Mineração de Padrões Freqüentes Ortogonais e sua Aplicação em Classificação Associativa

Leandro Souza Costa
Orientador: Wagner Meira Jr.

Departamento de Ciência da Computação
Universidade Federal de Minas Gerais

Defesa de Dissertação de Mestrado
9 de abril de 2008

Introdução

- Era da Informação;

Introdução

- Era da Informação;
- Sistemas de Gerenciamento de Banco de Dados (SGBD);

Introdução

- Era da Informação;
- Sistemas de Gerenciamento de Banco de Dados (SGBD);
- *Data Warehouse*;

Introdução

- Era da Informação;
- Sistemas de Gerenciamento de Banco de Dados (SGBD);
- *Data Warehouse*;
- Mineração de Dados.

Mineração de Dados

- Descoberta de Conhecimento em Bases de Dados (KDD - *Knowledge Discovery in Databases*);

Mineração de Dados

- Descoberta de Conhecimento em Bases de Dados (KDD - *Knowledge Discovery in Databases*);
- Padrões Freqüentes;

Mineração de Dados

- Descoberta de Conhecimento em Bases de Dados (KDD - *Knowledge Discovery in Databases*);
- Padrões Freqüentes;
- Regras de Associação.

Ortogonalidade

Definição Matemática

Dois vetores x e y são ortogonais num espaço vetorial V se o produto interno $\langle x, y \rangle$ é zero. Esta situação é descrita por $x \perp y$.

Ortogonalidade

Definição Matemática

Dois vetores x e y são ortogonais num espaço vetorial V se o produto interno $\langle x, y \rangle$ é zero. Esta situação é descrita por $x \perp y$.

Definição Adotada

Estamos interessados no quanto os elementos de um conjunto contribuem com informações não redundantes para a solução de um problema. Considerando que seja possível medir esta contribuição, e chamá-la de **significância**, podemos definir **ortogonalidade** como a média das significâncias dos elementos do conjunto.

Objetivos

Explorar o problema de classificação associativa considerando ortogonalidade entre padrões freqüentes com a intenção de:

- Minimizar o número de padrões utilizados na geração das regras;

Objetivos

Explorar o problema de classificação associativa considerando ortogonalidade entre padrões freqüentes com a intenção de:

- Minimizar o número de padrões utilizados na geração das regras;
- Diminuir a redundância das regras geradas;

Objetivos

Explorar o problema de classificação associativa considerando ortogonalidade entre padrões freqüentes com a intenção de:

- Minimizar o número de padrões utilizados na geração das regras;
- Diminuir a redundância das regras geradas;
- Diminuir a ambigüidade das regras geradas;

Objetivos

Explorar o problema de classificação associativa considerando ortogonalidade entre padrões freqüentes com a intenção de:

- Minimizar o número de padrões utilizados na geração das regras;
- Diminuir a redundância das regras geradas;
- Diminuir a ambigüidade das regras geradas;
- Aumentar a efetividade das classificações.

Trabalhos Relacionados

- Compactação do Conjunto de Padrões Freqüentes:

Trabalhos Relacionados

- Compactação do Conjunto de Padrões Freqüentes:
  - Padrões Fechados e Maximais;

Trabalhos Relacionados

- Compactação do Conjunto de Padrões Freqüentes:
  - Padrões Fechados e Maximais;
  - Extração de sub-conjunto (top-k) de padrões;

Trabalhos Relacionados

- Compactação do Conjunto de Padrões Freqüentes:
  - Padrões Fechados e Maximais;
  - Extração de sub-conjunto (top-k) de padrões;
  - Representação alternativa do conjuntos de padrões.

Trabalhos Relacionados

- Compactação do Conjunto de Padrões Freqüentes:
  - Padrões Fechados e Maximais;
  - Extração de sub-conjunto (top-k) de padrões;
  - Representação alternativa do conjuntos de padrões.
- Diminuição de Redundância no Conjunto de Padrões Freqüentes:

Trabalhos Relacionados

- Compactação do Conjunto de Padrões Freqüentes:
  - Padrões Fechados e Maximais;
  - Extração de sub-conjunto (top-k) de padrões;
  - Representação alternativa do conjuntos de padrões.
- Diminuição de Redundância no Conjunto de Padrões Freqüentes:
  - Função objetivo relacionando significância e redundância;

Trabalhos Relacionados

- Compactação do Conjunto de Padrões Freqüentes:
  - Padrões Fechados e Maximais;
  - Extração de sub-conjunto (top-k) de padrões;
  - Representação alternativa do conjuntos de padrões.
- Diminuição de Redundância no Conjunto de Padrões Freqüentes:
  - Função objetivo relacionando significância e redundância;
  - Modelos baseados em agrupamentos e representantes;

Trabalhos Relacionados

- Compactação do Conjunto de Padrões Freqüentes:
  - Padrões Fechados e Maximais;
  - Extração de sub-conjunto (top-k) de padrões;
  - Representação alternativa do conjuntos de padrões.
- Diminuição de Redundância no Conjunto de Padrões Freqüentes:
  - Função objetivo relacionando significância e redundância;
  - Modelos baseados em agrupamentos e representantes;
  - ORIGAMI (α-ortogonalidade).

Trabalhos Relacionados

- Compactação do Conjunto de Padrões Freqüentes:
  - Padrões Fechados e Maximais;
  - Extração de sub-conjunto (top-k) de padrões;
  - Representação alternativa do conjuntos de padrões.
- Diminuição de Redundância no Conjunto de Padrões Freqüentes:
  - Função objetivo relacionando significância e redundância;
  - Modelos baseados em agrupamentos e representantes;
  - ORIGAMI (α-ortogonalidade).
- Classificação Associativa:

Trabalhos Relacionados

- Compactação do Conjunto de Padrões Freqüentes:
  - Padrões Fechados e Maximais;
  - Extração de sub-conjunto (top-k) de padrões;
  - Representação alternativa do conjuntos de padrões.
- Diminuição de Redundância no Conjunto de Padrões Freqüentes:
  - Função objetivo relacionando significância e redundância;
  - Modelos baseados em agrupamentos e representantes;
  - ORIGAMI (α-ortogonalidade).
- Classificação Associativa:
  - Estratégia *lazy*.

Modelos de Classificação

- Modelos propostos: redes neurais, estatísticos, árvores de decisão, algoritmos genéticos, etc.;

Modelos de Classificação

- Modelos propostos: redes neurais, estatísticos, árvores de decisão, algoritmos genéticos, etc.;
- Modelo baseado em árvores de decisão é um dos mais indicados para Mineração de Dados;

Modelos de Classificação

- Modelos propostos: redes neurais, estatísticos, árvores de decisão, algoritmos genéticos, etc.;
- Modelo baseado em árvores de decisão é um dos mais indicados para Mineração de Dados;
- Classificação Associativa produz resultados ainda melhores.

Classificação Associativa

- Dados de entrada: Coleção de registros;

Classificação Associativa

- Dados de entrada: Coleção de registros;
- Cada registro é caracterizado por um par (x, y), onde x é um conjunto de atributos comuns, e y é um atributo especial, designado como **classe**;

Classificação Associativa

- Dados de entrada: Coleção de registros;
- Cada registro é caracterizado por um par (x, y), onde x é um conjunto de atributos comuns, e y é um atributo especial, designado como **classe**;
- Classificação é o processo de se descobrir uma função f que realiza o mapeamento de cada conjunto de atributos x para uma das classes y pré-definidas.

