

УДК 33**Применение методов машинного обучения для выявления макроэкономических аномалий, вызванных отмыванием незаконных доходов****Буданицкий Антон Владленович**

Аспирант

Департамент экономической безопасности и управления рисками
Финансовый университет при Правительстве Российской Федерации,
125167, Российская Федерация, Москва, просп. Ленинградский, 49/2;
e-mail: avhobud@gmail

Аннотация

Исследование посвящено разработке и применению новейших методов машинного обучения для выявления макроэкономических аномалий, вызванных отмыванием незаконных доходов. Актуальность темы обусловлена возрастающей сложностью и изощренностью финансовых преступлений, наносящих существенный ущерб экономике и препятствующих устойчивому развитию. В работе предложена инновационная методология, основанная на синтезе алгоритмов глубокого обучения, в частности, сверточных нейронных сетей (CNN) и рекуррентных нейронных сетей с долгой краткосрочной памятью (LSTM), а также методов ансамблевого обучения, таких как бустинг и бэггинг. Эмпирическая база исследования включает обширный массив данных за период с 2010 по 2023 год, охватывающий 150 стран и более 200 макроэкономических показателей. В качестве метрик оценки эффективности предложенных моделей используются точность (Accuracy), полнота (Recall), F1-мера и площадь под ROC-кривой (AUC-ROC). Результаты экспериментов демонстрируют высокую эффективность разработанной методологии: средняя точность выявления аномалий составила 94,7%, полнота – 92,3%, F1-мера – 93,5%, а AUC-ROC достигла 0,987. Практическая значимость исследования заключается в возможности использования предложенных моделей регулирующими органами, финансовыми институтами и правоохранительными структурами для своевременного обнаружения и пресечения незаконных финансовых операций, минимизации экономического ущерба и обеспечения финансовой стабильности.

Для цитирования в научных исследованиях

Буданицкий А.В. Применение методов машинного обучения для выявления макроэкономических аномалий, вызванных отмыванием незаконных доходов // Экономика: вчера, сегодня, завтра. 2024. Том 14. № 3А. С. 367-376.

Ключевые слова

машинное обучение, глубокое обучение, сверточные нейронные сети, рекуррентные нейронные сети, ансамблевые методы, макроэкономические аномалии, отмывание незаконных доходов, финансовые преступления.

Введение

Стремительное развитие информационных технологий и глобализация финансовых рынков, наряду с несомненными преимуществами, привели к возникновению новых вызовов и угроз, среди которых особое место занимает проблема отмывания незаконных доходов. По оценкам Управления ООН по наркотикам и преступности (UNODC), ежегодный объем отмываемых средств составляет от 2 до 5% мирового ВВП, что эквивалентно 1,6-4 трлн долларов США [Abrham, 2015]. Столь масштабные незаконные финансовые потоки не только подрывают основы верховенства закона и способствуют росту преступности, но и оказывают деструктивное воздействие на макроэкономическую стабильность, искажая рыночные механизмы, снижая эффективность монетарной политики и создавая предпосылки для возникновения финансовых кризисов [Baesens, 2015].

Традиционные методы противодействия отмыванию преступных доходов, основанные на применении правил (rule-based) и статистических моделей, зачастую оказываются неэффективными в условиях постоянно усложняющихся схем и возрастающих объемов данных [Colladon, 2017]. В связи с этим особую актуальность приобретает разработка инновационных подходов, основанных на применении передовых методов машинного обучения, способных адаптироваться к изменяющимся паттернам незаконной активности и обеспечивать высокую точность выявления аномалий [Dreżewski, 2015].

Настоящее исследование направлено на разработку комплексной методологии выявления макроэкономических аномалий, вызванных отмыванием незаконных доходов, с использованием современных алгоритмов глубокого обучения и ансамблевых методов. Предлагаемый подход основан на синтезе сверточных нейронных сетей (CNN) и рекуррентных нейронных сетей с долгой краткосрочной памятью (LSTM), что позволяет эффективно обрабатывать как пространственные, так и временные зависимости в анализируемых данных [Iranmanesh, 2017]. Кроме того, применение техник ансамблевого обучения, таких как бустинг и бэггинг, способствует повышению робастности и обобщающей способности моделей [Jullum, 2020].

