УДК 004.891: 004.832.3 doi: 10.15827/0236-235X.150.297-304 2025. Т. 38. № 2. С. 297–304

Модификация аргументационного алгоритма в задачах классификации

А.А. Николаев [1], Н.А. Благосклонов [1]✉, Б.А. Кобринский [1]

[1] Федеральный исследовательский центр «Информатика и управление» РАН, г. Москва, 119333, Россия

Ссылка для цитирования
Николаев А.А., Благосклонов Н.А., Кобринский Б.А. Модификация аргументационного алгоритма в задачах классификации // Программные продукты и системы. 2025. Т. 38. № 2. С. 297–304. doi: 10.15827/0236-235X.150.297-304
Информация о статье
Группа специальностей ВАК: 1.2.2
Поступила в редакцию: 28.06.2024 После доработки: 06.09.2024 Принята к публикации: 19.11.2024

Аннотация. Объяснимость гипотез в интеллектуальных системах поддержки принятия решений определяется для пользователя их аргументированностью. В статье рассмотрены важность аргументации и различные подходы к рассуждениям в интеллектуальных системах. В условиях выдвижения ряда гипотез происходит декомпозиция правил вывода для сходных труднораспознаваемых заболеваний или патологических состояний. Это может быть обусловлено наличием частично пересекающихся признаков (факторов риска). Одновременно возникает проблема формирования ранжированного списка гипотез, сопровождаемого всеми аргументами, включая аргументы низшего уровня, то есть относящиеся к слабым гипотезам. В настоящей работе рассматривается модификация алгоритма аргументационных рассуждений, что обеспечивает мягкое уменьшение числа гипотез при выделении из них одной или более ведущих, соответствующих наличию более одного подкласса или группы болезней, и использован метод задания отношения порядка. Модификация аргументационного алгоритма представлена и обоснована в рамках ранее созданной базы знаний интеллектуальной рекомендательной системы для оценки рисков развития заболеваний. Система реализована на неоднородной семантической сети. Представлена коррекция шагов алгоритма. Решатель обеспечивает ранжирование выдаваемых гипотез при одновременном сохранении информации обо всех обнаруженных аргументах независимо от их значимости. Модифицированный решатель включает предотвращение возможной потери одной из релевантных гипотез при одновременном наличии ряда болезней, предоставление информации обо всех аргументах для множества гипотез различного ранга, повышение объяснимости выдаваемых гипотез на основе признаков, аргументирующих различные классифицируемые заболевания. Приводится сравнение модифицированного алгоритма с другими подходами, обеспечивающими интерпретацию выдаваемых решений. Практическая значимость работы определяется повышением объяснимости для пользователя ведущих гипотез при одновременном получении всего комплекса обнаруженных аргументов.

Ключевые слова: алгоритм аргументационных рассуждений, аргументационный решатель, объяснимость, неоднородная семантическая сеть, классификация, система поддержки принятия решений, рекомендательная система, медицинская интеллектуальная система

Введение. В настоящее время все более широкое применение в медицине находят методы искусственного интеллекта. В клинической практике используются консультативные системы для поддержки принятия врачебных решений и рекомендательные системы, ориентированные на советы в области здоровьесбережения и основанные на учете индивидуальных особенностей пользователей. Системы для поддержки принятия врачебных решений предполагают взаимодействие с врачом в качестве профессионального пользователя, что накладывает определенные требования на формулировку и объяснение выдаваемых гипотез (решений). Рекомендательные системы чаще ориентированы на неспециалиста в проблемной области – пациента, для которого формируются советы в соответствии с решаемой задачей здоровьесбережения, включая выявление предрасполагающих факторов риска (предикторов) болезни. В тех случаях, когда рекомендательные системы ориентированы на врача, они могут выдавать советы по профилактике, обследованию, лечению или реабилитации. Оба класса – системы для поддержки принятия врачебных решений и рекомендательные системы – сводятся к решению задачи классификации. Помимо медицины, подобные задачи решаются в социологии, экономике и в других областях.

В медицинских системах искусственного интеллекта возникают ситуации, когда выдается слишком много предупреждений или рекомендаций, включая несущественные, и пользователи начинают игнорировать их независимо от важности, что определяет необходимость выбора наиболее существенных. Приоритизация играет ведущую роль в принятии пользователем окончательного решения на основе предлагаемой гипотезы или альтернативных гипотез.

