

ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОСЕТЕЙ ДЛЯ АНАЛИЗА ВИЗУАЛЬНОГО КОНТЕНТА ВЕБ-ПЛАТФОРМЫ

Шныра Б.О., Андриевская Н.К.

Донецкий национальный технический университет, кафедра автоматизированных систем
управления;

E-mail: mi_shnyra@mail.ru

Аннотация:

Шныра Б.О., Андриевская Н.К. Использование сверточных нейросетей для анализа визуального контента веб-платформы. В современном мире развитие интернет-технологий и веб-платформ привело к необходимости создания эффективных инструментов обработки и анализа больших объемов графических данных, представленных в различных форматах. На основе анализа визуального контента, представленного пользователями, формируются персонализированные предложения для пользователей. В статье подробно описываются этапы обзора существующих разработок по теме исследований, анализа моделей сверточных нейросетей, разработки проекта, включающего в себя интеграцию созданной модели в веб-платформу, а также перспективы дальнейших исследований и разработок.

Ключевые слова: сверточные нейросети, веб-платформа, глубокое обучение, обработка изображения, прогнозирование.

Annotation:

Shnyra B.O., Andrievskaya N.K. Utilization of Convolutional Neural Networks for Analyzing Visual Content on a Web Platform. In the modern world, the development of Internet technologies and web platforms has led to the need to create effective tools for processing and analyzing large amounts of graphical data presented in various formats. Based on the analysis of the visual content presented by users, personalized offers for users are generated. The article describes in detail the stages of reviewing existing developments on the topic of research, analyzing models of convolutional neural networks, developing a project that includes integrating the created model into a web platform, as well as prospects for further research and development.

keywords: convolutional neural networks (CNNs), web platform, deep learning, image processing, prediction.

Постановка проблемы

В современном мире развитие интернет-технологий и веб-платформ позволяет пользователям не только получать информацию, но и активно участвовать в ее создании, распространении и обмене. Это привело к необходимости создания эффективных инструментов для обработки и анализа больших объемов графических данных, представленных в различных форматах. Одной из областей, где данные инструменты могут найти применение, является разработка инструментальных площадок для обмена идеями, где взаимодействие пользователей и обмен контентом между ними становится основой для инноваций и совместного развития.

На текущий момент подобные веб-платформы используют базовые методы обработки изображений, такие, как ручное аннотирование или простые алгоритмы машинного обучения. Однако данные методы не соответствуют современным требованиям по точности и масштабируемости, что делает исследования в данной области актуальными.

Обзор существующих разработок по теме исследований

Искусственный интеллект продолжает оставаться одним из главных технологических трендов 2024 года. Согласно данным международной исследовательской компании Markets and Markets, объем глобальных инвестиций в ИИ превысил \$100 млрд. По их же прогнозам, к 2025 году рынок искусственного интеллекта достигнет \$190 млрд благодаря активному распространению технологий машинного обучения, распознавания естественного языка, а также других передовых методов в таких секторах, как реклама, розничная торговля, финансы, здравоохранение и безопасность [1].

Компания «Яндекс» продолжает лидировать в применении технологий искусственного интеллекта в своих продуктах. Уже несколько лет она успешно внедряет ИИ в свои поисковые механизмы, что позволяет значительно повысить точность и релевантность результатов поиска, а также обеспечить высокое качество работы других сервисов компании, таких как автотранспортные, рекомендательные и голосовые системы [2].

Таким образом, использование искусственного интеллекта нейронных сетей поможет привести наше общество к высоким достижениям и решить глобальные проблемы, стоящие перед современным миром. Именно искусственный интеллект нейронных сетей, создаваемый и управляемый специалистами, владеющими этой высокотехнологичной, сверхсложной, суперсовременной технологией, способен в действительности решать жизненно важные и значимые вопросы современности [3].

В последние годы значительное внимание уделяется использованию сверточных нейросетей (CNN).

