агротехнічні, інженерно-технічні, ветеринарно-санітарні, лісотехнічні, лісогосподарські, меліоративні та інші заходи максимально знизять результати впливу вражаючих факторів воєнного часу на людей, сільськогосподарських тварин і створять сприятливі умови для швидкої ліквідації наслідків надзвичайної ситуації.

У випадку неможливості функціонування виробництва на певній території приймається рішення про оперативне його перенесення в більш безпечне місце. У пріоритеті – підприємства харчової промисловості та військово–промислового комплексу. За інформацією Тернопільської обласної військової адміністрації з початку воєнних дій станом на 18 квітня 2022 року Тернопільщина отримала 141 заявку на релокацію підприємств. 26 підприємств вже розпочали роботу – 4 швейних; 5 машинобудівних; 3 металообробних; 2 харчових; 2 підприємства з виробництва іншої неметалевої мінеральної продукції; 1 – поліграфія, ІТ компанія та 8 з інших галузей. Також 3 підприємства завершили переїзд і встановлюють обладнання. Ще 11 в процесі переїзду [62]. Аналогічна ситуація по всій Україні – в разі неможливості провадити виробництво у зоні бойових дій підприємства переносять на більш безпечні території.

Система постачання енергією, водою, паливом, устаткуванням і ремонтною базою є невід’ємною складовою стійкості об’єкта господарської діяльності. Для захисту основних систем життєзабезпечення господарського об’єкта важливо провести наступні заходи:

- заглибити комунально–енергетичних мережі, в такий спосіб вони будуть менш уразливі;
- подбати про створення альтернативних джерел постачання електрикою, газом, водою;
- подачу енергоносіїв здійснювати із різних напрямків (не менш двох);
- використовувати пристрій роздільних систем.

Аналіз напрямів військового ураження з боку Російської Федерації у війні проти України свідчить про масові атаки на заклади критичної інфраструктури – електростанції, нафтобази, продовольчі та військові склади тощо. Мета – послабити стійкість промислових та сільськогосподарських підприємств через зменшення запасів палива, електроенергії, запасних частин, продовольства.

Для забезпечення можливості подальшої експлуатації, відновлювальних робіт чи демонтажу об’єктів господарської діяльності Уряд затвердив механізм

проведення обстеження пошкоджених в результаті збройної агресії Росії будівель і споруд, зокрема критичної інфраструктури. 5 квітня 2022 року прийнято відповідну Постанову «Про внесення змін до Порядку проведення обстеження прийнятих в експлуатацію об'єктів будівництва», якою удосконалено алгоритм відновлення роботи підприємств в умовах військової агресії [57].

Враховуючи вищезазначене приходимо до висновку – для того, щоб забезпечити стійкість об'єктів господарської діяльності у воєнний час необхідно:

- максимально сприяти захисту персоналу та їх родин від можливих загроз,
- здійснити комплекс заходів щодо підвищення стану опірності інженерно-технічного комплексу, а у випадку необхідності
 - оперативного перенесення виробництва, налагодити дієву систему сировинного та енергетичного забезпечення, резервні джерела постачання, зберігати та розширювати систему виробничих і корпоративних зв'язків,
 - проводити навчальні заняття, здійснювати психологічну підготовку для командно - керівного складу цивільної оборони об'єкту з метою оперативних, чітких, злагоджених дій в умовах військового стану.

5.3 Висновки до п'ятого розділу

В цьому розділі розглянуто важливі питання охорони праці та безпеки в надзвичайних ситуаціях, зокрема проаналізовано шляхи підвищення стійкості роботи об'єктів господарської діяльності у воєнний час, що є надзвичайно актуальним в умовах російсько – української війни, масованих атак на об'єкти критичної інфраструктури, та висвітлено тематику охорони праці під час експлуатації газового обладнання, детально проаналізовано обов'язки роботодавця та чинні нормативно – правові акти щодо створення належних і безпечних умов праці для працівників, які обслуговують газове обладнання.

ВИСНОВКИ

У кваліфікаційній роботі апробовано 3 моделі нейронних мереж - одношаровий перцептрон, багатошаровий перцептрон та LSTM модель для прогнозування газоспоживання у м.Тернопіль з використанням метеофакторів.

Результатом роботи стало напрацювання алгоритмів застосування та принципів відбору нейронних мереж і вхідних даних для прогнозування газоспоживання.

На основі проведених досліджень відмічено не доцільність використання нейронної мережі багатошаровий перцептрон для розв'язання поставленої задачі, зокрема і прогнозування газоспоживання міста.

Кращий результати прогнозування досягнуто з використанням нейронних мереж одношаровий перцептрон, де вхідними даними були попередні дані газоспоживання та 8 метеофакторів, відібраних на основі кореляційного аналізу, але ця модель через погану кореляцію між даними теж не варто використовувати

Модель LSTM, для роботи якої використано тільки дані про газоспоживання чудово справилась з поставленим завданням і показала високу точність прогнозування на інтервалі 1 год.

Як підтвердили результати дослідження модель LSTM оптимально підходить для прогнозування газоспоживання, не потребує багатьох статистичних даних, та показує точні результати.

Проведена робота дозволить застосовувати успішно апробовані методики для прогнозування газоспоживання у м.Тернопіль і в інших населених пунктах України та ефективно планувати споживання газу, що сприятиме посиленню як економічної безпеки держави і міста, так і раціональному використанню фінансових ресурсів.

