https://doi.org/10.53360/2788-7995-2026-3(23)-6

IRSTI: 28.23.25

O. Umarov[1], Zh. Oralbekova[1]*, Sh. Tutkyshbayeva[1], D. Oralbekova[2], G. Tyulepberdinova[3]
[1]Astana IT University,
010000, Republic of Kazakhstan, Astana, 55/11 Mangilik El avenue, C1 block
[2]Institute of Information and Computational Technologies,
050010, Republic of Kazakhstan, Almaty, 28 Shevchenko street
[3]Al-Farabi Kazakh National University,
050040, Republic of Kazakhstan, Almaty, 71 Al-Farabi avenue
*e-mail: zh.oralbekova@astanait.edu.kz

STUDENT STRESS CLASSIFICATION BASED ON MACHINE LEARNING ALGORITHMS

Annotation: *Survey instruments remain the most practical approach for large-scale monitoring of student stress in educational settings, and the Perceived Stress Scale (PSS) continues to be validated across contexts. This study evaluates how modern, off-the-shelf machine-learning methods perform on a tabular, multi-class student survey with pronounced class imbalance. We conduct a two-stage evaluation. First, we benchmark Decision Tree (DT), Random Forest (RF), and k-Nearest Neighbors (KNN) across five train-test regimes (90/10, 80/20, 70/30, 60/40, 50/50) to probe stability under different sample sizes. Second, we add an imbalance-aware experiment on an 80/20 split using SMOTE-style oversampling, and compare RF, Gradient Boosting (GB), and XGBoost (XGB). We report macro-averaged metrics suited for skewed multi-class targets. Across the five-split benchmark, RF and KNN are consistently stronger than a single DT. These outcomes align with recent evidence that tree-based ensembles dominate on typical tabular data and with reports that ensembles are reliable for stress-related classification. On the 80/20 split, RF and XGBoost tie at macro-F1 ≈ 0.73, while an additional SVM (RBF) reaches the best macro-F1 of 0.86. Clustering and projections confirm that stress responses form a continuum rather than discrete groups.*

Key words: *student stress, survey classification, tabular data, Random Forest, Gradient Boosting, XGBoost, class imbalance, macro F1.*

Introduction

Educational institutions regularly conduct student surveys to monitor groups of students and offer help at an early stage. Proven self-assessment tools, in particular Perceived Stress Scale (PSS), are widely used and psychometrically evaluated [1]. As data volumes increase, the pertinent inquiry is whether conventional machine learning can convert survey responses into more timely and focused interventions.

Tree-based ensembles, particularly Random Forests (RF) [2] and XGBoost (XGB) [3], provide an appealing combination of precision, resilience, and manageable engineering demands for tabular surveys. Recent cross-task observations emphasize that, on “typical tabular” problems with mixed discrete/continuous predictors and heterogeneous signal, tree-based models often outperform deep learning [4-5]. Ensemble methods have shown great reliability in stress analytics [6].

This paper presents a two-stage comparative analysis of a three-class student stress classification (Eustress, Distress, No Stress). Stage A is a common baseline study that uses Decision Tree (DT) / RF / k-Nearest Neighbors (KNN) across five splits to see how stable and sensitive it is to the size of the training set. Stage B uses SMOTE to deal with imbalances on the 80/20 split and compares RF, Gradient Boosting (GB), and XGB. We prioritize macro-averaged metrics over accuracy [7] and discuss the advantages of precision-recall (PR)-centric diagnostics in skewed contexts [8].

Random Forests combine random decision trees to lower variance while keeping bias low [2]. XGBoost is an example of a boosting family. These families build additive models that fix residual errors stage by stage and get the best results on structured features [3]. On a larger methodological scale, tree-based models have consistently surpassed deep networks in tabular problems characterized by modest sample sizes and categorical signals [4-5]. A recent study validates the efficacy of ensembles in the particular field of stress classification [6].

©O. Umarov, Zh. Oralbekova, Sh. Tutkyshbayeva, D. Oralbekova, G. Tyulepberdinova, 2026

Raw accuracy can be misleading with skewed labels; macro-averaged precision, recall, and F1 weigh classes equally and provide a fairer view of minority performance [7, 9]. When the positive class is rare, PR is usually more useful than ROC for diagnostic curves [8].

Resampling is a useful choice. Using Synthetic Minority Oversampling Technique (SMOTE) – style oversampling on the training set, which often makes it easier to remember rare classes without messing up the test distribution [10].

