

Avaliação da Identificação do Sexo de Pessoas Não-Identificadas por meio de Radiografias Faciais com o uso de Métodos de Aprendizado de Máquina*

Thais Mayumi Oshiro^{1,2}, Suzana Papile Maciel Carvalho³, Arsenio Sales Peres⁴, Renato Tinós¹, Ricardo Henrique Alves da Silva⁵, José Augusto Baranauskas¹

¹Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo

²Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo

³Faculdade de Odontologia, Universidade de São Paulo

⁴Faculdade de Odontologia de Bauru, Universidade de São Paulo

⁵Faculdade de Odontologia de Ribeirão Preto, Universidade de São Paulo

thais_ibm4@yahoo.com.br, sumaciel@uol.com.br, arsenio@usp.br,

rtinos@usp.br, ricardohenrique@usp.br, augusto@usp.br

Abstract. *Identification corresponds to the set of different procedures to identify a person or object. The process of identifying individuals by manual radiographic comparison appears to be very slow requiring high skills and accuracy by the human expert. Applying computer systems to assist the identification can improve this process, making it faster and more practical. Therefore, using measurements obtained from human frontal sinus radiographs this study aims to use Machine Learning methods and techniques to evaluate if it is possible to establish patterns among genders, thereby assisting in the identification process, using proposed measures from specific literature. Our results indicate the proposed measures are more indicated for specialization rather than generalization.*

Resumo. *A identificação corresponde ao conjunto de procedimentos diversos para individualizar uma pessoa ou objeto. O processo de identificação de indivíduos por comparação manual radiográfica é um processo muito lento e que exige altos níveis de habilidade e precisão por parte do especialista humano. O emprego de sistemas computacionais para auxiliar o processo de identificação pode tornar esse processo mais rápido e prático. Assim, neste trabalho visa-se, com a utilização de medidas obtidas a partir de radiografias dos seios frontais humanos, utilizar métodos e técnicas de Aprendizado de Máquina para verificar se é possível estabelecer padrões entre os sexos, auxiliando assim, o processo de identificação, por meio de medidas propostas em literatura específica da área. Os resultados indicam que as medidas propostas são mais adequadas para especialização do que para generalização.*

*Projeto de pesquisa realizado com apoio financeiro do CNPq/FAPEAM — INCT ADAPTA.

1. Introdução

Alguns acidentes fatais, quando ocorrem, deixam como consequência vítimas desfiguradas e impossíveis de serem reconhecidas em um primeiro momento. Devido ao fato de todos os seres humanos possuírem uma identidade em vida, a sociedade requer que esta identidade seja reconhecida após a morte, seja para conforto dos familiares da vítima ou para a resolução de questões jurídicas [Sweet 2001]. A identificação corresponde ao conjunto de procedimentos diversos para individualizar uma pessoa ou objeto [Vanrell 2002]. Uma das formas de se realizar o processo de identificação consiste em utilizar as impressões digitais da vítima, caso isso seja possível. Caso contrário, a identificação pode ser feita por meio da comparação manual de radiografias dentárias do indivíduo. Todavia, esse processo é muito lento e exige muita habilidade do especialista humano responsável. Existe, assim, um interesse em desenvolver sistemas computacionais capazes de realizar tal comparação, de um modo mais preciso e rápido.

A comparação de registros dentários por meio de sistemas computacionais tem se mostrado um meio adequado para a manipulação de grandes volumes de informação, preservando precisão, consistência e baixo custo [Nassar and Ammar 2007]. Uma vez que há uma radiografia *antemortem* do indivíduo a ser identificado, pode-se comparar as duas radiografias, *antemortem* e *post-mortem*, e assim, identificá-lo. Porém, se tal radiografia não existir, a possibilidade de classificar o sexo do indivíduo já contribuiria muito no processo de identificação, uma vez que diminuiria o espaço de possibilidades para o especialista humano. Uma maneira de realizar isso consiste em primeiro classificar o sexo do indivíduo a partir da sua radiografia dos seios frontais da face e, posteriormente, classificar outras características que se mostrarem necessárias para o processo de identificação. A Odontologia Legal exerce atualmente importância fundamental na identificação de vítimas carbonizadas, em decomposição, mutiladas ou esqueletizadas, uma vez que os dentes e os arcos dentários podem fornecer subsídios de real valor para a resolução de problemas médico-legais e criminológicos. O uso da técnica odontológica provê bons resultados, uma vez que os dentes e os materiais odontológicos apresentam uma considerável resistência à ação do calor e do fogo. Segundo [Oliveira et al. 1998], a contribuição da Odontologia Legal nos processos de identificação humana *post-mortem* está presente desde os procedimentos iniciais (identificação geral): estimativas de sexo e idade, determinação de grupo étnico, cor da pele e outras características como estatura, no diagnóstico de manchas ou líquidos provenientes da cavidade bucal, ou nela contidos; ou mesmo na definição da causa e do tempo de morte até a irrefutável possibilidade de identificação individual.

