

Date: 11th December-2025

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: В современную эпоху стремительного развития цифровых технологий и социальных сетей пользователи получают мгновенный доступ к новостям, при этом не всегда обращая внимание на их достоверность. В результате объём фейковых новостей значительно увеличивается. Фейковый контент представляет собой серьёзную угрозу для общества, поскольку оказывает негативное влияние на политические процессы, экономическую стабильность и социальное взаимодействие. Наиболее интенсивное распространение недостоверная информация получает через социальные сети и другие цифровые платформы.

В данной обзорной статье представлены современные методы обнаружения фейковых новостей, основанные на технологиях машинного обучения и глубокого обучения. Проведён всесторонний анализ существующих подходов, выполнена их сравнительная оценка, а также рассмотрены существующие ограничения, нерешённые задачи и перспективы дальнейших исследований. В обзор включены публикации за период с 2018 по 2025 годы от ведущих научных издателей, таких как IEEE, Intelligent Systems, EMNLP, ACM, Springer, Elsevier, JAIR и других.

Ключевые слова: фейковые новости; обнаружение фейкового контента; машинное обучение; глубокое обучение; точность; инженерия признаков; алгоритмы.

Введение

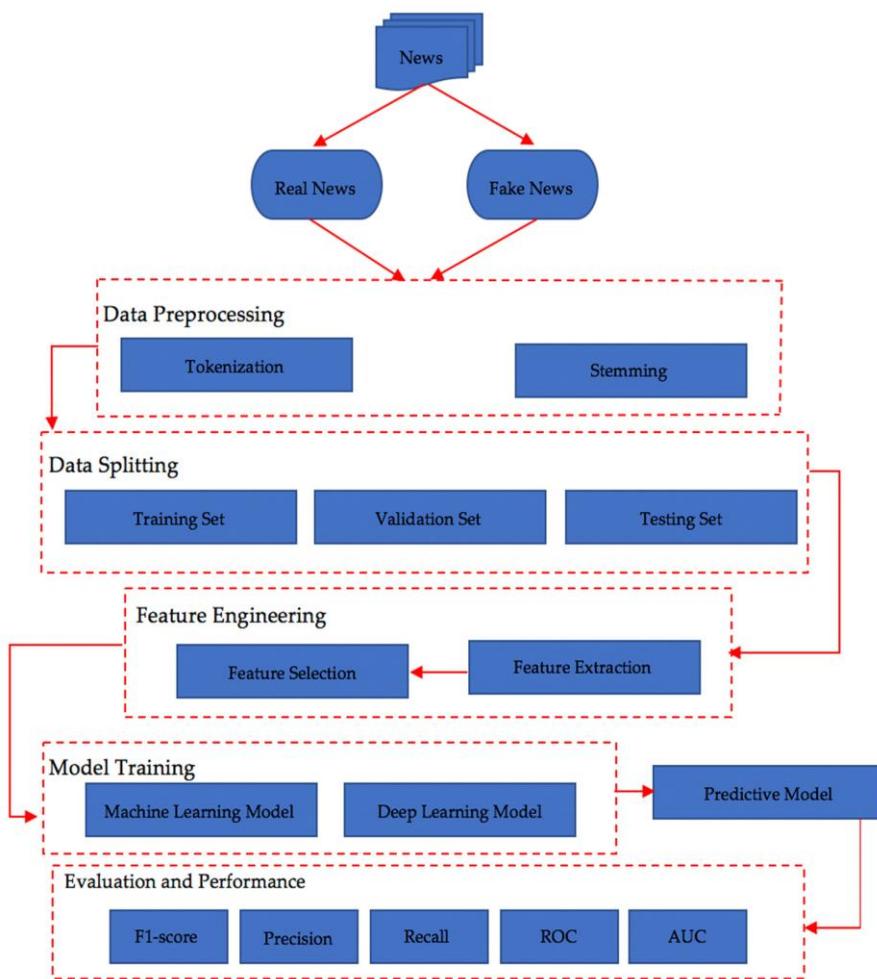
В последнее время мир стал очень динамичным. Поэтому такое быстрое развитие, особенно в цифровом мире, имеет ряд преимуществ и недостатков. Из-за легкости доступа к новостям без проверки их достоверности, распространение фейковых новостей увеличилось. Одним из основных недостатков цифровой эры является быстрое распространение дезинформации. Люди могут непреднамеренно или намеренно распространять фейковые новости, что может нанести вред или дискредитировать других людей или организаций. Кроме того, распространение фейковых новостей может служить инструментом пропаганды против отдельных лиц через различные онлайн-платформы [1–3]. Напротив, алгоритмы машинного обучения и глубокого обучения, которые являются частью искусственного интеллекта, в последнее время используются для обнаружения фейковых новостей

