Оригинальная статья

Original Paper

УДК 338.45:338.97 © Л.П. Дьяконова, А.Ф. Константинов, О.В. Китова✉, В.А. Китов, 2025

ФГБОУ ВПО «Российский экономический университет имени Г.В. Плеханова», 117997, г. Москва, Россия
✉ e-mail: Dyakonova.lp@rea.ru

UDC 338.45:338.97 © L.P. Dyakonova, A.F. Konstantinov, O.V. Kitova✉, V.A. Kitov, 2025

Plekhanov Russian University of Economics, Moscow, 117997, Russian Federation
✉ e-mail: Dyakonova.lp@rea.ru

Применение методов машинного обучения и искусственного интеллекта в сфере кредитования и обнаружения мошенничества*

Application of machine learning and artificial intelligence methods in credit scoring and fraud detection

DOI: http://dx.doi.org/10.18796/0041-5790-2025-10-63-69

Настоящее исследование посвящено вопросам построения и совершенствования классификационных моделей машинного обучения для задач сферы финансовых рисков. Проводится обзор работ в области основных подходов к минимизации финансовых рисков с помощью применения продвинутых методов искусственного интеллекта и машинного обучения, современных эффективных моделей, а также конкретных машинных экспериментов с наборами финансовых данных. Поднимается вопрос о необходимости применения методов объяснимого искусственного интеллекта для интерпретации сложных моделей. Авторами были выбраны восемь перспективных моделей машинного обучения, в качестве базовой модели была принята эффективная модель градиентного бустинга CatboostClassifier. В работе проводилось сравнение ее показателей качества с показателями пяти моделей, представляющих модификации базовой Catboost, а также с моделями IsolationForest и wrapper-модели. Для тестирования всех моделей использовались открытые данные, включая четыре набора данных по розничному кредитованию и один набор с транзакционными данными по операциям с кредитными картами. Наши расчеты не выявили существенных различий между модификациями модели CatboostClassifier по характеристике ROC_AUC. В то же время модели IsolationForest и wrapper показали значимое улучшение этой характеристики на данных по кредитованию.

Ключевые слова: *финансовые риски, кредитный скоринг, обнаружение мошенничества, машинное обучение, методы объяснимого искусственного интеллекта, Catboost, SHAP.*

***Для цитирования*:** Применение методов машинного обучения и искусственного интеллекта в сфере кредитования и обнаружения мошенничества / Л.П. Дьяконова, А.Ф. Константинов, О.В. Китова и др. // Уголь. 2025;(10):63-69. DOI: 10.18796/0041-5790-2025-10-63-69.

ДЬЯКОНОВА Л.П.
Канд. физ.-мат. наук, доцент, доцент кафедры информатики ВШ КМиС РЭУ им. Г.В. Плеханова, ФГБОУ ВПО «Российский экономический университет имени Г.В. Плеханова», 117997, г. Москва, Россия, e-mail: Dyakonova.lp@rea.ru

КОНСТАНТИНОВ А.Ф.
Аспирант кафедры информатики ВШ КМиС РЭУ им. Г.В. Плеханова, ФГБОУ ВПО «Российский экономический университет имени Г.В. Плеханова», 117997, г. Москва, Россия, e-mail: konstantinovaf@gmail.com

** Данное исследование выполнено в рамках государственного задания в сфере научной деятельности Министерства науки и высшего образования РФ на тему «Модели, методы и алгоритмы искусственного интеллекта в задачах экономики для анализа и стилизации многомерных данных, прогнозирования временных рядов и проектирования рекомендательных систем», номер проекта FSSW-2023-0004.*

КИТОВА О.В.
Доктор экон. наук, доцент,
заведующий кафедрой информатики
ВШ КМиС РЭУ им. Г.В. Плеханова,
ФГБОУ ВПО «Российский экономический
университет имени Г. В. Плеханова»,
117997, г. Москва, Россия,
e-mail: Kitova.OV@rea.ru

КИТОВ В.А.
Канд. техн. наук, доцент,
доцент кафедры информатики
ВШ КМиС РЭУ им. Г.В. Плеханова,
ФГБОУ ВПО «Российский экономический
университет имени Г.В. Плеханова»,
117997, Москва, Россия,
e-mail: Kitov.VA@rea.ru

