**TESTES DE ALGORITMOS DO WEKA PARA AUTOMAÇÃO E PRECISÃO NA ANÁLISE CARDÍACA: AVANÇOS NA CLASSIFICAÇÃO DE ELETROCARDIOGRAMAS**

Autor Principal: Carlos Augusto de Sousa Alves ¹

Coautor 1: Marília Vargas Silva ²

Coautor 2: Raphaela Silva de Andrade Machado 3

**RESUMO**

*O projeto busca demonstrar a eficácia da aplicação de aprendizado de máquina na detecção de diagnósticos em eletrocardiogramas (ECGs). O objetivo final é fornecer uma visão abrangente sobre a aplicação do aprendizado de máquina com o uso do software Weka, destacando suas potenciais contribuições para o campo da cardiologia e melhorando a precisão e eficiência no diagnóstico de doenças cardíacas. Contudo, não foi possível fazer uma avaliação completa de todos os modelos encontrados no Weka, logo com uma breve pesquisa no Google com intuito de procurar os melhores exemplares para o presente trabalho foram encontrados os seguintes modelos: Zeror, Randomtree, Oner, Randomforest e j48.*

**Palavras-chave:** Aprendizagem de máquina. Diagnóstico. Eletrocardiograma (ECGs). Doenças cardíacas. *Software Weka*.

**1 INTRODUÇÃO**

As doenças cardiovasculares (DCV) causaram um terço das mortes globais em 2019, conforme os dados do estudo Global Burden Diseases (GBD). Elas foram a principal causa de DALYs (Anos de vida ajustados por incapacidade). Essa mesma tendência ocorre no Brasil, onde a prevalência de DCV é estimada em 6,1% da população adulta, com uma incidência de 475 casos para cada 100.000 habitantes (Santana Júnior, 2023).

A inteligência artificial (IA) tem ganhado destaque como uma ferramenta promissora na área médica, trazendo soluções inovadoras para a previsão, diagnóstico e tratamento de doenças. Na cardiologia, a IA está sendo utilizada de diversas formas, como no desenvolvimento de algoritmos para análise de imagens médicas e na criação de sistemas de apoio à decisão clínica. Por exemplo, softwares de IA podem auxiliar na interpretação de exames cardíacos, como ecocardiogramas e angiografias, detectando anomalias e contribuindo para o diagnóstico precoce de doenças cardiovasculares. Além disso, sistemas baseados em IA conseguem analisar grandes volumes de dados clínicos, identificando padrões e correlações entre variáveis, o que possibilita uma avaliação mais precisa do risco cardiovascular em cada paciente (Barbosa et al., 2024).

Desenvolver métodos mais acessíveis que possam simplificar o rastreamento de DCV é fundamental para a saúde pública. Apesar de a ecocardiografia ser o método de referência para detectar DCVs, trata-se de um exame caro e demorado. Por isso, é importante adotar tecnologias mais avançadas para a triagem de DCVs, que ofereçam precisão diagnóstica sem custos elevados ou grande demanda de tempo. O uso de um ECG com integração de IA pode ser uma solução viável para aumentar a detecção precoce de pacientes com DCV (Guimarães; Neves; Souza, 2024).

Weka (Waikato Environment for Knowledge Analysis) é uma interface gráfica que reúne uma variedade de algoritmos de mineração de dados para tarefas como classificação, regressão, agrupamento e associação. Foi criada pelo Departamento de Ciências da Computação da Universidade de Waikato, na Nova Zelândia. Desenvolvida em Java, sua execução é possível em diferentes sistemas operacionais, garantindo alta portabilidade. Além disso, Weka é distribuída sob a licença General Public License (GPL), criada por Richard Matthew Stallman em 1989 como parte do projeto GNU da Free Software Foundation, permitindo a modificação de seu código-fonte (Campos Neto, 2016).

Nesse cenário, é fundamental explorar como essa abordagem pode superar as restrições dos métodos tradicionais, proporcionando uma alternativa mais atual, eficiente e precisa para o diagnóstico e monitoramento dessas doenças.

**2 OBJETIVO**

O principal objetivo desse trabalho é avaliar a acurácia dos modelos de algoritmos do WEKA, um software que possui algoritmos de aprendizado de máquina de dados escritos em JAVA. Nele existem modelos que podem ser treinados para classificações de instâncias (amostras) de ECGs. Com análise desses algoritmos o trabalho visa apresentar quais os modelos com maior desempenho para automatizar as avaliações de eletrocardiogramas, com maior velocidade e precisão, desempenhando o papel de auxiliar nos diagnósticos de doenças cardíacas.

