

**UNIVERSIDADE FEDERAL DE ALAGOAS**  
**INSTITUTO DE COMPUTAÇÃO**  
**COORDENAÇÃO DE PÓS-GRADUAÇÃO EM INFORMÁTICA**

**UM SISTEMA INTELIGENTE PARA A AVALIAÇÃO DE  
RISCO DA DRC E ENCAMINHAMENTO DE PACIENTES EM  
EMERGÊNCIA PARA UNIDADES DE SAÚDE**

**MESTRANDA**

**ANDRESSA CARVALHO MELO DA SILVEIRA QUEIROZ**

**ORIENTADOR(A)**

**LEANDRO DIAS DA SILVA**

**COORIENTADOR(A)**

**ÁLVARO ALVARES DE CARVALHO CÉSAR SOBRINHO**

**MACEIÓ, AL**  
**AGOSTO - 2020**

**ANDRESSA CARVALHO MELO DA SILVEIRA QUEIROZ**

**ORIENTADOR(A)**

**LEANDRO DIAS DA SILVA**

**COORIENTADOR(A)**

**ÁLVARO ALVARES DE CARVALHO CÉSAR SOBRINHO**

**MACEIÓ, AL**  
**AGOSTO - 2020**

**Catálogo na fonte**  
**Universidade Federal de Alagoas**  
**Biblioteca Central**  
**Divisão de Tratamento Técnico**

Bibliotecário: Marcelino de Carvalho Freitas Neto – CRB-4 - 1767

Q3s      Queiroz, Andressa Carvalho Melo da Silveira.  
Um sistema inteligente para a avaliação de risco da DRC e  
encaminhamento de pacientes em emergência para unidades de saúde /  
Andressa Carvalho Melo da Silveira Queiroz. – 2020.  
73 f. : il.

Orientador: Leandro Dias da Silva.  
Co-orientador: Álvaro Álvares de Carvalho César Sobrinho.  
Dissertação (mestrado em Informática) - Universidade Federal de  
Alagoas. Instituto de Computação. Maceió, 2020.

Bibliografia: f. 61-73.

1. Aprendizado do computador. 2. Insuficiência renal crônica. 3. Países  
em desenvolvimento. I. Título.

CDU: 004.89

## Agradecimentos

Primeiramente, agradeço a Deus por ser meu alicerce em toda minha vida, pela força que me dá diariamente pra seguir meu caminho, lutar pelos meus sonhos e por me ajudar a conseguir realizar mais um objetivo.

Aos meus pais Ilda e Vladimir, pelos seus cuidados em todos esses anos de companheirismo, em todos os aspectos da minha vida. A minha irmã Adryelly, amiga e cúmplice, por sempre estar ao meu lado me dando forças pra seguir e principalmente por me ajudar a cuidar do meu filho para que eu pudesse estudar. Não tenho palavras para agradecer por todo apoio, amor, incentivo e por vocês sempre acreditarem em mim quando muitos duvidaram. E principalmente por me ensinarem que as pessoas podem me tirar tudo, menos o conhecimento. E que o estudo é o único bem que realmente é seu.

Aos meus orientadores, Professor Leandro Dias e Professor Álvaro Sobrinho, que com a serenidade própria da sabedoria despertaram em mim a disciplina. Por toda paciência diante de minhas dificuldades, limitações, todo apoio e direcionamento. Seus ensinamentos e conselhos foram fundamentais para a minha formação como pesquisadora. Muito obrigada por tudo.

A toda minha família, minhas avós (in memoriam), mas que foram meu exemplo de simplicidade e humildade. A minha tia e segunda mãe Nice por toda ajuda, amor e incentivo. Aos amigos que fiz no mestrado, Ially e João, que levarei por toda minha vida. A minha prima e irmã Maysa, que sempre teve um ombro amigo quando eu precisei. A Juscel por todo apoio e incentivo no início dessa caminhada. E a todos que aqui não foram citados, mais que foram e sempre serão essenciais na minha vida.

À banca de avaliadores, que muito prontamente e com grande entusiasmo aceitaram fazer parte desse momento tão importante da minha vida, obrigada por toda a contribuição com este trabalho e com minha formação acadêmica.

Aos alunos de graduação da Universidade Federal Rural do Semi-Árido, campus Pau dos Ferros, que colaboraram com a realização da revisão sistemática da literatura apresentada no Capítulo 4 desta dissertação. A nefrologista Maria Eliete, por seu apoio e colaboração com este trabalho.

À FAPEAL e a CAPES, pelo apoio financeiro para realização deste trabalho de pesquisa

e a à UFAL pela infraestrutura fornecida.

Dedico este trabalho ao meu filho Leidson, que chegou pra me completar e mesmo sendo tão pequeno me ensina tanto. Descobri que estava grávida um mês depois de iniciar o mestrado, não foi fácil conciliar a maternidade com os estudos, mas foi por você que não desistir e que luto diariamente. Você foi meu grande milagre e é a razão do meu viver. Amo você meu filho!

A todos vocês o meu eterno e profundo Obrigada!

## Resumo

A alta incidência e prevalência de Doença Renal Crônica (DRC), frequentemente causada por diagnósticos tardios, é um problema crítico de saúde pública, principalmente em países em desenvolvimento como o Brasil. As terapias de tratamento da DRC, como diálise e transplante renal, aumentam as taxas de morbimortalidade, além dos custos com saúde pública. Inicialmente, neste estudo, foi analisado o uso de técnicas de aprendizado de máquina para auxiliar no monitoramento da DRC em países em desenvolvimento. Análises comparativas qualitativas e quantitativas foram, respectivamente, realizadas executando uma revisão sistemática da literatura e um experimento com técnicas de aprendizado de máquina, com o método de validação cruzada *k-fold*, baseado no software Weka<sup>©</sup> e um conjunto de dados da DRC. A partir das análises, foi possível discutir a adequação das técnicas de aprendizado de máquina para a avaliação de risco de DRC, concentrando-se em ambientes de baixa renda e de difícil acesso em países em desenvolvimento, devido aos problemas específicos enfrentados, como, por exemplo, atendimento primário inadequado. Com base nos resultados do estudo, foi possível observar que a árvore de decisão J48 é uma técnica de aprendizado de máquina adequada para a avaliação de risco em países em desenvolvimento, devido à fácil interpretação de resultados de classificação, com 95,00% de precisão, alcançando concordância quase perfeita com a opinião de um nefrologista experiente. Por outro lado, as técnicas de floresta aleatória, *naive Bayes*, máquina de vetores de suporte, *perceptron* multicamada e *k* vizinho mais próximo, respectivamente, apresentaram 93,33%, 88,33%, 76,66%, 75,00% e 71,67% de precisão, com pelo menos concordância moderada com o nefrologista, à custa de uma interpretação mais difícil dos resultados da classificação. Com esta conclusão, a árvore de decisão J48 foi usada para desenvolver um sistema inteligente para avaliar o risco de DRC em países em desenvolvimento. Além disso, quando o paciente com DRC está fora de seu município e ocorre uma emergência, o sistema recomenda que o paciente compareça a uma unidade de saúde apropriada, dependendo da situação clínica, para evitar cuidados de saúde tardios ou inadequados.

**Palavras-chave:** Aprendizado de Máquina, Doença Renal Crônica, Países em Desenvolvimento.

## Abstract

The high incidence and prevalence of chronic kidney disease (CKD), often caused by late diagnoses, is a critical public health problem, especially in developing countries such as Brazil. CKD treatment therapies, such as dialysis and kidney transplantation, increase the morbidity and mortality rates, besides the public health costs. Firstly, this study analyses the usage of machine learning techniques to assist in the early diagnosis of CKD in developing countries. Qualitative and quantitative comparative analyses are, respectively, conducted using a systematic literature review and an experiment with machine learning techniques, with the k-fold cross-validation method based on the Weka<sup>©</sup> software and a CKD dataset. These analyses enable a discussion on the suitability of machine learning techniques for screening for CKD risk, focusing on low-income and hard-to-reach settings of developing countries, due to the specific problems, e.g., inadequate primary health care. The study results show that the J48 decision tree is a suitable machine learning technique for such screening in developing countries, due to the straightforward interpretation of its classification results, with 95.00% accuracy, reaching a nearly perfect agreement with an experienced nephrologist's opinion. Conversely, random forest, naive Bayes, support vector machine, multilayer perceptron, and k-nearest neighbor techniques, respectively, yield 93.33%, 88.33%, 76.66%, 75.00%, and 71.67% accuracy, presenting at least moderate agreement with the nephrologist, at the cost of a more difficult interpretation of the classification results. With this conclusion, J48 decision tree was used to develop an intelligent system to evaluate the CKD risk in developing countries. In besides, when the CKD patient is outside his/her county and an emergency occurs, it is proposed that the system recommends the patient to attend an appropriate health unit depending on the clinical situation to prevent late or inadequate health care.

**Keywords:** Machine Learning, Chronic Kidney Disease, Web Application, Developing Countries.

# Lista de Figuras

2.1	Exemplo de árvore de decisão, adaptado de [99]. . . . .	9
2.2	Rede Naive Bayes [70]. . . . .	10
2.3	Arquitetura Rede MLP, adaptado de [50]. . . . .	11
2.4	Ilustração da lógica relacionada ao algoritmo da floresta aleatória [65]. . . . .	13
2.5	A interface do usuário do Weka Explorer. . . . .	14
2.6	Interface de classificação do Weka Explorer. . . . .	15
2.7	Algoritmo Cross-Validation em pseudocódigo [78]. . . . .	16
2.8	Algoritmo Hold-Out em pseudo-código [78]. . . . .	17
2.9	Arquitetura simplificada do <i>Expert Sinta</i> [95]. . . . .	20
4.1	Visão geral da abordagem para a revisão sistemática da literatura, incluindo pesquisa em banco de dados, pesquisa manual e a técnica <i>snowballing</i> . . . . .	29
4.2	Fonte dos 30 trabalhos resultantes da busca inicial guiada pela abordagem utilizada na revisão sistemática da literatura. . . . .	30
4.3	Procedimento de avaliação usando classificadores de AM e opiniões dos nefrologistas. . . . .	38
4.4	(a) Árvore de decisão para avaliação de risco de DRC feita pelo algoritmo J48. . . . .	47
4.5	(b) Amostra das árvores aleatórias para avaliação de risco de DRC feita pelo algoritmo de RF. . . . .	49
5.1	GUI do sistema para registrar um novo exame. . . . .	51
5.2	GUI principal do sistema. . . . .	52
5.3	Resumo das primeiras decisões do sistema especialista. . . . .	53
5.4	Lista de medicamentos relacionados com cada patologia. . . . .	54

5.5	Relacionamentos entre perguntas para identificar o risco de hiperglicemia. .	55
5.6	Relacionamentos entre perguntas para identificar o risco de hipoglicemia. .	56
5.7	Relacionamentos entre perguntas para identificar o risco de hipercalemia. .	56
5.8	Relacionamentos entre perguntas para identificar o risco de hipocalemia. . .	57
5.9	GUI do sistema especialista para DM e álcool. . . . .	57

# Lista de Tabelas

2.1	Estagiamento da DRC proposta pelo grupo de trabalho do K/DOQI [8]. . . . .	18
4.1	Critérios de inclusão e exclusão da revisão sistemática da literatura. . . . .	28
4.2	Resumo dos atributos, técnicas e desempenho de acordo com o QP1.2 e QP1.3. . . . .	35
4.3	Amostra do conjunto de dados da DRC coletada no Hospital Universitário Prof. Alberto Antunes. Cada ID representa um assunto específico do conjunto de dados. . . . .	36
4.4	Resultados da avaliação de qualidade para os classificadores e o nefrologista mais experiente usando as métricas de desempenho. . . . .	40
4.5	Resultados da avaliação de qualidade para os classificadores e o nefrologista mais experiente usando a estatística Kappa. . . . .	41
4.6	Resultados da avaliação da qualidade dos classificadores e dos dois nefrologistas menos experientes que usam a estatística Kappa. . . . .	42
4.7	Resultados da avaliação da qualidade dos classificadores e dos três nefrologistas que usam a estatística Kappa. . . . .	42
5.1	Amostra de cenários simulados de hiperglicemia usados para avaliar o sistema. . . . .	58
5.2	Resultados da primeira etapa da avaliação do sistema especialista. . . . .	58

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Motivação . . . . .	3
1.2	Problemática . . . . .	4
1.3	Objetivos . . . . .	5
1.4	Organização do Documento . . . . .	6
<b>2</b>	<b>Fundamentação Teórica</b>	<b>7</b>
2.1	Aprendizado de Máquina . . . . .	7
2.1.1	Árvore de Decisão J48 . . . . .	8
2.1.2	Naive Bayes . . . . .	9
2.1.3	Rede Neural Perceptron Multicamadas . . . . .	10
2.1.4	K-vizinhos Mais Próximos . . . . .	10
2.1.5	Máquina de Vetores de Suporte . . . . .	12
2.1.6	Floresta Aleatória . . . . .	12
2.2	Weka . . . . .	13
2.3	Doença Renal Crônica . . . . .	16
2.4	Sistemas Especialistas . . . . .	18
2.5	Considerações Finais . . . . .	19
<b>3</b>	<b>Trabalhos Relacionados</b>	<b>21</b>
3.1	Saúde Inteligente . . . . .	21
3.2	Aplicações para a Doença Renal Crônica . . . . .	23
3.3	Aplicações para o Encaminhamento para Unidades de Saúde . . . . .	24
3.4	Considerações Finais . . . . .	25

<b>4</b>	<b>Análise Comparativa</b>	<b>26</b>
4.1	Revisão Sistemática da Literatura . . . . .	27
4.1.1	Protocolo . . . . .	27
4.1.2	Resultados . . . . .	28
4.2	Experimento com o Weka . . . . .	34
4.2.1	Conjunto de Dados . . . . .	34
4.2.2	Procedimento . . . . .	37
4.2.3	Análise de Dados . . . . .	39
4.2.4	Discussão . . . . .	42
4.3	Considerações Finais . . . . .	48
<b>5</b>	<b>Desenvolvimento</b>	<b>50</b>
<b>6</b>	<b>Conclusões</b>	<b>59</b>

# Capítulo 1

## Introdução

O objetivo principal com o Sistema Único de Saúde (SUS), criado em 1998, é proporcionar o acesso aos serviços de saúde para os cidadãos brasileiros. Segundo o Ministério da Saúde (MS), o SUS é composto por três níveis de atenção à saúde: primário, secundário e terciário [28]. As unidades de Atenção Básica em Saúde (ABS) constituem o primeiro nível, tendo como propósito ser a porta de entrada ao SUS, onde são agendadas consultas e exames, e efetuados procedimentos menos complexos, como, por exemplo, vacinação e curativos. No nível secundário estão as clínicas e Unidades de Pronto Atendimento (UPAs), bem como os hospitais escola, caracterizados pelo atendimento de casos de média complexidade e realização de tratamentos em casos de doenças crônicas. No nível terciário, encontram-se hospitais de grande porte. Neste caso, o propósito é receber pacientes que correm risco à vida, com a realização de procedimentos mais invasivos.

Diante de diferentes tipos de enfermidades que necessitam de constante monitoramento, tem-se a alta incidência da Doença Renal Crônica (DRC), que tem alertado a comunidade científica mundial. Neste contexto, é preocupante observar que, em alguns casos, existe inadequabilidade dos cuidados de saúde oferecidos durante o tratamento da DRC [9]. Por se tratar de uma doença assintomática, o diagnóstico da DRC geralmente ocorre em estágios mais avançados. Com isso, medidas de controle muitas vezes não podem ser realizadas, diminuindo a qualidade de vida do paciente ao causar complicações renais.

Em um estudo sobre Terapia Renal Substitutiva (TRS), realizado em janeiro de 2009, foi constatado que 77.589 pacientes encontravam-se em diálise no Brasil, e, que, a predominância e a ocorrência de Doença Renal em Estágio Terminal (DRET) representavam cerca

de 405 e 144 por milhão na população, respectivamente [92]. Em um relatório técnico, a Organização Mundial da Saúde (OMS) destacou que a DRC é a maior causa de morte entre os adultos em países de baixa e média renda e que dos 58 milhões de mortes que ocorreram no mundo em 2005, aproximadamente 35 milhões foram resultado de doenças crônicas [75].

Quando a DRC é identificada de maneira tardia, perde-se a oportunidade de ter acesso a prevenção primária, secundária e terciária em estágios menos avançados. Um dos motivos é a falta de informação sobre a definição e classificação das fases da doença, bem como a falta de acesso a exames para a detecção e avaliação da doença [8]. As etapas essenciais para o tratamento de pacientes com DRC são: (1) a investigação precoce da doença e (2) a orientação imediata para procurar um nefrologista. Ao realizar estas etapas, é possível proporcionar a educação pré-diálise e a realização de atitudes que previnam e retardem (ou interrompam) a evolução para os estágios mais avançados da DRC [10].

No estudo apresentado por Roubicek *et al.* (2000) [17], foram comparados pacientes com Encaminhamento Precocemente (EP) com pacientes que tiveram o Encaminhamento Tardio (ET). Com isso, observou-se que, pacientes com EP, quando comparado com os pacientes com ET, permaneciam menos dias no hospital depois do início da diálise, possuíam menos chance de necessitar de diálise de urgência, possuíam um controle da pressão arterial melhor e menos casos de edema agudo de pulmão.

Portanto, a utilização de sistemas de software para o auxílio no tratamento e identificação de doenças crônicas pode ser uma forma de diminuir os altos índices de mortalidade pelo monitoramento de doenças crônicas como a DRC [98]. Tal situação remete ao uso de informações e tecnologias, de forma inteligente e eficaz, para orientar àqueles que eventualmente serão atendidos no SUS. A identificação precoce da DRC auxiliada por computador pode auxiliar a população que vive em ambientes rurais e de difícil acesso a receberem atendimentos primários. Além disso, aplicativos móveis de saúde (*i.e.*, *mHealth*), que geram registros de saúde pessoais (*Personal Health Records* - PHR), podem ser usados para reduzir problemas (*i.e.*, armazenando o histórico médico completo de um paciente com diagnósticos, medicamentos administrados, planos de tratamento, datas de vacinação, alergias) relacionados à atenção primária à saúde em locais remotos.

Com o intuito de fornecer recursos para auxiliar os usuários a realizarem o automonitoramento de risco de DRC em países em desenvolvimento, em um estudo desenvolvido por

Sobrinho *et al.* (2018) [98], os autores apresentaram um protótipo em Android de um sistema denominado MultCare. O sistema MultCare foi desenvolvido baseado em diretrizes médicas, especificações formais, avaliação de eficácia e testes de usabilidade. Nesta dissertação, resultados obtidos por Sobrinho *et al.* (2018) [98] são estendidos com o objetivo de aumentar a confiança em avaliações de risco da DRC, melhorar usabilidade e adicionar novas funcionalidades.

## 1.1 Motivação

Uma das doenças que mais tem causado problemas a saúde pública mundial é a DRC. Pacientes com DRC possuem risco de mortalidade cardiovascular 8 a 10 vezes maior que a população geral [15]. A Sociedade Brasileira de Nefrologia (SBN) realizou uma pesquisa, detectando que, aproximadamente 13 milhões dos brasileiros tem algum grau de problema renal [100]. Por se tratar de uma doença assintomática, quando diagnosticada, geralmente, cerca de 50% dos rins possuem algum tipo de comprometimento da capacidade de excreção. Com isso, aproximadamente 70% das mortes ocorrem antes do diagnóstico da DRC [103]. O diagnóstico precoce da DRC é ainda mais difícil para cidadãos que residem em áreas remotas e de difícil acesso, dado que, muitas vezes, a atenção primária pode ser precária [97].

Alguns pacientes demonstram suscetibilidade aumentada para DRC e encontram-se em grupos de risco: hipertensos, diabéticos, idosos, pacientes com doença cardiovascular (DCV), familiares de pacientes portadores de DRC e pacientes em uso de medicações nefrotóxicas [8]. Um dos problemas detectados no tratamento da DRC é a dificuldade na detecção precoce, que geralmente ocorre em fases mais avançada (Estágios 4 e 5) [98]. O diagnóstico da DRC em estágios iniciais de desenvolvimento é relevante para auxiliar na redução das taxas de mortalidade, bem como na redução de custos com saúde pública nos países em desenvolvimento [97]. A falta de tratamento em estágios iniciais da doença resulta na necessidade de diálise e transplante renal, elevando também os custos com a saúde pública.

Como exemplo do uso de computação para reduzir o diagnóstico tardio, o conhecimento de especialistas experientes pode ser utilizado para o treinamento de classificadores baseados em Aprendizado de Máquina (AM), possibilitando suporte aos médicos que realizam atendimentos primários em comunidades carentes, algumas vezes em locais remotos e de

difícil acesso. O baixo poder aquisitivo em parte considerável da população, atendimentos precários à saúde (ou falta de atendimentos) e o elevado número de pessoas residindo em locais remotos e de difícil acesso são características comumente encontradas em países em desenvolvimento, que motivam o estudo sobre aplicação de AM para o auxílio ao diagnóstico e monitoramento da DRC.

Além disso, um dos principais problemas enfrentados pelo SUS é o da superlotação nos hospitais públicos. Técnicos do tribunal de contas da união fiscalizaram em 2014 cerca de 116 hospitais e prontos-socorros do Brasil, verificando que 64% encontravam-se constantemente superlotados, os 36% restante sofreram com essa situação com menos regularidade [37]. Por exemplo, em uma pesquisa realizada no Hospital Geral do Estado de Alagoas (HGE) foi constatado que 77% dos atendimentos efetuados nessa unidade possuíam o setor de clínico geral como porta de entrada. Porém, muitos desses pacientes atendidos pelo maior hospital do estado de Alagoas possuíam doenças que poderiam ser tratadas nas unidades de atenção secundária ou terciária [5].

## 1.2 Problemática

O diagnóstico tardio é um problema crítico, especialmente em países em desenvolvimento (*e.g.*, Brasil), que geralmente sofrem com a precariedade da atenção primária à saúde, por exemplo, em locais de difícil acesso e rurais. Neste estudo, países em desenvolvimento representam regiões de baixa e média renda, enquanto os países desenvolvidos são regiões de alta renda, como os Estados Unidos da América. O baixo poder aquisitivo em parte considerável da população, atendimentos precários à saúde (ou falta de atendimento) e o número elevado de pessoas residindo em locais remotos e de difícil acesso foram as características comumente encontradas em países em desenvolvimento, que motivaram a realização deste estudo.

Outro problema é a falta de informações necessárias para identificar qual unidade de saúde o paciente deve se dirigir de acordo com a sua condição clínica, principalmente para pacientes que residem em zona rural ou que estão em trânsito, (*i.e.*, pacientes que estão fora da sua cidade natal) [38]. Com isso, na maioria dos casos, pacientes que poderiam ser atendidos em ABS ou UPA, se dirigem para os grandes hospitais, sobrecarregando o nível

terciário, que é o que mais sofre com superlotações.

Apesar de já existirem sistemas de software e trabalhos relacionados com o monitoramento e diagnóstico de DRC, pode-se observar ao menos uma lacuna de pesquisa. Até o momento, em nenhum dos estudos identificados foi disponibilizada uma análise comparativa que aborde o uso de técnicas de AM para o diagnóstico de DRC sob a perspectiva de problemas específicos em países em desenvolvimento.

