

УДК 004.8

DOI 10.30837/bi.2021.2(97).05

В.М. Горбенко¹, К.Г. Онищенко², І.В.Афанасьєва³, Р.В. Каменєв⁴¹Студент, кафедра програмної інженерії,Харківський національний університет радіоелектроніки,
vladyslav.horbenko@nure.ua, ORCID iD: 0000-0002-4988-3268²Асистент кафедри Програмної інженерії, Харківський національний університет радіоелектроніки,
kostiantyn.onyshchenko@nure.ua, ORCID iD: 0000-0002-7746-4570³Доцент кафедри програмної інженерії, кандидат технічних наук,
Харківський національний університет радіоелектроніки,
iryana.afanasieva@nure.ua, ORCID iD: 0000-0003-4061-0332⁴Студент, кафедра програмної інженерії, Харківський національний університет радіоелектроніки,
roman.kameniev@nure.ua, ORCID iD: 0000-0002-8263-8862

АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ТА МОДЕЛЕЙ ГЛИБИННОГО НАВЧАННЯ В ЗАДАЧАХ ОБРОБКИ ПРИРОДНОЇ МОВИ

Розглянуто підходи та методи глибокого навчання природної мови. Глибокі нейронні мережі створюються як мережі прямого поширення, але дослідження дуже успішно застосували рекурентні нейронні мережі до таких задач, як моделювання мов. Згорткові глибокі нейронні мережі застосовуються в комп'ютерному зорі та природній мові. Згорткові глибокі нейронні мережі також було застосовано до акустичного моделювання для автоматичного розпізнавання мовлення. Гнучка нейронна мережа, дані функції використовуються для описання прямокутних об'єктів у вигляді двійкової маски в природній мові, а також для обрахування масштабної регресії для покращення точності визначеного положення. Багатозмінна логістична функція, що використовуються для описання прямокутних об'єктів у вигляді двійкової маски, а також для обрахування масштабної регресії для покращення точності визначеного положення. Глибока мережа переконань, що є одним із типів глибоких нейронних мереж, що складений з декількох шарів прихованих змінних, що мають з'єднання, але не між вузлами всередині кожного шару. Нейронні мережі зберігання та вибірки великої пам'яті, її особливістю є те що вона є швидкою нейронною мережею з багатьма шарами які можуть використовуватись одночасно. Глибоке машинне навчання. Складені автокодувальники, ціль автокодувальника, що продиктована поняттям «доброго представлення». Розглянуто існуючі моделі та методи для вирішення поставлених задач. Було проаналізовано різні методи та проведено порівняння, позитивних та негативних факторів вибору того чи іншого методу. В кінці проведені висновки, в яких було підсумковано як позитивні так і негативні сторони оглянутих методів в розрізі застосування в природній мові.

ГЛИБИННЕ НАВЧАННЯ, ПРИРОДНОЇ МОВИ, ІНСТРУМЕНТИ МАШИННОГО НАВЧАННЯ,
МАШИННЕ НАВЧАННЯ, НАВЧАННЯ З УЧИТЕЛЕМ, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ

Горбенко В.Н., Онищенко К.Г., Афанасьєва І.В., Каменєв Р.В. Анализ существующих методов и моделей глубокого обучения в задачах обработки природного языка. Рассмотрены подходы и методы глубокого обучения естественному языку. Глубокие нейронные сети создаются как сети прямого распространения, но исследования очень успешно применили рекуррентные нейронные сети к таким задачам, как моделирование языков. Сверточные глубокие нейронные сети используются в компьютерном зрении и естественной речи. Сверточные глубокие нейронные сети также были применены к акустическому моделированию для автоматического распознавания речи. Гибкая нейронная сеть, данные функции используются для описания прямоугольных объектов в виде двоичной маски в естественной речи, а также для вычисления масштабной регрессии для улучшения точности определенного положения. Многосменная логистическая функция, используемая для описания прямоугольных объектов в виде двоичной маски, а также для вычисления масштабной регрессии для улучшения точности определенного положения. Глубокая сеть убеждений является одним из типов глубоких нейронных сетей, составленным из нескольких слоев скрытых переменных, имеющих соединения, но не между узлами внутри каждого слоя. Нейронные сети хранения и выборки большой памяти, ее особенностью является то, что она является быстрой нейронной сетью со многими слоями, которые могут использоваться одновременно. Глубокое машинное обучение. Составлены автокодировщики, цель автокодировщика, продиктованная понятием «хорошего представления». Рассмотрены существующие модели и методы решения поставленных задач. Были проанализированы различные методы и проведены сравнения, положительных и негативных факторов выбора того или иного метода. В конце проведены выводы, в которых были итигированы как положительные так и отрицательные стороны рассмотренных методов в разрезе применения в естественном языке.