Materials and methods

We analyzed dataset [11] of 843 respondents, each represented by 26 numerical characteristics (25 input features and one target variable) that cover both demographic attributes (such as sex and age) and a variety of behavioral and psychological indicators including sleep quality, headaches frequency, the level of anxiety, workload perception, and academic concentration. These variables collectively provide a multidimensional view of stress-related symptoms and lifestyle factors.

The target variable is a three-class categorical sign distinguished between Eustress (positive, motivating stress), Distress (negative, harmful stress) and No Stress. The data set shows a strong imbalance of classes, most of which belong to Eustress. According to stratified section 80/20, the test set contains 154 Eustress, 6 Distress and 9 No Stress instances (a total of = 169), which makes the minority categories particularly difficult for generalizing the model. This imbalance requires careful selection of metrics and re-sampling strategies to ensure fair evaluation of models in all groups.

Preprocessing. The data set is already enumerated, requiring minimal transformation before modeling. For stage A (5-part baseline study) we apply an easy pre-processing channel and standardize the functions only for the KNN algorithm, which relies on distance calculations and is sensitive to different scales. Decision Tree and Random Forest models are used on raw feature scales because they are scalable.

Baseline study across five split regimes**.** We benchmark DT, RF, KNN on five stratified splits: 90/10, 80/20, 70/30, 60/40, 50/50. Each model is trained on the training portion of the split and evaluated on its hold-out. We report Accuracy, macro-Precision, macro-Recall, and macro-F1 [7]. KNN uses standardized features; DT/RF do not require scaling.

Imbalance-aware ensembles on 80/20. On an 80/20 stratified split, we apply SMOTE-style oversampling [10] to the training set only, then fit:

- Random Forest (RF) with 200 trees and class_weight = ‘balanced’ [2];
- Gradient Boosting (GB) from a standard library implementation;
- XGBoost (XGB) with default multi-class objective [3]. We again report macro-averaged metrics and per-class scores.

KNN was implemented with k = 5. Decision Tree and Random Forest used the default scikit-learn settings. SVM employed an RBF kernel with C = 1.0 and gamma = ‘scale’. XGBoost used the multi-class objective with eval_metric = 'mlogloss' and random_state = 42.

Following guidance in [7] and [9], we emphasize macro-F1 as the main scalar summary. Under severe imbalance, we additionally recommend PR-curve analyses for minority classes [8]. Macro-F1 as the main scalar summary.

All experiments were conducted in Python 3.12 using Google Colab. The main libraries included scikit-learn 1.6, xgboost, pandas, numpy and imbalanced-learn. To ensure reproducibility, stratified train–test splitting was performed with random_state = 42, and all stochastic models were initialized using the same random seed where applicable. Feature standardization for KNN and SVM was performed using StandardScaler.

Results and Discussion

Stage A – Five-split baseline benchmark (DT, RF, KNN)

Table 1 replicates our split-wise results. Metrics are macro-averaged.

Observations: (1) RF and KNN always do better than DT on macro-F1. (2) KNN does better when it is scaled; its best macro-F1 (0.768) is at 70/30. (3) RF is stable across splits, which is what the literature says about ensemble stability [2, 4, 6]. (4) DT has a higher variance and a lower macro-F1, which is what you would expect from a single-tree learner on targets that are noisy and unbalanced.

Table 1 – Train results

#	Algorithm	Iteration	Features	Targets	Train %	Test %	Accuracy	Precision	Recall	F1
0	Decision Tree	1	25	3	90	10	0.882353	0.506410	0.506410	0.506410
1	Decision Tree	2	25	3	80	20	0.863905	0.565382	0.692690	0.611121
2	Decision Tree	3	25	3	70	30	0.893281	0.600405	0.607915	0.590070
3	Decision Tree	4	25	3	60	40	0.881657	0.544913	0.557766	0.547879
4	Decision Tree	5	25	3	50	50	0.886256	0.564035	0.610006	0.579058
5	Random Forest	1	25	3	90	10	0.929412	0.813008	0.523504	0.598611
6	Random Forest	2	25	3	80	20	0.934911	0.977505	0.616667	0.705026
7	Random Forest	3	25	3	70	30	0.936759	0.978320	0.528205	0.618425
8	Random Forest	4	25	3	60	40	0.946746	0.981763	0.558608	0.652132
9	Random Forest	5	25	3	50	50	0.938389	0.978862	0.560606	0.647204
10	KNN	1	25	3	90	10	0.929412	0.816667	0.602564	0.654008
11	KNN	2	25	3	80	20	0.923077	0.854036	0.598977	0.676543
12	KNN	3	25	3	70	30	0.948617	0.874824	0.708473	0.767914
13	KNN	4	25	3	60	40	0.940828	0.857564	0.577371	0.659551
14	KNN	5	25	3	50	50	0.938389	0.869912	0.603456	0.686206

Stage B – Imbalance-aware ensembles on 80/20 (SMOTE on train)
Below (Figure 1-3), per-class outcomes on the held-out test of the 80/20 split.