ПЕРЕЛІК ДЖЕРЕЛ

1. Akpinar M. and Yumusak N., Naive forecasting of household natural gas consumption with sliding window approach, *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 25, pp. 30-45, 2017.
2. Akpinar M. and Yumusak N., Year Ahead Demand Forecast of City Natural Gas Using Seasonal Time Series Methods," *Energies*, vol. 9, no. 9, p. 727, 2016.
3. Akpinar M. and Yumusak N/, Forecasting household natural gas consumption with ARIMA model: A case study of removing cycle, in *Proceedings of 7th International Conference on Application of Information and Communication Technologies (AICT)*, Baku, 2013.
4. Akpinar M. and Yumusak N., Day-ahead natural gas forecasting using nonseasonal exponential smoothing methods, in *Conference Proceedings - 2017 17th IEEE International Conference on Environment and Electrical Engineering and 2017 1st IEEE Industrial and Commercial Power Systems Europe*, Milan, 2017.
5. Akpinar M. and Yumusak N., Estimating household natural gas consumption with multiple regression: Effect of cycle, in *Proceedings of the 2013 International Conference on Electronics, Computer and Computation (ICECCO)*, Ankara, 2013.
6. Akpinar M., Adak F. and Yumusak N., Day-Ahead Natural Gas Demand Forecasting Using Optimized ABC-Based Neural Network with Sliding Window Technique: The Case Study of Regional Basis in Turkey, *Energies*, vol. 10, no. 6, p. 781, 2017.
7. Aras N., Forecasting Residential Consumption of Natural Gas Using Genetic Algorithms, *Energy Exploration & Exploitation*, vol. 26, no. 4, pp. 241-266, 2008
8. Azadeh A., Asadzadeh S. and Ghanbari A., An adaptive network-based fuzzy inference system for short-term natural gas demand estimation: Uncertain and complex environments, *Energy Policy*, vol. 38, no. 3, pp. 1529-1536, 2010.
9. Azadeh A., Zarrin M., Rahdar Beik H. and Bioki A., A neuro-fuzzy algorithm for improved gas consumption forecasting with economic, environmental and IT/IS indicators, *Journal of Petroleum Science and Engineering*, vol. 133, pp. 716-739, 2015.

10. Azadeh A., Asadzadeh S. and Ghanbari A., An adaptive network-based fuzzy inference system for short-term natural gas demand estimation: Uncertain and complex environments, *Energy Policy*, vol. 38, no. 3, pp. 1529-1536, 2010.
11. Azadeh A., Asadzadeh S.M., Saberi M., Nadimi V., A. Tajvidi and M. Sheikalishahi, A Neuro-fuzzy-stochastic frontier analysis approach for long-term natural gas consumption forecasting and behavior analysis: The cases of Bahrain, Saudi Arabia, Syria, and UAE, *Applied Energy*, vol. 88, no. 11, pp. 3850-3859, 2011.
12. Bai Y. and Li C., Daily natural gas consumption forecasting based on a structure-calibrated support vector regression approach, *Energy and Buildings*, vol. 127, no. 1, pp. 571-579, 2016.
13. Bhardwaj A. What is a Perceptron? Basics of Neural Networks. [Электронный ресурс] – Режим доступа: <https://towardsdatascience.com/what-is-a-perceptron-basics-of-neural-networksc4cfea20c590#:~:text=How%20do%20they%20work%3F,together%20along%20with%20the%20bias-> (дата звернения 11.04. 2022).
14. Boran F., Forecasting Natural Gas Consumption in Turkey Using Grey Prediction, *Energy Sources*, vol. 10, no. 2, pp. 208-213, 2015.
15. Debnath, K.B., Mourshed, M., 2018. Forecasting methods in energy planning models. *Renew. Sustain. Energy Rev.* 88, 297–325. <https://doi.org/10.1016/j.rser.2018.02.002>.
16. Demirel O., Zaim S., Caliskan A. and Ozuyar P., Forecasting natural gas consumption in Istanbul using neural networks and multivariate time series methods, *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 20, no. 5, pp. 695-711, 2012.
17. Elragal H., Improving neural networks prediction using fuzzy-genetic model, in *Proceedings of the Twenty First National Radio Science Conference (NRSC 2004)*, Cairo, 2004.

18. Feng X, J. Zhang J.S., Zou S.H. and Bao W.Y.B., Study on natural gas demand prediction model in China, in Proceedings of the 3rd International Conference on Energy, Environment and Sustainable Development, Shanghai, 2014.
19. Forouzanfar M., Doustmohammadi A., Menhaj M. and Hasanzadeh S., Modeling and estimation of the natural gas consumption for residential and commercial sectors in Iran, *Applied Energy*, vol. 87, no. 1, pp. 268- 274, 2010.
20. Gil S. and Deferrari J., Generalized Model of Prediction of Natural Gas Consumption, *Journal of Energy Resources Technology*, vol. 126, no. 2, pp. 90-98, 2004.
21. Hribar, R., Potočnik, P., Šilc, J., Papa, G., 2019. A comparison of models for forecasting the residential natural gas demand of an urban area. *Energy* 167, 511–522. <https://doi.org/10.1016/j.energy.2018.10.175>.
22. Izadyar N., Ong H., Shamshirband S., Ghadamian H. and Tong C., Intelligent forecasting of residential heating demand for the District Heating System based on the monthly overall natural gas consumption, *Energy and Buildings*, vol. 104, no. 1, pp. 208-214, 2015.
23. Jang J., ANFIS: adaptive-network-based fuzzy inference system, *IEEE transactions on systems, man, and cybernetics*, vol. 23, no. 3, pp. 665-685, 1993.
24. Kaynar O., Yilmaz I. and Demirkoparan F., Forecasting of natural gas consumption with neural network and neuro fuzzy system, *Energy Education Science and Technology Part A: Energy Science and Research*, vol. 26, no. 2, pp. 221-238, 2011.
25. Kaynar O., Yilmaz I. and Demirkoparan F., Forecasting of natural gas consumption with neural network and neuro fuzzy system, *Energy Education Science and Technology Part A: Energy Science and Research*, vol. 26, no. 2, pp. 221-238, 2011.
26. Kumar U. and Jain V.K., Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India, *Energy*, vol. 35, no. 4, pp. 1709-1716, 2010.