Existem algumas técnicas de padronização de medidas já desenvolvidas para serem utilizadas no processo de identificação. Uma delas é o método de [Ribeiro 1993], o qual se baseia em quatro medidas obtidas a partir das radiografias dos seios frontais. Essas medidas podem ser obtidas a partir das radiografias produzidas na prática profissional tanto em Odontologia quanto na prática médica. Neste sentido, o objetivo deste estudo consiste em avaliar a viabilidade das medidas dos seios frontais obtidas método de [Ribeiro 1993] para uso tanto no reconhecimento (especialização) quanto na classificação (generalização) de gêneros.

Este documento encontra-se organizado da seguinte forma: na Seção 2 é descrita a metodologia empregada neste estudo para obter as medidas de seios frontais da face;

com base nestas medidas foram realizados experimentos utilizando diversos paradigmas de Aprendizado de Máquina, cujos resultados são reportados e discutidos na Seção 3. Finalmente, na Seção 4 são descritas as conclusões obtidas neste estudo.

2. Metodologia

O trabalho de [Ribeiro 1993] consiste em um método de padronização de medidas feitas em radiografias. Esse método consiste em uma linha traçada sobre os bordos superiores das cavidades orbitárias (linha base) e quatro linhas projetadas perpendicularmente a esta, que é ilustrado à esquerda da Figura 1 (esse desenho é feito em papel manteiga, o qual é sobreposto às radiografias, apoiado sobre um negatoscópio), obtendo-se assim quatro medidas ilustradas à direita na Figura 1. A medida A corresponde ao diâmetro maior dos seios frontais; a medida B corresponde a distância entre os pontos mais altos; a medida C corresponde a distância entre o limite lateral e o ponto mais alto do seio frontal direito e a medida D corresponde a distância entre o limite lateral e o ponto mais alto do seio frontal esquerdo. A primeira linha projetada passa pelo limite lateral máximo do seio frontal direito (linha E). A segunda, pelo ponto mais alto (mais distante da linha de base) do seio maxilar direito (linha F). A terceira, pelo ponto mais alto do seio maxilar esquerdo (linha G), e a quarta, pelo limite lateral máximo do seio frontal esquerdo (linha H).

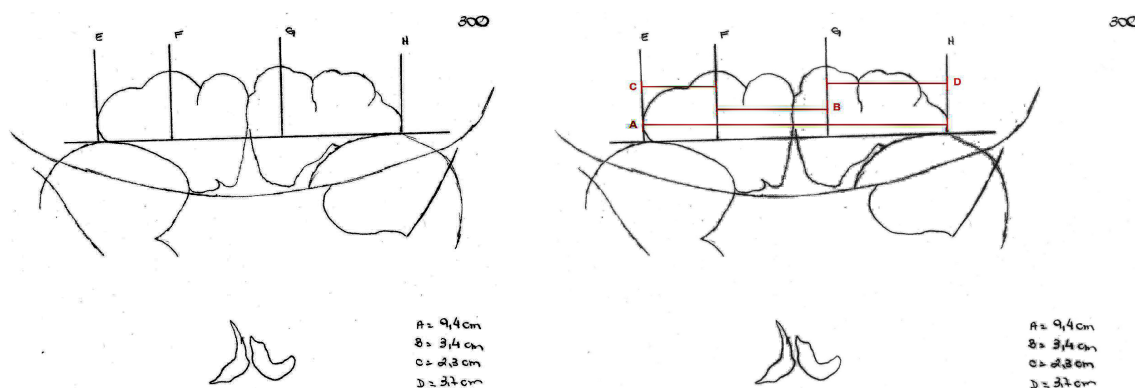


Figura 1. Linhas projetadas e suas respectivas medidas

Neste estudo foram utilizadas 184 radiografias dos seios frontais de indivíduos adultos (com idade superior a 20 anos), sendo que destas, 148 pertenciam a mulheres e 36 pertenciam a homens¹. As radiografias foram cedidas pela cidade de Bauru e avaliadas segundo o método de [Ribeiro 1993] para a obtenção das medidas propostas por cirurgiões dentistas e profissionais da área de Odontologia. Dessa forma, com base nessas quatro medidas e utilizando algoritmos de Aprendizado de Máquina é possível construir um classificador para realizar uma identificação desejada, nesse caso, a identificação do sexo do indivíduo. Neste sentido, há duas perguntas que este estudo procura responder: (1) *É possível identificar o sexo de um indivíduo previamente conhecido?*; (2) *É possível identificar o sexo de um indivíduo previamente desconhecido?* Em geral, a primeira pergunta está relacionada à capacidade das quatro medidas serem utilizadas no reconhecimento do sexo dos indivíduos conhecidos e uma possível medida é taxa de erro (acerto) de *especialização*, ou seja, a taxa de erro (acerto) medida utilizando-se próprio conjunto

¹O presente projeto tem aprovação pelo Comitê de Ética em Pesquisa.

de treinamento. A segunda pergunta está relacionada à capacidade de generalização que essas quatro medidas podem proporcionar; ou seja, se é possível identificar o sexo de um indivíduo nunca antes visto pelos algoritmos de aprendizado. Esta segunda pergunta pode ser mensurada pela taxa de erro (acerto) de *generalização* medida em exemplos não vistos previamente pelo classificador [Domingos 1999]. Neste sentido, é importante salientar que medidas adequadas para responder a primeira pergunta não são, necessariamente adequadas para responder a segunda [Mitchell 1997, Bishop 2006].