ГЛИБИННЕ ОБУЧЕННЯ, ЕСТЕСТВЕННИЙ ЯЗЫК, ІНСТРУМЕНТИ МАШИННОГО ОБУЧЕННЯ,
МАШИННЕ ОБУЧЕННЯ, ОБУЧЕННЯ С УЧИТЕЛЕМ, СВЕРТОЧНІ НЕЙРОННІ МЕРЕЖІ

Vladyslav Horbenko, Kostiantyn Onyshchenko, Iryna Afanasieva, Roman Kameniev. Analysis of existing methods and models of in-depth learning in the problems of natural language processing. Approaches and methods of deep learning of natural language are considered. Deep neural networks are being created as direct propagation networks, but research

has very successfully applied recurrent neural networks to tasks such as language modeling. Convolutional deep neural networks are used in computer vision and natural language. Convolutional deep neural networks have also been applied to acoustic modeling for automatic speech recognition. Flexible neural network, these functions are used to describe rectangular objects in the form of a binary mask in natural language, as well as to calculate scale regression to improve the accuracy of a particular position. Multivariate logistic function used to describe rectangular objects in the form of a binary mask, as well as to calculate scale regression to improve the accuracy of a particular position. A deep network of beliefs, which is a type of deep neural network made up of several layers of hidden variables that have connections but not between nodes within each layer. Neural networks store and sample large memory, its feature is that it is a fast neural network with many layers that can be used simultaneously. Deep machine learning. Compound car encoders, the purpose of the car encoder, which is dictated by the concept of «good performance». The existing models and methods for solving the tasks are considered. Different methods were analyzed and a comparison of positive and negative factors of choosing a method was made. In the end, the conclusions were summarized, which summarized both the positive and negative aspects of the reviewed methods in terms of application in natural language.

IN-DEPTH LEARNING, NATURAL LANGUAGE, MACHINE LEARNING TOOLS, MACHINE LEARNING, TEACHER LEARNING, CONVOLVED NEURAL NETWORKS.

Вступ

В наш час, коли інформаційний простір не обмежується територіально і для її отримання не потрібно багато часу, ми маємо велику кількість не перевіреної інформації, що може містити пагубний контент на осіб, які психологічно слабкі до негативу, або особи, що знайомляться інформацією, є неповнолітніми. Тому потрібні інструменти які можуть опрацьовувати велику кількість текстової інформації та знаходити в ньому харасмент однієї особи до іншої, домагання та расизм.

Обробка природної мови (NLP – Natural Language Processing) – це підрозділ інформаційних технологій, штучного інтелекту та лінгвістики, метою якого є вивчення проблем комп'ютерного аналізу та синтезу природної мови. Повне розуміння та відтворення сенсу мови – надзвичайно складне завдання, оскільки людська мова має цілий ряд особливостей.

Людська мова – це складний механізм, набутий в процесі еволюції, що використовує звук або писемність для передачі інформації від однієї людини до іншої або до групи людей. Мова це складна система слів, що у купі створює речення, а ті в свою чергу надають послідовну інформацію, яка сприймається опонентом і він може отримати інформацію, будь-якого характеру, це можуть бути знання, попередження, оповідь, будь-що.

