

Desenvolvimento de um Sistema para Auxílio ao Diagnóstico de Diabetes empregando Redes Neurais Artificiais (SADD)

João Paulo Vieira¹, Fabio José Parreira², Sidnei Renato Silveira²

¹Curso de Bacharelado em Sistemas de Informação, ²Departamento de Tecnologia da Informação – Universidade Federal de Santa Maria (UFSM/CESNORS) - Frederico Westphalen – RS – Brasil

vieira.jpvieira@gmail.com, fabiojparreira@gmail.com,
sidneirenato.silveira@gmail.com

Resumo. *Este artigo aborda o desenvolvimento de um sistema cuja principal função é o auxiliar os profissionais da área da saúde no diagnóstico da diabetes, utilizando a capacidade de aprendizagem das Redes Neurais Artificiais (RNAs). Foram utilizadas duas RNAs, cujo treinamento é realizado por meio da base de dados Pima Indians. Esta base de dados contém as informações referentes a uma comunidade indígena que é conhecida por apresentar a maior taxa de incidência de diabetes no mundo. As RNAs possuem objetivos diferentes, sendo que uma foi treinada para a busca de casos normais e a outra para diagnosticar casos de diabetes.*

Palavras-Chave: *Redes Neurais Artificiais, Diabetes, Base de Dados Pima Indians.*

Abstract. *This paper presents the development of a system whose main function is to assist health professionals in the diagnosis of diabetes, using the learning ability of Artificial Neural Networks. Training of the RNAs was done by Pima Indians database. Neural networks was trained differently where one will be trained to search for normal cases and the second to diagnose diabetes cases where further samples will be tested.*

Keywords: *Artificial Neural Networks, Diabetes, Pima Indians Database.*

1. Introdução

Nos últimos anos a incidência do diabetes tem aumentado rapidamente em nível mundial. Recentemente, a Organização Mundial de Saúde (OMS) reconheceu que a doença é epidêmica. As estatísticas apontam que o número de casos, em todo o mundo, atualmente chega a 246 milhões. Até 2025, esse número deve chegar a 350 milhões, de acordo com a Federação Internacional de Diabetes (IDF). No Brasil, segundo dados do Ministério da Saúde, estima-se que existam cerca de 11 milhões de portadores de diabetes, sendo que 7,5 milhões já sabem que tem a doença (OLIVEIRA; VENCIO, 2014).

Alguns dos fatores que favorecem esse crescimento alarmante do diabetes são a obesidade, hereditariedade, hábitos alimentares inadequados e o sedentarismo. Estudos demonstram que 46,6% dos pacientes diagnosticados desconheciam o fato de serem portadores da doença, e dos pacientes sabidamente diabéticos, 22,3% não faziam nenhum tipo de tratamento (BOTELHO, 2003).

Embora a detecção de diabetes esteja melhorando, o tempo para diagnosticá-lo pode ser superior a 10 anos, a contar do início da doença, até a concretização do diagnóstico. Para diagnosticar um paciente com diabetes, o médico tem que analisar diversos fatores e isto dificulta o seu trabalho. Para auxiliar estes profissionais, da área de saúde, são utilizados sistemas computacionais com a finalidade de apoiar o diagnóstico. Tais sistemas processam informações, com mais detalhe e em menor tempo, quando comparados com aos seres humanos, proporcionando uma melhora na qualidade dos serviços médicos, além de contribuir para a difusão de conhecimentos especializados (KAYAER; YILDIRIM, 2003).

Neste sentido, este artigo apresenta o desenvolvimento de um sistema inteligente denominado SADD (Sistema de Auxílio ao Diagnóstico de Diabetes), com capacidade para aprender e fazer análise de novos dados com o conhecimento adquirido, empregando Redes Neurais Artificiais (RNAs) multicamadas. As RNAs foram implementadas em uma aplicação *web*, adotando o *design* responsivo, para que fosse possível o acesso à aplicação por meio da maioria dos dispositivos existentes no mercado atual, tais como *smartphones*, *tablets* e computadores. Por meio dessa aplicação são manipuladas as informações necessárias para apoiar o diagnóstico de diabetes. O treinamento das RNAs foi realizado a partir da base de dados *PimaIndians*. As informações para compor esta base foram coletadas na comunidade indígena *Pima*, que vive perto de Phoenix, Arizona, Estados Unidos. Diversas pesquisas foram realizadas nesta comunidade, pois ela é conhecida por apresentar a maior taxa de incidência de diabetes do mundo (UCI, 2010).

Desta forma, este artigo apresenta, na seção 2, um referencial teórico sobre as áreas envolvidas (destacando conceitos de Diabetes, Redes Neurais, a Base de Dados *PimaIndians* e o *Design* Responsivo na *web*), situa o estado da arte (seção 3) e apresenta a solução implementada (seção 4). Encerrando o artigo são apresentadas as considerações finais, destacando os resultados obtidos e as referências empregadas.

2. Referencial Teórico

Esta seção apresenta um breve referencial teórico sobre as áreas envolvidas no desenvolvimento deste trabalho, destacando conceitos de Diabetes, Redes Neurais Artificiais, a Base de Dados *Pima Indians* e o Design Responsivo na *web*.

2.1 Diabetes

Uma pessoa diagnosticada com diabetes sofre de hiperglicemia, que por definição é o excesso de açúcar no sangue. Este excesso ocorre porque o pâncreas, órgão responsável pela produção da insulina, não consegue produzir a quantidade necessária deste hormônio para que a glicose possa entrar na célula, onde seria consumida, o que resulta no acúmulo desta no sangue. O diabetes não afeta só a capacidade do organismo de consumir açúcar, mas também a capacidade de utilizar outras fontes de energia como as proteínas e as gorduras (BAIER, HANSON, 2004). Basicamente, existem três tipos principais de diabetes: o diabetes tipo 1, que afeta principalmente os jovens, o diabetes tipo 2, que, geralmente, afeta as pessoas adultas e idosas e o diabetes gestacional, que pode ocorrer durante a gestação (MAZZAFERRI, 1988).

Para o diabetes tipo 1 existem diversas denominações: diabetes insulino dependente, diabetes infanto-juvenil ou diabetes imunomediado. A característica marcante deste tipo de diabetes é a produção insuficiente de insulina, pois suas células sofrem o que se chama de destruição autoimune. Neste tipo de diabetes a produção de insulina do pâncreas é insuficiente, pois suas células sofrem o que se chama de destruição autoimune. Os portadores de diabetes tipo 1 necessitam de injeções diárias de insulina para manterem a glicose no sangue em valores normais. Embora este tipo de diabetes ocorra em qualquer idade, é mais comum em crianças, adolescentes ou adultos jovens (OLIVEIRA; VENCIO, 2014).

Já o diabetes tipo 2, denominado também de diabetes não insulino dependente ou diabetes do adulto, corresponde a 90% dos casos de diabetes e geralmente ele é detectado em pessoas obesas, com mais de 40 anos de idade. Atualmente este quadro vem sofrendo mudanças, pois há grande ocorrência em jovens, em virtude de maus hábitos alimentares, sedentarismo, *stress* da vida urbana e do uso de agrotóxicos nos alimentos. Os agrotóxicos possuem um leque extenso de discussão acerca do seu uso, os efeitos nocivos que podem causar à saúde humana e ao meio ambiente. A ação sobre a saúde humana pode ser percebida horas após a exposição (essa exposição pode ser durante a aplicação, pelo ar que respiramos ou até mesmo pela ingestão de produtos contaminados) aos agrotóxicos, como também semanas, meses e anos após, sendo que os principais problemas percebidos são: náuseas, tonturas, dores de cabeça, alergias e até mesmo doenças mais graves como cânceres, alterações genéticas, doença de Parkinson e diabetes (ORTIZ, ZANETTI, 2001).

Os herbicidas e pesticidas estão mais ligados à incidência de diabetes do tipo 2, entre os humanos expostos aos agrotóxicos. As pesquisas sobre este tema ainda estão em fase inicial, mas estudos já apontam ligação entre a exposição a estes produtos e casos de diabetes tipo 2 (RIGOTTO, 2014). Neste tipo de diabetes, embora exista a presença de insulina, a sua ação é dificultada pela obesidade, o que é conhecido como resistência insulínica. Por ser pouco sintomático, este tipo, na maioria das vezes, permanece por muitos anos sem diagnóstico e sem tratamento o que favorece a ocorrência de suas complicações no coração e no cérebro (BOTELHO, 2003).

