

# Classificação de Alto Nível Utilizando Grafo K-associados Ótimo

Murillo G. Carneiro \*, João L. G. Rosa, Alneu A. Lopes, and Liang Zhao

Instituto de Ciências Matemáticas e Computação,  
Universidade de São Paulo, São Carlos - Brasil  
{carneiro, joaoluis, alneu, zhao}@icmc.usp.br

**Resumo** O grafo k-associados ótimo caracteriza-se por não ser paramétrico, possuir estrutura de otimização local baseada na medida de pureza e ser bastante robusto a ruídos e *outliers*. O classificador baseado no grafo k-associados emprega um método probabilístico (classificador bayesiano ótimo) que não é apto a considerar os padrões de formação das classes, uma vez que é constituído apenas a partir de atributos dos dados. Por outro lado, a classificação de alto nível utiliza medidas de rede para capturar as relações espaciais, funcionais, topológicas e outras, existentes em muitos conjuntos de dados, de modo a classificá-los de maneira eficiente. Uma fusão dos conceitos relacionados ao grafo k-associados ótimo e à classificação de alto nível é apresentada neste artigo: CAN-KAOG (Classificação de Alto Nível utilizando o Grafo K-Associados Ótimo). Foram realizados experimentos com bases de dados artificiais e reais e os resultados apontam que CAN-KAOG é uma técnica promissora para classificação em conjuntos de dados com padrões de formação.

**Palavras Chaves:** grafo k-associados ótimo, classificação de alto nível, redes complexas, aprendizado supervisionado

## 1 Introdução

Redes complexas são grafos cuja organização e topologia não são triviais [1]. Elas possuem aplicabilidade em vários ramos da ciência e podem ser úteis nas análises dos mais diversos aspectos de redes e com os mais variados propósitos. Além disso, reúnem conceitos de estatística, sistemas complexos e teoria dos grafos e possuem a capacidade de capturar as relações espaciais, funcionais, topológicas e outras, existentes em muitos conjuntos de dados.

No contexto de aprendizado de máquina, redes complexas têm sido bastante utilizadas em abordagens do tipo não supervisionado e semi-supervisionado, porém existem poucos trabalhos na literatura que abordam-nas no contexto supervisionado [2]. Apesar disso, os resultados apresentados por estes trabalhos mostram que o aprendizado supervisionado baseado em grafos possui aplicação prática em diversos cenários tais como redes sociais e reconhecimento de padrões [3,2,4].

---

\* Agradecimentos à FAPESP pelo suporte financeiro.

Em geral, a grande maioria das técnicas de aprendizado, quer baseadas em grafos ou não, são dependentes de seleção de modelos (*tunning*). Contudo, recentemente, foi proposta uma técnica não paramétrica, denominada classificador baseado no grafo k-associados ótimo (KAOG) [2]. Outros aspectos importantes desta técnica são: (i) estrutura de otimização local baseada na medida de pureza e (ii) mecanismo robusto para lidar com ruídos e *outliers* no conjunto de dados.

A classificação, após construído o grafo k-associados ótimo, é feita utilizando o classificador bayesiano ótimo (CBO). Todavia, um ponto negativo do CBO e de várias outras técnicas (árvore de decisão e redes neurais artificiais, por exemplo) é que elas geralmente realizam a classificação exclusivamente a partir das características físicas dos dados (proximidade). Essa forma limitada de fazer a classificação dos dados é conhecida como classificação de baixo nível [4]. Por outro lado, um tipo diferente de classificação que considera o padrão de formação das classes e a estrutura topológica da rede é apresentado em [4] e denominado classificação de alto nível. Assim, a hipótese de pesquisa investigada neste trabalho é o emprego de classificação de alto nível utilizando o grafo k-associados ótimo. De fato, as duas técnicas apresentam aspectos importantes para a tarefa de classificação e um modelo capaz de unificá-las pode atingir bons resultados.

A organização do artigo é exibida a seguir: a Seção 2 aborda o grafo k-associados ótimo e a classificação de alto nível; a Seção 3 apresenta CAN-KAOG, o novo método proposto; a Seção 4 mostra os resultados experimentais obtidos; e a Seção 5 traz as considerações finais do trabalho.

