

Гришин Николай Артемович

студент

Дедов Станислав Витальевич

студент

Научный руководитель

Пронина Наталья Андреевна

канд. пед. наук, доцент, доцент

ФГБОУ ВО «Тульский государственный

педагогический университет им. Л.Н. Толстого»

г. Тула, Тульская область

НЕЙРОСЕТИ-МАНИПУЛЯТОРЫ: КАК АЛГОРИТМЫ УПРАВЛЯЮТ ПОВЕДЕНИЕМ И ПОДРЫВАЮТ ПСИХИКУ

Аннотация: современные технологии искусственного интеллекта, особенно нейронные сети, открыли новую эру в создании реалистичного мультимедийного контента. Однако их стремительное развитие порождает серьезные риски. В статье анализируются ключевые угрозы: манипуляция потребительским выбором через персонализированную рекламу, усиление социальной поляризации из-за «информационных эхо-камер», опасности дипфейков и целенаправленной дезинформации.

Ключевые слова: искусственный интеллект, нейросеть, психологическая безопасность.

Развитие технологий искусственного интеллекта (ИИ), в частности искусственных нейронных сетей (ИНС), представляет собой закономерный этап эволюции компьютерных технологий [1, с. 52]. Начиная с третьего десятилетия XXI века наблюдается экспоненциальный рост их возможностей. За сравнительно короткий промежуток времени ИНС добились впечатляющих успехов в создании мультимедийного контента. От генерации текстовых материалов, которые по стилистике неотличимы от созданных человеком, до производства фото-видео и

аудиоконтента с невероятной реалистичностью. Данный технологический прогресс порождает ряд серьезных социокультурных и этических рисков. Ключевая проблема заключается в том, что, обладая способностью к имитации человеческого творчества, нейронные сети могут производить контент, содержащий недостоверную информацию или деструктивные нарративы. Это создает предпосылки для манипуляции как сознанием отдельного человека, так и общественным сознанием.

Доминирующей сферой применения технологий искусственного интеллекта (ИИ) в настоящее время является коммерческий сектор, в частности в области рекламы [9, с. 62]. В условиях постиндустриального общества, где ключевым фактором экономического успеха выступает не столько качество товара, сколько эффективность его продвижения, алгоритмы ИИ приобретают решающее значение. Они позволяют осуществлять персонализированное воздействие на потребителя путем анализа больших данных и формирования индивидуального рекламного контента. Однако данная функциональность порождает серьезную социально-гуманитарную проблему. Создавая персонализированную информационную среду, ИИ не просто отражает предпочтения пользователя, но и активно формирует их, целенаправленно направляя его интересы в определенные рамки. Это подрывает базовые принципы свободного потребительского выбора. Указанный риск обуславливает необходимость разработки этических и правовых норм, регламентирующих применение ИИ в маркетинговых коммуникациях с целью защиты когнитивного суверенитета потребителя.

Второй значимой проблемой, связанной с применением искусственного интеллекта (ИИ), выступает усиление социальной поляризации и формирование так называемых «информационных эхо-камер» [2, с. 272]. Механизм данной проблемы обусловлен действием алгоритмов персонализации, которые формируют вокруг пользователя индивидуальное информационное пространство. Отбирая контент, соответствующий сложившимся предпочтениям и убеждениям индивида, система создаёт однородный информационный поток. Следствием этого

2 <https://phsreda.com>

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

является усиление когнитивных искажений, в частности склонности к подтверждению своей точки зрения. Индивид систематически сталкивается лишь с информацией, укрепляющей его существующие установки, что минимизирует возможность столкновения с альтернативными точками зрения. Данная тенденция ведёт к фрагментации общественного дискурса, радикализации мнений и росту социальной напряженности. Исходя из этого алгоритмическая персонализация новостных потоков, осуществляемая системами ИИ, представляет собой существенный вызов для формирования целостного информационного пространства и конструктивного диалога, требуя разработки мер по повышению информационной прозрачности и медиаграмотности.

Третьей критически важной проблемой, связанной с развитием искусственного интеллекта, является его применение для создания дипфейков (deepfake) [5, с. 2]. Сущность данной угрозы заключается в создании мощного инструментария для атак социальной инженерии. Дипфейки активно используются в мошеннических схемах, шантаже и целенаправленной компрометации лиц. Особую опасность представляет несанкционированное использование биометрических данных пользователей (изображения лица, голоса) для генерации компрометирующего контента, что ведет к уничтожению их цифровой репутации [6, с. 98]. Это порождает состояние хронической цифровой уязвимости и психологического дискомфорта у индивида, который осознает, что его собственная биометрическая идентичность более не принадлежит ему полностью и может быть использована сторонними лицами. Учитывая, что масштабное использование дипфейков способно подрывать общественное доверие и угрожать национальной безопасности, данная проблема становится предметом регулирования на законодательном уровне. Логическим продолжением этой угрозы, подрывающей доверие на индивидуальном уровне, является четвертая значимая проблема – использование ИИ для генерации целенаправленной дезинформации и фейковых новостей в массовом масштабе [7, с. 34]. Если дипфейки наносят точечный удар по репутации человека, то тексты и мультимедиа оказывают массированное воздействие на общественное сознание. Современные генеративные модели достигли

уровня развития, при котором они способны продуцировать контент, обладающий высокими показателями правдоподобия при поверхностном восприятии. Хотя критический анализ позволяет выявить внутренние противоречия в подобных материалах, определенные социальные группы остаются уязвимыми к их воздействию. Ключевая общественная опасность заключается в том, что синтезированная дезинформация может использоваться для целенаправленного манипулирования общественным мнением, формирования заданных настроений и провоцирования социальной напряженности. В данной ситуации основной стратегией противодействия для индивида является развитие медиаграмотности. Однако, учитывая масштаб и потенциальные последствия угрозы, данного уровня защиты недостаточно, что актуализирует необходимость разработки комплексных систем распознавания искусственного контента на технологическом и нормативно-правовом уровнях.

