
UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE ESTATÍSTICA

Anny Kerollayny Gomes Rodrigues

COMPARAÇÃO ENTRE REDES BAYESIANAS DO TIPO *NAIVE*
BAYES BASEADAS NA DISTRIBUIÇÃO EXPONENCIAL

João Pessoa, 07 de Junho de 2017

Anny Kerollayny Gomes Rodrigues

COMPARAÇÃO ENTRE REDES BAYESIANAS DO TIPO *NAIVE BAYES*
BASEADAS NA DISTRIBUIÇÃO EXPONENCIAL

Monografia apresentada ao Curso de Bacharelado em Estatística da Universidade Federal da Paraíba, como requisito parcial para obtenção do Grau de Bacharel. Área de Concentração: Estatística Aplicada.

Orientador: Prof. Dr. RONEI MARCOS DE MORAES.

João Pessoa, 07 de Junho de 2017

Catálogo na publicação
Biblioteca Setorial do CCEN/UFPB
Josélia M.O. Silva – CRB-15/113

R696c Rodrigues, Anny Kerollayny Gomes.
Comparação entre redes bayesianas do tipo Naive Bayes baseadas na
distribuição exponencial / Anny Kerollayny Gomes Rodrigues. – João
Pessoa, 2017.
50 p. : il. color.

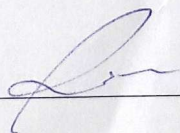
Monografia (Bacharelado em Estatística) – Universidade Federal da
Paraíba.

Orientador(a): Prof^o. Dr^o. Ronei Marcos de Moraes.

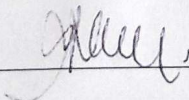
1. Estatística - Computação. 2. Avaliação de treinamento. 3. Redes
Naive Bayes. 4. Distribuição exponencial. I. Título.

Aos sete dias do mês de junho de dois mil e dezessete, às quatorze horas, na Sala 20 do Centro de Ciências Exatas e da Natureza da Universidade Federal da Paraíba, reuniram-se os membros da Banca Examinadora constituída para avaliar o Trabalho de Conclusão Curso intitulado "COMPARAÇÃO ENTRE REDES BAYESIANAS DO TIPO NAIVE BAYES BASEADAS NA DISTRIBUIÇÃO EXPONENCIAL" de autoria de ANNY KEROLLAYNY GOMES. A Banca Examinadora foi composta pelos professores: Prof. Dr. Ronci Marcos de Moraes (DE-UFPB, orientador), Profa. Ma. Ana Cláudia Oliveira de Melo (DE-UFPB, examinadora), Profa. Dra. Izabel Cristina Alcântara de Souza (DE-UFPB, examinadora) e Prof. Dr. Marcelo Rodrigo Portela Ferreira (DE-UFPB, examinador). Dando início aos trabalhos, o presidente da banca cumprimentou os presentes, comunicou aos mesmos a finalidade da reunião e passou à palavra à discente para que se fizesse, oralmente, a exposição de seu trabalho de monografia. Concluída a apresentação, a discente foi arguida pela Banca Examinadora que sugeriu algumas alterações até o dia 12 de junho de 2017, de acordo com a Resolução No. 02/2014 do Colegiado do Curso de Bacharelado em Estatística da UFPB. Uma vez entregue a versão final do Trabalho de Conclusão de Curso à Coordenação do Bacharelado em Estatística, com as alterações solicitadas pela Banca Examinadora dentro do prazo estabelecido, a discente será aprovada com nota (9,5) Nove e meio, que é a média aritmética das notas atribuídas pelos membros da Banca Examinadora.

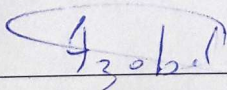
João Pessoa, 07 de junho de 2017.



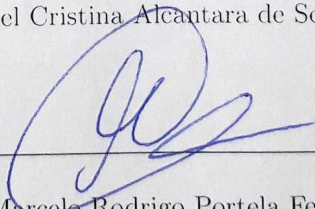
Prof. Ronci Marcos de Moraes (Orientador)



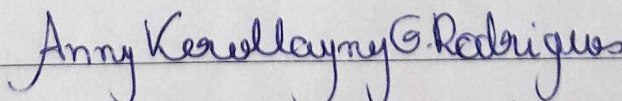
Prof. Ana Cláudia de Oliveira Melo (Examinadora)



Profa. Izabel Cristina Alcântara de Souza (Examinadora)



Prof. Dr. Marcelo Rodrigo Portela Ferreira (Examinador)



Anny Kerollayny Gomes (Discente)

Agradecimentos

Agradeço primeiramente aos meus pais que sempre incentivaram meus estudos, e sempre me apoiaram nas horas difíceis. As minhas irmãs Anna e Maria que sempre me trouxeram alegria nos momentos de estresse.

Agradeço ao meu namorado, Reinaldo, que sempre esteve ao meu lado durante os quatro anos de graduação, sempre cuidando e acreditando em mim, mesmo quando eu não acreditava, mergulhando na estatística comigo.

Agradeço a minha amiga, Tainara que sempre me deu os melhores conselhos sobre a vida, você sempre será minha irmã de coração. Ao meu velho amigo, Adriano, você é muito especial pra mim, obrigada por todo suporte que você sempre me deu e pelas risadas escancaradas. A minha amiga Nayara, obrigada pelos conselhos, incentivos e boas conversas, você é demais.

Agradeço aos meus amigos e colegas de sala, obrigada por tudo, vocês são demais. A Adenice pela amizade e cumplicidade e por me mostrar sempre o lado bom das coisas, obrigada. Agradeço a André por ter sido um exemplo a ser seguido durante os quatro anos de graduação, com a sua organização e companheirismo sempre, obrigada. A Clarissa pela amizade e por sempre apresentar as melhores soluções para os problemas. Agradeço a Diogo pelas palavras sábias e por ser aquela pessoa que podemos sempre contar. A José pelas ótimas risadas e por ter tornado nossos dias na graduação mais alegres.

Aos colegas de curso de Ligia, Diego, Danilo e Lucas obrigada por sempre estarem próximos.

Aos meus amigos Felipe, Tássio e Nelson, pelas alegrias e companheirismo. A Jodavid pelo incentivo e ajuda durante a graduação, obrigada.

Aos colegas do LABPIG, Thiago, Elaine, Ives, Danielly, Laisa e Luciana pela ajuda, amizade e paciência comigo sou muito grata a vocês.

Agradeço ao Professor Ronei, por seus ensinamentos, paciência e confiança nesses

três anos sob sua orientação, aprendi muito com o senhor.

À professora Ana Flávia pela amizade, carinho, cuidado e ensinamentos, obrigada pelo acolhimento, a senhora é demais.

Aos professores do DE pelo incentivo e apoio nas disciplinas, em especial a Tarciana, Tatiene, Luiz, Marcelo, João Agnaldo, Rodrigo, Hemílio e Izabel.

À banca examinadora pelas correções e incentivos.

Ao CNPQ pelo apoio financeiro durante meus anos de graduação.

*”... Pedras no caminho?
Guardo todas um dia vou construir um castelo ... ”
Fernando Pessoa*

Resumo

Este trabalho compara e avalia a precisão de quatro avaliadores do tipo Naive Bayes baseados na distribuição Exponencial. Os resultados obtidos foram analisados de acordo com o Acerto Percentual, o Coeficiente Kappa, sua variância e o Tempo Computacional de cada método. Para analisar a performance dos métodos foram geradas dois conjuntos de amostras denominadas Amostras para Treinamento e Teste, que possuem dados de 1 a 4 dimensões com quatro classes de desempenho distintas. Para a amostra de treinamento foi gerada 16 amostras com 10000 observações cada e para amostra de teste foram geradas 16 amostras cada uma com 30000 observações. Como resultados obteve-se que o Método de Avaliação de treinamento na Rede *Fuzzy Exponencial Naive Bayes* (MATRENB) foi significativamente melhor que os demais. Em relação ao tempo computacional o MATRENB conseguiu avaliar dados que seguem uma Distribuição Exponencial, com quatro dimensões, com um baixo custo computacional. Vale ressaltar, que a partir de três ou mais dimensões a maioria dos avaliadores conseguiram obter um percentual de acertos, com um grau de concordância moderado, de acordo com o coeficiente Kappa.

Palavras-Chaves: Avaliação de Treinamento, Redes *Naive Bayes*, Distribuição Exponencial.

Abstract

This paper compares and evaluates the precision of four evaluators of Naive Bayes based on the Exponential Distribution. The acquired results were analysed according to the percentage of correct classification, the coefficient Kappa, its variance and the time of CPU consumed for each method. To analyse the performance of the methods, two groups of samples called training and testing samples, each having data from 1 to 4 dimensions with four different classes of performance. For the training samples with 10.000 observations were created for each and for the test samples were generated 16 samples with 30.000 observations. As result, the Method of Assessment and Training in the Fuzzy Exponential Naive Bayes Network was significantly better than the others. In relation to the time of CPU consumed, MATRENB managed to evaluate data that followed the Exponential Distribution, with four dimensions, therefore is a low computational cost application. It is worth to mention that from three or more dimensions most of the evaluators managed to acquire a percentual hits, with a moderate degree of substancial agreement, according to the Kappa coefficient.

Keywords: Training Assessment, Naive bayes Network, Exponential Distribution.