Random Forest Results:
Accuracy: 0.9467455621301775

	precision	recall	f1-score	support
Distress (Negative Stress) - Stress that causes anxiety and impairs well-being.	0.75	0.50	0.60	6
Eustress (Positive Stress) - Stress that motivates and enhances performance.	0.95	0.99	0.97	154
No Stress - Currently experiencing minimal to no stress.	1.00	0.44	0.62	9
accuracy			0.95	169
macro avg	0.90	0.65	0.73	169
weighted avg	0.95	0.95	0.94	169

Figure 1 – Random forest results

RF achieves strong, balanced performance with relatively stable recall for No Stress and competitive recall for Distress. Reducing variations of the ensemble (packaging + sub-sampling of features) limits the readjustment to over-analyzed minority models, which helps to keep the generalization on an unaffected set of tests. Typical errors are Eustress ↔ No Stress-confusions when respondents exhibit moderate symptoms without clear threshold crossings.

Gradient Boosting Results:
Accuracy: 0.9171597633136095

	precision	recall	f1-score	support
Distress (Negative Stress) - Stress that causes anxiety and impairs well-being.	0.57	0.67	0.62	6
Eustress (Positive Stress) - Stress that motivates and enhances performance.	0.94	0.97	0.96	154
No Stress - Currently experiencing minimal to no stress.	0.50	0.22	0.31	9
accuracy			0.92	169
macro avg	0.67	0.62	0.63	169
weighted avg	0.91	0.92	0.91	169

Figure 2 – Gradient boosting results

GB follows mainly because of the low recall on No Stress, pointing to a conservative boundary solution that favors the majority class. Although the amplification improves separability for some severe cases of Eustress/Distress, it seems more sensitive to residual noise in minority regions. This manifests as an under-definition of No Stress and a drop in the F1 macro, despite the good precision of Eustress.

XGBoost Accuracy: 0.9408284023668639

	precision	recall	f1-score	support
Distress (Negative Stress) - Stress that causes anxiety and impairs well-being.	0.62	0.83	0.71	6
Eustress (Positive Stress) - Stress that motivates and enhances performance.	0.96	0.98	0.97	154
No Stress - Currently experiencing minimal to no stress.	1.00	0.33	0.50	9
accuracy			0.94	169
macro avg	0.86	0.72	0.73	169
weighted avg	0.95	0.94	0.93	169

Figure 3 – XGBoost results

RF and XGB tie on macro-F1 (~0.73). XGB achieves the best recall on Distress (0.83), while RF is stronger on No Stress (F1 0.62). GB lags mainly due to poor No Stress recall (0.22). These behaviors mirror the broader evidence that ensembles handle tabular surveys well but still struggle on tiny minority classes, even with SMOTE [4-6].

The questionnaire's signal is heterogeneous and nonlinear. Tree-based methods efficiently carve the predictor space into partitions that align with such local interactions (e.g., sleep issues × headaches × workload). Randomization in RF reduces variance [2]. Boosting focuses on hard cases by iteratively correcting residuals [3]. Across typical tabular problems, this bias–variance profile often outperforms deep learning without heavy feature engineering [4-5]. Our findings-RF/XGB with macro-F1 ≈ 0.73-are squarely in line with that picture and with stress-classification literature emphasizing ensembles [6].

Because the test distribution is 154/6/9, raw accuracy ≈ 0.94-0.95 is over-optimistic. Macro-F1 [7] reflects the real difficulty on Distress and No Stress (≈ 0.63-0.73 overall). For triage use cases-e.g., flagging potential Distress for human review-PR curves and recall at chosen thresholds for the minority class are more informative than ROC [8]. In practice, it can be beneficial to tune a higher-recall operating point for Distress, accepting more false positives but reducing the risk of misses.

To investigate the presence of natural clusters in the student stress data set, we applied a density-based DBSCAN clustering algorithm to a standardized feature matrix of 843 student responses. The output of the algorithm shown in Figure 4 illustrates that, despite the change in the region radius (eps = 0.3-0.5) and min_samples = 5, all configurations created a single dense area in which all points were classified as noise (843 noise points, 0 cluster, silhouette = None).