**3 MÉTODO**

Foi realizada uma pesquisa bibliográfica em bases acadêmicas, como Google Acadêmico, Pubmed e Science Direct com as seguintes palavras-chave “aprendizado em máquina”, “diagnósticos”, “eletrocardiogramas”, no recorte temporal de 2014 a 2024. Os critérios de exclusão foram aqueles que fugiam ao tema ou estavam fora do período definido.

Para a execução desse trabalho foi utilizado o WEKA que é um *software* de código aberto, para mineração de dados que possui um conjunto de algoritmos de classificação, (https://sourceforge.net/projects/weka/.). Para fazer as classificações dos algoritmos contidos nele tem como relação uma *dataset* de arritmias cardíacas contendo 452 amostras em formato ARFF, (arff-datasets/classification/arrhythmia.arff at master · renatopp/arff-datasets · GitHub).

Para iniciar os testes foi feito o *download* do *software* e do arquivo logo em seguida esse arquivo foi aberto na área de classificação, escolhido o modo de *use trainning set* que usa o conjunto de dados de treinamento para testar e treinar o modelo que utiliza validação cruzada k-fold onde o conjunto é dividido, treinado e testado em k partes. Isso ajuda a avaliar a performance do modelo e garantir que ele generalize bem para novos dados. Em seguida inicia-se o processo podendo verificar os resultados na tabela ao lado. Abaixo a tabela mostra os tipos de classificadores e os algoritmos aplicados.

|  |  |
| --- | --- |
| **CLASSIFICADORES** | **ALGORITMOS** |
| Rules(regras) | Zeror |
| Tress(árvore de decisão) | Randomtree |
| Rules(regras) | Oner |
| Tress(árvore de decisão) | Randomforest |
| Tress(árvore de decisão) | J48 |
| Tabela 1 dados Weka. | |

**4 RESULTADOS E DISCUSSÃO**

No modelo (Randomtree) a árvore possui um total de 311 nós. O modelo foi construído em 0,1 segundos, o modelo foi avaliado no conjunto de treinamento e obteve uma taxa de classificação correta de 100%, classificando corretamente todas as 452 instâncias (amostras) desse conjunto. As métricas de avaliação, como erro médio absoluto, erro relativo absoluto e erro quadrático relativo, foram todas calculadas como 0%. Isso indica que o modelo obteve uma correspondência perfeita com os rótulos do conjunto de treinamento. A matriz de confusão mostra a distribuição das classificações do modelo em relação aos rótulos corretos. Observamos que, para a maioria das classes (a-p), o modelo classificou de forma correta todas as instâncias, exceto para as classes 14 (n) e 15 (o), onde houve 4 e 5 classificações incorretas, respectivamente. As métricas de taxa de verdadeiros positivos, taxa para falso positivos, precisão recall e F1-score são todas iguais a 1 para todas as classes (a-p), indicando que o Randomtree classificou corretamente todas as instâncias em cada classe. A área sob a curva ROC também obteve resultado igual a 1 para todas as classes, indicando um desempenho perfeito do modelo na classificação.