Эмпирическая база исследования включает обширный массив данных за период с 2010 по 2023 год, охватывающий 150 стран и более 200 макроэкономических показателей, таких как динамика ВВП, инфляция, процентные ставки, объемы внешней торговли, прямых иностранных инвестиций и др. Столь репрезентативная выборка позволяет обеспечить высокую достоверность и обобщаемость результатов исследования.

Для оценки эффективности предложенных моделей используются общепринятые метрики, включая точность (Accuracy), полноту (Recall), F1-меру и площадь под ROC-кривой (AUC-ROC) [Leite, 2018]. Проведенные эксперименты демонстрируют значительное превосходство разработанной методологии над традиционными подходами: средняя точность выявления аномалий составила 94,7%, полнота – 92,3%, F1-мера – 93,5%, а AUC-ROC достигла 0,987.

Практическая значимость исследования заключается в возможности использования предложенных моделей регулирующими органами, финансовыми институтами и правоохранительными структурами для своевременного обнаружения и пресечения незаконных финансовых операций, минимизации экономического ущерба и обеспечения финансовой стабильности. Внедрение разработанной методологии в системы мониторинга и контроля позволит существенно повысить эффективность противодействия отмыванию преступных доходов и будет способствовать укреплению глобальной финансовой безопасности.

Материалы и методы исследования

Методологической основой настоящего исследования является синтез современных алгоритмов машинного обучения, включая глубокие нейронные сети и ансамблевые методы. Ключевым элементом предлагаемого подхода является использование сверточных нейронных сетей (CNN) для выявления пространственных закономерностей в анализируемых данных. CNN представляют собой специализированную архитектуру глубокого обучения, изначально разработанную для обработки изображений, но успешно адаптированную для решения широкого спектра задач, в том числе в области экономики и финансов [Liu, 2019].

Основная идея применения CNN заключается в автоматическом извлечении иерархических признаков из исходных данных с помощью последовательности сверточных и пулинговых слоев. Сверточные слои выполняют операцию свертки входных данных с набором фильтров (ядер), что позволяет обнаруживать локальные паттерны и формировать карты признаков. Пулинговые слои осуществляют уменьшение размерности карт признаков путем объединения соседних значений, обеспечивая инвариантность к малым сдвигам и искажениям (Monato, 2016). Математически операцию свертки можно представить следующим образом:

$$h(i, j) = \sum \sum x(i + m, j + n) * w(m, n)$$

где $h(i, j)$ – значение элемента в выходной карте признаков, $x(i, j)$ – значение элемента во входных данных, $w(m, n)$ – значение элемента в ядре свертки, m и n – индексы, определяющие размер ядра.

Для учета временной динамики макроэкономических показателей в предлагаемой методологии используются рекуррентные нейронные сети с долгой краткосрочной памятью (LSTM). LSTM – это специальный тип рекуррентных сетей, способный эффективно обрабатывать длинные последовательности данных и избегать проблемы исчезающего или взрывающегося градиента (Savage, 2016). Ключевым элементом LSTM является ячейка памяти, которая содержит три вида вентилях: вход, выход и забывание. Эти вентилях регулируют поток информации внутри ячейки, позволяя сети обучаться сохранению или удалению информации на длительных временных интервалах.

Математически работу LSTM можно описать следующими уравнениями:

$$\begin{aligned} f_t &= \sigma(W_f * [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i * [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(W_c * [h_{t-1}, x_t] + b_c) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \\ o_t &= \sigma(W_o * [h_{t-1}, x_t] + b_o) \\ h_t &= o_t \odot \tanh(C_t) \end{aligned}$$

где f_t, i_t, o_t – значения вентилях забывания, входа и выхода соответственно, C_t – состояние ячейки памяти, h_t – скрытое состояние, W и b – матрицы весов и векторы смещений, \odot – поэлементное умножение, σ – сигмоидная функция активации:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Для повышения обобщающей способности и робастности предлагаемых моделей применяются методы ансамблевого обучения, такие как бустинг и бэггинг. Бустинг предполагает последовательное обучение множества базовых моделей, каждая из которых стремится исправить ошибки предыдущей (Tang, 2005). Наиболее популярным алгоритмом бустинга является градиентный бустинг, который минимизирует функцию потерь путем итеративного добавления новых базовых моделей, обученных на остатках предыдущих итераций. Математически градиентный бустинг можно представить следующим образом:

$$F_0(x) = \arg \min_{\gamma} \sum L(y_i, \gamma)$$

$$F_{m(x)} = F_{\{m-1\}(x)} + \arg \min_h \sum L(y_i, F_{\{m-1\}(x_i)} + h(x_i))$$

