

Висновок

Застосування звичайної GRU для розпізнавання тональності тексту на базі набору даних STT виявилось ефективним та практично здійсненим підходом. Виділення ключових методичних аспектів, аналіз предметної галузі та створення детального опису нейронної мережі дозволили створити модель, яка здатна до точного та швидкого розпізнавання цифр.

Результати експериментів та оцінки моделі на тестовому наборі відображені в розділі 4 цієї роботи, підтверджують достатньо високу ефективність застосованого підходу. Точність класифікації та інші метрики підтверджують здатність навченої мережі правильно розпізнавати рукописні цифри з задовільною точністю.

Здобуті знання та результати можуть служити основою для подальших досліджень у галузі розпізнавання образів та розвитку більш складних моделей. Дана робота відкриває нові можливості для впровадження подібних систем у сучасних технологічних рішеннях в сферах розваг, освіти, медицини та інших, де важливим є оцінити стан та настрої клієнта за текстом.

Список джерел

1. Головна сторінка документації Python-telegram-bot.: [Електронний ресурс]. – Режим доступу: <https://docs.python-telegram-bot.org/en/v20.7/>
2. Головна сторінка документації Tensorflow.: [Електронний ресурс]. – Режим доступу: https://www.tensorflow.org/api_docs/python/tf
3. Загальний огляд Keras API.: [Електронний ресурс]. – Режим доступу: <https://www.tensorflow.org/guide/keras>
4. Каллан, Р. Нейронні мережі: Короткий довідник / Р. Каллан. - М.: Вільямс І.Д., 2017. - 288 с
5. Опис набору даних Stanford sentiment treebank.: [Електронний ресурс]. – Режим доступу: <https://nlp.stanford.edu/sentiment/>

ДОСЛІДЖЕННЯ ТА АНАЛІЗ ВИКОРИСТАННЯ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ РОЗПІЗНАВАННЯ РУКОПИСНИХ ЦИФР

Халимон В.К.

Харківський національний автомобільно-дорожній університет

У сучасному світі нейронні мережі стали неот'ємною частиною багатьох сфер, включаючи обробку зображень та розпізнавання образів. Однією з ключових задач є розпізнавання рукописних цифр, що може бути вирішено за допомогою нейронних мереж. У даній курсовій роботі буде розглянуто застосування звичайної MLP (Multilayer Perceptron) для розпізнавання цифр на основі відомого набору даних MNIST.

Розвиток нейронних мереж та їх використання в різних областях життя в останні роки значно зростає. Однією з ключових областей, де виявляється великий потенціал нейронних мереж, є розпізнавання образів. У цьому контексті особливий інтерес представляє завдання розпізнавання рукописних цифр.

З впровадженням технологій машинного навчання стає можливим вирішення цього завдання за допомогою нейронних мереж, зокрема, Multi-Layer Perceptron (MLP) або звернутої мережі. Одним із стандартних наборів даних для такого роду завдань є MNIST (Modified National Institute of Standards and Technology), що містить рукописні зображення цифр від 0 до 9.

Розпізнавання рукописних цифр має широкі застосування в сучасному світі, починаючи від систем автоматичного сортування пошти та закінчуючи розпізнаванням текстів на медичних зображеннях. Використання нейронних мереж у цьому контексті дозволяє створювати ефективні та точні системи, які можуть забезпечити високий рівень автоматизації у розпізнаванні рукописних символів.

Характеристика об'єкта дослідження

Для реалізації розпізнавання рукописних цифр буде використано набір даних MNIST, який містить 60 тисяч тренувальних та 10 тисяч тестових зображень рукописних цифр. Описаний підхід використовує нейронну мережу зі стандартними шарами (вхідний, прихований, вихідний) та функцією активації для кожного шару.

MNIST включає чорно-білі зображення розміром 28x28 пікселів, представлені у вигляді матриці пікселів.



Рисунок 1 – Приклад набору MNIST

Набір даних MNIST (Modified National Institute of Standards and Technology) став фундаментальним об'єктом для багатьох досліджень у галузі машинного навчання та комп'ютерного бачення. Створений для ефективно оцінки та порівняння алгоритмів розпізнавання образів, він включає 70 тисяч чорно-білих зображень рукописних цифр, які розміром 28x28 пікселів. Цей набір даних відзначається своєю простотою та однообразністю, що робить його ідеальним об'єктом дослідження для вивчення та розробки моделей машинного навчання.

Характеристики зображень

Кожне зображення у наборі MNIST представлене у форматі матриці пікселів, де кожен піксель визначає інтенсивність чорного або білого кольору. Розмір кожного зображення складає 28 на 28 пікселів, утворюючи область 784 пікселів. Така структура дозволяє представити кожен цифру у вигляді вектора, який може бути подано нейронній мережі в якості вхідного сигналу для навчання та тестування.