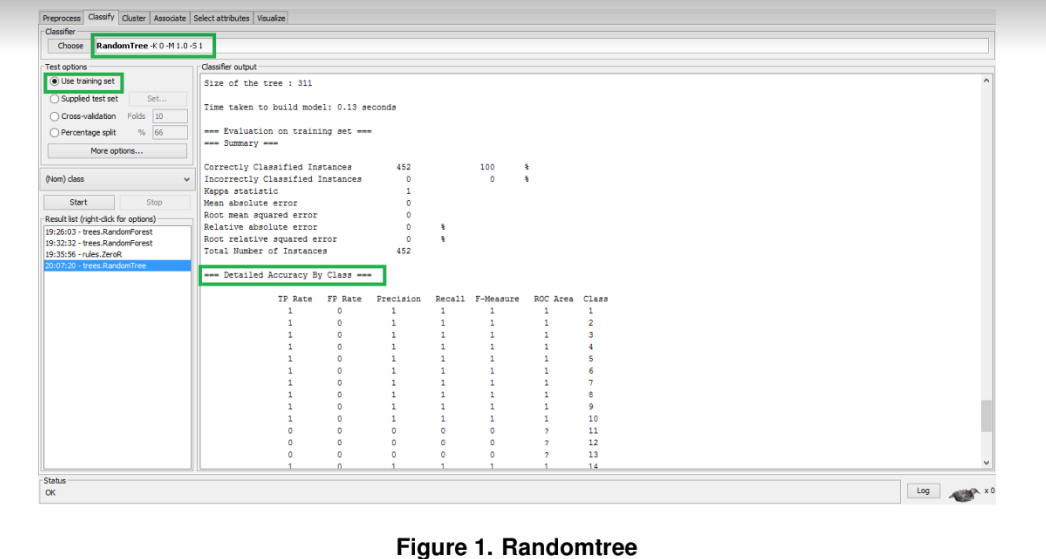


Figura 1. Randomtree

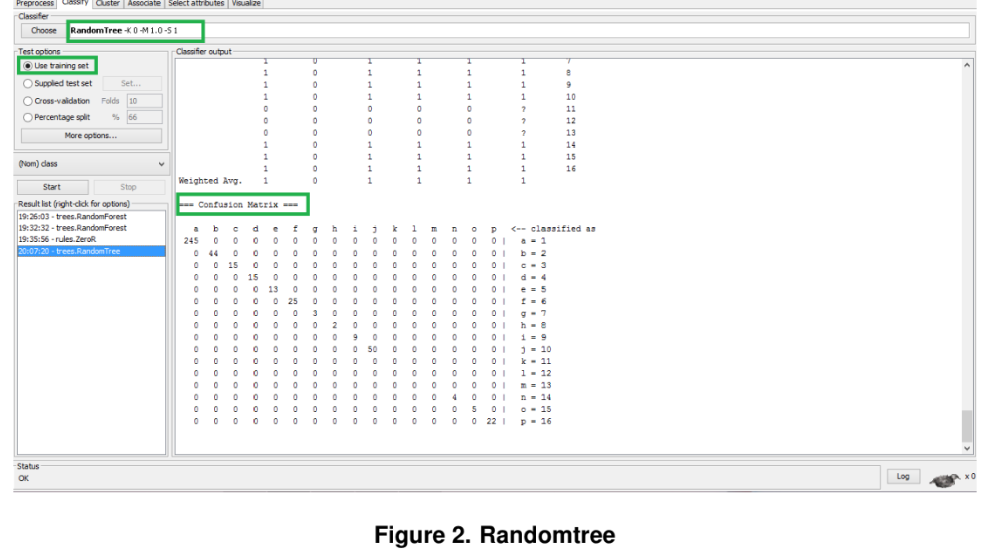
****

Figura 2. Randomtree

No modelo (Zeror), é um classificador simples que prevê a classe majoritária de todo o conjunto de dados. Nesse caso, o Zeror está prevendo a classe 1 para todos os casos, a taxa de classificação correta foi de 54,2035%, o que significa que o modelo que classificou corretamente 245 instâncias do total de 452. A taxa de classificação incorreta foi 45,7965%, indicando classificação incorreta de 207 instâncias.

O coeficiente Kappa foi calculado como 0, o que sugere que a concordância das classificações do modelo e as classificações reais ocorreram por acaso. O erro absoluto médio foi calculado como 0.855% e o erro quadrático médio foi 0.2055%. Esses valores indicam a diferença média entre as classificações do modelo e as classificações reais. A matriz de confusão mostra que todas as instâncias foram classificadas como classe 1 (a), e nenhuma instância foi classificada como qualquer outra classe, isso indica que o modelo não teve um desempenho muito bom em relação ao Randomtree.

