

$$S = \langle B, E, O, U, Z, H, F, \Phi, \Omega \rangle, \quad (2)$$

где B – множество городских автобусов; E – контролируемые элементы; O – набор возможных технических воздействий; U – факторы эксплуатации; Z – диагностические данные; H – история выполненных работ и отказов; F – модели оценки остаточного ресурса; Φ – процедуры группировки операций; Ω – критерии и ограничения планирования.

Таким образом, эффективная периодичность технического обслуживания городских автобусов определяется не единым усредненным интервалом, а сочетанием эксплуатационных, диагностических и организационных данных. Дифференцированный подход позволяет уменьшить преждевременные вмешательства, снизить внеплановые простои и рационально использовать ресурсы депо. Структурная модель прогнозируемого технического обслуживания служит основой для разработки алгоритмов автоматизированного планирования периодичности и состава операций.

Литература

1. Общесоюзные нормы технологического проектирования предприятий автомобильного транспорта : утв. протоколом концерна «Росавтотранс» от 07.08.1991 № 3. ОНТП-01-91. РД 3107938-0176-91. – URL: <https://neo-park.ru/wp-content/uploads/2021/08/ontp-01-91.-rd-3107938-0176-91.-obshhesoyuznye-normy-tehnologicheskogo-proektirovaniya-atp.pdf> (дата обращения 20.03.2026).
2. Ahmad, R. An overview of time-based and condition-based maintenance in industrial application / R. Ahmad, S. Kamaruddin // Computers & Industrial Engineering. – 2012. – Vol. 63, N 1. – P. 135–149. – DOI 10.1016/j.cie.2012.02.002
3. Jardine, A. K. A review on machinery diagnostics and prognostics implementing condition-based maintenance / A. K. Jardine, D. Lin, D. Banjevic // Mechanical Systems and Signal Processing. – 2006. – Vol. 20, N 7. – P. 1483–1510. – DOI 10.1016/j.ymssp.2005.09.012
4. Machinery health prognostics: A systematic review from data acquisition to RUL prediction / Y. Lei, N. Li, L. Guo [et al.] // Mechanical Systems and Signal Processing. – 2018. – Vol. 104. – P. 799–834. – DOI 10.1016/j.ymssp.2017.11.016
5. Explainable AI Using On-Board Diagnostics Data for Urban Buses Maintenance Management: A Study Case / B. Tormos, B. Pla, R. Sanchez-Marquez, J. L. Carballo // Information. – 2025. – Vol. 16, N 2. – Article 74. – P. 1–24. – DOI 10.3390/info16020074
6. Dimensioning reserve bus fleet using life cycle cost models and condition-based/predictive maintenance: a case study / H. Raposo, J. Torres Farinha, L. Ferreira, D. Galar // Public Transport. – 2018. – Vol. 10. – P. 169–190. – DOI 10.1007/s12469-017-0167-x

ОСОБЕННОСТИ ПРОГРАММНОЙ РЕАЛИЗАЦИИ ПОТОКОВОГО АНАЛИЗА ТОНАЛЬНОСТИ ФИНАНСОВЫХ НОВОСТЕЙ С ПОВТОРНЫМ ОБУЧЕНИЕМ НА ОСНОВЕ ПОЛЬЗОВАТЕЛЬСКОЙ РАЗМЕТКИ

Г. Б. Никифоров

*Федеральное государственное бюджетное образовательное учреждение
высшего образования «Псковский государственный университет»,
Российская Федерация*

Научный руководитель Д. А. Андреев

Предложена архитектура программного комплекса потокового анализа тональности финансовых новостей, ориентированного на функционирование в условиях ограниченного объема размеченных данных. Комплекс реализует автоматический сбор новостных сообщений, первичную оценку их тональности и накопление пользовательской разметки в форме попарных сравнений. Результаты сравнений агрегируются в количественные оценки

тональности, используемые для формирования обучающей выборки. Предусмотрены процедуры первичного и повторного обучения моделей анализа тональности по мере накопления разметки, а также оценка качества моделей на основе стандартных метрик. Для учета изменения распределения данных новостного потока обучение осуществляется на скользящем временном окне с исключением устаревших наблюдений. Предложенный подход обеспечивает последовательное формирование обучающей выборки и обновление моделей анализа тональности в процессе эксплуатации программного комплекса.

