

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

Факультет комп'ютерно-інформаційних систем і програмної інженерії

(повна назва факультету)

Кафедра комп'ютерних наук

(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

бакалавр

(назва освітнього ступеня)

на тему: **Розробка інформаційної системи з використанням методів машинного навчання в задачах для підбору цільової аудиторії інтернет магазинів.**

Виконав: студент **IV** курсу, групи **СН-41**

спеціальності **122 Комп'ютерні науки**

(шифр і назва спеціальності)

(підпис)

Матвієнко Т.В.

(прізвище та ініціали)

Керівник

(підпис)

Литвиненко Я.В.

(прізвище та ініціали)

Нормоконтроль

(підпис)

Шимчук Г.В.

(прізвище та ініціали)

Завідувач кафедри

(підпис)

Боднарчук І.О.

(прізвище та ініціали)

Рецензент

(підпис)

Пастух О.А.

(прізвище та ініціали)

Тернопіль
2021

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

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

Кафедра комп'ютерних наук
(повна назва кафедри)

ЗАТВЕРДЖУЮ
 Завідувач кафедри
 _____ Боднарчук І.О.
(підпис) (прізвище та ініціали)
 «__» _____ 2021 р.

**ЗАВДАННЯ
 НА КВАЛІФІКАЦІЙНУ РОБОТУ**

на здобуття освітнього ступеня _____ Бакалавр
(назва освітнього ступеня)

за спеціальністю _____ 122 Комп'ютерні науки
(шифр і назва спеціальності)

Студенту _____ Матвієнко Тарас Володимирович
(прізвище, ім'я, по батькові)

1. Тема роботи Розробка інформаційної системи з використанням методів машинного навчання в задачах для підбору цільової аудиторії інтернет магазинів.

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

Затверджені наказом ректора від «__» _____ 2021 року № _____

2. Термін подання студентом завершеної роботи _____ 2021р.

3. Вихідні дані до роботи _____

4. Зміст роботи (перелік питань, які потрібно розробити)
 Вступ.

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

АНОТАЦІЯ

Розробка інформаційної системи з використанням методів машинного навчання в задачах для підбору цільової аудиторії інтернет магазинів // Кваліфікаційна робота освітнього рівня «Бакалавр» // Матвієнко Тарас Володимирович// Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем і програмної інженерії, кафедра комп'ютерних наук, група СН-41 // Тернопіль, 2021 // С. , рис. _ , табл. _ , кресл. _ , додат. _ , бібліогр. _ .

Ключові слова: машинне навчання, штучний інтелект, браузер, моделі машинного навчання.

Кваліфікаційна робота присвячена розробці інформаційної системи для підбору цільової аудиторії на основі моделі машинного навчання.

Мета роботи: демонстрація головних особливостей машинного навчання у браузері використовуючи мову програмування JavaScript.

В першому розділі кваліфікаційної роботи розглянуто історію машинного навчання, основні методи машинного навчання, та способи їх застосування.

В другому розділі кваліфікаційної роботи розглянуто спосіб використання машинного навчання, розроблено застосунок для підбору цільової аудиторії інтернет магазину.

ANNOTATION

Development of information system using machine learning methods in tasks for selection of target audience of online stores // Qualification work of educational level "Bachelor" // Matvienko Taras Vladimirovich // Ternopil National Technical University named after Ivan Pulyuy, Faculty of Computer Information Systems and Software Engineering , Department of Computer Science, group SN-41 // Ternopil, 2021 // C., fig. _, table. _, chair. _, added. _, bibliogr. _.

Keywords: machine learning, artificial intelligence, browser, machine learning models.

Qualification work is devoted to the development of an information system for the selection of the target audience based on the model of machine learning.

Purpose: to demonstrate the main features of machine learning in a browser using the JavaScript programming language.

The first section of the qualification work considers the history of machine learning, the main methods of machine learning, and methods of their application.

In the second section of the qualification work the method of using machine learning is considered, the application for selection of target aduty of online store is developed.

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ПЗ програмне забезпечення

UI (User Interface) інтерфейс користувача

ШІ штучний інтелект

МН машинне навчання

API (Application Programming Interface) набір визначень взаємодії
різнотипного програмного забезпечення

ЗМІСТ

ВСТУП.....	7
РОЗДІЛ 1. ШТУЧНИЙ ІНТЕЛЕКТ ТА ЙОГО РОЛЬ У СВІТІ.....	8
1.1. Історія штучного інтелекту.....	8
1.2. Застосування машинного навчання	12
1.3. Методи машинного навчання.....	17
1.4. Цільова Аудиторія в інтернет магазинах.....	24
1.5. Висновок до першого розділу.....	26
РОЗДІЛ 2. ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ В ЗАДАЧАХ ДЛЯ ПІДБОРУ ЦІЛЬОВОЇ АУДИТОРІЇ ІНТЕРНЕТ МАГАЗИНІВ.....	27
2.1. Проектування навчальної моделі.....	27
2.2. Підготовка даних до навчання	33
2.3. Підготовка моделі до навчання та прогнозування.....	35
2.4 Висновки до другого розділу.....	39
РОЗДІЛ 3. БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОХОРОНА ПРАЦІ.....	41
3.1 Психологічні причини нещасних випадків і травматизму.....	41
3.2 Санітарно-гігієнічні вимоги до умов праці.....	44
ВИСНОВКИ	48
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	49

ВСТУП

Глобалізація світової економіки та формування соціальноорієнтованого середовища вимагають значних обчислювальних і телекомунікаційних ресурсів, систем накопичення, оброблення, аналізу і збереження великих обсягів даних. З цією метою мережі провайдерів, корпорацій, комунальні та державні ІТ-системи інтегрувалися в єдине розподілене обчислювальне середовище, ІТ-інфраструктура та сервіси якого утворили інфокомунікаційні системи різного призначення.

При цьому особливої актуальності набуває задача інформаційного аналізу і синтезу систем керування плануванням завдань та розподілом ресурсів в розподіленому обчислювальному середовищі з метою забезпечення енергозбереження при високій якості обслуговування користувачів інфокомунікаційних сервісів.

Основним напрямком розв'язання цієї важливої науково-практичної задачі є застосування ідей і методів машинного навчання та розпізнавання образів.

Але відомі методи інтелектуального аналізу даних характеризуються невисокою функціональною ефективністю машинного навчання через науковометодологічні ускладнення, пов'язані із довільними початковими умовами формування вхідних навчальних матриць, багато вимірністю словника ознак і алфавіту класів розпізнавання, суттєвим апріорним перетином класів розпізнавання, впливом неконтрольованих збурюючих факторів тощо.

Основна перевага методів, створених в рамках ІЄІ-технології, полягає в тому, що вони ґрунтуються на максимізації в процесі машинного навчання інформаційної спроможності системи, яка визначає її функціональну ефективність.

РОЗДІЛ 1. ШТУЧНИЙ ІНТЕЛЕКТ ТА ЙОГО РОЛЬ У СВІТІ

1.1 Історія Штучного Інтелекту

Ідея, надія та мрія мати розумні машини з'явилися разом із винаходом сучасних комп'ютерів у 20 столітті.

Перший випадок нейронних мереж був у 1943 році, коли нейрофізіолог Уоррен Мак-Каллок та математик Уолтер Піттс написали статтю про нейрони та їх роботу. Вони вирішили створити модель за допомогою електричної схеми, і так нейронна мережа народилася.

У 1950 році Алан Тьюрінг створив всесвітньо відомий тест Тьюрінга. Цей тест досить простий - щоб комп'ютер пройшов, він повинен переконати людину, що це людина, а не комп'ютер.

У 1952 році з'явилася перша комп'ютерна програма, яка могла вчитися під час роботи. Це була гра, в яку грали в шашки, створену Артуром Самуелем.

Френк Розенблатт спроектував першу штучну нейронну мережу в 1958 році під назвою Персептрон. Головною метою цього було розпізнавання зразків і фігур.

Ще один надзвичайно ранній приклад нейронної мережі був у 1959 році, коли Бернард Удроу та Марсіан Хофф створили дві їх моделі в Стенфордському університеті. Перший називався ADELINЕ, і він міг виявляти двійкові закономірності. Наприклад, у потоці бітів він може передбачити, яким буде наступний. Наступне покоління називалося MADELINE, і воно могло усунути луну на телефонних лініях, тому було корисним. Використовується і сьогодні.

Незважаючи на успіх MADELINE, до кінця 1970-х років з багатьох причин не було великого прогресу,[30] головним чином за популярністю архітектури фон Неймана. Це архітектура, де інструкції та дані зберігаються в одній пам'яті, яку, можливо, простіше зрозуміти, ніж нейронна мережа, і так багато людей будували програми на основі цього.

Термін штучний інтелект виник в 1956 році, коли Джон Маккарті запропонував його. Першу програму для штучного інтелекту, теоретик логіки, продемонстрували Герберт Саймон та інші. Після цього програмісти почали розробляти програмне забезпечення, яке могло блискуче виконувати міркування. Були великі надії досягти справжнього людського інтелекту на комп'ютері; проте це не вдалося. Природний інтелект та варіації мовних навичок виявилося важким для наслідування. Програмістам було дуже важко дублювати людський інтелект, і тому обмежилися комп'ютерними програмами з великою інформацією (експертними системами) у спеціалізованих областях. До таких комп'ютерів належали ті, які могли проводити медичну діагностику. Вони включали MYCIN та Internist. Найдосконаліший штучний інтелект був реалізований в 1997 році, коли IBM Deep Blue, як зображено на рисунку 1.1 вдалося перевершити найкращого шахіста у світі.



Рисунок 1.1 - «Шахматний суперкомп'ютер «Deep Blue»

З початку 21 століття багато підприємств зрозуміли, що машинне навчання збільшить потенціал для розрахунку. Ось чому вони досліджують в ньому більше, щоб випередити конкуренцію.

Деякі великі проекти включають:

GoogleBrain (2012) - це глибока нейронна мережа, створена Джеффом Діном з Google, яка зосереджувалась на виявленні шаблонів у зображеннях та відео. Він зміг використовувати ресурси Google, що зробило його незрівняним із значно меншими нейронними мережами. Пізніше його використовували для виявлення об'єктів у відео YouTube.

AlexNet (2012) - AlexNet з великим відривом переміг у конкурсі ImageNet у 2012 році, що призвело до використання графічних процесорів та згорткових нейронних мереж у машинному навчанні. Вони також створили ReLU, який є функцією активації, що значно покращує ефективність CNN.

DeepFace (2014) - це глибока нейронна мережа, створена Facebook, яка, за їхніми твердженнями, може розпізнавати людей з такою ж точністю, як і людська.

DeepMind (2014) - цю компанію придбав Google і може грати в основні відеоігри на тих самих рівнях, що і люди. У 2016 році йому вдалося перемогти професіонала в грі Go, яка вважається однією з найскладніших настільних ігор у світі.

OpenAI (2015) - це некомерційна організація, створена Ілоном Маск та іншими, для створення безпечного штучного інтелекту, який може принести користь людству.

Платформа машинного навчання Amazon (2015) - це частина Amazon Web Services і показує, як більшість великих компаній хочуть брати участь у машинному навчанні. Вони кажуть, що це керує багатьма їх внутрішніми системами - від регулярно використовуваних сервісів, таких як рекомендації щодо пошуку та Alexa, до більш експериментальних, таких як Prime Air та Amazon Go.

