

Utilizando *Reservoir Computing* para realizar diagnóstico precoce de Mal de Alzheimer

Lara Dantas
Escola Politécnica de Pernambuco
Universidade de Pernambuco
Brasil
Email: ldc@ecom.poli.br

Mêuser Valença
Escola Politécnica de Pernambuco
Universidade de Pernambuco
Brasil
Email: meuser@ecom.poli.br

Resumo—A expectativa de vida da população dos países mais desenvolvidos está crescendo a cada dia e, conseqüentemente, há o aumento de várias doenças relacionadas à idade. No Brasil, cerca de 1,2 milhão de pessoas possuem a Doença de Alzheimer (DA), hoje considerada o tipo de demência mais frequente na população brasileira. Apesar de ser uma doença degenerativa e irreversível, se diagnosticada precocemente, medicamentos podem ser administrados a fim de retardar a progressão dos sintomas e uma melhor qualidade de vida para o paciente. Ray et al. e Moscato realizaram estudos com classificadores contidos no software Weka e, a partir de uma base de dados com valores de 120 proteínas sanguíneas, conseguiram classificar se o paciente podia ou não ser diagnosticado com DA com uma taxa de acerto de 91% e 93% respectivamente. Dessa forma, esse trabalho visa utilizar uma nova abordagem conexionista denominada *Reservoir Computing* (RC) para realizar o diagnóstico precoce de um paciente com ou sem DA e também comparar esses resultados com os obtidos através de uma Rede Neural *Multi-Layer Perceptron* (MLP), com algumas técnicas implementadas no WEKA e com os disponíveis na literatura.

I. INTRODUÇÃO

Os países mais desenvolvidos estão passando por uma grande mudança demográfica. Os segmentos mais velhos da população estão crescendo a uma taxa mais rápida, e por isso, há um aumento das doenças relacionadas à idade, principalmente os distúrbios demenciais progressivos. Descrita pela primeira vez pelo psiquiatra Alois Alzheimer em 1907, a Doença de Alzheimer (DA), é, hoje, a causa mais comum de demência da população idosa.

Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE) e a Organização Mundial de Saúde (OMS), há 1,2 milhões de pessoas no Brasil com DA. Acredita-se que apenas 5% dos portadores desenvolveram a doença precocemente, ou seja, antes dos 65 anos de idade. Nos portadores em que a DA iniciou após os 65 anos, estima-se que entre 10% e 30% dos casos iniciaram após os 85 anos [1].

A DA é uma doença degenerativa e irreversível que causa a morte de várias células cerebrais, os neurônios. O paciente que sofre dessa doença possui um cérebro com microscópicas lesões anatomopatológicas, denominadas placas neuríticas, e emaranhados neurofibrilares [2]. Além disso, o cérebro de uma pessoa com Mal de Alzheimer é bem menor que o cérebro de uma pessoa saudável.

Essa doença se desenvolve em cada paciente de forma única, porém, há vários sintomas comuns a todos eles, como

por exemplo, a perda de memória, distúrbios na linguagem, depressão, agressividade, entre outros.

Inicialmente, o paciente perde a memória episódica, ou seja, a memória que contém informações de eventos e suas relações espaço-temporais. Dessa forma, os fatos mais antigos e os que acabaram de acontecer são facilmente esquecidos.

Com o avanço da doença, a memória semântica também é perdida, ou seja, o conhecimento vocabular, regras, símbolos são esquecidos e o paciente começa a perder a sua identidade cultural [3].

Apesar de ser uma doença irreversível, se descoberta em sua fase inicial, medicamentos podem ser administrados a fim de retardar a progressão dos sintomas e prolongar o bem-estar do paciente [4]. Dessa forma, nota-se que é de extrema importância que sejam desenvolvidos mecanismos para a previsão de DA em toda a população.

Na literatura, Ray et al [5]. desenvolveu um estudo utilizando uma base de dados com amostras de 120 proteínas contidas no plasma de diversos pacientes. Ele concluiu em sua pesquisa que uma combinação de 18 das 120 disponíveis possibilitava a realização de um diagnóstico precoce de DA com uma taxa de acerto de 91% utilizando um conjunto de testes com dados de 92 pacientes que foram diagnosticados ou não com DA.

