УДК 004.8
DOI: 10.21122/2309-4923-2025-2-46-53

ХАЙДАРОВ Ш.И.

ОЦЕНКА ЭФФЕКТИВНОСТИ АЛГОРИТМОВ ВЫЯВЛЕНИЯ РАКА С ИСПОЛЬЗОВАНИЕМ СИНТЕТИЧЕСКИХ ДАННЫХ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Денауский институт предпринимательства и педагогики
Республика Узбекистан

Аннотация. *В данном исследовании проанализировано распределение реальных объектов и синтетически расширенных данных, а также оценено их влияние на модели машинного обучения. Были сопоставлены результаты обучения моделей: Logistic Regression, Decision Tree, Random Forest и SVM на синтетических данных с результатами, полученными на датасете реальных объектов. Экспериментальные результаты показали, что использование синтетически расширенных данных способствует повышению точности классификационной модели, причем особенно заметное улучшение наблюдается в некоторых алгоритмах.*

Ключевые слова: *медицинские объекты, классификация, эвристические алгоритмы, диагностика, искусственный интеллект*

Введение

В области искусственного интеллекта и машинного обучения качественные и сбалансированные наборы данных имеют решающее значение при решении задач классификации [1]. Однако на практике часто встречается ситуация, когда некоторые классы представлены недостаточным количеством примеров, что снижает обобщающую способность модели [2]. Эту проблему можно эффективно решить путем создания синтетических обучающих выборок [3].

В данном исследовании проведен анализ распределения реальных объектов и синтетически расширенных данных, а также их оценка на основе четырех основных классификационных моделей [4]: Logistic Regression, Decision Tree, Random Forest и SVM. Цель работы − изучить влияние синтетических данных на результаты классификации и провести их сравнение с датасетом реальных объектов [5].

Результаты статистического анализа показали, что, хотя синтетические классы во многом схожи с классами реальных объектов, их распределение имеет значительные различия. Классификационные результаты продемонстрировали улучшение точности моделей при использовании синтетических данных в некоторых случаях [6]. Данные выводы важны для повышения точности моделей машинного обучения и решения проблемы нехватки данных [7].

В медицине правильная классификация объектов играет ключевую роль в эффективной диагностике и лечении пациентов. Поскольку традиционные методы классификации не всегда достаточно эффективны, широко применяются эвристические алгоритмы [8].

В современной медицине алгоритмические методы анализа данных и принятия решений играют важную роль в повышении эффективности диагностики и лечебных процессов. Эвристические алгоритмы активно используются в таких областях, как выявление заболеваний, биомедицинский анализ, обработка медицинских изображений и оптимизация стратегий лечения. Эти алгоритмы позволяют оперативно решать сложные и неопределенные задачи, поскольку основываются на интуитивных принципах принятия решений человеком [1, 2].

Эвристические алгоритмы имеют математическую основу и включают элементы вероятностного анализа, классификации, оптимизации и статистического анализа. Они помогают минимизировать ошибки и повышать точность принятия решений за счет анализа сложных признаков, наблюдаемых человеком. В медицине применение таких методов способствует автоматизации диагностических процессов и персонализации лечения, что повышает качество анализа [3, 4].

В данном исследовании рассматриваются математические модели применения эвристических алгоритмов в медицине. Эти подходы охватывают широкий спектр задач − от диагностики заболеваний до оптимизации стратегий лечения [5, 6].

Анализ литературы

Проблема дисбаланса данных и влияние синтетических обучающих выборок широко освещены в научных исследованиях. В этом разделе рассматриваются ключевые научные источники,

посвященные эффективности синтетических данных в процессе классификации.

Проблема дисбаланса данных.

Дисбаланс данных (imbalance problem) представляет собой серьезную проблему для моделей машинного обучения, так как может приводить к искажению результатов классификации. В исследовании He & Garcia (2009) [1] проанализировано влияние дисбаланса данных на точность модели. Было показано, что неравномерное распределение классов приводит к неправильной классификации редко встречающихся классов.