или прогнозирования. Алгоритмы сначала обучаются с помощью обучающего набора данных, который содержит как фейковые новости, так и достоверные новости. После обучения ранее обученные модели проходят валидацию и тестирование. Затем модели развертываются для выполнения других задач, таких как прогнозирование или выявление подсказок, которые помогают идентифицировать фейковые новости [1–5]. Онлайн-платформы уделяют приоритетное внимание доставке новостей удобным, доступным и быстрым способом. Однако эта скорость и легкость доступа также создают больше возможностей для распространения фейковых новостей. В результате отдельные лица и организации предпринимают усилия по проверке и разоблачению ложной информации. Обнаружение фейковых новостей остается серьезной проблемой. Многие исследователи занимаются решением этой проблемы, используя алгоритмы машинного обучения и глубокого обучения, обучая эти модели выявлять фейковый контент. После надлежащего обучения эти алгоритмы могут автоматически обнаруживать фейковые новости с определенной степенью точности [6–8].

Для правильного функционирования классификатора необходимо следить за его точностью в обнаружении фейковых новостей, поскольку неспособность обнаружить фейковые новости может нанести вред разным людям. Ниже приведены некоторые популярные классификаторы, которые используются для этой цели в машинном обучении: naïve Bayes классификатор, методы опорных векторов (SVM), случайные леса, k-ближайших соседей (KNN), деревья решений и логистическая регрессия. Некоторые распространенные алгоритмы глубокого обучения, используемые для этой цели, — это сверточные нейронные сети (CNN), двунаправленные сети с длинной краткосрочной памятью (BI-LSTM), рекуррентные нейронные сети (RNN) и графовые нейронные сети (GNN) [9–16]. На рисунке 1 показана концепция обнаружения фейковых новостей с помощью алгоритмов машинного или глубокого обучения.

Date: 11th December-2025

Рисунок 1. Обнаружение фейковых новостей с помощью алгоритмов машинного или глубокого обучения.



Ответы на вопросы исследования, поставленные в обзоре литературы, будут даны с упором на машинное обучение и глубокое обучение для обнаружения фальшивых новостей. В них также будет рассмотрено, как машинное обучение и глубокое обучение могут быть использованы для обнаружения фальшивых новостей, на основе анализа соответствующих работ в литературе.[\[17-21\]](#).

Машинное обучение

Aphiwongsophon и Chongstitvatana [\[1\]](#) использовали три метода машинного обучения для обнаружения фальшивых новостей: naïve Bayes алгоритм, нейронные сети и методы опорных векторов (SVM). Кроме того, с помощью API Twitter они извлекли двадцать две характеристики. В результате naïve Bayes алгоритм достиг точности более 96%, а нейронные сети и SVM - 99,90%.

В этом исследовании были использованы методы обработки естественного языка (NLP) для отличия реальных новостей от «фейковых новостей», которые поступают из ненадежных источников. Авторы полагались на построение модели на основе вектора подсчета (с использованием статистики слов) или матрицы TF-IDF (частота термина — обратная частота документа) (статистика слов по частоте их

использования в других статьях в данном наборе данных). Однако эти модели не учитывают такие важные характеристики, как организация слов и контекст. Поэтому вероятность того, что две статьи с одинаковым количеством слов могут иметь совершенно разное значение, высока. Набор данных, использованный в этой модели, — это «Fake News Challenge» от Kaggle. Таким образом, в предложенной работе был выполнен предварительный анализ набора данных о фейковых и реальных новостях статей и использован naïve Bayes классификатор для построения бинарной модели на основе слов, чтобы правильно классифицировать новости. В результате была достигнута точность 92,20 % [2].

В исследовании Ni et al. [3] были изучены особенности фейковых новостей с целью выявления резких изменений в контексте новостей с помощью метода сопоставления по коэффициенту склонности (PSM) для извлечения характеристик частоты документов, включающих все переменные, с целью смягчения влияния нежелательных переменных. Экспериментальные данные были взяты из открытого источника FakeNewsNet, который состоит из данных PolitiFact и GossiCop, и результаты показали, что PSM более применим к фейковым новостям, чем только сырой PSM, который также работает лучше, чем использование сырой частоты для выбора характеристик. Они достигли точности 68%. С использованием набора данных PolitiFact были рассмотрены различные классификаторы фейковых новостей, включая логистическую регрессию, случайный лес и машину поддержки векторов, для оценки производительности и наблюдения за улучшениями [3].

