



## Uma ferramenta baseada em computação evolutiva para o apoio ao diagnóstico da cardiopatia isquêmica

A based on evolutionary computation tool to support the diagnosis of heart disease

Una herramienta basada en la computación evolutiva para apoyar el diagnóstico de la cardiopatía isquémica

Ubiratan Roberte C. Passos<sup>1</sup>, Ítalo de Oliveira Matias<sup>2</sup>, Ubiratan Rosa Passos<sup>3</sup>, Marlito Andrade Pereira<sup>1</sup>, Vinícius Machado Carvalho<sup>4</sup>

### RESUMO

**Descritores:** Algoritmos; Técnicas de apoio para a decisão; Doença das coronárias; Inteligência artificial

**Objetivo:** O trabalho em questão vem apresentar a proposta de uma ferramenta desenvolvida com técnicas de computação evolutiva e reconhecimento baseado em casos para auxílio no diagnóstico da cardiopatia isquêmica. **Método:** Análises bibliográficas e aplicação combinada das técnicas de Algoritmos Genéticos (AG's) e reconhecimento baseado em casos (RBC's) e derivações da função de distância euclidiana. **Resultado:** Os testes realizados na ferramenta mostraram que esta possui grande grau de acertos em suas indicações de diagnóstico, chegando a 97,01% de acertos nas etapas de treinamento com acurácia, especificidade e sensibilidades superiores a 92%. **Conclusão:** A escolha das tecnologias citadas bem como dos métodos aplicados para evolução do algoritmo dentro das técnicas de AG, proporcionaram a ferramenta grande capacidade de decisão levando-a grandes taxas de acertos em sua indicação diagnóstica.

### ABSTRACT

**Keywords:** Algorithms; Decision support techniques; Coronary disease; Artificial intelligence

**Objective:** The work in question is presenting the proposal of a tool developed with evolutionary computation techniques and recognition based on cases to aid in the diagnosis of ischemic heart disease. **Method:** literature analysis and the combined use of techniques of Genetic Algorithms (GAs) and recognition based on cases (RBC's) and derivations of the function of Euclidean distance. **Result:** The tests performed on the tool showed that it has high degree of accuracy for their diagnostic indications, reaching 97.01 % accuracy in the training stages with accuracy, specificity and sensitivity higher than 92 %. **Conclusion:** The choice of technologies cited as well as the methods used for the evolution of the algorithm within the AG techniques, provided a great capacity for decision-taking tool to major hit ratios in the diagnostic statement.

### RESUMEN

**Descriptores:** Algoritmos; Técnicas de apoyo para la decisión; Enfermedad coronaria; Inteligencia artificial

**Objetivo:** La obra en cuestión es la presentación de la propuesta de una herramienta desarrollada con técnicas de computación evolutiva y el reconocimiento basado en casos para ayudar en el diagnóstico de la cardiopatía isquémica. **Método:** análisis de la literatura y el uso combinado de técnicas de algoritmos genéticos (GAs) y reconocimiento basado en casos (RBC) y derivaciones de la función de distancia euclídea. **Resultado:** Las pruebas realizadas en la herramienta demostró que tiene alto grado de respuestas correctas en sus indicaciones diagnósticas, alcanzando una precisión 97,01 % en las etapas de formación con precisión, especificidad y sensibilidad superior al 92 %. **Conclusión:** La elección de las tecnologías citadas, así como los métodos utilizados para la evolución del algoritmo dentro de las técnicas de AG, a condición de una gran capacidad para la toma de decisiones herramienta para los cocientes de ataque en el estado de diagnóstico.

<sup>1</sup> Mestrando da Universidade Cândido Mendes - UCAM, Campos dos Goytacazes (RJ), Brasil.

<sup>2</sup> Professor da Universidade Cândido Mendes - UCAM, Campos dos Goytacazes (RJ), Brasil.

<sup>3</sup> Cardiologista Especialista pela Sociedade Brasileira de Cardiologia, Brasil.

<sup>4</sup> Bacharelado pela UNES – Faculdade do Espírito Santo, Cachoeiro de Itapemirim (ES), Brasil.

## INTRODUÇÃO

O processo de diagnóstico de uma doença é além de completo de extrema importância, pois os problemas resultantes de falhas nesta atividade podem ser gravíssimos. No entanto é impossível afirmar que todo diagnóstico é realizado sem que tais falhas ocorram<sup>(1-4)</sup>. Um agravante para isto é o fato de, em muitas localidades ser comum a ausência de médicos de todas as especialidades, e ainda que muito capacitado, um médico não é capaz de atender todas as especialidades. Assim sendo, o fornecimento de meios que reduzam a complexidade de processo de diagnóstico torna-se vantajoso, não somente para a classe médica, mas também a sociedade em geral<sup>(5)</sup>.

É comum ao se buscar na literatura, encontrar trabalhos que apresentam os mais variados métodos relacionados às técnicas de Inteligência Artificial (IA) aplicados à área da saúde<sup>(1,5-7)</sup>. Neste trabalho será apresentado alguns dos vários conceitos de IA aplicados, sobretudo ao auxílio no diagnóstico médico, além de questões relacionadas às técnicas de Reconhecimento Baseado em Casos com utilização de Algoritmos Genéticos, bem como a análise de uma ferramenta desenvolvida com estas técnicas para este propósito.