Abstract
This study is devoted to the issues of constructing and improving classification models of machine learning for problems in the field of financial risks. A review of works in the field of the main approaches to minimizing financial risks using advanced methods of artificial intelligence and machine learning, modern effective models, as well as specific machine experiments with sets of financial data is provided. The issue of the need to apply methods of explainable artificial intelligence to interpret complex models is raised. The authors selected 8 promising machine learning models, and the effective gradient boosting model CatboostClassifier was adopted as the basic model. In the work, its quality indicators were compared with the indicators of 5 models representing modifications of the basic Catboost, as well as with the IsolationForest and wrapper models. Open data were used to test all models, including 4 data sets on retail lending and 1 set with transactional data on credit card operations. Our calculations did not reveal significant differences between the modifications of the CatboostClassifier model in terms of the ROC_AUC characteristic. At the same time, the IsolationForest and wrapper models showed a significant improvement in this characteristic on the lending data.

Keywords
FInancial risks, credit scoring, fraud detection, machine learning, explainable artificial intelligence methods, Catboost, SHAP.
Acknowledgements
This research was performed in the framework of the state task in the field of scientific activity of the Ministry of Science and Higher Education of the Russian Federation, project "Models, methods, and algorithms of artificial intelligence in the problems of economics for the analysis and style transfer of multidimensional datasets, time series forecasting, and recommendation systems design", grant no. FSSW-2023-0004.
For citation
Dyakonova L.P., Konstantinov A.F., Kitova O.V., Kitov V.A. Application of machine learning and artificial intelligence methods in credit scoring and fraud detection. *Ugol'*. 2025;(10):63-69. (In Russ.). DOI: 10.18796/0041-5790-2025-10-63-69.

ВВЕДЕНИЕ

В настоящее время модели ИИ и машинного обучения активно успешно применяются финансовыми учреждениями для управления финансовыми рисками. Это связано как с огромным ростом финансовых операций, в том числе онлайн-транзакций, так и с постоянным совершенствованием моделей и алгоритмов ИИ, превосходящих традиционные подходы. Новые подходы позволяют банкам и другим финансовым организациям сокращать затраты с одновременным расширением операционной деятельности, способствуют повышению лояльности клиентов и доходности.

Традиционным применением ИИ в банках является кредитный скоринг, который позволяет снижать риски кредитования благодаря выявлению ненадежных заемщиков. Алгоритмы машинного обучения применяются в системах кредитного конвейера, что позволяет автоматизировать и оптимизировать обработку заявок на кредит.

Другим направлением применения механизмов ИИ является обнаружение мошенничества с кредитными картами благодаря способности таких моделей на основе анализа больших объемов данных выявлять аномальные транзакции.

Хотя методы и модели ИИ вывели на новый уровень бизнес-процессы и механизмы мониторинга и снижения финансовых рисков, сами модели вносят риски при их автоматизированном или автоматическом встраивании в процессы принятия решений, так как представляют собой «черные ящики» и не обладают наглядностью и объяснимостью более простых

моделей. Поэтому важной задачей являются развитие и применение к построенным моделям методов объяснимого искусственного интеллекта (XAI).

В нашем исследовании мы ставим целью изучить возможности современных моделей МО и ИИ для снижения кредитных и операционных рисков. Мы провели машинные эксперименты по построению классификаторов с данными, относящихся к розничному кредитованию (три набора данных), к потребительскому кредиту (один набор данных) и операциям с кредитными картами (один набор данных). Для каждого набора данных применялись восемь наиболее продвинутых моделей МО, зарекомендовавших себя в финансовых приложениях.