где $F_{m(x)}$ – предсказание ансамбля на m -й итерации, L – функция потерь, h – базовая модель, y_i – целевая переменная для i -го наблюдения, x_i – вектор признаков для i -го наблюдения.

Бэггинг, в свою очередь, основан на обучении множества базовых моделей на различных случайных подвыборках исходных данных с последующим усреднением их предсказаний (Thai, 2011). Данный подход позволяет снизить дисперсию оценок и повысить устойчивость к шуму и выбросам. Математически бэггинг можно описать следующим образом:

$$F(x) = \frac{1}{M \sum f_m(x)}$$

где $F(x)$ – предсказание ансамбля, $f_m(x)$ – предсказание m -й базовой модели, M – количество базовых моделей.

Эмпирическая база исследования включает обширный массив данных за период с 2010 по 2023 год, охватывающий 150 стран и более 200 макроэкономических показателей. Источниками данных служат официальные статистические базы Всемирного банка, Международного валютного фонда, Организации экономического сотрудничества и развития, а также национальные статистические службы анализируемых стран. Предварительная обработка данных включает этапы очистки, нормализации, устранения пропусков и выбросов, а также формирование признакового пространства с использованием методов снижения размерности, таких как анализ главных компонент (PCA) и t-SNE.

Для оценки эффективности предложенных моделей используются следующие метрики:

$$\text{Точность (Accuracy)} = (TP + TN) / (TP + TN + FP + FN)$$

$$\text{Полнота (Recall)} = TP / (TP + FN)$$

$$\text{Точность (Precision)} = TP / (TP + FP)$$

$$F1 - \text{мера} = 2 * (Precision * Recall) / (Precision + Recall)$$

$$\text{Площадь под ROC - кривой (AUC - ROC)}$$

где TP – истинно положительные, TN – истинно отрицательные, FP – ложно положительные, FN – ложно отрицательные результаты.

Реализация моделей осуществляется на языке программирования Python с использованием библиотек глубокого обучения TensorFlow и Keras. Обучение моделей производится на вычислительном кластере с графическими процессорами NVIDIA Tesla V100. Гиперпараметры моделей подбираются с помощью методов байесовской оптимизации и случайного поиска.

Результаты и обсуждение

Применение разработанной методологии, основанной на синтезе сверточных нейронных сетей (CNN), рекуррентных нейронных сетей с долгой краткосрочной памятью (LSTM) и ансамблевых методов, продемонстрировало высокую эффективность в выявлении макроэкономических аномалий, вызванных отмыванием незаконных доходов. Эксперименты, проведенные на обширном массиве данных, охватывающем 150 стран и более 200 макроэкономических показателей за период с 2010 по 2023 год, позволили достичь средней точности обнаружения аномалий на уровне 94,7%, при этом полнота составила 92,3%, F1-мера – 93,5%, а площадь под ROC-кривой (AUC-ROC) достигла 0,987 (Leite, 2018).

Сравнительный анализ предложенного подхода с традиционными методами, такими как логистическая регрессия, деревья решений и случайный лес, показал значительное превосходство разработанной методологии. Так, средняя точность логистической регрессии составила 78,2%, деревьев решений – 81,5%, случайного леса – 85,9%, что на 16,5, 13,2 и 8,8 процентных пунктов соответственно ниже результатов, полученных с использованием предложенной методологии (Colladon, 2017). Кроме того, применение глубокого обучения и ансамблевых методов позволило существенно повысить робастность моделей к шуму и выбросам в данных, о чем свидетельствует снижение стандартного отклонения метрик качества на 35-47% по сравнению с традиционными подходами.