27. Ma X. and Liu Z., Application of a novel time-delayed polynomial grey model to predict the natural gas consumption in China, *Journal of Computational and Applied Mathematics*, vol. 324, pp. 17-24, 2017.

28. Ma Y. and Li Y., Analysis of the supply-demand status of China's natural gas to 2020, *Petroleum Science*, vol. 7, no. 1, pp. 132-135, 2010.

39. Manchanda P. Libraries in Python [Электронный ресурс] – Режим доступа : <https://www.geeksforgeeks.org/libraries-in-python/#:~:text=Python%20standard%20library,in%20the%20C%20programming%20language> – (дата звернення 16.04. 2022).

30. Nazarevych O.B. Information technology for monitoring of city gas consumption. Thesis for the degree of candidate of technical sciences: 05.13.06 – Information Technology / O.B. Nazarevych — Ternopil, 2015. — 160 p.

31. Olah C. Understanding LSTM Networks. [Электронный ресурс] – Режим доступа: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/> - (дата звернення 21.03. 2022).

32. Olgun V., Ozdemir G. and Aydemir E., Forecasting of Turkey's natural gas demand using artificial neural networks and support vector machines, *Energy, Education, Science and Technology*, vol. 30, no. 1, pp. 15-20, 2012.

33. Panapakidis I. and Dagoumas A., Day-ahead natural gas demand forecasting based on the combination of wavelet transform and ANFIS/genetic algorithm/neural network model, *Energy*, vol. 118, no. 1, pp. 231-245, 2017.

34. Potočnik P, Thaler M., Govekar E., Grabec I. and Poredoš A., Forecasting risks of natural gas consumption in Slovenia, *Energy Policy*, vol. 35, no. 8, pp. 4271-4282, 2007.

35. Sabo K., Scitovski R., Vazler I. and Zekić Sušac M., Mathematical models of natural gas consumption, *Energy Conversion and Management*, vol. 52, no. 3, pp. 1721-1727, 2011.

36. Soldo B., Forecasting natural gas consumption, *Applied Energy*, vol. 92, pp. 26-37, 2012.

37. Soldo B., Potočnik P., Šimunović G/, Šarić T. and Govekar E., Improving the residential natural gas consumption forecasting models by using solar radiation, *Energy and Buildings*, vol. 69, pp. 498-506, 2014.
38. Szoplik J., Forecasting of natural gas consumption with artificial neural networks, *Energy*, vol. 85, no. 1, pp. 208-220, 2015.
39. Taspinar F., Celebi N. and Tutkun N., Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods, *Energy and Buildings*, vol. 56, pp. 23-31, 2013.
40. Tonković Z., Zekić Sušac M. and Somolanji M., Predicting natural gas consumption by neural networks, *Technical Gazette*, vol. 16, no. 3, pp. 51-61, 2009.
41. Viet N.H. and Mandziuk J, Neural and fuzzy neural networks for natural gas consumption prediction, *Neural, Parallel and Scientific Computations*, vol. 13, no. 3-4, pp. 265-286, 2005.
42. Wu L., Liu S., Chen H. and Zhang N/, Using a Novel Grey System Model to Forecast Natural Gas Consumption in China, *Mathematical Problems in Engineering*, vol. 2015, pp. 1-7, 2015.
43. Xie Y. and Li M., Research on Prediction Model of Natural Gas Consumption Based on Grey Modeling Optimized by Genetic Algorithm, in *Proceedings - 2009 IITA International Conference on Control, Automation and Systems Engineering, CASE 2009, Zhangjiajie, 2009*.
44. Zekić Sušac M., Frajman Jakšić A. and Drvenkar N., Neuronske mreže i stabla odlučivanja za predviđanje uspješnosti studiranja, *Ekonomski vjesnik*, vol. 22, no. 2, pp. 314-327, 2009.
45. Zeng B. and Li C., Forecasting the natural gas demand in China using a self-adapting intelligent grey model, *Energy*, vol. 112, no. 1, pp. 810-825, 2016.
46. Zhang W. and Yang J., Forecasting natural gas consumption in China by Bayesian Model Averaging, *Energy Reports*, vol. 1, no. November, pp. 216-220, 2015.

47. Zhu L., Li M., Wu Q. and Jiang L., Short-term natural gas demand prediction based on support vector regression with false neighbours filtered, *Energy*, vol. 80, pp. 428-436, 2015.

