

UM SISTEMA CONEXIONISTA DE APOIO A DECISÃO PARA TRIAGEM DE PACIENTES EM UNIDADES BÁSICAS DE SAÚDE

Anderson Palácio de Queiroz (Universidade Estadual do Ceará - Brazil),
Necio Veras (Instituto Federal de Educação, Ciência e Tecnologia do Ceará - Brazil),
Gustavo Campos (Universidade Estadual do Ceará - Brazil)

Abstract: *The non-prioritization of clinical cases by basic health unit teams is a problem that restricts people's access to services. This difficulty arises from the excessive number of people under the responsibility of the teams, which is often higher than its service capacity, creating a great competition and an overcrowding of the posts. In this context, the objective of this paper is to present a decision support system, based on artificial neural networks, which perform triage of patients from their health complaints, making possible to identify the different risk gradations, directing prioritized cases to the appropriate health professionals.*

Resumo: A não priorização dos casos clínicos pelas equipes das unidades básicas de saúde é um problema que restringe o acesso da população aos serviços. Essa dificuldade surge a partir do número excessivo de pessoas sob a responsabilidade das equipes, que muitas vezes é superior à sua própria capacidade de atendimento, gerando uma grande concorrência e a superlotação dos postos. Nesse contexto, o objetivo deste trabalho é apresentar um sistema de apoio a decisão, baseado em redes neurais artificiais, que realize triagens de pacientes a partir de suas queixas de saúde, possibilitando identificar as diferentes gradações de risco, direcionando os casos priorizados para os profissionais de saúde adequados.

1. Introdução

A triagem (ou classificação) de pacientes é um processo complexo que envolve múltiplos atores, como profissionais, pacientes, familiares, órgãos municipais, estaduais e federais da área da saúde. Consiste na etapa inicial do atendimento ao paciente e determina o tempo máximo que este poderá aguardar pelo atendimento médico sem que esta espera cause prejuízos a sua saúde⁶.

Em uma Unidade Básica de Saúde (UBS) é preciso tratar cada indivíduo de acordo com sua necessidade, evitando diferenciações injustas. Para isso, é importante manter uma estratégia que garanta o acesso justo ao serviço frente às filas, distribuição de senhas e atendimentos agrupados por ordem crescente em relação à chegada dos usuários à UBS. Uma estratégia fundamental para garantir um acesso eficiente deve considerar a adoção da avaliação de riscos, identificando diferentes gradações e/ou situações de urgência para, a partir disso, realizar as devidas priorizações. É relativamente fácil identificar sinais de riscos mais evidentes, como um paciente com dificuldade de respirar ou mesmo um idoso com dor no peito, no entanto, nem sempre o risco é facilmente reconhecível, variando de acordo com o modo de manifestação da doença entre os pacientes.

Por isso, a avaliação de riscos não somente indica o tipo de intervenção necessária, mas também o momento que o atendimento deve ocorrer. Na atenção básica à saúde, diferentemente de uma emergência hospitalar, a intervenção imediata só se faz necessária em situações de alto risco, no entanto, é necessário priorizar o atendimento de alguns casos, sob a pena de manter uma pessoa em sofrimento por tempo prolongado¹.

A classificação geral dos casos de demanda em uma UBS ocorrem de duas formas: situação não-aguda e situação aguda ou crônica agudizada². A situação não aguda diz respeito a casos menos graves

nos quais a intervenção pode ser remarcada para um outro momento. Já a situação aguda ou crônica agudizada é subdividida em três outras classes: (1) atendimento imediato (alto risco), (2) atendimento prioritário (risco moderado) ou (3) atendimento no dia (risco baixo ou ausência de risco)^{1,2}.

O presente artigo objetiva apresentar um sistema de apoio a decisão que correlacione, de modo automático, as queixas de saúde dos pacientes para três doenças específicas, comuns em unidades básicas de saúde, com a avaliação adequada de riscos. O sistema, baseado em Redes Neurais Artificiais (sistema conexcionista), reconhece padrões que o habilita a realizar triagens de pacientes a partir de uma base de dados de queixas de saúde, fornecida por profissionais atuantes na atenção básica de um município brasileiro. Portanto, o processo objetiva possibilitar a identificação de diferentes gradações de riscos visando direcionar os casos prioritizados para as intervenções adequadas.