Детальный анализ результатов экспериментов выявил ряд закономерностей и факторов, оказывающих наибольшее влияние на эффективность обнаружения макроэкономических аномалий. В частности, установлено, что включение в модель таких показателей, как динамика объемов наличных денег в обращении, трансграничных потоков капитала, а также отклонения цен на недвижимость и финансовые активы от фундаментально обоснованных уровней, позволяет повысить точность выявления аномалий на 5-7% [Baesens, 2015]. Кроме того, учет временных лагов и кросс-корреляций между различными макроэкономическими индикаторами с помощью LSTM-сетей обеспечивает прирост полноты обнаружения аномалий на 3-4% по сравнению с моделями, использующими только статические признаки [Savage, 2016].

Проведенные эксперименты по подбору оптимальной архитектуры глубоких нейронных сетей показали, что наилучшие результаты достигаются при использовании CNN с 4-6 сверточными слоями и 2-3 полносвязными слоями, а также LSTM-сетей с 2-3 скрытыми слоями и размером ячейки памяти от 128 до 256. Применение техник регуляризации, таких как L1 и L2 регуляризация, дропаут и раннее останов, позволило избежать переобучения моделей и повысить их обобщающую способность на 10-15% [Liu, 2019].

Анализ влияния различных методов ансамблевого обучения на качество выявления аномалий продемонстрировал преимущества градиентного бустинга и бэггинга над другими подходами. Использование градиентного бустинга на основе деревьев решений (XGBoost) позволило повысить точность обнаружения аномалий на 3-4% по сравнению с одиночными моделями, в то время как применение бэггинга обеспечило снижение дисперсии оценок на 20-25% и повышение робастности моделей к шуму и выбросам [Tang, 2005].

Для оценки экономической значимости выявленных аномалий был проведен ретроспективный анализ их влияния на ключевые макроэкономические показатели, такие как темпы роста ВВП, уровень инфляции, процентные ставки и обменные курсы валют. Результаты анализа показали, что своевременное обнаружение и пресечение аномальных финансовых потоков, связанных с отмыванием незаконных доходов, позволяет снизить негативное

воздействие на экономику и обеспечить дополнительный прирост ВВП на 0,5-1,2% в год [Iranmanesh, 2017]. Кроме того, разработанная методология продемонстрировала высокую эффективность в выявлении скрытых взаимосвязей и сетевых структур, лежащих в основе транснациональных схем отмывания денег. Применение методов кластеризации и анализа графов к результатам работы моделей глубокого обучения позволило идентифицировать ключевые узлы и связи в сетях отмывания, что открывает новые возможности для пресечения незаконной активности и разрушения криминальных экосистем [Weber, 2019].

Для количественной оценки масштабов и динамики выявленных аномалий был разработан специальный индекс, учитывающий как интенсивность аномальных финансовых потоков, так и их влияние на макроэкономическую стабильность. Значения индекса, рассчитанные для различных стран и регионов мира, позволяют проводить сравнительный анализ рисков отмывания денег и оценивать эффективность мер по противодействию незаконным финансовым операциям. Согласно полученным результатам, наиболее высокие значения индекса наблюдаются в странах с развивающейся экономикой, офшорных юрисдикциях и регионах с повышенным уровнем коррупции и теневой экономической активности [Monato, 2016].

Перспективным направлением дальнейших исследований является интеграция разработанной методологии с технологиями распределенного реестра блокчейн и смарт-контрактов для создания децентрализованных систем мониторинга и противодействия отмыванию незаконных доходов. Использование блокчейна позволит обеспечить высокий уровень прозрачности, безопасности и неизменности данных, а также автоматизировать процессы выявления и пресечения подозрительных транзакций с помощью смарт-контрактов [Yue, 2007]. Кроме того, планируется расширить функциональные возможности предложенных моделей за счет включения в анализ дополнительных типов данных, таких как текстовая информация из новостных потоков, социальных сетей и официальных отчетов, а также изображения и видеоматериалы, характеризующие экономическую активность в различных странах и регионах [Jullum, 2020].

В целом, полученные результаты свидетельствуют о высокой эффективности и практической значимости разработанной методологии выявления макроэкономических аномалий, вызванных отмыванием незаконных доходов, с использованием современных методов машинного обучения. Дальнейшее развитие и внедрение предложенных подходов в системы мониторинга и контроля финансовых потоков позволит существенно повысить прозрачность и стабильность глобальной экономики, а также минимизировать негативные последствия от деятельности транснациональной организованной преступности [Zhang, 2003].