ResNet (2015) - це було значним прогресом у CNN, і більше інформації можна знайти на сторінці Вступ до CNN.

U-net (2015) - це архітектура CNN, що спеціалізується на біомедичній сегментації зображень. Він ввів однакову кількість шарів підвищення та

зменшення вибірки, а також пропустив з'єднання. Більше інформації про те, що це означає, можна знайти на сторінці Семантична сегментація.

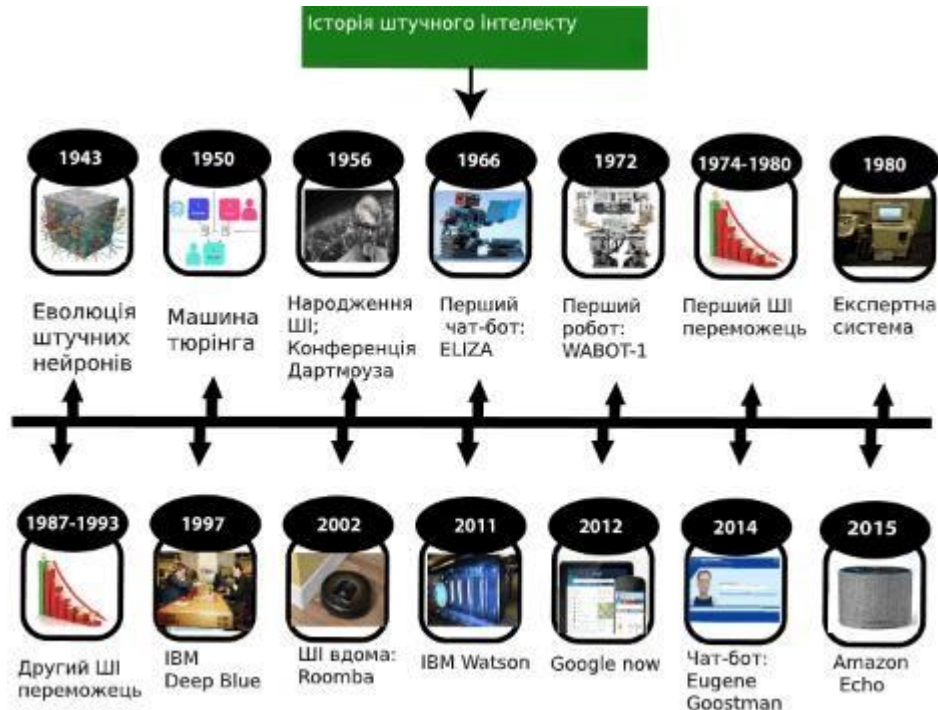


Рисунок 1.2 – Історія штучного інтелекту

Також є цікаві думки на рахунок майбутнього машинного навчання:

Удосконалення алгоритмів навчання без нагляду:

У майбутньому ми побачимо більше зусиль, спрямованих на вдосконалення некерованих алгоритмів машинного навчання, які допоможуть робити прогнози на основі немаркованих наборів даних. Ця функція стане все більш важливою, оскільки дозволяє алгоритмам виявляти цікаві приховані закономірності або групування в наборах даних та допомагати компаніям краще розуміти свій ринок або клієнтів.

Зростання квантових обчислень:

Одне з основних застосувань тенденцій машинного навчання полягає в квантових обчисленнях, які можуть перетворити майбутнє цієї галузі. Квантові комп'ютери призводять до швидшої обробки даних, покращуючи здатність алгоритму аналізувати та отримувати значущі уявлення з наборів даних.

Зосередьтеся на когнітивних послугах: Програмні додатки стануть більш інтерактивними та інтелектуальними завдяки когнітивним послугам, керованим машинним навчанням. Такі функції, як розпізнавання зору, розпізнавання мови та розуміння мови, будуть простішими у впровадженні. Ми побачимо, що на ринку з'являться інтелектуальні програми, що використовують когнітивні послуги.

1.2 Застосування машинного навчання

Машинне навчання - це програма штучного інтелекту (ШІ), що надає системам можливість автоматичного навчання та вдосконалення на основі досвіду без явного програмування. Машинне навчання фокусується на розробці комп'ютерних програм, які можуть отримувати доступ до даних та використовувати їх для самостійного навчання.

Безсумнівно, машинне навчання є одним із найбільш захоплюючих підмножин штучного інтелекту. Він виконує завдання вивчення даних із конкретними входами до машини. Важливо зрозуміти, що змушує працювати машинне навчання, а отже, як його можна використовувати в майбутньому.

Процес машинного навчання починається з введення навчальних даних у вибраний алгоритм. Навчальні дані - відомі або невідомі дані для розробки остаточного алгоритму машинного навчання. Тип введення навчальних даних впливає на алгоритм, і ця концепція буде розглянута далі.

Щоб перевірити, чи правильно працює цей алгоритм, нові вхідні дані подаються в алгоритм машинного навчання. Потім перевіряються прогноз та результати.

Якщо прогнозування не відповідає очікуваному, алгоритм повторно тренується кілька разів, поки не буде знайдено бажаний результат. Це дозволяє алгоритму машинного навчання постійно вчитися самостійно і давати найбільш оптимальну відповідь, яка з часом поступово збільшується в точності.

Зазвичай результати програм машинного навчання які ми бачимо включають результати веб-пошуку, оголошення в реальному часі на веб-сторінках та мобільних пристроях, фільтрацію спаму електронною поштою, виявлення вторгнень у мережу та розпізнавання малюнків та зображень. Все це побічні продукти використання машинного навчання для аналізу великих обсягів даних.

Традиційно аналіз даних базувався на спробах та помилках, що стає неможливим, коли масиви даних великі та неоднорідні. Машинне навчання пропонує розумні альтернативи аналізу великих обсягів даних. Розробляючи швидкі та ефективні алгоритми та керовані даними моделі для обробки даних у реальному часі, машинне навчання може дати точні результати та аналіз.

Згідно з відповідним звітом Мак-Кінсі, "Оскільки все більше аналогового світу оцифровується, наша здатність вчитися на даних шляхом розробки та тестування алгоритмів стане лише важливішою для того, що зараз розглядається як традиційний бізнес". Також цитується головний економіст Google Хел Варіан, який називає це "комп'ютерним кайдзен" і додає, "як масове виробництво змінило спосіб збирання продуктів, а постійне вдосконалення змінило спосіб виробництва ... настільки постійні (і часто автоматичні) експерименти покращити спосіб оптимізації бізнес-процесів у наших організаціях".

Через нові обчислювальні технології машинне навчання сьогодні не схоже на машинне навчання минулого. Це народилося завдяки розпізнаванню образів та теорії, згідно з якою комп'ютери можуть вчитися, не будучи запрограмованими на виконання конкретних завдань; дослідники, які цікавляться штучним інтелектом, хотіли перевірити, чи можуть комп'ютери вчитися на даних. Ітераційний аспект машинного навчання важливий, оскільки, коли моделі потрапляють до нових даних, вони можуть самостійно адаптуватися. Вони вчаться на попередніх обчисленнях для отримання надійних, повторюваних рішень та результатів. Це наука не нова, але така, що набрала нових обертів.

Більшість галузей, що працюють з великими обсягами даних, визнали цінність технології машинного навчання. Отримуючи інформацію з цих даних - часто в реальному часі - організації можуть працювати ефективніше або отримати перевагу над конкурентами.

Банки та інші компанії у фінансовій галузі використовують технологію машинного навчання для двох ключових цілей: виявлення важливих даних у даних та запобігання шахрайству. Аналіз може визначити інвестиційні можливості або допомогти інвесторам знати, коли торгувати. Видобуток даних також може ідентифікувати клієнтів з профілями високого ризику або використовувати кібернагляд, щоб визначити попереджувальні ознаки шахрайства.

Державні установи, такі як громадська безпека та комунальні послуги, мають особливу потребу в машинному навчанні, оскільки вони мають безліч джерел даних, які можна видобути для отримання знань. Наприклад, аналізуючи дані датчиків, визначаються шляхи підвищення ефективності та економії грошей. Машинне навчання також може допомогти виявити шахрайство та мінімізувати крадіжки особистих даних.

Машинне навчання - це швидкозростаюча тенденція у галузі охорони здоров'я завдяки появі носимих пристроїв та датчиків, які можуть використовувати дані для оцінки стану здоров'я пацієнта в режимі реального часу. Ця технологія також може допомогти медичним експертам проаналізувати дані, щоб виявити тенденції або червоні позначки, які можуть призвести до поліпшення діагнозів та лікування.

Веб-сайти, що рекомендують товари, які можуть вам сподобатися на основі попередніх покупок, використовують машинне навчання для аналізу історії покупок. Роздрібні продавці покладаються на машинне навчання для збору даних, їх аналізу та використання для персоналізації досвіду покупок, реалізації маркетингової кампанії, оптимізації цін, планування постачання товарів та отримання інформації про клієнтів.

Пошук нових джерел енергії. Аналіз мінералів у землі. Прогнозування несправності датчика НПЗ. Впорядкування розподілу нафти для підвищення ефективності та економічності. Кількість випадків використання машинного навчання для цієї галузі величезна - і все ще зростає.

Аналіз даних для виявлення закономірностей та тенденцій є ключовим для транспортної галузі, яка покладається на підвищення ефективності маршрутів та прогнозування потенційних проблем для збільшення прибутковості. Аналіз даних та аспекти моделювання машинного навчання є важливими інструментами для компаній-постачальників, громадського транспорту та інших транспортних організацій.

Автоматизовані машини повинні мати можливість швидко і надійно реагувати на навколишнє середовище. Ці навички зміцнюються завдяки машинному навчанню. Однак програми ШІ не працюють належним чином як такі. Помилки у виборі відповідних навчальних даних, а також при генерації та обробці даних можуть призвести до небезпечних несправностей системи, які сам ШІ не може розпізнати та запобігти.

Особливо в критичних для безпеки програмах важливо, щоб система, в яку вбудований штучний інтелект, працювала абсолютно безпечно та надійно. Проблеми, пов'язані із забезпеченням машинного навчання на основі даних, сильно відрізняються від проблем із традиційно програмованим програмним забезпеченням. Використані навчальні дані відіграють важливу роль для якості генерованої нейронної мережі. Якщо дані не є репрезентативними для багатьох ситуацій, з якими система пізніше стикається, модель недостатньо хороша і приймає неправильні рішення[27]. Щоб модель також застосовувалась до даних, які не вивчені, модель повинна бути надійною та абстрактною. Отже, він не повинен бути надто близьким до навчальних даних, що призводить до переобладнання, а модель недостатньо абстрактна для нових даних. З іншого боку, не повинно бути недооборудованя тобто моделі, яка є занадто простою і яка недостатньо точно описує структуру даних.

Наприклад, точний аналіз даних та контроль процесів є важливими для автономного водіння. Транспортні засоби повинні мати можливість розпізнавати оточення, точно інтерпретувати їх, а потім оптимізувати свої дії.