48. Грабовський М. Довготерміновий прогноз попиту на нафту й газ в Україні / М. Грабовський // *Нафта & газ. Енергетичний бюлетень*. – 2005. – №7. – С.12–15.

49. Гречка О.М. Перспективний аналіз продажу природного газу споживачам України. / О.М. Гречка // *Нафтова і газова промисловість*. – 2005. – №5. – С.37–39.

50. Загородна, Н. В., Литвиненко, Я. В., Фриз, М. Є. (2010) Розробка методу короткотермінового прогнозу добового газоспоживання в опалювальний період на основі регресійного аналізу. *Вісник Тернопільського національного технічного університету*. Тернопіль, №4 (15). С. 130–140.

51. Закон України «Про охорону праці». [Електронний ресурс] – Режим доступу: <https://zakon.rada.gov.ua/laws/show/2694-12> - (дата звернення 06.02.2022).

52. Мацюк О.В. Оцінка кореляційного взаємозв'язку періодичних випадкових процесів та її використання в задачах енергетики /О.В. Мацюк, М.В. Приймак // *Розвідка та розробка нафтових і газових родовищ*. – 2005. – №2(15).– С.39–42.

53. Назаревич О.Б. Інформаційна технологія моніторингу газоспоживання міста. Дисертація на здобуття наукового ступеня кандидата технічних наук: 05.13.06 – інформаційні технології / О.Б. Назаревич — Тернопіль, 2015. — 160 с.

54. Наказ Міністерства Надзвичайних ситуацій України «Про загальні вимоги стосовно забезпечення роботодавцями охорони праці працівників». [Електронний ресурс] – Режим доступу:<https://zakon.rada.gov.ua/laws/show/z0226-12#Text-> (дата звернення 04.02. 2022).

55. Норми та вказівки по нормуванню витрат палива та теплової енергії на опалення житлових та громадських споруд, а також на господарсько-побутові потреби в Україні; КТМ 204 України 244-94. – К., 1998. – 376с.

56. «Положення про порядок забезпечення працівників спеціальним одягом, спеціальним взуттям та іншими засобами індивідуального захисту». [Електронний ресурс] – Режим доступу: <https://zakon.rada.gov.ua/laws/show/z0674-15#Text> - (дата звернення 03.01.2022).

57. Постанова Кабінету Міністрів України «Про внесення змін до Порядку проведення обстеження прийнятих в експлуатацію об'єктів будівництва». [Електронний ресурс] – Режим доступу: <https://zakon.rada.gov.ua/laws/show/423-2022-%D0%BF#Text> - (дата звернення 12.05.2022).

58. Скуржанський Т.А. Одношаровий перцептрон як інструмент для аналізу газоспоживання // Інформаційні моделі, системи та технології: Праці ІХ наук.-техн. конф. (08-09 грудня 2021 р.), Тернопіль, 2021. – С. 88.

59. Указ Президента України «Про Концепцію захисту населення і територій у разі загрози та виникнення надзвичайних ситуацій». [Електронний ресурс] – Режим доступу: <https://zakon.rada.gov.ua/laws/show/284/99#Text> - (дата звернення 18.12.2021).

60. Червоний Монах. Рейтинг мов програмування RedMonk: червень 2021 року. [Електронний ресурс] – Режим доступу: <https://redmonk.com/sogrady/2021/08/05/language-rankings-6-21> - (дата звернення 21.03. 2022).

61. Шишко Г.Г. Учет расхода газа / Г.Г. Шишко, М.П. Енин. – К.: Урожай, 1993. – 310 с.

62. Як бізнес з інших міст рятується від війни на Тернопільщині.[Електронний ресурс] – Режим доступу: <https://oda.te.gov.ua> - (дата звернення 19.04.2022).

