colloquium exactarum

USO DO APRENDIZADO DE MÁQUINA NO DIAGNÓSTICO MÉDICO DE PATOLOGIAS

USE OF MACHINE LEARNING IN THE MEDICAL DIAGNOSIS OF PATHOLOGIES

Aline Miki Takakura[1]; Danillo Roberto Pereira[1]; Francisco Assis da Silva[1]; Mario Augusto Pazoti[1]; Leandro Luiz de Almeida[1]; Helton Molina Sapia[1]

[1]Universidade do Oeste Paulista – UNOESTE, Faculdade de Informática – FIPP, Presidente Prudente, SP. E-mail: fipp@fipp.unoeste.br

RESUMO – A tecnologia se desenvolve rapidamente em diferentes áreas do conhecimento, uma destas áreas é a medicina. Este trabalho apresenta uma proposta de automatização de diagnóstico de patologias utilizando técnicas e métodos classificadores de Aprendizagem de Máquina (AM). Através dos métodos que serão explanados e implementados, as informações contidas nas bases de dados serão analisadas e classificadas, gerando resultados. Com a utilização dessas técnicas, através de máquinas para o diagnóstico médico, patologias poderão ser detectadas num estágio menos avançado da doença e com maior precisão, quando comparado a diagnóstico inteiramente humano sofre influências de fatores externos, o que pode afetar no diagnóstico do paciente. Neste trabalho busca-se analisar dados e classifica-los de acordo com os métodos à serem citados, e por fim pode ser definido o método aplicado mais viável e eficaz.

Palavras-chave: aprendizagem de máquina; diagnóstico; patologias; classificadores; base de dados.

ABSTRACT – Technology develops rapidly in different areas of knowledge, one of these areas is medicine. This paper presents a proposal for the automation of diagnosis of pathologies using techniques and methods of Machine Learning (AM) classification.
Through the methods that will be explained and implemented, the information contained in the databases will be analyzed and classified, generating results. With the use of techniques for medical diagnosis, pathologies are detected at the less advanced stage of the disease and more accurately, when compared to a fully human diagnosis it is influenced by external individuals, which can affect any diagnosis of the patient. This work seeks to analyze data and classify according to the methods to be cited, and finally can be defined the most feasible and effective applied method.

Keywords: machine learning; diagnosis; pathologies; classifiers; database.

Recebido em:03/04/2017
Revisado em: 20/08/2017
Aprovado em: 04/10/2017

1. INTRODUÇÃO

O Aprendizado de Máquina (AM) é uma subárea da Inteligência Artificial (IA), que estuda o desenvolvimento de métodos capazes de extrair informações e conhecimento a partir de amostras de dados. São utilizados uma vasta gama de algoritmos diferentes com a capacidade de classificar conjuntos de elementos. Por classificação entende-se o processo de atribuir a um determinado dado, o rótulo da classe à qual ele pertence. Neste sentido, as técnicas de AM são empregadas na indução, a partir de um conjunto de exemplos de um classificador, que deve ser capaz de prever a classe de novos dados quaisquer do domínio em que ele foi treinado (ESTEVES et al., 2009).

As técnicas de Aprendizado de Máquina empregam um princípio de inferência denominado indução, o aprendizado indutivo que pode ser dividido em duas formas básicas de AM: Aprendizado Supervisionado e Não Supervisionado (LORENA et al., 2007).

No aprendizado supervisionado um agente externo informa um conjunto padrão de entradas que contém suas respectivas saídas, a resposta fornecida pelo método é então verificada e comparada à resposta esperada. Se houver qualquer erro o método utilizado irá fazer reajustes para evitar discrepâncias.

Ao contrário, no aprendizado não supervisionado, não é necessário um agente externo para comparação de resultados finais. O método utilizado processa as entradas e analisa suas regularidades e características de modo a definir um padrão e os classifica automaticamente (FERNEDA, 2006).

A área de aprendizagem de máquina possui métodos que se baseiam em um pequeno conjunto de dados para a detecção de um padrão, possibilitando a classificação dos dados. Essas abordagens podem ser utilizadas para prever diagnósticos e consequentemente mitigar os erros médicos (BENNET et al., 2013). Assim, permite-se precisão no diagnóstico de uma determinada doença, e em qual estagio ela se encontra, o que possibilita a realização de um tratamento mais adequado (NEILL, 2013).

Diversas doenças patológicas dependem de um diagnóstico médico, para detectar e identificar seu estágio de desenvolvimento. Com o aprendizado de máquina aplicado ao diagnóstico patológico pode-se identificar a presença de uma patologia através dos métodos classificadores de dados.

O ser humano é amplamente influenciado por fatores externos, decorrente disso, pode-se afirmar que o médico está sujeito a tais influências, e consequentemente, seus diagnósticos. Dois médicos com mesma formação, expostos a ambientes diferentes, terão diagnósticos com um grau de discrepância para um mesmo caso analisado. Enquanto que um computador devidamente ensinado a diagnosticar, fornecerá o mesmo resultado. Visto que uma máquina realiza a análise sem influência do emocional e com ausência da cognição humana, pautado exclusivamente em algoritmos, assim qualquer erro presente no diagnóstico, será proveniente de falhas na programação.