O sistema de software MultCare [98] é um exemplo de sistema de software utilizado para auxiliar no monitoramento de pacientes com DRC em países em desenvolvimento [96]. A resposta final do MultCare é uma análise de risco de desenvolvimento da DRC e um alerta de encaminhamento pra um nefrologista, caso necessário [98]. Entretanto, ao analisar os resultados de análise de risco, os autores identificaram baixa concordância entre o MultCare e três nefrologistas, problema que ocorre devido ao uso apenas de diretrizes médicas. Já nos testes de usabilidade, foram identificados problemas com o uso de muitos componentes de interfaces gráficas de usuário (*i.e.*, falta de simplificação). Um dos problemas identificados foi a necessidade de inclusão manual de exames médicos pelo usuário. Por exemplo, uma solução para reduzir a necessidade de inclusão manual de um grande número de entrada de dados é a possibilidade de utilização de arquivos XML para a leitura e cadastro automatizado de exames. Como descrito anteriormente, nesta dissertação, resultados preliminares obtidos com o desenvolvimento do sistema de software MultCare são estendidos, considerando o foco em países em desenvolvimento.

### 1.3 Objetivos

O objetivo principal com este trabalho é utilizar AM no desenvolvimento de um sistema de software para auxiliar o monitoramento da DRC em países de baixa e média renda. Para contemplar o objetivo principal destacado, os seguintes objetivos específicos devem ser considerados:

- identificar a técnica de AM mais adequada para a avaliação de risco da DRC em países em desenvolvimento;
- desenvolver sistema Web para o monitoramento da DRC, incluindo gerenciamento de

dados clínicos e avaliação de risco usando AM;

- modelar quatro situações clínicas (hipoglicemia, hiperglicemia, hipocalemia e hipercalemia) a partir de conhecimento disponível na literatura, com a validação de um profissional de saúde. Com isso, o objetivo é possibilitar o projeto de uma funcionalidade para o encaminhamento para unidades de saúde;
- desenvolver sistema Web para encaminhar indivíduos em monitoramento para atendimento adequado quando uma situação de risco é identificada, como uma nova funcionalidade do sistema de software MultCare;
- validar os sistemas desenvolvidos.

Portanto, ao contemplar o objetivo principal e os objetivos específicos, as principais contribuições com este trabalho incluem:

- resultados de uma análise comparativa sobre técnicas de AM para o diagnóstico da DRC em países em desenvolvimento;
- disponibilização de um sistema Web para a identificação e monitoramento da DRC usando AM;
- disponibilização de um sistema Web para o encaminhamento de pacientes em situações de emergência.

## 1.4 Organização do Documento

No primeiro capítulo desta dissertação foi apresentada uma introdução. O restante do documento está organizado da seguinte maneira. No Capítulo 2 é apresentada a fundamentação teórica, composta por conceitos de AM, classificadores, Weka<sup>®</sup>, DRC e sistemas especialistas. Estes conceitos foram utilizados durante o desenvolvimento deste trabalho. No Capítulo 3 são apresentados trabalhos relacionados com o tema de pesquisa. No Capítulo 4 é apresentada uma análise comparativa de AM, no qual é identificada a técnica de AM mais adequada para a avaliação de DRC em países em desenvolvimento. No Capítulo 5 é apresentado o desenvolvimento da versão Web do software MultCare. No Capítulo 6 são apresentadas conclusões e trabalhos futuros. Por fim, são apresentadas as referências bibliográficas.

# Capítulo 2

## Fundamentação Teórica

Neste capítulo são definidos conceitos fundamentais utilizados na solução proposta, contextualizando-os na literatura. Mais especificamente, são descritos conceitos fundamentais sobre AM, classificadores baseados em AM, Weka©, DRC e sistemas especialistas.

### 2.1 Aprendizado de Máquina

Aprendizado de Máquina (AM) é uma sub-área da Inteligência Artificial (IA). Um dos objetivos é possibilitar que computadores sejam programados para aprender com experiências passadas [18, 69]. A habilidade de aprender é um dos atributos mais importantes do comportamento inteligente [46]. Um sistema de aprendizado toma decisões fundamentado em experiências acumuladas com base em conclusões bem sucedidas de problemas anteriores. Estes sistemas possuem particularidades e atributos comuns que possibilitam sua classificação quanto à linguagem de descrição, modo, paradigma e forma de aprendizado aplicado [69].

Um dos principais métodos para gerar um novo conhecimento e prever eventos futuros é o aprendizado indutivo. Neste caso, a inferência lógica obtém previsões genéricas em um conjunto particular de exemplos [69]. O aprendizado indutivo é um dos mais úteis, por permitir a obtenção de novos conhecimentos a partir de exemplos (ou casos) particulares previamente observados [11]. Este tipo de aprendizado pode ser definido como aprendizado supervisionado e não supervisionado.

Por um lado, o aprendizado supervisionado consiste em buscar uma função para otimizar

a habilidade de generalização de um modelo matemático gerado com base em instâncias utilizadas para treinamento de um algoritmo de AM [84]. O objetivo é construir um modelo preditivo que consiste em rotular novos exemplos [90]. Por outro lado, o aprendizado não supervisionado tem como propósito encontrar semelhanças usando instâncias com o intuito de descobrir similaridades e anomalias, produzindo um modelo que busca por regularidades, construindo agrupamentos ou *clusters* de exemplos com características similares. Com isso, o conjunto de dados aplicado não deve conter nenhuma espécie de rótulo [11].

Dado o grande número de classificadores disponíveis, é desafiador identificar qual classificador é provavelmente o melhor para uma nova instância ou classe de instâncias de problemas. Portanto, para avaliar o desempenho ao detectar a Doença Renal Crônica (DRC), foram avaliados cinco classificadores neste estudo, que foram selecionados baseado nas questões de pesquisa apresentadas na revisão sistemática do Capítulo 3. Mais especificamente, os seguintes classificadores foram abordados: a árvore de decisão (J48), Naive Bayes (NB), Rede Neural Perceptron Multicamadas (*Multilayer Perceptron* (MLP)), k-vizinhos mais próximos (*K-Nearest Neighbors* (KNN)), máquina de vetores de suporte (*Support-Vector Machines* (SVM)) e floresta aleatória (*Random Forest* (RF)).

### 2.1.1 Árvore de Decisão J48

O classificador de árvore de decisão é uma técnica baseada na estratégia de dividir e conquistar, que consiste em nós de decisão e nós folha. O nó de decisão é usado para especificar um teste sobre um dos atributos e um nó folha é usado para representar o valor da classe [87]. Na Figura 2.1 é ilustrado um exemplo de árvore de decisão simples, no qual é usado um banco de dados de aluno em uma turma. Os alunos são classificados em diferentes categorias com base em desempenhos. Portanto, são obtidas diferentes classes de alunos e também a contagem do número de alunos em cada classe [99].

Existem classificadores diferentes para gerar a árvore de decisão, como, por exemplo, ID3, J48, FT, BFTree e LMT. Para este estudo, foi utilizado o algoritmo J48, por possui maior taxa de precisão [52]. O classificador J48 é uma recodificação do algoritmo C4.5. Na versão recodificada é utilizada a taxa de ganho de informações pelo critério de divisão. O propósito é gerar uma árvore de decisão baseada em um conjunto de dados de treinamento. O modelo é utilizado para classificar as instâncias no conjunto de teste com base em uma

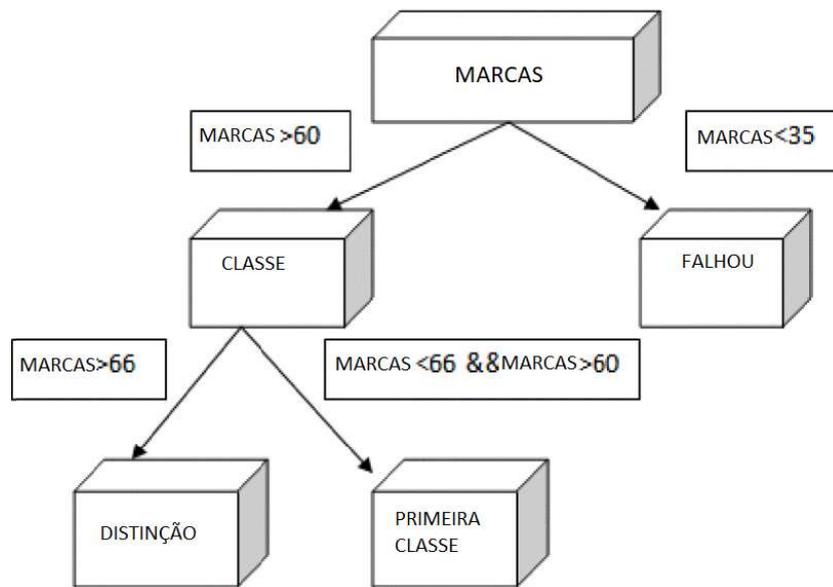


Figura 2.1: Exemplo de árvore de decisão, adaptado de [99].

árvore binária [77, 99].

Nesta técnica são utilizadas variáveis dependentes e independentes. Variável dependente é o atributo a ser previsto, enquanto que a variável independente é o atributo que ajuda a prever o valor da variável dependente. Para classificar novas amostras, uma árvore de decisão com base nos valores de atributos no conjunto de treinamento deve ser criada [71].

### 2.1.2 Naive Bayes

O classificador Naive Bayes é baseado em um método probabilístico, no qual é considerado um grupo de probabilidades, computando as combinações de valores e frequência em um conjunto de dados. Neste classificador é utilizado o teorema de Bayes, assumindo que dado o valor da variável de classe, qualquer atributo é independente [77]. O classificador Naive Bayes tem sido utilizado em aplicações de diversas áreas, como, por exemplo, classificação de texto, diagnóstico médico e gerenciamento de desempenho de sistemas. Na Figura 2.2 é ilustrada uma rede do classificador Naive Bayes. Neste exemplo é capturada a principal suposição relacionada ao algoritmo bayesiano. Cada folha da rede é independente do restante dos atributos, dado o estado da variável de classe (a raiz da rede) [70].

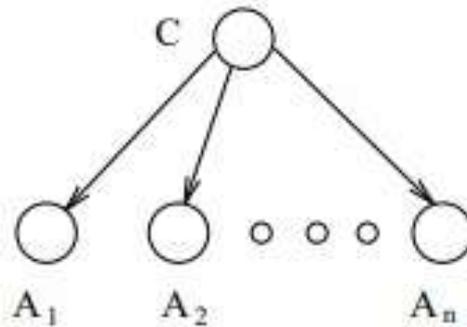


Figura 2.2: Rede Naive Bayes [70].

### 2.1.3 Rede Neural Perceptron Multicamadas

O classificador MLP foi uma solução de AM famosa na década de 80, que utiliza a abordagem de redes neurais artificiais e pode ser visto como o tipo mais simples de rede neural *feedforward*. Nenhuma saída de um neurônio de uma camada  $k$  será início de entrada para um neurônio de uma camada inferior ou igual a  $k$ . A rede é absolutamente conectada e cada neurônio oferece sua resposta para cada item da camada seguinte [50, 106]. Este classificador utiliza a técnica de aprendizado supervisionado chamada de *backpropagation*, cujo objetivo é calcular um gradiente necessário para realizar o ajuste dos pesos que são utilizados na rede [73].

Este classificador baseia-se em um bloco de unidades sensoriais (Figura 2.3) que constitui três tipos de categorias: camada de entrada, camadas ocultas e camadas de saída. A primeira categoria é composta pelos sinais de entrada. A segunda categoria é composta por camadas de neurônios que se encontram entre as camadas de entrada e de saída. Por fim, a terceira categoria é composta por neurônios que apresentam as saídas da rede. Uma única camada oculta é suficiente para o classificador MLP computar uma aproximação para um dado conjunto de treinamento. A quantidade total de neurônios da camada de saída sempre são conhecidos e iguais a quantidade de saídas que se deseja prever [40].

### 2.1.4 K-vizinhos Mais Próximos

KNN é um classificador de aprendizado não paramétrico e "tardio", dado que não constrói o classificador no momento do treinamento. A construção é atrasada até o momento da

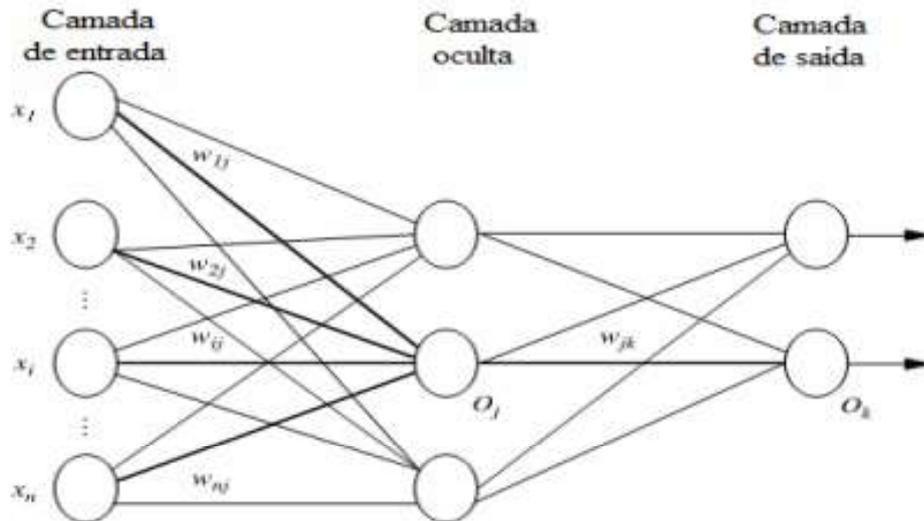


Figura 2.3: Arquitetura Rede MLP, adaptado de [50].

classificação [101]. É uma das técnicas de classificação mais simples, com menos ou nenhum conhecimento prévio de distribuição de dados. Normalmente, a classificação envolve particionar amostras em atividades de treinamento e teste [12].

Huang et al. (2018) [48] descreveu que, quando uma amostra de teste é fornecida, o classificador KNN pesquisa o espaço padrão  $n$ -dimensional dos dados de treinamento e localiza as  $k$  amostras de treinamento mais próximas da amostra para serem classificadas por uma determinada medida de distância. Posteriormente, a categoria é considerada a classe que tem os vizinhos mais próximos de  $k$ -vizinhos mais próximos [48].

Uma das principais dificuldades na classificação é que cada dado da amostra recebe igual importância quando o rótulo da classe deve ser atribuído aos dados de entrada. Isso causa problemas quando as amostras se sobrepõem. Outra dificuldade é que, uma vez que uma amostra de entrada recebe um rótulo de classe, não há informações sobre sua força de associação nessa classe [101]. KNN é um classificador baseado em aprendizado por analogia, no qual um registro de teste é comparado com um registro de treinamento, considerando similaridades [14].

### 2.1.5 Máquina de Vetores de Suporte

O classificador SVM tem sido usado desde 1992, quando havia a necessidade de ferramentas de classificação e regressão com base em previsões. Esta técnica foi apresentada por Vapnick, Guyon e Boser no COLT-92 [99]. O classificador SVM é um método de aprendizado supervisionado que, quando utilizado, pode evitar vários problemas relacionados com métodos não supervisionados (*e.g.*, *cluster* hierárquico e mapas de auto-organização) [16].

Este classificador pode ser definido como uma ferramenta de previsão, na qual procura-se uma linha ou limite de decisão específico (*i.e.*, hiperplano) que separa conjuntos de dados ou classes, evitando o excesso de ajuste nos dados [99]. Entretanto, ao usar o classificador, requisitos de computação e armazenamento aumentam rapidamente com o número de treinamentos de vetores [42].

Para separar qualquer dado, certas classes são definidas, e, dependendo da complexidade dos conjuntos de dados, a classificação é definida como classificação linear ou não linear. O classificador SVM é eficiente e escalável para problemas muito grandes, com centenas de milhares de vetores de treinamento. Ao invés de analisar o conjunto de treinamento em uma única etapa de otimização, os dados são divididos em subconjuntos e otimizados separadamente com vários SVMs. Neste contexto, o espaço de hipótese de um espaço linear em um espaço de característica de alta dimensão é utilizado. É possível também classificar dados não lineares, usando funções de kernel [99, 42].

### 2.1.6 Floresta Aleatória

O classificador floresta aleatória (*Random Forest* - RF) distingue-se de outros classificadores devido, principalmente, às seguintes características: (1) algoritmo mais robusto, quando comparado exclusivamente a uma árvore de decisão; (2) alta taxa de acerto, se for experimentado em diferentes conjuntos de dados; (3) técnica exata; (4) evita *overfitting*; (5) menos sensível a ruídos; e (6) classificação aleatória das árvores sem intervenção humana [65].

O classificador RF é utilizado para combinar árvores de decisão, de forma que cada árvore necessita de valores de vetores apresentados ao acaso e de forma independente, repartidos igualmente para todas as árvores [4]. Na Figura 2.4 é apresentado o método de classificação floresta aleatória, onde, partindo de uma base de dados, o classificador gera

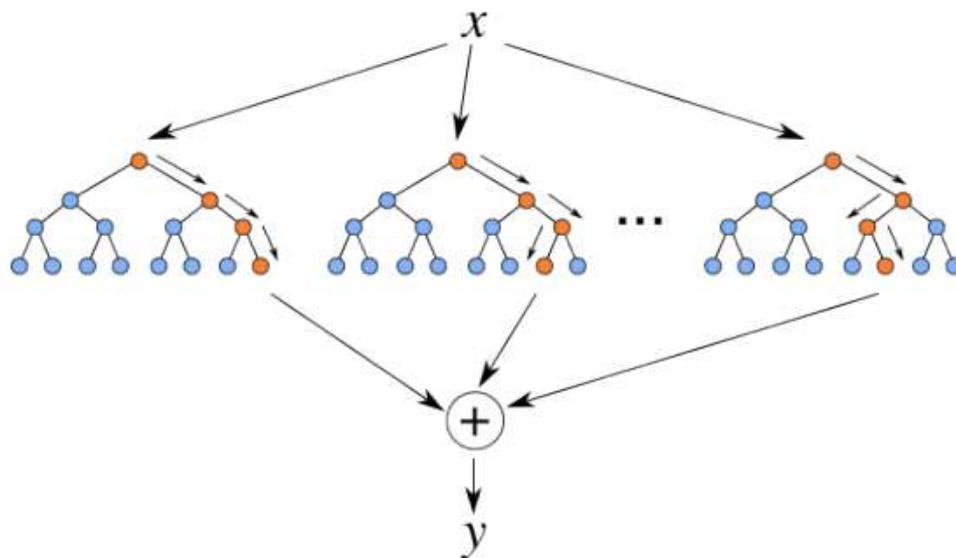


Figura 2.4: Ilustração da lógica relacionada ao algoritmo da floresta aleatória [65].

diversas árvores aleatórias que serão utilizadas para obter um resultado final. Assim, regras são geradas, possibilitando a descoberta de novos padrões que podem se tornar decisivos na tomada de decisão correta [65]. Neste método, é escolhido um subconjunto distinto de dados de treinamento para treinar cada árvore, onde os dados de treinamento restantes servem para estimar o erro e a importância da variável [12]. Com isso, pode-se determinar qual árvore de decisão possui as regras mais precisas para a mineração. Posteriormente, o classificador é aplicado nos dados, alcançando um resultado  $Y$  [65].

## 2.2 Weka

O Weka© é um ambiente de trabalho usado para auxiliar na aplicação de técnicas de AM a uma variedade de problemas do mundo real [45], distribuído sobre os termos da GNU *General Public License* e desenvolvido em Java. O Weka© foi desenvolvido em 1992, motivado pela necessidade de disponibilizar acesso a técnicas de AM. Este ambiente possui ampla aceitação na academia e nos círculos de negócios, tornando-se uma ferramenta amplamente usada para pesquisa de mineração de dados [43].

O Weka© pode ser utilizado da seguinte maneira: interface gráfica, linha de comando e *Application Programming Interface* (API). As interfaces gráficas de usuário (*Graphical User*

*Interface* (GUI) são fornecidas para a análise de dados, visualização de resultados, vinculação de banco de dados, validação cruzada, comparação de conjuntos de regras, configuração de experimentos em larga escala em plataformas de computação distribuída e projeto de configurações para processamento de dados em fluxo [35, 45]. A linha de comando é aplicada com o intuito de agilizar processos e facilitar o acesso a funcionalidades com diversos passos a serem executados. Por fim, o acesso por meio da API destina-se para desenvolvedores Java, fornecendo, de forma prática, o uso das funcionalidades implementadas no Weka© [29].

Na GUI, os dados podem ser inseridos a partir de várias fontes, incluindo arquivos, *Uniform Resource Locator* (URL) e bancos de dados. Os arquivos podem estar no formato *Attribute-Relation File Format* (ARFF), *Comma-Separated Values* (CSV), LibSVM e C4.5 [43]. A GUI do Weka© possui um conjunto de painéis, que podem ser usados para executar uma determinada tarefa. O painel de pré-processamento, apresentado na Figura 2.5, recupera dados de um arquivo. Uma limitação é que todos os dados são mantidos na memória principal. Portanto, subamostragem pode ser necessária para conjuntos de dados muito grandes [36].

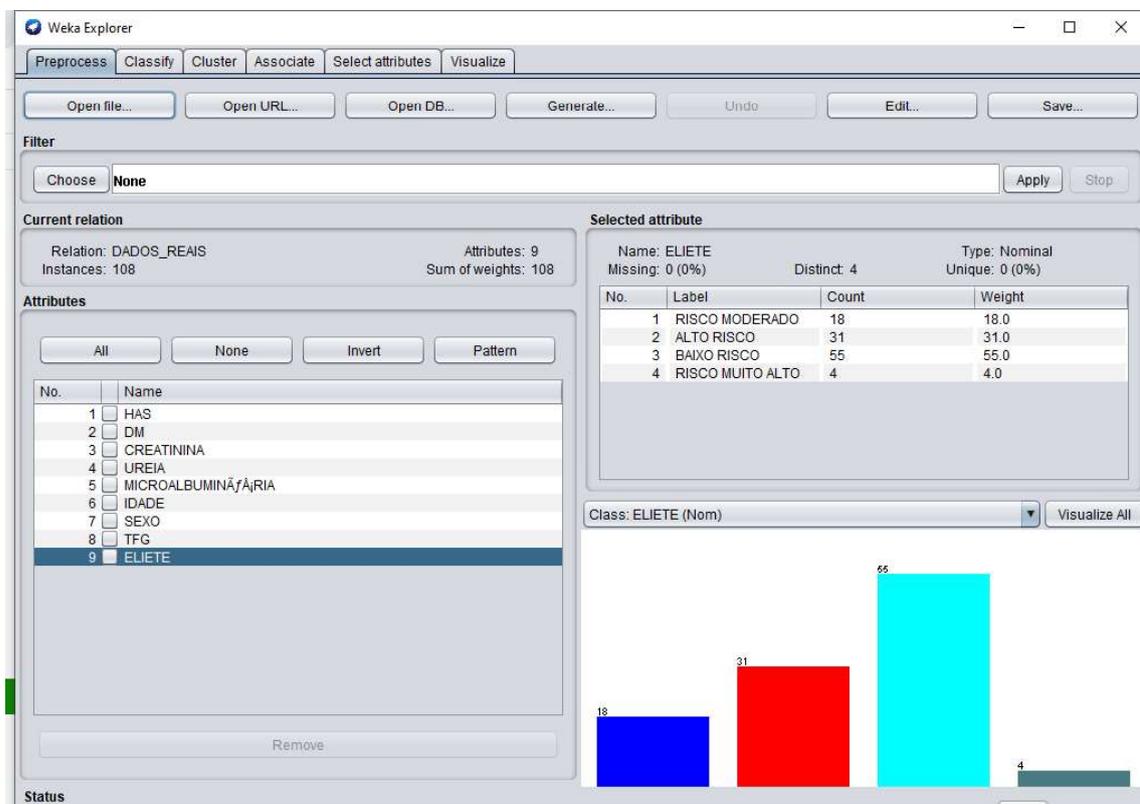


Figura 2.5: A interface do usuário do Weka Explorer.