Sumário

Agradecimentos	4
Resumo	7
Abstract	8
Lista de Figuras	11
Lista de Tabelas	12
1 Introdução	1
2 Objetivos	4
3 Referencial teórico	5
3.1 Sistemas de Treinamento em Realidade Virtual	5
3.2 Avaliação de Treinamento	6
3.3 Avaliação <i>on-line</i> e <i>off-line</i>	8
3.4 Conjunto <i>Fuzzy</i> e Cortes- α	11
3.5 Números <i>Fuzzy</i> e Aritmética <i>Fuzzy</i>	12
3.6 Probabilidade Clássica	14
3.7 Probabilidade <i>Fuzzy</i>	14
3.8 Probabilidade Buckley	15
3.9 Distribuição Exponencial com Parâmetros <i>Fuzzy</i>	15
3.10 Método de Avaliação de Treinamento baseado na Rede <i>Naive Bayes</i>	16
3.11 Método de avaliação de Treinamento Baseado na Rede Exponencial <i>Naive Bayes</i>	17

	10
3.12 Método de Avaliação de Treinamento baseado na Rede <i>Fuzzy</i> Exponencial <i>Naive Bayes</i>	18
3.13 Método de Avaliação de Treinamento Baseado na Rede <i>Naive Bayes</i> Ex- ponencial com Parâmetros <i>Fuzzy</i>	19
3.14 Coeficientes de Concordância	20
3.14.1 Acerto Percentual	20
3.14.2 Coeficiente Kappa	21
4 Materiais	23
4.1 Amostras Para Treinamento e Teste	23
4.1.1 Simulação	24
5 Resultados e Discussão	25
5.1 Método de Avaliação de Treinamento baseado na Rede <i>Naive Bayes</i>	26
5.2 Método de avaliação de Treinamento Baseado na Rede Exponencial <i>Naive</i> <i>Bayes</i>	27
5.3 Método de Avaliação de Treinamento baseado na Rede <i>Fuzzy</i> Exponencial <i>Naive Bayes</i>	28
5.4 Método de Avaliação de Treinamento Baseado na Rede <i>Naive Bayes</i> Ex- ponencial com Parâmetros <i>Fuzzy</i>	29
5.5 Comparação entre os avaliadores	30
6 Conclusão	34
Referências Bibliográficas	35

Lista de Figuras

3.1	Arquitetura genérica de um sistema de avaliação de treinamento baseado em realidade virtual (SANTOS, 2010).	7
3.2	Diagrama de casos de uso do Siteg (SANTOS, 2010).	11
3.3	Números <i>Fuzzy</i> Triângular \overline{M}	13
5.1	Amostra aleatória da distribuição Exponencial (linhas são as dimensões e colunas são as classes de desempenho)	26
5.2	Comparação entre os avaliadores de acordo com o Acerto Percentual por dimensão.	33
5.3	Comparação entre os métodos avaliação de acordo com o Coeficiente Kappa e o Tempo Computacional para quatro dimensões.	33

Lista de Tabelas

3.1	Matriz de Confusão.	21
3.2	Interpretação do coeficiente Kappa.	22
4.1	Parâmetros da distribuição Exponencial.	24
5.1	Matriz de Confusão com quatro dimensões para o MATRNB.	27
5.2	Matriz de Confusão com quatro dimensões para o MATRENB.	28
5.3	Matriz de Confusão com quatro dimensões para o MATRFENB.	29
5.4	Matriz de Confusão com quatro dimensões para o MATRNBEPF.	30
5.5	Comparação entre os avaliadores para a primeira dimensão.	32
5.6	Comparação entre os avaliadores para primeira e segunda dimensões.	32
5.7	Comparação entre os avaliadores para primeira, segunda e terceira dimensões.	32
5.8	Comparação entre os avaliadores para todas as quatro dimensões.	32
5.9	Tempo Computacional por métodos e dimensão, em segundos.	32

Capítulo 1

Introdução

O advento da Realidade Virtual (RV) trouxe novas possibilidades de aplicações para diversos campos da ciência além da computação, como na área da Estatística, Física, Ciências Cognitivas, Matemática, Design, entre outras (MORAES; MACHADO, 2010b). Em 1991, Rheingold (RHEINGOLD, 1991) publicou um dos primeiros trabalhos abordando essa nova tecnologia revolucionária chamada de RV. Essa nova tecnologia se refere a ambientes artificiais criados por computadores e dispositivos eletrônicos, que incluem recursos imersivos e interativos. Esses ambientes permitem aos usuários fazer parte das simulações virtuais, com seus sentidos explorados por meio de dispositivos de interação (RHEINGOLD, 1991). Um foco importante da RV é sua utilização para fins de treinamento para tarefas que envolvem riscos na sua execução real, como por exemplo na área da aviação comercial (MORAES; MACHADO, 2014), cirurgias médicas (HARDERS, 2008), indústria petrolífera (BURDEA; COIFFET, 2003) entre outras. Nesses sistemas, o usuário está imerso em um mundo virtual para realizar treinamento realista com interações realistas.

No entanto é importante avaliar o treinamento dos usuários e conhecer suas habilidades, visando corrigir seus erros e aprimorar seus reflexos, condições psicológicas e destreza, para a realização da tarefa real (MORAES; MACHADO, 2009a). Esse monitoramento das ações do usuário é suportado por algum tipo de método de avaliação. Basicamente, existem dois métodos de avaliação a forma *off-line* e a *on-line*. A forma *off-line* é muito utilizada na área médica, em que a avaliação é feita por um ou mais especialistas e se baseiam na análise de gravações em vídeo (MACHADO *et al.*, 2000).

Uma desvantagem desse método é que o tempo de resposta da avaliação não é

imediatamente e como consequência o usuário pode esquecer os detalhes do procedimento. No entanto a avaliação *on-line* possui um tempo de resposta imediato inferior de 1 segundo, por observação, que é o tempo máximo permitido para um bom avaliador *on-line*. Logo, como vantagem após o término da simulação o usuário pode identificar seus erros e corrigi-los no próximo treinamento (MORAES; MACHADO, 2010b).

A avaliação de treinamento *on-line* vem sendo muito utilizada no campo da ciência, entre elas técnicas envolvendo avaliadores estatísticos utilizando redes probabilistas, redes possibilistas, Lógica *Fuzzy*, agregados a simuladores baseados em RV (MORAES; MACHADO, 2010b; MORAES; MACHADO, 2012; MORAES; MACHADO, 2004; MORAES; MACHADO, 2009a).

Em geral, os avaliadores estatísticos possuem desempenho diretamente relacionado às distribuições estatísticas para as quais eles foram delineados (MORAES; MACHADO, 2016). Outros métodos foram delineados sem levar em consideração a distribuição estatística dos dados, como por exemplo as redes neurais (BISHOP, 2007). Alguns desses métodos podem obter bons resultados para diferentes distribuições, mas podem ser inadequados para outras (MORAES, 2013) (FERREIRA *et al.*, 2014).

Na literatura, diversos trabalhos foram propostos usando diferentes aplicações para a distribuição Exponencial: como em Controle Estatístico de Qualidade (CAETANO; LOUZADA-NETO, 2007), na área de Meteorologia (DUAN *et al.*, 1998) (CATALUNHA *et al.*, 2002), na área de classificação (RODRIGUES *et al.*, 2016) (MORAES; MACHADO, 2016) entre outros.

Dados provenientes da distribuição Exponencial são tradicionalmente difíceis de avaliar. O forte formato assimétrico dessa distribuição propicia grandes intersecções entre as diversas classes, causando esse problema (MORAES; MACHADO, 2016) (RODRIGUES *et al.*, 2016). Mesmo avaliadores que possuem bons resultados para dados provenientes de distribuições da família exponencial, como o *Fuzzy Gaussian Naive Bayes* (FGNB) (FERREIRA *et al.*, 2014) não foram capazes de proporcionar bons resultados para essa distribuição. Durante as interações do usuário com o simulador diversas variáveis são monitoradas ao mesmo tempo. Logo, é pertinente que o avaliador possa mensurar a contagem do tempo de cada interação, através de um avaliador com base na distribuição Exponencial.

Neste trabalho foi estudado e implementado quatro métodos de avaliação para

treinamento *on-line* para Redes Bayesianas, baseadas na distribuição Exponencial. Os resultados da avaliação foram analisados e comparados no que diz respeito a precisão e performance dos avaliadores e segundo aos coeficientes de concordância: Acerto Percentual, Kappa e sua variância e o Tempo Computacional.

Este trabalho foi estruturado da seguinte maneira: no Capítulo 2, estão apresentados o Objetivo Geral e os Objetivos Específicos desse trabalho. O Capítulo 3 apresenta o Referencial Teórico, que aborda temas como Realidade Virtual, Avaliação de Treinamento, Métodos de Avaliação e Coeficientes de Concordância. O Capítulo 4 apresenta os materiais utilizados para se atingir o propósito dessa pesquisa, como geração das amostras para treinamento e teste e a simulação. No Capítulo 5 estão apresentados os Resultados e Discussões. Por fim, o Capítulo 6 apresenta as conclusões.

Capítulo 2

Objetivos

Objetivo Geral:

- Comparar diferentes avaliadores baseado na Rede *Naive Bayes*, avaliar seus desempenhos em classificar dados que seguem uma distribuição Exponencial.

Objetivos Específicos:

- Realizar uma implementação computacional dos Métodos de Avaliação de Treinamento para a Rede *Naive Bayes*, Rede Exponencial *Naive Bayes*, Rede *Fuzzy Exponencial Naive Bayes* e Rede *Naive Bayes Exponencial* com Parâmetros *Fuzzy*.
- Mensurar o desempenho dos avaliadores através de medidas estatísticas.
- Analisar o desempenho dos avaliadores quanto ao tempo computacional.

Capítulo 3

Referencial teórico

3.1 Sistemas de Treinamento em Realidade Virtual

A Realidade Virtual (RV) refere-se a sistemas em tempo real modelados por computação gráfica que permitem a interação do usuário com o ambiente virtual (BURDEA; COIFFET, 2003). Pesquisadores e desenvolvedores de *software* sempre tentam defini-la com base nas suas próprias experiências, gerando diversas definições na literatura. O termo RV pode ser definido como uma ciência que une o conhecimento de diversas áreas, como informática, eletrônica, física e cognição, ao lado de outros, a fim de oferecer sistemas adicionais que integram características de imersão e de interatividade para simular ambientes reais (MORAES; MACHADO, 2010b).

