

Секция «Нейронные сети в научной работе студента (совместно с Сибирским государственным университетом науки и технологий имени М.Ф. Решетнева)»

Нейронная сеть DeepSeek и ее применение

Научный руководитель – Зуева Анна Сергеевна

Кулешова Юлия Игоревна

Студент (бакалавр)

Московский государственный университет имени М.В.Ломоносова, Высшая школа
государственного аудита, Москва, Россия

E-mail: uliakulesova504@gmail.com

Аннотация: В данной статье рассматривается нейронная сеть DeepSeek, ее создание, основные черты и преимущества использования этой нейронной сети. Также рассматривается актуальность и перспективы развития данной технологии.

Ключевые слова: DeepSeek, нейронная сеть, искусственный интеллект, анализ данных, обучение, модель, Российский рынок.

Нейронная сеть DeepSeek разработана китайской компанией DeepSeek AI. Идея возникла еще в 2020 году в ответ на существующие модели, такие как GPT и BERT, которые хотя и мощные, но имеют ограничения в плане гибкости и скорости обучения. Первый прототип был представлен в 2022 году, а уже к 2025 году компания представила официальную нейросеть, которая доступна как в веб-версии, так и в приложении. Таким образом, DeepSeek показала феноменальные результаты по производительности, при этом затратив всего 5,6 млн. долларов, что, по сравнению с затратами OpenAI и других корпораций в сфере ИИ, является очень маленькой суммой, что подтверждают данные Statista. При этом DeepSeek публикует модель в open-source режиме и стоимость тарифов на использование 1 млн. токенов составляет 0,14\$, в то же время Open AI стоит 7,5\$.

DeepSeek основана на нейронных сетях, таких как Deep Neural Networks, DNN, что позволяет создать внутри нее несколько слоев, которые будут обеспечивать лучшую обработку информации. Архитектура данного искусственного интеллекта включает в себя следующие слои:

- Рекуррентные слои, которые позволяют нейросети обрабатывать последовательные данные, например временные, исходя из контекста.
- Механизмы внимания, позволяющие модели фокусироваться на важных деталях входных данных
- Сверточные слои, обеспечивающие анализ изображений, структурированных данных и т.п.
- Трансформеры, которые способствуют быстрому обучению модели.

Тем самым, данная модель искусственного интеллекта работает на основе следующих принципов:

- 1) Сбор данных
- 2) Предобработка данных. Модель преобразует полученные данные в более удобный формат для обучения.
- 3) Обучение модели. В процессе обучения модель использует такие методы, как метод обратного распространения ошибки и оптимизации.

- 4) Тестирование и валидация. После обучения модель тестируется на новых данных для оценки ее точности и надежности.

Чем же данная нейронная сеть отличается от других не менее популярных моделей ИИ, таких как OpenAI (GPT-3.5 и GPT-4)? Как было выше сказано, DeepSeek более экономичный в использовании, так как тарифы более доступны для бизнеса и частных пользователей. Также DeepSeek обладает более гибкими возможностями для настройки под конкретные задачи, так называемый “fine-tuning”. OpenAI больше ориентирована на облачные сервисы, в то время как DeepSeek лучше адаптирована для интеграции с локальными системами и корпоративными решениями. Наконец, DeepSeek обладает более высокой скоростью обработки данных из-за менее сложной системы. DeepSeek может предлагать более простой и доступный API для разработчиков, с меньшим количеством ограничений. В конце концов, рассматриваемая модель лучше оптимизирована для работы на определенных языках, таких как русский и китайский, что актуально для локальных рынков России и Китая соответственно. Таким образом, можно отметить, что DeepSeek является более гибкой и удобной для применения для частных пользователей, что и обеспечивает привлекательность данной нейросети на рынке ИИ.