Singh и др. [4] сравнили модели ансамблевого обучения для сортировки фейковых новостей путем анализа качества репортажа и установления достоверности новостей. Целью статьи было использование алгоритмов обработки естественного языка (NLP) и машинного обучения (ML) для обнаружения фейковых новостей на основе контекста новостей. В качестве классификаторов они использовали деревья решений, случайный лес, классификацию AdaBoost и XGBoost. В качестве признаков, которые подаются в вышеупомянутые классификаторы, они использовали TF, TF-IDF и встраивание слов. Таким образом, было разработано веб-приложение, чтобы уменьшить трудности, с которыми сталкиваются пользователи при различении фейковых новостей. В данной статье авторы опирались на анализ фейковых новостей как двумерный подход к классификации с использованием характеристик содержания и контекста [8]. Поэтому эксперименты проводились на основе деревовидной ансамблевой структуры машинного обучения (градиентное усиление) с полным моделированием на основе содержания для обнаружения фейковых новостей. Результаты экспериментов продемонстрировали более высокую точность по сравнению с существующими эталонными показателями, при этом алгоритм градиентного усиления (ансамблевая структура машинного обучения) достиг 86% точности в классификации фейковых новостей [8].

Albahr и Albahr [9] исследовали несколько традиционных алгоритмов

машиинного обучения, а именно случайные леса, naïve Bayes алгоритм, нейронные сети и деревья решений, чтобы проверить эффективность классификации при обнаружении фальшивых новостей на основе однограммных, двуграммных и трехграммных признаков. Обучение проводилось на одном из популярных наборов данных, известном как LIAR, и результаты показали, что naïve Bayes алгоритм значительно превосходит свои аналоги, достигая точности 99,0%.

Goldani и др. [10] сосредоточились на использовании капсулльных нейронных сетей в процессе обнаружения фейковых новостей. Были использованы различные модели встраивания с разной длиной. В случае коротких новостей использовались фиксированные встраивания слов и n-граммные характеристики, но для новостей средней или большой длины использовались нефиксированные встраивания слов, поддерживающие прогрессивное обучение. Кроме того, для извлечения характеристик применялись различные уровни n-грамм. Для процесса оценки они были обучены на двух недавно известных наборах данных в этой области, а именно ISOT и LIAR. Исследование продемонстрировало высокую эффективность: новые методы прошли 7,8% на ISOT, достигнув аналогичной эффективности на наборе данных LIAR с более чем 3% на наборе валидации и 1% на наборе тестирования.

Биунда и Деви [13] использовали модель текстовых характеристик подлинных и фальшивых новостных текстов, основанную на уравнении частоты терминов. Для расчета рейтинга достоверности источников они опирались на характеристики URL-адреса веб-сайта и домена верхнего уровня. Комбинируя TF-IDF, site_URL и текстовые характеристики с рейтингом достоверности нескольких источников, они оценивали достоверность новостей. Экспериментальный набор данных, собранный из Kaggle, содержит 2050 новостных статей. Модель была применена к классификаторам машинного обучения (ML) для проверки ее эффективности в обнаружении фальшивых новостей. Результаты эксперимента показали, что предложенная модель достигла максимальной эффективности примерно 99,5%.

Mugdha и др. [14] продемонстрировали модель, способную обнаруживать фальшивые новости на основе заголовков новостей, создав новый набор данных для бенгальского языка. Используя гауссовый naïve Bayes алгоритм, модель достигла приемлемой производительности. Этот алгоритм использовал текстовую характеристику на основе TF-IDF и дополнительный древовидный классификатор для выбора характеристик. Точность достигла 87%, что относительно лучше, чем любой другой алгоритм, используемый в этой модели.

Jardaneh и др. [16] использовали новые функции, связанные с текстом, содержащим мнение пользователей, для обнаружения фейковых новостей на арабском языке. Анализ мнений усовершенствовал процесс прогнозирования. Для обучения классификационных моделей было использовано несколько алгоритмов машинного обучения, в том числе случайные леса, деревья решений, AdaBoost и логистическая регрессия. В результате они продемонстрировали, что

система способна обнаруживать фейковые новости с точностью 76 %.

Tiwari и Jain [22] сравнили несколько алгоритмов машинного обучения, используя классификацию деревьев решений, классификаторы случайных лесов и логистическую регрессию с набором данных HSpam14, содержащий 400 000 твитов и семантические характеристики. Результаты продемонстрировали точность в идентификации выбранных новостей с точностью от 98 до 99 %.

