

Методы и алгоритмы интеллектуального анализа данных в решении медицинских задач

*Сенчилов Владислав Владимирович
Смоленский государственный университет
к.ф.-м.н., доцент кафедры информатики*

*Григорьева Галина Михайловна
Смоленский государственный университет
студент*

*Ходченков Валерий Юрьевич
Смоленский государственный университет
студент*

Аннотация

В статье рассмотрен выбор методов и алгоритмов анализа данных в медицинской практике. Приведена структура и возможный метод обработки данных из системы электронного ведения историй болезни. Предложен вариант СУБД для хранения данных, и дальнейшей реализации системы компьютерной поддержки принятия решений в вопросах врачебной практики.

Ключевые слова: анализ данных в медицине, большие данные, интеллектуальный анализ данных.

Methods and algorithms for data mining in solving medical problems

*Senchilov Vladislav Vladimirovich
Smolensk State University
candidate of physico-mathematical sciences, associate professor of Informatics*

*Grigoryeva Galina Mihaylovna
Smolensk State University
student*

*Khodchenkov Valeriy Yur'yevich
Smolensk State University
student*

Abstract

This article discusses the choice of methods and algorithms for data analysis in medical practice. The structure and the possible method of data processing from the electronic management of case histories are given. A version of a DBMS for

data storage and the further implementation of a computer-assisted decision-making support system for medical practice have been proposed.

Keywords: data analysis in medicine, big data, data mining.

Главная проблема инфекционных заболеваний – наличие осложнений, которые могут приводить к летальным исходам. Поэтому для благоприятного исхода лечения возникает необходимость дополнительного контроля пациентов, находящихся в группе риска по формированию осложнений. Это требует поиска способов распознавания проблем со здоровьем на ранних этапах лечения и в амбулаторных условиях.

Таким образом, выбор методов и алгоритмов интеллектуального анализа данных, которые способны решать конкретные медицинские задачи, и разработки на их основе специализированных систем поддержки принятия решений – актуальная задача, а развитие медицинской кибернетики и современной вычислительной техники открывает для её решения новые возможности.

В статье рассматривается задача выбора методов и алгоритмов ИАД для решения конкретных медицинских задач на примере определения пациентов в группу риска по формированию осложнений при заболеваниях корью. Данная проблема имеет решение в рамках поставленной задачи при условии, что извлечение и структурирование информации из истории болезни уже проведено.

Также, в качестве гипотезы можно предположить, что подобная система анализа данных в дальнейшем может быть использована и для других инфекционных вирусных заболеваний.

Данные представляют собой три связанных таблицы, хранящихся в общей базе данных. Распределение данных по таблицам (рис.1) соответствует распределению данных по классификационным разделам. Связь между таблицами осуществляется через идентификатор пациента (ID), что позволяет обрабатывать данные отдельно, а затем выбрать результат по каждому пациенту в одну таблицу. Распределение данных по таблицам происходит в соответствии с логическими разделами.

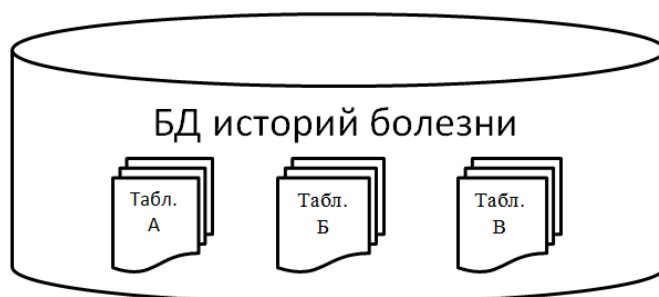


Рисунок 1 – Предполагаемая структура базы данных

Раздел первый «Temp list» - раздел, в котором содержится информация о колебаниях температуры пациента, находящегося на стационарном лечении. Структура раздела:

1. Идентификатор пациента – строковый
2. Дата и время – числовой
3. Показатель температуры – числовой
4. Приём жаропонижающих – логический
5. Наличие гипертермии – логический

Раздел второй «Lab information» - раздел, содержащий данные клинической лабораторной диагностики. Структура раздела:

1. Идентификатор пациента – строковый
2. Анализ «Аланинаминотрансфераза (АЛТ)» – строковый
3. Анализ «Аспартатаминотрансфераза (АСТ)» – строковый
4. Анализ «Скорость оседания эритроцитов (СОЭ)» – строковый
5. Анализ «Креатинин» – строковый
6. Анализ «Мочевина» – строковый
7. Анализ «Лейкоциты» – строковый

Раздел третий «Main» - данный раздел включает в себя сразу несколько разделов электронной истории болезни: Anamnesis Vitae, Anamnesis Morbi - описание течения заболевания со слов пациента, данные осмотра лечащим врачом и динамика клинических данных в процессе лечения - раздел, в котором находятся данные о протекании заболевания у пациента, при нахождении в стационаре

1. Идентификатор пациента – строковый
2. Возраст – числовой
3. Наличие вакцинации – логический
4. Перенесённые заболевания – строковый
5. Период начала заболевания – числовой
6. Гипертермия – логический
7. Рвота (при гипертермии) – логический
8. Головная боль (при гипертермии) – логический
9. Одышка – логический
10. Кашель – логический
11. Мокрота – логический
12. Менингеальные знаки – логический
13. Выраженная гиперемия зева – логический
14. Гнойный налёт – логический
15. Конъюнктивит – логический
16. Гнойное отделяемое (при конъюнктивите) - логический
17. Прибывает в стационаре – логический
18. Дата поступления в стационар – числовой
19. Интоксикационный синдром – логический
20. Экзантемный синдром – логический
21. Предполагаемое начало экзантемного синдрома – числовой
22. Катарально-респираторный синдром – логический

Данный набор основан на медицинской карте стационарного больного по учётной форме №003/у и медицинской карте пациента, получающего медицинскую помощь в амбулаторных условиях, по учётной форме №025/у.

В разрабатываемой базе данных в качестве полей используется набор диагностических факторов, который в дальнейшем будут параметрически задавать основу для статического анализа.

В дальнейшем планируется разработка СКИАМД, которая представляет собой графическую оболочку и нейросеть.

Предполагаемая структура нейронной сети (рис.2) включает учёт возможных смещений, которые необходимо производить перед выводом данных.

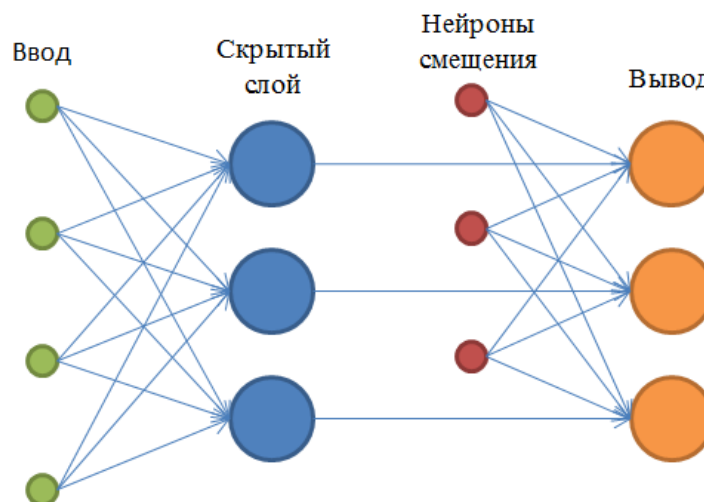


Рисунок 2 – Предполагаемая структура нейросети

Для хранения исходных данных и передачи их в программу предложена реляционная система управления базами данных PostgreSQL, так как эта СУБД очень хорошо приспособлена для связи с программами на различных языках.

Предлагаемый метод открывает возможность для конструктивного построения системы компьютерного интеллектуального анализа медицинских данных. Подобная система может быть полезна для лечащего врача при работе с пациентами, попадающими в группу риска по развитию осложнений вирусных инфекций на примере кори.

Библиографический список

1. Андрейчиков А.В. Интеллектуальные информационные системы. М.: Финансы и статистика, 2006. 424 с.
2. Анохин П.К. Узловые вопросы теории функциональной системы. М.: Наука, 1980. 197с.
3. Кузнецова А. В., Сенько О. В. Возможности использования методов Data Mining при медико-лабораторных исследованиях для выявления закономерностей в массивах данных // Врач и информационные технологии. 2005. №2.