УДК 004.8

Т.А. ЛЫСАКОВА

T.A. LYSAKOVA

**АНАЛИЗ МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, ПРИМЕНИМЫХ ДЛЯ ОБРАБОТКИ СЛАБОСТРУКТУРИРОВАННЫХ ДАННЫХ В МЕДИЦИНЕ**

**ANALYSIS OF ARTIFICIAL INTELLIGENCE METHODS APPLIED FOR PROCESSING SEMI-STRUCTURED DATA IN MEDICINE**

*В данной статье автор освещает проблему использования методов искусственного интеллекта в области здравоохранения. Рассматриваются достоинства и недостатки таких методов искусственного интеллекта, как знание-ориентированный подход, деревья решений, нейронные сети и генетические алгоритмы.*

*Ключевые слова: искусственный интеллект, нейронные сети, экспертные системы, деревья решений, генетические алгоритмы.*

*In given article authors shine a problem of using artificial intelligence methods in the field of healthcare. The advantages and disadvantages of such artificial intelligence methods as knowledge-based approach, decision trees, neural networks and genetic algorithms are considered.*

*Keywords: artificial intelligence, neural networks, expert systems, decision trees, genetic algorithms*

В развитии цифрового здравоохранения одним из наиболее перспективных направлений является искусственный интеллект. В нашей стране постоянно появляются все новые и новые разработки и исследования, предлагающие реализацию этих технологий для медицины и здравоохранения.

Рассмотрим некоторые методы искусственного интеллекта, применимые для обработки слабоструктурированных данных в медицине, такие как знание-ориентированный подход, деревья решений, нейронные сети и генетические алгоритмы.

Знание-ориентированный подход. Данный подход может быть использован для решения неформализованных проблем, к которым относятся задачи, обладающие следующими характеристиками:

- сложность или невозможность представления задач и данных в числовой форме;

- неточность и неоднозначность исходных данных и знаний о предметной области;

- отсутствие конкретного алгоритма для решения определенной медицинской задачи.

Задачи здравоохранения можно охарактеризовать всеми вышеперечисленными свойствами, так как в большинстве своем они содержат большой объем многомерных, запутанных и противоречивых клинических данных. Основная проблема при создании знание-ориентированных систем – построение базы знаний. Главная трудность при создании баз знаний заключается в практической сложности получения знаний от врачей-экспертов, а также в том, чтобы заложить в систему «здравый смысл», которым зачастую руководствуется эксперт. Кроме того, в отличие от экспертов, такие системы сложно адаптируются к изменению среды. Так же минусом является невозможность восприятия сенсорной информации.

Если таких ограничений будет слишком много, это сильно затруднит создание знание-ориентированной системы. В этом случае знание-ориентированную систему можно рассматривать не как точный инструмент, а как систему поддержки принятия решений. Работа с ней происходит следующим образом: система задает вопросы пользователю, по мере получения ответов механизм логического вывода движется по базе знаний, определяет, какие ещё вопросы нужно задать и в конце делает вывод. Причём, знание-ориентированная система может объяснить всю цепочку рассуждений, что также является преимуществом при медицинской диагностике.

Деревья решений. Использование этого метода предполагает занесение в программу последовательно задаваемых врачом вопросов, необходимых для решения определенной диагностической проблемы. Эти вопросы структурируются в виде дерева, в котором вершинами являются вопросы, задаваемые больному, а ветками – альтернативные ответы на вопросы, которые, в свою очередь, ведут к следующим вопросам. Программа осуществляет переход от вопроса к вопросу до тех пор, пока не будет найдено решение или исчерпаны возможные переходы. Проблемы в использовании деревьев решений начинаются при усложнении решаемых диагностических задач. Деревья решений для таких задач становятся слишком сложными и имеют высокую степень ветвления. Всё это приводит к проблемам в анализе работы дерева и интерпретации его результатов. Кроме того, при необходимости внесения изменений в логику программы – приходится заново строить дерево и полностью перепрограммировать всю задачу.