Na prática, a RV permite que o usuário observe e navegue em um mundo tridimensional (3D) em tempo real com seis ou mais graus de liberdade (NETTO *et al.*, 2002). Para isso, exige-se a capacidade do *software* de definir e a do *hardware* de reconhecer movimentos como para frente/para trás, acima/abaixo, esquerda/direita, inclinação para cima/para baixo, angulação à esquerda/à direita e rotação à esquerda/à direita. Para suportar esse tipo de interação do usuário com o mundo virtual, pode-se utilizar dispositivos não convencionais, como luvas de dados chamadas *datagloves*, capacetes de visualização e controle. Entretanto, esses dispositivos possuem suas limitações, como por exemplo: o capacete é restrito apenas a visualização e algumas luvas de dados só é possível sentir o objeto mas não é possível exercer nenhuma força sobre ele (MORAES; MACHADO, 2010b).

Além destes, existem os chamados dispositivos hápticos, cuja função é estimular

sensações com o tato, como a tensão muscular e a temperatura, esses dispositivos são úteis em simulações que não exigem informação visual. Um exemplo de um dispositivo háptico é o *Phantom Omni*, onde sua tecnologia fornece ao usuário uma realimentação física, tornando-se possível sentir, tocar e exercer força sobre um objeto, uma característica que os dispositivos não convencionais oferecem (MORAES; MACHADO, 2012).

O uso desse dispositivo proporciona ao usuário a impressão de que a aplicação está funcionando em um mundo 3D real. Esse dispositivo pode ser utilizado para avaliações de treinamento baseado em RV (MORAES; MACHADO, 2012).

Uma vantagem de se utilizar RV para treinamento é a possibilidade de monitorar as ações do usuário e a capacidade do sistema armazenar os dados de entrada e saída (MORAES; MACHADO, 2010a). Além disso, os dados coletados no treinamento em RV podem ser utilizados para avaliar o desempenho do usuário e permitir que eles identifiquem seus erros e melhorem suas habilidades, não havendo desgaste do material e o treinamento poderá ser realizado quantas vezes forem necessárias (MORAES; MACHADO, 2009a).

3.2 Avaliação de Treinamento

Os primeiros trabalhos na área de avaliação de treinamento foram propostos por Dinsmore *et al.* (1997) que utilizou um questionário para avaliar os usuários de um ambiente em RV para identificar tumores subcutâneos, esse questionário continha questões relacionadas ao diagnóstico do tumor. Paralelamente, vários grupos de pesquisadores desenvolveram métodos para avaliar as habilidades cirúrgicas. Posteriormente, várias metodologias foram propostas para efetuar essa avaliação (MACHADO *et al.*, 2000; WEBSTER *et al.*, 2001; MORAES; MACHADO, 2004; WEISS *et al.*, 2004; BASDOGAN *et al.*, 2004).

A avaliação de treinamento deve monitorar os movimentos e ações do usuário sobre o ambiente virtual e fornecer um *feedback* sobre seu desempenho na simulação (MORAES; MACHADO, 2012). Estes se constituem em parâmetros que possibilitam saber se o usuário realizou o treinamento com o desempenho necessário para se executar a tarefa real. O sistema precisa coletar informações sobre a posição espacial relativa do usuário como: forças, torque, resistências, velocidades, acelerações, posição de visualização e etc. (MORAES; MACHADO, 2008). As interações entre o simulador e o usuário são monitoradas

e as informações são enviadas ao avaliador que as analisa e emite um relatório sobre o desempenho do usuário ao final do treinamento. A avaliação é realizada comparando os parâmetros definidos por um ou mais especialistas com os parâmetros obtidos durante o treinamento do usuário.

Um sistema de treinamento em RV é formado por 3 subsistemas: geração das imagens, interação usuário/máquina e a avaliação. Como qualquer subsistema de um simulador, a visualização, interação e avaliação devem funcionar num mesmo sistema computacional, a simulação deve sempre garantir o tempo real de execução (MORAES; MACHADO, 2009b). Assim, um subsistema de avaliação deverá fornecer uma medida de adequabilidade sobre o desempenho do usuário.

Entretanto, vários treinamentos não podem ser simplesmente classificados como bons ou ruins devido a complexidade que cada treinamento apresenta. A existência de uma ferramenta de avaliação de treinamento acoplada a um simulador virtual é de extrema importância para prover uma melhora no desempenho e na avaliação do usuário (SANTOS *et al.*, 2010). A Figura 3.1 apresenta uma arquitetura genérica de um sistema baseado em RV (SANTOS, 2010), vale ressaltar que o simulador de treinamento em RV e o seu subsistema de avaliação são interdependentes, porém agem de forma simultâneas. As interações entre o usuário e o simulador são monitoradas e as informações são enviadas ao seu subsistema de avaliação que as analisa e emite um relatório sobre o desempenho do usuário.

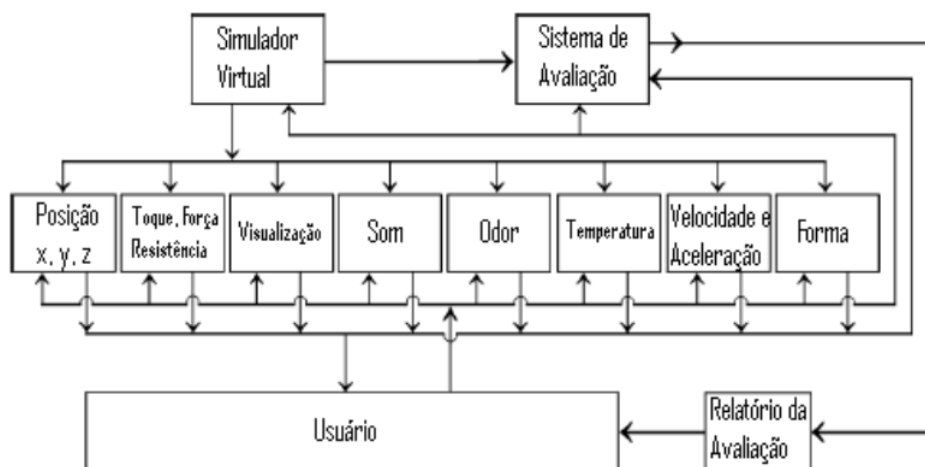


Figura 3.1: Arquitetura genérica de um sistema de avaliação de treinamento baseado em realidade virtual (SANTOS, 2010).

O subsistema de avaliação deve funcionar sem comprometer o desempenho do simulador e o seu grau de realismo, pois simuladores muito realistas tendem a demandar um alto custo computacional, (MORAES; MACHADO, 2009b). Na pesquisa de subsistemas de avaliação para treinamentos complexos, como cirurgias médicas (HARDERS, 2008), nos quais é necessário monitorar um grande número de variáveis simultaneamente é dependente de quatro fatores: a) complexidade computacional do ambiente virtual; b) complexidade computacional do método de avaliação; c) acurácia do método e d) o sistema computacional disponível para executar o ambiente para treinamento e o subsistema de avaliação (SANTOS *et al.*, 2010).

O método de avaliação de treinamento pode ser classificado em duas formas: a forma de *off-line* tradicional que usava as informações coletadas durante o procedimento para uma avaliação posterior podendo ser computadorizada ou não e a forma *on-line*, na qual a avaliação é realizada imediatamente ao término da simulação, com um tempo de resposta imperceptível para quem opera o sistema (MORAES; MACHADO, 2009b) (MORAES; MACHADO, 2012) (SANTOS *et al.*, 2010).

3.3 Avaliação *on-line* e *off-line*

A primeira proposta de avaliação de treinamento *on-line* em sistemas em RV foi apresentada por Machado *et al.* (2000), esse sistema era baseado em lógica booleana que comparava os diagnósticos fornecidos pelos usuários com os armazenados pelo sistema. Entretanto, os primeiros computadores do tipo *PC* não foram capazes de executar simultaneamente ambientes virtuais, simuladores e sistemas de avaliação *on-line* quando várias variáveis de interação foram monitoradas. A avaliação de treinamento é considerada *on-line* quando ela oferece ao usuário um tempo de resposta imediato como citado anteriormente, o usuário não corre o risco de perder informações sobre o treinamento. Uma das vantagens de se utilizar a avaliação de treinamento *on-line* é a de emissão de relatórios técnicos sobre o desempenho do usuário. Uma habilidade muito importante agregada ao simulador, pois esse tipo de avaliação permite que o utilizador aumente ou diminua seu grau de dificuldade de acordo com suas necessidades.

Por outro lado, em uma avaliação de treinamento *off-line* o tempo de resposta é perceptível pelo usuário (MORAES; MACHADO, 2012). Na área médica, a avaliação

de treinamento era realizada posteriormente por médicos especialistas onde se baseavam-se em análise feita em cima de gravações de vídeos. Essa avaliação caracteriza a forma *off-line*. Uma grande desvantagem desse método é o tempo de espera dos resultados da avaliação feita posteriormente podendo durar até semanas de acordo com o grau de dificuldade do treinamento. O usuário receberá um relatório sobre seu desempenho dias após a realização do treinamento e as lembranças dos detalhes podem ser perdidas.

Recentemente, várias metodologias foram apresentadas utilizando sistemas de avaliação *on-line* para procedimentos cirúrgicos utilizando: máxima Verossimilhança (MORAES; MACHADO, 2005), redes bayesianas (MORAES; MACHADO, 2013) (MORAES; MACHADO, 2007), redes possibilistas (MORAES; MACHADO, 2014) (MORAES; MACHADO, 2009b) e métodos estatísticos *Fuzzy* (FERREIRA *et al.*, 2014)(MORAES; MACHADO, 2012) (MORAES; MACHADO, 2004), entre outros. Machado e Moraes (2006) propuseram uma metodologia embasada em sistemas especialistas baseados em lógica *Fuzzy* para um sistema de avaliação *on-line* acoplado a um simulador de exame ginecológico.