## 2 Aprendizado supervisionado baseado em grafos

### 2.1 Grafo k-associados ótimo para classificação de dados

O algoritmo 1 apresenta os passos para construção do grafo k-associados, necessário para obtenção do grafo k-associados ótimo. Ele recebe como entrada um conjunto de dados e um parâmetro  $K$ , relacionado ao número de vizinhos a serem considerados. Nas linhas um e dois, inicializam-se os componentes e o grafo k-associados. Entre as linhas três e seis, tem-se um laço que determina as conexões para cada vértice/instância  $v_i$ , a princípio verificando os  $K$  vizinhos mais próximos de  $v_i$  ( $A_{v_i, K}$ ) e a igualdade entre as classes. Na linha sete há uma função para encontrar cada componente do grafo. O restante do algoritmo adiciona cada um dos componentes encontrados ao grafo k-associados e calcula uma medida de pureza sobre eles, retornando tais informações como saída do algoritmo.

Para determinar a medida de pureza de um componente, é preciso considerar o percentual de ligações que ocorrem dentro do componente em relação ao número total de conexões possíveis (soma das conexões que ocorrem e que não ocorrem). Assim, a fórmula para cálculo da pureza de um componente é dada por  $\Phi_\alpha = \frac{D_\alpha}{2K}$ , onde  $D_\alpha$  representa o grau médio de todas as instâncias que pertencem ao componente  $C_\alpha$ . Ver [2] para maiores detalhes.

Em relação ao grafo k-associados, é importante notar que raramente um grafo obtido por um único valor de  $K$  produzirá a melhor configuração de instâncias

**Algoritmo 1** Construção do grafo k-associados (Kac)**Require:** K e conjunto de dados X

- 
- 1:  $C \leftarrow \emptyset$
  - 2:  $G^{(K)} \leftarrow \emptyset$
  - 3: **for all**  $v_i \in V$  **do**
  - 4:  $\Delta_{v_i, K} \leftarrow \{v_j | v_j \in A_{v_i, K} \text{ and } classe_j = classe_i\}$
  - 5:  $E \leftarrow E \cup \{e_{i,j} | v_j \in \Delta_{v_i, K}\}$
  - 6: **end for**
  - 7:  $C \leftarrow encontraComponentes(V, E)$
  - 8: **for all**  $C_\alpha \in C$  **do**
  - 9:  $\Phi_\alpha \leftarrow calculaPureza(C_\alpha)$
  - 10:  $G^{(K)} \leftarrow G^{(K)} \cup \{(C_\alpha(V', E'); \Phi_\alpha)\}$
  - 11: **end for**
  - 12: **return** Grafo k-associados  $G^{(K)}$
- 

nos componentes para um dado conjunto de dados, pois estes componentes normalmente diferem entre si em termos de organização e disposição no espaço, ou seja, limitar a estrutura e a pureza dos componentes a um único valor de K, na maioria das vezes, não reflete as características de conjuntos de dados reais. Para lidar com este problema, tem-se o grafo k-associados ótimo.

O Algoritmo 2 apresenta um esboço da técnica para construção do grafo k-associados ótimo. O algoritmo não utiliza nenhum parâmetro e utiliza como entrada apenas o conjunto de dados de treino para construção do classificador. Nas primeiras linhas, K é inicializado com valor 1 e o grafo 1-associado é tomado como o grafo ótimo ( $G^{opt}$ ). A partir daí, inicia-se um laço cuja condição de término é uma heurística que visa garantir a continuação do algoritmo até que ele alcance o grafo com grau máximo de pureza [2]. Enquanto tal condição de parada não é satisfeita K é incrementado. Entre as linhas sete e doze, o algoritmo verifica, para cada um dos componentes deste grafo k-associados ( $C_\beta^{(K)}$ ), se a condição dada pela linha oito do algoritmo é satisfeita. Caso seja, na linha nove, tem-se uma operação para remover do grafo ótimo cada um dos componentes que compõe  $C_\beta^{(K)}$ , ou seja,  $C_\alpha^{(K)} \subseteq C_\beta^{(K+z)}$ . Na linha seguinte, o novo componente é acrescentado ao  $G^{opt}$ . Por fim, o algoritmo retorna cada um dos componentes obtidos, com seus respectivos valores de K e medidas de pureza.

