УДК 004.032.26

АНАЛИЗ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ С ИСПОЛЬЗОВАНИЕМ МАЛЫХ ЯЗЫКОВЫХ МОДЕЛЕЙ

К. В. АНДРЕНКО[1)]***, А. А. КРОЩЕНКО***[1)]***, В. А. ГОЛОВКО***[1)]***, Ю. Б. КРАПИВИН***[1)]

[1)]*Брестский государственный технический университет,*
ул. Московская, 267, 224017, г. Брест, Беларусь

Аннотация. Предложен метод решения задачи автоматической оценки качества управленческих решений по текстовым ответам на русском языке, соотнесенным с проблемными ситуациями, на базе нейросетевой модели DistilBERT. Новизна исследования заключается в разработке алгоритма контекстно зависимой классификации, при котором входные данные формируются как конкатенация описания ситуации и ответа респондента, а также в использовании методики постобработки предсказаний, предполагающей взаимосвязь результатов с конкретными управленческими компетенциями или красными флагами. Достигнуты высокие показатели (accuracy = 0,986 8, F1-score = 0,987 0) при оценке управленческих решений на русском языке с интерпретируемым выводом зон профессионального и личностного роста, в то время как известные работы ориентированы на общий анализ текста или его тональности. Предложенный подход открывает возможность применения указанных алгоритма и методики при разработке систем автоматизированной оценки руководящих работников и подбора персонала, а также персонализированного развития управленческих компетенций в практике найма и корпоративном обучении.

Ключевые слова: малая языковая модель; DistilBERT; анализ управленческих решений; нейронная сеть; дистилляция знаний.

Образец цитирования:
Андренко КВ, Крощенко АА, Головко ВА, Крапивин ЮБ. Анализ управленческих решений с использованием малых языковых моделей. *Журнал Белорусского государственного университета. Математика. Информатика.* 2026;1:121–129 (на англ.).
EDN: WVIPLU

For citation:
Andrenko KV, Kroshchanka AA, Golovko VA, Krapivin YuB. Managerial decision analysis using small language models. *Journal of the Belarusian State University. Mathematics and Informatics.* 2026;1:121–129.
EDN: WVIPLU

Авторы:
Ксения Владимировна Андренко – преподаватель-стажер кафедры интеллектуальных информационных технологий факультета электронно-информационных систем.
Александр Александрович Крощенко – кандидат технических наук; доцент кафедры интеллектуальных информационных технологий факультета электронно-информационных систем.
Владимир Адамович Головко – доктор технических наук, профессор; заведующий кафедрой интеллектуальных информационных технологий факультета электронно-информационных систем.
Юрий Борисович Крапивин – кандидат технических наук; доцент кафедры интеллектуальных информационных технологий факультета электронно-информационных систем.

Authors:
Ksenia V. Andrenko, trainee lecturer at the department of intelligent information technologies, faculty of electronic and information systems.
ksenia.andrenko2002@gmail.com
https://orcid.org/0009-0001-0976-2613
Aliaksandr A. Kroshchanka, PhD (engineering); associate professor at the department of intelligent information technologies, faculty of electronic and information systems.
kroschenko@gmail.com
https://orcid.org/0000-0003-3285-3545
Vladimir A. Golovko, doctor of science (engineering), full professor; head of the department of intelligent information technologies, faculty of electronic and information systems.
vladimir.golovko@gmail.com
https://orcid.org/0000-0003-2615-289X
Yury B. Krapivin, PhD (engineering); associate professor at the department of intelligent information technologies, faculty of electronic and information systems.
ybox@list.ru
https://orcid.org/0009-0004-8489-6912

Журнал Белорусского государственного университета. Математика. Информатика. 2026;1:121–129
Journal of the Belarusian State University. Mathematics and Informatics. 2026;1:121–129

MANAGERIAL DECISION ANALYSIS USING SMALL LANGUAGE MODELS

K. V. ANDRENKO[a], *A. A. KROSHCHANKA*[a], *V. A. GOLOVKO*[a], *Yu. B. KRAPIVIN*[a]

[a]*Brest State Technical University, 267 Maskowskaja Street, Brest 224017, Belarus*

Corresponding author: K. V. Andrenko (ksenia.andrenko2002@gmail.com)

Abstract. This paper proposes a method for the automatic assessment of managerial decision quality based on free-form textual responses to Russian-language business cases using a fine-tuned DistilBERT model. The novelty of the study lies in a context-dependent classification algorithm with concatenated input (case description + response) and a post-processing technique that explicitly links predictions to specific managerial competencies or red flags. For the first time, high-accuracy evaluation of managerial decisions in Russian (accuracy = 0.986 8, F1-score = 0.987 0) with interpretable competency-based feedback has been achieved, unlike prior work focused on general text or sentiment analysis. The approach enables the development of automated systems for managerial assessment, personnel selection and personalised leadership development in human resources and corporate training.

Keywords: small language model; DistilBERT; management decision analysis; neural network; knowledge distillation.

