

Власник документа:
Бевза Олег Миколайович

ID перевірки:
1004052688

Дата перевірки:
15.06.2020 16:27:09 EEST

Тип перевірки:
Doc vs Internet + Library

Дата звіту:
15.06.2020 16:28:06 EEST

ID користувача:
90740

Назва документа: 2020-bachelor-EDD_Колесник_Нейронна мережа_fch

ID файлу: 1004065575 Кількість сторінок: 21 Кількість слів: 7656 Кількість символів: 56192 Розмір файлу: 83.38 KB

5.36% Схожість

Найбільша схожість: 0.72% з джерело бібліотеки. ID файлу: 1004044647

2.13% Схожість з Інтернет джерелами 24 Page 23

4.44% Текстові збіги по Бібліотеці акаунту 132 Page 23

0% Цитат

Не знайдено жодних цитат

0% Вилучень

Вилучений текст відсутній

Підміна символів

Заміна символів 1

Колесник Олександрі Юріївні
Нейронна мережа з пошкодженими зв'язками

АНОТАЦІЯ

В дипломній роботі представлено огляд науково-технічної літератури по нейронних мережах, їх властивостях та структурах.

Створено нейронну мережу для прогнозування температури навколишнього середовища, проведено тренування, використовуючи дані на інтервалі 4,5 місяці, здійснено редукцію окремих нейронів мережі. Проведено аналіз на основі отриманих залежностей, наведені відповідні графіки.

Згідно отриманих залежностей та результатів дослідження похибок прогнозу температури навколишнього середовища, в дипломній роботі наведено алгоритм створення та навчання нейронної мережі. Описано результати залежності похибок прогнозування значень температури навколишнього середовища від кількості нейронів у мережі, дослідження залежностей виконано для двох випадків: при пошаровому та при комбінаційному пошкодженні шарів. Нейронна мережа перевірена на даних для тренування.

Ключові слова: нейрона мережа; LSTM; прогнозування; пошкодження нейронів.

ANNOTATION

The thesis presents a review of scientific and technical literature on neural networks, their properties and structure.

A neural network was created to predict the ambient temperature, training was performed using data at an interval of 4.5 months, and individual neurons were removed. Afterwards, the analysis was performed on the basis of the obtained dependences, the corresponding graphs are given. The forecast error was also investigated.

According to the obtained dependences and the results of the study of ambient temperature forecast errors, the algorithm of neural network creation and training was presented in the thesis. The results of the dependence of the prediction error of the ambient temperature value on the number of neurons in the network are described. The dependences of the temperature prediction error on the number of neurons in the layer were investigated, performed for two cases: in case of layer - by - layer damage and in case of combination damage of the layer. The neural network was tested on training data.

Keywords: neural network; LSTM; forecasting; damage of neurons;

ВСТУП

Важливою сферою досліджень останні кілька років є прогнозування температури навколишнього середовища. Складність аналізу полягає в мінливому характері погодних даних, а оскільки дані про погоду хаотичні, то штучна нейронна мережа (ШНМ) стала однією з найкращих технік для виявлення залежностей між різними чинниками, що впливають на значення температури навколишнього середовища. У дипломній роботі розглядається застосування ШНМ для прогнозу температури навколишнього середовища, а також порівнюється точність розроблених моделей, які відрізняються між собою кількістю нейронів в шарах [1]. Передбачення про температуру є важливими прогнозами, тому що вони використовуються для захисту життя та власності.

Прогнозування температури - це застосування науки та техніки для прогнозування стану температури на майбутній час та певного місця. Прогнози температури робляться шляхом збору кількісних даних про поточний стан атмосфери. У цій роботі представлений алгоритм прогнозування температури на основі нейронної мережі

На рис.1.1 зображені професіональні сфери, в роботі яких застосовуються штучні нейронні мережі. В медицині, ШНМ використовують для визначення діагнозу та прогнозу протікання процесів, необхідних для одужання пацієнта. Завдяки можливості якісно працювати з великим обсягом математичних даних, ШНМ застосовуються в економіці. Вони дають оцінку ризику неповернення кредитів, передбачення банкрутств, оцінку вартості нерухомості.

В промисловості, завдяки нейронним мережам, здійснюється оптимізація режимів виробничого процесу, комплексна діагностика якості продукції, моніторинг і візуалізація багатовимірної диспетчерської інформації, попередження аварійних ситуацій.

Рис.1.1. Галузі застосування штучних нейронних мереж

Через можливість аналізувати закономірності масових процесів, нейронні мережі здатні прогнозувати стан трафіку на дорогах та визначати швидший маршрут в логістиці. Аналізуючи рис 1.1, прогнозування — це задача, яка вирішується майже в усіх галузях. Енергетика — це найбільша технологічна сфера, яка використовує ШНМ. Енергозбереження — один із актуальних напрямків застосування передбачення.

Прогнозування температури навколишнього середовища — це застосування технічних засобів, для передбачення майбутнього стану температури для певного регіону в заданий час. Прогнозування значень температури відбувається шляхом збору кількісних даних про поточний стан атмосфери, таких як: вологість повітря, атмосферний тиск, швидкість вітру, освітленість та видимість. Дані, зібрані з метеостанцій та супутників, які використовуються для аналізу та прогнозу температури, застосовуються для покращення ефективності використання енергетичних ресурсів. Завдяки

передбаченню температури, можливий контроль систем кондиціонування та нагріву, що значно підвищує енергоефективність об'єкту.

На помилку прогнозу впливають декілька чинників, а саме: хаотичний характер зміни температурних значень та неточне вимірювання початкових даних, які використовуються для тренування нейронної мережі. Внаслідок чого, прогнози стають менш точними. З розвитком науки та технологій, можливо отримати крайнє точні прогнози значень температури незалежно від пори року або локації, для яких проводиться передбачення.

В дипломній роботі, рекурентна нейронна мережа з довгою короткостроковою пам'яттю (LSTM) використовується, як програмна технологія для прогнозування та аналізу майбутнього значення температури навколишнього середовища на основі даних за 4,5 місяці.

1. ЗАСТОСУВАННЯ НЕЙРОННИХ МЕРЕЖ У СИСТЕМАХ SMARTGRID

1.1 Нейронні мережі в інтелектуальних системах

SmartGrid - це новітня всесвітня технологія, яка розглядає процес покращення технічних показників в сфері електроенергетики. Комплект цифрових інновацій для покращення енергетичної інфраструктури притаманний, як для споживачів, так і для постачальників електроенергії та послуг. Технологія започаткувалась в США на початку минулого століття, а на протязі останнього десятиліття впроваджується майже у всіх країнах світу.

Деякі фактори стали передумовою для появи SmartGrid, а саме: зменшення природних енергетичних ресурсів, екологічні зміни навколишнього середовища, збільшення потреб користувачів та кількості використовуваної електроенергії.

Історичним поштовхом для розвитку SmartGrid стало неякісне перебіжне електропостачання в стратегічних містах США наприкінці минулого століття. Технологічною передумовою для модернізації енергетичної інфраструктури були наукові досягнення в області комп'ютерних технологій, електроніки та енергетики.

Для побудови аналізуючої та керуючої частини в розумній мережі SmartGrid використовуються штучні нейронні мережі (ШНМ). ШНМ — це математична модель, або ж її апаратне втілення, яка базується на принципах функціонування біологічної нейронної мережі в людському мозку.