Rampurkar и D.R [23] предварительно обработали входные тексты, чтобы определить их особенности. Для оценки важности слов в каждой статье использовалась концепция TF-IDF. Затем новостные статьи были сегментированы с помощью наивного байесовского алгоритма, чтобы отличить настоящие новости от фальшивых. Набор данных ISOT содержит 23 481 запись. Этот алгоритм рассчитывает вероятность классификации статьи, предполагая, что слово является условно независимым. Затем эффективность используемых алгоритмов была определена с помощью матрицы путаницы для оценки достоверности модели. Результаты показали, что логистическая регрессия хорошо справляется с обнаружением фальшивых новостей с точностью 98,31 %.

Mutri и др. [24] сосредоточились на разработке метода обнаружения фальшивых новостей путем сортировки и анализа прошлых данных с использованием машинного обучения. Были использованы различные методы машинного обучения, в том числе предложенные алгоритмы KNN и SVM как эффективное решение для обнаружения фальшивых новостей. KNN — это алгоритм машинного обучения, который классифицирует тексты на основе близости к известным данным по таким характеристикам, как категория и дата/время. Этот метод был использован из-за его способности обрабатывать нелинейные данные и простоты использования. Применение KNN может повысить эффективность выявления фейковых новостей за счет использования характеристик близлежащего текста. В исследовании, проведенном с использованием набора данных FakeNewsDetection, алгоритм KNN показал лучшие результаты, чем другие модели, достигнув средней абсолютной ошибки (MAE) 0,011, которая измеряет средний размер ложных обнаружений в наборе прогнозов без учета их направления, и среднеквадратичной ошибки (RMSE) 0,077, которая показывает квадратный корень из средней квадратичной разницы между прогнозируемыми и наблюдаемыми результатами данных.

Глубокое обучение

Gereme et al. [6] представили несколько моделей, в том числе модель обнаружения фальшивых новостей на амхарском языке, набор данных на амхарском языке (GPAC), набор данных ETH_FAKE и функции встраивания слов Amharic FastText (AMFTWE). Таким образом, модель, разработанная с использованием набора данных ETH_FAKE, достигла превосходной точности, превышающей 99% при использовании 300- и 200-мерного встраивания.

Обнаружение фальшивых новостей является сложной задачей для многих исследователей, особенно когда новости распространяются через социальные сети. Это помогает выявлять ложные и вводящие в заблуждение истории в социальных сетях. Одной из ключевых проблем в этой области исследований является ограниченная доступность данных для обучения моделей обнаружения. Новый метод автоматического генерации вводящих в заблуждение (и, возможно, фальшивых) новостей на арабском языке был представлен Nagoudi et al. [25]. Были использованы функции маркировки частей речи (POS) и встраивания слов. Для облегчения будущих исследований это требование будет полностью устранено путем предоставления готового к использованию набора данных под названием AraNews. В конечном итоге были разработаны модели для обнаружения фальшивых новостей на арабском языке, достигающие точности более 70 % [25].

Hamed и др. [26] сосредоточились на извлечении признаков, в частности для анализа тональности новостных статей, который включает комментарии пользователей к этим новостям и признаки анализа эмоций. Эти признаки, наряду с признаками содержания новостей, были добавлены в двунаправленную модель краткосрочной памяти для обнаружения фальшивых новостей. Для обучения и тестирования предложенной модели был использован стандартный набор данных Fakeddit с опубликованными заголовками. Точность обнаружения была высокой - 96,77 %, что является самым высоким показателем по сравнению с другими недавними исследованиями. Verma et al. [27] предложили двухэтапную стандартную модель под названием WELFake, основанную на встраивании слов (WE), путем внедрения лингвистических характеристик для обнаружения фейковых новостей с помощью классификации машинного обучения. На первом этапе выполняется предварительная обработка набора данных и проверка содержания новостей с помощью лингвистических характеристик. Второй этап заключается во встраивании наборов лингвистических характеристик с помощью WE и применить классификацию голосования. Чтобы подтвердить эффективность своего подхода, был выбран новый набор данных WELFake, состоящий из почти многих статей, который содержит различные наборы данных для получения объективной классификации. В результате модель WELFake продемонстрировала точность 96,73% в обнаружении фальшивых новостей.