Introduction

In the era of big data, the integration of artificial intelligence systems into all spheres of life, including economics, education, culture and management, has become particularly significant [1–5]. Managerial decision-making requires processing large volumes of both qualitative and quantitative information at various stages (for instance, such as during recruitment or performance appraisal to identify areas for professional growth). The sheer amount of unstructured data that cannot be handled manually has created a need for tools capable of interpreting free-form responses in the context of solving real-life managerial tasks [6]. The automatic evaluation of the quality of managerial decisions in the form of textual responses when solving case studies is therefore especially relevant.

Existing research primarily focuses on general text analysis rather than the specific assessment of managerial decision quality [7–9]. For example, the importance of applying advances in artificial intelligence and natural language processing, such as structural topic modelling, in human resources (HR) context is highlighted [10–16].

One of the most promising directions in natural language processing is the use of transformer-based models, such as BERT (*bidirectional encoder representations from transformers*) and its distilled variant, DistilBERT. The advantages of BERT are high accuracy, context awareness, and the ability to fine-tune only a single output layer to create models that solve the task of text sequence classification (for example, the classification of managerial decisions) [17–21].

DistilBERT preserves the performance of the original BERTModel while requiring substantially fewer computational resources. These features make DistilBERT a suitable model for practical applications demanding both high accuracy and low latency[1] [22; 23]. The authors previously investigated the use of distilled language models for sentiment analysis, demonstrating their effectiveness in that domain [24].

The aim of this study is to develop and evaluate a novel method for the automatic assessment of managerial decision quality based on free-form textual responses to Russian-language cases. To achieve this aim, the following tasks were addressed:

1) construction of a specialised Russian-language dataset, comprising 760 annotated examples of managerial cases with paired high- and low-quality responses, enriched with structured competency scores and red flags;

2) fine-tuning of a DistilBERT model for context-dependent binary classification by concatenating case descriptions and responses as input sequences;

3) comprehensive evaluation of the proposed approach using accuracy, F1-score and stratified cross-validation metrics, alongside comparisons with existing solutions.

In the present study, the DistilBERT model was fine-tuned for the binary classification of managerial decisions while explicitly accounting for case-specific context. The input sequence was formed by concatenating the case description and the respondent's textual response, thereby enabling the model to capture causal relationships and pragmatic dependencies between the situation and the proposed solution. The achieved performance metrics

[1]DistilBERT // Hugging Face : website. URL: https://huggingface.co/docs/transformers/model_doc/distilbert (date of access: 19.12.2024).

(accuracy = 0.986 8, F1-score = 0.987 0) demonstrates that the proposed context-aware approach delivers excellent classification accuracy and can serve as a foundation for deployable intelligent decision-support systems in real-world managerial and HR practices.

Materials and methods

Dataset preparation. A dedicated Russian-language dataset was constructed for the binary classification of managerial decisions. It comprises 760 annotated examples corresponding to 380 unique managerial cases (`Case_ID 1 – Case_ID 380`). For each case, two contrasting responses were created: one high-quality response (`Is_Good = 1`) and one low-quality response (`Is_Good = 0`), yielding a perfectly balanced dataset with no class imbalance.

The responses were initially generated using three large language models (Claude 3.5 Sonnet, Grok 3 and Gemini 2.0 Pro) and subsequently underwent thorough manual revision and stylistic refinement by a professional linguist to ensure linguistic naturalness, stylistic diversity, and alignment with real-world managerial discourse.

The cases cover a broad spectrum of managerial domains typically encountered by executives at municipal and corporate levels, including infrastructure (street lighting, sidewalks, bike lanes, waste containers), social policy (support for families, elderly citizens and youth), healthcare (service conditions, medical equipment, telemedicine), education (textbooks, extracurricular activities, career guidance), public safety (vandalism, accidents, theft), ecology (waste separation, urban greening), culture and leisure (festivals, museums, parks), and transport and housing services (traffic congestion, snow removal, elevator maintenance).

Representative examples include the situation «В районе участились случаи вандализма на детских площадках» («Vandalism on children's playgrounds has recently increased in the district») and the situation «В компании сотрудники жалуются на отсутствие гибкого графика» («Employees are complaining about the lack of flexible working hours»).

The dataset is stored in CSV format (semicolon-separated) with the following fields:

- `Case_ID` – unique case identifier (integer);
- `Situation` – textual description of the managerial problem or situation (string, mean length ≈78 tokens);
- `Response` – free-form textual response proposing a managerial decision (string);
- `Is_Good` – binary label (1 – good response, 0 – poor response);
- `Scores` – dictionary containing two competency scores (scale 3–5 points) for good responses and two competency scores (scale 1–2 points) for poor responses; stored as string representation of Python dictionary (e. g., `"{'Инициатива и ответственность': 4, 'Планирование и организация': 5}"`);
- `Red_Flags` – dictionary containing two red flag scores: low values (1–2 points) indicating minimal shortcomings for good responses, or high values (3–5 points) indicating severe shortcomings for poor responses; stored as string representation of Python dictionary (e. g., `"{'Бездействие и перекладывание ответственности': 4, 'Игнорирование жалоб граждан': 3}"`).

