

# Um Sistema de Recomendação Híbrido para Gestão em Saúde Pública

Caroline Lucena, Marcelo Lacerda, Paulo Cordeiro, Rafael Caldas, Fernando Buarque  
Universidade de Pernambuco  
Escola Politécnica de Pernambuco  
Pernambuco, Brasil  
cvl,mgpl,prgc,rrc,fbln@ecom.poli.br

**Resumo**—A relação entre a medicina e os problemas sociais é tema de pesquisas científicas há mais de um século em todo o mundo. Visando o estudo de determinantes sociais em saúde (DSS), a Organização Mundial de Saúde (OMS) criou uma comissão especial com tal objetivo. Tais determinantes englobam fatores relacionados à ocorrência de problemas em saúde pública, podendo estes serem culturais, econômicos, sociais, entre outros tipos. Contudo, a complexidade da relação entre os determinantes sociais e os indicadores em saúde implica dificuldade no processo de tomada de decisões. Ademais, deve-se levar em consideração os custos agregados à cada ação. Nesse contexto, o presente trabalho propõe a criação de um sistema de recomendação híbrido para auxílio em gestão em saúde pública. Este é composto por dois módulos: (i) aproximador de funções e (ii) otimizador multiobjetivo. No módulo (i), uma Multilayer Perceptron é utilizada. Para o módulo (ii), uma das versões multiobjetivo do algoritmo Particle Swarm Optimization foi utilizada, o MOPSO-CDR. O sistema possui como objetivo retornar um conjunto de sugestões de ações a um gestor de saúde pública, visando a redução da ocorrência de determinadas doenças em sua região de atuação, porém levando em consideração o custo de cada ação, o qual deve ser minimizado. Para a execução dos testes, o sistema foi alimentado com dados reais colhidos na base de dados do DataSUS através do Sinan (Sistema de Informação de Agravos de Notificação) e do IBGE (Instituto Brasileiro de Geografia e Estatística). Considerando a aplicação do sistema em um cenário real, este obteve um resultado satisfatório, cumprindo com os objetivos previstos.

**Keywords**—*Determinantes sociais de saúde, Sistemas de Recomendação, Otimização Multiobjetivo, Redes Neurais Artificiais, Inteligência de Enxames*

## I. INTRODUÇÃO

Virchow (1821 - 1902), considerado o pai da medicina social, afirmava que a “ciência médica é intrínseca e essencialmente uma ciência social”. Durante o Movimento pela Reforma Médica na Alemanha do século XIX, Virchow detectou a existência de relações entre a medicina e os problemas sociais; para ele, as condições econômicas e sociais da população exercem um efeito importante sobre o processo saúde-doença e tais relações devem ser submetidas à pesquisa científica [1][2].

As diferenças ou desigualdades no quadro de saúde são naturais e dependem das características de cada indivíduo ou de cada grupo ou comunidade. No entanto, o que não é natural são aquelas diferenças na situação de saúde que mantém relação com os determinantes sociais de saúde. Estes, segundo a Comissão Nacional sobre os Determinantes Sociais da Saúde, podem ser compreendidos como os fatores sociais,

econômicos, culturais, étnicos/raciais, psicológicos e comportamentais que influenciam a ocorrência de problemas de saúde e seus fatores de risco na população. Em oposição às diferenças e desigualdades naturais, essas desigualdades são injustas e inaceitáveis, e por isso são denominadas como iniquidades [3][4].

Os sanitaristas brasileiros realizam, desde o início do século passado, estudos aprofundados sobre as relações entre os determinantes socioeconômicos e a situação de saúde. Baseadas nesse conhecimento e com o objetivo de combater as iniquidades em saúde, as ações são desenvolvidas, uma vez que o Brasil está entre os países com maiores indicadores relativos a esse tipo de desigualdade, pois o país apresenta grandes disparidades entre os diversos estratos sociais e econômicos da população brasileira [5].

A Comissão Nacional sobre Determinantes Sociais da Saúde (CNDSS) foi criada pelo Decreto Presidencial de 13/3/2006 e se insere num processo global, desencadeado pela Organização Mundial da Saúde (OMS), quando esta criou sua Comissão sobre Determinantes Sociais em março de 2005. Apesar da riqueza e diversidade de enfoques e abordagens relativos aos estudos das relações entre os Determinantes Sociais de Saúde com os indicadores de saúde, há ainda muito que se avançar no conhecimento da dinâmica de relações e mediações entre esses, pois trata-se de uma relação complexa [1].