Estratégias *eager* e *lazy*

Estratégia *eager*

Gera um conjunto de regras a partir da base de treinamento, e, para cada instância de teste, utiliza a melhor regra do conjunto para classificá-la.

Estratégias *eager* e *lazy*

Estratégia *eager*

Gera um conjunto de regras a partir da base de treinamento, e, para cada instância de teste, utiliza a melhor regra do conjunto para classificá-la.

Estratégia *lazy*

Para cada instância de teste, gera um conjunto de regras a partir de uma projeção da base de treinamento que possui apenas transações relacionadas com a instância de teste.

Padrões Freqüentes

- Seja $\mathcal{I}$ um conjunto de itens;

Padrões Freqüentes

- Seja $\mathcal{I}$ um conjunto de itens;
- Um conjunto $X = \{i_1, \cdots, i_k\} \subseteq \mathcal{I}$ é chamado de *itemset* (ou padrão);

Padrões Freqüentes

- Seja $\mathcal{I}$ um conjunto de itens;
- Um conjunto $X = \{i_1, \cdots, i_k\} \subseteq \mathcal{I}$ é chamado de *itemset* (ou padrão);
- Uma transação sobre $\mathcal{I}$ é um par $T = (tid, I)$ onde *tid* é o identificador da transação e I é um *itemset*;

Padrões Freqüentes

- Seja $\mathcal{I}$ um conjunto de itens;
- Um conjunto $X = \{i_1, \cdots, i_k\} \subseteq \mathcal{I}$ é chamado de *itemset* (ou padrão);
- Uma transação sobre $\mathcal{I}$ é um par $T = (tid, I)$ onde tid é o identificador da transação e I é um *itemset*;
- Dizemos que uma transação $T = (tid, I)$ é coberta por um *itemset* $X \subseteq \mathcal{I}$, se $X \subseteq I$;

Padrões Freqüentes

- Uma base de dados de transações $\mathcal{D}$ sobre $\mathcal{I}$ é um conjunto de transações sobre $\mathcal{I}$;

Padrões Freqüentes

- Uma base de dados de transações $\mathcal{D}$ sobre $\mathcal{I}$ é um conjunto de transações sobre $\mathcal{I}$;
- A freqüência de um *itemset* X em $\mathcal{D}$ é o número de transações cobertas por X em $\mathcal{D}$;

Padrões Freqüentes

- Uma base de dados de transações $\mathcal{D}$ sobre $\mathcal{I}$ é um conjunto de transações sobre $\mathcal{I}$;
- A freqüência de um *itemset* X em $\mathcal{D}$ é o número de transações cobertas por X em $\mathcal{D}$;
- O suporte de um *itemset* X em $\mathcal{D}$ é a probabilidade de X ocorrer em uma transação $T \in \mathcal{D}$;

Padrões Freqüentes

- Uma base de dados de transações $\mathcal{D}$ sobre $\mathcal{I}$ é um conjunto de transações sobre $\mathcal{I}$;
- A freqüência de um *itemset* X em $\mathcal{D}$ é o número de transações cobertas por X em $\mathcal{D}$;
- O suporte de um *itemset* X em $\mathcal{D}$ é a probabilidade de X ocorrer em uma transação $T \in \mathcal{D}$;
- Um *itemset* é freqüente se o seu suporte é maior ou igual a um dado valor relativo mínimo σ, com $0 \leq \sigma \leq 1$.

Padrões Freqüentes

Definição

Seja $\mathcal{D}$ uma base de dados de transações sobre um conjunto de itens $\mathcal{I}$, e σ um valor mínimo de suporte. A coleção de *itemsets* freqüentes em $\mathcal{D}$ em relação a σ é dado por:

$$\mathcal{F}(\mathcal{D}, \sigma) := \{X \subseteq \mathcal{I} | suporte(X, \mathcal{D}) \geq \sigma\} .$$

Regras de Associação

- Uma regra de associação é uma implicação da forma $X \Rightarrow Y$, onde X é um conjunto de itens em $\mathcal{I}$, e Y é um único item em $\mathcal{I}$ que não está presente em X;

Regras de Associação

- Uma regra de associação é uma implicação da forma $X \Rightarrow Y$, onde X é um conjunto de itens em $\mathcal{I}$, e Y é um único item em $\mathcal{I}$ que não está presente em X;
- A regra $X \Rightarrow Y$ é satisfeita no conjunto de transações T com confiança $0 \leq c \leq 1$ se, e somente se, pelo menos c% das transações em T que satisfazem X também satisfazem Y;

Regras de Associação

- Uma regra de associação é uma implicação da forma $X \Rightarrow Y$, onde X é um conjunto de itens em $\mathcal{I}$, e Y é um único item em $\mathcal{I}$ que não está presente em X;
- A regra $X \Rightarrow Y$ é satisfeita no conjunto de transações T com confiança $0 \leq c \leq 1$ se, e somente se, pelo menos c% das transações em T que satisfazem X também satisfazem Y;
- O suporte de uma regra $X \Rightarrow Y$ em $\mathcal{D}$ é o suporte de $X \cup Y$ em $\mathcal{D}$, e a freqüência da regra é a freqüência de $X \cup Y$;

Regras de Associação

- Dizemos que uma regra de associação é freqüente se o seu suporte excede um determinado valor mínimo σ;

Regras de Associação

- Dizemos que uma regra de associação é freqüente se o seu suporte excede um determinado valor mínimo σ;
- A confiança de uma regra de associação $X \Rightarrow Y$ em $\mathcal{D}$ é a probabilidade condicional de encontrar Y numa transação, dado que esta contém X;

Regras de Associação

- Dizemos que uma regra de associação é freqüente se o seu suporte excede um determinado valor mínimo σ;
- A confiança de uma regra de associação $X \Rightarrow Y$ em $\mathcal{D}$ é a probabilidade condicional de encontrar Y numa transação, dado que esta contém X;
- Dizemos que a regra é de confiança se $P(Y|X)$ excede um determinado valor mínimo de confiança γ, com $0 \leq \gamma \leq 1$.

Regras de Associação

Definição

Seja $\mathcal{D}$ uma base de dados de transações sobre um conjunto de itens $\mathcal{I}$, σ um valor mínimo para suporte e γ um valor mínimo para confiança, o conjunto de regras de associação freqüentes e de confiança considerando σ e γ é dado por:

$$\mathcal{R}(\mathcal{D},\sigma,\gamma) := \{X \Rightarrow Y | X, Y \subseteq \mathcal{I}, X \cap Y = \{\}, X \cup Y \in \mathcal{F}(\mathcal{D},\sigma), \\ \mathit{confianca}(X \Rightarrow Y, \mathcal{D}) \geq \gamma\}.$$

Métricas Alternativas

- Convicção: Definida como $conviccao(X \Rightarrow Y) = \frac{P(X) \times P(\neg Y)}{P(X \wedge \neg Y)}$, compara a probabilidade de X aparecer sem Y com a freqüência real do aparecimento de X sem Y;

Métricas Alternativas

- Convicção: Definida como $conviccao(X \Rightarrow Y) = \frac{P(X) \times P(\neg Y)}{P(X \wedge \neg Y)}$, compara a probabilidade de X aparecer sem Y com a freqüência real do aparecimento de X sem Y;
- Leverage: Definida como $leverage(X \Rightarrow Y) = P(X \wedge Y) - (P(X) \times P(Y))$, mede a diferença de X e Y aparecendo juntos na base de dados e o que seria esperado se X e Y fossem estatisticamente dependentes;

