

УДК 004.8

МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ ВЕРОЯТНОСТИ ЗАБОЛЕВАНИЯ COVID-19 ПО ПЕРВИЧНЫМ ПРИЗНАКАМ

Красюкова Ю.И.

магистрант Южный Университет ИУБиП, гр. ИМ-101

e-mail: krasukova84@mail.ru

Вахрушева Т.А.

магистрант Южный Университет ИУБиП, гр. ИМ-101

e-mail: vakhrusheva.1997@mail.ru

Аннотация: В данной статье рассматривается создание алгоритма машинного обучения для возможности постановки диагноза COVID-19 на основании первичного опроса пациента. Пандемия является одной из самых тревожных и актуальных событий современного общества. А выявление признаков заболевания на ранних стадиях – важная составляющая предотвращения эпидемии.

Ключевые слова: машинное обучение, нейронные сети, короновирс, бинарная система счисления, двоичный код, бинарное кодирование, COVID-19, искусственный интеллект, интеллектуальные информационные системы

MACHINE LEARNING MODEL FOR DETERMINING THE PROBABILITY OF COVID-19 DISEASE BY PRIMARY SIGNS

Krasyukova Yu. I.

Vachrusheva T.A.

Abstract: This article discusses the creation of a machine learning algorithm for the possibility of making a COVID-19 diagnosis based on an initial patient survey. The pandemic is one of the most disturbing and urgent events in modern society. And the detection of signs of the disease in the early stages is an important component of preventing the epidemic.

Keywords: machine learning, neural networks, coronavirus, binary number system, binary code, binary coding, COVID-19, artificial intelligence, intelligent information systems

Распространение COVID-19, его быстрая мутация и схожая с другими заболеваниями симптоматика – пожалуй, одна из самых острых тем 2020-2021 гг. Выявить симптомы на ранней стадии – главная задача для скорейшего выздоровления пациента и, что не маловажно, для предотвращения распространения вируса среди контактирующих с заболевшим.

Изобретения с использованием искусственного интеллекта - новое и стремительно развивающееся направление [1-4]. Машинное обучение на основе нейронных сетей во многом облегчает задачу выявления заболевания [5].

Уже известно о таких проектах как КТ-калькулятор и распознавание наличия вируса по кашлю через приложение в смартфоне. Несмотря на то, что диагноз, получаемый таким способом, носит исключительно информационный характер, эти технологии оказывают существенную помощь специалистам, подтверждая достоверность диагноза более чем в 85% случаев [6,7].

В данной статье предложен один из вариантов построения модели по определению вероятности заболевания COVID-19 на основе первичного опроса пациента.

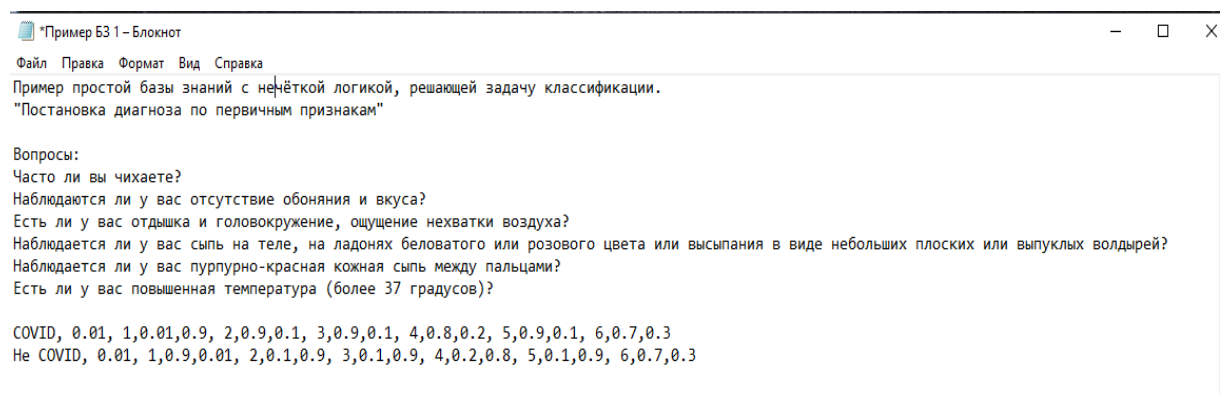
Строится модель на базе программы MiniExpertSystem, которая представляет собой простую экспертную систему, использующую байесовскую систему логического вывода и предназначенную для проведения консультации с целью определения вероятностей возможных исходов [1,8].

На первом этапе необходимо создать базу знаний, которая включает в себя набор вопросов (симптомов заболевания) и набор вероятных решений.

Нам известно, что одни и те же симптомы могут быть характерны для разных заболеваний и степень вероятности их возникновения может быть различной. Например, чихание характерно при простуде, гриппе, аллергии, но не характерно для коронавируса. Сухой лающий кашель является одним из основных признаков коронавируса, но также он может возникать при гриппе или аллергии.

Исходя из этого, применяется система построения решений на основе нечеткой логики. А именно ответы на вопросы ранжируются по шкале от 5 до -5, где 5 – однозначное «Да», а -5 – однозначное «Нет».