Ключевые слова: анализ тональности, финансовые новости, потоковый режим обработки данных, попарные сравнения, пользовательская разметка.

Информационный поток финансовых новостей влияет на динамику фондовых рынков и инвестиционные решения участников рынка. Публикации о корпоративных событиях, макроэкономических показателях и регуляторных изменениях формируют информационный фон, воздействующий на ожидания инвесторов и краткосрочные колебания цен активов [1].

Анализ тональности финансовых новостей используется в системах мониторинга информационного пространства и может служить источником сигналов для моделей прогнозирования динамики финансовых активов [2].

Разработка систем анализа тональности в финансовой предметной области сталкивается с рядом ограничений. Основным из них является недостаток размеченных данных для обучения моделей анализа тональности. Формирование размеченных наборов данных требует значительных затрат времени и экспертных ресурсов [3]. Дополнительную сложность представляет изменение распределения данных новостного потока во времени, связанное с изменением экономического контекста и содержания публикуемых новостей [4].

В этих условиях актуальной является разработка программного комплекса, позволяющего постепенно формировать и обновлять обучающую выборку непосредственно в процессе эксплуатации и использовать ее для первоначального и повторного обучения моделей анализа тональности по мере накопления пользовательской разметки.

Целью работы является разработка архитектуры программного комплекса потокового анализа тональности финансовых новостей, обеспечивающего накопление пользовательской разметки и повторное обучение моделей анализа тональности по мере ее накопления.

Архитектура программного комплекса. Программный комплекс ориентирован на обработку непрерывного потока новостных сообщений, получаемых из интернет-источников. Сбор данных осуществляется автоматически путем извлечения финансовых новостей из открытых интернет-источников, таких как RSS-ленты информационных агентств и специализированные новостные агрегаторы.

Архитектуру комплекса определяют следующие функциональные модули (рис. 1):

1) модуль сбора новостных данных, предназначенный для автоматического получения финансовых новостей из интернет-источников и извлечения из них текста, временных меток и дополнительных метаданных публикаций;

2) модуль автоматического анализа тональности, формирующий первичную количественную оценку тональности финансовой новости на основе используемых моделей анализа тональности;

3) модуль пользовательских попарных сравнений, обеспечивающий сбор пользовательской разметки в форме выбора более положительной или более отрицательной финансовой новости из предложенной пары;

4) модуль построения обучающей выборки, предназначенный для преобразова-

ния результатов попарных сравнений в количественные оценки тональности, используемые для обучения моделей анализа тональности;

5) модуль обучения моделей анализа тональности, предназначенный для первичного и повторного обучения моделей по мере накопления пользовательской разметки и расширения обучающей выборки;

6) модуль оценки качества и управления обновлением моделей, предназначенный для вычисления метрик качества моделей анализа тональности и принятия решения о замене используемой модели на новую при улучшении значений выбранных метрик.



Рис. 1. Архитектура программного комплекса потокового анализа тональности финансовых новостей

Модульная архитектура предполагает возможность использования различных моделей анализа тональности и их метрик качества. В качестве моделей анализа тональности могут использоваться как традиционные методы машинного обучения, так и современные нейросетевые модели. В качестве метрик качества могут использоваться коэффициент детерминации, средняя квадратичная ошибка или корреляционные метрики.

Процесс обработки новостного потока. Программный комплекс функционирует в потоковом режиме обработки данных. Новостные сообщения автоматически извлекаются из интернет-источников и обрабатываются в порядке их публикации.

На начальном этапе обработки новостей, до накопления достаточного объема пользовательской разметки, используются методы анализа тональности, не требующие обучения на размеченных данных, включая предварительно обученные большие языковые модели. Одновременно с этим осуществляется сбор пользовательской разметки. Так как назначение пользователями количественных оценок тональности может быть затруднительно, для разметки используется процедура попарных сравнений финансовых новостей.

