

UNIVERSIDADE FEDERAL FLUMINENSE
INSTITUTO DE COMPUTAÇÃO
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

Matheus Folly de Paula Peixoto

DiseaseDx - Sistema Especialista para Auxílio no Diagnóstico Médico de Doenças
Raras

Trabalho de Conclusão de Curso submetido ao Departamento de Ciência da Computação da Universidade Federal Fluminense como requisito parcial para a obtenção do título de Bacharel em Ciência da Computação.

Niterói
2025

Matheus Folly de Paula Peixoto

DiseaseDx - Sistema Especialista para Auxílio no Diagnóstico Médico de Doenças Raras

Trabalho de Conclusão de Curso submetido ao Departamento de Ciência da Computação da Universidade Federal Fluminense como requisito parcial para a obtenção do título de Bacharel em Ciência da Computação.

Aprovado em 23 de julho de 2025:

Prof. Leonardo Gresta Paulino Murta
Orientador - UFF

Profa. Vanessa Braganholo Murta
Co-orientadora - UFF

Prof. Luiz Eduardo da Costa Oliveira
Co-orientador - UFF

Prof. Troy Costa Kohwalter
UFF

Heleno de Souza Campos Junior
UFF

Niterói
2025

Ficha catalográfica automática - SDC/BEE
Gerada com informações fornecidas pelo autor

P324d Paula Peixoto, Matheus Folly de
DiseaseDx : Sistema Especialista para Auxílio no
Diagnóstico Médico de Doenças Raras / Matheus Folly de
Paula Peixoto. - 2025.
34 f.: il.

Orientador: Leonardo Gresta Paulino Murta.
Coorientador: Vanessa Braganholo Murta; Luiz Eduardo da
Costa Oliveira.

Trabalho de Conclusão de Curso (graduação)-Universidade
Federal Fluminense, Instituto de Computação, Niterói, 2025.

1. Sistema especialista (Computação). 2. Inteligência
artificial. 3. Diagnóstico. 4. Doença. 5. Produção
intelectual. I. Paulino Murta, Leonardo Gresta, orientador.
II. Murta, Vanessa Braganholo, coorientadora. III. Costa
Oliveira, Luiz Eduardo da, coorientador. IV. Universidade
Federal Fluminense. Instituto de Computação. V. Título.

CDD - XXX

DiseaseDx - Sistema Especialista para Auxílio no Diagnóstico
Médico de Doenças Raras

*DiseaseDx – Expert System to Assist Medical Diagnosis
of Rare Diseases*

Matheus Folly de Paula Peixoto¹

Leonardo Gresta Paulino Murta²

Vanessa Braganholo Murta³

Luiz Eduardo da Costa Oliveira⁴

Resumo

Doenças raras representam um relevante problema de saúde pública que têm sido negligenciado ao longo do tempo, sobretudo em razão dos desafios que impõem à medicina. Devido à raridade dessas doenças, grande parte dos médicos possui pouco conhecimento sobre elas, o que contribui para subdiagnósticos, diagnósticos demorados ou até mesmo diagnósticos incorretos. Nesse contexto, surge a necessidade de ferramentas computacionais capazes de apoiar os médicos no processo de diagnóstico dessas enfermidades. Este estudo aborda o projeto e desenvolvimento de um sistema especialista, cujo propósito é auxiliar os médicos no diagnóstico de doenças raras. O objetivo do sistema é que, dado um conjunto de sintomas e resultados de exames como entrada, seja avaliada a convergência de possíveis diagnósticos que satisfaçam às operações lógicas contendo esse conjunto, concedendo uma maior pontuação para diagnósticos mais promissores.

Palavras-chaves: Sistema Especialista; Inteligência Artificial; Diagnóstico; Doenças Raras.

Abstract

Rare diseases represent a significant public health issue that has been neglected over time, mainly due to the challenges they pose to medicine. Because of the rarity of these diseases, most physicians have limited knowledge about them, which contributes to underdiagnosis, delayed diagnosis, or even misdiagnosis. In this context,

¹ Graduando do Curso de Ciência da Computação - UFF, matheusfolly@id.uff.br

² Orientador - Instituto de Computação - UFF, leomurta@ic.uff.br

³ Co-orientadora - Instituto de Computação - UFF, vanessa@ic.uff.br

⁴ Co-orientador - Departamento de Medicina Clínica - UFF, luizeduardocostaoliveira@id.uff.br

there is a need for computational tools capable of supporting physicians in the diagnostic process of these conditions. This study presents the design and development of an expert system whose purpose is to assist physicians in diagnosing rare diseases. The system's goal is to evaluate the convergence of possible diagnoses that satisfy the logical operations containing a given set of symptoms and test results as input, assigning a higher score to more promising diagnoses.

Keywords: Expert System; Artificial Intelligence; Diagnosis; Rare Diseases.

Aprovado em: 23/07/2025. **Versão Final em:** 25/07/2025

1. INTRODUÇÃO

Doenças raras são uma importante questão de saúde pública que tem sido bastante negligenciada ao longo da história, principalmente devido ao seu desafio para a medicina (SCHIEPPATI *et al.*, 2008). As definições de doenças raras podem variar para cada país, sendo definida para a União Europeia como uma doença que afete não mais que 50 a cada 100.000 indivíduos (SCHIEPPATI *et al.*, 2008). Já para os Estados Unidos, foi definido que essa proporção não seria maior que 86 a cada 100.000, enquanto para outros países essa proporção pode variar de 5 a 76 para cada 100.000 (WAKAP *et al.*, 2020). Apesar de individualmente afetarem poucas pessoas, é estimado que existam em torno de 7.000 doenças raras, aumentando significativamente o número de indivíduos que são afetados por alguma delas. De acordo com a Organização Mundial da Saúde (OMS), esse número de indivíduos é de aproximadamente 400 milhões (SCHAAF *et al.*, 2020). Em média, 1 a cada 17 pessoas é afetada por alguma doença rara ao longo de sua vida (RONICKE *et al.*, 2019).

Desse conjunto de doenças raras, aproximadamente 80% são de origem genética e afetam predominantemente crianças (SCHAAF *et al.*, 2020). A maioria dos médicos tem um conhecimento limitado sobre essas doenças, dada a raridade da sua natureza, o que leva a um subdiagnóstico⁵ ou diagnóstico demorado. Além do mais, mesmo que o médico suspeite que o paciente porte alguma doença rara, ainda há uma grande possibilidade de ele errar o diagnóstico devido à sobreposição de sintomas comuns com várias outras doenças raras (FAVIEZ *et al.*, 2020).

Segundo Ronicke *et al.* (2019), pacientes que portam alguma dessas doenças geralmente esperam em média 6 anos, a partir da manifestação dos sintomas, para obter um diagnóstico preciso. Durante essa espera, esses pacientes acabam tendo uma maior perda em qualidade de vida do que pacientes que portam alguma doença crônica comum, além dos gastos com consultas e exames que impactam monetari-

⁵ Diagnóstico insuficiente ou incompleto.

amente tanto o indivíduo quanto o sistema de saúde. Estima-se que, antes de se obter um diagnóstico correto, o paciente visita em média 7,3 médicos.

Em geral, o diagnóstico dessas doenças se torna um grande desafio, dadas as possibilidades de combinações de genes e sintomas. Esse grande número de combinações possíveis e a baixa incidência de casos, consequentemente, acarretam na falta de conhecimento e diagnósticos errados. Na medicina, existe um fenômeno chamado de *fechamento prematuro*, que ocorre quando o médico deixa de considerar outras alternativas após o diagnóstico inicial. Esse fenômeno é um dos erros mais comuns, e pode-se assumir que tenha um grande peso no caso das doenças raras (RONICKE *et al.*, 2019).

Diante desse cenário, há uma demanda por sistemas computacionais que possam auxiliar os médicos no diagnóstico dessas doenças. Um estudo conduzido por Faviez *et al.* (2020) identificou tipos de sistemas através de suas abordagens de implementação, como, por exemplo, sistemas *baseados em conhecimento* (derivados do conhecimento humano), *orientados a dados* (baseados em dados e aprendizado de máquina) e *híbridos* (orientados a dados, mas com algum tipo de conhecimento prévio). Segundo Faviez *et al.* (2020), os sistemas *orientados a dados* foram todos baseados em aprendizado de máquina, o que trouxe uma incerteza dada a falta de explicabilidade dos modelos. Já os sistemas *baseados em conhecimento* tinham como material conceitos fenotípicos, sendo a maioria deles implementados utilizando técnicas de similaridade simples, como presença ou frequência dos fenótipos. Os sistemas restantes foram classificados como “manuais”, que consistiam em árvores de decisão feitas manualmente, mas voltados apenas para doenças individuais ou grupos de doenças.

De acordo com Mehnen *et al.* (2023), muitas pessoas costumam procurar na internet a possível causa dos sintomas que estão sentindo, vide a existência de sites e aplicativos voltados para essa finalidade. O avanço das técnicas de inteligência artificial (IA) generativa e sua disponibilização ao público, como o ChatGPT, também despertou um forte interesse tanto pela comunidade científica quanto pela população em geral para esse uso. Entretanto, apesar de a IA conseguir retornar o diagnóstico correto em alguns casos, pequenas alterações em seus parâmetros podem fazer com que a IA alucine, dando respostas erradas para os mesmos casos previamente respondidos corretamente (PAL; UMAPATHI; SANKARASUBBU, 2023).

Esse trabalho tem como objetivo desenvolver um sistema especialista, denominado DiseaseDX, para auxiliar os médicos no diagnóstico de doenças raras, utilizando uma estrutura lógica de diagnóstico obtida no estudo de Gattorno *et al.* (2019). Diferentemente dos trabalhos baseados em aprendizagem de máquina ou mesmo IA generativa, o DiseaseDX tem como principal motivação ser determinístico e explicável, permitindo assim que o médico possa compreender a recomendação dada e tomar a decisão de diagnóstico mais apropriada. Esperamos não

apenas propor um sistema, mas instigar e contribuir para possíveis novas formas de estruturar e definir os diagnósticos de forma lógica. O sistema proposto representa doenças raras como expressões lógicas que relacionam sintomas e resultados de exames, permitindo que o médico informe esses valores e visualize, em tempo real, os diagnósticos que contenham os valores informados, sendo ordenados por relevância.

Para demonstrar o uso do sistema, desenvolvemos um exemplo baseado nas doenças descritas por Gattorno *et al.* (2019), onde o diagnóstico dessas doenças é descrito por uma expressão que combina elementos lógicos como *E*, *Ou* e *Ao Menos*, com elementos textuais como sintomas e resultados de exame, resultando em um diagnóstico com determinado grau de sensibilidade, especificidade e acurácia.

O restante do texto está organizado da seguinte forma: a Seção 2 trata da fundamentação teórica, detalhando os conceitos de sistemas especialistas e da representação de diagnósticos. A Seção 3 mostra os trabalhos relacionados, indicando a carência de sistemas determinísticos e explicáveis para o diagnóstico de doenças. A Seção 4 detalha a concepção, *design* e implementação do sistema DiseaseDX. A Seção 5 fornece uma demonstração do uso do sistema DiseaseDX, mostrando como um médico o utilizaria para obter um auxílio no diagnóstico de doenças raras. Finalmente, a Seção 6 apresenta a conclusão deste trabalho e aborda trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

Nesta seção, são explicados os conceitos fundamentais que compõem o trabalho. A Subseção 2.1 explica a origem, tipos, vantagens e desvantagens dos sistemas especialistas. A Subseção 2.2 explica a abordagem que utilizamos para configurar os diagnósticos no sistema.

2.1. SISTEMAS ESPECIALISTAS

De acordo com Liao (2005), o campo de estudo da Inteligência Artificial (IA) é bastante abrangente, sendo Sistemas Especialistas uma ramificação do mesmo, desenvolvido pela comunidade de IA nos meados da década de 1960. A ideia básica por trás desse tipo de sistema é transferir o conhecimento humano específico de uma tarefa para o computador. Com isso, as pessoas poderiam consultar esse sistema para obter uma conclusão ou um conselho relativo a essa tarefa. Já o sistema, dado alguns dados de entrada, avaliaria esses dados em sua base de conhecimento e chegaria a conclusões, além de que também conseguiria explicar a lógica por trás delas, caso necessário, assim como um consultor ou especialista faria. Conforme mencionado por Saibene, Assale e Giltri (2021), “o poder de um sistema especialista deve residir na sua capacidade de imitar o processo de decisão humano, especialmente quando aplicado no campo da medicina”.