Иванцова и др. [28] сосредоточились на обнаружении фейковых новостей из новостных статей на словацком языке на основе Word2Vec, GloVe и морфологического анализа. Был создан набор данных для обучения моделей на политических новостях. Две архитектуры, CNN и нейронные сети LSTM, были обучены на сгенерированных учебных данных. Первая модель (модель 1) представляла собой CNN, которая достигла общей точности 92,38%. Вторая модель (модель 2) представляла собой рекуррентную нейронную сеть, в которой слой LSTM, содержащий 128 нейронов, питался выходом встраиваемого слоя. Эта модель достигла точности 93,56 % на словацком наборе данных.

Wang et al. [11] представили SemSeq4FD, новую модель нейронной сети на основе графов, предназначенную для раннего обнаружения фейковых новостей с использованием модифицированных текстовых структур. SemSeq4FD использует графы для моделирования глобальных семантических представлений предложений, а глобальные представления предложений обучаются с помощью графовой сверточной сети. Были рассмотрены характеристики предложений с использованием одномерной сверточной сети для обучения внутренних классификаторов предложений с использованием данных SLN и LUN. Для оптимизированных предложений была использована сеть на основе LSTM, которая создавала окончательное представление документа для распознавания фейковых новостей с использованием обучающих данных как на английском, так и на китайском языках. Достигнута точность 92,6%.

Subramanian и др. [29] обнаружили фальшивый контент на малаяламском языке в социальных сетях. Процесс проверки состоит из двух подзадач: первая классифицирует контент как фальшивый или нефальшивый с помощью контекстного встраивания и последовательных характеристик, а во второй подзадаче классификация была расширена до пяти категорий (ложное, наполовину правдивое, в основном ложное, частично ложное и в основном правдивое) с использованием многоязычных характеристик контекстного встраивания. Для первой задачи были использованы методы машинного обучения, такие как SVM, naïve Bayes алгоритм и SGD, а также алгоритмы на основе BERT. Среди этих алгоритмов XLM-RoBERTa достиг высокой производительности в 89,80 %. Для второй задачи были использованы модели, использующие LSTM, GRU, XLM-RoBERTa и SVM. XLM-RoBERTa снова показал хорошие результаты по сравнению с другими алгоритмами, достигнув самого высокого общего показателя F1 в 62,83%.

Jingyuan и др. [30] сосредоточились на улучшении обнаружения графиков за счет значительных усовершенствований языковых моделей, фреймворков и моделей обучения в литературе по фейковым новостям. Опираясь на несколько успешных подходов, будет подчеркнут потенциал обнаружения фейковых новостей в реальном времени на разных платформах. Контекстные и Symantec-особенности были использованы для интеграции знаний по обнаружению дезинформации, обнаружения фейковых новостей с помощью мультимодальных крупных языковых моделей, адаптивного к домену обнаружения фейковых новостей с помощью нескольких примеров и фреймворка обнаружения, не зависящего от стиля. Все эти модели были построены на основе графовых нейронных сетей (GNN). Кроме того, в их эксперименте были использованы наборы данных FakeNewsNet, PolitiFact fact, PAN2020 и COVID-19. Обнаружение фейковых новостей с использованием крупных мультимодальных языковых моделей на наборе данных PolitiFact дало высокую точность 95,10 %.

Tan и Bakir [31] представили модель, основанную на алгоритме трансформатора, который имеет несколько применений для более надежной обработки длинных текстов. В модели был реализован гибридный двунаправленный

блок обработки длинных текстов с алгоритмом трансформатора. Чтобы облегчить идентификацию фальшивых твитов

(TruthSeeker), исследователи добавили в набор данных балансирующий фактор на уровне классов с использованием векторных представлений слов (word embedding). Для улучшения качества прогнозирования был использован алгоритм TomekLinks. Для достижения этой цели был рассмотрен набор параметров и выполнен поиск по сетке для определения параметров, дающих оптимальные результаты. Что касается результатов тестирования, модель продемонстрировала высокую производительность, достигнув точности 99,91%.

Alsuwat, E. и Alsuwat, H. [32] сосредоточили свое внимание на новом предложении по обнаружению фейковых новостей, получившем название «Мультимодальное обнаружение фейковых новостей» (MM-FND). В своих экспериментах они опирались на

трех наборами данных, а именно набором данных о фейковых новостях ISOT, набором данных LIAR и набором данных о фейковых новостях COVID-19. Для генерации признаков они использовали Word2Vec и частоту терминов — обратную частоту документов (TF-IDF) для извлечения временных признаков. Bi-LSTM использовался для извлечения временных признаков с помощью двунаправленных сетей с длинной краткосрочной памятью. Кроме того, пространственные характеристики были извлечены с помощью распознавания именованных сущностей (NER) в сочетании с глобальными векторными вложениями для представления слов (GloVe). Результаты показали, что предложение достигло 96,3% точности при тестировании на наборе данных ISOT. На наборе данных LIAR алгоритм достиг 95,6% точности. На наборе данных о фальшивых новостях COVID-19 алгоритм достиг 97,1% точности.

