

Здравоохранение в информационном обществе

ПРИМЕНЕНИЕ ЛОГИЧЕСКОГО АНАЛИЗА ДАННЫХ В ПРОГРАММНОМ КОМПЛЕКСЕ ДЛЯ ДИАГНОСТИКИ ГАСТРИТОВ

Статья рекомендована к публикации главным редактором Т. В. Ершовой 25.11.2023.

Лютикова Лариса Адольфовна

Кандидат физико-математических наук

Кабардино-Балкарский научный центр РАН, Институт прикладной математики и автоматизации

Нальчик, Российская Федерация

lylarisa@yandex.ru

Аннотация

Предлагается программный комплекс для компьютерной диагностики гастрита на основе доступных показателей и опыта специалистов. Он основан на логическом подходе анализа данных и предназначен для определения наиболее вероятных типов заболеваний с заданной точностью. Входные данные включают показатели обследования пациентов, их диагнозы и опыт специалистов в постановке диагноза, накопленный в медицинской практике. Для представления данных используются системы многозначной логики предикатов. Результат работы программного комплекса заключается в выборе наиболее подходящих типов заболеваний с заранее заданной точностью, основываясь на данных диагностики пациентов. Если точность постановки диагноза не может быть достигнута с заданной точностью, предлагается либо изменить точность решения, либо пройти дополнительное обследование.

Ключевые слова

диагностика, связи, многозначная логика, входные данные, анализ, скрытые закономерности

Введение

Медицинская диагностика в машинном обучении является активной и перспективной областью исследований. Методы машинного обучения могут быть эффективно применены для анализа медицинских данных и поддержки процесса диагностики.

Одним из основных преимуществ машинного обучения в медицинской диагностике является его способность обрабатывать большие объемы данных и выявлять скрытые закономерности, которые могут помочь в прогнозировании заболеваний и принятии решений о диагнозе. Вместо традиционного ручного кодирования правил, алгоритмы машинного обучения могут автоматически извлекать признаки и строить модели, основанные на обучающих данных.

Существует несколько подходов к медицинской диагностике с использованием машинного обучения:

Классификация: Модели машинного обучения могут быть обучены классифицировать пациентов на основе их симптомов, медицинских изображений, результатов тестов и других данных. Например, алгоритмы классификации, такие как метод опорных векторов (SVM), наивный Байесовский классификатор и нейронные сети, могут применяться для диагностики конкретных заболеваний.

Кластеризация: Алгоритмы кластеризации могут группировать пациентов с похожими характеристиками или симптомами. Это может помочь в выявлении подтипов заболеваний и индивидуальном подборе лечения. Кластерный анализ, алгоритмы снижения размерности, такие как метод главных компонент (PCA), и алгоритмы кластеризации, такие как k-средних, могут быть использованы для этой цели.

© Лютикова Л. А., 2024

Производство и хостинг журнала «Информационное общество» осуществляется Институтом развития информационного общества.

Данная статья распространяется на условиях международной лицензии Creative Commons «Атрибуция — Некоммерческое использование — На тех же условиях» Всемирная 4.0 (Creative Commons Attribution – NonCommercial - ShareAlike 4.0 International; CC BY-NC-SA 4.0). См. <https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode.ru>

https://doi.org/10.52605/16059921_2024_04_78

Прогнозирование и регрессия: Машинное обучение может быть применено для прогнозирования хода заболевания, оценки риска развития определенных состояний и определения оптимального лечения. Методы, такие как регрессионный анализ, случайные леса и градиентный бустинг, могут быть использованы для прогнозирования.

Обработка изображений и сигналов: Машинное обучение может быть применено для анализа медицинских изображений (например, рентгеновских снимков, МРТ, УЗИ) и сигналов (например, ЭКГ, ЭЭГ). Глубокое обучение и сверточные нейронные сети демонстрируют впечатляющую эффективность в обработке и анализе медицинских изображений.

Однако, необходимо учитывать, что успешное применение машинного обучения в медицинской диагностике требует качественных и разнообразных данных, а также тщательной валидации и проверки результатов, прежде чем они будут использованы в клинической практике.