O principal objetivo deste trabalho é fornecer aos profissionais médicos um Sistema de Suporte à Decisão (SSD), também conhecidos como Sistemas Especialistas (SE) capazes de reduzir a complexidade do processo de diagnóstico da cardiopatia isquêmica. Os resultados obtidos ao final deste trabalho reforçam os argumentos de Almeida et al<sup>(7)</sup>, o qual afirma que com o auxílio deste grupo de ferramentas espera-se como resultado diagnósticos mais ágeis e precisos. Com análise dos resultados dos testes realizados utilizando esta ferramenta, pode-se notar uma sensível melhora dos resultados de processos de diagnósticos com base em dados clínicos que podem ser auxiliados aumentando sua taxa de acertos quando acompanhados de resultados de alguns exames complementares, como se poderá observar nos testes de sensibilidade, especificidade, curvas de ROC e evolução do algoritmo.

### Doenças Crônicas não Transmissíveis e as Doenças Cardiovasculares

No Brasil, aproximadamente 72% do total de óbitos são atribuídos a doenças crônicas não transmissíveis (DCNT), das quais pode-se citar: doenças cardiovasculares, doenças respiratórias crônicas, diabetes, câncer, doenças renais entre outras destes 33% são causados por doenças cardiovasculares (DCV) 10, as quais na maioria dos casos observados havia presença de aterosclerose, ou seja, o acúmulo de placas de gorduras nas artérias ao longo dos anos, o que impede a passagem do sangue<sup>(5)</sup>.

### Doença Obstrutiva Coronariana

Segundo informações do Boletim Brasileiro de Avaliação de Tecnologias em Saúde<sup>(8)</sup>, “a doença arterial coronariana é resultante do processo de estreitamento ou

oclusão das artérias coronarianas por aterosclerose, uma doença que afeta o revestimento endotelial das grandes e médias artérias”. Neste tipo de enfermidade em 90% dos casos a obstrução da coronariana é causada pela formação de placa aterosclerótica. Estas placas podem produzir uma grave redução do lume da artéria coronariana numa proporção superior a 50% do diâmetro desta (ou superior a 75% da área transversal do lume), o que pode levar a uma importante redução do fluxo sanguíneo causando assim um desequilíbrio na distribuição da oferta e consumo de oxigênio (isquemia miocárdica). Com o passar do tempo a má distribuição do oxigênio para os tecidos pode causar lesões irreversíveis, trata-se de uma enfermidade que pode evoluir de forma assintomática, insidiosa, durante décadas, vindo a se manifestar clinicamente das mais diversas formas<sup>(8)</sup>.

### Inteligência Artificial

As técnicas de Inteligência Artificial (IA) ou Inteligência Computacional (IC) são originadas da junção das mais diversas áreas de conhecimento, que permitem assim a criação de tecnologias capazes de simular a capacidade de pensar e tomar decisões de forma semelhante às realizadas pelos seres humanos<sup>(9)</sup>.

A palavra inteligência é proveniente do latim *inter* (entre) e *legere* (escolher), ou seja, escolher entre uma coisa e outra, sendo a inteligência uma forma de resolver esta questão. Logo o mesmo considera inteligência artificial uma forma de inteligência produzida pelo homem para beneficiar as máquinas, permitindo assim que estas possam de alguma forma simular algumas das habilidades de escolhas humanas<sup>(7)</sup>.

Devido a suas possibilidades tem sido cada vez mais comuns pesquisas e desenvolvimento de sistemas que fazem utilização das tecnologias relacionadas a IA na área de saúde problemas<sup>(10-11)</sup>. O trabalho em questão sugere a utilização do aprendizado de máquina (AM), através das técnicas de Computação Evolutiva (CE) e Algoritmos Genéticos (AG) associados a Raciocínio Baseado em Casos (RBC), conforme será apresentados a seguir.

### Aprendizado de Máquina

Para alguns autores aprendizado de máquina nada mais é do que o aprendizado por experiência, o que significa que, a medida em que uma tarefa é executada, a máquina aprende a melhor forma de resolver o problema, além de estruturar o conhecimento existente para levar a um entendimento do aprendizado<sup>(12)</sup>.

As técnicas de aprendizado de máquina empregam princípios de inferência denominados indução, no qual é possível obter conclusões genéricas com base em um conjunto particular de exemplos. Tais técnicas de aprendizado indutivo são divididas em principalmente em: **aprendizado supervisionado** (aplicado neste trabalho) e **aprendizado não supervisionado**<sup>(4,13)</sup>.

**Aprendizado supervisionado** é a técnica de aprendizado onde é fornecida ao sistema uma referência do resultado a ser alcançado, o sistema é treinado com base no conhecimento prévio do ambiente. O treinamento é realizado com base em um conjunto de

exemplos de entradas e uma saída esperada. Desta forma, espera-se que, as representações geradas sejam capazes de produzir saídas corretas para as novas entradas até então desconhecidas pelo sistema.