Машинное обучение и глубокое обучение

Jiang et al. [5] применили два подхода. Во-первых, были оценены пять моделей машинного обучения, а во-вторых, были протестированы три модели глубокого обучения. Для оценки была проведена перекрестная валидация с использованием двух наборов данных о фейковых новостях, имеющих совершенно разные размеры. Кроме того, в качестве входных данных для моделей машинного обучения и глубокого обучения были извлечены характеристики частоты терминов и обратной частоты документов (TF-IDF) и вложения слов. Затем они предложили модель стекирования, которая при тестировании на наборах данных ISOT и KDnugget достигла точности почти 99,95 % и 96 % соответственно.

Pardamean и Pardede [7] работали над выявлением неточных новостей с помощью Bidi-rectional Encoding Representations from Transformers (BERT). BERT — это языковая модель глубокого обучения, которая очень эффективна в обработке языка. Эксперименты показали, что представления с использованием гиперпараметрических характеристик могут достигать точности 99,23% по набору данных Kaggle.

Date: 11th December-2025

Mouratidis и др. [33] провели сравнительный эксперимент по традиционным классификаторам машинного обучения, включая наивные байесовские, SVM и случайные леса, в дополнение к моделям глубокого обучения, таким как CNN, LSTM и BERT. В ходе исследования были сгенерированы характеристики, включая TF-IDF, Word2Vec и контекстные вложения. Кроме того, они провели различные тесты на основе нескольких наборов данных. Исследователи обнаружили, что модели на основе BERT демонстрируют высокую производительность, что выражается в повышении точности обнаружения фейковых новостей. При применении алгоритма BERT они достигли производительности 98,40 %.

Al-Tarawneh и др. [34] обнаружили, что TF-IDF может потенциально извлекать из контента характеристики, демонстрирующие дискриминационные особенности.

Кроме того, TF-IDF улучшает CNN, эффективно извлекая локальные характеристики и паттерны из текстового контента при использовании набора данных Truthseeker, который содержит новостные статьи и социальные блоги, помеченные для этой цели. С другой стороны, они продемонстрировали, что вложения Word2Vec и FastText не показали хороших результатов в улавливании семантических и синтаксических нюансов, что не всегда выгодно для традиционных моделей машинного обучения, включая многослойный перцептрон (MLP) или SVM. Это исследование подчеркивает важность тщательного выбора подходящих методов вложения на основе алгоритма модели для достижения высокой эффективности предсказания при обнаружении фальшивых новостей. Для встраивания TF-IDF CNN 1 и CNN 3 продемонстрировали сопоставимую производительность с точностью 98,77% и 98,99% соответственно, что свидетельствует о необходимости использования этих двух моделей для встраивания.

Shen et al. [35] разработали GAMED, мультимедийный алгоритм моделирования, который в первую очередь генерирует отчетливые и отличительные особенности посредством сортировки медиафайлов для усиления взаимосвязанности, тем самым улучшая общую эффективность обнаружения. Для извлечения отличительных и дискриминационных признаков и включения семантических знаний в GAMED используются несколько параллельных экспертных сетей. Затем распределение признаков систематически корректируется. GAMED объясняет сложные решения и выполняет новую классификацию для динамического управления вкладами из разных медиа. Экспериментальные результаты на наборах данных Fakeddit и Yang показывают, что GAMED показал лучшие результаты, чем современные модели, с точностью 93,90%.

Заключение

В данном исследовании представлен обзор алгоритмов машинного и глубокого обучения для обнаружения фейковых новостей. Также были описаны наборы данных, использованные в работах, и признаки, применяемые для извлечения значимой информации. В обзоре обозначены пробелы, выявленные в каждом исследовании, и предложены пути их устранения. Некоторые исследования,

Date: 11thDecember-2025

в которых сочетались методы машинного и глубокого обучения, а также использовались достаточные и релевантные наборы данных с тщательно отобранными признаками, продемонстрировали особенно высокие показатели точности. Также была проведена оценка производительности и качества каждого исследования. В завершение, в обзоре представлено обсуждение текущих проблем и выделены перспективные направления для будущих исследований в области обнаружения фейковых новостей.