В условиях ограниченной исходной информации особое значение имеют аргументирован-

ность и объяснимость. Однако именно этот актуальный аспект в системах поддержки принятия решений не нашел полноценного решения. Предлагаемый в настоящей статье подход к формированию решателя интеллектуальной системы позволяет при классификации (дифференциальной диагностике, прогнозировании) не только формировать и аргументировать более одной гипотезы (дифференциальный ряд), но и выделять ведущую гипотезу, сопровождаемую признаками, характеризующими весь спектр первично сформированных гипотез, что является отличием от общепринятого подхода и имеет существенное практическое значение.

Аргументация в интеллектуальных системах

С самого начала исследования в области искусственного интеллекта характеризовались фундаментальным логическим подходом [1]. Аргументация может показать шаг за шагом, как система искусственного интеллекта принимает решение, предоставить рассуждение о неопределенности и находить решения при выявлении противоречивой информации [2, 3]. Значительный вклад в теорию и практику аргументированных рассуждений внесли работы [4, 5], в которых утверждается, что решатели должны использовать (в зависимости от потребности) логики как объективного, так и субъективного (экспертного) знания [6]. Применительно к слабоструктурированной медицинской предметной области логика используется для представления знаний в понятном человеку виде и с аргументированием посредством правил, полученных от экспертов и/или моделей принятия решений, созданных на основе данных [7].

Использование теории аргументации для объяснения того, почему началось событие или что привело к решению, является способностью рассуждения, которую может предложить Argumentation Framework [3]. В контексте систем поддержки принятия клинических решений для помощи в рассуждениях о методах лечения возможно использование схем аргументации [8]. Такой подход обеспечивает механизм принятых в практике врачей клинических рассуждений, отвечающих понятию формальной аргументации. В работе [9] сделан аргументационный анализ для оценки негативных последствий применения медикаментов при коморбидных (сочетанных) заболеваниях. Решая задачу диагностики, интеллектуальная система будет стремиться обнаружить наличие признаков, являющихся положительными аргументами для гипотезы, и отсутствие признаков, являющихся отрицательными аргументами [10]. Аргументация позволяет делать выводы на противоречивых и неполных наборах утверждений, которые могут быть пересмотрены на более поздних этапах рассуждений при поступлении новых знаний [11].

При рассмотрении точной эпистемологии, позволяющей реализовать упорядоченное множество стратегий правдоподобных рассуждений, автор [12] выделяет среди прочего вывод следствий из посылок, способность к объяснению и аргументацию при принятии решений.

Алгоритм аргументационных рассуждений

В слабоструктурированной медицинской предметной области особенно целесообразно построение интеллектуальной системы, учитывающей отношения признаков и опирающейся на систему аргументов. Важным аспектом консультативных и рекомендательных систем является аргументированная поддержка гипотез, на основе которых пользователь должен принимать окончательное решение. В интеллектуальной рекомендательной системе по здоровьесбережению [13] был использован алгоритм аргументационных рассуждений [14], который обеспечивал возможность выдвижения гипотез (уровни риска развития болезни и рекомендации по их предупреждению), подтверждая их аргументами (факторами риска). Данный алгоритм характеризуется рядом особенностей, существенных для функционирования системы и понимания принципа ее работы.

– Шаги алгоритма близки к человеческим рассуждениям, соответствуют характеру клинического мышления врача в процессе выдвижения и обоснования диагноза/прогноза, когда он в процессе рассуждения, рассматривая и анализируя теоретически возможные гипотезы, отсеивает менее вероятные с использованием аргументов за и против.

– Каждая гипотеза сопровождается перечнем аргументов, то есть фактически предлагается объяснение, на основе каких аргументов система приняла конкретное решение. Это важно для медицинских (и не только) задач, так как иначе у пользователя не возникает достаточного доверия к предлагаемым системой гипотезам.

– Алгоритм, учитывая аргументы и контраргументы, обеспечивает правдоподобный вывод вследствие того, что доверие к гипотезе

представляет собой сумму масс свидетельств, поддерживающих гипотезу, а правдоподобие гипотезы отличается от абсолютной достоверности на сумму масс всех свидетельств, противоречащих гипотезе, и что такие оценки вполне соответствуют практике здравого смысла и позволяют интерпретировать доверие и правдоподобие как границы интервала возможного значения истинности гипотезы: доверие ≤ какая-то мера истинности ≤ правдоподобие. В силу разнообразия проявлений болезней и индивидуальных особенностей пациентов в клинической практике часто встречается противоречивая картина течения заболевания. Рассматриваемый в статье алгоритм способен оперировать нечеткими и даже противоречивыми признаками, характеризующими риск развития болезни.

– Алгоритм способен работать с гипотезами, аргументы которых имеют определенные пересечения. То есть в ситуациях, когда уровни риска нескольких заболеваний имеют один и тот же аргумент (фактор риска), не происходит потери гипотез.