Цель работы состоит в разработке метода анализа данных и создании программного комплекса для диагностики гастрита. Предлагаемый метод позволяет находить закономерности и строить классы на основе общих признаков при небольшом объеме данных. Также метод позволяет отбирать наиболее важные свойства для принятия решений.

1 Анализ современных методов медицинской диагностики

В области медицинской диагностики с использованием машинного обучения проведено множество исследований. Вот некоторые примеры областей исследований и их результатов:

Диагностика рака: Машинное обучение применяется для диагностики различных типов рака, включая рак груди, рак легких, рак кожи и др. Исследования показывают, что модели машинного обучения могут достичь высокой точности при классификации образцов раковых и нормальных тканей на основе медицинских изображений, биомаркеров и генетических данных.

Диагностика заболеваний сердца: Машинное обучение применяется для диагностики различных заболеваний сердца, включая ишемическую болезнь сердца, аритмию и сердечную недостаточность. Исследования показывают, что модели машинного обучения могут использоваться для анализа ЭКГ сигналов, изображений сердца и клинических данных для диагностики и прогнозирования заболеваний сердца.

Диагностика неврологических заболеваний: Машинное обучение применяется для диагностики неврологических заболеваний, таких как болезнь Альцгеймера, болезнь Паркинсона и эпилепсия. Исследования показывают, что модели машинного обучения, использующие данные МРТ, ЭЭГ и другие нейровизуальные данные, могут помочь в диагностике и прогнозировании этих заболеваний.

Диагностика инфекций: Машинное обучение применяется для диагностики инфекционных заболеваний, таких как COVID-19, малярия и туберкулез. Исследования показывают, что модели машинного обучения могут анализировать клинические данные, результаты тестов и медицинские изображения для диагностики и классификации инфекций с высокой точностью.

Прогнозирование и риск-стратификация: Машинное обучение может использоваться для прогнозирования хода заболевания, риска развития осложнений и эффективности лечения. Например, модели машинного обучения могут помочь в прогнозировании прогноза пациента с раком, оценке риска сердечно-сосудистых событий и определении оптимального лечения для индивидуального пациента [1-4].

Это лишь некоторые примеры исследований, проведенных в области медицинской диагностики с использованием машинного обучения. С каждым годом появляются новые исследования и методы, которые улучшают эффективность и точность диагностики в медицине.

Машинная диагностика гастрита - это область исследований, где машинное обучение применяется для поддержки диагностики и классификации гастрита. Гастрит может иметь различные причины и проявления.

Исследования в области машинной диагностики гастрита включают следующие аспекты:

- **Анализ симптомов и клинических данных:** Машинное обучение может использоваться для анализа симптомов и клинических данных пациента, таких как боли в желудке, тошнота, рвота, изменения аппетита и другие характеристики. Модели машинного обучения могут обрабатывать эти данные и помогать в классификации гастрита и оценке его тяжести.

-Анализ результатов эндоскопии: Эндоскопия является важным методом для диагностики гастрита. Машинное обучение может применяться для анализа изображений, полученных во время эндоскопии, с целью автоматического обнаружения и классификации признаков гастрита. Это может включать оценку степени воспаления слизистой оболочки, обнаружение язв и других аномалий.

-Использование биомаркеров и лабораторных данных: Машинное обучение может помочь в анализе биомаркеров и лабораторных данных, таких как уровень кислотности желудочного сока, наличие *Helicobacter pylori* (бактерии, связанной с гастритом), и других показателей. Алгоритмы машинного обучения могут помочь в классификации гастрита на основе этих данных и оценке его характеристик.

Интеграция различных источников данных: Машинное обучение может быть использовано для объединения различных типов данных, таких как клинические данные, результаты обследований, генетическая информация и другие, для более точной диагностики гастрита. Это может помочь в выявлении паттернов, связанных с определенными подтипами гастрита или прогнозировании риска осложнений.

В целом, машинное обучение может значительно улучшить диагностику гастрита, обеспечивая более точные и объективные методы анализа симптомов, изображений, лабораторных данных и других факторов. Однако, важно отметить, что машинное обучение не заменяет консультацию и участие опытного врача, а служит вспомогательным инструментом в процессе диагностики и принятия решений в области медицины [5-7].

