

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CÂMPUS CORNÉLIO PROCÓPIO
DIRETORIA DE GRADUAÇÃO E EDUCAÇÃO PROFISSIONAL
DEPARTAMENTO DE COMPUTAÇÃO
ENGENHARIA DE COMPUTAÇÃO**

RAFAEL WILSON DANTAS DA SILVA

**CONSTRUÇÃO DE ALGORITMOS BAYESIANOS UTILIZANDO ALGORITMOS
EVOLUTIVOS**

TRABALHO DE CONCLUSÃO DE CURSO

**CORNÉLIO PROCÓPIO
2017**

RAFAEL WILSON DANTAS DA SILVA

**CONSTRUÇÃO DE ALGORITMOS BAYESIANOS UTILIZANDO ALGORITMOS
EVOLUTIVOS**

Trabalho de Conclusão de Curso apresentada ao Departamento de Computação da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de “Bacharel em Engenharia de Computação”.

Orientador: Prof. Dr. Danilo Sipoli Sanches

Co-orientador: Prof. Dr. Carlos Nascimento Silla Jr.

CORNÉLIO PROCÓPIO

2017

RESUMO

SILVA, Rafael. **Construção de algoritmos bayesianos utilizando algoritmos evolutivos**. 2017. 21 f. Trabalho de Conclusão de Curso – Engenharia de Computação, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2017.

Atualmente os classificadores bayesianos tem recebido bastante destaque no campo de Mineração de Dados pelo seu ótimo desempenho. A independência entre seus atributos faz com que o processo de mineração torne-se rápido. Porém esta independência não pode ser aplicada a problemas do mundo real. Assim surgiram modificações aos classificadores Bayesianos que sugerindo possíveis relações de dependência entre os atributos. Porém encontrar o melhor classificador a se utilizar gera um problema Np-Completo. Uma das possíveis abordagens, que tem-se utilizado para solucionar este problema, é a utilização de algoritmos evolutivos na busca para uma solução boa.

Palavras-chave: Redes Bayesianas. Classificadores Bayesianos. Classificadores Bayesianos K-dependentes

ABSTRACT

SOBRENOME, Nome. **Construction of Bayesian algorithms using evolutionary algorithms.** 2017. 21 f. Master Thesis – Electrical Engineering Graduate Program, Federal University of Technology - Paraná. Cornélio Procópio, 2017.

Nowadays the Bayesian classifiers have received a lot of attention in the field of Data Mining for their excellent performance. The independence of its attributes makes the mining process fast. But this independence can not be applied to real-world problems. Hence modifications appeared to the Bayesian classifiers that suggest possible dependency relations between the attributes. However finding the best classifier to use generates an Np-Complete problem. One of the possible approaches that has been used to solve this problem is the use of evolutionary algorithms in the search for a good solution

Keywords: Bayes Network. Bayes Classifiers. K-Dependence Bayes classifiers

LISTA DE FIGURAS

FIGURA 1 – Estrutura da rede do Naive Bayes	12
FIGURA 2 – Algumas possíveis formas geradas pelo KDBC com 3 Atributos e $K = 2$ (2-BDC).	13
FIGURA 3 – Fluxograma de um Típico Algoritmo Evolutivo	15
FIGURA 4 – Demonstração da capacidade da colônia de formigas em encontrar o melhor caminho entre dois pontos <i>Nest</i> e <i>Food</i>	17

LISTA DE TABELAS

TABELA 1 – Atividades	19
TABELA 2 – Cronograma de atividades	19

LISTA DE SIGLAS

ACO	Ant Colony Optimization
AE	Algoritmos Evolucionários
AG	Algoritmos Genéticos
GRASP	<i>Greedy Randomized Adaptive Search Procedures</i>
KDBC	<i>K-Dependence Bayesians Classifier</i>
NB	<i>Naive Bayes</i>
SA	inglês Simulated Annealing
VNS	<i>Variable Neighborhood Search</i>

SUMÁRIO

1	INTRODUÇÃO	8
1.1	ORGANIZAÇÃO DO TEXTO	8
1.2	PROBLEMA	8
1.3	JUSTIFICATIVA	9
2	CLASSIFICADORES BAYESIANOS	10
2.1	TEOREMA DE BAYES	10
2.2	REDES BAYESIANAS	11
2.3	CLASSIFICADORES BAYESIANOS	11
2.4	K-DEPENDENCE BAYESIAN CLASSIFIER	12
3	HEURÍSTICAS E META-HEURÍSTICAS	14
3.1	ALGORITMOS EVOLUCIONÁRIOS	14
3.1.1	ALGORITMOS GENÉTICOS	14
3.2	ARREFICIMENTO SIMULADO	15
3.3	VNS	16
3.4	GRASP	16
3.5	OTIMIZAÇÃO POR COLÔNIA DE FORMIGAS	17
4	PROPOSTA	18
4.1	ATIVIDADES PLANEJADAS	18
	REFERÊNCIAS	20

1 INTRODUÇÃO

Nos tempos atuais estamos sobrecarregados de dados. A quantidade de dados no mundo e em nossas vidas tende a crescer cada vez mais e não há nenhum sinal de que esta estimativa reduza. Os computadores atuais tornam o processo de armazenamento de dados muito simples para descartarmos qualquer que seja o dado produzido no cotidiano de nossos dias. Há uma grande diferença entre a quantidade de dados que produzimos e a quantidade de informação que conseguimos retirar dela (WITTEN; FRANK; HALL, 2017).

Mineração de dados é sobre resolver problemas por meio da análise de dados já presentes na base de dados e um dos seus objetivos é a descoberta de conhecimento através de técnicas computacionais, que são capazes de explorar um grande conjunto de dados evidenciando padrões e auxiliando na descoberta de conhecimento. Este desenfreado crescimento de base de dados, traz a mineração de dados para o primeiro plano das novas tecnologias (WITTEN; FRANK; HALL, 2017).