Для примера был составлен список вопросов и вариантов вероятностей исходов (рис 1).



```
*Пример БЗ 1 – Блокнот
Файл Правка Формат Вид Справка
Пример простой базы знаний с нечёткой логикой, решающей задачу классификации.
"Постановка диагноза по первичным признакам"

Вопросы:
Часто ли вы чихаете?
Наблюдаются ли у вас отсутствие обоняния и вкуса?
Есть ли у вас одышка и головокружение, ощущение нехватки воздуха?
Наблюдается ли у вас сыпь на теле, на ладонях беловатого или розового цвета или высыпания в виде небольших плоских или выпуклых волдырей?
Наблюдается ли у вас пурпурно-красная кожная сыпь между пальцами?
Есть ли у вас повышенная температура (более 37 градусов)?

COVID, 0.01, 1,0.01,0.9, 2,0.9,0.1, 3,0.9,0.1, 4,0.8,0.2, 5,0.9,0.1, 6,0.7,0.3
Не COVID, 0.01, 1,0.9,0.01, 2,0.1,0.9, 3,0.1,0.9, 4,0.2,0.8, 5,0.1,0.9, 6,0.7,0.3
```

Рис.1 – Листинг программы

Первое число в атрибутах базы фактов экспертной системы означает [9] априорную вероятность того, что любой наугад выбранный пациент может оказаться больным этой болезнью. Далее следует порядковый номер вопроса. Следующие два числа означают вероятность ответа «Да» и «Нет», что и будет определять вероятную степень заболевания. Например, при ответе на второй вопрос, больной коронавирусной инфекцией скорее всего поставит 4 или 5, что будет близко к значению 0.9, а значит увеличит вероятность подтверждения диагноза. И, напротив, при оценке со знаком «минус» система распознает это как отрицательный ответ и снизит

вероятность заболеваемости данным вирусом при выставлении диагноза[10,11].

Выводы

Данная модель может выступать как дополнение к уже существующей медицинской базе знаний или использоваться как отдельный инструмент опроса пациентов и выставления первичного диагноза. Она проста в реализации, и, при качественном составлении базы знаний (вопросов и вероятностей), может помочь не только выявить вероятность заболевания коронавирусной инфекцией, но и определить к какому штамму оно относится. Все это может значительно облегчить работу медиков, ведущих первичный прием пациентов, а для выезжающих на вызовы врачей-терапевтов подобное приложение может стать основой для принятия решения о дальнейшей тактике лечения.

Библиографический список

1. Акперов И.Г., Храмов В.В. Управление социально-экономическими системами региона - становление цифровой экономики // Интеллектуальные ресурсы – региональному развитию. – 2020. – № 2. – С. 36-47. – URL: <https://www.elibrary.ru/item.asp?id=43033189> (Дата обращения 09.04.2021)/
2. Попов Э.В., Фоминых И.Б., Кисель Е.Б. Статические и динамические экспертные системы (классификация, состояние, тенденции): методические материалы. – М.: Центральный рос. дом знаний, 1995. – 126 с.
3. Фрэнк Дж. Бартос. Искусственный интеллект: принятие решений в сложных системах управления.// Мир компьютерной автоматизации. – 1997. – № 4. – С. 2-27.
4. Нейлор К. Как построить свою экспертную систему / Пер. англ. – М.: Энергоатомиздат, 1991. – 286 с.
5. Храмов В.В. Интеллектуальные информационные системы. Часть 1: Представление знаний в информационных системах: учебно-методическое пособие / В.В.Храмов; Рост. гос. ун-т путей сообщения, – Ростов н/Д, 2010. – 108 с.
6. Бернштейн, Л.С. Нечеткие модели принятия решений: дедукция, индукция, аналогия./Л.С. Бернштейн, А.В. Боженюк. – Таганрог: Изд-во ТРТУ, 2001.
7. Акперов И.Г., Храмов В.В. Мягкие модели функциональной связности измерений геоинформационного пространства региона // Интеллектуальные ресурсы – региональному развитию. – 2020. – № 2. – С. 54-59. – URL: <https://www.elibrary.ru/item.asp?id=43033191>(Дата обращения 09.04.2021).
8. Храмов В.В. Агрегирование информации как проблема личностной самоорганизации // Российский психологический журнал. – 2007. – Т. 4, № 4. – С. 9-21. – URL: <https://www.elibrary.ru/item.asp?id=16974213> (Дата обращения 09.04.2021).

9. Храмов В.В., Гвоздев Д.С. Интеллектуальные информационные системы: интеллектуальный анализ данных. учебное пособие. – Ростов-на-Дону, 2012 – 98 с. – URL: <https://www.elibrary.ru/item.asp?id=32762296> (Дата обращения 09.04.2021).

10. Акперов I.G., Khramov V.V., Viktor L., Mityasova O.Yu. Fuzzy methods and algorithms in data mining and formation of digital plan-schemes in earth remote sensing // Procedia Computer Science. – 2017. – С. 120-125. DOI: 10.1016/j.procs.2017.11.218

11. Чернышев Ю.О., Храмов В.В. Особенности агрегирования качественных признаков опорных ориентиров в системах технического зрения // Известия ТРТУ. – 2001. – № 3 (21). – С. 55. – URL: <https://www.elibrary.ru/item.asp?id=12886334>.