Os estudos que correlacionam determinantes sociais de saúde permitem identificar e relacionar esses, e pode ser ferramenta de auxílio à gestão em saúde indicando como devem ser feitas as intervenções, objetivando a redução das iniquidades em saúde, ou seja, os pontos mais sensíveis onde tais intervenções podem provocar maior impacto [1]. No entanto, devido à tal complexidade, estudos e processos de tomada de decisão, quando executados por gestores humanos, podem se tornar altamente custosos. Além disso, é importante observar a existência inerente dos custos financeiros de cada ação a ser tomada no combate às iniquidades, o que torna tal processo de decisão ainda mais complexo, levando em consideração que se deve minimizar tanto as iniquidades, quanto os recursos.

Diversos trabalhos publicados propuseram sistemas computacionais de recomendação e suporte a decisão em gestão de saúde. Em sua maioria, uma técnica de busca metaheurística é utilizada para, através de um modelo matemático/computacional pré-definido de um problema de alocação de recursos em gestão de saúde, sugerir ao gestor

as melhores decisões a serem tomadas, visando minimizar custos e maximizar a satisfação dos indivíduos beneficiados pelo serviço prestado. Em 2010, Helm et al. [6] propuseram a aplicação de Algoritmos Genéticos para a otimização dos parâmetros de um consolidado sistema de suporte a decisão, utilizado para auxiliar o gestor de um hospital na alocação de recursos para o atendimento dos pacientes. O objetivo era reduzir o trabalho manual necessário para o ajuste de tais parâmetros e possibilitar a criação de um *pareto* de soluções para melhor informar o usuário sobre os trade-offs existentes entre as métricas de avaliação da qualidade do atendimento. Em 2012, Shariff et al. [7] utilizaram Algoritmos Genéticos para otimizar a alocação de locais de atendimento à saúde em uma determinada região, buscando minimizar os custos sociais relacionados à manutenção da saúde e maximizando os benefícios à população.

Um grande esforço também vem sendo aplicado no desenvolvimento de modelos matemáticos/computacionais de tais cenários, objetivando o uso destes na elaboração dos ambientes de simulação necessários em trabalhos como os citados anteriormente. Alguns exemplos podem ser citados. Em 2008, Tai e Williams [8] modelaram o problema de alocação de recursos em hospitais sob a perspectiva de linha de produção, onde os pacientes que precisavam ser tratados eram consumidores dos serviços de saúde oferecidos. Em 2011, Zheng et al. [9] propuseram um modelo computacional de eventos discretos capaz de simular o fluxo de pacientes em salas de operação de hospitais e a alocação de recursos necessários para tais operações, a fim de detectar ineficiências no processo de alocações de tais recursos em hospitais. Em 2012, Rodriguez-Verjan et al. [10] introduziram novas variáveis no problema de alocação de recursos em serviços de saúde oferecidos no lar do paciente, sendo estas locações de facilidades em uma determinada região, atribuição de demanda a empresas do ramo, autorização de prestação de serviços específicos e alocações de recursos gerais, incrementando o modelo previamente existente. Contudo, nota-se que em todos estes trabalhos nos quais são feitas modelagens matemáticas/computacionais de problemas de alocação de recursos em gestão de saúde, os modelos criados são restritos a situações específicas, não havendo capacidade de adaptação.

Nesse contexto, o presente trabalho objetiva a construção de sistema de recomendação para auxílio a gestão em saúde pública capaz de, a partir de dados reais relacionados a determinantes sociais em saúde de uma região, recomendar ao gestor ações a serem tomadas para, através da melhoria destes determinantes, reduzir a ocorrência de doenças na área. No entanto, os custos destas ações de gestão política devem ser levados em consideração pelo sistema, sendo necessário que o sistema retorne um conjunto de soluções equivalentes em custo/benefício. É importante ressaltar que, dessa forma, o sistema deve ser capaz de ser utilizado em qualquer local ou situação, pois este é capaz de adaptar-se aos dados fornecidos, o que não é possível fazer com modelos rígidos, sem capacidade de adaptação.

O artigo está organizado da seguinte maneira: a Seção II trás uma rápida revisão da literatura. A Seção III detalha a abordagem proposta neste trabalho. A Seção IV aborda a metodologia que foi utilizada no desenvolvimento dos experimentos. Na Seção V estão dispostos os resultados dos experimentos

além de discussões acerca do mesmo. Finalmente, a Seção VI as considerações finais acerca deste trabalho estão presentes, além de possíveis trabalhos futuros.