O diagnóstico automatizado envolve a detecção pelo computador de padrões equivalentes nos dados, que auxilia o médico no diagnóstico de uma patologia. Assim quando diagnosticado algo de errado com o paciente, este deverá realizar um exame especifico. A automação de diagnósticos, propicia a detecção da doença num estágio menos avançado, através de técnicas e métodos computacionais, onde dados serão analisados e classificados, indicando se o paciente contém alguma patologia.

A aplicação das técnicas e métodos de Aprendizado de Máquina nas bases de dados médicas geram resultados que podem ser utilizados para o auxílio de diagnósticos. Assim o diagnóstico médico será realizado com maior precisão e a patologia detectada em menor tempo.

Este documento está organizado em 6 seções. A seção 2 apresenta a declaração dos objetivos deste trabalho. Uma breve revisão bibliográfica é apresentada na seção 3. A seção 4 traz a apresentação de trabalhos relacionados. Na seção 5 são apresentadas as metodologias relativas a este trabalho. Os resultados e conclusões são apresentados na seção 6. E por fim a seção 7 apresenta as referências bibliográficas utilizadas.

2. OBJETIVOS

Através da base de dados de doenças patológicas, os métodos da AM extraem as informações das bases de dados pré-processadas para classifica-las pós analise. Os métodos ou classificadores propostos a ser comparados são: SVM (*Support Vector Machine* ou Máquina de Vetores de Suporte), Redes Neurais (*Neural Network*), OPF (*Optimum-Path Forest*), KNN (*K Nearest Neighbor*) e Bayes.

Este trabalho tem como objetivo realizar o estudo comparativo da aplicação destes métodos no contexto da medicina. Os resultados dos métodos aplicados serão analisados comparativamente, evidenciando o mais eficaz na extração de dados das regiões de interesse da base de dados.

3. REVISÃO BIBLIOGRAFICA

3.1. Métodos de Aprendizado Supervisionados

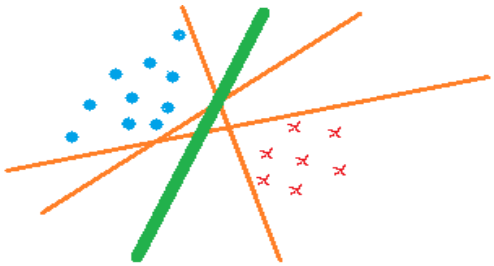
3.1.1. *Support Vector Machine*

As Maquinas de Vetores de Suporte (SVM, do inglês *Support Vector Machine*), é um dos métodos mais utilizados na Aprendizagem de Máquina. Possui uma teoria mais complexa em relação aos outros podendo ser utilizado de diferentes formas para diversas aplicações (JUNIOR, 2010). Essa teoria estabelece conceitos a serem seguidos para a obtenção de classificadores, definindo sua capacidade de prever a classe de novos dados (LORENA et al., 2007).

O SVM constrói um classificador de acordo com um conjunto de padrões identificados por ele.

Figura 1: contém um conjunto de classificadores ou hiperplanos que dividem duas classes, azul (bolinha) e vermelha (x). Os classificadores laranjas são as possibilidades de separação que existe entre as duas classes e o classificador verde representa o hiperplano com margem máxima, que é considerado ótimo.

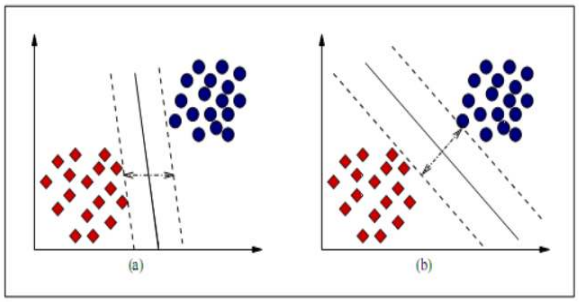
Figura 1. Possíveis hiperplanos de separação (linha laranja e mais fina) e hiperplano ótimo (linha verde e mais grossa), onde se encontra duas classes (bolinha azul e x vermelho).



Fonte: Autor

Figura 2: é mostrado dois hiperplanos que separam duas classes (azul e vermelha). O hiperplano (a) apresenta uma margem pequena de divisão de classes, ao contrário do hiperplano (b) que mostra uma margem máxima, considerada ótima.

Figura 2. (a) Hiperplano com margem pequena. (b) Hiperplano com margem máxima.



Fonte: (JUNIOR, 2010).

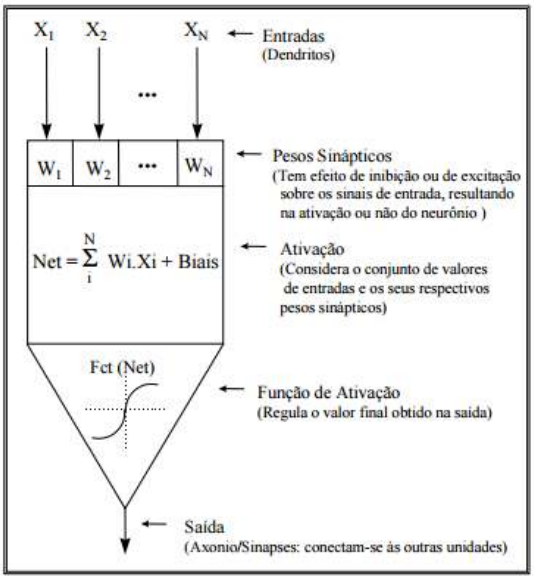
Funcionamento: sobre um conjunto de dados pertencentes as classes, o SVM define um hiperplano que separe esses dados de forma que a maioria dos dados de uma mesma classe estejam do mesmo lado (MARINS, 2008). Ele pode utilizar três tipos de Kernel: RBF, *Sigmoid* e o Linear.