No contexto desse trabalho, a ideia por trás do sistema é transferir, para o computador, o conhecimento de diagnosticar doenças. O médico, nesse caso, consultaria esse sistema informando alguns dados para obter um diagnóstico. O sistema, então, avaliaria os dados utilizando seu conhecimento armazenado e atualizando os diagnósticos que contêm os dados em questão, considerando tanto diagnósticos que permanecem como indeterminados quanto diagnósticos convergidos. Caso solicitado, explicaria as decisões tomadas para cada diagnóstico, assim como um médico o faria.

Conforme mencionado por Liao (2005), existem algumas abordagens diferentes desse tipo de sistema e suas aplicações, sendo algumas delas:

- **Baseado em Conhecimento (*Knowledge-Based*):** Surgiram na tentativa de entender o conhecimento humano e traduzi-lo para o computador. Engloba todos os outros tipos de sistemas especialistas que contêm uma base de conhecimento. Exemplos de aplicações: tratamento médico (ALONSO-AMO *et al.*, 1995), apoio à decisão (O'KEEFE; PREECE, 1996) e gerenciamento estratégico (VOLBERDA; RUTGES, 1999).
- **Baseado em Regras (*Rule-Based*):** Representa as informações na forma de *SE-ENTÃO*. Ao receber uma informação o sistema checa sua base de conhecimento para saber qual regra aplicar para chegar a uma conclusão. Exemplos de aplicações: planejamento de agricultura (PLANT; VAYSSIÈRES, 2000), planejamento de produção (HAMADA *et al.*, 1995) e controle de sensores (VALENZUELA; BENTLEY; LORENZ, 2004).
- **Nebuloso (*Fuzzy*):** Utiliza lógica nebulosa que lida com incerteza. É uma abordagem para casos onde a questão não envolva apenas verdadeiro ou falso, podendo também existir o talvez. Exemplos de aplicações: diagnóstico médico (MEESAD; YEN, 2003), agendamento online (CHANG; THIA, 1996) e casamento de ofertas de emprego (DRIGAS *et al.*, 2004).
- **Orientado a Objetos (*Object-Oriented*):** Adiciona aos objetos de dados os procedimentos relativos à aquele dado. Ao chamar o mesmo procedimento para objetos de dados diferentes, a forma como esse mesmo procedimento é executado pode ser diferente. Exemplos de aplicações: representação de conhecimento (VRANES; STANOJEVIC, 1995), diagnóstico industrial (BATANOV; CHENG, 1995) e gerenciamento de sistemas de energia (KAWAHARA *et al.*, 1998).
- **Baseado em Casos (*Case-Based*):** A ideia é adaptar soluções já conhecidas de casos passados para solucionar novos casos. O sistema resgata em sua base de conhecimento um caso antigo que tenha parâmetros parecidos com o

do caso novo. Exemplos de aplicações: planejamento médico (ABIDI; MANICKAM, 2002), diagnóstico de falha (YANG; HAN; KIM, 2004) e aplicação médica (MONTANI; BELLAZZI, 2002).

De acordo com o Tunmibi *et al.* (2013), um dos primeiros sistemas especialistas foi o DENDRAL, desenvolvido na Universidade de Stanford em 1967, que deduzia a estrutura molecular química de uma substância orgânica a partir de sua fórmula. Segundo Lindsay *et al.* (1993), o DENDRAL foi o primeiro sistema baseado em regras aplicado a um problema do mundo real. Em seguida, foi desenvolvido o sistema INTERNIST na Universidade de Pittsburgh em 1974, um sistema especialista baseado em regras para o diagnóstico de problemas complexos em medicina interna geral, que é uma especialidade generalista focada em pacientes hospitalizados e doenças complexas em adultos. Esse sistema abrangia 80% do conhecimento da medicina interna na época, mas foi criticado pela superficialidade de seu conhecimento. De acordo com Miller *et al.* (1986) que realizou um relatório de status sobre o INTERNIST, o sistema desempenhava como clínicos que atendiam pacientes em um hospital universitário, mas a habilidade do sistema de chegar ao diagnóstico correto não era tão boa quanto a dos especialistas. A conclusão do estudo foi que existiam algumas deficiências para que o sistema pudesse ser usado como um consultor generalista de diagnóstico, sendo a principal delas a base de conhecimento incompleta.

2.2. REPRESENTAÇÃO DE DIAGNÓSTICOS

Estudos como o do An, Marwaha e Laxer (2024) representam o diagnóstico através da análise clínica, considerando a prevalência e a frequência dos sintomas na doença em questão, como podemos ver na Tabela 1. Já Gattorno *et al.* (2019) representam o diagnóstico através de uma estrutura lógica (usando *E*, *Ou* e *Ao Menos*) com os dados clínicos, conforme pode ser visto na Tabela 2. Percebe-se, nessa última forma, que o texto dos diagnósticos se assemelha à estrutura de uma árvore lógica. Por exemplo, para a doença *Familial Mediterranean Fever (FMF)*, o diagnóstico pode ser representado conforme a Figura 1.

Como ilustrado na Figura 1, o diagnóstico dessa árvore do FMF pode ser lido como “*ou o paciente tem a presença confirmatória do MEFV e tem ao menos 1 dos sintomas listados, ou, não tem a presença confirmatória do MEFV e tem ao menos 2 dos sintomas listados*”. De acordo com Gattorno *et al.* (2019), o diagnóstico dessa doença, conforme essa expressão lógica, tem sensibilidade de 0,94, especificidade de 0,95 e acurácia de 0,98. Dessa forma, é possível não só representar os diagnósticos fornecidos por Gattorno *et al.* (2019), como também seria possível representar outros diagnósticos nesse padrão, se beneficiando da capacidade preditiva que essa classificação fornece.

Tabela 1 – Mapa de calor de algumas doenças auto-inflamatórias. Traduzido de: An, Marwaha e Laxer (2024).

Categoria	Condição Autoinflamatória	Gene	Febre	Articulação	Exantema	Úlcera	Olhos	GI	SNC	Ossos	Outros
Inflamassomopatias	CAPS	<i>NLRP3</i>	X		X		X		X	X	
Inflamassomopatias	FMF	<i>MEFV</i>	X	X	X			X			
Outras	TRAPS	<i>TNFRSF1A</i>	X	X	X		X	X			
Deficiências enzimáticas	MKD	<i>MVK</i>	X	X	X	X		X			IgD elevada

Observação: Algumas colunas de sintomas foram removidas em relação ao original para melhor redimensionar a tabela, já que não faziam parte das condições autoinflamatórias listadas.

Abreviações: CAPS: Síndromes Periódicas Associadas à Criopirina; FMF: Febre Familiar do Mediterrâneo; TRAPS: Síndrome Periódica Associada ao Receptor do Fator de Necrose Tumoral; MKD: Deficiência de Mevalonato Quinase; GI: Trato Gastrointestinal; SNC: Sistema Nervoso Central; IgD: Imunoglobulina D.

Tabela 2 – Novos critérios de classificação Eurofever/PRINTO para febres hereditárias recorrentes e seu desempenho em 281 pacientes com consenso. Traduzido de: Gattorno *et al.* (2019).

CAPS	FMF	TRAPS	MKD
<p>Presença de um <i>genótipo NLRP3</i> confirmatório* e ao menos um dos seguintes:</p> <ul style="list-style-type: none"> - Exantema urticariforme. - Olho vermelho (conjuntivite, episclerite, uveíte). - Perda auditiva neurosensorial. <p>OU</p> <p>Presença de <i>genótipo NLRP3</i> não confirmatório† e ao menos dois dos seguintes:</p> <ul style="list-style-type: none"> - Exantema urticariforme. - Olho vermelho (conjuntivite, episclerite, uveíte). - Perda auditiva neurosensorial. 	<p>Presença de um <i>genótipo MEFV</i> confirmatório* e ao menos um dos seguintes:</p> <ul style="list-style-type: none"> - Duração dos episódios 1–3 dias. - Artrite. - Dor torácica. - Dor abdominal. <p>OU</p> <p>Presença de <i>genótipo MEFV</i> não confirmatório‡ e ao menos dois dos seguintes:</p> <ul style="list-style-type: none"> - Duração dos episódios 1–3 dias. - Artrite. - Dor torácica. - Dor abdominal. 	<p>Presença de um <i>genótipo TNFRSF1A</i> confirmatório* e ao menos um dos seguintes:</p> <ul style="list-style-type: none"> - Duração dos episódios ≥ 7 dias. - Mialgia. - Exantema migratório. - Edema periorbitário. - Parentes afetados. <p>OU</p> <p>Presença de <i>genótipo TNFRSF1A</i> não confirmatório† e ao menos dois dos seguintes:</p> <ul style="list-style-type: none"> - Duração dos episódios ≥ 7 dias. - Mialgia. - Exantema migratório. - Edema periorbitário. - Parentes afetados. 	<p>Presença de um <i>genótipo MVK</i> confirmatório* e ao menos um dos seguintes:</p> <ul style="list-style-type: none"> - Sintomas gastrointestinais. - Linfadenite cervical. - Estomatite aftosa.
Sensibilidade: 1	Sensibilidade: 0,94	Sensibilidade: 0,95	Sensibilidade: 0,98
Especificidade: 1	Especificidade: 0,95	Especificidade: 0,99	Especificidade: 1
Acurácia: 1	Acurácia: 0,98	Acurácia: 0,99	Acurácia: 1

Observação: Um paciente com (1) evidência de elevação de reagentes de fase aguda (VHS, PCR ou SAA) correspondente aos surtos clínicos e (2) avaliação cuidadosa de possíveis doenças confundidoras (neoplasias, infecções, condições autoimunes, outros erros inatos da imunidade) e um período razoável de atividade recorrente da doença (pelo menos 6 meses) é classificado como portador de febre hereditária recorrente se os critérios forem atendidos.

Símbolos: * Variantes patogênicas ou provavelmente patogênicas (heterozigose em doenças autossômicas dominantes, homozigose ou em trans (ou bialélicas) composto heterozigoto em doenças autossômicas recessivas).

† Variante de Significado Incerto (VUS). Variantes benignas e provavelmente benignas devem ser excluídas.

‡ Composto heterozigoto em trans para uma variante MEFV patogênica e uma VUS, ou bialélica VUS, ou heterozigose para uma variante MEFV patogênica. Ver tabela suplementar online 7 para glossário.

Abreviações: AD, Autossômico Dominante; AR, Autossômico Recessivo; CAPS, Síndromes Periódicas Associadas à Criopirina; PCR, Proteína C Reativa; VHS, Velocidade de Hemossedimentação; FMF, Febre Familiar do Mediterrâneo; MKD, Deficiência de Mevalonato Quinase; MVK, Mevalonato Quinase; PRINTO, Organização Internacional de Ensaios em Reumatologia Pediátrica; SAA, Amiloide Sérico A; TRAPS, Síndrome Periódica Associada ao Receptor do Fator de Necrose Tumoral.

3. TRABALHOS RELACIONADOS

Um estudo realizado por Schaaf *et al.* (2020) identificou 19 estudos sobre Sistemas de Apoio à Decisão Médica (CDSSs) diferentes, dos quais 12 utilizavam análise ou comparação de dados genéticos e fenotípicos. Os 7 CDSSs restantes utilizavam ou aprendizado de máquina ou recuperação de informação. Apesar de

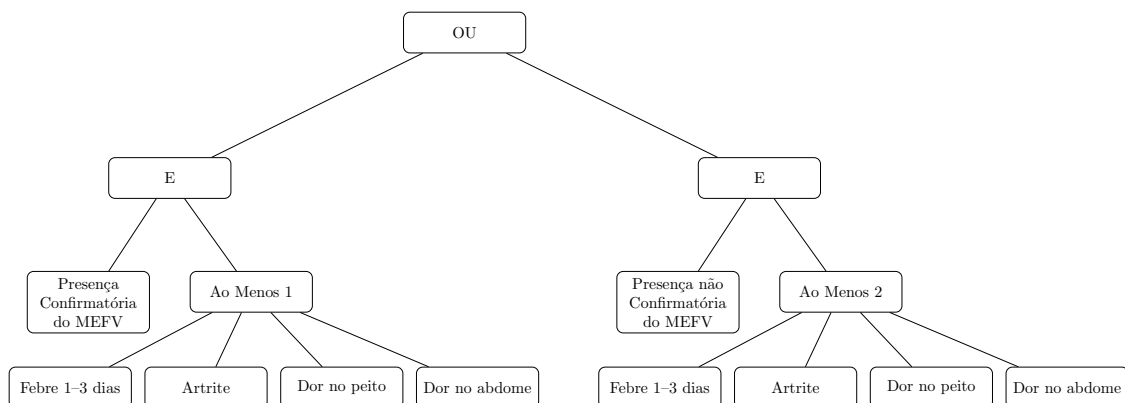


Figura 1 – Árvore de decisão para representar o diagnóstico de FMF, com base na presença de variantes confirmatórias do gene MEFV e em sintomas clínicos.

alguns estudos focarem em uma doença específica ou grupo de doenças, os 12 estudos que utilizavam análise ou comparação cobrem todo o espectro das doenças raras.

