

Разработка модели прогнозирования уровня отмывания преступных доходов

Анисимов Ефим Сергеевич

Студент, ORCID: 0000-0002-2632-4439, e-mail: EAnisimov_Sci@mail.ru

Бекетнова Юлия Михайловна

Канд. техн. наук, доц. Департамента информационной безопасности
ORCID: 0000-0002-1005-6265, e-mail: beketnova@mail.ru

Финансовый Университет при Правительстве Российской Федерации (Финуниверситет), г. Москва, Россия

Аннотация

В статье предложена модель прогнозирования уровня легализации доходов, полученных преступным путем, на основе данных Министерства внутренних дел Российской Федерации по состоянию преступности экономической направленности в России с начала 2011 г. С помощью сезонной интегрированной модели авторегрессии – скользящего среднего (SARIMA) – сравнены различные регрессионные модели для задач исследования (линейная регрессия, логистическая регрессия, авторегрессия и SARIMA). Подчеркнута необходимость учета сезонных закономерностей в структуре отмывания преступных доходов, выбора модели SARIMA с наименьшими отклонениями от действительных значений. В статье представлены результаты анализа данных с применением метода наименьших квадратов (англ. least squares), подсчета среднеквадратической ошибки – MSE (англ. mean squared error). Отмечена высокая точность получаемых краткосрочных прогнозов: отклонение от действительного числа составляет около трех дел (при среднем количестве в 68 преступлений за последние 10 лет). Модель прогнозирования можно рекомендовать для внедрения в аналитические комплексы субъектов финансового мониторинга и надзорных органов.

Ключевые слова

Финансовый мониторинг, противодействие отмыванию доходов, экономические преступления, прогнозирование правонарушений, статистический анализ

Для цитирования: Анисимов Е.С., Бекетнова Ю.М. Разработка модели прогнозирования уровня отмывания преступных доходов // Вестник университета. 2022. № 5. С. 136–143.

Development of a model for predicting money laundering rate

Efim S. Anisimov

Student, ORCID: 0000-0002-2632-4439, e-mail: EAnisimov_Sci@mail.ru

Julia M. Beketnova

Cand. Sci. (Tech.), Assoc. Prof. at the Information Security Department
ORCID: 0000-0002-1005-6265, e-mail: beketnova@mail.ru

Financial University, Moscow, Russia

Abstract

The article suggests model for predicting the level of money laundering on the basis of data from the Ministry of Internal Affairs of the Russian Federation on the state of economic crime in Russia since the beginning of 2011. Using a seasonally integrated autoregressive moving average (SARIMA) model, it compares different regression models for the research tasks (linear regression, logistic regression, autoregressive and SARIMA). The necessity of taking into account seasonal regularities in the structure of money laundering was underlined, and the SARIMA model with the lowest deviations from the actual values was chosen. The necessity of taking into account seasonal regularities in the structure of money laundering was underlined, and the SARIMA model with the lowest deviations from the actual values was chosen. The article presents the results of data analysis using the method of least squares, calculating the mean squared error (MSE). High accuracy of short-term forecasts was noted: the deviation from the actual number of cases is about three cases (with the average number of cases being 68 over the last 10 years). The forecasting model can be recommended for implementation in the analytical complexes of financial monitoring and supervisory authorities.

Keywords

Financial monitoring, anti-money laundering, economic crime, crime forecasting, statistical analysis

For citation: Anisimov E.S., Beketnova J.M. (2022) Development of a model for predicting money laundering rate. *Vestnik universiteta*, no. 5, pp. 136–143.



ВВЕДЕНИЕ

Стандартами Группы разработки финансовых мер по борьбе с отмыванием денег (далее – ФАТФ; от англ. FATF, Financial Action Task Force) заложены принципы риск-ориентированного подхода при работе в области противодействия отмыванию доходов, полученных преступным путем, и финансированию терроризма (далее – ПОД/ФТ). Важную роль играет прогнозирование развития событий с целью проведения превентивных мероприятий. ПОД/ФТ окажется более результативным, если финансовая разведка будет готова к динамике изменения активности рассматриваемого преступления. Один из вариантов решения задачи готовности заключается в прогнозировании уровня преступлений легализации преступных доходов. Поддержка финансового мониторинга в виде предсказания вероятного количества противоправных действий позволит: повысить выявляемость предикатных преступлений, оптимизировать перераспределение ресурсов подразделений финансовых разведок, формировать связи активности с внешними факторами (среди которых требования законодательства, налоговые периоды в различных юрисдикциях, результаты финансовой деятельности предприятий по секторам экономики).