This finding indicates that the stress survey data do not form special regions with high density – in accordance with their continuous and overlapping nature at all levels of stress. As shown in Figure 4, students' responses form a smooth continuum rather than discrete behavioral clusters, confirming that the intensity of stress varies gradually between people. DBSCAN's sensitivity to global density explains the lack of separate cluster concentrations for Eustress, Distress and No Stress.

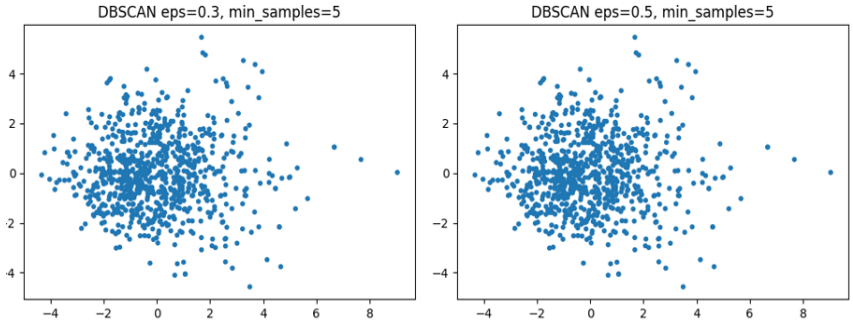


Figure 4 – DBSCAN results at (a) eps = 0.3, (b) eps = 0.5 showing all samples classified as a single dense region without cluster boundaries

To extend the comparison of baseline data beyond tree ensembles, we evaluated support vector machines (SVM) with a Radial Basis Function (RBF) core and XGBoost using the same 80/20 stratified and standardized elements as in stage B.

Both models were trained on coded target labels (Distress, Eustress, No Stress).

The comparative results are summarized in Table 2.

Table 2 – Comparative analysis of models' accuracy

Model	Accuracy	Macro F1	Notes
Random Forest (baseline)	0.947	0.73	Strong ensemble stability
SVM (RBF)	0.970	0.86	Excellent overall; slight under-recall on minorities
XGBoost	0.959	0.80	Strong performance; better Distress recall, weaker No Stress recall

Per-class metrics:

– SVM: Distress F1 = 0.80, Eustress F1 = 0.98, No Stress F1 = 0.80.

– XGBoost: Distress F1 = 0.91, Eustress F1 = 0.98, No Stress F1 = 0.50.

SVM achieved the highest overall accuracy (97%) but missed several Distress and No Stress cases due to class imbalance.

The XGBoost efficiency (95.8%) remained consistent with the Random Forest base, with a slightly higher response for Distress (0.83), but weaker for No Stress (0.33).

These results confirm that the marginal and incentive approaches provide comparable effectiveness in all tabular survey data, supporting their use in multiclass stress classification.

The comparative visualization in Figure 5 highlights the proximity of the three models' performance, reinforcing the statistical results summarized above.

```
Label mapping: {'Distress': np.int64(0), 'Eustress': np.int64(1), 'No Stress': np.int64(2)}
SVM Accuracy: 0.9704142011834319
              precision    recall  f1-score   support

    Distress       1.00      0.67      0.80         6
    Eustress       0.97      1.00      0.98       154
   No Stress       1.00      0.67      0.80         9

    accuracy                           0.97       169
   macro avg       0.99      0.78      0.86       169
weighted avg       0.97      0.97      0.97       169

Confusion matrix (SVM):
 [[  4   2   0]
 [  0 154   0]
 [  0   3   6]]
XGBoost Accuracy: 0.9585798816568047
              precision    recall  f1-score   support

    Distress       1.00      0.83      0.91         6
    Eustress       0.96      1.00      0.98       154
   No Stress       1.00      0.33      0.50         9

    accuracy                           0.96       169
   macro avg       0.99      0.72      0.80       169
weighted avg       0.96      0.96      0.95       169

Confusion matrix (XGBoost):
 [[  5   1   0]
 [  0 154   0]
 [  0   6   3]]
```

Figure 5 – Accuracy comparison between Random Forest, XGBoost, and SVM

To visually study the high-dimensional stress response structure, we projected a standardized object space into two dimensions using PCA and t-SNE.

The PCA projection (Figure 6) showed overlapping, but somewhat expanded clusters.

Electrical models dominate in the central region, while distress and absence of stress appear at the periphery, indicating that linear variations divide levels of stress partially but not cleanly.

The introduction of t-SNE (Figure 7) led to the formation of clearer local groupings.

Although the global structure remained mixed due to small minority classes, t-SNE retained subtle divisions between Distress (lower area) and No Stress (upper clusters), emphasizing its ability to capture non-linear local similarities.

• PCA effectively preserves global differences, but unites minority classes.