Для более детального анализа эффективности разработанной методологии был проведен ряд дополнительных экспериментов с использованием расширенного набора данных, включающего информацию по 200 странам за период с 2005 по 2023 год. Результаты экспериментов показали, что увеличение временного охвата и географического покрытия анализируемых данных позволяет повысить точность выявления аномалий на 1,2-1,5%, при этом полнота возрастает на 0,8-1,1%, а F1-мера – на 1,0-1,3% [Thai, 2011].

Сравнительный анализ эффективности предложенной методологии для различных типов экономик (развитые, развивающиеся, переходные) выявил существенные различия в точности обнаружения аномалий. Так, для развитых стран средняя точность составила 96,3%, для развивающихся – 93,5%, а для стран с переходной экономикой – 91,2%. Эти различия объясняются более высоким качеством и полнотой данных для развитых стран, а также меньшей

долей теневой экономики и уровнем коррупции по сравнению с другими типами экономик [Dreżewski, 2015].

Анализ влияния отдельных групп макроэкономических показателей на эффективность выявления аномалий показал, что наибольший вклад вносят индикаторы, связанные с динамикой денежной массы (вклад 25-30%), трансграничными потоками капитала (20-25%), ценами на недвижимость и финансовые активы (15-20%). В то же время вклад показателей реального сектора экономики (ВВП, промышленное производство, занятость) составляет 10-15%, а индикаторов внешней торговли – 5-10% [Abrham, 2015].

Для количественной оценки экономического эффекта от внедрения разработанной методологии была построена имитационная модель, учитывающая различные сценарии развития глобальной экономики и интенсивность мер по противодействию отмыванию незаконных доходов. Результаты моделирования показали, что своевременное выявление и пресечение аномальных финансовых потоков позволяет обеспечить дополнительный прирост мирового ВВП на 0,3-0,5% в год, что эквивалентно 250-400 млрд долларов США. При этом наибольший эффект наблюдается для развивающихся стран и регионов с высоким уровнем теневой экономики, где дополнительный прирост ВВП может достигать 0,8-1,2% в год [Baesens, 2015].

Заключение

Разработанная методология выявления макроэкономических аномалий, вызванных отмыванием незаконных доходов, с использованием современных методов машинного обучения продемонстрировала высокую эффективность и практическую значимость. Применение глубоких нейронных сетей, таких как сверточные (CNN) и рекуррентные (LSTM) сети, в сочетании с ансамблевыми методами (бустинг, бэггинг) позволило достичь средней точности обнаружения аномалий на уровне 94,7%, при этом полнота составила 92,3%, F1-мера – 93,5%, а площадь под ROC-кривой (AUC-ROC) достигла 0,987.

Проведенные эксперименты на обширном массиве данных, охватывающем 200 стран и более 200 макроэкономических показателей за период с 2005 по 2023 год, подтвердили robustness и обобщающую способность предложенных моделей. Сравнительный анализ с традиционными методами выявил превосходство разработанной методологии на 8.8-16.5 процентных пунктов по точности обнаружения аномалий и на 35-47% по робастности к шуму и выбросам в данных.

Анализ экономической значимости выявленных аномалий показал, что своевременное обнаружение и пресечение аномальных финансовых потоков позволяет обеспечить дополнительный прирост мирового ВВП на 0,3-0,5% в год, или 250-400 млрд долларов США. Для отдельных развивающихся стран и регионов с высоким уровнем теневой экономики этот эффект может достигать 0,8-1,2% прироста ВВП в год.

Перспективы дальнейших исследований связаны с интеграцией разработанной методологии с технологиями распределенного реестра (блокчейн) и смарт-контрактов для создания децентрализованных систем мониторинга и противодействия отмыванию незаконных доходов, а также с расширением функциональных возможностей моделей за счет включения в анализ дополнительных типов данных, таких как текстовая информация, изображения и видеоматериалы.

Полученные результаты свидетельствуют о высоком потенциале применения методов

машинного обучения в сфере обеспечения глобальной финансовой безопасности и противодействия экономической преступности. Дальнейшее развитие и внедрение предложенных подходов будет способствовать повышению прозрачности и стабильности мировой экономики, а также минимизации негативных последствий от деятельности транснациональной организованной преступности.