Neste estudo, para responder as duas perguntas anteriores foram utilizados algoritmos de Aprendizado de Máquina englobando quatro paradigmas [Monard and Baranauskas 2003]: Árvores de Decisão, Redes Neurais, Vizinhos mais Próximos e Bayesiano a fim de buscar por modelos capazes de diferenciar entre os sexos feminino e masculino. Foi utilizada a biblioteca Weka [Witten and Frank 2005], que contempla uma variedade de algoritmos de Aprendizado de Máquina. Tal biblioteca consegue lidar com valores desconhecidos, tanto para o aprendizado quanto para a classificação; porém, neste estudo, não há ocorrência de valores desconhecidos, uma vez que os profissionais responsáveis pelas medições conseguiram obter todas as medidas necessárias. Para o desenvolvimento deste trabalho, foram utilizados os indutores de Árvores de Decisão (J48), Perceptron Multicamadas (*Multilayer Perceptron* - MLP) e K-vizinhos mais próximos (*K-nearest Neighbors* - IBK) e Naive Bayes da biblioteca Weka. Para o indutor J48, os valores do parâmetro 'número mínimo de objetos por folha' utilizados neste estudo foram 2, 3 e 5 (indicados por J48(2), J48(3) e J48(5), respectivamente). Para o algoritmo de Vizinhos mais Próximos, IBK, variou-se o parâmetro 'número de vizinhos mais próximos' K em 3, 5 e 7 (indicados por IBK(3), IBK(5) e IBK(7), respectivamente). No caso de Redes Neurais MLP, variou-se o parâmetro da 'semente do gerador de números aleatórios' que determina os pesos iniciais da MLP em 0, 1 e 2 (indicadas por MLP(0), MLP(1) e MLP(2), respectivamente). O algoritmo Naive Bayes (NB) foi utilizado com seus parâmetros *default* assim como todos os demais parâmetros dos demais algoritmos.

Em um estudo preliminar, o conjunto original contendo 148 mulheres e 36 homens foi analisado sendo que não foi possível obter bons resultados (não mostrados). Uma vez que a classificação errada de todos ou da maior parte dos exemplos masculinos pode ser explicada devido a diferença significativa entre o número de exemplos femininos e o número de exemplos masculinos, 148 (80.43%) e 36 (19.57%), respectivamente, esse fato pode induzir os algoritmos de aprendizado a entenderem que, por haver mais exemplos de uma determinada classe, essa classe é mais importante, dando assim maior relevância para a mesma. Portanto, a solução foi obter amostras desse conjunto, com quantidades iguais de cada classe, pois segundo [Batista et al. 2004], classificadores induzidos a partir de classes balanceadas artificialmente, geralmente apresentam resultados melhores do que aqueles com distribuição natural das classes. Assim, a partir do conjunto de dados original contendo 184 exemplos, foram definidos 10 subconjuntos (T1, T2, ..., T10) contendo 72 exemplos cada — Figura 2. O subconjunto T1 contém todos os 36 exemplos do sexo masculino e uma amostra de 36 exemplos do sexo feminino; o subconjunto T2 contém também todos os 36 exemplos do sexo masculino e uma outra amostra de 36 exemplos do sexo feminino e assim sucessivamente para T3, ..., T10. Portanto, cada subconjunto contém todos os exemplos do sexo masculino e uma amostra dos exemplos do sexo feminino.

Na seção seguinte são apresentados os resultados dos experimentos realizados neste estudo. Foram realizados dois experimentos: o primeiro visando analisar se as medidas obtidas segundo o método de [Ribeiro 1993] poderiam responder a primeira pergunta (especialização), ou seja, identificar o sexo de um indivíduo previamente conhecido; o segundo experimento visando responder a segunda pergunta (generalização), ou seja, analisar se tais medidas permitem classificar o sexo de indivíduos nunca antes vistos pelos algoritmos de aprendizado.