На відміну від класичних алгоритмів, проблема машинно навчених програм полягає в тому, що окремі етапи навчання люди не можуть інтерпретувати. Завдяки автоматичному регулюванню зважувань у нейронних мережах, лише вхідні дані та результат залишаються доступними для людського контролю. Таким чином, мета дослідження під загальною фразою «Пояснюваний ШІ» полягає в тому, щоб зрозуміти нейромережі. Оскільки процес прийняття рішень ШІ є непрозорим, безпека та надійність ШІ не можуть бути оцінені без зайвих сумнівів.

Поки що машинне бачення (сприйняття) ШІ ще не було настільки надійним, щоб воно було придатним для критичного використання безпеки в автономних транспортних засобах на дорогах загального користування[18]. Перш за все, потрібно знайти способи зробити невизначеність ШІ кількісною, щоб мати можливість оцінити сприйняття значущим чином. Інститут когнітивних систем Фраунгофера IKS працює над вирішенням цих проблем та створенням надійно надійних систем шляхом кількісної оцінки невизначеності ШІ. Таким чином, раніше непрозора класифікація ШІ може стати керованою. Тому що лише якщо системи штучного інтелекту зрозумілі для людей, вони настільки безпечні, що їх можна використовувати, наприклад, для автономної їзди по дорогах загального користування.

Одним із підходів Fraunhofer IKS є доповнення ШІ розширеною архітектурою програмного забезпечення. Це контролює штучний інтелект і перевіряє прийняті рішення на ймовірність. У той же час динамічний підхід до управління безпекою надає ШІ більше свободи, ніж класичні підходи до безпеки, які завжди починаються з найгіршого сценарію. Таким чином можна використовувати переваги швидкої обробки даних за допомогою машинного навчання і одночасно вловлювати можливі неправильні рішення. Це особливо

важливо, якщо, як і при автономному керуванні автомобілем, неправильні рішення ШІ можуть поставити під загрозу життя людей.

Навіть при оцифровці промисловості, так званій Індустрії 4.0, машинний інтелект забезпечує оптимізоване планування та кращі прогнози. Автоматизовані та мережеві машини розпізнають своє оточення і можуть самостійно адаптувати свої дії до нього. Співпраця людина-робот без захисної огорожі можлива лише за допомогою машинного навчання та штучного інтелекту. Тут також ШІ повинен бути захищений, щоб не загрожувати життю людей.

Процеси машинного навчання також можуть бути використані у виробництві для оптимізації процесів, для визначення ймовірності відмов компонентів та для усунення несправностей. На мережевих промислових підприємствах несправність підсистеми може означати збій усього виробництва. Виправлення неполадок і простої трудомісткі та дорогі. Інститут когнітивних систем Фраунгофера IKS розробив інструмент аналізу програмного забезпечення DANA, щоб уникнути простоїв та полегшити пошук несправностей. За допомогою процесів машинного навчання він створює віртуальну модель виробництва. Це корисно при інтеграції нових частин системи та при прогнозуванні помилок. DANA виявляє та виправляє помилки, перш ніж система зупиниться. Таким чином можна запобігти виробничим збоям та захистити працівників виробництва від небезпечних несправностей.

1.3 Методи машинного навчання

Світ з кожним днем стає «розумнішим», і, щоб не відставати від сподівань споживачів, компанії все частіше використовують алгоритми машинного навчання, щоб полегшити ситуацію. Ви можете бачити, як вони використовуються на пристроях кінцевих користувачів (за допомогою розпізнавання обличчя для розблокування смартфонів) або для виявлення

шахрайства з кредитними картками (наприклад, спрацьовування сповіщень про незвичайні покупки).

В рамках штучного інтелекту (ШІ) та машинного навчання існує два основних підходи: навчання під контролем та навчання без нагляду, як зображено на рисунку 1.3 Основна відмінність полягає в тому, що одна використовує марковані дані для прогнозування результатів, а інша - ні. Однак є два нюанси між двома підходами та ключовими сферами, в яких один перевершує інший.

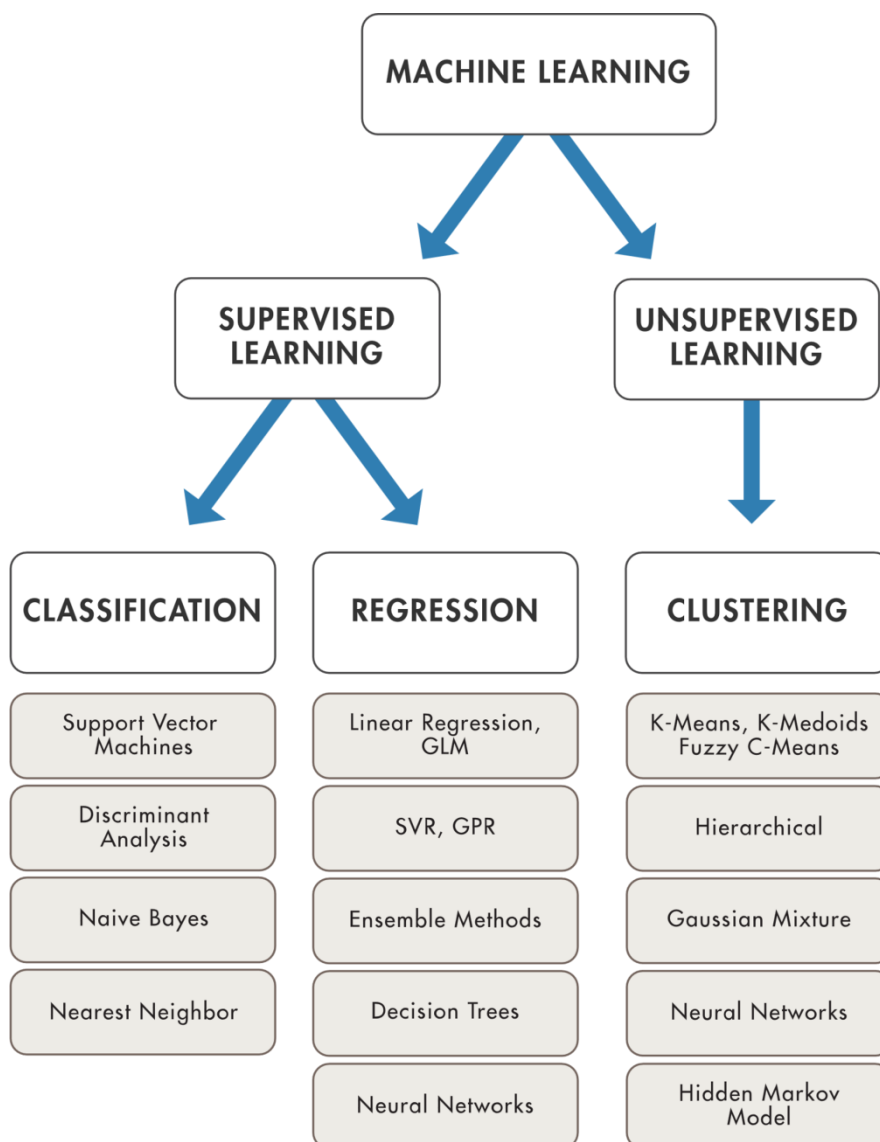


Рисунок 1.3 – Основні підходи ШІ

Контрольоване навчання - це підхід до машинного навчання, який визначається використанням маркованих наборів даних. Ці набори даних призначені для навчання або „нагляду” алгоритмів у класифікації даних або точному прогнозуванні результатів.[13] Використовуючи позначені входи та виходи, модель може виміряти свою точність і навчитися з часом.

Навчане під контролем можна розділити на два типи[25] проблем при аналізі даних: класифікація та регресія:

Завдання класифікації використовують алгоритм для точного віднесення даних тестів до конкретних категорій, наприклад, відокремлення яблук від апельсинів.[16] Або, в реальному світі, керовані алгоритми навчання можна використовувати для класифікації спаму в окремій папці з папки "Вхідні". Лінійні класифікатори, опорні векторні машини, дерева рішень та випадковий ліс - це загальні типи алгоритмів класифікації.

Регресія - це ще один тип контрольованого методу навчання, який використовує алгоритм для розуміння взаємозв'язку між залежними та незалежними змінними. Регресійні моделі корисні для прогнозування числових значень на основі різних точок даних, таких як прогнози доходу від продажу для даного бізнесу. Деякі популярні алгоритми регресії - це лінійна регресія, логістична регресія та поліноміальна регресія.

Навчання без нагляду використовує алгоритми машинного навчання для аналізу та кластеризації немічених наборів даних. Ці алгоритми виявляють приховані закономірності в даних без потреби втручання людини (отже, вони "не контролюються").

Моделі навчання без нагляду використовуються для трьох основних завдань: кластеризації, асоціації та зменшення розмірності:

Кластеризація - це техніка видобутку даних для групування немаркованих даних на основі їх подібності чи відмінності. Наприклад, алгоритми кластеризації K-засобів призначають подібні точки даних у групи, де значення K відображає розмір групування та деталізацію. Цей прийом корисний для сегментації ринку, стиснення зображень тощо.

Асоціація - це ще один тип неконтрольованого методу навчання, який використовує різні правила для пошуку зв'язків між змінними в даному наборі даних. Ці методи часто використовуються для аналізу ринкових кошиків та механізмів рекомендацій, відповідно до рекомендацій "Клієнти, які придбали цей товар".

Зменшення розмірності - це техніка навчання, яка використовується, коли кількість функцій (або розмірів) у даному наборі даних занадто велика. Це зменшує кількість введених даних до керованого розміру, одночасно зберігаючи цілісність даних. Часто цей[28] прийом використовується на етапі попередньої обробки даних, наприклад, коли автокодери видаляють шум із візуальних даних для поліпшення якості зображення.

Основна відмінність між двома підходами полягає у використанні маркованих наборів даних. Простіше кажучи, контрольоване навчання використовує позначені вхідні та вихідні дані, тоді як неконтрольований алгоритм навчання - ні.

При контрольованому навчанні алгоритм «навчається» із набору навчальних даних шляхом ітераційного прогнозування даних та коригування правильної відповіді. Хоча моделі навчання під наглядом, як правило, є більш точними, ніж моделі навчання без нагляду, вони потребують втручання людини заздалегідь, щоб належним чином позначити дані.[29] Наприклад, контрольована модель навчання може передбачити, як довго буде їхати ваша поїздка на дорогу, залежно від часу доби, погодних умов тощо. Але спочатку вам доведеться його тренувати, щоб знати, що дощова погода збільшує час водіння.

Моделі навчання без нагляду, навпаки, працюють самостійно, щоб виявити властиву структуру немічених даних. Зверніть увагу, що для перевірки вихідних змінних вони все ще потребують втручання людини. Наприклад, модель без нагляду може визначити, що покупці в Інтернеті часто купують групи товарів одночасно. Однак аналітику даних доведеться

підтвердити, що має сенс механізм рекомендацій групувати дитячий одяг за замовленням підгузників, яблучного пюре та чашок sipru.

Інші ключові відмінності між контрольованим та неконтрольованим навчанням:

- Цілі: Під час контрольованого навчання метою є прогнозування результатів для нових даних. Ви заздалегідь знаєте, яких результатів очікувати. Завдяки неконтрольованому алгоритму навчання мета полягає в тому, щоб отримати уявлення про великі обсяги нових даних. Машинне навчання саме визначає, що відрізняється чи цікаве від набору даних.