**Aprendizado não supervisionado** consiste no aprendizado sem a utilização de referências, ou seja, não ocorre um treinamento com conhecimento prévio do ambiente. “O algoritmo de aprendizado de máquina não supervisionado aprende a representar (ou agrupar) as entradas submetidas, segundo medidas de similaridade”<sup>(13)</sup>.

**Algoritmos Genéticos**

AG's consistem basicamente de estratégias de buscas onde estados são gerados a partir de estados anteriores provenientes de uma geração automática e aleatória, as quais, após uma série de “interferências genéticas” são capazes de encontrar soluções consideradas satisfatórias para os mais variados problemas<sup>(11-12)</sup>. Coello<sup>(14)</sup> propõe uma sequencia considerada suficiente para o bom entendimento dos processos realizados por AG's, onde dado um problema tem-se nas populações iniciais e correntes as soluções candidatas as quais, após serem submetidas a um critério de aptidão são passadas por um critério de seleção, sendo os mais comuns **roleta** e **elitismo** e posteriormente por a uma série de operadores genéticos, tais como **cruzamento** e **mutação**, dando início assim a uma nova população de possíveis soluções repetindo este procedimento até que um determinado critério de parada seja atendido.

**Roleta**, cada indivíduo da geração é escolhido para fazer parte da próxima geração através de um sorteio de roleta, conforme é ilustrado na Figura 1.



Figura 1 - Método da Roleta. Adaptada de Lopes<sup>(15)</sup>

Cada indivíduo da população é representado na roleta de forma proporcional ao seu índice de aptidão, desta forma indivíduos com alta aptidão ocupam uma maior parcela da roleta, enquanto os indivíduos com

menor aptidão ocupam uma fatia menor da roleta. Sendo assim, quando a roleta é girada um determinado número de vezes, a depender do tamanho da população e vão para a próxima geração os que foram sorteados na roleta.

**Elitismo** é o método de seleção onde, o indivíduo com melhor aptidão é preservado para a próxima geração. Desta forma procura-se garantir que as melhores soluções não se percam durante o processo.

**Cruzamento** realiza a recombinação das características dos indivíduos pais durante o processo de reprodução, com isso as próximas gerações podem herdar estas características. A Figura 2 ilustra o cruzamento de um ponto. Os tipos mais comuns de cruzamento são, **um ponto**, **multipontos** e **uniforme**.

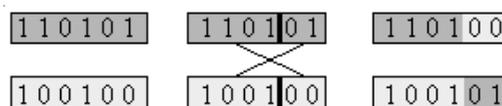


Figura 2 - Cruzamento de um ponto. Adaptada de Lopes<sup>(15)</sup>

**Cruzamentos de um ponto** consistem na escolha de um ponto a partir do qual as informações genéticas dos pais serão trocadas, assim sendo, as informações anteriores de um dos pais são ligadas as informações posteriores dos outro.

**Cruzamento multipontos** é basicamente uma generalização da do modelo anterior, onde a troca de informações genéticas ocorre não somente em um, mas vários pontos.

**Cruzamento uniforme** que no lugar de pontos de cruzamento, determina através de algum parâmetro global a probabilidade de cada variável ser trocada entre os pais.

**Mutação**, geralmente aplicada a uma taxa de probabilidade de mutação  $P_m$  muito baixa, consiste em trocar os valores de alguns genéticos do indivíduo, representada na Figura 3, permitindo desta forma uma leve alteração no mecanismo de busca e a diversidade genética dos indivíduos.

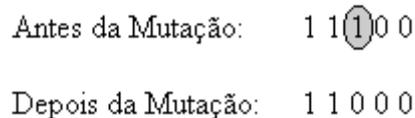


Figura 03 - Mutação. Adaptada de Lopes<sup>(15)</sup>

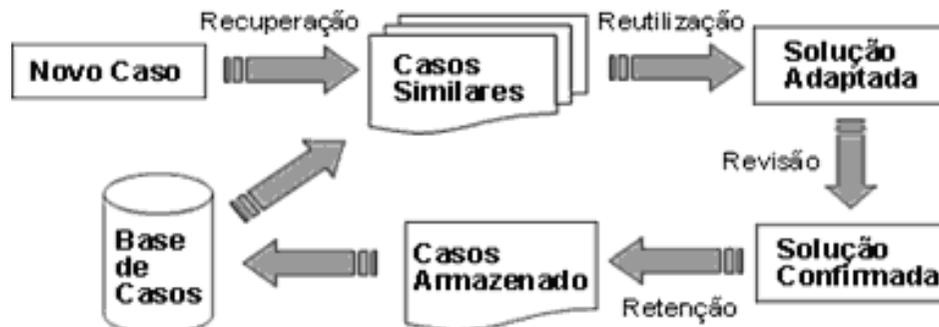


Figura 4 - Ciclo do RBC. Adaptada de Wangenheim, Wangenheim<sup>(5)</sup>

### Raciocínio Baseado em Casos - RBC

O processo de tomada de decisão realizada pelo cérebro é frequentemente baseado em conhecimentos passados, adquiridos sobre determinada situação, desta forma quando exposto a uma nova situação este procura em sua memória a solução mais adequada para o novo problema, caso este ainda não esteja armazenado, definida a nova solução esta é armazenada para futuras consultas. Semelhantemente agem os sistemas de Raciocínio Baseado em Casos (RBC)<sup>5</sup>. A Figura 4, apresenta uma proposta capaz de facilitar a compreensão do ciclo do RBC.