Métricas Alternativas

- Convicção: Definida como $conviccao(X \Rightarrow Y) = \frac{P(X) \times P(\neg Y)}{P(X \wedge \neg Y)}$, compara a probabilidade de X aparecer sem Y com a freqüência real do aparecimento de X sem Y;
- Leverage: Definida como $leverage(X \Rightarrow Y) = P(X \wedge Y) - (P(X) \times P(Y))$, mede a diferença de X e Y aparecendo juntos na base de dados e o que seria esperado se X e Y fossem estatisticamente dependentes;
- Lift: Definida como $lift(X \Rightarrow Y) = \frac{P(X \wedge Y)}{P(X) \times P(Y)}$, mede quantas vezes X e Y ocorrem juntos a mais que o esperado se eles fossem estatisticamente independentes. Uma das desvantagens do *lift* é ser susceptível a ruídos em pequenas bases de dados;

Métricas Alternativas

- Jaccard: O coeficiente de Jaccard é uma medida estatística utilizada para comparar similaridade e diversidade entre conjuntos, definida pela razão entre a interseção e a união entre dois conjuntos. Esta métrica é obtida pela expressão $jaccard(X \Rightarrow Y) = \frac{P(X \wedge Y)}{P(X)+P(Y)-P(X \wedge Y)}$;

Métricas Alternativas

- Jaccard: O coeficiente de Jaccard é uma medida estatística utilizada para comparar similaridade e diversidade entre conjuntos, definida pela razão entre a interseção e a união entre dois conjuntos. Esta métrica é obtida pela expressão $jaccard(X \Rightarrow Y) = \frac{P(X \wedge Y)}{P(X)+P(Y)-P(X \wedge Y)}$;
- Laplace: Definida como $laplace(X \Rightarrow Y) = \frac{frequencia(X \wedge Y)+1}{frequencia(X)+c}$, onde c é o número de classes do domínio;

Métricas Alternativas

- Jaccard: O coeficiente de Jaccard é uma medida estatística utilizada para comparar similaridade e diversidade entre conjuntos, definida pela razão entre a interseção e a união entre dois conjuntos. Esta métrica é obtida pela expressão $jaccard(X \Rightarrow Y) = \frac{P(X \wedge Y)}{P(X)+P(Y)-P(X \wedge Y)}$;
- Laplace: Definida como $laplace(X \Rightarrow Y) = \frac{frequencia(X \wedge Y)+1}{frequencia(X)+c}$, onde c é o número de classes do domínio;
- Kulc: Definida como $kulc(X \Rightarrow Y) = \frac{P(X \wedge Y)}{2}\left(\frac{1}{P(X)} + \frac{1}{P(Y)}\right)$, a medida *Kulczynski* é muito utilizada na área química;

Métricas Alternativas

- Jaccard: O coeficiente de Jaccard é uma medida estatística utilizada para comparar similaridade e diversidade entre conjuntos, definida pela razão entre a interseção e a união entre dois conjuntos. Esta métrica é obtida pela expressão $jaccard(X \Rightarrow Y) = \frac{P(X \wedge Y)}{P(X)+P(Y)-P(X \wedge Y)}$;
- Laplace: Definida como $laplace(X \Rightarrow Y) = \frac{frequencia(X \wedge Y)+1}{frequencia(X)+c}$, onde c é o número de classes do domínio;
- Kulc: Definida como $kulc(X \Rightarrow Y) = \frac{P(X \wedge Y)}{2}\left(\frac{1}{P(X)} + \frac{1}{P(Y)}\right)$, a medida *Kulczynski* é muito utilizada na área química;
- Cosseno: Esta métrica, bastante utilizada como medida de similaridade para textos, é definida como $cosseno(X \Rightarrow Y) = \frac{P(X \wedge Y)}{\sqrt{P(X) \times P(Y)}}$;

Métricas Alternativas

- Sensitividade: Definida como $sensitividade(X \Rightarrow Y) = P(X|Y)$, sensitividade (ou *recall*) é bastante utilizada em sistemas de recuperação de informação;

Métricas Alternativas

- Sensitividade: Definida como $sensitividade(X \Rightarrow Y) = P(X|Y)$, sensitividade (ou *recall*) é bastante utilizada em sistemas de recuperação de informação;
- Especificidade: Definida como $especificidade(X \Rightarrow Y) = P(\neg Y|\neg X)$, esta métrica representa a proporção de verdadeiro-negativos sobre os casos negativos da regra.

Padrões Freqüentes

- Largamente utilizados em diversas aplicações, incluindo regras de associação, classificação, agrupamento, indexação, etc.;

Padrões Freqüentes

- Largamente utilizados em diversas aplicações, incluindo regras de associação, classificação, agrupamento, indexação, etc.;
- Minimizar o conjunto-solução ainda é um desafio:

Padrões Freqüentes

- Largamente utilizados em diversas aplicações, incluindo regras de associação, classificação, agrupamento, indexação, etc.;
- Minimizar o conjunto-solução ainda é um desafio:
  - Padrões freqüentes obedecem à propriedade de antimonotonia;

Padrões Freqüentes

- Largamente utilizados em diversas aplicações, incluindo regras de associação, classificação, agrupamento, indexação, etc.;
- Minimizar o conjunto-solução ainda é um desafio:
  - Padrões freqüentes obedecem à propriedade de antimonotonia;
  - Soluções propostas minimizam o conjunto-solução apenas sob a perspectiva do suporte, não considerando a semântica dos dados.

Padrões Freqüentes

- Largamente utilizados em diversas aplicações, incluindo regras de associação, classificação, agrupamento, indexação, etc.;
- Minimizar o conjunto-solução ainda é um desafio:
  - Padrões freqüentes obedecem à propriedade de antimonotonia;
  - Soluções propostas minimizam o conjunto-solução apenas sob a perspectiva do suporte, não considerando a semântica dos dados.
- Diminuir a redundância no conjunto-solução é outro desafio:

Padrões Freqüentes

- Largamente utilizados em diversas aplicações, incluindo regras de associação, classificação, agrupamento, indexação, etc.;
- Minimizar o conjunto-solução ainda é um desafio:
  - Padrões freqüentes obedecem à propriedade de antimonotonia;
  - Soluções propostas minimizam o conjunto-solução apenas sob a perspectiva do suporte, não considerando a semântica dos dados.
- Diminuir a redundância no conjunto-solução é outro desafio:
  - Poucos estudos têm se dedicado a obter sub-conjuntos de alta significância e baixa redundância ao mesmo tempo.

Padrões Freqüentes

Padrões Ortogonais

O objetivo da aplicação de ortogonalidade no problema da mineração de padrões freqüentes é desenvolver uma técnica capaz de extrair um sub-conjunto de padrões com tanto alta significância quanto baixa redundância entre seus elementos.

Métricas de Ortogonalidade

- É necessário definir métricas de ortogonalidade capazes de avaliar um possível conjunto solução;

Métricas de Ortogonalidade

- É necessário definir métricas de ortogonalidade capazes de avaliar um possível conjunto solução;
- O complemento do coeficiente de **Jaccard** aplicado à cobertura da base de dados pode ser considerado como uma métrica de ortogonalidade entre dois padrões:

$$D(p_1, p_2) = 1 - \frac{|TS(p_1) \cap TS(p_2)|}{|TS(p_1) \cup TS(p_2)|},$$

onde $TS(p)$ é o conjunto de transações cobertas pelo padrão p.