## II. REVISÃO DA LITERATURA

### A. Otimização multiobjetivo

Um problema de otimização multiobjetivo pode ser definido como:

$$\text{minimizar } \vec{f}(\vec{x}) = [f_1(\vec{x}), f_2(\vec{x}), \dots, f_k(\vec{x})], \quad (1)$$

sujeito as seguintes restrições:

$$g_i(\vec{x}) \leq 0 \quad i = 1, 2, \dots, m \quad (2)$$

$$h_j(\vec{x}) = 0 \quad j = 1, 2, \dots, m \quad (3)$$

No qual  $\vec{x} = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$  é o vetor no espaço de busca;  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \dots, k; j = 1, \dots, k$  são as funções objetivo e  $g_i, h_j : \mathbb{R}^n \rightarrow \mathbb{R}$  são as restrições as quais a função está submetida.

Outro conceito presente em otimização multiobjetivo é a dominância entre as soluções. Sendo dois vetores  $\vec{x}, \vec{y} \in \mathbb{R}^k$ , é dito que  $\vec{x}$  domina  $\vec{y}$  se  $f_i(\vec{x}) \leq f_i(\vec{y})$  para pelo menos um valor de  $i$ , quando trata-se de problemas de minimização. A dominância de  $\vec{x}$  sobre  $\vec{y}$  é representada por  $\vec{x} \prec \vec{y}$ . O conjunto das soluções sem dominâncias entre si é denominado de *pareto*. Já o conjunto de soluções não dominadas encontradas no processo de otimização é denominado *pareto front*.

Diversos algoritmos metaheurísticos baseados em enxames de partículas foram propostos para solucionar problemas multiobjetivos [11]. Dentre estes algoritmos, pode-se citar o MOPSO-CDR (*Multiple Objective Particle Swarm Optimization Approach using Crowding Distance and Roulette Wheel*) [12], uma das versões multiobjetivo do algoritmo *Particle Swarm Optimization* [13]. O MOPSO-CDR é um aprimoramento do MOPSO-CDLS (*Multi-Objective Particle Swarm Optimization with Crowding Distance and Local Search*) proposto por Tsou et al. [14]. Ambos utilizam o conceito de arquivo externo para armanezar as melhores soluções encontradas durante as iterações dos algoritmos.

### B. Redes neurais artificiais

Redes neurais artificiais são modelos computacionais bem difundidos advindos da área de inteligência computacional. Tais modelos são inspirados nas funcionalidades do cérebro humano, o qual possui os neurônios como unidades de processamento e sinapses, *i.e.* pesos, mediando as conexões entre neurônios. A estrutura básica de uma rede neural artificial é composta por diversos neurônios conectados entre si, assim como um número ainda maior de sinapses. Cada conexão recebe um peso (chamado de peso sináptico) que faz alusão aos neurotransmissores biológicos. Atribui-se os créditos do primeiro modelo de rede neural artificial aos cientistas McCulloch e Pitts [15].

Diversas arquiteturas de redes neurais artificiais já foram desenvolvidas, cada qual objetivando a solução de problemas de diferentes classes [16]. Dentre as diversas arquiteturas desenvolvidas, a MLP (*Multilayer Perceptron*) está entre as mais utilizadas [17]. Esta é composta por uma camada de neurônios

de entrada conectada a uma ou mais camadas escondidas (estas conectadas entre si em sequência), sendo a última camada escondida conectada a uma camada de saída. Um exemplo de uma arquitetura MLP com duas entradas, uma camada escondida com três neurônios e uma camada de saída com dois neurônios, está representada na Figura 1. Entre os diversos métodos existentes utilizados para realizar o treinamento de redes MLP, o algoritmo *Backpropagation* é o mais conhecido [18] [19].

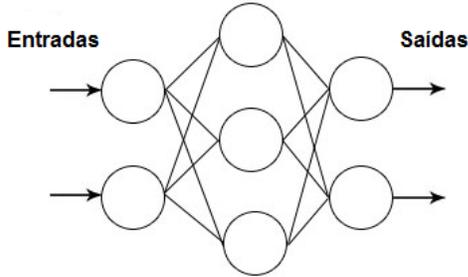


Figura 1. Arquitetura de uma rede MLP, com duas entradas, uma camada escondida e duas saídas.