ОБЗОР ЛИТЕРАТУРЫ

Мл и риск-менеджмент

В работе [1] приводится краткий обзор основных методов искусственного интеллекта (ИИ) и машинного обучения, применяемых в финансовых организациях. Рассматриваются успешные результаты применения этих моделей в основных сферах финансовых рисков. Авторы выделяют некоторые существенные практические проблемы, связанные с такими моделями. Важнейшей проблемой авторы считают недостаточность или недоступность качественных данных, необходимых для обучения моделей. Отсутствие в достаточном количестве квалифицированных специалистов в области ИИ отмечается как следующая по важности проблема. Авторы отмечают, что руководители финансовых организаций должны понимать, что разработка и внедрение моделей ИИ представляют собой непрерывный процесс, требующий постоянного мониторинга и оценки. Необходимо также учитывать риски, связанные с непрозрачностью сложных моделей и этические проблемы.

В обзоре [2] дается систематический анализ появляющихся во все большем числе исследований по применению машинного обучения для управления финансовыми рисками. Авторами выделены наиболее значимые за последнее десятилетие публикации. Авторы дают классификацию задач управления финансовыми рисками и сопоставляют с этими задачами соответствующие методы МО. Для каждого класса задач выделяются зоны, связанные с основными проблемами: отсутствием общедоступных наборов данных, сложностью алгоритмов, отсутствием объяснимости работы моделей и возможной дискриминационностью моделей глубокого обучения

В качестве одного из важных направлений развития методов МО в финансовом секторе авторы рассматривают федеративное обучение, в котором для сохранения конфиденциальности и безопасности данных используются распределенные методы обучения. Федеративное обучение – это новая тенденция, которая позволяет обучать модели МО на нескольких серверах, без перемещения данных.

Недавние исследования показывают, что объединение нескольких классификаторов, т.е. ансамблевое обучение, дает по сравнению с отдельными моделями лучшие характеристики качества и точности. В работе [3] проводилось изучение как статистических методов, так и методов искусственного интеллекта применительно к задаче кредитного скоринга. Авторы провели сравнительную оценку производительности ансамблевых методов с такими базовыми моделями МО, как логистическая регрессия, деревья решений, искусственные нейронные сети и опорные векторы. Их эксперименты показывают превосходство ансамблевых методов над базовыми по значениям средней точности, ошибкам I и II типа. Аналогичные результаты получены при решении задачи кредитного скоринга на реальных наборах данных [4].

В статье [5] дается достаточно полный обзор применения гибридных моделей и мягких вычислений для прогнозирования банкротства.

Explainable Ai

Современные модели машинного обучения обладают высокой сложностью и представляют собой черный ящик. Применение таких моделей в финансовой сфере вызывает недоверие к результатам. Для повышения доверия необходимо использовать методы объяснимого искусственного интеллекта (XAI). Наиболее полно эти методы описаны в [6]. Методы XAI подразделяются на моделеспецифичные и независимые от модели (модельно-агностические). Модельно-независимые методы делятся на глобальные и локальные. Глобальные объясняют принципы работы модели на всем наборе данных, локальные фокусируются на индивидуальных предсказаниях. Наиболее широко применяются локальные методы LIME и SHAP.

В методе LIME (Local interpretable model-agnostic explanations) для конкретного примера генерируются возмущенные значения, и строится простая интерпретируемая модель. Полученная модель может объяснять предсказания исходной модели локально [7].

Метод SHAP (SHapley Additive exPlanations) основан на теории кооперативных игр и использует концепцию Шепли для распределения важности признаков. Он предоставляет вклад каждого признака в конкретный прогноз [8, 9]. SHAP-значения показывают, насколько признак изменил итоговый прогноз по сравнению с базовым значением. Подходы в задачах управления кредитным риском рассмотрены в [10, 11, 12].

Подход Partial Dependence Plots (PDP) отображает влияние изменения одного признака на прогноз, учитывая средние значения остальных признаков [13].

В [14] проводится сравнение методов интерпретации на примере данных по продажам. Автор выделяет высокую эффективность методов PDP и SHAP. В [15] изучается применение машинного обучения в банках на трех задачах: кредитные риски, кредитование физлиц и транзакционные риски. Самым эффективным оказался локальный метод SHAP.

Некоторые модели (случайные леса, градиентный бустинг) предоставляют встроенные метрики важности признаков [16]. Метод регуляризации LASSO может уменьшать коэффициенты до нуля, что дает понимание вклада признаков в модель [17].