2. Métodos

Na presente pesquisa foi utilizada uma abordagem dividida em três fases como forma de explicitar o modelo da solução para o problema da triagem de pacientes: (a) Extração de atributos, (b) Otimização paramétrica e (c) Classificação. Durante a preparação dos dados, ocorreu a normalização e extração de atributos e, em seguida, a fase de otimização dos parâmetros internos das redes, ocorrendo de forma anterior à etapa de classificação.

A fase inicial, de Extração de atributos, foi necessária para remover as informações redundantes ou irrelevantes para o processo de classificação, além de aumentar a acurácia e o desempenho do classificador. A entrada desta etapa é, normalmente, um conjunto de dados brutos e as saídas são os dados com dimensionalidade reduzida^{5,4}, contendo os atributos que carregam a maior parte da informação relevante, mantendo assim, o máximo poder discriminativo do conjunto. Com a fase de Otimização, buscou-se minimizar a combinação do erro global de treinamento e do erro de generalização, identificando os melhores parâmetros e topologias da rede através da meta-heurística de *Genetic Algorithm* (GA). A partir da redução de dimensionalidade dos dados e da identificação das melhores redes, a probabilidade de melhorar o desempenho da classificação é significativamente maior^{4,5}. A dinâmica do funcionamento da abordagem utilizada nesta pesquisa é ilustrada na Figura 1.

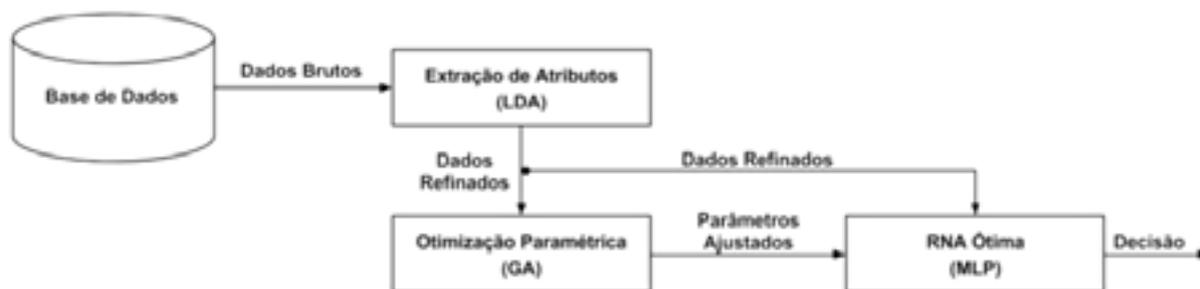


Figura 1: Dinâmica da abordagem proposta. Fonte: Primária.

Feita a extração de atributos, iniciou-se a fase de otimização paramétrica, que correspondeu ao treinamento das redes, onde o GA recebeu como entrada os dados refinados e, para cada iteração, gerou-se uma população de 50 cromossomos. Para cada solução paramétrica é instanciada uma rede MLP (*Multilayer Perceptron*) capaz de ser executada durante 100 (cem) épocas de treinamento. A melhor solução é então avaliada através da validação cruzada de *k-partições* e o seu desempenho analisado. Por tanto, a rede neural escolhida para a fase de classificação desta pesquisa é do tipo multicamadas (MLP).

Basicamente, a rede utilizou dados redimensionados e agrupados pela Análise Discriminante Linear (LDA) de forma a generalizar e aprender a classificar os casos clínicos de três doenças comuns

em postos de saúde². A saída do modelo neural supervisionado generalizou casos de teste corretamente entre os tipos possíveis de situações agudas, ilustradas na Figura 2, encaminhando os pacientes para os procedimentos de saúde adequados.

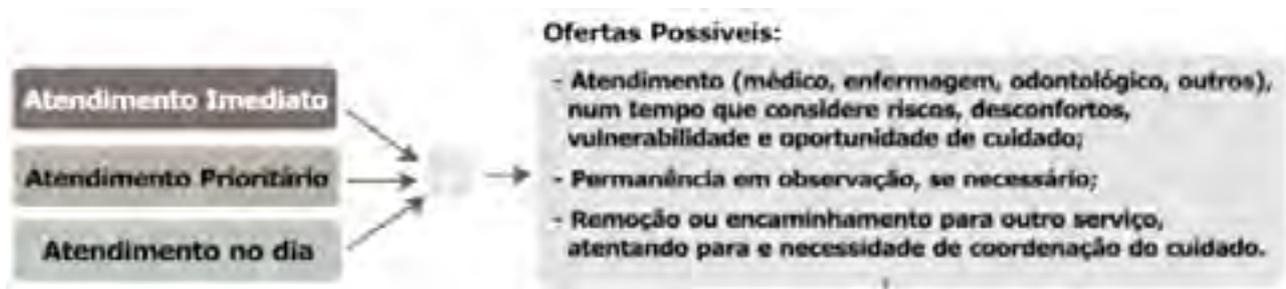


Figura 2: Estratificação de risco para situações agudas².