Пятой значимой проблемой выступает недостаточный контроль за контентом, генерируемым искусственным интеллектом, что создает серьезные риски для психического развития несовершеннолетних [8, с. 1024]. Алгоритмы ИИ способны производить непроверенный мультимедийный контент, зачастую содержащий сцены насилия и деструктивные элементы. Широкая доступность подобных материалов в открытом цифровом пространстве, включая социальные сети и видеохостинги, усугубляет их потенциальное воздействие. Систематическое потребление такого контента детской аудиторией может провоцировать повышенную тревожность, деформацию поведенческих норм и причинять значительный психологический ущерб. Соответственно отсутствие эффективных механизмов цензуры и фильтрации контента, созданного ИИ, представляет собой угрозу, требующую разработки превентивных мер на технологическом и регуляторном уровнях.

Подводя итог, можно констатировать, что технологии искусственного интеллекта (ИИ) обладают значительным трансформационным потенциалом в сфере человеческой деятельности. Важно подчеркнуть, что ошибочно рассматривать ИИ исключительно как деструктивную силу; его применение в социальном

4 <https://phsreda.com>

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

ориентированных областях демонстрирует выраженный позитивный эффект. В качестве ключевых благоприятных направлений можно выделить:

- психологическую поддержку: Алгоритмы ИИ лежат в основе чат-ботов и виртуальных ассистентов, способных оказывать когнитивно-поведенческую поддержку. Данные системы представляют особую ценность для лиц, страдающих тревожными и депрессивными расстройствами, в условиях ограниченной доступности традиционной психологической помощи;
- анализ и диагностику: Современные нейросетевые модели показывают высокую эффективность в анализе поведенческих паттернов и речевых маркеров для выявления психических отклонений. Это позволяет осуществлять раннюю диагностику заболеваний и повышать эффективность терапевтических интервенций при работе специалистов [3, с. 2–3]. В конечном итоге воплощение потенциала искусственного интеллекта определяется не самим свойством технологии, а социокультурными и этическими установками человеческого общества.

Список литературы

1. Ефимова С.А. Развитие искусственного интеллекта / С.А. Ефимова // Цифровая наука. – 2020. – №6. – С. 49–57. EDN ZLKAHF
2. Жуци Я. Применение машинного обучения и искусственного интеллекта для анализа медиапотребления аудитории: возможности и вызовы / Я. Жуци // Управление образованием: теория и практика. – 2024. – Т. 14. №5–1. – С. 269–276.
3. Игнатьева И.А. Искусственный интеллект и нейросети как фактор риска и ресурс в обеспечении психологической безопасности / И.А. Игнатьева // Психологически безопасная образовательная среда: проблемы проектирования и перспективы развития: материалы VI Международной научно-практической конференции. – Чебоксары: Среда, 2024. – С. 178–179. – EDN EMPMZH
4. Генеративный искусственный интеллект в образовании: дискуссии и прогнозы / Л.В. Константинова, В.В. Ворожихин, А.М. Петров, Е.С. Титова [и др.] // Открытое образование. – 2023. – №2. – С. 12–21. DOI 10.21686/1818-4243-2023-2-36-48. EDN VPMIZK

5. Лыкова А.Б. Генеративный искусственный интеллект и дипфейк / А.Б. Лыкова, А.А. Хохлов, Л.И. Никонорова // Наука и образование. – 2024. – Т. 7. №2.
6. Мосечкин И.Н. Дипфейк-технологии и биометрические данные: направления уголовно-правового регулирования / И.Н. Мосечкин // Вестник Санкт-Петербургского университета. Право. – 2025. – Т. 16. №1. – С. 95–110. DOI 10.21638/spbu14.2025.107. EDN QGPADN
7. Мун Д.В. «From fake to deepfake»: угрозы и риски развития и распространения технологий искажения реальности в глобальном информационном пространстве / Д.В. Мун, В.В. Попета // Культура: теория и практика. – 2020. – №1 (34). – С. 32–43.
8. Ходаев А.С. Искусственный интеллект и социализация подростков: риски влияния / А.С. Ходаев, Л.Н. Макарова // Вестник Тамбовского университета. Серия: Гуманитарные науки. – 2024. – Т. 29. №4. – С. 1011–1021. – DOI 10.20310/1810-0201-2024-29-4-1011-1021. – EDN RMUFXY
9. Использование возможностей искусственного интеллекта в рекламе / Д.А. Шевченко, Е.М. Крюкова, В.В. Зеленов, В.В. Галстян // Практический маркетинг. – 2024. – №1. – С. 60–64. – DOI 10.24412/2071-3762-2024-1319-60-64. – EDN INPYMK