Além disso, ele utilizou também outro conjunto de testes contendo dados de 47 pacientes com diagnóstico de Déficit Benigno da Memória ou *Mild Cognitive Impairment* (MCI). Para esse conjunto, a taxa de acerto foi de 81%. Esses valores foram calculados a partir da média das taxas de acerto encontradas por todos os classificadores utilizados na realização dos testes para ambos os conjuntos [6].

Posteriormente, Moscato realizou um estudo utilizando 20 classificadores diferentes disponíveis no software Weka e criou várias assinaturas com 18, 10, 6 e 5 proteínas todas elas contidas no conjunto descrito por Ray et al. A assinatura de 5 proteínas foi a que obteve o melhor resultado alcançando uma taxa de acerto de 93% utilizando o conjunto de testes de DA e 65% para o conjunto de testes de MCI. Essas taxas de acerto também foram calculadas a partir da média dos valores encontrados pelos 20 classificadores utilizados [6].

Uma vez que os resultados disponíveis na literatura utilizam apenas técnicas de classificação disponibilizadas pelo

software Weka, neste trabalho vamos utilizar uma nova abordagem conexionista denominada *Reservoir Computing* (RC) para realizar a classificação de um paciente, dado o conjunto das 5 e 10 proteínas definidas por Moscato.

Afim de realizar uma comparação com o resultado obtido utilizando o RC, foi utilizada uma rede neural *Multi-Layer Perceptron* (MLP), ou rede neural de múltiplas camadas, topologia essa amplamente utilizada e com um bom desempenho para problemas de classificação. Além da MLP, e com o objetivo de possuir mais resultados para comparação e conclusão da melhor forma de solucionar o problema abordado, também foram usadas os seguintes classificadores implementados no software WEKA [7]:

- *Bayes Net*: Representa a relação probabilística entre atributos.
- *Sequential Minimal Optimization* (SMO): Eficiente algoritmo para resolver problemas de otimização a partir do treinamento das Máquinas de Vetores de Suporte.
- *Multiclass Classifier*: Algoritmo padrão para classificação de atributos em pelo menos duas classes distintas.
- *Naive Bayes*: Classificador probabilístico baseado no Teorema de Bayes.

II. Reservoir Computing

As Redes Neurais Recorrentes (RNR) foram criadas a fim de possibilitar a solução de problemas dinâmicos. Isto é obtido através de uma realimentação de um neurônio de uma camada i para um que se encontra em alguma camada anterior, $i-j$. Essa topologia de rede neural apresenta uma melhor semelhança com o funcionamento e comportamento do cérebro humano [8].

Em 2001, surgiu uma nova abordagem para o projeto de treinamento de uma RNR proposta por Wolfgang Mass denominada *Liquid State Machine* (LSM) [9]. Ao mesmo tempo, porém independentemente, a mesma abordagem foi descrita por Herbert Jaeger e foi denominada de *Echo State Networks* (ESN). [10].

Ambas as redes ESN e LSM possuem a Propriedade do Eco, do inglês *Echo State Property* (ESP) [11], ou seja, devido às conexões recorrentes da rede, as informações das entradas anteriores são armazenadas. Porém, esses dados não ficam guardados por um período de tempo infinito, assim como o cérebro humano, as informações mais antigas devem ser esquecidas com o tempo. Dessa forma, a rede neural possui um rico conjunto de informações do presente e do passado reforçando assim, a sua aplicabilidade para sistemas dinâmicos [12].

Em 2007, Verstraeten criou o termo *Reservoir Computing* (RC) que unificava os conceitos descritos na LSM e ESN. Desde então, esse termo passou a ser utilizado na literatura para exemplificar sistemas de aprendizagem que são representados por uma rede neural recorrente dinâmica [13].

O RC é composto por três partes: uma camada de entrada que, assim como a MLP, representa as variáveis de entrada

do problema, um *reservoir*, que pode ser entendido como uma grande RNR distribuída e dinâmica com pesos fixos, e uma camada de saída linear denominada *readout*.

A Figura 1 representa a topologia RC com dois neurônios na camada de entrada, três no *reservoir*, e um neurônio na camada de saída.

