

Título do Trabalho

Aluno: Vitor de Carvalho Melo Lopes

Orientador: Prof. Dr. Antônio Mauro Barbosa de Oliveira

Programa de Pós-Graduação em Ciência da Computação (PPGCC)
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

3 de março de 2017

Agenda

- 1 Introdução
- 2 Objetivos Geral e Específicos
- 3 Fundamentação Teórica
- 4 Metaheurísticas
- 5 Métodos de Treinamento de SVMs
- 6 Simulações Computacionais
- 7 Conclusões e Trabalhos futuros

Introdução

- Envelhecimento da população e as abordagens para o tratamento clínico;

Motivação

- Motivação;

Descrição do problema

- Descrição do problema;

Objetivos Geral e Específicos

Objetivo Geral

- Oferecer uma solução que auxilie, através de sistemas de informática, o cuidador, o doente e à equipe médica no tratamento do paciente em domicílio.

Objetivos Específicos

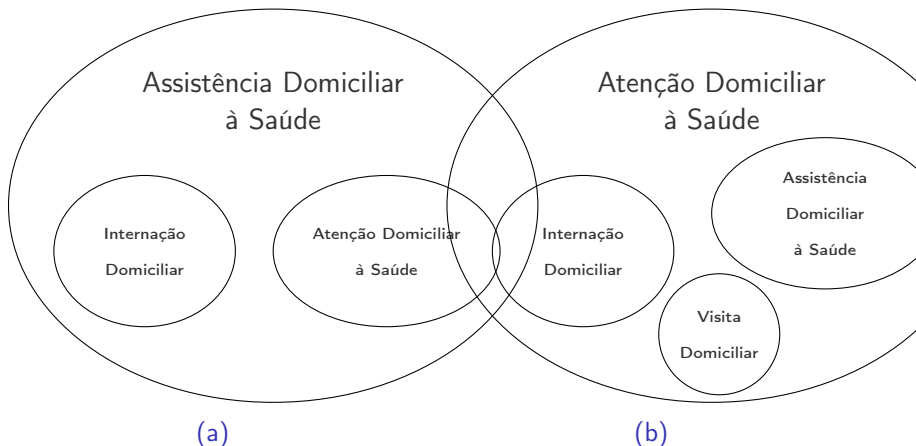
- Permitir um acompanhamento constante da equipe médica ao paciente;
- Facilitar as atividades diárias do cuidador em relação aos procedimentos destinados ao doente;

Aspectos de Saúde

- Hospitalização no Brasil
- Assistência Domiciliar à Saúde:
 - Terminologia
 - Envolvidos;

Terminologia ilustrada

Figura 1: (a) Representação gráfica das categorias defendida por Tavolari, Fernandes e Medina; (b) Representação gráfica das categorias defendida por Giacomozzi.



Implementação típica

Figura 3: Fluxograma canônico de um SA.

Implementação típica

- Representação da solução;
- Parâmetros de configuração do SA:
 - Função de energia $E(\theta)$;
 - Função de agendamento de temperaturas $W(T)$;
 - Temperatura inicial T_0 e final T_f ; e
 - Estrutura de vizinhança.
- Soluções infactíveis:
 - Penalização;
 - Eliminação;
 - Reparo das soluções; e
 - Métodos híbridos.

Conceitos iniciais

- Método computacional de busca probabilística;
- Proposto por ??) e popularizado por ??);
- Inspirado nos mecanismos de evolução natural e na genética;
- Metaheurística populacional.

Implementação típica

Figura 4: Fluxograma canônico de um GA.

Implementação típica

- Componentes necessários:
 - Representação genética do problema;
 - Função de aptidão;
 - Mecanismo de seleção;
 - Operadores de mutação e reprodução e
 - Critério de parada.

Métodos de Treinamento de SVMs

- Programação quadrática(QP)
 - Métodos matemáticos clássicos (??);
 - Ponto interior, conjunto ativo, Lagrangiano aumentado, etc.
 - Requer o armazenamento de uma matriz de dimensões $l \times l$.
- *Sequential Minimal Optimization* (SMO)
 - Método iterativo, proposto por ??);
 - Particiona o problema de otimização nos menores subproblemas;
 - Menor subproblema possível envolve dois multiplicadores;
 - Número excessivo de combinações: $l(l-1)$.