2. Постановка задачи

Гастрит — распространенное заболевание, от которого страдает до 20% взрослого населения развитых стран и наблюдается высокий уровень заболеваемости среди детей.

Для достоверного определения хронического гастрита необходимо проводить гистологическое исследование биоптатов слизистой оболочки желудка, поскольку это позволяет обнаружить морфологические изменения, в то время как клиническая картина может быть неинформативной.

В 1990 году была принята Сиднейская система классификации гастритов на IX Международном конгрессе гастроэнтерологов. В 1996 году был опубликован окончательный вариант модифицированной Сиднейской системы, называемый "Классификация и градация гастрита. Модифицированная Сиднейская система". Согласно этой классификации, выделяются три основные категории гастритов: острый, хронический и особые формы.

Для создания обучающей выборки мы использовали результаты гистологического исследования гастробиопсий 132 пациентов, проведенного в период с 2019 по 2022 год в Патологоанатомическом бюро, принадлежащем Государственному учреждению здравоохранения.

Эти данные представляют собой ценный исходный материал для нашего исследования. Они позволяют нам анализировать и изучать различные патологические состояния желудка, выявлять связи между различными факторами и патологиями, и разрабатывать модели и алгоритмы диагностики на основе этих данных.

Использование результатов гистологического исследования как основы для обучающей выборки обеспечивает нам возможность создать надежную и информативную модель, которая может быть применена в будущих исследованиях и клинической практике для более точной и эффективной диагностики патологий желудка.

Таким образом, цель исследования состоит в создании модели машинного обучения, которая будет способна классифицировать пациентов и определять тип гастрита на основе имеющихся симптомов и результатов обследований. Это позволит помочь врачам в более точной и эффективной диагностике гастрита, а также позволит распространить полученные результаты на других пациентов с неизвестным диагнозом.

Учитывая, что в задаче относительно небольшой объем данных, было решено использовать логические методы анализа, а не нейросетевые подходы.

Логические методы могут быть эффективными при интерпретации и понимании причинно-следственных связей между симптомами и диагнозом гастрита. Это может дать возможность моделировать вероятность принадлежности пациента к определенному типу гастрита на основе

комбинации симптомов и обследований. Модель может быть обучена на имеющихся данных, чтобы выявить значимые признаки и создать алгоритм диагностики.

Выбор логических методов для данной задачи может обеспечить понятность и интерпретируемость модели, что особенно важно в медицинской области. Однако следует отметить, что выбор метода зависит от специфики данных и требований проекта, и иногда комбинация различных методов может привести к лучшим результатам.

На рисунке 1 показан фрагмент анкеты, которую должен заполнить врач для получения диагноза.

Рис. 1. Фрагмент анкеты вводимых данных

Математическая формулировка следующая: функция $Y = f(x_1, x_2, \dots, x_{28})$ из 28 переменных, что определено в 132 точках, область определения каждой переменной имеет разброс от 2 до 4 вариантов. Необходимо восстановить значение функции в других запрошенных точках.

Постановка этой проблемы сводится к постановке проблемы на основе прецедентов.

Пусть $X = \{x_1, x_2, \dots, x_{28}\}$, $x_i \in \{0,1,2,3\}$ – комплекс симптомов, диагностируемых заболеваний. $Y = \{y_1, y_2, \dots, y_{132}\}$ – множество диагнозов, каждый диагноз характеризуется соответствующим набором симптомов $x_1(y_i), x_2(y_i), \dots, x_{28}(y_i): y_i = f(x_1(y_i), x_2(y_i), \dots, x_{28}(y_i))$. Это может быть представлено в следующем виде:

$$\begin{pmatrix} x_1(y_1) & x_2(y_1) & \dots & x_n(y_1) \\ x_1(y_2) & x_2(y_2) & \dots & x_n(y_2) \\ \dots & \dots & \dots & \dots \\ x_1(y_m) & x_2(y_m) & \dots & x_n(y_m) \end{pmatrix} \rightarrow \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{pmatrix}$$

В задаче диагностики разновидности гастрита опытный специалист может предоставить ценную информацию о возможностях диагностики, которая может быть полнее и объективнее, чем результаты, полученные при использовании предложенного метода логического анализа данных.