3.1.2. *Neural Network*

As Redes Neurais Artificiais (ANN, do inglês *Artificial Neural Network*) são conhecidas como métodos conexionistas, por se basear no cérebro humano e seu funcionamento. As redes conexionistas são formas de representação do conhecimento, que estão relacionadas a conexão entre os neurônios e na comunicação por meio de ligações. Os conhecimentos da rede são representados por valores numéricos que descrevem as conexões e o comportamento da rede (OSÓRIO et al., 2000).

Neurônio Artificial (Figura 3): Através da lógica e matemática o neurônio artificial faz a simulação do comportamento e funções de um neurônio biológico. Onde os dentritos são as entradas, e as sinapses são ligações chamadas peso, as entradas são então processadas pela função de soma e o limiar de disparo passa a ser uma função de transferência (BRUMATTI, 2005).

Figura 3. Neurônio Artificial.



Fonte: (OSÓRIO et al., 2000).

3.1.3. Classificadores Bayes

Um classificador Bayesiano é uma rede simples Bayesiana que tem a função de classificação. Ele possui um nó C que representa a classe de características e um nó Xi para cada uma delas. A rede

Bayesiana permite a computação da probabilidade P (C = ck|X = x) de uma amostra pertencer a qualquer uma das classes existentes, através de uma instancia x dada, para cada classe ck dela (Sahami et. al., 1998). Ou seja, P (C = ck|X = x) é a probabilidade de P (C = ck) ocorrer contando que a probabilidade de P (X = x) tenha ocorrido de antemão (SILVA, 2006).

$$P(X = x) = \frac{P(C=c)P(C=c_k)}{P(X=x)} \quad (1)$$

As Redes Bayesianas misturam a teoria dos grafos com a teoria da probabilidade, ela é composta de uma parte qualitativa e outra quantitativa (OLIVEIRA et al., 2003).

3.1.4. *Optimun Path Forest*

A classificação baseada em floresta de caminhos ótimos ou como é conhecida OPF (do inglês, *Optimum-Path Forest*), faz a modelagem das amostras utilizando nós em grafos, onde os nós são amostras representadas pelos seus atributos, e os arcos fazem a relação entre as amostras, que podem ter aplicação de diversas funções de custo sobre eles, e consequentemente o grafo pode ser dividido em uma árvore com possíveis caminhos ótimos (QUINTA et al., 2012).

Um caminho da árvore é uma sucessão de dados de amostras n, e a cada caminho n, o custo é dado por uma função f (n). O caminho n é dito ótimo se a função de n for menor ou igual à função de qualquer caminho T, denotado como f (T), no qual n e T terminam na mesma amostra (FREITAS et al., 2010):

$f_{max\ (\langle s \rangle)} = \{0\ se\ s \subset S + \infty\ caso\ contrário$ (2)

$f_{max(\pi.\langle s,t \rangle)} = max\{f_{max}(\pi, d(s,t))\}$ (3)

O OPF faz a interpretação de um grafo completo, considerando as distâncias entre os nós de atributos, e atribui a qualquer caminho o peso máximo do arco ao longo deste. Cedido uma base de dados de treinamento com amostras de várias classes, um conjunto de recursos representa cada amostra (PAPA et al., 2012).

3.1.5. *K-Nearest Neightbors (KNN)*

Para a classificação de um novo elemento, o KNN busca K elementos vizinhos mais próximos a ele. É então verificada a classe dos vizinhos mais próximos ao elemento, a classe com maior quantidade de vizinhos mais próximos é atribuída ao elemento sem classe. A distância dos k-vizinhos pode ser calculada a partir de uma dessas equações: Euclidiana, *Manhattam* e *Minkowski* (SILVA et al., 2005):

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_1 - y_1)^2 + \cdots + (x_n - y_n)^2} \quad (4)$$

$$d(x,y) = |x_1 - y_1| + |x_2 - y_2| + \cdots + |x_n - y_n| \quad (5)$$

$$d(x,y) = (|x_1 - y_1|^q + |x_2 - y_2|^q + \cdots + |x_n - y_n|^q)^{1/q} \quad (6)$$

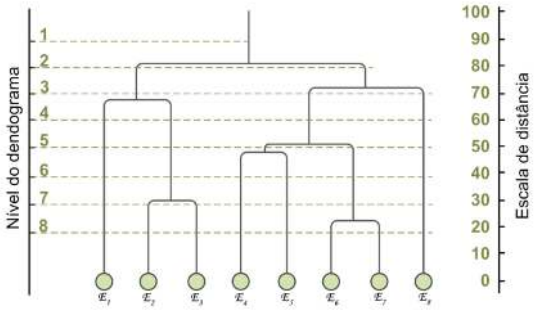
3.2. Métodos de Aprendizado Não-Supervisionados

3.2.1. *Hierarquical Clustering*

Hierachical Clustering realiza o agrupamento de dados em clusters através da variável de distância entre os dados obtidos na base de dados. Através de um método implementado ele calcula a menor distância entre os dados e vai agrupando-os de modo a formar uma árvore hierárquica (METZ, 2006).

