

МЕСТО СРЕДСТВ DATA MINING В СИСТЕМЕ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЯ В ОБЛАСТИ МЕДИКО-ПСИХОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ

Шагарова М.Д., Мокина Е.Е.

Томский политехнический университет
mds1@tpu.ru

Введение

Одной из методик исследования качества жизни пациентов является применение специальных опросников, содержащих варианты ответов на стандартные вопросы и составленных для подсчета значений по определенным показателям. Такие опросники помогают оценить качество жизни. Методика позволяет зарегистрировать и количественно оценить изменения качества жизни у больных с определенным видом заболевания на протяжении конкретного периода стационарного лечения, а также выделить составляющие, которые вносят наиболее весомый вклад в обусловленные лечением изменения качества жизни.

Необходимо решение, предназначенное для анализа показателей в динамике, выявления влияния конкретных вопросов на конечный результат по данным результатам проведенного опроса, определения риска появления заболевания у пациента и рекомендаций по улучшению показателей; также решение, позволяющее сократить время на поиск и выбор совокупности данных по одному или нескольким параметрам, для дальнейшего исследования.

Проектированию системы поддержки принятия решений

Архивы данных, накопленные по какой-либо сфере человеческой деятельности, содержат огромный запас сведений, явлений, факторов, относящихся к данной области.

Применительно к медицинской сфере, обработанные накопленные данные о состоянии здоровья пациента, представляют собой важный материал для изучения различных случаев каждого конкретного заболевания.

Извлечение закономерностей из массива данных – одна из задач многих исследований медицинской тематики.[1].

Выявленные зависимости и закономерности в накопленных данных играют не последнюю роль при принятии решения по диагностированию, определения возникновения риска заболевания у пациента, улучшению методики лечения заболевания пациента. Поэтому важным шагом является выбор способа обработки большого объема информации. По причине наличия в массиве данных «скрытых», неочевидных закономерностей, целесообразным является применение технологии Data Mining.

Технологию Data Mining достаточно точно определяет один из основателей этого направления Григорий Пиатецкий-Шапиро: «Data Mining – это процесс обнаружения в сырых

данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности». Суть и цель, характеризующую технологию Data Mining можно представить как поиск в больших объемах данных неочевидных, объективных и полезных на практике закономерностей [1].

Выявленные закономерности представляют собой особый интерес для определения риска появления заболевания у пациента при изучении новых данных. В нашем случае, по показателям состояния здоровья, полученных при проведении тестирования по опроснику здоровья.

Данный процесс должен быть автоматизирован. Информационная система должна поддерживать процесс прохождения тестирования пользователем и функцию поддержки принятия решений для исследователя, по полученным в ходе тестирования значениям показателей опросников здоровья.

Одной из важных задач является процесс организации правил выдачи рекомендаций по новым введенным данным о пациенте, формирование решений, как по суммарным показателям, так и по каждому вопросу в отдельности. То есть, в систему должны быть заложены правила вывода рекомендаций.

Другой задачей является сохранение полученных результатов в системе для организации анализа показателей в динамике, определения влияния конкретного вопроса на конечный результат.

В соответствии с выявленными задачами и требованиями при проведении медико-психологических исследованиях необходимо:

- Разработать алгоритм для выдачи рекомендаций и риска появления заболевания у пациента, определения динамики и характеристики изменения показателей по результатам проведенного опроса;
- автоматизировать процесс обработки полученных данных, введенных пользователем (опрашиваемым) и предоставлять результат после каждого пройденного опроса;
- обеспечить хранение полученных результатов по пройденному тесту в систематизированном виде;
- обеспечить доступ (исследователю) к хранящимся результатам для выборки данных по необходимым параметрам.

