Инженерный вестник Дона, №12 (2025)
ivdon.ru/ru/magazine/archive/n12y2025/10577

Системный анализ жанрово-событийной типологии финансовой информации без предварительного обучения

И.Р. Мусин

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Аннотация: Предлагается комплексный метод системного анализа и обработки финансовой информации без предварительного обучения по пяти взаимодополняющим таксономиям (жанр, тип события, тональность, уровень влияния, временность) с одновременным извлечением сущностей. Метод основан на ансамбле из трех специализированных инструкций для локальной модели искусственного интеллекта с адаптированным алгоритмом мажоритарного голосования и двухуровневым механизмом объяснимых отказов. Протокол валидирован сравнительным тестированием 14 локальных моделей на 100 экспертно размеченных единицах информации, при этом модель достигла 90 % точности обработки. Система реализует принципы самосогласованности и селективной классификации, воспроизводится на стандартном оборудовании и не требует обучения на размеченных данных.
Ключевые слова: системный анализ, классификация без предварительного обучения, обработка финансовой информации, ансамблевые методы, мажоритарное голосование, селективная классификация, модели искусственного интеллекта, объяснимые отказы, оптимизация, статистические методы.

Введение

Актуальность исследования. Автоматическая классификация финансовых новостей представляет приоритетную задачу для информационных систем финансового рынка, обеспечивая структурирование потоков неформализованных текстов для аналитики и поддержки инвестиционных решений. Оценка эмоциональной окраски материалов позволяет исследователям выявлять общественные настроения и отношение аудитории к событиям и тематикам [1]. Подходы, использующие обучение на размеченных данных с дообучением трансформеров (BERT, RoBERTa), требуют масштабных размеченных корпусов и регулярного переобучения при смене предметной области [2], при этом сохраняется высокий уровень чувствительности к изменениям финансового контекста.

Классификация без предварительного обучения с применением больших языковых моделей представляет альтернативный метод [3]. При

использовании в финансовой предметной области возникают проблемы: нестабильность инструкций, необходимость многотаксономической классификации, требования к надежности и вычислительные ограничения при работе с большими корпусами на стандартном оборудовании.

Проблематика. Существующие методы классификации текста с нулевого уровня для финансовых новостей характеризуются рядом фундаментальных ограничений. Незначительные изменения формулировок инструкций вызывают значительную вариативность откликов языковых моделей, особенно при работе с текстами, насыщенными отраслевыми жаргонизмами и аббревиатурами. Типовые реализации не предоставляют количественных оценок достоверности предсказаний, при этом отсутствие таких метрик неприемлемо для финансовых систем, где ошибки классификации могут привести к значительным экономическим издержкам.

Большинство исследований сосредоточено на классификации по единственной таксономии (жанр либо сентимент), тогда как прикладные задачи требуют параллельного определения нескольких атрибутов новостного контента [4]. Последовательное использование независимых классификаторов неэффективно с вычислительной точки зрения и игнорирует взаимозависимость между таксономическими категориями.

Дополнительную проблему составляет низкая производительность существующих решений: применение проприетарных интерфейсов прикладного программирования, таких как крупномасштабная мультимодальная модель GPT-4, большие языковые модели Claude, при обработке крупных корпусов сопровождается высокими затратами и рисками для конфиденциальности, а локальные открытые модели требуют оптимизации под стандартное оборудование. В публикациях отсутствуют комплексные протоколы валидации с системным сравнением моделей разных архитектур при решении задач классификации финансовых новостей.

Цель исследования – разработка и валидация комплексного метода классификации финансовых новостей без предварительного обучения по множественным взаимодополняющим таксономиям с механизмами надежности, интерпретируемости и вычислительной эффективности для практического применения на стандартном оборудовании.

Для достижения цели нужно решить следующие задачи:

1. Разработать ансамблевый протокол классификации по пяти взаимодополняющим таксономиям с использованием трех специализированных инструкций, мажоритарного голосования и трехуровневой резервной логики.

2. Реализовать механизм объяснимых отказов с глобальной и селективной классификацией для оптимизации соотношения между точностью и охватом.

3. Выполнить валидацию на 14 локальных моделях различных архитектур с применением экспертно размеченного набора и расчетом метрик по каждой таксономии.

4. Провести итеративную оптимизацию протокола на корпусе из 32 272 статей с обеспечением вычислительной эффективности на стандартном оборудовании.