3. Experimentos, Resultados & Discussão

Experimento 1. De forma a analisar se é possível identificar o sexo de um indivíduo previamente conhecido os 10 subconjuntos (T1, T2, ..., T10) foram fornecidos aos algoritmos de aprendizado mencionados na seção anterior. Neste sentido o desempenho dos algoritmos foi avaliado no próprio conjunto de treinamento, ou seja, foi obtida a taxa de acerto de especialização (também conhecida como acerto aparente). Acertos aparentes têm a tendência de serem otimistas mas é importante salientar que a análise de especialização (reconhecimento do sexo de indivíduos) é uma tarefa distinta daquela de classificação (generalização), na qual um modelo mais genérico é obtido a partir dos dados [Hastie et al. 2001]. Assim, neste experimento 1.1, quanto maior é a taxa de acerto, entende-se que é maior a capacidade do algoritmo em reconhecer (especializar) o sexo dos indivíduos previamente conhecidos. Os resultados encontram-se na Tabela 1. Em média, as árvores de decisão acertaram 56.34%, vizinhos mais próximos 69.12% e redes neurais 68.15% sendo que Naive Bayes obteve desempenho de 61.66%. Esses resultados indicam que, para os dados utilizados e as quatro medidas propostas por [Ribeiro 1993] que vizinhos mais próximos e redes neurais são os algoritmos mais indicados. Também são mostrados nessa tabela os *ranks* médios obtidos pelo teste de Friedman-Nemenyi [Demsar 2006]. Como pode ser observado na Tabela 1, o algoritmo de melhor *rank* médio é o IBK utilizando 3 vizinhos mais próximos. Como a análise dos resultados envolve múltiplas comparações foi utilizado o teste de Friedman-Nemenyi com 95% de confiança; a hipótese nula H_0 do teste de Friedman assume que todos os classificadores têm igual desempenho. Nos resultados obtidos, a hipótese nula foi rejeitada; assim, pode-se prosseguir com o teste *post-hoc* de Nemenyi para detectar entre quais classificadores há diferenças significativas. Com isso constatou-se que IBK(3) foi melhor do que IBK(7), J48(2), J48(3), J48(5) e NB; IBK(5) foi melhor do que J48(2) e J48(3); MLP(0) e MLP(1) foram melhores do que J48(2) e J48(3).

Para comprovar estes resultados, este experimento foi estendido para incluir outros parâmetros dos classificadores (experimento 1.2): J48* utilizando ‘número mínimo de objetos por folha’ igual a um e sem poda; IBK* utilizando ‘número de vizinhos mais próximos’ $K = 1$ e MLP* com tempo de treinamento de 500.000 épocas (1000 vezes o valor *default*). Os resultados do algoritmo NB não foram mostrados pois são idênticos àqueles da Tabela 1. O objetivo deste experimento para consistiu em ajustar os parâmetros dos algoritmos em questão para que houvesse o máximo de ajuste (*overfitting*) dos classificadores extraídos [Domingos 1999], ou seja, que a especialização fosse máxima. Os resultados encontram-se na Tabela 2. Observa-se que o classificador IBK* obteve 100% de acerto para todos os subconjuntos, o que é esperado para este algoritmo quando não há exemplos conflitantes. Nas Figuras 3 e 4 encontram-se a melhor e a segunda melhor árvore, respectivamente, obtidas nesta extensão do experimento. Estes resultados

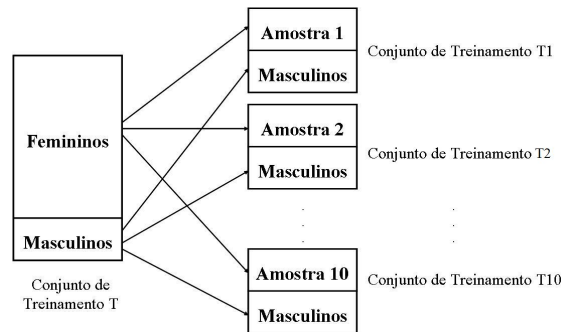


Figura 2. Partição do conjunto original em 10 subconjuntos

Tabela 1. Experimento 1.1 Taxa de acerto de especialização dos classificadores utilizando os atributos originais

	J48(2)	J48(3)	J48(5)	IBK(3)	IBK(5)	IBK(7)	MLP(0)	MLP(1)	MLP(2)	NB
T1	50.00	50.00	50.00	75.00	62.50	58.33	70.83	66.67	68.06	63.89
T2	58.33	58.33	58.33	79.17	69.44	58.33	59.72	59.72	54.17	55.56
T3	50.00	50.00	50.00	77.78	68.06	68.06	72.22	69.44	70.83	58.33
T4	61.11	61.11	61.11	80.56	68.06	55.56	63.89	66.67	65.28	63.89
T5	50.00	50.00	50.00	72.22	66.67	56.94	68.06	65.28	50.00	56.94
T6	50.00	50.00	50.00	75.00	68.06	58.33	68.06	72.22	68.06	61.11
T7	70.83	70.83	75.00	79.17	72.22	61.11	73.61	68.06	73.61	62.50
T8	69.44	69.44	69.44	72.22	72.22	68.06	69.44	73.61	68.06	69.44
T9	50.00	50.00	56.94	66.67	68.06	63.89	68.06	73.61	75.00	65.28
T10	50.00	50.00	50.00	80.56	76.39	75.00	77.78	73.61	70.83	59.72
Média	55.97	55.97	57.08	75.83	69.17	62.36	69.17	68.89	66.39	61.66
Desvio Padrão	2.68	2.68	2.87	1.42	1.18	1.98	1.59	1.44	2.57	1.33
Rank Médio	8.10	8.10	7.50	1.55	3.65	7.20	3.40	3.65	5.25	6.60