ДОДАТКИ

ДОДАТОК А

Тези конференції

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ

МАТЕРІАЛИ

ІХ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

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



8–9 грудня 2021 року

ТЕРНОПІЛЬ
2021

Н.С. Таванець, В.В. Никитюк СПОСІБ ГОЛОСОВОЇ ІДЕНТИФІКАЦІЇ КОРИСТУВАЧА N. Tavanets, V. Nykytyuk THE METHOD OF VOICE IDENTIFICATION OF THE USER	84
М. Тимків, О. Яскілка АНАЛІЗ ЗАСТОСУНКІВ ДЛЯ ВИВЧЕННЯ ІНОЗЕМНИХ МОВ M. Tymkiv, O. Yaskilka ANALYSIS OF APPLICATIONS FOR THE STUDY OF FOREIGN LANGUAGES	85
Д.Ю. Дзюба, Л.П. Дмитроца ПРИСТРОЇ З ФІЗИЧНОЮ НЕКЛОНОВАНОЮ ФУНКЦІЄЮ (PUF) D. Dziuba, L. Dmytrotsa DEVICES WITH PHYSICAL NON-CLONED FUNCTION (PUF)	86
Ж.Ж. Захем, В.Б. Савків РОЗРОБКА ТА ДОСЛІДЖЕННЯ СИСТЕМИ ВІРТУАЛЬНОГО ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ПРАЦІВНИКІВ ВИРОБНИЦТВ НА БАЗІ VR.AR ТА AI J. Zakhem, V. Savkiv DEVELOPMENT AND RESEARCH OF A VIRTUAL SECURITY SYSTEM FOR PRODUCTION WORKERS BASED WITH THE HELP OF VR.AR AND AI	87
Т. Скуржанський, О.Б. Назаревич ОДНОШАРОВИЙ ПЕРЦЕПТРОН ЯК ІНСТРУМЕНТ ДЛЯ АНАЛІЗУ ГАЗОСПОЖИВАННЯ T. Skurzhanskyi, O. Nazarevych SINGLE LAYER PERCEPTRON AS A TOOL FOR GAS CONSUMPTION ANALYSIS	88
Д. Корж, Д. Радчук, М. Тимків, А. Колесник, Т. Зошук РІЗНИЦЯ МІЖ «ТРАДИЦІЙНИМИ» ТА «РОЗУМНИМИ» МІСТАМИ D. Korzh, D. Radchuk, M. Tymkiv, A. Kolesnyk, T. Zoshchuk THE DIFFERENCE BETWEEN "TRADITIONAL" AND "SMART" CITIES	90
Д. Корж, Д. Радчук, О. Лішук, А. Колесник, Т. Зошук РОЗУМНА СИСТЕМА ЕЛЕКТРОННОГО ЗДОРОВ'Я ДЛЯ ВІДСТЕЖЕННЯ ТА МОНІТОРИНГУ ПАЦІЄНТІВ, ПЕРСОНАЛУ В РЕЖИМІ РЕАЛЬНОГО ЧАСУ D. Korzh, D. Radchuk, O. Lishchuk, A. Kolesnyk, Zoshchuk T. SMART ELECTRONIC HEALTH SYSTEM FOR TRACKING AND MONITORING OF PATIENTS, PERSONNEL IN REAL TIME	92
М. Тимків, О. Яскілка АНАЛІЗ ЗАСТОСУНКІВ ДЛЯ ВИВЧЕННЯ ІНОЗЕМНИХ МОВ M. Tymkiv, O. Yaskilka ANALYSIS OF APPLICATIONS FOR THE STUDY OF FOREIGN LANGUAGES	91
О.І. Тимчак, І.Ю. Дедів АЛГОРИТМ ВИДІЛЕННЯ ТА РОЗПІЗНАВАННЯ ОБЛИЧЧЯ O.I. Tymchak, Deditiv. ALGORITHM OF SELECTION AND FACE RECOGNITION	93

УДК 004.891

Т. Скуржанський – ст. гр. СНм-61, О.Б. Назаревич, к.т.н., доцент кафедри комп'ютерних наук
(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

ОДНОШАРОВИЙ ПЕРЦЕПТРОН ЯК ІНСТРУМЕНТ ДЛЯ АНАЛІЗУ ГАЗОСПОЖИВАННЯ

UDC 004.891

T. Skurzhanyskiy, O. Nazarevych, Ph.D., Associate Professor of Computer Science

SINGLE LAYER PERCEPTRON AS A TOOL FOR GAS CONSUMPTION ANALYSIS

Ключові слова: Нейронна мережа, одношаровий перцептрон.
Key words: Neural network, single-layer perceptron.

З розвитком та вдосконаленням енергетичного сектору методи прогнозування стали його невід'ємною частиною. Прогнозування з допомогою нейронних мереж почало використовуватися для оптимізації, покращення планування, управління та ефективності багатьох систем. Однією з таких систем є газоспоживання. Оптимізація газопостачання є складовою енергетичної системи України, а газоспоживання в рамках новітніх світових тенденцій стало однією з актуальних тем сьогодення.

Сучасні автори, зокрема Hřibar R., Potočník P., Šilc J., Pára G., Назаревич О., Загородна Н., Литвиненко Я., Фриз М. застосовували різноманітні методи з різними вхідними параметрами для більш точного прогнозування споживання природного газу [1-4]. Було зроблено прогнози для наступних сфер застосування – установка, місто, область та часових періодів – годинний, щоденний, річний.

Ключовим завданням даного дослідження є застосування одношарового перцептрона для прогнозування споживання природного газу на основі даних, розроблених або застосованих за останні кілька років з можливістю використання метеофакторів, як головного чинника зміни газоспоживання.

Для проведення дослідження використовуються наступні параметри: кількість спожитого газу (одиниця виміру – м³), температура повітря (градуси Цельсія) на висоті 2 метри, атмосферний тиск, відносна вологість повітря, загальна хмарність, кількість опадів, висота снігового покриву та інші.

На основі результатів тренування моделей отримуємо матрицю з даними. Вона представлена на рисунку 1.

```
Epoch 1/50  
20320/20320 - 7s - loss: 0.0163  
Epoch 2/50  
20320/20320 - 8s - loss: 0.0160  
Epoch 3/50  
20320/20320 - 7s - loss: 0.0159  
Epoch 4/50  
20320/20320 - 8s - loss: 0.0158  
Epoch 5/50  
20320/20320 - 8s - loss: 0.0158  
Epoch 6/50  
20320/20320 - 8s - loss: 0.0157
```

Рисунок 1. Вихідні дані результатів тренування моделі по епохах

Як результат використання нейронної мережі Одношаровий перцептрон отримуємо регресію, яка відображає співвідношення прогнозованих показників до реальних.