Приклади задач, які нейронні мережі можуть виконувати в інтелектуальних мережах SmartGrid: 1) Керування системою чи окремими її секціями; 2) Розпізнавання режиму роботи; 3) Передбачення різного виду навантаження; 4) Аналіз даних для передбачення аварій; 5) Аналіз минулих аварійних станів для їх уникнення в майбутньому; 6) Регулювання певних показників (напруги, температури, тощо); 7) Виявлення помилок та несправностей;

Баланс виробництва і споживання електроенергії - це основа технологічної стійкості енергосистеми, його порушення позначається на якості електроенергії (відбувається деградація частоти і напруги в мережі), що знижує

ефективність роботи обладнання. Крім того, правильний прогноз дозволяє забезпечити оптимальний розподіл навантаження між об'єктами енергосистеми [7].

1.2 Типи нейронних мереж

Нейронні мережі - це розділ штучного інтелекту, в якому для обробки сигналів використовуються явища, аналогічні тим, що відбуваються в нейронах живих істот. Найважливіша особливість мережі, яка свідчить про її широкі можливості і величезний потенціал, полягає в паралельній обробці інформації усіма ланками, що дозволяє значно прискорити процес обробки інформації. Крім того, при великому числі міжнейронних з'єднань, мережа набуває стійкості до помилок, які виникають на деяких етапах.

Інша, не менш важлива властивість — здатність до навчання і узагальнення накопичених знань. Нейронна мережа має риси штучного інтелекту. Натренована на обмеженій множині даних, мережа здатна узагальнювати отриману інформацію і показувати хороші результати на даних, що не використовувалися при її навчанні

Існує велика кількість видів нейронних мереж, всі вони відрізняються між собою структурою, властивостями та призначенням. Але їх усіх об'єднує історичне походження. Ідея створення штучних нейронних мереж ґрунтувалась на роботі біологічної нейронної мережі в людському мозку [13]. Базуючись на спостереженнях роботи мозку, американські вчені Уоррен С. Мак-Каллок та Вальтер Піттс створили роботу, яка вперше провела аналогію між біологічною та математичною моделлю нейрона. На рис.1.2. зображена класифікація ШНМ.

Рис.1.2 Класифікація штучної нейронної мережі

Штучні нейронні, за своєю структурою, бувають одношарові та багатшарові. На відміну від одношарових мереж, багатшарові містять в собі декілька прихованих шарів між входом та виходом. Мережі, які не мають зворотних зв'язків звуться мережами прямого поширення, ті, які мають зворотні зв'язки, називаються рекурентними або мережами зі зворотними зв'язками.

Нейронні мережі діляться за принципом навчання: з вчителем, без вчителя та змішаний. Нейронна мережа навчається з вчителем, коли вихідний сигнал мережі, порівнюється с заданим прикладом. В процесі навчання мережа самоорганізовується, щоб виводити сигнали якомога ближчі до прикладу. Мережа прогнозує вихідні сигнали, не базуючись на раніше заданому прикладі. Під час використання змішаного методу, комбінуються два вищесказаних метода[11].

Існують задачі, які в змозі вирішити тільки людина, такі завдання також вирішують ШНМ. Найпопулярніші задачі, для виконання яких, потребується використання нейронних мереж [9]:

1. Розпізнавання та детекція об'єктів.

Нейронна мережа здатна розпізнати картинки, букви, символи, цифри, звуки та інші об'єкти. Наразі, детекція та розпізнавання найпопулярніша галузь використання ШНМ.

2. Класифікація.

Розподіл сигналів здійснюється за заданими параметрами.

3. Кластеризація.

Розподіл значень, заданих на вході, на класи. Порівняння та знаходження спільного між різноманітними сигналами. При виконанні задачі кластеризації, ШНМ визначає до якого класу відноситься сигнал, або ж не відноситься, що також дуже корисно, тому що це означає що об'єкт унікальний.

4. Прийняття рішень і управління.

Автоматично включає в себе задачу класифікації. Заздалегідь відсортовані критерії виходу слугують для порівняння результати з ними, від чого залежить прийняття рішення.

5. Прогнозування.

Завдяки зможі знаходити та об'єднувати приховані залежності між сигналами входу і виходу, нейронна мережа здатна передбачати майбутній стан опираючись на минулий досвід.

6. Апроксимація.

Нейронна мережа здатна наближувати будь-яку безперервну функцію з наперед заданою точністю.

7. Стиснення даних і асоціативна пам'ять.

Так як нейронні мережі здатні виявляти схожість між різними ознаками, це дає можливість представити дані більш компактно, за умови що дані пов'язані між собою. Асоціативна пам'ять — процес відновлення вихідного набору даних по частині інформації, що дозволяє також відновлювати вихідний сигнал, пошкоджені вхідні дані та навіть зашумлений сигнал.

Види найпопулярніших нейронних мереж [6]: 1) Нейронна мережа прямого поширення; 2) Радіально-базисна нейронна мережа; 3) Нейронна мережа Хопфілда; 4) Рекурентна нейронна мережа (RNN); 5) Згортова нейронна мережа (CNN); 6) Розгорткова нейронна мережа.

1.2.1. Властивості та структура нейронних мереж

На рис. 1.3. представлена структура штучного нейрону, на якому представлено вектор вхідних сигналів, позначений як x_i , що надходить на штучний нейрон. Вектор сигналів x_i множиться на вектор вагів w_i , і надходить на блок-суматор, позначений Σ . Кожна вага відповідає «силі» однієї біологічної синаптичного зв'язку.

Рис.1.3 Структура штучного нейрону

Блок-суматор, що відповідає тілу біологічного елемента, складає зважені входи алгебраїчно, створюючи вихід, який має назву S . В свою чергу, отримане значення S служить аргументом для функції активації. Саме ця функція визначає значення вихідного сигналу нейрона, що потім надсилається на головний вихід нейрона y . Отже, штучні нейрони об'єднуються шляхом з'єднання виходів та входів між собою. Нейрон складається з трьох основних елементів:

1. ваги (множники), які характеризують важливість зв'язку між нейронами в системі;

2. суматор, що проводить операцію додавання вхідних сигналів, перед цим помножених на відповідні ваги;

3. функція активації, яка зв'язує вхід та вихід системи.

Теоретично описавши, можливо математично зобразити будову нейрона наступним чином:

де x_i — вектор вхідних даних, w_i — вага, b_i — зсув, y — вихід нейрона.

Функція активації представлена у формулі (1.2) як $f(S)$. Нижче наведені головні види функцій та їх графічні зображення на рис.1.4, 1.5, 1.6, 1.7, 1.8.

1. Ступінчата функція (рис. 1.4).

Рис.1.4. Графічне зображення ступінчатої функції

2. Лінійна функція (рис. 1.5).

Рис.1.5. Графічне зображення лінійної функції

3. Сигмоподібна логістична функція.

Дана функція (рис. 1.6) є однією з найпоширеніших, що застосовуються в ШНМ. Вона описується наступним чином:

де a — параметр, зміна якого, дозволяє будувати функцію на графіку з різною крутизною нахилу. Область значень цієї функції лежить в діапазоні від 0 до 1.

Рис.1.6. Графічне зображення сигмоподібної функції.

Однак, іноді потребується область значень в діапазоні від -1 до +1, в цьому випадку, функція визначається наступним чином:

Ця функція називається **сігнум**, в безперервному диференціальному вигляді вона буде мати форму гіперболічного тангенса:

де $f(S)$ — шукане значення елемента, S — зважена сума входів, a — параметр, для зміни нахилу.

4. Кусково-лінійна функція.

Рис.1.7. Графічне зображення кусково-лінійної функції

5. Випрямлена лінійна функція активації (ReLU).