Помимо отмеченной практической значимости исследования актуальность темы подкрепляется текущими мировыми событиями по направлению ПОД/ФТ. В отчетах ФАТФ можно проследить тенденцию к расширению многообразия применяемых для нелегальных целей инструментов, а также заметить обновления серого списка стран, который на регулярной основе подвергается изменениям. Последний факт говорит в пользу тезиса о периодических обострениях проблем легализации преступных доходов в тех или иных регионах. По обозначенным причинам исследуемая тема является актуальной и практически значимой.

ОБЗОР НАУЧНОЙ ЛИТЕРАТУРЫ

Существуют разные толкования термина «отмывание (легализация) доходов», однако все они сводятся к одному краткому характерному признаку: операции над доходами или имуществом, полученными преступным путем, проводятся с целью скрытия их реального происхождения [1]. В Российской Федерации и в большей части стран это явление – уголовно наказуемое деяние ввиду серьезного дестабилизирующего влияния, оказываемого субъектами схем отмывания на элементы государственной безопасности. Отечественные ученые, например А.Е. Шалагин и А.Д. Идиятуллов [2], подчеркивают особую значимость предиктивной аналитики, прогнозирования уровня преступности и предупреждения. Распиряются практические приложения по прогнозированию преступлений, которые основаны на различных методах: такие разработки были изучены на примерах работ зарубежных авторов D. Robertson и W. Goodridge, использовавших скрытую марковскую модель, оценку апостериорного максимума, корреляцию Спирмена и основанную на плотности пространственную кластеризацию [3]; G. Hajela, Dr. M. Chawla и Dr. A. Rasool [4], обратившихся к двумерному анализу горячих точек и кластерному анализу.

В рекомендациях Базельского комитета по поводу политики «Знай своего клиента» (англ. know your customer), изданных еще в 2001 г., подчеркивается необходимость чувствительности постоянного мониторинга к рискам [5]. Эффективные процедуры проверки в рамках реализации данного принципа могут основываться на риск-ориентированном подходе с динамическим отнесением деятельности клиентов к категориям риска. Такое динамическое отнесение, по мнению авторов, должно оказывать поддержку в случаях отклонения поведения клиента, подпадающего чаще в категорию низкого риска, по причине чего в его отношении проводится усеченный комплекс проверочных процедур. Отмеченный принцип на основе прогнозирования будет в периоды предполагаемой повышенной недобросовестной активности расширять группу клиентов финансовых организаций с высоким риском, позволяя при этом потенциально повысить предупреждение совершения экономических преступлений и повысить уровень их раскрываемости.

Центральный банк Российской Федерации в настоящее время проводит тестирование платформы «Знай своего клиента», предназначенной для снижения затрат банков на выполнение требований ПОД/ФТ [6]. Данный информационный сервис на базе Банка России позволит повысить эффективность лежащей проверки клиентов участниками российской банковской системы за счёт централизованности. Хотя присоединение кредитных организаций не является обязательным, значительная доля игроков банковского рынка показала интерес к получению дополнительной информации от платформы.

Примеры отражают исключительную роль информационно-аналитических механизмов в противодействии отмыванию доходов, полученных преступным путем. Информационно-аналитическая деятельность

содержит элементы прогнозирования – это подтверждение практической значимости настоящего исследования. Прогнозирование может основываться на различных методах и данных. Например, у исследователей в работе [3] прогнозирование уровня преступности строилось по месяцам в разрезе отдельных территориальных округов, а в научной статье [4] отмечаются подходы анализа на основе географического профилирования.

При построении модели предсказания отмывания преступных доходов авторы после визуального анализа временного ряда исторических данных (подробное рассмотрение приведено далее в разделе «Подготовка данных эмпирического исследования») считают важным учет сезонности в динамике легализации доходов. Исследование различных научных работ показало, что одной из подходящих моделей является сезонная интегрированная модель авторегрессии скользящего среднего (SARIMA). Примеры успешного использования этой модели в различных сферах были рассмотрены у ряда авторов [7; 8; 9].