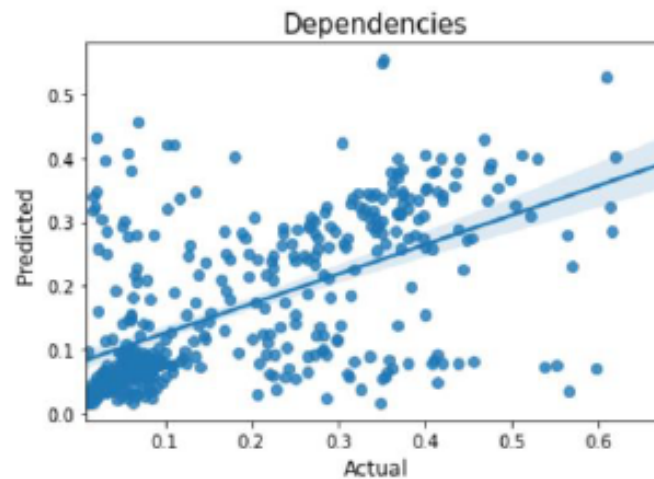


Рисунок 2. Результати регресійного аналізу

Таким чином результати даного дослідження свідчать про те, що за допомогою нейронних мереж, зокрема одношарового перцептрона на основі параметрів метеодосліджень можна прогнозувати потребу в газоспоживанні, що дозволить оптимізувати ринок газопостачання.

Література.

1. Назаревич О.Б. Інформаційна технологія моніторингу газоспоживання міста. Дисертація на здобуття наукового ступеня кандидата технічних наук: 05.13.06 – інформаційні технології / О.Б. Назаревич – Тернопіль, 2015. – 160 с.
2. Nazarevych O.B. Information technology for monitoring of city gas consumption. Thesis for the degree of candidate of technical sciences: 05.13.06 – Information Technology / O.B. Nazarevych – Ternopil, 2015. – 160 p.
3. Hribar, R., Potočník, P., Šilc, J., Papa, G., 2019. A comparison of models for forecasting the residential natural gas demand of an urban area. *Energy* 167, 511–522. <https://doi.org/10.1016/j.energy.2018.10.175>.
5. Debnath, K.B., Mourshed, M., 2018. Forecasting methods in energy planning models. *Renew. Sustain. Energy Rev.* 88, 297–325. <https://doi.org/10.1016/j.rser.2018.02.002>.

Додаток Б

```
# -*- coding: utf-8 -*-
```

```
"""Mag.ipynb
```

Automatically generated by Colaboratory.

Original file is located at

<https://colab.research.google.com/drive/1nXSCbBNmbmFkHMIXBPT57zFnYXiKarI->

```
"""
```

```
from google.colab import drive
```

```
drive.mount("/content/drive", force_remount=True)
```

```
import pandas as pd
```

```
from pathlib import Path
```

```
import tensorflow as tf
```

```
from tensorflow import keras
```

```
from tensorflow.keras.models import Sequential
```

```
from tensorflow.keras.layers import Activation, Dense, Dropout, Embedding, LSTM, GlobalMaxPooling1D, SpatialDropout1D
```

```
from tensorflow.keras.optimizers import Adam
```

```
from tensorflow.keras.metrics import categorical_crossentropy
```

```
import numpy as np
```

```
import seaborn as sns
```

```
import matplotlib.pyplot as plt
```

```
from sklearn.preprocessing import MinMaxScaler
```

```
from sklearn.metrics import mean_squared_error
```

```

from math import sqrt

# Seed RNG to ensure experiments are consistent
np.random.seed(42)

gas_file = Path("/content/drive/MyDrive/2006-
2019(+day_of_week)_t_gaz_sum_hours_full.csv")
wether_file = Path("/content/drive/MyDrive/2006-2019(3hour)_meteo_rp5ua.csv")

gas_df = pd.read_csv(gas_file)
gas_df.head()

dataset_gas = gas_df
dataset_gas["Month"] = pd.to_datetime(gas_df["TS"]).dt.month
dataset_gas["Year"] = pd.to_datetime(gas_df["TS"]).dt.year
dataset_gas["Date"] = pd.to_datetime(gas_df["TS"]).dt.date
dataset_gas["Time"] = pd.to_datetime(gas_df["TS"]).dt.time
dataset_gas["Week"] = pd.to_datetime(gas_df["TS"]).dt.week
dataset_gas["Day"] = pd.to_datetime(gas_df["TS"]).dt.day_name()
dataset_gas = gas_df.set_index("TS")
dataset_gas.index = pd.to_datetime(dataset_gas.index)
dataset_gas.head(1)

# How many Unique Year do we Have in Dataset
print(gas_df.Year.unique(), "\n")
print("Total Number of Unique Year", gas_df.Year.nunique(), "\n")

#Lets us see the energy consumption Each Year

```

```

from matplotlib import style
fig = plt.figure()
ax1 = plt.subplot2grid((1,1), (0,0))

style.use('ggplot')

sns.boxplot(x=dataset_gas["Year"], y=dataset_gas["Q"], data=gas_df)
sns.set(rc={'figure.figsize':(15,6)})

plt.title("Energy consumptionnin Year ...")
plt.xlabel("Date")
plt.ylabel("Energy in MW")
plt.grid(True)
plt.legend()

for label in ax1.xaxis.get_ticklabels():
    label.set_rotation(90)

plt.title("Energy Consumption According to Year")

#Lets us see the energy consumption Each Year
from matplotlib import style
fig = plt.figure()
ax1 = plt.subplot2grid((1,1), (0,0))

style.use('ggplot')