Принципы построения БЗ рекомендательной интеллектуальной системы

Для решения задачи выявления рисковых ситуаций развития хронических неинфекционных заболеваний (артериальной гипертонии, инфаркта миокарда, инсульта, депрессии) была создана интеллектуальная рекомендательная система ИИ-ГИППОКРАТ, обеспечивающая формирование гипотез с персональными рекомендациями по здоровьесбережению [13]. Для оценки уровней риска разработана пятиранговая шкала с диапазоном значений от очень низкого до очень высокого, что позволило структурировать и в последующем формализовать нечеткие характеристики отклонений в состоянии здоровья.

На основе неоднородной семантической сети, которая содержит знания о факторах риска, уровнях риска и профилактических рекомендациях, реализована БЗ. Концептуально она состоит из разделов, узлов неоднородной семантической сети, свойств узлов и связей между узлами [14]. Разделы БЗ – это узлы, сгруппированные на основе их логической взаимосвязи. Принадлежность узла к определенному разделу базы не влияет на алгоритм выдвижения и подтверждения или отклонения гипотез, но может использоваться в последующем на этапе постобработки выдаваемого решения. Узлы неоднородной семантической сети – основные составляющие БЗ. В качестве узлов выступают факторы риска заболеваний (включающие характеристики состояния здоровья пользователя, факты об образе жизни, личностные факторы, ситуации), уровни риска и рекомендации по профилактике. Узлы представлены семантически полноценными понятиями и могут быть одного из четырех видов: узел-признак (FEATURE_NODE), целевой узел или узел-гипотеза (TARGET_NODE), сервисный узел полной активации или узел «И» (FULL_ACTIVATION_SERVISE_NODE) и сервисный узел частичной активации или узел «ИЛИ» (PARTIAL_ACTIVATION_SERVISE_NODE). Узлы-гипотезы – это уровни риска и рекомендации. Узлы-признаки – факторы риска. Сервисные узлы обеспечивают выполнение логических операций И/ИЛИ.

Связи между узлами представлены четырьмя основными типами, характеризующими состояние гипотезы.

RS – слабая положительная связь, соответствует выражению «узел со входящей связью может наблюдаться, когда наблюдается узел с исходящей связью»; связь типа RS указывает на возможность или увеличение вероятности гипотезы.

TRA – сильная положительная связь, соответствует выражению «узел со входящей связью наблюдается всегда, когда наблюдается узел с исходящей связью»; связь типа TRA указывает на совместное наличие аргументов или обязательное наличие гипотезы.

S – сильная отрицательная связь, соответствует выражению «узел со входящей связью никогда не наблюдается при наличии узла с исходящей связью»; связь типа S указывает на невозможность наличия гипотезы.

SN – слабая отрицательная связь, соответствует выражению «узел с входящей связью, может не наблюдаться, когда наблюдается узел с исходящей связью»; связь типа SN указывает на уменьшение вероятности гипотезы.

Такая структурная организация БЗ выбрана, поскольку позволяет оперировать наполненными смыслом понятиями, что способствует адекватной интерпретации гипотез, то есть предлагаемых решений, которые должны быть понятны пользователям.

Основания для модификации решателя

В ходе разработки системы ИИ-ГИППОКРАТ, в которой предполагалась выдача множества

гипотез о риске нескольких заболеваний с их объяснением, изначально был реализован описанный алгоритм аргументационных рассуждений [14]. Однако в процессе эксплуатации системы выявились определенные проблемы, послужившие причиной для внесения изменений в работу аргументационного алгоритма. Обнаружилось, что в процессе его использования часть гипотез может быть потеряна. При детальном анализе стало ясно, что причина в особенности как поставленной задачи, так и самого алгоритма.

Выяснилось, что при работе аргументационного алгоритма гипотезы теряются на этапе редукции множества первично сформированных гипотез. Например, на этапе выдвижения и подтверждения гипотез в числе множества гипотез мог быть риск как инфаркта, так и инсульта, а после редукции мог остаться, например, только риск инфаркта. Предполагалось, что для каждой болезни в случае выявленных рисков должна была выводиться гипотеза, соответствующая обнаруженному риску (то есть если для какой-то патологии был выявлен риск, то на выходе должно быть отражено наличие риска в отношении данного патологического состояния), однако бывали ситуации, когда это требование не соблюдалось (нарушалось). Разрешить их можно было путем создания отдельных БЗ для каждой из болезней. Однако такой подход имеет ощутимый недостаток, заключающийся в том, что огромное количество узлов нужно дублировать в разных БЗ и часть из них синхронизировать, а это крайне трудоемко. Особенно ухудшалась бы ситуация при расширении БЗ за счет включения новых проблемных областей, которые будем именовать подклассами, например, подклассы ишемическая болезнь сердца, нарушения ритма сердца и пороки сердца в классе заболеваний сердечно-сосудистой системы.