Для здійснення більш точного наближення, система потребує достатньої кількості слоїв та має бути нелінійною. Сигмоподібна або тангенціальна функції є нелінійними, але в процесі тренування їм притаманні багато недоліків. Застосовують простіший варіант — функцію ReLU:

Рис.1.8. Графічне зображення ReLU

1.2.2. Мережа прямого поширення

Нейрофізіолог Френк Розенблатт започаткував **схему пристрою, що моделює процес людського сприйняття, і назвав його «перцептроном»** (від латинського *perceptio* - сприйняття).

Перцептрон є однією із перших моделей нейромереж, на основі якого, побудований перший у світі нейрокомп'ютер «Марк-1».

Мережа прямого поширення (перцептрон) — це штучна нейронна мережа, в якій сигнал поширюється строго від вхідного шару до вихідного. У зворотному напрямку сигнал не поширюється. Після проходження вхідного сигналу через систему генерується вихідний сигнал, який є реакцією мережі на вхідний сигнал. Вони бувають одношарові та багатошарові. Багатошарові

мережі прямого поширення мають 3 або більше шарів. Найчастіше, їх використовують для вирішення задач розпізнавання та аналізу;

1.2.3. Рекурентні нейронні мережі (RNN)

Рекурентна нейронна мережа — це модель глибокого навчання, яка використовується для прогнозування, розпізнавання, тощо. На відміну від традиційних нейронних мереж, рекурентні мережі використовують свою пам'ять (стани) для прогнозування результатів послідовностей, тобто, рекурентна нейронна мережа використовується у випадку передбачення майбутнього результату на основі попередніх послідовних вхідних сигналів. Ключовою особливістю цієї архітектури є обмін інформацією між кроками по часу.

Рекурентна нейронна мережа здатна працювати над послідовними даними, використовуючи зворотні зв'язки, які роблять вихідні значення не тільки залежними від чергових вхідних сигналів, але і від прихованих станів мережі з попереднього часового кроку. [1, ст. 18-20].

RNN представляється в якості безлічі копій одної нейронної мережі, де кожна копія передає повідомлення наступній. Рекурентна мережа володіє природною архітектурою для оперування таким типом даних, як списки чи послідовності.

На рис. 1.9 зображено модель рекурентної мережі, де $h(t)$ – модифікований стан, $h(t-1)$ – попередній стан, тоді як $x(t)$ – сигнал поточних даних, A – комірка пам'яті, $t = 0, 1, 2, \dots, n$ – часовий крок.

Рис.1.9. Модель рекурентної мережі в розгорнутому вигляді

Алгоритм роботи рекурентної нейронної мережі:

1. $X(t)$ подається в мережу.
2. Обчислюється поточний стан сигналу, використовуючи комбінацію поточного вводу та попереднього стану, тобто обчислюється $h(t)$.
3. Поточний $h(t)$ стає $h(t-1)$ для наступного кроку (проходження часових кроків здійснюється в такому обсязі, якого цього вимагає проблема).
4. Здійснюється проходження всіх попередніх кроків для обчислення виходу.
5. Результат порівнюється з фактичним результатом і генерується помилка.
6. Помилка передається в мережу для оновлення ваг і мережа навчається.

У звичайних рекурентних мережах існують певні труднощі з вирішенням проблем, де довгострокові часові залежності є вирішальними. Причина полягає в тому, що під час зворотного розповсюдження, вплив попередніх прихованих шарів занепадає експоненціально з кожним часовим кроком.

Це явище зветься зникаючим градієнтом, і його неможливо подолати, просто збільшуючи вплив попередніх кроків часу, оскільки, в кінцевому підсумку, це призведе до вибуху градієнту, де вплив зв'язку зворотного зв'язку зростає сильніше, ніж вплив вхідного сигналу. [1, ст.31-32].

1.2.4. Мережа LSTM

Проблема градієнта, що зникає, вирішується шляхом заміни нейронів RNN на більш вдосконалені блоки LSTM. Блок LSTM складається з комірки пам'яті та трьох шлюзів: входу, виходу та забуття. Рис.1.10 демонструє основний склад моделі LSTM.

Рис.1.10. Структура моделі LSTM

Крім того, кожен нейрон має періодично зв'язану лінійну одиницю, яку називають "карусель помилок", і її активація називається "стан комірки". Стан комірки виступає як транспортна магістраль, яка передає відносну інформацію всю дорогу по ланцюгу послідовностей. Кожен із шлюзів має власну матрицю вагів, яка тренується аналогічно звичайному штучному нейрону зі застосуванням градієнтного спуску.

Завдяки цим шлюзам, блок LSTM здатний дізнатися, які значення зберігати у своїй комірці пам'яті, коли застосовувати ці значення та коли їх забувати. Замінюючи нейрони RNN на одиниці LSTM, мережа здатна вивчити часові залежності всередині даних, навіть коли між важливими подіями є великі розриви. [1, ст. 31-36] [2].

В шлюзі входу, перед створенням нового стану комірки, перевіряється чи потребує система збереження нового повідомлення. Цей шлюз працює як і з попереднім, так і з наступним сигналом. В результаті, одержується значення. На рис.1.11 продемонстровано процес проходження через шлюз забування, попередній прихований стан $h(t-1)$ та сигнал поточного входу $x(t)$ додаються та передаються через сигмоподібну функцію. Значення виходять в діапазоні між 0 і 1. Ближче до 0 — інформація забувається, а ближче до 1 — зберігається.

Рис.1.11. Графічне зображення операції проходження через шлюз забування

В шлюзі забування вирішується, яку інформацію слід викинути або зберегти. Визначається, скільки інформації передавати в стан комірки через функцію активації [5].

На рис.1.12. зображено процес проходження через вхідний шлюз.

Рис.1.12. Графічне зображення операції проходу через шлюз входу

Для оновлення стану комірок служить вхідний шлюз. На рис.1.10 зображений процес проходу через шлюз входу. Спочатку сума попереднього прихованого стану $h(t-1)$ і сигналу поточного входу $x(t)$ поступає в сигмоподібну функцію, яка визначає, які значення будуть оновлені, шляхом перетворення значень на 0 і 1. На 0 перетворюються неважливі значення, а на 1 важливі.

Паралельно сума $h(t-1)$ і $x(t)$ передається на функцію гіперболічного тангенсу \tanh . Функція \tanh розташовує значення між -1 і 1, тим самим регулюючи вихід нейронної мережі. Виходи з σ і \tanh перемножуються, в результаті створюється кандидат на вихід — C^* .

На рис.1.13. зображено процес оновлення стану комірки.

Спочатку, стан комірки $C(t-1)$ та вихід шлюзу забуття $f(t)$ перемножуються. Далі, обчислений добуток і кандидат на вихід C^* додаються, стан комірки оновлюється на нове значення, яке нейронна мережа вважає релевантним. Створюється оновлений стан комірки $C(t)$.

Рис.1.13. Графічне зображення процесу оновлення стану комірки

Рис.1.14 демонструє процес проходження через шлюз виходу.

Рис.1.14. Графічне зображення процесу проходження через вихідний шлюз

У вихідному шлюзі вирішується, яким має бути наступний прихований стан. Прихований стан містить інформацію про попередні входи та використовується для прогнозування. Сума $h(t-1)$ і $x(t)$ поступає в сигмоподібну функцію. Оновлений стан клітини $C(t)$ поступає на функцію \tanh . Вихід сигмоподібної функції $o(t)$ множиться на вихід функції \tanh , в результаті, отримується прихований стан $h(t)$. Новий стан комірки та новий прихований переносяться на наступний крок часу.

1.2.5. Вибір моделі

При виборі моделі, розглядалися головні види ШНМ, ґрунтуючись на задачі прогнозування. Для прогнозування температури навколишнього середовища потрібна мережа, яка має змогу ідентифікувати непередбачуваний спад чи навпаки підйом температури.