Однако, важно иметь общий подход, основанный на логическом анализе данных, который позволяет формально выявить наиболее значимые правила для постановки диагноза.

Важно понимать, что врачи и специалисты имеют ценный опыт и знания в данной области, и их экспертиза должна быть учтена при разработке модели. Комбинация логического анализа данных и опыта специалистов может привести к лучшим результатам в диагностике и обеспечить более точное понимание возможностей диагностики гастрита [8].

3. Методы решения

Для реализации модели мы предложили Карту гистологического исследования (КГИ), которая представляет собой инструмент для диагностики и анализа патологических состояний. КГИ состоит из двух частей, каждая из которых содержит определенные данные.

Первая часть КГИ содержит информацию о пациенте, такую как фамилия, инициалы, пол, возраст, а также дата и номер исследования. Она также включает 28 диагностических признаков, которые устанавливаются во время гистологического исследования. Эти признаки организованы в определенной последовательности и представляют собой вариант первоначального алгоритма диагностики. Врач, проводящий исследование, фиксирует значения этих признаков в КГИ в строгой последовательности.

Вторая часть КГИ, называемая "Диагноз", содержит основные целевые признаки. Главный целевой признак - это сам диагноз, который может быть одним из трех значений: "норма",

"хронический поверхностный гастрит (ХПГ)" или "хронический атрофический гастрит (ХГА)". В зависимости от выбранного диагноза, возможно включение до 9 дополнительных признаков, таких как топография, этиология и активность.

КГИ позволяет систематизировать и стандартизировать процесс диагностики гистологических образцов и обеспечивает единый подход к оценке и классификации патологических состояний. Он предоставляет врачам полную информацию о пациентах и их гистологических данных, что способствует более точному и надежному диагнозу. КГИ также может быть использован в исследованиях и анализе данных для выявления связей между различными признаками и патологическими состояниями.

Каждый признак, включенный в Карту гистологического исследования (КГИ), представляет собой бинарный или многозначный предикат с различной значимостью. Для облегчения ввода данных в компьютер был использован код, который расположен слева от каждого значения рисунков 2.

1. Отдел желудка			2. Общая структура слизистой			13. Просвет желез		
<input type="checkbox"/> 0	<input type="checkbox"/>	кардиальный	<input type="checkbox"/> 0	<input type="checkbox"/>	сохранена	<input type="checkbox"/> 0	<input type="checkbox"/>	не определяется
<input type="checkbox"/> 1	<input type="checkbox"/>	фундальный	<input type="checkbox"/> 1	<input type="checkbox"/>	нарушена	<input type="checkbox"/> 1	<input type="checkbox"/>	расширен
<input type="checkbox"/> 2	<input type="checkbox"/>	антральный	<input type="checkbox"/> 2	<input type="checkbox"/>	резко нарушена			
<input type="checkbox"/> 3	<input type="checkbox"/>	все отделы						

Рис. 2 фрагмент КГИ

Таким образом, мы имеем формальную постановку задачи диагностики хронических гастритов, данную выше.

При решении диагностических задач мы сталкиваемся с неполными, неточными и неоднозначными данными. Однако полученные решения должны соответствовать закономерностям, явным и неявным, присутствующим в рассматриваемых данных.

Использование логических методов позволяет анализировать исходные данные, выделять значимые и незначимые признаки, определять минимальный набор правил, необходимых для полного восстановления исходных закономерностей. В результате можно получить более компактное и достоверное представление исходной информации, которое обладает большей надежностью и обрабатывается быстрее.

Будим считать, что если система правил, построенная на основе логического анализа данных, способна вывести все решения в рассматриваемой области, то она считается полной.

Группу диагнозов, выявленных на основе определенного признака или группы симптомов, можно называть классом [9].

Каждый диагноз может быть представителем одного или нескольких классов, причем каждый класс определяется набором схожих симптомов.