Для детального рассмотрения причин, по которым требование сохранения гипотез для каждой из патологий нарушалось, приведем описание шагов аргументационного алгоритма. Введем определения, связанные с алгоритмом аргументационных рассуждений. Пусть E – заданное множество узлов. Введем атомарные формулы:

1. $O(e)$ – узел e имеет место;
2. $\neg O(e)$ – узел e не имеет места;
3. $M(e)$ – узел e, возможно, имеет место;
4. $H(e)$ – узел e является гипотезой;
5. $\neg H(e)$ – узел e не является гипотезой;
6. $S(e)$ – событие e является решением (подтвержденной гипотезой);
7. $\neg S(e)$ – событие e не является решением (подтвержденной гипотезой).

Введем следующие правила вывода:

П1. $(O(e_1), (e_1,e_2) \in TRA) \rightarrow S(e_2)$;

П2. $(O(e_1), (e_1,e_2) \in RS) \rightarrow H(e_2)$;

П3. $(H(e_2), O(e_1), (e_1,e_2) \in S) \rightarrow \neg H(e_2)$;

П4. $(H(e_1), (e_1,e_2) \in RS) \rightarrow M(e_2)$;

П5. $(H(e_1), \neg O(e_2), (e_1,e_2) \in TRA) \rightarrow \neg H(e_1)$.

Рассмотрим более подробно шаги алгоритма.

1. Порождение множества гипотез.

Для всех e_1, таких, что для них имеет место $O(e)$ и $(e_1, e_2) \in TRA$, применить правило П1: $(O(e_1), (e_1, e_2) \in TRA) \rightarrow S(e_2)$; выполнить $H := H \cup \{e_2\}$.

2. Расширение множества аргументов.

Ко всем e_1, таким, что $H(e_1)$, $(e_1, e_2) \in RS$, применить правило П4 и построить множество $M := M \cup \{e_2\}$ всех e_2, таких, что $M(e_2)$.

3. Тестирование аргументов.

Для каждого e_1, такого, что $M(e_1)$, применяем процедуру Q; если $Q(e) = O(e)$, то $O := O \cup \{e\}$. Переходим к шагу 1.

Шаги 1–3 алгоритма являются циклом, и цикл продолжается до стабилизации множеств O и H.

4. Уменьшение множества гипотез по отвергающим аргументам.

Для всех e, таких, что имеет место $H(e)$, $O(e_1)$ и $(e_1, e) \in S$, полагаем $H := H / \{e\}$.

5. Уменьшение множества гипотез по обусловленным аргументам.

Для всех e, таких, что $H(e)$, $(e, e_1) \in TRA$ и выполняется $\neg O(e_1)$, в соответствии с правилом П5 заключаем, что $\neg H(e)$. Полагаем $H := H / \{e\}$.

6. Если мощность множества гипотез $H \vee$ в результате оказалась меньше или равной единице, полагаем $S = H$, и алгоритм завершает работу.

7. Удаление вложенных гипотез.

Если $|H| > 1$ и в H найдутся две гипотезы h_1 и h_2, множества аргументов которых $Arg(h_1)$ и $Arg(h_2)$ строго вложены одно в другое, то есть $Arg(h_1) \subset Arg(h_2)$, то гипотеза h_1 удаляется из множества гипотез. Эта процедура применяется попарно ко всем таким гипотезам. Результирующее множество гипотез H называется объясняющим множеством.

8. Минимизация объясняющего множества.

Если $Arg(h)$ – множество аргументов е гипотезы h, таких, что $O(e)$, и если $|H| > 2$ и в H найдутся гипотезы h_1, h_2, …, h_n, такие, что $Arg(h_1) \subset Arg(h_2) \cup \ldots \cup Arg(h_n)$, то h_1 удаляется из H. Повторяется до исчерпания множества таких гипотез.

9. $S = H$. Завершение работы.

Любой полученной в ходе работы аргументационного алгоритма гипотезе H_i соответствует некоторое множество аргументов $Arg(H_i)$.