Интеграция CNN в системы рекомендаций может значительно улучшить их способность анализировать и интерпретировать визуальные данные, делая предложения более релевантными и персонализированными.

Однако, несмотря на потенциальные преимущества, интеграция сверточных нейросетей в системы рекомендаций также сталкивается с рядом проблем и вызовов. Одной из главных проблем является сложность обучения и настройки нейросетей на больших объемах данных, а также поддержание актуальности данных и их эффективности в условиях быстро меняющихся трендов и предпочтений пользователей.

Другой важной проблемой является необходимость разработки специализированных алгоритмов и методов анализа визуального контента, которые бы учитывали специфику пользовательских предпочтений и поведения в сети. Такие алгоритмы должны быть способны не только эффективно обрабатывать и классифицировать изображения, но и учитывать контекст и привычки пользователя для предоставления наиболее подходящих рекомендаций. Такой подход требует не только разработки новых технологий, но и исследований в области психологии и социологии, чтобы лучше понять потребности и ожидания пользователей.

Также важно учитывать и точность, и объективность результатов, получаемых с помощью нейросетей, чтобы избежать искажения информации и предвзятости в рекомендациях. Это требует разработки адекватных методов оценки качества работы алгоритмов и внедрения механизмов контроля за их работой.

Следовательно, несмотря на перспективы и потенциал использования сверточных нейросетей в системах рекомендаций, существует ряд вопросов и задач, которые требуют решения и дальнейших исследований для успешной интеграции этих технологий. Устранение этих проблем может значительно улучшить качество и персонализацию рекомендаций, а также повысить удовлетворенность пользователей и эффективность платформ для обмена идеями [4].

Разработка модели сверточной нейросети

Одним из ключевых аспектов обработки информации в системе анализа визуального контента является сбор данных, связанных с созданием пинов и досок. Эти данные не только представляют визуальный и текстовый контент, созданный пользователями, но и формируют

основу для тренировки моделей и дальнейшего функционирования системы рекомендаций на основе сверточных нейросетей.

Эффективный сбор и обработка данных о созданных пинах и досках – фундаментальные компоненты для обеспечения высокой производительности и точности рекомендательной системы на основе сверточных нейросетей [5-6]. Улучшение этих процессов способствует повышению удовлетворенности пользователей и расширению функциональных возможностей веб-платформы обмена идеями.

Для сбора данных о созданных пинах и досках на веб-платформе используются различные методы, включая мониторинг активности пользователей, анализ визуального и текстового контента, а также сбор метаданных о каждом созданном элементе.

Одним из способов сбора данных является отслеживание действий и дальнейшая фиксация результатов обработки данных пользователей на платформе, таких, как создание новых пинов, добавление комментариев или оценок, сохранение элементов в досках и т.д. Эти функции позволяют понять предпочтения и интересы пользователей, что в свою очередь помогает улучшить рекомендательную систему.

Для анализа визуального контента используются алгоритмы компьютерного зрения, которые позволяют идентифицировать объекты, людей, места и другие элементы на изображениях. Текстовый контент анализируется с помощью методов обработки естественного языка для определения ключевых слов и тематик, связанных с каждым пином или доской.

Собранные данные затем используются для тренировки моделей сверточных нейронных сетей, которые выявляют скрытые закономерности и паттерны в поведении пользователей. Эти модели помогают предлагать персонализированные рекомендации, основанные на индивидуальных предпочтениях каждого пользователя.

Усовершенствование процессов сбора и обработки данных играет ключевую роль в повышении качества и эффективности рекомендательной системы на веб-платформе, что способствует улучшению пользовательского опыта и развитию функционала платформы.

Для создания модели анализа изображений построили нейронную сеть, состоящую из сверточных слоев (Conv2D), активационной функции ReLU, пулинговых слоев (MaxPooling2D) и полносвязных слоев (Dense).

