Classificação intervalar com aplicação em imagens coloridas

Rodrigo Augusto Dias Faria Orientador: Prof. Dr. Roberto Hirata Jr

> Instituto de Matemática e Estatística Universidade de São Paulo

> > São Paulo, 2016

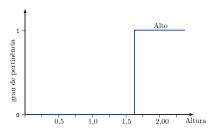
Agenda

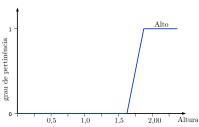
- 1 Introdução
- 2 Fundamentação Teórica
 - Modelos de cores
 - Teoria fuzzy
 - Classificadores
- 3 Experimentos Preliminares
 - Conjunto de dados UCI
 - Conjunto de dados SFA
 - Primeiro experimento
 - Segundo experimento
 - Terceiro experimento
- 4 Plano de Trabalho

Motivação

- Em muitos problemas, não há dificuldade em determinar se um dado elemento é ou não parte de um grupo
- $7 \in \mathbb{N}$ e $-7 \notin \mathbb{N}$
- Diversos fenômenos na natureza não podem ser classificados com conjuntos clássicos
- Relação de pertinência não é bem definida (Pedrycz e Gomide, 1998)
- Incerteza e imprecisão nos conjuntos de dados
- Explorar a capacidade de conjuntos fuzzy de expressar transições graduais de pertinência e não pertinência

Motivação





Trabalhos relacionados

Classificação com conjuntos fuzzy

- Algoritmo de árvore de decisão fuzzy (Umano et al., 1994), adaptado do ID3 clássico proposto por Quinlan (1986)
- Bhatt et al. (2009) segmentaram regiões de pele no espaço RGB; cinco clusters com fuzzy c-means (Bezdek et al., 1984)
- FuzzyDT proposta por Cintra et al. (2013), baseado no C4.5 (Quinlan, 1993)
- Formulação geral de *kernel* sobre conjuntos *fuzzy* (Guevara *et al.*, 2014)

Trabalhos relacionados

Detectores de pele

- Regra de decisão Bayesiana com um modelo de histograma 3-dimensional; histogramas de tamanho 32 mostraram o melhor desempenho com uma taxa de erro de 88% (Jones e Rehg, 2002)
- Classificação com regras no modelo de cores YCbCr; taxa de verdadeiro positivo de 90,66% (Kovac et al., 2003)
- Yogarajah et al. (2011) desenvolveram uma técnica onde os limiares definidos nas regras são adaptados dinamicamente

Trabalhos relacionados

Comparação do modelo de cores

- Desempenho ótimo dos classificadores de pele é independente do modelo de cores (Albiol et al., 2001)
- Abordagens Gaussiana e histograma em 805 imagens coloridas em 9 espaços de cores distintos; SCT, HSI e CIELab com abordagem de histograma (Jayaram et al., 2004)
- 10 espaços de cor com base no *k-means* em 15 imagens do AR; YCgCr, YDbDr e **HSV** (Chaves-González *et al.*, 2010)
- Kaur e Kranthi (2012) similar ao proposto por Kovac et al.
 (2003) com operações morfológicas e de filtragem no YCbCr e
 CIELab, ignorando o componente de luminância em ambos
- Técnica similar implementada em Shaik et al. (2015) e
 Kumar e Malhotra (2015) nos espaços de cores HSV e YCbCr

Objetivos

- Estudo de conjuntos e números fuzzy
- Modelagem de conjuntos fuzzy para classificação
- Estudo de classificação fuzzy
- Escolher uma aplicação real para aplicar a modelagem *fuzzy*
- Compreender a influência do espaço de cores para a modelagem dos dados

Introdução

- Os seres humanos têm a capacidade de discernir milhares de tonalidades e intensidades
- A percepção humana das cores se dá pela ativação de células nervosas que enviam mensagens ao cérebro sobre brilho (brightness), matiz (hue) e saturação (saturation)
- As cores podem ser especificadas por modelos matemáticos em tuplas de números em um sistema de coordenadas