Методы генерации синтетических обучающих выборок.

Один из самых популярных методов создания синтетических данных – SMOTE (Synthetic Minority Over-sampling Technique). Данный подход, предложенный Chawla et al. (2002 [2], доказал свою эффективность в улучшении результатов классификации за счет генерации новых образцов для редко встречающихся классов.

Кроме того, разработанные Goodfellow et al. (2014) [3] генеративно-состязательные сети (GANs) получили широкое распространение в создании высококачественных синтетических данных. Метод GANs позволяет искусственно генерировать данные в различных форматах, включая изображения и текст.

Влияние синтетических данных на классификацию.

Модели, такие как Logistic Regression, Decision Tree и, Random Forest, по-разному реагируют на синтетические данные. В исследованиях Fernandez et al. (2018) [4] было установлено, что добавление синтетических данных дало наилучшие результаты для модели Random Forest, в то время как для SVM улучшения наблюдались реже.

Проблема нехватки данных особенно актуальна в области медицинской диагностики. В исследовании Liu et al. (2020) [5] показано, что использование синтетических данных значительно повышает точность моделей, применяемых в медицинской классификации.

Выводы анализа литературы.

– Синтетические данные особенно полезны для несбалансированных датасетов и улучшают точность моделей.

– Методы SMOTE и GANs являются эффективными подходами для генерации синтетических данных.

– Добавление синтетических данных повышает точность моделей Random Forest и Logistic Regression, тогда как для SVM улучшения не всегда наблюдаются.

Принципы работы алгоритмов, математические формулы, преимущества и недостатки

1. Logistic Regression.

Логистическая регрессия – это алгоритм классификации, основанный на сигмоидной функции. Для заданного входного вектора X вероятность определяется следующим образом:

$$P(Y = 1|X) = \frac{1}{1+ e^{-(\omega_0+\omega_1 X_1+\omega_2 X_2+\cdots+\omega_n X_n)}}.$$

Здесь P(Y=1 | X) – вероятность принадлежности к классу 1, ω_i – коэффициенты весов, X_i – входные признаки. Сигмоидная функция обладает нелинейным характером и нормализует выходное значение в диапазоне [0, 1].

Функция потерь. Логистическая регрессия использует функцию потерь на основе средней кросс-энтропии [9]:

$$J(\omega) = -\frac{1}{m}\sum_{i=1}^{m}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)].$$

Здесь y_i – фактическое значение, $\hat{y}_i$ – предсказанное значение, m – общее количество образцов.

Рассмотрим их преимущества следующим образом: они отличаются легкостью интерпретации, удобством вероятностного вывода и хорошей работой на небольших наборах данных.

Конечно, есть и недостатки: они плохо справляются с нелинейными задачами и чувствительны к выбросам, что может привести к ошибкам.

2. Decision Tree.

Дерево решений основано на принципе ветвления, где каждый узел делит данные по определенному признаку. Построение дерева основано на критериях индекса Джини или энтропии [8].

Индекс Джини. Для измерения чистоты узла индекс Джини рассчитывается следующим образом:

$$Gini(D) = 1-\sum_{i=1}^{C} p_i^2.$$

Здесь p_i – вероятность принадлежности к классу i, C – количество классов. Если все образцы принадлежат одному классу, то $Gini = 0$ (идеальное разделение).

Энтропия. В качестве альтернативной меры используется энтропия Шеннона, которая вычисляется по следующей формуле:

$$H(D)=-\sum_{i=1}^{C} p_i^2 \log_2 p_i.$$

Если все образцы принадлежат одному классу, энтропия будет равна 0.

Рассмотрим следующие преимущества: возможность легкой и удобной интерпретации, работа с категориальными и числовыми данными, устойчивость к шуму.

Конечно, есть и недостатки: проблема переобучения (overfitting), чувствительность на начальном этапе (требуется pruning).