Обробка природної мови (NLP – Natural Language Processing) – це підрозділ інформаційних технологій, штучного інтелекту та лінгвістики, метою якого є вивчення проблем комп'ютерного аналізу та синтезу природної мови. Повне розуміння та відтворення сенсу мови – надзвичайно складне завдання, оскільки людська мова має цілий ряд особливостей.

Технології машинного навчання все глибше потрапляють у наш життєвий побут і можуть використовуватись у таких продуктах як камери або смартфони. Машинне навчання використовується, щоб виділити об'єкти в зображенні, перетворити мовлення в текст або надати релевантні результати пошуку. Все частіше для вирішення даної задачі використовують таку галузь машинного навчання, як глибинне навчання. Воно стало на зміну звичайному машинному

навчанню, що потребувало створення складних систем, які б перетворювали сирі дані у представлення, що може бути оброблене системою. Отже, глибинне навчання – це набір методів, що використовує багаточисельні штучні нейронні мережі, щоб забезпечити найсучаснішу точність виконання завдань.

1. Постановка задачі

Необхідно розглянути методи глибинного навчання та їх використання в обробці природної мови, огляд позитивних та негативних сторін у порівнянні одного методу з іншим.

Різні архітектури глибинного навчання, такі як глибинні нейронні мережі, згорткові глибинні нейронні мережі, глибинні мережі переконань та рекурентні нейронні мережі застосовувалися в таких областях, як комп'ютерне бачення, автоматичне розпізнавання мовлення, обробка природної мови, розпізнавання звуків та біоінформатика, де вони, як було показано, представляють передові результати в різноманітних задачах.

В кінці зробимо висновки, в яких задачах вибрати той чи інший підхід, що він може нам запропонувати, та в яких ситуаціях може вирішити ту чи іншу задачу.

2. Огляд існуючих моделей та методів для вирішення поставленої задачі

Є дуже багато варіантів архітектури глибинного навчання. В основному, вони відгалужені від батьківських архітектур. Порівняння ефективності різної архітектури не завжди є можливим, адже не всі з них оцінювались на однакових вхідних наборах даних. Глибинне навчання на даний час, дуже швидко розвивається і нових варіантів або алгоритмів з'являються дуже швидко.

Глибинна нейронна мережа – це штучна нейронна мережа з декількома прихованими шарами вузлів між вхідним та вихідним шарами. Подібно до плоских штучних нейронних мереж, глибинні нейронні мережі можуть моделювати складні нелінійні відношення [6].

Архітектури глибинних нейронних мереж, наприклад, для виявлення об'єктів та граматичного аналізу,

породжують композиційні моделі, де об'єкт виражається як шарувата композиція примітивів зображення. Додаткові шари дозволяють композиції включати ознаки з нижчих шарів, забезпечуючи потенціал для моделювання складних даних меншою кількістю вузлів, ніж настільки ж ефективна пласка мережа.

Глибинні нейронні мережі створюються як мережі прямого поширення, але дослідження дуже успішно застосували рекурентні нейронні мережі, до таких задач, як моделювання мов. Згорткові глибинні нейронні мережі застосовуються в комп'ютерному зорі та природній мові. Згорткові глибинні нейронні мережі також було застосовано до акустичного моделювання для автоматичного розпізнавання мовлення.

Гнучка нейронна мережа, тренується за допомогою одного з стандартних алгоритму зворотного поширення, або іншими словами метод навчання багат шарового перцептронну. Основи безперервного зворотного поширення були створені в розрізі теорії керування Генрі Келлі та Артуром Брайсоном в 1960-1961 роках, принципи були застосовані в динамічному програмуванні. В 1962 році Стюарт Дрейфус розробив простіше рішення на основі ланцюгового правила, це було подано як багатоетапний метод оптимізації динамічної системи. Знаходження вагових коефіцієнтів відбувається стохастичним найшвидшим спуском, ітеративний метод оптимізації градієнтного спуску за допомогою стохастичного наближення, використовується таке рівняння [1].