Modelo de Munsell

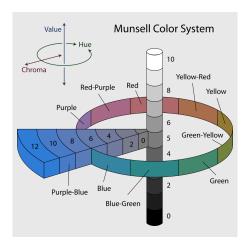
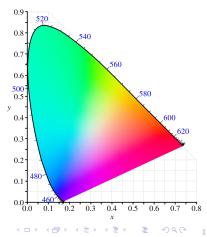


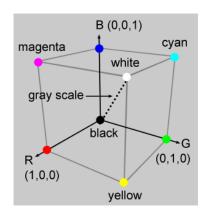
Diagrama de cromaticidade CIE 1931

- Primeiro modelo matemático de especificação numérica da cor
- Componente de luminância Y; X e Z de cromaticidade (tristímulus)
- Derivações do CIE XYZ:
 CIE 1976 L*u*v* e 1976
 CIE L*a*b*



Modelo RGB

- Modelo de cores aditivo
- Baseado na teoria tricromática de Thomas Young e Hermann Helmholtz em meados do século 19



Modelos da família YUV

- Y = luminância, U = Azul Y, V = Vermelho Y
- Utilizado em sistemas de transmissão analógica de televisão nos padrões PAL e SECAM
- YCbCr é um modelo desta família e é largamente utilizado em vídeos digitais

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Modelos da família HSI

■ (I)ntensidade é decomposta da informação de crominância

$$H = \begin{cases} 60\frac{(G-B)}{M-m}, & \text{se } M = R \\ 60\frac{(B-R)}{M-m} + 120, & \text{se } M = G \\ 60\frac{(R-G)}{M-m} + 240, & \text{se } M = B \end{cases}$$

$$S = \begin{cases} \frac{(M-m)}{M}, & \text{se } M \neq 0 \\ 0, & \text{caso contrário} \end{cases}$$

$$V = M$$

Introdução

- Visão tradicional e alternativa da ciência sobre a incerteza (Klir e Yuan, 1995)
- Com base na ideia moderna de que a incerteza é algo útil na ciência, Zadeh propôs a teoria de conjuntos fuzzy
- Capacidade de conjuntos fuzzy expressarem transições graduais de pertinência e não pertinência
- Representação significativa e poderosa da medida de incerteza
- Forma de expressar conceitos vagos em linguagem natural

Da teoria de conjuntos clássicos tem-se a função característica:

$$\mu_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases}$$

Que é um mapeamento dos elementos de U no conjunto binário $\{0,1\}$:

$$\mu_A = U \to \{0, 1\}$$

 $\forall x \in U$, se $\mu_A(x) = 1$, então $x \in A$, se $\mu_A(x) = 0$, então $x \notin A$.

Em conjuntos fuzzy, generalização aplicada no intervalo [0,1]:

$$\mu_A = U \rightarrow [0,1]$$

Definição de conjuntos fuzzy

Um conjunto fuzzy A é um subconjunto do conjunto universo U formado por pares ordenados de um elemento qualquer x e seu grau de pertinência dado por $\mu_A(x)$, da forma:

$$A = \{(x, \mu_A(x)) \mid x \in U\}$$

As noções de inclusão, união, intersecção, complemento, relação, convexidade, etc., oriundas da teoria de conjuntos clássica, são estendidas a esses conjuntos (Zadeh, 1965).

- U pode ser composto por elementos discretos ou ser um espaço contínuo
- A mesma implicação vale para o subconjunto A
- Quando U é um conjunto discreto e finito, tal que $U = \{x_1, x_2, x_3, \dots, x_n\}$, pode-se simplesmente enumerar os seus elementos, juntamente com seus graus de pertinência:

$$A = \frac{\mu_A(x_1)}{x_1} + \frac{\mu_A(x_2)}{x_2} + \ldots + \frac{\mu_A(x_n)}{x_n} = \sum_{i=1}^n \frac{\mu_A(x_i)}{x_i}$$

Definição de altura

A altura de A, denotado por altura(A), corresponde ao limite superior do codomínio da sua função de pertinência, da forma:

$$altura(A) = \{ \mu_A(x) \mid x \in U \}$$

Definição de suporte

O suporte de um conjunto fuzzy A em U, denotado por suporte(A), é o conjunto dado por:

$$suporte(A) = \{x \in U \mid \mu_A(x) > 0\}$$

Definição de α -corte ou α -nível

Um α -corte ou α -nível, é o subconjunto clássico de elementos cujo grau de pertinência é maior ou igual a um valor α , formalmente:

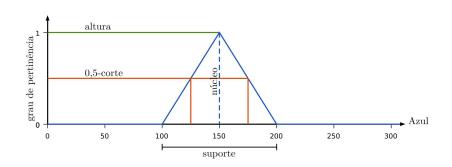
$$\alpha - corte(A) = \{x \in U \mid \mu_A(x) \ge \alpha\}$$

Definição de núcleo ou kernel

O núcleo ou kernel de um conjunto fuzzy A em U, \acute{e} o conjunto de elementos pertencentes inteiramente à A, da forma:

$$n\'ucleo(A) = \{x \in U \mid \mu_A(x) = 1\}$$

Representação gráfica das principais propriedades dos conjuntos *fuzzy*



Números *fuzzy*

Um número fuzzy é um tipo especial de conjunto fuzzy definido no conjunto $\mathbb R$ dos números reais, da forma (Klir e Yuan, 1995):

$$A:\mathbb{R}\to [0,1]$$

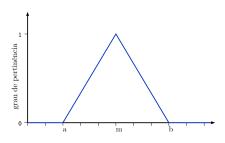
Para que A seja, de fato, um número fuzzy, o conjunto universo no qual μ_A está definida deve ser \mathbb{R} e as seguintes propriedades devem ser satisfeitas (Barros e Bassanezi, 2006):

- (i) todos os α -corte de A são não vazios, com $0 \le \alpha \le 1$
- (ii) todos os α -corte são intervalos fechados de $\mathbb R$
- (iii) suporte(A) = $\{x \in U \mid \mu_A(x) > 0\}$

Funções de pertinência

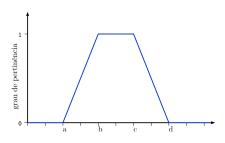
Um número fuzzy A é dito triangular se sua função de pertinência, denotada por $\mu_A(x)$, é da forma:

$$\mu_{\mathcal{A}}(x) = egin{cases} rac{x-a}{m-a}, & ext{se } a < x \leq m & ext{polymer} \ rac{b-x}{b-m}, & ext{se } m < x < b & ext{op} \ 0, & ext{c.c.} \end{cases}$$



Funções de pertinência

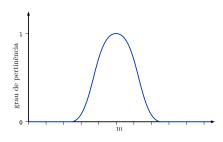
Um número fuzzy A é dito trapezoidal se sua função de pertinência, denotada por $\mu_A(x)$, é da forma:



Funções de pertinência

Um número fuzzy A é dito gaussiano se sua função de pertinência, denotada por $\mu_A(x)$, é da forma:

$$\mu_{A}(x) = \begin{cases} \exp\left(-\frac{(x-m)^{2}}{\sigma}\right), & \text{exp of } (x-m) = 0, \\ \sin(mx) - \sin(mx) = 0, \\ \cos(mx) - \cos(mx) = 0, \\ \cos(mx) -$$



Introdução

Seja o conjunto de dados de treinamento com N amostras da forma:

$$D = (x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$$

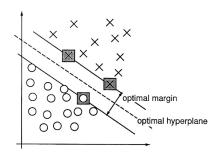
onde i = 1, 2, ..., N, $y_i \in Y$ e $Y = \{+1, -1\}$, e cada x_i é um vetor d-dimensional da forma:

$$x = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{bmatrix}$$

onde $x \in X$ e X é o espaço de entrada, ou seja, todos os x vetores possíveis.

- SVMs têm a habilidade de gerar um hiperplano ou conjunto de hiperplanos num espaço de alta ou infinita dimensionalidade (Duda et al., 2012)
- Assumindo que D é linearmente separável (Lorena e Carvalho, 2003):

$$w \cdot x + b = 0$$



Outros dois hiperplanos paralelos ao hiperplano ótimo podem ser obtidos:

$$\begin{cases} w \cdot x + b = +1 \\ w \cdot x + b = -1 \end{cases}$$

Restrições são definidas para evitar que não existam pontos entre $w \cdot x + b = 0$ e $w \cdot x + b = \pm 1$ (Lorena e Carvalho, 2003):

$$\begin{cases} w \cdot x_i + b \ge +1, & \text{se } y_i = +1 \\ w \cdot x_i + b \le -1, & \text{se } y_i = -1 \end{cases}$$

ou, equivalentemente:

$$y_i(w \cdot x_i + b) \geq 1$$

- A minimização de ||w|| maximiza a margem e, sendo assim, tem-se um problema de otimização
- w e b ótimos que resolvem este problema definem o classificador e podem ser obtidos por multiplicadores de Lagrange (Campbell, 2000):