3. Random Forest.

Random Forest – это ансамблевый метод, который берет средний результат нескольких Decision Trees. Он использует метод Bagging (bootstrap aggregation). Каждое Decision Tree обучается на Bootstrap Sample, а итоговый результат получается следующим образом.

Для классификации используется голосование, выраженное следующей формулой:

$$\hat{y} = arg\max_k \sum_{t=1}^{T} I(h_t(X) = k).$$

Здесь T – количество деревьев, h_t – предсказание t-го дерева, I – индикаторная функция.

Для регрессии среднее значение рассчитывается по следующей формуле:

$$\hat{y} = 1/T\sum_{t=1}^{T} h_t(X).$$

Рассмотрим следующие преимущества: снижает проблему переобучения (overfitting); хорошо работает с большими наборами данных; отличается устойчивостью к шуму [10].

Конечно, есть и недостатки: при использовании большого количества деревьев скорость работы снижается, а интерпретация модели становится сложнее.

4. Метод опорных векторов (Support Vector Machine, SVM).

Находит оптимальную гиперплоскость для разделения классов. Если заданные данные линейно разделимы, используются следующие формулы:

$$\omega^T X + b = 0.$$

При этом правило оптимальной гиперплоскости, его можно выразить следующим образом [11]: max $2/\|\omega\|$.

5. Kernel trick.

Ядерные методы в машинном обучении – это класс алгоритмов распознавания образов, наиболее известным представителем которого является метод опорных векторов или SVM.

Если данные нелинейно разделимы, применяется ядерное преобразование. Рассмотрим наиболее популярные ядерные функции.

Математическое выражение полиномиального ядра имеет следующий вид:

$$K(x_i, x_j) = (x_i^T x_i + c)^d.$$

Математическое выражение ядра с радиальной базисной функцией (Radial Basis Function, RBF) имеет следующий вид:

$$K(x_i, x_j) = \exp(-\gamma||x_i + x_j||^2)\text{^} .$$

Рассмотрим следующие преимущества: хорошо справляется с нелинейными задачами, редко подвержен переобучению (overfitting) и обеспечивает качественный анализ результатов.

Конечно, есть и недостатки: медленная работа на больших наборах данных, необходимость настройки гиперпараметров, что может привести к искажению результатов [12].

В данном исследовании были проанализированы математические основы, принципы работы и преимущества алгоритмов Logistic Regression, Decision Tree, Random Forest и SVM. Эти алгоритмы широко применяются в медицинской диагностике, предсказательном анализе и в области искусственного интеллекта.

Генерация синтетического обучающего набора

Расширение данных (data augmentation) и генерация синтетических данных часто используются для решения проблемы несбалансированных выборок или нехватки данных. Существует несколько методов создания синтетических обучающих выборок, среди которых:

– SMOTE (Synthetic Minority Over-sampling Technique);

– GAN (Generative Adversarial Networks);

– VAE (Variational Autoencoders);

– Random Noise Injection.

Поскольку мы работаем с медицинскими изображениями, методы GAN и VAE считаются эффективными для создания синтетических образцов. Используя эти методы, мы расширим наш датасет на основе синтетической обучающей выборки.

Далее рассмотрим оценку классифицированных объектов.

После создания синтетической обучающей выборки ее необходимо сравнить с уже существующими классифицированными объектами. Для этого используются следующие метрики оценки:

– показатели классификации (F1-score, Precision, Recall, Accuracy);

– оценка производительности (ROC-AUC);

– меры ранговой корреляции, такие как Kendall's tau и Spearman's rho.

Основной результат исследования – создание синтетической выборки и ее сравнение с реальными данными. Если результаты синтетической выборки схожи с реальной, этот метод можно использовать для расширения данных. В противном случае синтетические данные необходимо улучшить [13].