Para a implementação deste algoritmo podem ser usados várias técnicas, onde elas se diferenciam nos métodos para identificar os pares de cluster mais similares e a heurística a ser escolhida para otimizar os resultados. Um conjunto de dados contém diversos clusters e estes podem possuir sub-clusters que ainda podem ser formados por outros sub-clusters (agrupamento de clusters pequenos). E para a representação gráfica desses agrupamentos e sub-agrupamentos de clusters é utilizado um Dendrograma (Figura 4), uma representação em uma estrutura de árvore (METZ, 2006).

Figura 4. Dendograma.



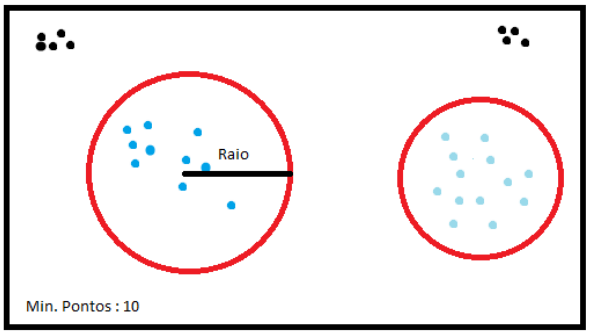
Fonte: (Metz, 2006).

3.2.2. DBSCAN

Do inglês *Density Based Spatial Clustering of Application with Noise* (Clusterização Espacial Baseada em Densidade de Aplicações com Ruído), é um método significativamente efetivo para identificar clusters de formato arbitrário, diferentes tamanhos, identificar e separar os ruídos dos dados e detectar clusters "naturais" e seus arranjos dentro do espaço de dados sem qualquer informação sobre o conjunto de dados. O algoritmo se aplica para espaços euclidianos de duas e três dimensões, ele é aplicável a qualquer base de dados com função de distância para pares de objetos, podendo ser qualquer função de distância a ser utilizada, sendo ela escolhida de acordo com a sua aplicação (ESTER et al., 1996).

Figura 5: existe um raio e um número mínimo de pontos, onde um cluster é formado pelo número mínimo de amostras dentro desse raio, e existem também os ruídos que são as amostras que não puderam ser agrupados por não estar dentro desses dois parâmetros de agrupamento.

Figura 5. Exemplo DBSCAN.



Fonte: Autor

Ruído: Tendo um conjunto de C clusters em uma base de dados D, em relação ao raio e ao número mínimo de pontos, é definido como ruído objetos de D que não estão agrupados em nenhum dos clusteres de C é dito ruído (CASSIANO, 2014).

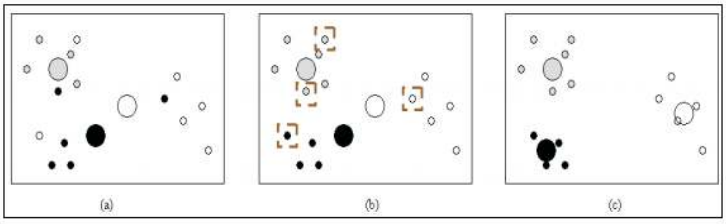
3.2.3. KMEANS

O K-Means é um dos métodos de agrupamento de dados não hierárquico, que tenta minimizar a distância entre os dados de um dataset e um centro escolhido ou dado por um valor aleatório. O algoritmo do K-Means pode ser descrito da forma (LINDEN, 2009):

a. Escolher valores distintos ou gerar valores aleatórios para ser o centro dos grupos.
b. Relaciona cada ponto ao centro mais próximo.
c. Realizar um cálculo para definir os novos centros.
d. Repetir os passos b-c até que não tenha mais pontos mudando de lugar.

O referido algoritmo é extremamente veloz, em poucas iterações ele faz o agrupamento até que se obtenha uma estabilidade, no qual não haja elementos pertencentes a um cluster onde o seu centro não é o mais próximo dele. Figura 6: r pode ser visto um exemplo da execução do algoritmo do K-Means (LINDEN, 2009):

Figura 6. Exemplo de execução do algoritmo de K-Means. (a) Cada elemento foi designado para um dos três grupos aleatoriamente e os seus centros representados pelos círculos maiores na figura, foram calculados para gerar outros centros (b) Os elementos foram direcionados para os grupos cujos centros calculados na etapa (a) estão mais próximos (c) Os centros foram recalculados e os grupos já estão em sua forma final, pois todos os elementos já estão agrupados e se não estivesse as etapas (b) e (c) seriam executados até que que estivessem.



Fonte: (LINDEN, 2009).