Por fim, o diabetes Gestacional, ocorre quando há a presença de glicose elevada no sangue, durante a gravidez. Ao final da gravidez, após o parto, geralmente a glicose no sangue, se normaliza. No entanto, as mulheres que apresentam ou apresentaram diabetes gestacional, possuem maior risco de desenvolverem diabetes tipo 2 tardiamente, o mesmo ocorrendo com os seus filhos (MAZZAFERRI, 1988).

2.2. Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) são um paradigma de processamento de informação inspirado no sistema neural biológico. As RNAs são sistemas massivamente paralelos e distribuídos, formados por unidades de processamento simples, que calculam determinadas funções, normalmente não lineares. Essas unidades são distribuídas em camadas, sendo interligadas por conexões, as quais se associam a pesos. Estes pesos armazenam o conhecimento representado na rede, servindo para ponderar as entradas recebidas por unidade constituinte (HAYKIN, 2001).

Uma das características mais importantes de uma RNA é a capacidade de aprender através de exemplos, denominados de conjunto de treinamento, sendo estes apresentados a ela durante a etapa de treinamento. Durante o aprendizado a rede consegue extrair os padrões existentes neste conjunto de dados e, após finalizar a etapa de treinamento, ela aprende o suficiente para generalizar o problema. Ao final, na etapa de teste, a RNA consegue produzir respostas satisfatórias para dados desconhecidos, que fazem parte do mesmo problema, para o qual ela foi treinada (HAYKIN, 2001).

Segundo Palma Neto e Nicoletti (2005), os algoritmos de treinamento das RNAs podem ser agrupados em supervisionados e não-supervisionados. Nos supervisionados, além das informações, cada exemplo de treinamento traz a classe a qual ele pertence. Nos grupos não supervisionados a RNA é treinada sem considerar as informações da classe a qual cada exemplo está associado.

Existem várias arquiteturas para a implementação de uma RNA, mas estas podem ser divididas, basicamente, quanto ao número de camadas, e pelos tipos de conexões entre os neurônios (FAUSETT, 1994). De acordo com o número de camadas, a arquitetura de uma RNA pode ser denominada de camada única ou de múltiplas camadas. As RNAs com múltiplas camadas distinguem-se de redes com camada única pela presença de camadas ocultas. Além disso, são utilizadas em problemas mais complexos, que necessitam analisar várias combinações das entradas da rede, como por exemplo, o diagnóstico de diabetes.

Quanto aos tipos de conexões, que definem a direção do fluxo de ativações, as RNAs são definidas como *feedforward* e *feedback*. As RNAs *feedforward*, também conhecidas como acíclicas ou não recorrentes, são estruturadas em camadas, e não possuem realimentação de suas saídas para as entradas. Estas podem ter de uma ou mais camadas em sua configuração. Dentre os algoritmos para treinamento de redes *feedforward*, o *back-propagation* está entre os que proporcionam uma maior eficiência. Já as RNAs *feedback*, também denominadas de cíclicas ou recorrentes, diferenciam-se

das não recorrentes por terem em sua estrutura a realimentação das saídas para as entradas (HAYKIN, 2001).

Para que uma rede neural tenha sucesso na resolução de problemas práticos, é necessário que ela consiga representar funções lineares e não-lineares¹, assim como o algoritmo de treinamento deve ser suficientemente robusto e genérico, não restringindo a eficácia na resolução de problemas. As redes neurais de múltiplas camadas, com o algoritmo para treinamento *back-propagation*, contemplam essas características (MASSAD et. al, 2004).

2.3 Pima Indians (PID)

Para treinar e testar a aplicação da RNA no sistema que será desenvolvido, pretende-se utilizar a base de dados da Universidade da Califórnia (UCI, 2010), denominada *Pima Indians Diabetes* (PID). As informações para compor esta base foram coletadas na comunidade indígena *Pima*, que vive perto de Phoenix, Arizona, Estados Unidos. Todas as pessoas, cujos dados foram coletados, são mulheres, com idade igual ou superior a 21 anos.

Diversas pesquisas foram realizadas nesta comunidade, pois ela é conhecida por apresentar a maior taxa de incidência de diabetes do mundo. Entre os adultos 50% são diabéticos, além de apresentarem uma alta prevalência de obesidade (BAIER; HANSON, 2004). Devido a estas características, os pesquisadores resolveram criar a base de dados PID, que é composta por 768 pessoas, onde 268 destas foram classificadas como tendo diabetes, do tipo 2, o que resulta em 34,9% da base. Os outros 65,1% não possuem diabetes e são classificados como normais.

Para cada pessoa que compõe a base, é informado: o número de vezes que a paciente ficou grávida, teste oral de tolerância à glicose, pressão arterial diastólica (mm Hg), espessura da dobra cutânea do tríceps (mm), idade (anos), índice de massa corporal (IMC) e a classificação (possui ou não diabetes).

2.4 Design Responsivo

Design Responsivo é considerado aquele design capaz de responder às características do dispositivo que está tentando acessá-lo, adaptando-se à expansão e contração que os mais variáveis dispositivos submetem. Para isso são usados itens fundamentais, tais como as tecnologias de *grid* fluído, imagem e mídias flexíveis e *media queries*. Os pilares de sustentação que mantêm o Design Responsivo são o HTML (*HyperText Markup Language*), *Java Script* e CSS3 (*Cascade Style Sheets*) (SÁNCHEZ, 2014).

¹ Funções Lineares são aquelas que podem ser resolvidas utilizando uma reta ou hiperplano como fronteira de decisão. Tem-se, como exemplo de função linear, a porta AND (RUSSELL; NORVIG, 2004). Funções Não-lineares são aquelas em que a solução contempla, por exemplo, quatro dimensões, sendo impossível resolvê-la por um hiperplano (por exemplo, a porta XOR) (RUSSELL; NORVIG, 2004).

As *medias queries* são uma ferramenta implementada no CSS3. A partir das mesmas, é possível enviar uma consulta ao navegador *web*, para identificar qual a mídia ou dispositivo que o usuário está utilizando para acessar o site. Desta forma, é possível realizar as modificações necessárias para o site se adaptar ao dispositivo. No caso da imagem e mídia flexíveis, as imagens e vídeos são adaptadas com relação ao tamanho da tela que o dispositivo está utilizando. Sendo assim, as mesmas devem ter uma boa resolução para não perder qualidade. Já o *grid* fluído trata-se da forma como se organiza a tela, onde os *grids* de conteúdos são adaptados conforme a porcentagem repassada pela tela do dispositivo que está acessando a aplicação (SILVA, 2014).

Para a implementação de um design responsivo existem boas práticas ou metodologias, tais como a *mobile-first*, que prevê primeiramente o desenvolvimento focado para dispositivos móveis. As ferramentas ainda estão em fases muito iniciais para o desenvolvimento deste tipo de design, mas existem alguns *frameworks* que apoiam este desenvolvimento (SILVA, 2014). Neste trabalho será adotado o *Bootstrap* que é um *framework* que contém todas as ferramentas necessárias para o desenvolvimento desta aplicação.

O *Bootstrap* é um *framework front-end* que foi desenvolvido pela equipe de engenharia do *Twitter*, com a intenção de melhorar o desenvolvimento de software interno. Em 2011 a ferramenta foi disponibilizada na comunidade de software livre através do *github* e acabou ganhando versões novas. Este *framework front-end* contém os códigos-fonte em *Javascript* e CSS prontos para o design responsivo, auxiliando no desenvolvimento de páginas *web* que serão utilizadas por vários dispositivos (GETBOOTSTRAP.COM, 2015).

Na aplicação apresentada neste artigo, o design responsivo foi utilizado para que os usuários possam acessá-la por meio de qualquer dispositivo com acesso à Internet, o que proporciona, aos profissionais da área de saúde, ampla acessibilidade aos serviços implementados no SADD, por meio dos diversos dispositivos existentes no mercado.