Outros autores destacam a capacidade de armazenar e poder utilizar conhecimentos previamente adquiridos como a principal vantagem dos sistemas RBC frente a outras tecnologias que fazem uso de regras estruturadas, quando expostas a problemas não estruturados, como àquelas que utilizam lógica fuzzy<sup>(16)</sup>.

### FERRAMENTA PROPOSTA

Este trabalho propõe apresentar uma ferramenta desenvolvida com base nas tecnologias acima citadas capaz de auxiliar no diagnóstico da cardiopatia isquêmica e outras doença, desde que previamente treinadas<sup>(17)</sup>. Ainda que seja possível encontrar ferramentas semelhantes na literatura a proposta de utilização de AG's combinada a RCB deu-se pela grande capacidade adaptativa destas tecnologias<sup>(8)</sup>, e como forma de contribuir com novas possibilidades para a comunidade científica em geral.

Cabe então ao AG aqui desenvolvido encontrar as melhores combinações para a sequencia de pesos atribuídos a cada atributo apresentado ao sistema, denotado por  $W_i$ , buscando a menor distância possível entre os casos armazenados e os casos apresentados através da função *fitness* ou aptidão.

Desta forma, ao realizar a junção das técnicas de AG com RBC, tem-se que, sempre que for necessário encontrar uma solução adequada para um novo problema, o AG será ativado para que encontre a melhor adaptação para aproximação dos casos registrados com o novo caso existente, vide Figura 5. Contudo, uma vez treinado as combinações dos pesos para cada atributo são registrados no sistema e um novo treinamento faz-se necessário somente para adaptação da ferramenta para enfermidades

que a mesma ainda não contemple.

Para cada ciclo do treinamento várias combinações de pesos  $W$  são testadas, procurando assim a melhor adequação da função para o problema em questão. Ao encontrar os valores mais adequados para melhor aptidão da função estes são armazenados pelo sistema para futura utilização. O AG desenvolvido para a solução proposta implementa como método de seleção de indivíduos (soluções) mais aptas, as funções de **roleta** e **elitismo**, ambas adaptadas aumentando assim sua capacidade de encontrar máximos globais. Para operadores genéticos, que buscam realizar alterações genéticas nos indivíduos buscando melhorar suas aptidões (adaptatividade) para o problema em questão, foram utilizados o **crucamento** e a **mutação**. As adaptações realizadas nos métodos buscam resolver alguns dos problemas apontados por Lopes<sup>15</sup>, os demais métodos foram desenvolvidos segundo as práticas observadas em outros estudos<sup>(18)</sup>.

Para seu treinamento a ferramenta proposta conta com uma base de dados contendo 1034 registros de prontuários médicos gentilmente cedida por Almeida et al<sup>(8)</sup>, que servirá de base de conhecimento para o sistema, fornecendo ao mesmo todo conhecimento necessário para que este possa calibrar os pesos da equação tornando-se assim apto a realizar seus processos de tomadas de decisões.

O propósito geral da ferramenta é, não somente fornecer apoio ao processo de diagnóstico da cardiopatia isquêmica, mas também servir de ferramenta de suporte operacional para atividades ambulatoriais, mantendo cadastros e registros de pacientes, gerando históricos dos mesmos, permitindo desta forma que a ferramenta acumule em sua base conhecimentos suficientes para que seja posteriormente submetida a treinamentos que lhe permita diagnosticar diversas outras enfermidades.

### MATERIAIS E MÉTODOS

O trabalho em questão foi confeccionado a partir de uma série de estudos bibliográficos que confirmam o grande interesse da comunidade pesquisadora sobre o tema aqui abordado, como pode ser observado em outros trabalhos<sup>(2)</sup>.

A ferramenta proposta foi criada com base nas técnicas de AG's desenvolvidos com as seguintes características:

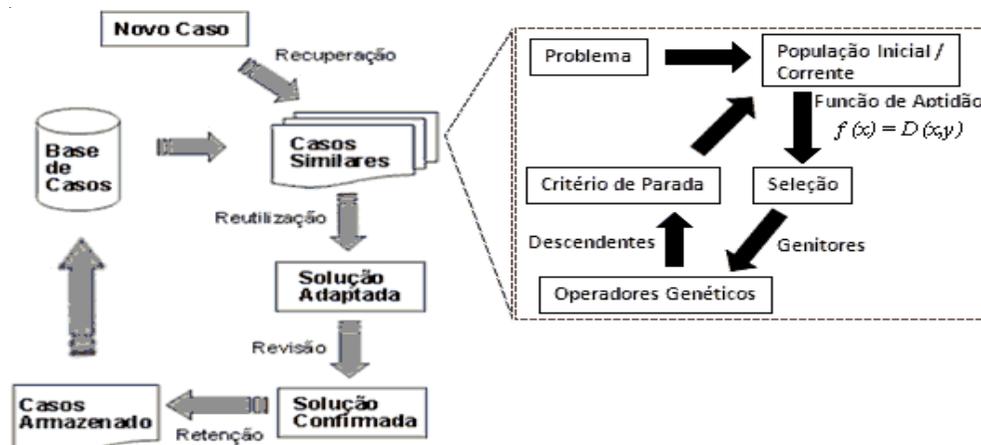


Figura 5 - Relação RBC x AG adaptada de Wangenheim, Wangenheim e Coello<sup>(5,14)</sup>.