No painel é apresentado um histograma do atributo que está atualmente selecionado e algumas estatísticas sobre histogramas para todos os atributos, que podem ser apresentados simultaneamente em uma janela separada. Os dados podem ser pré-processados usando uma das ferramentas de filtragem do Weka©. Por exemplo, é possível excluir todas as instâncias (*i.e.*, linhas) nos dados para qual determinado atributo (*i.e.*, coluna) possui um valor específico. Na Figura 2.6 é apresentado o painel de classificação, fornecendo um conjunto de algoritmos de AM e pré-processamento de dados. É permitido que os usuários testem e comparem diferentes métodos de AM usando um conjuntos de dados [45]. Nesta caso, são apresentados os resultados das classificações, onde é possível verificar quais foram as instâncias classificadas corretamente e incorretamente, a matriz de confusão e detalhes sobre acurácia.

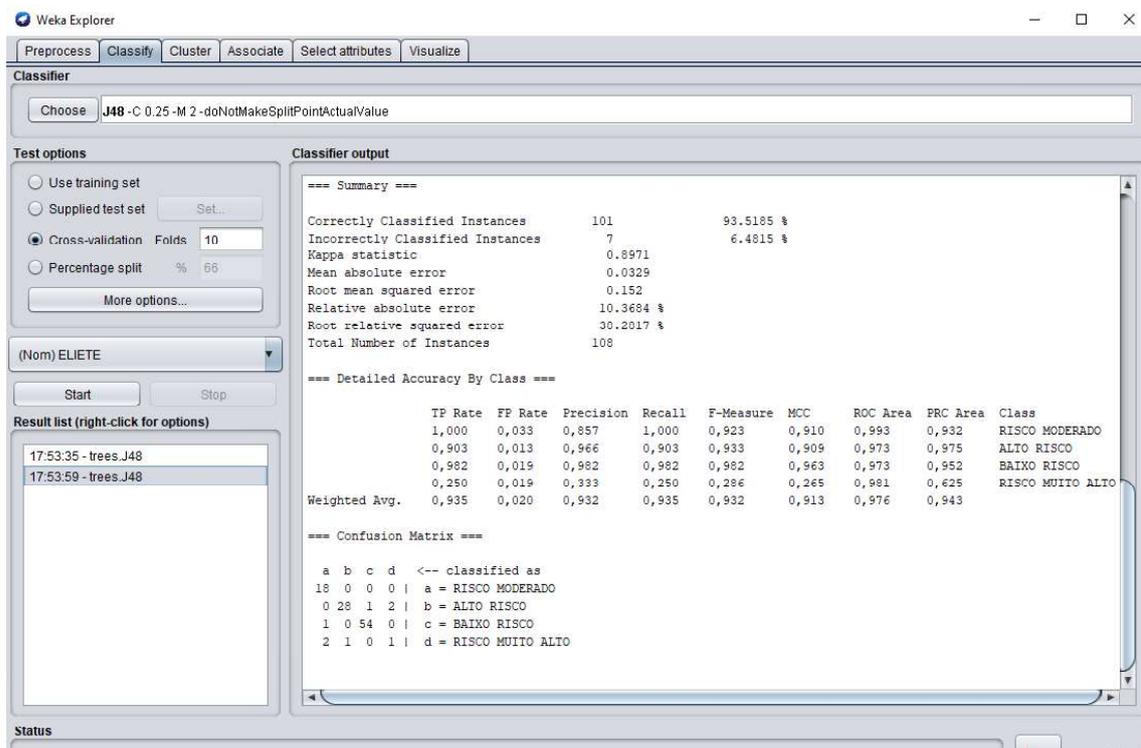


Figura 2.6: Interface de classificação do Weka Explorer.

O Weka© fornece quatro opções de configuração para a realização de testes. Na primeira opção (*Use training set*), os dados são divididos em duas partes: treinamento e teste. Os dados de treinamento são usados para induzir conhecimento, já os de teste são aplicados para analisar qualidade. Na segunda opção (*Supplied test set*), é necessário informar

o conjunto de teste. Na terceira opção, é utilizada a técnica de validação cruzada (*Cross Validation*). No pseudocódigo apresentado na Figura 2.7 é descrita a lógica básica e o funcionamento convencional da validação cruzada. Nesta técnica, cada dado do conjunto de teste é experimentado. Com isso, uma série de dados estatísticos é utilizada, como, por exemplo, a acurácia (taxa de acerto). Após dois laços de repetição, a validação cruzada é finalizada, produzindo dados estatísticos sobre o conhecimento induzido [78].

```
CrossValidation(numeroFolds      Inteiro,      minerador
AlgoritmoDeMineracao)

De i=1 até numeroFolds faça
  Cria subconjunto de treinamento Tri
  Classificador C = Aplica minerador em Tri
  Cria subconjunto de teste Tei

  De j=1 até (Numero de instancias de Tei) faça
    Testa (prediz) a instância j de Tei em C
  Fim laço.
Fim laço.
Gera dados estatísticos a partir da avaliação realizada.

Fim CrossValidation
```

Figura 2.7: Algoritmo Cross-Validation em pseudocódigo [78].

Por fim, ainda como parte da terceira opção do Weka©, é possível utilizar a técnica *Hold-Out*, também conhecida como *Percent split*. Nesta técnica, são aceitos como entrada os dados referentes ao percentual no qual estes devem ser separados em treinamento e teste, além do algoritmo de mineração a ser utilizado. Na Figura 2.8 é apresentado o pseudocódigo do algoritmo relacionado [78]. Portanto, o usuário pode escolher entre a validação cruzada e o *Hold-Out* durante a configuração de um experimento.

## 2.3 Doença Renal Crônica

Um dos órgãos essenciais para o corpo humano e manutenção da homeostase são os rins. Com isso, a perda gradativa da função renal acarreta no comprometimento de vários órgãos [8]. Segundo a definição da *National Kidney Foundation - Kidney Disease Outcomes Quality Initiative (US NKF-KDOQI)*, a DRC é reconhecida quando alguns dos seguintes critérios

```

HoldOut(percentagem      Inteiro,      minerador
AlgoritmoDeMineracao)

Gera subconjunto de treinamento Tr com percentagem% da
totalidade da amostra
Gera subconjunto de teste Te com o restante dos dados da
amostra, (100-percentagem)%
Classificador C = Aplica minerador em Tr
De j=1 até (Numero de instancias de Te) faça
    Testa (prediz) a instância j de Te em C
Fim laço.
Gera dados estatisticos a partir da avaliação realizada.

Fim HoldOut.

```

Figura 2.8: Algoritmo Hold-Out em pseudo-código [78].

são identificados: (1) redução da função renal identificada por Taxa de Filtração Glomerular (TFG) < 60 mL/min/1,73 m<sup>2</sup> e/ou (2) existência de marcador de comprometimento renal (Razão Albumina/Creatinina (RAC) > 30 mg/g), ainda que da causa subjacente da DRC, superior a três meses [15].

Os rins exercem muitas funções no organismo humano. A função renal pode ser analisada pela redução da TFG, que está relacionada com a perda das funções regulatórias, excretórias e endócrinas do rim. Quando o resultado atinge valores menores do que 15 mL/min/1,73m<sup>2</sup>, denomina-se de Falência Funcional Renal (FFR), ou seja, o nível mais alto de redução funcional progressiva examinada na DRC [8].

Quando a DRC é diagnosticada em estágios menos avançados, é possível reduzir a necessidade de diálise e transplante renal, melhorando a qualidade de vida dos pacientes [96]. Portanto, é necessário um modelo de atendimento com um planejamento estruturado, em que o médico da atenção primária e o especialista colaborem para a detecção e gerenciamento da DRC [9]. O diagnóstico e o monitoramento da DRC são tarefas complicadas porque os nefrologistas devem analisar fatores de risco que influenciam o desenvolvimento e a progressão da doença durante as avaliações de risco [96].

O processo para resolver os problemas referentes à DRC é complexo e necessita de três ações: (1) O diagnóstico antecipado da DRC; (2) O direcionamento imediato para acompanhamento especializado; e (3) O diagnóstico e a correção das complicações e comorbidades

Tabela 2.1: Estagiamento da DRC proposta pelo grupo de trabalho do K/DOQI [8].

Estágio	Descrição	TFG
1	Lesão renal com TFG normal ou aumentada	$\geq 90$
2	Lesão renal com TFG levemente diminuída	60-89
3	Lesão renal com TFG moderadamente diminuída	30-59
4	Lesão renal com TFG severamente diminuída	15-29
5	FFR estando ou não em terapia renal substitutiva	$<15$

essenciais da DRC, bem como o encaminhamento do paciente para a Terapia Renal Substitutiva (TRS) [9].

Dessa forma, pode-se identificar a DRC usando a lesão do parênquima renal e/ou pela redução da funcional renal identificada no período acima de três meses. Neste contexto, a classificação (estagiamento) para a DRC é apresentada na Tabela 2.1 [8]. O diagnóstico tardio é um problema crítico, principalmente em países em desenvolvimento, que geralmente sofrem com cuidados de saúde primários precários em contextos específicos, como, por exemplo, ambientes rurais e de difícil acesso.

## 2.4 Sistemas Especialistas

Os Sistemas Especialistas (SE) possuem como objetivo substituir ou auxiliar um perito em uma determinada área do conhecimento, como, por exemplo, no diagnóstico médico. São fundamentados em conhecimento construído, principalmente, usando regras que representam o conhecimento do especialista [68]. Segundo Moyyad *et al.* (2016) [39], SE estão relacionados com quatro tipos de atividades:

- aquisição de conhecimento (o conhecimento é coletado de especialistas humanos e documentado para a próxima etapa).
- representação do conhecimento (formatar o conhecimento adquirido como um conjunto de regras que podem ser facilmente codificadas no programa).

- inferência de conhecimento (executa a função de inferência lógica para alcançar a decisão ou recomendação final).
- transferência de conhecimento (a decisão final, que deve corresponder à decisão do especialista humano, é informada ao usuário).

Os SE podem contribuir com a diminuição de enganos em avaliações médicas realizadas por profissionais de saúde, dada a combinação sensata de conhecimento adquirido de muitos especialistas humanos [39]. Entre as principais desvantagens, em comparação com especialistas humanos, está o fato que um especialista humano pode apresentar resultados criativos para casos especiais, enquanto que os SE apenas cumprem com regras pré-programadas [39]. Uma aplicação que auxilia na tomada de decisão representa as reais necessidades de usuários, dando-lhes orientação, reproduzindo cenários e proporcionando exatidão e maior confiança nos resultados [68].

Neste contexto, o *Expert Sinta* é uma das ferramentas usadas para simplificar o projeto deste tipo de sistema. As principais características do *Expert Sinta* incluem: (1) utilização do encadeamento para trás (*backward chaining*); (2) utilização de fatores de confiança; (3) ferramentas de depuração; e (4) chance de acrescentar ajudas on-line em cada base.

Na Figura 2.9 é apresentada a arquitetura do *Expert Sinta*, composta pelos seguintes componentes:

- a base de conhecimentos (representa a informação (fatos e regras) que um especialista utiliza, representada computacionalmente);
- o editor de bases (permite a implementação das bases desejadas);
- a máquina de inferência (responsável pelas deduções sobre a base de conhecimentos);
- o banco de dados global (composto por evidências apontadas pelo usuário durante uma consulta).

## 2.5 Considerações Finais

Neste capítulo foram apresentados, brevemente, alguns termos e conceitos utilizados nesta pesquisa. Procurou-se descrever o método de aprendizado mais utilizado (aprendi-

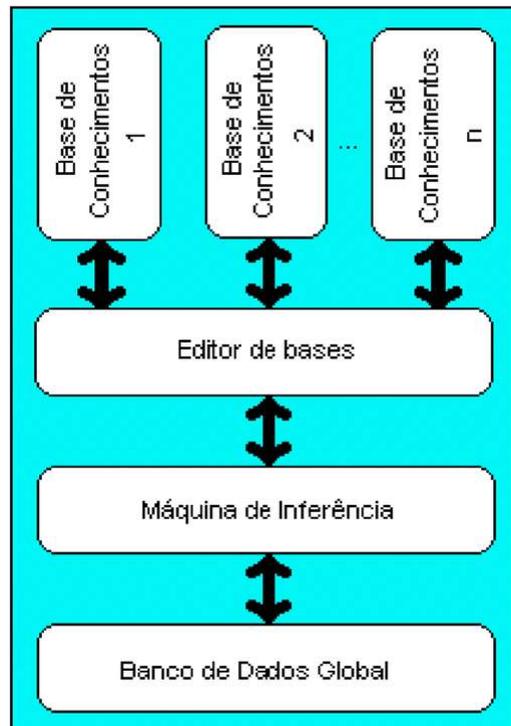


Figura 2.9: Arquitetura simplificada do *Expert Sinta* [95].

zado indutivo). Porém, outros três tipos de aprendizado podem ser encontrados na literatura: aprendizado semissupervisionado, aprendizado por reforço e aprendizado ativo. O aprendizado semissupervisionado pode ser utilizado quando dados não estão rotulados, mas são conhecidos ou encontram-se em grupos diferentes [18]. No aprendizado por reforço, o agente aprendiz aprende de forma autônoma ativamente por experimentação direta, sem ser ensinado por meio de exemplos fornecidos por um supervisor [13]. No aprendizado ativo, é permitido que o algoritmo de aprendizado determine, dentro do conjunto de exemplos não rotulados, os melhores exemplos a serem rotulados [22].

Foi também apresentado o ambiente de trabalho Weka©, utilizado neste trabalho para realizar o experimento de AM. Procurou-se também descrever os conceitos sobre DRC, ressaltando a importância do diagnóstico precoce para reduzir a necessidade de diálise e transplante renal.

# Capítulo 3

## Trabalhos Relacionados

Neste capítulo são apresentados alguns trabalhos relacionados ao conceito de saúde inteligente, diagnóstico e monitoramento de DRC e aplicações para o encaminhamento de pacientes para unidades de saúde.

### 3.1 Saúde Inteligente

O ambiente inteligente tem ganhado foco na área de saúde, com o propósito de promover uma melhor qualidade de vida para seus usuários, auxiliar na detecção de um cenário de emergência e proporcionar uma vida mais independente. Pesquisadores têm discutido sobre ambientes inteligentes para, por exemplo, monitoramento em tempo real de pessoas com problemas de saúde, utilizando alertas inteligentes para: monitoramento de pressão arterial [33]; monitoramento de pacientes com Alzheimer [85]; monitoramento de frequência cardíaca [56]; monitoramento contínuo [24]; e monitoramento visual do cenário da emergência em tempo real [41].

As casas inteligentes, que oferecem instalações de jogos interativos, têm um efeito positivo na saúde dos idosos [76], visto que os jogos de computador ajudam os pacientes a estimular o cérebro e reduzir os sintomas da doença [110], bem como o uso fisioterapêutico para pacientes com Acidente Vascular Cerebral (AVC) [33] e para uma melhor adesão à medicação [24]. Encontram-se também serviços robóticos para providenciar companheirismo [32], usados para reduzir os sentimentos de solidão, tédio e isolamento social entre os idosos, soluções baseadas em vídeo que são muito eficazes para melhorar a sua satisfação mental e

a comunicação social [76].

A inovação permite o desenvolvimento de aplicações úteis na computação ubíqua utilizando a tecnologia *Radio-Frequency Identification* (identificação por radiofrequência) nos campos da saúde e da habitação. Uma das aplicações citadas na área médica é o implante de *tags* (etiquetas) em humanos contendo as informações de um paciente, facilitando o trabalho médico e que chegue ao hospital [33].

Um dos avanços tecnológicos de maior destaque dos últimos anos é a aplicação da Internet das Coisas (*Internet of Things* (IoT)) na saúde. Uma rede que conecta o dispositivo sensor para dispositivos médicos ou saúde para possibilitar que esses dispositivos se comuniquem com sistemas ou outros dispositivos na rede [86]. A integração de vários componentes de IoT para sistemas médicos torna-se cada vez mais popular, principalmente para eventos como detecção de quedas e detecção de crises, onde os cuidadores podem fornecer um melhor tratamento e tomar medidas imediatas para evitar uma situação perigosa durante uma convulsão, por exemplo [1]. A IoT também vem sendo utilizada para gestão hospitalar, com o objetivo de prevenir infecções hospitalares, determinar um plano completo para educar pacientes e gerenciar situações de emergência e sistemas logísticos [41, 1].

Na área de Saúde Eletrônica (*e-health*), os dispositivos médicos são conectados à Internet para realizar vários serviços de telessaúde, como tele-monitoramento, supervisão de idosos, teleconsultas e cirurgia assistida por robô [1]. Outra área que vem atraindo pesquisadores são os agentes inteligentes na área da saúde. Os sistemas multi-agentes são aplicados desde sistema médico, baseado no conhecimento, até sistemas complexos, baseados em múltiplos componentes, como a unidade de cuidados de saúde completos [85]. Atualmente, boa parte dos estudos tem abordado sistemas de suporte à decisão multi-agente clínico [67], sistemas de saúde domiciliares, simulação de triagem e otimização de um sistema de recuperação de dados clínicos baseados em agentes [85].

O suporte a decisão clínica é um importante direcionador para a *Health information technology* (HIT), porque pode apoiar a tomada de decisões clínicas [66], dar apoio à decisão em terapia intensiva neonatal [67], dar suporte à decisão médica em aplicações de oftalmologia [31], melhorar os resultados dos pacientes e reduzir erros e despesas desnecessárias [66].

## 3.2 Aplicações para a Doença Renal Crônica

Nesta seção são apresentados trabalhos relacionados diretamente com a solução proposta nesta dissertação. Sistemas de software tem sido desenvolvidos para auxiliar os médicos durante o monitoramento e diagnóstico da Doença Renal Crônica (DRC). Por exemplo, CKD-Go<sup>1</sup> é um aplicativo da Web que permite a visualização de um plano de ação personalizado, com base na taxa de TFG e creatinina de albumina na urina. eGFR<sup>2</sup> auxilia os profissionais a estimarem a função renal utilizando cinco calculadoras TFG separadas. Isto inclui uma relação de referências fáceis de serem utilizadas e outras informações para ajudar os profissionais a identificarem os fatores de risco, avaliar a DRC e gerenciar a progressão usando estratégias baseadas em evidências das diretrizes KDOQI. A calculadora eGFR estima a TFG do usuário usando as equações propostas por Cockcroft-Gault [26], modificação de dieta em doença renal [60], colaboração em epidemiologia da doença renal crônica [61] e Schwartz [91]. O aplicativo assistente para doenças renais<sup>3</sup> permite que os usuários registrem os resultados de exames laboratoriais e visualizem sua função renal. Por fim, o aplicativo Android MultCare fornece recursos para ajudar os usuários a realizar o auto-monitoramento de seus riscos de DRC nos países em desenvolvimento, com base em diretrizes médicas [96].

Técnicas relacionadas com AM também tem sido aplicadas para possibilitar que sistemas de software sejam usados para monitorar e diagnosticar doenças crônicas [63]. Por exemplo, Topuz *et al.* (2018) [104] aplicaram um algoritmo de rede de crenças bayesiana para prever a sobrevivência do enxerto no transplante renal. O método combina seleção exaustiva de variáveis com técnicas de AM para permitir melhores habilidades preditivas e identificação de fatores importantes que contribuem para o risco de falha do enxerto e as dependências probabilísticas. Tem como objetivo criar uma estrutura de previsão e exploração multinomial que possa gerar informações adicionais considerando três classes de risco.

Jahantigh *et al.* (2017) aplicou lógica nebulosa para projetar um sistema especialista para diagnosticar doenças renais que empregava um conjunto de funções e regras de associação difusa [51]. O protótipo do sistema especialista forneceu um alto nível de precisão (63,3%) para cálculos renais e níveis mais baixos para todas as outras infecções suspeitas.

<sup>1</sup>[kidney.org.au/health-professionals/detect/calculator-and-tools/ckd-goe](http://kidney.org.au/health-professionals/detect/calculator-and-tools/ckd-goe)

<sup>2</sup><https://www.kidney.org/apps/professionals/egfr-calculator>

<sup>3</sup>[itunes.apple.com/br/app/kidney-disease-assistant/id830127960?mt=8](https://itunes.apple.com/br/app/kidney-disease-assistant/id830127960?mt=8)

Neves *et al.* (2015) aplicou Redes Neurais Artificiais (RNAs) para identificar lesão renal aguda e DRC, desenvolvendo um sistema híbrido de apoio à decisão em termos de representação do conhecimento e procedimentos de raciocínio baseados na programação lógica, que permite considerar informações incompletas, desconhecidas e até contraditórias. O objetivo foi pesar o grau de confiança que se tem em tal acontecimento [72].

Polat *et al.* (2017) [79] aplicou a técnica de máquina de vetores de suporte (*Support-Vector Machine (SVM)*) e o *wrapper* e filtro do método de seleção de dois recursos para identificar a DRC nos estágios iniciais. Utilizaram o avaliador de subconjunto de classificador com mecanismo de pesquisa passo a passo ganancioso e o avaliador de subconjunto de invólucro com o mecanismo de pesquisa *Best First* que obteve a taxa de precisão mais alta (98,5%) no diagnóstico de DRC utilizando os 24 atributos do repositório de AM UCI [80].

### 3.3 Aplicações para o Encaminhamento para Unidades de Saúde

Nesta seção são apresentados trabalhos relacionados com a proposta de sistemas de software para o encaminhamento de emergência para unidades de saúde. Por exemplo, Alan e Rajib (2014) [30] apresentam um trabalho desenvolvido em Bangladesh, com o intuito de auxiliar pessoas que precisem das informações dos centros de saúde mais próximos (por exemplo, hospitais, clínicas e centros de diagnóstico) logo após a ocorrência dos incidentes. Teve como justificativa o aumento dos acidentes rodoviários e o fato de que informação instantânea pode salvar a vida das vítimas desses acidentes. Assim, propuseram um sistema de assistência médica de emergência baseado em localização usando o *OpenStreetMap (OSM)*, onde os centros de saúde são mapeados, obtendo *waypoints* [30].

Ganesan e Chamundeeswari (2020) [38] abordaram o problema de encontrar a rota otimizada para os serviços de saúde disponíveis mais próximos. Destaca-se a importância de ter informações como a disponibilidade de instalações de saúde, por exemplo, farmácia, hospital, clínica e banco de sangue. Propõem utilizar algoritmos para o processo de filtragem e verificação do caminho e distância para encontrar o caminho otimizado. As informações são obtidas usando o servidor da Web do Google (dados em tempo real) usando a *Application Programming Interface (API)* do Google.

Na Malásia, foi realizada uma pesquisa para detectar as informações necessárias para os pacientes, como a disponibilidade de conhecimento dos hospitais do governo, que não estão disponíveis com facilidade. Mohamad *et al.* (2017) [64] visaram resolver o problema realizando uma análise de informação com base no Sistema de Informações Geográficas (SIG) para que os pacientes possam descobrir as informações sobre o tratamento necessário em um hospital com base nos serviços especializados fornecidos. O sistema apresenta a visão gráfica dos hospitais na superfície da terra. A partir dessa informação, o usuário tem uma melhor compreensão para encontrar os hospitais a partir de sua localização, ao invés de possuir apenas o endereço dos hospitais do governo como referência [64].