Научная новизна работы заключается в разработке комплексного метода классификации без обучения по пяти взаимодополняющим таксономиям с адаптированным алгоритмом мажоритарного голосования, содержащим трехуровневую резервную логику принятия решений.

Обзор литературы

В обработке естественного языка классификация без предварительного обучения развивалась с появлением крупных предобученных языковых моделей [3, 5], способных выполнять задачи без примеров обучения по естественноязычным инструкциям. Крупномасштабная мультимодальная

модель GPT-3 достигла результатов, сопоставимых с базовыми методами, использующими размеченные данные; при этом выявлена высокая чувствительность к формулировке инструкций. В финансовой предметной области подход применялся преимущественно к анализу тональности [6], ограничиваясь однотаксономической классификацией.

Систематизация методов построения инструкций для языковых моделей выделяет несколько стратегий. Инструкционный подход представляет задачу в форме прямой инструкции [7, 8]. Подход с цепочкой рассуждений формирует промежуточные рассуждения, повышая точность при многошаговых выводах [9]. Метод самосогласованности применяет мажоритарное голосование по множественным откликам, снижая вариативность. Ансамблевые схемы демонстрируют преимущество объединения ответов по нескольким парафразированным инструкциям над единичными формулировками.

Селективная классификация описывает сценарий, при котором классификатор воздерживается от предсказания при высокой неопределенности, обеспечивая баланс между охватом и точностью [5]. В задачах без предварительного обучения отказ формируется на основе консенсуса между несколькими инструкциями. Концепция объяснимых отказов дополняется требованием интерпретируемости причин, критически значимым для финансовой предметной области [10].

Классификация финансовых новостей

Автоматическая классификация финансовых новостей традиционно ограничивалась анализом тональности. С появлением трансформеров разработаны специализированные модели FinBERT [6]. Жанровая классификация рассмотрена на материале общего новостного корпуса. Многотаксономическая классификация (жанр, сентимент, события) применительно к финансовым новостям практически не исследована.

Развитие открытых языковых моделей с квантизацией обеспечило возможность локального вывода на стандартном оборудовании [8]. Модели с объемом параметров в пределах 3–7 млрд обеспечивают приемлемую производительность, однако систематические сравнительные оценки на задачах финансовой классификации отсутствуют. Вычислительная оптимизация достигается за счет параллелизации запросов и кэширования.

Таким образом, анализ литературы выявляет следующие пробелы:

1. отсутствие многотаксономической классификации финансовых новостей без предварительного обучения по жанру, событию и сентименту одновременно;

2. нехватка систематизированных протоколов валидации с сопоставлением нескольких моделей;

3. ограниченное использование ансамблевых методов с механизмами оценки уверенности и объяснимых отказов;

4. недостаточная проработка вычислительной эффективности при использовании на стандартном оборудовании.

Методы

Предложенный метод классификации финансовых новостей без предварительного обучения представляет комплексную систему обработки информации, основанную на ансамблевом проектировании инструкций и статистической агрегации предсказаний [3, 5, 9]. Система использует селективную классификацию [5] для повышения надежности при высокой неопределенности.

Пусть корпус включает N финансовых новостных статей, каждая из которых описывается заголовком и основным текстом.

$$N = \{d_1, d_2, \ldots, d_n\}. \quad (1)$$

Пусть N – корпус статей, каждая статья d_1состоит из заголовка h_1 и содержания c_1. Задача заключается в построении отображения статей в пространство таксономических классификаций:

$$T = T_m \times T_e \times T_s \times T_i \times T_t \times P(E), \quad (2)$$

где T – полное пространство таксономических классификаций;

T_m – множество жанров (event, analysis, opinion, press-release);

T_e – множество типов событий;

T_s – множество уровней влияния;

T_i – множество жанров;

T_t – множество временностей;

$P(E)$ – множество подмножеств пространства сущностей (токены, персоны, организации, концепции).

Для повышения стабильности предсказаний и снижения вариативности, характерной для вывода без предварительного обучения, используется ансамбль из трех специализированных инструкций, реализующих различные стратегии классификации. Выбор нечетного количества основан на балансе между вычислительной эффективностью и качеством мажоритарного голосования, исключающим ситуацию равных голосов. При объединении нескольких методов с умеренной индивидуальной результативностью и обучении на взаимной коррекции ошибок общее качество системы существенно превышает показатели каждого метода по отдельности [11].