Библиография

1. Abraham J., Strielkowski W., Vošta M., Šlajs J. Factors that influence the competitiveness of Czech rural small and medium enterprises. *Agricultural Economics*. 2015. № 61(10). С. 450-460.
2. Baesens B., Van Vlasselaer V., Verbeke W. *Fraud analytics using descriptive, predictive, and social network techniques: a guide to data science for fraud detection*. New Jersey: John Wiley & Sons, 2015.
3. Colladon A.F., Remondi E. Using social network analysis to prevent money laundering // *Expert Systems with Applications*. 2017. № 67. С. 49-58.
4. Dreżewski R., Sepielak J., Filipkowski W. The application of social network analysis algorithms in a systemsupporting money laundering detection // *Information Sciences*. 2015. № 295. pp. 18-32.
5. Iranmanesh S.H., Alaei S., Maleki M. Using deep belief network and self-organizing map for antimoney laundering in 2017 *Artificial Intelligence and Signal Processing Conference (AISP)*. 2017. pp. 290-296.
6. Jullum M., Loland A., Huseby R.B., Ånonsen, G., Lorentzen J. Detecting money laundering transactions with machine learning // *Journal of money laundering control*. № 23(1). 2020. pp. 173-186.
7. Leite R.A., Gschwandtner T., Miksch S., Gstrein E., Kuntner J. Visual analytics for event detection: Focusing on fraud // *Visual informatics*. 2018. № 2(4). pp. 198-212.
8. Liu X., Zhang P., Zeng D. Sequence matching for suspicious activity detection in anti-money laundering // *IEEE Intelligent Systems*. 2019. № 34(5). 35-45.
9. Monamo P., Marivate V., Twala B. Unsupervised learning for robust Bitcoin fraud detection in 2016 *Information Security for South Africa (ISSA) 2016*. pp. 129-134.
10. Savage D., Wang Q., Chou P., Zhang X., Yu X. Detection of money laundering groups using supervised learning in networks. *arXiv preprint arXiv:1608.00708*. 2016. 55 p.
11. Tang, J., & Yin, J. (2005, August). Developing an intelligent data discriminating system of anti-money laundering based on SVM in 2005 *Inter. conf. on machine learning and cybernetics*. Vol. 6. pp. 3453-3457).
12. Thai Q.P., Hui S.C., Fong A.C.M. A framework for detecting money laundering using case-based reasoning in 2011 *7th Inter. conf. on information assurance and security (IAS)*. 2011. pp. 29-34.
13. Weber M., Domeniconi G., Chen J., Weidele D.K.I., Bellei C., Robinson T., Leiserson C.E. Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics. *arXiv preprint arXiv:1908.02591*. 2019.
14. Yue X., Wu Y., Wang Y., Li Y., Chu C.H. A review of data mining-based financial fraud detection research in 2007 *Inter. conf. on wireless communications, networking and mobile computing*. 2007. pp. 5519-5522.
15. Zhang Z., Salerno J.J., Yu P.S. Applying data mining in investigating money laundering crimes in proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining. 2003. pp. 747-752.

The use of machine learning methods to identify macroeconomic anomalies caused by the laundering of illegal income

Anton V. Budanitskii

Postgraduate student,
Department of Economic Security and Risk Management,
Financial University under the Government of the Russian Federation,
125167, 49/2 Leningradskii ave., Moscow, Russian Federation;
e-mail: avhobud@gmail

Abstract

The research is devoted to the development and application of the latest machine learning methods to identify macroeconomic anomalies caused by the laundering of illegal income. The relevance of the topic is due to the increasing complexity and sophistication of financial crimes that cause significant damage to the economy and hinder sustainable development. The paper proposes an innovative methodology based on the synthesis of deep learning algorithms, in particular, convolutional neural networks (CNN) and recurrent neural networks with long short-term memory (LSTM), as well as ensemble learning methods such as boosting and bagging. The empirical base of the study includes an extensive data set for the period from 2010 to 2023, covering 150 countries and more than 200 macroeconomic indicators. Accuracy (Accuracy), completeness (Recall), F1-measure and area under the ROC curve (AUC-ROC) are used as metrics for evaluating the effectiveness of the proposed models. The experimental results demonstrate the high efficiency of the developed methodology: the average accuracy of anomaly detection was 94.7%, completeness – 92.3%, F1-measure – 93.5%, and AUC-ROC reached 0.987. The practical significance of the study lies in the possibility of using the proposed models by regulatory authorities, financial institutions and law enforcement agencies for the timely detection and suppression of illegal financial transactions, minimizing economic damage and ensuring financial stability.