Таким образом, обзор научных взглядов на различные элементы выбранного направления исследования позволил найти подтверждение актуальности и практической значимости информационно-аналитической деятельности в ПОД/ФТ и, в том числе, создания прогнозов динамики противоправной активности. Помимо этого, анализ научных статей предоставил методологическую базу для разрабатываемой модели.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ

ПОДГОТОВКА ДАННЫХ ЭМПИРИЧЕСКОГО ИССЛЕДОВАНИЯ

Для построения прогнозной модели использованы данные официальной статистики Министерства внутренних дел Российской Федерации (далее – МВД России). Число выявленных в отчетном периоде преступлений, классифицируемых как «легализация (отмывание) денежных средств или иного имущества, приобретенных лицом в результате совершения им преступления либо приобретенных другими лицами преступным путем», из отчетов МВД России «Состояние преступности в России» анализировалось за период 2011–2022 гг. [10]. Данный показатель отражает динамику преступлений по ст. 174 и 174.1 Уголовного кодекса Российской Федерации.

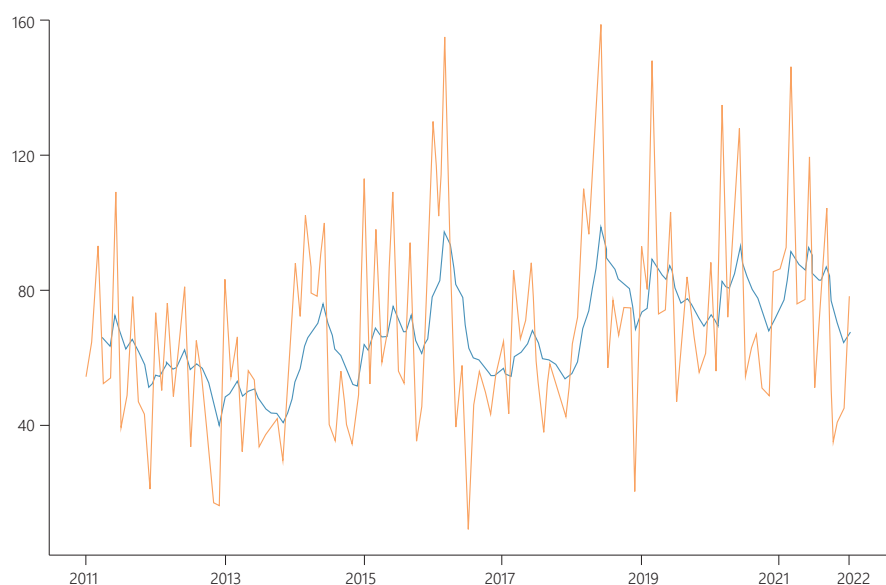
Рассмотрим порядок решения задачи прогнозирования, применяемый в рамках исследования. Анализ временного ряда, сформированного в ходе обзора 133 отчетов МВД России (с января 2011 г. по январь 2022 г.), состоял из следующих стадий:

- сглаживание ряда;
- исследование ряда на стационарность и принятие решения о приведении его к стационарному виду;
- исследование сезонности ряда;
- выбор регрессионной

модели: построение и вычисление статистик об определении значимости модели;

- построение прогноза.

Подробно опишем методологическую основу каждой из перечисленных стадий. Сглаживание временного ряда осуществлялось при помощи экспоненциального сглаживания на четырех предшествующих значениях с параметром $\alpha = 0,2$, что представляется формулой (1). В результате временной ряд изменился, как показано на рисунке 1 (сглаженный набор показан синим цветом).