- Програми: Моделі навчання під контролем ідеально підходять для виявлення спаму, аналізу настроїв, прогнозування погоди та прогнозування цін, серед іншого. Навпаки, безконтрольне навчання чудово підходить для виявлення аномалій, механізмів рекомендацій, персональних даних клієнтів та медичних зображень.

- Складність: Контрольоване навчання - це простий метод машинного навчання, який, як правило, обчислюється за допомогою таких програм, як R або Python. При неконтрольованому навчанні потрібні потужні інструменти для роботи з великими обсягами некласифікованих даних. Моделі навчання без нагляду є складними в обчислювальному плані, оскільки їм потрібен великий навчальний набір для отримання запланованих результатів.

- Недоліки: Моделі навчання під наглядом можуть зайняти багато часу, а мітки для змінних введення та виведення вимагають знань. Тим часом, некеровані методи навчання можуть мати надзвичайно неточні результати, якщо у вас не буде людського втручання для перевірки вихідних змінних.

Тож який метод можна назвати кращим саме для Вас?!

Вибір правильного підходу для вашої ситуації залежить від того, як ваші науковці оцінюють структуру та обсяг ваших даних, а також від випадку використання. Щоб прийняти рішення, обов'язково виконайте такі дії:

- Оцініть свої вхідні дані: це марковані чи немарковані дані? Чи є у вас експерти, які можуть підтримати додаткове маркування?
- Визначте свої цілі: Чи маєте вирішити постійну, чітко визначену проблему? Або алгоритму потрібно буде передбачати нові проблеми?
- Перегляньте свої варіанти алгоритмів: Чи існують алгоритми з однаковою розмірністю, яка вам потрібна (кількість ознак, атрибутів чи характеристик)? Чи можуть вони підтримувати обсяг та структуру даних?

Класифікація великих даних може бути справжньою проблемою у навчанні під контролем, але результати є дуже точними та надійними. Навпаки, навчання без нагляду може обробляти великі обсяги даних у режимі реального часу. Однак бракує прозорості щодо кластеризації даних та підвищується ризик отримання неточних результатів. Тут з'являється навчання під наглядом.

Не можете вирішити, застосовувати під наглядом чи без нагляду? Напівконтроль - це щасливий засіб, коли ви використовуєте набір навчальних даних як з маркованими, так і без маркованих даних. Це особливо корисно, коли важко витягти відповідні функції з даних - і коли у вас великий обсяг даних.

Напівконтроль ідеально підходить для медичних зображень, де невелика кількість навчальних даних може призвести до значного підвищення точності. Наприклад, рентгенолог може позначити невелику підмножину КТ на наявність пухлин або захворювань, щоб апарат міг точніше передбачити, яким пацієнтам може знадобитися більше медичної допомоги.

Напівконтроль використовується для тих самих програм, що і навчання під наглядом. Але для навчання використовується як марковані, так і немічені дані - як правило, невелика кількість маркованих даних із великою кількістю немаркованих даних (оскільки немічені дані є менш дорогими та вимагають менших зусиль для їх отримання). Цей тип навчання можна використовувати з такими методами, як класифікація, регресія та прогнозування. Навчання під наглядом є корисним, коли витрати, пов'язані з маркуванням, занадто високі,

щоб забезпечити повністю маркований навчальний процес. Ранні приклади цього включають виявлення обличчя людини у веб-камері.

Підкріплене навчання часто використовується для робототехніки, ігор та навігації.[17] Під час навчання з підкріпленням алгоритм виявляє шляхом спроб і помилок, які дії приносять найбільші винагороди. Цей тип навчання має три основні компоненти: агент (той, хто навчається або хто приймає рішення), середовище (усе, з чим агент взаємодіє) та дії (що агент може робити). Мета полягає в тому, щоб агент вибирав дії, які максимізують очікувану винагороду протягом заданого періоду часу. Агент досягне мети набагато швидше, дотримуючись належної політики. Тож метою навчального підкріплення є вивчення найкращої політики.

Окрім вже озвучених, вдосконалюються також і інші напрямки розвитку ШІ: активне, багатозадачне, різноманітне, трансферне... Протягом останнього часу успішно використовується «глибоке навчання», яке однаково може поєднувати в собі навчання з вчителем і навчання без вчителя. Досить добре цей метод зарекомендував себе у випадках обробки великих обсягів даних. Наприклад фотографії свійських тварин з ярликами: це кінь, а то корова.

Треба розробити такий алгоритм, щоб машина могла по запропонованій фотографії, яку раніше «не показували», визначити що там зображено: кінь, або корова. «Вчителем» буде людина, яка наперед розставила мітки. А ШІ вже сам вибере критерії, згідно яких буде обробляти запит. Також в майбутньому можна швидко змінити основну задачу і цей алгоритм сам переналаштується на розпізнавання інших об'єктів.

Ці алгоритми навчання натхнені зв'язками нервових клітин у мозку людини. Мозок обробляє інформацію через нейрони та синапси. Подібним чином, штучні нейронні мережі складаються з декількох рядів вузлів даних, які з'єднані між собою за допомогою зважених з'єднань.

Нейромережа навчається шляхом неодноразового подання їй даних. Завдяки цьому повторенню нейронна мережа вчиться класифікувати дані

щоразу точніше. Це працює, постійно коригуючи зважування для окремих зв'язків між нейронними шарами. Тоді модель, сформована в ході навчальних занять, також може бути застосована до даних, про які Штучний інтелект ще не дізнався під час навчання.

Якщо нейронні мережі мають приховані нейронні шари, які не пов'язані безпосередньо з вхідним або вихідним шаром, вони називаються "глибокими нейронними мережами". Глибокі нейронні мережі можуть мати сотні тисяч або мільйони шарів нейронів. Це означає, що дедалі складніші проблеми можна вирішити за допомогою так званого «глибокого навчання».

1.4 Цільова Аудиторія в інтернет магазинах

Сьогодні Машинне навчання дуже часто використовується для опрацювання цільової аудиторії інтернет магазинів. Цільова аудиторія це власне група відвідувачів сайту, які з найбільшою імовірністю можуть бути зацікавленими у асортименті товарів і(або) послуг та здійснять покупку.[2]

Саме цю групу людей вважають найбільш цінною для інтернет магазину. Їхні смаки враховуються найбільше при розробці та просуванні сайту, що в подальшому дозволить підвищити ефективність ресурсу та досягти потрібного ефекту.[1]

Виявити ЦА є важливим з таких причин:[3]

- отримані дані дозволять зробити більш направлену рекламу в пошуковиках та соціальних мережах;
- опрацювання даних дозволить надати асортимент, який буде користуватись більшим попитом;
- знання ЦА дасть змогу значно збільшити трафік клієнтів;
- добре продумана ЦА дозволить знизити затрати в часі та коштах на просування інтернет-магазину.

Щоб якнайкраще визначити потенційних покупців, треба визначити їх за такими критеріями:

- основні характеристики – вік, місце проживання, сімейний стан;
- інтереси – така інформація зазвичай береться з їхніх акаунтів у соціальних мережах;
- рівень платоспроможності – формується на основі вартості товарів;
- завдання – визначає які потреби покупця задовільнить продукт.

Однією з найбільших помилок помилок, які часто дупускають, при підборі ЦА є занадто широке охоплення людей. Витрачені кошти на таку рекламу окупаються дуже рідко.[4]

Зазвичай таку широку аудиторію поділяють на підгрупи, які відрізняються інтересами.

Наприклад онлайн магазин меблів розрахований на осіб різного віку і з різними смаками.[5] Це особи від 18 до 60 років, різних верств населення з різними ціновими запитам та різними смаками. Таку аудиторію можна розділити на декілька менших підгруп[6] і там вже виявити конкретні запити:

- 18-25 років. Молоді люди, не одружені проживають в орендованом житлі.
- 25-40 років. Одружені, купили власне житло і роблять ремонт. Мають дітей. Хочуть отримати якісні меблі.
- 40-60 років. Особи які проживають окремо від дітей. Старі меблі уже втратили гарний зовнішній вигляд, або в поганому стані.

Чим конкретніший опис опис ЦА, тим вущє коло соіб можна охопити, які були б дійсно зацікавленні у покупці.

Так можна визначити і ціновий сегмент для кожної підгрупи і пропонувати саме той товар який користується попитом серед інших клієнтів з цієї самої групи.[7]

Незвлежно від етапу на якому знаходиться веб-ресурс визначення ЦА є дуже важливою задачею. Погано складений портрет призведе до того, що бізнес не зможе знайти покупців для своєї продукції.[8]

1.5 Висновок до першого розділу

Машинне навчання – обширний підрозділ штучного інтелекту, який вивчає побудову алгоритмів, які можуть навчатись самостійно. Існує два основні типи такого навчання: навчання з претендентами, яке базується на закономірностях, виявлених у емпіричних даних. Та навчання дедуктивне, яке використовує знання експертів і заносить їх у комп'ютер у вигляді баз даних. Також багато методів індуктивного навчання були розроблені як альтернатива класичним статичним підходам.

Для того щоб розробити інформаційну систему для підбору ЦА інтернет магазинів буде використана модель на основі TensorFlow.js так як вона працює через браузер і не потребує встановлення додаткових драйверів зі сторони користувача. Також бібліотека TensorFlow.js дозволяє працювати з різними мовами програмування, що спрощує роботу розробників і суттєво економить час.

РОЗДІЛ 2. ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ В ЗАДАЧАХ ДЛЯ ПІДБОРУ ЦІЛЬОВОЇ АУДИТОРІЇ ІНТЕРНЕТ МАГАЗИНІВ

2.1 Проектування навчальної моделі.

Модель машинного навчання - це функція з вивченими параметрами, що відображає вхідні дані до бажаного результату. Оптимальні параметри отримуються шляхом навчання моделі на даних.

Багато моделей було впроваджено[24] в Python, оскільки Python є популярним вибором серед науковців даних і має найкращу підтримку з точки зору функцій. Однак широке поширення глибокого навчання у всіх типах додатків залучило розробників з різних мов програмування. Крім того, практики впровадження моделей стали краще зрозумілими та широко доступними, що дозволяє більшій кількості розробників створювати власну модель, яка краще відповідає їх застосуванню.

На щастя, TensorFlow[14] розроблений для підтримки прив'язок різних мов, зокрема, програмування на Python, C, R[22], JavaScript та Java™. Оскільки кожна мова надає свій набір переваг, розробники мають свої причини для вибору мови програмування. Тому важливо дозволити розробникам залишатися зі своїми звичними середовищами програмування, а не вимагати від них вивчення нової мови.

Концептуально нейронна мережа складається з безлічі шарів ваг разом із обчисленнями, які представлені у вигляді вузлів та ребер на графіку. Платформи програмування підтримують реалізацію цих графіків по-різному. У попередніх підходах ви працювали б на низькому рівні, чітко розподіляючи тензори та кодуючи індивідуальні обчислення на тензорах. Оскільки технологія швидко розвивалася протягом останніх кількох років, з'явилися загальні шаблони, і ці моделі вбудовані в абстракції програмування вищого рівня, де

цілі шари в нейронних мережах доступні як API на платформі. Моделі легко будувати, складаючи ці шари між собою, значно спрощуючи реалізацію[15].