Outro estudo realizado por Faviez *et al.* (2020) identificou 61 estudos, dos quais 19 utilizavam alguma base de conhecimento. Os 44 restantes utilizavam ou aprendizado de máquina ou um híbrido entre os dois. Desses 19 que utilizavam base de conhecimento, 5 sistemas eram de árvores de decisão feitas manualmente. Nos 14 sistemas restantes, os diagnósticos eram representados por conceitos fenotípicos que correspondiam aos sinais e sintomas da doença, sendo que, desses 14, 3 sistemas incluíam conhecimento genético.

Desses sistemas, alguns não estavam mais disponíveis, outros incluíam informações genéticas, mas não contemplavam todas as doenças ou vice-versa. Dos sistemas que exibiam um exemplo da interface ou do funcionamento, não foi observado o uso de uma expressão lógica como obtida na tabela do estudo levantado por Gattorno *et al.* (2019).

Há também estudos de sistemas específicos. Um exemplo é o estudo realizado por Ronicke *et al.* (2019) a respeito do Ada DX, um CDSS desenvolvido como um protótipo de pesquisa pela empresa *Ada Health*. O sistema permite que os médicos construam e acompanhem o caso do paciente ao longo de suas consultas, adicionando ou removendo “achados” (sintomas, sinais, resultados de exame) como presentes ou ausentes. O sistema fornece uma lista de sugestões de possíveis achados, ordenada por sua relevância em relação ao conjunto de achados atual, sendo atualizada a cada adição ou remoção de um achado. Baseado na entrada de achados do paciente, o sistema gera duas listas ranqueadas, uma de probabilidade, que calcula estimativas levando em conta idade, região e outros fatores, e outra de *fit*, que ignora a probabilidade e considera apenas os achados adicionados. Ambas as listas incluem tanto doenças comuns quanto doenças raras. A Figura 2 mostra um exemplo do funcionamento desse sistema. Segundo Ronicke *et al.* (2019), dentro dos 93 casos

estudados, o sistema sugeriu a doença correta na primeira consulta do paciente, entre as 5 melhores sugestões, em um terço dos casos, e como principal sugestão em aproximadamente um sexto dos casos. Entretanto, sua base de conhecimento não é baseada em nenhum conteúdo de ontologia médica publicamente disponível, nem em bases de dados pré-existentes.

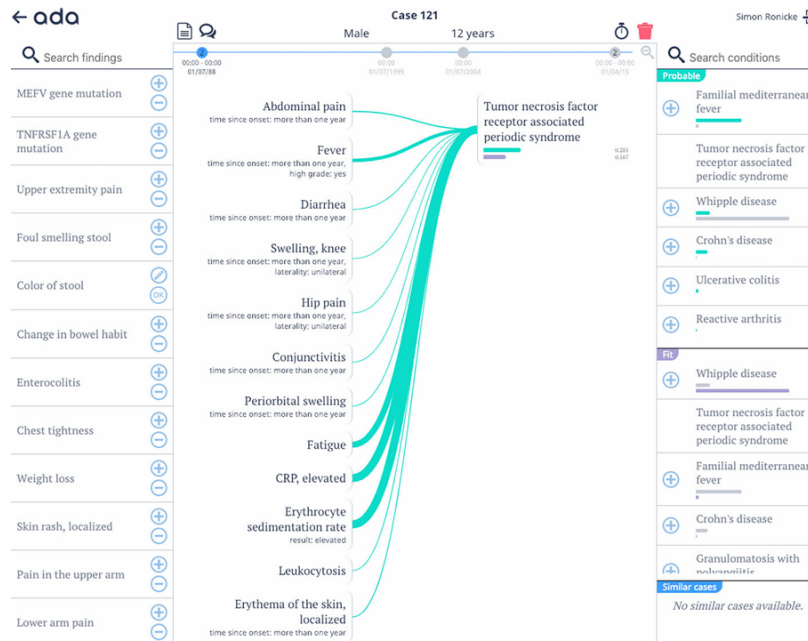


Figura 2 – Captura de tela do protótipo do Ada DX mostrando o caso de um paciente com a doença TRAPS. Fonte: Ronicke *et al.* (2019)

Com o avanço da inteligência artificial na última década e a popularização do uso das IAs generativas, foram levantados alguns estudos sobre o potencial dessas ferramentas para o diagnóstico médico. Zelin *et al.* (2024) menciona que o RareDxGPT, seu modelo do ChatGPT 3.5 aprimorado com o data set do RareDis Corpus através de Geração Aumentada via Recuperação (RAG), conseguiu obter uma acurácia de 43% nos diagnósticos. Já Mehnen *et al.* (2023) realizou uma avaliação da utilização do ChatGPT 4, informando o caso clínico e pedindo para ele sugerir os 10 diagnósticos mais prováveis para o paciente em questão. No conjunto das 10 doenças raras utilizadas na avaliação, 40% foram corretamente diagnosticadas no 1º diagnóstico dos 10 mais prováveis e que, ao considerar até os 8 primeiros dos 10, a acurácia sobe para 90% como ilustrado na Figura 3.

Curiosos com a possibilidade, realizamos um teste simples com o ChatGPT 4.1 (atual versão paga mais recente), desativando a memória dos chats e sem personalizar com instruções, perguntando tanto em inglês quanto em português. A tentativa é de obter, dados os sintomas, possíveis diagnósticos ou sugestões de exames genéticos que seriam interessantes, caso ele não fornecesse o diagnóstico correto na primeira tentativa. Para o FMF, como mostrado na Tabela 2, ele indicou o diagnóstico esperado quando fornecido o prompt da Figura 4.

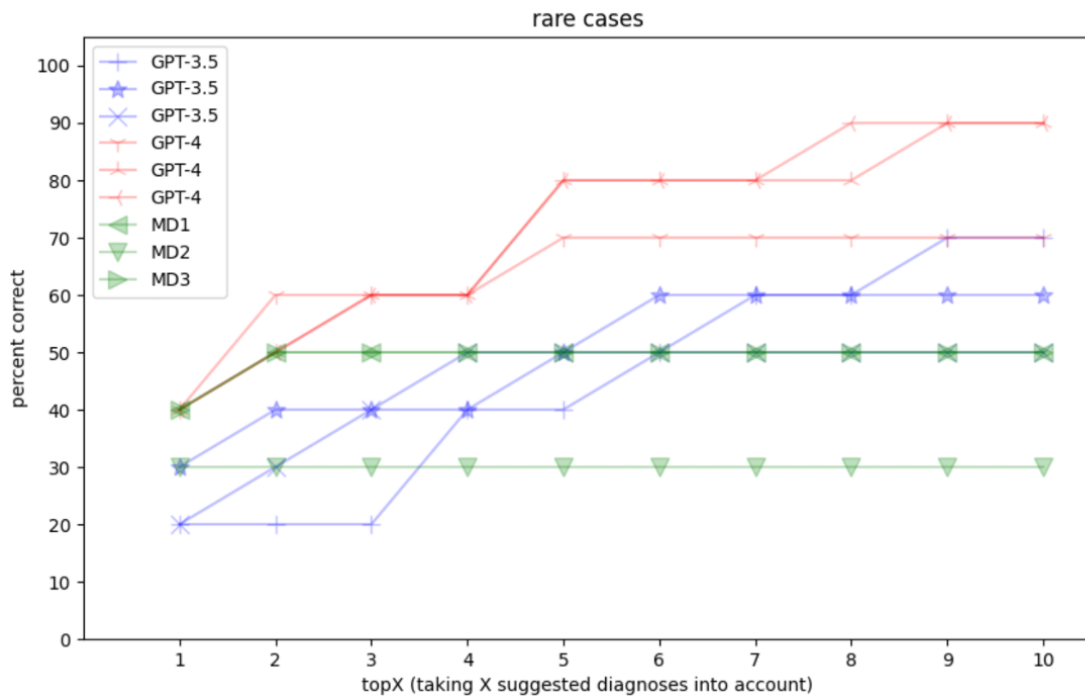


Figura 3 – Acurácia do diagnóstico do ChatGPT 4 e 3 médicos em 10 doenças raras, permitindo um máximo de 10 sugestões de diagnósticos mais prováveis. Fonte: Mehnen *et al.* (2023)

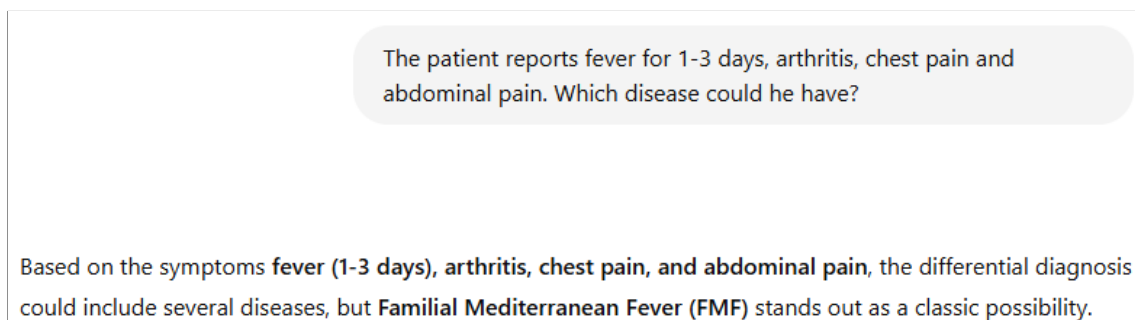


Figura 4 – Prompt em inglês para verificar o possível diagnóstico do FMF dado os sintomas, usando a versão paga do ChatGPT 4.1.

Por outro lado, para o CAPS, também presente na Tabela 2, ele não retornou o diagnóstico correto na primeira tentativa, conforme pode-se observar na Figura 5. Entretanto, ao perguntar por um gene interessante para se fazer um teste genético, ele sugeriu o gene que poderia identificar o CAPS, como ilustrado na Figura 6

Variamos os prompts para analisar a constância e, ao testar os prompts em português, o FMF foi diagnosticado como esperado, porém para o CAPS foi sugerido outro diagnóstico, como ilustrado na Figura 7. Ao perguntar por um gene interessante para se fazer um exame, nesse caso, não foi mencionado o NLRP3 do CAPS, como mostra a Figura 8. Além do mais, ao variar o final da pergunta de “Qual seria um possível diagnóstico” para “Qual doença ele pode ter”, conforme exibido na Figura 9, vemos que foi sugerida outra doença diferente da primeira.

Esses testes, apesar de informais, reforçam o grande potencial latente do uso

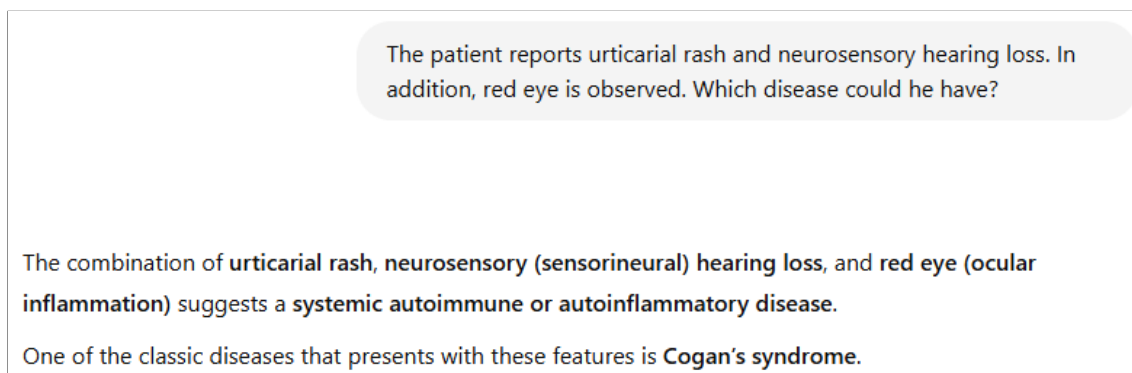


Figura 5 – Prompt em inglês para verificar o possível diagnóstico do CAPS dado os sintomas, usando a versão paga do ChatGPT 4.1.