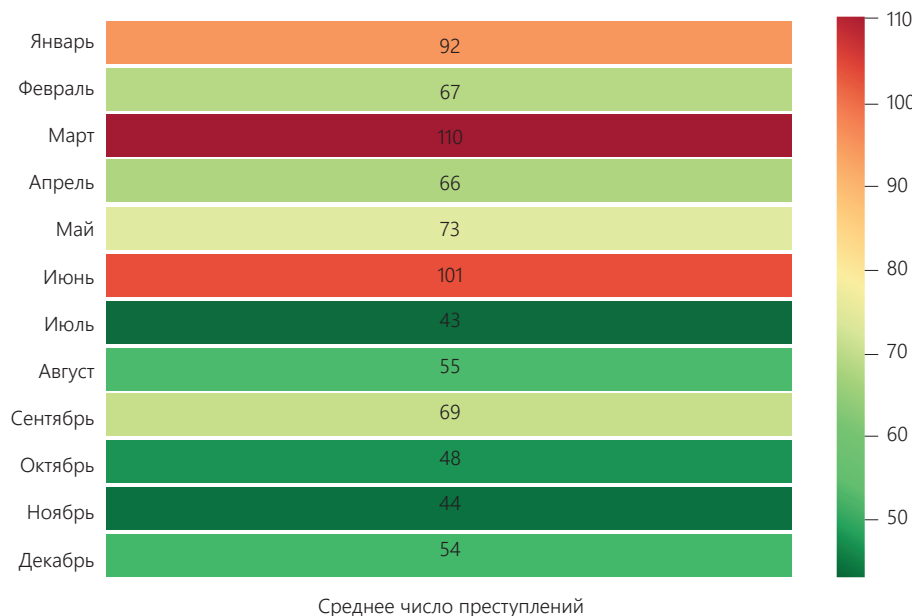
Составлено авторами по материалам источника [10]

Рис. 1. Экспоненциальное сглаживание ряда

$$x_{\text{exp}_i} = \begin{cases} \frac{x_1 + x_2 + x_3 + x_4}{4}, i = 1 \\ \alpha \cdot x_{i+3} + (1 - \alpha) \cdot x_{\text{exp}_{i-1}}, i > 1 \end{cases} \quad (1)$$

Сглаженная выборка использовалась для исследования на стационарность при помощи теста Дики–Фуллера. Его проведение показало, что временной ряд нестационарен. При помощи взятия первых разностей был получен стационарный временной ряд, что подтвердил тест Дики–Фуллера.

Данное действие по приведению ряда к стационарному виду, в том числе, было необходимо для формирования значений нескольких параметров в выбранной далее модели временных рядов.



Составлено авторами по материалам источника [10]

Рис. 2. Тепловая карта сезонности преступлений

чительных месяцах первого и второго кварталов. Июнь при этом не только завершает второй квартал, но и первое полугодие в целом. Наименьшая активность стабильно наблюдается в завершении года — четвертом квартале.

Подобные сезонные закономерности необходимо было учесть в построении прогноза, поэтому в качестве модели исследования была выбрана сезонная интегрированная модель авторегрессии — скользящего среднего (SARIMA). Определение всех параметров модели получено при помощи анализа полной и частичной функций автокорреляции (ACF и PACF), которые строились на основе составленного ранее стационарного ряда. Тренировка модели позволила сформировать тестовый прогноз для сравнения истинных и полученных прогнозных значений в рамках временного интервала от января 2020 г. до января 2022 г. Прогноз на 2022 г. был сформирован после обучения модели на полных данных. Перед анализом составленного прогноза отметим качество созданной модели SARIMA с помощью коэффициента детерминации. Вычисления производились на прогнозных значениях модели для 25 последних заданных статистических данных. Коэффициент детерминации составил 0,88, что для модели с таким малым объемом исторических обучающих данных является качественным показателем, предоставляющим основание для признания созданной модели значимой.

