

УДК 004.93:343

**ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ
МОНИТОРИНГА СОСТОЯНИЯ ЗДОРОВЬЯ ЛИЦ,
СОДЕРЖАЩИХСЯ В УЧРЕЖДЕНИЯХ УГОЛОВНО-
ИСПОЛНИТЕЛЬНОЙ СИСТЕМЫ**

Д.Л. Барковская¹⁾, В.А. Новиков²⁾

1) студентка 1 курса АНОО ВО «Кубанский институт профессионального образования» г. Краснодар, Россия

2) канд. технич. наук, доцент АНОО ВО «Кубанский институт профессионального образования» г. Краснодар, Россия; vsenovik@bk.ru

Аннотация. В статье предлагается концепция применения машинного обучения для мониторинга здоровья лиц в учреждениях уголовно-исполнительной системы. В среде Loginom Community реализуется аналитический конвейер: от очистки данных до построения модели логистической регрессии, позволяющей выявлять аномальные изменения показателей и классифицировать наблюдения по признаку вероятной симуляции заболевания. Практическая значимость — повышение объективности медицинского контроля и стандартизация оценки рисков.

Ключевые слова: машинное обучение, Loginom Community, мониторинг здоровья, уголовно-исполнительная система, логистическая регрессия, выявление аномалий.

**APPLICATION OF MACHINE LEARNING FOR HEALTH
MONITORING OF PERSONS HELD IN PENAL INSTITUTIONS**

D.L. Barkovskaya¹⁾, V.A. Novikov²⁾

1) 1st-year student, Autonomous Non-Commercial Higher Education Institution “Kuban Institute of Professional Education”, Krasnodar, Russia 930copper@tiffincrane.com

2) Candidate of Technical Sciences, Associate Professor, Autonomous Non-Commercial Higher Education Institution “Kuban Institute of Professional Education”, Krasnodar, Russia; vsenovik@bk.ru

Annotation. The paper proposes a concept of applying machine learning for health monitoring of persons held in penal institutions. In the Loginom Community environment, an analytical pipeline from data cleaning to logistic regression is implemented to detect anomalous changes in medical indicators and classify observations by the likelihood of malingering. The practical significance lies in improving objectivity of medical supervision and standardizing risk assessment.

Keywords: machine learning, Loginom Community, health monitoring, penal system, logistic regression, anomaly detection.

Введение

Цифровизация уголовно-исполнительной системы создает предпосылки для накопления структурированных медицинских данных, пригодных для аналитической обработки. Это позволяет использовать методы машинного обучения для раннего выявления критических состояний, обнаружения аномальной динамики показателей и сопоставления исходных и последующих медицинских данных.

Практическая актуальность такого подхода обусловлена тем, что решения о специализированном обследовании, изменении условий содержания, отсрочке исполнения наказания либо освобождении по болезни зависят от полноты и достоверности медицинской информации. При отсутствии унифицированного входного профиля состояния здоровья возрастает риск пропуска опасных состояний и появления спорных случаев, когда существенное изменение диагноза фиксируется уже на более поздних стадиях уголовного преследования или исполнения наказания.

Цель исследования состоит в теоретическом обосновании применения машинного обучения для мониторинга состояния здоровья заключённых и в описании аналитической модели, реализуемой в Loginom Community и предназначенной для выявления аномалий в медицинских данных.

Теоретические и эмпирические предпосылки. Пенитенциарная медицина страдает от задержек диагностики и фрагментарности записей, что создаёт основу для применения предиктивной аналитики (исследования Edge et al., 2025; Digitalization of Prison Records, 2025) [1-3].

Отдельное направление исследований связано с выявлением симуляции заболеваний. К числу возможных маркеров относятся внутренние противоречия в описании жалоб, несоответствие между заявляемыми симптомами и объективными данными, а также аномальные сочетания признаков, нехарактерные для подтвержденных случаев заболевания. Статистические методы, включая логистическую регрессию, могут использоваться для отделения подозрительных наблюдений от типичных при условии качественной подготовки признаков и обязательной экспертной интерпретации результатов [4-5].

Эмпирическая применимость модели определяется возможностью накопления стандартной базы наблюдений, включающей результаты первичного осмотра, данные о хронических заболеваниях, лабораторные показатели, результаты инструментальной диагностики, сведения о повторных обращениях, параметрах двигательной активности, сна, изменении массы тела и противоречиях между последовательными медицинскими заключениями. Формирование исходного цифрового профиля состояния здоровья на входе в систему создает основу для

Материалы и методы. Материалом исследования в рамках предлагаемой концепции выступает обезличенный массив медицинских и сопутствующих регистрационных данных лиц, содержащихся в

учреждениях уголовно-исполнительной системы. В структуру выборки включаются сведения первичного медицинского осмотра, данные о хронических заболеваниях, лабораторных показателях, результатах инструментальных обследований, повторных обращениях за медицинской помощью, изменениях массы тела, показателях двигательной активности и последующих диагнозах.

В качестве инструментальной среды предлагается использовать платформу Loginom Community. Базовая цепочка в Loginom Community включает импорт, фильтрацию, группировку, заполнение пропусков, удаление дубликатов, редактирование выбросов и логистическую регрессию.

После загрузки данных выполняется исключение технических и неполных записей, не относящихся к целевой категории анализа. Затем формируются агрегированные характеристики по каждому лицу либо эпизоду наблюдения: число обращений за медицинской помощью, частота повторяющихся жалоб, интервалы между осмотрами и число пересмотров диагноза. На их основе рассчитываются производные признаки, включая коэффициент повторяемости жалоб, темп изменения массы тела, индекс нестабильности показателей, частоту расхождений между жалобами и объективными результатами обследований, а также индикаторы появления новых тяжёлых диагнозов после исходного осмотра.