Спершу розглядалися моделі ШНМ, які працюють за принципом прямого розподілення, але так як в такій системі сигнали проходять через мережу лише в одному напрямку, а циклічний зворотний зв'язок неможливий, даний вид нейронної мережі не підходить для вирішення завдання прогнозу. Тому для цієї задачі обрано RNN мережу.

В рекурентній нейронній мережі рух сигналів відбувається в обох напрямках, створюючи цикли. Завдяки цьому, мережа є потужною та динамічною, і може приймати складну структуру. Один із різновидів рекурентних мереж є штучна мережа довгої короткострокової пам'яті, завдяки своїй структурі, майже бездоганно оперує багатоваріантними даними, що ідеально підходить для передбачення такого мінливого параметру, як температура.

1.3. Пошкодження у нейронних мережах

Глибокі нейронні мережі складаються з безлічі окремих нейронів, які поєднують складні та інтуїтивні способи вирішення широкого кола складних завдань. Ця складність надає нейронним мережам свою силу, але із-за своїх властивостей, нейронні мережі вважають чорними ящиками.

Для побудови нових більш складних архітектур, проводяться різноманітні дослідження у ролі окремих нейронів. Задля покращення надійності системи, проводиться дослідження впливу шкоди на нейронну мережу, для визначення важливості окремих нейронів та зв'язків між ними. На рис.1.15. наведено структуру нейронної мережі з відключеним нейроном.

Рис.1.15. Відключення нейрону

Виконується штучне пошкодження слів нейронної мережі, а саме:

1. Пошарове пошкодження. Редукція нейронів в одному шарі.

2. Комбінаційне пошкодження. Редукція нейронів в декількох шарів одночасно. Рис.1.16 демонструє структуру нейронної мережі для пошкодження нейронів.

Рис.1.16. Структура моделі з новоствореними шарами Lambda.

При програмному розробленні, прописуються 2 шари, функції активації яких, є Lambda. Функція активації Lambda:

де — вихідний сигнал попереднього шару;

Вихідний сигнал попереднього шару потрапляє на шар Lambda та обнуляється. В процесі програмної розробки, можливо виділити певний процент нейронів або специфічний нейрон для пошкодження. В розробленій існує можливість виключення нейрона с певного місця.

Для оцінки важливості нейрона, вимірюється зміна точності мережі та аналізується її продуктивність, при видаленні нейрона. Якщо нейрон є дуже важливим, його видалення має бути дуже пошкоджуючим і суттєво знизити продуктивність мережі, тоді як видалення незначного нейрона має незначний вплив.

Висновки до розділу 1

В першому розділі, наведено історію виникнення штучної нейронної мережі. Розглянуто найпоширеніші галузі застосування ШНМ. Описується основні види нейронних мереж, їх властивості та структури. Представлені пояснення до основних функцій активації, наведені графічні зображення. Обґрунтований вибір моделі нейронної мережі прогнозування температури навколишнього середовища. Наведено основні принципи реалізації пошкодження нейронної мережі, а саме нейронного вимкнення.

РОЗДІЛ 2. РОЗРОБКА НЕЙРОННОЇ МЕРЕЖІ

2.1. Призначення нейронної мережі

Завдання прогнозування, в загальному випадку, зводиться до отримання оцінки майбутніх значень, упорядкованих у часі, на основі вже проведеного аналізу, а також, при необхідності, тенденції зміни факторів, що впливають. Іноді виникає необхідність не в прогнозі значень часового ряду на заданому інтервалі, а в прогнозі ймовірності того, що вони будуть вести себе тим чи іншим чином (зростати, спадати, перебувати в деяких межах і т.д.). Задача прогнозування є комплексною, тому підлягає ретельному розгляду. [26,27].

На рис.2.1 наведені основні етапи розробки штучної нейронної мережі.

Рис.2.1. Основні етапи розробки штучної нейронної мережі

На першому етапі чітко формулюються вимоги до виконання. На другому, усуваються несуттєві дані, які не впливають на прогноз.

При необхідності, також, відновлюється пропущена інформація, усуваються аномальні викиди, прибираються високочастотні шуми. Вибирається найбільш підходяща структура нейронної мережі, а також алгоритм і параметри її навчання. Ретельно проведена попередня обробка даних, дозволяє значно поліпшити якість прогнозу.

Проаналізувавши поставленні вимоги до ШНМ, визначивши її архітектуру, відбувається перехід до процесу навчання. Навчання включає в себе декілька процесів: проходження вхідного сигналу через ШНМ, порівняння виходу системи з заданим результатом, обчислення помилки прогнозу, корекція вагів системи [8]. У теорії, кількість шарів і нейронів у шарі може бути вибрано довільним чином, але на практиці воно обмежується ресурсами комп'ютера, на

якому будуть проводитися обчислення і часом, за яке мережа повинна виробляти необхідні обчислення. Однак, чим складніше нейронна мережа, тим масштабніші завдання вона може вирішувати.

Процес функціонування нейронної мережі залежить від величин синаптичних зв'язків (вага мережі), тому, після визначення структури нейронної мережі, для вирішення будь-якої задачі, необхідно знайти оптимальні значення всіх параметрів, що настраюються. Цей процес називається навчанням нейронної мережі, від того, наскільки точно будуть підібрані значення параметрів, залежить здатність мережі вирішувати поставлені перед нею завдання.

Якщо ШНМ працює некоректно, відбувається повернення до кроку вибору архітектури моделі. Проблема навчання перетворюється на завдання оптимізації (мінімізації помилок) з метою мінімізувати функцію втрат, налаштовуючи параметри нейронної мережі. Алгоритм оптимізації, що використовується для навчання нейронної мережі, називається градієнтним спуском.

Спуск градієнта включає обчислення градієнтів функції втрат по відношенню до початкових параметрів. Метод, що використовується для обчислення градієнтів, називається зворотним поширенням [8]. Процес навчання виконується ітераційно. Передача даних під час навчання називається епохою і після кожної епохи, параметри наближаються до оптимальних значень, що мінімізує функцію втрат.

За умови великого розміру набору даних, обчислення втрат та градієнта для всього набору даних є надто повільним та обчислювально нездійсненним. Таким чином, на практиці зазвичай використовується варіант схилу градієнта, який називається стохастичним градієнтним спуском (SGD). У SGD дані поділяються на підмножини, що називаються дозами batches, і параметри оновлюються після обчислення функції втрат протягом однієї дози. Інші популярні варіанти: RMSprop, AdaGrad, Adam.

У деяких із цих варіантів додатковий параметр, який називається розпадом decay, використовується для поступового зниження швидкості навчання, коли параметри наближаються до оптимальних значень.

Найпоширеніша проблема у навчанні нейронних мереж - це перенавчання. Перенавчання виникає тоді, коли модель намагається вписати шум у навчальні дані. Це є результатом використання більш складної моделі, ніж потрібно. Перенавчена модель добре працює на даних, на яких навчалася, а на даних для перевірки погано.

Існує кілька способів запобігти надмірному навчанню. Один із них, рання зупинка. Після кожної епохи значення функції втрати на навчальному наборі порівнюється зі значенням на наборі для перевірки. Якщо втрати на валідаційному наборі починають збільшуватися, хоча втрати на навчальному наборі зменшуються, це є ознакою перевиконання, і тренування моделі можна припинити.

Для боротьби з перенавчанням застосовується проріджування (Dropout), що полягає у видаленні (обнулення) 10% випадково обраних виходів попереднього шару. На етапі тестування проріджування не проводиться.

Важливо зазначити, що мережеві параметри (ваги та зсуви) встановлюються за алгоритмом навчання. Такі параметри, як швидкість навчання, розмір навчальної партії, розпад — параметри алгоритму навчання, встановлюються розробником. Ці параметри, в сукупності, називаються гіперпараметрами.