Однако рассмотренная базовая реализация алгоритма, включающая девять шагов, не подходила для решения задачи оценки уровней риска заболеваний. Оказалось, что шаг 7 алгоритма не может быть использован из-за того, что при имеющейся постановке задачи могла сложиться следующая ситуация. Пусть в результате работы первых пяти шагов алгоритма множество H содержит гипотезы h_1, h_2, известно, что $ARG(h_1) \subset ARG(h_2)$ и h_1, h_2 относятся к различным заболеваниям. В случае применения шага 7 алгоритма гипотеза h_1 будет исключена из множества H. Это приводит к потере риска для одного из заболеваний.

Схожая ситуация возникает в случае применения шага 8 алгоритма. Пусть имеются гипотезы h_1, h_2, h_3, относящиеся к различным заболеваниям, и $ARG(h_1) \subset ARG(h_2) \cup ARG(h_3)$, при этом $ARG(h_1)$ не является подмножеством $ARG(h_2)$ и $ARG(h_3)$. В случае применения шага 8 алгоритма гипотеза h_1 была бы исключена, что привело бы к потере риска для одной из болезней.

С точки зрения диагностики (распознавания) основного заболевания такой подход является правильным и допустимым, если не стоит задача выявления сопутствующих заболеваний, однако при задаче выявления уровней риска нескольких заболеваний алгоритм может исключить критически важную оценку – очень высокий уровень риска некоторого заболевания. В то же время с медицинской точки зрения важно выделить в первую очередь гипотезы с высоким уровнем риска заболевания при одновременном получении информации обо всех факторах риска, релевантных для гипотез более низкого уровня.

Модификация аргументационного алгоритма

Для решения этой проблемы была проведена реструктуризация задачи, то есть выделен новый тип задач, и для их решения модифицирован аргументационный алгоритм. Опишем новый тип задач. Пусть семантическая сеть содержит n гипотез H_1, ..., H_n, гипотезы разделены на m групп (группы могут иметь пересечения). Внутри каждой из групп задано строгое отношение порядка, то есть определена некоторая функция *rang*, которая присваивает каждой гипотезе из группы определенное число – чем выше число, тем более высокий ранг имеет гипотеза. Задачей для каждой из групп является получение гипотезы с наивысшим рангом в случае, если она подтверждена, а для информирования пользователя необходимо вывести гипотезы с их аргументами для более низких рангов. Оказалось, что для такого типа задач нужно в аргументационном алгоритме заменить шаг редукции множества гипотез. Взамен каждая группа гипотез будет рассматриваться отдельно, из нее будет выделяться гипотеза с наивысшим рангом, а аргументы остальных гипотез добавляться к ней в качестве информации, которая может оказаться полезной в силу наличия уникальных аргументов, характеризующих выявленную патологию, то есть повысит объяснимость выдаваемой гипотезы. Аргументация естественным образом связана с объяснением. В [3] было обращено внимание на то, что использование аргументации для предоставления объяснений делает систему с искусственным интеллектом более дружелюбной и заслуживающей большего доверия у пользователя.

Пусть R_1, …, R_m – группы рисков. Определена функция $rang_j(h)$, которая каждой гипотезе из группы R_j ставит в соответствие некоторое числовое значение. Гипотезы, описывающие более высокие уровни рисков в рамках одной группы, должны получать более высокие значения, таким образом, функция *rang* устанавливает порядок гипотез в рамках группы.

Как уже было сказано, в результате работы шагов 1–5 алгоритма получено множество гипотез H. Для каждой группы R_j выбираются все гипотезы из H, принадлежащие рассматриваемой группе. Обозначим полученное множество $H_j = \{h : h \in R_j \wedge h \in H\}$. Если $|H_j| \leq 1$, то оставляем все без изменений. В случае, когда $|H_j| \geq 2$, выберем гипотезу из множества H_j с максимальным значением функции *rang* и обозначим ее $\hat{h_j}$. Гипотеза $\hat{h_j} \in H_j$ будет отобрана как основная для данной группы, гипотезы с более низкими рангами $H_j/\{\hat{h_j}\}$ будут отобраны для возможности учета уникальных аргументов для данной группы.

Таким образом, в качестве основного решения будут получены гипотезы, каждая из которых имеет наивысший ранг в своей группе.

Следует отметить, что предложенный в ходе работы способ модификации аргументационного алгоритма из медицинской сферы может быть применен в других предметных областях.

Дискуссия

Рассмотрим некоторые подходы к аргументации в системах искусственного интеллекта. Так, ДСМ-рассуждения осуществляют взаимодействие индукции для порождения предгипотез о причинах эффектов и аналогии для порождения предгипотез о предсказаниях этих эффектов [15]. В представленном в настоящей работе алгоритме в поиске решения использована дедукция на основе сопоставления объективного знания нового конкретного случая с субъективными (экспертными) знаниями базы правил неоднородной семантической сети.