The inclusion of structured `Scores` and `Red_Flags` dictionaries enables not only supervised classification but also direct post-processing of model predictions into interpretable competency-based feedback and personalised development recommendations – a feature rarely found in existing text-classification datasets.

The key distinguishing feature of the dataset is the integration of structured competency scores and red flags, which extends its utility far beyond conventional binary classification datasets. This dual structure enables (i) standard supervised training, (ii) direct post-processing of model predictions into human-readable competency profiles and personalised development recommendations, and (iii) detailed interpretability analysis. These capabilities are rarely encountered in publicly available text classification corpora.

Input representation. To ensure context-aware classification, the input sequence was constructed by concatenating the case description and the respondent's textual response, separated by the `[SEP]` token:

```
input_text = [CLS] + Situation + [SEP] + Response + [SEP].
```

Tokenisation was performed using the original DistilBERT tokeniser from the Hugging Face Transformers library (version 4.41.2) with the following settings: `maximum sequence length = 512 tokens, truncation = True, padding = "max_length"`. This approach enables the model to capture causal and pragmatic relationships between the situation and the proposed solution. For illustration, the response «I will establish a task force» may represent a good decision for the case «Residents complain about poor street lighting» but a poor decision for the case «Urgent evacuation due to fire».

Training protocol. The DistilBERT model (66 mln parameters) was fine-tuned in Google Colab using a single NVIDIA T4 GPU (VRAM – 16 GB). This hardware configuration provided sufficient memory and computational performance for the selected hyperparameters.

The dataset (760 instances) was stratified split into training (608 instances, or 80 %), validation (76 instances, or 10 %) and test (76 instances, or 10 %) sets.

All random seeds (PyTorch, NumPy, Python random) were fixed at 42 to ensure full reproducibility.

Fine-tuning was performed for three epochs using AdamW optimiser and the following hyperparameters: learning rate $2 \cdot 10^{-5}$ (linear scheduler with warmup over the first 10 % of steps), batch size 16, weight decay 0.01, gradient clipping 1.0.

Cross-entropy loss was used as the loss function, which is minimised during training to predict the `Is_Good` labels:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log\left(\tilde{y}_i\right) + \left(1-y_i\right)\log\left(1-\tilde{y}_i\right)\right],$$

where N is the number of samples; y_i is the ground-truth label; $\tilde{y}_i$ is the predicted probability of the positive class.

Model architecture. The fine-tuned model consists of the pre-trained DistilBERT encoder (6 transformer layers: hidden size 768, 12 attention heads) followed by a binary classification head: a dropout layer ($p = 0.1$) and a fully connected layer with softmax activation (fig. 1).