• t-SNE, although stochastic and slow, reflects nuanced psychological differences in screening patterns.

These additions support controlled results by showing that Eustress dominates a large homogeneous cluster, while Distress and No Stress occupy rarer, better pockets.

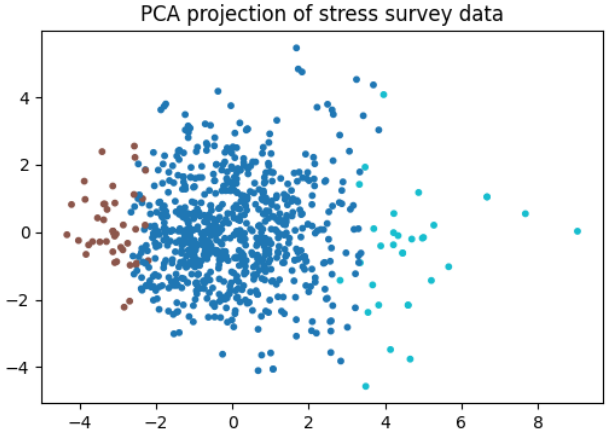


Figure 6 – PCA projection of student stress survey data

The PCA forecast shows that most deviations can be generalized in a small number of linear directions, which explains the wide overlapping structure in Figure 6. In this space, the Eustress responses occupy a central mass, while the Distress and No Stress appear on the periphery - assuming that the linear combinations of questionnaire points (e.g., sleep quality, headache frequency, workload) share the groups only partially. This is consistent with the expectation that stress responses form a continuum, not sharply separated groups.

In contrast, t-SNE in Figure 7 focuses on the conservation of local areas. As a result, slightly narrower micro-groupings appear at the edges of the Eustress collector. These focal points correspond to respondents sharing similar patterns of high stress, and help to identify a fine-grained local structure that smooths out the PGSs.

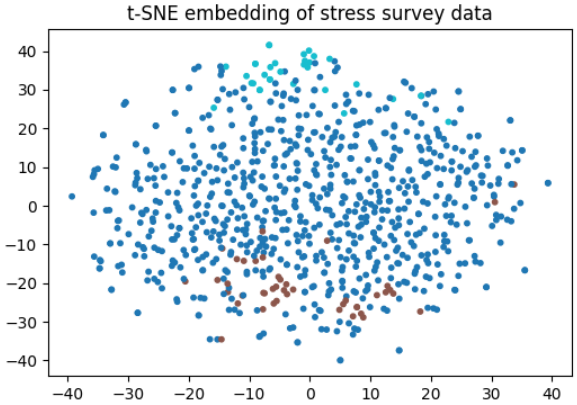


Figure 7 – t-SNE embedding of student stress survey data (perplexity = 30)

Although global geometry is less reliable in t-SNE, the method identifies locally agreed subgroups that may be useful for qualitative review (e.g., students with associated sleep problems and workload overload). In practice, the two views are complementary: PCA is an accurate global summary; t-SNE emphasizes subtle local similarities that can guide targeted interventions.

Conclusion

We combined a five-fold baseline study (DT/RF/KNN) with an assessment of imbalance-aware ensembles (RF/GB/XGB + SMOTE). In the five splits, RF and KNN consistently outperform DT; and in the 80/20 split with SMOTE, RF and XGBoost tie at macro-F1 ≈ 0.73 (accuracy 0.94-0.95), while GB lags due to weak No Stress recall.

These results are consistent with the broader observation that tree-based ensembles continue to be the strongest standard for typical tabular data [4-5], as well as literature highlighting stress classification [6].

Additional comparative and exploration methods were implemented to extend the analysis: the Support Vector Machine achieved 97% accuracy with improved macro-F1 (0.86), while XGBoost maintained strong balanced performance (precision 0.96, macro 0.80). Unsupervised clustering

(DBSCAN) did not reveal clear clusters based on density, confirming that stress response varies over a continuous spectrum. Finally, the dimensionality-reduction techniques (PCA and t-SNE) provided visual evidence that Eustress dominates in latent space, while Distress and No Stress form rarer peripheral regions.

Taken together, these results confirm the conclusion that ensemble and margin models are very effective for structured stress survey data, while visualization and clustering analyses increase model interpretability and reliability.