Construído o grafo k-associados ótimo, um classificador bayesiano ótimo (CBO) é empregado para fazer a classificação de novas instâncias utilizando o grafo. Assim, supondo uma nova instância  $v_y$ , é possível classificá-la calculando a sua probabilidade de pertencer a cada um dos componentes. Dessa forma, considerando a teoria de Bayes, a probabilidade a posteriori de  $v_y$  pertencer ao componente  $C_\alpha$  dado todos os vizinhos observados deste novo caso ( $A_{v_y}$ ), é representada pela Equação 1.

$$P(v_y \in C_\alpha | A_{v_y}) = \frac{P(A_{v_y} | v_y \in C_\alpha) P(v_y \in C_\alpha)}{P(A_{v_y})} \quad (1)$$

**Algoritmo 2** Construção do grafo k-associados ótimo**Require:** Conjunto de dados X

---

```

1:  $K \leftarrow 1$ 
2:  $G^{(Ot)} \leftarrow Kac(K, X)$ 
3: repeat
4:    $ultimoGrauMedio \leftarrow D^{(K)}$ 
5:    $K \leftarrow K + 1$ 
6:    $G^{(K)} \leftarrow Kac(K, X)$ 
7:   for all  $C_\beta^{(K)} \subset G^{(K)}$  do
8:     if  $\Phi_\beta^{(K)} \geq \Phi_\alpha^{(Ot)}$  para todo  $C_\alpha^{(Ot)} \subseteq C_\beta^{(K)}$  then
9:        $G^{(Ot)} \leftarrow G^{(Ot)} - \cup_{C_\alpha^{(Ot)} \subseteq C_\beta^{(K)}} C_\alpha^{(Ot)}$ 
10:       $G^{(Ot)} \leftarrow G^{(Ot)} \cup \{C_\beta^{(K)}\}$ 
11:     end if
12:   end for
13: until  $D^{(K)} - ultimoGrauMedio < D^{(K)}/K$ 
14: return Grafo k-associados ótimo  $G^{(Ot)}$ 

```

---

De acordo com [2], a probabilidade da vizinhança  $A_{v_y}$  condicionada ao fato de que  $v_y \in C_\alpha$  é dada pela Equação 2, o termo de normalização  $P(A_{v_y})$  pode ser obtido conforme apresentado na Equação 3 e a probabilidade a priori  $P(v_y \in C_\alpha)$  é expressa pela Equação 4. Além disso, caso o número de componentes seja maior que o número de classes, tem-se a Equação 5. Ao final, escolhe-se a classe com maior probabilidade a posteriori.

$$P(A_{v_y}|v_y \in C_\alpha) = \frac{|A_{v_y, K_\alpha} \cap C_\alpha|}{K_\alpha} \quad (2)$$

$$P(A_{v_y}) = \sum_{N_{v_y, C_\beta} \neq 0} P(A_{v_y}|v_y \in C_\alpha)P(v_y \in C_\alpha) \quad (3)$$

$$P(v_y \in C_\alpha) = \frac{\Phi_\alpha}{\sum_{N_{v_y, C_\beta} \neq 0} \Phi_\beta} \quad (4)$$

$$P(v_y|\omega_i) = \sum_{class(C_\alpha)=\omega_i} P(v_y \in C_\alpha|A_{v_y}) \quad (5)$$

**2.2 Classificação de alto nível**

A classificação de alto nível considera não apenas os atributos físicos mas também as características semânticas globais dos dados. Matematicamente, a associação da instância de teste  $x_i \in X_{test}$  da classe  $j \in L$  produzida pelo classificador de alto nível  $C_i^{(j)}$  é dada pela Equação 6, onde  $\alpha(u) \in [0, 1], \forall u \in \{1, \dots, K\}$  é um coeficiente controlado pelo usuário que indica a influência de cada medida de rede no processo de classificação e  $f_i^{(j)}(u)$  é uma função que depende das medidas

de rede aplicadas na classe  $j$ . Para o classificador  $C$  ser válido, é necessário que  $\sum_{u=1}^K \alpha(u) = 1$  [4].

$$C_i^{(j)} = \frac{\sum_{u=1}^K \alpha(u)[1 - f_i^{(j)}(u)]}{\sum_{g \in L} \sum_{u=1}^K \alpha(u)[1 - f_i^{(g)}(u)]} \quad (6)$$