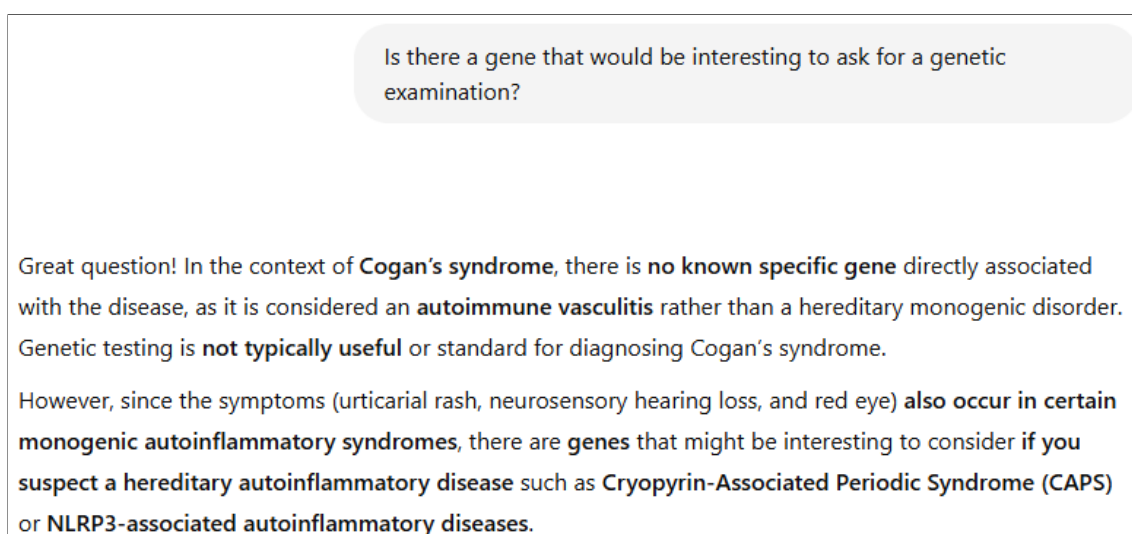


Figura 6 – Prompt em inglês para verificar a sugestão de exame genético para diagnosticar o CAPS, usando a versão paga do ChatGPT 4.1.

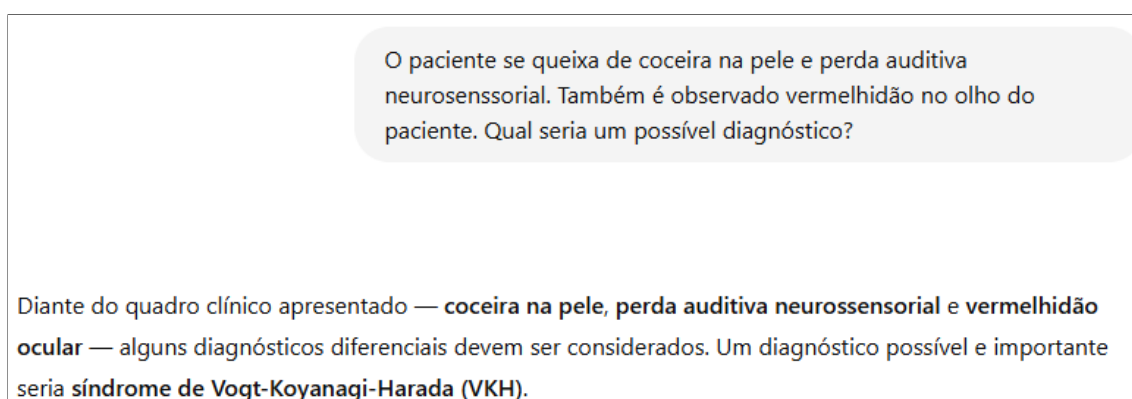


Figura 7 – Prompt em português para verificar o possível diagnóstico do CAPS dado os sintomas, usando a versão paga do ChatGPT 4.1.

das IAs. Entretanto, também evidenciam o fato de que a IA, nesse caso o ChatGPT 4.1, não é determinística. Considerando uma área tão crítica como a da saúde, é necessário que qualquer sistema de apoio a diagnósticos possua uma abordagem determinística e explicável em que os médicos possam embasar as suas decisões.

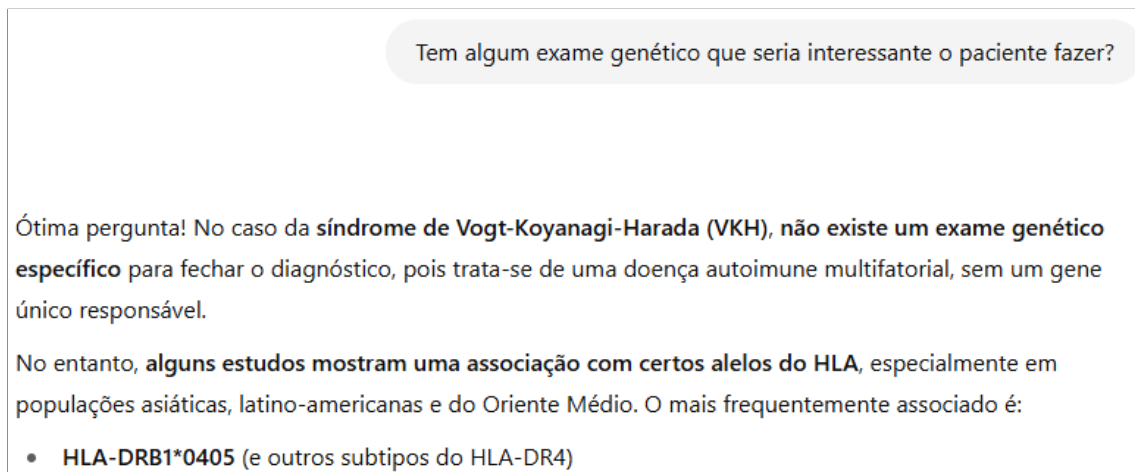


Figura 8 – Prompt em português para verificar a sugestão de exame genético para diagnosticar o CAPS, usando a versão paga do ChatGPT 4.1.

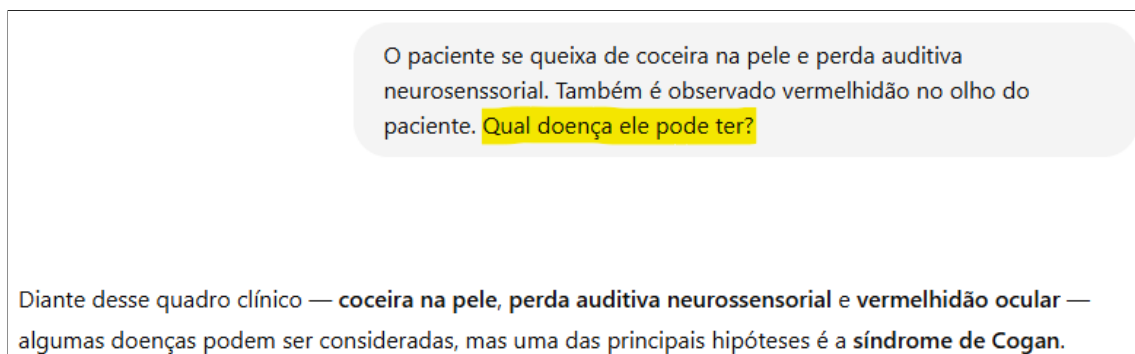


Figura 9 – Prompt em português com pequena alteração para verificar o possível diagnóstico do CAPS dado os sintomas, usando a versão paga do ChatGPT 4.1.

4. DISEASEDX

Nesta seção, são explicados os conceitos, os componentes e a proposta do trabalho, que chamamos de DiseaseDx. A Subseção 4.1 fornece uma visão geral do sistema antes de detalhar seus principais conceitos e componentes. A Subseção 4.2 trata do modelo do sistema através de um diagrama de classes simplificado para a contextualização das principais entidades do sistema e seus papéis. A Subseção 4.3 fala sobre a lógica tri-valorada que será utilizada para calcular e verificar o estado das expressões lógicas que compõem o diagnóstico de uma doença. A Subseção 4.4 explica o raciocínio e a lógica utilizados para estabelecer pontuações (*scores*) para os diagnósticos, a fim de criar uma ordenação por relevância, com os diagnósticos mais promissores aparecendo primeiro. A Subseção 4.5 apresenta o funcionamento do sistema, combinando os componentes das subseções anteriores. Por fim, a Subseção 4.6 explica as ferramentas e conceitos mais técnicos utilizados para a implementação do sistema.

4.1. VISÃO GERAL

Conforme ilustrado na Figura 10, o DiseaseDx possui uma base de conhecimento, que armazena as fórmulas lógicas dos diagnósticos (envolvendo sintomas e resultados de exames) e uma interface gráfica onde o médico informa sintomas e resultados presentes e ausentes. Com base nessas informações, o DiseaseDx indica os diagnósticos mais prováveis de acordo com o resultado da avaliação das expressões lógicas dos diagnósticos e de um sistema de pontuação.

A base de conhecimento do DiseaseDx é pré-populada com os diagnósticos obtidos na Tabela 2, mas outros diagnósticos podem ser incluídos pelo médico especialista. Atualmente, esse cadastro é feito diretamente pelo código que gera a base pré-populada, pois, apesar do código que permite esse cadastro estar parcialmente implementado, a interface acabou sendo despriorizada em detrimento de outras funcionalidades que ainda restavam implementar. Ao iniciar o sistema, essa base é lida e as árvores dos diagnósticos são exibidas e atualizadas em tempo real conforme os dados de entrada (sintomas e resultados de exames) vão sendo preenchidos como presentes ou ausentes. Essa atualização ocorre pela avaliação das expressões contidas na base de conhecimento, resultante do modelo apresentado na Seção 4.2, junto com os fatos obtidos de entrada, utilizando uma lógica tri-valorada para lidar com a ideia do “desconhecido”, explicada na Seção 4.3.

Junto com essa atualização, é utilizada uma estratégia de pontuação que é abordada na Subseção 4.4, fornecendo ao médico a explicabilidade do porquê de algum diagnóstico estar se destacando dos demais. Além disso, o sistema também identifica nos diagnósticos a frequência dos sintomas e resultados de exame, sugerindo os de maior frequência como os mais comuns. A Figura 10 fornece uma visão geral do sistema com seus principais atores e componentes.

4.2. MODELO

Conforme previamente mencionado na Subseção 2.2, o diagnóstico das doenças pode ser descrito por uma árvore que representa uma expressão lógica envolvendo *Sintomas*, *Resultados*, e os operadores lógicos *E*, *Ou* e *Ao Menos*. Para a FMF, por exemplo, o diagnóstico pode ser representado pela expressão “*ou o paciente tem a presença confirmatória do MEFV e tem ao menos 1 dos sintomas listados, ou, não tem a presença confirmatória do MEFV e tem ao menos 2 dos sintomas listados*”.

O modelo do DiseaseDx representa essas expressões lógicas como pode ser observado de forma simplificada na Figura 11, onde as classes *Sintoma*, *Resultado*, *E*, *Ou* e *Ao Menos* implementam a interface *Expressão*, fazendo com que a expressão atrelada a um diagnóstico seja a combinação de objetos dessas classes e permitindo a construção da árvore que representa a expressão lógica do diagnóstico.

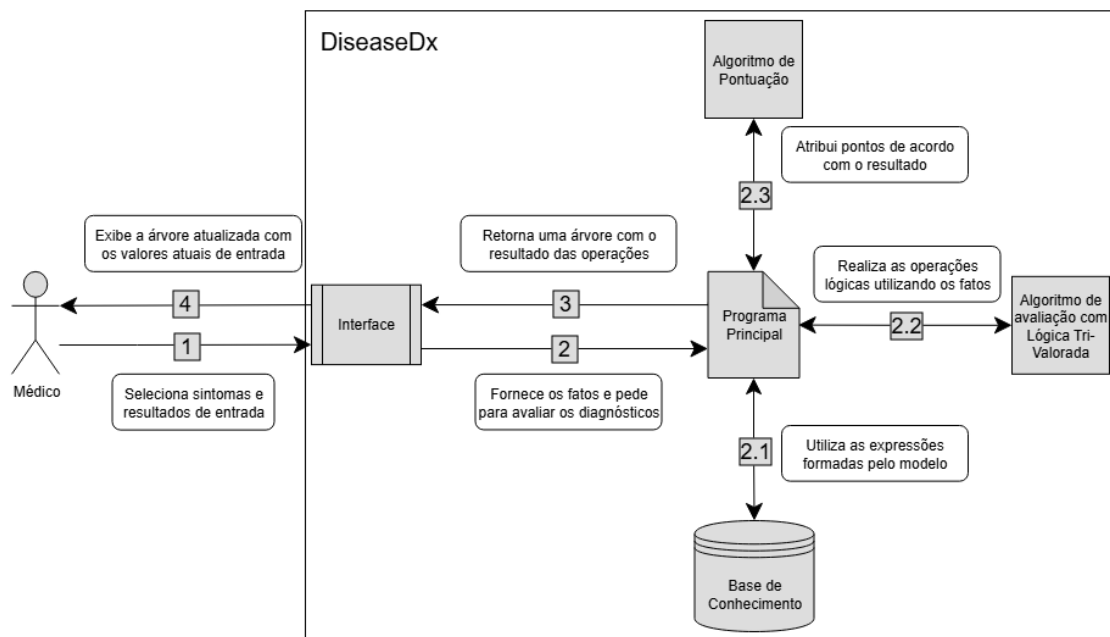


Figura 10 – Visão geral do DiseaseDx.

