УДК 621.3
DOI: 10.53015/18159958_2025_21_3_57

АНАЛИЗ КАСКАДНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ МНОГОУРОВНЕВОЙ КЛАССИФИКАЦИИ ТЕКСТОВ

ANALYSIS OF CASCADE NEURAL NETWORKS FOR MULTILEVEL CLASSIFICATION OF TEXTS

©2025

Погодаев Анатолий Кирьянович[1], доктор технических наук, профессор кафедры прикладной математики и системного анализа
Бухтояров Дмитрий Викторович[2], аспирант кафедры прикладной математики и системного анализа
Жбанова Наталья Юрьевна[3], кандидат технических наук, доцент кафедры прикладной математики и системного анализа
Pogodaev Anatoly Kiryanovich[1], Doctor of Technical Sciences, Professor of the Department of Applied Mathematics and systems analysis
Buchtoyarov Dmitry Viktorovich[2], postgraduate student of the Department of Applied Mathematics and systems analysis
Zhbanova Natalia Yuryevna[3], Candidate of Technical Sciences, Associate Professor of the Department of Applied Mathematics and systems analysis

Липецкий государственный технический университет, Липецк (Россия)
Lipetsk State Technical University, Lipetsk (Russia)

E-mail: pogodaev_ak@mail.ru [1]ORCID: https://orcid.org/0000-0002-4049-2199
E-mail: buhtoyarow.dim@yandex.ru [2]ORCID: https://orcid.org/0009-0001-0238-4509
E-mail: zhbanova_nyu@stu.lipetsk.ru [3]ORCID: https://orcid.org/0000-0001-7324-9841

Аннотация: Исследование посвящено анализу каскадных нейронных сетей (CNN, BERT, LSTM, GRU) для многоуровневой классификации текстов в иерархических структурах. На корпусе писем CNN достигла точности 75 %, остальные модели – 20 %. Выявлены ограничения, включая накопление ошибок. Предложены рекомендации для гибридных архитектур, адаптированных к иерархическим задачам, что повышает точность классификации.

Ключевые слова: каскадные нейронные сети, многоуровневая классификация, обработка текстов, CNN, BERT, LSTM, GRU, иерархические структуры, накопление ошибок, гибридные архитектуры.

Для цитирования: Погодаев, А.К. Анализ каскадных нейронных сетей для многоуровневой классификации текстов / А.К. Погодаев, Д.В. Бухтояров, Н.Ю. Жбанова // Вести высших учебных заведений Черноземья. – 2025. – Т. 21, № 3 (79). – С. 57–62. DOI 10.53015/18159958_2025_21_3_57.

Abstract: This study analyzes cascade neural networks (CNN, BERT, LSTM, GRU) for multi-level text classification in hierarchical structures. Evaluated on a letter corpus, CNN achieved 75 % accuracy, while others reached 20 %. Limitations, such as error accumulation, were identified. Recommendations for hybrid architectures tailored to hierarchical tasks are proposed to enhance classification accuracy.

Keywords: cascade neural networks, multi-level classification, text processing, CNN, BERT, LSTM, GRU, hierarchical structures, error accumulation, hybrid architectures.

For citation: Pogodaev A.K. Analysis of cascade neural networks for multilevel classification of texts / A.K. Pogodaev, D.V. Buchtoyarov, N.Yu. Zhbanova // Vesti of Higher Educational Institutions of the Chernozem region. – 2025. – Vol. 21, № 3 (79). – P. 57–62. DOI 10.53015/18159958_2025_21_3_57.

ВВЕДЕНИЕ

Современные организации сталкиваются с необходимостью обработки значительных объемов текстовой информации, поступающей в виде писем, запросов и отчетов [1]. В условиях иерархических структур традиционные подходы к классификации текстов, основанные на ручном распределении документов, демонстрируют низкую эффективность, что приводит к временным задержкам и увеличению числа ошибок [2]. В последние годы активно исследуются методы машинного обучения, в частности нейронные сети, для автоматизации таких процессов [3]. Среди них каскадные архитектуры представляют собой перспективное направление благодаря их способности моделировать иерархические зависимости в данных [4].