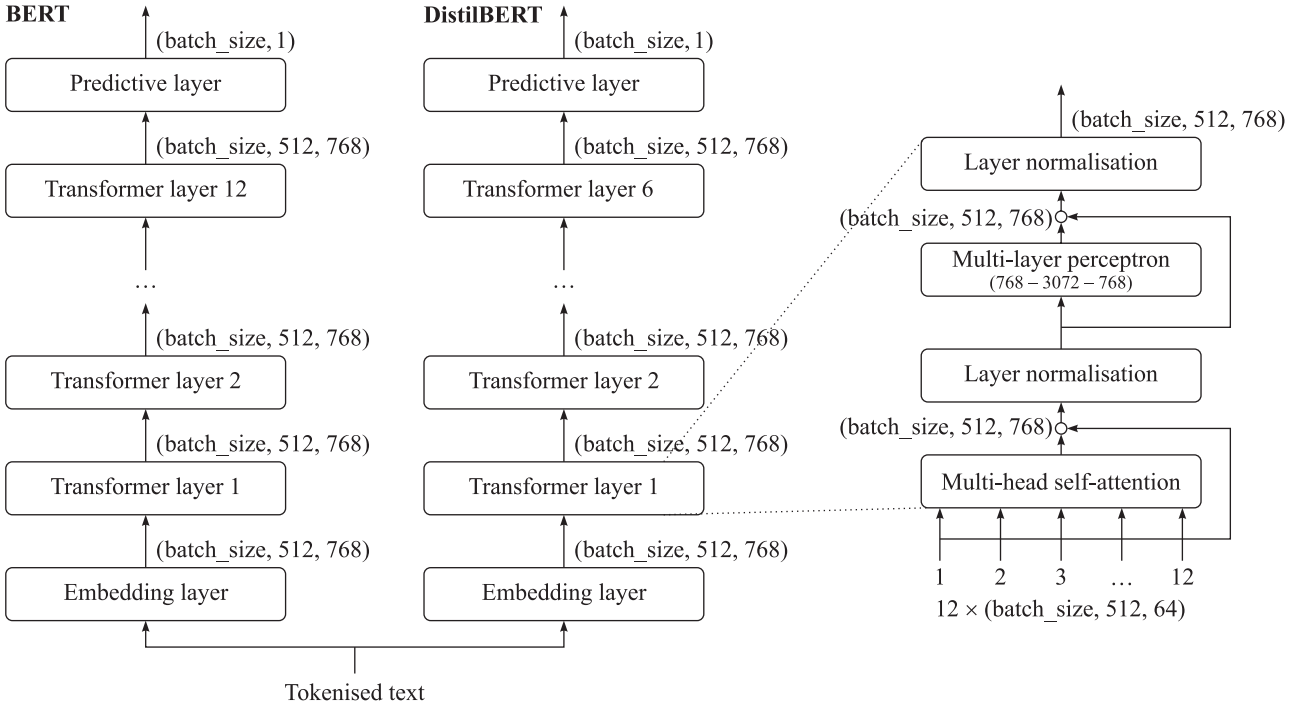

Fig. 1. Architectural comparison of BERT and DistilBERT

Each transformer layer includes the standard components: multi-head self-attention, multi-layer perceptron, layer normalisation and residual connections.

Compared with BERT-base (12 transformer layers, 110 mln parameters), DistilBERT retains approximately 96 % of the original performance while being 40 % smaller and 45 % faster in inference.

The predictive layer is a fully connected layer with one neuron and a softmax activation function, which makes the final decision on the evaluation of the decision made by the manager.

Interpretability via post-processing. After inference, predictions were enriched through deterministic post-processing:

- if `Is_Good = 1`, then the corresponding `Scores` dictionary is returned, highlighting two demonstrated competencies with high scores (3–5 points) from the universal competency framework (e. g., `"{'Инициатива и ответственность': 4, 'Планирование и организация': 5}"`);
- if `Is_Good = 0`, then the corresponding `Red_Flags` dictionary is returned, highlighting two critical shortcomings with high severity scores (3–5 points) from the red flags taxonomy (e. g., `"{'Бездействие и перекладывание ответственности': 4, 'Игнорирование жалоб граждан': 3}"`).

Thus, this rule-based post-processing layer transforms raw probabilistic outputs into immediately actionable, human-readable competency profiles and personalised development recommendations, significantly enhancing practical utility and interpretability of the system.

The classification head consists of a single linear layer applied to the `[CLS]` token representation, followed by softmax activation. This produces probabilities for the two classes (`Is_Good = 1 "good decision"; Is_Good = 0 "poor decision"`).

To evaluate the model, two metrics were employed – accuracy (0.986 8 in our case) and F1-score (0.987 0) calculated using the following formula:

$$\mathrm{F1} = 2 \cdot \frac{\mathrm{precision} \cdot \mathrm{recall}}{\mathrm{precision} + \mathrm{recall}},$$

where $\mathrm{precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$; $\mathrm{recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$; TP, FP and FN denote the number of true positives, false positives and false negatives, respectively.

The experiments utilised 5-fold stratified cross-validation with data shuffling and a fixed random state (`random_state=42`) to ensure reproducibility. The folds were stratified by the target class `Is_Good`, maintaining an equal proportion of labels 1 and 0 in each fold. This approach enables a reliable assessment of the model's generalisation ability and mitigates the risk of overfitting.

Results and discussion

On the independent test set (76 instances), the model achieved the following metrics: accuracy = 0.986 8, F1-score = 0.987 0. The confusion matrix (fig. 2) shows only one misclassification (one false positive), confirming the high accuracy of the model. The receiver operating characteristic (ROC) curve (fig. 3) has an area under the curve (AUC) of 0.998 6, significantly outperforming the random classifier line (0.5), which indicates excellent discriminative ability.