3 Estado da Arte

Esta seção apresenta alguns trabalhos relacionados ao desenvolvimento do sistema implementado para diagnóstico de diabetes, sistemas estes que estão disponíveis para acesso via dispositivos móveis e também para computadores do tipo *desktop*. No final desta seção apresenta-se um quadro comparativo com os trabalhos estudados e o trabalho desenvolvido.

3.1 Um modelo híbrido para o apoio ao diagnóstico de diabetes

Menezes (2013) apresenta um modelo híbrido para o diagnóstico de diabetes do tipo 2, utilizando um Sistema Especialista probabilístico, aplicando Redes Bayesianas, Metodologias Multicritério de Apoio à Decisão (*Multi-Criteria Decision Analysis - MCDA*) e representações de conhecimento estruturadas em regras de produção.

Para a aplicação dessa metodologia foram definidos 5 passos para que a mesma obtivesse sucesso: 1) identificar os parâmetros para o diagnóstico de Diabetes; 2) aplicar Redes Bayesianas, a fim de representar as relações entre os parâmetros para o diagnóstico e estabelecer a probabilidade de cada parâmetro; 3) aplicar a Metodologia Multicritério de Apoio à Decisão para selecionar quais desses parâmetros proporcionam um maior impacto no diagnóstico de Diabetes; 4) construir um Sistema Especialista (SE), onde os parâmetros selecionados como de maior impacto no diagnóstico irão compor a base de conhecimento; 5) complementar a base de dados com os resultados obtidos pelo SE para recalculas as probabilidades.

Inicialmente foi realizado um levantamento para obter os parâmetros necessários para o diagnóstico da diabetes. Após este levantamento, foi desenvolvido o modelo para aplicação da metodologia MCDA, usando como base de dados a *Statistics Canada*. A partir dessa base foram estabelecidos sete parâmetros para o diagnóstico da diabetes: idade, presença de complicações crônicas, Índice de Massa Corporal (IMC), atividade física, fumo, consumo de álcool e consumo diário de frutas/vegetais.

Um *software* foi desenvolvido para a o cadastro destes dados e a geração automática de *Redes Bayesianas*, utilizando a API (*Application Program Interface*) *Netica-J* e a geração das Matrizes Juízo de Valor, que são matrizes contendo os dados das *Redes Bayesianas*. Este *software*, denominado SDD (Sistema de Diagnóstico de Diabetes), foi desenvolvido usando a linguagem de programação Java e o banco de dados *PostgreSQL*. Logo após realizar os cadastramentos na base de dados do SDD são geradas as redes Bayesianas onde, para cada parâmetro, são calculadas as probabilidades, utilizando a API *Netica-j*. As *Redes Bayesianas* geradas podem ser salvas em arquivo e abertas posteriormente pela interface gráfica da aplicação. Essas redes são responsáveis por apontar quais parâmetros têm mais relevância no diagnóstico da doença.

A aplicação da metodologia multicritério de apoio à decisão foi realizada por meio do *software* MACBETH, que permite a identificação dos parâmetros que possuem mais ou menos influência no diagnóstico da doença. Para isso foi realizada uma sequência de passos que são: estruturação, avaliação e recomendação.

A fase de estruturação busca identificar, caracterizar e organizar os fatores considerados relevantes no processo de apoio à decisão. Na segunda fase é realizada a avaliação dos fatores levantados pelas fases anteriores. Na recomendação são feitas análises de sensibilidade e robustez, para verificar se mudanças nos parâmetros do modelo de avaliação interferem no resultado final.

A construção da base de conhecimento foi realizada por meio da ferramenta *ExpertSINTA*, onde são definidos os valores e variáveis. Em um segundo momento deve-se selecionar o objetivo sendo, neste caso, o diagnóstico da doença. Definiu-se, também, a interface do usuário, onde são realizadas perguntas ao paciente para que o mesmo possa interagir com o sistema. O passo seguinte envolve a criação das regras na base de conhecimento (regras estas que farão as sugestões de diagnósticos).

Essa metodologia foi aplicada em duas bases de dados diferentes, a *Statistics Canada* (Canadá) e a *CAMED* (Brasil), onde parâmetros são checados como diferentes, pois na base canadense são encontrados sete parâmetros (descritos anteriormente) e na brasileira foram encontrados oito, que são: Idade, Histórico familiar, circunferência abdominal, atividade física, fumo, álcool, hábitos alimentares saudáveis e hipertensão.

Como conclusão do trabalho a autora menciona que obteve sucesso na aplicação, pois o grau de confiança da rede *Bayesiana* foi considerado adequado pelo *software Hiview* que é uma ferramenta de modelagem de decisão que permite apoiar os tomadores de decisão na avaliação das opções, a fim de escolher o melhor caminho a seguir. Como trabalhos futuros a mesma ressalta a aplicação do sistema em bases diferentes das apresentadas e, também, com um maior volume de dados do que foi utilizado no trabalho. Ressalta, também, o possível desenvolvimento de uma aplicação *web* para melhorar a interação do usuário com o sistema.

3.2 Diagnosing diabetes using neural networks on small mobile devices (Diagnósticos de diabetes usando redes neurais e tecnologia móvel)

O trabalho apresentado por Karan (2012) foi desenvolvido na *Haliç University* em Istanbul, Turquia, tratando-se de uma aplicação móvel para o diagnóstico de diabetes, empregando RNAs. Essa aplicação utiliza a arquitetura cliente-servidor, pois a rede neural é criada e treinada em um servidor que tem capacidade de processamento muito maior do que o dispositivo móvel que foi usado no desenvolvimento da aplicação. Depois, a rede neural é enviada para a aplicação que realiza os diagnósticos. O dispositivo móvel não precisa estar conectado ao servidor no momento do teste, pois o mesmo, quando recebe a rede atualizada, armazena a última rede que foi treinada no servidor, armazenando também os dados do que foram fornecidos pelo usuário. Ao restabelecer a conexão entre ambos é realizada uma sincronização, momento em que o servidor recebe os dados coletados pelo dispositivo móvel e este recebe a rede treinada mais recentemente.

A rede neural desenvolvida apresenta três camadas: na primeira, para a entrada de dados, são usadas onze entradas para os testes e o mesmo número para o treinamento, que são: idade, atividade física, gravidez, diabetes na família, índice de massa corporal, espessura da dobra cutânea, colesterol, pressão arterial, teste de insulina em jejum, taxa de glicose no sangue, ascendência de diabetes. A segunda camada (camada oculta) é composta por 12 neurônios, onde é realizado o processamento das informações. Por último existe a terceira camada (camada de saída) onde existem somente duas saídas: diabético ou não diabético.

Para o treinamento da rede foi utilizada uma base de dados com informações extraídas de um hospital, contendo um conjunto de dados com 456 observações utilizadas como dados de treinamento. As primeiras 228 são pessoas consideradas normais e as demais possuem a doença. Ainda na fase de treinamento foram realizados testes aplicando a metodologia de tentativa e erro. Neste sentido, foram realizados

vários treinamentos de 1000 até 7000 interações, sendo observado que com 5000 interações é possível chegar a um nível adequado de confiabilidade para a rede .

Na aplicação móvel existe um formulário onde são coletadas as informações de entrada para as variáveis necessárias ao funcionamento do sistema. A partir desta entrada é realizado o teste, aplicando-se a versão da rede neural mais atual, pois se o dispositivo móvel estiver conectado em uma rede com acesso à Internet, a aplicação se encarrega de atualizar a rede neural. Após realizado o teste, o dispositivo envia os dados para o servidor para que os mesmos possam ser usados em futuros treinamentos. A aplicação foi desenvolvida para *Windows Mobile 6.5* utilizando o *Compact Framework* da plataforma .NET. O aplicativo do servidor foi desenvolvido utilizando-se o *Framework 3.5* desta mesma plataforma.

Como conclusão o autor menciona o diferencial da aplicação, já que o treinamento da rede neural é realizado em dispositivo com maior capacidade de processamento e seus resultados são enviados para o dispositivo móvel, que tem menos poder de cálculo. O autor concluiu, também, que esse tipo de arquitetura de sistema pode ser utilizado para auxiliar no diagnóstico de outros tipos de doenças que possuam características semelhantes a diabetes.