indicam que as medidas anatômicas utilizadas parecem ser apropriadas para reconhecer indivíduos previamente conhecidos, sendo que o classificador IBK* mostrou-se o mais indicado para os dados sob análise. Como pode ser observado na Tabela 2, o algoritmo de melhor *rank* médio é o IBK*. Novamente, utilizando o teste Friedman-Nemenyi com 95% de confiança constatou-se que IBK* foi melhor do que J48*.

Experimento 2. A seguir, foi realizado um experimento para verificar se é possível identificar o sexo de um indivíduo desconhecido, ou seja, se é possível classificar corretamente o sexo de indivíduos nunca antes vistos pelos algoritmos de aprendizado. Neste caso, o desempenho dos algoritmos foi avaliado por meio de validação

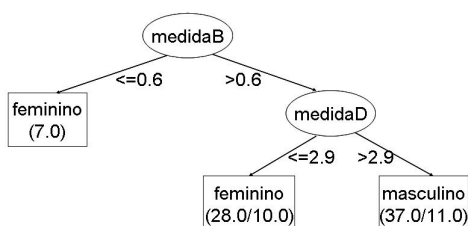


Figura 3. Melhor árvore obtida, com taxa de acerto igual a 70.83%

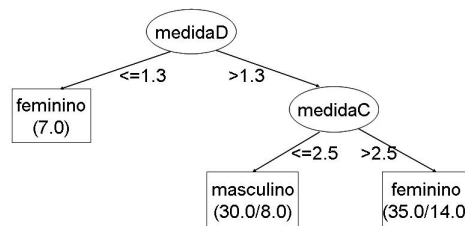


Figura 4. Segunda melhor árvore obtida, com taxa de acerto igual a 69.44%

Tabela 2. Experimento 1.2 Taxa de acerto de especialização dos classificadores (parâmetros adicionais) utilizando os atributos originais

	J48*	IBK*	MLP*
T1	50.00	100.00	75.00
T2	58.33	100.00	70.83
T3	50.00	100.00	80.56
T4	61.11	100.00	81.94
T5	50.00	100.00	79.17
T6	50.00	100.00	77.78
T7	70.83	100.00	76.39
T8	69.44	100.00	73.61
T9	50.00	100.00	73.61
T10	50.00	100.00	80.56
Média	55.97	100.00	76.94
Desvio Padrão	2.68	0	1.16
Rank Médio	3	1	2

cruzada em 10 partições estratificadas (*10-fold stratified cross-validation*) em cada um dos 10 subconjuntos (T1, T2, ..., T10) [Kohavi 1995]. Portanto, neste caso a taxa de acerto de generalização foi mensurada. Inicialmente (experimento 2.1), foram utilizadas as 4 medidas propostas por [Ribeiro 1993], cujos desempenhos dos classificadores encontram-se na Tabela 3 e os maiores valores obtidos, para cada algoritmo analisado, estão destacados. Como é possível notar, em média, as árvores de decisão acertaram 47.63%, vizinhos mais próximos 45.79% e redes neurais 51.70% sendo que Naive Bayes obteve desempenho de 51.87%. Como todos dos conjuntos encontram-se com as mesmas proporções de indivíduos dos sexos feminino e masculino, esses resultados indicam que nenhum paradigma de aprendizado analisado conseguiu obter uma boa separação do sexo dos indivíduos. Como pode ser observado na Tabela 3, o algoritmo de melhor *rank* médio é o MLP utilizando semente igual a 0. Utilizando o teste Friedman-Nemenyi com 95% de confiança constatou-se que MLP(0) e NB foram melhores do que IBK(7).