### III. ABORDAGEM PROPOSTA

O sistema proposto neste artigo tem como objetivo recomendar a um gestor de saúde possíveis ações a serem realizadas no âmbito dos determinantes sociais de saúde, visando a redução da ocorrência de determinadas doenças em sua região de atuação. A abordagem é composta por dois módulos: aproximador de função e otimizador multiobjetivo. Esses módulos são utilizados em duas fases distintas. Na primeira fase ocorre o treinamento do aproximador de função, em seguida o mesmo é utilizado pelo módulo do otimizador como função objetivo. O módulo do aproximador de função tem como objetivo aproximar a função que relaciona determinantes sociais de saúde a ocorrências de doenças. O otimizador multiobjetivo possui o papel de recomendar soluções ao gestor. Este módulo, por sua vez, utilizará tal aproximador como avaliador das soluções geradas. Em outras palavras, o aproximador de função retornará ao otimizador multiobjetivo o número de ocorrências de doenças de um local, dado os valores atribuídos aos determinantes sociais de saúde.

Cada possível ação gestora possui um custo financeiro associado, pois reduzir a ocorrência de doenças possui custos associados. Estes custos são determinados por uma função, a qual deve ser definida por um especialista. É importante ressaltar que este trabalho não visa discutir acerca do cálculo destes custos, sendo os mesmos calculados empiricamente através da Equação 4, onde  $D$  é o número de determinantes sociais de saúde considerados,  $|\Delta d_i|$  é o módulo da diferença do determinante  $i$  antes e após a execução das ações recomendadas pelo sistema e  $\alpha_i$  representa o custo das melhorias do mesmo determinante. Os valores de  $\alpha_i$  estão definidos em uma tabela de custos. Neste trabalho, tal equação é utilizada apenas para considerar a existência de um custo para as ações, podendo esta ser mais complexa se definida por um especialista, o que exige estudos mais profundos sobre o assunto.

$$C = \sum_{i=0}^D \alpha_i |\Delta d_i| \quad (4)$$

A Figura 2 exibe um diagrama onde pode ser visualizado o sistema em módulos e fluxo de dados. No processo de otimização, dois objetivos são considerados: a minimização da ocorrência de doenças em uma região e a minimização dos custos financeiros decorrentes das ações políticas necessárias para tal redução. Portanto, o otimizador multiobjetivo deve gerar um *pareto* ótimo de soluções equivalentes em custo/benefício, o qual é o conjunto de recomendações retornado ao gestor.



Figura 2. Sistema proposto representado em módulos.

Neste trabalho, utilizamos uma rede neural com arquitetura de múltiplas camadas (MLP) como aproximador de função. Optou-se por tal escolha devido a sua capacidade da rede MLP aproximar funções não-lineares. Como algoritmo de otimização multiobjetivo, o algoritmo MOPSO-CDR foi utilizado. Para o treinamento da MLP, foi utilizado o algoritmo *Backpropagation*. Foram utilizadas como funções de ativação a sigmóide logística para a camada intermediária e a função degrau para a camada de saída.

A base de dados BD utilizada para treinar a MLP, representada pela Equação 5, é composta por  $m$  instâncias. Cada instância é representada pela  $n$ -upla especificada pela Equação 6.

$$BD = (e_1, e_2, \dots, e_{m-1}, e_m) \quad (5)$$

$$e_i = (DSS_1, DSS_2, \dots, DSS_{D-1}, DSS_D, d_1, d_2, \dots, d_{n-1}, d_n) \quad (6)$$

onde  $e_i$  é a instância (elemento)  $i$  da base de dados,  $DSS_k$  é o determinante social em saúde  $k$  e  $d_k$  é a doença de índice  $k$ .  $D$  e  $n$  são o números de determinantes sociais em saúde e o número de doenças consideradas no problema, respectivamente. As variáveis  $DSS_s$  foram apresentadas na entrada na MLP, enquanto as variáveis  $d$  serão utilizadas para comparação com a saída da rede. Desta forma, esta terá  $D$  neurônios de entrada e  $n$  neurônios de saída. Cada instância representa a situação (número de ocorrência de determinadas doenças) de um determinado local dados os determinantes sociais de saúde relativos ao mesmo.

#### IV. METODOLOGIA

O banco de dados deste estudo contempla doenças de notificação compulsória, as quais devem ser registradas obrigatoriamente e de forma universal, no intuito de controlar imediatamente eventos que necessitam de pronta intervenção, e os determinantes sociais em saúde que foram selecionados de acordo com as seguintes variáveis: percentuais de residências com tratamento de água por filtragem, água fervida ou água clorada; percentual de residências de baixa renda, taxa de analfabetismo, percentuais de residências com tratamento de lixo por coleta ou lixo queimado, taxa de cidadãos com 2º grau completo ou mais; percentuais de residências com tratamento de resíduos em esgoto ou fossa.