4. TRABALHOS RELACIONADOS

4.1 *Câncer*

O câncer é uma doença que pode ocorrer em diferentes partes do corpo, sendo assim existem vários tipos de câncer, onde as células do corpo crescem e se multiplicam descontroladamente. Para a detecção do câncer são utilizados sistemas automatizados de diagnósticos, onde os pacientes com câncer são classificados através de um classificador não linear usando o SVM ou Máquina de Vetores de Suporte (SWEILAM et al., 2010).

4.2 *Nefrite Aguda*

A Nefrite Aguda tem origem pelve renal com ocorrência maior em mulheres, seu primeiro sintoma é uma febre imediata que possui alto grau e calafrios, e dores lombares muito fortes. Os dados recebidos são os sintomas da doença que são utilizados para montar a Rede Neural, que irá executar o diagnóstico da patologia, onde o classificador faz a distinção entre a pessoa infectada e não infectada (AL-SHAYEA, 2011).

4.3 *Parkinson*

Parkinson é uma doença que atualmente afeta muitas pessoas, onde ocorre um pouco a perda de movimentos. É realizado o diagnóstico automatizado através de técnicas baseadas em visão computacional aplicadas em um conjunto de dados de vários pacientes. Esses conjuntos de dados são compostos de escritas de pacientes, que são analisados pelo Processamento de Imagens e métodos de Aprendizagem de Máquina (PEREIRA et al., 2015).

4.4 *Pterígio*

O Pterígio é uma doença ocular, que ainda não possui uma causa exata, podendo levar a cegueira. É aplicado o OPF juntamente com outras técnicas de AM buscando a identificação do Pterígio. O dataset é dividido em duas partes: 50% para treinamento e 50% para a classificação, num ciclo repetido 10 vezes onde a cada ciclo é gerado aleatoriamente conjuntos de treinamento e classificação (PAGNIN, 2011).

4.5 *Clusters de fatores de risco cardiovascular e a ocorrência de hipertensão arterial em adultos*

Muitas pessoas hoje são sedentárias, o que faz com que elas estejam mais propensas a doenças cardiovasculares, obesidade e doenças metabólicas. Sendo assim com uma grande quantidade de amostras de adultos sedentários foram coletados e analisados atributos relacionados

a doenças cardiovasculares. Como resultados foram obtidos que não houve muita diferença entre ambos os sexos (CARVALHO et al., 2013).

5. METODOLOGIA

Para análise e comparação de avanços nos resultados dos dados, através do método k-Fold Validação Cruzada houve a aplicação dos métodos supervisionados sobre os arquivos gerados de cada base de dados, onde o código foi executado 10 vezes sobre o arquivo, gerando assim 10 arquivos decorrente de teste e treinamento de 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% e 90%, calculando a Acurácia, Precisão e Revocação, onde essa variação tem como função a verificação de uma possível otimização nos resultados dos métodos.

O k-Fold Validação Cruzada (ou, do inglês k-Fold Cross Validation), é uma técnica que utiliza todos os dados disponíveis como exemplo de treinamento e teste, fazendo o treinamento repetidamente k vezes (Stone, 1974).

As bases de dados de patologias utilizadas estão disponíveis na UCI - *Machine Learning Repository*, os dados deste acervo tiveram a aplicação dos métodos descritos e assim classificados de acordo com os padrões de cada técnica.

O método utilizado Hold-Out divide o conjunto de dados em dois subconjuntos, um utilizado para o treinamento e o outro para teste (KOHAVI, 1995). O Hold-Out apresenta uma taxa de erro do conjunto de teste que aproxima a taxa de erro verdadeira (BATISTA, 1998).

Foi aplicado também às bases de dados os métodos de Seleção de Atributos, esses métodos fazem uma seleção dos atributos considerados mais relevantes de acordo com o que cada método especifica. Os métodos aplicados foram:

a. *Variance Threshould*: Ele remove todos os atributos que não estão abaixo de um limiar definido, caso o limiar não seja definido, por padrão ele elimina atributos que não tenham variância de valores.
b. KBEST: É definido um valor K, onde são então selecionados K % atributos que possuírem maior pontuação entre todos. Foi utilizado o KBEST com 25%, 50% e 75% dos atributos.

Tabela 1: representa as características de algumas das bases selecionadas da UCI - *Machine Learning Repository*.

Tabela 1. Bases.

Patologias	Classes	Atributos	Instancias
Arrhythmia	16	279	452
Audiology (Standardized)	24	69	226
Breast Cancer Wisconsin (Diagnostic)	2	32	569
Breast Cancer Wisconsin (Original)	2	10	699
Contraceptive Method Choice	3	9	1473
Dermatology	6	33	366
Liver Disorders	2	7	345
Lung Cancer	3	56	32
Pima Indians Diabetes	2	8	768
Post-Operative Patient	3	8	90
Statlog (heart)	2	13	270

Fonte: Autor

6. RESULTADOS E CONCLUSÕES

Este trabalho realizou a análise e comparação dos resultados utilizando métodos de Aprendizagem de Máquina (AM) para o diagnóstico médico de patologias, métodos supervisionados e não supervisionados.