Аргументация и принятие решений представлены в [16] для рекомендации лечения пациентам с множественными хроническими заболеваниями. Авторы предлагают подход к обоснованию решения, предоставляя два разных способа объяснения: при первом, называемом пассивным, пациент принимает все аргументы, указывающие конкретное лечение, при втором, называемом активным, пациент может попросить объяснение о рекомендованном лечении. Предложенный и рассмотренный авторами модифицированный аргументационный механизм вывода имеет сходство с рассмотренным в [16], но вместо поиска оптимальных путей для предотвращения неблагоприятных взаимодействий при подборе лечения пользователю представляются однозначно значимые аргументы, указывающие на основную гипотезу, в то время как дополнительные аргументы соответствуют гипотезе(ам) более низкого уровня. Это напоминает пассивный вариант объяснения, но одновременно включает аргументы, потенциально отвечающие на вопрос о дополнительных признаках для каждой из выдвигаемых гипотез.

В [17] аргументативный диалектический процесс рассуждения и ассоциация аргументов, которую он конструирует, представлены для объяснения. Такие пояснения могут иметь атрибутивный элемент, исходящий из начальных аргументов, поддерживающих вывод, в то время как дополнительные аргументы усиливают объяснительный контраст предлагаемого решения в противовес контраргументам. Объяснение в рассматриваемом случае при сочетанном риске нескольких заболеваний или диагностике нескольких заболеваний включает признаки разного уровня значимости. Рекомендации выдаются в краткой и развернутой формах.

Заключение

В определенных ситуациях (для определенного класса задач) при выдаче множества гипотез об одновременно выявленных нескольких состояниях (заболеваниях) может иметь место потеря отдельных гипотез. Это послужило причиной для модификации алгоритма аргументационных рассуждений. Модифицированный аргументационный алгоритм позволяет осуществлять мягкое уменьшение множества гипотез. На этапе постобработки была реализована функция выбора гипотезы наивысшего риска заболевания с ее аргументами. Выдача ведущей или ведущих (для каждого подкласса, группы) гипотез сопровождается предоставлением пользователю информации о признаках, соответствующих аргументам гипотез более низкого уровня (более низкого уровня риска или сопутствующих неактивных хронических болезней). Это одновременно является основой для объяснимости выдаваемых решений, которая достигается на основе предоставления пользователю совокупности аргументов – факторов риска или признаков, подтверждающих гипотезу.

Список литературы

1. Amendola G. Special issue on logic-based artificial intelligence. Algorithms, 2023, vol. 16, no. 2, art. 106. doi: 10.3390/a16020106.

2. Cyras K., Rago A., Albini E., Baroni P., Toni F. Argumentative XAI: A survey. Proc. IJCAI, 2021, pp. 4392–4399. doi: 10.24963/ijcai.2021/600.

3. Vassiliades A., Bassiliades N., Patkos T. Argumentation and explainable artificial intelligence: A survey. The Knowledge Engineering Review, 2021, vol. 36, art. e5. doi: 10.1017/S0269888921000011.

4. Финн В.К. Об эвристиках ДСМ-исследований (дополнения к статьям) // НТИ. Сер. 2: Информ. процессы и системы. 2019. № 10. С. 1–34. doi: 10.36535/0548-0027-2019-10-1.

5. Финн В.К. ДСМ-рассуждения и обнаружение знаний: амплиативные выводы, распознавание причинности и три вида полноты // НТИ. Сер. 2: Информ. процессы и системы. 2022. № 4. С. 4–36. doi: 10.36535/0548-0027-2022-04-2.

6. Финн В.К. Об эмпирических закономерностях в ДСМ-методе автоматизированной поддержки исследований // НТИ. Сер. 2: Информ. процессы и системы. 2023. № 12. С. 14–33. doi: 10.36535/0548-0027-2023-12-2.

7. Calegari R., Ciatto G., Denti E., Omicini A. Logic-based technologies for intelligent systems: State of the art and perspectives. Information, 2020, vol. 11, no. 3, art. 167. doi: 10.3390/info11030167.

8. Sassoon I., Kökciyan N., Modgil S., Parsons S. Argumentation schemes for clinical decision support. Argument & Computation, 2021, vol. 12, no. 3, pp. 329–355. doi: 10.3233/AAC-200550.

9. Walton D., Oliveira T., Satoh K., Mebane W. Argumentation analytics for treatment deliberations in multimorbidity cases: An introduction to two artificial intelligence approaches. Topoi, 2021, vol. 40, pp. 373–386. doi: 10.1007/s11245-020-09701-6.