### 3.4 Considerações Finais

Neste capítulo foram apresentados alguns trabalhos relacionados com a solução proposta. No entanto, estudos existentes apresentam algumas limitações, que indicam pelo menos uma lacuna na pesquisa: até o momento, não foi identificado nenhum estudo abordando o uso de técnicas de AM para o diagnóstico de DRC a partir da perspectiva de problemas específicos nos países em desenvolvimento, como altos níveis de pobreza e ambientes de difícil acesso.

No próximo capítulo são apresentados relacionados com análises comparativas qualitativas e quantitativas, que foram, respectivamente, realizadas usando uma Revisão Sistemática da Literatura (RSL) e um experimento com classificadores baseados em AM, usando o software Weka© e um conjunto de dados DRC. Os trabalhos apresentados na RSL são também considerados como trabalhos relacionados com a solução apresentada nesta dissertação.

# Capítulo 4

## Análise Comparativa

Neste capítulo é apresentada uma análise da técnica de Aprendizado de Máquina (AM) mais adequada, discutindo como os sistemas de software podem ser usados para auxiliar no diagnóstico precoce da Doença Renal Crônica (DRC) nos países em desenvolvimento, devido aos problemas específicos enfrentados por essas populações, como a atenção primária à saúde inadequada. Nesse sentido, a análise comparativa foi guiado por duas Questões de Pesquisa Principais (QPP).

- **QPP 1:** Qual é o classificador baseado em AM mais adequado para o diagnóstico de DRC em países em desenvolvimento?
- **QPP 2:** Quais são os atributos mais adequados para o diagnóstico de DRC em países em desenvolvimento?

Portanto, esta análise compreende três contribuições principais: (1) os resultados de uma RSL focada no diagnóstico de DRC auxiliado por AM; (2) os resultados de um experimento baseado em técnicas de AM usando um conjunto de dados de DRC coletado de um hospital de um país em desenvolvimento e analisado por nefrologistas; e (3) uma discussão sobre a adequação das técnicas de AM nos países em desenvolvimento. Resultados relacionados com a análise comparativa foram publicados no periódico IEEE Access (Sobrinho *et al.*, 2020) [97], que possui *Journal Citation Reports* 2019 de 3.745<sup>1</sup>.

---

<sup>1</sup><https://ieeaccess.ieee.org/>

## 4.1 Revisão Sistemática da Literatura

O software Start© foi utilizado para definir o protocolo da RSL considerando três fases: (1) Questões de Pesquisa (QP), (2) processos de busca e (3) critérios de inclusão e exclusão. O protocolo foi estruturado com base nas diretrizes de RSL em engenharia de software propostas por Kitchenham *et al.* (2007) [55] e Wohlin (2014) [108].

### 4.1.1 Protocolo

O objetivo principal com esta RSL foi responder a seguinte QP principal: "Qual é o classificador baseado em AM mais adequado para o diagnóstico da DRC?". Com base na QPP, três QP secundárias foram definidas:

- **QP 1.1:** Quais são os classificadores baseados em AM atualmente aplicados no diagnóstico de DRC?
- **QP 1.2:** Quais são os atributos atualmente considerados ao usar classificadores baseados em AM para o diagnóstico de DRC?
- **QP 1.3:** Qual é o classificador baseado em AM com o mais alto nível de desempenho no diagnóstico da DRC em aplicações gerais?

Oito palavras-chave foram definidas para a busca de artigos com o objetivo de responder as QP: *chronic kidney disease, computer-aided, software system, software, system, machine learning, medical diagnosis, e diagnosis*. A seguinte sequência de pesquisa orientou a pesquisa inicial: `chronic kidney disease ^ (computer-aided v software system v software v system) ^ machine learning ^ (medical diagnosis v diagnosis)`.

Além disso, foram definidos dois critérios gerais de seleção: (1) os artigos foram revisados por pares e (2) o inglês foi o idioma usado para redigir o artigo. Os bancos de dados IEEE Xplore Library, ACM Digital Library, Scopus e PubMed foram utilizados para realizar as pesquisas iniciais com base nos critérios gerais de seleção e nos critérios de inclusão e exclusão. Na Tabela 4.1 são apresentados os três critérios de inclusão e cinco critérios de

exclusão definidos para a RSL. Durante a RSL, foram buscados estudos experimentais apresentando resultados quantitativos. Para melhorar a qualidade da pesquisa inicial, a RSL foi realizada também utilizando a técnica *snowballing* e buscas manuais no Google Scholar®.

Tabela 4.1: Critérios de inclusão e exclusão da revisão sistemática da literatura.

<b>Tipo</b>	<b>Descrição</b>
Inclusão	O artigo refere-se ao uso de AM.
Inclusão	O artigo refere-se ao uso de recursos e tecnologias para diagnóstico médico.
Inclusão	O artigo completo está disponível.
Exclusão	O artigo está duplicado.
Exclusão	Artigos curtos e pôsteres.
Exclusão	O artigo não está diretamente relacionado aos diagnósticos de DRC.
Exclusão	O artigo foi publicado antes de 2005.
Exclusão	O artigo compreende uma revisão de literatura.

Finalmente, o protocolo da RSL foi composto por 9 campos de extração de dados para orientar a resposta da QP principal e secundárias. Mais especificamente, campos de extração incluíram a data de publicação (2005-2019), a fonte do artigo (biblioteca digital IEEE Xplore, Biblioteca Digital ACM, Scopus ou PubMed), o tipo de pesquisa (empírica, experimental, laboratorial ou pesquisa de campo), o método de pesquisa (quantitativa ou qualitativa), o classificador baseado em AM, os atributos, o conjunto de dados e o tipo de publicação (conferência ou periódico/revista).

### 4.1.2 Resultados

Na Figura 4.1 é apresentada uma visão geral da abordagem RSL que integra a diretriz proposta por Kitchenham *et al.* [55] e a diretriz da técnica *snowballing* apresentada por Wohlin [108].

Na primeira etapa, a RSL foi realizada usando bases de dados e pesquisas manuais, resultando em 30 artigos (Figura 4.2). Após a leitura dos títulos e resumos dos 30 artigos e da aplicação dos critérios de exclusão, 15 artigos foram aceitos, 12 artigos foram rejeitados e 3 foram identificados como duplicados.

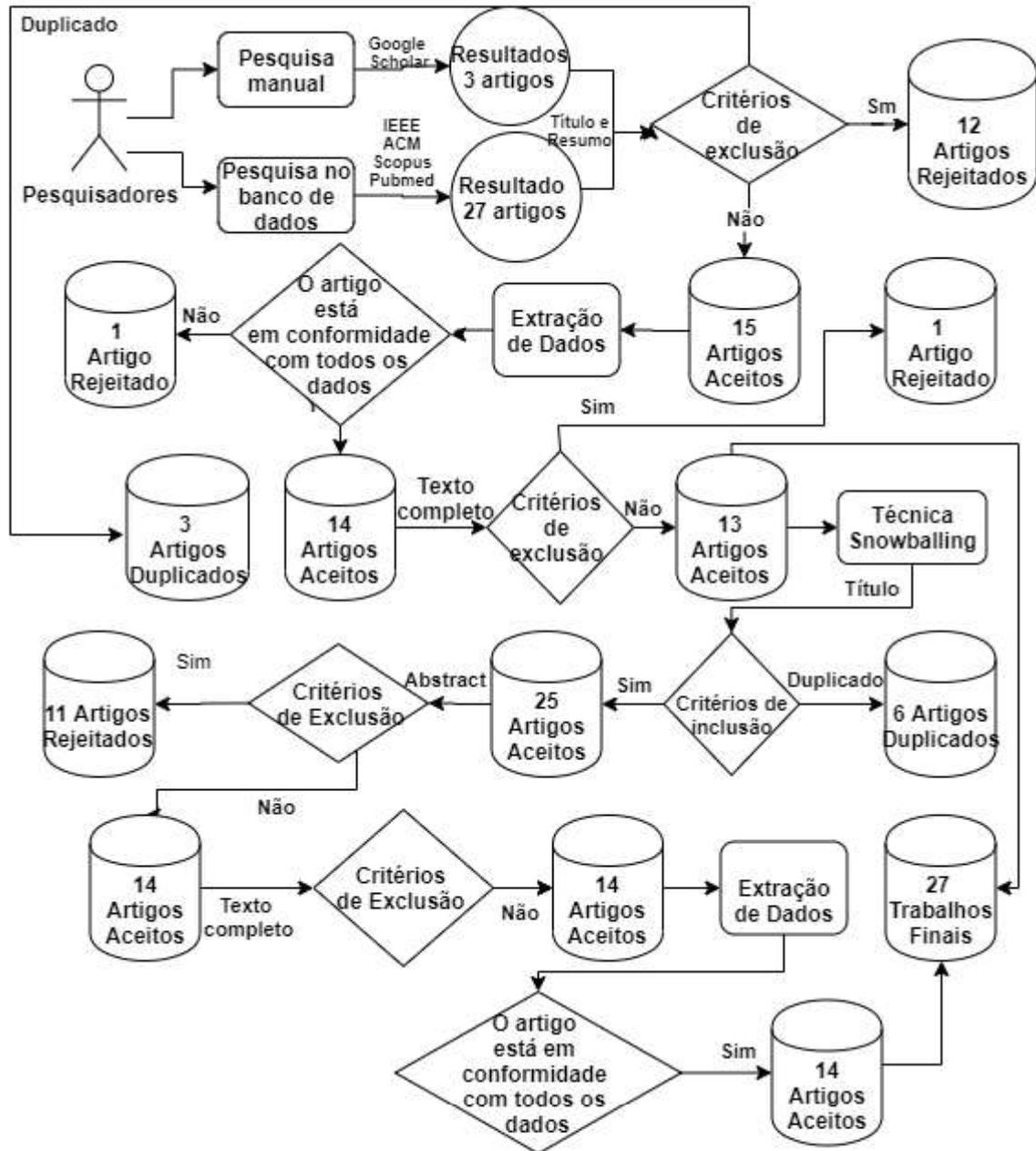


Figura 4.1: Visão geral da abordagem para a revisão sistemática da literatura, incluindo pesquisa em banco de dados, pesquisa manual e a técnica *snowballing*.

A segunda etapa envolveu a extração de dados, considerando a data de publicação, fonte de pesquisa, objetivos de pesquisa, método de pesquisa, técnicas, resultados de classificação, atributos, bases de dados e tipo de publicação. Nesta etapa, apenas 1 artigo foi rejeitado, resultando em 14 artigos aceitos. Na terceira etapa, os pesquisadores leram todos os artigos, analisando os critérios de exclusão, o que resultou em 13 artigos restantes.

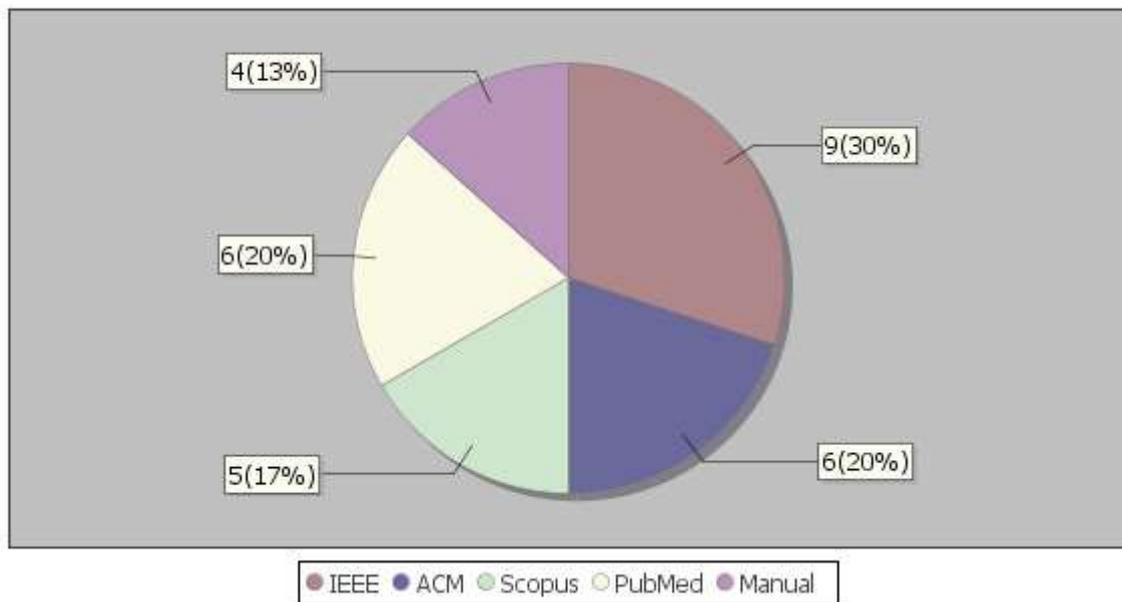


Figura 4.2: Fonte dos 30 trabalhos resultantes da busca inicial guiada pela abordagem utilizada na revisão sistemática da literatura.

Na quarta etapa, para melhorar a confiança nos resultados da pesquisa, foi aplicada a técnica *backward snowballing*, incluindo 25 artigos com base em pesquisas de referência dos 13 artigos aceitos inicialmente, lendo seus títulos. Nesta etapa, os critérios de inclusão foram aplicados para decidir o conjunto inicial de artigos. Em seguida, os resumos desses artigos foram analisados para aplicação dos critérios de exclusão, resultando em 14 artigos aceitos. A extração final dos dados foi realizada em 27 artigos aceitos.

Considerando os 27 artigos aceitos, a maior frequência de publicação por ano foi de 2016 com 8 artigos, seguida de 2017 com 5 artigos, 2015 com 4 artigos, 2013 com 2 artigos, 2018 com 2 artigos, 2010 com 1 artigo, 2012 com 1 artigo e 2019 com 1 artigo. A frequência de publicação por ano diminuiu de 8 para 1 artigo entre 2016 e 2019. Menos de 50% dos artigos foram publicados em periódicos (9 artigos), apenas 3 foram publicados em periódicos com fator de impacto (JCR).

Com base nos resultados da RSL, ao longo dos anos, foi possível observar o desinteresse no desenvolvimento de novos estudos com a aplicação de classificadores baseados em AM para o diagnóstico da DRC. Isto pode ser atribuído ao fato de que os estudos realizados, em conjunto, cobriram a análise de desempenho dos classificadores baseados em AM mais

conhecidos na literatura (detalhado nos próximos parágrafos). Na maioria dos estudos, os mesmos 24 atributos (ou seu subconjunto) do repositório de AM da Universidade da Califórnia Irvine (UCI) foram utilizados (18 dos 27 artigos aceitos) [79], [107] - [6] [53] [3] [74] [94] [23] [19] [54] [82] [7] [58] [57] [83] [21] [44] [44] [44] [89]. No entanto, os estudos selecionados apresentam algumas limitações, que indicam pelo menos uma lacuna de pesquisa: falta de análise comparativa dos classificadores baseados em AM considerando características específicas de países em desenvolvimento, como, por exemplo, altos níveis de pobreza.

Dos 27 artigos, em apenas 1, os autores abordaram o uso de AM para auxiliar no diagnóstico da DRC considerando o caso específico de países em desenvolvimento, com foco na população chinesa [109]. Entretanto, os autores apresentaram uma abordagem para estimar a TFG, sem considerar o diagnóstico final da DRC. Portanto, em nenhum dos trabalhos selecionados, foram consideradas as implicações de características específicas enfrentadas para realizar um diagnóstico completo da DRC em países em desenvolvimento ao aplicar classificadores baseados em AM.

Além disso, na maioria dos estudos, os autores não consideraram os custos associados ao uso de atributos necessários para realizar o diagnóstico da DRC. Os altos níveis de pobreza em países em desenvolvimento exigem uma abordagem de baixo custo no diagnóstico da DRC, analisando o conjunto mínimo de atributos bem aceitos pela comunidade de nefrologistas. Em apenas um dos artigos selecionados, os autores discutiram, de maneira resumida, os custos dos atributos da DRC durante o diagnóstico usando AM [89]. Salekin e Stan-kovic (2016) [89] demonstraram que o uso dos 24 atributos da base de dados UCI, custou aproximadamente \$ 451,36 USD ( *United States Dollar*), enquanto que o uso de 5 atributos de DRC, custou aproximadamente \$ 45,05 USD, apresentando um desempenho semelhante com os mesmos classificadores. Claramente, o primeiro cenário (\$ 451,36 USD) não é viável para populações de baixa renda. No entanto, os autores não discutiram se os atributos escolhidos são comumente usados na prática clínica por médicos da atenção básica que avaliam populações de baixa renda.

Em nenhum dos artigos, os autores discutiram a interpretação dos resultados dos classificadores baseados em AM pelos médicos da atenção básica. Esta é uma questão crítica a ser considerada em países em desenvolvimento devido à necessidade de uma avaliação

final do estado do paciente, que geralmente é realizada por médicos não especializados que trabalham em ambientes rurais e de difícil acesso. A comunidade médica geralmente se sente mais confiante em resultados de avaliação automatizada, fornecidos por sistemas de software, quando é possível interpretar e confirmar com facilidade os resultados antes de um diagnóstico final [102].

### **Respondendo a QP 1.1**

Dentre os classificadores baseados em AM abordados nos 27 estudos selecionados, com foco no diagnóstico da DRC, foram identificados: Redes Neurais Artificiais (RNA); *Support-Vector Machine* (SVM); *Naive Bayes* (NB); *linear discriminant analysis*; *k-nearest neighbor* (KNN); árvore de decisão; *random subspace*, *adaptive boosting*; REPTree; BFTree; árvore de decisão J48; *logistic regression*; *decision table*; *conjunctive rules*; *wide and deep learning*; XGBoost; *bagging*; *extra trees*; *gradient boosting*; *Random Forest* (RF); abordagens *non-temporal*, *stached-temporal*, *and multitask-temporal*; *sequential minimal optimization*; *instance-based learner*; *soft independent modeling of class analogy*; e *simple logistic*.

Na maioria dos estudos, algum tipo de RNA foi utilizada (13 dos 27 estudos aceitos). Classificadores SVM foram utilizados em 12 estudos; classificadores NB em 8 estudos; classificadores baseados em árvores em 7 estudos; classificadores KNN em 6 estudos; e AM baseado em conjuntos e regressão logística em 4 estudos. Os demais classificadores relatados nos trabalhos foram utilizados apenas por, no máximo, um ou dois estudos.

### **Respondendo a QP 1.2**

O conjunto de dados com o maior número de atributos usados para identificar a DRC foi composto por idade, pressão arterial, gravidade específica, albumina, açúcar, glóbulos vermelhos, células pus, aglomerados de células pus, bactérias, glicose no sangue, uréia, creatinina, sódio, potássio, hemoglobina, volume de células compactadas, contagem de glóbulos brancos, contagem de glóbulos vermelhos, hipertensão, DM, doença arterial coronariana, apetite, edema do pedal e anemia. Os conjuntos mínimos de atributos foram compostos por imagens do rim ultrassônicas ou valores de TFG.

Outros conjuntos de dados incluíram atributos como índice de massa corporal, histórico familiar, doenças renais associadas, fatores de risco primários, fatores de risco secundários, obesidade, reação em cadeia do polímero, colesterol, altura, peso, hematúria, proteinúria, ferritina sérica, cálcio, proteína total e fósforo.

### **Respondendo a QP 1.3**

Os classificadores NB, RF, árvore de decisão bagging/J48 e o subespaço aleatório/árvore de decisão J48 apresentaram maior precisão (100%) na análise dos 24 atributos relacionados com DRC do repositório UCI [7] - [58] [57]. Considerando os mesmos 24 atributos, o classificador *Multilayer Perceptron* (MLP) apresentou precisão de 99,75% [83], enquanto classificadores SVM e AdaBoost apresentaram precisão de 99% [21, 44]. O classificador RF apresentou precisão de 99,8% na análise de 23 atributos da DRC [89]. O classificador baseado em árvore de decisão apresentou precisão de 92,2% na análise de 15 atributos [2]. A precisão do classificador SVM ainda foi alta com apenas 12 atributos (98,5%) [79]. O classificador RNA apresentou precisão de 95,1% na análise de 11 atributos [81].

A rede de retropropagação apresentou precisão de 94,75% na análise de 10 atributos [25], enquanto o classificador SVM apresentou precisão de 76,32% na análise de 5 atributos [105]. Usando apenas um atributo, por exemplo, imagem de ultrassom, o classificador SVM apresentou precisão de 93,82%. Na Tabela 4.2 são resumidos os resultados do QP 1.2 e do QP 1.3, mostrando que o desempenho da técnica de AM geralmente depende do número de atributos usados durante a classificação da DRC.

### **Respondendo QP Principal**

Portanto, com base nas respostas obtidas para a QP 1.1, QP 1.2 e QP 1.3, não é possível afirmar qual é o classificador baseado em AM mais adequado para o diagnóstico da DRC sem considerar o número e o tipo de atributos disponíveis. A resposta para a QPP depende do contexto relacionado ao diagnóstico da DRC. Por exemplo, os classificadores baseados em AM, denominados NB, RF, árvore de decisão em *bagging/J48* e o método de subespaço aleatório/J48, apresentaram os melhores resultados quando vários atributos estão disponíveis

para o diagnóstico da DRC.

## 4.2 Experimento com o Weka

Um experimento focado no diagnóstico da DRC em países em desenvolvimento foi realizado para melhorar a qualidade da análise comparativa e solucionar a lacuna de pesquisa identificada na RSL. Os classificadores baseados em AM foram selecionados a partir de um subconjunto das técnicas identificadas ao responder as QP da RSL: RF, NB, árvore de decisão J48, KNN, SVM e MLP. Esse subconjunto de classificadores baseados em AM também foi selecionado de acordo com a resposta da QP1.3.

### 4.2.1 Conjunto de Dados

Os atributos do conjunto de dados foram escolhidos principalmente com base em diretrizes médicas amplamente aceitas pela comunidade geral de nefrologistas, especificamente, a diretriz KDIGO [59], o instituto nacional de diretrizes de excelência em saúde e assistência [34] e a diretriz KDOQI [49]. Além disso, três nefrologistas brasileiros foram entrevistado para confirmar a relevância dos atributos pré-selecionados no contexto dos países em desenvolvimento. O conjunto final de atributos da DRC com foco nos países em desenvolvimento incluía hipertensão, DM, creatinina, uréia, albuminúria, idade, sexo e TFG.

Uma amostra do conjunto de dados usado para conduzir as classificações é mostrada na Tabela 4.2 e compreende um subconjunto dos atributos totais identificados pela resposta da QP1.2. Na Tabela 4.3, Presente, representa a existência de DM e/ou hipertensão, enquanto Ausente, a inexistência de DM e hipertensão. Cada ID da tabela está associado a um assunto no conjunto de dados. Crea e Alb representam abreviações para creatinina e albumiúria, respectivamente.

Foi considerado um conjunto de dados contendo indivíduos diagnosticados ou não com DRC, os dados médicos (60 prontuários médicos do mundo real) foram coletados a partir de prontuários físicos de indivíduos adultos (idade  $\geq 18$ ), sob tratamento do Hospital Universitário Prof. Alberto Antunes, da UFAL, Brasil. A coleta de dados dos prontuários mantidos em formato não eletrônico no hospital foi aprovada pelo comitê de ética da UFAL e realizada em um trabalho anterior, entre 2015 e 2016 [98].

Tabela 4.2: Resumo dos atributos, técnicas e desempenho de acordo com o QP1.2 e QP1.3.