Для побудови моделі візьмемо онлайн магазин італійських меблів. Асортимент досить обширний і підійде для кожного покупця. Аудиторія магазину також велика, тому її ми поділили на підгрупи:

- 18-25 років. Молоді люди, не одружені проживають в орендованому житлі.
- 25-40 років. Одружені, купили власне житло і роблять ремонт. Мають дітей. Хочуть отримати якісні меблі.
- 40-60 років. Особи які проживають окремо від дітей. Старі меблі уже втратили гарний зовнішній вигляд, або в поганому стані.

Модель будемо навчати на аудиторії 25-40 років. Вхідні дані візьмемо з

Ніхто з нас не любить довго затримуватись на сайтах де ми не бачимо того що шукали, або того що нам може сподобатись.

Щоб уникнути небажаного впливу відвідувачів потрібно зробити так, щоб при першому перегляді сайту одразу показувались ті товари, які з найбільшою вірогідністю сподобаються клієнтові. Для цього ми підключимо машинне навчання.[19]

Передовсім нейронну мережу потрібно навчити.

Тренінг включає кілька етапів:

- Отримання пакету даних у модель.
- Ми просимо модель зробити прогноз.
- Порівняння цього прогнозу з “справжнім” значенням.
- Вирішення, наскільки змінювати кожен параметр, щоб модель могла краще передбачити майбутнє для цієї партії.

Добре навчена модель забезпечить, щоб вхідні дані точно відповідали бажаним результатам. [20]

Розпочати слід з того щоб створити HTML файл і підключити до нього JavaScript-файл script.js., як зображено на рисунку 2.1

```

<!DOCTYPE html>
<html>
<head>
  <title>TensorFlow.js Tutorial</title> <!-- Import
TensorFlow.js -->
  <script
src="https://cdn.jsdelivr.net/npm/@tensorflow/tfjs@1.0.0/dist/tf.
mi
n.js"></script>
  <!-- Import tfjs-vis -->
  <script
src="https://cdn.jsdelivr.net/npm/@tensorflow/tfjs-
vis@1.0.2/dist/tfjs-vis.umd.min.js"></script> <!-- Import the
main script file -->
  <script src="script.js"></script></head>
<body>
</body>
</html> |

```

Рисунок 2.1 – Початковий HTML код

Всі початкові дані будемо брати з файлу JSON.

```

{
  "Avg. Visitors Income": 79545.45857,
  "Avg. Visitors Age": 25.682861322,
  "AvgNumbersofItems": 7.009188143,
  "Total Check": 1059033.558,
}
{
  "Avg. Visitors Income": 79248.64245,
  "Avg. Visitors Age": 26.002899808,
  "AvgNumbersofItems": 6.730821019,
  "Total Check": 1505890.915,
}

```

Рисунок 2.2 – Дані з файлу JSON

Вище на рисунку 2.2 наведено приклад даних у файлі JSON. З усіх цих методів ми використовуємо методи "AvgNumbersofItems" та "Total Check", які вказують кількість номерів та ціну відповідно.[23]

Також до файлу script.js додамо метод getData(), як зображено на рисунку 1.3.

```

    async function getData() {
const visitorDataReq = await
fetch('https://raw.githubusercontent.com/meetnandu05/ml1/master/vi
r.json');
const visitorData = await
visitorDataReq.json(); const cleaned =
visitorData.map(visitor => ({ total
check: visitor.totalcheck,
check: visitor.AvgNumbersofItems,
}))
.filter(visitor => (visitor.totalcheck!= null &&
visitor.AvgNumbersofItems!= null)); return cleaned;
}

```

Рисунок 2.3 – Метод getData()

Цей метод допоможе відсіяти всі наявні записи у яких не зазначено кінцевий чек, або кількість товару. [26]

Після цього треба внести дані, які використовуються для тренування на точкову діаграму. Щоб зробити це потрібно додати у script.js ще одну асинхронну функцію run().

```

async function run() {
  діаграму const data = await
  getData(); const values =
  data.map(d => ({ x: d.items,
  y: d.price,
  })); tfvis.render.scatterplot
  {name: 'No.of items v Price'},

  {values},
  {
    xLabel: 'No. of
  items', yLabel:
  'Price', height:
  300
  }
  );
}document.addEventListener('DOMContentLoaded', run);

```

Рисунок 2.4 – Функція run()

Якщо перезавантажити браузер то з'явиться панель з точковою діаграмою, як зображено на рисунку 2.5

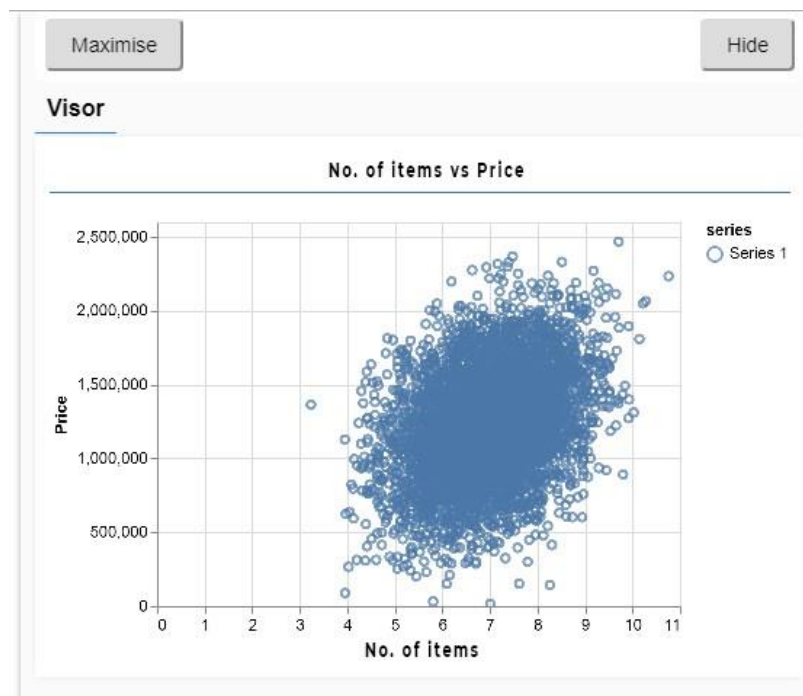


Рисунок 2.5 – Точкова діаграма

Зазвичай при роботі з великими даними хороша ідея це знайти можливість візуалізувати дані. Сама візуалізація дає поняття про те чи дані структуровані і чи зможе модель її вивчити. На рисунку 2.5 спостерігається позитивна кореляція між кількістю товарів та ціною. Отже коли збільшується кількість товарів в чеку, то збільшується і остаточна ціна в чекові.

Перед тим як передати моделі вхідні дані та отримати вихідні слід побудувати нейронну мережу. Для цього треба у файл script.js додати функцію createModel().

```
function createModel() {
  // створення послідовної моделі
  const model = tf.sequential();

  // Додавання одиночного прихованого шару
  model.add(tf.layers.dense({inputShape: [1], units: 1,
    useBias: true}));

  // Додавання вихідного шару
  model.add(tf.layers.dense({units: 1, useBias:
    true}));

  return model;
}
```

Рисунок 2.6 – Функція createModel()

Дана модель являється послідовною, так як її входи прямують прямо до виходу.

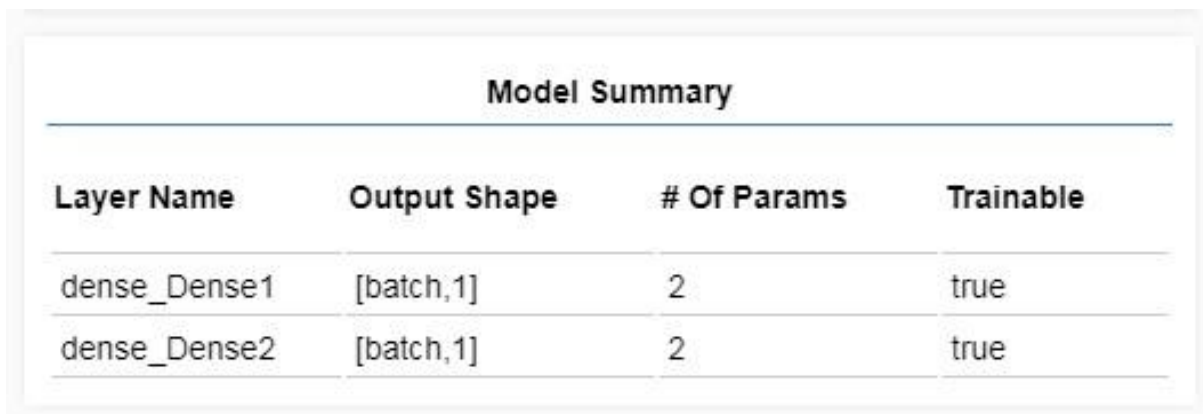
```
model.add(tf.layers.dense({inputShape: [1], units: 1,
  useBias: true}));
```

Строка вказана вище додає прихований шар. Так як це перший рівень , потрібно оприділити вхідну форму .inputShape є [1], так як число 1 наш вхід.

Також функцію run() треба іще доповнити:

```
// створення моделі const model = createModel();
tfvis.show.modelSummary({name: 'Model Summary'}, model);
```


Після цього з'явиться варіант моделі, який виведе зведення шарів на вебсторінці. Екземпляр зображений на рисунку 2.7.



The image shows a screenshot of a web browser displaying a 'Model Summary' table. The table has four columns: 'Layer Name', 'Output Shape', '# Of Params', and 'Trainable'. There are two rows of data representing dense layers.

Layer Name	Output Shape	# Of Params	Trainable
dense_Dense1	[batch,1]	2	true
dense_Dense2	[batch,1]	2	true

Рисунок 2.7 – Екземпляр моделі

2.2 Підготовка даних до навчання

Щоб отримати всі переваги TensorFlow.js, потрібно перетворити всі дані в тензори. Для цього треба у файл script.js дописати функцію convertToTensor(), як зображено на рисунку 2.8

```

function convertToTensor(data) { return tf.tidy(()
=> { // Step 1. Змішання даних
tf.util.shuffle(data);

// Step 2. Конвертування даних в
тензори
const inputs = data.map(d => d.items)
const labels = data.map(d => d.price);
const inputTensor = tf.tensor2d(inputs,
[inputTensor.length,
1]); const labelTensor = tf.tensor2d(labels,
[labelTensor.length, 1]);
//Step 3. Нормалізація даних за в ряд 0 - 1 з використанням
мін-мак масштабування
const inputMax = inputTensor.max();
const inputMin =
inputTensor.min(); const
labelMax = labelTensor.max();
const labelMin =
labelTensor.min();

const normalizedInputs =
inputTensor.sub(inputMin).div(inputMax.sub(inputMin))
; const normalizedLabels =
labelTensor.sub(labelMin).div(labelMax.sub(labelMin));
return {
inputs: normalizedInputs,
labels: normalizedLabels,
//Повернення мін/мак границь для подальшого
використання.
inputMax, inputMin,
labelMax, labelMin,
}
});
}

```

Рисунок 2.8 – Перетворення в тензори

Перший крок це власне перемішування даних. Він є важливим, бо модель не повинна отримувати однакові дані.