ОБЗОР РАБОТ ПО КРЕДИТОВАНИЮ

В обзоре [18] отмечается важность изучения экономических и финансовых последствий машинного обучения при оценке кредитного риска. Рассматриваются традиционные методы оценки кредитного риска и их ограничения. Алгоритмы машинного обучения предлагают расширенные возможности благодаря учету сложных закономерностей в данных и демонстрируют превосходную производительность при обработке нелинейных и многомерных данных. Однако продвинутые модели МО требуют применения дополнительных методов для достижения интерпретируемости. Автор уделяет внимание работам, применяющим методы SHAP и LIME для повышения интерпретируемости сложных моделей МО.

В исследовании [19] оценивается производительность нескольких моделей машинного обучения — логистической регрессии, случайного леса, градиентного усиления (XGBoost), опорных векторных машин (SVM) и нейронных сетей при прогнозировании кредитного риска. Результаты показывают, что градиентный спуск (XGBoost) превзошел другие модели по всем метрикам качества. Нейронные сети и SVM показали хорошие результаты по точности, но были более ресурсоемкими и менее интерпретируемыми. Исследователи делают вывод, что Gradient Boosting (XGBoost) является наиболее подходящей моделью для управления крупномасштабными кредитными рисками.

В статье [20] предлагается подход с использованием стекированного классификатора в сочетании с методом выбора признаков (FS) на основе фильтра. Предлагаемая модель включает базовые оценщики: случайный лес (RF), усиление градиента (GB) и экстремальное усиление градиента (XGB). Метод FS основан на теории прироста информации (IG). Экспериментальные результаты показывают, что предлагаемый stacked-классификатор превосходит отдельные оценщики и другие существующие методы.

В работе [21] для прогнозирования дефолта используются модели логистической регрессии, дерева решений, XGBoost и LightGBM. Результаты показывают, что LightGBM и XGBoost превосходят модели логистической регрессии и дерева решений. Авторы использовали локальный интерпретируемый подход к модельно-агностическим объяснениям. Результаты показывают, что срок кредита, кредитный рейтинг и сумма кредита существенно влияют на прогнозируемые результаты.

В статье [22] рассматривается проектирование модели машинного обучения в финансовом трейдинге и прогнозировании рисков. Представлены история количественных инвестиций и применение машинного обучения в финансовой количественной оценке. Демонстрируется процесс построения и оценки модели случайного леса. Хотя случайный лес немного менее точен, чем XGBoost, он имеет высокую производительность. Авторы отмечают, что случайные леса могут превзойти XGBoost при работе с большими наборами данных, учитывая низкую сложность модели и высокую скорость обучения.

МАТЕРИАЛЫ И МЕТОДЫ

В работе применялись модели для оценки финансовых рисков:

1. CatboostClassifier.
2. SHAP_col_weights.
3. bagg_temp_08.
4. scale_pos_weights.
5. RFE.
6. RFE_SHAP.
7. IsolationForest IF_outlier_model.
8. wrapper_model.

1. Модель градиентного бустинга CatboostClassifier (базовая). Ансамблевые методы бустинга основаны на последовательном построении моделей (деревья решений), где каждая следующая модель строится на основе результатов предыдущих. CatboostClassifier разработана компанией Яндекс на основе градиентного бустинга с использованием градиентов второго порядка, обработки категориальных признаков и поддержки больших наборов данных. Принята в качестве базовой модели. Во избежание переобучения применялась досрочная остановка при отсутствии улучшения качества на протяжении 100 итераций.

2. SHAP_col_weights SHAP (SHapley Additive exPlanations) – модельно-независимая методология для интерпретации моделей. SHAP-значения показывают вклад каждого признака в прогноз. В модели значения признаков учитываются в соответствии с весами Shapley values. Обучается временная базовая модель, затем с помощью TreeExplainer рассчитываются веса признаков.

3. Модель баггинга bagg_temp_08 Баггинг (bootstrap aggregating) делит тренировочные данные случайным образом и обучает несколько моделей, объединяя прогнозы усреднением. В эксперименте параметр «bagging_temperature» снижен до 0,8 с 1,0 по умолчанию.