<b>Técnica</b>	<b>Atributos</b>	<b>Desempenho</b>
RNAs.	Idade (I), albuminúria, IMC, cistacina C, Creatinina (C), DCV, dislipidemia, histórico, Sexo (S), TFG, uréia, glomerulonefrite, anti-inflamatórios, neoplasias, nefropatias, rins policísticos, fatores de risco, lúpus, tabaco e obstrução de urinário.	Sensibilidade (Se) (entre 93,1% e 94,9%) [72].
Rede neural associativa.	I, colesterol, DM, S, TFG, hemoglobina, hipertensão, obesidade, reação em cadeia do polímero e C.	Coefficiente de correlação (0,951) [81].
Rede de retro-propagação.	I, nitrogênio da uréia, C, peso, S, TFG, Glicose (G), hematuria, proteinúria e PA.	Precisão (P) (94,75%) [25].
Bagging/J48, NB, RF e subespaço aleatório/J48.	I, albumina, anemia, apetite, bactérias, G, PA, DAC, C, DM, hemoglobina, hipertensão, volume celular compactado, edema de pedal, potássio, células pus, aglomerados de células pus, glóbulos vermelhos, células sanguíneas, sódio, gravidade específica, açúcar, uréia e células brancas.	P (100%)[7, 58, 57].
Árvore de decisão.	I, albumina, PA, nitrogênio da uréia, cálcio, C, açúcar no sangue, S, hemoglobina exame de sangue, fósforo, potássio, ferritina, sódio, proteína total e peso.	P (92,2%) [2].
SVM.	I, apetite, PA, uréia, DAC, volume de células compactadas, aglomerados de células pus, urina de células pus, glóbulos vermelhos, C, gravidade específica e contagem de glóbulos brancos.	P (98,5%) [79].
SVM.	I, creatinina, S, TFG e uréia.	P (76,32%) [105].
SVM.	Imagem de ultrassom.	Se (93,82%) [47].

O conjunto de dados compreende 16 sujeitos sem lesão renal, sendo 14 diagnóstica-

Tabela 4.3: Amostra do conjunto de dados da DRC coletada no Hospital Universitário Prof. Alberto Antunes. Cada ID representa um assunto específico do conjunto de dados.

ID	DM	HAS	Crea	Urea	Alb	Idade	Sexo	TFG	Risco
1	Ausente	Presente	1.00	27.0	0.4	50	F	89.0	Baixo
2	Ausente	Presente	1.27	28.0	9.0	62	M	52.0	Moderado
3	Presente	Presente	0.52	23.1	3.3	54	F	164.0	Alto
4	Ausente	Ausente	0.70	34.0	6.0	69	M	59.8	Moderado
5	Ausente	Ausente	1.04	14.9	9.5	39	F	93.5	Baixo
6	Ausente	Ausente	0.68	19.0	2.0	65	F	59.7	Moderado
7	Ausente	Ausente	1.11	12.0	357.0	62	F	49.7	Muito Alto
8	Ausente	Ausente	0.53	23.0	4.0	59	F	104.6	Baixo

dos apenas com DRC, e 30 indivíduos diagnosticados com DRC, hipertensão e/ou DM. De maneira geral, a amostra incluiu sujeitos com idades entre 31 e 79 anos; aproximadamente 94,5% dos indivíduos foram diagnosticados com hipertensão e 58,82% foram diagnosticados com DM.

O método *k-fold Cross Validation* (com  $k = 10$ ) foi usado para avaliar os classificadores com base em 114 registros, incluindo os 60 dados de prontuários de pacientes do HUPAA e 54 aumentados. O aumento de dados foi realizado apenas no conjunto de treinamento para cada uma das 10 dobras. Portanto, o número total, para cada conjunto de treinamento e teste, foi composto por 108 e 6 registros, respectivamente. As amostras foram aumentadas duplicando os registros de prontuários, modificando cuidadosamente os atributos (aumentando cada biomarcador da DRC em 0,5). A validade dos dados aumentados foi verificada por um nefrologista com mais de 30 anos de experiência no tratamento e diagnóstico da DRC em países em desenvolvimento, que analisou cada registro quanto à classificação de risco adequada.

Os 60 prontuários médicos também foram avaliados pelo nefrologista experiente. Além disso, dois nefrologistas menos experientes avaliaram os 60 registros médicos, como parte do experimento. Os nefrologistas classificaram os sujeitos considerando as classes de baixo risco, risco moderado, alto risco e risco muito alto, de acordo com a diretriz KDIGO CKD “*evolution not revolution*” [59] e os atributos disponíveis. A etapa de treinamento foi condu-

zida com base nas avaliações de risco da DRC fornecidas pelo nefrologista mais experiente. O nefrologista mais experiente foi escolhido como base, pelo maior alinhamento com a literatura especializada atual. Concluída a etapa de treinamento, os dados foram utilizados para a classificação de risco, ou seja, a etapa de teste.

### 4.2.2 Procedimento

Uma amostra das avaliações de risco é apresentada na última coluna da Tabela 4.3. Dois nefrologistas menos experientes também avaliaram os 60 registros médicos do mundo real como parte do experimento. A etapa de treinamento foi conduzida com base na classificação do nefrologista mais experiente. Concluída a etapa de treinamento, os dados foram utilizados para a classificação de risco, ou seja, a etapa de teste. As etapas de treinamento e teste foram realizadas usando o método de validação cruzada *k-fold*, aplicando os classificadores da árvore de decisão J48, RF, NB, SVM, MLP e KNN com o software Weka©.

O procedimento de avaliação utilizado neste estudo está resumido na Figura 4.3. Primeiro, os classificadores de AM e o conjunto de treinamento foram usados para gerar os modelos de aprendizado usando o software Weka©. Em seguida, o conjunto de testes foi fornecido aos classificadores da árvore de decisão J48, RF, NB, SVM, MLP e KNN, e os resultados foram comparados com as opiniões dos três nefrologistas que avaliaram os prontuários médicos e forneceram suas conclusões sobre as situações de saúde dos indivíduos considerando o risco de desenvolver DRC. A classificação do nefrologista mais experiente foi comparada com os resultados dos classificadores de AM usando as seguintes métricas de desempenho [88]:

- **Instâncias classificadas corretamente (*Correctly Classified Instances (CCIs)*):** a porcentagem de indivíduos classificados corretamente.
- **Instâncias classificadas incorretamente (*Incorrectly Classified Instances (ICIs)*):** a porcentagem de indivíduos classificados incorretamente.
- **Erro absoluto médio (*Mean Absolute Error (MAE)*) e erro quadrático médio da raiz (*Root Mean Squared Error (RMSE)*):** a diferença existente entre as saídas obtidas e as desejadas.

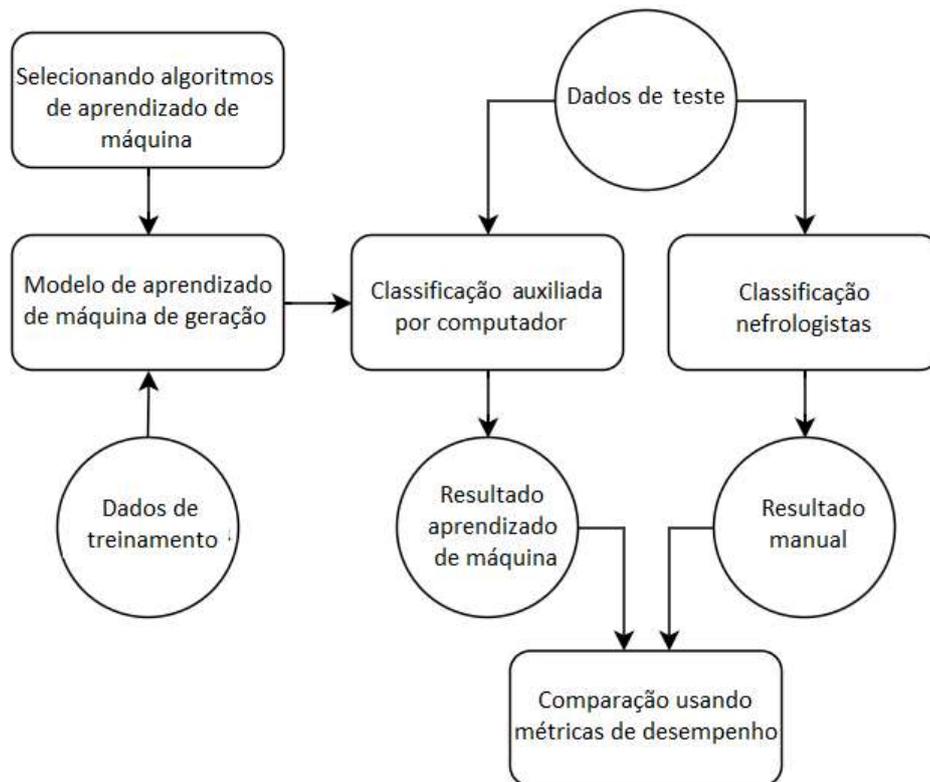


Figura 4.3: Procedimento de avaliação usando classificadores de AM e opiniões dos nefrologistas.

- **Taxa de verdadeiro positivo (*True Positive (TP)*):** o número de classificações definidas como positivas que são realmente negativas.
- **Taxa de falso positivo (*False Positive (FP)*):** o número de classificações definidas como positivas que são realmente negativas.
- **Precisão (P):** a razão entre o total de instâncias de uma classe e o total de instâncias classificadas.
- **Área da curva de precisão de recuperação (*Precision-Recall Curve (PRC)*):** a razão entre P e sensibilidade (ou seja, taxa TP).
- **Área de característica operacional do receptor (*Receiver Operating Characteristic (ROC)*):** a razão entre especificidade (isto é, negativos verdadeiros corretamente identificados) e sensibilidade.

- **Estatística kappa de Cohen (k):** o nível de concordância entre os juízes independentes.

Além disso, os resultados da classificação dos dois nefrologistas menos experientes foram comparados com os classificadores baseados em AM usando a estatística *kappa de Cohen* para medir a concordância entre observadores independentes.

A estatística *kappa de Cohen* foi medida usando o software Epidat©7 4.11 calculando a concordância bruta e o índice de concordância kappa (k) com intervalos de confiança de 95%, sem ajustar o viés e a prevalência. Na análise, foram considerados os seguintes índices k:  $k < 0$ , sem concordância; k entre 0 e 0,19, má concordância; k entre 0,20 e 0,39, baixa concordância; k entre 0,40 e 0,59, concordância moderada; k entre 0,60 e 0,79, concordância substancial; e k entre 0,80 e 1, concordância quase perfeita.

Finalmente, os resultados dos classificadores também foram comparados com as opiniões combinadas dos três nefrologistas. O objetivo com as últimas duas comparações foi realizar uma discussão mais detalhada sobre o uso dos classificadores no contexto de países em desenvolvimento.

### 4.2.3 Análise de Dados

Cinco iterações de 10-fold *cross validation* foram realizadas para cada classificador usando o conjunto de dados da DRC, apresentando estabilidade. Os classificadores apresentaram estabilidade porque os modelos substitutos produziram a mesma previsão para os mesmos dados de entrada. Na Tabela 4.4 são apresentados os resultados do valor médio das métricas de desempenho usadas para comparar as classificações do nefrologista mais experiente e os classificadores baseados em AM com base nos atributos hipertensão, DM, creatinina, uréia, albuminúria, idade, sexo e TFG. Mais especificamente, as métricas de desempenho CCI, ICI, MAE, RMSE, taxa TP, taxa FP, área P, ROC e PRC foram aplicadas.

Os classificadores baseados em AM com melhor desempenho foram J48 e RF com 95,00% e 93,33% para ICC (ou seja, um total de 57 e 56 instâncias, respectivamente). O classificador J48 apresentou P de 0,97, área ROC de 0,96, área PRC de 0,94, taxa FP de 0,01 e erros médios e absolutos de 0,09 e 0,03; enquanto que o RF apresentou precisão de 0,99, área ROC de 0,99, área PRC de 0,97, taxa de FP de 0,01 e erros médios e absolutos de

0,17 e 0,09. Os classificadores J48 e RF foram seguidos por NB (CCI = 88,3%, ou seja, 53 instâncias), SVM (CCI = 76,66%, ou seja, 46 instâncias), MLP (CCI = 75,00%, ou seja, 45 instâncias) e KNN (CCI = 71,67%, ou seja, 43 instâncias). Portanto, apesar dos classificadores baseados em AM NB, SVM, MLP e KNN apresentarem algum ICI, o percentual foi baixo quando comparado à árvore de decisão J48 e aos classificadores de RF.

Tabela 4.4: Resultados da avaliação de qualidade para os classificadores e o nefrologista mais experiente usando as métricas de desempenho.

	<b>J48</b>	<b>RF</b>	<b>NB</b>	<b>SVM</b>	<b>MLP</b>	<b>KNN</b>
<b>CCI (%)</b>	95.00	93.33	88.33	76.66	75.00	71.67
<b>ICI (%)</b>	5.00	6.67	11.67	23.33	25.00	28.33
<b>MAE</b>	0.03	0.09	0.11	0.28	0.14	0.15
<b>RMSE</b>	0.09	0.17	0.21	0.36	0.31	0.33
<b>TP</b>	0.95	0.93	0.88	0.77	0.75	0.72
<b>FP</b>	0.01	0.01	0.05	0.22	0.28	0.16
<b>P</b>	0.97	0.99	0.95	0.78	0.78	0.79
<b>ROC</b>	0.96	0.99	0.97	0.79	0.83	0.88
<b>PRC</b>	0.94	0.97	0.96	0.67	0.82	0.86

Na Tabela 4.5 são ilustrados os resultados de kappa obtidos ao comparar a opinião do nefrologista mais experiente com os classificadores baseados em AM, k foi calculado para cada classe de risco de DRC: risco baixo, risco moderado, risco alto e risco muito alto. SVM, MLP e KNN apresentaram a menor quantidade de concordância k com o nefrologista experiente; ou seja, concordância moderada com  $k = 0,5807$ ,  $k = 0,5418$  e  $k = 0,5112$ , respectivamente; seguido por NB ( $0,8170$ ), RF ( $k = 0,8950$ ) e J48 ( $k = 0,9221$ ), que apresentaram concordância quase perfeita com o nefrologista mais experiente.

Para todos os classificadores, a classe de "risco muito alto" teve o impacto mais negativo na diminuição dos valores de k. Por exemplo, de acordo com a amostra de avaliação apresentada na Tabela 4.3 (última coluna), o classificador RF discordou do nefrologista mais experiente na avaliação do paciente identificado com ID 7, afirmando um "risco moderado" em vez de "muito alto". A classe de "baixo risco" e a classe de "alto risco" apresentaram os maiores níveis de concordância.

Tabela 4.5: Resultados da avaliação de qualidade para os classificadores e o nefrologista mais experiente usando a estatística Kappa.

	<b>J48</b>	<b>RF</b>	<b>NB</b>	<b>SVM</b>	<b>MLP</b>	<b>KNN</b>
<b>Baixo risco</b>	1.0000	0.9333	0.9000	0.5833	0.5567	0.5982
<b>Risco moderado</b>	0.8958	0.8887	0.8268	0.0510	0.1091	0.4372
<b>Alto risco</b>	0.9565	0.9179	0.8295	0.9148	0.8695	0.5402
<b>Risco muito alto</b>	0.3739	0.4828	0.2982	-0,0256	-0.0256	-0.0345
<b>kappa Global</b>	0.9221	0.8950	0.8170	0.5807	0.5418	0.5112

Para melhorar a discussão relacionada ao experimento realizado neste estudo, os resultados dos classificadores de AM, treinados com base na opinião do nefrologista mais experiente, foram comparados à opinião de dois nefrologistas menos experientes usando a estatística *kappa de Cohen*.

Na Tabela 4.6 são apresentados os resultados kappa dos classificadores baseados em AM e a avaliação dos dois nefrologistas, k também foi calculado para as classes de baixo risco, risco moderado, alto risco e risco muito alto. Além disso, foi apresentado um kappa global para cada classificador baseado em AM. Os classificadores J48, RF, NB, SVM, MLP e KNN apresentaram valores de kappa globais de 0,4619, 0,4743, 0,44443, 0,4365, 0,4028 e 0,4105, respectivamente.

Como esperado, o desempenho dos classificadores diminuiu, apresentando concordância moderada com os dois nefrologistas. As classes de “risco moderado” e “risco muito alto” contribuíram significativamente para a diminuição dos valores globais de kappa, com alguns classificadores apresentando concordância ruim, baixa ou inexistente com os nefrologistas menos experientes. A classe de “baixo risco” apresentou os maiores níveis de concordância.

Por fim, na Tabela 4.7 é apresentada a comparação completa, considerando os resultados dos classificadores e as opiniões dos três nefrologistas. Considerando os resultados dos classificadores e as opiniões dos três nefrologistas, os valores de k aumentaram.

O classificador J48 apresentou um kappa global de 0,5578, enquanto que os classificadores RF, NB, SVM, MLP e KNN apresentaram 0,5884, 0,5425, 0,4877, 0,4706 e 0,4646. Entretanto, nesta comparação, a concordância entre o J48, RF, NB, SVM, MLP e KNN e os nefrologistas continuou moderada. Além disso, como na comparação anterior, as classes

Tabela 4.6: Resultados da avaliação da qualidade dos classificadores e dos dois nefrologistas menos experientes que usam a estatística Kappa.

	<b>J48</b>	<b>RF</b>	<b>NB</b>	<b>SVM</b>	<b>MLP</b>	<b>KNN</b>
<b>Baixo risco</b>	0.6429	0.5982	0.5988	0.4866	0.4886	0.5333
<b>Risco moderado</b>	0.0051	-0.0651	0.0317	-0.0169	-0.0465	-0.0405
<b>Alto risco</b>	0.4344	0.4848	0.4135	0.4360	0.4118	0.3925
<b>Risco muito alto</b>	0.1771	0.3220	0.1771	-0.0112	-0.0112	-0.0227
<b>kappa Global</b>	0.4619	0.4743	0.4443	0.4365	0.4028	0.4105

de “risco moderado” e “risco muito alto” contribuíram significativamente para a diminuição dos valores kappa globais. A classe de “baixo risco” apresentou os maiores níveis de concordância, seguida pela classe de “alto risco”.

Tabela 4.7: Resultados da avaliação da qualidade dos classificadores e dos três nefrologistas que usam a estatística Kappa.

	<b>J48</b>	<b>RF</b>	<b>NB</b>	<b>SVM</b>	<b>MLP</b>	<b>KNN</b>
<b>Baixo risco</b>	0.7215	0.7384	0.6828	0.5767	0.5833	0.6164
<b>Risco moderado</b>	0.2707	0.2466	0.2661	0.0392	0.0664	0.1273
<b>Alto risco</b>	0.5220	0.5718	0.5139	0.5180	0.4997	0.4406
<b>Risco muito alto</b>	0.3624	0.4302	0.3624	0.2511	0.2511	0.2023
<b>kappa Global</b>	0.5578	0.5884	0.5425	0.4877	0.4706	0.4646

#### 4.2.4 Discussão

Considerando todos os resultados obtidos na análise comparativa, os resultados da SLR indicaram que a escolha do classificador baseado em AM e dos atributos da DRC depende do contexto de aplicação. Em geral, os estudos estão relacionados com a classificação de indivíduos com DRC, ou não. A análise dos resultados de classificação durante a RSL demonstraram que, geralmente, um número maior de atributos resulta em um melhor desempenho para os classificadores baseados em AM. Entretanto, a necessidade do uso de vários biomarcadores da DRC na realização de classificações aumenta o custo para os usuários que

se beneficiam de diagnósticos auxiliados por computador. O alto custo é um problema crítico para países em desenvolvimento, que enfrentam altas taxas de pobreza, bem como altas taxas de mortalidade de indivíduos com DRC.

Embora que, em um dos estudos, os autores tenham relatado alta precisão usando apenas um atributo da DRC [93], essa classificação requer dados temporais dos registros eletrônicos de saúde (*Electronic Health Records* (EHRs)), ou seja, uma quantidade considerável de registros da TFG coletados ao longo de anos de monitoramento. A disponibilidade de EHRs não é realista na maioria dos países em desenvolvimento, especialmente em ambientes rurais e de difícil acesso. Nessas regiões, muitos registros médicos ainda são escritos em papel e mantidos em armazenamento físico devido ao orçamento disponível, que é geralmente limitado. Além disso, alguns países em desenvolvimento, como o Brasil, são de tamanho continental, tornando a disponibilidade de EHRs problemática.

Em outros estudos, os autores consideram a classificação da DRC baseada em imagens de ultrassom usando classificadores baseados em AM [20], [47]. Para usar imagens de ultrassom, a disponibilidade de máquinas de ultrassom precisa ser suficiente para o atendimento de localidades com altas incidências de DRC, incluindo localidades rurais e de difícil acesso. No entanto, essa também não é a realidade atual de países em desenvolvimento.

Portanto, embora que, em vários estudos publicados sobre o diagnóstico da DRC, os autores considerem uma grande quantidade de atributos da DRC e classificadores baseados em AM, até o momento, não existe consenso na literatura sobre qual abordagem é a mais adequada para auxiliar o diagnóstico da DRC em países em desenvolvimento.

A RSL e o experimento apresentados neste capítulo complementaram o estado da arte, considerando o caso específico da classificação de risco de DRC em países em desenvolvimento. O uso de diretrizes médicas bem aceitas e a opinião de um nefrologista com mais de 30 anos de experiência no tratamento de pacientes com DRC em países em desenvolvimento, para selecionar os atributos e treinar os classificadores, é uma abordagem confiável, permitindo uma discussão focada nos problemas específicos relacionados ao diagnóstico de DRC auxiliado por computador.

Os resultados experimentais indicaram que um alto nível de precisão pode ser obtido pela maioria dos classificadores usando hipertensão, DM, creatinina, uréia, albuminúria, idade, sexo e TFG.

Com base nos resultados obtidos, durante o monitoramento, assumindo o diagnóstico prévio de DM, o usuário de uma solução auxiliada por computador com base nos classificadores J48, RF, NB, SVM, MLP e KNN, precisa realizar periodicamente dois exames de sangue: creatinina e uréia. Por outro lado, a albuminúria é medida usando um teste de urina, enquanto que a TFG pode ser calculada com a equação de *Cockcroft-Gault* [27], comumente aplicado em países em desenvolvimento com base em creatinina, sexo, idade e peso. Portanto, com base nestes resultados, as classificações usando classificadores baseados em AM exigem apenas que os usuários gastem recursos financeiros com dois exames de sangue e um exame de urina.

Mesmo com uma taxa de FP superior a zero nos seis classificadores, três ainda apresentaram concordância moderada (0,40 k 0,59) com o nefrologista experiente (SVM, MLP e KNN); enquanto que o J48, RF e NB apresentaram concordância quase perfeita (0,80 k 1). Utilizando as classes de baixo risco, risco moderado, risco alto e risco muito alto, é possível realizar o encaminhamento precoce de uma pessoa com risco de DRC para um nefrologista, visando uma avaliação clínica mais detalhada. Os classificadores SVM, MLP e KNN exibiram o desempenho mais baixo ao classificar com base no conjunto de dados considerado. Em outros estudos, SVM, MLP e KNN apresentaram altos níveis de precisão ao considerar, por exemplo, 24 atributos.

Além disso, quando os dois nefrologistas menos experientes que tratam pacientes com DRC em países em desenvolvimento foram considerados durante as comparações, os classificadores J48, RF e NB diminuíram a concordância de níveis quase perfeitos para níveis moderados, enquanto que o SVM, KNN e MLP continuaram apresentando níveis de concordância moderados. Isto pode ser atribuído ao fato de que ocorreram algumas discordâncias entre os dois nefrologistas menos experientes e o mais experiente. No entanto, a opinião do nefrologista mais experiente se torna mais confiável, devido ao grande número de anos do especialista no tratamento de pacientes com DRC em países em desenvolvimento, juntamente com sua experiência de ensino e pesquisa durante o mesmo período e maior alinhamento com a literatura especializada atual.