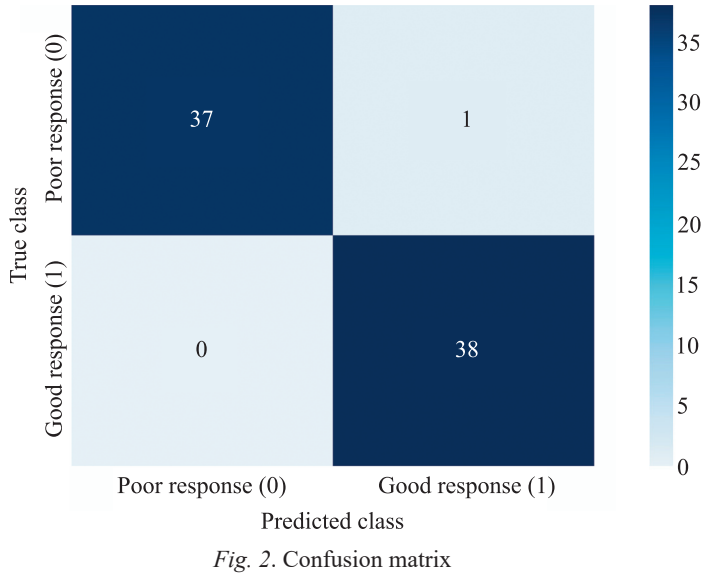

Fig. 2. Confusion matrix

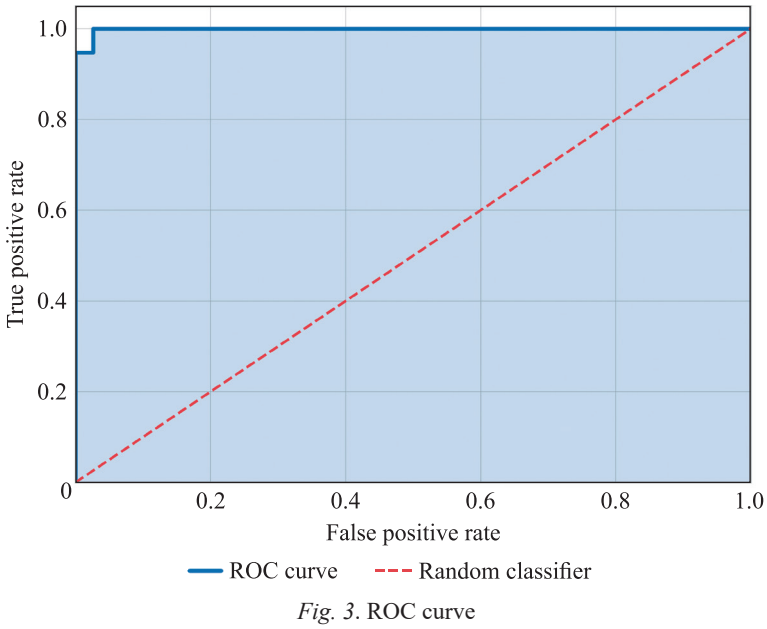

Fig. 3. ROC curve

The average metrics of the 5-fold stratified cross-validation are shown in the table 1. Detailed per-fold metrics are shown in fig. 4. The lowest performance is observed in the third fold (accuracy = 0.927 6, F1-score = 0.922 0), possibly due to the distribution of more challenging examples in that subset. Nevertheless, the overall stability across folds confirms the model's strong generalisation capability.

Table 1

The average metrics of the 5-fold stratified cross-validation

Metrics	Value
Accuracy	0.971 1 ± 0.023 0
F1-score	0.970 0 ± 0.025 2
Precision	0.981 9 ± 0.013 0
Recall	0.960 5 ± 0.054 6

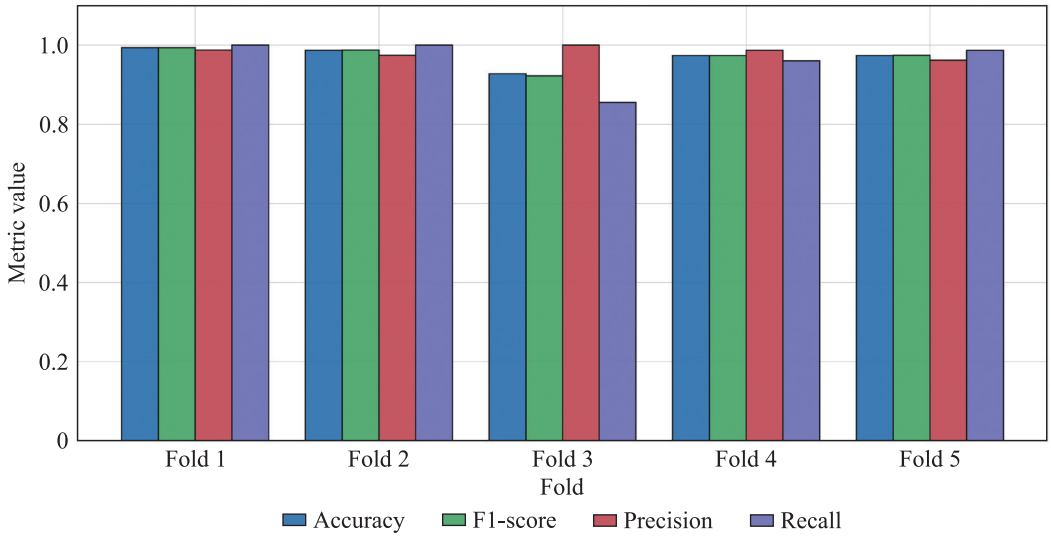

Fig. 4. Metrics across each fold

The histogram of predicted probabilities for the «Good response» class (fig. 5) demonstrates clear class separation: most good responses received probabilities >0.95, while poor responses received probabilities <0.05. This indicates high model confidence in predictions and a low risk of ambiguous classifications.

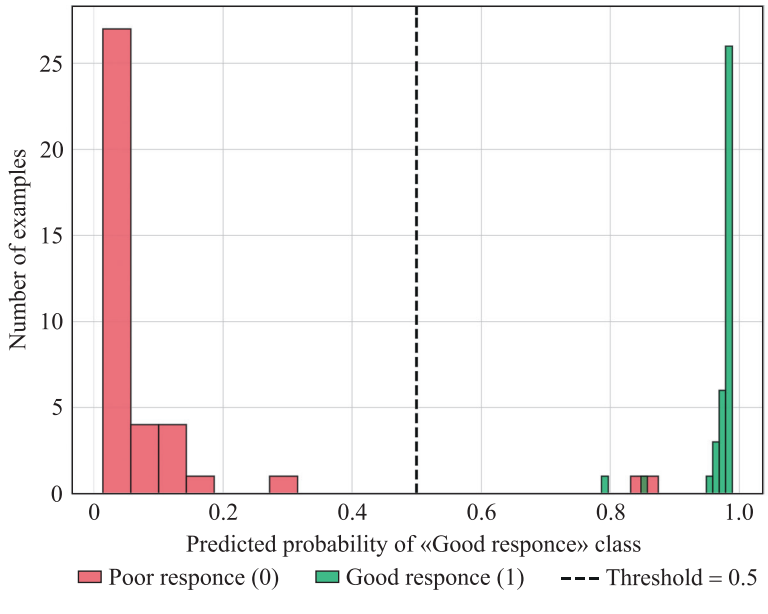

Fig. 5. Distribution of predicted probabilities

Example of model inference for `Case_ID 71` (good response)