Visando uma possível melhora nos resultados obtidos este experimento foi estendido (experimento 2.2), gerando-se atributos derivados a partir dos 4 atributos originais (medida A, medida B, medida C e medida D). É importante salientar que estes atributos derivados foram gerados heurísticamente sem supervisão de especialistas no domínio. O objetivo da inclusão de atributos derivados é que alguns dos algoritmos utilizados não têm a capacidade de compor relacionamentos ou combinações entre os mesmos. Assim, sua inclusão poderia resultar em melhor desempenho dos algoritmos [Lee et al. 2000]. Os atributos derivados gerados foram a razão e o produto entre os atributos originais: A/B, A/C, A/D, B/C, B/D, C/D, A*B, A*C, A*D, B*C, B*D, C*D e A*B*C*D. Os resultados obtidos encontram-se na Tabela 4; os maiores valores da taxa de acerto obtidos para cada classificador estão destacados na tabela. Em média, as árvores de decisão acertaram 48.05%, vizinhos mais próximos 49.82% e redes neurais 52.78% sendo que Naive Bayes obteve desempenho de 53.93%. Como antes, esses resultados indicam que nenhum paradigma de aprendizado analisado conseguiu obter uma boa separação do sexo dos indivíduos. Como pode ser observado na Tabela 4, o algoritmo de melhor *rank* médio é o MLP utilizando semente igual a 2. Aplicando o teste Friedman-Nemenyi com 95% de confiança constatou-se que MLP(2) foi melhor do que IBK(3), J48(3) e J48(5). Comparando-se as Tabelas 3 e 4 é possível observar a inclusão de atributos derivados teve uma tendência a melhorar os resultados obtidos, exceto para IBK(3) e MLP(0).

Tabela 3. Experimento 2.1 Taxa de acerto e desvio padrão dos classificadores, utilizando os 4 atributos originais propostos por [Ribeiro 1993]

	J_{48} (2)	J_{48} (3)	J_{48} (5)	IBK (3)	IBK (5)	IBK (7)	MLP (0)	MLP (1)	MLP (2)	NB
T1	45.71 (4.01)	45.71 (4.01)	48.21 (4.97)	39.82 (4.17)	40.18 (3.70)	45.36 (4.26)	55.18 (4.39)	56.43 (5.65)	52.32 (5.55)	49.64 (4.65)
T2	50.00 (2.13)	50.00 (2.13)	50.00 (4.76)	53.93 (7.79)	45.71 (6.84)	42.68 (7.40)	41.61 (6.62)	38.57 (5.95)	38.75 (7.46)	46.25 (5.46)
T3	45.71 (1.58)	45.71 (1.58)	45.71 (1.58)	52.68 (6.80)	41.61 (6.26)	44.46 (7.86)	51.25 (6.69)	48.39 (8.48)	48.57 (6.64)	41.61 (6.82)
T4	58.04 (4.43)	58.04 (4.43)	58.04 (4.43)	54.29 (6.60)	51.25 (5.77)	48.39 (7.01)	46.79 (5.75)	46.96 (2.29)	43.75 (5.20)	58.04 (5.16)
T5	44.29 (0.95)	44.29 (0.95)	44.29 (0.95)	38.39 (8.69)	33.04 (5.36)	36.07 (4.77)	53.93 (4.46)	45.18 (6.15)	52.50 (4.45)	44.46 (5.29)
T6	44.29 (0.95)	44.29 (0.95)	44.29 (0.95)	48.39 (7.03)	44.29 (3.67)	35.89 (3.99)	52.68 (7.29)	47.14 (5.46)	44.11 (7.98)	52.86 (3.56)
T7	54.11 (4.24)	54.11 (4.24)	52.86 (4.55)	50.00 (6.99)	41.61 (4.07)	45.89 (4.24)	58.39 (4.07)	56.07 (7.26)	52.68 (7.62)	54.64 (5.32)
T8	50.00 (3.69)	50.00 (3.69)	50.00 (3.69)	48.39 (6.14)	52.32 (5.12)	48.57 (3.66)	54.11 (6.01)	58.39 (6.26)	65.18 (4.89)	60.00 (5.45)
T9	38.93 (2.63)	41.79 (2.92)	37.50 (4.62)	41.43 (6.48)	47.32 (8.45)	47.14 (7.84)	61.07 (6.22)	52.86 (7.02)	51.43 (6.74)	55.54 (4.06)
T10	44.29 (2.33)	44.29 (2.33)	44.29 (2.33)	55.36 (5.56)	52.50 (4.57)	46.79 (7.53)	59.82 (6.03)	58.39 (6.62)	58.57 (4.36)	55.71 (5.62)
Média	47.54	47.82	47.52	48.27	44.98	44.12	53.48	50.84	50.79	51.87
Desvio Padrão	1.33	1.33	1.65	1.22	1.51	1.80	1.09	1.61	1.36	0.89
Rank Médio	6.1	5.9	5.8	5.9	7.2	8.0	3.1	4.5	4.8	3.7

Tabela 4. Experimento 2.2 Taxa de acerto e desvio padrão dos classificadores testados, utilizando os atributos originais e derivados