Métricas de Ortogonalidade

- É necessário definir métricas de ortogonalidade capazes de avaliar um possível conjunto solução;
- O complemento do coeficiente de **Jaccard** aplicado à cobertura da base de dados pode ser considerado como uma métrica de ortogonalidade entre dois padrões:

$$D(p_1, p_2) = 1 - \frac{|TS(p_1) \cap TS(p_2)|}{|TS(p_1) \cup TS(p_2)|},$$

  onde $TS(p)$ é o conjunto de transações cobertas pelo padrão p.
- Estamos interessados em definir métricas aplicáveis a conjuntos de qualquer tamanho.

Considerando Estrutura dos Padrões

Motivação

Dois padrões são ortogonais se eles não possuem itens em comum, ou seja, pode-se dizer que os padrões ABC e DEF são ortogonais, mas ABC e CDE não o são, já que o item C está presente nos dois padrões. O mesmo pode ser aplicado a conjuntos maiores, por exemplo, os padrões AB, CD e EF são ortogonais, mas os padrões AB, BC e CD não o são.

Considerando Estrutura dos Padrões

- Seja $\mathcal{I}$ um conjunto de itens, $\mathcal{D}$ uma base de dados de transações em $\mathcal{I}$, $\mathcal{F}$ o conjunto de padrões freqüentes em $\mathcal{D}$, e $\mathcal{F}'$ um sub-conjunto de $\mathcal{F}$ ($\mathcal{F}' \subseteq \mathcal{F}$);

Considerando Estrutura dos Padrões

- Seja $\mathcal{I}$ um conjunto de itens, $\mathcal{D}$ uma base de dados de transações em $\mathcal{I}$, $\mathcal{F}$ o conjunto de padrões freqüentes em $\mathcal{D}$, e $\mathcal{F}^{'}$ um sub-conjunto de $\mathcal{F}$ ($\mathcal{F}^{'} \subseteq \mathcal{F}$);
- Chamamos de $\mathcal{I}^{'} \subseteq \mathcal{I}$ o sub-conjunto itens que aparecem em, pelo menos, um dos padrões de $\mathcal{F}^{'}$;

Considerando Estrutura dos Padrões

- Seja $\mathcal{I}$ um conjunto de itens, $\mathcal{D}$ uma base de dados de transações em $\mathcal{I}$, $\mathcal{F}$ o conjunto de padrões freqüentes em $\mathcal{D}$, e $\mathcal{F}^{'}$ um sub-conjunto de $\mathcal{F}$ ($\mathcal{F}^{'} \subseteq \mathcal{F}$);
- Chamamos de $\mathcal{I}^{'} \subseteq \mathcal{I}$ o sub-conjunto itens que aparecem em, pelo menos, um dos padrões de $\mathcal{F}^{'}$;
- Para cada item $i \subseteq \mathcal{I}^{'}$ é dado um peso:

$$w_i = \frac{|\mathcal{F}^{'}| - |\mathcal{F}_i^{'}|}{|\mathcal{F}^{'}| - 1},$$

onde $\mathcal{F}_i^{'} \subseteq \mathcal{F}^{'}$ é o sub-conjunto de padrões de $\mathcal{F}^{'}$ que contém o item i;

Considerando Estrutura dos Padrões

- A ortogonalidade baseada na estrutura dos padrões do conjunto é dada por:

$$O_e = \frac{\sum_{i \subseteq \mathcal{I}'} w_i}{|\mathcal{I}'|}.$$

Considerando Cobertura de Transações

Motivação

Dois padrões são ortogonais se eles cobrem áreas diferentes da base de dados, ou seja, se os conjuntos de transações cobertas por cada padrão não possuem elementos em comum.

Considerando Cobertura de Transações

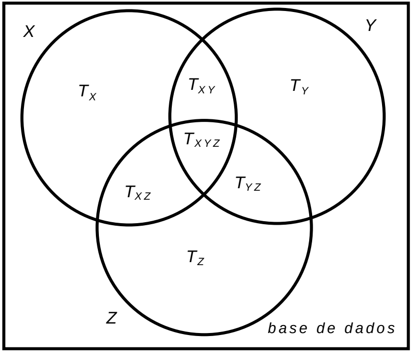


Figura: Visualização de Cobertura de Transações na Base de Dados

Considerando Cobertura de Transações

- Seja $\mathcal{I}$ um conjunto de itens, $\mathcal{D}$ uma base de dados de transações em $\mathcal{I}$, $\mathcal{F}$ o conjunto de padrões freqüentes em $\mathcal{D}$, e $\mathcal{F}'$ um sub-conjunto de $\mathcal{F}$ ($\mathcal{F}' \subseteq \mathcal{F}$);

Considerando Cobertura de Transações

- Seja $\mathcal{I}$ um conjunto de itens, $\mathcal{D}$ uma base de dados de transações em $\mathcal{I}$, $\mathcal{F}$ o conjunto de padrões freqüentes em $\mathcal{D}$, e $\mathcal{F}^{'}$ um sub-conjunto de $\mathcal{F}$ ($\mathcal{F}^{'} \subseteq \mathcal{F}$);
- Chamamos de $\mathcal{D}^{'} \subseteq \mathcal{D}$ o sub-conjunto transações cobertas por, pelo menos, um dos padrões de $\mathcal{F}^{'}$;

Considerando Cobertura de Transações

- Seja $\mathcal{I}$ um conjunto de itens, $\mathcal{D}$ uma base de dados de transações em $\mathcal{I}$, $\mathcal{F}$ o conjunto de padrões freqüentes em $\mathcal{D}$, e $\mathcal{F}'$ um sub-conjunto de $\mathcal{F}$ ($\mathcal{F}' \subseteq \mathcal{F}$);
- Chamamos de $\mathcal{D}' \subseteq \mathcal{D}$ o sub-conjunto transações cobertas por, pelo menos, um dos padrões de $\mathcal{F}'$;
- Para cada transação $t \subseteq \mathcal{D}'$ é dado um peso:

$$w_t = \frac{|\mathcal{F}'| - |\mathcal{F}'_t|}{|\mathcal{F}'| - 1},$$

onde $\mathcal{F}'_t$ é o sub-conjunto de padrões de $\mathcal{F}'$ que cobrem a transação t;

Considerando Cobertura de Transações

- A ortogonalidade baseada em cobertura de transações do conjunto é dada por:

$$O_t = \frac{\sum_{t \subseteq \mathcal{D}'} w_t}{|\mathcal{D}'|}.$$

.

Considerando Cobertura de Classes

Motivação

Dois padrões são ortogonais se são encontrados em transações de classes distintas na base de dados, ou seja, os conjuntos de transações cobertas por cada um dos padrões não devem possuir classes em comum.