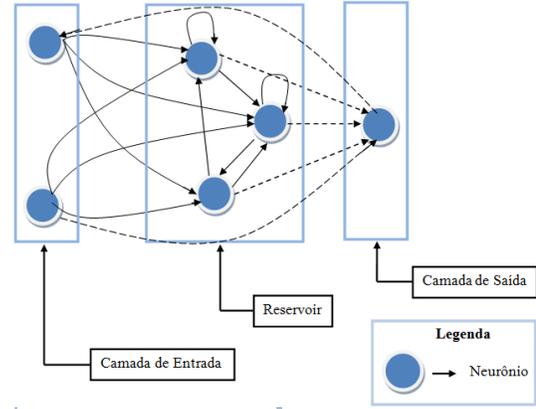


Figura 1. Arquitetura de um *Reservoir Computing*. As linhas tracejadas representam quais os pesos que deverão ser ajustados durante o treinamento da rede.

A. Construção e Simulação do RC

Primeiramente deve-se definir a arquitetura do RC, ou seja, quantos neurônios serão utilizados nas camadas de entrada e saída, além de vários parâmetros do RC que também devem ser fixados. Por ser uma metodologia recente, não há estudos que comprovem quantos neurônios no *reservoir* são necessários para que a rede neural tenha um melhor desempenho, nem qual a taxa de conectividade entre esses neurônios. Portanto, para esse trabalho, foram definidos valores baseados em alguns testes empíricos realizados.

Determinada a arquitetura da rede neural, o próximo passo é gerar as matrizes de pesos que conectam a camada de entrada ao *reservoir*, W_{in} , e a matriz com os pesos entre os neurônios do próprio *reservoir*, W_{res} , ambas com valores aleatórios gerados entre -1 e 1.

Pesquisas afirmam que a matriz W_{res} deve ter um raio espectral igual a 1 para proporcionar uma maior estabilidade numérica [14], ou seja, ao inicializar a W_{res} , ela deve ter seus valores alterados da seguinte forma:

- Inicialmente ela deverá ser decomposta em valores singulares;
- Em seguida, W_{res} deverá ter seus valores alterados até que o máximo valor da diagonal principal da matriz de autovalores seja menor ou igual a 1.

Para realizar a simulação do RC, a base de dados a ser utilizada foi dividida em 3 conjuntos: treinamento, utilizado para realizar a atualização dos estados dos neurônios do *reservoir*, validação cruzada, utilizada para parar o treinamento da rede neural, e o conjunto de testes, utilizado para calcular a taxa de classificação do RC [15].

Os estados dos neurônios do *reservoir* devem ser inicializados com o valor zero. Já que se trata de uma rede recorrente e o RC armazena em uma matriz todos os seus estados (M_{est}), é necessário que os valores finais encontrados pela rede não sejam tão influenciados por essa inicialização. Dessa forma, a literatura sugere que antes de iniciar o treinamento, um conjunto de ciclos de *warm up* sejam executados a fim de realizar atualizações de estados dos neurônios do *reservoir* e desprezar a influência do valor inicial [13]. Os estados são atualizados de acordo com a equação 1:

$$x[k+1] = f(W_{res}x[k] + W_{in}u[k]) \quad (1)$$

Onde, $W_{in}u[k]$ representa a matriz contendo o resultado do produto dos valores oriundos da camada de entrada pelos pesos que conectam esses neurônios ao *reservoir* em um instante k e $W_{res}x[k]$ representa a matriz com os estados dos neurônios do RC no mesmo instante k . O resultado que será atribuído a $x[k+1]$, ou seja, o estado do neurônio do RC em um instante a frente, será o resultado do cálculo da função de ativação do neurônio a partir do somatório das duas parcelas descritas acima. No caso deste trabalho, a função de ativação utilizada foi a Tangente Hiperbólica de acordo com a equação 2.

$$f(net_i) = \frac{e^{net_i} - e^{-net_i}}{e^{net_i} + e^{-net_i}} \quad (2)$$

Finalizado o período de *warm up*, o treinamento do RC pode ser inicializado. O primeiro passo deve ser carregar o conjunto de treinamento e realizar a atualização dos estados do *reservoir*, ressaltando que as matrizes W_{in} e W_{res} não devem ser alteradas. Elas são geradas aleatoriamente durante a construção do RC, conforme foi descrito na seção anterior, e não devem ser ajustadas.