3.3 An Intelligent Expert Based System Neural Network For The Diagnosis Of Type2 Diabetes Patients (Um sistema Inteligente Baseado Rede Neural para o diagnóstico de pacientes com diabetes tipo 2)

O trabalho apresentado por Sujatha (2014) foi desenvolvido em uma universidade da Índia, país em que o índice de diabetes é um dos maiores do mundo. Neste trabalho foi desenvolvido um comparativo entre duas técnicas que podem ser utilizadas para o diagnóstico de diabetes que são a regressão² e as RNAs. Para o desenvolvimento deste comparativo o autor utilizou a base de dados *Pima Indians (PID)*, a partir da qual pôde testar e validar a aplicação desenvolvida.

Para fazer essa avaliação o autor desenvolveu uma rede neural e também um sistema de regressão, a partir do qual, utilizando a base PID, obteve um índice de acerto de 57% na regressão. Aplicando a RNA, o mesmo obteve um resultado de 74% de acerto. Sendo assim o mesmo concluiu que a RNA tem um grau maior de acerto no diagnóstico de diabetes.

3.4 Quadro comparativo

Após o estudo dos trabalhos relacionados à proposta apresentada neste artigo, traçou-se um quadro comparativo (Quadro 1) onde se pode analisar algumas características destes trabalhos.

² Regressão são modelos matemáticos que relacionam o comportamento de uma variável Y com outra X (REGAZZI.; SILVA, 2010).

Comparando os três trabalhos apresentados com o trabalho implementado, o SADD (Sistema de Auxílio ao Diagnóstico de Diabetes), verificam-se muitas características semelhantes, tais como aplicação de RNAs e o uso da base PID. Destaca-se, também, que os trabalhos apresentados por Menezes (2013) e Karan (2012) têm, como propósito, diagnosticar a diabetes do tipo 2, o mesmo propósito do SADD.

O diferencial do SADD é a implementação de duas RNAs, fazendo com que o índice de certeza aumente gradativamente em relação aos outros trabalhos. Outra característica importante que o SADD apresenta é a sua capacidade de ser executado na maioria dos dispositivos que estão no mercado atual, já que o mesmo foi projetado com *design* responsivo.

Quadro 1: Comparativo entre Trabalhos

Característica Analisada	Trabalho proposto por Sujatha (2014)	Trabalho proposto por Karan (2012)	Trabalho proposto por Menezes (2013)	Solução Implementada SADD
Tecnologias utilizadas	Java	.NET(<i>Compact Framework, Framework3.5</i>)	Java, MACBETH	Java, CSS, JavaScript, HTML5
Base de Dados	PID	Base própria	<i>Statistics Canada</i> (Canadá), CAMED (Brasil)	PID
Plataforma Desktop	Sim	Não	Sim	Sim
Plataforma móvel	Não	Sim	Não	Sim
Técnicas empregadas para apoiar o diagnóstico	RNAs e Métodos de Regressão	RNAs	Redes <i>Bayesianas</i>	RNAs
Banco de dados	Arquivo texto	<i>MySql</i>	<i>PostgreSQL</i>	<i>MySql</i>
Design responsivo	Não	Não	Não	Sim

4 Solução Implementada

A solução apresentada neste trabalho envolve a implementação de um Sistema de Informação para o apoio ao diagnóstico de diabetes, empregando RNAs, o SADD (Sistema de Auxílio ao Diagnóstico de Diabetes). Inicialmente uma aplicação *web* foi sendo desenvolvida para que os usuários (profissionais da área de saúde) pudessem

interagir com o sistema. Esta aplicação é responsável por receber os dados dos pacientes e submetê-los as redes neurais para que o teste seja realizado.

A escolha de uma aplicação *web* deveu-se ao fato de que a mesma permite a padronização das informações a serem salvas e, também, pela mobilidade, pois a mesma é responsiva, podendo ser acessada por meio de qualquer *smartphone*, *tablet* ou computador *desktop*. Essa mesma aplicação é usada, posteriormente, para que os profissionais da área de saúde realizem a inserção dos dados pessoais e obtenham apoio para realizarem o diagnóstico de diabetes.

4.1 Arquitetura e Funcionamento do Sistema

A Figura 1 apresenta a arquitetura do SADD. Nesta figura também é possível observar o fluxo de dados do sistema. A aplicação *web* é responsável por receber os parâmetros dos pacientes e enviá-los para a base de dados, onde os mesmos ficam armazenados para futuros treinamentos das RNAs. Esses dados são utilizados pelas RNAs para que seja realizado o teste da amostra.

Logo após esse teste, as RNAs enviam os resultados obtidos para um quantificador de certeza, que faz a análise entre os resultados obtidos por meio das RNAs, devolvendo o resultado, ou seja, a mensagem contendo o pré-diagnóstico.

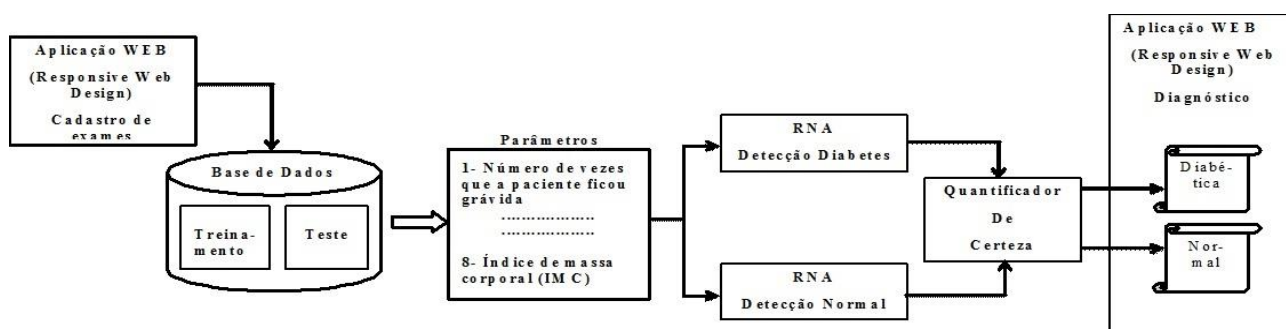


Figura 1: Arquitetura do sistema Fonte: dos autores.

Os parâmetros são enviados para as RNAs para que as mesmas realizem a análise das informações e as enviem para o quantificador de certeza, responsável por realizar a análise entre as duas redes e definir a porcentagem de certeza do diagnóstico. Após realizada a quantificação da amostra, o resultado é enviado para a aplicação *web* em forma de uma mensagem “DIABÉTICO % DE CERTEZA” ou “NÃO DIABÉTICO % DE CERTEZA”.

O uso de RNAs foi escolhido devido a sua habilidade de adaptação e aprendizagem, já que podem lidar com dados imprecisos e situações não totalmente definidas. Uma rede treinada corretamente tem a habilidade de generalizar quando é

apresentada a entradas que não estão presentes em dados já conhecidos pela mesma. Isso pode acontecer com dados relacionados a diabetes, pois não há uma uniformidade dos dados a serem analisados (HAYKIN, 2001).

Estudos realizados apontam que as RNAs têm um grau maior de acerto no diagnóstico de diabetes. O trabalho apresentado por Sujatha (2014) realizou uma comparação entre redes neurais e regressão matemática, atingindo um índice de 17% a mais de acertos usando RNA. Esse teste foi realizado utilizando a mesma base de dados que será usada pela aplicação aqui apresentada, a PID.

Na base de dados do SADD são armazenadas todas as informações que o levantamento coletar, sendo que a mesma é dividida em 2 partes: uma para fazer o treinamento das duas redes neurais e a segunda para fazer o teste das mesmas. Essa divisão é de 30% da base para o teste e 70% para o treinamento (KAYAER; YILDIRIM, 2003).

Foram desenvolvidas duas RNAs, sendo que as duas foram treinadas com as mesmas informações. Entretanto, uma das RNAs realiza a detecção da diabetes e a outra busca os casos normais, onde não é encontrada a doença. Cada uma das RNAs fornece uma porcentagem de certeza que é calculada pelo algoritmo da RNA, por meio da implementação de um quantificador de certeza.

O quantificador de certeza tem, como objetivo, a partir das porcentagens geradas pelas duas RNAs, quantificar uma porcentagem geral, classificando se a amostra treinada é ou não diabética. Esse resultado é devolvido para a aplicação *web* que, por sua vez, mostra uma mensagem informando se esse paciente é ou não diabético e a porcentagem de certeza que o sistema conseguiu atingir.