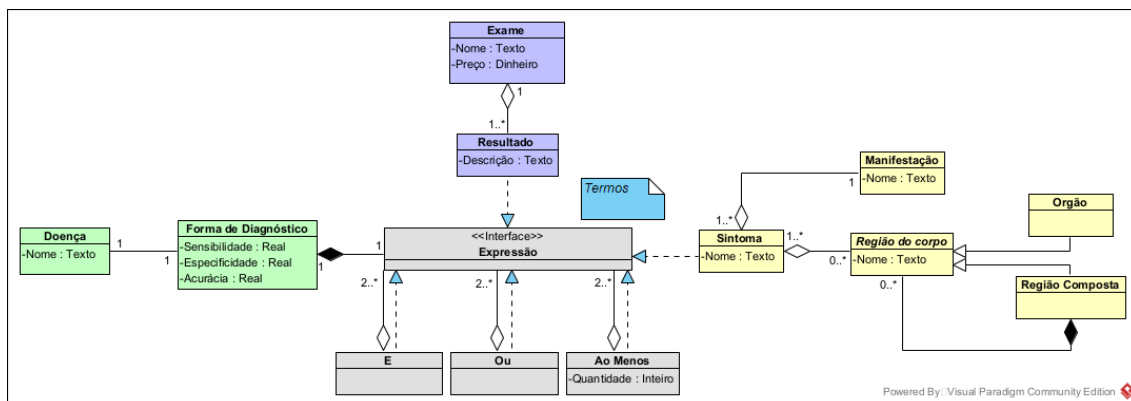


Figura 11 – Diagrama de classes simplificado do DiseaseDx.

Na parte cinza da Figura 11, temos as classes relativas à expressão lógica, com adição das classes *Resultado* e *Sintoma*. A classe *E* agrega duas ou mais *Expressões* quaisquer, e é verdadeira somente quando todas as suas *Expressões* são verdadeiras. De forma análoga, a classe *Ou* somente é verdadeira se uma das suas *Expressões* for verdadeira. A classe *Ao Menos* é uma generalização da classe *Ou*, onde é possível indicar o número mínimo de *Expressões* verdadeiras para que o resultado seja verdadeiro. Ou seja, *Ou* equivale ao *Ao Menos* quando a quantidade é 1.

A parte roxa da Figura 11 representa as classes relativas aos resultados de exames. A classe *Exame* agrega um ou mais *Resultados*, da mesma forma que um ou mais *Resultados* são agregados por um *Exame*. Portanto, um *Exame* genético para o gene MEFV, por exemplo, pode ter como *Resultado* tanto a presença do gene quanto alguma variante de significado desconhecido.

Na parte amarela da Figura 11, temos as classes relativas aos sintomas. A classe *Sintoma* agrega uma *Manifestação* e também agrega zero ou mais *Regiões do Corpo*, tornando possível a criação de sintomas como “febre”, que é uma manifestação, mas sem uma região do corpo específica. Tanto a classe *Manifestação* quanto a classe *Região do Corpo* são agregadas por um ou mais sintomas, permitindo que uma manifestação “dor”, por exemplo, possa fazer parte de diversos sintomas como “dor no peito” e “dor no abdome”, assim como também permite que uma região do corpo como “braço”, por exemplo, possa fazer parte de diversos sintomas como “dor no braço” e “formigamento no braço”. Por fim, as classes *Orgão* e *Região Composta* permitem criar regiões dentro de regiões, pois *Região Composta* é composta por zero ou mais *Regiões do Corpo*, que podem ser tanto *Orgãos* quanto outras *Regiões Compostas*.

Por fim, na parte verde da Figura 11, temos as classes relativas aos diagnósticos das doenças. A classe *Doença* tem uma *Forma de Diagnóstico* assim como se tem a *Forma de Diagnóstico* de uma *Doença*. A classe *Forma de Diagnóstico* é composta por uma *Expressão*, assim como uma *Expressão* compõe uma *Forma de Diagnóstico*. Portanto, a doença FMF, por exemplo, tem uma forma de diagnóstico que é composta por uma expressão, sendo essa expressão uma combinação de todas as classes que implementam a interface *Expressão*, mostradas na parte cinza da Figura 11, com adição das classes *Resultado* e *Sintoma* que também a implementam. Apesar do foco deste trabalho ser nas doenças raras, essas classes previamente mencionadas possuem uma estrutura genérica, permitindo tanto o cadastro de doenças raras quanto de doenças comuns.

Um exemplo de instância dessas classes pode ser observado através do diagrama de objetos da Figura 12, onde é representado o diagnóstico da FMF como visto na Tabela 2. As cores dos objetos nesse diagrama correspondem às cores no diagrama de classes, mostrando, através desses objetos, as relações previamente discutidas no diagrama de classes. É importante ressaltar também que na árvore formada para representar a expressão lógica de um diagnóstico, os objetos de *Sintomas* e *Resultados* são sempre nós-folhas.

4.3. LÓGICA TRI-VALORADA

Uma lógica tri-valorada é uma lógica que possui 3 valores, geralmente verdadeiro, falso e indeterminado (NAVATHE; ELMASRI, 2015). Ela é útil para modelar problemas que lidam com dados que nem sempre são apenas verdadeiros ou falsos, como é o caso, por exemplo, da análise de sintomas de um paciente. Até o momento de ser conhecido o fato de o paciente ter ou não um sintoma, o sintoma permanece em um estado indeterminado, pois não se têm informações ainda sobre ele. O poder que essa lógica traz é o de permitir realizar operações lógicas sobre esses valores.

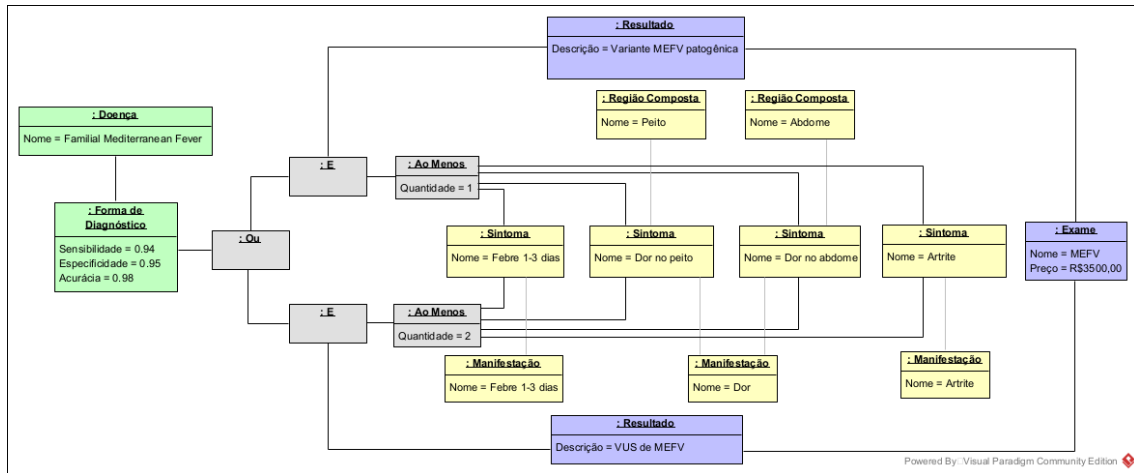


Figura 12 – Exemplo de diagrama de objetos do DiseaseDx para a doença FMF.

Tabela 3 – Tabela verdade de lógica tri-valorada. *Not* ocultado pois sintomas e resultados ausentes vão para uma lista de ausente ao invés de serem a negação do presente.

A	B	A e B	A ou B
Verdadeiro	Verdadeiro	Verdadeiro	Verdadeiro
Verdadeiro	Falso	Falso	Verdadeiro
Verdadeiro	Indeterminado	Indeterminado	Verdadeiro
Falso	Verdadeiro	Falso	Verdadeiro
Falso	Falso	Falso	Falso
Falso	Indeterminado	Falso	Indeterminado
Indeterminado	Verdadeiro	Indeterminado	Verdadeiro
Indeterminado	Falso	Falso	Indeterminado
Indeterminado	Indeterminado	Indeterminado	Indeterminado

No contexto desse trabalho, essas expressões lógicas envolvem os operadores *E*, *Ou* e *Ao Menos* obtidos da representação do diagnóstico da Tabela 2. O *Ao Menos* é uma expressão que combina a expressão do *E* com o *Ou*, pois para um *Ao Menos 2* de um conjunto $[A, B, C]$ ser verdadeiro é avaliado se ' $A \text{ e } B \text{ ou } A \text{ e } C \text{ ou } B \text{ e } C$ ' e, portanto, só aparecem na tabela verdade as operações de *E* e *Ou*. A Tabela 3 mostra a tabela verdade para essa lógica tri-valorada.

4.4. PONTUAÇÃO DOS DIAGNÓSTICOS

Para aperfeiçoar o destaque dos diagnósticos promissores, foi adotada uma estratégia de pontuação (*score*). A ideia é que diagnósticos com sintomas ou resultados presentes se destaquem, aparecendo primeiro na lista de árvores de avaliação devido a uma maior pontuação em relação aos demais. Em contrapartida, diagnósticos com sintomas ou resultados ausentes vão para o final da fila, devido a uma menor pontuação.

Para atribuir essa pontuação a um termo, antes é preciso saber o estado dele,

se é verdadeiro, falso ou indeterminado, e atribuir um valor de acordo com esse estado. Esse estado é calculado utilizando a lógica tri-valorada vista na Subseção anterior, através da Tabela 3. Para cada um dos termos, o cálculo da pontuação foi estabelecido a partir do seguinte raciocínio:

- **Sintoma:** +1 para verdadeiro, -1 para falso e 0 para indeterminado. Para o indeterminado foi escolhido o 0 pois, até ter como entrada algum fato sobre sintoma ou resultado, todos os diagnósticos são indeterminados e com a mesma pontuação de 0. Para o verdadeiro e falso, foram escolhidos o +1 e -1 pois, se um ‘sintoma pertencente a um diagnóstico’ é verdadeiro, sua pontuação incrementa deixando o diagnóstico mais promissor, se for falso, sua pontuação decrementa deixando-o menos promissor.
- **Resultado:** +1 para verdadeiro, -1 para falso e 0 para indeterminado. O raciocínio é o mesmo do Sintoma.
- **E:** Média da pontuação dos seus termos. Foi escolhida a média, já que cada um dos seus termos tem o poder de anular o resultado como um todo, pois um E com n termos onde $n-1$ sejam verdadeiros, basta que o n -ésimo seja falso para fazer com que o E se torne falso. Por exemplo, um E onde seus termos A, B e C tenham respectivamente a pontuação 1, 0 e 0 ficaria com uma pontuação de 0,33, já que, apesar de A ser verdadeiro, basta B ou C serem falsos para tornar o E falso. Uma alternativa seria usar a pontuação do menor dos seus termos, porém essa alternativa seria muito conservadora e não conseguiria refletir bem a evolução do preenchimento do E .
- **Ou:** Maior valor entre a pontuação dos seus termos. Foi escolhido o maior dos seus termos, já que basta que um dos seus termos seja verdadeiro para a expressão se tornar verdadeira, pois um Ou com n termos onde $n-1$ sejam falsos, se o n -ésimo termo for verdadeiro, o Ou se torna verdadeiro. Por exemplo, um Ou onde seus termos A e B tenham, respectivamente, a pontuação -1 e 0,33, ficaria com uma pontuação de 0,33. Um bom exemplo da motivação desse raciocínio é que, uma árvore de diagnóstico com o Ou como raiz pode ter vários ramos falsos, porém, se existe um ramo que ainda não é falso, o diagnóstico ainda é possível e esse ramo vai ter uma pontuação maior sobre os demais.
- **Ao Menos n:** Média das n maiores pontuações dos seus termos. Para a pontuação desse operador, foi escolhida a lógica do E e Ou combinadas, pois avaliar um *Ao Menos 2* de uma lista de sintomas ‘[A, B, C]’, por exemplo, seria equivalente a avaliar se ‘A e B ou A e C ou B e C’. Além disso, quando

a quantidade do *Ao Menos* é 1, sendo equivalente a um *Ou*, o critério de pontuação segue consistente com o do *Ou*.