Quando os três nefrologistas foram considerados durante as comparações, o desempenho das classificações aumentou novamente, demonstrando que, em geral, os classificadores J48, RF, NB, SVM, MLP e KNN apresentaram quase a mesma concordância moderada com

os nefrologistas. O classificador J48 apresentou concordância quase substancial com os nefrologistas com um kappa global de 0,5578. Somente a classe de “baixo risco” apresentou concordância substancial para os classificadores J48, RF, NB e KNN, com  $k = 0,7215$ ,  $k = 0,7384$ ,  $k = 0,6828$  e  $k = 0,6164$ , respectivamente.

O aumento no número de juízes também aumenta a confiança na avaliação dos classificadores, demonstrando que, na maioria dos casos, existe, ao menos, uma concordância moderada entre os classificadores baseados em AM e os nefrologistas que trabalham em situações do mundo real em países em desenvolvimento. O resultado mais relevante apresentado neste capítulo é a comparação dos classificadores com o nefrologista mais experiente; no entanto, as outras comparações aumentam a confiança no experimento, demonstrando diferentes cenários de avaliação. A discordância entre os três nefrologistas entrevistados também indica a relevância do uso de soluções para o diagnóstico precoce da DRC auxiliado por computador em países em desenvolvimento para auxiliar nefrologistas menos experientes em tomadas de decisões clínicas. Por exemplo, para a amostra do conjunto de dados apresentado na Tabela 4.3, os dois nefrologistas menos experientes discordaram do nefrologista mais experiente na análise de risco do paciente com ID 2, afirmando um risco alto em vez de risco moderado, e aqueles com ID 4 e 6, afirmando alto risco em vez de risco moderado, enquanto apenas um nefrologista menos experiente discordou do nefrologista mais experiente na avaliação do paciente com ID 7, afirmando apenas alto risco em vez de risco muito alto.

Considerando a QP sobre o classificador baseado em AM mais adequado para o diagnóstico de DRC em países em desenvolvimento, os resultados sugerem o J48 e a RF. Os classificadores apresentaram 95,00% e 93,33% de precisão para o conjunto de dados da DRC coletado para o experimento, seguido por NB com 88,33%, SVM com 76,66%, MLP com 75,00% e KNN com 71,67%. Por exemplo, SVM e MLP discordaram do nefrologista mais experiente, afirmando baixo risco em vez de risco moderado para pacientes com ID 2 e ID 4. O KNN também discordou do nefrologista mais experiente, afirmando alto risco em vez de risco moderado para pacientes com ID 2. No paciente com ID 6, SVM e KNN discordaram do nefrologista mais experiente, também declarando baixo risco em vez de risco moderado, enquanto a árvore de decisão J48, NB, SVM e MLP discordaram do nefrologista mais experiente, afirmando risco moderado em vez de risco muito alto para paciente com ID 7.

Considerando o conjunto completo de avaliações, 23 classificações forneceram baixo risco em vez de risco moderado. Esse é um problema crítico, porque um paciente geralmente é encaminhado a um nefrologista quando existe um risco moderado ou mais alto; portanto, a subestimação da situação de risco de um paciente pode resultar em diagnóstico tardio da DRC. Os classificadores J48 e o RF não foram responsáveis por nenhuma das 23 subestimações. Por outro lado, 6 classificações forneceram baixo risco em vez de alto risco, a árvore de decisão J48 e a RF também não foram responsáveis por nenhuma delas, enquanto 5 classificações apresentaram risco moderado em vez de baixo risco (RF, MLP e KNN), 3 classificações apresentaram alto risco ao invés de baixo risco (NB e KNN) e 2 apresentaram baixo risco ao invés de risco muito alto (SVM e KNN). Portanto, o classificador de árvore de decisão J48 não se refere a nenhuma das 39 classificações enganosas críticas totais, seguidas de RF com 1. As classificações enganosas restantes foram menos prejudiciais porque ainda resultaram no encaminhamento do paciente em avaliação e até superestimaram/subestimaram o risco.

Na prática clínica, os resultados dos classificadores RF, NB, SVM, MLP e KNN geralmente não são tão fáceis de interpretar quanto os da árvore de decisão J48. Por exemplo, a classificação com RF usou 8 árvores aleatórias para cada dobra para alcançar 93,33% de precisão. A fácil interpretação das análises de risco da DRC é um fator crítico para a reutilização dos resultados da classificação por nefrologistas e médicos de cuidados primários que precisam realizar exames adicionais para confirmar a situação clínica de um paciente. Nos últimos anos, a comunidade médica se preocupou com a interpretação dos resultados da classificação fornecidos pelas técnicas de AM [93].

Além disso, o uso de classificadores baseados em AM que apresentam resultados de simples interpretação pode aumentar a confiança dos médicos que realizam atendimento da DRC em países em desenvolvimento, aumentando também a reutilização dos resultados em cenários reais. O classificador J48, uma abordagem de fácil interpretação baseada em árvores de decisão, apresentou precisão de 95,00% e concordância quase perfeita com o nefrologista mais experiente. No contexto da identificação de riscos de DRC nos países em desenvolvimento, os resultados da árvore de decisão J48 são mais fáceis de interpretar pelos médicos. Por exemplo, na Figura 4.4a é apresentada a árvore de decisão gerada pelo classificador da árvore de decisão J48, que compreende cada biomarcador de DRC considerado e a classifi-

cação relacionada definida pelo software Weka©.

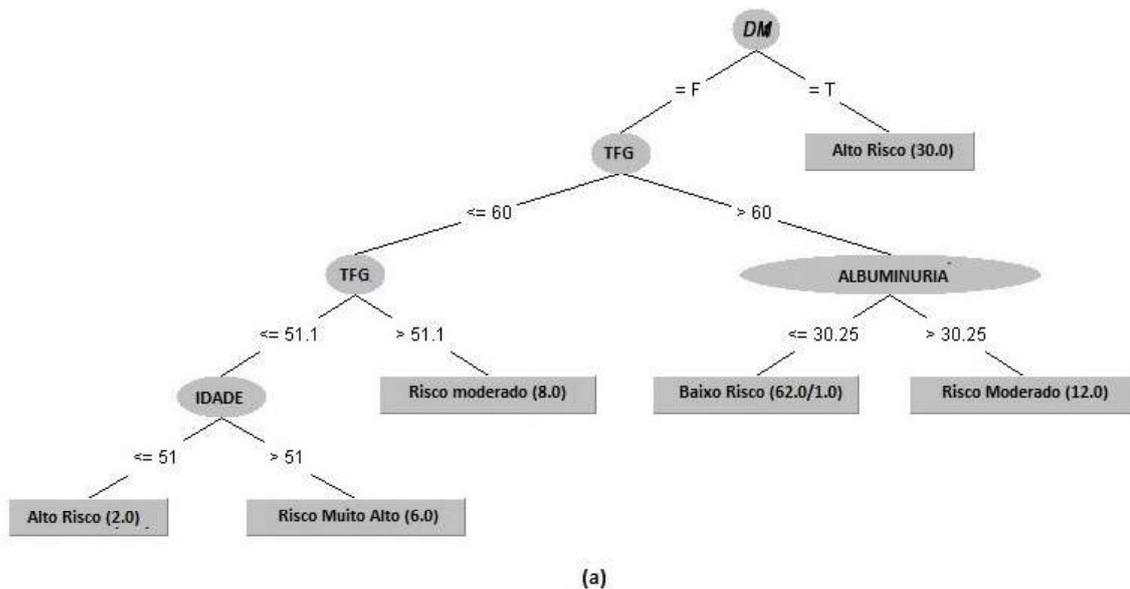


Figura 4.4: (a) Árvore de decisão para avaliação de risco de DRC feita pelo algoritmo J48.

Neste caso, é necessário somente que o médico de atendimento primário analise as árvore de decisões para interpretar a lógica da classificação apresentado pelo sistema de suporte à decisão clínica. Dos 8 atributos da DRC na Tabela 4.3, o classificador J48 usou apenas 4 para classificar o risco (isto é, DM, TFG, albuminúria e idade), exigindo 1 exame de sangue e 1 exame de urina quando o DM já está avaliado, ao custo de 4 ICIs. Os atributos DM (1), TFG (2) e albuminúria (3) apresentaram o mesmo poder de previsão para o nefrologista experiente e a árvore de decisão J48, considerando uma escala que varia de 1 (prioridade mais alta) a 8 (prioridade mais baixa); no entanto, o nefrologista priorizou e usou os atributos creatinina (4), uréia (5), sexo (6), hipertensão (7) e idade (8).

A tomada de decisão do nefrologista considera idade, sexo, valores de referência de biomarcadores e fatores de risco, ou seja, não se restringe à classificação de risco da diretriz KDIGO. Além disso, apresentar o resultado da avaliação auxiliada por computador como uma árvore de decisão pode simplificar o diagnóstico final de um médico de cuidados primários. Por outro lado, na Figura 4.5b é apresentada uma amostra das árvores aleatórias geradas pela RF, mostrando que a interpretação dos resultados exige que o médico da atenção primária siga toda as árvore de decisões para entender completamente a lógica da classificação.

No entanto, com trabalho extra, também é possível criar intérpretes dos classificadores

baseados em AM para ajudar os médicos a entender os resultados de outros classificadores. Portanto, os classificadores baseados em AM mais adequado para o diagnóstico de DRC nos países em desenvolvimento pode ser tratado considerando duas possibilidades: (1) RF junto com resultados interpretados ou (2) árvore de decisão J48. Essa decisão de projeto deve ser tomada com cuidado, pois uma classificação de risco falso negativo pode levar a um diagnóstico tardio da DRC, aumentando a progressão da doença e a probabilidade de diálise e transplante renal.

### 4.3 Considerações Finais

Neste capítulo pode-se observar que os classificadores J48 e RF exibiram o melhor desempenho usando o conjunto de dados da DRC, incluindo os atributos hipertensão, DM, creatinina, uréia, albuminúria, idade, sexo e TFG. Esses atributos são comumente usados pelos nefrologistas para diagnosticar DRC em países em desenvolvimento. O classificador baseado em AM, denominado RF, geralmente realiza avaliações mais complexas, dificultando a interpretação dos resultados da classificação pelos médicos. A aplicação de algoritmos (intérpretes) é necessária para interpretar os resultados antes de apresentá-los aos médicos da atenção primária. Por outro lado, o classificador baseado em AM J48 aborda o problema de interpretação que ainda apresenta concordância quase perfeita com um nefrologista experiente que trata pacientes com DRC em países em desenvolvimento há mais de 30 anos. Além disso, as classificações enganosas críticas não são apresentadas pelo classificador da árvore de decisão J48 ao avaliar os sujeitos do conjunto de dados da DRC.

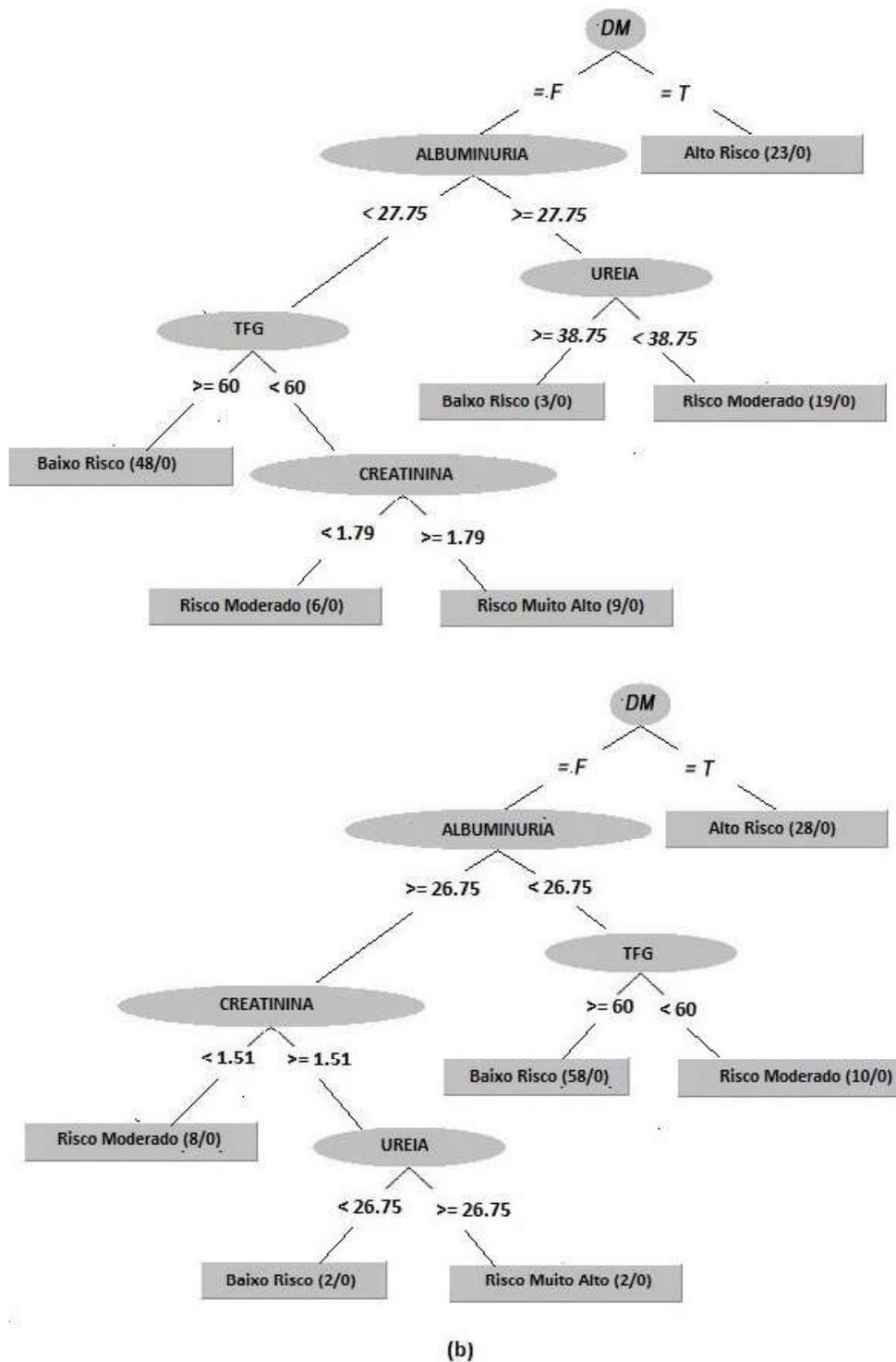


Figura 4.5: (b) Amostra das árvores aleatórias para avaliação de risco de DRC feita pelo algoritmo de RF.

# Capítulo 5

## Desenvolvimento

Com base na conclusão obtida pela realização do estudo apresentado no Capítulo 4, neste capítulo, é apresentado o desenvolvimento de um sistema inteligente para a avaliação de risco da DRC, além do projeto e desenvolvimento de um sistema de encaminhamento de pacientes em situações de emergência para unidades de saúde.

Um protótipo inicial para Android do sistema MultCare foi desenvolvido [96], incluindo especificações formais, avaliação de eficácia e testes de usabilidade. Com base nos resultados dos testes de usabilidade, foi possível identificar algumas limitações que motivaram o uso de tecnologias Web. Também com base na análise comparativa, um sistema Web inteligente de suporte à decisão clínica, denominado MultCare Paciente, foi desenvolvido. A versão Web do sistema foi utilizada para aprimorar a análise de risco da DRC usando o classificador baseado em Aprendizado de Máquina (AM) J48. Outra questão relevante foi o projeto de um recurso para encaminhar o paciente em situações de emergência para a um serviço de saúde adequado por meio de um sistema especialista, ao visitar um local desconhecido.

O *backend* e o *frontend* do sistema foram isolados para facilitar a manutenção, mantidos em dois repositórios de controle de versão. Por um lado, o *back-end* do sistema foi implementado, principalmente, usando a linguagem de programação Java, serviços web e as interfaces de programação de aplicação Weka© e Dools©. Por outro lado, o *front-end* do sistema foi implementado usando, principalmente, HTML 5, Bootstrap, JavaScript e Vue.js.

As principais funcionalidades implementadas incluem <sup>1</sup>: (1) controle de acesso; (2) gerenciamento de exames, alergias e medicamentos; (3) realização de análise de risco; (4)

---

<sup>1</sup>Principais funcionalidades: [youtu.be/icdpVJuUw00](https://youtu.be/icdpVJuUw00) e [youtu.be/nuyGMpEdN0s](https://youtu.be/nuyGMpEdN0s).

monitoramento de DM e hipertensão; e (5) encaminhamento em situações de emergência

Na Figura 5.1 é ilustrada a interface gráfica do usuário (*Graphical User Interface (GUI)*) para registrar o resultado de um novo exame. Além disso, o usuário pode fazer upload de um arquivo XML contendo os resultados dos exames para adicionar um grande número de entradas manuais. Depois que o paciente fornece os resultados atuais da classificação, a GUI principal do sistema é atualizada, mostrando os resultados disponíveis para o mês atual.

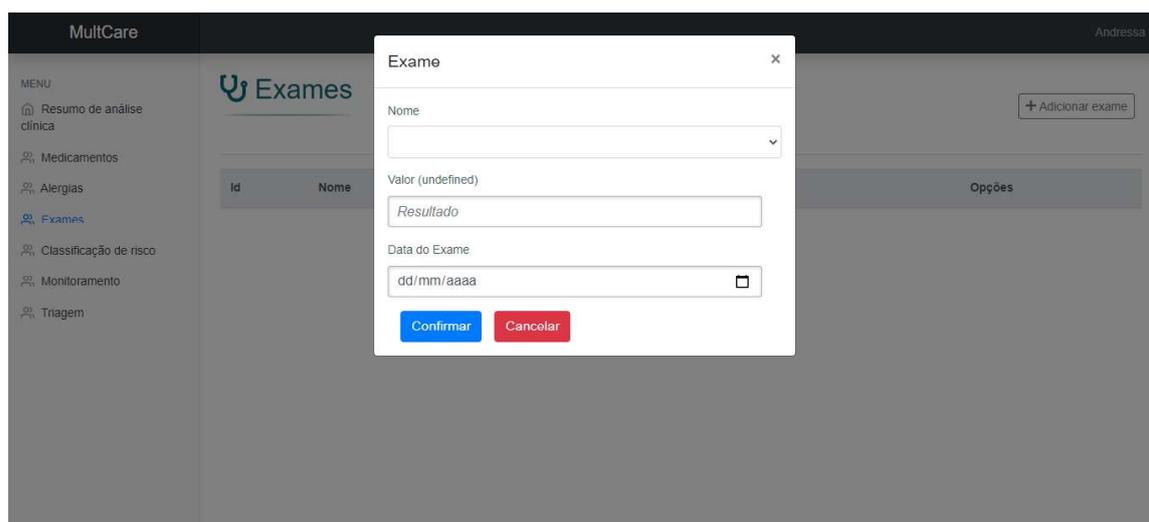


Figura 5.1: GUI do sistema para registrar um novo exame.

Na Figura 5.2 é ilustrada a GUI principal do sistema, com a apresentação da creatinina, uréia, albuminúria e TFG. Neste estudo, é reduzido o número de resultados de exames necessários para realizar a análise de risco de DRC de 5 para 4, quando comparado com a pesquisa publicada anteriormente [98]. Isso é crítico para populações de baixa renda, facilitando o uso do sistema. O sistema fornece uma nova análise de risco de DRC quando o paciente digita todos os atributos necessários. A redução foi realizada com base na aplicação do algoritmo J48.

Durante a análise de risco, realizada com base na presença/ausência de DM e hipertensão, creatinina, uréia, albuminúria, idade, sexo, e TFG; o classificador J48 é usado para analisar a situação do paciente considerando quatro classes de risco: baixo risco, risco moderado, alto risco e risco muito alto. O conjunto completo de dados da DRC, usado durante a análise comparativa, foi considerado no treinamento do classificador J48. A interface de programação de aplicação do Weka© foi usada na implementação do sistema.

Nos casos de risco moderado, risco alto ou risco muito alto o sistema, um documento

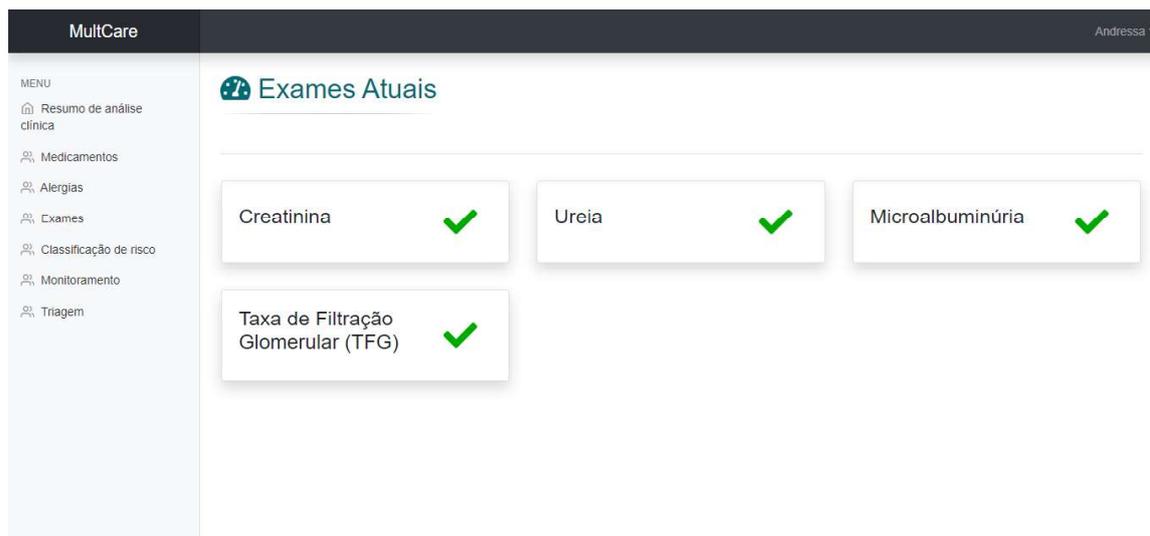


Figura 5.2: GUI principal do sistema.

clínico, estruturado usando o padrão HL7 CDA, é enviado por e-mail ao médico responsável pelo monitoramento do paciente. O CDA é um arquivo XML que contém os dados da análise de risco, uma árvore de decisão da análise de risco e o registro pessoal do paciente. Em caso de baixo risco, o sistema registra apenas os resultados da análise de risco para acompanhar a situação clínica do paciente e não envia o alerta ao médico. Isso automatiza a análise de risco e o compartilhamento solicitado anteriormente pelos usuários por eventos de botão [98].

Neste trabalho, como outra nova contribuição a partir dos resultados originalmente apresentados em [98] é proposto e desenvolvido um sistema especialista (como parte do MultCare Paciente) para encaminhar o paciente com DRC e fatores de risco a um serviço de saúde adequado em uma situação de emergência. Neste recurso é considerado o cenário em que o paciente está fora de sua cidade e não conhece as instalações corretas para receber tratamento, de acordo com a sua situação atual de saúde.

Com base em entrevistas semi-estruturas com o nefrologista mais experiente, o projeto inicial de uma funcionalidade de encaminhamento em emergência foi implementado com o software Expert Sinta. Os seguintes tópicos foram abordados:

- locais de atendimento em emergência;
- patologias a serem identificadas;
- sintomas;

- medicamentos associados com patologias.