A construção de um classificador é uma tarefa básica na análise de dados e reconhecimento de padrões na Mineração de Dados. Simplificadamente, um classificador é uma função que assemelha com base em um conjunto de atributos um valor ao rótulo da classe (FRIEDMAN; GEIGER; GOLDSZMIDT, 1997).

As Redes Bayesianas (serão explicadas na seção 2.2) tem se destacado como uma das abordagens mais promissoras no processo de descoberta de conhecimento em base de dados e uma das suas implementações mais conhecidas é o classificador *Naive Bayes* (NB) (SEBASTIANI; ABAD; RAMONI, 2010). O Naive Bayes é um classificador probabilístico e tem como principal pressuposto a independência entre seus atributos.

1.1 ORGANIZAÇÃO DO TEXTO

Na seção 1.2 serão abordados os problemas que motivaram o desenvolvimento deste trabalho.

A justificativa para a execução deste trabalho encontra-se na seção 1.3.

Nas seções 2, 3.1 e 3.5 serão introduzidos conceitos sobre os algoritmos que envolvem este trabalho.

Na seção 4 será abordado detalhes da proposta deste trabalho, como cronograma e atividades a serem realizadas.

1.2 PROBLEMA

No situações reais a independência de atributos é irreal. Diante da limitação de independência entre os atributos, diversos trabalhos foram realizados sobre possíveis modificações na estrutura do NB. Um dos métodos de melhoria dos resultados obtidos pelo NB é extensão da estrutura do classificador, que consiste na representação de dependências entre os atributos (JIANG et al., 2007).

Uma forma de extensão dos classificadores Bayesianos foi apresentada por Sahami (1996), a fim de lidar com as suposições de independência entre os atributos do NB, onde ele introduz o termo *K-Dependence Bayesian Classifier* (KDBC), onde o valor K representa o

número máximo de nós pais que um atributo pode ter além da classe. Dessa forma um 0-KDBC seria o equivalente a um Naive Bayes (FLORES et al., 2011).

Em (COOPER, 1990) foi provado que tentar encontrar a melhor rede bayesiana gerada pelo KDBC é um problema que se encontra no domínio dos NP-Completo, ou seja, é inviável computacionalmente tentar encontrar a melhor solução.

1.3 JUSTIFICATIVA

A procura por algoritmos meta-heurísticos, que são capazes de encontrar boas soluções realizando uma busca no espaço das possíveis soluções, tem aumentado para resolver os complexos problemas existentes no mundo real (LARRANNAGA et al., 2013).

Os algoritmos evolutivos são um dos mais bem sucedidos a conseguir resultados em diversos campos de atuação. Aplicando seus mecanismos inspirados na evolução natureza, por exemplo, a sobrevivência do mais apto ou cruzamento genético e mutação, em uma população de soluções candidatas, abordagens evolutivas como algoritmos genéticos foram capazes de realizar uma busca mais eficaz e diversificada no espaço de soluções para problemas difíceis (LARRANNAGA et al., 2013).

Com isso, visto o notável sucesso dos algoritmos evolutivos em problemas considerados difíceis, este trabalho propõe a utilização de algoritmos Evolutivos e seus derivados como uma possível solução para o problema dos classificadores bayesianos gerados pelo algoritmo KDBC.

2 CLASSIFICADORES BAYESIANOS

2.1 TEOREMA DE BAYES

Se A , B e C são eventos, a probabilidade de que estes eventos ocorram, $P(A)$, $P(B)$ e $P(C)$ respectivamente, pode ser representada por um número real entre 0 e 1. A probabilidade de que um evento está ligada diretamente a quantidade de vezes em que ele ocorre em relação a quantidade de vezes em que um experimento é realizado. Logo, se um evento E ocorre N vezes em um experimento realizado M vezes, então temos que $P(E) = M/N$ (DOUGHERTY, 2012).

Para entender melhor o Teorema de Bayes que será explicado neste capítulo considere as seguintes definições:

- **Eventos mutuamente exclusivos**

Eventos mutuamente exclusivos são eventos que não podem ocorrer ao mesmo tempo. Se E e F são mutuamente exclusivos, então a probabilidade da união entre eles é dada pela seguinte equação:

$$P(E \cup F) = P(E) + P(F) \quad (1)$$

Da mesma forma podemos ter eventos que não sejam mutuamente exclusivos, ou seja, podem ocorrer ao mesmo tempo. Neste caso eles são representados pela seguinte equação:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F) \quad (2)$$

Onde, $P(E \cap F)$ é a verificação de que os eventos sejam independentes, definida pela equação 3:

$$P(E \cap F) = P(E) \cdot P(F) \quad (3)$$

- **Probabilidade Condicional**

A probabilidade condicional é a probabilidade de um evento E ocorrer em função da ocorrência de outro evento F . A sua representação é dado por $P(E|F)$ e é lida como “a probabilidade de E, dado que F é verdadeiro” (DOUGHERTY, 2012). Uma das formas de calcular a probabilidade condicional entre os eventos E e F é dada pela seguinte equação:

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \quad (4)$$

E da mesma forma para calcularmos $P(F|E)$, temos:

$$P(F|E) = \frac{P(E \cap F)}{P(E)} \quad (5)$$

- **Regra do Produto de Probabilidades**

Se aplicarmos a multiplicação cruzada nas equações 4 e 5, obtemos a seguinte equação:

$$P(E \cap F) = P(E|F) \cdot P(F) = P(F|E) \cdot P(E) \quad (6)$$

Dadas as considerações das a cima, manipulando os dois termos mais a direita da equação 6 chegamos enfim na equação do **Teorema de Bayes** como pode ser visto a seguir:

$$P(E|F) = \frac{P(F|E) \cdot P(E)}{P(F)} \quad (7)$$

Onde a probabilidade de $P(E|F)$ é conhecida como **probabilidade posteriori** de um determinado evento E ocorrer dado a ocorrência do evento F a partir do conhecimento de estimativas das probabilidades $P(E)$ e $P(F)$ e do conhecimento da probabilidade condicional $P(F|E)$. A probabilidade posteriori pode ser expressa informalmente como:

$$posteriori = \frac{verosimilhanca \cdot priori}{evidencia} \quad (8)$$

Com base no Teorema de Bayes surgiram os classificadores Bayesianos que serão explicados nas seções 2.2 e 2.3.