Сверточные слои применяют фильтры к исходному изображению для извлечения различных признаков. Фильтры перемещаются по изображению, вычисляя скалярное произведение между пикселями изображения и весами фильтра. Это позволяет выявить различные характеристики, такие как грани, текстуры и образцы.

Математически сверточный слой можно выразить следующим образом: [7]

$$Conv(I) = \sum_{i=1}^n (I * W_i) + b_i, \quad (1)$$

где I - входное изображение, W_i - веса фильтра, b_i - смещение, и N - количество фильтров.

ReLU (Rectified Linear Unit) применяется после сверточных слоев для добавления нелинейности в сеть. Она просто пропускает положительные значения через себя, оставляя отрицательные нулями. Это помогает сети улавливать сложные зависимости в данных.

Математически ReLU выглядит следующим образом:

$$ReLU(x) = \max(0, x), \quad (2)$$

Пулинговые слои уменьшают пространственные размеры предыдущего слоя, уменьшая количество параметров и вычислений в сети. MaxPooling2D выбирает максимальное значение из небольшого окна (обычно 2x2) и уменьшает размерность предыдущего слоя.

Формула MaxPooling2D:

$$\text{MaxPooling}(I) = \max(\text{window}), \quad (3)$$

Полносвязные слои принимают высокоуровневые признаки, извлеченные из предыдущих слоев, и объединяют их для классификации или регрессии. В этом слое каждый нейрон соединен со всеми нейронами предыдущего слоя.

Формула полносвязного слоя:

$$\text{Dense}(I) = \text{activation}(I * W + b), \quad (4)$$

где I - входной вектор, W - матрица весов, b - вектор смещений, и activation - активационная функция, обычно используется ReLU или сигмоида.

Самая очевидная и часто обсуждаемая проблема — это вопросы приватности. Сверточные нейросети, обученные на огромных массивах данных, могут включать в себя и чувствительные данные, такие, как личные изображения, которые пользователи не предназначали для анализа. Для решения этой проблемы необходимо разработать строгие политики по обработке и защите персональных данных, включая анонимизацию данных и разработку механизмов для соблюдения законодательства, например, GDPR в Европейском Союзе.

Еще одна значимая проблема связана с предвзятостью в обучающих данных, которая может привести к неравномерному качеству услуг среди различных демографических групп. Например, если данные для тренировки нейросетей недостаточно разнообразны, результаты анализа могут быть менее точными для некоторых групп пользователей. Это требует внедрения методов проверки и корректировки обучающих наборов данных для гарантии их разнообразия и представительности.

Проблемы интерпретируемости результатов работы сверточных нейросетей также играют важную роль. Важно, чтобы пользователи и разработчики могли понимать, как ИИ делает те или иные выводы. Это обеспечивает не только доверие к системе, но и возможность корректировать и улучшать алгоритмы, а также контролировать за соблюдением этических норм. Решением может служить создание более прозрачных моделей и разработка методик, обеспечивающих объяснимость их решений.[8]

Разработка проекта

Укрупненная архитектура разработанной веб-платформы приведена на рисунке 1.