4.2 Levantamento de Requisitos

Nesta seção são descritos os principais requisitos que foram considerados na implementação do SADD, destacando-se que o sistema foi implementado de forma gradativa:

- Importação automática das informações contidas na base *Pima*: a base de dados PID está disponível na *web*, no formato TXT (texto), portanto, foi preciso importá-la para a base de dados *MySQL* do SADD, sendo que os dados nela contidos devem manter a sua integridade;
- Treinamento e teste das RNAs: O administrador do sistema deve treinar as redes neurais sempre que houver alterações significativas nos parâmetros e/ou nas amostras referentes a diabetes. O treinamento deve se dar com os 70% da base de dados *pima*, usando cerca de 538 amostras para o treinamento, os outros 30 % da amostra são utilizados para o teste das redes neurais, onde são usadas 230 amostras;

- Teste das amostras: após o treinamento das RNAs, as mesmas devem passar por um teste, para que se tenha a certeza de que o conhecimento adquirido pelas mesmas é o suficiente para realizar um diagnóstico seguro da doença;
- Cadastro das informações via interface *web*: as informações dos usuários são cadastradas por meio de uma aplicação *web*; a mesma deve conter um formulário, onde são informados os parâmetros sobre a doença; logo após o usuário entrar com os parâmetros para teste, a aplicação deve submeter essa amostra ao teste e devolver o resultado para o usuário novamente.

4.3 Modelagem do Sistema

4.3.1 Diagrama de Casos de Uso

De acordo com o levantamento de requisitos, o SAAD possui, basicamente, quatro funcionalidades: 1) importação das informações contidas na base de dados PID; 2) cadastro das informações dos pacientes; 3) treinamento das RNAs e 4) teste das redes. Para realizar essas tarefas são necessários dois atores principais, que são o administrador do sistema e o usuário final que, na maioria das vezes, é um profissional da área da saúde. A Figura 2 mostra o diagrama de casos de uso, construído com a ferramenta *Visual Paradigm for UML11.0* (VISUAL-PARADIGM.COM, 2015), apresentando a interação dos atores com as funcionalidades do SADD.

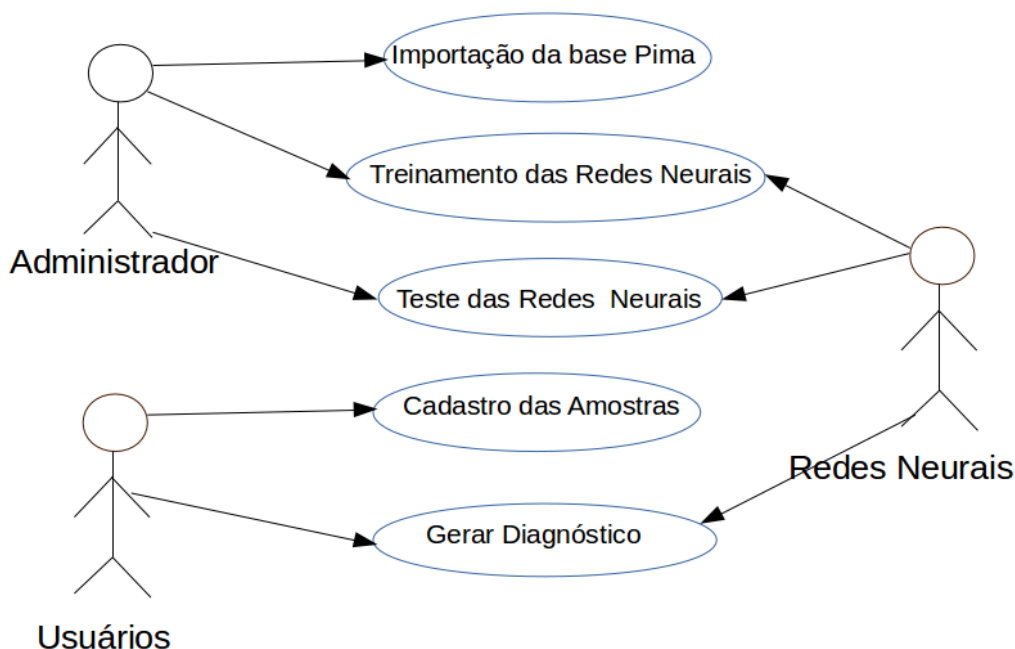


Figura 2: Diagrama de Casos de Uso (Fonte: Dos autores)

Optou-se por incluir as RNAs como atores, já que são partes inteligentes do sistema. A interação do ator RNA se dá no treinamento e testes das redes, bem como para mostrar o diagnóstico para os usuários, como se vê na Figura 2.

4.3.2 Modelagem da Base de Dados

A base de dados do SADD foi desenvolvida usando a linguagem SQL (*Structured Query Language*) por meio do SGBD (Sistema Gerenciador de Bancos de Dados) *MySQL*. Esta base de dados possui 13 tabelas sendo uma para armazenar os dados da base *Pima*, uma para armazenar os dados dos usuários, uma para armazenar os dados das amostras recebidas (parâmetros dos pacientes, informados pelos usuários), seis tabelas utilizadas para armazenar os resultados dos treinamentos das RNAs, onde são salvos os pesos dos neurônios das RNAs, e quatro tabelas para armazenar dados dos testes onde são armazenadas as saídas dos testes antes do quantificador e após o quantificador. A Figura 3 apresenta o modelo E-R (Entidade-Relacionamento) da base de dados do SADD, construído por meio da ferramenta *MySQL Workbench* (MYSQL.COM, 2015).