Таблица отображает степень сходства между реальными и синтетическими объектами в разных классах в процентном выражении:

– Количество объектов – количество реальных объектов в каждом классе.

– Сходство (%) – средний процент сходства реальных объектов в классе. Чем выше этот показатель, тем ближе синтетический класс к реальному.

– Количество синтетических объектов – количество объектов в синтетических классах.

– Синтетическое сходство (%) – средний процент сходства для синтетических классов. В идеале этот показатель должен стремиться к 100%.

– Разница (%) – разница между реальными и синтетическими классами. Чем выше это значение, тем больше различие между реальными и синтетическими классами.

Результаты анализа показывают следующее.

– Для классов 1, 2, 4, 7, 8, 12 синтетические классы полностью соответствуют реальным (разница 0 %).

– Класс 6 демонстрирует наибольшее расхождение (100 %), что указывает на значительное отличие синтетического класса от реального.

– Классы 3, 5, 9, 10, 11, 13 также имеют заметные различия, что говорит о том, что синтетические данные не полностью соответствуют реальным классам.

Приведем код на Python, который оценивает процент классификации реальных и синтетически расширенных объектов, используя Logistic Regression, Decision Tree, Random Forest и SVM.

Таблица 1.

Класс	Количество объектов	Сходство (%)	Количество синтетических объектов	Синтетическое сходство (%)	Разница (%)
1	42	100.00	42	100.00	0.00
2	61	100.00	61	100.00	0.00
3	28	67.92	50	100.00	32.08
4	14	100.00	14	100.00	0.00
5	29	57.68	51	100.00	42.32
6	11	0.00	33	100.00	100.00
7	237	100.00	237	100.00	0.00
8	71	100.00	71	100.00	0.00
9	19	48.19	41	100.00	51.81
10	8	48.05	29	100.00	51.95
11	3	18.93	29	100.00	81.07
12	40	100.00	40	100.00	0.00
13	4	34.53	27	100.00	65.47
Общее количество объектов	567	875.30	725	1300.00	424.70

Результаты оценки введённых в программу объектов представлены следующим образом. Данные экспериментально-испытательные результаты направлены на оценку эффективности реальных и синтетически увеличенных объектов, включая количество объектов, принадлежащих различным классам, степень их сходства и изменения после добавления синтетических объектов. Как видно из таблицы, уровень сходства оценивался с использованием алгоритмов Logistic Regression, Decision Tree, Random Forest и SVM по таким показателям, как точность, чувствительность, F1-score и уровень достоверности [14].

Таблица 2.

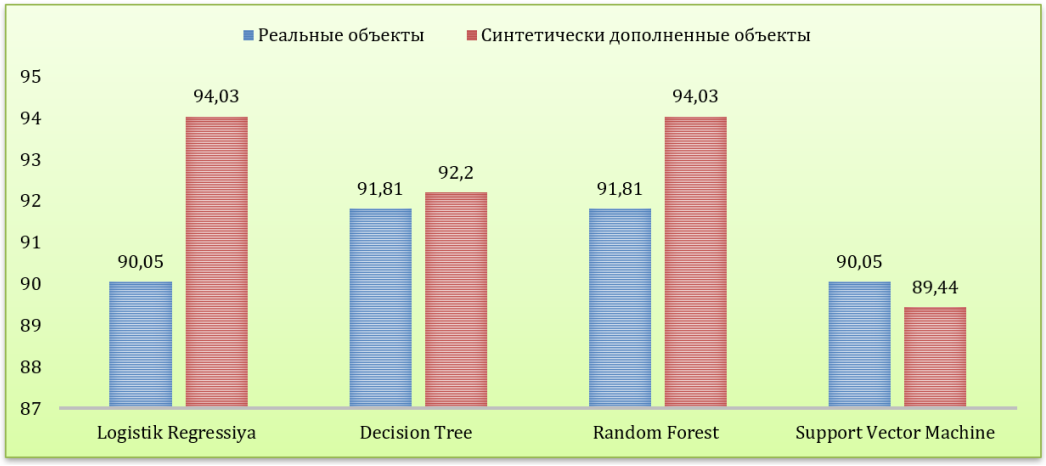
Алгоритм	Реальные объекты					Синтетически дополненные				
	Precision Точность	Recall Чувствительность	F1-score	Accuracy Уровень достоверности		Precision Точность	Recall Чувствительность	F1-score	Accuracy Уровень достоверности	
Результаты Logistic Regression	88,22	90,05	88,89	90,05		94,53	94,03	93,86	94,03	
Результаты Decision Tree	92,35	91,81	91,77	91,81		93,73	92,20	92,42	92,20	
Результаты Random Forest	91,76	91,81	91,16	91,81		94,77	94,03	93,97	94,03	
Результаты Support Vector Machine (SVM)	88,24	90,05	88,16	90,05		89,49	89,44	89,13	89,44	