É importante acrescentar que o modelo neural MLP foi escolhido devido sua ampla utilização, simplicidade e estabilidade, sendo aplicado, principalmente, em classificação, análise e interpolação de dados, o que o torna particularmente adequado para a modelagem de processos em geral^{4,5}.

2.1 Preparação dos dados

Os dados utilizados nesta pesquisa foram relativos a queixas comuns em postos de saúde coletados por uma equipe de enfermeiros que trabalham em diferentes Unidades Básicas de Saúde de um município brasileiro. Especificamente, considerou-se casos clínicos de três patologias: cefaleia, diabetes e diarreia. Os casos foram apresentados ao processo de triagem conexcionista de forma aleatória, simulando, assim, o acolhimento à demanda espontânea na atenção básica. Para tanto, foram coletados 81 casos clínicos de cefaleia, distribuídos conforme ilustrado na Tabela 1.

Tabela 1: Quantidade de casos de cefaleia por classe de risco.

Classes	Quantidade de casos
Atendimento não urgente	14
Atendimento no dia	19
Atendimento prioritário	30
Atendimento imediato	18

Além disso, foram coletados mais 83 casos clínicos de diabetes divididas em quatro classes em conformidade com a estratificação de riscos¹. Na Tabela 2 é possível perceber a distribuição dos casos para diabetes.

Tabela 2: Quantidade de casos de diabetes por classe de risco.

Classes	Quantidade de casos
Atendimento não urgente	11
Atendimento no dia	16
Atendimento prioritário	31
Atendimento imediato	25

Finalmente, a Tabela 3 mostra a distribuição por classe de estratificação de risco de 82 casos clínicos para diarreia e vômito, igualmente coletados pela pesquisa.

Tabela 3: Quantidade de casos de diarreia e/ou vômito por classe.

Classes	Quantidade de casos
Atendimento não urgente	11
Atendimento no dia	25
Atendimento prioritário	28
Atendimento imediato	18

Sabe-se que a aquisição e construção da base de conhecimento de um sistema reconhecedor de padrões com o auxílio de especialistas é uma tarefa que exige um considerável tempo para ser executada. Por essa razão, a base de conhecimento extraída para esta pesquisa contém uma quantidade relativamente baixa de exemplos de treinamento e teste, porém suficientes para a geração de dados de forma a permitir que conclusões sejam elencadas. Observa-se também que as três bases possuem uma dimensionalidade elevada, o que pode, na prática, degradar o desempenho da classificação. Quando se dispõe de um conjunto de treinamento limitado, deve-se selecionar apenas um pequeno número de atributos relevantes³. Para tanto, foi realizado o pré-processamento (normalização e redução de dimensionalidade) dos dados de entrada antes do início do processo de classificação.

A normalização correspondeu à ajustes nos dados, garantindo que todas as amostras estivessem em uma mesma escala, viabilizando a entrada dos dados e um treinamento eficaz pela rede neural. Uma técnica comum para isso é escalar as entradas em média zero e variância um, também conhecida em estatística por normalização *z-score*. Com o intuito de reduzir as dimensões do vetor de características obtido a partir das três bases coletadas, utilizou-se a Análise Discriminante Linear (LDA). Com isso, objetivou-se preservar o máximo possível a informação discriminatória das classes durante a redução da dimensionalidade dos dados.

Além de reduzir o custo operacional e simplificar o problema, o LDA fez com que os padrões de uma mesma classe permanecessem agrupados o mais densamente possível e os agrupamentos de classes distintas estivessem o mais separados possível. Com isso, além da extração de atributos, o LDA também realizou a clusterização das amostras, facilitando ainda mais o processo de discriminação do classificador.