Considerando Cobertura de Classes

- Seja $\mathcal{I}$ um conjunto de itens, $\mathcal{D}$ uma base de dados de transações em $\mathcal{I}$, $\mathcal{F}$ o conjunto de padrões freqüentes em $\mathcal{D}$, e $\mathcal{F}^{'}$ um sub-conjunto de $\mathcal{F}$ ($\mathcal{F}^{'} \subseteq \mathcal{F}$). Chamamos de $\mathcal{D}^{'} \subseteq \mathcal{D}$ o sub-conjunto transações cobertas por, pelo menos, um dos padrões de $\mathcal{F}^{'}$;

Considerando Cobertura de Classes

- Seja $\mathcal{I}$ um conjunto de itens, $\mathcal{D}$ uma base de dados de transações em $\mathcal{I}$, $\mathcal{F}$ o conjunto de padrões freqüentes em $\mathcal{D}$, e $\mathcal{F}^{'}$ um sub-conjunto de $\mathcal{F}$ ($\mathcal{F}^{'} \subseteq \mathcal{F}$). Chamamos de $\mathcal{D}^{'} \subseteq \mathcal{D}$ o sub-conjunto transações cobertas por, pelo menos, um dos padrões de $\mathcal{F}^{'}$;
- Seja $\mathcal{C}$ um conjunto de classes associadas às transações de $\mathcal{D}$. Chamamos de $\mathcal{C}^{'} \subseteq \mathcal{C}$ o sub-conjunto de classes associadas às transações de $\mathcal{D}^{'}$;

Considerando Cobertura de Classes

- Para cada classe $c \subseteq \mathcal{C}'$ é dado um peso:

$$w_c = \frac{|\mathcal{F}'| - |\mathcal{F}'_c|}{|\mathcal{F}'| - 1},$$

onde $\mathcal{F}'_c$ é o sub-conjunto de padrões de $\mathcal{F}'$ que cobrem uma quantidade de transações de classe $c \subseteq \mathcal{C}'$ maior que 90% da média esperada;

Considerando Cobertura de Classes

- A ortogonalidade baseada em cobertura de classes é dada por:

$$O_c = \frac{\sum_{c \subseteq \mathcal{C}'} w_c}{|\mathcal{C}'|}.$$

Utilização da ortogonalidade no LAC

- Para cada instância de teste, o LAC cria uma projeção da base de treinamento apenas com as transações que possuem itens em comum com a instância;

Utilização da ortogonalidade no LAC

- Para cada instância de teste, o LAC cria uma projeção da base de treinamento apenas com as transações que possuem itens em comum com a instância;
- A partir desta projeção, a obtém um conjunto de padrões freqüentes, de acordo com determinado suporte fornecido pelo usuário;

Utilização da ortogonalidade no LAC

- Para cada instância de teste, o LAC cria uma projeção da base de treinamento apenas com as transações que possuem itens em comum com a instância;
- A partir desta projeção, a obtém um conjunto de padrões freqüentes, de acordo com determinado suporte fornecido pelo usuário;
- Com estes padrões, gera as regras de associação utilizadas durante a tarefa de classificação.

Utilização da ortogonalidade no LAC

- Neste trabalho, a ortogonalidade foi utilizada para se extrair, do conjunto de padrões freqüentes, um sub-conjunto de padrões ortogonais;

Utilização da ortogonalidade no LAC

- Neste trabalho, a ortogonalidade foi utilizada para se extrair, do conjunto de padrões freqüentes, um sub-conjunto de padrões ortogonais;
- As regras de associação foram geradas a partir do sub-conjunto de padrões ortogonais obtido.

Heurística de Obtenção de Conjuntos Ortogonais

- O problema de se encontrar o sub-conjunto de padrões com maior métrica de ortogonalidade, dado o conjunto de padrões freqüentes, é não polinomial;

Heurística de Obtenção de Conjuntos Ortogonais

- O problema de se encontrar o sub-conjunto de padrões com maior métrica de ortogonalidade, dado o conjunto de padrões freqüentes, é não polinomial;
- Foi desenvolvida uma heurística gulosa que inicia com um conjunto ortogonal de dois elementos, e, iterativamente, tenta obter um novo conjunto com um elemento a mais, acrescentando padrões candidatos e realizando modificações para que a métrica de ortogonalidade seja maximizada.

Heurística de Obtenção de Conjuntos Ortogonais

Algoritmo 1 OLAC

Require: $\mathcal{D}, \sigma$

$\mathcal{F} \leftarrow FindFrequentPatterns(\mathcal{D}, \sigma)$
$Sort(\mathcal{F})$
$\mathcal{O} \leftarrow GetFirstAvailablePattern(\mathcal{F})$
repeat
  $rate \leftarrow GetOrthogonalityRate(\mathcal{O})$
  $\mathcal{O}_{try} \leftarrow \mathcal{O} \cup GetFirstAvailablePattern(\mathcal{F})$
  $rate_{try} = GetOrthogonalityRate(\mathcal{O}_{try})$
  for $P \in \mathcal{F}, P \notin \mathcal{O}_{try}$ **do**
    $S \leftarrow GetMoreSimilar(\mathcal{O}, P)$
    $\mathcal{O}_{try} \leftarrow \mathcal{O}_{try} \cup P \setminus S$
    $rate_{tmp} = GetRate(\mathcal{O})$
    if $rate_{tmp} \leq rate_{try}$ **then**
      $\mathcal{O}_{try} \leftarrow \mathcal{O}_{try} \cup S \setminus P$
    else
      $rate_{try} \leftarrow rate_{tmp}$
    end if
  end for
  if $rate_{try} \geq rate$ **then**
    $\mathcal{O} \leftarrow \mathcal{O}_{try}$
  end if
until $rate_{try} < rate$
$\mathcal{R} \leftarrow \mathcal{O}$

Contextualização

O **ORIGAMI** é um algoritmo para mineração de grafos encontrado na literatura, onde os autores introduzem a definição de conjuntos α-ortogonais e β-representativos, e apresentam o novo paradigma de mineração de conjuntos de grafos ortogonais com foco nos padrões, e não nas transações.

Definição de α-ortogonalidade

- Seja $\mathcal{F}$ o conjunto de todos os sub-grafos freqüentes de uma coleção;

Definição de α-ortogonalidade

- Seja $\mathcal{F}$ o conjunto de todos os sub-grafos freqüentes de uma coleção;
- Seja *sim* : $\mathcal{F} \times \mathcal{F} \rightarrow [0, 1]$ uma função binária e simétrica que retorna a *similaridade* entre dois grafos;

Definição de α-ortogonalidade

- Seja $\mathcal{F}$ o conjunto de todos os sub-grafos freqüentes de uma coleção;
- Seja $sim : \mathcal{F} \times \mathcal{F} \rightarrow [0, 1]$ uma função binária e simétrica que retorna a *similaridade* entre dois grafos;
- Dada uma coleção de grafos $\mathcal{G}$, e um limite superior para similaridade $\alpha \in [0, 1]$, dizemos que o sub-conjunto de grafos $\mathcal{R} \subseteq \mathcal{G}$ é α-**ortogonal** em relação a $\mathcal{G}$ se, e somente se, para quaisquer $G_a, G_b \in \mathcal{R}$, $sim(G_a, G_b) \leq \alpha$ e para qualquer $G_a \in \mathcal{R}$ e qualquer $G_b \in \mathcal{G} \backslash \mathcal{R}$, $sim(G_a, G_b) > \alpha$;

Definição de α-ortogonalidade

- Dada uma coleção de grafos $\mathcal{G}$,um conjunto α-ortogonal $\mathcal{R} \subseteq \mathcal{G}$ e um limite inferior para similaridade $\beta \in [0,1]$, dizemos que $\mathcal{R}$ **representa** um grafo $G \in \mathcal{G}$ se existe algum $G_a \in \mathcal{R}$ tal que $sim(G_a, G) \geq \beta$. Seja $\Upsilon(\mathcal{R}, \mathcal{G}) = \{G \in \mathcal{G} : \exists G_a \in \mathcal{R}, sim(G, G_a) \geq \beta\}$, dizemos que $\mathcal{R}$ é um conjunto β-representativo para $\Upsilon(\mathcal{R}, \mathcal{G})$;

Definição de α-ortogonalidade

- Dada uma coleção de grafos $\mathcal{G}$ e o seu conjunto α-ortogonal e β-representativo $\mathcal{R}$, chamamos de **resíduo** de $\mathcal{R}$ o conjunto de padrões não representados em $\mathcal{G}$, dado como $\Delta(\mathcal{R},\mathcal{G}) = \mathcal{G} \setminus \{\mathcal{R} \cup \Upsilon(\mathcal{R},\mathcal{G})\}$, o *resíduo* de $\mathcal{R}$ é definido como a cardinalidade do seu conjunto resíduo $|\Delta(\mathcal{R},\mathcal{G})|$. Finalmente, definimos a média de similaridade do resíduo de $\mathcal{R}$ como $ars(\mathcal{R},\mathcal{G}) = \frac{\sum_{G_b \in \Delta(\mathcal{R},\mathcal{G})} max_{G_a \in \mathcal{R}}\{sim(G_a,G_b)\}}{|\Delta(\mathcal{R},\mathcal{G})|}$.