2.2 REDES BAYESIANAS

Segundo (SEBASTIANI; ABAD; RAMONI, 2010), atualmente as redes Bayesianos são uma das abordagens mais promissoras para o processo de descoberta do conhecimento.

As redes bayesianas pertencem a uma classe mais geral de modelos chamados de modelos probabilísticos gráficos que surgiram da combinação da teoria de grafos e da teoria da probabilidade. O seu sucesso se deve a sua capacidade de lidar com modelos probabilísticos complexos através da decomposição em componentes menores e acessíveis. Um modelo probabilístico gráfico é definido por uma grafo onde os nós representam variáveis estocásticas e os arcos representam as dependências entre tais variáveis. Esses arcos são marcados pela probabilidade de distribuição de interação entre as variáveis vinculadas (SEBASTIANI; ABAD; RAMONI, 2010).

Uma rede Bayesiana é um modelo gráfico probabilístico onde o grafo conectando suas variáveis é um Grafo Acíclico Dirigido (do inglês, DAG - *Directed Acyclic Graph*). Este grafo gerado representa suposições de independência condicional entre os atributos que são usados para fatorar a distribuição de probabilidade conjunta das variáveis da rede. Assim o processo de aprendizagem em bancos de dados, com grandes quantidades de dados torna-se possível. (SEBASTIANI; ABAD; RAMONI, 2010).

2.3 CLASSIFICADORES BAYESIANOS

Um dos classificadores bayesianos mais conhecidos é o Naive Bayes (NB). Este classificador aprende do conjunto de treinamento a probabilidade condicional de cada atributo A devida ao rótulo da classe C . A classificação é feita então aplicando o teorema de Bayes para calcular a probabilidade da classe C , dada a instância de A_1, \dots, A_n e então prediz a classe com a maior

probabilidade a posteriori. Este cálculo é baseado em uma suposição de independência onde os atributos A_i são condicionalmente independentes dado o valor da classe C (FRIEDMAN; GEIGER; GOLDSZMIDT, 1997).

Quando representado por uma rede Bayesiana a estrutura do NB pode ser visto na figura 1. Essa rede consegue representar o principal pressuposto por trás do classificador NB, de que os atributos são independentes entre si e dependem apenas da classe.

Naive Bayes é um dos mais efetivos classificadores, levando em conta o seu desempenho, atualmente. O desempenho do NB é surpreendente dado que a suposição de independência entre os atributos é quase sempre irreal (FRIEDMAN; GEIGER; GOLDSZMIDT, 1997).

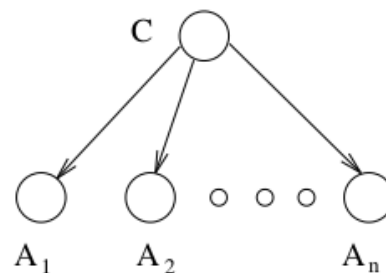


Figura 1 – Estrutura da rede do Naive Bayes

Fonte: (FRIEDMAN; GEIGER; GOLDSZMIDT, 1997).

O desempenho do NB pode ser prejudicado quando a suposição de independência entre os atributos é violada. Sendo assim diversos trabalhos tem surgido afim aperfeiçoar os classificadores Bayesianos. As técnicas aplicadas nestes trabalhos podem ser divididas da seguinte maneira (JIANG et al., 2007):

- **Seleção de Características:** Seleciona e agrupa os atributos em conjuntos nos quais os atributos satisfazem a relação de independência dos mesmo;
- **Extensão da Estrutura:** Estende a estrutura do NB afim de representar as possíveis dependências entre os atributos;
- **Aprendizagem Local:** Aplica o conceito de treinamento local a fim de encontrar um sub-conjunto local e utilizar o classificador NB;
- **Expansão de Dados:** Expandindo os dados de treinamento e construir o NB utilizando os dados de treinamento expandidos.

O método que iremos analisar neste trabalho será o de Extensão da Estrutura, em especial o algoritmo KDBC que será explicado na seção 2.4.

2.4 K-DEPENDENCE BAYESIAN CLASSIFIER

Estendendo a estrutura do NB e utilizando arcos para explicitamente representar as dependências entre os atributos é uma forma direta de relaxar a suposição de que os atributos devem ser independentes entre si.

Em (SAHAMI, 1996) introduz o conceito do algoritmo KDBC, mostrando que a probabilidade de cada atributo é condicionada a classe e até K outros atributos, ou seja, cada atributo poderá ter de 0 a K dependências com outros atributos (FLORES et al., 2011). Aplicar esta extensão dos classificadores bayesianos, pode gerar um número inacessível computacionalmente de classificadores.

Na figura 2 é mostrado um exemplo onde temos $k = 2$ para um problema com 3 atributos. Neste caso cada atributo poderá depender, além da classe, de mais 0, 1 ou 2 atributos. Na demonstração da figura 2 não foram desenhadas todas as possíveis redes pra este exemplo.