Ainda no período de treinamento, a matriz de pesos que conecta os neurônios da camada de entrada aos de saída e (W_{inout}) do *reservoir* à saída (W_{out}) deve ser calculada através da pseudo-inversa de Moore-Penrose. Como se tratam de uma matrizes não quadradas e seus determinantes podem se aproximar de zero, se faz necessário o cálculo da pseudo-inversa.

Ao fim de cada ciclo de treinamento, deve ser iniciado um ciclo de validação cruzada. Esse processo deve se repetir até que os critérios de parada definidos sejam atingidos e o treinamento seja finalizado. Durante o processo de validação cruzada, as matrizes W_{inout} e W_{out} devem continuar sendo reajustadas.

Finalizado o processo de treinamento, é iniciada a etapa de testes. O conjunto de testes é apresentado ao RC e, nesse momento, todas as matrizes de pesos, W_{in} , W_{res} , W_{inout} e W_{out} , devem permanecer inalteradas, assim como a matriz M_{est} . Nesse momento, é calculado o erro de classificação. Esses valores serão utilizados no futuro para realizar as comparações necessárias.

O comportamento do RC pode ser melhor visualizado através do pseudocódigo descrito no Algoritmo 1:

Algoritmo 1: Pseudocódigo do RC

```

1 Definir a quantidade de neurônios na camada de
  entrada;
2 Definir a quantidade de neurônios no reservoir;
3 Definir a quantidade de neurônios na camada de saída;
4 Gerar aleatoriamente os pesos da matriz  $W_{in}$  entre -1 e
  1;
5 Gerar aleatoriamente os pesos da matriz  $W_{res}$  entre -1 e
  1;
6 Normalizar os pesos de  $W_{res}$  para que o raio espectral
  da matriz fique menor ou igual a 1;
7 while até o final da quantidade de ciclos de warm up
  do
8   | atualiza os estados dos neurônios do RC;
9 end
10 while até que o critério de parada seja atingido do
11   for cada valor do conjunto de entrada do
12     | atualiza os estados dos neurônios do RC;
13   end
14   Calcula a pseudo-inversa de Moore-Penrose para
     encontrar a matriz de pesos que conecta o RC à
     camada de saída;
15   Calcula a pseudo-inversa de Moore-Penrose para
     encontrar a matriz de pesos que conecta a camada
     de entrada à de saída;
16   for cada valor do conjunto de validação cruzada do
17     | atualiza os estados dos neurônios do RC;
18   end
19   Calcula os valores de saída do RC;
20   Calcula o EMQ;
21   Verifica se o critério de parada já foi atingido;
22 end
23 for cada valor do conjunto de testes do
24   | atualiza os estados dos neurônios do RC;
25 end
26 Calcula os valores de saída do RC;
27 Calcula a taxa de acerto;

```

III. METODOLOGIA

A. Bases de Dados

A base de dados utilizada no desenvolvimento deste trabalho foi a mesma exposta por Moscato em sua publicação. Ela contém os valores de 120 proteínas encontrados através da análise de amostras sanguíneas de diversos pacientes. O objetivo final da base é classificar se um paciente pode ser diagnosticado ou não com DA ou MCI [6].

Em seu trabalho, Moscato subdividiu a base de dados em 3 conjuntos. O primeiro deles continha 83 amostras sanguíneas com os valores das 120 proteínas. Dessa quantidade de amostras, 68 delas foram reservadas para o treinamento dos classificadores utilizados. As 15 restantes foram destinadas para o processo de validação cruzada, critério de parada utilizado para finalizar o treinamento das técnicas de classificação [16].

O segundo conjunto da base de dados continha informações de 92 pacientes diagnosticados com ou sem DA. Já o terceiro,

continha amostras sanguíneas de 47 pacientes que apresentavam ou não MCI.

A Tabela I exibe quais das 120 proteínas estavam contidas nas assinaturas definidas por Moscato e que serão utilizadas neste trabalho.

Tabela I. REPRESENTAÇÃO DAS PROTEÍNAS CONTIDAS EM CADA UMA DAS ASSINATURAS UTILIZADA.