$$w_{ij}(t+1) = w_{ij}(t) + \eta \frac{\partial C}{\partial w_{ij}},$$

де h – темп навчання, C – функція витрат. Вибір функції витрат буде залежить від того який тип навчання було використано: спонтанне – при вирішенні яких випробовувана система спонтанно навчається виконувати поставлене завдання, без втручання з боку експериментатора, з підкріплення, виконують програмні агенти в певному середовищі задля максимізації деякого уявлення про сукупну винагороду, кероване, в ході якого випробовувана система примусово навчається за допомогою наявної множини прикладів «стимул-реакція» з метою визначення «реакції» для «стимулів», які не належать до наявної множини прикладів, та функції активації – залежність вихідного сигналу штучного нейрона від вхідного.

Багатозмінна логістична функція, функція розраховується за допомогою рівняння,

$$p_j = \frac{\exp(x_j)}{\sum_k \exp(x_k)},$$

де p_j – є ймовірністю класу, а x_j та x_k – сумарний вхід до вузлів j та k на одному й тому ж рівні.

Перехресна ентропія розраховується за формулою

$$C = -\sum_j d_j \log(p_j).$$

В даній формулі d_j – це цільові ймовірності для вихідних вузлів j , p_j – вихідна ймовірність для j після активаційної функції.

Дані функції використовуються для описання прямокутних об'єктів у вигляді двійкової маски в природній мові, а також для обрахування масштабної регресії для покращення точності визначеного положення.

Як із глибинним нейронним навчанням так із штучним нейронним навчанням, в процесі роботи може виникати багато проблем, якщо їх тренувати наївно. Основними проблемами є перенавчання та тривалість обчислення.

Глибинна мережа переконань – є одним із типів глибинних нейронних мереж, що складений з декількох шарів прихованих змінних, що мають з'єднання, але не між вузлами всередині кожного шару.

В ході тренування на наборі прикладів спонтанним чином, даний метод може навчитися краще відбудовувати свої входи. Кожний шар являє собою детектор ознак, набір методик, що дозволяє системі автоматично виявляти представлення, необхідні для виявлення ознак, на входах. Глибинна мережа переконань розглядається як суміш простих та спонтанних мереж, де попередній прихований шар слугує явним шаром наступному. Це веде до швидкої пошарової процедури спонтанного тренування, в якій порівняльна розбіжність застосовується до кожної підмережі по черзі, починаючи з найнижчої пари шарів.

Глибинна мережа переконань працює наступним чином:

1) натренувати обмежену машину Больцмана (рис. 1) на X , щоби отримати матрицю її вагових коефіцієнтів, W . Використовувати її як матрицю вагових коефіцієнтів між двома нижніми шарами мережі;

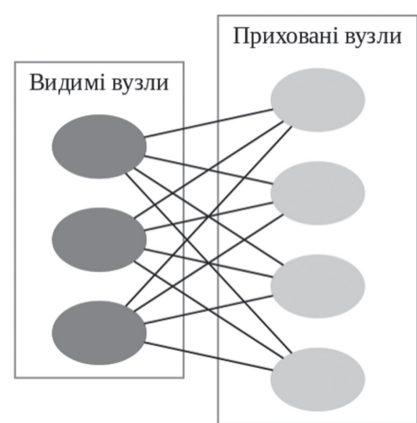


Рис. 1. Обмежена машина Больцмана

2) перетворити X за допомогою цієї ОМБ для отримання нових даних X' , або шляхом вибірки, або обчисленням середньої активації прихованих вузлів;

3) повторювати цю процедуру з $X \leftarrow X'$ для наступної пари шарів, поки не буде досягнуто двох верхніх шарів мережі;

4) здійснити тонке налаштування всіх параметрів цієї глибинної архітектури по відношенню до того,

що виконує роль логарифмічної правдоподібності ГМП, або по відношенню до критерію керованого тренування (після додавання додаткових механізмів навчання для перетворення навченого представлення на керовані передбачення, наприклад, лінійного класифікатора).