10. Соколов И.А. Теория и практика применения методов искусственного интеллекта // Вестн. РАН. 2019. Т. 89. № 4. С. 365–370. doi: 10.31857/S0869-5873894365-370.

11. Antipov S.G., Vagin V.N., Morosin O.L., Fomina M.V. Protection of information in networks based on methods of machine learning. In: CCIS. Proc. RCAI, 2018, vol. 934, pp. 273–279. doi: 10.1007/978-3-030-00617-4_25.

12. Финн В.К. Точная эпистемология и искусственный интеллект // НТИ. Сер. 2: Информ. процессы и системы. 2020. № 6. С. 1–36. doi: 10.36535/0548-0027-2020-06-1.

13. Kobrinskii B.A., Grigoriev O.G., Molodchenkov A.I., Smirnov I.V., Blagosklonov N.A. Artificial intelligence technologies application for personal health management. IFAC-PapersOnLine, 2019, vol. 52, no. 25, pp. 70–74. doi: 10.1016/j.ifacol.2019.12.448.

14. Осипов Г.С. Методы искусственного интеллекта. М.: ФИЗМАТЛИТ, 2015. 296 с.

15. Финн В.К. Об эмпирических закономерностях ранга R в ДСМ-методе автоматизированной поддержки исследований // НТИ. Сер. 2: Информ. процессы и системы. 2024. № 1. С. 11–33.

16. Shaheen Q., Toniolo A., Bowles J.K.F. Dialogue games for explaining medication choices. In: LNPSE. Proc. RuleML+RR, 2020, vol. 12173, pp. 97–111. doi: 10.1007/978-3-030-57977-7_7.

17. Dietz E., Kakas A., Michael L. Argumentation: A calculus for human-centric AI. Frontiers in Artificial Intelligence, 2022, vol. 5, art. 955579. doi: 10.3389/frai.2022.955579.

Software & Systems doi: 10.15827/0236-235X.150.297-304 2025, 38(2), pp. 297–304

Modifying the reasoning algorithm in classification tasks

Artem A. Nikolaev [1], Nikolay A. Blagosklonov [1]✉, Boris A. Kobrinsky [1]

[1] Federal Research Center “Computer Science and Control” of RAS,
Moscow, 119333, Russian Federation

For citation
Nikolaev, A.A., Blagosklonov, N.A., Kobrinsky, B.A. (2025) ‘Modifying the reasoning algorithm in classification tasks’, *Software & Systems*, 38(2), pp. 297–304 (in Russ.). doi: 10.15827/0236-235X.150.297-304
Article info
Received: 28.06.2024 After revision: 06.09.2024 Accepted: 19.11.2024

Abstract. The user determines the explainability of hypotheses in intelligent decision support systems in terms of their reasoning. The paper discusses the importance of reasoning and different approaches to reasoning in intelligent systems. Given a set of hypotheses, there is a decomposition of inference rules for similar difficult-to-recognize diseases or pathological conditions. This may be due to partially overlapping features (risk factors). Simultaneously, there arises the problem of forming a ranked list of hypotheses accompanied by all arguments, including arguments of the lowest level, i.e., those related to weak hypotheses. This paper considers a modification of the reasoning algorithm for argumentation reasoning. This provides a gentle reduction in the number of hypotheses by selecting one or more leading ones corresponding to the presence of more than one subclass or group of diseases. In addition, the authors use the method of assigning an order relation. The authors present and justify the reasoning algorithm modification within the framework of the previously created knowledge base of the intelligent recommendation system for disease risk assessment. The system is implemented on a heterogeneous semantic network. The algorithm steps are corrected. The solver ranks issued hypotheses while storing information about all detected arguments regardless of their relevance. The modified solver includes prevention against possible loss of one of the relevant hypotheses when there is a number of diseases present at the same time. It provides information about all reasons for multiple hypotheses of different ranks. The solver enhances the explainability of the issued hypotheses based on features that give reasons for different classified diseases. The authors compare the modified algorithm with other approaches that interpret the issued solutions. The practical significance of the work is in increasing the explainability for the user of the leading hypotheses while simultaneously retrieving the entire set of detected arguments.
Keywords: argumentation reasoning algorithm, argumentation solver, explainability, heterogeneous semantic network, classification, decision support system, recommendation system, medical intelligent system

References

1. Amendola, G. (2023) ‘Special issue on logic-based artificial intelligence’, *Algorithms*, 16(2), art. 106. doi: 10.3390/a16020106.

2. Cyras, K., Rago, A., Albini, E., Baroni, P., Toni, F. (2021) ‘Argumentative XAI: A survey’, *Proc. IJCAI*, pp. 4392–4399. doi: 10.24963/ijcai.2021/600.