Em relação a função  $f_i^{(j)}(u)$ , esta possui uma forma mais geral dada pela Equação 7, onde  $\Delta G_i^{(j)}(u) \in [0, 1]$  é a variação que ocorre na medida da rede sempre que um novo caso  $x_i$  é inserido e  $p^{(j)} \in [0, 1]$  é a proporção das instâncias que pertencem a classe  $j$ , conforme é exibido na Equação 8 .

$$f_i^{(j)}(u) = \Delta G_i^{(j)}(u)p^{(j)} \quad (7)$$

$$p^{(j)} = \frac{1}{V} \sum_{u=1}^V \mathbf{1}_{\{y_u=j\}} \quad (8)$$

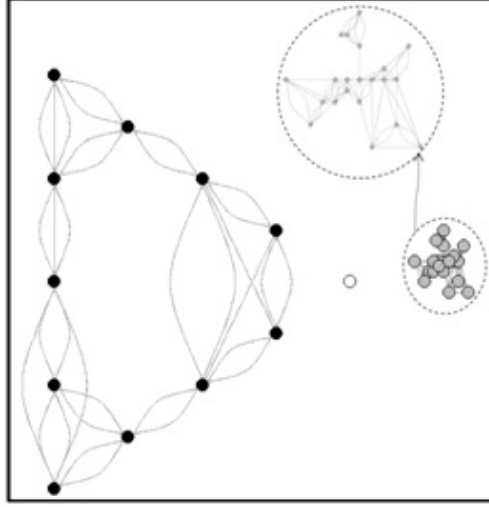
As medidas de rede ( $\Delta G_i^{(j)}$ ) têm o objetivo de verificar se o novo exemplo inserido está em conformidade com o padrão de formação da classe. Assim, a classificação de alto nível examina se a inserção de uma nova instância causa (ou não) uma grande variação nas medidas do componente. Em outras palavras, se houver uma pequena mudança nas medidas de rede, a instância testada está em conformidade com todos os outros casos do componente ( $f_i^{(j)}$  recebe valor pequeno), enquanto que se houver uma variação significativa nas medidas de rede do componente ( $f_i^{(j)}$  recebe valor grande), a instância testada provavelmente não pertence àquela classe. Considerando a Equação 6, a nova instância é classificada na classe/componente que apresenta maior valor de  $C_i^{(j)}$ . Exemplos de medidas de rede consideradas são: assortatividade, coeficiente de aglomeração e grau médio [1]. Em [4] é exibido em maiores detalhes o processo de classificação.

### 3 CAN-KAOG

#### 3.1 Motivação

O classificador bayesiano ótimo é baseado em probabilidade e não considera os padrões de formação das classes, pois é constituído a partir de atributos físicos dos dados. Um exemplo pode ser visto na Figura 1 que apresenta um conjunto de instâncias rotuladas em duas classes: cinza e preto. Uma nova instância  $v_y$  é inserida (sem rótulo) e deseja-se classificá-la. Os valores  $k$  encontrados pelo algoritmo para cada componente foram:  $K_{preto} = 3$  e  $K_{cinza} = 4$ ; e as purezas associadas a cada componente ( $C_{preto}$  e  $C_{cinza}$ ):  $\Phi_{preto} = 1$  e  $\Phi_{cinza} = 1$ .

A princípio, é possível perceber na Figura 1 que as instâncias pretas formam um triângulo, enquanto que as cinzas apresentam um comportamento indefinido onde os casos estão relativamente próximos. Além disso, as arestas entre as instâncias representam as ligações realizadas pelo grafo  $k$ -associados ótimo. A partir



**Figura 1.** Primeiro experimento: grafo  $k$ -associados ótimo obtido em um conjunto de instâncias rotuladas em duas classes: cinza e preto.

delas, o classificador bayesiano ótimo é utilizado para realizar a classificação da nova instância. Inicialmente, é verificada a distância do novo caso para os demais. Na figura, as cinco instâncias mais próximas são, respectivamente, das classes cinza, cinza, preto, preto e cinza. Assim, de acordo com a Equação 4, o classificador bayesiano ótimo calcula as probabilidades a priori de  $v_y$  pertencer ao componente preto  $P(v_y \in C_{preto}) = 0.5$  e cinza  $P(v_y \in C_{cinza}) = 0.5$ , conforme Equação 4.