Тестування нейронної мережі проводиться, використовуючи початкові дані, які були заздалегідь виділені окремо для цього процесу. Зазвичай, приблизно третина вхідних даних виділяється для процесу тестування. Для оцінки коректності роботи ШНМ, розраховується функція втрати. Існують декілька видів функції втрат, найчастіше вживані функції наведені нижче [12]:

1. Проста:

визначається як різниця отриманого значення і фактичного.

де y — значення виходу мережі, яке повинно бути отримано в ідеальному випадку, y^* — фактичний вихід мережі.

2. Квадратична:

де C — константа.

3. Двійкова:

Експоненціальна (функція втрат з насиченням):

Мета прогнозу — це створення моделі, яка здатна передбачити майбутнє та оцінити тенденції в тих чи інших факторах. Якість прогнозу залежить від декількох факторів: наявності попереднього значення передбачуваного параметру та точності виміряних значень.

Класифікація задач прогнозування:

1. За властивостями даних явищ прогнозу :

1) дані представлені у вигляді визначених часових рядів. Завдання такого типу вирішуються за допомогою математичного аналізу.

2) дані представлені у вигляді невизначених часових рядів. Завдання такого типу вирішуються за допомогою теорії ймовірностей.

2. За кількістю параметрів, які мають вплив на явище дослідження:

1) одновимірна задача.

2) багатовимірна задача. Вплив на об'єкт дослідження мають одразу декілька чинників.

Головна задача розробленої нейронної мережі — прогнозування температури навколишнього середовища. Система представляє собою єдиний набір модулів, кожен із яких, вирішує окрему задачу. Модулі можна розділити основні та допоміжні. Основні модулі: навчання та прогнозування. До допоміжних модулів відносять: обробка даних, завантаження фактичної температури, завантаження прогнозованої температури та аналіз помилки моделі.

В якості інструментів для розробки штучної нейронної мережі було вибрано мову програмування Python та середовище розробки Google Colaboratory.

Програма починається з під'єднання до Google Drive, для синхронізації початкових та результуючих файлів Excel. Відбувається запуск бібліотек, які потрібні для функціонування нейронної мережі: Pandas, Keras, Scikit-learn, matplotlib [5].

На стадії обробки, дані для тренування представлені у виді таблиці в середовищі Excel, форматування не потребується. Враховуючи різновид даних, масштабування необхідне, оскільки LSTM мережа працює з представленими числами в діапазоні 0-1. Виклик функції фільтрації даних для видалення пустих комірок в датасеті.

Для збереження якості навчання в нейронній мережі, необхідний належний розподіл даних для навчання, перевірки та тестування. В роботі, набір даних розділений на три частини: навчання (60%), перевірка (20%) і тестування (20%). Тренування нейронної мережі відбувається на навчальному наборі даних, а перевірка на виділених даних для тесту. Під час навчання мінімізується функція середньої квадратичної помилки.

Для виводу значень, використовується обернене нормування даних. Компілювання. Перехід до GoogleDisk, де у вигляді файлу Excel, зберігається новостворений набір спрогнозованих значень. Згідно отриманих табличних значень, будуються графіки кривих прогнозу та похибки.

Мережа з довгої короткостроковою пам'яттю (LSTM) вирішує проблему втрати інформації в рекурентних нейронних мережах, використовуючи шлюзи і явно задану комірку пам'яті. Кожен нейрон має комірку пам'яті і три шлюзи: вхідний, вихідний і забування. Мета шлюзів - захистити інформацію.

Вхідний шлюз визначає, скільки інформації з попереднього шару буде зберігатися в клітці. Вихідний шлюз визначає, скільки інформації отримають наступні шари. Шлюз забування контролює міру збереження значення в пам'яті. Мережі LSTM здатні створювати складні структури, але для цього вони вимагають велику кількість ресурсів.

Для дипломного проєкту було побудовано нейронну мережу для прогнозування температури навколишнього середовища. На рис.2.2 зображена структура створеної мережі, яка складається з трьох шарів. Перший шар складається з 64 нейронів LSTM. Другий щільний шар містить 32 нейрони з функцією активації ReLU, третій, складається лише з одного нейрону і являє собою суматор. Структура розробленої нейронної мережі наведена нижче на рис.2.2.

Рис.2.2. Структура розробленої нейронної мережі

Перший шар складається з нод LSTM, вихідні сигнали яких, поступають на вхід, створюючи зворотній зв'язок. LSTM мережа є універсальною. При достатній кількості елементів в мережі вона може виконувати будь-які обчислення. При навчанні і тренуванні не відбувається розмивання і зникнення даних. Завдяки здатності знаходити довгострокові залежності і нелінійну

динаміку в даних часових рядів, цей тип мереж знайшов застосування в різних додатках, включаючи прогнозування.

Другий щільний шар складається з нейронів, які працюють за випрямленою лінійною функцією активації ReLU. Нейрони в другому шарі мають безліч входів, але лише один вихід. Кількість виходів прямо пропорційно залежить від кількості нейронів.

Вихідний щільний шар складається лише з одного нейрона, який за походженням ідентичний нейронам з другого шару, має лише один вихід та безліч входів. На виході мережі повинно бути лише одне значення температури, тому одного нейрону в третьому шарі цілком достатньо. Дана нейронна мережа є добре пристосованою до завдань прогнозування.

2.2. Опис роботи нейронної мережі

Нейронна мережа тренується на даних, які зібрані з метеостанції авіабази Києврес, Бельгія. Завантажені із загальнодоступної бази наборів даних із Надійного прогнозу, gr5.ru. Отримані дані, зняті інтервалом в 10 хвилин 4,5 місяці починаючи з 01.01.2019 року. В базі даних присутні такі параметри: температура наколишнього середовища [°C], атмосферний тиск [мм.рт.ст.], вологість у [%], освітлювальність [лм], швидкість вітру [м/с], видимість [%]. Дані для тренування наведені в Додатку А.

В якості мови був обраний Python. Дана мова є інтерпретується. Інтерпретатор мови представлений на всіх основних платформах: MacOS, Windows, Linux. Python є високорівневою, що дозволяє вести швидко розробку, але при цьому не погіршує читаність коду. Дана мова володіє динамічною типізацією, механізмом обробки винятків, автоматичним управлінням пам'яттю. Мова має чіткий і послідовний синтаксис, продуману модульність і масштабованість, завдяки чому, вихідний код, що написаний на Python легко читаємий.

Нейронну мережу розроблено в середовищі GoogleColaboratory. Colaboratory - це хмарна платформа від Google для просування технологій машинного навчання, яка надає безкоштовну віртуальну машину з встановленими популярними бібліотеками TensorFlow, Keras, sklearn, pandas і т.п. Найзручніше, що на Colaboratory можна запускати блокноти, для написання коду. Блокноти зберігаються на Google Drive, можна їх поширювати і навіть організувати спільну роботу. В даному середовищі можна створювати, навчати, використовувати мережу, а також варіювати її параметри.