Portanto, as seguintes patologias foram definidas para o projeto do sistema especialista: hiperglicemia, hipoglicemia, hipercalemia e hipocalemia. Para estas patologias, o paciente pode ser encaminhado para unidades de pronto atendimento ou emergências de hospitais, variando com base na situação de saúde atual do paciente.

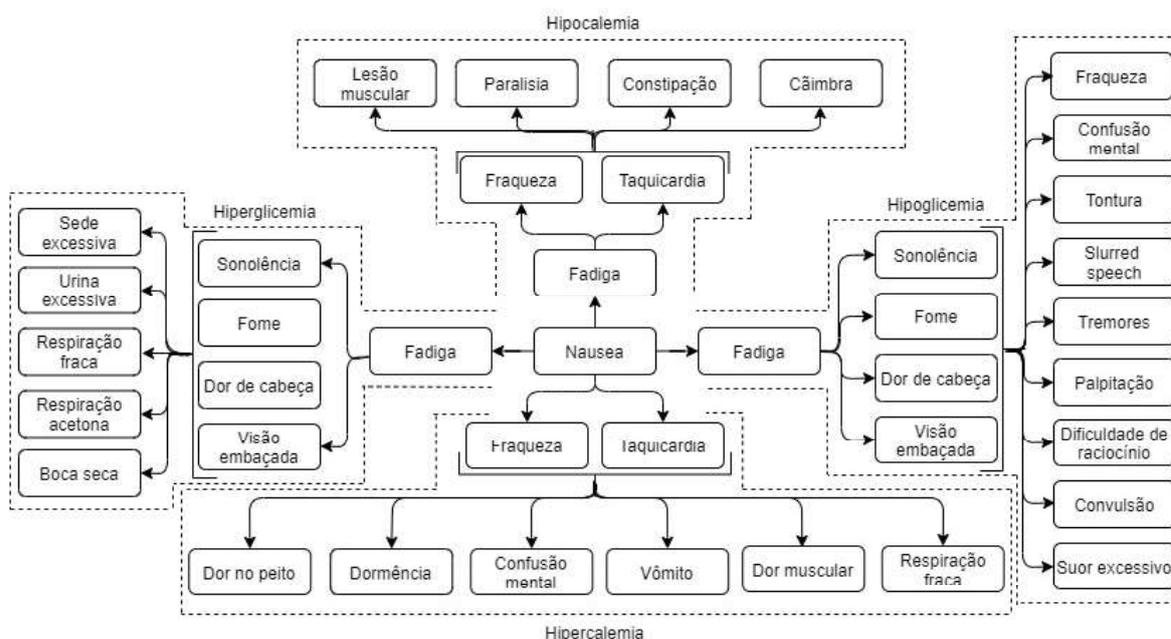


Figura 5.3: Resumo das primeiras decisões do sistema especialista.

A base de conhecimento do sistema foi projetada também com base em informações obtidas por meio de diretrizes médicas (*e.g.*, diretrizes da sociedade brasileira de diabetes<sup>2</sup>). A base foi elaborada levando em consideração informações sobre a condição clínica atual do paciente e considerando que outros fatores podem causar essas patologias, como, por exemplo, o uso de algumas medicações. Na Figura 5.3 são descritas as primeiras decisões usadas para identificar o risco de hiperglicemia, hipoglicemia, hipercalemia ou hipocalemia. A náusea é um sintoma compartilhado por todas as condições clínicas e, cada uma incluindo o sintoma, auxilia na identificação de uma condição específica. Além disso, o álcool, se consumido em excesso, e a insulina utilizada no tratamento do diabetes, podem aumentar o risco de hipoglicemia.

<sup>2</sup><https://www.diabetes.org.br/profissionais/images/2017/diretrizes/diretrizes-sbd-2017-2018.pdf>

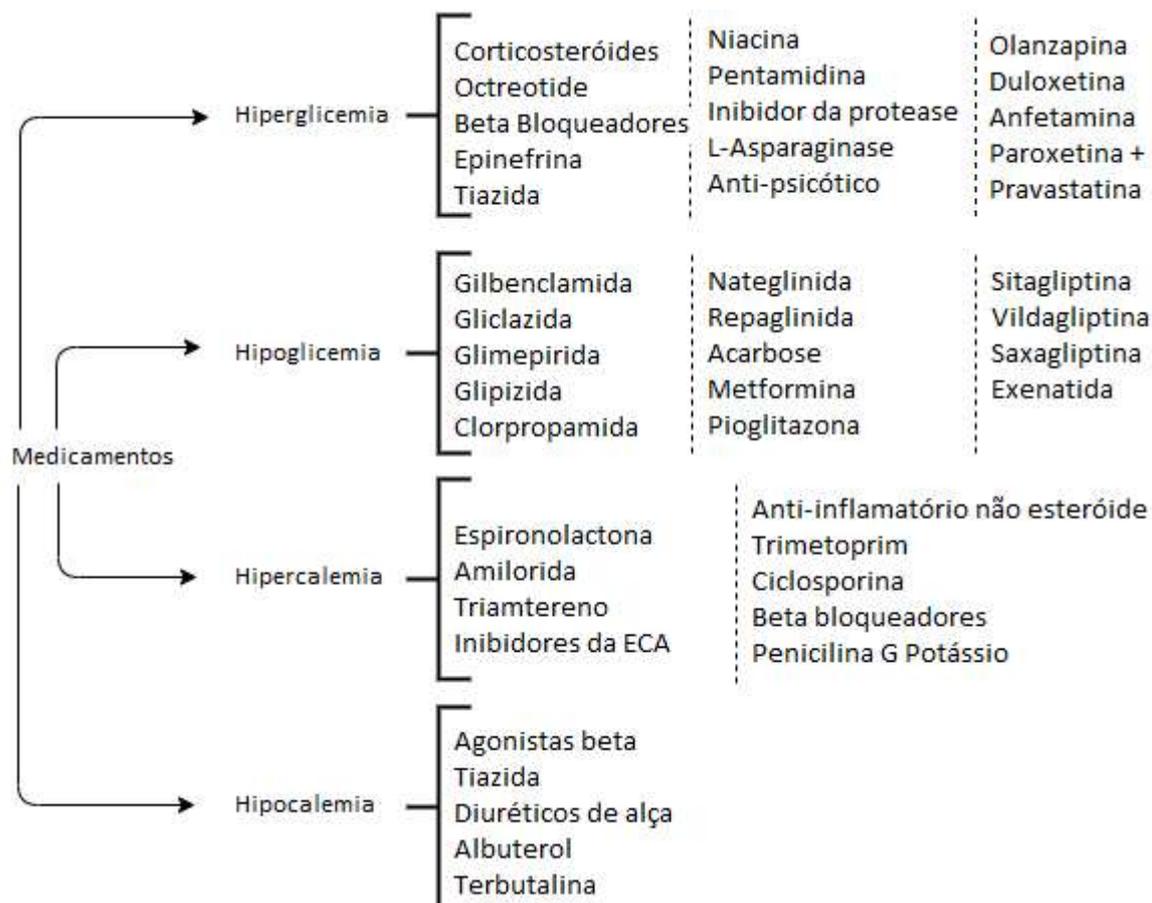


Figura 5.4: Lista de medicamentos relacionados com cada patologia.

A utilização de medicamentos específicos também pode resultar nas condições clínicas consideradas [62]. A possível ingestão de um medicamento é uma indicação relevante do risco de uma condição clínica específica. Na Figura 5.4 são descritos os medicamentos comumente ingeridos que podem levar a uma dessas patologias.

O sistema especialista foi projetado inicialmente utilizando o *Expert Sinta*. O objetivo foi simplificar o trabalho de implementação de sistemas especialistas [95]. Após a validação da base de regra do projeto inicial, o sistema especialista foi implementado, como parte do código-fonte do MultCare Paciente, usando a ferramenta *Dools*©, um sistema de regras de produção desenvolvido pela comunidade JBoss.

Para simplificar a apresentação das regras que compõem a base de conhecimento, na Figura 5.5 são apresentados, como uma árvore de decisão, os relacionamentos entre as perguntas que serão consideradas durante a identificação de hiperglicemia. A hiperglicemia, hipoglicemia, hipercalemia e hipocalemia são condições clínicas comum em pacientes que



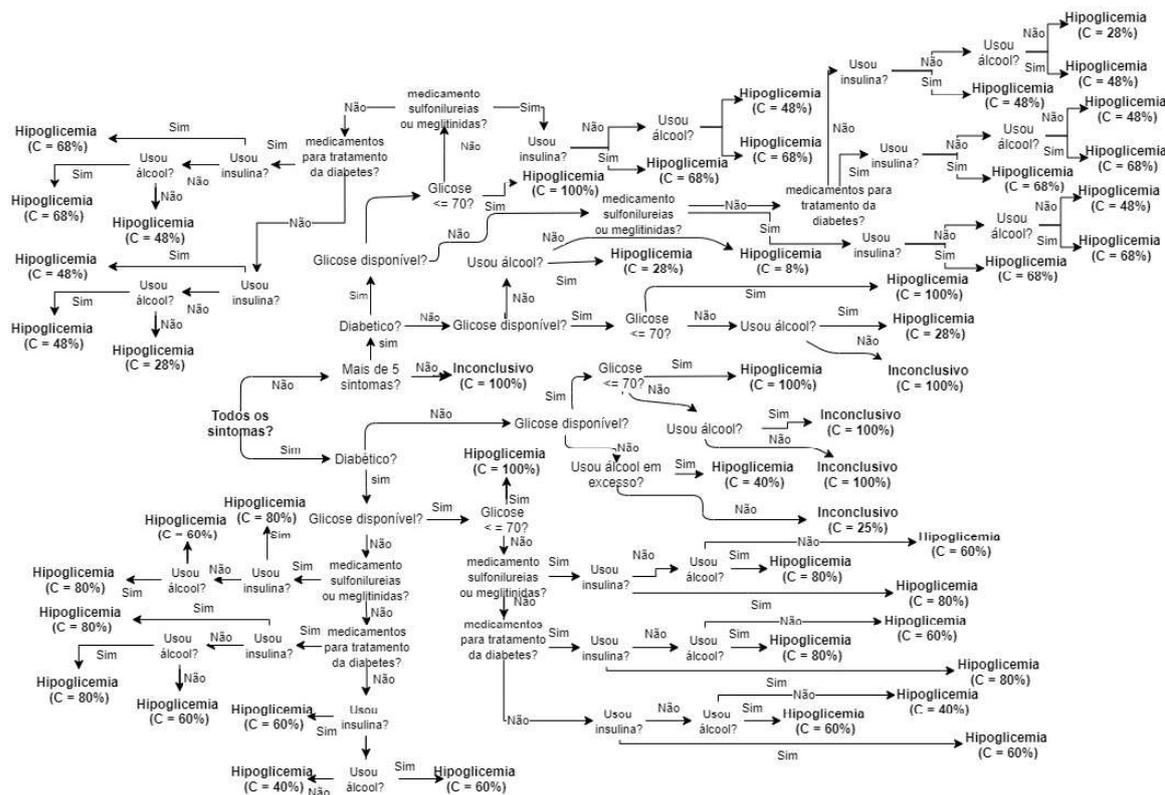


Figura 5.6: Relacionamentos entre perguntas para identificar o risco de hipoglicemia.

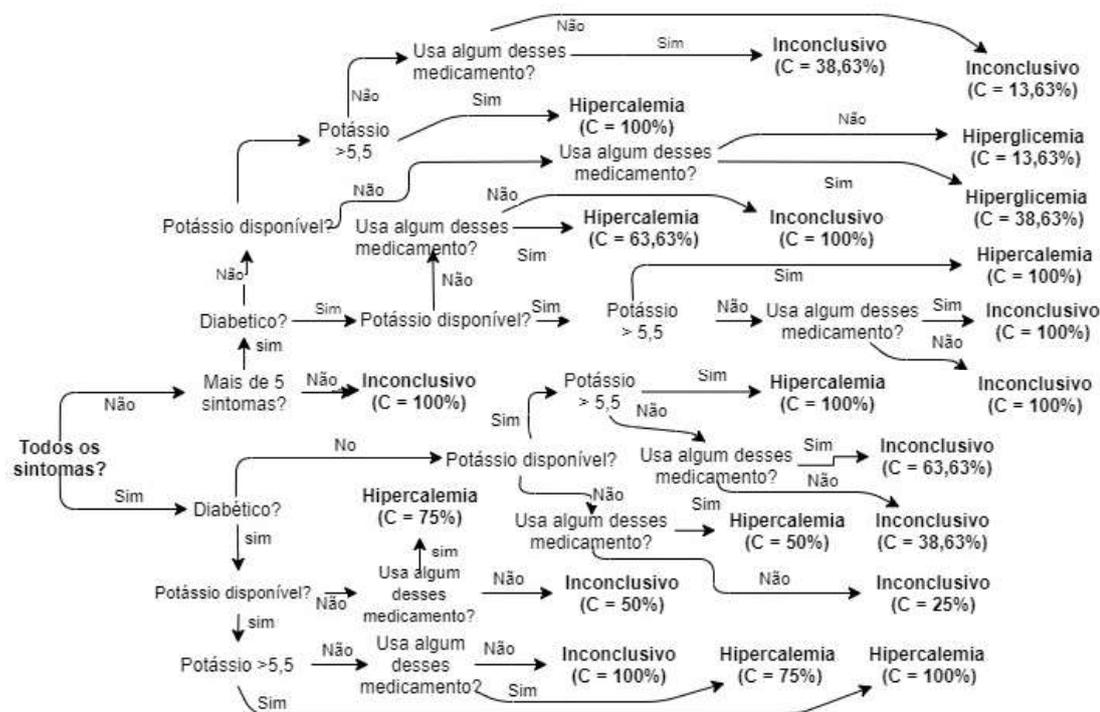


Figura 5.7: Relacionamentos entre perguntas para identificar o risco de hipercalcemia.

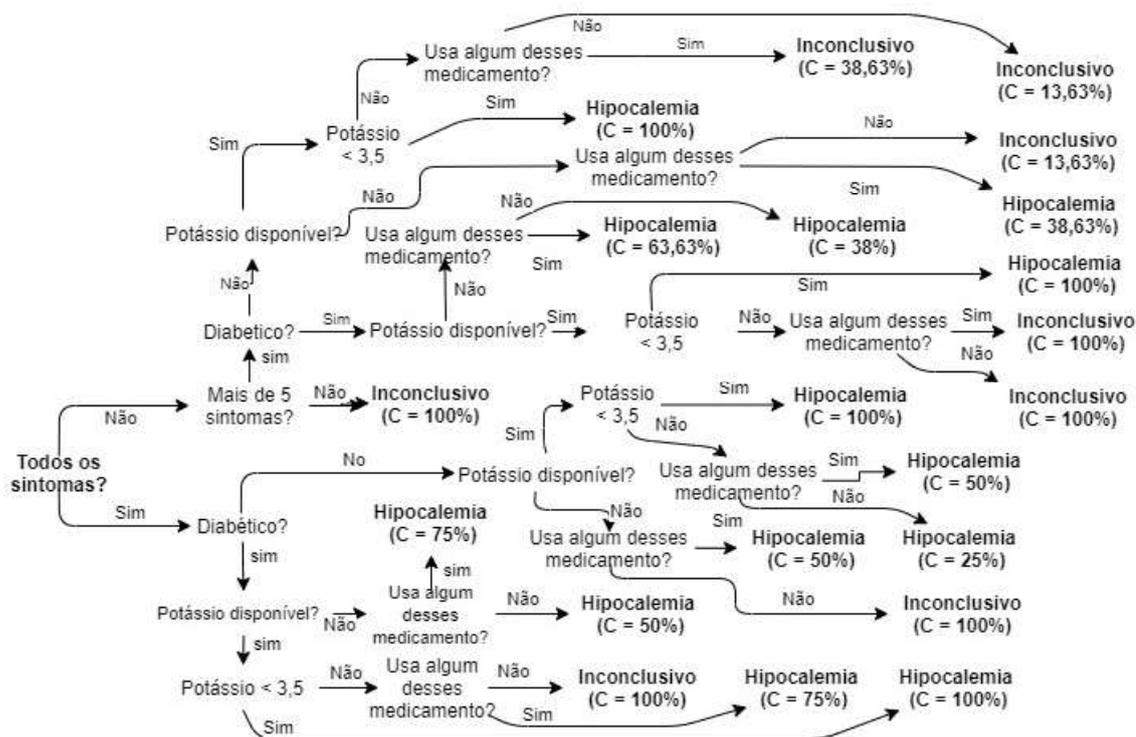


Figura 5.8: Relacionamentos entre perguntas para identificar o risco de hipocalemia.

se fez uso de insulina, fatores que influenciam nas patologias. O sistema verifica as regras e apresenta para o paciente a possível patologia indicando a unidade de referência apropriada.

The screenshot shows a web interface for a triage system. The header includes 'MultCare' and 'Andressa'. The main content area is titled 'Triagem' and contains three questions with radio button options:
 

- Você é diabético? (Sim, Não)
- Faz uso de insulina? (Sim, Não)
- Você consumiu álcool? (Sim, Não)

 A blue 'Próximo' button is located below the questions. A sidebar on the left lists menu items: 'Resumo de análise clínica', 'Medicamentos', 'Alergias', 'Exames', 'Classificação de risco', 'Monitoramento', and 'Triagem'.

Figura 5.9: GUI do sistema especialista para DM e álcool.

A base de conhecimento também foi apresentada ao nefrologista mais experiente para avaliar o sistema especialista, que revisou e analisou dados simulados. Os mesmos dados foram analisados usando o sistema para comparar avaliações de risco usando a estatística kappa. Na Tabela 5.1 é apresentada uma amostra do total de 112 cenários simulados usados,

representando 3 dos 21 caminhos da Figura 5.5 (indivíduos fictícios com hiperglicemia). Os 112 cenários abrangem todos os caminhos possíveis na base de conhecimento completa.

Tabela 5.1: Amostra de cenários simulados de hiperglicemia usados para avaliar o sistema.

<b>Id</b>	<b>Sintomas</b>	<b>Medicamentos</b>	<b>DM</b>	<b>Glicose</b>	<b>Encaminhamento</b>
1	Todos os sintomas	Não	-	-	UPA
2	Todos os sintomas	Sim	Não	-	UPA
3	Todos os sintomas	Sim	Não	120	UPA

Foram realizadas duas etapas de avaliação com o nefrologista experiente, usando o kappa de Cohen. Na primeira etapa (Tabela 5.2), o sistema especialista obteve concordância substancial ( $k = 0,6821$ ) e moderada ( $k = 0,5962$ ) com o nefrologista para classificação de risco e encaminhamento. O desacordo ocorreu porque o nefrologista considerou alguns dos cenários de risco de hipercalemia e hipocalemia como inconclusivos (Tabela 5.2, coluna 2). Na segunda etapa, a base de conhecimento foi corrigida e a avaliação resultou em 100% de concordância, em ambos, classificação de risco e encaminhamento.

Tabela 5.2: Resultados da primeira etapa da avaliação do sistema especialista.

<b>Risco</b>	<b>Kappa (Risco)</b>	<b>Encaminhamento</b>	<b>Kappa (Encaminhamento)</b>
Hiperglicemia	0.9361	UPA	0.5880
Hipoglicemia	0.8886	Hospital	0.5547
Hipercalemia	0.5089	Inconclusivo	1.0000
Hipocalemia	0.7404		
Inconclusivo	0.0290		
<b>Kappa Global</b>	<b>0.6903</b>		<b>0.5962</b>

# Capítulo 6

## Conclusões

Neste trabalho foi realizada uma análise comparativa a fim de identificar qual o classificador baseado em Aprendizado de Máquina (AM) e quais os atributos mais adequados para o diagnóstico de Doença Renal Crônica (DRC) nos países em desenvolvimento. Os classificadores baseados em AM apresentam diferentes níveis de precisão para o diagnóstico da DRC, dependendo do número de atributos considerados durante a classificação.

Os classificadores J48 e o *Random Forest* (RF) exibiram o melhor desempenho, apresentando concordância quase perfeita, usando o conjunto de dados da DRC, incluindo os atributos hipertensão, DM, creatinina, uréia, albuminúria, idade, sexo e Taxa de Filtração Glomerular (TFG). Esses atributos são comumente usados por nefrologistas para diagnosticar a DRC em países em desenvolvimento.

Com relação à implementação da nova versão do sistema MultCare, conseguiu-se projetar e avaliar, obtendo os resultados esperados nas funcionalidades projetadas, gerando desta forma as recomendações esperadas. Vale ressaltar que o Sistema Especialista (SE) não substitui a avaliação de um médico, mas visa auxiliar no monitoramento de pacientes com DRC. Este tipo de sistema é relevante para auxiliar na redução de atendimentos inadequados e na redução também de superlotações em hospitais. As superlotações de hospitais são problemas críticos, principalmente em estado de calamidade pública, como é o caso de pandemias. Muitos atendimentos que poderiam ser realizados em unidades de cuidado à saúde alternativas são, muitas vezes, realizados em hospitais, pela falta de conhecimento de pacientes sobre o local mais adequado para atendimento.

Existem algumas limitações na validade dos resultados deste trabalho. Por exemplo, o

---

conjunto de dados foi considerado pequeno. No entanto, não há conjuntos de dados disponíveis publicamente que apresentem avaliações de risco de DRC de indivíduos, considerando diretrizes médicas bem aceitas e nefrologistas com experiência em países em desenvolvimento. Os dados simulados foram aumentados garantindo a permanência na mesma classe de risco de DRC. Para melhorar a confiança, o método de validação cruzada *k-fold* foi usado para conduzir o experimento, e um nefrologista experiente analisou a validade dos dados aumentados (usados apenas no conjunto de treinamento). Portanto, os dados aumentados não impactaram negativamente a fase de teste dos classificadores de AM.

Outra limitação foi o fato de que os nefrologistas foram solicitados a analisar cada sujeito, o que é uma tarefa demorada, mesmo para apenas 60 sujeitos. Além disso, o tamanho do conjunto de dados exigia o aumento dos dados, modificando os valores dos biomarcadores da DRC para melhorar a confiança nas classificações. No sistema especialista, como limitação, está o fato da avaliação atual do sistema especialista não considerar a opinião dos pacientes. As entrevistas com o nefrologista experiente foram relevantes para diminuir os impactos negativos dessa limitação.

Como trabalhos futuros, tem-se: (1) a ampliação para outras patologias relacionadas com a DRC; (2) realizar testes de usabilidade; (3) analisar a interpretação do profissional apresentando o resultado dado por uma rede neural versus um resultado produzido por um modelo de fácil interpretação como a árvore J48; (4) realizar novos testes utilizando combinação de classificadores; e (5) aprimorar a avaliação do sistema inteligente de suporte à decisão. Testar a avaliação de risco requer melhorar o número de sujeitos no conjunto de dados da DRC. Além disso, testar o sistema especialista também requer entrevistas com os pacientes para avaliar a interface gráfica de usuário projetada, considerando a usabilidade e as percepções do usuário.