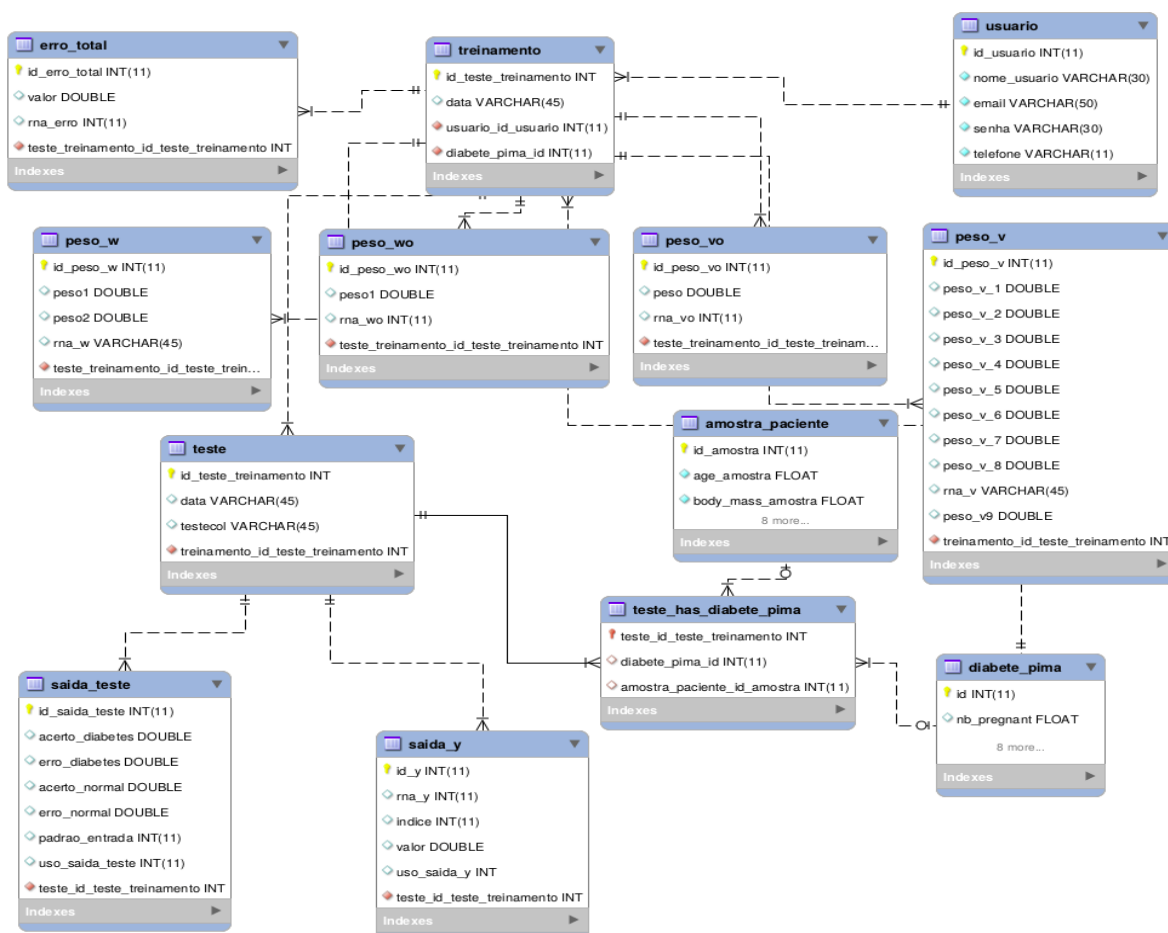


Figura 3: Modelo ER da base de dados do SADD Fonte: Dos autores

De acordo com a Figura 3, destaca-se que a tabela *diabete_pima* serve para armazenar os dados referentes à comunidade indígena *Pima*, com base nos parâmetros estabelecidos na PID disponível na *web* (UCI, 2010). Já a tabela *amostra_paciente*, armazena as informações referentes aos dados coletados pelos profissionais da área de saúde, por meio de uma interface *web*.

Na tabela *usuário* são salvos alguns dados sobre os usuários que irão administrar o sistema, sendo armazenados dados como nome, *e-mail*, senha e telefone. Nas tabelas *treinamento*, *peso_w*, *peso_wo*, *peso_vo*, *peso_v* e *erro_total* são armazenados os dados referentes ao treinamento das RNAs. A tabela *treinamento* é usada como uma tabela geral onde é salvo o administrador que realizou o treinamento, a data do treinamento e os dados que foram utilizados para o treinamento. A tabela *peso_w* e *peso_wo* armazenam os dados referente à camada de saída das RNAs. As tabelas *peso_vo* e *peso_v* são responsáveis por armazenar os pesos da camada oculta, pesos estes usados pela aplicação para realizar os testes das amostras submetidas pelos usuários.

Nas tabelas *saida_teste*, *teste* e *saida_y* são armazenados os dados referentes aos testes realizados nas RNAs, testes estes que demonstram a eficiência das redes neurais bem como do quantificador de certeza. Esses dados de teste são armazenados para posteriormente mostrar os índices em forma de gráfico na aplicação *web*.

4.4 Tecnologias Empregadas

Para desenvolver esse projeto foram utilizadas tecnologias *open source*, disponíveis gratuitamente, podendo-se destacar o HTML5 (*Hypertext Markup Language*), que é a linguagem de marcação mais usada da atualidade para a implementação de *sites* (SILVA, 2011). Para o desenvolvimento das RNAs utilizou-se a linguagem de programação Java, que proporciona uma performance estável e, também, pela facilidade na implementação de RNAs com esta linguagem, já que a linguagem Java possui instruções que facilitam processamento de cálculos, que é a base das redes neurais (GONÇALVES, 2006). Para a base de dados utilizou-se o SGBD *MySQL*, que é muito utilizado em aplicações *web* de pequeno porte.

4.5 Funcionalidades Implementadas

Foram implementadas todas as funcionalidades descritas na seção de levantamento de requisitos (4.2) tornando, assim, todos esses requisitos válidos. Nas próximas seções as funcionalidades implementadas serão detalhadas.

4.5.1 Importação da base Pima

Como descrito anteriormente, a base de dados *Pima* é disponibilizada no *site* da Universidade da Califórnia, em formato de arquivo do tipo texto (TXT). Para facilitar o desenvolvimento da aplicação e possibilitar que o treinamento das RNAs fosse realizado

de forma mais rápida, esta base foi transferida para uma base de dados MySQL, de forma automatizada, por meio de um aplicativo desenvolvido em linguagem de programação Java. Estes dados foram utilizados no treinamento e testes das RNAs.

4.5.2 Página Principal do SADD

Por meio da página principal do SADD, os usuários (administrador e usuário final) dispõem de acesso às funcionalidades do sistema. O SADD está disponível no *link* 200.132.38.216:8080 para o uso da população. A Figura 4 apresenta a página inicial do SADD.



Figura 4: Página Principal do SADD Fonte: Dos autores

Na parte superior da página principal do SADD (Figura 4) existe um menu com informações sobre alguns parceiros da aplicação (*links* para os *sites da* UFSM e do CESNORS), além da SBD – Sociedade Brasileira de Diabetes). Por meio do menu “SADD” os usuários têm acesso às funcionalidades do sistema, que são: Cadastro de novos exames, gerar gráfico das redes neurais e uma página *Sobre*, com informações sobre o sistema.

4.5.3 Cadastro de novos exames

Na Figura 5 pode-se observar onde é realizado o cadastro das amostras de dados dos pacientes, por meio de um formulário. Cada um dos campos é validado, para que não seja deixado em branco ou com caracteres que não dão características de uma amostra válida. Por exemplo, quando forem digitadas letras em um campo que deve ser

numérico, o sistema apresenta uma mensagem mostrando que os caracteres encontrados não fazem parte de uma amostra válida. Esse formulário conta com um botão na parte inferior da tela onde os dados são validados e, se estiverem dentro do padrão de entrada, são submetidos às redes neurais para serem avaliados, bem como são armazenados na base de dados.

TESTE - Mozilla Firefox
localhost:8080/TGSI/Form-teste.jsp

Realizar Teste

Para realizar o teste tenha todas as informações em mãos, pois não serão permitidos campos em branco
A consistência dos dados aplicados ao teste é de responsabilidade do usuário
Aplicação desenvolvida em caráter acadêmico. Os diagnósticos estabelecidos pela mesma devem ser validados por um Médico

Passe os dados:

Número de vezes que ficou grávida
Quantidade de vezes que ficou grávida

Concentração de glicose no Plasma em teste de tolerância de glicose oral
Concentração de glicose no Plasma

Pressão sanguínea Diastólica (mm Hg)
Pressão sanguínea

Dobras na pele do tríceps (mm)
Dobras na pele

2-Horas de insulina de soro (mu U/ml)
Insulina de soro

Índice de massa corpórea (peso em kg/(altura em m2))
Massa corpórea

Figura 5: Formulário de coleta dos dados do SADD Fonte: Dos autores

Ao submeter esses dados, os mesmos seguem o fluxo apresentado na Figura 1 e o SADD retorna a mensagem mostrando o diagnóstico que o sistema encontrou, como se pode observar na Figura 6.

Resultado - Mozilla Firefox
localhost:8080/TGSI/resultado.jsp

A amostra é Normal com 72,5% de Certeza

Diabetes Home

Figura 6: Resultado do Teste Fonte: dos autores

O resultado final mostra a classificação das amostras submetidas ao SADD. A mensagem mostra a porcentagem de certeza obtida por meio de uma média dos pesos dos neurônios da camada de saída e uma classificação atribuída pelo quantificador de certeza: Normal, Diabética ou Indefinida. No caso da amostra ser Normal, ou seja, o paciente não possuir a doença, a mensagem será mostrada na cor verde; caso o

quantificador a classifique como indefinida, identificando que não foi possível obter um resultado, a mensagem será mostrada na cor laranja e, se a amostra for classificada como diabética, a mensagem mostrada será na cor vermelha.

4.5.4 Treinamento das RNAs

Na Figura 7 pode-se observar a funcionalidade que permite que o administrador do sistema limpe os dados de um treinamento anterior e realize novamente o treinamento das redes.

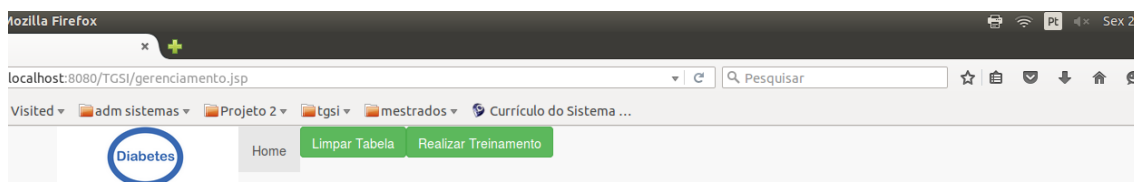


Figura 7: Gerenciamento das Redes Neurais

O acesso à página demonstrada pela Figura 7 se dá por meio da página principal do SADD (Figura 2) onde o administrador do sistema deve clicar no menu “Acesso Restrito” e, posteriormente, em “Manutenção”. O usuário será submetido a uma tela solicitando *login* e senha. Após a autenticação do usuário, o mesmo visualizará as opções mostradas na Figura 7.