	J48(2)	J48(3)	J48(5)	IBK(3)	IBK(5)	IBK(7)	MLP(0)	MLP(1)	MLP(2)	NB
T1	49.64 (5.01)	49.64 (5.01)	49.64 (5.01)	41.25 (5.65)	51.07 (5.53)	51.25 (5.77)	42.86 (3.38)	48.57 (3.51)	55.54 (2.72)	48.04 (6.15)
T2	45.89 (4.24)	45.89 (4.24)	45.89 (4.24)	41.43 (7.96)	51.25 (7.02)	51.43 (7.36)	42.86 (7.53)	48.57 (6.97)	47.14 (5.85)	53.04 (4.61)
T3	44.46 (4.06)	44.46 (4.06)	45.89 (2.99)	52.14 (6.97)	47.14 (5.46)	49.82 (3.91)	49.46 (7.65)	54.11 (4.24)	50.18 (6.36)	44.64 (4.29)
T4	55.36 (4.68)	55.36 (4.68)	55.36 (4.68)	49.82 (6.52)	51.43 (7.04)	51.43 (5.61)	52.68 (7.89)	51.43 (4.62)	55.36 (5.56)	56.96 (4.11)
T5	44.29 (0.95)	44.29 (0.95)	44.29 (0.95)	41.25 (7.69)	45.89 (7.22)	47.14 (5.13)	47.68 (6.66)	50.00 (5.66)	47.50 (5.38)	46.07 (4.34)
T6	40.54 (2.79)	40.54 (2.79)	43.21 (1.70)	45.89 (3.37)	44.46 (4.58)	48.21 (5.90)	53.04 (6.44)	48.04 (8.63)	46.07 (6.78)	50.89 (7.20)
T7	51.96 (6.15)	51.96 (6.15)	51.96 (6.15)	48.39 (7.47)	52.86 (6.23)	45.89 (6.01)	58.04 (6.84)	56.96 (6.80)	57.14 (5.95)	61.61 (6.41)
T8	48.39 (4.91)	48.39 (4.91)	45.71 (4.01)	56.96 (4.36)	55.18 (5.73)	63.93 (6.03)	59.82 (5.22)	68.21 (5.48)	71.07 (4.20)	62.68 (3.33)
T9	47.32 (4.86)	47.32 (4.86)	45.89 (5.20)	41.25 (5.23)	55.89 (8.80)	53.04 (7.28)	47.68 (5.12)	53.04 (6.08)	49.11 (7.34)	59.82 (7.84)
T10	54.11 (3.67)	52.68 (3.81)	51.25 (4.92)	55.54 (3.45)	55.18 (6.48)	48.21 (4.97)	54.11 (2.99)	58.57 (8.02)	58.75 (6.21)	55.54 (5.70)
Média	48.19	48.05	47.91	47.39	51.03	51.03	50.82	53.75	53.78	53.93
Desvio Padrão	1.43	1.42	1.63	1.72	1.18	1.02	1.74	1.64	1.32	1.48
Rank Médio	7.25	7.40	7.25	7.65	5.20	4.95	5.15	3.65	3.05	3.45

Na Figura 5 são ilustrados *boxplots* dos Experimentos 1.1, 2.1 e 2.2. No Experimento 1.1 todos os classificadores apresentaram precisão maior ou igual do que 50%, sendo que IBK(3) apresentou o melhor resultado. No Experimento 2.1 a precisão da maioria dos classificadores ficou abaixo de 50%. Por último, no Experimento 2.2, a maioria ficou acima de 50%, porém um pouco pior do que os resultados do Experimento 1.1.

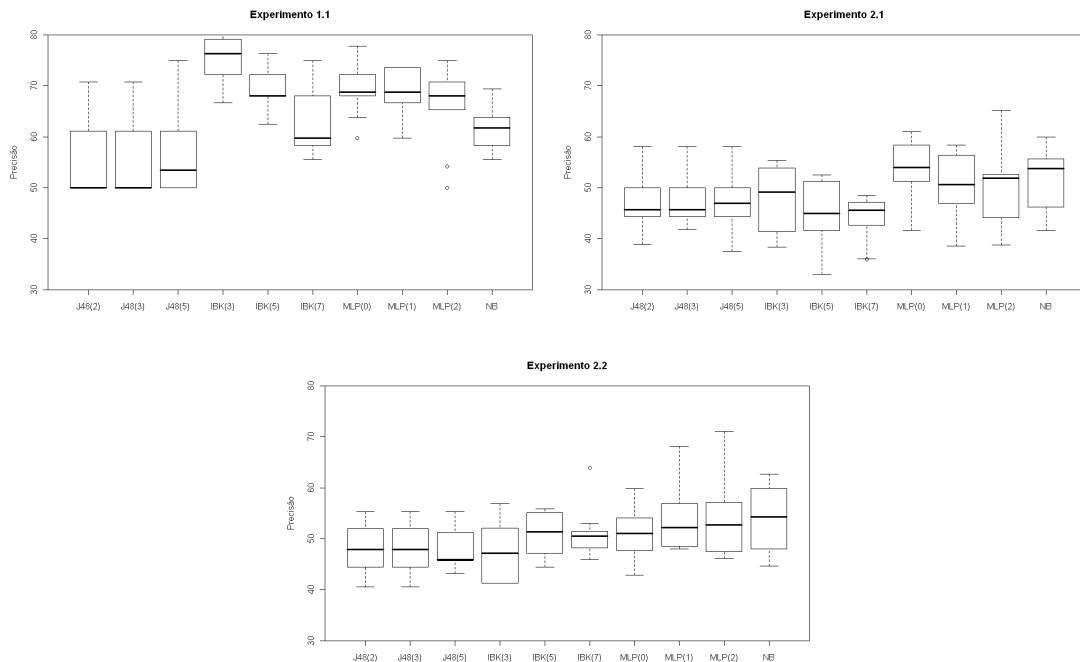


Figura 5. Boxplots dos Experimentos 1.1 (superior esquerda), 2.1 (superior direita) e 2.2 (inferior centro)