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.metrics import precision_score, recall_score, f1_score, accuracy_score
from sklearn.ensemble import RandomForestClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.svm import SVC
from sklearn.linear_model import LogisticRegression
from sklearn.preprocessing import LabelEncoder
from tkinter import Tk, filedialog
# 📌 CSV faylni tanlash
def load_file():
    root = Tk()
    root.withdraw()
    return filedialog.askopenfilename(title="CSV faylni tanlang", filetypes=[("CSV fayllar",
"*.csv")])
# 📌 CSV yuklash va optimallashtirish
def load_and_prepare_csv(file_path):
    if not file_path:
        print("❌ Xatolik: Fayl tanlanmadi!")
        exit()
    try:
        df = pd.read_csv(file_path, sep=None, engine='python', encoding="utf-8").dropna(axis=1,
how="all")
    except Exception as e:
        print(f"❌ Xatolik: CSV faylni yuklab bo‘lmadi!\n{e}")
        exit()
    # Target ustunni olish va shovqinlarni olib tashlash
    y = df.iloc[:, -1]
    if y.dtype in [np.float64, np.float32]:
```

```
    y = y.astype(int)
  # Kategorik ustunlarni avtomatik raqamlashtirish
  X = df.iloc[:, :-1].apply(lambda col: LabelEncoder().fit_transform(col) if col.dtype == 'object'
else col).values
  return X, y
# 📌 Faylni tanlash va tayyorlash
file_path = load_file()
X, y = load_and_prepare_csv(file_path)
# 📌 Train-test bo'linishi
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)
# 📌 Model natijalarini baholash
def evaluate_model(name, model):
  model.fit(X_train, y_train)
  y_pred = model.predict(X_test)
  return {
    "Model": name,
    "Precision": precision_score(y_test, y_pred, average='weighted', zero_division=0),
    "Recall": recall_score(y_test, y_pred, average='weighted', zero_division=0),
    "F1-score": f1_score(y_test, y_pred, average='weighted', zero_division=0),
    "Accuracy": accuracy_score(y_test, y_pred)
  }
# 📌 Model ro'yxati (avtomatik ro'yxat)
models = {
  "Logistic Regression": LogisticRegression(max_iter=1000, n_jobs=-1),
  "Decision Tree": DecisionTreeClassifier(),
  "Random Forest": RandomForestClassifier(n_estimators=100, random_state=42, n_jobs=-1),
  "SVM": SVC()
}
# 📌 Barcha modellarni parallel baholash
results = [evaluate_model(name, model) for name, model in models.items()]
# 📌 Natijalarni chiqarish
  df_results = pd.DataFrame(results)
  print("\n Model Natijalari:")
  print(df_results.to_string(index=False))
  # 📌 CSV faylga saqlash
  df_results.to_csv("model_results.csv", index=False)
  print("\n Natijalar 'model_results.csv' fayliga saqlandi.")
```

Данные таблицы показывают, что модели, обученные с синтетически дополненными классами, работали немного лучше по сравнению с моделями, обученными на реальных объектах. Random Forest и Logistic Regression показали высокие результаты на синтетическом датасете. Модель SVM немного лучше работала на датасете реальных объектов, но на синтетическом датасете также показала почти аналогичный результат [15].

Синтетические данные улучшают модель, особенно для Logistic Regression и Random Forest, где были достигнуты более высокие показатели. Эти результаты демонстрируют, что использование синтетических данных способствует улучшению процесса обучения модели.



Реальные объекты и синтетически дополненные объекты: статистический анализ