References

1. Brown T.A. Further validation of the Personality Disorder Severity for *ICD-11* (PDS-*ICD-11*) scale in a community mental health sample / T.A. Brown, M. Sellbom // Psychological Assessment. – 2023. – Vol. 35(8). – P. 706-714. https://doi.org/10.1037/pas0001253.
2. Breiman L. Random forests / L. Breiman // Machine Learning. – 2001. – Vol. 45. – P. 5-32. https://doi.org/10.1023/A:1010933404324.
3. Chen T. XGBoost: A scalable tree boosting system / T. Chen, C. Guestrin // KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – 2016. – P. 785-794. https://doi.org/10.1145/2939672.2939785.
4. Grinsztajn L. Why do tree-based models still outperform deep learning on typical tabular data? / L. Grinsztajn, E. Oyallon, G. Varoquaux // 36th Conference on Neural Information Processing Systems (NeurIPS 2022) Track on Datasets and Benchmarks. – 2022. – Vol. 35. – P. 507-520. URL: https://proceedings.neurips.cc/paper_files/paper/2022/file/0378c7692da36807bdec87ab043cdadc-Paper-Datasets_and_Benchmarks.pdf (date of access: 3.01.2026).
5. Shwartz-Ziv R. Tabular data: Deep learning is not all you need / R. Shwartz-Ziv, A. Armon // Information Fusion. – 2022. – Vol. 81. – P. 84-90. https://doi.org/10.1016/j.inffus.2021.11.011.
6. Enhancing diagnostic decision making: Ensemble learning techniques for reliable stress level classification / R.V. Anand et al // Diagnostics. – 2023. – Vol. 13(22). – P. 3455. https://doi.org/10.3390/diagnostics13223455.
7. Grandini M. Metrics for multi-class classification: An overview / M. Grandini, E. Bagli, G. Visani // arXiv:2008.05756. – 2020. – URL: https://arxiv.org/abs/2008.05756 (date of access: 3.01.2026).
8. Saito T. The precision–recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets / T. Saito, M. Rehmsmeier // PLOS ONE. – 2015. – Vol. 10(3). – P. e0118432. https://doi.org/10.1371/journal.pone.0118432.
9. Raschka S. Model evaluation, model selection, and algorithm selection in machine learning / S. Raschka // arXiv:1811.12808. – 2018. – URL: https://arxiv.org/abs/1811.12808 (date of access: 3.01.2026).
10. Lemaître G. Imbalanced-learn: A Python toolbox to tackle the curse of imbalanced datasets in machine learning / G. Lemaître, F. Nogueira, C.K. Aridas // Journal of Machine Learning Research. – 2017. – Vol. 18. – P. 1-5. https://www.jmlr.org/papers/volume18/16-365/16-365.pdf (date of access: 3.01.2026).
11. Kaggle. Student Stress Monitoring Datasets. – 2025. https://www.kaggle.com/datasets/mdsultanulislamovi/student-stress-monitoring-datasets (date of access: 15.09.2025).

О. Умаров[1], Ж. Оралбекова[1]*, Ш. Туткышбаева[1], Д. Оралбекова[2], Г. Тюлепбердинова[3]

[1]Astana IT University, 010000, Республика Казахстан, Астана, проспект Мәңгілік Ел, 55/11, блок С1

[2]Институт информационных и вычислительных технологий,
050010, Республика Казахстан, Алматы, ул. Шевченко, 28

[3]Казахский национальный университет им. Аль-Фараби,
050040, Республика Казахстан, Алматы, проспект Аль-Фараби, 71

*e-mail: zh.oralbekova@astanait.edu.kz

КЛАССИФИКАЦИЯ СТРЕССА СТУДЕНТОВ НА ОСНОВЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Инструменты опроса остаются наиболее практичным подходом для крупномасштабного мониторинга стресса студентов в образовательных учреждениях. А шкала воспринимаемого стресса продолжает подтверждать свою эффективность в различных контекстах. В данном исследовании оценивалась эффективность современных стандартных методов машинного

обучения на табличном многоклассовом опросе студентов с выраженным дисбалансом классов. Проведена двухэтапная оценка. Во-первых, сравнены деревья решений, случайный лес и метод k-ближайших соседей (KNN) в пяти режимах обучения-тестирования (90/10, 80/20, 70/30, 60/40, 50/50), чтобы проверить стабильность при различных размерах выборки. Во-вторых, добавлен эксперимент с учетом дисбаланса на разбиении 80/20 с использованием передискретизации в стиле SMOTE и сравнены случайный лес, градиентный бустинг и XGBoost. Представлены макроусредненные метрики, подходящие для многоклассовых целевых значений с асимметричным распределением. В пяти вариантах сравнения алгоритмы случайный лес и KNN неизменно показывали лучшие результаты, чем алгоритм деревья решений. Эти результаты согласуются с недавними данными о том, что ансамбли на основе деревьев решений доминируют на типичных табличных данных, а также с сообщениями о том, что ансамбли надежны для классификации, связанной со стрессом. При разделении 80/20 RF и XGBoost показывают одинаковый результат по макро-F1 ≈ 0,73, в то время как дополнительный SVM (RBF) достигает наилучшего значения макро-F1 0,86. Кластеризация и проекции подтверждают, что реакции на стресс образуют континуум, а не дискретные группы.