```

```

sns.barplot(x=dataset_gas["Month"], y=dataset_gas["Q"], data=gas_df)
sns.set(rc={'figure.figsize':(15,6)})

plt.title("Energy consumptionnin Year ...")
plt.xlabel("Date")
plt.ylabel("Energy in MW")
plt.grid(True)
plt.legend()

for label in ax1.xaxis.get_ticklabels():
    label.set_rotation(90)

plt.title("Energy Consumption According to Year")

sns.distplot(dataset_gas["Q"])
plt.title("Ennergy Distribution")

"""#Prepering data for LSTM model"""

# Resampeling
#print(dataset_gas)

NewDataSet = dataset_gas.resample('H').mean()
print(NewDataSet)

print("Old Dataset ",dataset_gas.shape )
print("New Dataset ",NewDataSet.shape )

```

```

TestData = NewDataSet.tail(100)

Training_Set = NewDataSet.iloc[:,0:1]

Training_Set = Training_Set[:-60]

print("Training Set Shape ", Training_Set.shape)
print("Test Set Shape ", TestData.shape)

Training_Set = Training_Set.values
sc = MinMaxScaler(feature_range=(0, 1))
Train = sc.fit_transform(Training_Set)

X_Train = []
Y_Train = []

# Range should be fromm 60 Values to END
for i in range(60, Train.shape[0]):

    # X_Train 0-59
    X_Train.append(Train[i-60:i])

    # Y Would be 60 th Value based on past 60 Values
    Y_Train.append(Train[i])

# Convert into Numpy Array
X_Train = np.array(X_Train)

```

```

Y_Train = np.array(Y_Train)

print(X_Train.shape)
print(Y_Train.shape)

# Shape should be Number of [Datapoints , Steps , 1 )
# we convert into 3-d Vector or #rd Dimesnsion
X_Train = np.reshape(X_Train, newshape=(X_Train.shape[0], X_Train.shape[1], 1))
X_Train.shape

#MODEL

regressor = Sequential()

# Adding the first LSTM layer and some Dropout regularisation
regressor.add(LSTM(units = 50, return_sequences = True, input_shape =
(X_Train.shape[1], 1)))
regressor.add(Dropout(0.2))

# Adding a second LSTM layer and some Dropout regularisation
regressor.add(LSTM(units = 50, return_sequences = True))
regressor.add(Dropout(0.2))

# Adding a third LSTM layer and some Dropout regularisation
regressor.add(LSTM(units = 50, return_sequences = True))
regressor.add(Dropout(0.2))

# Adding a fourth LSTM layer and some Dropout regularisation

```

```
regressor.add(LSTM(units = 50))
regressor.add(Dropout(0.2))

# Adding the output layer
regressor.add(Dense(units = 1))

# Compiling the RNN
regressor.compile(optimizer = 'adam', loss = 'mean_squared_error')

regressor.fit(X_Train, Y_Train, epochs = 50, batch_size = 32)

TestData.head(2)

TestData.shape

NewDataSet.shape

Df_Total = NewDataSet[["Q"]]

Df_Total.shape

inputs = Df_Total[len(Df_Total) - len(TestData) - 60:].values
inputs.shape

inputs = Df_Total[len(Df_Total) - len(TestData) - 60:].values
```



```

# We need to Reshape
inputs = inputs.reshape(-1,1)

# Normalize the Dataset
inputs = sc.transform(inputs)

X_test = []
for i in range(60, 160):
    X_test.append(inputs[i-60:i])

# Convert into Numpy Array
X_test = np.array(X_test)

# Reshape before Passing to Network
X_test = np.reshape(X_test, (X_test.shape[0], X_test.shape[1], 1))

# Pass to Model
predicted_Q = regressor.predict(X_test)

# Do inverse Transformation to get Values
predicted_Q = sc.inverse_transform(predicted_Q)

True_M3 = TestData["Q"].to_list()
Predicted_M3 = predicted_Q
dates = TestData.index.to_list()

Machine_Df = pd.DataFrame(data={
    "Date":dates,

```

```

    "TrueM3": True_M3,
    "PredictedM3": [x[0] for x in Predicted_M3 ]
})

#Future Predicted
Machine_Df

True_M3 = TestData["Q"].to_list()
Predicted_M3 = [x[0] for x in Predicted_M3]
dates = TestData.index.to_list()

fig = plt.figure()

ax1= fig.add_subplot(111)

x = dates
y = True_M3

y1 = Predicted_M3

plt.plot(x,y, color="green",label = 'real')
plt.plot(x,y1, color="red",label = 'predicted')
# beautify the x-labels
plt.gcf().autofmt_xdate()
plt.xlabel('Dates')
plt.ylabel("Gas consumption in m3 ")
plt.title("Machine Learned the Pattern Predicting Future Values ")
plt.legend()

```

```

# cleanup gas data by converting TS column to timestamp
gas_df = pd.read_csv(gas_file).assign(TS=lambda df: pd.to_datetime(df.TS))

# cleanup the weather data by removing empty columns and converting time to
datetime objects
weather_df = pd.read_csv(wether_file) \
    .assign(Time=lambda df: pd.to_datetime(df.Time)) \
    .drop(columns=['DD', 'N', 'WW', 'W1', 'W2', 'Cl', 'Nh', 'H', 'Cm', 'Ch', 'E', 'E1'])

weather_df.head(10)

weather_df.info()

gas_df.info()

gas_df.rename(columns = {'TS':'Time'}, inplace = True)

gas_df.head()

n = gas_df['Time']

# simple merge of the dataframes
df = pd.merge(gas_df,weather_df, on="Time", how="outer")

df = df.drop('Time', axis = 1)