Os dados deste banco foram obtidos a partir do DataSUS (Sinan - Sistema de Informação de Agravos de Notificação), da Secretaria de Saúde do Governo do Estado de Pernambuco e dos dados do último Censo do Instituto Brasileiro de Geografia e Estatísticas (IBGE). Diante do exposto, escolheu-se o ano de 2010 pela maior expressividade das informações.

Neste trabalho optou-se por utilizar cada instância da base de dados fazendo correspondência aos dados reais de um determinado município para a alimentação da rede neural. Com isso, foi desenvolvida a hipótese de que com a criação de um conjunto contendo todos os municípios do Estado de Pernambuco, a MLP seria capaz de representar, com uma precisão mínima aceitável, a função que relaciona os determinantes sociais de saúde com as doenças em qualquer cidade do estado. Os autores optaram pela abordagem da criação da base de dados relativa ao conjunto dos municípios pela pouca quantidade de dados referentes apenas a uma cidade. A escolha feita para o presente trabalho é validada na seção dos experimentos. Porém não fica excluída a possibilidade de um usuário do sistema realizar o treinamento da rede neural com a granularidade dos dados que achar mais viável para o seu domínio. É importante observar que o desempenho da rede neural depende da representatividade dos dados escolhidos para o cenário real de aplicação.

Os custos relacionados a cada determinante social de saúde estão descritos na Tabela I, lembrando que estes foram definidos empiricamente.

Tabela I. CUSTOS ASSOCIADOS PARA OS DETERMINANTES SOCIAIS DE SAÚDE

Determinante social de saúde	Valor associado ( $\alpha_i$ )
Tratamento de água Tipo 1 (Água Filtrada)	200
Tratamento de água Tipo 2 (Água Fervida)	50
Tratamento de água Tipo 3 (Água Clorada)	100
Residências com baixa renda	2000
Taxa de analfabetismo	2000
Tratamento de Lixo Tipo 1 (Lixo Coletado)	200
Tratamento de Lixo Tipo 2 (Lixo Queimado)	100
Taxa de cidadãos com 2º grau completo ou mais	2000
Tratamento de resíduos Tipo 1 (Fezes e Urina em Esgoto)	400
Tratamento de resíduos Tipo 2 (Fezes e Urina em Fossa)	200

Todas as variáveis, de entrada e saída, foram normalizadas entre 0,15 e 0,85. O método de validação utilizado foi validação cruzada. Foram utilizados 25% dos dados da base para a validação, 25% para teste e 50% para o treinamento da rede. O critério de parada para o treinamento da rede foi a estabilização do erro de validação. A MLP e o MOPSO-CDR foram configurados, para todos os experimentos, de acordo

com as Tabelas II e III, respectivamente. Experimentos prévios foram realizados com o objetivo de selecionar as melhores configurações de parâmetros.

Tabela II. PARÂMETROS DA MLP

Parâmetros	Valores
Taxa de aprendizado	0.2
Momentum	0.2
Quantidade neurônios na camada de entrada	10
Quantidade neurônios na camada escondida	número de doenças do experimento
Quantidade neurônios na camada escondida	19

Tabela III. PARÂMETROS DO MOPSO-CDR

Parâmetros	Valores
Tamanho da população	20
Tamanho máximo do arquivo externo	200
Taxa de mutação	0,5
Fator de inércia (decaimento linear)	0,9 - 0,4
Fator cognitivo	1.49445
Fator social	1.49445
Quantidade de iterações	1600

Neste trabalho, as soluções retornadas pelo MOPSO-CDR representam o percentual que deve ser acrescido ou decrescido em cada determinante. Por exemplo, uma solução  $\vec{S} = (0, 2; 0, 1; 0, 0; 0, 1; 0, 0; 0, 0; 0, 0; 0, 5; 0, 2; 0, 1; 0, 0)$  representa acréscimo de 20 pontos percentuais na taxa de residências localizadas no município que possuem água tratada por filtragem, 110 pontos percentuais na taxa de residências localizadas no município que possuem água tratada por fervura, nenhuma alteração no tratamento de água através do cloro, redução de 10 pontos percentuais na taxa de residências de baixa renda e assim por diante. A decisão de acréscimo ou decréscimo depende dos determinantes utilizados e do conhecimento prévio acerca dos mesmos.

Para o uso do sistema, este deve ter conhecimento da situação atual da região onde o sistema será usado. No cenário de teste criado para os experimentos realizados neste trabalho, foi criado um município fictício, sendo o seu vetor dos valores dos determinantes sociais de saúde definido por  $\vec{M} = (0, 3; 0, 02; 0, 6; 0, 4; 0, 2; 0, 7; 0, 1; 0, 5; 0, 4; 0, 4)$ .