4.5. FUNCIONAMENTO

Inicialmente, sem ter informação de sintomas ou resultados de exames de um paciente sendo avaliado, não podemos nem afirmar nem negar a possibilidade do diagnóstico, o que pode ser representado pela Figura 1 mostrada na Subseção 2.2. Para lidar com essas possibilidades e incertezas dos sintomas e resultados desconhecidos até o momento da análise clínica do paciente, é utilizada a lógica tri-valorada como foi vista na Subseção 4.3.

Se o médico, ao interagir com o paciente, descobre que é fato que o paciente teve febre por 1 a 3 dias, mas não teve artrite, nem dor no peito e nem dor no abdome, pode consultar o sistema e observar as árvores de diagnóstico atualizarem, sendo uma delas a árvore da FMF, como pode ser observado na Figura 13.

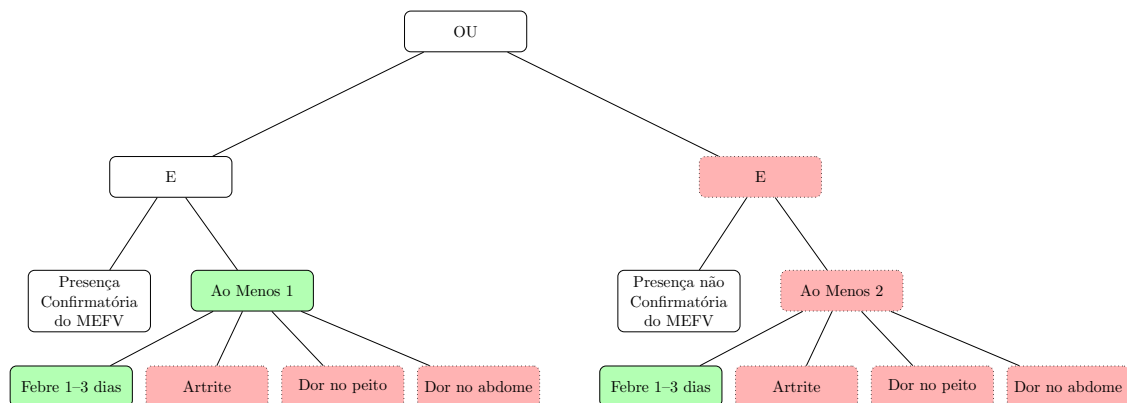


Figura 13 – Árvore de decisão para representar o diagnóstico de FMF. Os nós pintados de verde indicam que o nó é verdadeiro e os nós pintados de vermelho indicam que o nó é falso.

Como agora é conhecido que *febre de 1 a 3 dias* é um sintoma presente, pode-se observar que o nó *Ao Menos 1* no lado esquerdo da árvore, conseqüentemente se torna verdadeiro, pois ao menos 1 dos sintomas listados é verdadeiro, conforme descrita pela expressão lógica lembrada no início dessa seção. Em contrapartida, no lado direito, o nó *E* se torna falso, pois o nó *Ao Menos 2* também é falso, dado que não é possível ter ao menos 2 sintomas verdadeiros em uma lista com 4 sintomas onde 3 são conhecidamente falsos.

Acontece que, ainda não é possível determinar se o diagnóstico é verdadeiro ou não, pois, apesar do lado direito da árvore ser falso, o lado esquerdo ainda é uma possibilidade. Se o médico solicitar um exame genético ao paciente para verificar a presença do gene MEFV, e o resultado for a presença confirmatória desse gene, o diagnóstico seria verdadeiro, pois o nó *E* da esquerda da árvore seria verdadeiro e,

consequentemente, o nó raiz *Ou* seria verdadeiro, conforme pode ser observado na Figura 14.

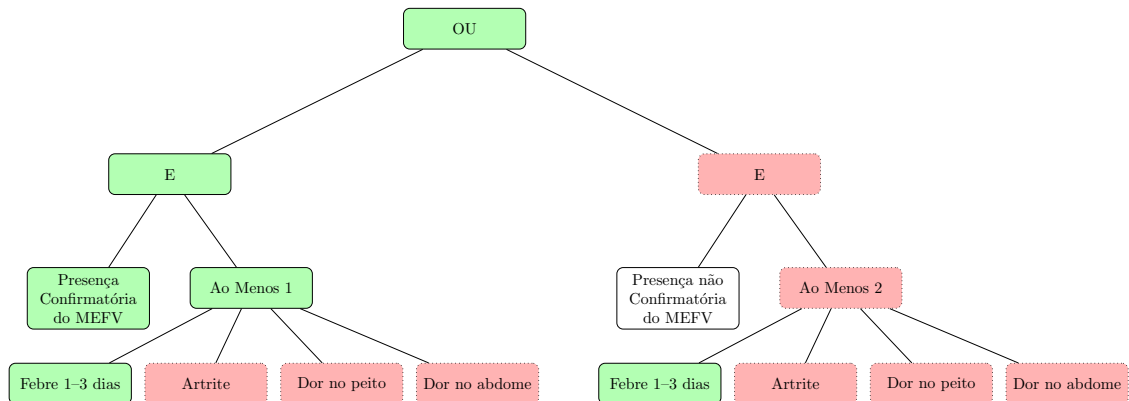


Figura 14 – Árvore de decisão para representar o diagnóstico de FMF. Os nós pintados de verde indicam que o nó é verdadeiro e os nós pintados de vermelho indicam que o nó é falso.

Com isso, o sistema informa ao médico que os sintomas do paciente de febre de 1 a 3 dias e a presença confirmatória do gene MEFV são compatíveis com o diagnóstico da FMF com uma sensibilidade de 0,94, especificidade de 0,95 e acurácia de 0,98, conforme é mostrado na Tabela 2.

4.6. PROJETO E IMPLEMENTAÇÃO

As avaliações dos estados das expressões lógicas ocorrem por meio da adoção do padrão de projeto do *Interpreter* que, de acordo com Gamma *et al.* (1994), é um padrão de projeto comportamental onde, para uma determinada linguagem, define-se uma representação gramatical e um interpretador para ela. Esse interpretador utiliza essa representação da gramática para ler frases da linguagem e conseguir interpretá-las.

Nesse sistema, o interpretador é acionado por meio de uma função chamada *avalia()* pertencente às classes que implementam a interface da classe *Expressão*. Para o interpretador conseguir avaliar essas expressões, é passado para ele um contexto que, nesse caso, é um dicionário de fatos a respeito dos sintomas e resultados de exames. O dicionário tem como conteúdo o estado (verdadeiro, falso ou indeterminado) de cada sintoma e resultado. A função *avalia()*, através do polimorfismo, retorna o resultado da avaliação por um nó *AvaliaNode* que contém sua *expressão*, seu *resultado*, seus *filhos*, sua *classe* e sua *pontuação*. Esse nó *AvaliaNode* é o que serve como estrutura de dados para poder gerar a árvore e iterar sobre ela.

Então, considerando a expressão lógica do FMF previamente mencionada na Seção 2.2, ao chamar o *avalia()* para ela, o sistema vai inicialmente executar o *avalia()* do *Ou*. O *Ou*, por sua vez, antes de realizar qualquer cálculo do resultado de seu estado ou pontuação, vai chamar recursivamente a função *avalia()* para cada um

dos seus filhos, até chegar em um *sintoma* ou *resultado* que são folhas, retornando o estado e a pontuação de acordo com o dicionário de fatos. Tendo, por fim, o estado e a pontuação dos nós de seus filhos, é calculado o estado e a pontuação do *Ou* em questão.

Como exemplo concreto, vamos assumir que os seguintes sintomas e resultados de exames foram indicados pelo paciente como presentes: febre de 1 a 3 dias e variante MEFV patogênica. Além disso, vamos assumir que os seguintes sintomas e exames foram indicados como ausentes: dor no peito e dor no abdome. Ao avaliar a doença FMF, temos que febre e variante MEFV patogênica são verdadeiras (+1 cada) e que dor no peito e dor no abdome são falsos (-1 cada). Com isso, temos que o *Ao Menos 2* é a média das 2 maiores pontuações, logo, o 1 da febre e o 0 da artrite que segue indeterminado, sendo a média entre eles, portanto 0,5. Já o *Ao Menos 1* é o maior entre eles, funcionando como um *Ou* e, portanto, com pontuação de 1 da febre. O *E* da direita fica com pontuação 0,25, dada a média entre o 0,5 do *Ao Menos 2* e o 0 do resultado indeterminado. Já o *E* da esquerda fica com pontuação 1, dada a média entre o 1 do *Ao Menos 1* e o 1 do resultado do exame. O *Ou* por fim, fica com pontuação de 1, pois é o maior valor entre eles, fazendo o diagnóstico da FMF subir na lista. Podemos observar essa árvore com scores na Figura 15.

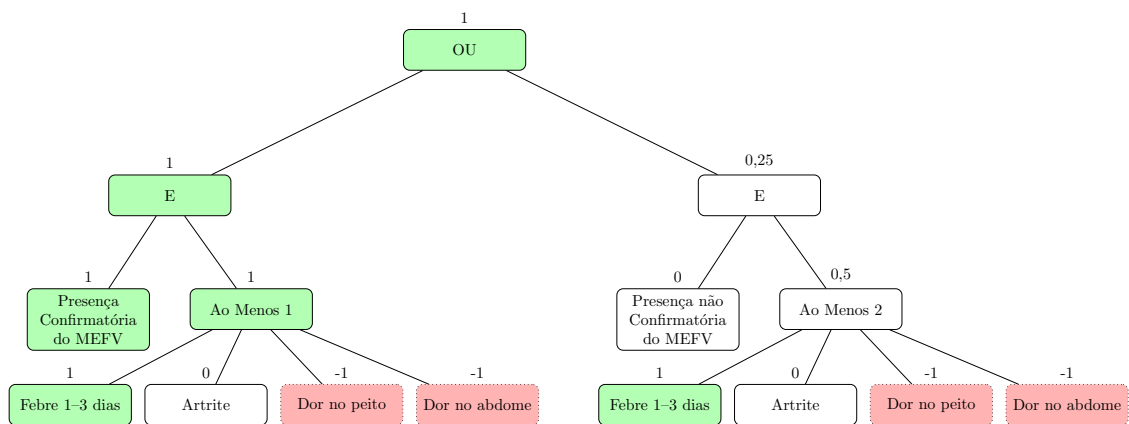


Figura 15 – Árvore de decisão para representar o diagnóstico de FMF. Os nós pintados de verde indicam que o nó é verdadeiro e os nós pintados de vermelho indicam que o nó é falso.

O DiseaseDX foi implementado na linguagem Python, utilizando principalmente as bibliotecas SQL Alchemy, Tribol e Streamlit. Para o mapeamento objeto-relacional, utilizamos o SQL Alchemy. Essa escolha permitiu que os objetos pudessem ser guardados e recuperados da base de dados sem precisar recriá-los, pois, ao recuperar os dados da base, o SQL Alchemy remonta o objeto. Para poder operar com a lógica tri-valorada, utilizamos a biblioteca Tribol. Para a interação com o usuário, utilizamos o Streamlit. Essa escolha foi feita para possibilitar a prototipação rápida, pois essa biblioteca é voltada para aplicação de dados, além de permitir fazer a interface pelo Python.

5. EXEMPLO DE UTILIZAÇÃO

O paciente, ao chegar no consultório médico e ser atendido pelo médico, relata que está se sentindo febril. O médico, por sua vez, além de verificar que o paciente de fato está com febre, percebe que o paciente apresenta vermelhidão nos olhos. Instigado pelos sintomas apresentados, o médico decide abrir o DiseaseDx e consultar pelos sintomas.

Ao abrir a página principal do DiseaseDx, o médico se depara com a tela da Figura 16 onde há 3 campos principais: as entradas de sintomas e resultados de exames presentes (na esquerda da região amarela) e ausentes (na direita da região amarela), a sugestão de sintoma e resultado mais comuns, assim como tabelas auxiliares, na região verde e, por fim, as árvores de avaliação dos diagnósticos na região azul.

The screenshot displays the DiseaseDx interface with the following components:

- Top Section (Yellow Highlight):** Contains four dropdown menus for selecting symptoms and results, categorized as 'presentes' (present) and 'ausentes' (absent).
- Middle Section (Green Highlight):** Displays common suggestions.
 - Sintoma mais comum:** Erupcao no(a) Pele
 - Resultado mais comum:** Variante NLRP3 patogênica
- Tables (Green Highlight):** Two tables showing the association between symptoms/results and diseases.