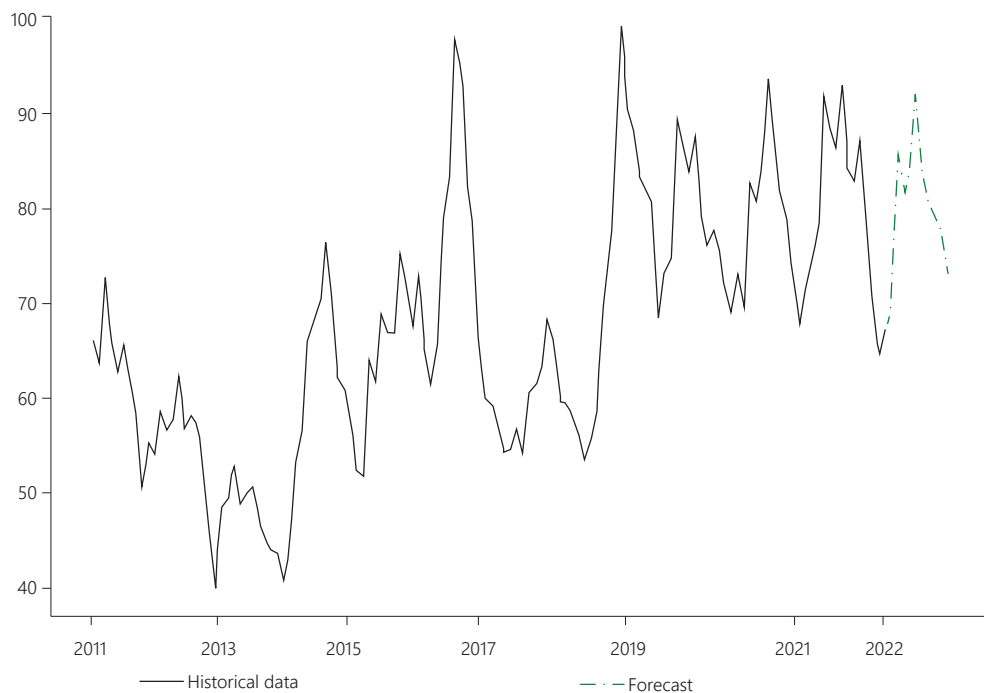
Среднеквадратическая ошибка (англ. mean squared error) MSE составила 8,76, это позволило заключить, что отклонение в почти три преступления при построении прогноза является отличным показателем (4 % от среднего числа преступлений за последние 10 лет). Таким образом, предсказания модели являются хорошим ориентиром при формировании краткосрочных оперативных прогнозов, показывая в среднем небольшие отклонения от фактических значений.

АНАЛИЗ ДАННЫХ И СОЗДАНИЕ МОДЕЛИ

Следующим этапом был анализ сезонности, для чего рассматривалась спектральная плотность, а также средние значения выявляемых преступлений по месяцам. На основе средних значений была построена тепловая карта (рис. 2), которая хорошо отражает динамику легализации доходов, полученных преступным путем в рамках календарного года.

Результат построения тепловой карты позволил выделить месяцы, характеризующиеся наивысшей активностью: в начале года (в январе), а также в заклю-

График, включающий исторические сглаженные данные количества преступлений, квалифицируемых как «легализация (отмывание) денежных средств или иного имущества, приобретенных лицом в результате совершения им преступления либо приобретенных другими лицами преступным путем» и их прогноз на оставшиеся 11 месяцев 2022 г., представлен на рисунке 3 (построенные нижняя и верхняя границы соответствуют 95 %-му доверительному интервалу).



Составлено авторами по материалам исследования

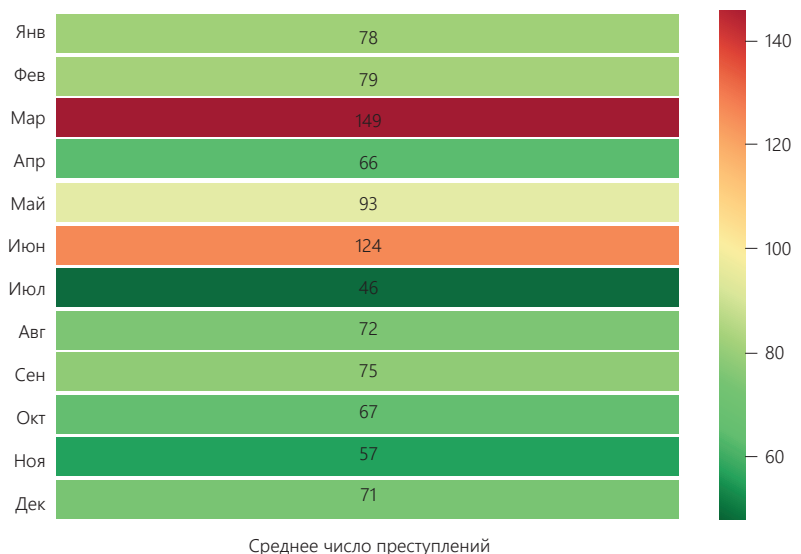
Рис. 3. Прогноз вместе с исторической динамикой числа преступлений