Существенное значение имеет предварительная очистка массива. Пропущенные значения обрабатываются с учетом природы признака, одновременно проводится поиск дубликатов и выявление противоречивых записей. После этого выполняется редактирование выбросов, поскольку резкое отклонение показателей от типового профиля может свидетельствовать как о реальном ухудшении состояния, так и об аномалии данных.

Для итоговой классификации наблюдений используется логистическая регрессия, где зависимая переменная принимает значение 1 при наличии признаков вероятной симуляции либо атипичного сочетания медицинских и поведенческих характеристик и 0 при их отсутствии. Формирование целевой переменной предполагает предварительную экспертную разметку обучающей выборки. Тем самым модель предназначается не для самостоятельного медицинского заключения, а для раннего выявления наблюдений повышенного риска, требующих углублённой проверки. Качество классификации может оцениваться с использованием показателей *accuracy*, *precision*, *recall* и *AUC-ROC*.

Операциональное определение признака вероятной симуляции. В целях демонстрации работоспособности предлагаемого аналитического контура была сформирована синтетическая выборка 50 наблюдений, имитирующая обезличенные записи первичного медицинского осмотра и последующего наблюдения за лицами, содержащимися в учреждениях уголовно-исполнительной системы. В состав такой демонстрационной базы данных включены показатели, описывающие: возраст, число жалоб за

фиксированный период наблюдения (30 дней), изменение массы тела, количество обращений за медицинской помощью, число выявленных клинических и документарных противоречий в медицинских записях, наличие нового тяжёлого диагноза после входного осмотра, степень отклонения лабораторных показателей от референсных значений и факт снижения двигательной активности. Для каждого наблюдения дополнительно фиксируется бинарный класс, присвоенный экспертами.

Целевая переменная определяется операционально как признак наличия совокупности индикаторов вероятной симуляции заболевания либо иного атипичного поведения. В качестве базовых критериев разметки используются: клинические противоречия (заявляемые жалобы не подтверждаются данными осмотра, лабораторных и инструментальных исследований), документарные противоречия (взаимоисключающие сведения в медицинской документации без клинического объяснения), появление тяжёлого диагноза после входного осмотра без подтверждённой предшествующей динамики, а также чрезмерная частота однотипных жалоб при ограниченной объективной динамике. Значение 1 присваивается наблюдению при достижении пороговой суммы критериев (например, трёх и более), при этом окончательное решение принимается врачебной комиссией либо двумя независимыми экспертами с последующим согласованием спорных случаев; значение 0 присваивается при отсутствии совокупности указанных индикаторов. Тем самым целевая переменная получает чёткое операциональное определение и опирается на формализованный набор признаков, а не на интуитивную оценку.

Логистическая регрессия, обученная на такой демонстрационной выборке, используется исключительно для иллюстрации принципиальной применимости предложенного конвейера и не претендует на эмпирическую валидацию модели. Вместе с тем уже на уровне концепции возможно задать ориентиры качества, к которым должна стремиться пилотная реализация: с учётом приоритета минимизации пропуска потенциально значимых случаев целесообразно ориентироваться на значение recall не ниже 0,80 при AUC-ROC не ниже 0,75, даже если при этом значительная часть наблюдений будет отнесена к категории повышенного риска и направлена на углублённую экспертную проверку.

Ограничения исследования. Исследование носит теоретико-концептуальный характер и не предполагает апробации модели на реальном массиве данных уголовно-исполнительной системы, вследствие чего в его рамках не оцениваются фактическая точность классификации, устойчивость модели на независимой выборке и воспроизводимость результатов в различных категориях учреждений; дополнительным ограничением выступает специфика предметной области, поскольку медицинские данные лиц, содержащихся в учреждениях уголовно-исполнительной системы, относятся к числу чувствительных, а их качество зависит от полноты первичного осмотра, регулярности наблюдения, уровня цифровизации документации и доступности инструментальной диагностики, что означает

возможность воспроизведения моделью дефектов исходного массива, включая пропуски, несогласованность записей и институциональные смещения; при этом признак вероятной симуляции не может трактоваться как юридически или клинически окончательный вывод, поэтому логистическая регрессия в рассматриваемом контексте допустима только как вспомогательный инструмент предварительного отбора наблюдений, требующих углублённой врачебной и экспертной оценки.

Выводы и перспективы внедрения. Предложенная концепция применения логистической регрессии в Loginom Community для выявления аномалий в медицинских данных лиц в УИС теоретически обоснована и имеет практическую значимость. Приоритетным направлением является пилотная апробация на обезличенных данных с достижением *recall* $\geq 0,80$. Модель — вспомогательный инструмент, не заменяющий врачебную комиссию.

Список использованных источников:

1. Edge C., Hayward A., Whitfield A., Hard J. Prison health data collection: building an evidence base for improved health outcomes in custodial settings // Social Science and Medicine. 2025.
2. Allahyari E, Moshtagh M. Predicting mental health of prisoners by artificial neural network. Biomedicine (Taipei). 2021 Mar 1;11(1):26-33. doi: 10.37796/2211-8039.1031. PMID: 35223392; PMCID: PMC8823470.
3. The quality of prison primary care: cross-sectional cluster-level analyses of routinely collected electronic health record data // EClinicalMedicine. 2023.
4. Digitalization of Prison Records Supports Artificial Intelligence in Correctional Health Care // Criminal Justice and Behavior. 2025.
5. Calder I., May J. Artificial intelligence and the digital right to health in prisons // Prison Insider / Health through Walls analytical materials. 2026.