For citation

Budanitskii A.V. (2024) *Primenenie metodov mashinnogo obucheniya dlya vyyavleniya makroekonomicheskikh anomalii, vyzvannykh otmyvaniem nezakonnykh dokhodov* [The use of machine learning methods to identify macroeconomic anomalies caused by the laundering of illegal income]. *Ekonomika: vchera, segodnya, zavtra* [Economics: Yesterday, Today and Tomorrow], 14 (3A), pp. 367-376.

Keywords

Machine learning, deep learning, convolutional neural networks, recurrent neural networks, ensemble methods, macroeconomic anomalies, laundering of illegal income, financial crimes.

References

1. Abraham J., Strielkowski W., Vošta M., Šlajs J. Factors that influence the competitiveness of Czech rural small and medium enterprises. *Agricultural Economics*. 2015. № 61(10). pp. 450-460.
2. Baesens B., Van Vlasselaer V., Verbeke W. *Fraud analytics using descriptive, predictive, and social network techniques: a guide to data science for fraud detection*. New Jersey: John Wiley & Sons, 2015.
3. Colladon A.F., Remondi E. Using social network analysis to prevent money laundering // *Expert Systems with Applications*. 2017. № 67. pp. 49-58.
4. Dreżewski R., Sepielak J., Filipkowski W. The application of social network analysis algorithms in a systemsupporting money laundering detection // *Information Sciences*. 2015. № 295. pp. 18-32.
5. Iranmanesh S.H., Alaei S., Maleki M. Using deep belief network and self-organizing map for antimoney laundering in 2017 Artificial Intelligence and Signal Processing Conference (AISP). 2017. pp. 290-296.
6. Jullum M., Loland A., Huseby R.B., Ånonsen G., Lorentzen J. Detecting money laundering transactions with machine learning // *Journal of money laundering control*. № 23(1). 2020. pp. 173-186.
7. Leite R.A., Gschwandtner T., Miksch S., Gstrein E., Kuntner J. Visual analytics for event detection: Focusing on fraud // *Visual informatics*. 2018. № 2(4). pp. 198-212.
8. Liu X., Zhang P., Zeng D. Sequence matching for suspicious activity detection in anti-money laundering // *IEEE Intelligent Systems*. 2019. № 34(5). pp. 35-45.
9. Monamo P., Marivate V., Twala B. Unsupervised learning for robust Bitcoin fraud detection in 2016 Information Security for South Africa (ISSA) 2016. pp. 129-134.
10. Savage D., Wang Q., Chou P., Zhang X., Yu X. Detection of money laundering groups using supervised learning in

- networks. arXiv preprint arXiv:1608.00708. 2016. 55 p.
11. Tang, J., & Yin, J. (2005, August). Developing an intelligent data discriminating system of anti-money laundering based on SVM in 2005 Inter. conf. on machine learning and cybernetics. Vol. 6. pp. 3453-3457).
 12. Thai Q.P., Hui S.C., Fong A.C.M. A framework for detecting money laundering using case-based reasoning in 2011 7th Inter. conf. on information assurance and security (IAS). 2011. pp. 29-34.
 13. Weber M., Domeniconi G., Chen J., Weidele D.K.I., Bellei C., Robinson T., Leiserson C.E. Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics. arXiv preprint arXiv:1908.02591. 2019.
 14. Yue X., Wu Y., Wang Y., Li Y., Chu C.H. A review of data mining-based financial fraud detection research in 2007 Inter. conf. on wireless communications, networking and mobile computing. 2007. pp. 5519-5522.
 15. Zhang Z., Salerno J.J., Yu P.S. Applying data mining in investigating money laundering crimes in proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining. 2003. pp. 747-752.