Инструкция 1. Строгий криптовалютный классификатор с якорными лексиконами. Реализует инструкционный подход с фокусом на якорных терминах криптовалютной сферы (Bitcoin, Ethereum, Binance, SEC, regulation, hack и др.) [9]. Функция классификации задается через проверку наличия якорных терминов в заголовке и тексте статьи. При отсутствии релевантных терминов присваивается метка нерелевантности.

Инструкция 2. Финансовый аналитик с многошаговым выводом. Применяет цепочку рассуждений [9], декомпозируя задачу на пять этапов: проверка релевантности → определение жанра → классификация события → извлечение сущностей → оценка влияния. Каждый этап формирует промежуточный результат для следующего шага.

Инструкция 3. Пошаговый эксперт с явным структурированием. Комбинирует инструкционный подход с нумерацией шагов и обоснованием решений. Генерирует не только итоговые классификации, но и текстовые объяснения по каждой таксономической категории.

Каждая инструкция применяется к статье через локальную языковую модель LLaMA-3.2-3B (квантизация Q4_K_M, параметры: τ = 0.3, top-K = 40, top-P = 0.8, max tokens = 256) [8], генерируя структурированный JSON-ответ (JavaScript Object Notation, текстовый формат обмена данными, основанный на JavaScript). Критически важным является использование грамматики JSON Schema для обеспечения валидности выходных данных и предотвращения галлюцинаций модели.

Для сокращения задержки вывода используется параллельное выполнение трех инструкций через асинхронную конкурентную обработку (реализация через Promise.all в среде Node.js). Время обработки уменьшается за счет того, что продолжительность определяется максимальным временем выполнения одной инструкции, а не суммой длительностей. В экспериментах зафиксировано ускорение с 7715 мс до 3995 мс (коэффициент 1.93×).

$$T_{seq} = \sum(i = 3)t_j, \quad (3)$$

$$T_{par} = max(t_1, t_2, t_3), \quad (4)$$

где T_{seq} – общее время последовательного выполнения трех инструкций;

T_{par}– общее время параллельного выполнения трех инструкций;

t_j– время выполнения инструкции p_j.

Параллельное выполнение инструкций сокращает время обработки за счет того, что общая длительность определяется максимальным временем одной инструкции, а не суммой времен в пределах одной загруженной модели с созданным контекстом. В экспериментах зафиксировано сокращение времени с 7715 мс до 3995 мс (коэффициент ускорения – 1.93).

Рассмотрим статистическую агрегацию и мажоритарное голосование.

Для каждой таксономической категории (жанр, событие, сентимент, влияние, временность) применяется адаптированный алгоритм мажоритарного голосования с резервным механизмом [12]. Из трех предсказаний инструкций отбираются валидные (не отказные) ответы. Для каждого уникального значения вычисляется частота, после чего победителем становится вариант с наибольшим числом голосов.

Показатель уверенности вычисляется как отношение количества голосов за победившее значение к общему числу валидных ответов:

$$conf_k(d_i) = \frac{n_{v^*}}{\left|R_k^{valid}\right|}, \quad (5)$$

где n_{v^*}– количество голосов за победившее значение v^*, а $\left|R_k^{valid}\right|$ – общее число валидных ответов.

Финальная функция с трехуровневой логикой принятия решений:

$$f_k(d_i) = \begin{cases} v^*, & \text{если } conf_k(d_i) \geq 0.67 \\ v^*, & \text{если } 0 < conf_k(d_i) < 0.67, \\ null, & \text{если отсутствуют ответы} \end{cases} \quad (6)$$

где $f_k(d_i)$ – финальное решение для категории k на статье d_i;

v^*– победившее значение (с максимумом голосов);

0.67 – порог высокой уверенности (2 из 3 инструкций согласны).

Трехуровневая логика устранила неопределенные значения, зафиксированные в первой итерации (40,14 % для жанров), за счет резервного механизма, выбирающего наиболее частотный ответ даже при отсутствии строгого большинства. При трех инструкциях возможны

следующие конфигурации: единогласие (3,0,0) соответствует conf = 1.0; большинство (2,1,0) – conf = 0.67; полное разногласие (1,1,1) – conf = 0.33. Результат возвращается при наличии хотя бы одного валидного ответа.