4. Учет дисбаланса классов (scale_pos_weights). Финансовые данные не сбалансированы – доля положительного класса варьируется от 0,1 до 10%. В модели учтен дисбаланс классов заданием параметра scale_pos_weight в обратном соответствии доли положительного класса.

5. Модель RFE Рекурсивное исключение признаков (Recursive Feature Elimination). Начинается с двух самых важных переменных, и рекурсивно добавляются переменные с меньшей важностью до прекращения роста показателей качества. Важность определяется встроенным методом «model.feature_importances_».

6. Модель RFE_SHAP аналогична RFE, но важность признаков определяется из Shapley values. Используется TreeExplainer для преобразования данных в Shapley values с расчетом модуля суммы по каждому столбцу.

7. Isolation Forest (IF). Алгоритм обнаружения аномалий на основе ансамблевой модели случайного леса. Объекты удаленного класса выделяются за меньшее число итераций. Имеет линейную временную сложность и работает с большими многомерными данными.

8. Модель обертка wrapper_model. Методы wrapper оптимизируют работу моделей выбором подмножества

Таблица 1

Перечень наборов данных, используемых в исследовании

List of data sets used in the study

№	Наименование	Отрасль	Количество строк	Количество столбцов	Доля положительного класса
1	Transactions Data Bank. Fraud Detection	Финансы	1 048 316	11	0,17
2	Loan Defaulter	Финансы	307 511	122	0,08
3	Loan Default Prediction Dataset	Финансы	255 347	18	0,12
4	Bank Account Fraud Dataset Suite (NeurIPS 2022)	Финансы	1 000 000	32	0,01
5	Automobile Loan Default Dataset	Финансы	121 856	40	0,08

признаков, выявляя взаимодействия между признаками. Для обучения модели обертки используются валидационные данные.

ОПИСАНИЕ НАБОРОВ ДАННЫХ

Для сравнительного анализа отобраны пять наборов данных с www.kaggle.com (*табл. 1*). Исключены наборы с высокими значениями ROC-AUC 95-99% из-за вероятности утечки данных.

Набор 1 [23]. Финансовые транзакции для обнаружения мошенничества с атрибутами транзакций, демографическими данными клиентов и паттернами использования.

Набор 2 [24]. Более 300 тысяч записей по кредитованию с более чем 100 признаками.

Набор 3 [25]. Coursera Loan Default Prediction Challenge – 255 347 строк и 18 столбцов.

Набор 4 [26]. NeurIPS 2022 – шесть синтетических наборов данных по мошенничеству с банковскими счетами, основанных на реальных данных. Имеют крайне низкую распространенность положительного класса (доля 0,01).

Набор 5 [27]. Данные небанковской финансовой компании по более чем 120000 автокредитов.

Для увеличения достоверности полученных результатов была выбрана следующая схема валидации: для каждого из пяти наборов данных наблюдения разделялись с помощью метода кросс-валидации StratifiedShuffleSplit на тренировочные (50%) и тестовые(50%) данные. Для каждого набора тренировочных данных пять раз проводилось разбиение на тренировочный и тестовый наборы данных с помощью метода кросс-валидации kFold (20% – на тестовые данные, 80% – данные для обучения). Таким образом, для каждого эксперимента было проведено 125 замеров показателей качества. Всего для этой работы обучено более 1625 моделей машинного обучения и проведены замеры их показателей качества. Это позволило вычислить статистически значимые различия в показателях качества базовой модели и моделей с применением предлагаемых методик.

При проведении экспериментов проводился замер Area Under the Receiver Operating Characteristic Curve (ROC_AUC). Данный показатель применяется для бинарных классификаторов при наличии несбалансированных классов. Выбор данного показателя произведен исходя из особенностей финансовых данных, для которых характерна несбалансированность положительного класса.

РЕЗУЛЬТАТЫ

Всего было проведено 1625 замеров показателей качества. На каждом эксперименте и наборе данных считались среднее значение ROC_AUC, доверительные интервалы и p-value. Результаты экспериментов представлены в *табл. 2*.

В ходе исследования были выявлены два метода (IF_outlier_second_model, wrapper_model), которые во всех наборах данных приводят к статистически значимому улучшению показателей качества в финансовых данных. В ходе дальнейших исследований необходимо рассмотреть механизмамы оптимизации этих методов.