Что касается областей применения данной нейросети, то сейчас DeepSeek используется во многих сферах. Например, она используется для анализа текстов, документов, переводов – т.е. обработки естественного языка (NLP). Также модель применяется для прогнозирования временных рядов, анализа больших данных и выявления закономерностей. В области автоматизации задач DeepSeek помогает в обработке запросов в службе поддержки или управление ресурсами. Надо сказать, что и в медицине DeepSeek используется для анализа изображений и прогнозирования заболеваний. Таким образом, DeepSeek обладает широким полем применения, и, учитывая перспективы ее развития, такие как увеличение точности моделей, снижение затрат на обучение, интеграция с другими технологиями и другие, она сможет быть доступна и для остальных сфер деятельности.

Применительно к российскому рынку, актуальность DeepSeek стремительно растет. К 2025 году нейросеть уже широко используется в таких сферах, как:

1. Экономика и бизнес

DeepSeek активно используется в бизнесе для оптимизации процессов и повышения эффективности. Например:

- Финансовый сектор: Банки и страховые компании применяют DeepSeek для анализа рисков, прогнозирования рынков и автоматизации обработки клиентских запросов.
- Ритейл: Модель помогает анализировать потребительское поведение, оптимизировать цепочки поставок и персонализировать маркетинговые кампании.
- Промышленность: В промышленности DeepSeek используется для прогнозирования сбоев оборудования и оптимизации производственных процессов.

2. Наука и образование

Российские научные институты и университеты активно внедряют DeepSeek в исследовательские проекты. Модель применяется для:

- Анализа больших данных: В таких областях, как климатология, генетика и физика, где требуется обработка огромных объемов информации.
- Разработки образовательных программ: DeepSeek помогает создавать персонализированные курсы и автоматизировать проверку заданий.
- Исследований в области ИИ: Российские ученые используют DeepSeek для разработки новых алгоритмов и улучшения существующих моделей.

3. Государственное управление и социальная сфера

- Управление городской инфраструктурой: Модель помогает анализировать данные с датчиков и оптимизировать транспортные потоки, энергопотребление и логистику.

- Здравоохранение: В России DeepSeek используется для анализа медицинских данных, диагностики заболеваний и прогнозирования эпидемий.

- Социальные программы: Модель помогает анализировать данные о населении, чтобы улучшать качество социальных услуг и поддерживать vulnerable группы.

Таким образом, из-за сложившейся экономической ситуации, российский рынок стремится к независимости, а DeepSeek, благодаря своей гибкости и открытости, способствует развитию отечественных решений в области ИИ. Ее адаптивность позволяет использовать модель для решения специфических задач, характерных для российского контекста, таких как обработка текстов на русском языке или анализ данных в условиях низкой цифровизации.

В заключение, нейронная сеть DeepSeek представляет собой мощный инструмент для решения широкого спектра задач в области искусственного интеллекта. Ее архитектура, основанная на глубоких нейронных сетях и механизмах внимания, обеспечивает высокую точность и эффективность. Применение DeepSeek в различных сферах, таких как NLP, анализ данных и автоматизация, демонстрирует ее универсальность и потенциал для дальнейшего развития. Кроме того, DeepSeek активно используется российскими стартапами, которые разрабатывают инновационные продукты на основе ИИ. Благодаря своей доступности и простоте интеграции, модель становится основой для создания новых приложений в области медицины, образования, логистики и других сфер.

Источники и литература

- 1) А.А. Садковкин, М.В. Зорина. Распределение экономического потенциала между странами на рынке искусственного интеллекта.
- 2) The Extreme Cost of Training AI Models // Statista. – [Электронный ресурс]. – Режим доступа: <https://www.statista.com/chart/33114/estimated-cost-of-training-selected-ai-models/>
- 3) A shocking Chinese AI advancement called DeepSeek is sending US stocks plunging // CNN. – [Электронный ресурс]. – Режим доступа: <https://edition.cnn.com/2025/01/27/tech/deepseek-stocks-ai-china/index.html>
- 4) Официальная документация DeepSeek: <https://deepseek.com>