Maximizar:
$$\sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j x_i \cdot x_j$$

Sujeito a:
$$\begin{cases} \alpha_i \geq 0 \\ \sum_{i=1}^N \alpha_i y_i = 0 \end{cases}$$

SVMs lineares podem ser estendidas:

$$D' = (\Phi(x_1), y_1), (\Phi(x_2), y_2), \dots, (\Phi(x_N), y_N)$$

A forma do hiperplano ótimo agora é definida por:

$$w \cdot \Phi(x) + b = 0$$

O problema de otimização pode ser resolvido como (Lorena e Carvalho, 2003):

Maximizar:
$$\sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j \Phi(x_i) \cdot \Phi(x_j)$$

Kernels são funções que têm a finalidade de projetar os vetores de entrada num espaço de características com número de dimensões exponencial ou infinito (Shawe-Taylor e Cristianini, 2004):

$$k(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$$

Alguns dos kernels mais utilizados são (Lorena e Carvalho, 2003):

Kernel linear

$$k(x_i,x_j)=(x_i\cdot x_j)$$

Kernel polinomial

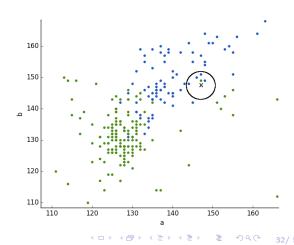
$$k(x_i, x_i) = (\gamma x_i \cdot x_i + r)^g$$

■ Kernel gaussiano ou RBF

$$k(x_i, x_j) = \exp\left(-\gamma ||x_i - x_j||^2\right)$$

k-Vizinhos Mais Próximos (k-NN)

- k-NN é um algoritmo baseado em instâncias
- Rotula uma nova amostra x com a classe de maior frequência dentre as k mais próximas
- Decisão por maioria de votos (Duda et al., 2012)



k-Vizinhos Mais Próximos (k-NN)

A distância entre duas amostras x_i e x_j quaisquer pode ser obtida em termos da distância Euclidiana:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^{d} (a_r(x_i) - a_r(x_j))^2}$$

Para classificar uma nova amostra x_q (Mitchell, 1997):

$$g(x_q) = \underset{y \in Y}{\operatorname{argmax}} \sum_{i=1}^k \delta(y, f(x_i))$$

onde

$$\delta(y, f(x_i)) = \begin{cases} 1, & \text{se } y = f(x_i) \\ 0, & \text{caso contrário} \end{cases}$$

k-Vizinhos Mais Próximos (k-NN)

- Uma variação da função é a atribuição de pesos a cada um dos k vizinhos, conforme sua distância
- Implica que pontos mais próximos de x_q têm maior influência na sua rotulação

$$g(x_q) = \underset{y \in Y}{\operatorname{argmax}} \sum_{i=1}^k w_i \delta(y, f(x_i))$$

onde

$$w_i = \frac{1}{d(x_a, x_i)^2}$$

No caso em que $d(x_q, x_i) = 0$, $g(x_q)$ pode assumir o mesmo valor de $f(x_i)$ (Mitchell, 1997).

Árvores de decisão

- Processo iterativo onde um atributo (nó) é escolhido como raiz até algum nó folha, onde a classe é atribuída à amostra
- ID3 avalia cada atributo através de um teste estatístico para determinar o quão bem ele, por si só, classifica as amostras de treinamento (Quinlan, 1986)
- O melhor atributo é selecionado como o nó raiz da árvore
- Um ramo descendente do nó raiz é criado para cada valor possível deste atributo (Mitchell, 1997)
- Algoritmo de busca guloso (Mitchell, 1997)
- C4.5 estende o ID3 para possibilitar o uso de atributos contínuos, dados ausentes e poda da árvore Quinlan (1993)