Як тільки обмежену машину Больцмана було натреновано, на неї накладається інша, беручи свої входи із завершального вже натренованого рівня. Значенням нового вхідного видимого шару встановлюється тренувальний вектор, а значення вузлів уже натренованих шарів встановлюються із застосуванням поточних вагових коефіцієнтів та зсувів. Потім нова обмежена машинна Больцмана тренується за наведеною вище процедурою. Весь цей процес повторюється до досягнення бажаного критерію зупинки.

Хоча наближення порівняльної розбіжності до максимальної правдоподібності і є дуже грубим, було показано, що порівняльна розбіжність не слідує градієнтові будь-якої функції, було емпірично показано, що вона є ефективною в тренуванні глибинних архітектур.

Одним із нових досягнень у глибинному навчанні є застосування згорткових глибинних мереж переконань, що мають структуру, дуже схожу до згорткових нейронних мереж, і тренуються схоже до глибинних мереж переконань. Тому, вони використовують двовимірну структуру зображень, так само як це роблять згорткові глибинні мережі переконань, і використовують попереднє тренування, як глибинні мережі переконань. Загальна структура, що може застосовуватися в багатьох задачах обробки зображень та сигналів. за останній час, було багато результатів на стандартних наборах зображень, таких як CIFAR, було отримано із застосуванням згорткової глибинної мережі переконань.

Нейронні мережі зберігання та вибірки великої пам'яті – є одним із видів глибинного навчання, її особливістю є те що вона є швидкою нейронною мережею з багатьма шарами які можуть використовуватись одночасно. Фільтри можуть бути нелінійними, логічними та не стаціонарними і мають неперервне навчання.

Данна нейронна мережа слугує динамічною нейронною мережею в просторово-часовій області визначення та в обох. Дана швидкість забезпечується геббовими (це теорія в нейронауці, яка пропонує пояснення пристосування нейронів мозку під час процесу навчання, описуючи основний механізм синаптичної пластичності, в якому підвищення синаптичної дієвості виникає в результаті повторюваного й постійного стимулювання пресинаптичною клітиною постсинаптичної) ваговими коефіцієнтами з'єднань, що об'єднує між собою не схожі по своїй суті фільтрів, функції попередньої обробки, в її багатьох шарах, для динамічного ранжування ваги різних шарів та функцій по відношенню до заданої задачі глибинного навчання. Вона повністю прозораю,

це досягається завдяки своїм ваговим коефіцієнтам з'єднань.

Метод нейронних мереж зберігання та вибірки великої пам'яті, застосовували в медичних та фінансових прогнозах та адаптивному фільтруванні зашумленого мовлення в невідомому шумі, розпізнаванні нерухомих зображень, безпеці програмного забезпечення, адаптивному керуванні нелінійними системами та ін. У порівняльному дослідженні з розпізнавання символів, дана методика, мала значно вищу швидкість обрахування і меншу похибку, в порівнянні із згортковою нейронною мережею на основі фільтрування функціями ReLU та максимальною підвибіркою.

Застосунки на базі даного методу показують приховані аспекти даних, що є прихованими від мереж поверхневого навчання, та навіть людської інтуїції, що покладається на очі та слух.

Метод нейронних мереж зберігання та вибірки великої пам'яті, був запропонований в 1996 році і вдосконалено Даніелем Граупе та Губертом Кордилевським.

Глибинне машинне навчання – це один із видів двійкового парного марковського випадкового поля з шарами латентних випадкових змінних [6]. Що є мережею симетрично спарованих випадкових двійкових вузлів. В даному випадку складається з набору видимих вузлів $v \in \{0, 1\}^D$ та ряду латентних вузлів $h^{(1)} \in \{0, 1\}^{F_1}$, $h^{(2)} \in \{0, 1\}^{F_2}$, ..., $h^{(L)} \in \{0, 1\}^{F_L}$. Зав'язків між однорівневими вузлами відсутні. Ймовірність глибинності машини Больцмана розраховується за формулою.