Ключевые слова: *стресс у студентов, классификация опросов, табличные данные, случайный лес, градиентный бустинг, XGBoost, дисбаланс классов, макро F1.*

О. Умаров[1], Ж. Оралбекова[1]*, Ш. Туткышбаева [1], Д. Оралбекова[2], Г. Тюлепбердинова[3]
[1]Astana IT University, 010000, Қазақстан Республикасы, Астана, Мәңгілік Ел даңғылы, 55/11, С1 блогы
[2]Ақпараттық және есептеуіш технологиялар институты,
050010, Қазақстан Республикасы, Алматы, Шевченко көшесі, 28
[3]Әл-Фараби атындағы Қазақ ұлттық университеті,
050040, Қазақстан Республикасы, Алматы, Әл-Фараби даңғылы, 71
*e-mail: zh.oralbekova@astanait.edu.kz

МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІНЕ НЕГІЗДЕЛГЕН СТУДЕНТТЕРДІҢ СТРЕСІНІҢ ЖІКТЕУІ

Сауалнама құралдары білім беру мекемелеріндегі студенттердің стресін кең көлемді мониторингтеу үшін ең тәжірибелік тәсіл болып табылады. Ал қабылданған стрес шкаласы әртүрлі жағдайларда өзінің тиімділігін дәлелдеуді жалғастыруда. Бұл зерттеу класстағы теңгерімсіздігі айқын, кестелік, көп сыныптық студенттердің сауалнамасындағы заманауи стандартты машиналық оқыту әдістерінің тиімділігін бағалайды. Екі кезеңді бағалау жүргізілді. Біріншіден, әртүрлі үлгі өлшемдеріндегі орнықтылықты тексеру үшін бес оқыту-сынақ режимінде (90/10, 80/20, 70/30, 60/40, 50/50) шешім ағаштары, кездейсоқ орман және k-жақын көршілер (KNN) салыстырылды. Екіншіден, SMOTE стиліндегі қайта іріктеуді қолдана отырып, 80/20 бөлінуіне теңгерімсіздікке бейімделген эксперимент жасалып, кездейсоқ орман, градиентті бустинг пен XGBoost салыстырылды. Қиғаш үлестірімдері бар көп класты нысандарға жарамды макро-орташаланған метрикалар ұсынылды. Бес салыстыру бойынша кездейсоқ орман және KNN алгоритмдері шешім ағашы алгоритмінен асып түсті. Бұл нәтижелер шешім ағашына негізделген ансамбльдердің типтік кестелік деректерде басым екендігі туралы соңғы зерттеулермен, сондай-ақ ансамбльдердің стреске байланысты жіктеу үшін сенімді екендігі туралы есептермен сәйкес келеді. 80/20 бөлінуінде RF және XGBoost макро-F1 ≈ 0,73 кезінде теңеседі, ал қосымша SVM (RBF) 0,86 макро-F1 ең жақсы мәніне жетеді. Кластерлеу және проекциялау нәтижелері студенттердің стресс реакцияларының дискретті топтардан гөрі үздіксіз континуум түрінде қалыптасатынын растады..

Түйін сөздер: *студенттер стресі, сауалнама жіктемесі, кестелік деректер, кездейсоқ орман, градиентті бустинг, XGBoost, класс теңгерімсіздігі, макро F1.*

Сведения об авторах

Олжас Аслзадаевич Умаров – студент 3 курса, Astana IT University, Казахстан; e-mail: 230017@astanait.edu.kz. ORCID: https://orcid.org/0009-0001-8683-2216.

Жанар Орымбаевна Оралбекова* – PhD, ассоциированный профессор, и.о. профессора Школы ИИ и науки о данных, Astana IT University, Казахстан; e-mail: zh.oralbekova@astanait.edu.kz. ORCID: https://orcid.org/0000-0002-3726-8882.

Шырын Сарыбайевна Туткышбаева – PhD, ассистент профессора Школы программной инженерии, Astana IT University, Казахстан; e-mail: sh.tutkyshbayeva@astanait.edu.kz. ORCID: https://orcid.org/0000-0002-1419-2571.

