

**ХЕРСОНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ДИЗАЙНУ
КАФЕДРА ІНФОРМАТИКИ І КОМП'ЮТЕРНИХ НАУК**

Пояснювальна записка
до дипломної бакалаврської роботи

на тему:

**Кластеризація багатовимірних часових рядів
в енергетиці та метеорології**

Виконав: студент 2 курсу, групи 2КНс
спеціальності 122 «Комп'ютерні науки»

М. О. Ламов

Керівник: С. В. Вишемирська

Рецензент: О. Є. Огнева

Факультет	<u>Інформаційних технологій та дизайну</u>
Кафедра	<u>Інформатики і комп'ютерних наук</u>
Рівень вищої освіти	<u>перший (бакалаврський) рівень</u>
Галузь підготовки	<u>12 «Інформаційні технології»</u> (шифр і назва)
Освітньо-професійна програма	<u>Комп'ютерні науки</u> (назва)
Спеціальність	<u>122 «Комп'ютерні науки»</u> (шифр і назва)

ЗАТВЕРДЖУЮ
Завідувач кафедри ІКН,
професор

_____ **В.І. Литвиненко**
 «_____» _____ **2021 року**

З А В Д А Н Н Я
НА ДИПЛОМНУ РОБОТУ СТУДЕНТА

_____ **Ламова Максима Олександровича**

(прізвище, ім'я, по батькові)

1. Тема роботи: Кластеризація багатовимірних часових рядів в енергетиці та метеорології

керівник роботи Вишемирська Світлана Вікторівна, кандидат технічних наук, доцент кафедри інформатики і комп'ютерних наук

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

затверджені наказом ХНТУ від 26.11.2020 р. № 643-с

2. Строк подання студентом роботи 04.06.2021 р.

3. Вихідні дані до роботи _____

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити): Вступ. Розділ 1. Аналітичний огляд літературних та інших джерел. Розділ 2. Теоретична частина. Аналіз предметної області. Розділ 3. Практична частина. Розробка нейронної мережі. Розділ 4. Охорона праці в комунальному підприємстві Херсонської обласної ради «Херсонські Авіалінії»

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)
 Таблиць – 2,
 Формул – 13,
 Рисунків – 36.

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
4	к.с.-г.н., доцент Малєєв В.А.		

7. Дата видачі завдання 02.02.2021 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломної роботи	Строк виконання етапів роботи	Примітка
1	Аналітичний огляд літературних джерел за тематикою бакалаврської роботи	10.02.2021- 01.03.2021	
2	Обґрунтування проблеми та аналіз методів і засобів її вирішення	02.03.2021- 15.03.2021	
3	Системний аналіз об'єкта дослідження та предметної області	16.03.2021- 01.04.2021	
4	Обґрунтування вибору мови програмування та програмних засобів для побудови нейронної мережі	02.04.2021- 10.04.2021	
5	Розробка та валідація нейронної мережі для передбачення погоди	11.04.2021- 21.04.2021	
6	Проведення порівняльного аналізу та оформлення висновків	01.05.2021- 10.05.2021	
7	Аналіз можливих небезпечних і шкідливих факторів, охорона праці	11.05.2021- 21.05.2021	
8	Написання та оформлення пояснювальної записки до дипломної роботи	22.05.2021- 04.06.2021	

Студент _____ **М.О. Ламов** _____

(підпис) (прізвище та ініціали)

Керівник роботи _____ **С.В. Вишемирська** _____

(підпис) (прізвище та ініціали)

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ	7
ВСТУП.....	8
РОЗДІЛ 1. АНАЛІТИЧНИЙ ОГЛЯД ЛІТЕРАТУРНИХ ТА ІНШИХ ДЖЕРЕЛ.....	10
1. 1 Типи даних та методи їх обробки.....	10
1.2 Нейронні мережі	14
1.3 Мова програмування Python	19
1.4 Бібліотеки Python для роботи з даними	20
1.5 Середовище розробки PyCharm	23
Висновки до розділу 1	25
РОЗДІЛ 2 ТЕОРЕТИЧНА ЧАСТИНА. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ.....	26
2.1 Постановка задачі	26
2.2 Системний аналіз та обґрунтування проблеми	26
2. 3 Кластеризація як інструмент підвищення якості навчання нейронних мереж.....	29
2. 4 Методи та засоби розв’язання поставленої задачі.....	34
Висновки до розділу 2	37
РОЗДІЛ 3 ПРАКТИЧНА ЧАСТИНА. РОЗРОБКА НЕЙРОННОЇ МЕРЕЖІ	38
3.1 Комп’ютерна система. Технічні характеристики комп’ютера.....	38
3.2 Реалізація обробки даних та побудова теплової карти кореляції функцій	38
3.3 Створення нейронної мережі та її навчання	41
3.4 Програмування та тестування	50