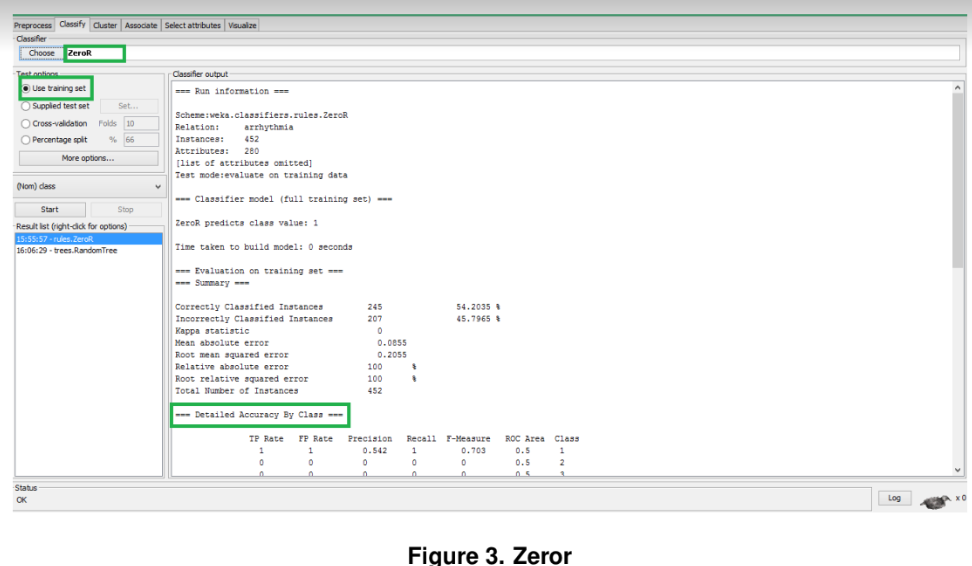


Figura 3. Zeror

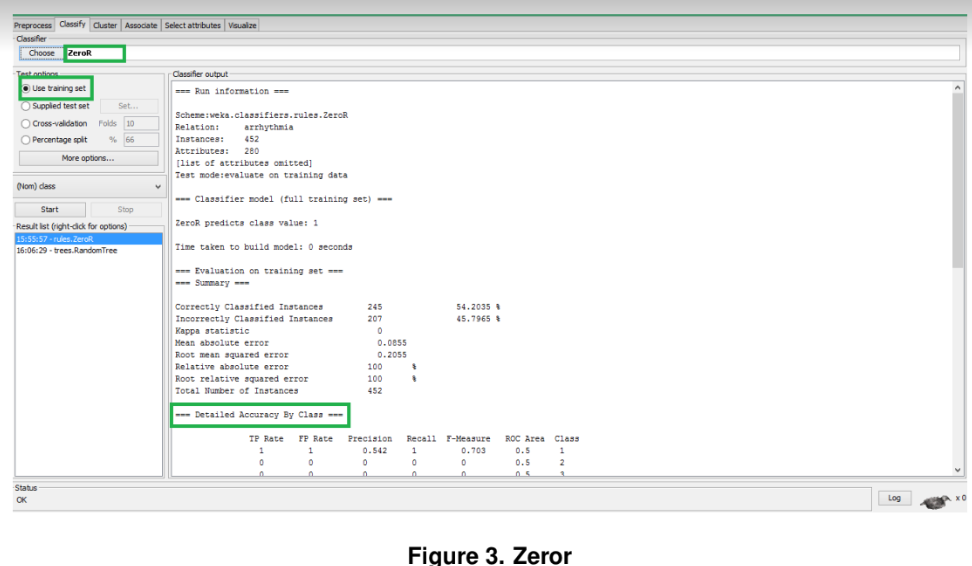


Figura 4. Zeror

Já o modelo (Oner) tem como base de dados 280 atributos (características) que foram empregados para treinar o classificador. O classificador classificou corretamente 277 instâncias (61,28%) e incorretamente 175 instâncias (38,72%) do conjunto de dados do treinamento, a matriz de confusão mostra como as instâncias de cada classe foram classificadas pelo modelo. No caso apresentado, a classe (a) rotulada com 1 teve 242 instâncias corretamente classificadas, enquanto as classes (b) a (p) rotuladas com 2 a 16 tiveram poucas ou nenhuma instância corretamente classificada, algumas medidas de avaliação são fornecidas, como a taxa de verdadeiros positivos (FP Rate), precisão, recall, F-measure e área sob a curva ROC.

Essas medidas são apresentadas para cada classe individualmente, bem como a média ponderada (Weighted Avg.) para todas as classes, é importante observar que algumas classes têm uma quantidade muito baixa de instâncias corretamente classificadas (0), o que indica que o modelo pode não ser capaz de distinguir essas classes com base nos atributos fornecidos, os valores “?” aparecem em algumas células da matriz de confusão e das medidas de avaliação. Isso mostra que não há informações disponíveis para essas classes especificas ou que as instâncias dessas classes não foram corretamente classificadas.