$$p(v) = \frac{1}{Z} \sum_h e^{\sum_{ij} W_{ij}^{(1)} v_i h_j^{(1)} + \sum_{jl} W_{jl}^{(2)} h_j^{(1)} h_l^{(2)} + \sum_{lm} W_{lm}^{(3)} h_l^{(2)} h_m^{(3)}}.$$

Глибинній мережі переконань верхні два шари створюють обмежену машину Больцмана, ті що знизу шари створюють орієнтовану створюючу модель. Зв'язки є симетричними, так як вона є графовою моделлю, ймовірнісна модель, для якої умовні залежності між випадковими змінними виражено графом, слої що знаходяться на знизу створюють орієнтовану створюючу модель.

Що глибинні мережі переконань так глибинна машина Больцмана, вони можуть отримати навчання складних та абстрактних внутрішніх представлень входу в ситуаціях при використанні розпізнавання об'єктів та мовлення, застосовуючи обмежені мічені дані для настроювання тонкого представлення, зробленого з великого представлення немічених сенсорних даних на вхід [2].

Швидкість роботи глибинної машини Больцмана, зменшує її продуктивність та функціональність. Для даної методики є непіддатливим навчання максимальної точної правдоподібності, тому виконується навчання наближеної максимальної правдоподібності. Також можна використати осереднене поле для

оцінювання залежних від даних очікувань, наближено очікуваної достатньої моделі Монте-Карло марковських ланцюгів (рис. 2), це клас алгоритмів для вибірки з розподілу ймовірностей на базі побудови такого ланцюга Маркова, що має бажаний розподіл як свій рівноважний розподіл

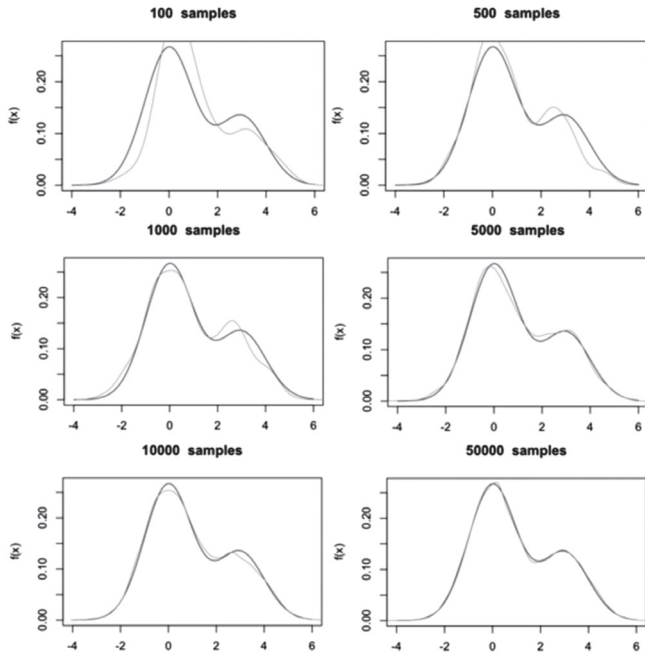


Рис. 2. Приклад Монте-Карло марковських ланцюгів

Воно є наближеними, що мусить бути здійснено для кожного перевіреного входу, є у діапазоні 25-50 разів є повільним за єдиним висхідним проходом у глибокій машині переконань. Що робить спільну оптимізацію дуже напратчною у використанні для великих наборів даних, що в свою чергу дуже сильно обмежує застосування глибокої машини Больцмана у ситуаціях представлення ознак.

Складені автокодувальники, кодувальник – детермінське відображення f_0 , перетворює вектор що на вході x на приховане представлення y , $O = \{W, b\}$, W є ваговою атрицею, а b є вектором зсуву.

Декодувальник – повертає назад приховане представлення y на відтворений вхід z через g_0 [6].

Головна ціль авдокодувальника продиктована поняттям «доброго представлення». Як прикладом можна розглянути класифікатор добре представлення може бути признане, як ефективний класифікатор.