4.5.5 Funcionamento do sistema

O SADD foi implementado para ter um funcionamento prático e dinâmico. Para tanto, utilizou-se o *design* responsivo da aplicação *web*, para que a mesma possa ser acessada por meio de diferentes dispositivos. A aplicação *web* é responsável por realizar o treinamento das RNAs e, também, o teste das amostras submetidas. Quando o treinamento das redes é disparado, ocorre de forma paralela entre as RNA1 (para diagnóstico de casos diabéticos) e a RNA2, que tem por objetivo descobrir os casos normais. Para esse treinamento os dados são submetidos a um algoritmo de normalização, onde os mesmos são dispostos em uma faixa de -1×10^{-6} a 1×10^6 para serem submetidos às redes neurais. Essa faixa de valor é utilizada para que não haja um intervalo muito grande entre os dados passados para as RNAs. Após a normalização, os dados são passados para as RNAs. Cada uma das RNAs possui três camadas: entrada, oculta e de saída, como mostra a Figura 8.

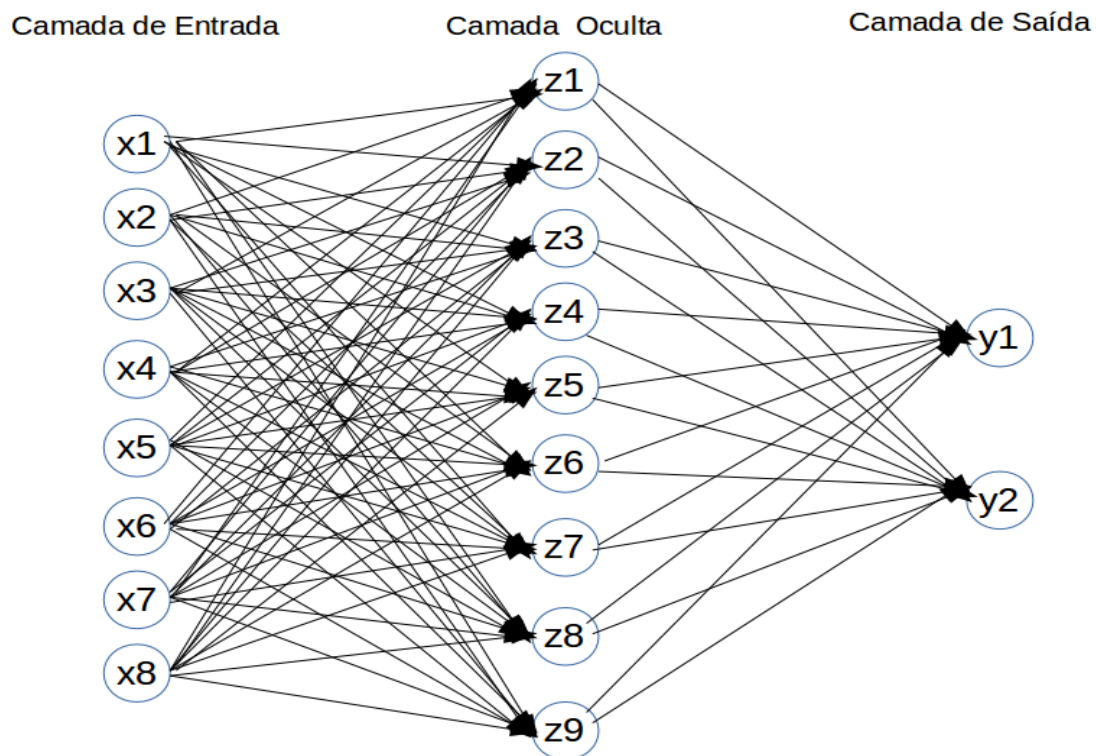


Figura 8: Camadas das Redes Neurais do SADD: Fonte: Dos autores

Como se pode ver na Figura 8, existem 8 neurônios na camada de entrada, correspondentes aos parâmetros que são nativos da base que está sendo utilizada (base *Pima*), parâmetros estes que são passados para a camada oculta. Na camada oculta existem 9 neurônios que são treinados pelas amostras, por meio dos parâmetros passados pela camada anterior. Esses neurônios recebem uma espécie de peso que é salvo no banco de dados para posteriormente serem utilizados nos testes. Por fim, a camada de saída possui dois neurônios (que indicam os possíveis resultados), cujos pesos também são salvos na base de dados.

Nos testes das RNAs, tanto utilizando a base de dados nativa como na entrada de novas amostras, os parâmetros de teste são submetidos à normalização. Após a normalização os dados atribuídos aos neurônios são buscados na base de dados e é realizado o teste da amostra, teste esse que atribui quatro valores, sendo dois para cada rede neural.

O quantificador de certeza (conforme arquitetura apresentada na Figura 1) recebe os dados das saídas das RNAs e, por meio de seu algoritmo, analisa os quatro valores que são obtidos por meio do teste ao que a amostra foi submetida e aplica métricas para que seja calculada a saída final que será submetida para a aplicação *web* e mostrada para o usuário que realizou o teste o resultado.

4.6 Testes e Validação

A aplicação foi testada e validada por meio da base de dados PID, utilizando um fator de certeza de no mínimo 75% (LORENZI; SILVEIRA 2011). A base PID já foi testada e validada por um número expressivo de pessoas especializadas na área de diabetes (UCI, 2010). Para os testes realizados pelo SADD, utilizando os dados da base PID, obteve-se um resultado satisfatório após 83 treinamentos, já que o índice de acerto ficou dentro dos parâmetros esperados. A Figura 9 apresenta os gráficos que mostram os índices de acerto relacionados ao treinamento das RNAs e, também, do quantificador de certeza.

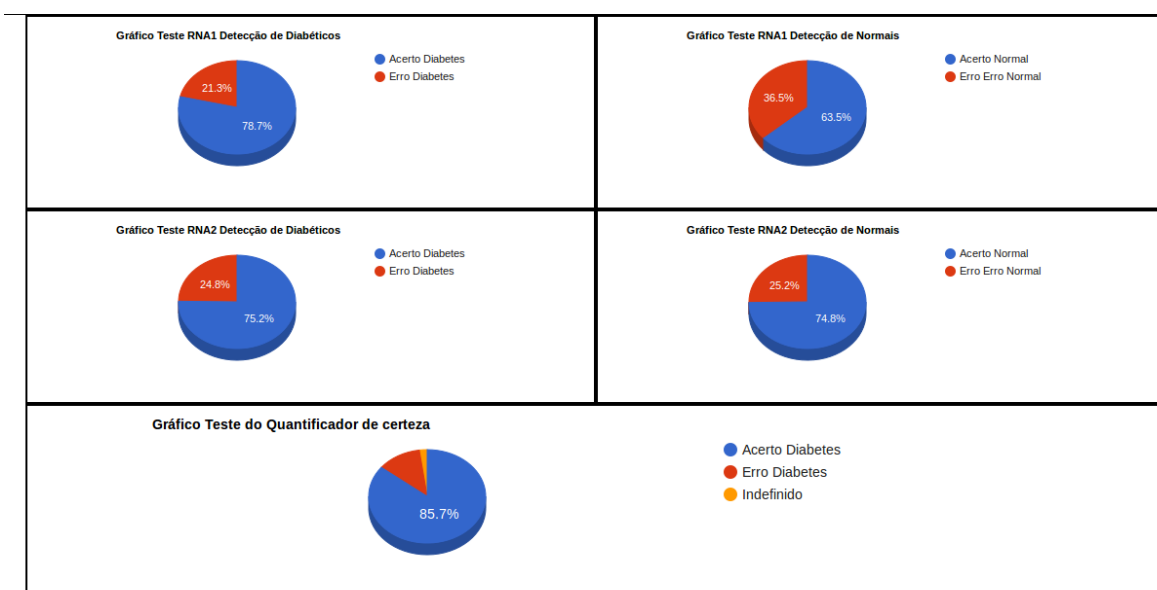


Figura 9: Gráficos dos Testes Fonte: Dos Autores

Como se pode observar na Figura 9, os resultados dos testes das RNAs foram apresentados em 4 gráficos principais, sendo 1 para cada saída das redes, além de um gráfico que se refere ao quantificador de certeza. Nos dois primeiros gráficos pode-se observar os testes da RNA1, rede que é responsável pela detecção dos casos diabéticos. Verifica-se que, no primeiro neurônio de saída, obteve-se um índice de 78,7% de acerto e, no segundo neurônio, o índice foi de 63,5% de acerto. Nos dois gráficos localizados no meio da figura têm-se os resultados dos testes da RNA2 que está programada para detectar casos normais (aqueles que não possuem a doença). Verifica-se, no primeiro neurônio, que se obteve um índice de acerto de 75,2% de acerto e, no segundo neurônio o índice de acerto ficou em 74,8%.