В отличие от категориальных таксономий, для извлечения сущностей используется операция объединения множеств без применения мажоритарного голосования. Итоговое множество формируется как объединение всех уникальных сущностей, извлеченных тремя инструкциями:

$$\mathcal{E}(d_i) = E_1(d_i), E_2(d_i), E_3(d_i), \quad (7)$$

где $\mathcal{E}(d_i)$ – это список всех сущностей (например, имена, места, объекты), которые нашлись в статье d_i;

$E_j(d_i)$ – сущности, которые нашла каждая из трех инструкций (p_1, p_2, p_3) в этой статье.

Подход максимизирует полноту извлечения за счет агрегации всех уникальных упоминаний из трех инструкций, обеспечивая надежную основу для последующего конструирования признаков в задачах прогнозирования цен токенов. Теоретико-множественное свойство объединения гарантирует, что полнота ансамбля не ниже, чем у наиболее результативной индивидуальной инструкции.

Система реализует иерархический механизм отказа от классификации, оптимизирующий компромисс между охватом и точностью [5].

Уровень 1 (глобальный отказ). Статья d_i отклоняется как нерелевантная, если большинство промптов устанавливают переменную refusal как:

$$GlobalRefusal(d_i) = \begin{cases} 1 \\ 0 \end{cases}. \quad (8)$$

Другими словами, для статьи d_i смотрим, сколько инструкций p_j, где $j = 1,2,3$ отказались ее классифицировать $\text{refusal}_j(d_i) = \text{true}$ или посчитали нерелевантной $r_j^e(d_i) = \text{irrelevant}$. При двух или трех отказах или метках

нерелевантности статья отклоняется. Такой фильтр отсекает явно неподходящие статьи, исключая траты ресурсов на их обработку. Если статья отклонена – дальнейшая обработка не выполняется, причина отказа сохраняется; если принята – направляется на этап селективной классификации.

Уровень 2: Селективная классификация полей

Для статей, которые прошли первый уровень $GlobalRefusal(d_i)$, система решает, какие категории (k, например, жанр или тональность) можно классифицировать, а какие нет:

$$f_k(d_i) = \text{null, если } \text{conf}_k(d_i) < \theta_{\text{low}} \text{ или } |\, R_k^{\text{valid}}(d_i) \,| = 0. \quad (9)$$

Если уверенность $\text{conf}_k(d_i)$ в классификации категории k ниже порога θ_{low} или нет валидных ответов $|\, R_k^{\text{valid}}(d_i) \,| = 0$, то для этой категории возвращается нулевой ответ null. Это позволяет системе отказаться от ответа, если он ненадежен, вместо того чтобы делать ошибочное предположение.

Результаты

Эксперименты выполнены на стандартном оборудовании (Intel Core i5-12450H, NVIDIA RTX 4050, 16 GB RAM) с использованием локальной модели через библиотеку node-llama-cpp. Конфигурация включает автоматическое определение GPU (graphics processing unit, графический процессор), частичное размещение модели на GPU, подгонку контекста до 4096 токенов, квантизацию Q4_K_M (4-bit), среду выполнения Node.js и PostgreSQL для хранения корпуса. Время классификации определяется максимальной длительностью обработки одной инструкции в параллельном режиме, что позволяет снизить задержку на 48,2 % за три итерации оптимизации. Дополнительные меры: ограничение длины контента до 500 символов, кэширование контекста между запросами. Метрики: совокупная точность для совместной классификации, компромисс охват–точность при

порогах уверенности 0.5, 0.67, 0.8, анализ распределений уровней уверенности.

Для проверки надежности протокола выполнен сравнительный анализ 14 локальных языковых моделей на размеченном наборе из 100 новостей (50 криптовалютных, 50 нерелевантных) с использованием ансамбля из трех инструкций и мажоритарного голосования при пороге 2/3. Тестирование заняло 14 часов 35 минут с учетом загрузки моделей, 2-секундных пауз между статьями и 10-секундных пауз между моделями.

В группу моделей с точностью ≥85 % вошли Llama-3.2-3B, Phi-3-Mini и Qwen2.5-3B. Llama-3.2-3B достигла 90 % общей точности классификации и 97,96 % точности по жанру при сбалансированной производительности. Доменно-специализированные модели показали неожиданно низкие результаты (Phinance-Phi-4: 73 %, WiroAI-Finance: 59–53 %), что подчеркивает значение общих языковых способностей при классификации без предварительного обучения (таблица 1).

Таблица №1.