Задача розпізнавання рукописних цифр

Головною задачею у цьому дослідженні є розпізнавання рукописних цифр за допомогою нейронної мережі. Кожній цифрі від 0 до 9 призначено відповідна мітка, і модель навчається класифікувати цифри на основі зразків, які представлені у наборі даних. Це завдання становить виклик для нейронних мереж, оскільки вони повинні навчитися визначати унікальні особливості кожної цифри та генерувати відповіді, що максимально наближені до дійсних значень.

Опис обраної нейронної мережі

Обрана нейронна мережа є мережею MLP з одним вхідним шаром (784 нейрони, що відповідають розміру зображення), одним або кількома прихованими шарами та вихідним шаром (10 нейронів, оскільки є 10 класів цифр). Використовується функція активації ReLU для прихованих шарів та softmax для вихідного шару.

Архітектура MLP (Multilayer Perceptron)

Обрана нейронна мережа базується на звичайній архітектурі Multilayer Perceptron (MLP), яка визначається своєю здатністю до навчання та адаптації до складних взаємозв'язків у вхідних даних. MLP складається з трьох типів шарів: вхідний, прихований і вихідний.

Вхідний шар

Вхідний шар має 784 нейрони, що відповідають розміру зображення, яке подається на вхід нейронній мережі. Кожен нейрон в цьому шарі представляє піксель зображення розміром 28x28 пікселів, яке плоско розглядається.

Приховані шари

Мережа може містити один або декілька прихованих шарів, де кожен нейрон в прихованому шарі взаємодіє з кожним нейроном попереднього та наступного шарів. Використовуються функції активації ReLU (Rectified Linear Unit) для прихованих шарів, які надають можливість моделі виявляти та вивчати нелінійні залежності в даних.

Вихідний шар

Вихідний шар має 10 нейронів, оскільки задача розпізнавання рукописних цифр полягає в класифікації на 10 категорій (цифри від 0 до 9). Для отримання ймовірностей кожного класу використовується функція активації softmax, яка перетворює виходи нейронів в ймовірності.

Застосування функцій активації

Функції активації ReLU для прихованих шарів дозволяють уникнути проблеми ванішінга градієнта, забезпечуючи ефективне навчання мережі. Функція активації softmax вихідного шару перетворює вихідні значення в ймовірності та робить модель придатною для многокласової класифікації.

Загальна концепція

Отже, обрана архітектура нейронної мережі створює просторову модель для вивчення складних взаємодій у вхідних даних та ефективної класифікації рукописних цифр. Підбір параметрів та функцій активації є ключовим для досягнення високої точності розпізнавання в межах даної задачі.

Реалізація створеної нейронної мережі

Для реалізації нейронної мережі використовується бібліотека TensorFlow з використанням Keras API. Код розбитий на етапи, включаючи завантаження даних, побудову моделі, навчання та оцінку результатів. Застосовується звичайна процедура тренування за допомогою стохастичного градієнтного спуску та зворотнього поширення помилки.

Для практичної реалізації створеної нейронної мережі використовується потужна бібліотека TensorFlow, яка забезпечує ефективні інструменти для створення, навчання та експериментування з нейронними мережами. Реалізація проводиться через високорівневий інтерфейс Keras API, який дозволяє створювати та конфігурувати нейронні мережі з легкістю.

Основні етапи включають:

1. Завантаження даних: В цьому етапі використовуються інструменти TensorFlow для завантаження набору даних MNIST, які використовуються для тренування та тестування моделі.

2. Побудова моделі: Використовуючи Keras API, створюється архітектура нейронної мережі з вхідним, прихованим та вихідним шарами. Параметри моделі, такі як кількість шарів, кількість нейронів та функції активації, конфігуруються в цьому етапі.

3. Навчання моделі: Використовуючи стохастичний градієнтний спуск та зворотне поширення помилки, модель навчається на тренувальному наборі даних. Ваги нейронів оновлюються так, щоб мінімізувати функцію втрат.

4. Оцінка результатів: Після завершення тренування проводиться оцінка результатів на тестовому наборі даних. Використовуються різні метрики, такі як точність класифікації та матриця плутанини, для об'єктивного визначення ефективності нейронної мережі.



Рисунок 2 – Виконання запиту

Висновок

У даній курсовій роботі було розглянуто застосування нейронних мереж для розпізнавання рукописних цифр, проведено глибокий аналіз та практичну реалізацію нейронних мереж для розпізнавання рукописних цифр, зосереджуючись на використанні набору даних MNIST та звичайної архітектури MLP. Отримані результати та виявлені в ході експериментів особливості мають значущий вплив на розвиток та впровадження систем розпізнавання образів у різноманітних областях застосування.

Список джерел

1. Шуклін Д. Є. Моделі семантичних нейронних мереж та їх застосування в системах штучного інтелекту: 05.13.23.. Дис.канд. техн. наук. Харків, 2003. 196 с.
2. Schmidhuber, J. Deep Learning in Neural Networks: An Overview // Neural Networks. – 2015. – Vol. 61. – P. 85–117.