Произведен поиск и анализ программных средств, реализующих выявленные требования. Например, доступен программный продукт SF36

по опроснику «Краткая форма здоровья – MOS 36» (Medical Outcomes Study-Short Form), сервис для проведения психологических тестирований по опросникам изучения качества жизни в соответствии со шкалами тревоги и депрессии [2]. Данные программные решения предназначены, главным образом, только для подсчета показателей и представляют собой числовые значения, требующие дальнейшей интерпретации и исследования. Таким образом, данные программные продукты не выполняют все необходимые функции.

Актуальность исследования обусловлена отсутствием универсальной системы проведения медицинских исследований по данным медицинским показателям пациента (по данным результата опросников о качестве жизни, связанных со здоровьем).

Анализ показал, что рассмотренные программные решения не полностью соответствуют предъявляемым требованиям. Поэтому необходим программный продукт, который в полной мере соответствовал бы требованиям, предъявляемым к функционалу.

Важным компонентом системы поддержки принятия решения являются те знания, которые она в себе содержит. Знания могут быть представлены в виде базы знаний, наборов правил. Выходными данными системы поддержки принятия решения являются результаты применения алгоритмов логического вывода к формализованным данным.

База знаний может быть представлена набором правил, на основе которых алгоритм логического вывода определяет выходные данные. Такие правила позволяют представить знания в виде предложений типа *Если (условие), то (действие)*.

Таким образом, можно получить, например, степень риска появления заболевания у пациента в процентном или лингвистическом отношении.

В соответствии с тем фактом, что в наличии имеется база накопленных данных о состоянии здоровья пациентов и результаты проведенных тестирований по опросникам здоровья, подходом к разработке системы поддержки принятия решения выбран подход, основанный на правилах логического вывода. После обработки накопленного массива данных можно получить выявленные зависимости и на основе этих зависимостей построить правила вывода.

Для проектирования компонента системы поддержки принятия решения, для реализации задач, выявленных в ходе исследования, нужно иметь следующую информацию:

1. Накопленные данные (диагноз, показатели качества жизни) о пациентах;
2. Выявленные скрытые закономерности в имеющихся данных по технологии Data Mining, обуславливающие риск появления заболевания;

3. Построенные правила логического вывода по результатам применения средств Data Mining;
4. Построенные правила выдачи сообщений о риске появления заболевания и рекомендаций по улучшению показателей здоровья;
5. Разработанные логические правила вывода результатов прохождения опросов на языке программирования.

Применяя данный процесс к новым данным о состоянии здоровья пациента на основе опросников здоровья, введенных в информационную систему можно получить:

- сообщения о риске возникновения заболевания;
- динамику изменения показателей, рекомендации по улучшению показателей состояния здоровья.

С другой стороны, сохраненные данные в системе можно использовать в других исследованиях.

В соответствии с рассмотренным подходом можно сделать вывод о том, что главным процессом является установление зависимостей и написание правил логического вывода, применяемых в программном решении, так как от них зависит адекватность и ценность вывода рекомендаций, результаты работы системы.

Заключение

Предложен подход к проектированию модуля системы поддержки принятия решений с целью прогнозирования состояния здоровья пациента, определения риска возникновения заболевания, выдачи рекомендаций, анализа показателей в динамике.

Практическая значимость исследования определяется тем, что выполненные разработки в виде конечного приложения применимы для широкого круга медицинских работников, исследующих качество жизни больных.

Список использованных источников

1. Марухина О.В., Мокина Е.Е., Берестнева Е.В. Применение методов Data Mining для выявления скрытых закономерностей в задачах анализа медицинских данных // *Фундаментальные исследования*. – 2015. – № 4. С. 107-113.
2. Калькулятор SF-36. Электронный ресурс. – [Режим доступа]: <https://sites.google.com/site/71microsurgery/home/sf36>, <http://atio-irk.ru/oprosnik-sf-36.html>
3. Мещеряков Р.В., Конев А.А. К вопросу об исследовании биологических параметров человека в защищенных системах // *Доклады Томского государственного университета систем управления и радиоэлектроники*. 2010. Т. 1. № 1. С. 131-136.