O exame ginecológico é considerado um dos mais importantes para a saúde da mulher, através dele é possível identificar doenças como: câncer do colo do útero, Herpes e HPV (ROBBINS *et al.*, 2001). Esse exame é dividido em cinco etapas: a Anamnese, consiste em coletar os dados de identificação do paciente (idade, sexo, número de filhos), suas queixas e seu histórico médico; o Exame das Mama, nessa etapa é possível identificar a presença de nódulos; o Exame do Abdmômen, essa etapa consiste na inspeção (descrição do tecido gorduroso e presença de cicatrizes), palpação superficial (identificação de áreas doloridas) e a palpação profunda (identificação de tumores ginecológicos e verificação do baço e fígado). A quarta etapa é o Exame dos Órgãos Genitais Externos, onde é realizada a inspeção estática, dinâmica e palpação do ânus, o médico estuda a distribuição dos pêlos, pesquisa alterações anatômicas, averigua os músculos levantadores do ânus, entre outras características. A última etapa é a inspeção dos órgãos genitais internos do paciente, no qual é realizado o exame especular e tátil, observa-se as paredes vaginais, o conteúdo vaginal e as características do colo do útero (ROBBINS *et al.*, 2001).

O Sistema Interativo para Treinamento em Exame Ginecológico (Siteg), é um exemplo de um sistema para treinamento em RV, no qual permite identificar através da simulação doenças relacionadas ao colo útero (SANTOS, 2010). O desenvolvimento do

Siteg é oriundo das dificuldades que estão relacionadas a realização do exame ginecológico real, onde suas particularidades dificultam seu treinamento nas universidades. Como por exemplo, o constrangimento do paciente na realização do exame em hospitais universitários às vistas dos estudantes e a dificuldade do médico residente identificar por meios de fotografias diversos tipos de patologias sem a inspeção tátil no paciente.

O Siteg proporciona ao usuário a realização do exame ginecológico virtual onde riscos aos pacientes e problemas éticos associados a utilização de cobaias são desconsiderados, permitindo a emissão de um relatório sobre o desempenho do usuário ao final da simulação. Devido às restrições impostas por um nível aceitável de complexidade computacional, as atividades simuladas pelo Siteg deve fornecer situações suficientes para transmissão de conhecimento para o aluno-usuário.

O Siteg apresenta as seguintes etapas do exame ginecológico: Anamnese, Exame dos órgãos Genitais Externos (Inspeção Estática), Exame dos órgãos Genitais Internos (Exame Especular e Tátil). Devido ao grande número de patologias que podem ser identificadas pelo exame ginecológico, para a simulação escolheu-se as doenças Herpes e HPV, devido ao fato de suas ocorrências estarem ligadas ao câncer do colo do útero. No Siteg, essas patologias são encontradas em diferentes estágios e permitem o usuário conhecer e experimentar várias situações. A Figura 3.2 apresenta as interações que ocorrem no Siteg com o usuário.

O tempo do exame ginecológico nas diferentes fases de execução do treinamento são variáveis de avaliação que seguem uma distribuição Exponencial, como por exemplo: *tempo do exame visual*, *o tempo de realização do exame tátil*, *o tempo de coleta do material para o exame laboratorial* e *o tempo total de realização do exame ginecológico*. Essas variáveis podem ser associadas ao método de avaliação utilizado no simulador (SANTOS, 2010) e permitem auxiliar na diferenciação do usuário e seu nível de treinamento.



Figura 3.2: Diagrama de casos de uso do Siteg (SANTOS, 2010).

3.4 Conjunto *Fuzzy* e Cortes- α

Zadeh (1965) introduziu um novo conceito sobre conjuntos, onde seus elementos não constituem um conjunto no sentido matemático usual, o qual denominou de Conjuntos *Fuzzy*. Sua aplicação vem da incerteza sobre os elementos que o compõe e da necessidade de uma quantificação dessa incerteza que é dado por um “grau de pertinência”. Formalmente segue-se a definição (ZADEH, 1965).

Definição 1. Seja X um espaço de pontos em \mathbb{R}^n com $n \geq 1$, um elemento genérico de X denotado por x . Um conjunto *Fuzzy* de A em X é caracterizado por uma função de pertinência $\mu_A(x)$ que associa cada ponto em X a um número real no intervalo de $[0, 1]$, o qual representa para x o seu grau de pertinência em A (ZADEH, 1965).

É interessante escrever os elementos que pertencem a um conjunto *Fuzzy* com seu grau de pertinência, da mesma forma os números *Fuzzy* podem ser expressos usando-os, em ambos casos podemos utilizar cortes- α (ZADEH, 1965). Os cortes- α (ou α -cortes) são cortes feitos em números *Fuzzy* produzindo números regulares (não *Fuzzy*), seu uso facilita as operações matemáticas com números *Fuzzy*. Segue-se a definição.

Definição 2. Um conjunto *Fuzzy* A com função de pertinência de $\mu_A(x)$, pode

ser expresso pelo conjunto de seus cortes- α , denotados por A_α , como (ZADEH, 1965):

$$A_\alpha = \{x \in X | \mu_A(x) \geq \alpha\}. \quad (3.1)$$

A função de pertinência do conjunto *Fuzzy* A pode ser representada em termos de seus cortes $-\alpha$, de acordo com a equação (MORAES, 2015):

$$\mu_\alpha = \sup_{\alpha \in [0,1]} \min\{\alpha, \mu_{A_\alpha}(x)\} \quad (3.2)$$

em que $\mu_{A_\alpha} = 1$, se $x \in A$ ou $\mu_{A_\alpha} = 0$ para outros casos.

Devido à incerteza sobre os elementos que pertencem a um conjunto *Fuzzy*, os números *Fuzzy* também devem refletir essa incerteza sobre os elementos que os compõe. Na literatura existe vários formatos de números *Fuzzy* mas devido à manutenção de sua forma ao realizar operações matemáticas os números *Fuzzy* trapezoidais e triangulares são mais utilizados (BUCKLEY, 2005).

3.5 Números *Fuzzy* e Aritmética *Fuzzy*

Definição 3. Um número *Fuzzy* pode ser composto por uma forma geométrica convexa, onde a mais usual é a forma triangular. Um número triangular *Fuzzy* \bar{M} é representado por três números reais a, b, c de forma que $a \leq b \leq c$ onde $[a, c]$ é a base do triângulo e b é o vértice superior, cujo valor de pertinência é o máximo.

Assim, Buckley (2005) define um número *Fuzzy* em função dos seus cortes- α como:

$$\bar{M}_{[\alpha]} = [m_1(\alpha), m_2(\alpha)], \text{ com } m_1(\alpha) \leq m_2(\alpha). \quad (3.3)$$

para todo $\alpha, 0 \leq \alpha \leq 1$.

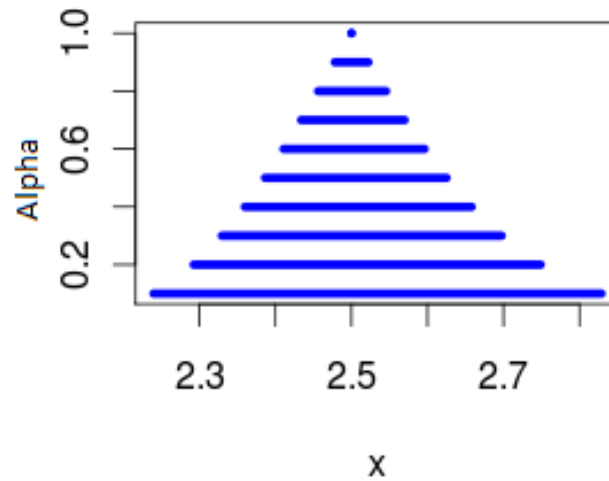


Figura 3.3: Números *Fuzzy* Triângular \bar{M} .

Assim como o conjunto de números reais o conjunto de números *Fuzzy* também possui operações aritméticas *Fuzzy* básicas (adição, subtração, multiplicação e divisão) que foram necessárias para a implementação desses novos avaliadores. Sejam dois números *Fuzzy* \bar{M} e \bar{N} e admita que seus cortes α são fechados e limitados nos intervalos $\bar{M}_{[\alpha]} = [m_1(\alpha), m_2(\alpha)]$, $\bar{N}_{[\alpha]} = [n_1(\alpha), n_2(\alpha)]$. Então, se $\bar{O} = \bar{M} + \bar{N}$ temos (BUCKLEY, 2005)

$$\begin{aligned}\bar{O}_{[\alpha]} &= \bar{M}_{[\alpha]} + \bar{N}_{[\alpha]} \\ &= [m_1(\alpha) + n_1(\alpha), m_2(\alpha) + n_2(\alpha)]\end{aligned}\tag{3.4}$$

Definindo $\bar{O} = \bar{M} - \bar{N}$ para todo $\alpha \in [0, 1]$

$$\begin{aligned}\bar{O}_{[\alpha]} &= \bar{M}_{[\alpha]} - \bar{N}_{[\alpha]} \\ &= [m_1(\alpha) - n_2(\alpha), m_2(\alpha) - n_1(\alpha)]\end{aligned}\tag{3.5}$$

para todo α no intervalo $[0, 1]$. No caso da multiplicação, ou seja, para $\bar{O} = \bar{M}\bar{N}$, temos

(BUCKLEY, 2005)

$$\bar{O}_{[\alpha]} = \bar{M}_{[\alpha]} \bar{N}_{[\alpha]} = [\delta, \beta], \quad (3.6)$$

em que

$$\delta = \min\{m_1 n_1, m_1 n_2, m_2 n_1, m_2 n_2\} \quad (3.7)$$

e

$$\beta = \max\{m_1 n_1, m_1 n_2, m_2 n_1, m_2 n_2\}. \quad (3.8)$$

para o *corte* $-\alpha$ definido. Quando $\bar{O} = \bar{M}/\bar{N}$ temos:

$$\bar{O}_{[\alpha]} = \bar{M}_{(\alpha)}/\bar{N}_{(\alpha)} = [m_1, m_2] * \left[\frac{1}{n_2}, \frac{1}{n_2}\right], \quad (3.9)$$

desde que zero não pertence ao $\bar{N}_{[\alpha]}$ para todo α (BUCKLEY, 2005).