Na parte inferior da Figura 9 pode-se observar o gráfico do teste do quantificador de certeza, onde se verifica uma melhora no índice de acerto, pois o mesmo obteve um percentual de 85,7% de acertos nos testes realizados, tendo-se, também, um índice de 2,2% de indefinidos, ou seja, amostras em que o quantificador de certeza, pela proximidade dos valores dos neurônios de saída, não obteve um diagnóstico. O índice de erro após a aplicação do quantificador de certeza ficou em 12,2%. Considera-se este

um índice baixo, comparando-se com outros trabalhos já desenvolvidos usando RNAs, como se pode observar no trabalho desenvolvido por Sujatha (2014).

5 Considerações Finais

Acredita-se que os objetivos do trabalho foram alcançados, já que o SADD encontra-se em funcionamento e a validação comprovou os resultados esperados. As Redes Neurais estão sendo treinadas e sendo usadas para o diagnóstico das amostras cadastradas pelos profissionais da saúde e pelo público em geral. Além disso, no teste das redes, o índice de acerto de amostras que não eram conhecidas pela RNA foi de 85,7%, apresentando um resultado satisfatório para o trabalho realizado.

Neste contexto, este trabalho teve como principal contribuição o desenvolvimento da ferramenta SADD, que pode ser utilizada por profissionais da área da saúde no auxílio ao diagnóstico da diabetes tipo 2. Cabe destacar que essa ferramenta ficará disponível com o código *open source* para quem quiser utilizar e adaptá-la para apoiar a análise de diagnósticos de outras doenças comuns na população mundial.

Como trabalhos futuros pode-se destacar a construção de uma base de dados da população brasileira permitindo, assim, que o SADD use parâmetros mais próximos da realidade brasileira. Outra possibilidade é utilizar outros parâmetros, tais como a porcentagem de agrotóxicos aplicados aos alimentos. Acredita-se que, com uma base de dados propriamente construída a partir de dados da população brasileira, pode-se chegar a uma maior precisão no diagnóstico da diabetes tipo 2.

Para utilizar essa nova base de dados, as RNAs precisam ser adaptadas, fornecendo, assim, um treinamento mais voltado para a realidade brasileira, permitindo aumentar a precisão do teste das amostras coletadas na população.

Referências

- BAIER, L. J.; HANSON, R. L. (2004) **Genetic studies of the etiology of type 2 diabetes in Pima Indians**. *Diabetes*, 53, 1181–1186. Disponível em: <diabetes.diabetesjournals.org/content/53/5/1181>. Acesso em: Janeiro de 2015.
- BOTELHO, J. M. A. (2003) **Aspectos Clínicos e Epidemiológicos do Diabetes Mellitus**: Coordenadoria de Doenças e Agravos Não Transmissíveis - Superintendência de Epidemiologia/SES SUS – MG.
- FAUSETT, L. (1994) **Fundamentals of Neural Networks**: architectures, algorithms and applications. Prentice Hall International.
- GETBOOTSTRAP.COM. (2015) **Bootstrap is the most popular HTML, CSS and JS framework for developing responsive, mobile first projects on the web**. Disponível em: <<http://getbootstrap.com>>. Acesso em 03 jun. 2015.
- GONÇALVES, E. (2006) **Dominando NetBeans**. Rio de Janeiro: Ciência Moderna.

- HAYKIN, S. (2001) **Redes Neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman.
- KAYAER, K.; YILDIRIM T. (2003) Medical diagnosis on Pima Indian diabetes using general regression neural networks. In: **Proceedings of the International Conference on Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP)** (pp. 181–184).
- KARAN, O. (2012) **Diagnosing diabetes using neural networks on small mobile devices**. Expert Systems with Applications. Istanbul, Turquia. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417411009468>>. Acesso em: Abril de 2015.
- LORENZI, F.; SILVEIRA, S. R. (2011) **Desenvolvimento de Sistemas de Informação Inteligentes**. Porto Alegre: UniRitter.
- MASSAD, E; MENESES, R. X.; SILVEIRA, P. S. S; ORTEGA, N. R. S. (2004) **Métodos Quantitativos em Medicina**. Barueri, SP: Manole.
- MAZZAFERRI, E. L. (1988) **Endocrinologia**. Columbus, Ohio: Guanabara.
- MYSQL.COM. (2015) **MySQL Workbench**. Disponível em: <<https://www.mysql.com/products/workbench/>>. Acesso em: 17 de junho 2015.
- MENEZES, A. C. (2013) **Um modelo híbrido para apoio ao diagnóstico de diabetes**. Dissertação (Mestrado) – Universidade de Fortaleza. Disponível em: <http://www2.unifor.br/tede//tde_busca/arquivo.php?codArquivo=907681>. Acesso em: maio de 2015.
- PALMA NETO, L. G.; NICOLETTI, M. C. (2005) **Introdução às Redes Neurais Construtivas**. São Carlos, SP: Editora da Universidade Federal de São Carlos.
- ORTIZ M. C. A.; ZANETTI M. L. (2001) **Levantamento dos fatores de risco para diabetes mellitus tipo 2 em uma instituição de ensino superior**. Revista Latino Americana de Enfermagem Maio 2001; 9(3):58-63.
- OLIVEIRA, J. E. P.; VENCIO, S. (2014) **Diretrizes da Sociedade Brasileira de Diabetes**. São Paulo: AC Farmacêutica.
- REGAZZI, J. A.; SILVA, O. H. C. (2010) Teste para verificar a igualdade de parâmetros e a identidade de modelos de regressão não-linear em dados de experimento com delineamento em blocos casualizados. **Revista Ceres**, v. 57. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-737X2010000300005>. Acesso em junho de 2015.
- RIGOTTO, R. M. (2014) **Uso de agrotóxicos no Brasil e problemas para a saúde pública**. Cadernos de Saúde Pública, Rio de Janeiro, 30(7):1-3, jul, 2014.
- RUSSELL, S.; NORVIG, P. (2004) **Inteligência Artificial**. Rio de Janeiro: Elsevier.

SÁNCHEZ, I. (2014) **Mais fundo no design responsivo**. Imasters, 13/01/2014. Disponível em: <imasters.com.br/design-ux/design-responsivo/mais-a-fundo-no-design-responsivo/?trace=1519021197&source=admin>. Acesso em: 10 de abril 2015.

SILVA, M. S. (2014) **Web Design Responsivo**. São Paulo: Novatec.

SILVA, M. S. (2011) **HTML5: A Linguagem de Marcação que Revolucionou a WEB**. São Paulo: Novatec.

SUJATHA, V. (2014) **An Intelligent Expert Based System Neural Network For The Diagnosis Of Type2 Diabetes Patients**. School of Advanced Sciences VIT University, India. Disponível em: <<http://ijirae.com/images/downloads/vol1issue2/ACS10106.April14.17.pdf>>. Acesso em: junho 2015.

UCI. University of California. (2010). **Machine Learning and Intelligent System**, School of Information and Computer Science. Disponível em: <<https://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes>>. Acesso em: 17 ago. 2014.

VISUAL-PARADIGM.COM. **What can we help you with? You may find a solution from the following channels**. Disponível em <<http://www.visual-paradigm.com/>>. Acesso em: 10 de jun 2015.