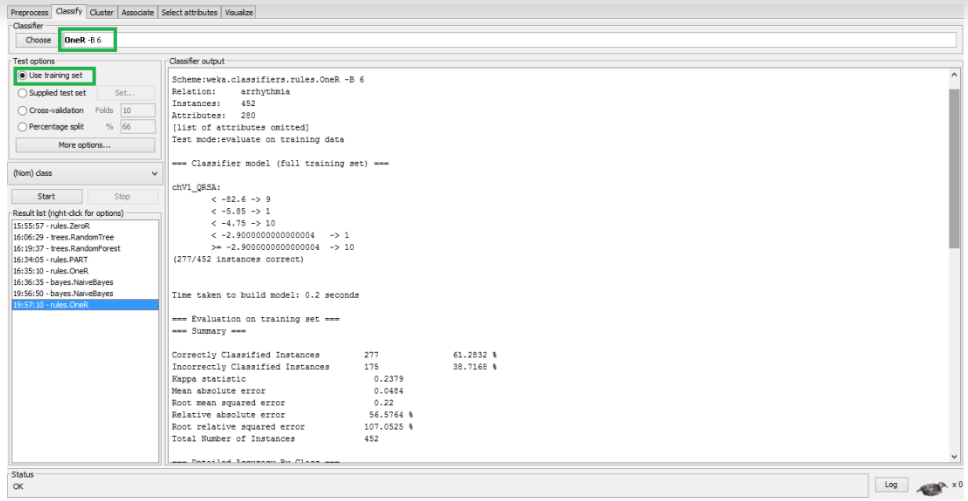


Figura 5. Oner

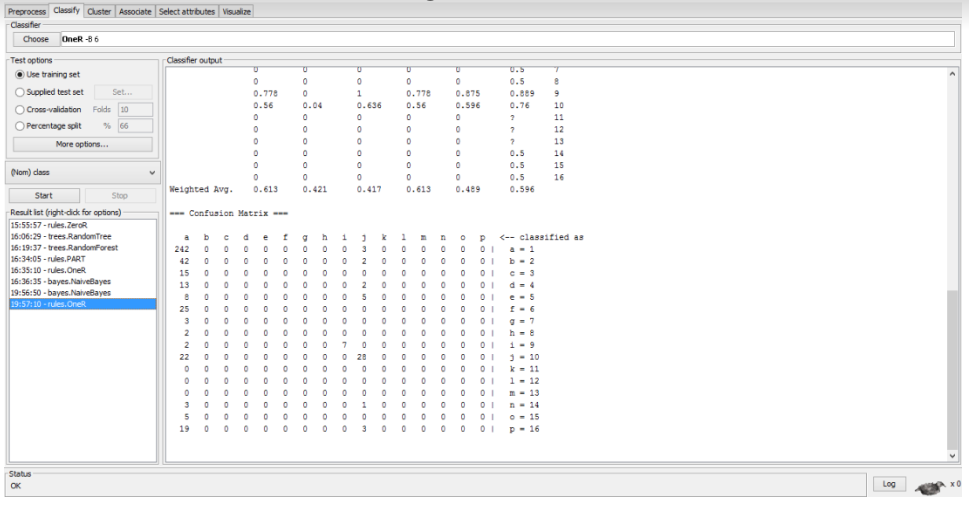


Figura 6. Oner

O modelo de classificação (RandomForest) foi construído usando 100 árvores, sendo cada árvore construída considerando 9 atributos aleatórios, a taxa de erro fora do conjunto é de 0.3319%, o que significa que, em média, aproximadamente 33,19% das instâncias foram classificadas incorretamente pelo modelo, na avaliação do conjunto de treinamento, todas as instâncias foram classificadas corretamente, resultando em uma taxa de classificação correta de 100%, não houve instâncias classificadas incorretamente. O coeficiente Kappa é 1, indicando uma concordância perfeita entre as classificações previstas pelo modelo e as classificações reais, os valores de erro absoluto médio e erro quadrático médio são baixos, indicando que o modelo tem um bom desempenho em termos de precisão, a matriz de confusão mostra a contagem de instâncias classificadas para cada classe. Observa-se que a maioria das instâncias foram corretamente classificadas para todas as classes, exceto para as classes 14 (n) e 15 (o), onde houve algumas instâncias classificadas incorretamente.