3.6 Probabilidade Clássica

Definição 4. Seja o espaço de probabilidade definido por $(\mathbb{R}^n, P, \mathbf{B})$, onde \mathbf{B} é o conjunto de todos os subconjuntos possíveis do espaço amostral em \mathbb{R}^n ; e P é uma medida de probabilidade mensurada no \mathbb{R}^n . Seja A um evento em \mathbf{B} , a probabilidade de A é definida como (MORAES, 2015):

$$P(A) = \int_{A \subset \mathbb{R}^n} P(x) dx. \quad (3.10)$$

3.7 Probabilidade Fuzzy

Em 1968, Zadeh introduziu uma medida de probabilidade para eventos *Fuzzy* (ZADEH, 1968). **Definição 5.** Seja F em \mathbf{B} um evento *Fuzzy* com uma função de pertinência $\mu_F : \mathbb{R}^n \rightarrow [0, 1]$, a probabilidade de um evento *Fuzzy* F é definida pela integral de Lebesgue-Stieltjes:

$$\mathbf{P}(F) = \int_{\mathbb{R}^n} dP = \int_{\mathbb{R}^n} \mu_F(x) dP = E(\mu_F). \quad (3.11)$$

A probabilidade de um evento *Fuzzy* é a esperança matemática, também conhecido como o valor esperado da sua função de pertinência (ZADEH, 1968). A Equação (3.11) pode ser reescrita como:

$$\mathbf{P}(F) = \int_{R^n} \mu_F(x)P(x)dx. \quad (3.12)$$

3.8 Probabilidade Buckley

Buckley (2005) define probabilidade *Fuzzy* para um evento C em \mathbf{B} , em que os parâmetros são números *Fuzzy* triangulares ou trapezoidal. Esses parâmetros são utilizados para descrever a incerteza do verdadeiro valor de um parâmetro específico. Como produto dessa definição as distribuições de probabilidade podem ser generalizadas e seus parâmetros estimados por números *Fuzzy*.

Definição 6. Seja X em \mathbb{R}^n uma variável aleatória *crisp* (não *Fuzzy*) com densidade de probabilidade D , cuja a função de pertinência é dada por $f_x(C, \theta)$, onde θ é vetor de parâmetros para esta função densidade de probabilidade com $\theta = \{\theta_1, \theta_2, \dots, \}$. A probabilidade do evento C em B é definido por (MORAES, 2015):

$$p(C)[\alpha] = \left\{ \int_{R^n} f_x(C, \theta)dx \mid \theta \in \theta_f[\alpha] \right\} \quad (3.13)$$

onde $0 \leq \alpha \leq 1$, θ_f são números *Fuzzy* e $\theta_f[\alpha]$ seus respectivos *cortes* – α definidos por (3.3).

Como o vetor de parâmetros θ é desconhecido, uma maneira de estimá-los é usando Intervalos de Confiança (IC), variando-se os níveis de confiança (BUCKLEY, 2005).

3.9 Distribuição Exponencial com Parâmetros Fuzzy

A distribuição Exponencial com Parâmetros *Fuzzy* pode ser generalizada por (3.13) onde seu parâmetro $\bar{\lambda}$ será um número *Fuzzy*, que pode ser calculado por meio de um IC utilizando o nível de significância β que corresponde a cada *cortes* – α de $\bar{\lambda}$. O nível de significância β é a probabilidade calculada a partir de uma normal $(0, 1)$ já que para um n suficientemente grande podemos aproximar a distribuição Exponencial usual pela distribuição Normal.

Para $0 < \beta \leq 1$, para efeitos computacionais será utilizado $0,01 \leq \beta \leq 1$. O cálculo a ser realizado para a determinação de $\bar{\lambda}$ é expresso por:

$$\bar{\lambda}[\beta] = \left[\frac{\sqrt{n\bar{x}}}{z_{\beta/2} + \sqrt{n}}, \frac{\sqrt{n\bar{x}}}{\sqrt{n} - z_{\beta/2}} \right] \quad (3.14)$$

onde \bar{X} será a média de cada amostra utilizada no estudo. Portanto, a distribuição Exponencial *Fuzzy* será expressa por:

$$f_x(x, \bar{\lambda}) = \bar{\lambda} \exp(-\bar{\lambda}x), \quad (3.15)$$

onde $\bar{\lambda} \in \bar{\lambda}[\alpha]$ (BUCKLEY, 2005).

3.10 Método de Avaliação de Treinamento baseado na Rede *Naive Bayes*

Avaliadores *Naive Bayes* (NB) são conhecidos por serem simples, mas produzirem resultados satisfatórios. São baseados no teorema de *Bayes*, tomando-se a suposição *Naive* (ingênuo) de que as variáveis aleatórias que descrevem suas características são independentes entre si. Por apresentar fácil implementação, precisão e rapidez, os avaliadores NB são aplicados em diversas áreas da literatura (FRANK *et al.*, 2002).

Definição 7. Seja M o total de classes num espaço de decisão $\Omega = \{1, \dots, M\}$ e X uma variável aleatória que descreve n características distintas definidas por $X = \{X_1, X_2, \dots, X_n\}$ e ω_i , $i \in \Omega$ é uma classe no espaço de decisão para o vetor de X (MORAES; MACHADO, 2012). Assim, a probabilidade X ser atribuído a classe ω_i , pode ser estimada a partir do Teorema de *Bayes*, simplificado pela hipótese *Naive Bayes* como:

$$P(\omega_i|X) = \frac{P(X|\omega_i)P(\omega_i)}{P(X)} = \frac{P(X_1, X_2, \dots, X_n|\omega_i)P(\omega_i)}{P(X)}. \quad (3.16)$$

O cálculo da Equação (3.17) apresenta uma grande complexidade à medida que se incrementa um número k de variáveis. Logo assumindo a hipótese *NB*, temos que todo

X_k é condicionalmente independente de todas as outras características X_i , para todo $k \neq l \leq n$ e o avaliador poderá classificar os dados para quais não estava treinado (RAMONIA; SEBASTIANI, 2001). Assim, a menos de um fator de escala S do que depende de X_1, X_2, \dots, X_n a Equação (3.16) pode ser escrita por (MORAES; MACHADO, 2012):

$$P(\omega_i|X_1, X_2, \dots, X_n) = \left(\frac{1}{S}\right)P(\omega_i) \prod_{k=1}^n P(X_k|\omega_i) \quad (3.17)$$

onde $P(*)$ é a probabilidade clássica dada na Equação (3.10).

Logo, a regra de avaliação para o Método de Avaliação de Treinamento baseado na Rede *Naive Bayes* (MATRNB) é dada por (MORAES; MACHADO, 2012):

$$X \in \omega_i, \text{ se } P(\omega_i|X_1, X_2, \dots, X_n) > P(\omega_j|X_1, X_2, \dots, X_n) \quad (3.18)$$

para todo $i \neq j$.

3.11 Método de avaliação de Treinamento Baseado na Rede Exponencial *Naive Bayes*

A partir da Equação (3.17) é possível assumir distribuição Exponencial para cada X_i onde (MORAES; MACHADO, 2016):

$$P(X_k = x|\omega_i) = \lambda_{ki} \exp(-\lambda_{ki}x_k), \quad (3.19)$$

onde o parâmetro λ_{ki} é a média da variável X_k , para uma amostra de dados D numa classe i (FELLER, 1971).

A partir da Equação (3.17) é possível utilizar o logaritmo a fim de reduzir a complexidade computacional na fórmula da distribuição Exponencial na Equação (3.19), substituindo as multiplicações por adições, temos que:

$$\begin{aligned} g(\omega_i|X_1, X_2, \dots, X_n) &= \log[P(\omega_i|X_1, X_2, \dots, X_n)] = \\ &= \log\left(\frac{1}{S}\right) + \log[P(\omega_i)] + \sum_{k=1}^n \log[P(X_k|\omega_i)] \end{aligned} \quad (3.20)$$

onde g é a função de avaliação, S é um fator de escala e $P(X_k|\omega_i)$ é dada por (3.19).

Logo, podemos reescrever $\log[P(X_k|\omega_i)]$ na Equação 3.20:

$$\log[P(X_k|\omega_i)] = \log[\lambda_{ki} \exp(-\lambda_{ki}X_k)] = \log(\lambda_{ki}) - \lambda_{ki}X_k, \quad (3.21)$$

A regra de avaliação para o Método de Avaliação de Treinamento Baseado na Rede Exponencial Naive Bayes (MATRENB) é dado por:

$$X \in \omega_i, \text{ se } g(\omega_i|X_1, X_2, \dots, X_n) > g(\omega_j|X_1, X_2, \dots, X_n) \quad (3.22)$$

para todo $i \neq j$ com a função g dada pela Equação (3.20).

3.12 Método de Avaliação de Treinamento baseado na Rede *Fuzzy Exponencial Naive Bayes*

A partir da Equação (3.12) pode-se assumir que X_1, X_2, \dots, X_n são variáveis *Fuzzy* (KLIR; BO, 2008) e para cada função de pertinência $\mu_{\omega_i}(X_k)$ é válido que $k \leq n$. Logo, com base na probabilidade de eventos *Fuzzy* (ZADEH, 1968) dado pela Equação (3.12) o Método de Avaliação de Treinamento baseado na Rede *Fuzzy Exponencial Naive Bayes* (MATRFENB) é dado por (MORAES; MACHADO, 2012):

$$\begin{aligned} g_f(\omega_i|X_1, X_2, \dots, X_n) &= \log[\mathbf{P}(\omega_i|X_1, X_2, \dots, X_n)] = \\ &= \log\left(\frac{1}{S_f}\right) + \log[\mathbf{P}(\omega_i)] + \sum_{k=1}^n \log[\mu_{\omega_i} \mathbf{P}(X_k|\omega_i)] \end{aligned} \quad (3.23)$$

onde g_f é a função de avaliação, S_f é um novo fator de escala e $\log[P(X_k|\omega_i)]$ é dado pela Equação (3.21).