Sintoma	Doenças	Count
0 Erupcao no(a) Pele	Cryopyrin-Associated Periodic Syndromes	1
1 Vermelhidão no(a) Olho	Cryopyrin-Associated Periodic Syndromes	1
2 Perda no(a) Audição	Cryopyrin-Associated Periodic Syndromes	1
3 Febre	Familial Mediterranean Fever	1
4 Dor no(a) Peito	Familial Mediterranean Fever	1
5 Dor no(a) Abdome	Familial Mediterranean Fever	1
6 Artrite no(a) Corpo	Familial Mediterranean Fever	1
7 Dor no(a) Músculo	TNFRSF1A-Associated Periodic Syndrome	1
8 Erupção Migratória no(a) Pele	TNFRSF1A-Associated Periodic Syndrome	1
9 Edema no(a) Periorbital	TNFRSF1A-Associated Periodic Syndrome	1

Resultado	Doenças	Count
0 Variante NLRP3 patogênica	Cryopyrin-Associated Periodic Syndromes	1
1 VUS de NLRP3	Cryopyrin-Associated Periodic Syndromes	1
2 Variante MEFV patogênica	Familial Mediterranean Fever	1
3 VUS de MEFV	Familial Mediterranean Fever	1
- Right Section (Blue Highlight):** Displays diagnostic evaluation trees.
 - Possíveis Doenças:** Cryopyrin-Associated Periodic Syndromes | Score = 0.00
 - Familial Mediterranean Fever:** | Score = 0.00

Figura 16 – Tela inicial do *Auxiliar no Diagnóstico* sem interação do usuário.

O médico então adiciona os sintomas de febre e vermelhidão nos olhos que foram observados e nota as árvores de avaliação que os contêm sofrerem alterações visuais que mostram o estado de convergência daquele diagnóstico em questão, conforme mostra a Figura 17. Conforme o médico observa as possibilidades de diagnósticos, percebe que tanto a FMF quanto a CAPS são relevantes e decide perguntar ao paciente um pouco mais sobre seu histórico.

O paciente conversando com o médico sobre seu histórico e consultas anteriores, diz que já realizou um exame genético para o gene NLRP3 e que não foi encontrada nenhuma variante patogênica. Além do mais, devido a sintomas que já sentiu outras vezes no passado, o paciente relata que não teve perda de audição

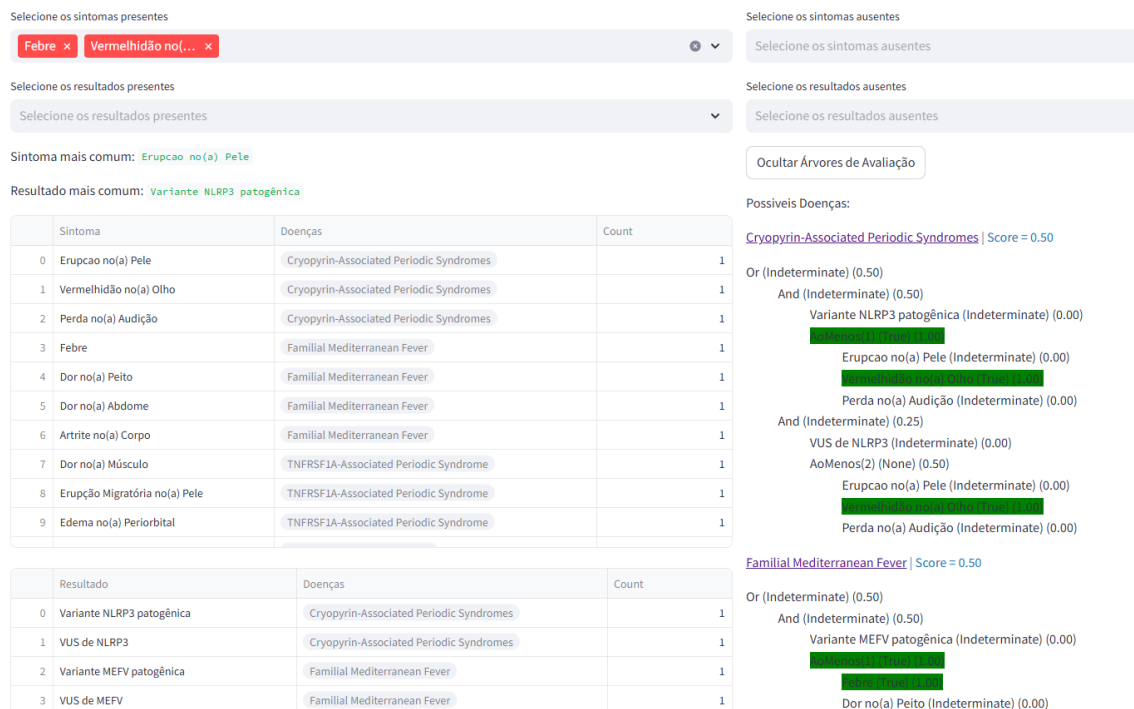


Figura 17 – Exemplo com os sintomas e resultados iniciais observados e selecionados.

nem artrite. O médico então seleciona os sintomas e resultados de exames relatados pelo paciente e vê as árvores de avaliação dos diagnósticos atualizarem novamente. Conforme podemos ver na Figura 18, tanto a FMF quanto a CAPS ainda estão convergindo, sendo a pontuação da FMF mais alta pois a CAPS perdeu pontuação ao confirmar a ausência de uma variante NLRP3 patogênica, fazendo com que a FMF apareça em primeiro lugar.

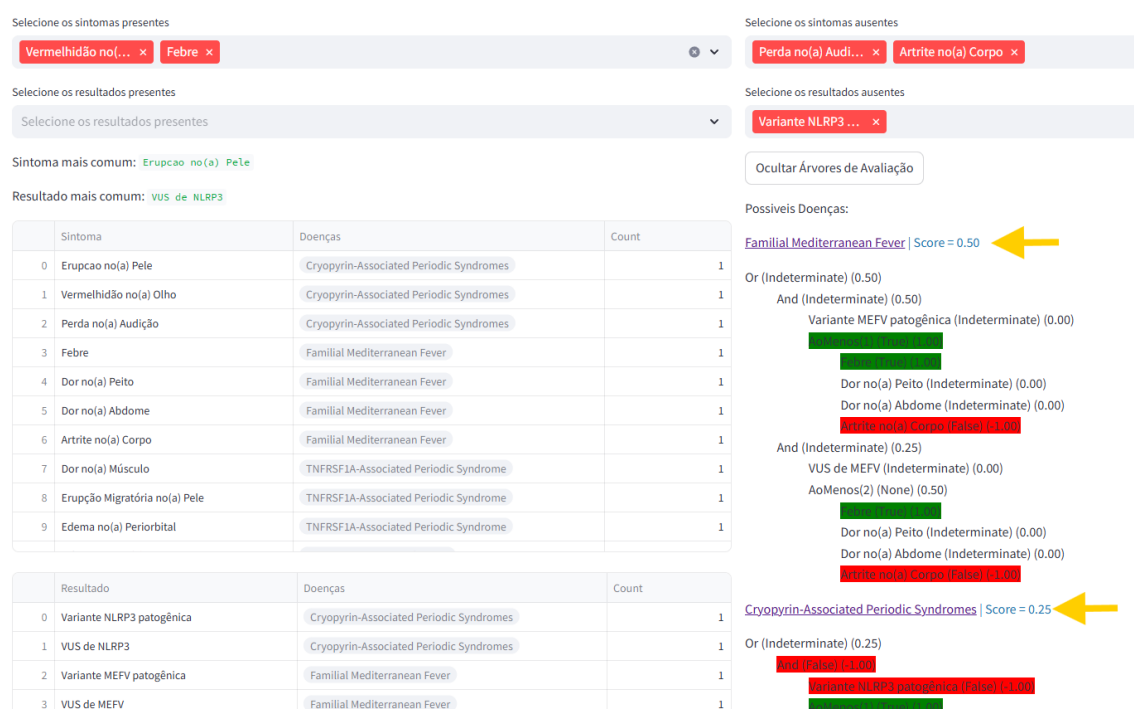


Figura 18 – Exemplo com mais alguns sintomas e resultados selecionados.

O médico, ao analisar as árvores dos diagnósticos, percebe que para a FMF convergir, atual diagnóstico mais promissor, bastaria informar uma variante patogênica do MEFV. O médico, ao conversar um pouco mais com o paciente, decide que seria interessante solicitar um exame genético para o paciente especificando o gene MEFV, pois, conforme ele observou na árvore, tendo esse resultado como verdadeiro, os sintomas e resultados seriam compatíveis com o diagnóstico da FMF, levando em conta a sensibilidade, especificidade e acurácia desse diagnóstico.

Após passar um tempo, o paciente retorna ao consultório do médico informando o resultado do exame genético que acusou uma variante patogênica do MEFV. Com esse resultado em mãos, o médico selecionou esse resultado no sistema e viu a notificação do sistema de que o diagnóstico da FMF é compatível com os sintomas e resultados de exames selecionados, conforme mostra a Figura 19. Com isso, o médico tem um forte indício de que o paciente possa ter a FMF e, portanto, decide montar um plano de acompanhamento com ele, além de analisar e indicar os tratamentos adequados à doença rara em questão.

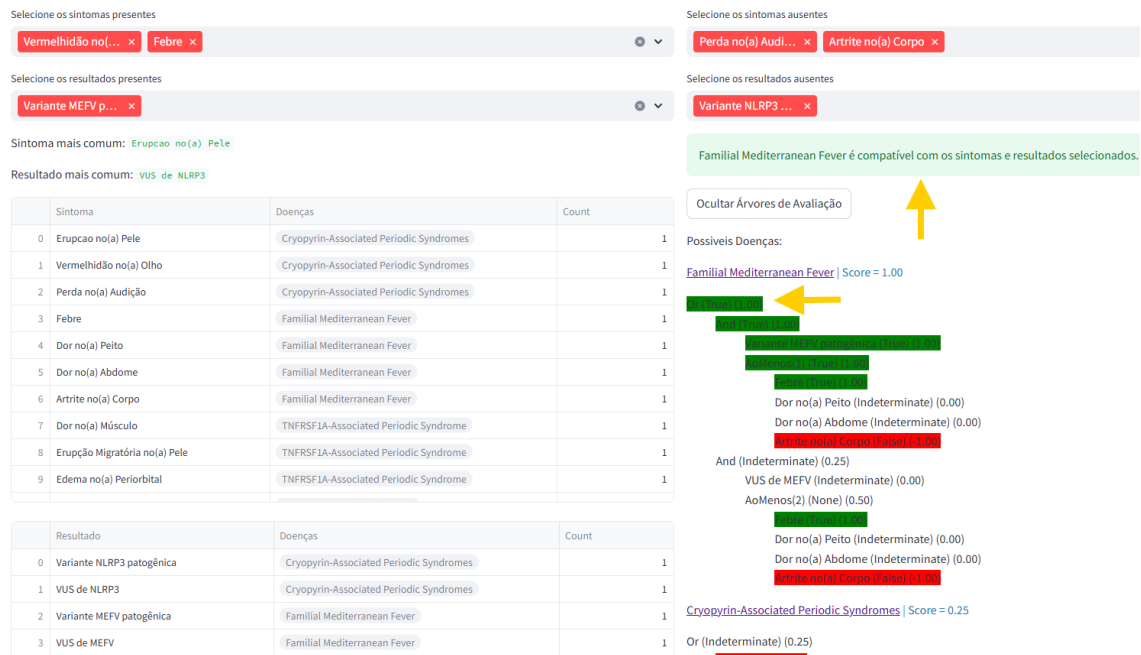


Figura 19 – Exemplo de diagnóstico convergido.

6. CONCLUSÃO

Neste trabalho, desenvolvemos um sistema chamado *DiseaseDx*, cujo propósito é auxiliar médicos no diagnóstico de doenças raras. Esse sistema utiliza uma estrutura lógica para o diagnóstico observada no estudo do Gattorno *et al.* (2019) e representa essa estrutura de forma genérica, possibilitando também o cadastro de outras doenças. Com isso, podemos configurar um sistema interativo que consegue avaliar a convergência de diagnósticos, dados os sintomas e resultados de exame

como entrada. O sistema consegue exibir a árvore lógica de um diagnóstico em tempo real, conforme os dados vão sendo preenchidos como presentes ou ausentes, fornecendo ao médico a explicabilidade do porquê de algum diagnóstico estar se destacando dos demais.

Uma das principais contribuições do sistema está na modelagem das classes que possibilitam representar a estrutura do diagnóstico, conforme exemplificada na Tabela 2. Ademais, também destacamos um critério de avaliação adequado à lógica tri-valorada, gerando uma pontuação capaz de ordenar a relevância dos diagnósticos e exibir os mais promissores primeiro. O sistema também é capaz de informar os sintomas e resultados de exames mais comuns, verificando a ocorrência dos mesmos em outras doenças. Desta forma, este trabalho contribui para o campo da tecnologia voltada para a área da saúde, proporcionando novas possibilidades para o uso de inteligência artificial e sugerindo também que abordagens baseadas em sistemas especialistas podem ainda ser valiosas para campos em que o determinismo e a explicabilidade são extremamente necessárias.

