

Разработка модели поддержки принятия врачебных решений для прогнозирования риска развития сердечной недостаточности

Иванёнок Н.Д.^{1,2}

Научный руководитель

Смирнова Д.А.¹

¹ФГАОУ ВО БФУ им. И. Канта, Калининград, Россия

²МАОУ СОШ г. Зеленоградска, Зеленоградск, Россия

Email: nadezhdaivanenok@yandex.ru

Работа представляет собой практико-ориентированный проект, направленный на разработку системы поддержки принятия врачебных решений для прогнозирования риска развития сердечной недостаточности на основе клинических показателей пациента. Актуальность темы обусловлена высокой распространенностью сердечно-сосудистых заболеваний и их ведущим вкладом в смертность, необходимостью раннего выявления сердечной недостаточности для снижения летальности и повышения качества жизни пациентов, а также возможностями применения методов статистики и машинного обучения к клиническим данным для персонализированной диагностики. Цель работы заключалась в выявлении наиболее значимых клинических факторов и создании прогностической модели раннего риска сердечной недостаточности. Для достижения цели решались задачи: анализ и предварительная обработка медицинского набора данных; статистическая оценка влияния признаков; обучение и сравнение алгоритмов машинного обучения; выбор наилучшей модели и оценка её диагностических характеристик.

Исследование выполнено на основе базы данных из 918 пациентов, включающей 11 клинических признаков: возраст, пол, тип боли в груди, артериальное давление в покое, уровень холестерина и глюкозы натощак, показатели ЭКГ в покое, максимальная частота сердечных сокращений, наличие стенокардии при нагрузке, депрессия сегмента ST и наклон сегмента ST. На первом этапе проведены проверка распределений и статистическая оценка различий между группами с использованием непараметрического критерия Манна-Уитни, что позволило установить значимое влияние ряда признаков на целевую переменную. На втором этапе решалась задача бинарной классификации: выполнено обучение и сравнительная оценка моделей (логистическая и гребневая регрессия, k-ближайших соседей, дерево решений, случайный лес, градиентный бустинг) с подбором гиперпараметров методом поиска по сетке и кросс-валидацией.

Главным результатом является разработка прогностической модели на основе метода k-ближайших соседей, достигшей точности 86,92% ($\pm 2,34\%$), чувствительности 0,83 и специфичности 0,86, что свидетельствует о высокой способности корректно выявлять как пациентов с сердечной недостаточностью, так и без неё. В ходе работы показано, что комбинация показателей сердечного ритма и характеристик сегмента ST обладает наибольшей прогностической ценностью, а применение методов машинного обучения к стандартным клиническим данным позволяет эффективно прогнозировать риск сердечной недостаточности на ранних стадиях.

Возьмите на заметку:

Для прогнозирования сердечной недостаточности достаточно учитывать всего несколько ключевых показателей: наклон и депрессию сегмента ST, холестерин, и максимальную ЧСС. Эти признаки обеспечивают основу надежной и простой модели ранней диагностики, способной выявлять пациентов с высоким риском до появления явных симптомов.