Quantidade de proteínas	Proteínas
5	IL-1a, IL-3, EGF, TNF-a and G-CSF
10	CCL7/MCP-3, CCL15/MIP-1d, EGF, G-CSF, IL-1a, IL-3, IL-6, IL-11, PDGF-BB, TNF-a

Neste trabalho, foram preparadas 4 bases de dados para reproduzir os experimentos descritos por Moscato, utilizando a MLP, o RC e os classificadores do WEKA previamente citados. Foram elas:

- 1 base para testar a assinatura de 5 proteínas com o conjunto de testes DA;
- 1 base para testar a assinatura de 5 proteínas com o conjunto de testes MCI;
- 1 base para testar a assinatura de 10 proteínas com o conjunto de testes DA;
- 1 base para testar a assinatura de 10 proteínas com o conjunto de testes MCI;

Todas as bases descritas acima mantiveram a organização utilizada por Moscato no que concerne a divisão de valores para o conjunto de treinamento, validação cruzada e testes.

1) *Pré-processamento dos dados*: Para realizar o treinamento da rede neural se faz necessário que os seus dados estejam normalizados, ou seja, os valores que serão inseridos na rede neural devem estar contidos em um mesmo intervalo numérico. Isso é importante visto que valores muito diferentes podem influenciar no treinamento e causar uma perda na capacidade de generalização da rede neural [16].

Uma das técnicas de normalização mais utilizadas na literatura é a transformação linear e foi ela a escolhida para este trabalho. A Equação 3 representa a fórmula utilizada para normalizar os valores da base de dados.

$$y = ((b - a) \times \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right)) + a \quad (3)$$

Na Equação 3, a e b representam os valores máximos e mínimos que os dados devem assumir. Neste trabalho, foi utilizado o valor de -0,85 para a e 0,85 para b, já que a função de ativação escolhida para essa rede neural foi a Tangente Hiperbólica. Portanto, os valores contidos na base de dados devem estar entre -1 e 1.

B. Simulações

Conforme descrito anteriormente, a fim de comparar os resultados obtidos com o RC e os classificadores foi utilizada uma rede neural MLP. A Tabela II mostra quais os parâmetros escolhidos para realizar as simulações com o RC e MLP. Eles foram obtidos através de testes empíricos e as configurações que apresentaram os menores erros de validação cruzada foram escolhidas.

Tabela II. REPRESENTAÇÃO DOS PARÂMETROS UTILIZADOS PARA AS SIMULAÇÕES COM O RC E MLP

Parâmetro	Valor para o RC	Valor para a MLP
Conectividade do RC	20%	Não se aplica
Quantidade de neurônios na camada de entrada	5 ou 10 (a depender da assinatura de proteínas que estiver sendo utilizada)	5 ou 10 (a depender da assinatura de proteínas que estiver sendo utilizada)
Quantidade de neurônios no RC	4	Não se aplica.
Quantidade de neurônios na camada escondida	Não se aplica	20
Quantidade de neurônios na camada de saída	2	2
Quantidade de ciclos de warm up	100	Não se aplica.
Função de ativação dos neurônios do <i>reservoir</i> ou da camada escondida	Tangente Hiperbólica	Tangente Hiperbólica
Função de ativação dos neurônios da camada de saída	Linear	Linear
Taxa de aprendizagem	Não se aplica.	0,7
Taxa de momento	Não se aplica.	0,4

Após definir as configurações do RC e da MLP, foram realizadas 30 simulações com cada uma das bases de dados e em cada uma das topologias de rede neural escolhidas neste trabalho. Esse número é considerado ideal para realizar comparações estatísticas mais significantes [17].

Além disso, foram realizadas também a mesma quantidade de simulações com cada um dos classificadores escolhidos na ferramenta WEKA. Contudo, optou-se por apenas realizar uma análise estatística nos resultados encontrados pelo RC e pela MLP.

C. Análise Estatística

Finalizadas todas as simulações, fez-se necessário realizar uma sequência de testes estatísticos com o objetivo de validar cientificamente os resultados encontrados. Para isso, utilizou-se o software matemático R, já que ele contém todas as implementações dos testes que foram utilizados. Esse software utiliza como padrão um nível de significância (α) previamente definido no valor de 0,05.

Antes de utilizar um teste paramétrico em um conjunto de dados é necessário verificar se as amostras são normalmente distribuídas e se elas possuem variâncias estatisticamente iguais. Caso essas duas premissas sejam validadas, pode-se aplicar um teste paramétrico, caso contrário, deverá ser utilizado um teste não paramétrico.