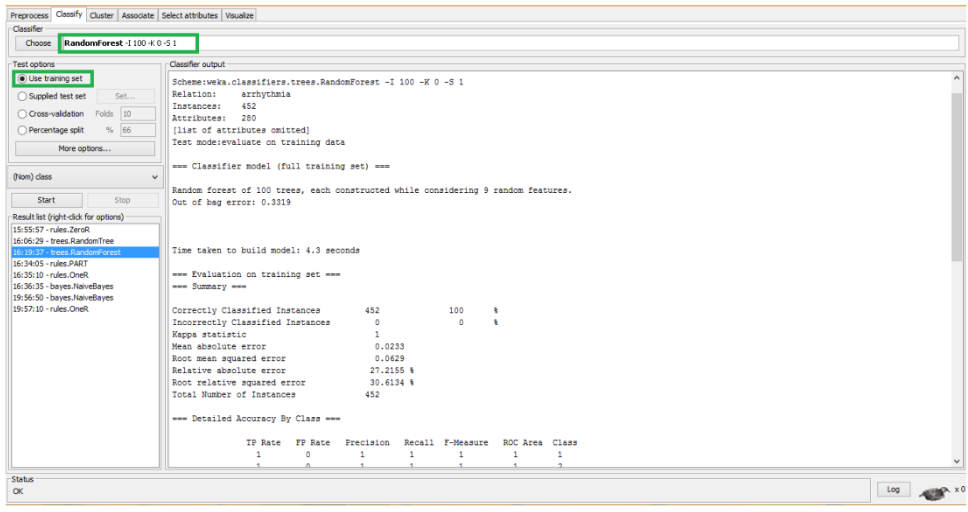


Figura 7. Randomforest

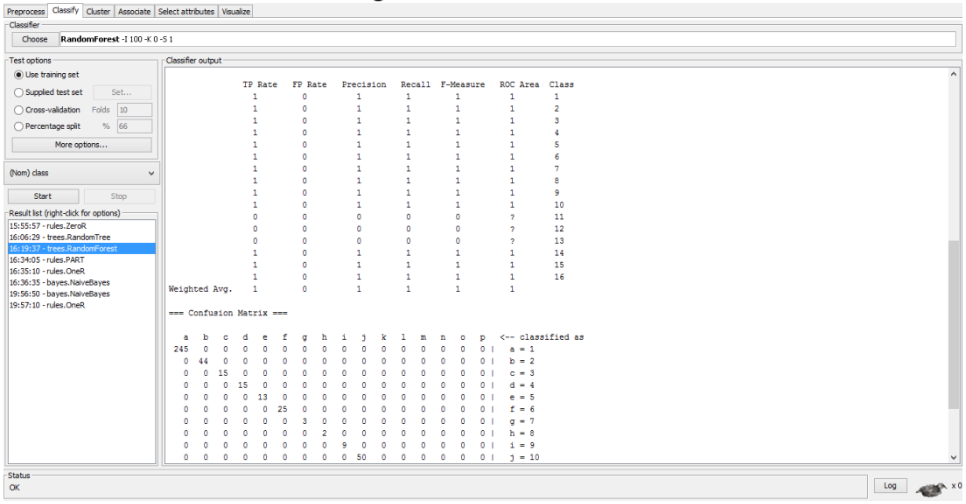


Figura 8. Randomforest

O modelo J48 obteve uma taxa de classificação correta de 92,26%, acertando 417 das 452 instâncias, enquanto errou em 35 casos, resultando em uma taxa de classificação incorreta de 7,74%. O coeficiente Kappa foi de 0,8836, indicando uma forte concordância entre as previsões do modelo e as classes reais, o que sugere um desempenho muito superior ao acaso. O erro absoluto médio foi de 0,0138, refletindo uma pequena discrepância entre as previsões e as classes reais, e o erro quadrático médio foi de 0,083, mostrando a magnitude média dos erros quadráticos. O erro relativo absoluto foi de 16,15%, enquanto o erro quadrático relativo foi de 40,37%, oferecendo uma visão sobre a precisão e a variação dos erros em relação às classes reais.

A matriz de confusão revela que o modelo teve um desempenho excepcional em várias classes, como 1, 2 e 3, com taxas de acerto muito altas, mas apresentou dificuldades com algumas classes específicas, como 7 e 8, que tiveram taxas de acerto mais baixas. Apesar dessas dificuldades, o modelo J48 demonstrou uma performance geral muito boa, com alta taxa de acerto e boa concordância com as classes reais, embora ainda haja espaço para melhorias na classificação de certas classes.