```

```
df = df.interpolate(method='linear', limit_direction='backward', axis=0)
```

```
df["Time"] = n
```

```
df = pd.merge(n,df, on="Time", how="inner")
```

```
df = df.dropna()
```

```
9# convert time to seconds since epoch and fill missing values with zeroes
```

```
# TODO: check number of missing values beforehand
```

```
# TODO: use better strategy to fill missing values
```

```
data = df.assign(timestamp=lambda df: df.Time.apply(lambda t:  
t.timestamp())).fillna(0).drop(columns=["Time"])
```

```
data.info()
```

```
data.tail()
```

```
data = data.drop(columns = ['timestamp','DAYOFWEEK(TS)','id'])
```

```
# output basic information about gas consumption dataset
```

```
print("="*50)
```

```
print("First Five Rows ", "\n")
```

```

print(gas_df.head(2),"\n")

print("="*50)
print("Information About Dataset", "\n")
print(gas_df.info(), "\n")

print("="*50)
print("Describe the Dataset ", "\n")
print(gas_df.describe(), "\n")

print("="*50)
print("Null Values t ", "\n")
print(gas_df.isnull().sum(), "\n")

data.corr()

# Increase the size of the heatmap.
plt.figure(figsize=(32, 12))

# Store heatmap object in a variable to easily access it when you want to include
more features (such as title).

# Set the range of values to be displayed on the colormap from -1 to 1, and set the
annotation to True to display the correlation values on the heatmap.

heatmap = sns.heatmap(data.corr(), annot=True)

# Give a title to the heatmap. Pad defines the distance of the title from the top of the
heatmap.

heatmap.set_title('Correlation Heatmap', fontdict={'fontsize':12}, pad=12);

data = data.drop(columns=['Po', 'Tx', 'Tn', 'Tg', 'ff3', 'ff10', 'Td', 'Pa'])

```

```

# Increase the size of the heatmap.
plt.figure(figsize=(32, 12))

# Store heatmap object in a variable to easily access it when you want to include
more features (such as title).

# Set the range of values to be displayed on the colormap from -1 to 1, and set the
annotation to True to display the correlation values on the heatmap.

heatmap = sns.heatmap(data.corr(), annot=True)

# Give a title to the heatmap. Pad defines the distance of the title from the top of the
heatmap.

heatmap.set_title('Correlation Heatmap', fontdict={'fontsize':12}, pad=12);

""""#Preparing data for SLP and MLP""""

df =
df.drop(columns=['DAYOFWEEK(TS)', 'Po', 'Tx', 'Tn', 'Tg', 'ff3', 'ff10', 'Td', 'Pa', 'id', 'Time'
])

#print (df)

df.info()

from sklearn.model_selection import train_test_split

X = df.iloc[:,1:]
X = X.values
y = df.iloc[:,0:1]
y = y.values

scaler = MinMaxScaler()

```

```

# transform data
X = scaler.fit_transform(X)
y = scaler.fit_transform(y)

train_X, test_X, train_y, test_y = train_test_split(X, y, train_size=0.999,
test_size=0.001)

type(train_X)
train_X.shape
train_y.shape

# create model
model_8 = Sequential()
model_8.add(Dense(8, input_shape=(8,),kernel_initializer='normal',
activation='linear'))
model_8.add(Dense(4, activation='linear'))
model_8.add(Dense(1, activation='linear'))
# Compile model
model_8.compile(loss='mean_absolute_error', optimizer='adam')

model_8.summary()

model_8.fit(train_X, train_y, epochs=20, batch_size=32,
verbose=2)

"""#Prediction MLP"""

```

```
predicted_y = model_8.predict(test_X)
predicted_y = scaler.inverse_transform(predicted_y)
```

```
import math
```

```
MSE = np.square(np.subtract(test_y, predicted_y)).mean()
```

```
RMSE = math.sqrt(MSE)
```

```
print("Root Mean Square Error:\n")
```

```
print(RMSE)
```

```
mean_squared_error(test_y, predicted_y)
```

```
model_2 = Sequential()
```

```
model_2.add(Dense(8, input_dim=8, activation='linear'))
```

```
model_2.add(Dense(1))
```

```
model_2.compile(loss='mean_squared_error', optimizer='adam')
```

```
model_2.summary()
```

```
model_2.fit(train_X, train_y, epochs=20, batch_size=2, verbose=2)
```

```
""""#Prediction SLP""""
```



```

predicted_y2 = model_2.predict(test_X)
predicted_y2 = scaler.inverse_transform(predicted_y2)

predicted_y = predicted_y.tolist()
predicted_y2 = predicted_y2.tolist()

REAL = scaler.inverse_transform(test_y)

Results_MLP_SLP = pd.DataFrame(data={
    "REAL": [x[0] for x in REAL ],
    "MLP": [x[0] for x in predicted_y ],
    "SLP": [x[0] for x in predicted_y2 ]
})

Results_MLP_SLP.head(5)

MSE = np.square(np.subtract(test_y, predicted_y2)).mean()

RMSE = math.sqrt(MSE)
print("Root Mean Square Error:\n")
print(RMSE)

MSE = np.square(np.subtract(test_y, predicted_y)).mean()

RMSE = math.sqrt(MSE)
print("Root Mean Square Error:\n")
print(RMSE)

```