operador de seleção do tipo **roleta** a qual foi adaptada realizando ordenação por ordem crescente de aptidão dos indivíduos, amenizando desta forma um dos problemas descritos por Lopes (2006)<sup>(15)</sup> que levavam frequentemente ao encontro de máximos locais, foi aplicado também o **elitismo**, com uma pequena alteração permitindo assim definir a taxa de elitismo, garantindo desta forma que um número maior dos mais aptos indivíduos fossem preservados nas próximas gerações. Foram aplicados também os operadores genéticos **cruzamento multiponto** e a **mutação**, ambos sem qualquer alteração e aplicando as taxas de cruzamento de 75% e mutação de 3%, garantindo assim grande diversidade genética mas também preservação das melhores características dos indivíduos, seguindo as práticas observadas na literatura<sup>(18)</sup>.

A função de aptidão, equação 01, utilizada pelo algoritmo para medir a semelhança local e global entre os casos testados é uma adaptação da função de cálculo de distância conhecida como distância euclidiana quadrática e distância euclidiana quadrática ponderada<sup>(19)</sup>.

Quando menor a distância entre os casos, maior a similaridade entre os mesmos. Devido as grandes

#### Equação 01

$$\min Z = D(x,y)$$

Sendo

$$D(x,y) = \sqrt{\frac{\sum_{i=1}^n w_i (x_i - y_i)^2}{n}}$$

Sujeito às restrições

$$0 \leq \sqrt{\sum_{i=1}^n w_i (x_i - y_i)^2} \leq 1$$

$$0 \leq W_i \leq 1$$

$$y_i, x_i \geq 0 \quad i = \{i \in \mathbb{R}^+\}$$

semelhanças observadas entre as diversas enfermidades cardíacas a taxa mínima de aptidão definida para que um caso seja levado ao teste de similaridade para resolução do problema é de 85%, o que garante que somente aqueles casos que realmente apresentam bom grau de solução para o problema sejam testados como soluções aptas<sup>(9)</sup>.

O treinamento e teste realizados levaram em consideração a base de dados contendo prontuários médicos de vários pacientes, anteriormente citados, tendo sido realizados comparações registros por registros, ou seja, para cada registro do banco de dados foi realizado comparação com todos os 534 registros selecionados.

O algoritmo foi configurado para trabalhar com uma população de 30 indivíduos por geração, tendo de encontrar o melhor resultado possível em um total de 20 gerações, tendo sido observado estagnação do mesmo a partir deste total. O Gráfico 1 representa a evolução do algoritmo durante a etapa de treinamento.

Observando o gráfico de evolução do algoritmo nota-se as altas taxas de acerto que o mesmo obteve em suas indicações de diagnósticos, em alguns casos chegou-se a taxa de 99,9% de acertos, contudo para efeito de estudos e análises foi considerado a taxa média de acertos que

Onde

**x** = Valor do atributo do caso de controle

**y** = Valor do atributo do caso de teste

**w** = Peso atribuído ao atributo

**i** = Índice do atributo testado

**n** = Número de atributos testados

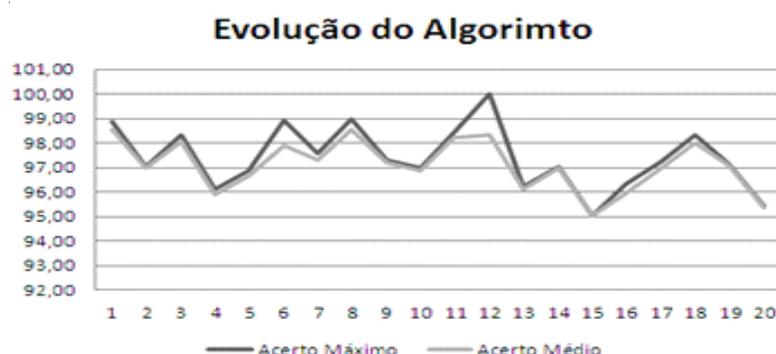


Gráfico 1- Evolução do algoritmo na fase de treinamento

nos testes em questão ficou em 97,01%. Entretanto, novas análises foram realizadas para levantamento das taxas de sensibilidade e especificidade do algoritmo, bem como sua acurácia, uma vez que, a alta taxa de acerto apresentada leva em consideração somente os casos levados a teste mas não casos que não apresentaram grau mínimo de

semelhança com os casos de controle.

## RESULTADOS

Uma vez calibrado o sistema foi submetido a novos testes de aptidão, nos quais foram utilizados 500 registros

da mesma base de dados utilizada para seu treinamento, entretanto os registros selecionados para os novos testes não participaram do processo de treinamento. Chama-se atenção para a importância de uma série de testes para validação de resultados de análises clínicas laboratoriais, mas que podem ser aplicadas a resultados obtidos por ferramentas de análises e suporte ao diagnóstico médico, como a proposta neste artigo<sup>(20)</sup>.