Далі дані поділяються у два масиви: один для вхідних даних, другий для вихідних значень. Після цього перетворюємо кожен масив у тензор.

Після цього дані нормалізуються в діапазоні від 0 до 1. Цей етап також є важливим бо TensorFlow.js розроблена під роботу не з великими числами.

Цей принцип треба зберегти щоб у подальшому була можливість «ненормалізувати» результати, і надати їм початкового масштабу і нормалізувати дані подібним способом.

2.3 Підготовка моделі до навчання та прогнозування

Тренувальний цикл запускаємо за допомогою коду зобраеного на рисунку 2.9

```

        return model.fit(inputs, labels, {
            batchSize,
            epochs,
            callbacks: tfvis.show.fitCallbacks(
                { name: 'Training Performance' },
                ['loss', 'mse'],
                { height: 200,
                callbacks:
                ['onEpochEnd']
                }
            )
        });

```

Рисунок 2.9 - Код для запуску тренувального циклу

`model.fit` – асинхронна функція, яку потрібно виконати, щоб абонент міг визначити коли навчання завершилось.

Щоб стежити за прогресом навчання деякі зворотні дзвінки передаються до `model.fit`.

Знову треба звернутись до функцій, які були визначені у `run()`. (Див. рисунок 2.10)

```

// Конвертування даних до форми для
тренування const tensorData =
convertToTensor(data); const {inputs, labels}
= tensorData;
// Навчання моделі
await trainModel(model, inputs, labels);
console.log('Done Training');

```

Рисунок 2.10 – Звернення до функції run()

Якщо оновити сторінку то появиться графік продуктивності навчання, яке зображено на рисунку 2.11

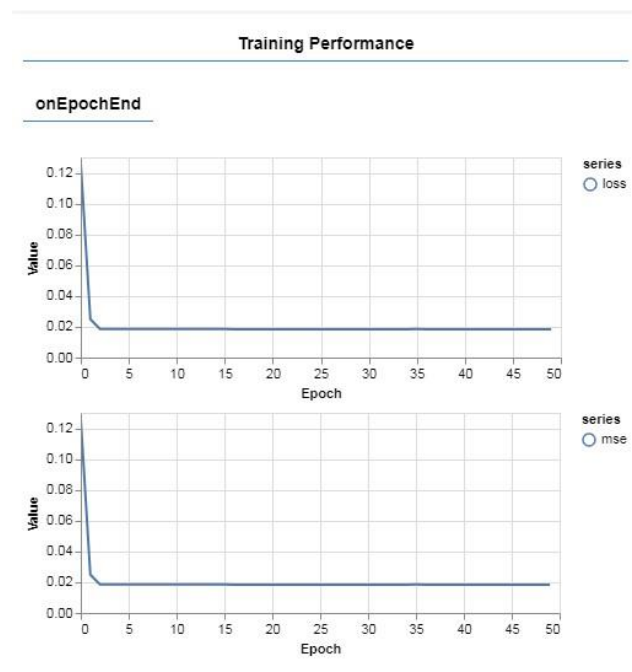


Рисунок 2.11 – Графіки продуктивності навчання

Саме вони виконуються завдяки зворотнім дзвінкам, які були створенні до того. Зображені графіки показують втрати та mse в кінці кожної епохи. Під час тренування моделі бажано зменшувати втрати. Так як наша метрика є мірою помилки, то потрібно щоб вона теж знижувалась.

Вже коли модель пройшла навчання треба зробити прогнози. Спершу треба оцінити, побачити її прогнози для рівномірного діапазону номерів. Для цього треба використати функцію `testModel()`, як зображено на рисунку 2.12

```
function testModel(model, inputData, normalizationData) {
  const {inputMax, inputMin, labelMin, labelMax} =
    normalizationData;
  // Генерування передбачень для ряду цифр між 0 та 1
  const [xs, preds] = tf.tidy(() => { const xs = tf.linspace(0,
    1, 100);

  const preds = model.predict(xs.reshape([100, 1]));
  const unNormXs = xs
    .mul(inputMax.sub(inputMin))
    .add(inputMin);
  const unNormPreds =
    preds
    .mul(labelMax.sub(labelMin))
    .add(labelMin);
  //Ненормалізація даних
  return [unNormXs.dataSync(), unNormPreds.dataSync()];
  });
  const predictedPoints = Array.from(xs).map((val, i) => {
  return {x: val, y: preds[i]}
  });
  const originalPoints = inputData.map(d
  => ({ x: d.items, y: d.price,
  }));
  tfvis.render.scatterplot(
  {name: 'Model Predictions vs Original Data'},
  {values: [originalPoints, predictedPoints], series:
  ['original',
  'predicted']},
  {
  xLabel: 'No. of
  items', yLabel:
  'Price', height:
  300 });}
```

Рисунок 2.12 – Використання функції `testModel()`

В цій частині генеруються 100 прикладів щоб передати їх для моделі. Щоб повернути дані у початковий вигляд діапазону (а не 0-1), потрібно використати звичайну інвертацію операції.

```
return [unNormXs.dataSync(), unNormPreds.dataSync()];
```

`dataSync()` – за допомогою цього методу можна отримати тип значень, які знаходяться в тензорі. Це якраз і дозволяє обробляти їх у JavaScript.

Тепер стає можливим використання `tfjs-vis` для побудови вихідних даних та прогнозів з моделі. В кінці ще треба звернутись до функції `run()`.

Після того можна оновити сторінку браузера і побачити передбачення, як тільки модель закінчить навчання. Саме це і зображено на рисунку 2.13

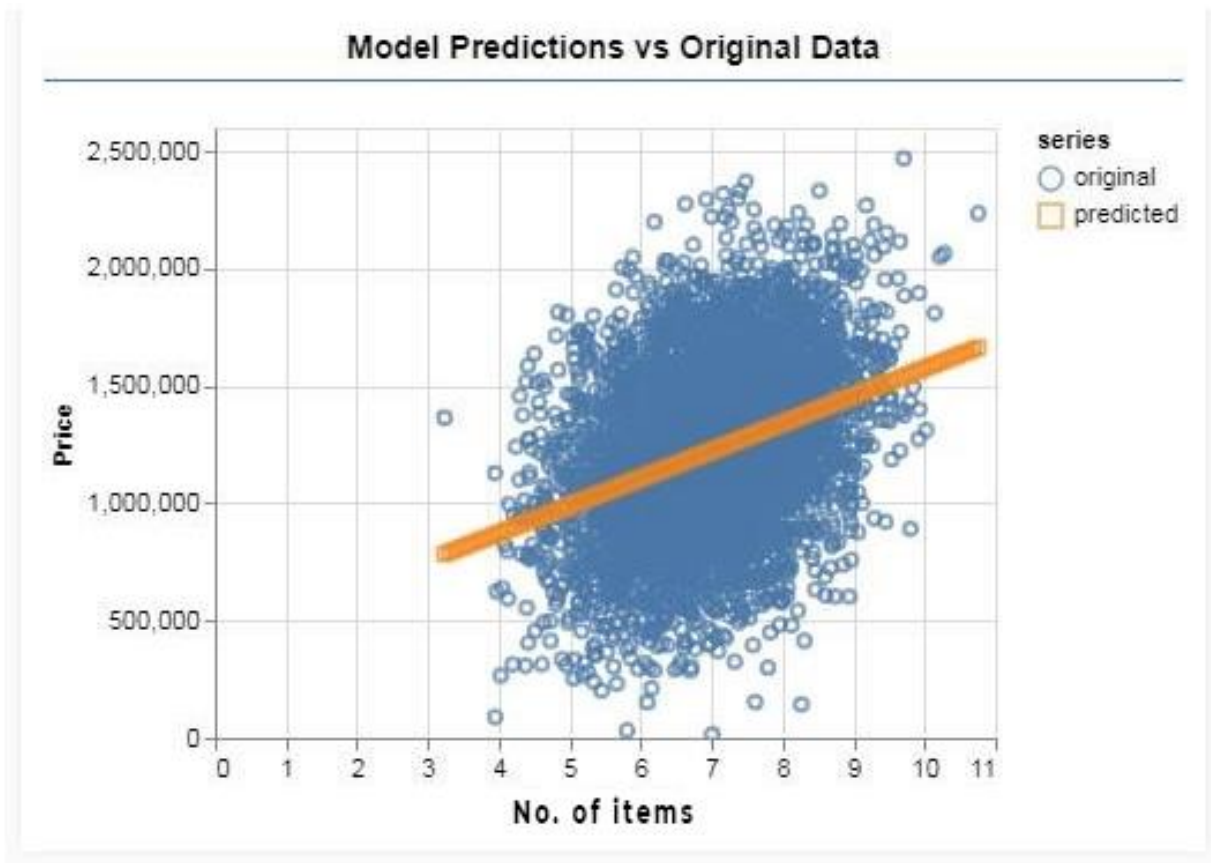


Рисунок 2.13 – Прогнозована модель

Цю ж саму модель машинного навчання можна перевчити для дослідження віку покупців і кількості товарів, які вони купують. Повторивши всі дії, які робили щоб визначити ціну залежно від кількості обраних товарів, можна зробити модель вік – кількість товарів.

Така модель зможе нам показати в якому віці покупці купують найбільшу кількість меблів. А отже які покупці витрачають найбільше коштів.