Дина Орымбаевна Оралбекова – PhD, ассоциированный профессор, старший научный сотрудник Института информационных и вычислительных технологий, Казахстан; e-mail: dinaoral@mail.ru, ORCID: https://orcid.org/0000-0003-4975-6493.

Гульнур Алпыскызы Тюлепбердинова – кандидат физико-математических наук, ассоциированный профессор кафедры искусственного интеллекта и Big Data, Казахский национальный университет имени аль-Фараби, Казахстан; e-mail: tiulepberdynova.g@kaznu.kz. ORCID: https://orcid.org/0000-0002-4322-8983.

Авторлар туралы мәліметтер

Олжас Аслзадаевич Умаров – 3-курс студенті, Astana IT University, Қазақстан; e-mail: 230017@astanait.edu.kz. ORCID: https://orcid.org/0009-0001-8683-2216.

Жанар Орымбаевна Оралбекова* – PhD, қауымдастырылған профессор, Жасанды интеллект және деректер ғылымы мектебінің профессоры м.а., Astana IT University, Қазақстан; e-mail: zh.oralbekova@astanait.edu.kz. ORCID: https://orcid.org/0000-0002-3726-8882.

Шырын Сарыбайевна Туткышбаева – PhD, Бағдарламалық инженерия мектебінің ассистент профессоры, Astana IT University, Қазақстан; e-mail: sh.tutkyshbayeva@astanait.edu.kz. ORCID: https://orcid.org/0000-0002-1419-2571.

Дина Орымбаевна Оралбекова – PhD, қауымдастырылған профессор, Ақпараттық және есептеуіш технологиялар институтының аға ғылыми қызметкері, Қазақстан; e-mail: dinaoral@mail.ru, ORCID: https://orcid.org/0000-0003-4975-6493.

Гульнур Алпыскызы Тюлепбердинова – физика-математика ғылымдарының кандидаты, Жасанды интеллект және Big Data кафедрасының қауымдастырылған профессоры, Әл-Фараби атындағы Қазақ ұлттық университеті, Қазақстан; e-mail: tiulepberdynova.g@kaznu.kz. ORCID: https://orcid.org/0000-0002-4322-8983.

Information about the authors

Olzhas Umarov – third-year student, Astana IT University, Kazakhstan; e-mail: 230017@astanait.edu.kz. ORCID: https://orcid.org/0009-0001-8683-2216.

Zhanar Oralbekova* – PhD, associate professor, acting Professor of the School of AI and Data Science, Astana IT University, Kazakhstan; e-mail: zh.oralbekova@astanait.edu.kz. ORCID: https://orcid.org/0000-0002-3726-8882.

Shyryn Tutkyshbayeva – PhD, assistant professor of the School of Software Engineering, Astana IT University, Kazakhstan; e-mail: sh.tutkyshbayeva@astanait.edu.kz. ORCID: https://orcid.org/0000-0002-1419-2571.

Dina Oralbekova – PhD, associate professor, Senior Researcher of Institute of information and computational technologies, Kazakhstan; e-mail: dinaoral@mail.ru, ORCID: https://orcid.org/0000-0003-4975-6493.

Gulnur Tyulepberdinova – candidate of physical and mathematical sciences, associate professor of the Department of Artificial Intelligence and Big Data, Al-Farabi Kazakh National University, Kazakhstan; e-mail: tiulepberdynova.g@kaznu.kz. ORCID: https://orcid.org/0000-0002-4322-8983.

Received 01.05.2026
Revised 23.06.2026
Accepted 24.06.2026

https://doi.org/10.53360/2788-7995-2026-3(23)-7

IRSTI: 81.96.00

CC BY 4.0
Check for updates

Ye. Seitkulov[*], R. Ospanov[1], B. Yergaliyeva[1], K. Utebayev[2]
[1]L.N.Gumilyov Eurasian National University, 010008, Republic of Kazakhstan, Astana
[2]Almaty branch of the National Research Nuclear University MEPhI,
050040, Republic of Kazakhstan, Almaty
[*]e-mail: yerzhan.seitkulov@gmail.com

IMPLEMENTATION AND COMPARATIVE ANALYSIS OF SPHINCS+-BASED DIGITAL SIGNATURE SCHEMES

Annotation: *In post-quantum cryptography, the focus is on the design of algorithms capable of withstanding threats posed by quantum computers, which requires the adoption of new approaches to constructing their internal architecture. Significant attention in this field is devoted to digital signatures, as they provide authentication and integrity protection for data in digital environments. One of the most promising*

©Ye. Seitkulov, R. Ospanov, B. Yergaliyeva, K. Utebayev, 2026