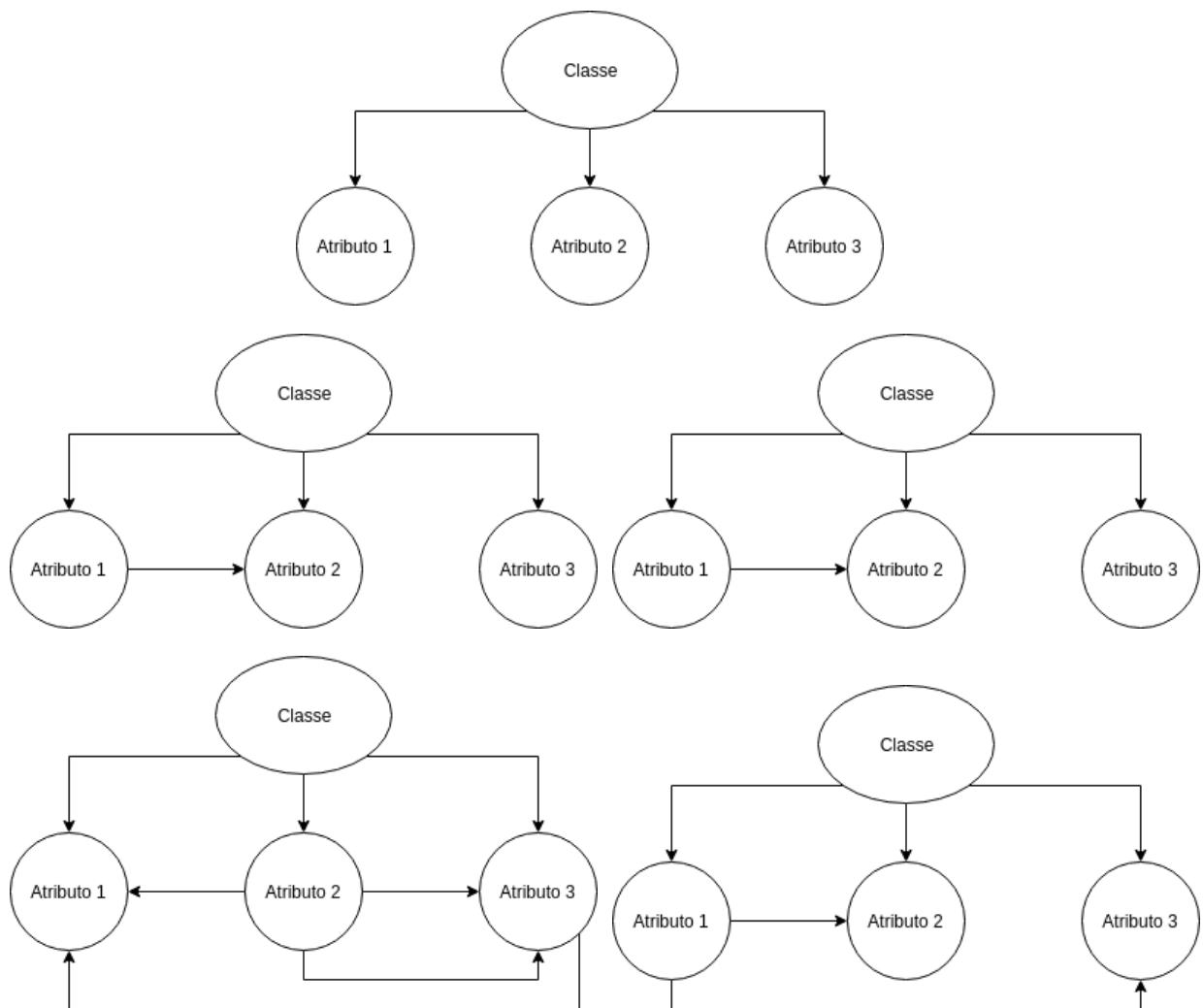


Figura 2 – Algumas possíveis formas geradas pelo KDBC com 3 Atributos e $K = 2$ (2-BDC).

Fonte: Autoria Própria

3 HEURÍSTICAS E META-HEURÍSTICAS

Heurísticas são abordagens aproximativas cujo objetivo é a melhora no desempenho de soluções que buscam resolver problemas NP-completos.

Um método Heurístico é uma abordagem cujo objetivo é encontrar uma solução localmente ótima para um determinado problema computacional. A busca heurística usa informações do problema para direcionar o algoritmo a locais promissores do espaço de busca. A principal característica dos métodos heurísticos consiste em encontrar “boas” soluções em tempo computacional razoável bem como minimizar o espaço ocupado pelas configurações e estados que um determinado algoritmo pode assumir (CAMPELLO; MACULAN, 1994).

As Heurísticas podem ser classificadas de três formas: construtivas, de melhoramento e meta-heurísticas.

Uma heurística é construtiva quando ela parte de uma solução vazia e considera a cada iteração apenas o próximo passo e após inserir uma partícula na solução, não é mais possível retirá-la. Já as de melhoramento partem de uma solução já construída e ela irá trabalhar no melhoramento da solução atual (ZANAKIS; EVANS, 1981).

As meta-heurísticas são métodos de busca que combinam outras formas de heurísticas para resolver determinado e particular problema.

3.1 ALGORITMOS EVOLUCIONÁRIOS

Algoritmos Evolucionários (AE) baseiam-se no processo de evolução natural proposto por Charles Darwin (1850) para criar modelos computacionais na resolução de problemas. Processo este que mantém uma população de indivíduos ou cromossomos que se comporta de forma semelhante à evolução das espécies. Cada indivíduo recebe uma avaliação que é quantificada em relação a solução do problema em questão, para então ser aplicado os operadores genéticos de forma a simular a sobrevivência dos indivíduos no meio. Os Operadores genéticos são aproximações de fenômenos vistos na natureza. Denomina-se gerações, todas as vezes que um indivíduo é exposto a essas etapas (LINDEN, 2008).