Definição de α-ortogonalidade

Objetivo

O objetivo é encontrar conjuntos de grafos α-ortogonais e β-representativos em relação ao conjunto de sub-grafos maximais $\mathcal{M}$.

O Algoritmo ORIGAMI

Algoritmo 2 ORIGAMI

Require: $\mathcal{D}, \sigma, \alpha, \beta$

1: $\mathit{EM} \leftarrow \mathit{EdgeMap}(\mathcal{D})$
2: $\mathcal{F}_1 \leftarrow \mathit{FindFrequentEdges}(\mathcal{D}, \sigma)$
3: $\widehat{\mathcal{M}} \leftarrow 0$
4: **while** $\neg \mathit{StopCondition}()$ **do**
5: $\quad M \leftarrow \mathit{RandomMaximalGraph}(\mathcal{D}, \mathcal{F}_1, \mathit{EM}, \sigma)$
6: $\quad \widehat{\mathcal{M}} \leftarrow \widehat{\mathcal{M}} \cup M$
7: **end while**
8: $\mathcal{R} \leftarrow \mathit{OrthogonalRepresentativeSets}(\widehat{\mathcal{M}}, \alpha, \beta)$

Adaptação do Algoritmo

- Foi implementada uma adaptação do ORIGAMI para o problema de Classificação Associativa;

Adaptação do Algoritmo

- Foi implementada uma adaptação do ORIGAMI para o problema de Classificação Associativa;
- Foi implementada uma heurística de obtenção de padrões maximais baseada no trabalho apresentado no artigo;

Adaptação do Algoritmo

- Foi implementada uma adaptação do ORIGAMI para o problema de Classificação Associativa;
- Foi implementada uma heurística de obtenção de padrões maximais baseada no trabalho apresentado no artigo;
- Foi implementada uma heurística de obtenção do conjunto ortogonal baseada no trabalho apresentado no artigo.

Heurística de Obtenção de Padrões Maximais

- O algoritmo inicia a execução com o conjunto-resultado vazio;

Heurística de Obtenção de Padrões Maximais

- O algoritmo inicia a execução com o conjunto-resultado vazio;
- A cada iteração, tenta obter o maior padrão freqüente possível, selecionando itens aleatoriamente;

Heurística de Obtenção de Padrões Maximais

- O algoritmo inicia a execução com o conjunto-resultado vazio;
- A cada iteração, tenta obter o maior padrão freqüente possível, selecionando itens aleatoriamente;
    - Se o algoritmo escolhe um item já utilizado, ou que produz um padrão não freqüente, um contador de tentativas é decrementado;

Heurística de Obtenção de Padrões Maximais

- O algoritmo inicia a execução com o conjunto-resultado vazio;
- A cada iteração, tenta obter o maior padrão freqüente possível, selecionando itens aleatoriamente;
  - Se o algoritmo escolhe um item já utilizado, ou que produz um padrão não freqüente, um contador de tentativas é decrementado;
  - A condição de parada para a geração do padrão maximal candidato é que o número de escolhas erradas do item não deve ser maior que o tamanho da instância de teste.

Heurística de Obtenção de Padrões Maximais

- Ao obter um novo padrão maximal, o algoritmo tenta inseri-lo no conjunto-solução;

Heurística de Obtenção de Padrões Maximais

- Ao obter um novo padrão maximal, o algoritmo tenta inseri-lo no conjunto-solução;
    - Se o padrão escolhido já existe no conjunto, o algoritmo incrementa um segundo contador de tentativas;

Heurística de Obtenção de Padrões Maximais

- Ao obter um novo padrão maximal, o algoritmo tenta inseri-lo no conjunto-solução;
    - Se o padrão escolhido já existe no conjunto, o algoritmo incrementa um segundo contador de tentativas;
    - A condição de parada para a obtenção de padrões maximais é que o número de padrões candidatos não maximais ou já conhecidos não deve ser superior ao tamanho da instância de teste.

Heurística de Obtenção do Conjunto Ortogonal

- O algoritmo inicia a execução com o valor de resíduo igual a 0 (zero);

Heurística de Obtenção do Conjunto Ortogonal

- O algoritmo inicia a execução com o valor de resíduo igual a 0 (zero);
- A cada iteração, tenta obter um novo conjunto ortogonal selecionando, aleatoriamente, padrões maximais encontrados na primeira fase do algoritmo, e adicionando-os ao conjunto-solução;

Heurística de Obtenção do Conjunto Ortogonal

- O algoritmo inicia a execução com o valor de resíduo igual a 0 (zero);
- A cada iteração, tenta obter um novo conjunto ortogonal selecionando, aleatoriamente, padrões maximais encontrados na primeira fase do algoritmo, e adicionando-os ao conjunto-solução;
    - Se, durante a obtenção dos padrões, o padrão selecionado já ter sido utilizado, ou não possuir similaridade menor que α para com todos os outros padrões do conjunto-solução, o algoritmo decrementa um contador de tentativas;

Heurística de Obtenção do Conjunto Ortogonal

- O algoritmo inicia a execução com o valor de resíduo igual a 0 (zero);
- A cada iteração, tenta obter um novo conjunto ortogonal selecionando, aleatoriamente, padrões maximais encontrados na primeira fase do algoritmo, e adicionando-os ao conjunto-solução;
    - Se, durante a obtenção dos padrões, o padrão selecionado já ter sido utilizado, ou não possuir similaridade menor que α para com todos os outros padrões do conjunto-solução, o algoritmo decrementa um contador de tentativas;
    - A condição de parada local para a geração de conjuntos ortogonais é que, durante este processo, o número máximo de escolhas erradas de padrões não pode ser maior que a quantidade de padrões maximais total.

Heurística de Obtenção do Conjunto Ortogonal

- Ao obter um novo conjunto ortogonal, o algoritmo calcula o valor do seu resíduo;

Heurística de Obtenção do Conjunto Ortogonal

- Ao obter um novo conjunto ortogonal, o algoritmo calcula o valor do seu resíduo;
- Se este valor é menor que o atual, o resíduo é atualizado, e o conjunto-solução passar a ser o conjunto ortogonal recém-encontrado;

Heurística de Obtenção do Conjunto Ortogonal

- Ao obter um novo conjunto ortogonal, o algoritmo calcula o valor do seu resíduo;
- Se este valor é menor que o atual, o resíduo é atualizado, e o conjunto-solução passar a ser o conjunto ortogonal recém-encontrado;
- A condição de parada para o algoritmo é que, durante todo o processo, o número máximo de conjuntos ortogonais candidatos que não melhoram o resultado não pode ser maior que a quantidade de padrões maximais total.

