

УДК 004.89

ПОДХОД К ОЦЕНИВАНИЮ ДОСТОВЕРНОСТИ ИНФОРМАЦИИ НА ОСНОВЕ АДАПТАЦИИ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

Мишин А. А. (ВКА им. А. Ф. Можайского), Сабилов Т. Р. (ВКА им. А. Ф. Можайского)

Научный руководитель – кандидат технических наук Сабилов Т. Р.

(ВКА им. А. Ф. Можайского)

Разработан подход, обеспечивающий автоматизированный анализ и оценку достоверности информации из новостных и информационных источников. Предложенная методология позволяет выявлять ключевые факты, классифицировать данные, а также оценивать степень правдивости сообщений с использованием моделей обработки естественного языка и внешних источников верификации. Подход может быть применён для повышения качества аналитических процессов в государственных структурах, СМИ и информационно-аналитических службах.

Введение. В современном информационном пространстве количество новостных сообщений и информационных публикаций стремительно растёт. Однако, значительная часть этих данных может содержать дезинформацию, недостоверные новости или манипулятивные интерпретации фактов. Это создаёт серьёзные трудности для аналитических центров, государственных служб и средств массовой информации. Традиционные методы верификации новостей требуют значительных временных и человеческих ресурсов, что делает их недостаточно эффективными в условиях постоянного увеличения информационных потоков. Разработка автоматизированной системы для анализа и оценки достоверности информации позволит минимизировать влияние дезинформации, повысить качество новостного контента и улучшить принятие решений в СМИ и государственных структурах.

Основная часть. В рамках данной работы рассматривается разработка программного модуля, позволяющего автоматизировать процесс проверки достоверности информации. Для этого решаются следующие задачи:

- 1) Задача сбора и предобработки информации. Автоматизированный сбор данных из различных новостных и информационных источников (агентства, социальные сети, официальные сайты) с последующей очисткой и нормализацией текста. Это позволяет устранить дубликаты, определить источник публикации и выделить первоисточник.
- 2) Задача выделения ключевых фактов и структуры сообщения. Извлечение именованных сущностей (Named Entity Recognition, NER) для идентификации участников событий, мест и временных характеристик. Анализ тематической направленности, выявление эмоциональной окраски и оценка риторики сообщения.
- 3) Задача проверки достоверности информации. Оценка фактов путём их сопоставления с внешними базами данных и агрегированными источниками. Использование больших языковых моделей (LLM) для анализа логических связей и выявления несоответствий. Разработка метода расчёта коэффициента достоверности на основе вероятностных моделей и контекстного анализа.
- 4) Задача классификации и суммаризации. Автоматическая категоризация сообщений по тематикам (политика, экономика, криминал, технологии и т. д.) с последующей генерацией краткого изложения (суммаризации). Это позволяет сократить время на анализ больших объёмов текстовой информации.
- 5) Задача представления и визуализации результатов. Генерация аналитического отчёта с наглядным представлением выявленных закономерностей. Использование графиков, таблиц и метрик достоверности для повышения удобства восприятия информации и поддержки принятия решений.

Для решения поставленных задач был разработан комплексный подход, включающий методы обработки естественного языка, машинного обучения и факт-чекинга. Гибкость

предложенной методологии позволяет адаптировать систему под различные аналитические задачи и интегрировать её в существующие информационные процессы.

Выводы. Разработка предложенного подхода позволит значительно улучшить процесс анализа информационных потоков, повысить оперативность выявления дезинформации и улучшить качество принимаемых решений. Внедрение такой системы в СМИ, государственные структуры и аналитические центры создаст новый уровень автоматизированного мониторинга и верификации новостей, что приведёт к снижению распространения фейковых данных и укреплению доверия к информационным источникам.

Список использованных источников:

1. Обнаружение фейковых новостей с помощью машинного обучения: [Электронный ресурс]. – URL: <https://arxiv.org/abs/2003.07724> (дата обращения: 31.01.2024);
2. Применение NLP в анализе достоверности новостных сообщений: [Электронный ресурс]. – URL: <https://arxiv.org/abs/2106.01712> (дата обращения: 31.01.2024);
3. Алгоритмы классификации и суммаризации текстов на основе трансформеров: [Электронный ресурс]. – URL: <https://arxiv.org/abs/1909.08593> (дата обращения: 31.01.2024);
4. Использование больших языковых моделей (LLM) для верификации новостей: [Электронный ресурс]. – URL: <https://arxiv.org/abs/2301.05678> (дата обращения: 31.01.2024).

Мишин А. А. (автор)

Сабиров Т. Р. (научный руководитель)