Целью данной работы является анализ применения четырех известных моделей нейронных сетей – CNN, BERT, LSTM и GRU – в задаче многоуровневой классификации текстов. Исследование сосредоточено на оценке эффективности этих моделей и выявлении их ограничений, а не на разработке новой архитектуры. Тем не менее полученные результаты позволяют обосновать направления для будущих исследований. Эксперименты проводились на основе реальных данных организации, включающих тексты писем в формате PDF и метки их распределения по уровням иерархии.

ТЕОРИЯ ВОПРОСА

Многоуровневая классификация текстов предполагает автоматическое распределение документов по иерархическим уровням исполнителей в организационной структуре. В данном исследовании иерархия представлена тремя уровнями: высший (например, руководитель), средний (например, заместитель руководителя) и нижний (например, исполнитель). Использованный набор данных состоит из текстов писем в формате PDF, снабженных метками, указывающими соответствующий уровень маршрутизации.

Каждая из четырех моделей (CNN, BERT, LSTM, GRU) была реализована в виде каскадной структуры, состоящей из трех последовательных уровней классификации. На каждом уровне модель предсказывала принадлежность текста к определенной категории, а результат передавался следующему уровню как входные данные. Такой подход позволяет моделировать иерархические зависимости, однако приводит к накоплению ошибок между уровнями [6].

CNN применяют сверточные фильтры для извлечения локальных признаков из текстовых данных, представленных в виде матрицы эмбеддингов слов. Операция свертки описывается как:

$$y_i = f\left(\sum_{j=1}^{k} W_j \cdot x_{i+j-1} + b\right), \qquad (1)$$

где y_i – выходной сигнал фильтра, W– веса фильтра размером k, x – входной вектор, b – смещение, а f – функция активации (в данном случае ReLU). За сверткой следует операция максимального пулинга, которая выделяет наиболее значимые признаки из сверточных карт, уменьшая размерность данных и повышая устойчивость к шуму [7].

Модель BERT, основанная на трансформерной архитектуре, использует двунаправленное представление текста с механизмом многослойного внимания:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (2)$$

где Q, K, V – матрицы запросов, ключей и значений соответственно, а d_k –размерность ключей. Модель дообучалась на задачу классификации с добавлением линейного слоя [8].

Сети LSTM сохраняют долгосрочные зависимости в последовательностях благодаря механизмам памяти, включающим забывающий, входной и выходной вентили [9]. За-

бывающий вентиль определяет, какую информацию из предыдущего состояния забыть:

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right), \quad (3)$$

где f_t – забывающий вентиль, определяет, какую информацию из предыдущего состояния C_{t-1} следует забыть.

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right), \quad (4)$$

где i_t – входной вентиль, определяет, какую новую информацию добавить в состояние.

$$\tilde{C}_t = \tanh\left(W_C \cdot \left[h_{t-1}, x_t\right] + b_C\right), \quad (5)$$

где $\tilde{C}_t$ – кандидат на обновление памяти, предлагает новое содержимое для ячейки памяти.

$$C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t, \quad (6)$$

где C_t – обновление памяти, комбинирует старое состояние с новым.

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right), \quad (7)$$

где o_t – выходной вентиль определяет, какая часть состояния будет выведена.

$$h_t = o_t \circ \tanh\left(C_t\right), \quad (8)$$

где h_t – выход, формирует скрытое состояние для передачи дальше.

Остальные обозначения: σ – сигмоидная функция, tanh – гиперболический тангенс, $\circ$ – поэлементное умножение, W и b – обучаемые веса и смещения. Обновление состояния происходит через вычисление нового состояния памяти C_t и скрытого состояния h_t по приведенным уравнениям.