Метод SHAP_col_weights показал статистически значимые отрицательные результаты. Это означает, что веса признаков не следует задавать пропорционально вкладу Shapley values.

Таблица 2

Результаты экспериментов. Значения метрики ROC_AUC

Experimental results. ROC_AUC metric values

№	Модель	Transactions Data Bank. Fraud Detection (1)	Loan Defaulter (2)	Loan Default Prediction Dataset (3)	Bank Account Fraud Dataset Suite (NeurIPS 2022)	Automobile Loan Default Dataset
1	CatboostClassifier (базовая)	0,885	0,757	0,755	0,896	0,745
2	SHAP_col_weights	0,885	0,747	0,750	0,883	0,734
3	bagg_temp_08	0,885	0,757	0,756	0,897	0,744
4	scale_pos_weight	0,885	0,755	0,756	0,894	0,738
5	RFE	0,885	0,757	0,754	0,895	0,744
6	RFE_SHAP	0,884	0,757	0,753	0,896	0,743
7	IsolationForest	0,847	**0,779**	**0,762**	**0,901**	**0,790**
8	wrapper_model	0,877	**0,774**	**0,765**	**0,899**	**0,769**

Два метода в среднем показали улучшение (bagg_temp_08, SHAP_col_interaction), но это улучшение статистически незначимо. В ходе дальнейших экспериментов мы планируем изменить параметры моделей с тем, чтобы получить статистически значимые результаты.

ОБСУЖДЕНИЕ

Анализ результатов показывает, что набор данных № 1 существенно отличается от остальных наборов данных, представляя собой транзакционные данные. У него самая большая доля положительного класса (17%). Этот случай реалистичен для микрофинансовых организаций. В ходе дальнейших исследований мы дополним машинные эксперименты расчетами на еще нескольких наборах данных с мошенническими транзакциями и проанализируем поведение рассматриваемых методов в разрезе доли положительного класса.

Замена встроенной в алгоритм важности признаков значениями Shapley values не привела к значимому повышению качества модели, что согласуется с выводами авторов исследования [19].

В методе SHAP_col_interaction использовалась информация о вкладе Shapley values в итоговый прогноз на валидационных данных. В дальнейших экспериментах для генерации наиболее связанных переменных необходимо брать информацию из параметра библиотеки SHAP explainer.shap_interaction_values.

Полученные результаты во многом совпадают с практическим опытом работы авторов в финансовом секторе.