# Bibliografia

- [1] H. Ahmadi, G. Arji, L. Shahmoradi, R. Safdari, M. Nilashi, and M. Alizadeh. The application of internet of things in healthcare: a systematic literature review and classification. *Universal Access in the Information Society*, pages 1–33, 2019.
- [2] A. Y. Al-Hyari, A. M. Al-Tae, and M. A. Al-Tae. Clinical decision support system for diagnosis and management of chronic renal failure. In *2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, pages 1–6. IEEE, 2013.
- [3] A. Al Imran, M. N. Amin, and F. T. Johora. Classification of chronic kidney disease using logistic regression, feedforward neural network and wide & deep learning. In *2018 International Conference on Innovation in Engineering and Technology (ICIET)*, pages 1–6. IEEE, 2018.
- [4] T. C. Alberto and T. A. Almeida. Aprendizado de maquina aplicado na detecção automatica de comentarios indesejados. *Anais do X Encontro Nacional de Inteligência Artificial e Computacional (ENIAC'13), Fortaleza, Brazil*, 2013.
- [5] T. Alves. Mais de setenta por cento dos atendimentos no hge são de casos clínicos. *Disponivel: <http://www.saude.al.gov.br/2019/06/04/mais-de-70-dos-atendimentos-no-hge-sao-de-casos-clinicos/>, Acesso: setembro, 2020.*
- [6] R. Ani, G. Sasi, U. R. Sankar, and O. Deepa. Decision support system for diagnosis and prediction of chronic renal failure using random subspace classification. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 1287–1292. IEEE, 2016.

- [7] M. D. Basar and A. Akan. Detection of chronic kidney disease by using ensemble classifiers. In *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*, pages 544–547. IEEE, 2017.
- [8] M. G. Bastos, R. Bregman, and G. M. Kirsztajn. Doença renal crônica: frequente e grave, mas também prevenível e tratável. *Revista da Associação Médica Brasileira*, 56(2):248–253, 2010.
- [9] M. G. Bastos, W. d. Carmo, R. R. Abrita, E. d. Almeida, D. Mafra, D. d. Costa, J. d. A. Gonçalves, L. A. d. Oliveira, F. R. d. Santos, and R. B. d. Paula. Doença renal crônica: problemas e soluções. *J Bras Nefrol*, 26(4):202–15, 2004.
- [10] M. G. Bastos and G. M. Kirsztajn. Doença renal crônica: importância do diagnóstico precoce, encaminhamento imediato e abordagem interdisciplinar estruturada para melhora do desfecho em pacientes ainda não submetidos à diálise. *Brazilian Journal of Nephrology*, 33(1):93–108, 2011.
- [11] G. E. d. A. P. Batista et al. *Pré-processamento de dados em aprendizado de máquina supervisionado*. PhD thesis, Universidade de São Paulo, 2003.
- [12] N. Bhargava, G. Sharma, R. Bhargava, and M. Mathuria. Decision tree analysis on j48 algorithm for data mining. *Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering*, 3(6), 2013.
- [13] R. A. d. C. Bianchi. *Uso de heurísticas para a aceleração do aprendizado por reforço*. PhD thesis, Universidade de São Paulo, 2004.
- [14] F. N. Bonifácio. Comparação entre as redes neurais artificiais mlp, rbf e lvq na classificação de dados. Master’s thesis, 2010.
- [15] J. Bravo-Zúñiga, J. Gálvez-Inga, P. Carrillo-Onofre, R. Chávez-Gómez, and P. Castro-Monteverde. Early detection of chronic renal disease: coordinated work between primary and specialized care in an ambulatory renal network of peru. *Brazilian Journal of Nephrology*, 41(2):176–184, 2019.

- [16] M. P. Brown, W. N. Grundy, D. Lin, N. Cristianini, C. W. Sugnet, T. S. Furey, M. Ares, and D. Haussler. Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proceedings of the National Academy of Sciences*, 97(1):262–267, 2000.
- [17] L. H. C. Roubicek, P. Brunet, F. L. X. Thirion, B. Dussol, K. Jaber, D. Andrieu, P. Ramananarivo, and Y. Berland. Timing of nephrology referral: influence on mortality and morbidity. *American journal of kidney diseases*, 36(1):35–41, 2000.
- [18] A. CARVALHO, K. Faceli, A. LORENA, and J. Gama. Inteligência artificial—uma abordagem de aprendizado de máquina. *Rio de Janeiro: LTC*, 2011.
- [19] A. Charleonnann, T. Fufaung, T. Niyomwong, W. Chokchueypattanakit, S. Suwanawach, and N. Ninchawee. Predictive analytics for chronic kidney disease using machine learning techniques. In *2016 Management and Innovation Technology International Conference (MITicon)*, pages MIT–80. IEEE, 2016.
- [20] C.-J. Chen, T.-W. Pai, H. Fujita, C.-H. Lee, Y.-T. Chen, K.-S. Chen, and Y.-C. Chen. Stage diagnosis for chronic kidney disease based on ultrasonography. In *2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pages 525–530. IEEE, 2014.
- [21] Z. Chen, X. Zhang, and Z. Zhang. Clinical risk assessment of patients with chronic kidney disease by using clinical data and multivariate models. *International urology and nephrology*, 48(12):2069–2075, 2016.
- [22] E. A. Cherman. *Aprendizado de máquina multirrótulo: explorando a dependência de rótulos e o aprendizado ativo*. PhD thesis, Universidade de São Paulo, 2013.
- [23] N. Chetty, K. S. Vaisla, and S. D. Sudarsan. Role of attributes selection in classification of chronic kidney disease patients. In *2015 International Conference on Computing, Communication and Security (ICCCS)*, pages 1–6. IEEE, 2015.
- [24] G. Chiarini, P. Ray, S. Akter, C. Masella, and A. Ganz. mhealth technologies for chronic diseases and elders: a systematic review. *IEEE Journal on Selected Areas in Communications*, 31(9):6–18, 2013.

- [25] R. K. Chiu, R. Y. Chen, S.-A. Wang, and S.-J. Jian. Intelligent systems on the cloud for the early detection of chronic kidney disease. In *2012 International Conference on Machine Learning and Cybernetics*, volume 5, pages 1737–1742. IEEE, 2012.
- [26] D. W. Cockcroft and H. Gault. Prediction of creatinine clearance from serum creatinine. *Nephron*, 16(1):31–41, 1976.
- [27] H. G. D. W. Cockcroft. Prediction of creatinine clearance from serum creatinine. *Nephron*, 16(1):31–41, 1976.
- [28] P. da Saúde. Sistema Único de saúde (sus): estrutura, princípios e como funciona. Disponível:<http://portalms.saude.gov.br/sistema-unico-de-saudeestrutura>, Acesso: setembro, 2020.
- [29] M. Damasceno. Introdução a mineração de dados utilizando o weka. Disponível: <http://connepi.ifal.edu.br/ocs/index.php/connepi/CONNepI2010/paper/viewFile/258/207>, Acesso: setembro, 2019.
- [30] R. C. Das and T. Alam. Location based emergency medical assistance system using openstreetmap. In *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*, pages 1–5. IEEE, 2014.
- [31] I. De la Torre-Díez, B. Martínez-Pérez, M. López-Coronado, J. R. Díaz, and M. M. López. Decision support systems and applications in ophthalmology: literature and commercial review focused on mobile apps. *Journal of medical systems*, 39(1):174, 2015.
- [32] J. DeFranco, M. Kassab, P. Laplante, and N. Laplante. The nonfunctional requirement focus in medical device software: a systematic mapping study and taxonomy. *Innovations in Systems and Software Engineering*, 13(2):81–100, 2017.
- [33] E. Fontana, D. Biduski, A. C. B. De Marchi, and R. Rieder. Smart environments using gesture-based interactions for health: a systematic review. In *2015 XVII Symposium on Virtual and Augmented Reality*, pages 170–175. IEEE, 2015.
- [34] A. Forbes and H. Gallagher. Chronic kidney disease in adults: assessment and management. *Clinical Medicine*, 20(2):128, 2020.

- [35] E. Frank, M. Hall, G. Holmes, R. Kirkby, B. Pfahringer, I. H. Witten, and L. Trigg. *Weka-A Machine Learning Workbench for Data Mining*, pages 1269–1277. Springer US, Boston, MA, 2010.
- [36] E. Frank, M. Hall, L. Trigg, G. Holmes, and I. H. Witten. Data mining in bioinformatics using Weka. *Bioinformatics*, 20(15):2479–2481, 04 2004.
- [37] G1. Mais de 60% dos hospitais públicos estão sempre superlotados. *Disponível: <http://glo.bo/1gthZIC>*, Acesso: setembro, 2020.
- [38] R. Ganesan and V. V. Chamundeeswari. Composite algorithm for pervasive healthcare system—a solution to find optimized route for closest available health care facilities. *Multimedia Tools and Applications*, 79(7):5125–5148, 2020.
- [39] M. H. Ghunaim, K. S. Alkhalaf, B. A. Altwaijri, and Y. M. Seddiq. Phewna: an expert system framework for children health care and awareness. In *2016 International Conference on Bio-engineering for Smart Technologies (BioSMART)*, pages 1–3. IEEE, 2016.
- [40] R. M. Gonçalves, L. S. Coelho, C. P. Krueger, and B. Heck. Modelagem preditiva de linha de costa utilizando redes neurais artificiais. *Boletim de Ciências Geodésicas*, 16(3), 2010.
- [41] E. Gonzalez, R. Peña, A. Avila, C. Vargas-Rosales, and D. Munoz-Rodriguez. A systematic review on recent advances in mhealth systems: deployment architecture for emergency response. *Journal of healthcare engineering*, 2017, 2017.
- [42] H. P. Graf, E. Cosatto, L. Bottou, I. Dourdanovic, and V. Vapnik. Parallel support vector machines: The cascade svm. In *Advances in neural information processing systems*, pages 521–528, 2005.
- [43] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: An update. *SIGKDD Explor. Newsl.*, 11(1):10–18, Nov. 2009.

- [44] K. Z. Hasan and M. Z. Hasan. Performance evaluation of ensemble-based machine learning techniques for prediction of chronic kidney disease. In *Emerging Research in Computing, Information, Communication and Applications*, pages 415–426. Springer, 2019.
- [45] G. Holmes, A. Donkin, and I. H. Witten. Weka: a machine learning workbench. In *Proceedings of ANZIIS '94 - Australian New Zealand Intelligent Information Systems Conference*, pages 357–361, 1994.
- [46] P. S. Horst. *Avaliação do conhecimento adquirido por algoritmos de aprendizado de máquina utilizando exemplos*. PhD thesis, Universidade de São Paulo, 1999.
- [47] J.-W. Hsieh, C.-H. Lee, Y.-C. Chen, W.-S. Lee, and H.-F. Chiang. Stage classification in chronic kidney disease by ultrasound image. In *Proceedings of the 29th International Conference on Image and Vision Computing New Zealand*, pages 271–276, 2014.
- [48] J. Huang, Y. Wei, J. Yi, and M. Liu. An improved knn based on class contribution and feature weighting. In *2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, pages 313–316, 2018.
- [49] L. A. Inker, B. C. Astor, C. H. Fox, T. Isakova, J. P. Lash, C. A. Peralta, M. K. Tamura, and H. I. Feldman. Kdoqi us commentary on the 2012 kdigo clinical practice guideline for the evaluation and management of ckd. *American Journal of Kidney Diseases*, 63(5):713–735, 2014.
- [50] J. P. J. Han, M. Kamber. 9 - classification: Advanced methods. In J. P. J. Han, M. Kamber, editor, *Data Mining (Third Edition)*, The Morgan Kaufmann Series in Data Management Systems, pages 393 – 442. Morgan Kaufmann, Boston, third edition edition, 2012.
- [51] F. F. Jahantigh, B. Malmir, and B. A. Avilaq. A computer-aided diagnostic system for kidney disease. *Kidney research and clinical practice*, 36(1):29, 2017.
- [52] Y. K. Jain. Upendra: An efficient intrusion detection based on decision tree classifier using feature reduction. *International Journal of scientific and research Publications*, 2(1), 2012.

- [53] R. Jeewantha, M. N. Halgamuge, A. Mohammad, and G. Ekici. Classification performance analysis in medical science: using kidney disease data. In *Proceedings of the 2017 International Conference on Big Data Research*, pages 1–6, 2017.
- [54] L. Jena and N. K. Kamila. Distributed data mining classification algorithms for prediction of chronic-kidney-disease. *International Journal of Emerging Research in Management & Technology*, 4(11):110–118, 2015.
- [55] S. Keele et al. Guidelines for performing systematic literature reviews in software engineering. Technical report, Technical report, Ver. 2.3 EBSE Technical Report. EBSE, 2007.
- [56] C. E. King and M. Sarrafzadeh. A survey of smartwatches in remote health monitoring. *Journal of healthcare informatics research*, 2(1-2):1–24, 2018.
- [57] M. Kumar. Prediction of chronic kidney disease using random forest machine learning algorithm. *International Journal of Computer Science and Mobile Computing*, 5(2):24–33, 2016.
- [58] V. Kunwar, K. Chandel, A. S. Sabitha, and A. Bansal. Chronic kidney disease analysis using data mining classification techniques. In *2016 6th International Conference-Cloud System and Big Data Engineering (Confluence)*, pages 300–305. IEEE, 2016.
- [59] E. J. Lamb, A. S. Levey, and P. E. Stevens. The kidney disease improving global outcomes (kdigo) guideline update for chronic kidney disease: evolution not revolution. *Clinical chemistry*, 59(3):462–465, 2013.
- [60] A. S. Levey, J. P. Bosch, J. B. Lewis, T. Greene, N. Rogers, and D. Roth. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. *Annals of internal medicine*, 130(6):461–470, 1999.
- [61] A. S. Levey, L. A. Stevens, C. H. Schmid, Y. L. Zhang, A. F. Castro, H. I. Feldman, J. W. Kusek, P. Eggers, F. Van Lente, T. Greene, et al. A new equation to estimate glomerular filtration rate. *Annals of internal medicine*, 150(9):604–612, 2009.

- [62] H. S. Libanes. Medicamentos e dieta errada estão entre as causas de hipoglicemia nos pacientes com diabetes. *Disponível: <https://hospitalsiriolibanes.org.br/sua-saude/Paginas/medicamentos-dieta-errada-estao-entre-causas-hipoglicemia-nos-pacientes-diabetes.aspx>, Acesso: março, 2020.*
- [63] L. Liu, Y. Ni, N. Zhang, et al. Mining patient-specific and contextual data with machine learning technologies to predict cancellation of children's surgery. *International journal of medical informatics*, 129:234–241, 2019.
- [64] M. T. Lokhman, A. R. A. Rasam, et al. Reforming health care facility using geographical information system. In *2012 IEEE Control and System Graduate Research Colloquium*, pages 267–270. IEEE, 2012.
- [65] C. D. C. Lorenzett and A. V. Telöcken. Estudo comparativo entre os algoritmos de mineração de dados random forest e j48 na tomada de decisão. *Simpósio de Pesquisa e Desenvolvimento em Computação (SPDC)*, 2(1), 2016.
- [66] S. R. Loya, K. Kawamoto, C. Chatwin, and V. Huser. Service oriented architecture for clinical decision support: A systematic review and future directions. *Journal of medical systems*, 38(12):140, 2014.
- [67] J. S. Malak, H. Zeraati, F. S. Nayeri, R. Safdari, and A. D. Shahraki. Neonatal intensive care decision support systems using artificial intelligence techniques: a systematic review. *Artificial Intelligence Review*, 52(4):2685–2704, 2019.
- [68] R. D. Mendes. Inteligência artificial: sistemas especialistas no gerenciamento da informação. *Ciência da Informação*, 26(1), 1997.
- [69] M. Monard and J. Baranauskas. Conceitos sobre aprendizado de máquina. sistemas inteligentes, reze, so, 2003.
- [70] M. G. N. Friedman, D. Geiger. Bayesian network classifiers. *Machine Learning*, 29, 1997.
- [71] F. A. Narudin, A. Feizollah, N. B. Anuar, and A. Gani. Evaluation of machine learning classifiers for mobile malware detection. *Soft Computing*, 20(1):343–357, 2016.

- [72] J. Neves, M. R. Martins, J. Vilhena, J. Neves, S. Gomes, A. Abelha, J. Machado, and H. Vicente. A soft computing approach to kidney diseases evaluation. *Journal of medical systems*, 39(10):131, 2015.
- [73] M. A. Nielsen. *Neural networks and deep learning*, volume 2018.
- [74] A. Ogunleye and Q.-G. Wang. Enhanced xgboost-based automatic diagnosis system for chronic kidney disease. In *2018 IEEE 14th International Conference on Control and Automation (ICCA)*, pages 805–810. IEEE, 2018.
- [75] W. H. Organization. Preventing chronic diseases : a vital investment : Who global report, 2005.
- [76] D. Pal, T. Triyason, and S. Funikul. Smart homes and quality of life for the elderly: A systematic review. In *2017 IEEE International Symposium on Multimedia (ISM)*, pages 413–419. IEEE, 2017.
- [77] T. R. Patil. Performance analysis of naive bayes and j48 classification algorithm for data classification. *Journal of Computer Science and Applications*, 6(2), 2013.
- [78] A. Pimenta, P. Valentim, D. Santos, and M. Neto. Weka-g: mineração de dados paralela em grades computacionais. *Revista de Sistemas de Informação da FSMA*, 4:2, 2009.
- [79] H. Polat, H. D. Mehr, and A. Cetin. Diagnosis of chronic kidney disease based on support vector machine by feature selection methods. *Journal of medical systems*, 41(4):55, 2017.
- [80] H. Polat, H. D. Mehr, and A. Cetin. Diagnosis of chronic kidney disease based on support vector machine by feature selection methods. *Journal of medical systems*, 41(4):55, 2017.
- [81] N. Rajalakshmi. Bharathi, “diagnosis and classification of level of kidney function using associative neural network and polynomial neural network”. *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, 4(4):724–738, 2013.

- [82] B. Ravindra, N. Sriraam, and M. Geetha. Classification of non-chronic and chronic kidney disease using svm neural networks. *International Journal of Engineering & Technology*, 7(1.3):191–194, 2018.
- [83] L. J. Rubini and P. Eswaran. Generating comparative analysis of early stage prediction of chronic kidney disease. *International Journal of Modern Engineering Research (IJMER)*, 5(7):49–55, 2015.
- [84] H. L. P. Rufino. *Algoritmo de aprendizado supervisionado-baseado em máquinas de vetores de suporte-uma contribuição para o reconhecimento de dados desbalanceados*. PhD thesis, Universidade Federal de Uberlândia, 2011.
- [85] M. A. S. Iqbal, W. Altaf and M. U. G. K. W. Mahmood. Application of intelligent agents in health-care: review. *Artificial Intelligence Review*, 46(1):83–112, Jun 2016.
- [86] B. H. S. Kajornkasirat, N. Chanapai. Smart health monitoring system with iot. In *2018 IEEE Symposium on Computer Applications Industrial Electronics (ISCAIE)*, pages 206–211, April 2018.
- [87] S. Sahu and B. M. Mehtre. Network intrusion detection system using j48 decision tree. In *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 2023–2026. IEEE, 2015.
- [88] T. Saito and M. Rehmsmeier. The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets. *PloS one*, 10(3), 2015.
- [89] A. Salekin and J. Stankovic. Detection of chronic kidney disease and selecting important predictive attributes. In *2016 IEEE International Conference on Healthcare Informatics (ICHI)*, pages 262–270. IEEE, 2016.
- [90] M. K. Sanches. *Aprendizado de máquina semi-supervisionado: proposta de um algoritmo para rotular exemplos a partir de poucos exemplos rotulados*. PhD thesis, Universidade de São Paulo, 2003.

- [91] G. J. Schwartz, A. Munoz, M. F. Schneider, R. H. Mak, F. Kaskel, B. A. Warady, and S. L. Furth. New equations to estimate gfr in children with ckd. *Journal of the American Society of Nephrology*, 20(3):629–637, 2009.
- [92] R. d. C. C. Sesso, A. A. Lopes, F. S. Thomé, J. R. Lugon, and E. A. Burdmann. Brazilian dialysis census, 2009. *Brazilian Journal of Nephrology*, 32(4):374–384, 2010.
- [93] A. Singh, G. Nadkarni, O. Gottesman, S. B. Ellis, E. P. Bottinger, and J. V. Guttag. Incorporating temporal ehr data in predictive models for risk stratification of renal function deterioration. *Journal of biomedical informatics*, 53:220–228, 2015.
- [94] P. Sinha and P. Sinha. Comparative study of chronic kidney disease prediction using knn and svm. *International Journal of Engineering Research and Technology*, 4(12):608–12, 2015.
- [95] L. E. Sinta. Expert sinta. uma ferramenta visual para criação de sistemas especialistas-versão 1.1-manual do usuário. *Laboratório de Inteligência Artificial/LIA-UFC, CE*, 1996.
- [96] A. Sobrinho, L. D. da Silva, A. Perkusich, M. E. Pinheiro, and P. Cunha. Design and evaluation of a mobile application to assist the self-monitoring of the chronic kidney disease in developing countries. *BMC medical informatics and decision making*, 18(1):7, 2018.
- [97] A. Sobrinho, A. C. D. S. Queiroz, L. D. Da Silva, E. D. B. Costa, M. E. Pinheiro, and A. Perkusich. Computer-aided diagnosis of chronic kidney disease in developing countries: A comparative analysis of machine learning techniques. *IEEE Access*, 8:25407–25419, 2020.
- [98] C. Sobrinho, Á. Á. de Carvalho, et al. Estudo e desenvolvimento de uma solução para o auxílio ao cuidado à saúde. 2013.
- [99] M. Somvanshi and P. Chavan. A review of machine learning techniques using decision tree and support vector machine. In *2016 International Conference on Computing Communication Control and automation (ICCUBEA)*, pages 1–7, 2016.

- [100] C. S. Souza, C. B. Zamboni, D. G. da SILVA, S. Metairon, et al. Uso de mini-espectrômetro de fluorescência de raios-x como alternativa para prática clínica de dialisados. 2020.
- [101] S. Taneja, C. Gupta, S. Aggarwal, and V. Jindal. Mfz-knn — a modified fuzzy based k nearest neighbor algorithm. In *2015 International Conference on Cognitive Computing and Information Processing(CCIP)*, pages 1–5, 2015.
- [102] L. R. M. The. Opening the black box of machine learning. *The Lancet. Respiratory medicine*, 6(11):801, 2018.
- [103] F. S. Thoma, R. C. Sesso, A. A. Lopes, J. R. Lugon, and C. T. Martins. Brazilian chronic dialysis survey 2017. *Brazilian Journal of Nephrology*, 41:208 – 214, 06 2019.
- [104] K. Topuz, F. D. Zengul, A. Dag, A. Almehti, and M. B. Yildirim. Predicting graft survival among kidney transplant recipients: A bayesian decision support model. *Decision Support Systems*, 106:97–109, 2018.
- [105] S. Vijayarani, S. Dhayanand, et al. Data mining classification algorithms for kidney disease prediction. *International Journal on Cybernetics & Informatics (IJCI)*, 4(4):13–25, 2015.
- [106] P. D. Wasserman and T. Schwartz. Neural networks. ii. what are they and why is everybody so interested in them now? *IEEE Expert*, 3(1):10–15, 1988.
- [107] M. S. Wibawa, I. M. D. Maysanjaya, and I. M. A. W. Putra. Boosted classifier and features selection for enhancing chronic kidney disease diagnose. In *2017 5th International Conference on Cyber and IT Service Management (CITSM)*, pages 1–6. IEEE, 2017.
- [108] C. Wohlin. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, pages 1–10, 2014.

- 
- [109] L. Xun, W. Xiaoming, L. Ningshan, and L. Tanqi. Application of radial basis function neural network to estimate glomerular filtration rate in chinese patients with chronic kidney disease. In *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, volume 15, pages V15–332. IEEE, 2010.
- [110] B. C. Zapata, J. L. Fernández-Alemán, A. Idri, and A. Toval. Empirical studies on usability of mhealth apps: a systematic literature review. *Journal of medical systems*, 39(2):1, 2015.