СПИСОК ЛИТЕРАТУРЫ:

1. Aphiwongsophon, S.; Chongstitvatana, P. Detecting Fake News with Machine Learning Method. In Proceedings of the 2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Rai, Thailand, 18–21 July 2018; pp. 528–531.
2. Krishna, I.; Kumar, S. Fake News Detection using Naïve Bayes Classifier. Int. J. Creat. Res. Thought (IJCRT) 2021, 9, e757–e761. Available online: <https://ijcrt.org/papers/IJCRT2106550.pdf> (accessed on 26 May 2025).
3. Ni, B.; Guo, Z.; Li, J.; Jiang, M. Improving Generalizability of Fake News Detection Methods using Propensity Score Matching. arXiv 2020, arXiv:2002.00838.
4. Singh, D.; Khan, A.H.; Meena, S. Fake News Detection Using Ensemble Learning Models. In Proceedings of the Data Analytics and Management. ICDAM 2023; Lecture Notes in Networks and Systems. Swaroop, A., Polkowski, Z., Correia, S.D., Virdee, B., Eds.; Springer: Berlin/Heidelberg, Germany, 2023; Volume 78, pp. 55–63.
5. Jiang, T.; Li, J.P.; Haq, A.U.; Saboor, A.; Ali, A. A novel stacking approach for accurate detection of fake news. IEEE Access 2021, 9, 22626–22639.
6. Gereme, F.; Zhu, W.; Ayall, T.; Alemu, D. Combating fake news in “low-resource” languages: Amharic fake news detection accompanied by resource crafting. Information 2021, 12, 20.
7. Pardamean, A.; Pardede, H.F. Tuned bidirectional encoder representations from transformers for fake news detection. Indones. J. Electr. Eng. Comput. Sci. 2021, 22, 1667–1671.
8. Kaliyar, R.K.; Goswami, A.; Narang, P. Multiclass Fake News Detection using Ensemble Machine Learning. In Proceedings of the 2019 IEEE 9th International Conference on Advanced Computing (IACC), Tiruchirappalli, India, 13–14 December 2019; pp. 103–107.
9. Albahr, A.; Albahar, M. An Empirical Comparison of Fake News Detection using different Machine Learning Algorithms. Int. J. Adv. Comput. Sci. Appl. 2020, 11, 146–152.
10. Goldani, M.H.; Momtazi, S.; Safabakhsh, R. Detecting fake news with capsule neural networks. Appl. Soft Comput. 2021, 101, 106991.
11. Wang, Y.; Wang, L.; Yang, Y.; Lian, T. Sem-Seq4FD: Integrating global semantic relationship and local sequential order to enhance text representation for fake news detection. Expert Syst. Appl. 2021, 166, 114090.

Date: 11thDecember-2025

12. Ozbay, F.A.; Alatas, B. A novel approach for detection of fake news on social media using metaheuristic optimization algorithms. *Elektron. Ir. Elektrotechnika* 2019, 25, 62–67.
13. Birunda, S.S.; Devi, R.K. A Novel Score-Based Multi-Source Fake News Detection using Gradient Boosting Algorithm. In Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 25–27 March 2021; pp. 406–414.
14. Mugdha, S.B.S.; Ferdous, S.M.; Fahmin, A. Evaluating machine learning algorithms for bengali fake news detection. In Proceedings of the 23rd International Conference on Computer and Information Technology (ICCIT), DHAKA, Bangladesh, 19–21 December 2020; pp. 1–6.
15. Al-Ahmad, B.; Al-Zoubi, A.M.; Abu Khurma, R.; Aljarah, I. An evolutionary fake news detection method for COVID-19 pandemic information. *Symmetry* 2021, 13, 1091.
16. Jardaneh, G.; Abdelhaq, H.; Buzz, M.; Johnson, D. Classifying Arabic tweets based on credibility using content and user features. In Proceedings of the 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), Amman, Jordan, 9–11 April 2019; pp. 596–601.
17. Alshuwaier, F.; Areshey, A.; Poon, J. Applications and Enhancement of Document-Based Sentiment Analysis in Deep learning Methods: Systematic Literature Review. *Intell. Syst. Appl.* 2022, 15, 200090.
18. Battal, B.; Yıldırım, B.; Dinçaslan, Ö.F.; Cicek, G. Fake News Detection with Machine Learning Algorithms. *Celal Bayar Univ. J. Sci.* 2024, 20, 65–83.
19. Kitchenham, B.; Brereton, O.; Budgen, D.; Turner, M.; Bailey, J.; Linkman, S. Systematic Literature Reviews in Software Engineering-A Systematic Literature Review; Elsevier: Amsterdam, The Netherlands, 2009; Volume 51, pp. 7–15.
20. Do, H.H.; Prasad, P.; Maag, A.; Alsadoon, A. Deep learning for aspect-based sentiment analysis: A comparative review. *Expert Syst. Appl.* 2019, 118, 272–299.
21. Toyer, S.; Thiebaux, S.; Trevizan, F.; Xie, L. Asnets: Deep learning for generalised planning. *J. Artif. Intell. Res. (JAIR)* 2020, 68, 1–68.
22. Tiwari, S.; Jain, S. Fake News Detection Using Machine Learning Algorithms. In Proceedings of the KILBY 100 7th International Conference on Computing Sciences 2023 (ICCS 2023), Phagwara, India, 5 May 2024.
23. Rampurkar, M.V.; Thirupurasundari, D.D. An Approach towards Fake News Detection using Machine Learning Techniques. *Int. J. Intell. Syst. Appl. Eng.* 2024, 12, 2868–2874.
24. Murti, H.; Sulastri, S.; Santoso, D.B.; Diartono, D.A.; Nugroho, K. Design of Intelligent Model for Text-Based Fake News Detection Using K-Nearest Neighbor Method. *Sinkron* 2025, 9, 1–7.
25. Nagoudi, E.M.; Elmadany, A.; Abdul-Mageed, M.; Alhindi, T.; Cavusoglu, H. Machine Generation and Detection of Arabic Manipulated and Fake News. In Workshop on Arabic Natural Language Processing; Association for Computational Linguistics: Stroudsburg, PA, USA, 2020.