2.2 Modelagem da Rede MLP

Nesta fase objetivou-se aplicar técnicas e algoritmos para melhorar a classificação dos dados. Durante o treinamento e teste da fase de modelagem das redes *Multilayer Perceptron*, utilizou-se os conjuntos de dados transformados a partir da fase de extração de atributos divididos em dois subconjuntos, conforme demonstrado na Tabela 4.

Tabela 4: Amostras de treinamento e teste para modelagem das redes MLP.

Subconjunto	Quantidade
Treinamento	162
Teste	81

O subconjunto de treinamento foi usado essencialmente no processo de aprendizado da rede, sintonizando os pesos e limiares de seus neurônios. Já o subconjunto de teste foi utilizado para verificar a generalização de soluções a partir de dados de entrada não apresentados no treinamento, possibilitando assim a validação da topologia assumida.

2.3 Evolução dos parâmetros para a Rede

Inicialmente foi necessário encontrar topologias de redes MLP capazes de exprimir a variância dos três conjuntos de dados e, para tanto, aplicou-se uma sistemática baseada

em computação evolutiva por meio de Algoritmo Genético para encontrar topologias ótimas para as redes de forma a discriminar apropriadamente os dados.

O Algoritmo Genético buscou evoluir os seguintes parâmetros internos das redes: (1) A Taxa de Aprendizagem (η); (2) O Termo *Momentum* (α); (3) A função de Ativação (g); (4) O número de camadas ocultas e (5) o número de neurônios contidos em cada camada oculta. Esses parâmetros foram codificados em sequências binárias de até 21 bits. A Tabela 5 exemplifica uma possível combinação para os parâmetros encontrados pelo GA.

Tabela 5: Exemplo de uma possível solução.

Representação	g(.)	h	a	1º camada	2º camada
Binária	1	011010	1010	10100	11011
Decimal	1	0,40625	0,625	20	27

Para a pesquisa utilizou-se duas possíveis funções para a Função de Ativação do Algoritmo Genético: sigmoide e tangente hiperbólica. Sobre o número de camadas ocultas, limitou-se a uma ou duas camadas e, quanto ao número de neurônios à 32 neurônios por camada. Após a codificação de cada solução candidata em uma topologia de rede, passou-se para a etapa de treinamento e, em seguida, a rede foi submetida ao conjunto de teste para verificação de capacidade de generalização. Dessa forma, foi possível analisar a capacidade da rede de prever padrões de saída desconhecidos. Finalmente, um valor é atribuído para a solução gerada de acordo com a função objetivo, construindo assim, uma nova solução.

2.4 Função Objetivo

A função objetivo utilizada pelo GA consistiu na combinação de dois critérios de erro e pode ser expressa conforme a Equação 1:

$$f = \frac{P_1 \times E_{\text{train}} + P_2 \times E_{\text{gen}}}{P_1 + P_2}, \text{ onde:} \quad (1)$$

E_{train} e E_{gen} são os erros de treinamento e generalização, respectivamente; e P_1 e P_2 são fatores de ajuste.

Os fatores P_1 e P_2 foram utilizados para ponderar o grau de importância atribuído aos erros, os quais $P_1, P_2 > 0$ e $P_1, P_2 \in \mathbb{R}$.

2.5 Topologias para a Redes MLP

A aplicação do Algoritmo Genético resultou em topologias de redes e parâmetros internos diferentes para cada conjunto de dados e as soluções encontradas pelo algoritmo foram consideradas boas, mesmo não tendo certeza de que são soluções ótimas (as melhores dentro do universo das soluções possíveis).

As Tabelas 6, 7 e 8 apresentam os cinco melhores resultados obtidos para a execução do algoritmo nas bases cefaleia, diabetes e diarreia ou vômito, respectivamente.

Tabela 6: Melhores parâmetros e topologias para a base da cefaleia.

Topologia	h	a	g(.)	Etrain	Egen
04 x 17 x 0 x 04	0,44	0,75	Sigmoide	0,059	0,088
04 x 26 x 0 x 04	0,53	0,69	Sigmoide	0,062	0,097
04 x 03 x 0 x 04	0,55	0,25	Sigmoide	0,086	0,063
04 x 14 x 0 x 04	0,45	0,12	Sigmoide	0,089	0,095
04 x 07 x 0 x 04	0,75	0,062	Sigmoide	0,091	0,081

Tabela 7: Melhores parâmetros e topologias para a base da diabetes.