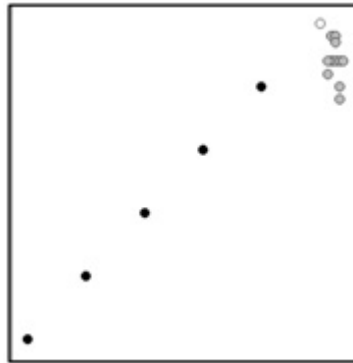
A probabilidade da vizinhança  $A_{v_y}$  dado que  $v_y \in C_\alpha$  (Equação 2) é:  $P(A_{v_y}|v_y \in C_{preto}) = \frac{1}{3}$  e  $P(A_{v_y}|v_y \in C_{cinza}) = \frac{2}{4}$ . Logo, o termo de normalização (Equação 3) pode ser calculado:  $P(A_{v_y}) \approx 0.416$ . Assim como na Equação 1, a probabilidade a posteriori do novo vértice  $v_y$  pertencer a um determinado componente é calculada:  $P(v_y \in C_{preto}|A_{v_y}) \approx \mathbf{0.40}$  e  $P(v_y \in C_{cinza}|A_{v_y}) \approx \mathbf{0.60}$ .

Como não há dois ou mais componentes associadas à mesma classe, a probabilidade a posteriori para cada classe é representada pelos valores em negrito. De acordo com o CBO, a nova instância seria classificada como pertencente à classe cinza, o que contraria o padrão de formação relacionado à classe preta.

Ao contrário das técnicas de classificação de baixo nível como o CBO, existem outras técnicas capazes de explorar padrões de conformidade das classes e a própria estrutura topológica da rede, sendo denominadas técnicas de classificação de alto nível. A hipótese deste trabalho contempla a modelagem e utilização de classificação de alto nível utilizando o grafo  $k$ -associados ótimo como meio de tornar mais eficiente o processo de classificação.

### 3.2 Modelo CAN-KAOG

Na modelagem adotada em [2], a vizinhança de uma nova instância  $v_y$  é calculada baseada no valor  $K_\alpha$  de cada componente  $C_\alpha$ . Por exemplo, um componente  $C_x$  possui  $K_x = n$ , assim para o cálculo da probabilidade deste componente são considerados os  $n$  vizinhos mais próximos de  $v_y$ . Logo, se pelo menos um destes vizinhos mais próximos pertencem ao componente  $C_x$  então há probabilidade de  $v_y$  receber a mesma classe de  $C_x$ . Contudo, este tipo de modelagem não é adequado para a classificação de alto nível, pois desconsidera uma de suas principais características: a exploração das propriedades da rede. Na Figura 2 o modelo de classificação proposto em [2] consideraria apenas a possibilidade da nova instância (sem rótulo) pertencer a classe cinza, pois para os valores de  $k$  ( $K_{preto}$  e  $K_{cinza}$ ) as instâncias pretas não seriam consideradas.



**Figura 2.** Exemplo de conjunto de dados onde a exploração topológica da rede é desconsiderada pelo modelo de classificação do KAOG.

Diante disto, a solução encontrada foi considerar os  $K_\alpha$  vizinhos mais próximos da nova instância pertencentes ao próprio componente  $C_\alpha$ . Por exemplo, considerando que um componente  $C_x$  possui  $K_x = 2$ , selecionam-se as duas instâncias pertencentes à  $C_x$  que são mais próximas da instância a ser classificada. Dessa forma, o padrão de formação de todos os componentes pode ser examinado pela técnica de classificação de alto nível. O método proposto foi denominado CAN-KAOG (Classificação de Alto Nível utilizando Grafo K-Associados Ótimo).