3. Vassiliades, A., Bassiliades, N., Patkos, T. (2021) ‘Argumentation and explainable artificial intelligence: A survey’, *The Knowledge Engineering Review*, 36, art. e5. doi: 10.1017/S0269888921000011.

4. Finn, V.K. (2019) ‘On the heuristics of JSM research (additions to articles)’, *Automatic Documentation and Mathematical Linguistics*, 53, pp. 250–282. doi: 10.3103/S0005105519050078.

5. Finn, V.K. (2022) ‘JSM reasoning and knowledge discovery: Ampliative reasoning, causality recognition, and three kinds of completeness’, *Automatic Documentation and Mathematical Linguistics*, 56, pp. 79–110. doi: 10.3103/S0005105522020066.

6. Finn, V.K. (2023) ‘On empirical regularities in the JSM method of automated research support’, *Automatic Documentation and Mathematical Linguistics*, 57, pp. 362–381. doi: 10.3103/S0005105523060055.

7. Calegari, R., Ciatto, G., Denti, E., Omicini, A. (2020) ‘Logic-based technologies for intelligent systems: State of the art and perspectives’, *Information*, 11(3), art. 167. doi: 10.3390/info11030167.

8. Sassoon, I., Kökciyan, N., Modgil, S., Parsons, S. (2021) ‘Argumentation schemes for clinical decision support’, *Argument & Computation*, 12(3), pp. 329–355. doi: 10.3233/AAC-200550.

9. Walton, D., Oliveira, T., Satoh, K., Mebane, W. (2021) ‘Argumentation analytics for treatment deliberations in multimorbidity cases: An introduction to two artificial intelligence approaches’, *Topoi*, 40, pp. 373–386. doi: 10.1007/s11245-020-09701-6.

10. Sokolov, I.A. (2019) ‘Theory and practice of application of artificial intelligence methods’, *Herald of the Russian Academy of Sciences*, 89, pp. 115–119. doi: 10.31857/S0869-5873894365-370.

11. Antipov, S.G., Vagin, V.N., Morosin, O.L., Fomina, M.V. (2018) ‘Protection of information in networks based on methods of machine learning’, in *CCIS. Proc. RCAI*, 934, pp. 273–279. doi: 10.1007/978-3-030-00617-4_25.

12. Finn, V.K. (2020) ‘Exact epistemology and artificial intelligence’, *Automatic Documentation and Mathematical Linguistics*, 54(3), pp. 140–173. doi: 10.3103/S0005105520030073.

13. Kobrinskii, B.A., Grigoriev, O.G., Molodchenkov, A.I., Smirnov, I.V., Blagosklonov, N.A. (2019) ‘Artificial intelligence technologies application for personal health management’, *IFAC-PapersOnLine*, 52(25), pp. 70–74. doi: 10.1016/j.ifacol.2019.12.448.

14. Osipov, G.S. (2015) *Artificial Intelligence Methods*. Moscow, 296 p. (in Russ.).

15. Finn, V.K. (2024) ‘On rank r empirical regularities in the JSM method of automated research supporta’, *Automatic Documentation and Mathematical Linguistics*, 58, pp. 10–31. doi: 10.3103/S0005105524010035.

16. Shaheen, Q., Toniolo, A., Bowles, J.K.F. (2020) ‘Dialogue games for explaining medication choices’, in *LNPSE. Proc. RuleML+RR*, 12173, pp. 97–111. doi: 10.1007/978-3-030-57977-7_7.

17. Dietz, E., Kakas, A., Michael, L. (2022) ‘Argumentation: A calculus for human-centric AI’, *Frontiers in Artificial Intelligence*, 5, art. 955579. doi: 10.3389/frai.2022.955579.

Авторы

Артем Александрович Николаев [1],
ведущий программист, nicepeopleproject@gmail.com
Николай Алексеевич Благосклонов [1],
научный сотрудник, nblagosklonov@frccsc.ru
Борис Аркадьевич Кобринский [1],
д.м.н., профессор, заведующий отделом,
kba_05@mail.ru

[1] Федеральный исследовательский центр
«Информатика и управление» РАН,
г. Москва, 119333, Россия

Authors

Artem A. Nikolaev [1], Leading Programmer,
nicepeopleproject@gmail.com
Nikolay A. Blagosklonov [1],
Research Associate, nblagosklonov@frccsc.ru
Boris A. Kobrinsky [1], Dr.Sci. (Medics),
Professor, Head of Department,
kba_05@mail.ru

[1] Federal Research Center “Computer Science
and Control” of RAS,
Moscow, 119333, Russian Federation