Сравнительные результаты языковых моделей

Модель	Параметры	Квантизация	Класс	Поля	Время (мс)
LLaMA-3.2-Instruct	3B	Q4_K_M	90.00%	60.34%	5620
Phi-3-Mini-Instruct	4K	Q4	90.00%	60.34%	5642
Qwen2.5-Instruct	3B	Q4_K_M	88.00%	66.67%	4012
Mistral-Instruct-v0.2	7B	Q4_K_M	85.00%	49.83%	11306
Qwen2-Instruct	7B	Q4_K_M	77.00%	63.78%	9863
Phinance-Phi-4-Mini	Finance	Q4_K_M	73.00%	51.72%	5201
Qwen3	0.6B	Q8_0	68.00%	53.45%	2140
Gemma-2-IT	2B	Q4_K_M	59.00%	68.89%	4479

Примечание: Остальные 6 моделей (варианты Unsloth и WiroAI-Finance) показали результаты 50-63% и опущены для краткости.

Эти результаты легли в основу выбора Llama-3.2-3B в качестве базовой модели для основного эксперимента.

Перейдем к результатам итеративной оптимизации.

Первая итерация представляла пилотное тестирование протокола на выборке из 759 новостных статей (по одной от каждого источника). Обработка заняла 117,6 минуты при среднем времени 7715 мс на статью, при этом высокая длительность связана с последовательным выполнением трех инструкций без параллельной обработки (таблица 2).

Таблица №2.

Общая статистика обработки первой итерации

Метрика	Значение
Обработано статей	759
Принято	142
Отказано	617
Процент отказов, %	81.29
Успешность классификации, %	18.71
Среднее время, мс	7715

Высокий процент отказов (81,29 %) выявил проблему релевантности контента. Среди 142 принятых статей 40,14 % остались без жанровой метки, 23,24 % – без метки типа события, отражая ограниченную согласованность инструкций. Средние показатели уверенности составили: genre – 0,673; event_type – 0,803; sentiment – 0,829; impact_level – 0,823; timelyness – 0,854. Распределение по жанрам показало преобладание категории event (29,58 %); по типам событий – listing (27,46 %) и regulation (21,83 %).

На основе анализа первой итерации реализован комплекс взаимосвязанных оптимизаций. Таксономия событий дополнена тремя категориями (*market* – рыночные движения, *partnership* – коллаборации, *upgrade* – обновления протоколов), что обеспечило классификацию 38,71 %

статей, ранее получавших метки неопределенности. Алгоритм мажоритарного голосования усовершенствован трехуровневой логикой: единогласие (conf = 1.0), большинство 2/3 (conf = 0.67) и резервный механизм при разногласии (conf = 0.33), устранивший неопределенные значения (с 40,14 % до 0 %).

Контекстное окно увеличено с 300 до 500 символов для более точной интерпретации сложных финансовых текстов. Механизм инструкций расширен якорными лексиконами для новых типов событий, используемыми как вспомогательные сигналы при классификационной неопределенности. Переход к параллельной обработке инструкций через Promise.all вместо последовательного выполнения снизил время обработки с 7715 мс до 5678 мс (–26,4 %).

Вторая итерация масштабировала оптимизированный протокол на 9662 статьи (обработка 35.7 часов, среднее время 5678 мс/статья, улучшение -26.4%) (таблица 3).

Таблица №3.

Общая статистика обработки второй итерации

Метрика	Значение	Δ от итерации 1
Обработано статей	9662	+1173.3%
Принято	2292	+1514.8%
Отказано	7370	+1094.5%
Процент отказов	76.28%	-5.01 п.п.
Успешность классификации	23.72%	+5.01 п.п.
Среднее время (мс)	5678	-26.4%

Успешность классификации увеличилась до 23,72 % (+5,01 п.п.), подтверждая эффективность модификаций. Существенное достижение – устранение неопределенных классификаций: только 1 из 2292 статей (0,04 %)

содержала пустой массив сущностей, все таксономические поля получили значения (0 % null для жанров против 40,14 % в первой итерации).

Показатели уверенности выросли: genre – 0,737 (+0,064), event_type – 0,834 (+0,031), sentiment – 0,825, impact_level – 0,835, timelyness – 0,851. Распределение по жанрам: event – 41,40 %, press-release – 25,61 %, analysis – 25,39 %, opinion – 7,59 %. Расширенная таксономия событий охватила 38,71 % корпуса: regulation – 22,60 %, market – 19,81 %, listing – 19,07 %, partnership – 11,13 %. Итерация выявила 251 источник с релевантностью ниже 20 %, ставших основой для оптимизации фильтрации.