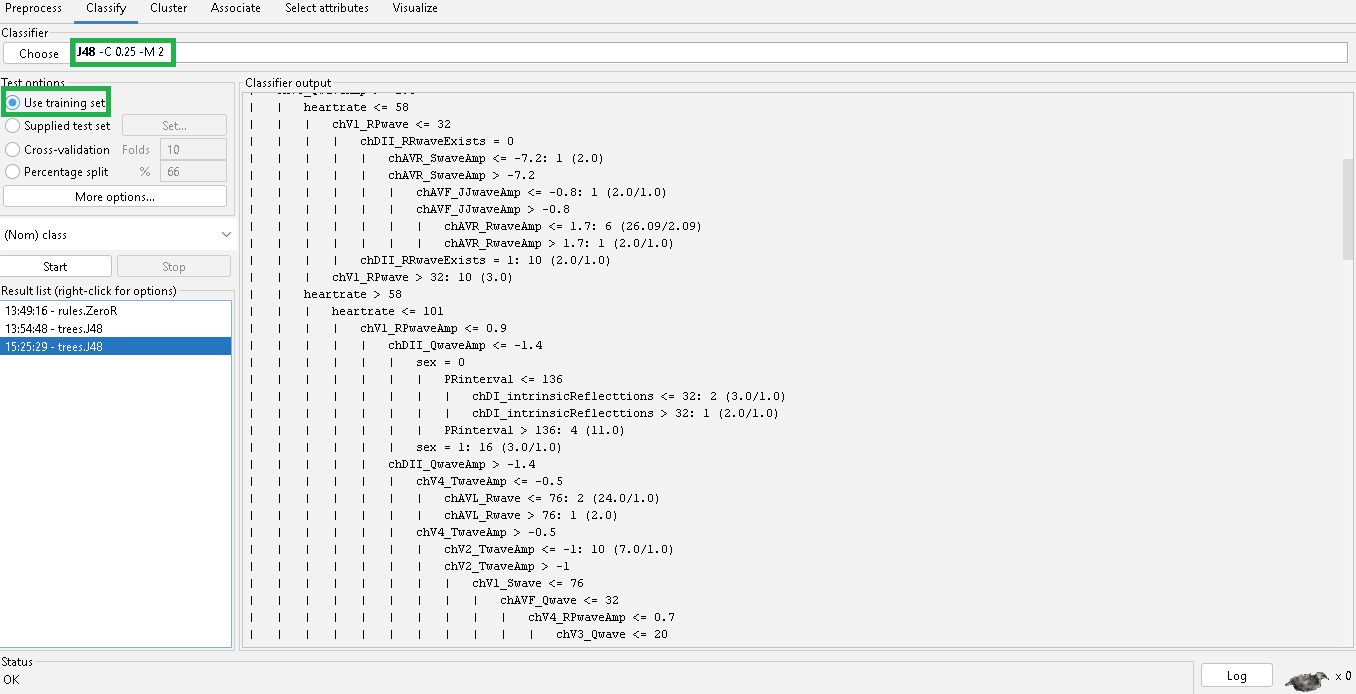


Figura 9. J48

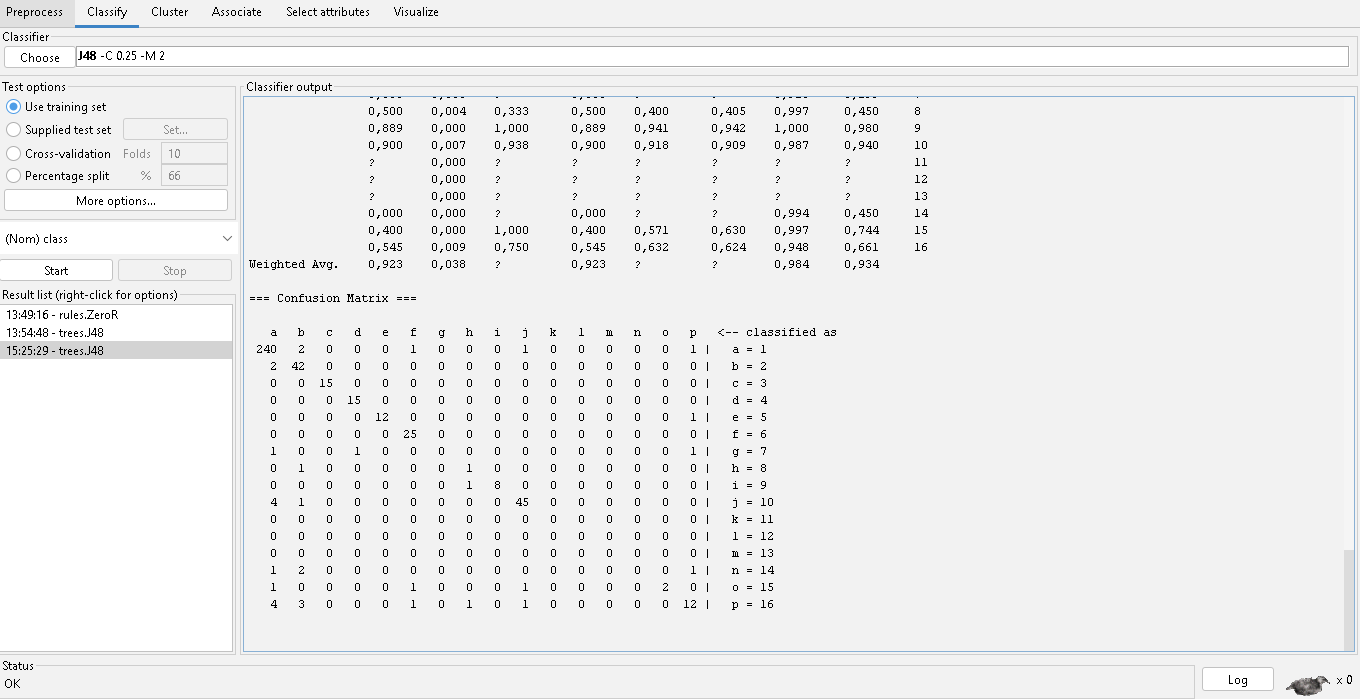


Figura 10. J48

**5 CONSIDERAÇÕES FINAIS**

Na avaliação dos modelos de classificação, os resultados revelam diferenças em desempenho. O modelo RandomForest destacou-se por sua taxa de classificação correta de 100% no conjunto de treinamento e coeficiente Kappa perfeito, indicando um potencial elevado para desempenho em ambientes práticos. Mas, é crucial que sua capacidade de generalização para novos dados seja verificada, dado que a taxa de erro fora do conjunto de treinamento foi de 33,19%. O J48 também demonstrou um desempenho muito bom, com uma taxa de classificação correta de 92,26% e um coeficiente Kappa de 0,8836, refletindo uma forte concordância entre as previsões e as classes reais. Apesar do modelo ter mostrado alguma dificuldade em classificar corretamente certas classes específicas, sua precisão geral e consistência torna-o uma escolha sólida para aplicações práticas, principalmente se ajustes forem feitos para melhorar a classificação das classes com taxa de erro.

Já os modelos Zeror e Oner apresentaram desempenhos baixos em relação aos outros. O Zeror, ao prever apenas a classe majoritária, demonstrou uma taxa de classificação correta de 54,2% e um coeficiente Kappa de 0, sugerindo um desempenho aleatório e inadequado entre classes. O Oner, apesar de ter uma taxa de acerto alta para a classe principal, falhou em distinguir corretamente entre outras classes, assim com um desempenho limitado e não confiável para classificações detalhadas.