Há uma grande variedade de modelos computacionais propostos na literatura, porém a grande maioria deles tem em comum a aplicação dos conceitos de seleção, mutação e reprodução na simulação da evolução das espécies. Estes processos dependem do desempenho dos indivíduos de cada espécie dentro do ambiente (LINDEN, 2008).

A figura 3 representa a mais comum representação de um algoritmo evolutivo.

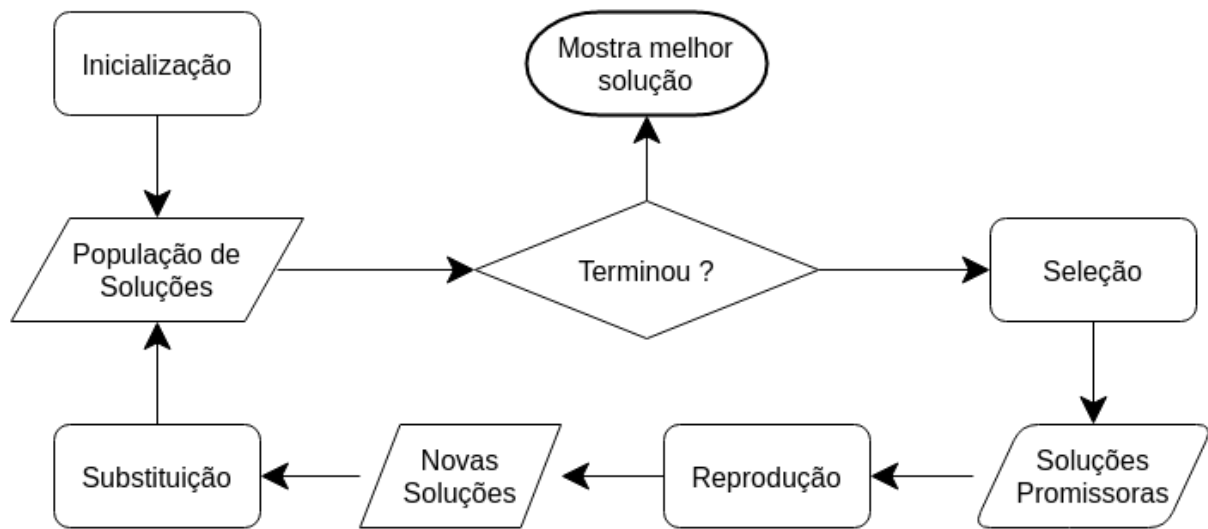
3.1.1 ALGORITMOS GENÉTICOS

Segundo (LINDEN, 2008), Algoritmos Genéticos (AG) são um ramo dos Algoritmos Evolucionários e como tal podem ser definidos como técnica de busca baseada numa metáfora do processo biológico natural. Os GAs por serem de otimização global, opõem métodos como o *hill climbing*, que segue a derivada de uma função ao ponto de encontrar o máximo de uma função.

Os algoritmos Genéticos são técnicas heurísticas de otimização global. Algoritmos Genéticos são eficientes em buscar, no espaço das soluções, as que sejam tão próximas da

Figura 3 – Fluxograma de um Típico Algoritmo Evolutivo

Fonte: Releitura de (LARRANNAGA et al., 2013)



solução ótimo quanto possível e isso quase sempre sem a interação humana. Portanto os algoritmos Genéticos são uma técnica adequada para problemas especialmente difíceis, como os problemas denominados NP completos (LINDEN, 2008).

O funcionamento básico de um AG é através da evolução de uma população de candidato soluções para o problema através de um número de gerações, a fim de obter melhores soluções. As soluções são geralmente representadas como sequências binárias. O algoritmo seleciona um subconjunto de soluções de adaptação a partir da população de acordo com um mecanismo de seleção. Estas soluções parentais são usadas para reproduzir novas soluções de prole aplicando operadores genéticos como crossover e mutação. As soluções recentemente geradas competem então com as soluções na população para a sobrevivência de acordo com sua aptidão (LARRANNAGA et al., 2013).

3.2 ARREFICIMENTO SIMULADO

O Arrefecimento simulado (inglês Simulated Annealing (SA)) é uma meta-heurística probabilística que segue uma linha de raciocínio baseado no sistema físico da termodinâmica. A técnica consiste em aquecer metais a altas temperaturas e permitir seu arrefecimento de forma lenta o suficiente para que seus átomos consigam se organizar em um estado de energia mínima. Com isso o material ganha rigidez e consistência (LAARHOVEN; AARTS, 1987).

Ao realizar uma comparação do modelo computacional do SA com a abordagem física do problema podemos ter as seguintes comparações:

- (1) Os possíveis estados do metal significam as soluções no espaço de busca;
- (2) A energia em cada estado física significa o valor da função objetivo;
- (3) A energia mínima ou máxima corresponde ao valor de uma solução ótima local, possivelmente global.

Ao considerar que o objetivo é a minimização, o SA tem o funcionamento básica da seguinte forma. A cada iteração é gerada um novo estado a partir do estado que atual, de forma aleatória. Então é comparado a energia entre os estados, como estamos em um problema de minimização, se o novo estado tiver energia menor que o atual ele será escolhido como o estado corrente. Caso a energia do atual seja menor do que a energia do estado gerado, é aplicado algum método à de determinar a probabilidade de mudança de estados.