GRU представляет собой упрощенную версию LSTM с двумя вентилями (обновляющим и сбрасывающим), что ускоряет обучение при сохранении способности моделировать последовательности [10]. Обновление состояния включает:

$$z_t = \sigma\left(W_z \cdot h_{t-1}, x_t\right] + b_z), \quad (9)$$

где z_t – обновляющий вентиль, определяет, сколько информации от предыдущего состояния h_{t-1} сохранить.

$$r_t = \sigma\left(W_r \cdot h_{t-1}, x_t\right] + b_r), \quad (10)$$

где r_t – сбрасывающий вентиль, контролирует, какую часть предыдущего состояния сбросить при вычислении кандидата.

$$\tilde{h}_t = \tanh\left(W_h \cdot \left[r_t \circ h_{t-1}, x_t\right] + b_h\right), \quad (11)$$

где $\tilde{h}_t$ – кандидат на состояние, предлагает новое скрытое состояние.

$$h_t = \sigma\left(1 - z_t\right) \circ h_{t-1} + z_t \circ \tilde{h}_t, \quad (12)$$

где h_t – новое состояние, линейно комбинирует старое и новое состояния. Обозначения аналогичны LSTM: $\sigma, \tanh, \circ, W, b$.

В работе [14] подчеркивается эффективность CNN для задач классификации текстов, но отмечается их ограниченная способность обрабатывать длинные последовательности. Исследование [15] демонстрирует высокую точность BERT в задачах обработки текстов, однако указывает на необходимость больших обучающих корпусов. Авторы [16] отмечают, что LSTM и GRU подходят для последовательных данных, но их производительность снижается в многоуровневых задачах. Настоящее исследование подтверждает эти выводы и выявляет дополнительные проблемы, связанные с каскадными архитектурами.

МЕТОДЫ ИССЛЕДОВАНИЯ

Качество предобработки данных существенно влияет на производительность моделей машинного обучения [5]. В рамках исследования применялись следующие этапы подготовки данных:

– извлечение текста из PDF-файлов;

– удаление служебных заголовков и приложений;

– очистка текста: удаление цифр, специальных символов и латинских букв;

– приведение текста к нижнему регистру;

– фильтрация слов по длине (исключение слов короче 5 символов).

Эксперименты проводились на корпусе из 500 писем, разделенных на обучающую (80 %) и тестовую (20 %) выборки. После начального тестирования корпус был расширен до 3000 примеров для дообучения лучшей модели. Метрика оценки – точность классификации (accuracy) на каждом уровне.

Использовались четыре модели:

– CNN, начальная точность составила 36 % , после дообучения – 75 %.

– BERT, начальная точность – 20%.

– LSTM, начальная точность – 20%.

– GRU, начальная точность – 20%.

CNN показала наилучшие результаты благодаря способности эффективно извлекать локальные признаки, однако теряет глобальный контекст в длинных текстах [11]. BERT требует значительных вычислительных ресурсов и большого объема данных для дообучения, что ограничивает ее применимость [12]. LSTM и GRU демонстрируют низкую точность на длинных последовательностях из-за ограничений памяти и накопления ошибок в каскадной структуре [13].

ЗАКЛЮЧЕНИЕ

Проведённый анализ показал, что CNN обладает наивысшей эффективностью (точность 75 % после дообучения) среди рассмотренных моделей для многоуровневой классификации текстов. Однако все протестированные архитектуры – CNN, BERT, LSTM и GRU имеют существенные ограничения: потеря контекста, высокая ресурсоёмкость и накопление ошибок. Эти выводы подчёркивают необходимость разработки гибридных архитектур, адаптированных к задачам иерархической классификации текстов.

Совершенствование архитектур через гибридные подходы улучшит качество классификации и повысит эффективность маршрутизации документов в иерархических организациях. Решение проблем, связанных с потерей контекста и ресурсоёмкостью, возможно за счёт интеграции гибридных подходов, таких как комбинация свёрточных и рекуррентных слоёв или внедрение механизмов внимания, что позволит моделям лучше адаптироваться к сложным иерархическим структурам.