Анализ 9662 статей второй итерации выявил два кластера источников: высокий уровень качества (501 источник с релевантностью >80 %) и крайне низкий (251 источник с релевантностью <20 %). Принято решение об отключении 251 источника, сократив активную базу с 759 до 508 (–33,1 %). Цели: повышение успешности классификации с 23,72 % до 30–35 %, снижение вычислительных затрат, сохранение охвата специализированного контента.

Финальная итерация применила оптимизированный протокол на 21851 статье из 447 источников (обработка 18.6 часов, среднее время 3995 мс/статья, улучшение -29.6% от итерации 2, -48.2% от итерации 1) (таблица 4).

Таблица №4.

Общая статистика обработки третьей итерации

Метрика	Значение	Δ от итерации 2
Обработано статей	21851	+126.2%
Принято	7950	+246.9%
Отказано	13901	+88.6%
Процент отказов	63.62%	-12.66 п.п.
Успешность классификации	36.38%	+12.66 п.п.

Среднее время (мс)	3995	-29.6%

Успешность достигла 36,38 % (+12,66 п.п., +53,4 %), превысив целевую планку 30–35 %. Показатели уверенности: genre – 0,739 (+0,002 к итерации 2), event_type – 0,852 (+0,018), sentiment – 0,833 (+0,008), impact_level – 0,825 (–0,010), timelyness – 0,868 (+0,017). Распределение по жанрам: event – 39,77 %, analysis – 32,83 % (+7,44 п.п.), press-release – 20,10 %, opinion – 7,30 %. Распределение по типам событий зафиксировало структурный сдвиг: market стал доминирующим (33,46 %, +13,65 п.п.), далее regulation – 20,60 %, listing – 15,01 %, partnership – 9,26 %.

Агрегированная сводка по итерациям представлена в таблицах 5-8.

Таблица №5.

Сравнительная статистика по трем итерациям

Метрика	Итерация 1	Итерация 2	Итерация 3	Улучшение
Обработано статей	759	9662	21851	+2779%
Принято	142	2292	7950	+5500%
Успешность	18.71%	23.72%	36.38%	+17.67 п.п.
Время (мс)	7715	5678	3995	-48.2%
Источников	759	759	447	-41.1%
Null (genre)	40.14%	0.00%	0.00%	-40.14 п.п.
Confidence (genre)	0.673	0.737	0.739	+0.066

Итеративная оптимизация дала следующие результаты: успешность классификации выросла с 18,71 % до 36,38 % (+94,5 %), время обработки снизилось с 7715 мс до 3995 мс (–48,2 %), неопределенные значения устранены полностью (40,14 % → 0 %), показатели уверенности стабилизировались выше 0,73 для всех таксономий.

Таблица №6.

Распределение по жанрам

Жанр	Итерация 1	Итерация 2	Итерация 3
event	42 (29.58%)	949 (41.40%)	3162 (39.77%)
analysis	31 (21.83%)	582 (25.39%)	2610 (32.83%)
press-release	9 (6.34%)	587 (25.61%)	1598 (20.10%)
opinion	3 (2.11%)	174 (7.59%)	580 (7.30%)
null	57 (40.14%)	0 (0.00%)	0 (0.00%)

Таблица №7.

Распределение по типам событий

Тип события	Итерация 1	Итерация 2	Итерация 3
regulation	31 (21.83%)	518 (22.60%)	1638 (20.60%)
market	Еще не введен	454 (19.81%)	2660 (33.46%)
listing	39 (27.46%)	437 (19.07%)	1193 (15.01%)
partnership	Еще не введен	255 (11.13%)	736 (9.26%)
hack	13 (9.15%)	209 (9.12%)	519 (6.53%)
funding	21 (14.79%)	140 (6.11%)	525 (6.60%)

Таблица №8.

Показатели уверенности по таксономиям

Поле	Итерация 1	Итерация 2	Итерация 3	Улучшение
genre	0.673	0.739	+0.066	0.066
event_type	0.803	0.834	0.852	0.049
sentiment	0.829	0.825	0.833	0.004
impact_level	0.823	0.835	0.825	0.002
timelyness	0.854	0.851	0.868	0.014

Обсуждение

Полученные результаты подтверждают гипотезу о возможности надежной многотаксономической классификации финансовых новостей без предварительного обучения с использованием ансамблевых методов. Существенное достижение – устранение нестабильности единичных инструкций: трехинструкционный ансамбль с адаптированным мажоритарным голосованием обеспечил стабильные показатели уверенности 0,73–0,87 по всем таксономиям при 100 % охвате.