3.3 VNS

Variable Neighborhood Search (VNS) é uma meta-heurística baseada na exploração de múltiplas definições de vizinhança impostas ao mesmo espaço de solução. O algoritmo funciona em cima de basicamente duas etapas, a etapa de agitação, onde um vizinha da solução atual é gerado aleatoriamente, e a etapa de busca local, onde é aplicada uma busca local em cima da solução gerada na etapa de agitação. O algoritmo vai gradativamente explorando vizinhanças cada vez mais distantes e focaliza a busca em torno de uma nova solução somente se um movimento de melhora é realizado (MLADENović; HANSEN, 1997).

O algoritmo VNS básico pode ser definido em aproximadamente 4 passos:

- **Passo 1.** seleciona-se um conjunto de estruturas de vizinhança $N_k = (k = 1, k = 2, \dots, k_{max})$ e defini-se uma solução inicial x ;
- **Passo 2.** gera-se uma solução $x \in N_k(x)$ aleatoriamente;
- **Passo 3.** aplica-se busca local em x' com objetivo de encontrar um ótimo local x'' ;
- **Passo 4.** se x'' for melhor que x , então x'' torna-se a solução atual.

A arquitetura do VNS apesar de simples e possuir poucos parâmetros pode ser implementada utilizando organizações mais sofisticadas ou mesmo hibridizadas com outras meta-heurísticas.

3.4 GRASP

Greedy Randomized Adaptive Search Procedures (GRASP) é uma meta-heurística que implementa uma hibridização entre um algoritmo semi-guloso, que está interessado nas melhores soluções de cada iteração, e um método de busca local. A cada iteração abrange basicamente de duas fases: construção e busca local. A fase de construção cria uma solução viável, cuja vizinhança é investigada até que um mínimo local seja encontrado durante a fase de busca local. A melhor solução global é mantida como resultado. GRASP e VNS são de alguma forma complementares, no sentido de que a randomização é aplicada no GRASP na fase de construção, enquanto que o VNS é aplicada na fase de busca local local (FEO; RESENDE, 1995). A fase construtiva é executada através da estratégia gulosa aleatorizada e pode ser realizada através de diferentes abordagens. Na fase de busca local, o GRASP pode empregar qualquer processo de busca heurística, no entanto, o algoritmo mais utilizado na literatura é o VNS.

O GRASP tenta solucionar o problema de diversidade que acompanha a busca gulosa ao mesmo tempo que permite a construção de soluções qualitativamente melhores quando

comparado à inicialização aleatória. Nem sempre é possível encontrar uma solução realizável, assim é necessário que seja realizado a invocação de algum método que seja capaz de reparar a solução de modo que ela se torne uma solução factível. Alternativamente este reparo pode ser simplesmente o descarte da solução e a aplicação da busca semi-gulosa novamente até que se ache uma solução factível (HIRSCH et al., 2007).

3.5 OTIMIZAÇÃO POR COLÔNIA DE FORMIGAS

Introduzido por Marico Dorigo e colegas no começos dos anos de 1990 (DORIGO; MANIEZZO; COLORNI, 1996)(DORIGO; BLUM, 2005), o Ant Colony Optimization (ACO)(em português, Otimização por Colônia de Formigas ou Otimização Colônia de Formigas) é baseado no comportamento das formigas ao se organizarem para encontrar o melhor percurso entre dois pontos, como por exemplo o lugar onde buscam comida e o ninho.

Inicialmente as formigas percorrem a área em busca de comida de forma randômica. Enquanto se movem as formigas deixam pelo caminho um rastro de feromônio, que pode ser percebido pelas outras formigas. Quando a formiga encontra uma fonte de alimento, ela avalia a qualidade e quantidade do alimento e carrega o alimento de volta ao ninho. No caminho de retorno o feromônio produzido pela formiga pode depender da qualidade e da quantidade de comida encontrada e como a formiga que encontrou o melhor caminho retornará antes, o rastro deixado por ela será mais forte. Probabilisticamente as formigas tendem a seguir o caminhos com grandes quantidades de feromônio. Assim as outras formigas serão guiadas pelo melhor caminho (DRÉO; SIARRY, 2002) (BLUM, 2005). A demonstração do problema pode ser visto na figura 4 .