Os resultados obtidos pela ferramenta foram qualificados segundo os seguintes critérios: **sensibilidade**, **especificidade**, **prevalência** e **acurácia**. Sendo estes apresentados na Tabela 1.

Os testes realizados acima permitem: definir qual a capacidade da ferramenta de realmente apresentar resultados positivos para indivíduos que realmente apresentam a enfermidade, **sensibilidade**; através da **especificidade**, pode-se medir a capacidade de o diagnóstico indicar verdadeiro negativo, ou seja, de indicar um indivíduo sadio quando realmente sadio, neste caso um indivíduo que realmente não apresente a cardiopatia isquêmica; pela **prevalência** mediu-se a quantidade de casos na população testada e por fim **acurácia**, medindo o grau de veracidade do que se esta sendo testado, indicando o quanto algo se aproxima ou se afasta da “verdade”<sup>(20)</sup>.

Os resultados obtidos neste trabalho foram confrontados com outros resultados obtidos através de diferentes formas de estudos<sup>(21)</sup>, os quais cita-se: **Análise Linear de Fisher (Fisher)**, **Árvores de Decisão (AD)**, **Regressão Logística (RL)** e **Redes Neurais Artificiais (RNA)**, além do método aqui proposto, **Algoritmos Genéticos com Reconhecimento Baseado em Casos (AG+RBC)**, sendo os mesmos disponibilizados no Gráfico 2.

Pode-se notar pelo Gráfico 2 que o método aqui

proposto apresenta taxas de sensibilidade bem próximas às apresentadas pelo modelo linear de Fisher, que, dos métodos em questão foi o que apresentou a melhor taxa. Quanto a especificidade as taxas ficaram bem próximas dos modelos de Árvores de Decisão, Regressão Logística e Redes Neurais Artificiais.

Sobre os resultados obtidos por cada um dos métodos estudados e o proposto neste trabalho, foi aplicado o teorema da probabilidade total<sup>(22)</sup>, aqui adaptado para atender a necessidade em questão, conforme é apresentado na equação 02.

#### Equação 02

$$\Pr(A) = \Pr(A/Enf)P(Enf) + \Pr(A/Sad)Pr(Sad)$$

Onde:

Pr = Probabilidade Total

A = Verdadeiros Positivos

Enf = Indivíduos Enfermos

Sad = Indivíduos Sadios

Uma vez aplicado o teorema especificado na equação 02, foi possível definir os valores preditivos positivos – VP(+), valores preditivos negativos – VP(-), falsos valores preditivos positivos - FVP(+) e falsos valores preditivos negativos – FVP(-) utilizados para a construção das curvas de ROC<sup>(23)</sup> de cada um dos métodos, apresentadas na Figura 6.

Analisando os gráficos de curvas de ROC gerados por meio dos resultados obtidos por cada um dos métodos analisados segundo as afirmações de Vasconcelos<sup>(23)</sup>, nota-se que os resultados obtidos pela ferramenta proposta, apresenta sensível melhora devido as curvas representadas em seu gráfico, contudo isto não é aqui tomado como base para discutir a qualidade do algoritmo aqui proposto, mas sim para comprovação de seus resultados positivos.

Tabela 1 - Sensibilidade, especificidade, prevalência e acurácia

Legenda	Valores	Observações
Total de Simulações	500	
Indicações Positivas	323	
Verdadeiros Positivos (VP)	289	
Falsos Positivos (FP)	13	
Indicações Negativas	177	
Verdadeiros Negativos (VN)	162	
Falsos Negativos (FN)	15	
Sensibilidade	95,065%	VP / (VP + FN)
Especificidade	92,571%	VN / (FP + VN)
Prevalência	63,465	(VP + FN) / (VP + FN + VN + FP)
Acurácia	96,868	(VP + VN) / (VP + FN + VN + FP)