O fato dos determinantes utilizados neste trabalho estarem representados em valores percentuais faz emergirem algumas restrições. Considerando  $d_i$  e  $\Delta d_i$  o determinante atual de índice  $i$  e a sua variação, respectivamente, e que  $S_i = d_i + \Delta d_i$ , tem-se que  $S_0 + S_1 + S_2 \leq 1, S_3 \leq 1, S_4 \leq 1, S_5 + S_6 \leq 1, S_7 \leq 1$  e  $S_8 + S_9 \leq 1$ . O MOPSO-CDR não é capaz de resolver problemas com restrições. Portanto, foi criado um mecanismo de penalização para soluções que não satisfazem pelo menos uma destas restrições. Neste mecanismo, estas soluções devem ter os valores dos seus dois objetivos penalizados através da soma dos seus valores normais (retornados pela MLP e pela função de cálculo de custo) com o produto da quantidade do percentual excedente da(s) restrição(ões) violada(s) com um valor numérico relativamente alto (i.e. 5000, para o caso deste trabalho). Dessa forma, um gradiente que decresce à medida que se aproxima das zonas permitidas do espaço de busca e aumenta ao distanciar-se de tal região, entrando cada vez mais nas zonas proibidas, permite que o algoritmo se afaste rapidamente das regiões com soluções não factíveis.

Foram realizados experimentos levando em consideração as seguintes quantidades de doenças: uma (Tuberculose), três

(Dengue, Tuberculose e Esquistossomose) e cinco doenças diferentes (Dengue, Tuberculose, Esquistossomose, Hanseníase e Hepatite), com o objetivo de avaliar a escalabilidade do sistema, tanto no contexto da rede neural tanto quanto no âmbito do otimizador de função. Todos os cenários foram testados para os mesmos 10 determinantes de saúde citados anteriormente. Para cada cenário, 30 simulações foram executadas. Os resultados dos experimentos estão detalhados na seção a seguir.

## V. RESULTADOS

A Figura 3 e a Tabela IV apresentam os resultados dos experimentos referentes ao módulo do aproximador de função. A Figura 3 é o gráfico do tipo *box-plot* representando o percentual de acerto da previsão da MLP, para as quantidades de doenças definidas na seção anterior, dados os determinantes sociais de saúde em um determinado município. Os valores de média e desvio padrão estão dispostos na Tabela IV. A partir da análise destas informações é possível notar que com o aumento do número de doenças a serem previstas, a rede neural perdeu eficácia, porém manteve uma taxa de acerto superior a 80%. Essa taxa de acerto pode ser considerada razoável, levando-se em consideração que os dados foram retirados de cenários reais referentes aos municípios do Estado de Pernambuco. Sendo assim, em todos os três casos de teste, a MLP foi capaz de aproximar de forma satisfatória a função que mapeia os determinantes sociais de saúde citados anteriormente nas doenças mencionadas na seção anterior. Sendo assim, a rede pôde assumir de forma eficaz o papel de uma das funções objetivo do MOPSO-CDR. Além disso, tais resultados validam a hipótese mencionada anteriormente de que o conjunto de dados referentes aos municípios do Estado de Pernambuco é representativo para qualquer município do estado.

Tabela IV. MÉDIA E DESVIO PADRÃO DAS TAXAS DE ACERTO DA MLP

Quantidade de doenças	Média e Desv. Padrão
Uma	87,9763 ± 0,5458
Três	83,7895 ± 0,3656
Cinco	80,31383579 ± 0,4983

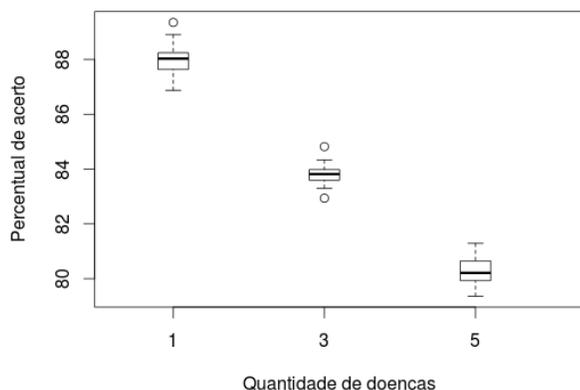


Figura 3. *Box-plot* com as taxas de acerto da MLP para os DSS para uma, três e cinco doenças.