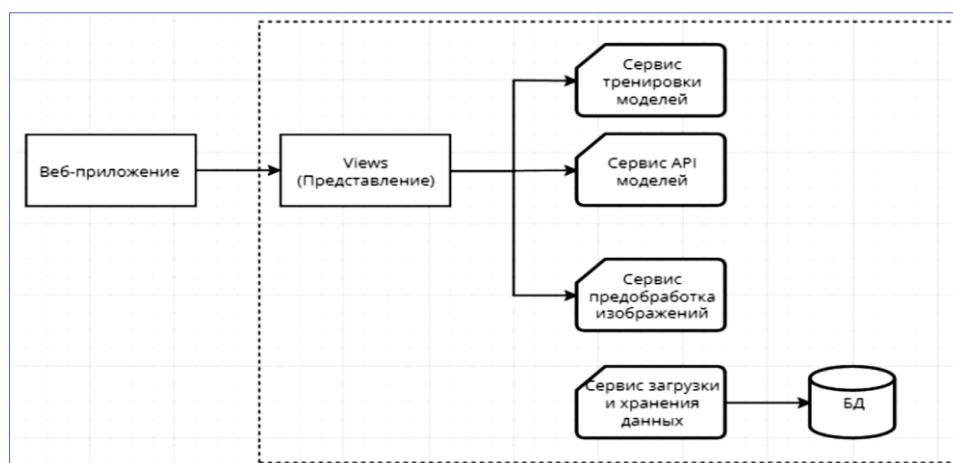


Рисунок 1 – Архитектура разработанной веб-платформы

Для разработки проекта по использованию сверточных нейросетей для анализа визуального контента веб-платформы были выбраны подходящие инструменты для разработки: СУБД PostgreSQL, язык программирования Python, библиотеки TensorFlow и Pillow для Python.

Python был выбран в качестве языка программирования, так как он является лидером в области машинного обучения и искусственного интеллекта. Этот язык отличается лаконичностью и читабельностью синтаксиса, что значительно упрощает процесс разработки и позволяет легко поддерживать код. Python предлагает богатый набор библиотек и фреймворков, которые используют для реализации сложных алгоритмов машинного обучения. Кроме того, активное сообщество разработчиков и обширная база знаний помогают быстро находить решения возникающих проблем [9].

TensorFlow выбран в качестве основной библиотеки для разработки и обучения нейронных сетей. Этот фреймворк предоставляет обширные возможности для разработки сверточных нейронных сетей и поддерживается Google, что гарантирует постоянное развитие и совершенствование. TensorFlow гибок и масштабируем, что позволяет применять его как для небольших экспериментов, так и для крупных производственных приложений.

Для предварительного анализа изображений была выбрана библиотека Pillow. Она позволяет легко обрабатывать и модифицировать изображения, что является важным этапом перед передачей данных в нейронные сети. Простота использования и поддержка множества форматов изображений делают Pillow привлекательным инструментом для предварительной обработки визуальных данных.

СУБД PostgreSQL, мощная объектно-реляционная система управления базами данных, используется для хранения и управления данными, необходимыми для машинного обучения. Эта СУБД известна своей надежностью и производительностью, что обеспечивает безопасное и эффективное хранение больших объемов данных. PostgreSQL поддерживает сложные запросы, индексы и транзакции, что облегчает управление данными и извлечение необходимой информации. Широкая поддержка и документация делают PostgreSQL удобным выбором для долгосрочных проектов.

Перспективы дальнейших исследований и разработок

Применение и развитие сверточных нейросетей (CNN) для анализа визуального контента веб-платформ является перспективным и динамично развивающимся направлением в области искусственного интеллекта. Однако, несмотря на значительный прогресс, существует ряд направлений для дальнейших исследований и улучшений. Ниже представлены ключевые аспекты, которые могут быть реализованы в будущем:

1. Интеграция с NLP для улучшения контекстного анализа. Сверточные нейросети могут быть интегрированы с технологиями обработки естественного языка (Natural Language Processing, NLP), что позволит улучшить понимание контекста изображений и видео. Например, анализ текстовых описаний видео или комментариев может помочь в более точной категоризации и аннотации визуального контента [10,11].

2. Расширенные рекомендательные системы. Использование CNN для анализа визуального контента может быть углублено путем интеграции с рекомендательными системами, что позволит предлагать пользовательский контент на основе визуальных предпочтений. Например, системы могут анализировать визуальные паттерны в предпочтениях пользователя и предлагать схожий контент.

3. Улучшение интерактивности в реальном времени. С развитием технологий обработки изображений в реальном времени, можно развивать методы, позволяющие пользователям взаимодействовать с визуальным контентом более активно. Например, можно разрабатывать функции, которые позволят пользователям "вырезать" интересующие объекты из видео-стримов для дополнительного анализа или манипуляций.