Apresentação do Aplicativo

- O aplicativo **olac** possui a implementação de três abordagens distintas de um classificador baseado em regras de associação:

Apresentação do Aplicativo

- O aplicativo **olac** possui a implementação de três abordagens distintas de um classificador baseado em regras de associação:
  - A abordagem LAC (*Lazy Associative Classifier*), é a abordagem *lazy* na sua versão original (e não-ortogonal);

Apresentação do Aplicativo

- O aplicativo **olac** possui a implementação de três abordagens distintas de um classificador baseado em regras de associação:
  - A abordagem LAC (*Lazy Associative Classifier*), é a abordagem *lazy* na sua versão original (e não-ortogonal);
  - A abordagem OLAC (*Orthogonal Lazy Associative Classifier*) é a modificação da abordagem *lazy* que considera a ortogonalidade dos padrões durante a tarefa de obtenção de regras;

Apresentação do Aplicativo

- O aplicativo **olac** possui a implementação de três abordagens distintas de um classificador baseado em regras de associação:
  - A abordagem LAC (*Lazy Associative Classifier*), é a abordagem *lazy* na sua versão original (e não-ortogonal);
  - A abordagem OLAC (*Orthogonal Lazy Associative Classifier*) é a modificação da abordagem *lazy* que considera a ortogonalidade dos padrões durante a tarefa de obtenção de regras;
  - A abordagem ORIGAMI é a implementação da adaptação apresentada para a estratégia ORIGAMI.

Opções de Execução

```
Usage: ./olac [options]
Options:
  -i, --training-file       Set the training file
  -t, --testing-file        Set the testing file
  -s, --support             Set the support
  -c, --confidence          Set the confidence
  -r, --run-mode            Set the run mode [c,o] [CLASSICAL, ORTHOGONAL]
  -p, --pattern-set         Set the pattern set type [f,m,r] [FREQUENT, MAXIMAL,
                              RANDOM MAXIMAL]
  -n, --min-num-rules       Set the minimum number of rules generated
  -l, --max-num-rank-rules  Set the maximum number of rules considered (rank size)
  -m, --min-rule-len        Set the minimum length of the rules
  -x, --max-rule-len        Set the maximum length of the rules
  -o, --orth-mode           Set the orthogonality mode [h,p,o] [HEURISTICAL,
                              POLYNOMIAL, ORIGAMI]
```

Opções de Execução

```
-e, --orth-metric         Set the orthogonality metric [s,c,l,a] [STRUCTURE,
                            TRANSACTION COVERAGE, CLASS COVERAGE, ALL]
-w, --orth-method         Set the way metrics are used [s,p,a] [SET, PAIR AVERAGE,
                            ALL]
-g, --orth-pat-ordering   Set the way patterns are ordered for heuristic
                            [s,r,i,z,n] [SORTED, REVERSE SORTED, SORTED BY SIZE,
                            REVERSE SORTED BY SIZE, NONE]
-u, --rule-measure        Set the rule measure used [s,c,j,k,o,n,e,p,l,i,v]
                            [SUPPORT, CONFIDENCE, JACCARD, KULC, COSINE,
                            CONVICTION, SENSITIVITY, SPECIFICITY, LAPLACE,
                            LIFT, LEVERAGE]
-a, --origami-alpha       Set the alpha parameter used by ORIGAMI
-b, --origami-beta        Set the beta parameter used by ORIGAMI
-d, --debug               Set the level of debug [-1,0,1,2,3,4] [SILENT, NO DEBUG,
                            LOW LEVEL, MEDIUM LEVL, HIGH LEVEL, MAX LEVEL]
-v, --verbose             Use verbose mode
-h, --help                Display this information
```

Metodologia

- Foram utilizadas 26 bases de dados do repositório **UCI** (*UC Irvine Machine Learning Repository*), amplamente referenciado em pesquisas na área de classificação em mineração de dados;

Metodologia

- Foram utilizadas 26 bases de dados do repositório **UCI** (*UC Irvine Machine Learning Repository*), amplamente referenciado em pesquisas na área de classificação em mineração de dados;
- Todas as bases utilizadas durante os testes foram reordenadas aleatoriamente e particionadas em dez sub-conjuntos, de onde foram criadas dez configurações de teste para cada uma delas;

Metodologia

- Foram utilizadas 26 bases de dados do repositório **UCI** (*UC Irvine Machine Learning Repository*), amplamente referenciado em pesquisas na área de classificação em mineração de dados;
- Todas as bases utilizadas durante os testes foram reordenadas aleatoriamente e particionadas em dez sub-conjuntos, de onde foram criadas dez configurações de teste para cada uma delas;
- Cada configuração de teste consiste de uma parte (um sub-conjunto da base) como arquivo de teste, e nove partes (os nove sub-conjuntos restantes da base) como arquivo de treinamento;

Metodologia

- Como resultados foram considerados a média das dez execuções diferentes para cada base de dados;

Metodologia

- Como resultados foram considerados a média das dez execuções diferentes para cada base de dados;
- Os parâmetros utilizados nos testes se encontram na tabela 1;

Metodologia

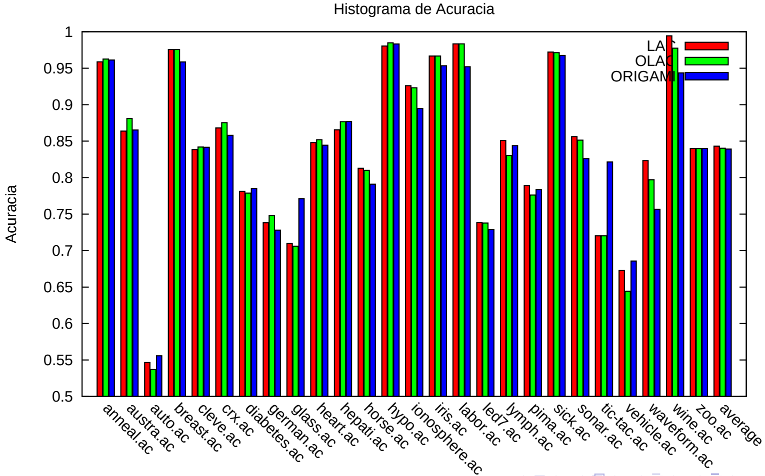
- Como resultados foram considerados a média das dez execuções diferentes para cada base de dados;
- Os parâmetros utilizados nos testes se encontram na tabela 1;
- Todas as combinações possíveis destes parâmetros foram realizadas, com exceção da combinação tamanho máximo de regra 1 e métrica de ortogonalidade *s* para o OLAC.