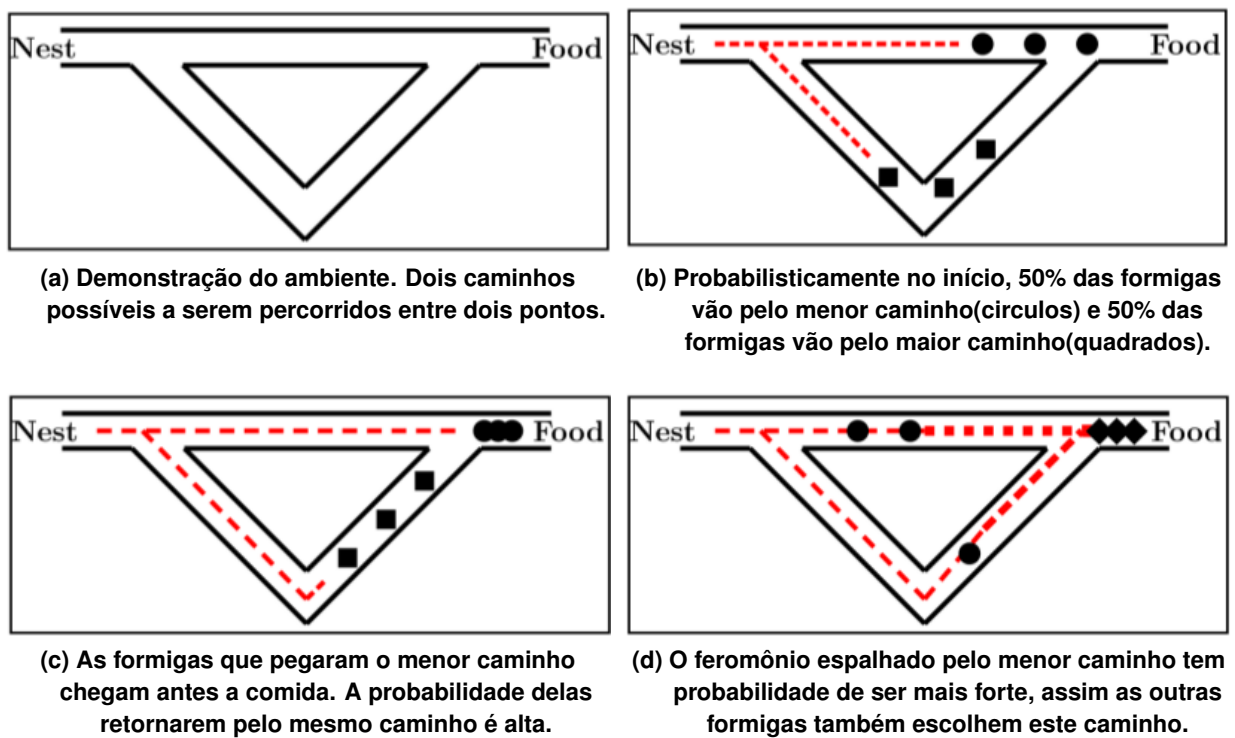


Figura 4 – Demonstração da capacidade da colônia de formigas em encontrar o melhor caminho entre dois pontos *Nest* e *Food*

Fonte: (BLUM, 2005)

4 PROPOSTA

Em (FLORES et al., 2011) propõem o uso de um tipo de AGs, que considera que todas as variáveis são independentes na hora da construção do modelo, para buscar por melhores representações de dependências entre os atributos preditores nas Redes Bayesianas.

Enquanto em (SIERRA; nAGA, 1998) é proposta o uso de AGs para encontrar a melhor relação entre os nós pais e filhos para atributos classe de problemas do mundo real. Eles compararam os resultados de seus classificadores com o NB e com classificadores bayesianos aprendidos por maximização de vizinhança. Quando comparados foi verificado que os classificadores gerados por seus métodos utilizando GAs obtiveram um número maior de acerto do que os outros classificadores. Este método foi utilizado em um esquema multi-classificador para classificar dados de pacientes da unidade de terapia intensiva.

Em (PENA; LOZANO; LARRANAGA, 2004) foi mostrado o uso de GAs na busca estruturas mais precisas para Redes Bayesianas em problemas de predição de Tromboembolismo (coágulo de sangue nas veias, mais conhecido popularmente como trombose).

As redes bayesianas são uma importante classe de modelos probabilísticos que se mostram muito eficazes em domínios incertos. Elas são estudadas ao longo das últimas três décadas e muitos métodos têm sido propostos para automatizar sua aprendizagem e inferência. No entanto, muitos destes métodos envolvem difíceis problemas de pesquisa combinatória que afetam diretamente seu uso generalizado, especialmente para grandes instâncias de problema, e assim requerem técnicas avançadas de busca como meta-heurísticas (LARRANNAGA et al., 2013).

Algoritmos evolutivos, são métodos de busca estocástica de propósito geral inspirados na evolução natural e frequentemente aplicado para resolver muitos problemas complexos do mundo real. Devido às suas vantagens, diferentes tipos de algoritmos evolutivos têm sido utilizados nas redes Bayesianas (LARRANNAGA et al., 2013).

Um dos pontos chaves da utilização de Algoritmos evolutivos combinados com redes bayesianas, é a escolha da representação das redes como indivíduos. Pois a representação do indivíduo pode impactar de maneira grande o desempenho computacional do algoritmos. Como consequência dessa necessidade vários trabalhos focam boa parte de seus estudos em maneiras de representação de grafos como indivíduos em Algoritmos Evolutivos.

Este trabalho pretende encontrar uma representação para redes bayesianas que se adeque aos algoritmos utilizados. Assim este trabalho propõe a aplicação de Algoritmos Evolutivos, métodos de busca heurísticas e meta-heurísticas para melhorar ou buscar pela melhor rede bayesiana. Este trabalho também abre para a possibilidade da utilização de algoritmos de busca local para melhorar as relações de dependência entre os atributos.

4.1 ATIVIDADES PLANEJADAS

Durante este trabalho pretende-se realizar as atividades descritas na tabela 1 bem como a ordem e o tempo em que ocorrerão descritos na tabela 2.