Entretanto, como S_f é um fator de escala não é necessário ser calculado nesse processo de maximização. Assim, a partir das Equações (3.23) e (3.21) temos que:

$$g_f(\omega_i|X_1, X_2, \dots, X_n) = \log[\mathbf{P}(\omega_i)] + \sum_{k=1}^n \log[\mu_{\omega_i}(X_k)] + \log(\lambda_{Ki}) - [\lambda_{ki}X_k] \quad (3.24)$$

Finalmente, a regra de avaliação para MATRFENB é dado por:

$$X \in \omega_i, \text{ se } g_f(\omega_i|X_1, X_2, \dots, X_n) > g_f(\omega_j|X_1, X_2, \dots, X_n) \quad (3.25)$$

para todo $i \neq j$ com a função g_f dada pela Equação (3.24).

3.13 Método de Avaliação de Treinamento Baseado na Rede *Naive Bayes* Exponencial com Parâmetros *Fuzzy*

A partir da Segunda Generalização do *Naive Bayes* (SGNB) proposta por Moraes (2015), sendo p a probabilidade cujo parâmetro é *Fuzzy*, na Equação (3.13) e S um fator de escala, pode-se utilizá-la para reescrever a Equação (3.17) como:

$$p(\omega_i|X_1, X_2, \dots, X_n)_{[\alpha]} = \left(\frac{1}{S_p}\right)p(\omega_i) \prod_{k=1}^n p(X_k|\omega_i)_{[\alpha]}. \quad (3.26)$$

Como no processo de avaliação não se utiliza as integrais, mas o cálculo diretamente sobre as densidades (DUDA *et al.*, 2000). Assim, pela SGNB dada pela Equação (3.26) temos:

$$X \in \omega_i, \text{ se } p(\omega_i|X_1, X_2, \dots, X_n) > p(\omega_j|X_1, X_2, \dots, X_n) \quad (3.27)$$

para todo $i \neq j$ com i e $j \in \Omega$ e cujos cálculos são realizados usando-se a Equação (3.26).

A partir da Equação (3.26) é possível assumir a distribuição Exponencial *Fuzzy* para X e computar seu parâmetro *Fuzzy* $\bar{\lambda}$, ou seja, sua média (RODRIGUES *et al.*, 2016). Por fim, aplicando o logaritmo, a função de avaliação g_p é definida por:

$$\begin{aligned}
g_p(\omega_i|X_1, X_2, \dots, X_n) &= \log[p(\omega_i|X_1, X_2, \dots, X_n)] = \\
&= \log\left(\frac{1}{S_p}\right) + \log[p(\omega_i)] + \sum_{k=1}^n \log[p(X_k|\omega_i)].
\end{aligned} \tag{3.28}$$

Como S_p é um fator de escala, não é necessário ser calculado nesse processo de maximização. Logo, aplicando (3.15) em (3.28), obtemos:

$$g_p(\omega_i|X_1, X_2, \dots, X_n) = \log[p(\omega_i)] + \sum_{k=1}^n \log[\bar{\lambda}_k] - \bar{\lambda}x, \tag{3.29}$$

onde g_p é a função de avaliação para o Método de Avaliação de Treinamento Baseado na Rede *Naive Bayes* Exponencial com Parâmetros *Fuzzy* (MATRNBEFP) e $\bar{\lambda}_k$ é estimado pela Equação (3.15).

A regra de avaliação do MATRNBEFP é dada por (RODRIGUES *et al.*, 2016):

$$X \in \omega_i, \text{ se } g_p(\omega_i|X_1, X_2, \dots, X_n) > g_p(\omega_j|X_1, X_2, \dots, X_n), \tag{3.30}$$

para todo $i \neq j$ com i e $j \in \Omega$.

3.14 Coeficientes de Concordância

3.14.1 Acerto Percentual

O Acerto Percentual é calculado através de uma matriz de Confusão, a qual é obtida após o término da execução do método. A sua estrutura é semelhante quando são avaliadas quatro classes de desempenho (Tabela 3.1) em que as linhas são as classes de desempenho reais e as colunas as classes de desempenho obtidas. Os resultados presentes na diagonal principal da Matriz de Confusão representam os dados que foram avaliados corretamente, já os demais correspondem aos termos que foram avaliados incorretamente. O acerto percentual é calculado da seguinte maneira:

$$AC = \frac{\sum_{i=1}^M n_{ii}}{\sum_{i=1}^M \sum_{j=1}^M n_{ij}} \tag{3.31}$$

em que $\sum_{i=1}^M n_{ii}$ é a soma da diagonal principal, $\sum_{i=1}^M \sum_{j=1}^M n_{ij}$ é a soma de todos os

Tabela 3.1: Matriz de Confusão.

classes de desempenho reais (c)	classes de desempenho obtidas (\hat{c})			
	1	2	3	4
1	n_{11}	n_{12}	n_{13}	n_{14}
2	n_{21}	n_{22}	n_{23}	n_{24}
3	n_{31}	n_{32}	n_{33}	n_{34}
4	n_{41}	n_{42}	n_{43}	n_{44}

elementos da matriz e o M é o total de classes (MORAES; MACHADO, 2014).

3.14.2 Coeficiente Kappa

O Coeficiente Kappa foi proposto por Cohen (1960), é uma medida robusta ponderada que leva em conta os acertos e erros de acordo com a Matriz de Confusão (MORAES; MACHADO, 2014). Esse coeficiente é expresso por:

$$K = \frac{(P_0 - P_c)}{(1 - P_c)}, \quad (3.32)$$

onde $P_0 = \sum_{i=1}^M (n_{ii})/N$, $P_c = \sum_{i=1}^M (n_{i+}n_{+i})/N^2$, onde n_{ii} é o total da diagonal principal da Matriz de Confusão; n_{i+} é o total da linha i na Matriz de Confusão; n_{+i} é o total da coluna da mesma matriz; M é o total de classes de desempenho na matriz e N é o número total de decisões presentes na matriz.

A variância do coeficiente Kappa é denotado por σ_K^2 é dado por Moraes e Machado (2014):

$$\sigma_K^2 = \frac{P_0(1 - P_0)}{N(1 - P_c)^2} + \frac{2(1 - P_0) + 2P_0P_c - \theta_1}{N(1 - P_c)^3} + \frac{(1 - P_0) + \theta_2 - 4P_c^2}{N(1 - P_c)^4}. \quad (3.33)$$

onde θ_1 e θ_2 são definidos por:

$$\theta_1 = \left[\sum_i^M n_{ii}(n_{i+}n_{+i}) \right] / N^2, \quad (3.34)$$

$$\theta_2 = \left[\sum_i^M n_{ii}(n_{i+}n_{+i})^2 \right] / N^3. \quad (3.35)$$

Este coeficiente é comumente utilizado para termos de avaliação de treinamento

(DUDA *et al.*, 2000). A partir dele pode-se distinguir em que situação se encontra a concordância das avaliações em relação a um sistema de referência. De acordo com Landis e Koch (1977), o coeficiente Kappa podem ser interpretado através da Tabela 3.2.

Tabela 3.2: Interpretação do coeficiente Kappa.

Coeficiente Kappa	Grau de Concordância
< 0,00	Concordância Pobre
0,00-0,20	Concordância Pequena
0,21-0,40	Concordância Regular
0,42-0,60	Concordância Moderada
0,61-0,80	Concordância Considerável
0,81-1,00	Concordância Quase perfeita

Através de testes de hipóteses é possível avaliar estatisticamente a diferença entre os coeficientes Kappa para os demais avaliadores presentes nesse estudo. Proposto por Congalton e Green (2008) a estatística de teste para diferença entre os coeficientes Kappa é dado por:

$$Z_c = \frac{|\hat{K}_1 - \hat{K}_2|}{\sqrt{\sigma_1^2 \hat{K}_1 + \sigma_2^2 \hat{K}_2}} \sim Normal(0, 1). \quad (3.36)$$

Com hipótese nula de $H_0 = K_1 - K_2 = 0$ e hipótese alternativa de $H_1 = K_1 - K_2 \neq 0$. Logo, rejeitamos a hipótese nula se $|Z_c| > Z_{\alpha/2}$, em que $Z_{\alpha/2}$ é o quantil da distribuição Normal (0, 1) .

Capítulo 4

Materiais

Recentemente foram propostos métodos de classificação utilizando a Rede *Naive Bayes* em conjunto com a distribuição Exponencial (RODRIGUES *et al.*, 2016) (MORAES; MACHADO, 2016). No entanto, esses classificadores não foram propostos para fins de Avaliação de Treinamento. Esse trabalho propõe-se utilizá-los desta forma, porém não foi realizada uma comparação entre esses métodos. Assim, houve a necessidade de utilizar simulações de Monte Carlo (GENTLE, 2005) para gerar dois conjuntos de amostras que seguem a distribuição Exponencial. Essas amostras denominam-se *Amostras para o Treinamento* e *Amostras para Teste* e serviram de base para os resultados obtidos com o MATRNB, MATRENB, MATRFENB e MATRNBEPF.

4.1 Amostras Para Treinamento e Teste

As *Amostras para Treinamento* servem para estimar os parâmetros que na prática também podem ser fornecidas por um ou mais especialistas. Para o conjunto de amostras de treinamento foram geradas 16 amostras com 30000 observações, porém as 20000 primeiras foram descartadas para evitar flutuações nos dados (GENTLE, 2005). O estudo utilizou apenas as 10000 últimas observações com dados de 1 a 4 dimensões, sendo as dimensões as variáveis relacionadas as diferentes fases do exame ginecológico.

O segundo bloco de amostras são denominadas como *Amostras para Teste* que são as amostras sobre as quais são realizadas as avaliações. Para esse conjunto de amostras foram geradas 16 amostras com 50000 observações cada, porém as 20000 primeiras observações também foram descartadas, obtendo 30000 observações com dados de 1 a 4

dimensões.