Árvores de decisão

Para medir a impureza de uma partição, o ID3 usa o conceito de entropia (Quinlan, 1986):

$$H(D) = -y_{\oplus} \log_2 y_{\oplus} - y_{\ominus} \log_2 y_{\ominus}$$

O teste estatístico, conhecido como ganho de informação, mede a efetividade de um atributo na classificação dos dados (Quinlan, 1986):

$$IG(D, a_r) = H(D) - \sum_{v \in V(a_r)} \frac{|D_v|}{|D|} H(D_v)$$

onde $V(a_r)$ é o conjunto de todos os possíveis valores do atributo a_r , D_v é o subconjunto de D no qual o atributo a_r tem o valor v.

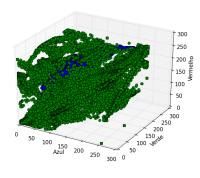
Repositório da Universidade da Califórnia em Irvine

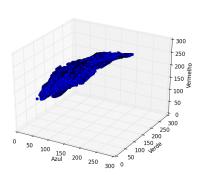
- Contém 245.057 amostras obtidas de imagens do PAL e FERET (Minear e Park, 2004; Phillips et al., 1996)
- 194.198 são de pixels não pele e 50.859 de pixels de pele
- Compostas por 3 atributos $x = [x_1, x_2, x_3], x \in \mathbb{R}^3$, que representam os canais do modelo de cores RGB
- Uma quarta coluna determina a classe a qual a amostra x pertence, denotada por y, sendo $y \in Y$ e $Y = \{+1, -1\}$

Amostras do conjunto de dados

В	G	R	Classe
74	85	123	1
207	215	255	1
74	82	122	1
202	211	255	1
54	72	125	1
166	164	116	-1
148	150	91	-1
29	26	5	-1
167	166	115	-1
180	177	133	-1

Visão 3-dimensional dos canais RGB





Banco de imagens de faces do FERET e AR

- 876 imagens de faces obtidas do FERET, criado por Phillips et al. (1996) e 242 do AR, proposto por Martínez e Benavente (1998)
- As imagens do AR têm fundo branco e pequenas variações de cor da pele e, portanto, o ambiente é mais controlado
- Três amostras de pele e cinco não pele foram geradas aleatoriamente considerando a máscara ground truth de cada imagem

Exemplos de imagens do banco de faces



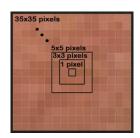




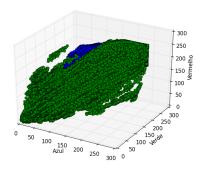


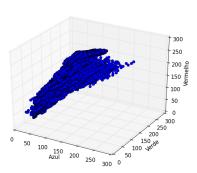
Estrutura das janelas

- Cada amostra é uma janela de tamanho $n \times n$, sendo n ímpar, que varia de 1×1 até 35×35
- O conjunto de dados foi gerado com janela 9 x 9, totalizando 724.464 amostras, sendo 271.674 de pele e 452.790 não pele



Visão 3-dimensional dos canais RGB





Primeiro experimento

- Realizado com *k*-NN e SVM no espaço de cores RGB
- Estratégia de validação cruzada 10-fold em ambos
- Os parâmetros ótimos foram:
 - **k-NN:** n_neighbors=3, weights=uniform no UCI e n_neighbors=15, weights=uniform no SFA
 - SVM: kernel=rbf, C=100 e gamma=1e-3 em ambos UCI e SFA

Tabela de busca

kernel			С			degree			
rbf	1	10	100	1000	1e-3	1e-4	1e-5		
poly	1	10	100	1000		1e-4	1e-5	3	4
linear	1	10	100	1000					

Tabela: Tabela de busca dos parâmetros do estimador ótimo na SVM.

n_ neighbors								weig	algorithm			
3	5	9	15	25	50	100	200	400	800	distance uniform		auto

Tabela: Tabela de busca dos parâmetros do estimador ótimo no k-NN.

Resultados

O treinamento foi executado com 10 tarefas em paralelo em ambos os classificadores e 30% dos dados, aleatoriamente, foram separados para teste.