Date: 11thDecember-2025

26. Hamed, S.K.; Ab Aziz, M.J.; Yaakub, M.R. Fake News Detection Model on Social Media by Leveraging Sentiment Analysis of News Content and Emotion Analysis of Users' Comments. *Sensors* 2023, 23, 1748.

27. Verma, P.K.; Agrawal, P.; Amorim, I.; Prodan, R. WELFake: Word Embedding Over Linguistic Features for Fake News Detection. *IEEE Trans. Comput. Soc. Syst.* 2021, 8, 881–893.

28. Ivancova, K.; Sarnovsky, M.; Krešňáková, V. Fake news detection in Slovak language using deep learning techniques. In Proceedings of the SAMI 2021, IEEE 19th World Symposium on Applied Machine Intelligence and Informatics, Herl'any, Slovakia, 21–23 January 2021; pp. 000255–000260.

29. Subramanian, M.; Premjith, B.; Shanmugavadivel, K.; Pandiyan, S.; Palani, B.; Chakravarthi, B. Overview of the Shared Task on Fake News Detection in Dravidian Languages-DravidianLangTech@NAACL 2025. In Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, Acoma, The Albuquerque Convention Center, Albuquerque, NM, USA, 3 May 2025; Association for Computational Linguistics (ACL): Stroudsburg, PA, USA, 2025; pp. 759–767.

30. Jingyuan, Y.; Zeqiu, X.; Tianyi, H.; Peiyang, Y. Challenges and Innovations in LLM-Powered Fake News Detection: A Synthesis of Approaches and Future Directions. *Comput. Lang.* 2025, 87–93.

31. Tan, M.; Bakır, H. Fake News Detection Using BERT and Bi-LSTM with Grid Search Hyperparameter Optimization. *Bilisim Teknolojileri Dergisi*. 2025, 18, 11–28.

32. Alsuwat, E.; Alsuwat, H. An improved multi-modal framework for fake news detection using NLP and Bi-LSTM. *J. Supercomput.* 2025, 81, 177.

33. Mouratidis, D.; Kanavos, A.; Kermanidis, K. From Misinformation to Insight: Machine Learning Strategies for Fake News Detection. *Information* 2025, 16, 189.

34. Al-Tarawneh, M.A.B.; Al-Irr, O.; Al-Maaitah, K.S.; Kanj, H.; Aly, W.H.F. Enhancing Fake News Detection with Word Embedding: A Machine Learning and Deep Learning Approach. *Computers* 2024, 13, 239.

35. Shen, L.; Long, Y.; Cai, X.; Razzak, I.; Chen, G.; Liu, K.; Jameel, S. GAMED: Knowledge Adaptive Multi-Experts Decoupling for Multimodal Fake News Detection. In Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining (WSDM '25), Hannover, Germany, 10–14 March 2025; pp. 586–595.

36. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Aki, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 2021, 372, n71.