Para conseguir o resultado almejado nesse estudo, foi necessário adquirir algumas técnicas referentes a área de computação, os dois conjuntos de amostras (treinamento e teste) foram geradas e validadas com o auxílio do *software R* (TEAM, 2014) e processadas por meio de programas em linguagem C.

4.1.1 Simulação

Como citado anteriormente foram geradas amostras para a distribuição estatística Exponencial onde utilizou-se o *software R*. Para a distribuição Exponencial utiliza-se a seguinte notação $X \sim Exp(\lambda)$ e seus parâmetros para a geração das amostras são encontrados na Tabela 4.1.

Tabela 4.1: Parâmetros da distribuição Exponencial.

Exponencial $X \sim \text{EXP}(\lambda)$	Classe 1	Classe 2	Classe 3	Classe 4
Dimensão 1	0,2	0,08	0,04	0,01
Dimensão 2	0,3	0,03	0,02	0,003
Dimensão 3	0,4	0,07	0,04	0,02
Dimensão 4	0,2	0,07	0,04	0,03

Os dois blocos de amostras (treinamento e teste) são divididos em classes de desempenho e dimensões, onde as classes de desempenho são os possíveis resultados do procedimento ou seja, os rótulos que os dados vão receber. As dimensões são as variáveis que foram escolhidas. Para quantificar o desempenho dos avaliadores MATRNB, MATRENB, MATRFENB e MATRNBEPF, na geração das amostras, todas as classes de desempenho foram forçadas a ter uma interseção entre si de no mínimo 20%, a fim de que pudesse verificar a acurácia dos métodos em avaliar valores interseccionados. Para a execução do método, foi variado o número de dimensões e armazenado em uma Matriz de Confusão. Os resultados foram analisados de acordo com os acertos percentuais, coeficiente Kappa, sua variância e o tempo computacional.

Capítulo 5

Resultados e Discussão

Neste capítulo estão descritos os resultados obtidos nesse estudo, utilizando os avaliadores MATRNB, MATRENB, MATRFENB e MATRNBEPF. As amostras de treinamento e de teste foram geradas a partir da metodologia descrita anteriormente. A Figura 5.1 representa os histogramas que foram obtidos a partir dos dados gerados para a distribuição Exponencial segundo os parâmetros na Tabela 4.1.

Assim, temos os conjuntos de amostras (treinamento e teste), em que as primeiras quatro linhas dos histogramas na Figura 5.1 representam as amostras de treinamento (amarelo), as quatro linhas seguintes são as amostras de testes (verde). As linhas são as dimensões estudadas e as colunas são as classes. No processo de avaliação foi utilizado um computador, com processador Intel Core(TM) *I3-3217U CPU 1.80GHz*, com uma memória RAM de *4GB*. O tempo médio de execução para a avaliação das 120000 observações, considerando os quatro métodos foram: com uma dimensão 13,47, com duas dimensões 16,35 segundos, com três dimensões 19,31 segundos e com quatro dimensões 21,90, segundos.

As sessões 5.1, 5.2, 5.3 e 5.4 apresentamos os resultados para os MATRNB, MATRENB, MATRFENB e MATRNBEPF, de acordo com o acerto percentual, coeficiente Kappa, sua variância e tempo computacional. A sessão 5.5 apresenta uma comparação dos avaliadores segundo o percentual de acerto e o tempo computacional.

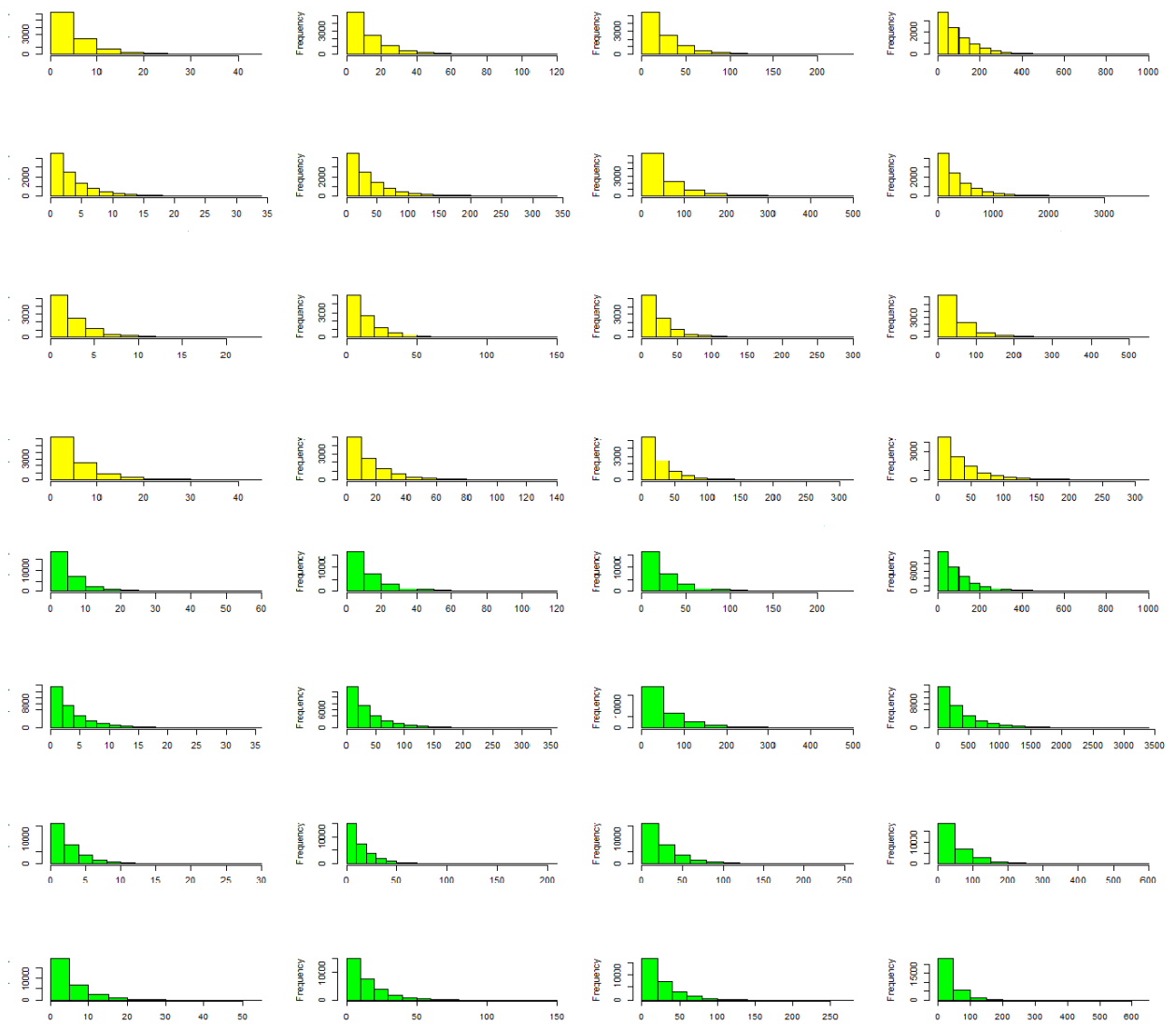


Figura 5.1: Amostra aleatória da distribuição Exponencial (linhas são as dimensões e colunas são as classes de desempenho)

5.1 Método de Avaliação de Treinamento baseado na Rede *Naive Bayes*

Na avaliação dos dados de teste que seguem a distribuição Exponencial utilizando o MATRNB foram encontrados os seguintes resultados para os coeficientes de concordância, utilizando uma dimensão, temos que o acerto percentual foi de 27,15%. Na mesma dimen-

são o coeficiente Kappa obtido foi de 2,87% e sua variância no valor de $8,52 \times 10^{-6}$, com um grau de concordância pequeno de acordo com a Tabela 3.2. O tempo computacional gasto pelo MATRNB para essa dimensão foi de 7,47 segundos, para avaliar 120000 observações.

Utilizando duas dimensões o acerto percentual foi de 27,35%, ou seja, o método avaliou 32820 observações corretamente. O coeficiente Kappa para essa mesma dimensão foi de 3,14% e uma variância de $7,65 \times 10^{-6}$, obtendo um grau de concordância pequeno, como definido na Tabela 3.2. Com três dimensões o acerto percentual foi de 27,34%. O coeficiente Kappa encontrado foi de 3,13% e sua variância no valor de $7,18 \times 10^{-6}$, com um grau de concordância considerado pequeno, dentro do intervalo de $[0,00; 0,20]$ na Tabela 3.2.

Para quatro dimensões o acerto percentual na Matriz de Confusão na Tabela 5.2 foi de 27,15%, com 87420 avaliadas incorretamente. O coeficiente Kappa foi de 2,86% e sua variância de $7,52 \times 10^{-6}$. O tempo de execução para essa dimensão foi de 17,90 segundos. Ao analisar o desempenho do avaliador MATRNB podemos notar que o método avalia dados com distribuição Exponencial com um grau de concordância pequeno, com uma ou mais dimensões de acordo com o coeficiente Kappa na Tabela 3.2.

Tabela 5.1: Matriz de Confusão com quatro dimensões para o MATRNB.

classes de desempenho reais (c)	classes de desempenho obtidas (\hat{c})			
	1	2	3	4
1	29980	7	9	4
2	27814	2167	13	6
3	23656	6230	111	3
4	10894	15483	3301	322

5.2 Método de avaliação de Treinamento Baseado na Rede Exponencial *Naive Bayes*

Na avaliação para dados com distribuição Exponencial utilizando o MATRENB com uma dimensão o acerto percentual foi de 51,17%. Analisando a mesma dimensão, o coeficiente Kappa encontrado foi de 34,90% e sua variância de $3,68 \times 10^{-6}$ obtendo uma grau de concordância regular, ou seja no intervalo de $[0,21; 0,40]$ na Tabela 3.2.