АНАЛИЗ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

Рассмотрим полученные прогнозные значения на 2022 г. более детально. В 2021 г. общее количество преступлений показало снижение по сравнению с 2020 г. на одно преступление (с 950 до 949 дел). В 2022 г. модель прогнозирует возможный незначительный рост годового количества до 977 случаев отмывания доходов (+2,95 % по сравнению с 949 преступлениями в 2021 г.). Вид сезонности сохранился: пики активности прогнозируются по-прежнему в конце первого квартала и первого полугодия. На рисунке 4 представлен прогноз на 2022 г. по месяцам, созданный при помощи пересчета сглаженных значений.

ЗАКЛЮЧЕНИЕ

Таким образом, в рамках исследования проведен анализ количественной структуры преступлений, квалифицируемых как отмывание (легализация) доходов, полученных преступным путем. Данный анализ предполагал рассмотрение сезонности, которая была успешно описана: больше всего попыток переброса черных денег и иного имущества в зеленую зону стабильно выявляется



Составлено авторами по материалам исследования

Рис. 4. Прогноз на 2022 г. по месяцам

правоохранительными органами в марте и июне. Подобная зависимость отражает связь действий схем легализации с окончаниями основных налоговых периодов и особенностями уплаты различных сборов и платежей, в том числе и в офшорных юрисдикциях. Помимо сезонности большое внимание уделено возможности прогнозирования, поскольку качественные оперативные прогнозы позволяют планировать мероприятия по предупреждению, пресечению и повышенному выявлению противоправной деятельности.

Позитивным наблюдением стала фиксация снижения количества преступлений в годовом выражении, что свидетельствует об успешных результатах работы Федеральной службы финансового мониторинга, ФСБ России, МВД России, Прокуратуры, регуляторов (например, Центрального банка Российской Федерации) в различных областях с повышенным риском вовлечения в процессы отмыwania доходов, а также других субъектов финансового мониторинга. Фиксация снижения не первый год является весомым практическим подтверждением высоких оценок отечественной антиотмывочной системы со стороны международного сообщества, включая ФАТФ.

Действия бизнеса формируют положительную основу для продолжения отмеченного вектора снижения уровня преступности по рассматриваемому направлению. Усилия государства, формирование эффективной законодательной базы позволяют комплексно предупреждать возникновение подозрительных операций. Ярким примером таких действий бизнеса является активная редомициляция вчерашних резидентов офшорных юрисдикций в российские специальные административные районы.

Авторы считают, что созданный механизм и модель в рамках исследования могут найти успешное применение в задачах прогнозирования преступлений экономической направленности (субъектами финансового мониторинга и надзорными органами), а также имеют свойство расширения сферы применения. Выразим уверенность, что национальные механизмы противодействия отмыванию доходов, полученных преступным путем, и финансированию терроризма не утратят своей актуальности, продолжат своё развитие, укрепляя основы государственной безопасности и устойчивого экономического развития страны.