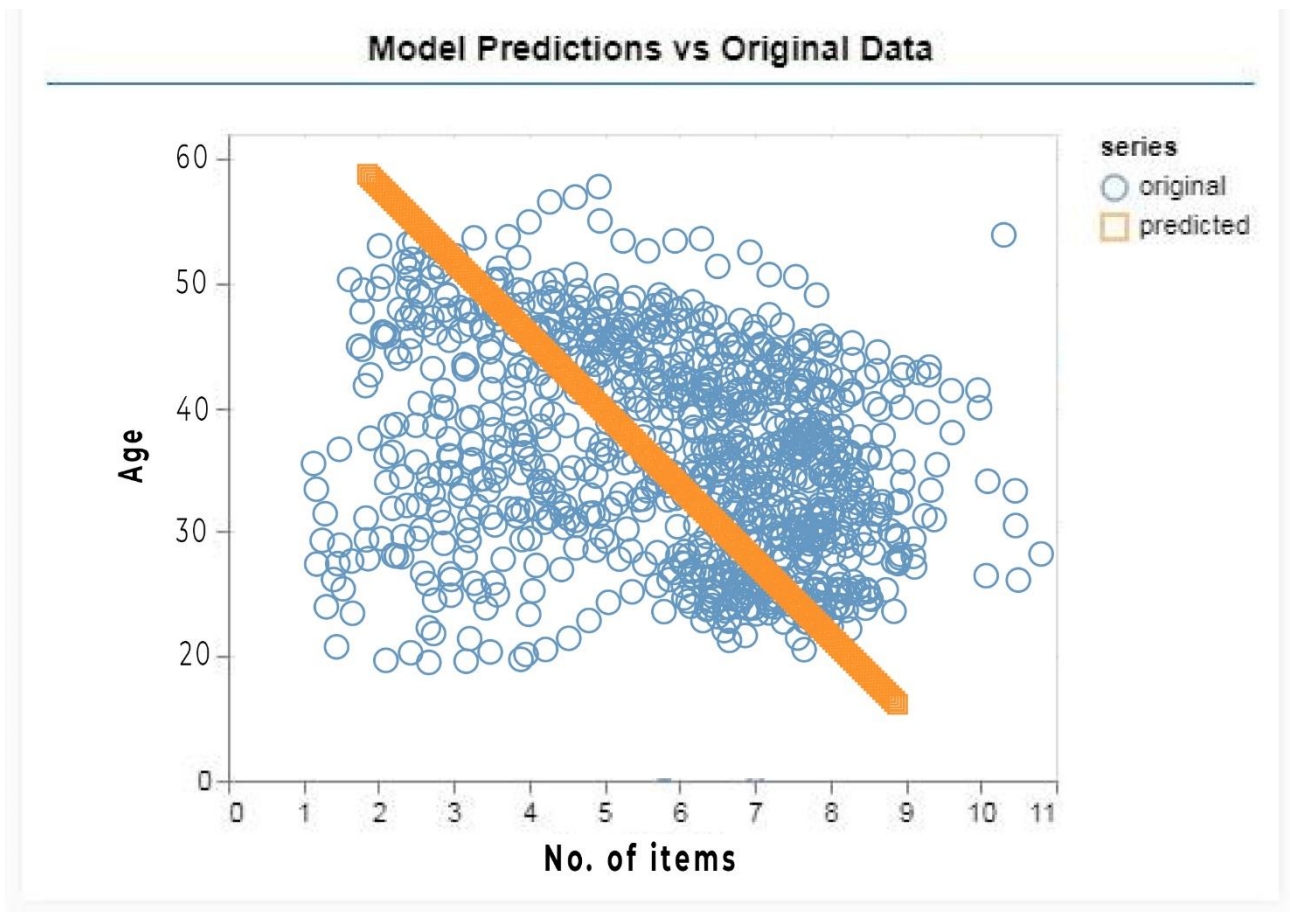


Рисунок 2.14 – Прогнозована модель вік – кількість товарів

На рисунках 2.13 ми бачимо прогноз того що, кінцева кількість товарів збільшує кінцеву суму в чекові.

Використовуючі ці дані можна краще зрозуміти свою ЦА. І відповідно запуснути рекламну кампанію акцентовану саме на таку категорію.

Також можна збирати та опрацьовувати інші дані відвідувачів, використовуючи даний метод. Чим більше і точніше оброблюватимуться данні, тим ефективніше їх можна використовувати щоб визначити ЦА і відповідно ефективніше робити рекламну кампанію.

2.4 Висновок до другого розділу

Запуск машинного навчання через браузер означає що з погляду користувача немає ніякої потреби встановлювати додаткові бібліотеки або

драйвери на свій пристрій. Крім того TensorFlow.js підтримує WebGL і автоматично виконує код. Користувач також може відкрити сайт через свій мобільний пристрій і в такому випадку модель може додатково використовувати гіроскоп, або акселерометр. Додатковим плюсом являється те що всі дані обробляються на стороні користувача, що робить його придатним для застосунків, що зберігають конфіденційність.

На сьогоднішній день можливості машинного навчання дозволяють максимально точно визначити ЦА для будь якого продукту. Чим точніше це зробити, тим краще можна опрацювати ці дані і вдало це реалізувати. Саме це ми бачили у другому розділі.

РОЗДІЛ 3. БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ ТА ОСНОВИ ОХОРОНИ ПРАЦІ

3.1. Психологічні причини нещасних випадків і травматизму

Психологія безпеки розглядає застосування психологічних знань для забезпечення безпеки діяльності людини. Психологія безпеки праці становить важливу ланку в структурі заходів щодо забезпечення безпечної діяльності людини. Проблеми аварійності та травматизму на сучасних виробництвах неможливо вирішувати тільки інженерними методами.

В основі аварійності та травматизму лежать організаційно–психологічні причини: низький рівень професійної підготовки з питань безпеки, недостатнє виховання, допуск до небезпечних видів робіт осіб з підвищеним ризиком травматизації, перебування людей в стані стомлення або інших психічних станах.

Таким чином, у психологічній класифікації причин виникнення небезпечних ситуацій і нещасних випадків можна виділити три класи: **Порушення мотиваційної частини дій.** Виявляється в небажанні виконувати певні дії (операції). Порушення може бути відносно постійним (людина недооцінює небезпеку, схильний до ризику, негативно відноситься до трудового та технічного регламенту) і тимчасовим (людина в стані депресії, алкогольного сп'яніння). **Порушення орієнтовної частини дій.** Виявляється в незнанні правила експлуатації сільськогосподарської техніки та технічних систем і норм безпеки праці та способів їх виконання.

Порушення виконавчої частини. Виявляється в невиконанні правил (інструкцій, розпоряджень, норм) внаслідок невідповідності психічних і фізичних можливостей людини вимогам роботи. Порушення можуть бути постійними (недостатня координація, погана концентрація уваги) і тимчасовими (перевтома, зниження працездатності, погіршення стану здоров'я, стрес, алкогольне сп'яніння).

Ці порушення створюють реальну можливість виникнення небезпечних ситуацій і нещасних випадків .

При експлуатації і ремонті техніки необхідно призначити групу профілактичних заходів в кожній частині:

- мотиваційна частина пропаганда і виховання;
- орієнтовна частина навчання, відпрацювання навичок;
- виконавча частина профвідбір, медичне обстеження.

Крім того, причини виникнення небезпечних ситуацій і виробничих травм пов'язаних з людським фактором, можна розкласти по різних рівнях, а саме:

- рівень індивідуума (вроджені чи набуті тимчасово або постійно психічні та фізіологічні характеристики організму);
- рівень ближній середовища (умови праці, порушення колегіальних відносин, незадовільний інструктаж з безпеки праці, житлові та матеріальні турботи);
- рівень суспільства (недостатня інформованість про професійні ризики та їх наслідки, вади в стратегії організації безпечної праці в галузі, порушення в умовах праці).

За обставин, однакових для всіх працюючих, що визначає значення у формуванні на виробництві поведінки кожної людини, мають його індивідуальні якості, що відображають сукупність соціально– психологічних і фізіологічних властивостей. Вони включають тип нервової системи, темперамент, характер, особливості мислення, освіту, досвід, виховання, здоров'я .Розглядаючи поведінку працюючого, можна виділити об'єктивні чинники виробничої обстановки, що створюють небезпечні дії і які спричиняють виникнення небезпечних ситуацій:

- відсутність з боку адміністрації належного контролю за дотриманням правил безпеки;

- конструктивні недоліки засобів захисту працюючих від впливу небезпечних і шкідливих виробничих факторів, які дають можливість проникнення в небезпечну зону;
- конструктивні недостачі блокуючих пристроїв, що дозволяють нейтралізувати їх дії і можливість ведення роботи небезпечними методами;
- недосконала технологія виробничих процесів, що допускає ведення операції з порушеннями правил безпеки;
- плинність кадрів, яка призводить до залучення до робіт у вихідні дні, надурочних робіт; погано поставлена пропагандистська і виховна робота;
- порушення правил виконання операцій в результаті відсутності інструменту, механізму, матеріалу, установки, на наявність яких виконавець мав право розраховувати при виконанні завдання;
- недостатня надійність. Інструмент, матеріал, верстат або установка раптово стають непридатними для подальшого використання, в той час як раніше вони знаходилися в робочому стані;
- недоліки в проектуванні обладнання. Помилки в проектуванні призводять до того, що інструменти, верстати або установки стає важко використовувати при виконанні виробничих завдань;
- особисті чинники. Недостатні знання методів виконання завдання, помилки інструктажу, погане навчання способів використання інструменту, неправильна або небезпечна поведінка під час роботи;
- неповне виконання елементів завдання. Елемент запланованого завдання не був виконаний вчасно або виконувався з помилками;
- небезпечні ситуації, які не залежать від завдання, є потенційною небезпекою (погані умови праці, відсутність або несправність засобів індивідуального захисту). Ризик, пов'язаний з такими ситуаціями, може бути, як правило, виключений за допомогою попереджувальних заходів технічного або організаційного характеру;

З метою попередження порушення правил безпеки необхідно застосовувати заходи організаційного та технічного характеру, що виключають можливість появи або створення умов для виконання небезпечних дій, позбавити працюючого можливості робити вибір між небезпечним і безпечним способом діяльності; посилити виховну, пропагандистську і навчальну діяльність, направлену на формування необхідної поведінки.

3.2.Санітарно-гігієнічні вимоги до умов праці

Санітарно-гігієнічні вимоги до умов праці під час виконання роботи відповідали визначеним нормативам:

Мікроклімат приміщення характеризується температурою, вологістю та швидкістю руху повітря, інтенсивністю радіації, переважно в інфрачервоній та ультрафіолетовій областях спектру електромагнітних випромінювань відповідав установленим нормам і параметрам. Параметри мікроклімату у приміщенні забезпечували комфортне самопочуття організму. Параметри мікроклімату закритих приміщень унормовані за санітарні норми Державних Санітарних Стандартів 3.3.6.042-99 «Санітарні норми мікроклімату виробничих приміщень».

Освітлення приміщень та робочих місць було забезпечене відповідно до встановлених вимог. Відносно вікна робоче місце було розміщено так, що природне світло було збоку, переважно з лівого та забезпечувало коефіцієнт природної освітленості не нижче 1,5 %.

Освітленість за штучного освітлення в площині робочої поверхні становила 300 – 500 Лк. Відношення яскравості робочих поверхонь було 3:1, а яскравість робочих поверхонь і стін (іншого обладнання) – 5:1.

Робоче місце, обладнане ПК було розташоване так, щоб уникнути потрапляння в очі прямого світла. Щоб уникнути світлових відблисків від екрану та клавіатури використане комп'ютерне обладнання з матовою поверхнею.

Для захисту очей від прямого сонячного світла чи джерел штучного освітлення застосовані захисні козирки та жалюзі на вікнах. Використана система вимикачів, що дозволяє регулювати інтенсивність штучного освітлення залежно від інтенсивності природного, а також дозволяє освітлювати тільки потрібні для роботи зони приміщення. Для забезпечення нормованих значень освітлення очищалося віконне скло та світильники не рідше, ніж 2 рази на рік, та своєчасно проведена заміна ламп, що перегоріли. Вікна приміщень мали регульовальні пристрої для відкривання.

Дотримані вимоги до рівнів шуму та вібрації. Шум часто є причиною зниження рівня працездатності, підвищення рівня загальної та професійної захворюваності, частоти виробничих травм.

Шум і стрес є загально-біологічними подразниками, які негативно впливають на усі органи і системи організму. У разі тривалого систематичного впливу шуму може виникнути патологія з переважним ураженням слуху, центральної нервової і серцево-судинної систем. Було дотримано допустимих рівнів звукового тиску в октавних смугах частот, еквівалентні рівні звуку на робочих місцях встановлені санітарними нормами виробничого шуму, ультразвуку та інфразвуку Державних Санітарних Стандартів 3.3.6.037-99 «Санітарні норми виробничого шуму, ультразвуку та інфразвуку».