В то же время такой подход крайне удобен, так как позволяет представить в программе логику составления последовательности вопросов врачом при решении диагностической задачи в клинических условиях. Этот подход позволяет имитировать процесс принятия решения врачом при постановке диагноза.

Нейронные сети. Нейронные сети хорошо подходят для решения трудноформализуемых задач, которые содержат большой объем разнородных данных, не имеющих математического описания, чем в большинстве случаев характеризуются медицинские данные.

Нейронные сети способны выявлять скрытые закономерности в многомерных данных, а так же открывают возможность уменьшения размерности входных данных, оставляя наиболее существенные. Кроме того, нейронные сети используют обучение на примерах и не требуют программирования правил вывода, что упрощает задачу.

В то же время, существуют определенные трудности, сопряженные с применением нейросетевых технологий в задачах здравоохранения. Так, в подавляющем большинстве случаев невозможно заранее определить архитектуру нейронной сети и её сложность для конкретно-поставленной задачи. Также для качественной работы нейронной сети необходимо провести работу по тщательному формированию релевантной обучающей выборки. Помимо этого, главным недостатком использования нейронных сетей в здравоохранении является отсутствие объяснения получения того или иного результата – нельзя проследить путь постановки диагноза.

Генетический алгоритм. Генетические алгоритмы представляют собой итерационные вероятностные эвристические алгоритмы поиска.

Основными преимуществами генетических алгоритмов является возможность использования данных различного вида и значительное упрощение трудоемких и долговременные задач. Кроме того, генетические алгоритмы саморазвиваются и не требуют постоянного перепрограммирования. Также генетические алгоритмы определяют не одно конкретное решение, а группу допустимых вариантов, что является полезным качеством в задачах диагностики.

Но, наряду с достоинствами, генетические алгоритмы обладают ограничениями в настройках генетических операторов. И кодирование решения представляет собой трудоемкий процесс. Ещё одной проблемой при использовании генетических алгоритмов становится отсутствие эффективных критериев для окончания работы алгоритма. Кроме того, генетические алгоритмы требуют достаточно больших вычислительных ресурсов.

Байесовский подход. В основе байесовского подхода лежит теорема Байеса, которая относится к теории вероятностей. Байесовский подход в задачах здравоохранения используется при вычислении частоты или вероятности наступления какого-либо события, например, заболевания.

При применении байесовского подхода возникает проблема определения всех взаимодействий в сложных системах, а так же необходимость использования экспертных оценок для определения множества условных вероятностей.

Но, в то же время, байесовский подход позволяет использовать только априорную информацию, что является несомненным преимуществом. Кроме того, он достаточно прост в понимании логически выведенных утверждений.

Все достоинства и недостатки вышерассмотренных методов приведены в таблице 1.

Таблица 1 – Сравнительная характеристика методов искусственного интеллекта

|  |  |  |
| --- | --- | --- |
| Название метода | Достоинства | Недостатки |
| Знание-ориентированный подход | - возможность объяснения всей цепочки рассуждений;  - могут объединять знания нескольких экспертов;  - возможность документации процесса решения | - практическая сложность получения знаний от врачей-экспертов;  - практическая сложность в том, чтобы заложить в систему «здравый смысл» эксперта;  - проблема адаптации экспертных систем к изменению среды;  - отсутствие восприятия сенсорной информации |
| Деревья решений | - не требуют подготовки данных;  - простота интерпретации полученных решений;  - способен работать с разными категориями данных | - проблема анализа больших деревьев для сложных диагностических задач;  - трудоемкость и сложность внесения изменений в логику программы |
| Нейронные сети | - выявление скрытых закономерностей в многомерных данных;  - возможность уменьшения размерности входных данных, оставляя наиболее существенные;  - используют обучение на примерах и не требуют программирования правил вывода | - невозможность заранее определить архитектуру нейронной сети и её сложность для конкретно-поставленной задачи;  - необходимость тщательного формирования обучающей выборки, которая определяет качество работы нейронной сети;  - отсутствие объяснения получения того или иного результата – нельзя проследить путь постановки диагноза |
| Генетические алгоритмы | - возможность использования данных различного вида;  - значительно упрощают трудоемкие и долговременные задачи;  - генетические алгоритмы саморазвиваются;  - генетические алгоритмы определяют не одно конкретное решение, а группу допустимых вариантов | - ограничения настройками генетических операторов;  - трудоемкость в кодировании решения;  - отсутствие эффективных критериев для окончания работы алгоритма;  - требуют достаточно больших вычислительных ресурсов |