Список литературы • References

1. Aziz S., Dowling M. Machine Learning and AI for Risk Management, in T. Lynn, G. Mooney, P. Rosati, and M. Cummins (eds.), Disrupting Finance: FinTech and Strategy in the 21st Century, Palgrave, 2019:33-50.
2. Mashrur A., Luo W., Zaidi N.A., Robles-Kelly A. Machine Learning for Financial Risk Management: A Survey. IEEE Access. 2020;(8):203203-203223. DOI:10.1109/access.2020.3036322 10.1109/access.2020.3036322.
3. Wang G., Hao J., Ma J., Jiang H. A comparative assessment of ensemble learning for credit scoring. *Expert Syst. Appl.* 2011;38(1):223-230. DOI: 10.1016/j.eswa.2010.06.048.
4. Wang G, Ma J., Huang L., Xu K. Two credit scoring models based on dual strategy ensemble trees. *Knowl.-Based Syst*. 2012;(26):61-68.
5. Verikas A., Kalsyte Z., Bacauskiene M., Gelzinis A. Hybrid and ensemble-based soft computing techniques in bankruptcy prediction: A survey. *Soft Comput.* 2010;14(9)995-1010. . https://www.researchgate.net/publication/220176584_Hybrid_and_ensemble-based_soft_computing_techniques_in_bankruptcy_prediction_A_survey.
6. Molnar Ch. Interpretable Machine Learning. [Electronic resource]. https://christophm.github.io/interpretable-ml-book.
7. Ribeiro M.T., Singh S., Guestrin C.. Why Should I Trust You?: Explaining the Predictions of Any Classifier. [Electronic resource]. https://arxiv.org/abs/1602.04938.
8. Official website of the Python library. Welcome to the SHAP documentation. [Electronic resource]. https://shap.readthedocs.io/en/latest/index.html.
9. Lundberg S.M., Lee S.-I. A Unified Approach to Interpreting Model Predictions. [Electronic resource]. https://arxiv.org/abs/1705.07874.
10. Bracke P., Datta A., Jung C., Sen S. Machine learning explainability in Fnance: An application to default risk analysis, Bank England, London, U.K., 2019, Working Paper 816/
11. Guidotti R., Monreale A., Ruggieri S., Turini F., Giannotti F., Pedreschi D. A survey of methods for explaining black box models. *ACM Comput. Surv.* 2019;51(5):1-42.
12. Adadi A., Berrada M. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access.* 2018;(6):52138-52160.
13. Goldstein A., Kapelner A., Bleich J., Pitkin E. Peeking Inside the Black Box: Visualizing Statistical Learning with Plots of Individual Conditional Expectation. [Electronic resource]. https://arxiv.org/abs/1309.6392.
14. Шевская Н.В. Объяснимый искусственный интеллект и методы интерпретации результатов. Моделирование, оптимизация и информационные технологии // Modeling, optimization and information technology. 2021;9(2). https://moitvivt.ru. DOI: 10.26102/2310-6018/2021.33.2.024.
    Shevskaya N.V. Explainable Artificial Intelligence and Methods of Result Interpretation. Modeling, Optimization, and Information Technologies. *Modeling, optimization and information technology.* 2021;9(2). https://moitvivt.ru. (In Russ.). DOI: 10.26102/2310-6018/2021.33.2.024.
15. Paolo Di Biasi, Rita Gnutti, Andrea Resti, Daniele Vergari. Machine Learning for Credit risk: three successful Case Histories. 25 Aug 2022. *Risk management magazine*. 2022;17(2):5-18. https://www.aifirm.it/wp-content/uploads/2022/08/RMM-2022-02-Excerpt-2.pdf.
16. Official article. Feature importances with a forest of trees. [Electronic resource]. https://scikit-learn.org/stable/auto_examples/ensemble/plot_forest_importances.html.
17. Official article. Lasso Regression. [Electronic resource].: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lasso.html.
18. Bello O.A. Machine Learning Algorithms for Credit Risk Assessment: An Economic and Financial Analysis. *International Journal of Management Technology.* 2023;10(1):109-133. https://eajournals.org/ijmt/wp-content/uploads/sites/69/2024/06/Machine-Learning-Algorithms.pdf.
19. Safayet Hossain, Ashadujjaman Sajal, Sakib Salam Jamee, Sanjida Akter Tisha, Md Tarake Siddique, Md Omar Obaid, MD Sajedul Karim Chy, Md Sayem Ul Haque. Comparative Analysis of Machine Learning Models for Credit Risk Prediction in Banking Systems. *The American Journal of Engineering and Technology.* 2025;7(04):22-33. https://doi.org/10.37547/tajet/Volume07Issue04-04.
20. Ileberi Emmanuel, Yanxia Sun, Zenghui Wang. A machine learning-based credit risk prediction engine system using a stacked classifier and a filter-based feature selection method. *Journal of Big Data.* 2024;(11):23. https://doi.org/10.1186/s40537-024-00882-0.
21. Xu Zhu a, Qingyong Chu a, Xinchang Song a, Ping Hu a, Lu Peng. Explainable prediction of loan default based on machine learning models. *Data Science and Management.* 2023;(6):123-133. https://doi.org/10.1016/j.dsm.2023.04.003.
22. Junliang Wang, Jiahao Xu, Qishuo Cheng, Rathna Kumar. Research on finance Credit Risk Quantification Model Based on Machine Learning Algorithm. *Academic Journal of Science and Technology*. 2024;10(1).
23. Dataset «Transactions Data Bank. Fraud Detection». Available at: https://www.kaggle.com/datasets/qusaybtoush1990/transactions-data-bank-fraud-detection (accessed 15.09.2025).