Библиографический список

1. Rocha-Salazar J.-d.-J., Segovia-Vargas M.-J., Camacho-Miñano M.-d.-M. Money laundering and terrorism financing detection using neural networks and an abnormality indicator. *Expert Systems with Applications*. 2021;169:114470. <https://doi.org/10.1016/j.eswa.2020.114470>
2. Шалагин А.Е., Идиятуллин А.Д. Трансформация преступности в XXI веке: особенности предупреждения и противодействия. *Вестник Казанского юридического института МВД России*. 2021;12(2):227–235. <https://doi.org/10.37973/KUI.2021.95.12.016>
3. Robertson D.L., Goodridge W.A. Predicting density of serious crime incidents using a Multiple-Input Hidden Markov Maximization a posteriori model. *Machine Learning with Applications*. 2022;7:100231. <http://dx.doi.org/10.1016/j.mlwa.2021.100231>
4. Hajela G., Chawla M., Rasool A. A Clustering Based Hotspot Identification Approach For Crime Prediction. *Procedia Computer Science*. 2020;167:1462–1470. <https://doi.org/10.1016/j.procs.2020.03.357>
5. Рекомендации Базельского комитета «Знай своего клиента». *Надлежащее отношение банков к клиентам (Стандарты Базельского комитета банковского надзора)*. <https://docs.cntd.ru/document/901934950> (дата обращения: 26.03.2022).
6. Платформа «Знай своего клиента» сократит нагрузку на добросовестных предпринимателей. <http://www.cbr.ru/press/event/?id=12411> (дата обращения: 06.03.2022).
7. ArunKumar K.E., Kalaga D.V., Mohan Sai Kumar Ch., Chilkoor G., Kawaji M., Brenza T.M. Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA). *Applied Soft Computing*. 2021;103:107161. <https://doi.org/10.1016/j.asoc.2021.107161>
8. Farsi M., Hosahalli D., Manjunatha B.R., Gad I., Atlam El-S., Ahmed A., Elmarhomy G., Elmarhoumy M., Ghoneim O.A. Parallel genetic algorithms for optimizing the SARIMA model for better forecasting of the NCDC weather data. *Alexandria Engineering Journal*. 2021;60(1):1299–1316. <https://doi.org/10.1016/j.aej.2020.10.052>
9. Mao Q., Zhang K., Yan W., Cheng C. Forecasting the incidence of tuberculosis in China using the seasonal auto-regressive integrated moving average (SARIMA) model. *Journal of Infection and Public Health*. 2018;11(5):707–712. <https://doi.org/10.1016/j.jiph.2018.04.009>
10. Министерство внутренних дел Российской Федерации. *Состояние преступности*. <https://мвд.рф/reports> (дата обращения: 06.03.2022).

References

1. Rocha-Salazar J.-d.-J., Segovia-Vargas M.-J., Camacho-Miñano M.-d.-M. Money laundering and terrorism financing detection using neural networks and an abnormality indicator. *Expert Systems with Applications*. 2021;169:114470. <https://doi.org/10.1016/j.eswa.2020.114470>
2. Shalagin A.E., Idiyatullo A.D. Transformation of Crime in the XXI Century: Features of Prevention and Counteraction. *Bulletin of the Kazan Law Institute of the Ministry of Internal Affairs of Russia*. 2021;12(2):227–235. <https://doi.org/10.37973/KUI.2021.95.12.016>
3. Robertson D.L., Goodridge W.A. Predicting density of serious crime incidents using a Multiple-Input Hidden Markov Maximization a posteriori model. *Machine Learning with Applications*. 2022;7:100231. <http://dx.doi.org/10.1016/j.mlwa.2021.100231>
4. Hajela G., Chawla M., Rasool A. A Clustering Based Hotspot Identification Approach For Crime Prediction. *Procedia Computer Science*. 2020;167:1462–1470. <https://doi.org/10.1016/j.procs.2020.03.357>
5. Basel Committee Recommendations “Know your customer “. *Customer due diligence for banks (Basel Committee on Banking Supervision standards)*. <https://docs.cntd.ru/document/901934950> (accessed: 26.03.2022).
6. “Know your customer” platform will reduce the burden on bona fide entrepreneurs. <http://www.cbr.ru/press/event/?id=12411> (accessed 06.03.2022).
7. ArunKumar K.E., Kalaga D.V., Mohan Sai Kumar Ch., Chilkoor G., Kawaji M., Brenza T.M. Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA). *Applied Soft Computing*. 2021;103:107161. <https://doi.org/10.1016/j.asoc.2021.107161>
8. Farsi M., Hosahalli D., Manjunatha B.R., Gad I., Atlam El-S., Ahmed A., Elmarhomy G., Elmarhoumy M., Ghoneim O.A. Parallel genetic algorithms for optimizing the SARIMA model for better forecasting of the NCDC weather data. *Alexandria Engineering Journal*. 2021;60(1):1299–1316. <https://doi.org/10.1016/j.aej.2020.10.052>
9. Mao Q., Zhang K., Yan W., Cheng C. Forecasting the incidence of tuberculosis in China using the seasonal auto-regressive integrated moving average (SARIMA) model. *Journal of Infection and Public Health*. 2018;11(5):707–712. <https://doi.org/10.1016/j.jiph.2018.04.009>
10. The Ministry of Internal Affairs of the Russian Federation. *Crime rate in Russia*. <https://мвд.рф/reports> (accessed 06.03.2022).