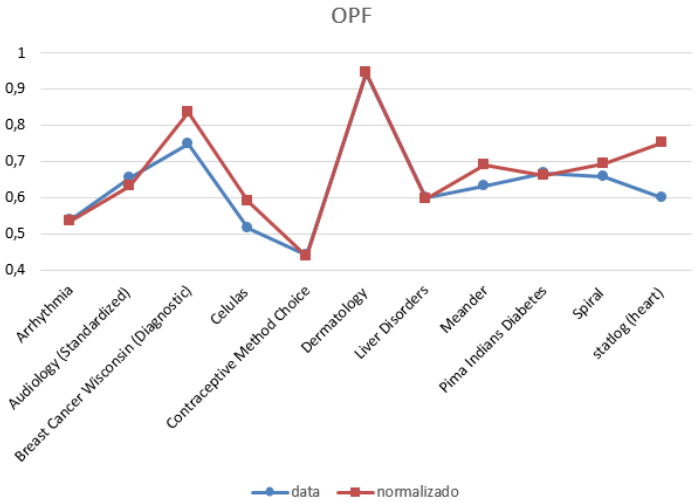
Na tentativa de melhores resultados foram aplicados junto aos métodos de AM o método de clareamento *Whitening*, onde os dados foram pré-processados para um melhor agrupamento, porém não houve diferenças nos resultados. Após essa tentativa, foram criados dados estendidos que são a aplicação dos métodos supervisionados sobre os resultados gerados pelos métodos não supervisionados (combinação de classificação com agrupamento), assim verificou-se resultados com uma leve melhoria, porém não sendo ainda os resultados desejados.

Foram aplicados às bases de dados os métodos de Normalização, Seleção de Atributos e a combinação da Seleção de Atributos com a Normalização. Os métodos de Seleção de Atributos utilizados foram: *Variance Threshould*, KBEST com 25%, 50% e 75% dos atributos da base de dados. A normalização aplicada trouxe um pequeno avanço nos resultados, e a aplicação da Seleção de Atributos e Seleção de Atributos com a Normalização melhorou bastante em relação a somente aplicação da Normalização, porém dos dois métodos de Seleção de Atributos que foram aplicados, o *Variance Threshold* não foi tão bom quanto o KBEST. Da Seleção de Atributos e Seleção de Atributos com Normalização o KBEST 50, KBEST 50 Normalizado e KBEST 75 Normalizado foram os que obtiveram resultados melhores em relação aos outros métodos.

As bases de dados foram treinadas com 9 faixas de treinamento, em que a faixa de 10% na maioria das bases teve um resultado melhor de acurácia e a faixa de 80% e 90 % foi pior devido ao *overfitting*, onde o método se ajusta demasiadamente ao padrão do conjunto de treino da base de dados. Dos métodos aplicados de classificação, analisando os resultados em média, o KNN foi o que alcançou melhores resultados em relação aos outros métodos e o SVM não trazia alterações em relação a classificação com a aplicação das metodologias estudadas.

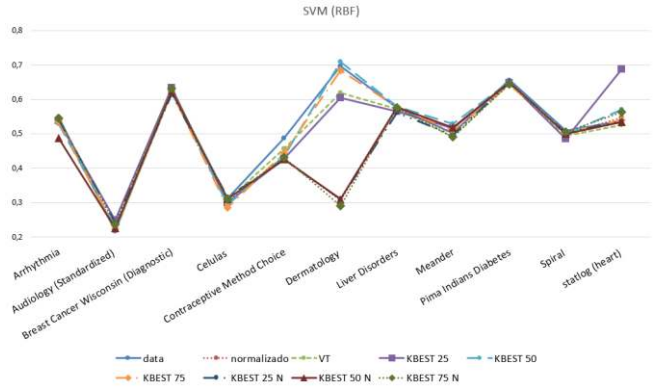
Pode-se verificar que os métodos entre si fazem uma classificação utilizando diversos parâmetros, portanto são geradas algumas classificações diferentes entre eles, onde uma tem mais eficácia ou não que a outra, e as base de dados que foram aplicados esses métodos também pode afetar os resultados.

Gráfico 1. Média OPF.



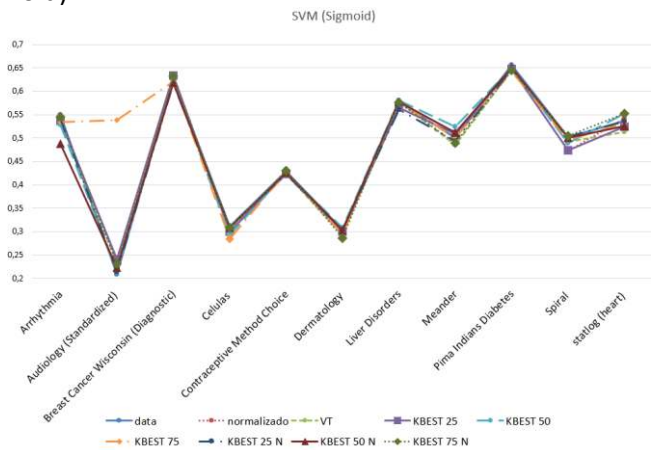
Fonte: Autor

Gráfico 2. Média SVM (RBF).