4. Кросс-модальные исследования. Изучение связей между визуальным контентом и другими модальностями, такими как аудио, текст или даже тактильные данные, может привести к созданию более комплексных мультимодальных систем. Эти системы будут лучше понимать и обрабатывать смешанный контент, что особенно важно для медиа и развлекательных платформ.

Каждое из этих направлений представляет собой обширное поле для исследований и инноваций, которое может значительно улучшить способы анализа визуального контента на веб-платформах и повысить пользовательский опыт на новом качественном уровне.

Выводы

Внедрение сверточных нейронных сетей в разработку веб-ориентированных информационных систем общественной платформы для обмена идеями является весьма обоснованным. Система, основанная на данных технологиях, демонстрирует улучшенные функциональные возможности по обработке больших объемов информации, что обеспечивает более глубокое понимание потребностей и интересов пользователей.

Таким образом, интеграция сверточных нейронных сетей в сферу веб-ориентированных информационных систем является весьма перспективным направлением, обеспечивая более высокую степень интерактивности и удовлетворенности пользователей, что подчеркивает необходимость продолжения исследований и разработок в этой области для дальнейшего улучшения коммуникационных процессов технологических платформ.

Литература

1. Искусственный интеллект в 2024: отчет AI Index [Электронный ресурс] \Режим доступа: <https://itentika.ru/news/iskusstvennyj-intellekt-2024>. Дата обращения: 10.05.24
2. «Яндекс» оценил эффект от внедрения ИИ в экономике [Электронный ресурс] /Режим доступа: https://www.rbc.ru/technology_and_media/06/12/2023/657059ee9a7947f8137408a4. Дата обращения: 12.05.24.
3. Искусственный интеллект на основе нейронных сетей [Электронный ресурс] /Режим доступа: file:///C:/Users/mi_sh/Downloads/iskusstvennyj-intellekt-na-osnove-neyronnyh-setey.pdf. Дата обращения: 12.05.24.
4. Учебник по машинному обучению, сверточные нейронные сети [Электронный ресурс] /Режим доступа: <https://education.yandex.ru/handbook/ml/article/svyortochnye-nejroseti> Дата обращения: 12.05.24.
5. Вовченко, В. О. Формирование датасета для решения задач машинного обучения / В. О. Вовченко, В. А. Светличная, Н. К. Андриевская // Информатика и кибернетика. – 2023. – № 2(32). – С. 5-12.
6. Андриевская, Н. К. Разработка алгоритмов предобработки информации для прогнозных моделей ИСППР управления закупками / Н. К. Андриевская, Т. В. Мартыненко // Информатика и кибернетика. – 2023. – № 3(33). – С. 11-18.
7. Сверточные нейронные сети. Глубокие остаточные сети [Электронный ресурс] /Режим доступа: <https://dzen.ru/a/ZPHDg7c8jSZx141>. Дата обращения: 12.05.24.
8. Разработка сверточной нейросети [Электронный ресурс] /Режим доступа: https://hpc-education.unn.ru/files/courses/intel-neon-course/Rus/Lectures/Presentations/4_CNN.pdf Дата обращения: 10.05.24.
9. Архитектура сверточных нейронных сетей [Электронный ресурс] /Режим доступа: <https://cyberleninka.ru/article/n/arhitektura-svyortochnyh-neyronnyh-setey/viewer>. Дата обращения: 15.05.24.
10. NLP для улучшения результатов поиска [Электронный ресурс] /Режим доступа: <https://svetsova.com/blog/> Дата обращения: 12.05.24.
11. Вовченко, В. О. Структурно-функциональная модель процесса анализа рекламаций / В. О. Вовченко, В. А. Светличная // Информатика, управляющие системы, математическое и компьютерное моделирование (ИУСМКМ-2022) : Материалы XIII Международной научно-технической конференции в рамках VIII Международного Научного форума Донецкой Народной Республики, Донецк, 25–26 мая 2022 года. – Донецк: Донецкий национальный технический университет, 2022. – С. 202-207.