4. Conclusões

Este estudo permite concluir que as medidas propostas por [Ribeiro 1993] são voltadas para especialização: elas podem ser utilizadas para identificar o sexo de um indivíduo desde que haja uma radiografia anterior para comparação. Assim, se houver radiografia *antemortem*, pode-se comparar essas medidas com as medidas referentes à nova radiografia e identificar o indivíduo. Nesta situação, o algoritmo de vizinhos mais próximos aparenta ser o mais indicado para este problema. Por outro lado, é altamente provável que somente com essas 4 medidas não seja possível classificar o sexo de um indivíduo desconhecido, já que os classificadores de paradigmas diferentes utilizados neste estudo não obtiveram uma boa taxa de acerto de generalização. Visando uma possível melhora nos resultados obtidos foram definidos, adicionalmente, atributos derivados a partir dos 4 atributos originais propostos por [Ribeiro 1993]. Estes atributos derivados foram gerados heurísticamente; trabalhos futuros podem utilizar medidas derivadas reportadas na literatura ou mesmo sob a supervisão de especialistas no domínio, o que poderia melhorar o desempenho aqui obtido. Por exemplo, em [Camargo 2000] foram analisadas medidas distintas daquelas propostas por [Ribeiro 1993]: as áreas direita e esquerda dos seios frontais, a maior altura e maior largura de ambos os lados com o objetivo de tentar estimar o sexo por meio destas medidas. Os resultados obtidos apontaram para a variável *área*

como a variável mais adequada para a diferenciação do sexo. Assim, trabalhos futuros podem testar algoritmos de Aprendizado de Máquina utilizando como atributos as áreas dos seios frontais para conseguir diferenciar os sexos.

Referências

- [Batista et al. 2004] Batista, G. E. A. P. A., Prati, R. C., and Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *SIGKDD Explorations*, 6(1):20–29.
- [Bishop 2006] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*.
- [Camargo 2000] Camargo, J. R. (2000). Estimativa do sexo através das características radiográficas dos seios frontais. Technical report, Universidade Estadual de Campinas, Piracicaba.
- [Demsar 2006] Demsar, J. (2006). Statistical comparison of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7:1–30.
- [Domingos 1999] Domingos, P. (1999). The role of occam’s razor in knowledge discovery. *Data Mining and Knowledge Discovery*, 3:409–25.
- [Hastie et al. 2001] Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The elements of statistical learning, data mining, inference and prediction*. Springer, Berlin.
- [Kohavi 1995] Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *IJCAI*, pages 1137–1145.
- [Lee et al. 2000] Lee, H. D., Monard, M. C., and Baranauskas, J. A. (2000). A practical approach for knowledge-driven constructive induction. *Argentine Symposium on Artificial Intelligence*, pages 71–85.
- [Mitchell 1997] Mitchell, T. M. (1997). *Machine Learning*. McGraw–Hill.
- [Monard and Baranauskas 2003] Monard, M. C. and Baranauskas, J. A. (2003). *Conceitos sobre Aprendizado de Máquina*, chapter 4, pages 89–114. In [Rezende 2003].
- [Nassar and Ammar 2007] Nassar, D. E. M. and Ammar, H. H. (2007). A neural network system for matching dental radiographs. *Pattern Recognition*, 40(1):65–79.
- [Oliveira et al. 1998] Oliveira, R. N., Daruge, E., Galvão, L. C. C., and Tumang, A. J. (1998). Contribuição da odontologia legal para a identificação post-mortem. 55(2):117–122.
- [Rezende 2003] Rezende, S. O., editor (2003). *Sistemas Inteligentes*. Manole.
- [Ribeiro 1993] Ribeiro, F. A. Q. (1993). Um método de padronização de medidas feitas em radiografias dos seios frontais para ser utilizado na identificação pessoal. Technical report, Universidade Federal de São Paulo, UNIFESP, São Paulo.
- [Sweet 2001] Sweet, D. (2001). Why a dentist for identification? *Dent Clin North Am*, 45(2):237–51.
- [Vanrell 2002] Vanrell, J. P. (2002). *Odontologia legal e antropologia forense*. Guanabara Koogan, Rio de Janeiro, 1 edition.
- [Witten and Frank 2005] Witten, I. H. and Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition*. Morgan Kaufmann.