Використані індивідуальні засоби і заходи захисту від шкідливого впливу виробничих чинників на здоров'я людини (засоби захисту органів дихання, органів зору, шкіри тощо). Облаштовуючи приміщення для роботи з ПК, передбачена припливно-витяжна вентиляція та кондиціонування повітря.

Надходження свіжого повітря регульоване, виходячи із відповідних нормативних умов якщо об'єм приміщення 20 м^3 , то потрібно подати не менш як $30 \text{ м}^3/\text{год}$ повітря; якщо об'єм приміщення у межах від 20 м^3 до 40 м^3 , то потрібно подати не менш як $20 \text{ м}^3/\text{год}$ повітря; якщо об'єм приміщення становить понад 40 м^3 , допускається природна вентиляція, у випадку, коли немає виділення шкідливих речовин.

Передбачений захист від шуму та вібрацій. Усунення шуму в приміщенні є однією з найскладніших проблем, оскільки джерела шуму різноманітні й потребують комплексу заходів технічного, організаційного і медичного характеру на всіх стадіях проектування, будівництва, експлуатації машин і устаткування.

Застосовані три головні напрямки зменшення впливу шуму на організм людини (зменшення рівня шуму у джерелі виникнення, застосування раціональних конструкцій, нових матеріалів і технологічних процесів; звукоізоляція устаткування за допомогою оздоблення стін, стелі, підлоги тощо; використання засобів індивідуального захисту). Рівні шуму та вібрації на робочих місцях осіб, що працюють з ПК, визначені відповідно до Державних санітарних Норм і Правил 3.3.2-007-98 «Гігієнічні вимоги до організації роботи з візуальними дисплейними терміналами електронно-обчислювальних машин»

Дотримані заходи особистої гігієни на робочому місці (підтримання чистоти, миття рук тощо). Заходи особистої гігієни на робочому місці передбачають щоденне вологе прибирання, утримання у чистоті робочого місця, наявність на робочому місці тільки необхідних для роботи засобів. На робочому місці необхідно дотримуватись вимог правил внутрішнього розпорядку, зокрема, заборонено приймати їжу, пити, курити.

Заходи безпеки під час експлуатації інших електричних приладів передбачають дотримання таких правил: постійно стежити за справним станом електромережі, розподільних щитків, вимикачів, штепсельних розеток, лампових патронів, а також мережевих кабелів живлення, за допомогою яких електроприлади під'єднують до електромережі; постійно стежити за справністю ізоляції електромережі та мережевих кабелів, не допускаючи їхньої експлуатації з пошкодженою ізоляцією; не тягнути за мережевий кабель, щоб витягти вилку з розетки; не закривати меблями, різноманітним інвентарем вимикачі, штепсельні розетки; не підключати одночасно декілька потужних електропристроїв до однієї розетки, що може викликати надмірне нагрівання провідників, руйнування їхньої ізоляції, розплавлення і загоряння полімерних

матеріалів; не залишати увімкненими електроприлади без нагляду; не допускати потрапляння всередину електроприладів крізь вентиляційні отвори рідин або металевих предметів, а також не закривати їх та підтримувати в належній чистоті, щоб уникнути перегрівання та займання приладу; не ставити на електроприлади матеріали, які можуть під дією теплоти, що виділяється, загорітися (канцелярські товари, сувенірну продукцію).

Враховані можливі на робочому місці аварійні ситуації техногенного характеру та загрози природного характеру, що можуть перерости у надзвичайні ситуації (пожежі, вибухи, аномальні гідрометеорологічні явища та медико-біологічні загрози). Вжито усіх заходів створення безпечних умов для роботи відповідно до нормативних документів. Передбачені заходи запобігання виникненню пожежі та вибуху, наявні первинні засоби пожежогасіння.

Наявна схема евакуації з приміщення, у якому виконувалася кваліфікаційна робота, зокрема, обґрунтований вибір шляхів евакуації, які забезпечують якнайшвидше і найбезпечніше виведення людей з небезпечних приміщень.

Для безпечної евакуації шляхи та виходи відповідають таким вимогам: евакуаційні шляхи і виходи вільні, не зашарашені та у разі потреби забезпечують евакуацію всіх людей, які перебувають у приміщенні; кількість та розміри евакуаційних виходів, їх конструктивні рішення, умови освітленості, забезпечення незадимленості, протяжність шляхів евакуації, їх оздоблення відповідають протипожежним вимогам будівельних норм; двері на шляхах евакуації відчиняються в напрямку виходу з будівель (приміщень) і замикаються лише на внутрішні запори, що можуть бути легко відімкнені.

ВИСНОВКИ

Машинне навчання, як і сам штучний інтелект, все частіше зустрічається у повсякденному житті людини. Саме тому ця галузь розробки програмного забезпечення є доволі трендовою, цікавою і являє собою широке поле для майбутніх досліджень. Розвиток штучного інтелекту формуватиме майбутні тенденції у розробці програмного забезпечення, і від нього залежатиме також розвиток суміжних галузей.

В даній роботі було описано тенденції та історія розвитку штучного інтелекту та машинного навчання. Розглянуто основні популярні алгоритми машинного навчання, основні принципи та проблеми машинного навчання.

Дана дипломна робота була націлена на проблематику браузерного машинного навчання – одного з напрямків використання машинного навчання. Було досліджено основні принципи розробки з використанням бібліотеки TensorFlow.js.

Загалом, ідея розробки веб-застосунків з використанням машинного навчання у браузері є доволі перспективною завдяки своїй відносно низькій енергозатратності, вседоступності, та захищеності даних.

Програми браузерного машинного навчання пишуться на популярних мовах програмування JavaScript або Typescript, що дозволяє швидко обробляти інформацію, повністю використовувати потенціал браузера, та розширень, які ним підтримуються.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Котлер Ф. Основы маркетинга: / Ф. Котлер. Пер. с англ. — М.: «Бизнескнига», «ИМА-Кросс. Плюс», ноябрь 1995. — 670 с.
2. Ландреви Ж., Леви Ж., Линдон Д. Меркатор. Теория и практика маркетинга: / Ж. Ландреви, Ж. Леви, Д. Линдон, Меркатор. Пер. с франц. В 2 т. — М.: МИФЭР, 2006. — 512 с.
3. Литовченко И.Л. Новые формы коммуникации в Интернетмаркетинге. / И.Л. Литовченко Вісник соціально-економічних досліджень: Зб. наук. праць. Вип. 33. — Одеса, 2008. — 193. — 198 с.
4. Литовченко І.Л. Інтернет — маркетинг : Навч. пос. Рекомендовано МОН України./ І.Л. Литовченко, В.П. Пилипчук. -Київ: Центр учбової літератури, 2008. -184 с.
5. Литовченко І.Л. Інтернет-маркетинг як інструмент антикризових програм підприємства. / І.Л. Литовченко // Маркетинг в Україні. — 2009. — № 5. — С. 49–53.
6. Литовченко І.Л. Інтернет-маркетинг. Маркетинг у секторах національної економіки../ І.Л. Литовченко / За ред. М.О. Окландера. — Одеса: Астропринт, 2004. 546 с.
7. Литовченко І.Л. Наукові засади ціноутворення в Інтернетмаркетингу І.Л. Литовченко // Вісник соціально-економічних досліджень. — Одеса: ОДЕУ, 2004. — Вип. 17. 976 с.
8. Литовченко І.Л. Удосконалення класифікації засобів маркетингової комунікативної діяльності../ І.Л. Литовченко // Вісник Національного ун-ту «Львівська політехніка». Серія: Логістика. — Львів: Видтво Нац. ун-ту «Львівська політехніка», 2009. — №649.
9. Литовченко І.Л., Цуркан Ю.В. Особливості маркетингових досліджень через Інтернет./ І.Л. Литовченко., Ю.В. Цуркан // Вісник Національного ун-ту «Львівська політехніка». Серія: Логістика. — Львів: Вид-тво Нац. ун-ту «Львівська політехніка», 2004. — №499. 545 с.

10. Макарова М.В. Електронна комерція. / М.В. Макарова Посібник для студентів вищих навчальних закладів. — К.:Видавничий центр «Академія», 2002. — 272 с.
11. Руделіус У., Азарян О.А та ін Маркетинг: Підручник / У Руделіус, О.А. Азарян та ін. — К.: Навчально-методичний центр «Консорціум із удосконалення менеджмент-освіти в Україні», 2005. 642 с.
12. Котлер Ф. Маркетинг-менеджмент. 11-е изд./Ф. Котлер. — СПб.: Питер, 2003 — 800 с.
13. Duchi J. Adaptive subgradient methods for online learning and stochastic optimization / J. Duchi, E. Hazan, Y. Singer // Journal of Machine Learning Research. — 2011. — P. 2121–2159.
14. Офіційний сайт бібліотеки tensorflow.js. – [Електронний ресурс] - Режим доступу <https://www.tensorflow.org/js>.
15. Офіційний сайт бібліотеки machinelearn.js. – [Електронний ресурс] - Режим доступу <https://www.machinelearnjs.com/>.
16. Моделирование процессов обучения в нейронных сетях. – Режим доступу: <http://old.exponenta.ru/soft/others/mvs/stud3/3.asp>.
17. Саттон Р.С Обучение с подкреплением / Саттон Р.С, Э. Г. Барто // БИНОМ, Лаборатория знаний, 2014 – С. 42-96.
18. Duchi J. Adaptive subgradient methods for online learning and stochastic optimization / J. Duchi, E. Hazan, Y. Singer // Journal of Machine Learning Research. — 2011. — P. 2121–2159.
19. Офіційний сайт бібліотеки brain.js. – [Електронний ресурс] - Режим доступу <https://brain.js.org/>.
20. Офіційний сайт бібліотеки math.js. – [Електронний ресурс] - Режим доступу <https://mathjs.org/>.
21. Sutskever I. Sequence to sequence learning with neural networks / I. Sutskever, O. Vinyals, Q.V. Le // In Advances in neural information processing systems. — 2014. — P. 3104–3112.

22. Офіційний сайт бібліотеки r.js. – [Електронний ресурс] - Режим доступу <https://requirejs.org/>.
23. Офіційний сайт бібліотеки stdlib-js. – [Електронний ресурс] - Режим доступу <https://stdlib.io/>.
24. Онлайн журнал engadget (Google DeepMind AI wins final Go match for 4-1 series win). – Режим доступу:
<https://www.engadget.com/2016/03/14/thefinal-lee-sedol-vs-alphago-match-isabout-to-start/>.
25. Aggarwal C.C., Charu C. Data Classification Algorithms and Applications. 2015: Chapman & Hall /CRC.
26. Manevitz L. M. Y.M. Document Classification on Neural Networks Using Only Positive Examples // SIGIR. 2000.
27. Breiman L. Random forests / L. Breiman // Machine learning. — Vol. 45, N 1. — 2001. — P. 5–32.
28. Schematic illustration of the convolutional neural network. – Режим доступу:
http://www.nature.com/nature/journal/v518/n7540/fig_tab/nature14236_F1.html.
29. Dunning T., Friedman E. Practical Machine Learning: A New Look at Anomaly Detection. O'Reilly Media, 2004
30. Beckman R., Cook R. Outliers // Technometrics, No. 25(2), 1983. pp. 119– 149