Assim, foi aplicado o Teste de Shapiro-Wilk para verificar se as amostras eram normalmente distribuídas e o Teste F para verificar se as amostras eram oriundas de uma mesma população, ou seja, se suas variâncias eram estatisticamente iguais.

Como nenhuma das 4 bases atendeu a essas duas premissas ao mesmo tempo, não foi possível realizar o Teste T-Student. Dessa forma, foi escolhido o Teste da Soma dos Postos de Wilcoxon, já que ele é um teste não paramétrico, ou seja, não faz suposições sobre a distribuição de probabilidade das amostras.

IV. RESULTADOS

Após serem realizadas todas as simulações com as bases de dados foi calculada a média aritmética para cada conjunto de simulações e a Tabela III exibe esses valores encontrados.

Tabela III. REPRESENTAÇÃO DA MÉDIA DAS TAXAS DE ACERTO APÓS OS 30 EXPERIMENTOS

Base de Dados	Média da taxa de acerto com RC	Média da taxa de acerto com MLP
Assinatura de 5 proteínas e conjunto de testes AD	94,34%	95,61%
Assinatura de 10 proteínas e conjunto de testes AD	86,62%	93,44%
Assinatura de 5 proteínas e conjunto de testes MCI	66,59%	69,14%
Assinatura de 10 proteínas e conjunto de testes MCI	69,29%	68,15%

Ao aplicar o teste Soma dos Postos de Wilcoxon para cada um dos 4 casos, nos 3 primeiros o resultado encontrado foi que a MLP possui um desempenho estatisticamente melhor do que o RC. Já para o último caso, ou seja, com a assinatura de 10 proteínas e conjunto de testes MCI, o teste indicou que não há diferenças estatísticas entre os resultados encontrados com as duas técnicas.

A Tabela IV resume os valores máximos encontrados com as simulações do RC, a média dessas simulações e os resultados obtidos no trabalho de Moscato [6].

Tabela IV. COMPARAÇÕES DOS RESULTADOS OBTIDOS COM O RC E COM OS DISPONÍVEIS NA LITERATURA.

Base de Dados	Valor máximo	Média	Resultado encontrado por Moscato
5 proteínas e conjunto de testes DA	96,73%	94,34%	93%
10 proteínas e conjunto de testes DA	91,30%	86,92%	89%
5 proteínas e conjunto de testes MCI	70,21%	66,59%	65%
10 proteínas e conjunto de testes MCI	74,46%	69,29%	66%

A partir da Tabela IV, pode-se concluir que O RC obteve resultados coerentes com os descritos por Moscato além de ter conseguido alcançar um valor máximo maior do que a média encontrada na literatura.

Ainda como resultado deste trabalho, a Tabela V exibe a comparação dos resultados encontrados após as simulações do RC e dos quatro classificadores utilizados.

V. CONCLUSÃO

A Doença de Alzheimer é hoje uma das doenças mais frequentes na população idosa mundial. Nos últimos anos, o número de pacientes cresceu bastante, visto que a expectativa de vida na maioria dos países desenvolvidos tem aumentado.

A DA é uma doença degenerativa, ou seja, as células cerebrais vão se deteriorando e não há como reverter o quadro da doença. Porém, quanto mais cedo os medicamentos forem administrados, melhor será a qualidade de vida do paciente, já que a medicação irá retardar a progressão dos sintomas.

Tabela V. COMPARAÇÕES ENTRE AS MÉDIAS DOS RESULTADOS OBTIDOS COM O RC E COM OS CLASSIFICADORES DISPONÍVEIS NO WEKA.

Base de Dados	RC	SMO	Bayes Net	Naive Bayes	Multiclass Classifier
5 proteínas e conjunto de testes DA	94,34%	95,65%	93,47%	96,73%	93,47%
10 proteínas e conjunto de testes DA	86,92%	91,3%	90,21%	84,78%	89,13%
5 proteínas e conjunto de testes MCI	66,59%	68,08%	68,08%	65,95%	70,21%
10 proteínas e conjunto de testes MCI	69,29%	68,08%	59,57%	63,82%	70,21%

Dessa forma, este trabalho teve como objetivo verificar o desempenho de uma nova abordagem conexionista de rede neural denominada *Reservoir Computing* para classificar, precocemente, se um paciente pode ser diagnosticado ou não com DA. Além disso, objetivou-se também realizar a comparação do desempenho do RC com a rede neural MLP, com alguns classificadores disponíveis no WEKA, e também com os resultados disponíveis na literatura.