Topologia	h	a	g(.)	Etrain	Egen
04 x 10 x 0 x 04	0,73	0,38	Sigmoide	0,075	0,046
04 x 22 x 0 x 04	0,58	0,5	Sigmoide	0,086	0,036
04 x 09 x 0 x 04	0,58	0,25	Sigmoide	0,089	0,053
04 x 17 x 0 x 04	0,58	0,69	Sigmoide	0,089	0,08
04 x 19 x 0 x 04	0,58	0,25	Sigmoide	0,093	0,055

Tabela 8: Melhores parâmetros e topologias para a base da diarreia e/ou vômito.

Topologia	h	a	g(.)	Etrain	Egen
04 x 14 x 0 x 04	0,17	0,81	Sigmoide	0,12	0,17
04 x 11 x 0 x 04	0,81	0,25	Sigmoide	0,12	0,18
04 x 18 x 0 x 04	0,84	0,19	Sigmoide	0,12	0,18
04 x 06 x 0 x 04	0,81	0,25	Sigmoide	0,12	0,18
04 x 13 x 0 x 04	0,81	0,69	Sigmoide	0,13	0,18

Em cada execução foi gerado, inicialmente, uma população de 50 cromossomos (soluções), com 20 cruzamentos e 10 mutações. Cada rede neural MLP foi executada ao longo da fase de treinamento com erro aceitável definido em 0,01 durante 100 épocas de treinamento. É importante observar que para todo o conjunto de treinamento e teste, somente uma camada neural intermediária foi suficiente para discriminar os dados entre as quatro classes desejadas e a função de ativação sigmoide foi a única selecionada entre as melhores soluções encontradas pelo algoritmo genético.

3. Resultados e Discussão

A avaliação das melhores soluções encontradas iniciou-se com a seleção das melhores topologias de redes e configurações de parâmetros em cada uma das três bases englobadas pela pesquisa. Em seguida, realizou-se os treinamentos e testes definitivos das redes selecionadas e avaliou-se os erros de treinamento e generalização, assim como os percentuais de acerto (acurácia) e as taxas de erro. O processo de validação foi realizado com base nas melhores soluções paramétricas e topológicas encontradas pelo Algoritmo Genético.

A presente pesquisa utilizou a validação cruzada de k -partições (*k-fold cross-validation*) para realizar a divisão do conjunto total de amostras em k partições, sendo que $(k - 1)$ delas foram usadas para compor o subconjunto de treinamento e a partição restante constituiu o subconjunto de teste. Todo o processo de aprendizado da rede repetiu-se k vezes até que todas as partições tivessem sido utilizadas como subconjunto de teste. O resultado experimental dos testes, considerando a união de todos os conjuntos de dados, está apresentado sob a forma de uma matriz de confusão (Tabela 9).

Tabela 9: Matriz de Confusão referente a totalidade dos dados.

	Atendimento não urgente	Atendimento no dia	Atendimento prioritário	Atendimento imediato
Atendimento não urgente	33	2	1	0
Atendimento no dia	2	52	4	0
Atendimento prioritário	1	2	86	0
Atendimento imediato	0	1	3	56

A partir da Matriz apresentada pela Tabela 9 é possível calcular as porcentagens dos casos clínicos que tiveram seus riscos corretamente reconhecidos pelo classificador (Tabela 10):

Tabela 10: Taxas de acerto e erro para cada classe de risco.

	Taxa de acerto	Taxa de erro
Atendimento não urgente	91,6%	8,3%
Atendimento no dia	91,2%	8,7%
Atendimento prioritário	91,4%	8,5%
Atendimento imediato	100%	0%

Logo, a acurácia média global é dada pela Equação 2:

$$\frac{33+52+86+56}{243} = \frac{227}{243} = 93.4 \% \quad (2)$$

Com isso, ao aplicar o método da validação cruzada de 3 partições, divide-se aleatoriamente os conjuntos de dados entre três ensaios, cada um com 100 épocas de treinamento, isto é, todas as amostras de treino são apresentadas a rede 100 vezes, totalizando 300 épocas de treinamento. Assim, é garantido que dois terços de cada base de dados foi utilizada para treino e um terço para teste.