O Randomtree, mostrou uma taxa de acerto de 100% e métricas perfeitas no treinamento, apresentou dificuldades com algumas classes. Apesar de seu desempenho, a eficácia prática do modelo pode ser comprometida se as classes com taxa de erro forem relevantes.

Em conclusão, para aplicações práticas no contexto da saúde, RandomForest e J48 são os modelos mais promissores, oferecendo alta precisão e boa capacidade de generalização. Para trabalhos futuros, uma validação adicional em dados novos e potenciais ajustes para melhorar a performance em classes específicas, garantindo assim que o modelo escolhido atenda efetivamente às necessidades dos profissionais de saúde.

**REFERÊNCIAS**

BARBOSA, P. A. M. *et al*. O papel da inteligência artificial na predição de doenças cardiovasculares. **Revista Corpus Hippocraticum**, v. 1, n. 1, 2024. Disponível em: https://revistas.unilago.edu.br/index.php/revista-medicina/article/view/1173

CAMPOS NETO, C. de M. **Análise inteligente de dados em um banco de dados de procedimentos em cardiologia intervencionista**. Repositório USP, 2016. Disponível em: https://www.teses.usp.br/teses/disponiveis/98/98131/tde-18102016-085650/publico/TeseCantidiodeMouraCamposNeto\_Corrigida.pdf

GUIMARÃES, T. B.; NEVES, N. S.; SOUZA, A. C. A. Diagnósticos de eletrocardiograma realizados pela inteligência artificial. **Revista Científica da UNIFENAS**, v. 6, n. 4, p. 44-48, 2024. Disponível em: https://revistas.unifenas.br/index.php/revistaunifenas/article/view/1080

SANTANA JÚNIOR, W. B. de. **Uso da inteligência artificial aplicada ao eletrocardiograma para diagnóstico de disfunção sistólica ventricular esquerda**. Repositório UFMG, 2023. Disponível em: https://repositorio.ufmg.br/handle/1843/61498