A partir das simulações e testes estatísticos, pode-se concluir que a MLP apresentou desempenho superior na maioria dos casos, apesar de o RC ter obtido valores máximos superiores à MLP e aos disponíveis na literatura. Isso pode ser explicado pelo fato de o RC ser mais adequado para o uso em sistemas dinâmicos já que ele apresenta uma boa capacidade de memorização.

Pode-se concluir também que os resultados encontrados pelo RC se aproximaram dos encontrados nas simulações com o *SMO*, *Bayes Net*, *Multiclass Classifier* e *Naive Bayes*.

Diferentemente do RC, a MLP apresenta uma melhor capacidade para aproximar funções não lineares e não contém a propriedade de armazenar os estados anteriores dos seus neurônios, ou seja, não possui recorrências.

Como trabalhos futuros, pretende-se realizar um estudo comparativo do RC entre a capacidade de aproximação não linear e a capacidade de memorização da rede, a fim de avaliar os parâmetros adequados para obtenção de melhores resultados.

Será realizado, também, um estudo aprofundado na parametrização do RC, como por exemplo, a definição do tamanho do raio espectral, quantos neurônios se devem colocar no *reservoir*, qual o grau de conectividade entre esses neurônios. Além disso, a fim de abordar problemas dinâmicos, será implementada a recorrência entre neurônios da camada de saída e desta com o reservatório.

REFERÊNCIAS

- [1] P. Saúde. (2012) Portal da saúde brasileira. Acessado em 15 de Julho de 2012. [Online]. Available: <http://portal.saude.gov.br/saude/>
- [2] R. Green, *Diagnóstico e Tratamento da Doença de Alzheimer e Outras Demências*. Editora de Publicações Científicas Ltda., 2001.
- [3] d. B. P. Marques, Roberto, *Demência tipo Alzheimer: diagnóstico, tratamento e aspectos sociais*. Editora de Publicações Científicas Ltda., 1997.
- [4] E. Giusti and V. Surdo, *Alzheimer, cuidados e aconselhamento familiar: necessidades e tratamento psicológico da demência*. Gryphus, 2010.

- [5] Herbert. C. Takeda-Uchimura. Y. Boxer A. Ray S. Britschgi M, "Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins" *Nat Med*, vol. v13, p. 1359 - 1362p, 2007.
- [6] M. Gómez, Ravetti and P. Moscato, "Identification of a 5-protein biomarker molecular signature for predicting alzheimer's disease," *PLoS One*, vol. v3, p. 12p, 2008.
- [7] I. Witten and E. Frank, *Data Mining: Pratical Machine Learning Tools and Techniques*. Elsevier, 2005.
- [8] M. Valença, *Aplicando Redes Neurais: Um Guia Completo*, 1st ed. Livro Rápido, 2005.
- [9] W. Maass, *Motivation, theory, and applications of liquid state machines*. Imperial College Press, 2011.
- [10] H. Jaeger, "The echo state approach to analysing and training recurrent neural networks," German National Resource Center for Information Technology, Tech. Rep., 2010.
- [11] M. Massar and S. Massar, "Mean-field theory of echo state networks," *PHYSICAL REVIEW E*, vol. 87, 2013.
- [12] M. Lukoševičius, *A practical guide to applying echo state networks*. Springer Berlin Heidelberg, 2012, pp. 659–686.
- [13] D. Verstraeten, "Reservoir computing : computation with dynamical systems," Ph.D. dissertation, Ghent University. Faculty of Engineering, Ghent, Belgium, 2009.
- [14] A. Ferreira, Araújo, "Um método para design e treinamento de reservoir computing aplicado à previsão de séries temporais," Ph.D. dissertation, Universidade de Federal de Pernambuco, 2001.
- [15] M. Kulkarni and C. Teuscher, "Memristor-based reservoir computing," in *Proceedings of the IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH'12)*, 2012.
- [16] M. Valença, *Fundamentos das Redes Neurais: exemplos em JAVA*, 2nd ed. Livro Rápido, 2009.
- [17] N. Juristo and A. Moreno, M., *Basics of Software Engineering Experimentation*. Kluwer Academic Publishers, 2001.