Таким образом, логический анализ данных позволяет выявить закономерности в данных и построить систему правил, которая обеспечивает полное восстановление исходных закономерностей. Это позволяет классифицировать диагнозы и разделить их на классы на основе схожих симптомов. Для каждого пациента мы можем составить решающее правило, которое устанавливает, что если у него одновременно присутствуют определенные симптомы, то это указывает на конкретный диагноз.

$$\&_{j=1}^{28} x_j(y_i) \rightarrow P(y_i), i = 1, \dots, l; x_j(y_i) \in \{0,1,2,3\}, \quad (1)$$

Это общее правило продукции, где предикат $P(y_i)$ принимает значение true, т.е. $P(y_i) = 1$ если $y = y_i$ и $P(y_i) = 1$, если $y \neq y_i$. Данное правило может быть записано в другой форме:

$$\bigvee_{i=1}^{28} \bar{x}(y_j) \vee P(y_j), \quad j \in [1, \dots, m]. \quad (2)$$

Можно сказать, что решающие функции для набора данных представляют собой конъюнкцию всех решающих правил.

$$f(X) = \&_{j=1}^{132} \left(\bigvee_{i=1}^{28} \bar{x}_i \vee P(y_j) \right). \quad (3)$$

Построенная функция позволяет исключить несущественные симптомы и разделить данные на классы, где один и тот же диагноз может быть характеризован разными симптомами. В итоге мы получим булеву функцию от $m+n$ переменных, где m - количество больных, а n - количество симптомов. Таким образом, данная функция позволяет установить любые правила, за исключением

отрицания тех правил, которые уже существуют. Это означает, что функция будет верна для всех комбинаций переменных, где все симптомы присутствуют и диагноз соответствует этим симптомам, и не будет верна только в случае, когда такой комбинации симптомов присутствует, но диагноз отрицается.

Эту функцию можно легко изменять, добавляя новые правила с использованием операции конъюнкции, которые могут включать уже существующие правила с некоторыми модификациями. Это позволяет расширять систему правил и учитывать новые данные или ситуации.

Кроме того, эту функцию можно представить в рекурсивной форме, где каждое правило может вызывать другие правила или подфункции для определения диагноза. Такая рекурсивная структура позволяет более гибко и эффективно обрабатывать данные и принимать решения на основе набора симптомов.

В результате, благодаря возможности изменения и расширения функции, а также использованию рекурсивной формы, мы можем создавать более гибкие и адаптивные системы диагностики, способные учитывать различные варианты симптомов и принимать более точные решения.

$$\begin{aligned}
 W(X) &= Z_k(q_k w_k X); \\
 Z_k(q_k w_k X_k) &= Z_{k-1} \& \left(\bigvee_{i=1}^{28} \bar{x}_k(w_i) \vee w_k \right) \vee q_{k-1} \& \left(\bigvee_{i=1}^{28} \bar{x}_k(w_i) \vee w_k \right); \\
 q_k &= q_{k-1} \& \left(\bigvee_{i=1}^{28} \bar{x}_k(w_i) \right); \\
 q_1 &= \bigvee_{i=1}^{28} \bar{x}_1(w_i); \quad j = 2, \dots, 123; \quad Z_1 = w_1.
 \end{aligned} \tag{4}$$

Где $W(X)$ – моделируемая функция, Z_j – характеристика объектов в текущий момент, q_j – состояние системы в текущий момент [12].

Если функция записана в стандартной дизъюнктивной нормальной форме (СДНФ) и успешно сокращена, она может представлять данные в компактной форме. В этой функции будут содержаться диагнозы, классы, объединяющие диагнозы на основе сходных симптомов, а также комбинации симптомов, которые не характерны для рассматриваемых диагнозов.

Однако, при работе с большими объемами данных такой подход может стать несколько громоздким. Поэтому предлагается следующий алгоритм реализации этого метода.

Такой алгоритм позволяет создать компактное и надежное представление данных, где диагнозы объединены в классы на основе симптомов. Это упрощает анализ и обработку данных, а также обеспечивает более быструю и точную диагностику пациентов.

Алгоритм.