As Figuras 4, 5, 6 mostram a disposição das soluções não dominadas presentes no Arquivo Externo na última iteração do MOPSO-CDR para os experimentos realizados levando em consideração, respectivamente uma, três e cinco doenças. Estas são as soluções que devem ser recomendadas ao gestor. Analisando os conjuntos de soluções em questão, percebe-se, primeiramente, formas similares de *paretos* com sugestões diversas. Além disso, é importante ressaltar que as soluções estão homoganeamente distribuídas ao longo destes, o que significa uma boa diversidade no conjunto de recomendações. Observa-se que o gestor, nestes experimentos, obteria desde soluções com custo zero e eficácia mínima (*i.e.* nenhuma alteração no número de ocorrência de doenças ou até aumento do mesmo) até soluções de alto custo e alta eficácia (*i.e.* grande redução de ocorrência de doenças). Dessa forma, as decisões poderiam ser tomadas de acordo com a disposição de recursos que o município detém. A tomada de decisão, por parte do gestor, estaria baseada nos custos de cada decisão e qual dessas escolhas teria a melhor relação custo-benefício, para o seu contexto. Isto é, uma ação que abrangesse a prevenção do maior número de doenças e que fosse financeiramente possível.

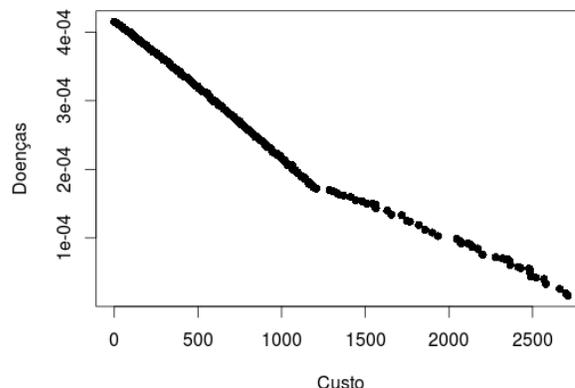


Figura 4. *Pareto front* com as soluções não dominadas encontradas pelo MOPSO-CDR para o experimento com uma doença.

## VI. CONSIDERAÇÕES FINAIS

Neste trabalho, um sistema de recomendação híbrido para gestão em saúde pública foi proposto. O algoritmo de otimização multiobjetivo MOPSO-CDR foi utilizado para sugerir ações a um gestor de saúde pública de forma que estas minimizem a ocorrência de determinadas doenças em uma região, mas levando em consideração o custo das mesmas. Uma MLP foi utilizada como função objetivo que estima o número de ocorrência de doenças dada uma determinada situação social. Para o cálculo dos custos, uma função determinada empiricamente foi utilizada. Como resultado, a MLP obteve uma precisão maior do que 80% para os três casos testados. Além disso, o MOPSO-CDR foi capaz de retornar uma diversificada gama de soluções ao gestor, levando em consideração a razão custo/benefício.

Nota-se que o sistema proposto neste trabalho obteve resultados interessantes nos cenários de teste utilizados, o que valida o uso desta abordagem no problema em questão.

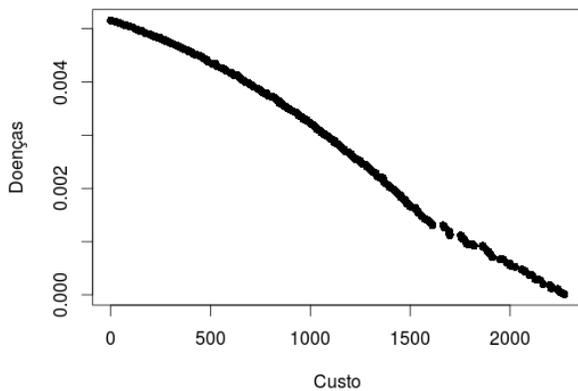


Figura 5. Pareto front com as soluções não dominadas encontradas pelo MOPSO-CDR para o experimento com três doenças.

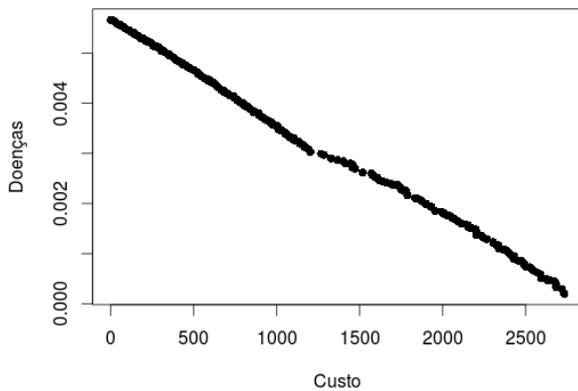


Figura 6. Pareto front com as soluções não dominadas encontradas pelo MOPSO-CDR para o experimento com cinco doenças.