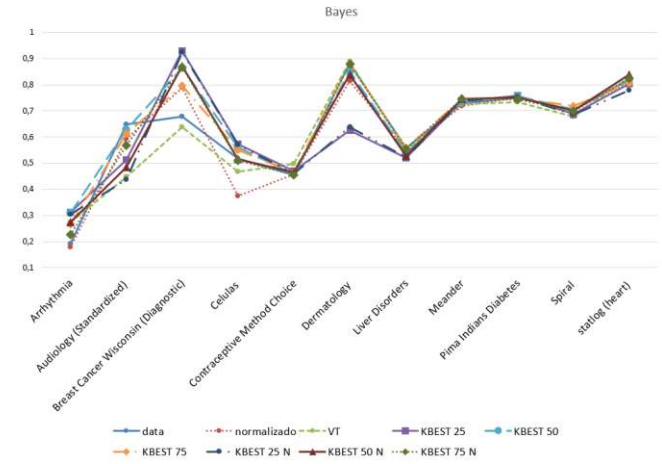
Fonte: Autor

Gráfico 3. Média SVM (Sigmoid).



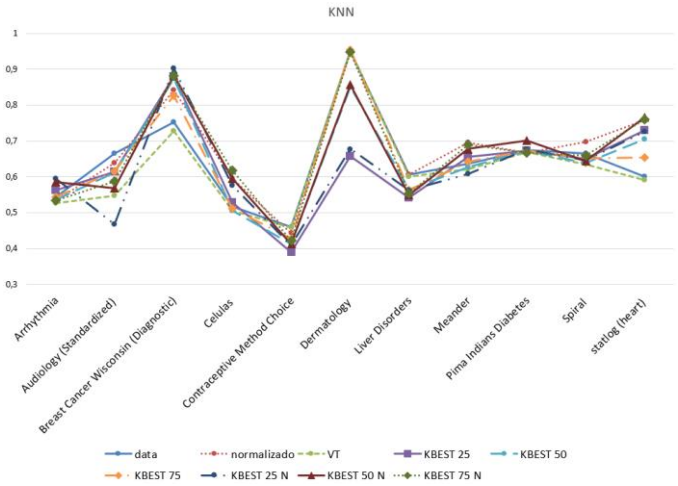
Fonte: Autor

Gráfico 4. Média Bayes.



Fonte: Autor

Gráfico 5. Média KNN.



Fonte: Autor

REFERÊNCIAS

AL-SHAYEA, Q. K. Artificial Neural Networks in Medical Diagnosis. IJCSI International Journal of Computer Science Issues, MIS Department, Al-Zaytoonah University of Jordan, v.8, n.2, Mar. 2011. Disponível em: <http://ijcsi.org/papers/IJCSI-8-2-150-154.pdf>

BATISTA, G. E. A. P. A.; MONARD, M. C. Utilizando Métodos Estatísticos de Resampling para Estimar a Performance de Sistemas de Aprendizado. 1998. Disponível em: <http://www.icmc.usp.br/~mcmonard/public/icmcwshG1998.pdf>

BENNET, C. C.; HAUSER, K. Artificial Intelligence Framework for Simulating Clinical Decision-Making: A Markov Decision Process Approach, Artificial Intelligence in Medicine, 2013. Disponível em: <http://arxiv.org/ftp/arxiv/papers/1301/1301.2158.pdf>

BRUMATTI, M. Redes neurais artificiais. Vitória: Universidade Federal do Espírito Santo, 2005.

CARVALHO, F. O. et al. Agregação de fatores de risco cardiovascular e ocorrência de hipertensão arterial em adultos sedentários. Revista Brasileira de Medicina do Esporte, v.19, n.6, Nov./Dec. 2013. Disponível em: <http://www.scielo.br/scielo.php?pid=S1517-86922013000600008&script=sci_arttext>

CASSIANO, K. M. Análise de Séries Temporais Usando Análise Espectral Singular (SSA) e

Clusterização de Suas Componentes Baseada em Densidade. Pontífica Universidade Católica do Rio de Janeiro – PU-Rio. Setembro, 2014. Disponível em: <http://www.maxwell.vrac.puc-rio.br/24787/24787_1.PDF>

ESTER, M. et al. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Institute for Computer Science, University of Munich, 1996. Disponível em: <https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf>

ESTEVES, S. R.; LORENA, A. C.; NASCIMENTO, M. Z. Aplicação de técnicas Aprendizado de Máquina na Classificação de Imagens Mamográficas. In: SIMPÓSIO DE INICIAÇÃO CIENTÍFICA DA UNIVERSIDADE FEDERAL DO ABC, 2. 2009. Disponível em: <http://ic.ufabc.edu.br/II_SIC_UFABC/resumos/paper_5_150.pdf>

FERNEDA, E. Redes neurais e sua aplicação em sistemas de recuperação de informação. Ci Inf. Brasília, v. 35, n. 1, abr. 2006. Disponível em: <http://www.scielo.br/pdf/%0D/ci/v35n1/v35n1a03.pdf>

FREITAS, G. M. et al. Estimativa de ocorrência de precipitação em áreas agrícolas utilizando floresta de caminhos ótimos. Revista Brasileira de Meteorologia, v.25, n.1, Mar. 2010. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-77862010000100002>

JUNIOR, G. M. O. Máquina de Vetores Suporte: estudo e análise de parâmetros para otimização de resultado. 2010. Disponível em: <http://www.cin.ufpe.br/~tg/2010-2/gmoj.pdf>

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE. Proceedings... v.14, p.1137–1145, 1995.