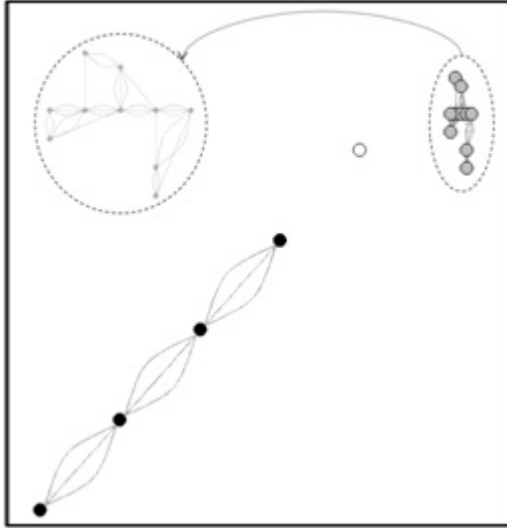
Entre as principais características de CAN-KAOG, destacam-se: (i) construção não paramétrica do classificador, (ii) mecanismo robusto a ruídos e *outliers*, (iii) estrutura de otimização local baseada na medida de pureza e (iv) classificação capaz de capturar as relações espaciais, funcionais e topológicas, bem como a formação de padrões nas classes através de medidas de rede como: assortatividade, coeficiente de aglomeração, grau médio, entre outras.

## 4 Resultados experimentais

Experimentos iniciais considerando algumas bases artificiais (com características de padrões de formação) e também base de dados reais foram realizados para avaliar o novo método. As medidas de rede utilizadas em CAN-KAOG foram: assortatividade, coeficiente de aglomeração e grau médio. O peso  $\alpha$  associado a cada uma delas foi, respectivamente 0.6, 0.3 e 0.1. O desempenho de CAN-KAOG foi comparado ao de KAOG (modelo que utiliza o CBO).

O primeiro experimento foi realizado com a base de dados apresentada na Figura 1. Vale destacar que tanto CBO quanto o classificador de alto nível utilizam o grafo k-associados ótimo para construção da rede. Além disso, pela própria estrutura de conexão da rede apresentada na Figura 1, é possível perceber a existência de um padrão de formação da classe preta. Conforme mostrado na Seção 3.1, o KAOG não foi capaz de reconhecer isto, porém o CAN-KAOG foi apto a fazê-lo, uma vez que considera relações complexas da rede através das medidas de rede empregadas.

O segundo experimento considerou um conjunto de dados pertencentes às classes preta e cinza, conforme mostrado na Figura 3. Uma nova instância  $v_y$  é inserida (sem rótulo) e deseja-se classificá-la. Os valores K encontrados pelo algoritmo foram:  $K_{preto} = 2$  e  $K_{cinza} = 3$ ; e as purezas associadas a cada componente ( $C_{preto}$  e  $C_{cinza}$ ) foram:  $\Phi_{preto} = 1$  e  $\Phi_{cinza} = 1$ .

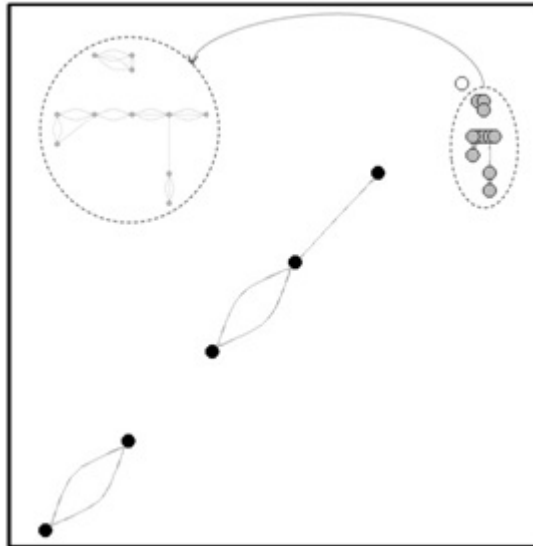


**Figura 3.** Segundo experimento: grafo k-associados ótimo obtido em um conjunto de instâncias rotuladas em duas classes: cinza e preto.



A princípio, é possível perceber na Figura 3 que as instâncias pretas formam uma reta, enquanto que as cinzas apresentam um comportamento indefinido onde os casos estão relativamente próximos. Novamente, as arestas entre as instâncias representam as ligações realizadas pelo grafo k-associados ótimo. Considerando o CBO, inicialmente tem-se que os três casos mais próximos da nova instância são, respectivamente, das classes: preto, cinza e cinza. Calculando as probabilidades têm-se:  $P(v_y \in C_{preto} | A_{v_y}) \approx 0.43$  e  $P(v_y \in C_{cinza} | A_{v_y}) \approx 0.57$ .

