

Прогнозирование ухудшения состояния здоровья пациентов с сердечно-сосудистыми заболеваниями на основе методов машинного обучения

М.С. Цыганова^{1}, Д.А. Романов¹, А.Г. Мелентьева¹, А.В. Макарихин²*

¹Тюменский государственный университет, Тюмень, Россия

²ООО «ЛАБРИКС», Тюмень, Россия

**m.s.cyganova@utmn.ru*

Аннотация. Настоящее исследование выполнялось в рамках проекта по разработке подсистемы прогнозирования ухудшения состояния пациентов с сердечно-сосудистыми заболеваниями на платформе медицинской информационной системы «1С:Медицина. Больница». Актуальность поставленной задачи обусловлена особенно высокой опасностью данной группы заболеваний и необходимостью принятия своевременных решений о госпитализации или назначении лечения при возникновении риска ухудшения состояния пациента. Целью данной работы было создание инструмента, позволяющего лечащему врачу оперативно получать обоснованную оценку риска ухудшения состояния пациентов на основе имеющихся медицинских показателей. В рамках данного исследования был выполнен анализ более 30 тыс. записей, содержащих показатели здоровья пациентов, выгруженных из региональной медицинской информационной системы. Набор данных был размечен в соответствии с имеющейся в системе информацией о принятых врачебных решениях (лечащими врачами поликлиники и стационара). Отсутствие стандартизированного ввода значений показателей здоровья в используемой медицинской системе потребовало значительного объема работы по предварительной обработке входных данных и их подготовке для целей моделирования. Подготовленные данные использовались для построения предсказательной модели с применением методов машинного обучения. По результатам проведенных вычислительных экспериментов в качестве алгоритма обучения был выбран градиентный бустинг, подобраны оптимальные параметры этого алгоритма. Качество прогнозирования обученных моделей проверялось на данных размеченного набора, не участвовавших в процессе обучения. Показатели качества лучшей модели на тестовых данных составили precision = 0,87; recall = 0,96; AUC-ROC = 0,97. Реализована интеграция обученных моделей с автоматизированным рабочим местом лечащего врача в системе «1С:Медицина. Больница». Таким образом, разработан алгоритм обработки показателей здоровья пациента от выгрузки первичных данных из медицинской учетной системы до получения прогноза, учитывающий особенности хранения данных в используемой системе и позволяющий врачу оперативно получать информацию о выявленных случаях риска после каждого обновления значений показателей в системе. Показано, что стандартизация значений вводимых в систему результатов медицинских исследований поможет повысить качество прогнозирования за счет повышения устойчивости модели к изменениям входных данных.

Ключевые слова: система поддержки принятия врачебных решений (СППВР), искусственный интеллект (ИИ), машинное обучение, прогнозирование риска, бинарная классификация, градиентный бустинг

Для цитирования: Цыганова М.С., Романов Д.А., Мелентьева А.Г., Макарихин А.В. Прогнозирование ухудшения состояния здоровья пациентов с сердечно-сосудистыми заболеваниями на основе методов машинного обучения // Прикладная информатика. 2024. Т. 19. № 4. С. 68–93. DOI: 10.37791/2687-0649-2024-19-4-68-93

Predicting the deterioration of the condition of patients with cardiovascular diseases based on machine learning methods

M. Tsyganova^{1*}, D. Romanov¹, A. Melentyeva¹, A. Makarikhin²

¹University of Tyumen, Tyumen, Russia

²Labrix LLC, Tyumen, Russia

*m.s.cyganova@utmn.ru

Abstract. This study was carried out as part of a project to develop a subsystem for predicting the deterioration of the condition of patients with cardiovascular diseases on the platform of the medical information system "1C: Medicine. Hospital". The relevance of this task is due to the particularly high danger of this group of diseases and the necessity to make timely decisions about hospitalization or treatment when there is a risk of deterioration of the patient's condition. The goal of this work was to create a tool that allows the attending physician to quickly obtain a reasonable assessment of the risk of deterioration of the patient's condition based on available medical indicators. As a part of this study, an analysis of more than 30 thousand records containing patient health indicators downloaded from the regional medical information system was performed. The data set was labeled in accordance with the available information about medical decisions made (by attending physicians at the clinic and hospital). The lack of a standardized input of health indicators into the medical system required a significant amount of work to pre-process the input data and prepare it for modeling purposes. The prepared data was used to build a predictive model applying machine learning methods. Based on the results of the computational experiments, gradient boosting was chosen as the learning algorithm; the optimal parameters of this algorithm were selected. The prediction quality of the trained models was tested on data from the labeled set that did not participate in the training process. The quality indicators of the best model on test data were precision = 0.87; recall = 0.96; AUC-ROC = 0.97. The integration of trained models with the attending physician's automated workstation in the 1C: Medicine. Hospital system was implemented. Thus, an algorithm for processing patient health indicators from downloading primary data from the medical accounting system to obtaining a forecast was developed, taking into account the peculiarities of data storage in the system and allowing the doctor to quickly receive information about identified risk cases after each update of indicator values in the system. It was shown that standardizing the values of medical research results entered into the system will help to improve the quality of forecasting by increasing the model's stability to changes in input data.

Keywords: clinical decision support system (CDSS), artificial intelligence (AI), machine learning, risk prediction, binary classification, gradient boosting

For citation: Tsyganova M., Romanov D., Melentyeva A., Makarikhin A. Predicting the deterioration of the condition of patients with cardiovascular diseases based on machine learning methods. *Prikladnaya informatika*=Journal of Applied Informatics, 2024, vol.19, no.4, pp.68-93 (in Russian). DOI: 10.37791/2687-0649-2024-19-4-68-93