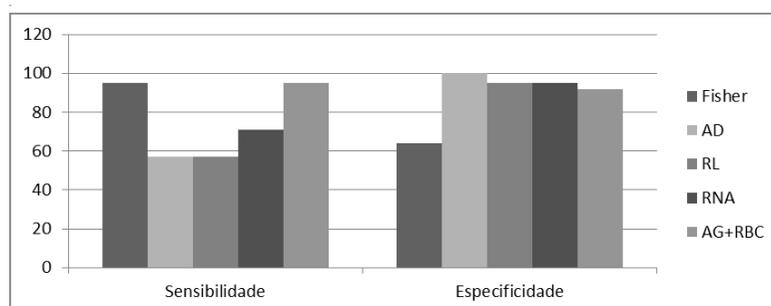


Gráfico 2 - Comparação de desempenho

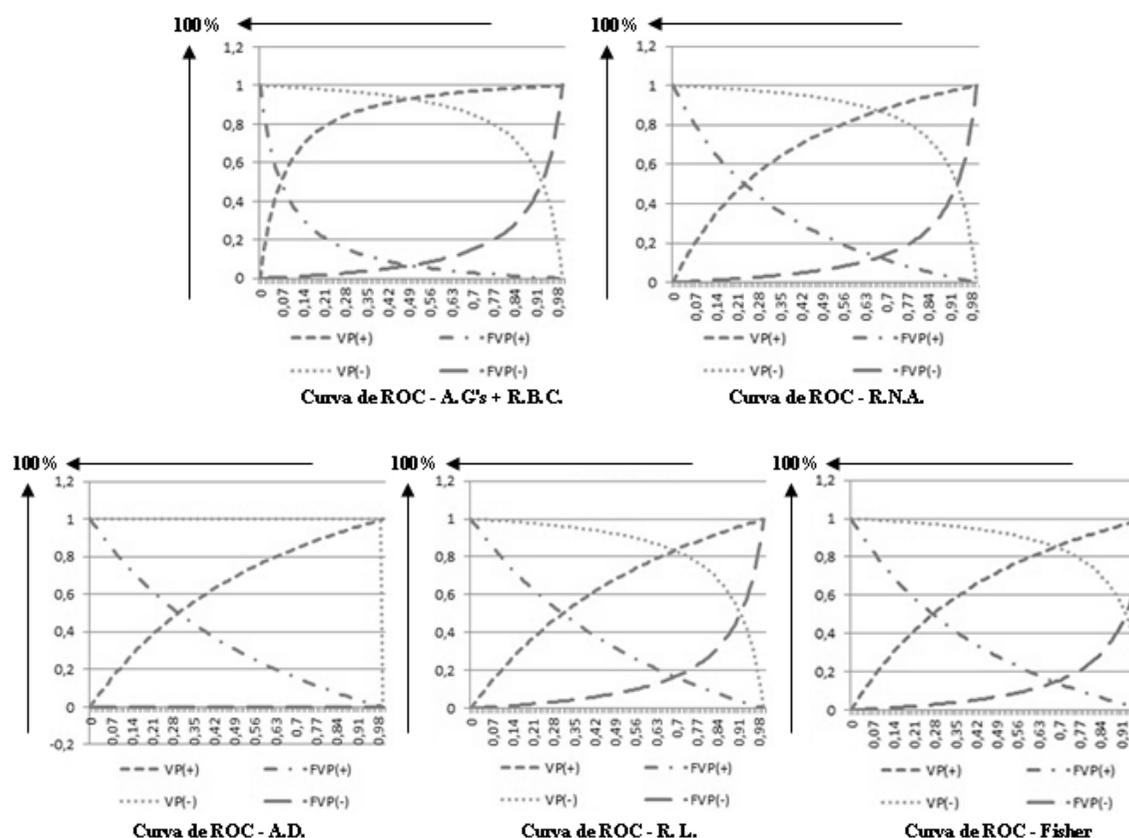


Figura 6 - Curvas de ROC, adaptadas de Vasconcelos<sup>(23)</sup>

## DISCUSSÃO

O principal objetivo do trabalho era apresentação de uma ferramenta de suporte a decisão específica para o diagnóstico da cardiopatia isquêmica, tal ferramenta apresenta-se como opção às soluções existentes, aplicando neste caso técnicas de algoritmos genéticos associados às técnicas de reconhecimento baseado em casos. Para calibragem do algoritmo foram considerados 17 atributos de cada indivíduo que foram analisados separadamente para medição de suas semelhanças e de forma agrupada para cálculo de sua similaridade global.

Os resultados obtidos e apresentados no trabalho mostram que a solução proposta mostra-se interessante para suporte ao profissional médico durante seu processo de consulta e sugestão diagnóstica, auxiliando-o na escolha ou encaminhamento do paciente para exames mais específicos, evitando desta forma encaminhamento do mesmo para exames que desnecessários retardando assim seu correto diagnóstico. É importante ressaltar que a proposta em questão não busca substituir ou reduzir a necessidade da realização de exames **ouros** capazes de diagnosticar com total precisão a enfermidade discutida neste trabalho, mas sim fornecer ao profissional médico parâmetros que o auxiliem na decisão de encaminhar ou não seu paciente para tais exames que muitas vezes contam

com reações indesejadas além de financeiramente inviáveis.

Deve-se, no entanto considerar que, quanto menor a quantidade de dados para se analisar, menos preciso será o diagnóstico sugerido pela ferramenta, desta forma, para casos onde, existem poucos casos registrados no sistema os diagnósticos sugeridos pelo especialista médico e por outras técnicas mostram-se mais eficientes. Entretanto devido a sua capacidade de aprendizado logo o sistema torna-se apto para o trabalho. Uma opção para otimização dos resultados é permitir que o sistema realize análise de imagens de eletrocardiogramas, oferecendo mais opções ao especialista, que em todos os casos torna-se **indispensável** para o bom funcionamento da aplicação.

## CONCLUSÃO

O trabalho aqui proposto surge, em um ambiente repleto de propostas, como uma solução simples, de baixo custo e confiável para o auxílio no processo de diagnóstico da cardiopatia isquêmica, podendo e devendo ser utilizada principalmente em centros médicos mais afastados das capitais, permitindo assim aos atendentes destes centros, maior agilidade no encaminhamento de seus pacientes para profissionais especializados que podem confirmar ou não a existência de tal enfermidade que pode evoluir rapidamente para o óbito.