Подтверждена ключевая роль предобработки данных в повышении точности моделей. Улучшение методов предобработки, таких как точное извлечение текста из PDF-файлов или применение продвинутых техник нормализации, способствует росту производительности.

Кроме того, гибридные подходы и механизмы внимания, используемые в трансформерных архитектурах, улучшат способность моделей выделять ключевые зависимости в тексте, сокращая временные затраты и ошибки при маршрутизации документов, что особенно важно при больших объёмах корреспонденции.

Перспективы будущих исследований включают:

– разработку гибридных моделей, объединяющих преимущества CNN и трансформеров для обработки как локальных, так и глобальных зависимостей в текстах;

– исследование методов регуляризации и оптимизации, направленных на снижение накопления ошибок в каскадных архитектурах;

– применение активного обучения для минимизации зависимости от больших размеченных корпусов данных;

– внедрение интерпретируемых механизмов, таких как внимание, для повышения прозрачности и доверия к моделям в реальных системах.

Адаптация моделей к особенностям организационных текстов (формальным конструкциям, шаблонам) и глубокий анализ структуры данных создадут основу для специализированных систем обработки текстов.

Данное исследование вносит вклад в развитие методов многоуровневой классификации текстов и открывает перспективы для создания более эффективных и устойчивых моделей, адаптированных к потребностям организаций с иерархическими структурами.

Полученные результаты и предложенные направления исследований служат основой для дальнейшего совершенствования технологий автоматизации, обеспечивая повышение производительности и точности в реальных приложениях.

СПИСОК ЛИТЕРАТУРЫ

1. Goodfellow I., Bengio Y., Courville A. Deep Learning. Cambridge : MIT Press, 2016. 800 p.
2. Manning C.D., Raghavan P., Schütze H. Introduction to Information Retrieval. Cambridge : Cambridge University Press, 2008. 496 p.
3. LeCun Y., Bengio Y., Hinton G. Deep learning. DOI 10.1038/nature14539 // Nature. 2015. Vol. 521, № 7553. P. 436-444.
4. Kim Y. Convolutional Neural Networks for Sentence Classification. DOI 10.3115/v1/D14-1181 // Proceedings of EMNLP, Doha, Qatar. 2014. P. 1746-1751.
5. Brownlee J. Data Preparation for Machine Learning. [S.l.] : Machine Learning Mastery, 2020. 398 p.
6. Bahdanau D., Cho K., Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate // International Conference on Learning Representations (ICLR), San Diego, USA. 2015. 15 p.
7. Zhang Y., Wallace B.C. A Sensitivity Analysis of Convolutional Neural Networks for Sentence Classification // Journal of Machine Learning Research. 2017. Vol. 18, No. 1. P. 1-32.
8. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin, M.-W. Chang, K. Lee, K. Toutanova. DOI 10.18653/v1/N19-1423 // Proceedings of NAACL-HLT, Minneapolis, USA. 2019. P. 4171-4186.
9. Hochreiter S., Schmidhuber J. Long Short-Term Memory. DOI 10.1162/neco.1997.9.8.1735 // Neural Computation. 1997. Vol. 9, No. 8. P. 1735-1780.
10. Learning Phrase Representations Using RNN Encoder-Decoder / K. Cho, B. van Merriënboer, C. Gulcehre [et al.]. DOI 10.3115/v1/D14-1179 // Proceedings of EMNLP, Doha, Qatar. 2014. P. 1724-1734.
11. Liu P., Qiu X., Huang X. Recurrent Neural Network for Text Classification with Multi-Task Learning // Proceedings of IJCAI, New York, USA. 2016. P. 2873-2879.
12. Attention Is All You Need / A. Vaswani, N. Shazeer, N. Parmar [et al.] // Advances in Neural Information Processing Systems (NeurIPS), Long Beach, USA. 2017. P. 5998-6008.
13. LSTM: A Search Space Odyssey / K. Greff, R.K. Srivastava, J. Koutník [et al.]. DOI 10.1109/TNNLS.2016.2582924 // IEEE Transactions on Neural Networks and Learning Systems. 2017. Vol. 28, No. 10. P. 2222-2232.
14. Very Deep Convolutional Networks for Text Classification / A. Conneau, H. Schwenk, L. Barrault, Y. Lecun. DOI 10.18653/v1/E17-1104 // Proceedings of EACL, Valencia, Spain. 2017. P. 1107-1116.
15. How to Fine-Tune BERT for Text Classification? / C. Sun, X. Qiu, Y. Xu, X. Huang // Proceedings of CCL, Beijing, China. 2019. P. 194-206.
16. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling / J.Chung, C. Gulcehre, K.H. Cho, Y. Bengio // NIPS Workshop on Deep Learning, Montreal, Canada. 2014. 9 p.