Input: "`В районе участились случаи граффити на общественных зданиях. [SEP] Организую конкурс уличного искусства, выделю легальные стены и запущу профилактическую работу с подростками.`"

Model output: `Is_Good = 1` (probability 0.987)

Interpretation: "`Good response. Competencies demonstrated: Профилактика и долгосрочный подход (5), Инновационность и оптимизация процессов (5).`"

Table 2

Test set results

Metrics	Value
Accuracy	0.986 8
F1-score	0.987 0
Precision	0.974 4
Recall	1.000 0
ROC-AUC	0.998 6

The obtained metrics significantly exceed typical values for contextual text classification tasks in Russian and approach the upper limits achievable on synthetic balanced datasets. The high accuracy (above 0.97 even in cross-validation) validates the effectiveness of the proposed context-aware approach with concatenation of the situation and response. The deterministic post-processing of predictions adds practical value, transforming binary classification into a tool for generating personalised competency-based feedback. The slight drop in the third fold highlights a potential improvement area – possibly through data augmentation or model ensembling.

Comparison with existing solutions

The task could also be addressed using the full-sized BERT model. While it may yield slightly higher accuracy, it demands significantly greater computational resources. Due to its larger parameter count (110 mln versus 66 mln in DistilBERT), both training time and memory consumption are substantially higher.

Modified BERT variants, such as RoBERTa (employs improved pre-training objectives) or ALBERT (relies on parameter factorisation and sharing), can offer competitive or superior performance. However, these models are more complex to fine-tune and generally more resource-intensive than DistilBERT [25; 26].

Traditional machine learning approaches, including support vector machines, random forest and naive Bayes, remain viable for text classification but require extensive manual feature engineering (e. g., TF-IDF vectorisation). In practice, these methods typically underperform transformer-based models (especially on complex contextual tasks such as evaluating managerial responses) due to their limited ability to capture long-range dependencies and semantic nuances [27; 28].

The key advantages of the proposed approach are as follows:

1) DistilBERT is approximately 45 % faster and 40 % lighter than BERT while retaining over 96 % of its performance, making it far more suitable for real-world deployment;

2) the model not only performs binary classification (`Is_Good = 0` or `Is_Good = 1`) but also extracts demonstrated competencies and critical red flags, thereby delivering actionable insights highly valuable to HR and training departments – an output rarely provided by standard classification models;

3) the achieved performance metrics (accuracy = 0.986 8, F1-score = 0.987 0, with cross-validation showing accuracy = 0.971 1 ± 0.023 0 and F1-score = 0.970 0 ± 0.025 2 across 5 folds) are among the highest reported for this type of task and surpass numerous results obtained with DistilBERT in the literature [29] (for instance, reference [22] reports a maximum accuracy of 99.41 % on certain benchmark datasets, which our solution approaches on a significantly more contextually complex managerial evaluation task);

4) concatenating the situational description and candidate response as a single input sequence (`Situation + Response`) enables the model to fully account for context, a factor of critical importance in assessing the quality of managerial decisions.

Among the limitations of the proposed solution are the requirement for a large volume of high-quality labelled training data and the maximum sequence length constraint of 512 tokens inherent to BERT-based architectures.

Conclusions

The present study demonstrates the high effectiveness of a fine-tuned DistilBERT model for the automatic evaluation of managerial decisions based on textual responses paired with situational descriptions. A key contribution is the creation of a specialised Russian-language dataset enriched with expert annotations of

demonstrated competencies and developmental areas (red flags). By incorporating a post-processing layer that maps model predictions to specific competency scores (e. g., `'Community engagement': 5`) and critical shortcomings (e. g., `'Inaction'`), the proposed solution transcends conventional binary classification and significantly enhances interpretability and practical utility of the results.

The achieved performance metrics (accuracy = 0.986 8, F1-score = 0.987 0) with robust cross-validation results (mean accuracy = 0.971 1 ± 0.023 0) establishes a strong benchmark and opens promising avenues for deploying automated decision-support systems in management training, personnel assessment and HR analytics. Future research directions include:

1) augmenting the original dataset with real-world managerial cases to improve ecological validity;

2) benchmarking other language models (including multi-lingual and larger Russian-specific models) on the proposed dataset;

3) extending the competency framework and red flags taxonomy to cover a broader range of managerial and leadership skills.