3.4.1 Програмування	50
3.4.2 Тестування	51
Висновки до розділу	52
РОЗДІЛ 4. ОХОРОНА ПРАЦІ В КОМУНАЛЬНОМУ ПІДПРИЄМСТВІ ХЕРСОНСЬКОЇ ОБЛАСНОЇ РАДИ «ХЕРСОНСЬКІ АВІАЛІНІЇ».....	53
4.1 Загальна характеристика робочого місця системного адміністратора в комунальному підприємстві Херсонської обласної ради «Херсонські Авіалінії»	53
4.2 Організація повітрообміну в приміщеннях на робочому місці системного адміністратора комунального підприємства Херсонської обласної ради «Херсонські авіалінії».....	57
Висновки до розділу 4	60
ВИСНОВКИ	61
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	63
ДОДАТОК А.....	68

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

Скорочення, термін, позначення	Пояснення
ОС	Операційна система
RNN	Recurrent Neural Networks
LTSM	Long short-term memory
ПЗ	Програмне забезпечення
Рис	Рисунок
SQL	Structure Query Language
CSV	Comma-Separated Values
SVM	Support vector machine
API	Application Programming Interface
ANN	Artificial Neural Network

ВСТУП

За останні десятиріччя всесвітня мережа Інтернет стала невід'ємною частиною нашого життя. З ростом Інтернету, зростає також і об'єм даних, які потрібно зберігати та обробляти. З кожним роком об'єм даних, що зберігається в Інтернет, зростає приблизно на 40%. Саме розвиток сучасних інформаційних технологій дозволяє і сприяє тому, що обсяги інформації, що зберігаються та обробляються, постійно зростають. Поряд із швидко зростаючими обсягами найрізноманітніших видів даних для їх обробки необхідно все більше ресурсів та більш складні програмні рішення [24].

У зв'язку з таким швидким збільшенням обсягу інформації та технологій велика увага приділяється автоматизованим способам аналізу та обробки інформації. Велику роль в автоматизованому аналізі та обробці даних відіграє штучний інтелект. Один з напрямків штучного інтелекту – це нейронні мережі [14]. Мета розвитку нейронних мереж полягає в максимальному наближенні до роботи людської нервової системи, здатність навчатися та діяти на підставі попереднього досвіду, роблячи з кожним разом все менше помилок.

Штучні нейронні мережі застосовуються в різних областях науки: від навчання розпізнавання рукописного тексту до класифікації різних видів раку. Для навчання нейронної мережі необхідна велика кількість інформації, оскільки неможливо досягти високої точності роботи штучних нейронних мереж на досить малій кількості даних. Так, наприклад, для аналізу зображень, Google в якості навчального набору даних використовував інформацію з сервісу YouTube. Що стосується виконання завдання розпізнавання мови, як навчальної вибірки, використовувалася серія звукозаписів з доданими до них описами. Перша представлена версія розпізнавання мови на основі нейронної мережі містила 25% помилок; через три роки результат був поліпшений і складав уже 8% помилок. На основі нейронних мереж проводяться дослідження задач прогнозування динаміки цін на фондовому ринку. Точність даних досліджень досягає 60-75% вірно передбачених значень ціни, а кількість вірно передбачених тенденцій ринку –

90%. Ще одним застосуванням штучних нейронних мереж є завдання в галузі геологорозвідки. Наприклад, компанія Amoco використовує штучні нейронні мережі для виділення характерних піків в показаннях сейсмічних датчиків. Надійність розпізнавання піків – 95% по кожній сейсмічній хвилі. У порівнянні з ручною обробкою швидкість аналізу даних збільшилася в 8 разів [14]. У деяких областях штучний інтелект досяг таких вершин, що став перевершувати професіоналів даної галузі.

Основним завданням дипломного проекту з теми «Кластеризація багатовимірних часових рядів в енергетиці та метеорології» є обробка даних, зібраних метеорологічної станцією за 7 років, та на основі цих даних навчання нейронної мережі з метою передбачення погоди на декілька днів вперед.

Актуальність теми. Розробка нейронних мереж почалася дуже давно, але в ті часи можливості обчислювальних системи були обмеженими. Лише в 90-х роках, коли комп'ютери стали достатньо потужні, нейронні мережі набули широкого поширення [31]. Нейронні мережі спочатку були створенні для того, щоб зрозуміти принципи роботи людського мозку. Однак, в порівнянні з людським мозком нейронна мережа сьогодні є вельми спрощеною моделлю, але не дивлячись на це вельми успішно використовуються при вирішенні найрізноманітніших завдань. Великим плюсом нейронних мереж є те, що нейронні мережі «не програмуються», а «навчаються», мережа вчиться виконувати завдання, а не програмується. В перспективі, в областях де необхідно обробляти та аналізувати дані, приймати рішення нейронні мережі можуть повністю замінити людину. На даний момент часу вже існують такі нейронні мережі, які можуть робити прогноз погоди з великою точністю.

Мета і задачі дослідження. Дана робота присвячена створенню програми для обробки даних з метеорологічної станції. На основі цих даних необхідно побудувати та навчати нейронну мережу. Після навчання нейронної мережі вона повинна робити прогноз погоди на декілька днів уперед. Програма буде виконана на мові програмування Python у середовищі розробки програмного забезпечення PyCharm.