Складені знешумлені автокодувальники частково спотворений вихід очищується. Дану систему було розвернуто в 2010 році Венсаном, разом з особливим підходом до доброго представлення, добре представлення є таким, що може бути надійно отримано зі спотвореного входу, і буде корисним для відновлення відповідного чистого входу [3]. Неявними в цьому визначенні є наступні ідеї:

- представлення вищого рівня є відносно стабільними й стійкими до спотворень входу;

- необхідно виділяти ознаки, що є корисними для представлення розподілу входу.

Даний підхід складається з декількох кроків, спочатку імовірне відображення, що є спотворювальним кроком, потім автокодування та відображається на приховане представлення, останнім кроком виконується мінімізація.

Щоб отримати глибоку архітектуру автокодування, їх накладають один на одного. Після знешумлення та навчання першої функції, можливо натренувати наступний рівень.

Після того як автокодувальник натреновано, його вихід може бути використано як «кероване навчання», (логістична регресія або метод опорних векторів).

Змішані ієрархічно-глибинні моделі komponують глибокі мережі з непараметричних баєсових моделей. Архітектуру які можна використати для навчання є глибокі машини баєса, глибокі автокодувальники та згорткові варіанти. Що забезпечує краще представлення, та швидке навчання, точну класифікацію із даними високої розмірності. Ці архітектури є слабкими в навчанні нових класів на декількох прикладах, так-як всі вузли мережі додані до представлення входу, і повинні бути скориговані разом [5].

Обмеження ступеню свободи знижує кількість параметрів для навчання, допомагаючи навчанню нових класів з кількох прикладів. Ієрархічні баєсові моделі забезпечують навчання з кількох прикладів, наприклад, для комп'ютерного бачення, статистики та когнітивної науки.

В своїй основній ідеї змішані ІГ-архітектури komponують характеристики глибоких мереж та ієрахії баєса. Всі рівні отримують навчання разом, використовуючи максимізації функції внеску, що показує наскільки чутливі функції правдоподібності, залежить від свого параметру.

3. Аналіз та порівняння

Є кілька способів, якими характеризувалася галузь глибокого навчання. Наприклад, 1986 року Ріна Дехтер ввела поняття глибокого навчання першого порядку та глибокого навчання другого порядку в контексті задоволення обмежень. Пізніше глибоке навчання було охарактеризовано як клас алгоритмів машинного навчання, які використовують каскад багатьох шарів вузлів нелінійної обробки для виділення ознак та перетворення. Кожен наступний шар використовує вихід із попереднього шару як вхід. Алгоритми можуть бути з керованим або спонтанним навчанням, а застосування включають розпізнавання образів та класифікацію.

Вони ґрунтуються на навчанні декількох шарів ознак або представлень даних. Ознаки вищих рівнів виводяться з ознак нижчих рівнів для формування ієрархічного представлення.

Глибинне навчання є частиною ширшої області машинного навчання з навчання представлень даних навчаються кільком рівням представлень, що відповідають різним рівням абстракції; ці рівні формують ієрархію понять.

Ці визначення мають спільними декілька шарів вузлів нелінійної обробки та кероване або спонтанне навчання представлень ознак у кожному з шарів, з формуванням шарами ієрархії від низькорівневих до високорівневих ознак. Побудова шару вузлів нелінійної обробки, що застосовується в алгоритмі глибинного навчання, залежить від розв'язуваної задачі.

Шари, що застосовувалися в глибинному навчанні, включають приховані шари штучної нейронної мережі та набори складних висловлень. Вони також можуть включати латентні змінні, організовані по шарах у глибинних породжувальних моделях, такі як вузли в глибинних мережах переконань та глибинних машинах Больцмана.

Алгоритми глибинного навчання перетворюють свої входи крізь більшу кількість шарів, ніж алгоритми поверхневого навчання. На кожному шарі сигнал перетворюється блоком обробки, таким як штучний нейрон, параметри якого «навчаються» шляхом тренування. Ланцюг перетворень від входу до виходу є шляхом передачі довіри.