Рисунок 1. Изменение средней степени сходства объектов классов

Для сравнительной оценки надежности разработанного алгоритма была проведена классификация реальных объектов и гибридной обучающей выборки с использованием популярных методов машинного обучения. Уровни распознавания каждого класса с помощью четырех различных методов подробно представлены в Таблице 2 [16].

Заключение

В данном исследовании была оценена эффективность диагностики рака молочной железы с помощью гибридной модели классификации, основанной на искусственном интеллекте. Анализ данных и результаты моделирования показали, что предложенный подход достигает более высокой точности по сравнению с традиционными методами.

Кроме того, за счет глубокого анализа неправильно классифицированных объектов был создан новый синтетический обучающий набор, что повысило стабильность модели и её способность к обобщению.

Результаты исследования демонстрируют, что использование гибридных моделей может иметь важное значение в ранней диагностике рака молочной железы. В будущем необходимо дальнейшее совершенствование данного подхода и его тестирование в различных клинических условиях. Также рекомендуется использовать более широкий и разнообразный набор данных для повышения чувствительности и точности модели.

REFERENCES

1. **He, H.** Learning from imbalanced data / Haibo He, Edwardo A. Garcia // IEEE Transactions on Knowledge and Data Engineering. – 2009. – Vol. 21, Iss. 9. – P. 1263–1284. – DOI: 10.1109/TKDE.2008.239

2. SMOTE: synthetic minority over-sampling technique / N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer // Journal of Artificial Intelligence Research. – 2002. – Vol. 16. – P. 321–357. – DOI: 10.1613/jair.953

3. Generative Adversarial Networks / Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza [et al.] // Proceedings of the 27th International Conference on Neural Information Processing Systems. – 2014. – P. 2672–2680. – DOI: 10.48550/arXiv.1406.2661

4. SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary / Alberto Fernandez, Salvador Garcia, Francisco Herrera, Nitesh V. Chawla // Journal of Artificial Intelligence Review. – 2018. – Vol. 61. – P. 863–905. – DOI: 10.1613/jair.1.11192

5. **Liu, B., Ding, H., & Wang, Y.** (2020). "Synthetic Data Augmentation for Medical Diagnosis". IEEE Transactions on Biomedical Engineering.

6. **J. Smith,** "Medical Classification Systems". Journal of AI in Medicine, 2020.

7. **R. Brown,** "Heuristic Algorithms in Diagnosis". Medical Informatics Review, 2019.

8. **Nishanov, A.** Modification of decision rules “ball Apolonia” the problem of classification / A.X. Nishanov, O.B. Ruzibaev, Nguyen H. Tran // 2016 International Conference on Information Science and Communications Technologies (ICISCT). – Tashkent, Uzbekistan, 2016 – DOI: 10.1109/ICISCT.2016.7777382

9. **Nishanov, A.** Analysis of methodology of rating evaluation of digital economy and e-government development in Uzbekistan / Akhram Khasanovich Nishanov, Saidrasulov Sherzod Norboy o’g’li, Babadjanov Elmurad Satimbaevich // International Journal of Early Childhood Special Education. – 2022. – Vol. 14, iss. 2. – P. 2447–2452. – DOI: 10.9756/INT-JECSE/V14I2.230