24. Dataset «Loan Defaulter». Available at: https://www.kaggle.com/datasets/gauravduttakiit/loan-defaulter (accessed 15.09.2025).
25. Dataset «Loan Default Prediction Dataset». Available at: https://www.kaggle.com/datasets/nikhil1e9/loan-default (accessed 15.09.2025).
26. Dataset «Bank Account Fraud Dataset Suite (NeurIPS 2022)». Available at: https://www.kaggle.com/datasets/sgpjesus/bank-account-fraud-dataset-neurips-2022 (accessed 15.09.2025).
27. Dataset «Automobile Loan Default Dataset». Available at: https://www.kaggle.com/datasets/saurabhbagchi/dish-network-hackathon (accessed 15.09.2025).

Authors Information

Dyakonova L.P. – PhD (Physics & Mathematics), Associate Professor, Associate Professor of the Department of Informatics, Higher School of Cybertechnologies, Mathematics and Statistics, Plekhanov Russian University of Economics, Moscow, 117997, Russian Federation, e-mail: Dyakonova.lp@rea.ru

Konstantinov A.F. – Graduate Student of the Department of Informatics Higher School of Cybertechnologies, Mathematics and Statistics, Plekhanov Russian University of Economics, Moscow, 117997, Russian Federation, e-mail: konstantinovaf@gmail.com

Kitova O.V. – Doctor of Economic Sciences, Associate Professor, Higher School of Cybertechnologies, Mathematics and Statistics, Plekhanov Russian University of Economics, Moscow, 117997, Russian Federation, e-mail: Kitova.OV@rea.ru

Kitov V.A. – PhD (Engineering), Associate Professor, Associate Professor of the Department of Informatics, Higher School of Cybertechnologies, Mathematics and Statistics, Plekhanov Russian University of Economics, Moscow, 117997, Russian Federation, e-mail: Kitov.VA@rea.ru

Информация о статье

Поступила в редакцию: 15.08.2025
Поступила после рецензирования: 18.09.2025
Принята к публикации: 29.09.2025

Paper info

Received August 15, 2025
Reviewed September 18, 2025
Accepted September 29, 2025

Якутуголь организовал экскурсию на разрез для экспертов молодежного форума

МЕЧЕЛ
ЯКУТУГОЛЬ

Компания «Якутуголь» (АО ХК «Якутуголь», входит в Группу «Мечел») организовала обзорную экскурсию на разрез «Нерюнгринский» для экспертов форума активной молодежи Нерюнгринского района «Главные смыслы». Посещение разреза стало частью программы форума, на котором обсуждались актуальные вопросы молодежной политики и социально-экономического развития района.

Экскурсия началась с обязательного вводного инструктажа по технике безопасности. Затем участники поднялись на смотровую площадку, откуда открывается вид на огромную чашу разреза. С такого расстояния огромные экскаваторы и самосвалы кажутся игрушечными на фоне бескрайних угольных горизонтов. Особый интерес вызвал диспетчерский пункт, откуда осуществляется круглосуточный контроль и координация всех производственных процессов. Гости разреза смогли увидеть, как с помощью передовых IT-решений осуществляется управление сложным процессом добычи, оптимизируется логистика и обеспечивается бесперебойная работа всего предприятия.

«После такого визита понимаешь, что в Нерюнгринском районе есть масса возможностей для самореализации, и наша молодежь должна знать об этом. Наглядно видно, что в современном горном производстве востребованы не только рабочие специальности, но и инженеры, IT-специалисты, аналитики», – отметила ***эксперт наставников сервиса «Росмолодежь. Гранты» нерюнгринка Наталья Угарова***.

«Я впервые посетил угольный разрез. И именно здесь ощущаются промышленные масштабы нашей страны», – поделился ***эксперт сервиса «Росмолодежь. Гранты» из Твери Константин Шадров***.

«Для компании «Якутуголь» важно, чтобы молодые люди видели реальные перспективы и возможности, которые открывает перед ними местная промышленность. Организация экскурсий - это часть общей стратегии Группы «Мечел» в области профориентации и привлечения молодых специалистов», – сказал ***генеральный директор АО ХК «Якутуголь» Олег Михайлов***.

Пресс-служба АО ХК «Якутуголь»