Алгоритм роботи мережі:

1) Завантаження бібліотек, які потрібні для функціонування нейронної мережі: Pandas, Keras, Scikit-learn, matplotlib

from math import sqrt — виклик функції квадрату

from sklearn.model_selection import train_test_split —виклик розподілення

даних на тренування, перевірку, тестування

import numpy as np — бібліотека масивів

from numpy import concatenate — викли функції з'єднання масивів вздовж

вказаної осі

```
from pandas import read_csv —читання формату csv
from pandas import DataFrame — виклик табличної структури даних
from pandas import concat — виклик функції об'єднання даних
from pandas import to_datetime — виклик функції вказування дати про
обробці даних
from sklearn.preprocessing import MinMaxScaler — нормування даних
from sklearn.preprocessing import LabelEncoder —
from sklearn.metrics import mean_squared_error — виклик функції
середньоквадратичної помилки
from keras.models import Sequential — виклик послідовної моделі
from keras.layers import LSTM — виклик моделі довгої короткочасної
пам'яті
from keras.layers import Dense — виклик щільного шару
from keras.layers import Dropout —виклик функції захисту від
перенавчання
from keras.layers import Lambda — виклик шару Lambda
from keras.losses import MeanSquaredError —виклик функції втрат
from matplotlib import pyplot — виклик побудови графіку
```

2) Обробка даних. Так як в базі даних заданий великий обсяг даних, в датасеті присутні пусті значення. Для очищення пустих комірок таблиці використовується метод заміни відсутніх значень, шляхом прямого заповнення їх попередніми моментами часу (наприклад, заповнення відсутнього значення температури останнім зафіксованим значенням температури). В програмній реалізації, ця процедура виконана за допомогою команди `dropna`.

3) Розподіл даних для нейронної мережі на три набори: для навчання 60%, перевірки 20% та тестування 20%. Дані для перевірки потрібні для оцінювання якості навчання. Для перевірки вже навченої нейронної мережі застосовують дані для тестування [12].

4) Нормування значень. Так як усі дані з бази для навчання лежать в різному діапазоні значень, проводиться нормалізація даних, тобто, всі значення масштабуються від 0 до 1, окрім вибірки для тестування. В програмній розробці це виконується викликом команди `MinMaxScaler`.

5) Перевірка якості прогнозу, з використання середньоквадратичної функції втрат, в програмі `MeanSquaredError`, яка розраховується за формулою:

На рис.3.1 зображені дві криві: крива блакитного кольору показує значення прогнозу, а помаранчева крива — заданого значення, в певний проміжок часу. Для кращої демонстрації, було взяті дані на інтервалі однієї ГОДИНИ.

Рис 3.1.Графік кривих прогнозу та заданих значень

Середньоквадратична помилка для розробленої нейронної мережі складає 0.08. Тобто, точність прогнозу складає приблизно 98,74% , що стверджує про коректну роботу нейронної мережі.

Висновки до розділу 2

Розглянуто етапи побудови моделі прогнозування температури навколишнього середовища, використовуючи реальні спостережувані погодні дані. Продемонстровано структуру розробленої мережі. Наведено процес обробки даних та алгоритм роботи мережі. Розраховано середньоквадратичну помилку, продемонстровано результати прогнозування.

РОЗДІЛ 3. ДОСЛІДЖЕННЯ НЕЙРОННОЇ МЕРЕЖІ З ПОШКОДЖЕННИМИ ЗВ'ЯЗКАМИ

3.1. Пошкодження першого шару

Пошкодження нейронної мережі проводиться з метою оцінювання надійності системи. Крім оцінювання похибки прогнозу, проводиться аналіз тенденції наслідування зміни кривої прогнозу нормально функціонуючої системи від пошкодженої. З теоретичних роздумів, чим більше ураження нейронної системи — тим недостовірніший прогноз моделі. Кількість уражених нейронів вибирається на власний розсуд.

Як описано в пункті 1.3, з метою імітації пошкодження нейронної мережі, прописуються два додаткові шари з функцією активації Lambda. Перший шар в розробленій нейронній мережі складається з 64 нейронів. В процесі виконання роботи, при пошкодженні першого шару, була здійснена поступова редукція кількості нейронів: 2, 16, 25, 32, 45, 56, 63. Запуски пошкоджених нейронних мереж проводилися при різних значеннях кількості нейронів, інші ж параметри нейронної мережі не були змінені. Розрахована середньоквадратична похибка для усіх пошкоджень, наведена в таблиці 3.1.

З рис. 3.1, не зважаючи на пошкодження 2х нейронів, отримані значення прогнозу мають відносно невелику MSE та слідує тенденції зміни кривої непошкодженої мережі. На рис.3.2., при ураженні 16-ти нейронів, мережа добре прогнозує, хоча похибка більша ніж при пошкодженні 2х нейронів. При ураженні 32 нейронів, мережа добре прогнозує локальні мінімуми та максимуми, а проміжні значення згладжує. Це свідчить про те, що при пошкодженні 50% відсотків шару, система буде якісно прогнозувати граничні значення.

З рис.3.5, видно, що система реагує лише на локальні мінімуми, що для деяких задач прогнозування, може бути прийнятним. На рис 3.6 система перестає працювати коректно, а при пошкодженні 63 нейронів, як на рис.3.7, мережа не працює зовсім.

Рис.3.2. Графік кривих після редукції 2х нейронів в першому шарі

Рис.3.3. Графік кривих після пошкодження 16 нейронів в першому шарі

Рис.3.4. Графік кривих після пошкодження 32х нейронів в першому шарі

Рис.3.5. Графік кривих після пошкодження 45 нейронів в першому шарі

Рис.3.6. Графік кривих після пошкодження 56 нейронів в першому шарі

Рис.3.7. Графік кривих після пошкодження 63 нейронів в першому шарі

В табл. 3.1 наведено середньоквадратичну похибку для пошкоджень в шарі.

Таблиця 3.1

Середньоквадратична похибка для пошкоджень в шарі

3.2. Пошкодження другого шару мережі

Пошкодження нейронів в другому шарі мережі відбувалось аналогічним чином, але так як другий шар мережі вдвічі менший, кількість пошкоджених нейронів була змінена на: 2, 8, 16, 31. Графічні зображення кривих наведені нижче на рис.3.10, 3.11, 3.12, 3.13.

Рис.3.10. Графік кривих після поразені 2 нейронів в другому шарі

Рис.3.11. Графік кривих після поразені 8 нейронів в другому шарі

Рис.3.12. Графік кривих після поразені 16 нейронів в другому шарі

Рис.3.13. Графік кривих після поразені 30 нейронів в другому шарі

В табл. 3.2. наведено вплив пошкодження на MSE.

Таблиця 3.2.

Таблиця впливу пошкодження на MSE.

Аналізуючи наведені графіки та проводивши аналогію з пошкодженням першого шару, виявлено, що другий шар більш чутливий до пошкоджень. Це обумовлено різною структурою нейронів в шарах, загальною кількістю нейронів і їх функцією активації. Функція активації прихованого шару— ReLU, а LSTM — гіперболічний тангенс, який має діапазон значень від -1 до 1.

З рис.3.10, видно, що пошкодження 2 нейронів майже не впливає на роботу мережі. На рис.3.11 видно, що при ураженні 8 нейронів, мережа не перестає виконувати роботу, крива прогнозу повторює криву діючих значень. Рис.3.12 демонструє ураження 16-ти нейронів, тобто 50% системи, з графіка видно, що при такому ураженні, система працює некоректно. На рис.3.13, результати пошкодження демонструють, що при 94% пошкоджень шару (30 нейронів), система перестає працювати зовсім, так як не реагує на будь-які зміни кривої діючих значень. Це підтверджує таблиця 2, при поступовому пошкодженні мережі, середньоквадратична похибка зростає.

3.3. Комбінаційне пошкодження

Під комбінаційним пошкодженням мається на увазі одночасне пошкодження першого і другого шару. Для дослідження впливу комбінаційного пошкодження експеримент розділяється на 3 частини: 1) однаковий відсоток пошкодження шарів; 2) ураження першого шару втричі більша ніж другого; 3) ураження другого шару втричі більша ніж першого.