Entretanto, o sistema possui algumas limitações, sendo a principal delas a necessidade de a informação de um diagnóstico precisar estar configurada nessa estrutura lógica para poder ser adicionada à base do sistema. Para o exemplo de utilização, utilizamos dados obtidos em Gattorno *et al.* (2019), onde os critérios lógicos do diagnóstico de quatro doenças são estudados. Contudo, para o povoamento da base de doenças da DiseaseDx, precisaríamos de ter acesso a mais estudos como esse, para outras doenças raras.

Para trabalhos futuros, sugere-se analisar o uso da inteligência artificial como forma de acelerar a criação de novos critérios para as doenças nessa estrutura esperada pelo sistema. Outro ponto interessante seria adicionar, para cada diagnóstico, o peso dos sintomas através de bases de sinais e sintomas como a da Orphanet ⁶, que classifica os sintomas relativos a uma doença como muito frequentes, frequentes, ocasionais e muito raros. Além disso, também seria proveitoso avaliar a interação dos usuários com o sistema, a fim de explorar interfaces mais intuitivas que possam fornecer o cadastro de novos diagnósticos de forma simples, a tradução mais natural das informações contidas nas árvores e uma melhor experiência do usuário ao utilizar o sistema. Por fim, também seria interessante elaborar testes de software para verificar e validar o sistema conforme o seu progresso, além de testar com os usuários e comparar com outras ferramentas para colher mais feedbacks.

⁶ <https://www.orpha.net/en/disease/signs>

Referências

- ABIDI, S. S. R.; MANICKAM, S. Leveraging xml-based electronic medical records to extract experiential clinical knowledge: An automated approach to generate cases for medical case-based reasoning systems. *International Journal of Medical Informatics*, Elsevier, v. 68, n. 1-3, p. 187–203, 2002.
- ALONSO-AMO, F. *et al.* An expert system for homeopathic glaucoma treatment (seho). *Expert Systems with applications*, Elsevier, v. 8, n. 1, p. 89–99, 1995.
- AN, J.; MARWAHA, A.; LAXER, R. M. Autoinflammatory diseases: a review. *The Journal of Rheumatology*, The Journal of Rheumatology, v. 51, n. 9, p. 848–861, 2024.
- BATANOV, D.; CHENG, Z. An object-oriented expert system for fault diagnosis in the ethylene distillation process. *Computers in Industry*, Elsevier, v. 27, n. 3, p. 237–249, 1995.
- CHANG, C.; THIA, B. Online rescheduling of mass rapid transit systems: fuzzy expert system approach. *IEE Proceedings-Electric Power Applications*, IET, v. 143, n. 4, p. 307–316, 1996.
- DRIGAS, A. *et al.* An expert system for job matching of the unemployed. *Expert Systems with Applications*, Elsevier, v. 26, n. 2, p. 217–224, 2004.
- FAVIEZ, C. *et al.* Diagnosis support systems for rare diseases: a scoping review. *Orphanet journal of rare diseases*, Springer, v. 15, p. 1–16, 2020.
- GAMMA, E. *et al.* *Design Patterns: Elements of Reusable Object-Oriented Software*. [S.l.]: Addison-Wesley Professional, 1994. 416 p. ISBN 978-0201633610.
- GATTORNO, M. *et al.* Classification criteria for autoinflammatory recurrent fevers. *Annals of the rheumatic diseases*, Elsevier, v. 78, n. 8, p. 1025–1032, 2019.
- HAMADA, K. *et al.* Hybridizing a genetic algorithm with rule-based reasoning for production planning. *IEEE Expert*, IEEE, v. 10, n. 5, p. 60–67, 1995.
- KAWAHARA, K. *et al.* A proposal of a supporting expert system for outage planning of electric power facilities retaining high power supply reliability, part 1: Outline of a supporting system and outage work allocation. *IEEE Power Engineering Review*, New York, NY: IEEE Power Engineering Society, Institute of Electrical and . . . , v. 18, n. 2, p. 37, 1998.
- LIAO, S.-H. Expert system methodologies and applications—a decade review from 1995 to 2004. *Expert Systems with Applications*, Elsevier, v. 28, n. 1, p. 93–103, 2005.
- LINDSAY, R. K. *et al.* Dendral: a case study of the first expert system for scientific hypothesis formation. *Artificial intelligence*, Elsevier, v. 61, n. 2, p. 209–261, 1993.

- MEESAD, P.; YEN, G. G. Combined numerical and linguistic knowledge representation and its application to medical diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, IEEE, v. 33, n. 2, p. 206–222, 2003.
- MEHNEN, L. *et al.* Chatgpt as a medical doctor? a diagnostic accuracy study on common and rare diseases. *MedRxiv*, Cold Spring Harbor Laboratory Press, p. 2023–04, 2023.
- MILLER, R. A. *et al.* The internist-1/quick medical reference project—status report. *Western Journal of Medicine*, v. 145, n. 6, p. 816, 1986.
- MONTANI, S.; BELLAZZI, R. Supporting decisions in medical applications: the knowledge management perspective. *International journal of medical informatics*, Elsevier, v. 68, n. 1-3, p. 79–90, 2002.
- NAVATHE, S.; ELMASRI, R. *Fundamentals of Database Systems*. [S.l.]: Pearson, 2015. 1280 p. ISBN 978-0133970777.
- O'KEEFE, R. M.; PREECE, A. D. The development, validation and implementation of knowledge-based systems. *European Journal of Operational Research*, Elsevier, v. 92, n. 3, p. 458–473, 1996.
- PAL, A.; UMAPATHI, L. K.; SANKARASUBBU, M. Med-halt: Medical domain hallucination test for large language models. *arXiv preprint arXiv:2307.15343*, 2023.
- PLANT, R. E.; VAYSSIÈRES, M. P. Combining expert system and gis technology to implement a state-transition model of oak woodlands. *Computers and Electronics in Agriculture*, Elsevier, v. 27, n. 1-3, p. 71–93, 2000.
- RONICKE, S. *et al.* Can a decision support system accelerate rare disease diagnosis? evaluating the potential impact of ada dx in a retrospective study. *Orphanet journal of rare diseases*, Springer, v. 14, p. 1–12, 2019.
- SAIBENE, A.; ASSALE, M.; GILTRI, M. Expert systems: Definitions, advantages and issues in medical field applications. *Expert Systems with Applications*, Elsevier, v. 177, p. 114900, 2021.
- SCHAAF, J. *et al.* Diagnosis of rare diseases: a scoping review of clinical decision support systems. *Orphanet journal of rare diseases*, Springer, v. 15, p. 1–14, 2020.
- SCHIEPPATI, A. *et al.* Why rare diseases are an important medical and social issue. *The Lancet*, Elsevier, v. 371, n. 9629, p. 2039–2041, 2008.
- TUNMIBI, S. *et al.* A rule based expert system for diagnosis of fever. *International Journal of Advanced Research*, v. 1, n. 7, 2013.
- VALENZUELA, L. A.; BENTLEY, J. M.; LORENZ, R. D. Expert system for integrated control and supervision of dry-end sections of paper machines. *IEEE Transactions on Industry Applications*, IEEE, v. 40, n. 2, p. 680–691, 2004.
- VOLBERDA, H. W.; RUTGES, A. Farsys: a knowledge-based system for managing strategic change. *Decision Support Systems*, Elsevier, v. 26, n. 2, p. 99–123, 1999.

VRANES, S.; STANOJEVIC, M. Integrating multiple paradigms within the blackboard framework. *IEEE Transactions on Software Engineering*, IEEE, v. 21, n. 3, p. 244–262, 1995.

WAKAP, S. N. *et al.* Estimating cumulative point prevalence of rare diseases: analysis of the orphanet database. *European journal of human genetics*, Springer International Publishing Cham, v. 28, n. 2, p. 165–173, 2020.

YANG, B.-S.; HAN, T.; KIM, Y.-S. Integration of art-kohonen neural network and case-based reasoning for intelligent fault diagnosis. *Expert Systems with Applications*, Elsevier, v. 26, n. 3, p. 387–395, 2004.

ZELIN, C. *et al.* Rare disease diagnosis using knowledge guided retrieval augmentation for chatgpt. *Journal of Biomedical Informatics*, Elsevier, v. 157, p. 104702, 2024.

APÊNDICE A - Lista de ilustrações

Figura 1 – Árvore de decisão para representar o diagnóstico de FMF, com base na presença de variantes confirmatórias do gene MEFV e em sintomas clínicos.	11
Figura 2 – Captura de tela do protótipo do Ada DX mostrando o caso de um paciente com a doença TRAPS. Fonte: Ronicke <i>et al.</i> (2019) .	12
Figura 3 – Acurácia do diagnóstico do ChatGPT 4 e 3 médicos em 10 doenças raras, permitindo um máximo de 10 sugestões de diagnósticos mais prováveis. Fonte: Mehnen <i>et al.</i> (2023)	13
Figura 4 – Prompt em inglês para verificar o possível diagnóstico do FMF dado os sintomas, usando a versão paga do ChatGPT 4.1.	13
Figura 5 – Prompt em inglês para verificar o possível diagnóstico do CAPS dado os sintomas, usando a versão paga do ChatGPT 4.1.	14
Figura 6 – Prompt em inglês para verificar a sugestão de exame genético para diagnosticar o CAPS, usando a versão paga do ChatGPT 4.1.	14
Figura 7 – Prompt em português para verificar o possível diagnóstico do CAPS dado os sintomas, usando a versão paga do ChatGPT 4.1.	14
Figura 8 – Prompt em português para verificar a sugestão de exame genético para diagnosticar o CAPS, usando a versão paga do ChatGPT 4.1.	15
Figura 9 – Prompt em português com pequena alteração para verificar o possível diagnóstico do CAPS dado os sintomas, usando a versão paga do ChatGPT 4.1.	15
Figura 10 – Visão geral do DiseaseDx.	17
Figura 11 – Diagrama de classes simplificado do DiseaseDx.	17
Figura 12 – Exemplo de diagrama de objetos do DiseaseDx para a doença FMF.	19
Figura 13 – Árvore de decisão para representar o diagnóstico de FMF. Os nós pintados de verde indicam que o nó é verdadeiro e os nós pintados de vermelho indicam que o nó é falso.	21
Figura 14 – Árvore de decisão para representar o diagnóstico de FMF. Os nós pintados de verde indicam que o nó é verdadeiro e os nós pintados de vermelho indicam que o nó é falso.	22
Figura 15 – Árvore de decisão para representar o diagnóstico de FMF. Os nós pintados de verde indicam que o nó é verdadeiro e os nós pintados de vermelho indicam que o nó é falso.	23
Figura 16 – Tela inicial do <i>Auxiliar no Diagnóstico</i> sem interação do usuário.	24
Figura 17 – Exemplo com os sintomas e resultados iniciais observados e selecionados.	25

Figura 18 – Exemplo com mais alguns sintomas e resultados selecionados. . .	25
Figura 19 – Exemplo de diagnóstico convergido.	26

APÊNDICE B - Lista de tabelas

Tabela 1	– Mapa de calor de algumas doenças auto-inflamatórias. Traduzido de: An, Marwaha e Laxer (2024).	10
Tabela 2	– Novos critérios de classificação Eurofever/PRINTO para febres hereditárias recorrentes e seu desempenho em 281 pacientes com consenso. Traduzido de: Gattorno <i>et al.</i> (2019).	10
Tabela 3	– Tabela verdade de lógica tri-valorada. <i>Not</i> ocultado pois sintomas e resultados ausentes vão para uma lista de ausente ao invés de serem a negação do presente.	19

Agradecimentos

Primeiramente, gostaria de agradecer aos meus orientadores, Leonardo Murta, Vanessa Braganholo e Luiz Eduardo Oliveira, por toda a orientação e apoio durante o processo. Gostaria também de agradecer aos meus colegas da faculdade, aos meus amigos, ao Rodolfo Albuquerque, Luigi Cavazzani, Sylvio Mello, Gustavo Lopes e Rodrigo Junger, por todo o apoio e aprendizado que pude adquirir com eles ao longo desse tempo. Por fim, gostaria de agradecer em especial a meus pais, Sônia Maria de Paula Pinto e José Luiz Moreira Peixoto, por todo o suporte que tive ao longo da vida e principalmente durante o meu período acadêmico. Reforço também o agradecimento à minha mãe por todo o esforço e suporte que me forneceu durante o período em que planejei mudar de curso. Agradeço a todos do fundo do coração por terem contribuído para a conclusão dessa etapa na minha vida.