Porém, os resultados obtidos podem ser melhorados e aplicados facilmente em situações reais. Como trabalhos futuros, os autores pretendem utilizar bases de dados de outras regiões do país, analisar funções de custos mais complexas e que se adequem melhor à realidade econômica da região. Além disso, pretende-se analisar o desempenho de diferentes algoritmos de aproximação de funções e de otimização multiobjetivo neste problema.

## REFERÊNCIAS

- [1] L. Penello, L. Lugarinho, and S. Rosario, "Parto humanizado e memória do nascimento: uma reflexão sobre a produção de saúde desde os seus primórdios," *Tempus Actas de Saúde Coletiva*, vol. 4, no. 4, 2010.
- [2] A. Pellegrini Filho, "Public policy and the social determinants of health: the challenge of the production and use of scientific evidence." *Cad Saude Publica*, vol. 27 Suppl 2, 2011.
- [3] Determinantes sociais da saúde ou por que alguns grupos da população são mais saudáveis que outros? Data de acesso: 20 de maio 2013. [Online]. Available: <http://www.determinantes.fiocruz.br>

- [4] M. Whitehead, "The concepts and principles of equity and health," *Health Promotion International*, vol. 6, no. 3, pp. 217–228, Jan. 1991.
- [5] "As causas sociais das iniquidades em saúde no Brasil," Comissão Nacional sobre Determinantes Sociais da Saúde, Tech. Rep., 2008.
- [6] J. E. Helm, M. Lapp, and B. D. See, "Characterizing an effective hospital admissions scheduling and control management system: a genetic algorithm approach," in *Proceedings of the Winter Simulation Conference*, ser. WSC '10. Winter Simulation Conference, 2010, pp. 2387–2398.
- [7] S. Shariff, N. H. Moin, and M. Omar, "Dynamic conditional location allocation model for healthcare facilities," in *Humanities, Science and Engineering Research (SHUSER), 2012 IEEE Symposium on*. IEEE, 2012, pp. 1389–1394.
- [8] G. Tai and P. Williams, "Product line management for health care system: Theoretic capacity planning over various resources," in *Industrial Engineering and Engineering Management, 2008. IEEM 2008. IEEE International Conference on*, 2008, pp. 690–694.
- [9] Q. Zheng, J. Shen, Z. qing Liu, K. Fang, and W. Xiang, "Resource allocation simulation on operating rooms of hospital," in *Industrial Engineering and Engineering Management (IE EM), 2011 IEEE 18th International Conference on*, vol. Part 3, 2011, pp. 1744–1748.
- [10] C. Rodriguez-Verjan, V. Augusto, T. Garaix, X. Xie, and V. Buthion, "Healthcare at home facility location-allocation problem," in *Automation Science and Engineering (CASE), 2012 IEEE International Conference on*, 2012, pp. 150–155.
- [11] C. A. C. Coello and M. Reyes-Sierra, "Multi-Objective Particle Swarm Optimizers: A Survey of the State-of-the-Art," *International Journal of Computational Intelligence Research*, vol. 2, no. 3, pp. 287–308, 2006.
- [12] R. A. Santana, M. R. Pontes, and C. J. A. Bastos-Filho, "A multiple objective particle swarm optimization approach using crowding distance and roulette wheel," in *Ninth International Conference on Intelligent Systems Design and Applications, 2009. ISDA '09*, december 2009, pp. 237–242.
- [13] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Neural Networks, 1995. Proceedings., IEEE International Conference on*, vol. 4. IEEE, Nov. 1995, pp. 1942–1948 vol.4.
- [14] C. Tsou, S. Chang, and P. Lai, *Using Crowding Distance to Improve Multi-Objective PSO with Local Search - Swarm Intelligence, Focus on Ant and Particle Swarm Optimization*. IN-Tech Education and Publishing, 2007.
- [15] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biology*, vol. 5, no. 4, pp. 115–133, Dec. 1943.
- [16] A. K. Jain, J. Mao, and K. Mohiuddin, "Artificial neural networks: A tutorial," *IEEE Computer*, vol. 29, pp. 31–44, 1996.
- [17] A. P. Engelbrecht, *Computational Intelligence: An Introduction*, 2nd ed. Wiley Publishing, 2007.
- [18] P. Werbos, "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences," Ph.D. dissertation, Harvard University, Cambridge, MA, 1974.
- [19] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Neurocomputing: foundations of research." Cambridge, MA, USA: MIT Press, 1988, ch. Learning representations by back-propagating errors, pp. 696–699.