Преимущество общецелевых языковых моделей (LLaMA-3.2, Phi-3-Mini, Qwen2.5) над доменно-специализированными (Phinance-Phi-4: 73 %, WiroAI-Finance: 53–59 %) подчеркивает значимость общих языковых способностей в задачах классификации без обучения. Этот вывод согласуется с результатами [3], где подчеркивается влияние масштаба предварительного обучения на эффективность в zero-shot сценариях.

Итеративная оптимизация выявила эффект нелинейного роста: фильтрация источников (итерация 2→3) обеспечила прирост +12,66 п.п., превысив эффект от модификаций протокола (итерация 1→2: +5,01 п.п.). Структурный сдвиг в распределении событий (market: +13,65 п.п. в итерации 3) отражает изменение состава корпуса после фильтрации.

Протокол валидирован на англоязычных криптовалютных новостях; требуется верификация применимости к материалам о традиционных финансовых рынках и многоязычным корпусам. Компромисс между охватом и точностью (36,38 % успешность при высокой точности полей) отражает консервативную стратегию отказов с возможностью настройки порогов под конкретные задачи. Вычислительная эффективность обеспечена контекстным ограничением в 500 символов, при этом ограничение снижает полноту обработки длинных аналитических публикаций.

Заключение

Разработан и экспериментально валидирован протокол классификации финансовых новостей без предварительного обучения по нескольким взаимодополняющим таксономиям, устраняющий основные ограничения существующих решений: нестабильность единичных инструкций, отсутствие оценки уверенности и вычислительную неэффективность.

Валидация на 14 локальных языковых моделях показала: LLaMA-3.2-3B достигает 90 % общей точности классификации и 97,96 % точности по жанру, превосходя альтернативные открытые модели сопоставимого объема. Итеративная оптимизация на корпусе из 32 272 статей увеличила успешность с 18,71 % до 36,38 % (+94,5 %), устранила неопределенные значения (с 40,14 % до 0 %) и снизила время обработки на 48,2 % (с 7715 мс до 3995 мс).

Механизм мажоритарного голосования с трехуровневой резервной логикой обеспечивает 100 % охват по всем таксономиям при сохранении показателей уверенности выше 0,73. Вычислительная эффективность достигнута за счет параллельной обработки инструкций и оптимизации параметров модели, что позволяет использовать протокол на стандартном оборудовании.

Работа развивает методологию селективной классификации для сценариев без предварительного обучения, демонстрируя возможность построения надежных классификационных систем на основе ансамблевых методов и формализованных механизмов отказа. Предложенный протокол обеспечивает воспроизводимую классификацию крупных финансовых корпусов, формируя основу для аналитических систем финансового рынка и автоматизированного мониторинга информационного пространства.

Направления дальнейшего развития включают расширение на многоязычные корпуса, адаптацию к потоковой обработке, интеграцию с

моделями прогнозирования цен криптовалютных активов и анализ временных лагов между классификационными признаками и рыночными движениями.

Литература

1. Баркович А.А., Петрова Н.С. Сентимент-анализ: прагматическая специфика // Труды БГТУ. Сер. 4, Принт- и медиатехнологии. 2023. № 2 (273). С. 40–46.

2. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // North American Chapter of the Association for Computational Linguistics. 2019. pp. 4171-4186.

3. Brown T., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D. M., Wu J., Winter C., Hesse C., Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner C., McCandlish S., Radford A., Sutskever I., Amodei D. Language Models are Few-Shot Learners // Neural Information Processing Systems. 2020. V. 33. pp. 1877-1901.

4. Li C., Gao F., Bu J., Xu L., Chen X., Gu Y., Sheng V. S., Zheng Z., Zhang Y., Yu Z. SentiPrompt: Sentiment Knowledge Enhanced Prompt-Tuning for Aspect-Based Sentiment Analysis URL: arxiv.org/abs/2109.08306.

5. Geifman Y., El-Yaniv R. Selective Classification for Deep Neural Networks // Neural Information Processing Systems. 2017. URL: proceedings.neurips.cc/paper_files/paper/2017/file/4a8423d5e91fda00bb7e46540e2b0cf1-Paper.pdf.