Dados	Classificador	Modelo de cores	Precision	Recall	F1-score
UCI	k-NN	RGB	0,9995	0,9995	0,9995
	SVM	RGB	0,9995	0,9995	0,9995
SFA	k-NN	RGB	0,9672	0,9669	0,9670
SFA	SVM	RGB	0,9643	0,9628	0,9638

Segundo experimento

- Realizado com k-NN e SVM usando o conjunto de dados SFA nos espaços de cores RGB, HSV, Lab e YCbCr
- O componente de luminância foi ignorado para que um teste somente com os componentes de crominância fosse realizado
- Estratégia escolhida de validação cruzada 10-fold
- Os parâmetros ótimos obtidos no primeiro experimento foram fixados aqui
- Treinamento com 10 tarefas em paralelo, 30% dos dados, aleatoriamente, foram separados para teste

Resultados

Modelo de cores	Classificador	Precision	Recall	F1-score	
RGB	k-NN	0,9672	0,9669	0,9670	
INGD	SVM	0,9643	0,9628	0,9638	
HSV	k-NN	0,9676	0,9673	0,9674	
1134	SVM	0,9718	0,9677	0,9679	
HS	k-NN	0,9215	0,9194	0,9199	
113	SVM	0,9305	0,9302	0,9302	
Lab	k-NN	0,9671	0,9660	0,9670	
Lau	SVM	0,9675	0,9665	0,9672	
ab	k-NN	0,9444	0,9439	0,9440	
au	SVM	0,9451	0,9447	0,9446	
YCbCr	k-NN	0,9679	0,9677	0,9677	
I CDCI	SVM	0,9635	0,9633	0,9632	
CbCr	k-NN	0,9487	0,9482	0,9483	
CDCI	SVM	0.9496	0.9492	0.9493	

Terceiro experimento

- Realizado com árvore de decisão fuzzy proposta por Cintra et al. (2013) usando o conjunto de dados SFA nos espaços de cores RGB, HSV, Lab e YCbCr
- Os parâmetros do FuzzyDT podem ser customizados para determinar a poda da árvore e o método de estimativa do número de conjuntos fuzzy por atributo
- Desenvolvido um algoritmo de tabela de busca para encontrar os parâmetros ótimos

Tabela de busca

Dados	Nível de confiança				ľ	∕létoc	# conj. fuzzy		
RGB	10	15	20	25	infogain	wm	rf	fixed	2-9
HSV	10	15	20	25	infogain	wm	rf	fixed	2-9
Lab	10	15	20	25	infogain	wm	rf	fixed	2-9
YCbCr	10	15	20	25	infogain	wm	rf	fixed	2-9

Resultados

Modelo de cores	Taxa de erro	Método	# conjuntos fuzzy		
RGB	3,00	fixed	4		
HSV	3,23	fixed	3		
Lab	2,84	fixed	6		
YCbCr	2,72	fixed	8		

Disciplinas cursadas

Disciplina	Término
Aprendizagem Computacional: Modelos, Algoritmos e Aplic.	1° Sem/2015
Introdução à Computação Gráfica	1° Sem/2015
Visão e Processamento de Imagens - Parte I	1° Sem/2015
Análise de Algoritmos	2° Sem/2015
Métodos de Aprendizagem em Visão Computacional	2° Sem/2015
Linguagens, Autômatos e Computabilidade	1° Sem/2016
Programação Orientada a Objetos	1° Sem/2016

Atividades previstas

- Revisar leituras adicionais
- Incorporar novos conjuntos de dados
- 3 Investigar características passíveis de uso em classificadores de dados intervalares
- 4 Desenvolver ferramentas para dar suporte aos experimentos subsequentes
- **5** Elaborar novos experimentos com base nas ferramentas desenvolvidas e conjuntos de dados estabelecidos
- 6 Analisar os resultados e reportá-los no projeto de pesquisa
- 7 Publicar resultados em artigos científicos
- Escrever a dissertação



Cronograma

		Meses 2016/2017										
Atividade	dez	jan	fev	mar	abr	mai	jun	jul	ago	set	out	
1	X	×	Х									
2	Х	Х										
3		Х	Х									
4			Х	Х	Х	Х						
5					Х	Х	Х					
6						Х	Х	х				
7							Х	х				
8								х	Х	X	Х	