REFERÊNCIAS

- BLUM, Christian. **Ant colony optimization: Introduction and recent trends**. 2005. Citado na página 17.
- CAMPELLO, R.E.; MACULAN, N. **Algoritmos e heurísticas: desenvolvimento e avaliação de performance**. EDUFF, 1994. ISBN 9788522801343. Disponível em: <<https://books.google.com.br/books?id=oUOBPgAACAAJ>>. Citado na página 14.
- COOPER, Gregory F. The computational complexity of probabilistic inference using bayesian belief networks (research note). **Artif. Intell.**, Elsevier Science Publishers Ltd., Essex, UK, v. 42, n. 2-3, p. 393–405, mar. 1990. ISSN 0004-3702. Disponível em: <[http://dx.doi.org/10.1016/0004-3702\(90\)90060-D](http://dx.doi.org/10.1016/0004-3702(90)90060-D)>. Citado na página 9.
- DORIGO, Marco; BLUM, Christian. Ant colony optimization theory: A survey. **Theoretical Computer Science**, v. 344, n. 2, p. 243 – 278, 2005. ISSN 0304-3975. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0304397505003798>>. Citado na página 17.
- DORIGO, M.; MANIEZZO, V.; COLORNI, A. Ant system: optimization by a colony of cooperating agents. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, v. 26, n. 1, p. 29–41, Feb 1996. ISSN 1083-4419. Citado na página 17.
- DOUGHERTY, Geoff. **Pattern Recognition and Classification: An Introduction**. [S.l.]: Springer Publishing Company, Incorporated, 2012. ISBN 1461453224, 9781461453222. Citado na página 10.
- DRÉO, Johann; SIARRY, Patrick. A new ant colony algorithm using the heterarchical concept aimed at optimization of multim minima continuous functions. In: _____. **Ant Algorithms: Third International Workshop, ANTS 2002 Brussels, Belgium, September 12–14, 2002 Proceedings**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. p. 216–221. ISBN 978-3-540-45724-4. Disponível em: <http://dx.doi.org/10.1007/3-540-45724-0_18>. Citado na página 17.
- FEO, Thomas A.; RESENDE, Mauricio G. C. Greedy randomized adaptive search procedures. **Journal of Global Optimization**, v. 6, n. 2, p. 109–133, 1995. ISSN 1573-2916. Citado na página 16.
- FLORES, M. Julia et al. Handling numeric attributes when comparing bayesian network classifiers: does the discretization method matter? **Applied Intelligence**, v. 34, n. 3, p. 372–385, 2011. ISSN 1573-7497. Disponível em: <<http://dx.doi.org/10.1007/s10489-011-0286-z>>. Citado 3 vezes nas páginas 8, 13 e 18.
- FRIEDMAN, Nir; GEIGER, Dan; GOLDSZMIDT, Moises. Bayesian network classifiers. **Machine Learning**, v. 29, n. 2, p. 131–163, 1997. ISSN 1573-0565. Disponível em: <<http://dx.doi.org/10.1023/A:1007465528199>>. Citado 3 vezes nas páginas 8, 11 e 12.
- HIRSCH, M. J. et al. Global optimization by continuous grasp. **Optimization Letters**, v. 1, n. 2, p. 201–212, 2007. ISSN 1862-4480. Citado na página 17.
- JIANG, Liangxiao et al. Survey of improving naive bayes for classification. In: _____. **Advanced Data Mining and Applications: Third International Conference, ADMA 2007 Harbin, China, August 6-8, 2007. Proceedings**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. p. 134–145. ISBN 978-3-540-73871-8. Disponível em: <http://dx.doi.org/10.1007/978-3-540-73871-8_14>. Citado 2 vezes nas páginas 8 e 12.

LAARHOVEN, Peter J. M. van; AARTS, Emile H. L. Simulated annealing. In: _____. **Simulated Annealing: Theory and Applications**. Dordrecht: Springer Netherlands, 1987. p. 7–15. ISBN 978-94-015-7744-1. Disponível em: <http://dx.doi.org/10.1007/978-94-015-7744-1_2>. Citado na página 15.

LARRANAGA, Pedro et al. A review on evolutionary algorithms in bayesian network learning and inference tasks. **Information Sciences**, v. 233, p. 109 – 125, 2013. ISSN 0020-0255. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0020025513000443>>. Citado 3 vezes nas páginas 9, 15 e 18.

LINDEN, Ricardo. **Algoritmos Genéticos**: Uma importante ferramenta de inteligência computacional. São Paulo: Brasport, 2008. Citado 2 vezes nas páginas 14 e 15.

MLADENOVÍĆ, N.; HANSEN, P. Variable neighborhood search. **Comput. Oper. Res.**, Elsevier Science Ltd., Oxford, UK, v. 24, n. 11, p. 1097–1100, nov. 1997. ISSN 0305-0548. Citado na página 16.

PENA, J. M.; LOZANO, J. A.; LARRANAGA, P. Unsupervised learning of bayesian networks via estimation of distribution algorithms: An application to gene expression data clustering. **Int. J. Uncertain. Fuzziness Knowl.-Based Syst.**, World Scientific Publishing Co., Inc., River Edge, NJ, USA, v. 12, n. 1 supp, p. 63–82, jan. 2004. ISSN 0218-4885. Disponível em: <<http://dx.doi.org/10.1142/S0218488504002588>>. Citado na página 18.

SAHAMI, Mehran. Learning limited dependence bayesian classifiers. In: **Proceedings of the Second International Conference on Knowledge Discovery and Data Mining**. AAAI Press, 1996. (KDD'96), p. 335–338. Disponível em: <<http://dl.acm.org/citation.cfm?id=3001460.3001537>>. Citado na página 12.

SEBASTIANI, Paola; ABAD, Maria M.; RAMONI, Marco F. Bayesian networks. In: _____. **Data Mining and Knowledge Discovery Handbook**. Boston, MA: Springer US, 2010. p. 175–208. ISBN 978-0-387-09823-4. Disponível em: <http://dx.doi.org/10.1007/978-0-387-09823-4_10>. Citado 2 vezes nas páginas 8 e 11.

SIERRA, B.; nAGA, P. LarraPredicting survival in malignant skin melanoma using Bayesian networks automatically induced by genetic algorithms. An empirical comparison between different approaches. **Artif Intell Med**, Department of Computer Science and Artificial Intelligence, University of the Basque Country, San Sebastian, Spain. ccpsiarb@si.ehu.es, v. 14, n. 1-2, p. 215–230, 1998. ISSN 0933-3657. Citado na página 18.

WITTEN, Ian H.; FRANK, Eibe; HALL, Mark A. **Data Mining: Practical Machine Learning Tools and Techniques**. 4rd. ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2017. ISBN 0123748569, 9780123748560. Citado na página 8.

ZANAKIS, Stelios H.; EVANS, James R. Heuristic optimization: Why, when, and how to use it. **Interfaces**, v. 11, n. 5, p. 84–91, 1981. Citado na página 14.