Алгоритм выбора правил, позволяющих охватить все рассматриваемые данные следующий:

Необходимо определить количество столбцов в таблице, которое будет соответствовать количеству вопросов по каждому симптому и количеству возможных ответов. В данном случае предполагается, что ответы могут быть от одного до четырех на каждый вопрос.

Количество строк в таблице будет соответствовать количеству пациентов, а также классов, которые могут быть обнаружены.

Данные каждого пациента записываются в таблицу, соблюдая указанный порядок. Каждому пункту (симптому) будет соответствовать столбец, а каждому пациенту будет соответствовать строка в таблице. Далее будет распределение каждого диагноза по соответствующим столбцам. Например, диагноз "y1" будет помещен в столбец, соответствующий результатам обследования данного пациента. Если у вас есть разделение по полу пациентов, то пол может иметь два столбца со значениями 0 и 1, и диагноз будет помещен в соответствующий столбец в зависимости от пола пациента.

Как уже отмечалось ранее, для диагностики гастрита используются предикаты с разными значениями, также известные как k-значные предикаты. Набор входных параметров и допустимых решений для каждого диагноза может быть представлен в виде системы продуктивных правил:

Набор симптомов 1 → Решение 1,

Набор симптомов 2 → Решение 2,

...

Набор симптомов m → Решение m .

Важно отметить, что одно и то же решение может произойти из разных наборов входных данных.

С использованием построенных логических функций такая система продуктивных правил может быть преобразована в оптимальное логическое выражение. Это позволяет убрать избыточную информацию и выделить все возможные подклассы эквивалентных решений, что помогает выявить скрытые закономерности.

Такая система продуктивных правил, или имплицитивных высказываний, является основой для построения логических нейронных сетей, которые могут быть обучены работать с соответствующей базой знаний [13].

Анализ наличия свободных знаний в полученных ответах позволяет определить необходимость дополнительного обучения. Это может включать пополнение исходной системы продуктивных правил новыми допустимыми решениями. Таким образом, реализуется процедура улучшения адаптивных свойств диагностики гастрита.

В результате, использование такой системы продуктивных правил и анализ наличия свободных знаний позволяет улучшить диагностику гастрита и реализовать процедуру обучения для ее дальнейшего совершенствования.

Описание программы: Данная программа реализует описанный ранее алгоритм и состоит из двух исполняемых модулей:

Модуль 1: Расшифровка базы данных и анализ результатов.

Этот модуль выполняет расшифровку базы данных с использованием словаря.

Загружает симптомы и диагнозы в форме вопросно-ответных пар.

Анализирует результаты с использованием описанного алгоритма.

Модуль 2: Создание базы знаний Knoledge.

Этот модуль отвечает за создание информации на основе данных из исходного файла или для уточнения существующей системы знаний.

Он может уменьшать размер базы данных в соответствии с заданным приблизительным значением. В этом случае, либо точность алгоритма уменьшается, либо добавляется информация для проверки правильности хранимых данных.

Для получения результата диагностики с заданной точностью необходимо заполнить все поля, указанные на рисунке 1.

Если заданная точность невозможна, это будет отображено в соответствующем сообщении, как показано на рисунке 2.

Таким образом, программа предоставляет возможность проводить диагностику на основе базы данных, а также создавать и обновлять систему знаний для улучшения точности и адаптивности диагностического процесса.

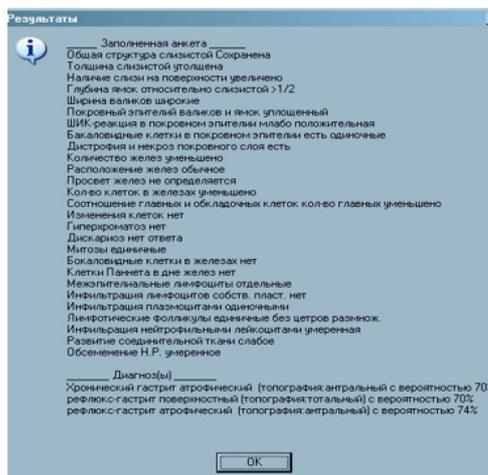


Рис. 3. Результат диагностики

В случае с предоставленными данными было проведено разбиение на две группы объектов: одна группа, обучаемая выборка, состояла из 80 больных, а другая группа, называемая контрольная группа, состояла из 40.