Metodologia

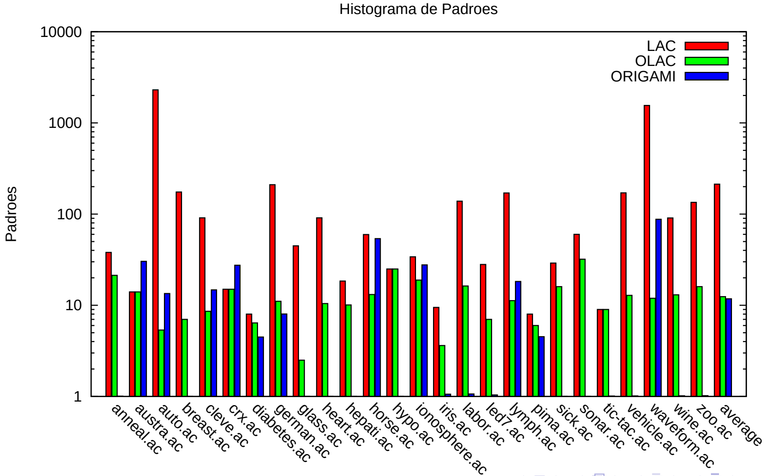
Parâmetros	Valores
support	$\{0.0001, 0.001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99, 1\}$
confidence	$\{0.0001, 0.001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99, 1\}$
min-num-rules	$\{1\}$
max-num-rank-rules	$\{1, 10, 100, 1000, 10000, 100000, 1000000\}$
min-rule-len	$\{1\}$
max-rule-len	$\{1, 2, 3\}$
rule-measure	$\{s, c, j, k, o, n, e, p, l, i, v\}$
orth-metric	$\{e, c, l, a\}$
orth-method	$\{s, p\}$
orth-pat-ordering	$\{s, r, i, z, n\}$
origami-alpha	$\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$
origami-beta	$\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$

Tabela: Parâmetros Utilizados Durante os Experimentos para Todas as Abordagens

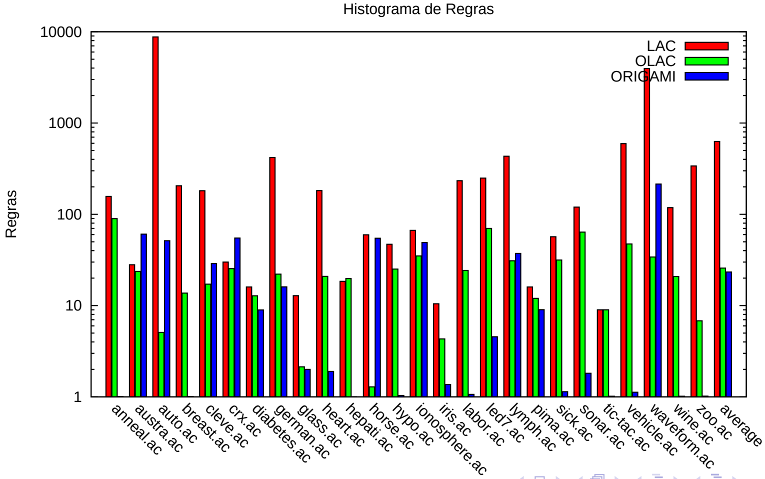
Acurácia


Histograma de Acuracia
Acuracia
LAC
OLAC
ORIGAMI
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
anneal.ac
austra.ac
auto.ac
breast.ac
cleve.ac
crx.ac
diabetes.ac
german.ac
glass.ac
heart.ac
hepati.ac
horse.ac
hypo.ac
ionosphere.ac
iris.ac
labor.ac
led7.ac
lymph.ac
pima.ac
sick.ac
sonar.ac
tic-tac.ac
vehicle.ac
waveform.ac
wine.ac
zoo.ac
average

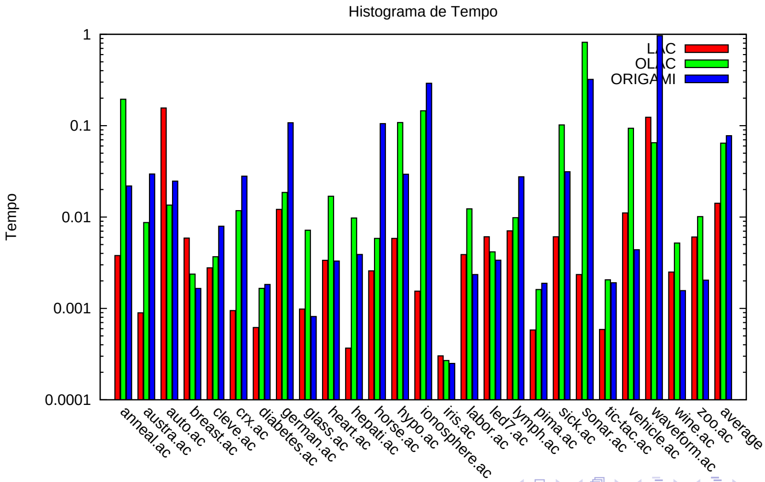
Número de Padrões



Número de Regras



Tempo de Classificação



Conclusões

- As abordagens baseadas em ortogonalidade obtiveram resultados semelhantes aos da abordagem clássica:

Conclusões

- As abordagens baseadas em ortogonalidade obtiveram resultados semelhantes aos da abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.843, 0.840 e 0.839;

Conclusões

- As abordagens baseadas em ortogonalidade obtiveram resultados semelhantes aos da abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.843, 0.840 e 0.839;
  - Considerando os melhores parâmetros para a média dos resultados, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.808, 0.813 e 0.782.

Conclusões

- As abordagens baseadas em ortogonalidade obtiveram resultados semelhantes aos da abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.843, 0.840 e 0.839;
  - Considerando os melhores parâmetros para a média dos resultados, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.808, 0.813 e 0.782.
- A quantidade de padrões utilizados na geração das regras nas abordagens ortogonais foi bem menor que na abordagem clássica:

Conclusões

- As abordagens baseadas em ortogonalidade obtiveram resultados semelhantes aos da abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.843, 0.840 e 0.839;
  - Considerando os melhores parâmetros para a média dos resultados, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.808, 0.813 e 0.782.
- A quantidade de padrões utilizados na geração das regras nas abordagens ortogonais foi bem menor que na abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as quantidades médias de padrões utilizados pelas abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 213, 12 e 12;

Conclusões

- As abordagens baseadas em ortogonalidade obtiveram resultados semelhantes aos da abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.843, 0.840 e 0.839;
  - Considerando os melhores parâmetros para a média dos resultados, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.808, 0.813 e 0.782.
- A quantidade de padrões utilizados na geração das regras nas abordagens ortogonais foi bem menor que na abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as quantidades médias de padrões utilizados pelas abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 213, 12 e 12;
  - Considerando os melhores parâmetros para a média dos

Conclusões

- As abordagens baseadas em ortogonalidade obtiveram resultados semelhantes aos da abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.843, 0.840 e 0.839;
  - Considerando os melhores parâmetros para a média dos resultados, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.808, 0.813 e 0.782.
- A quantidade de padrões utilizados na geração das regras nas abordagens ortogonais foi bem menor que na abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as quantidades médias de padrões utilizados pelas abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 213, 12 e 12;
  - Considerando os melhores parâmetros para a média dos

Conclusões

- As abordagens baseadas em ortogonalidade obtiveram resultados semelhantes aos da abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.843, 0.840 e 0.839;
  - Considerando os melhores parâmetros para a média dos resultados, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.808, 0.813 e 0.782.
- A quantidade de padrões utilizados na geração das regras nas abordagens ortogonais foi bem menor que na abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as quantidades médias de padrões utilizados pelas abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 213, 12 e 12;
  - Considerando os melhores parâmetros para a média dos

Conclusões

- As abordagens baseadas em ortogonalidade obtiveram resultados semelhantes aos da abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.843, 0.840 e 0.839;
  - Considerando os melhores parâmetros para a média dos resultados, as médias de acurácia obtidas para as abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 0.808, 0.813 e 0.782.
- A quantidade de padrões utilizados na geração das regras nas abordagens ortogonais foi bem menor que na abordagem clássica:
  - Considerando os melhores parâmetros para cada base, as quantidades médias de padrões utilizados pelas abordagens LAC, OLAC e ORIGAMI foram, respectivamente, 213, 12 e 12;
  - Considerando os melhores parâmetros para a média dos