Métodos de Treinamento de SVMs

- *Kernel Adatron (KA)*

- Método de aprendizado online, proposto por ??
- Topologia similar a arquitetura de uma rede neural;
- Utiliza apenas informação de primeira ordem da função custo;
- A restrição apresentada na Equação (??) não é respeitada.

- *Linear Particle Swarm Optimization (LPSO)*

- Baseado em no LPSO, proposto por ??;
- Decompoe o problema de otimização em um subconjunto de trabalho;
- Repete a otimização até que o valor da função objetivo não mude;
- A restrição apresentada na Equação (??) não é respeitada.

Proposta 1: *Genetic Algorithms for Training Easily*

- *Genetic Algorithms for Training Easily* (GATE) SVMs
- Algoritmos Genéticos mono objetivo
- Soluciona o problema de otimização quadrática
- As entradas do GATE são o conjunto de dados de treinamento e o parâmetro de regularização C (Similarmente ao QP, SMO, KA)

Algoritmo GATE

1. Gerar **população inicial** de soluções aleatoriamente
2. Aplicar o **algoritmo de ajuste** (se necessário)
3. Repetir até que o critério de parada for alcançado
 - i. Selecionar os melhores indivíduos através de **elitismo**
 - ii. Aplicar os operadores de **recombinação** e **mutação**
 - iii. Se necessário, aplicar o algoritmo de ajuste nos indivíduos mutados
4. Computar o **viés** a partir dos multiplicadores de Lagrange

Proposta 1: *Genetic Algorithms for Training Easily*

- Representação genética: direta
- População inicial: gerada aleatoriamente
- Função de aptidão: rearranjo da função objetivo
- Mecanismo de seleção: roleta padrão
- Operador de recombinação: reprodução aritmética
- Operador de mutação: *swap*
- Critério de parada: número de gerações e monitor de estagnação
- Cálculo do viés: $w_0 = \frac{1}{I} \sum_{i=1}^I \left(y_i + \sum_{j=1}^I y_j \alpha_j \mathbf{x}_i^T \mathbf{x}_j \right)$
- Violação das restrições:
 - População inicial e Operador de mutação
 - **algoritmo de ajuste**: redução iterativa dos valores dos multiplicadores

Algoritmo de ajuste

Figura 5: Exemplo de violação da restrição (??) para um indivíduo gerado aleatoriamente e seu reparo através da função de ajuste.

Proposta 2: *Simulated Annealing for Training Easily*

- *Simulated Annealing for Training Easily* (SATE) SVMs
- Baseado no método de Recozimento Simulado
- Soluciona o problema de otimização quadrática
- As entradas do SATE são o conjunto de dados de treinamento e o parâmetro de regularização C (Similarmente ao QP, SMO, KA)
- Requisitos de um algoritmo baseado em SA:
 - Representação da solução e solução inicial;
 - Função de energia; e
 - Função de vizinhança.
- Cálculo do viés (mesma forma que o GATE).

Proposta 2: *Simulated Annealing for Training Easily*

- Representação da solução: direta;
- Solução inicial: apenas dois multiplicadores;
- Função de energia: rearranjo da função objetivo;
- Função de vizinhança:
 - Baseada em três funções auxiliares: troca, adição e modificação;
 - Prioriza soluções esparsas.
- Cálculo do viés: similar ao GATE
- Violação das restrições:
 - Função de troca
 - **algoritmo de ajuste**: similar ao GATE

Experimento 1: prova de conceito

- Problema Artificial I (PAI)
 - Problema binário;
 - 100 instâncias em cada classe;
 - Linearmente separável.
- Problema Artificial II (PAII)
 - Problema binário;
 - 400 instâncias em cada classe;
 - Não linearmente separável.

Experimentos 2 e 3: conjuntos de dados reais

- Conjuntos de dados binários (tabelas ?? e ??);
- Normalização de média zero e desvio padrão igual a um;
- Avaliação feita através do método *holdout* com divisão 80% – 20%;
- Hiperparâmetros obtidos através de validação cruzada:
 - *k-fold*, com $k = 5$;
 - Variação de C no intervalo $\{2^{-5}, \dots, 2^{15}\}$.
- 30 execuções pareadas;
- Avaliação de equivalência realizada pelo teste de Friedman.