LINDEN, R. Técnicas de Agrupamento. Revista de Sistemas de Informação da FSMA, n.4, p.18-36, 2009. Disponível em: <http://www.fsma.edu.br/si/edicao4/FSMA_SI_2009_2_Tutorial.pdf>

LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma Introdução às Support Vector Machines. Revista de Informática Teórica e Aplicada, v.14, n.2, p.43-67, 2007.

MARINS, A. I. M. M. Implementação do FREx_SVM: máquinas de vetor suporte para classificação em múltiplas classes. VI Seminário de Iniciação Científica da Puc-Rio, Pontifícia Universidade Católica do Rio de Janeiro, 2008. Disponível em: <http://www.puc-rio.br/pibic/relatorio_resumo2008/relatorios/ctc/ele/ele_aarao_irving.pdf>

METZ, J. Interpretação de clusters gerados por algoritmos de clustering hierárquico. Serviço de pós-graduação do ICMC-USP, Instituto de Ciências Matemáticas e de Computação, Jun. 2006.

NEILL, B. D. Usina Artificial Intelligence to Improve Hospital Inpatient Care. IEEE Computer Society, H.J. Heinz III College, Carnegie Mellon University, 2013. Disponível em: <http://www.cs.cmu.edu/~neill/papers/ieee-is2013.pdf>

OLIVEIRA, F. L. Modelo de Interfaces Adaptativas utilizando Redes Bayesianas. Pós-Graduação em Ciências da Computação – CPGCC – UFSC – Universidade Federal de Santa Catarina. In: ENCONTRO DE ESTUDANTES DE INFORMÁTICA DO TOCANTINS, Palmas, TO. Anais... Out., 2003. p.139-148. Disponível em: <http://arquivo.ulbra-to.br/ensino/43020/artigos/anais2003/anais/interfacesadaptativas-encoinfo2003.pdf>

OSÓRIO, F.; BITTENCOURT, J. R. Sistemas Inteligentes baseados em Redes Neurais Artificiais aplicados ao Processamento de Imagens. In: WORKSHOP DE INTELIGÊNCIA ARTIFICIAL, 1. Anais... Universidade de Santa Cruz do Sul, 2000. Disponível em: <http://osorio.wait4.org/oldsite/wia-unisc/wia2000-full.pdf>

PAGNIN, A. F. Identificação e Caracterização de Pterígio Utilizando Floresta de Caminhos Ótimos e Técnicas de Otimização. São José do Rio Preto: Universidade Estadual Paulista "Júlio De Mesquita Filho", 2011. Disponível em: <http://www.ibilce.unesp.br/Home/Pos-Graduacao475/CienciadaComputacao/pdf-andre-franco-pagnin.pdf>

PAPA, J. P. et al. Efficient supervised optimum-path forest classification for large datasets. Pattern Recognition, v.45, n.1, p.512-520, Jan. 2012. https://doi.org/10.1016/j.patcog.2011.07.013

PEREIRA, C. R. et al. A Step Towards the Automated Diagnosis of Parkinson's Disease: Analysing Handwriting Movements
IEEE Computer Society, IEEE 28th International Symposium on Computer-Based Medical Systems (CBMS), 2015.

QUINTA, L. N. B. et al. Floresta de Caminhos Ótimos na Classificação de Pólen. In: WORKSHOP DE VISÃO COMPUTACIONAL, Centro-Oeste, Instituto de Informática, 8. 2012. Disponível em: <http://iris.sel.eesc.usp.br/wvc/Anais_WVC2012/pdf/98194.pdf>

SILVA, L. M. O.; SOUZA, R. C. Uma aplicação de decisão, redes neurais e knn para a identificação de modelos arma não-sazonais e sazonais. Pontifica Universidade Católica do Rio de Janeiro – PUC-RIO, 19 Dez. 2005. Disponível em: <http://www.maxwell.vrac.puc-rio.br/Busca_etds.php?strSecao=resultado&nrSeq=7587@1>

SILVA, R. L. S. Um modelo de Redes Bayesianas aplicado a sistemas de realidade aumentada. Tese – Universidade Federal do Rio de Janeiro, COPPE 2006. Disponível em: <http://wwwp.coc.ufrj.br/teses/doutorado/inter/2006/Teses/SILVA_RLS_06_t_D_int.pdf>

SWEILAM, N. H.; THARWAT, A. A.; MONIEM, N. K. A. Support vector machine for diagnosis cancer disease: a comparative study. Egyptian Informatics Journal, Cairo University, 2010. Disponível em: <http://ac.els-cdn.com/S1110866510000241/1-s2.0-S1110866510000241-main.pdf?_tid=823fa0f0-0277-11e6-9fd3-00000aacb360&acdnat=1460662479_26812601608e2a73ff3a7d71f565836b>