## REFERÊNCIAS

1. Carvalho EC de, Bachion MM, Ferraz AEP, Veiga EV, Ruffino MC, Robazzi MLCC. O processo de diagnosticar e seu ensino. Rev Esc Enferm USP. 1996;30(1):33-43.
2. Santos Neto L dos. O erro diagnóstico. 34ª ed. Brasília Médica: Brasília; 1997.
3. Carvalho M de, Vieira AA. Erro médico em pacientes hospitalizados. J Pediatr. 2002; 78(4):261-8.
4. Setiawan NA, Venkatachala PA, Hani AFM. Diagnosis of

- coronary artery disease using artificial intelligence based decision support system. Proceedings of the International Conference on Man-Machine Systems (ICoMMS); 2009 Oct 11-13; Penang, Malaysia.
5. Wangenheim CG, Wangenheim A. Raciocínio Baseado em Casos. São Paulo: Manole; 2003.
  6. Soni J, Ansari U, Sharma D. Intelligent and effective heart disease prediction system using weighted associative classifiers. *Int J Comput Sci Eng.* 2011; 3(6):2385-92.
  7. Almeida SM, Brasil L, Fernet E, Carvalho H. Algoritmo genético para determinação de relevância de sintomas no diagnóstico de cardiopatia isquêmica. *Anais do VII Workshop de Informática Médica*; 2007 Jun 25-26; Porto de Galinhas, PE. p.26-37.
  8. Boletim Brasileiro de Avaliação de Tecnologias em Saúde – BRATS. 2009;4(9). [citado 2013 Nov 28]. Disponível em: <http://200.214.130.94/rebrats/Publicações.html>
  9. Monard MC, Baranaukas JA. Aplicações de inteligência artificial: uma visão geral. São Carlos: Instituto de Ciências Matemáticas e de Computação de São Carlos; 2000.
  10. Poltosi LAC, Lemke N. Aplicação de algoritmos genéticos no projeto de transformadores. *Rev Liberato.* 2009; 10(14):167-74.
  11. Oliveira AG, Xexéo JAM. A aplicação de algoritmos genéticos no reconhecimento de padrões criptográficos [monografia]. Rio de Janeiro: Instituto Militar de Engenharia. Seção de Engenharia da Computação; 2011.
  12. Mitchell TM. *Machine learning*. New York: McGraw-Hill; 1997.
  13. Lorena AC, Carvalho ACPLF. Uma introdução às support vector machines. *Rev Inform Teor Aplic.* 2007; 14(2):43-67. [citado 2013 Nov 29]. Disponível em: [http://seer.ufrgs.br/index.php/rita/article/viewPDFInterstitial/rita\\_v14\\_n2\\_p43-67/3543](http://seer.ufrgs.br/index.php/rita/article/viewPDFInterstitial/rita_v14_n2_p43-67/3543)
  14. Coello CA. An empirical study of evolutionary techniques for multi objective optimization in engineering design [tese]. New Orleans: Tulane University; 1996.
  15. Lopes HS. *Fundamentos da computação evolucionária e aplicações*. Paraná: Bandeirantes; 2006.
  16. Laudon KC, Laudon JP. *Sistemas de informações gerenciais: administrando a empresa digital*. 5a ed. São Paulo: Prentice Hall; 2004.
  17. Tavares LG, Lopes HS, Lima CRE. Estudo comparativo de métodos de aprendizado de máquina na detecção de regiões promotoras de genes de *Escherichia coli*. *Anais do I Simpósio Brasileiro de Inteligência Computacional*; 2007 Out 8-11; Florianópolis, SC: Departamento de Automação e Sistemas, Programa de Pós-graduação em Engenharia de Automação e Sistemas, Centro Tecnológico, Universidade Federal de Santa Catarina. p. 8-11.
  18. Magalhães PMPA. Ferramenta educacional de simulação para algoritmos genéticos [dissertação]. Portugal: Universidade de Trás-os Montes e Alto Douro; 2008.
  19. Ferreira DF. *Estatística multivariada*. Minas Gerais: Universidade Federal de Lavras; 2008.
  20. Flores RE. A medicina baseada em evidências e o diagnóstico laboratorial. *NewsLab.* 2005;73:92-103.
  21. Tragante VO. Seleção de atributos por meio de algoritmos genéticos para diagnóstico auxiliado por computador em cardiopatia isquêmica [monografia]. Departamento de Física e Matemática da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto. Universidade de São Paulo; 2006.
  22. Freitas AML. *Teoria das probabilidades I*. Universidade Federal Fluminense. [citado 2013 Nov 18]. Disponível em: <http://www.professores.uff.br/anafarias/images/stories/meusarquivos/teoprobI-completa.pdf>
  23. Vasconcelos FDAGD, Cordeiro BA, Rech CR, Petroski EL. Sensitivity and specificity of the body mass index for the diagnosis of overweight/obesity in elderly. *Cad Saúde Pública.* 2010; 26(8):1519-27.