Experimento 2: conjuntos de dados reais¹

¹O método GATE não foi avaliado para os conjuntos de dados RIP e BKM.

Experimento 3: conjuntos de dados MNIST²

²O método GATE não foi avaliado para o conjuntos de dados MNIST.

Experimento 4 e 5:

- Conjuntos de dados HAB e BCW;
- Experimento 4: variação da divisão treino/teste;
- Experimento 5: variação do hiperparâmetro C .

Experimento 1: prova de conceito

(a) QP

(b) SMO

Figura 6: Superfícies de decisão e vetores suporte geradas pelos métodos de treinamento a partir de problema artificial I.

Experimento 1: prova de conceito

(a) KA

(b) LPSO

Figura 7: Superfícies de decisão e vetores suporte geradas pelos métodos de treinamento a partir de problema artificial I.

Experimento 1: prova de conceito

(a) GATE

(b) SATE

Figura 8: Superfícies de decisão e vetores suporte geradas pelos métodos de treinamento a partir de problema artificial I.

Experimento 1: prova de conceito

(a) QP

(b) SMO

Figura 9: Superfícies de decisão e vetores suporte geradas pelos métodos de treinamento a partir de problema artificial II.

Experimento 1: prova de conceito

(a) KA

(b) LPSO

Figura 10: Superfícies de decisão e vetores suporte geradas pelos métodos de treinamento a partir de problema artificial II.

Experimento 1: prova de conceito

(a) GATE

(b) SATE

Figura 11: Superfícies de decisão e vetores suporte geradas pelos métodos de treinamento a partir de problema artificial II.

Experimento 1: prova de conceito

Experimento 2: conjuntos de dados reais

Experimento 2: conjuntos de dados reais

Experimento 2: conjuntos de dados reais

Figura 12: Taxa de acerto (%) e # Vetores Suporte (%) do SATE, GATE, LPSO, SMO, QP e KA para os conjuntos de dados AUS, BCW, HAB e HEA.

Experimento 2: conjuntos de dados reais

Figura 13: Taxa de acerto (%) e # Vetores Suporte (%) do SATE, GATE, LPSO, SMO, QP e KA para os conjuntos de dados PID, SON e VCP.

Experimento 2: conjuntos de dados reais

Experimento 2: conjuntos de dados reais

Experimento 2: conjuntos de dados reais

Experimento 3: conjunto de dados MNIST

Experimento 3: conjunto de dados MNIST

Experimento 3: conjunto de dados MNIST

Experimento 4: variação da divisão treino/teste

(a)

(b)

Figura 14: Taxa de acerto (# VS) do SATE, SMO, KA, QP e LPSO *versus* percentagem de treino, para o conjunto de dados HAB no lado esquerdo (direito).

Experimento 4: variação da divisão treino/teste

(a)

(b)

Figura 15: Taxa de acerto (# VS) do SATE, SMO, KA, QP e LPSO *versus* percentagem de treino, para o conjunto de dados BCW no lado esquerdo (direito).

Experimento 5: variação do hiperparâmetro C

(a)

(b)

Figura 16: Taxa de acerto ($\#$ VS) do SATE, SMO, KA e QP *versus* o hiperparâmetro C para o conjunto de dados HAB no lado esquerdo (direito).

Experimento 5: variação do hiperparâmetro C

(a)

(b)

Figura 17: Taxa de acerto ($\#$ VS) do SATE, SMO, KA e QP *versus* o hiperparâmetros C para o conjunto de dados BCW no lado esquerdo (direito).

Conclusões e Trabalhos futuros

Conclusões

- Desenvolvimento de duas abordagens de treinamento para o SVM;
- Restrições do SVM foram incorporadas com êxito nos métodos;
- Desenvolvimento de um algoritmo de ajuste que eleva a esparsidade;
- GATE e SATE
 - Equivalentes ao SMO em termos de $\#VS$ e Taxa de acerto;
 - Mais esparsos que o KA, o QP e o LPSO.

Trabalhos futuros

- Geração de soluções iniciais mais próximas a solução ideal;
- Paralelização de alguns componentes do método GATE;
- Tornar os classificadores multiclass;
- Usar funções de *kernel* não lineares.

Referências I