Для експерименту 1, вибірка пошкоджених нейронів становить: 10%, 25%, 50%, 75%. Нижче наведені графічні зображення впливу пошкоджень на прогноз на рис. 3.16, 3.17, 3.18, 3.19.

Рис.3.16. Графік кривих після редуції 10% нейронів в обох шарах

Рис.3.17. Графік кривих після пошкодження 25% нейронів в обох шарах

Рис.3.18. Графік кривих після пошкодження 50% нейронів в обох шарах

Рис.3.19. Графік кривих після пошкодження 75% нейронів в обох шарах

В процесі проведення експерименту 1, виявлено, що комбінаційне пошкодження впливає на систему ефективніше ніж пошарове.

Рис.3.16 зображує пошкодження 10% першого і другого шару одночасно, з рисунку видно, що крива прогнозу майже співпадає з кривою діючих значень, тобто пошкодження майже не впливає на мережу.

При 25% відсотках ураження, система починає гірше прогнозувати локальні мінімуми, як видно з рис.3.17.

З рис.3.18, при пошкодженні 50% в обох шарах система перестає працювати коректно, крива прогнозу змінюється, але не суттєво.

Рис.3.19 демонструє, що при 75% пошкодженнь в обох шарах, система перестає працювати зовсім.

Для експерименту 2, наведені графічні зображення на рис. 3.20, 3.21, 3.22.

Рис.3.20. Графік кривих після пошкодження 30% в першому шарі і 10% в другому

Рис.3.21. Графік кривих після пошкодження 60% в першому і 20% в другому

Рис.3.22. Графік кривих після пошкодження 90% в першому шарі в 30% в другому

Результати другого експерименту, демонструють, що при такій комбінації пошкодженнь, система працює.

Рис.3.22 демонструє, що навіть при 90% пошкодженнь в першому шарі і 30% в другому, система реагує на вхідні сигнали, але якість прогнозу неприйнятна.

При проведенні експерименту 3, отримано графічні зображення, які наведені нижче на рис. 3.23, 3.24, 3.25.

Рис.3.23. Графік кривих після пошкодження 10% в першому шарі в 30% в другому

Рис.3.24. Графік кривих після пошкодження 20% в першому шарі в 60% в другому

Рис.3.25. Графік кривих після пошкодження 30% в першому шарі в 90% в другому

З рис.3.24 та рис.3.25. пошкодження мали не суттєвий вплив. Порівнюючи рис. 3.25 та рис.3.22, видно, що у випадку комбінації 30%-90% система чітко демонструє локальні максимуми та мінімуми, працездатність такої системи вища.

Висновки до розділу 3

В даному розділі було проведено два типи пошкодженнь: пошарове та комбінаційне. При пошаровому пошкодженні першого шару, система працювала задовільно навіть при 96% пошкодження. Здійснивши вимкнення нейронів у другому шарі та порівнявши результати з першим шаром, було виявлено, що другий шар більш чутливий до пошкодженнь ніж перший. Ураження в другому шарі продемонстрували високу чутливість системи.

При комбінаційному пошкодженні було проведено 3 експеримента, при пошкодженні однакового відсотка нейронів в мережі, спостерігається ефект, починаючи с 50% система не працює коректно. Під час проведення експериментів 2 і 3, було виявлено, що у разі пошкодження другого шару втричі більше першого, система працює краще ніж при ураженні першого втричі більше другого.

ВИСНОВКИ

В дипломній роботі реалізована нейронна мережа для прогнозування температури навколишнього середовища, вибір структури мережі обґрунтований, теоретичний опис розробленої нейронної мережі представлений. Реалізована мережа має гарну точність, близько 98%.

Проведено пошкодження мережі : 1-го шару, 2-го шару та комбінацій обох шарів. Досліджено вплив пошкодженнь на працездатність мережі. Продемонстровано граничні умови роботи мережі. Для оцінювання якості прогнозу використана середньоквадратична помилка. Наведено графічні

зображення впливу пошкодження на працездатність мережі, наведені таблиці середньоквадратичних помилок.

SUMMARY

Neural network with damaged connections

Diploma project of educational and qualification level "Bachelor" specialty 171 - Electronics, specialization - Electronic devices and systems. Kolesnyk Oleksandra Yuriyivna, KPI named after Igor Sikorsky. Faculty of Electronics, Department of Electronic Devices and Systems. DS-61 group. - К.: National Technical University of Ukraine Igor Sikorsky Kyiv Polytechnic Institute, 2020. - 60p., Ill. 43, table. 2.

Theme, purpose and nature of the work:

Thesis is devoted to the study of the effects of network damage on its performance.

Purpose: to develop a neural network to predict ambient temperature and simulate neuronal damage.

The work is applied.

Used methods and software products: recurrent neural network; calculation of the mean square error to assess the quality of forecasting; software development of the network and its damage.

The Python programming language and the Google Collaborative development environment were chosen as tools for developing an artificial neural network.

Summary: The first section describes the classification of neural networks and their applications. After analyzing these data, it was found that neural networks are widely used in many industries, such as economics, medicine, logistics, industry and energy.

The most common is the energy sector. Neural networks solve a class of tasks such as forecasting and monitoring power consumption, emergency forecasting, predicting changes in certain parameters (light, humidity and temperature. Neural networks are also used in smart grids Smart Grid.

Smart Grid is a technology that provides for the improvement of technical indicators in the field of energy. Smart Grid technology is used all over the world, and the precondition for this was the evolutionary demand for electricity and scientific advances in the fields of electronics and computer science. Artificial neural networks in intelligent networks are used to analyze data, predict accidents, control the load and control the system or its individual parts.

The classification of neural networks is described. Artificial neural networks are divided according to: structure, method of learning and tasks for which they are used. The structures of neural networks are considered. By type of connection, they are divided into direct distribution and reverse. In direct propagation, the signal goes from the input through the entire system to the output, and in reverse, loops are created, through which the signal returns to a certain block and a cycle is obtained.

Due to their analogy to a biological network, artificial neural networks are able to detect the relationship between different features, which allows you to present data more compactly and solve complex computational problems. Tasks that require the

use of artificial neural networks: classification, clustering, recognition and detection of objects, prediction, approximation. Types of the most popular neural networks: 1) neural network of direct propagation; 2) Hopfield neural network 3) recurrent neural network; 4) convolutional network; 5) scanning network.

The structure and principle of operation of an artificial neuron are described, its structure and mathematical description are given. An artificial neuron consists of three main blocks: weight, which characterizes the importance of a particular signal; an adder that adds input signals multiplied by weights; an activation function that connects the input and output of the network. The main types of activation functions and their graphic images are given. The most common functions: stepped, linear, sigmoid, piecewise linear, rectified linear.

A brief description of the direct distribution network, its historical origin and areas of application is given. The recurrent neural network is considered, the algorithm of its work and structure are given. The algorithm consists of three stages: preliminary processing of input data; determining the number of neurons in the hidden layer or layers; construction of a prediction neural network.

At the first stage, pre-processing of time series is performed, in particular the procedure of removing noise and fluctuation components in time series; normalization of data (reduction to intervals $[0 \dots 1]$ or $[-1 \dots 1]$), in order to increase the accuracy of information before the next stages of forecasting. In order for the obtained values at the output of the neural network to correspond to the real data, the inverse of the normalization of the transformation is performed. After processing, training, coaching and data validation are performed. Checking the quality of the forecast is performed using the root mean square loss function. The forecast error is influenced by several factors, namely: the chaotic nature of changes in temperature values and inaccurate measurement of the initial data used to train the neural network. As a result, predictions become less accurate.

A feature of the recurrent neural network is the use of network memory for prediction. A recurrent neural network is able to work on sequential data using feedback, which makes the output values not only dependent on the next input signals, but also on the hidden states of the network from the previous time step.