В настоящее время для более массового применения искусственного интеллекта в области здравоохранения существует два главных препятствия. Первое – необходим большой объем данных и, второе, профессиональный подход к построению систем искусственного интеллекта. Главную и серьёзную сложность при внедрении систем искусственного интеллекта в медицине представляют собой качественные, выверенные и структурированные данные, без которых искусственный интеллект не будет корректно работать. Так же без профессионального подхода к построению систем искусственного интеллекта для каждой конкретной медицинской задачи простое применение готовых алгоритмов к подготовленным данным также не будет давать корректный результат, т.к. искусственный интеллект требует определенной настройки для понимания и правильной интерпретации данных для решения конкретных прикладных задач.

Таким образом, все вышерассмотренные методы искусственного интеллекта имеют как достоинства, так и свои недостатки, применительно к обработке слабоструктурированных данных в медицине. Поэтому целесообразность применения того или иного метода будет зависеть от содержания конкретной прикладной задачи.

**СПИСОК ЛИТЕРАТУРЫ**

1. Жариков О.Г. Современные возможности использования некоторых экспертных систем в медицине/ О.Г.Жариков, В.А.Ковалев, А.А.Литвин// Врач и информационные технологии. – 2008, №5. – с. 24-30.

2. Головин П.А. Экспертные системы для классификации болезней в медицинской диагностике/ П.А.Головин, В.А.Нечаев, Д.А.Нечаев// Научно-технический вестник информационных технологий, механики и оптики. – 2006, №29. – с. 80-84.

3. Гусев А.В. Добриднюк С.Л. Искусственный интеллект в медицине и здравоохранении/ А.В.Гусев, С.Л.Добриднюк// Информационное общество. – 2017, №4-5. – с. 78-93.

4. Ле Н.В. Модель представления знаний при создании медицинской экспертной системы дифференциальной диагностики/ Н.В. Ле, В.А.Камаев, Д.П.Панченко, О.А.Трушкина// Известия Волгоградского государственного технического университета. – 2014, №6 (133). – с. 41-50.

5. Ле Н. В. Интеллектуальная медицинская система дифференциальной диагностики на основе экспертных систем// Вестник СГТУ. - 2014. №2 (75). – 167-179.

6. Волчек Ю.А. Положение модели искусственной нейронной сети в медицинских экспертных системах/ Ю.А.Волчек, О.Н.Шишко, О.С.Спиридонова, Т.В.Мохорт// Juvenis scientia. - 2017. №9. – с. 4-9.

7. Звягин Л.С. Метод байесовских сетей и ключевые аспекты байесовского моделирования/ Л.С.Звягин// XXII Международная конференция по мягким вычислениям и измерениям (SCM-2019). Сборник докладов. Санкт-Петербург. 23–25 мая 2019 г. - СПб.: СПбГЭТУ «ЛЭТИ». – с.30-34.

**Лысакова Татьяна Алексеевна**

Белгородский государственный национальный исследовательский университет, г. Белгород

Старший преподаватель кафедры прикладной информатики и информационных технологий

Тел.: +7 (4722) 301300 \* 2027

E-mail: lysakova@bsu.edu.ru