De acordo com o KAOG, a nova instância seria classificada como pertencente à classe cinza, o que de fato contraria o padrão de formação expresso pela classe preta. Utilizando o método de classificação de alto nível (CAN-KAOG), foi possível classificar corretamente a nova instância de acordo com o forte padrão que os vértices pretos apresentam, uma vez que a classificação de alto nível mede a conformidade da nova instância com a formação de padrões a partir dos dados de entrada, explorando as propriedades topológicas complexas da rede [4].



**Figura 4.** Terceiro experimento: grafo k-associados ótimo obtido em um conjunto de instâncias rotuladas em duas classes: cinza e preto.

O terceiro experimento utilizou a base de dados da Figura 4, que exibe um conjunto de instâncias também rotuladas em duas classes: preto e cinza. Uma nova instância  $v_y$  é inserida (sem rótulo) e deseja-se classificá-la. Os valores K encontrados pela técnica foram:  $K_{preto}^1 = 1$ ,  $K_{preto}^2 = 1$ ,  $K_{cinza}^1 = 2$  e  $K_{cinza}^2 = 2$ . Considerando os valores de K e calculando a distância de  $v_y$  para os demais casos, tem-se que os casos próximos pertencem apenas à classe cinza. Logo, o KAOG considera apenas esta classe no cálculo das probabilidades. Por outro

lado, mesmo com a proximidade de  $v_y$  com os casos cinzas, CAN-KAOG é capaz de reconhecer o padrão de formação nas classes pretas.

É importante notar na Figura 4 que caso  $K$  fosse incrementado ( $K = 3$ ), os dois componentes cinzas se tornariam um e a pureza continuaria máxima. Logo, a heurística de parada não garante o número mínimo de componentes considerando a pureza.

O último experimento foi feito com a base de dados Íris [5]. Esta base apresenta 150 instâncias rotuladas de maneira igual entre 3 classes. O método de amostragem utilizado foi o *leave-one-out*. O desempenho de CAN-KAOG obteve uma acurácia de 0.967 (classificou 145 instâncias corretamente) aproximando-se do KNN, técnica que apresentou melhor resultado em [2]. Utilizando o mesmo método de amostragem e incrementando o valor de  $K$  até 30, o KNN conseguiu acurácia máxima de 0.953 (classificou 143 instâncias corretamente).

## 5 Considerações finais

CAN-KAOG é uma técnica baseada em grafo para aprendizado supervisionado. Na fase de treinamento, grafo  $k$ -associados ótimo é utilizado para construir a rede. Na fase de classificação, utiliza-se a classificação de alto nível. Entre as principais características de CAN-KAOG, destacam-se: (i) construção não paramétrica do classificador, (ii) mecanismo robusto a ruídos e *outliers*, (iii) estrutura de otimização local baseada na medida de pureza e (iv) classificação capaz de capturar as relações espaciais, funcionais e topológicas, bem como a formação de padrões nas classes através de medidas de rede. Experimentos foram realizados para analisar CAN-KAOG em relação ao método de classificação original, onde um CBO é empregado. Os resultados mostraram que CAN-KAOG é uma técnica promissora, pois apresenta grande sensibilidade aos padrões de formação das classes, superando CBO-KAOG nos conjuntos de dados avaliados. Além disso, CAN-KAOG também obteve bom desempenho sobre a base de dados Íris, quando comparado ao algoritmo KNN.

Os próximos trabalhos incluem a realização de experimentos considerando mais bases reais e outras bases artificiais, investigações acerca da heurística de parada do KAOG e também a construção de refinamentos em CAN-KAOG para lidar com outros problemas relacionados a técnicas de classificação de baixo nível.

## Referências

1. Newman, M.: The structure and function of complex networks. *SIAM review* **45**(2) (2003) 167–256
2. Bertini, Jr., J.R., Zhao, L., Motta, R., Lopes, A.A.: A nonparametric classification method based on  $k$ -associated graphs. *Inf. Sci.* **181**(24) (2011) 5435–5456
3. Lu, Z., Savas, B., Tang, W., Dhillon, I.S.: Supervised link prediction using multiple sources. 2010 IEEE International Conference on Data Mining (2010) 923–928
4. Silva, T.C., Zhao, L.: Network-based high level data classification. *IEEE Transactions on Neural Networks* **23** (2012) 954–970
5. Frank, A., Asuncion, A.: UCI machine learning repository (2010)