Пользователь выбирает новость с более положительной или более отрицательной тональностью из предложенной пары. По мере накопления сравнений одной и той же новости в различных парах определяется ее относительная позиция на шкале тональности. На основе накопленных результатов попарных сравнений формируются количественные оценки тональности, пригодные для обучения моделей анализа тональности.

После накопления достаточного объема пользовательской разметки выполняется первоначальное обучение модели анализа тональности, а затем – ее периодическое повторное обучение по мере накопления новых размеченных пользователями данных.

Для повышения устойчивости программного комплекса к изменению распределения данных новостного потока предполагается использовать механизм исключения устаревших данных из обучающей выборки. Обучение модели выполняется на скользящем временном окне новостных сообщений, что позволяет учитывать актуальный информационный контекст и снижать влияние устаревших данных в обучающей выборке.

Решение о повторном обучении модели по мере накопления пользовательской разметки принимается на основе сравнения метрик качества моделей анализа тональности. Новая модель используется только в случае улучшения значений выбранных метрик по сравнению с текущей моделью. В противном случае программный комплекс продолжает использовать ранее обученную модель.

Ключевой особенностью реализации программного комплекса потокового анализа тональности финансовых новостей является преобразование пользовательских попарных сравнений в количественные оценки тональности, используемые для формирования обучающей выборки. Такой подход позволяет формировать пользовательскую разметку непосредственно в процессе эксплуатации программного комплекса и выполнять повторное обучение моделей по мере ее накопления.

Архитектура также предусматривает механизмы управления повторным обучением моделей и исключения устаревших данных из обучающей выборки, что позволяет учитывать изменение распределения данных новостного потока во времени.

Перспективами дальнейших исследований являются разработка методов отбора наиболее информативных пар финансовых новостей для пользовательской разметки, а также исследование эффективности различных моделей в условиях потокового анализа финансовых новостей.

Литература

1. Integrating sentiment analysis with graph neural networks for enhanced stock prediction: A comprehensive survey / N. Das, B. Sadhukhan, R. Chatterjee, S. Chakrabarti // Decision Analytics Journal. – 2024. – Vol. 10. – Article 100417. – P. 1–21. – DOI 10.1016/j.dajour.2024.100417
2. Market-Derived Financial Sentiment Analysis: Context-Aware Language Models for Crypto Forecasting / H. Moradi-Kamali, M.-H. Rajabi-Ghazlou, M. Ghazavi [et al.] // arXiv preprint. – 2025. – arXiv:2502.14897. – P. 1–13. – DOI 10.48550/arXiv.2502.14897
3. Liu, J. A Survey of Financial AI: Architectures, Advances and Open Challenges / J. Liu // arXiv preprint. – 2024. – arXiv:2411.12747. – PP. 1–19. – DOI 10.48550/arXiv.2411.12747
4. Quantifying Semantic Shift in Financial NLP: Robust Metrics for Market Prediction Stability / Z. Sun, C. Xiao, A. Harit, J. Yu // arXiv preprint. – 2025. – arXiv:2510.00205. – P. 1–8. – DOI 10.48550/arXiv.2510.00205

ПОДГОТОВКА ОБУЧАЮЩИХ МАСОК УТРАТ МОНУМЕНТАЛЬНОЙ ЖИВОПИСИ НА ОСНОВЕ ЭКСПЕРТНОЙ РАЗМЕТКИ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

Г. А. Хришкевич

*Федеральное государственное бюджетное образовательное учреждение
высшего образования «Псковский государственный университет»,
Российская Федерация*

Научный руководитель В. И. Пименов

Рассмотрен технологический подход по подготовке обучающих масок утрат монументальной живописи на основе экспертной разметки цифровых изображений. В отличие от работ, где основной акцент делается на анализе уже сегментированных дефектов, в настоящем сообщении центральной задачей является формирование воспроизводимого набора пар «изображение – маска» для дальнейшего обучения нейросетевых моделей. Предлагаемый подход включает загрузку исходного изображения в Label Studio, многоклассовую экспертную разметку утрат, экспорт растровых масок по классам, их бинаризацию, контроль ошибок и организацию файловой структуры датасета. На примере размеченного изображения фрески показано получение бинарных масок по классам: «трещины», «пятна» и «лакуны». Полученные результаты подтверждают, что