

Date: 9th December-2025

СТРУКТУРА МОДЕЛИ ОБНАРУЖЕНИЯ ФЕЙКОВОГО КОНТЕНТА С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Ташкентский университет информационных технологий имени Мухаммада
ал-Хоразмий

Доцент кафедры «кибербезопасность и криминалистика»

Юлдашева Нафиса Салимовна

Ташкентский университет информационных технологий имени Мухаммада ал-
Хоразмий

Магистрант кафедры «кибербезопасность и криминалистика»

Сайжанов Исмаил Абатович

Аннотация: Современное информационное пространство характеризуется высокой скоростью распространения данных и большим объемом контента, создаваемого пользователями. На этом фоне растет количество фейковых новостей, которые формируют искаженную картину мира, влияют на общественное мнение и подрывают доверие к медиа. В работе рассматриваются подходы к автоматическому обнаружению фейковой информации с использованием технологий искусственного интеллекта, включая методы машинного обучения, глубокие нейронные сети и обработку естественного языка. Представлена структура будущей модели, определены этапы ее разработки и ключевые параметры эффективности.

Ключевые слова: фейковые новости, искусственный интеллект, машинное обучение, классификация текста, нейронные сети, обработка данных

ВВЕДЕНИЕ

Проблема распространения фейкового контента в условиях стремительно развивающихся цифровых технологий приобретает все более серьезный характер. Недостоверная информация проникает практически во все сферы человеческой деятельности — от политики и экономики до науки, здравоохранения и образования. Это делает необходимым детальное изучение феномена фейковой информации, выявление факторов, способствующих её стремительному распространению в социальных сетях, а также анализ особенностей восприятия ложных материалов пользователями.

В цифровой среде фейковый контент становится одной из наиболее значимых угроз, поскольку включает в себя широкий набор манипулятивных, искаженных или полностью ложных сведений, распространяемых через интернет-ресурсы и медийные платформы. Его характерными признаками выступают отсутствие достоверных источников, недостаточная доказательная база, логические ошибки, эмоционально окрашенные и вводящие в заблуждение заголовки. Подобная информация способна оказывать негативное влияние на общественные процессы, усиливать социальное напряжение, провоцировать политические споры и снижать уровень доверия к традиционным медиа.



Date: 9th December-2025



В связи с этим анализ научных подходов и методов выявления фейкового контента приобретает особую значимость. Ученые и специалисты самых различных областей — включая компьютерные науки, обработку естественного языка, социальную аналитику и журналистику — активно развивают технологии, направленные на автоматическое обнаружение ложной информации. Цель таких исследований заключается в создании инструментов, способных эффективно противостоять нарастающему потоку недостоверных сообщений и обеспечивать более высокую информационную безопасность общества.

СТРУКТУРА МОДЕЛИ ОБНАРУЖЕНИЯ ФЕЙКОВОГО КОНТЕНТА

Одним из ключевых направлений исследований является использование алгоритмов машинного обучения и искусственного интеллекта для автоматического обнаружения фейкового контента. В этом контексте разработка и применение моделей глубокого обучения, таких как нейронные сети. Они могут анализировать текстовые данные, изображения и видео, учитывая лингвистические и визуальные признаки для обнаружения подозрительного контента.

Научные исследования и разработки в области обнаружения фейкового контента способствуют углублению знаний по этой проблеме и предоставляют инструменты для эффективного противодействия ей. Использование понимания механизмов распространения фейкового контента и разработка эффективных алгоритмов способны обеспечить большинство пользователей достоверной и надежной информацией. Продолжение исследований в этой области является неотъемлемой составляющей постоянного совершенствования и разработки новых методов.

Первым этапом использования моделей машинного обучения для распознавания фейкового контента является тщательный выбор соответствующего набора данных, который будет использоваться для тренировки и тестирования этих моделей. Важно, чтобы набор данных был репрезентативным и отражал разнообразие фейкового контента, с которым можно встретиться в реальном мире. Дополнительным преимуществом может быть наличие меток или оценок для каждого образца данных, указывающих, является ли он фейковым или правдивым. Это позволяет проводить тренировки моделей с учителем, где можно использовать методы обучения, основанные на обозначенных данных.

На следующем этапе было осуществлено объединение двух наборов данных в один.

Следующим этапом после отбора набора данных для проверки моделей машинного обучения процесс тренировки этих моделей. Этот процесс учитывает построение моделей, используя определенные алгоритмы машинного обучения, предназначенные для распознавания фейкового контента. Обычно для тренировки моделей используются методы обучения с учителем, предусматривающие использование обозначенных данных, где известно, является ли контент фейковым или правдивым. Аналогичные подходы могут использоваться для анализа

Date: 9th December-2025

изображений или видео. Получим модель, которая будет научена распознавать фейковый и правдоподобный контент.

После завершения процесса тренировки моделей важным шагом является оценка их точности и эффективности. Для этого используются метрики оценки, такие как точность, восстановление, F-мера и другие. Эти метрики позволяют объективно определить, насколько хорошо модели способны распознавать фейковый контент.

Оценка моделей может быть проведена на уединенном наборе данных, который не был использован во время тренировки. Этот подход позволяет оценить всеобщность моделей и их способность адаптироваться к новым входным данным. Кроме того, может быть использована перекрестная проверка во избежание переобучения моделей и получения более объективных результатов.

После оценки точности моделей может быть осуществлена их настройка и усовершенствование для достижения лучших результатов.

Было также установлено, что для выявления фейкового контента используются различные виды машинного обучения, учитывая классификацию, удостоверения, генеративные модели, кластеризацию и размерность.

Набор данных для тренировки моделей должен содержать как фейковый, так и правдивый контент, чтобы модели могли научиться различать их. Важно, чтобы набор данных был репрезентативным и отражал разнообразие фейкового контента, с которым можно встретиться в реальном мире.

СПИСОК ЛИТЕРАТУРЫ:

1. Акинин, М. В. Нейросетевые системы искусственного интеллекта в задачах обработки изображений / М. В. Акинин, М. Б. Никифоров, А. И. Таганов. - М.: РиС, 2020. - 152 с;
2. Косаренко Н. Н. Система искусственного интеллекта: понятие, теория, право и перспективы развития : монография / Н. Н. Косаренко. - Москва : РУСАЙНС, 2024.- 176 с;
3. Письменский, Г. И. Построение систем искусственного интеллекта на основе дискурса неопределенности сознания / Г. И. Письменский, Ж. Ю. Бакаева. Инновации в образовании. – 2023 – № 7 – 62-69 с;
4. Сидоркина, И. Г. Системы искусственного интеллекта / И. Г. Сидоркина. - М.: КноРус, 2022. - 167 с;
5. Эртель В. Введение в искусственный интеллект / В. Эртель , пер. с англ. А. Горман. - Москва : Эксмо, 2019. – 448 с.