References

1. Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network. arXiv:1503.02531 [Preprint]. 2015 [cited 2024 December 19]: [9 p.]. Available from: https://arxiv.org/abs/1503.02531.

2. Prema V, Elavazhahan V. Sculpting DistilBERT: enhancing efficiency in resource-constrained scenarios. In: Dwivedi RK, Saxena AK, Sharma R, Bhardwaj S, Khattri V, editors. *Proceedings of the 2023 12th International conference on system modeling & advancement in research trends (SMART)*; *2023 December 22–23; Moradabad, India*. [S. l.]: Institute of Electrical and Electronics Engineers; 2023. p. 251–256. DOI: 10.1109/SMART59791.2023.10428568.

3. Mukherjee A, Umme Salma M. Resume ranking and shortlisting with DistilBERT and XLM. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2024 IEEE International conference for women in innovation, technology and entrepreneurship (ICWITE); 2024 February 16–17; Bangalore, India*. [S. l.]: Institute of Electrical and Electronics Engineers; 2024. p. 301–304. DOI: 10.1109/ICWITE59797.2024.10502523.

4. Jingga K, Hidayaturrahman H. The utilization of DistilBERT features space for topics suggestion in the context of learning content. In: Jusman Y, Mat Isa NA, Sukamta, Ejaz W, Ismail AH, Bhattacharjya A, et al., editors. *Proceedings of the 2023 International workshop on artificial intelligence and image processing (IWAIIP); 2023 December 1–2; Yogyakarta, Indonesia*. [S. l.]: Institute of Electrical and Electronics Engineers; 2023. p. 243–247. DOI: 10.1109/IWAIIP58158.2023.10462787.

5. Islam MT, Parvin F, Sazan SA, Amir TB. Comparative analysis of sentiment classification on IMDB 50k movie reviews: a study using CNN, LSTM, CNN-LSTM, and BERT models. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2024 IEEE International conference on power, electrical, electronics and industrial applications (PEEIACON); 2024 September 12–13; Rajshahi, Bangladesh*. [S. l.]: Institute of Electrical and Electronics Engineers; 2024. p. 512–517. DOI: 10.1109/PEEIACON63629.2024.10800035.

6. Otani N, Bhutani N, Hruschka E. Natural language processing for human resources: a survey. In: Chen W, Yang Y, Kachuee M, Fu X-Y, editors. *Proceedings of the 2025 Conference of the nations of the Americas chapter of the Association for Computational Linguistics: human language technologies; April 2025; Albuquerque, USA. Volume 3, Industry track*. Kerrville: Association for Computational Linguistics; 2025. p. 583–597. DOI: 10.18653/v1/2025.naacl-industry.47.

7. Kang Y, Cai Z, Tan C-W, Huang Q, Liu H. Natural language processing (NLP) in management research: a literature review. *Journal of Management Analytics*. 2020;7(2):139–172. DOI: 10.1080/23270012.2020.1756939.

8. Schmiedel T, Müller O, vom Brocke J. Topic modeling as a strategy of inquiry in organizational research: a tutorial with an application example on organizational culture. *Organizational Research Methods*. 2019;22(4):941–968. DOI: 10.1177/1094428118773858.

9. Jia J, Liang W, Liang Y. A review of hybrid and ensemble in deep learning for natural language processing. arXiv:2312.05589v2 [Preprint]. 2024 [cited 2024 December 19]: [23 p.]. Available from: https://arxiv.org/abs/2312.05589v2.

10. Davenport T, Guha A, Grewal D, Bressgott T. How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*. 2020;48(1):24–42. DOI: 10.1007/s11747-019-00696-0.

11. Golestani A, Masli M, Shami NS, Jones J, Menon A, Mondal J. Real-time prediction of employee engagement using social media and text mining. In: Wani MA, Kantardzic M, Sayed-Mouchaweh M, Gama J, Lughofer E, editors. *Proceedings of the 17th IEEE International conference on machine learning and applications (ICMLA); 2018 December 17–20; Orlando, USA*. [S. l.]: Institute of Electrical and Electronics Engineers; 2018. p. 1383–1387. DOI: 10.1109/ICMLA.2018.00225.

12. Jamuna KM. Leveraging natural language processing to analyze employee feedback for enhanced HR insights. *International Journal of Scientific Research in Engineering and Management*. 2024;8(11):1–6. DOI: 10.55041/IJSREM39115.

13. Leidner JL, Stevenson M. Challenges and opportunities of NLP for HR applications: a discussion paper. arXiv:2405.07766 [Preprint]. 2024 [cited 2024 December 19]: [10 p.]. Available from: https://arxiv.org/abs/2405.07766.

14. Zhang M, Jensen KN, Sonniks SD, Plank B. SkillSpan: hard and soft skill extraction from English job postings. In: *Proceedings of the 2022 Conference of the North American chapter of the Association for Computational Linguistics: human language technologies; 2022 July 10–15; Seattle, USA*. Stroudsburg: Association for Computational Linguistics; 2022. p. 4962–4984. DOI: 10.18653/v1/2022.naacl-main.366.