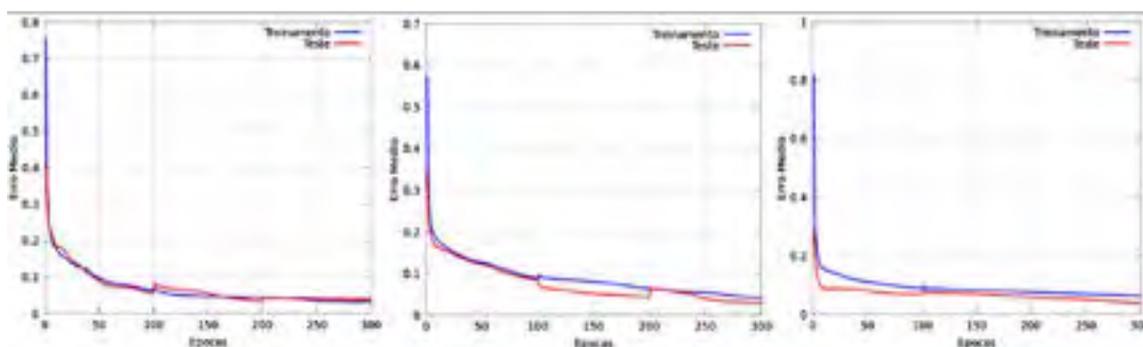


Figura 3: Decaimento do erro quadrático médio com o decorrer das épocas de treino e teste para as bases da cefaleia, diabetes e diarreia, respectivamente.

A Figura 3 apresenta os dados sobre o monitoramento do treinamento e teste dos modelos de redes durante a validação cruzada e o erro médio de classificação é gerado durante as iterações de treinamento (épocas) para as amostras de treino e validação. O decaimento dos erros de classificação e teste com o crescimento do número de épocas demonstra o aprendizado dos modelos para discriminar as

estratificações de risco das doenças, comprovando a capacidade de generalizar classificações para novos casos e, conseqüentemente, de aprender com o passado.

4. Conclusão

Neste trabalho foi apresentado um sistema de apoio à decisão baseado numa arquitetura de rede neural para auxiliar médicos e enfermeiros durante o processo de triagem de pacientes em unidades básicas de saúde. Os dados utilizados foram relativos aos sintomas de três doenças comuns em postos de saúde: cefaleia, diabetes e diarreia e/ou vômito². Estes dados foram coletados por uma equipe de profissionais da área da enfermagem que atuam em unidades básicas de saúde de município brasileiro. O objetivo do sistema foi classificar casos clínicos em quatro classes de riscos distintas, encaminhando o paciente para o atendimento adequado, auxiliando, dessa forma, o processo de triagem.

Após as fases de redução de dimensionalidade dos dados e otimização paramétrica, as melhores redes MLP foram selecionadas e submetidas ao processo de testes. O método da validação cruzada de *k-partições* foi aplicado para aferir a capacidade de generalização da abordagem e as soluções paramétricas das redes multicamadas obtiveram boas avaliações tanto no treinamento quanto na validação, sendo comprovado por meio da avaliação do decaimento dos erros quadráticos médios e das acurácias estimadas após os testes de validação cruzada.

Portanto, os resultados mostram que o sistema proposto pode atingir uma significativa acurácia (superior a 90%) de classificação, demonstrando real aplicabilidade no suporte a triagem de pacientes em unidades básicas de saúde. Como trabalhos futuros, sugere-se a aplicação da abordagem com outros tipos de enfermos considerando bases de dados relativamente maiores.

5. Referências

- [1] BRASIL, M. d. S. (2011). Acolhimento à demanda espontânea. Brasília: Ministério da Saúde - Cadernos de Atenção Básica n. 28, Volume I.
- [2] BRASIL, M. d. S. (2012). Acolhimento à demanda espontânea: queixas mais comuns na Atenção Básica. Brasília: Ministério da Saúde - Cadernos de Atenção Básica n. 28, Volume II.
- [3] Jain, A. K., Duin, R. P., and Mao, J. (2000). Statistical pattern recognition: A review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(1):4–37.
- [4] Khan, J., Wei, J. S., Ringner, M., Saal, L. H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Antonescu, C. R., Peterson, C., et al. (2001). Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nature medicine*, 7(6):673–679.
- [5] Olaniyi, E. O. and Adnan, K. (2014). Onset diabetes diagnosis using artificial neural network. *International Journal of Scientific and Engineering Research*, 5(10).
- [6] Schlatter, R. P. Análise do sistema de apoio à decisão para triagem de pacientes em serviço de emergência. Master's thesis, Universidade Federal do Rio Grande do Sul, Escola de Administração, Programa de Pós-graduação em Administração.