The main disadvantage of recurrent neural networks is the vanishing gradient when working with long-term time dependences. The solution of this problem, by replacement is considered. neurons of the recurrent network to advanced long-term memory nodes (LSTM).

The structure of the LSTM neuron, which consists of three main elements: input gateway, output gateway, and forget gateway, is considered. A detailed description of network operation and information processing operations through gateways is given.

Due to the described types of neural networks, the model for network development was selected. An artificial neural network of long-term memory (LSTM) has been chosen, which, due to its structure, handles multivariate data almost flawlessly, which is ideal for predicting such a variable parameter as temperature.

The basic principles of realization of network damage, namely neural shutdown are described.

The second section "Development of the neural network" describes the task of forecasting and presents the main stages of development of an artificial neural network. The classification of forecasting problems is considered. The main task of the network for development is to forecast the ambient temperature. Its structure is given.

The developed neural network consists of three layers: the first layer consists of 64 LSTM neurons, the second dense layer contains 32 neurons with ReLU activation function, the third consists of only one neuron and is an adder.

The Google Collaboratory platform is used to write the software. Python is the chosen programming language. There is a connection to Google Drive, to synchronize the initial and resulting Excel files.

A neural network for ambient temperature prediction has been developed. The implemented network has good accuracy, about 98%. Damage to the network: 1st layer, 2nd layer and combinations of both layers. Layer-by-layer damage was performed for: 2, 16, 32,45, 56, 63 for the first layer, 2,8,16,31 order of the number of removed neurons.

With layer-by-layer damage to the first layer, the system worked satisfactorily even at 96% damage. By switching off the neurons in the second layer and comparing the results with the first layer, it was found that the second layer is more sensitive to damage than the first. Lesions in the second layer showed high sensitivity of the system.

For combination damage, 3 experiments were performed: damage of the same percentage in both layers, damage to the first layer three times more than the second, damage to neurons in the second layer more than three times more than in the first. In Experiment 1, it was found that the combination damage affects the system more effectively than layer by layer. During experiments 2 and 3, it was found that in the case of damage to the second layer three times more than the first, the system works better than in the defeat of the first three times more than the second.

The influence of damages on network efficiency is investigated. When analyzing the results, it was found that the second layer of the neural network is more sensitive to damage. This is due to the different structure of neurons in the layers, the total number of neurons and their activation function.

In the thesis, a recurrent neural network with long short-term memory (LSTM) is used as a software technology to predict and analyze the future value of ambient temperature based on data for 4.5 months. The boundary conditions of the network are demonstrated. The standard error was used to assess the quality of the forecast. Graphical images of the impact of damage on network performance are given, tables of standard errors are given.

Схожість

Схожість із джерелами з Інтернету

24

2	https://knowledge.allbest.ru/programming/3c0a65635a2bc78a5d43b88421206d36_0.html	0.68%
3	https://nauka-online.com/wp-content/uploads/2019/06/Lyesna-Minakov.pdf	0.57%
5	http://masters.donntu.org/2009/kita/stikhar/diss/indexu.htm	0.38%
12	http://bukvar.su/jekonomiko-matematicheskoe-modelirovanie/213911-Neironnye-seti.html	0.27%
13	https://ukrbukva.net/76943-Neironnye-seti.html	0.26%
18	https://kivra.kpi.ua/wp-content/uploads/file/work/2018/Kovbych/Kovbych_magistr.pdf	0.21%
29	https://webometr.kpi.ua/en/citation-data	0.13%
31	https://www.datatechnotes.com/2019/03/classification-with-adaboost-classifier.html	0.12%
34	https://otherreferats.allbest.ru/medicine/00395747_0.html	0.1%
36	https://www.hindawi.com/journals/aav/2009/432340	0.1%
37	https://www.lnu.edu.ua/wp-content/uploads/2019/09/aref_bila.pdf	0.1%

Схожість по Бібліотеці акаунту

132

1	2020-bachelor-EDD_Natalych_muz_syntezator_fch	ID файлу: 1004044647	Institution: National Technical Uni	0.72%
4	2020-bachelor-EDS_Zelinskyy_pokryttya_fch	ID файлу: 1004030960	Institution: National Technical Universit	0.5%
6	Bahlaii-1-2019m	ID файлу: 1000776062	Institution: National Technical University of Ukraine "Kyiv Polytechnic Institu...	0.35%
7	2020-bachelor-EDD_Ilchenko_rezervne_zhyvlennya_fch	ID файлу: 1004048808	Institution: National Technical Universi...	0.33%
8	Диплом_Швець_ред13_12	ID файлу: 1000785614	Institution: National Technical University of Ukraine "Kyiv Polyte...	0.33%
9	ЧАЙКОВСЬКИЙ_Олександр_0d7ca8_main_part	ID файлу: 1000742264	Institution: National Technical Uni	0.31%
10	Студентська робота	ID файлу: 1000991491	Institution: National Aviation University	0.27%
11	Буханенко_КА82мп	ID файлу: 1000763737	Institution: National Technical University of Ukraine "Kyiv Poly	0.27%

14	Студентська робота	ID файлу: 5888328	Institution: Lviv Polytechnic National University	0.25%
15	Tilniak_Yurii	ID файлу: 5615485	Institution: National Technical University of Ukraine "Kyiv Polytechnic Institute"	0.25%
16	Третяк В А	ID файлу: 1000764626	Institution: National Technical University of Ukraine "Kyiv Polytechnic Institute"	0.24%
17	Диплом_Андржієвська_PA81мп	ID файлу: 1000763507	Institution: National Technical University of Ukra	2 Джерело 0.22%
19	2020-bachelor-EDD_Shut_vyavlennya_OEZ_fch	ID файлу: 1004048810	Institution: National Technical University of Uk...	0.21%
20	Khrolenko_2020b	ID файлу: 1004030617	Institution: National Technical University of Ukraine "Kyiv Polytechnic Instit...	0.21%
21	Студентська робота	ID файлу: 1004039246	Institution: Lviv Polytechnic National University	2 Джерело 0.2%
22	2020-bachelor-EDD_Salamatin_elektronna_harmata_fch	ID файлу: 1004040160	Institution: National Technica	8 Джерело 0.16%
23	Студентська робота	ID файлу: 1029890	Institution: Lviv Polytechnic National University	0.16%
24	Madojan_bachelor	ID файлу: 1000017951	Institution: National Technical University of Ukraine "Kyiv Polytec	2 Джерело 0.14%
25	Пікож_диплом_печать	ID файлу: 1000084261	Institution: National Technical University of Ukraine "Kyiv	3 Джерело 0.14%
26	Студентська робота	ID файлу: 3542473	Institution: Lviv Polytechnic National University	10 Джерело 0.13%
27	9 Нагорний	ID файлу: 1000105668	Institution: National Technical University of Ukraine "Kyiv Polytechnic	30 Джерело 0.13%
28	Студентська робота	ID файлу: 3555845	Institution: Vasyl Stus Donetsk National University	2 Джерело 0.13%
30	Гриценко_магістерська	ID файлу: 1000795272	Institution: National Technical University of Ukraine "Kyiv Polytec...	0.12%
32	Trachuk_bak_2020	ID файлу: 1003928804	Institution: National Technical University of Ukraine "Kyiv Polytec	5 Джерело 0.1%
33	Aleksandrova	ID файлу: 1000804732	Institution: National Technical University of Ukraine "Kyiv Polytechnic Institute"	0.1%
35	Студентська робота	ID файлу: 8263117	Institution: Vasyl Stus Donetsk National University	0.1%
38	Студентська робота	ID файлу: 9458942	Institution: National Aviation University	0.1%