Контрольная группа пациентов продемонстрировала диагноз, который соответствовал ей с точностью в пределах от 90% до 94%.

Однако, особый интерес вызывали пограничные случаи, один из таких показан на рисунке 3.

Такие случаи часто становятся источником ошибок для других методов машинного обучения. Однако, чтобы избежать возможных ошибок, в таких ситуациях на экран выводятся все возможные варианты диагноза вместе с соответствующими вероятностями для каждого варианта, выводится вся диагностика. Это позволяет представить всю информацию о пациенте и его состоянии в полном объеме, учитывая все данные, которые были внесены в систему. Возможно так же вывести на экран совокупность симптомов по которым эти диагнозы совпадают. Что облегчит дальнейшее уточнение диагноза специалисту. Для рассматриваемого случая это

Инфильтрация лимфоцитов собств. пласт. нет
Инфильтрация плазмощитами одиночными
Лимфотические фолликулы единичные без центров размнож.
Инфильтрация нейтрофильными лейкоцитами умеренная
Развитие соединительной ткани слабое

Рис. 4. Общие симптомы для пограничных случаев

Заключение

В результате проведенного исследования был разработан программный комплекс для диагностики гастрита, основанный на логическом анализе данных. Предложенный метод анализа позволяет обнаружить скрытые закономерности в данных, классифицировать их и выявить уникальные свойства каждого диагноза. В отличие от нейросетевых подходов, этот логический метод понятен и не требует переобучения. Он способен выявить наиболее значимые закономерности и упростить процесс принятия решений.

Можно утверждать, что для интеллектуального анализа данных логические алгоритмы являются эффективным инструментом. Они рассматривают исходные данные как набор общих правил, среди которых можно выделить те правила, которые достаточны для объяснения всех наблюдаемых данных. Эти правила, в свою очередь, являются генеративными для рассматриваемой области и помогают лучше понять природу изучаемых объектов.

Программный комплекс может быть полезным инструментом для врачей и специалистов в области гастроэнтерологии, помогая им принимать обоснованные решения на основе логического анализа данных, а также обосновывать более полное понимание рассматриваемой проблемы.

Литература

1. Журавлёв Ю. И. Об алгебраическом подходе к решению задач распознавания или классификации // Проблемы кибернетики. 1978. Т. 33. С. 5–68.

2. Шибзухов З.М. Корректные алгоритмы агрегирования операций // Распознавание образов и анализ изображений. 2014. № 3-24. С. 377-382.
3. Ashley I. Naimi, Laura B. Balzer Multilevel generalization: an introduction to super learning // European Journal of Epidemiology. 2018. Vol. 33. P. 459-464.
4. Haoxiang, Wang, Smith S. Big data analysis and perturbation using a data mining algorithm // Journal of Soft Computing Paradigm. 2021. №. 3 - 01. P. 19-28.
5. Joe MrK, Vijesh, Jennifer S. Raj User Recommendation System Dependent on Location-Based Orientation Context // Journal of Trends in Computer Science and Smart Technology. 2021. № 3-01. P. 14-23.
6. Grabisch M., Marichal J-L, Pap E. Aggregation functions // Cambridge University Press. 2009. Vol. 127.
7. Calvo T, Belyakov G. Aggregating functions based on penalties // Fuzzy sets and systems. 2010. № 10-161. P. 1420-1436.
8. Mesiar R, Komornikova M, Kolesarova A, Calvo T. Fuzzy aggregation functions: a revision // Sets and their extensions: representation, aggregation and models. Springer-Verlag, Berlin, 2008.
9. Yang F, Yang Zh, Cohen W.W. Differentiable learning of logical rules for reasoning in the knowledge base // Advances in the field of neural information processing systems. 2017. P. 2320-2329/
10. Флах П. Машинное обучение: наука и искусство построения алгоритмов, которые извлекают знания из данных. М.: ДМК Пресс, 2015. 400 с.
11. Akhlakur R., Sumaira T. Ensemble classifiers and their applications: a review // International Journal of Computer Trends and Technologies. 2014. Vol. 10. P. 31-35/
12. Дюкова Е.В., Журавлев Ю.И., Прокофьев П.А. Методы повышения эффективности логических корректоров // Машинное обучение и анализ данных. 2015. № 11-1. С. 1555-1583.
13. Lyutikova L. A., Shmatova E.V. Algorithm for constructing logical operations to identify patterns in data // E3S Web of Conferences, Moscow, 25-27 ноября 2020 года. - - Moscow, 2020. Vol. 224, P. 01009.