Описуючи потенційно причинні зв'язки між входом та виходом, і можуть мати змінну довжину. Для нейронної мережі прямого поширення довжина шляхів передачі довіри, і відтак глибина цієї мережі, є числом прихованих шарів плюс один. Для рекурентних нейронних мереж, в яких сигнал може поширюватися через якийсь шар більше одного разу, має потенційно необмежену довжину. Універсально узгодженого порогу глибини, що відділяв би поверхневе навчання від глибинного, не існує, але більшість дослідників у цій галузі погоджуються, що глибинне навчання має декілька нелінійних шарів (ШПД > 2), а Шмідгубер розглядає ШПД > 10 як дуже глибинне навчання.

Висновок

Розглядаючи різні методи та підходи глибинного навчання, можна зробити висновок в кожному з них є свої переваги та недоліки, це може бути можливість праці над великими об'ємами даних, та бути дуже повільними, або навпаки бути в роботі дуже швидкими в навчанні, але працювати з невеликими об'ємами даними.

В наш час коли передача природної мови є дуже швидким, основним методом якого є Інтернет, це може бути спілкування за допомогою додатків які забезпечують передачу мови від одної особи до іншої або групи осіб. Спілкування за допомогою набору коротких текстових повідомлень від однієї особи до іншої або групи людей.

Також може бути прослуховування або ознайомлення інформації на різних інтернет ресурсах.

Тому маючи велику кількість оточеної нас природної мови, потрібні інструменти у вигляді натренованих нейронних мереж, що можуть забезпечити обробку мови на виявлення неправильної поведінки однієї особи до іншої, харасмент, що може відбиватися на підставі сексуального характеру, расового та інших ознак.

Можливий один із варіантів розшифрування закодованої мови, або ж можливо навпаки закодувати, можливо використати метод кодування декодування мережі, де можливо навчити систему та використовувати дані інструменти для роботи, в тих чи інших цілях.

Опрацювання природної мови, може також нести вивчення тієї чи іншої ситуації для отримання інформації, для цього можливо застосувати один із методів глибинного навчання, за допомогою якого можна провести навчити нейронної мережі та застосувати для вирішення.

Моделі рекурсивного глибокого навчання можуть вирішувати багато мовних задач, що включають в себе передбачення на рівні слова та речення як безперервного, так і дискретного характеру. Одним із безлічі існуючих рішень для вирішення цієї задачі є використання рекурсивних або рекурентних методів для розрахунку подань рівня абзацу або документа.

Іншою реалізацією роботи для глибинних моделей – це логічні міркування першого порядку, які можуть бути використані для отримання правильної інформації з баз знань за допомогою питань природної мови. Переглянуті моделі в цій статті можуть бути розширені, щоб спільно моделювати мову, образи та бази знань в одній цілісній семантичній структурі.

Список літератури:

- [1] *J. Le.* (2018). The 7 NLP Techniques That Will Change How You Communicate in the Future (Part I) | [Online]. Available: <https://heartbeat.fritz.ai/the-7-nlp-techniques-that-will-change-how-you-communicate-in-the-future-part-i-f0114b2f0497>
- [2] *M. Bates* (1995). Models of natural language understanding | [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC40721/>
- [3] *R. Socher*, Recursive deep learning for natural language processing and computer vision | Stanford University, 2014, pp. 8-120.
- [4] *Afanasyeva I.* Data exchange model in the Internet of Things concept / I. Afanasyeva, N. Golian, O. Hnatenko, Y. Daniil, K. Onyshchenko // Telecommunications and Radio Engineering, New York, 2019. – 10(78). – p. 869-878
- [5] *Onyshchenko A.* Adaptive method of training neural networks / A. Onyshchenko, K. Onyshchenko // Technique and technology. Science, Research, Development #29. Gdansk, 2020. – p. 9-11.
- [6] Web-site - https://en.wikipedia.org/wiki/Deep_learning/

Надійшла до редколегії 28.10.2021