15. Uma VR, Velchamy I, Upadhyay D. Recruitment analytics: hiring in the era of artificial intelligence. In: Tyagi P, Chilamkurti N, Grima S, Sood K, Balusamy B, editors. *The adoption and effect of artificial intelligence on human resources management. Part A*. Bingley: Emerald Publishing Limited; 2023. p. 155–174 (Özen E, Grima S, editors. Emerald studies in finance, insurance, and risk management; volume 7).

16. França TJF, Mamede HS, Barroso JMP, dos Santos VMPD. Artificial intelligence applied to potential assessment and talent identification in an organisational context. *Heliyon*. 2023;9(4):e14694. DOI: 10.1016/j.heliyon.2023.e14694.

17. Devlin J, Chang M-W, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Burstein J, Doran C, Solorio T, Kumar R, Loukina A, Morales M, et al., editors. *Proceedings of the 2019 Conference of the North American chapter of the Association for Computational Linguistics: human language technologies; 2019 June 2–7; Minneapolis, USA*. *Volume 1, Long and short papers*. Stroudsburg: Association for Computational Linguistics; 2019. p. 4171–4186. DOI: 10.18653/v1/N19-1423.

18. Gardazi NM, Daud A, Malik MK, Bukhari A, Alsahfi T, Alshemaimri B. BERT applications in natural language processing: a review. *Artificial Intelligence Review*. 2025;58(6):166. DOI: 10.1007/s10462-025-11162-5.

19. Davoodi L, Mezei J, Heikkilä M. Aspect-based sentiment classification of user reviews to understand customer satisfaction of e-commerce platforms. *Electronic Commerce Research*. 2026;26(2):1417–1459. DOI: 10.1007/s10660-025-09948-4.

20. Abdel-Salam S, Rafea A. Performance study on extractive text summarization using BERT models. *Information*. 2022;13(2):67. DOI: 10.3390/info13020067.

21. Lan Z, Chen M, Goodman S, Gimpel K, Sharma P, Soricut R. ALBERT: a lite BERT for self-supervised learning of language representations. arXiv:1909.11942v6 [Preprint]. 2020 [cited 2024 December 19]: [17 p.]. Available from: https://arxiv.org/abs/1909.11942v6.

22. Barbon RS, Akabane AT. Towards transfer learning techniques – BERT, DistilBERT, BERTimbau, and DistilBERTimbau for automatic text classification from different languages: a case study. *Sensors*. 2022;22(21):8184. DOI: 10.3390/s22218184.

23. Sanh V, Debut L, Chaumond J, Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv: 1910.01108v4 [Preprint]. 2020 [cited 2024 December 19]: [5 p.]. Available from: https://arxiv.org/abs/1910.01108.

24. Andrenko KV, Kroshchanka AA, Golovko OV. The use of distilled large language models to determine the sentiment of a text. In: Golenkov VV, Azarov IS, Golovko VA, Gordey AN, Guliakina NA, Krasnoproshin VV, et al., editors. *Open semantic technologies for intelligent systems. Issue 9*. Minsk: Belarusian State University of Informatics and Radioelectronics; 2025. p. 229–234.

25. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: a robustly optimized BERT pretraining approach. arXiv: 1907.11692 [Preprint]. 2019 [cited 2024 December 19]: [13 p.]. Available from: https://arxiv.org/abs/1907.11692.

26. Jiao X, Yin Y, Shang L, Jiang X, Chen X, Li L, et al. TinyBERT: distilling BERT for natural language understanding. In: Cohn T, He Y, Liu Y, editors. *Findings of the Association for Computational Linguistics: EMNLP-2020; 2020 November 16–20*. Stroudsburg: Association for Computational Linguistics; 2020. p. 4163–4174. DOI: 10.18653/v1/2020.findings-emnlp.372.

27. Golovko V, Kroshchanka A, Treadwell D. The nature of unsupervised learning in deep neural networks: a new understanding and novel approach. *Optical Memory and Neural Networks*. 2016;25(3):127–141. DOI: 10.3103/S1060992X16030073.

28. Golovko V, Kroshchanka A, Turchenko V, Jankowski S, Treadwell D. A new technique for restricted Boltzmann machine learning. In: Institute of Electrical and Electronics Engineers. *Proceedings of the 2015 IEEE 8th International conference on intelligent data acquisition and advanced computing systems: technology and applications (IDAACS); 2015 September 24–26; Warsaw, Poland. Volume 1*. [S. l.]: Institute of Electrical and Electronics Engineers; 2015. p. 182–186. DOI: 10.1109/IDAACS.2015.7340725.

29. Andrenko KV, Kroshchanka AA, Golovko VA, Krapivin YuB. The use of neural networks for decision-making analysis. In: Ablameyko SV, Kazachonak VU, Kurbackij AN, Krasnoproshin VV, editors. *Information systems and technologies. Proceedings of the 11th International scientific congress on computer science (CSIST-2025); 2025 October 29–31; Minsk, Belarus. Part 2*. Minsk: Belarusian State University; 2025. p. 118–125. Russian.

Received 24.12.2025 / revised 19.01.2026 / accepted 29.01.2026.