APPLICATION OF LOGICAL DATA ANALYSIS IN A SOFTWARE PACKAGE FOR THE DIAGNOSIS OF GASTRITIS

Lyutikova, Larisa A.

Candidate of physical and mathematical sciences

Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences, Institute of Applied Mathematics and Automation

Nalchik, Russian Federation

lylarisa@yandex.ru

Abstract

A software package for computer diagnostics of gastritis is proposed based on available indicators and the experience of specialists. It is based on a logical approach of data analysis and is designed to determine the most likely types of diseases with a given accuracy. The input data includes patient examination indicators, their diagnoses and the experience of specialists in diagnosis accumulated in medical practice. Systems of multivalued predicate logic are used to represent data. The result of the work of the software package is to select the most appropriate types of diseases with a predetermined accuracy, based on the diagnostic data of patients. If the accuracy of the diagnosis cannot be achieved with the specified accuracy, it is proposed either to change the accuracy of the decision, or to undergo an additional examination.

Keywords

diagnostics, connections, multivalued logic, input data, analysis, hidden patterns

References

1. Zhuravlev Yu. I. Ob algebraicheskom podkhode k resheniyu zadach raspoznavaniya ili klassifikatsii // Problemy kibernetiki. 1978. T. 33. S. 5–68.
2. Shibzukhov Z.M. Korrektnyye algoritmy agregirovaniya operatsiy // Raspoznavaniye obrazov i analiz izobrazheniy. 2014. № 3–24. С. 377-382/
3. Ashley I. Naimi, Laura B. Balzer Multilevel generalization: an introduction to super learning // European Journal of Epidemiology. 2018. Vol. 33. P. 459–464.
4. Haoxiang, Wang, Smith S. Big data analysis and perturbation using a data mining algorithm // Journal of Soft Computing Paradigm . 2021. No. 3 – 01. P. 19-28.
5. Joe MrK, Vijesh, Jennifer S. Raj User Recommendation System Dependent on Location-Based Orientation Context // Journal of Trends in Computer Science and Smart Technology. 2021. No. 3-01. С. 14-23.
6. Grabisch M., Marichal J-L, Pap E. Aggregation functions // Cambridge University Press. 2009. Vol. 127.
7. Calvo T, Belyakov G. Aggregating functions based on penalties // Fuzzy sets and systems. 2010. No. 10-161. P. 1420-1436.
8. Mesiar R, Komornikova M, Kolesarova A, Calvo T. Fuzzy aggregation functions: a revision // Sets and their extensions: representation, aggregation and models. Springer-Verlag, Berlin, 2008.
9. Yang F, Yang Zh, Cohen W.W. Differentiable learning of logical rules for reasoning in the knowledge base // Advances in the field of neural information processing systems. 2017. P. 2320-2329
10. Flakh P. Mashinnoye obucheniye: nauka i iskusstvo postroyeniya algoritmov, kotoryye izvlekayut znaniya iz dannykh. M.: DMK Press, 2015. 400 s.
11. Akhlakur R., Sumaira T. Ensemble classifiers and their applications: a review //International Journal of Computer Trends and Technologies. 2014. Vol. 10. P. 31-35
12. Dyukova Ye.V., Zhuravlev Yu.I., Prokof'yev P.A. Metody povysheniya effektivnosti logicheskikh korrektorov // Mashinnoye obucheniye i analiz dannykh. 2015. № 11-1. S. 1555-1583.
13. Lyutikova L. A., Shmatova E.V. Algorithm for constructing logical operations to identify patterns in data // E3S Web of Conferences, Moscow, 25–27 ноября 2020 года. – Moscow, 2020. Vol. 224, P. 01009.