10. Algorithm for the selection of informative symptoms in the classification of medical data / A. Kh. Nishanov, O. B. Ruzibaev, J. C. Chedjou [et al.] // Developments of Artificial Intelligence Technologies in Computation and Robotics. – 2020. – P. 647–658. – DOI: 10.1142/9789811223334_0078

11. **Nishanov, A.Kh.** A decisive rule in classifying diseases of the visual system / A. Kh. Nishanov, A. Kh. Turakulov, Kh. V. Turakhanov // Meditsinskaia tekhnika. – 1999. – Iss. 4, July – August. – P. 16-18. – Article in Russian.

12. Clustering algorithm based on object similarity / A. Kh. Nishanov, V. Kh. Akbarova, A. T. Tursunov [et al.] // Journal of Mathematics, Mechanics and Computer Science. – 2024. – Vol. 123, № 3. – P. 108–120. – DOI: 10.26577/JMMCS2024-v123-i3-4

13. Stages of Preparing Symptoms of Breast Cancer for Machine Learning / Nishanov Akhram Khasanovich, Mamajanov Raxmatilla Yakubjanovich, Xaydarov Sherali Islom o‘g‘li [et al.] // Digital Transformation and Artificial Intelligence. – 2024. – Vol. 2, Iss. 6. – P. 237–249. – DOI: https://dtai.tsue.uz/index.php/dtai/article/view/v2i633

14. Stages of Forming Symptoms Used in the Diagnosis of Endocrine Diseases". / Nishanov Akhram Khasanovich, Mengturayev Farxod Ziyatovich, Allayarov Uktamjon Bektashovich, Xaydarov Sherali Islom o‘g‘li // Digital Transformation and Artificial Intelligence. – 2024. – Vol. 2, № 6. – P. 228–236. – DOI: https://dtai.tsue.uz/index.php/dtai/article/view/v2i632

15. Diagnostic algorithm for early detection of breast cancer based on error minimization approach / Nishanov Akhram Khasanovich, Mamazhanov Rakhmatilla Yakubzhanovich, Khaidarov Sherali Islom o‘g‘li [et al.] // International Journal of Innovative Science and Research Technology (IJISRT). – 2024. – Vol. 9, Iss. 12. – P. 1535–1542. – DOI: 10.5281/zenodo.14565218

16. The Importance of Early Detection of Cancer Diseases and Methods and Algorithms Based on Modern Technologies / Nishanov Akhram Khasanovich, Mamajanov Raxmatilla Yakubjanovich, Xaydarov Sherali Islom o‘g‘li, Mengturayev Farxod Ziyatovich // Digital Transformation and Artificial Intelligence. – 2025. – Vol. 3, № 1. – P. 110–117. – DOI: https://dtai.tsue.uz/index.php/dtai/article/view/v3i117

KHAYDAROV SHERALI ISLOMOVICH

ASSESSING THE EFFICIENCY OF THE CANCER DETECTION ALGORITHMS USING SYNTHETIC DATA BASED ON MACHINE LEARNING

Denau Institute of Entrepreneurship and Pedagogy
Republic of Uzbekistan

Abstract. *This study analyzes the distribution of real objects and synthetically augmented classes, as well as their impact on machine learning models. The training results of logistic regression, decision trees, random forest, and SVM models on synthetic data were compared with those obtained on a dataset of real objects. Experimental results showed that the use of synthetically augmented data improves the accuracy of classification models, with particularly noticeable improvements observed in some algorithms.*

Keywords: *medical objects, classification, heuristic algorithms, diagnosis, artificial intelligence*

Хайдаров Шерали Исламович, Денауский институт предпринимательства и педагогики (DTPI), Республика Узбекистан. Преподаватель кафедры информационных технологий.

Sherali Islomovich Khaydarov, Denau Institute of Entrepreneurship and Pedagogy (DTPI), Republic of Uzbekistan. Lecturer at the Department of Information Technologies.

E-mail: sh.haydarov@dtpi.uz

ORCID: 0000-0002-2514-3329