REFERENCES

1. Goodfellow I., Bengio Y., Courville A. Deep Learning. Cambridge : MIT Press, 2016. 800 p.
2. Manning C.D., Raghavan P., Schütze H. Introduction to Information Retrieval. Cambridge : Cambridge University Press, 2008. 496 p.
3. LeCun Y., Bengio Y., Hinton G. Deep learning. DOI 10.1038/nature14539 // Nature. 2015. Vol. 521, No. 7553. P. 436-444.
4. Kim Y. Convolutional Neural Networks for Sentence Classification. DOI 10.3115/v1/D14-1181 // Proceedings of EMNLP, Doha, Qatar. 2014. P. 1746-1751.

5. Brownlee J. Data Preparation for Machine Learning. [S.l.] : Machine Learning Mastery, 2020. 398 p.
6. Bahdanau D., Cho K., Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate // International Conference on Learning Representations (ICLR), San Diego, USA. 2015. 15 p.
7. Zhang Y., Wallace B.C. A Sensitivity Analysis of Convolutional Neural Networks for Sentence Classification // Journal of Machine Learning Research. 2017. Vol. 18, No. 1. P. 1-32.
8. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin, M.-W. Chang, K. Lee, K. Toutanova. DOI 10.18653/v1/N19-1423 // Proceedings of NAACL-HLT, Minneapolis, USA. 2019. P. 4171-4186.
9. Hochreiter S., Schmidhuber J. Long Short-Term Memory. DOI 10.1162/neco.1997.9.8.1735 // Neural Computation. 1997. Vol. 9, No. 8. P. 1735-1780.
10. Learning Phrase Representations Using RNN Encoder-Decoder / K. Cho, B. van Merriënboer, C. Gulcehre [et al.]. DOI 10.3115/v1/D14-1179 // Proceedings of EMNLP, Doha, Qatar. 2014. P. 1724-1734.
11. Liu P., Qiu X., Huang X. Recurrent Neural Network for Text Classification with Multi-Task Learning // Proceedings of IJCAI, New York, USA. 2016. P. 2873-2879.
12. Attention Is All You Need / A. Vaswani, N. Shazeer, N. Parmar [et al.] // Advances in Neural Information Processing Systems (NeurIPS), Long Beach, USA. 2017. P. 5998-6008.
13. LSTM: A Search Space Odyssey / K. Greff, R.K. Srivastava, J. Koutník [et al.]. DOI 10.1109/TNNLS.2016.2582924 // IEEE Transactions on Neural Networks and Learning Systems. 2017. Vol. 28, No. 10. P. 2222-2232.
14. Very Deep Convolutional Networks for Text Classification/ A. Conneau, H. Schwenk, L. Barrault, Y. Lecun. DOI 10.18653/v1/E17-1104 // Proceedings of EACL, Valencia, Spain. 2017. P. 1107-1116.
15. How to Fine-Tune BERT for Text Classification? / C. Sun, X. Qiu, Y. Xu, X. Huang // Proceedings of CCL, Beijing, China. 2019. P. 194-206.
16. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling / J.Chung, C. Gulcehre, K.H. Cho, Y. Bengio // NIPS Workshop on Deep Learning, Montreal, Canada. 2014. 9 p.