6. Araci D. FinBERT: Financial Sentiment Analysis with Pre-trained Language Models. URL: arxiv.org/abs/1908.10063.

7. Ouyang L., Wu J., Jiang X., Almeida D., Wainwright C. L., Mishkin P., Zhang C., Agarwal S., Slama K., Ray A., Schulman J., Hilton J., Kelton F., Miller

L., Simens M., Askell A., Welinder P., Christiano P., Leike J., Lowe R. J. Training language models to follow instructions with human feedback // Neural Information Processing Systems. 2022. V. 35. pp. 27730-27744.

8. Touvron H., Lavril T., Izacard G., Martinet X., Lachaux M.-A., Lacroix T., Rozière B., Goyal N., Hambro E., Azhar F., Rodriguez A., Joulin A., Grave E., Lample G. LLaMA: Open and Efficient Foundation Language Models URL: arXiv preprint arXiv: 2302.13971.

9. Wei J., Wang X., Schuurmans D., Bosma M., Ichter B., Xia F., Chi E., Le Q., Zhou D. Chain of Thought Prompting Elicits Reasoning in Large Language Models // Neural Information Processing Systems. 2022. V 35. pp. 24824-24837.

10. Varshney K. R., Alemzadeh H. On the safety of machine learning: Cyber-physical systems, decision sciences, and data products // Big data. 2017. V. 5. № 3. pp. 246–255.

11. Бутырский Е. Ю., Цехановский В. В., Жукова Н. А., Баймуратов И. Р., Куликов И. А. Машинное обучение. М.: Директ-Медиа, 2023. 368 с.

12. Kuncheva L.I. Combining Pattern Classifiers: Methods and Algorithms. John Wiley & Sons, 2004. 356 p.

References

1. Barkovich A.A., Petrova N.S. Trudy BGTU. Ser. 4, Print- i mediatekhnologii. 2023. № 2 (273). pp. 40–46.

2. Devlin J., Chang M.-W., Lee K., Toutanova K. North American Chapter of the Association for Computational Linguistics. 2019. pp. 4171-4186.

3. Brown T., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D. M., Wu J., Winter C., Hesse C., Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner C., McCandlish S., Radford A., Sutskever I., Amodei D. Neural Information Processing Systems. 2020. V. 33. pp. 1877-1901.

4. Li C., Gao F., Bu J., Xu L., Chen X., Gu Y., Sheng V. S., Zheng Z., Zhang Y., Yu Z. SentiPrompt: Sentiment Knowledge Enhanced Prompt-Tuning for Aspect-Based Sentiment Analysis URL: arxiv.org/abs/2109.08306.

5. Geifman Y., El-Yaniv R. Neural Information Processing Systems. 2017. URL: proceedings.neurips.cc/paper_files/paper/2017/file/4a8423d5e91fda00bb7e46540e2b0cf1-Paper.pdf.

6. Araci D. FinBERT: Financial Sentiment Analysis with Pre-trained Language Models URL: arxiv.org/abs/1908.10063.

7. Ouyang L., Wu J., Jiang X., Almeida D., Wainwright C. L., Mishkin P., Zhang C., Agarwal S., Slama K., Ray A., Schulman J., Hilton J., Kelton F., Miller L., Simens M., Askell A., Welinder P., Christiano P., Leike J., Lowe R. J. Neural Information Processing Systems. 2022. V. 35. pp. 27730-27744.

8. Touvron H., Lavril T., Izacard G., Martinet X., Lachaux M.-A., Lacroix T., Rozière B., Goyal N., Hambro E., Azhar F., Rodriguez A., Joulin A., Grave E., Lample G. LLaMA: Open and Efficient Foundation Language Models URL: arXiv preprint arXiv: 2302.13971.

9. Wei J., Wang X., Schuurmans D., Bosma M., Ichter B., Xia F., Chi E., Le Q., Zhou D. Neural Information Processing Systems. 2022. V 35. pp. 24824-24837.

10. Varshney K. R., Alemzadeh H. Big data. 2017. V. 5. № 3. pp. 246–255.

11. Butyrskij E. Yu., Cekhanovskij V. V., Zhukova N. A., Bajmuratov I. R., Kulikov I. A. Mashinnoe obuchenie [Machine learning]. M.: Direkt-Media, 2023. 368 p.

12. Kuncheva L.I. Combining Pattern Classifiers: Methods and Algorithms. John Wiley & Sons, 2004. 356 p.

Дата поступления: 17.10.2025
Дата публикации: 26.11.2025