Com duas dimensões o acerto percentual foi de 69,18%. O coeficiente Kappa obtido foi de 58,91%, com um grau de concordância moderado de acordo com a Tabela 3.2, a variância encontrada foi de $3,12 \times 10^{-6}$. O tempo computacional gasto pelo MATRENB foi de 9,51 segundos, para duas dimensões. Para três dimensões o acerto percentual foi de 73,91%. O coeficiente Kappa encontrado foi de 65,21% e sua variância no valor de $2,84 \times 10^{-6}$, que podemos avaliar através da Tabela 3.2 como uma avaliação considerável.

Com quatro dimensões o método avaliou 91220 observações corretamente segundo a Tabela 5.2, resultando em um acerto percentual de 76,01%. O grau de concordância obtido através da Tabela 3.2 é considerável, pois o valor do coeficiente Kappa de 68,02% está no intervalo de $[0,61; 0,80]$, a variância encontrada foi de $2,69 \times 10^{-6}$. Para quatro dimensões o tempo de execução foi de 14,96 segundos. Ao analisar o desempenho do avaliador MATRENB podemos notar que a partir de três dimensões o método avalia os dados com um grau de concordância considerável de acordo com o coeficiente Kappa (Tabela 3.2).

Tabela 5.2: Matriz de Confusão com quatro dimensões para o MATRENB.

classes de desempenho reais (c)	classes de desempenho obtidas (\hat{c})			
	1	2	3	4
1	28791	1206	3	0
2	2689	20252	6946	113
3	589	9892	17093	2426
4	16	564	4336	25084

5.3 Método de Avaliação de Treinamento baseado na Rede *Fuzzy Exponencial Naive Bayes*

Para o MATRFENB foi observado que para a primeira dimensão o acerto percentual foi de 46,23%. O coeficiente Kappa foi de 28,31% e sua variância de $3,60 \times 10^{-6}$, resultando em um grau de concordância regular de acordo com a Tabela 3.2.

Utilizando duas dimensões o acerto percentual foi de 58,72%. Para essa mesma dimensão o coeficiente Kappa foi de 44,96%, no intervalo de $[0,42; 0,60]$, tendo o resultado avaliado como moderado, segundo a Tabela 3.2, com variância de $3,49 \times 10^{-6}$. Com três dimensões o MATRFENB avaliou 48696 observações incorretamente tendo um acerto

percentual de 59,42%. O coeficiente Kappa encontrado foi de 45,90% e sua variância de $3,47 \times 10^{-6}$, obtendo um grau de concordância moderado na Tabela 3.2.

Para quatro dimensões o acerto percentual foi de 58,15%, que corresponde a 69780 observações avaliadas corretamente, de acordo com a Tabela 5.3. O coeficiente Kappa foi de 44,20% e sua variância de $3,50 \times 10^{-6}$, obtendo um grau de concordância moderado (Tabela 3.2). É possível observar que o MATRFENB avalia dados com distribuição Exponencial com grau moderado de acordo com o coeficiente Kappa a partir de duas dimensões, com um tempo computacional acima de 10,18 segundos.

Tabela 5.3: Matriz de Confusão com quatro dimensões para o MATRFENB.

classes de desempenho reais (c)	classes de desempenho obtidas (\hat{c})			
	1	2	3	4
1	18800	10218	982	0
2	1419	8458	16796	3327
3	838	2511	1389	12761
4	187	81	1099	28633

5.4 Método de Avaliação de Treinamento Baseado na Rede *Naive Bayes* Exponencial com Parâmetros *Fuzzy*

Na avaliação utilizando o MATRNBEPF foi obtido os seguintes resultados, para a primeira dimensão, ou seja ao avaliar a primeira linha em amarelo na Figura 5.1 o acerto percentual foi de 51,17%. O coeficiente Kappa 34,90% e sua variância de $3,68 \times 10^{-6}$, obtendo um grau de concordância regular (Tabela 3.2).

Com duas dimensões o acerto percentual foi de 69,18%, o coeficiente Kappa de 58,91%, com um grau de concordância moderado $[0,42; 0,60]$ na Tabela 3.2 e uma variância no valor de $3,12 \times 10^{-6}$. Para três dimensões o acerto percentual obtido foi de 72,83%, o coeficiente Kappa de 63,78% e sua variância de $2,91 \times 10^{-6}$, com um grau de concordância considerável (Tabela 3.2).

Avaliando quatro dimensões o MATRNBEPF obteve um acerto percentual de 72,99% que corresponde à 87588 observações avaliadas corretamente, segundo a Tabela

5.4. O resultado é considerado moderado de acordo com a Tabela 3.2, com um coeficiente Kappa de 63,98% e sua variância de $2,90 \times 10^{-6}$. É possível notar que a partir de três dimensões ou mais o MATRNBEPF é capaz de avaliar dados que seguem uma distribuição Exponencial com um grau de concordância moderado de acordo com o coeficiente Kappa.

Tabela 5.4: Matriz de Confusão com quatro dimensões para o MATRNBEPF.

classes de desempenho reais (c)	classes de desempenho obtidas (\hat{c})			
	1	2	3	4
1	27808	2156	36	0
2	3011	19765	7084	140
3	733	11122	14885	3260
4	19	1105	3744	25132

5.5 Comparação entre os avaliadores

Como citado anteriormente, foram usadas as amostras de treino para estimar os parâmetros utilizados pelos avaliadores e em seguida foram usadas as amostras de teste para avaliar o desempenho dos métodos. Uma maneira simplificada de conhecer a qualidade dos resultados obtidos com a avaliação com os MATRNB, MATRENB, MATRFENB e MATRNBEPF é a comparação utilizando os coeficientes de concordância. De modo geral, MATRENB foi significativamente melhor que os MATRNB, MATRFENB e MATRNBEPF, avaliando dados com distribuição Exponencial em 14,96 segundos, com quatro dimensões.

Comparando os resultados do MATRENB com os outros três avaliadores descritos nesse trabalho temos que o MATRNBEPF apresentou desempenho semelhante ao MATRENB com uma, duas e três dimensões (Tabelas 5.5, 5.6 e 5.7). Para quatro dimensões a diferença do MATRENB para o MATRNBEPF foi de 4,04% de acordo com o coeficiente Kappa Tabela 5.8, o que pode ser verificado também através dos acertos percentuais na Figura 5.2.

Essa diferença é confirmada através do teste de hipótese que fixando um nível de significância de 5%, rejeitamos a hipótese nula ($K_1 - K_2 = 0$), ou seja existe diferença entre os coeficientes Kappa do avaliador MATRENB para o MATRNBEPF, analisando quatro dimensões. Embora o percentual de acerto dos avaliadores MATRENB e MATRNBEPF

sejam semelhantes, a diferença dos tempos computacionais quando comparado por dimensões chega a 24,22 segundos, com quatro dimensões, sendo o MATRENB mais rápido Tabela 5.9.

Ao comparar o desempenho dos avaliadores MATRNB e MATRFENB, temos que para todas as dimensões estudadas o MATRFENB possui uma taxa de acerto superior ao MATRNB (Tabelas 5.5, 5.6, 5.7 e 5.8). A Figura 5.2 apresenta essa diferença, que é confirmada por meio do teste de hipótese, com um nível de significância de 5%, rejeitou a hipótese nula ($K_1 - K_2 = 0$) de que os coeficientes Kappa são iguais, para uma, duas, três e quatro dimensões.

Em relação aos tempos de execução o MATRFENB é mais rápido que o MATRNB, em todas as dimensões estudadas Tabela 5.9. Com quatro dimensões a diferença percentual dos avaliadores MATRNB e MATRFENB para o MATRENB foram 65,16% e 23,82% (Tabela 5.8), respectivamente, tomando como referência o coeficiente Kappa.

Analisando os avaliadores MATRFENB e MATRNBEPF, nos quais foram incrementadas a informação *Fuzzy* (Sessões 3.12 e 3.13) temos que para todas as dimensões, o coeficiente Kappa para o MATRNBEPF foi maior que o MATRFENB (Tabelas 5.5, 5.6, 5.7 e 5.8). Por meio do teste para diferença dos coeficientes Kappa, com 5% de significância, foi possível constatar que existe diferença entre os coeficientes do MATRNBEPF par o MATRFENB, pois a hipótese nula foi rejeitada. Entretanto, nas quatro dimensões o MATRNBEPF obteve um maior tempo computacional Tabela 5.9.

Com uma dimensão Tabela 5.6 o desempenho dos avaliadores alcançaram um percentual de acerto menor que as demais dimensões analisadas. Esses resultados são esperados para a primeira dimensão pois se analisa apenas uma característica dos dados. Na prática quanto mais dimensões forem analisadas, espera-se melhores resultados com a avaliação (MORAES; MACHADO, 2012) (MORAES; MACHADO, 2014). O percentual de acerto para o MATRNB foi o menor valor encontrado para todas as dimensões analisadas, quando comparado com os métodos MATRENB, MATRFENB e MATRNBEPF. É possível observar também que com o incremento de dimensões o percentual de acerto aumenta para a maioria dos avaliadores.

Tabela 5.5: Comparação entre os avaliadores para a primeira dimensão.

Avaliadores	Acerto Percentual	Coefficiente Kappa	Variância
MATRNB	27,15%	2,87%	$8,52 \times 10^{-6}$
MATRENB	51,17%	34,90%	$3,68 \times 10^{-6}$
MATRFENB	46,23%	28,31%	$3,60 \times 10^{-6}$
MATRNBEPF	51,17%	34,90%	$3,68 \times 10^{-6}$

Tabela 5.6: Comparação entre os avaliadores para primeira e segunda dimensões.

Avaliadores	Acerto Percentual	Coefficiente Kappa	Variância
MATRNB	27,35%	3,14%	$7,65 \times 10^{-6}$
MATRENB	69,18%	58,91%	$3,12 \times 10^{-6}$
MATRFENB	58,72%	44,96%	$3,49 \times 10^{-6}$
MATRNBEPF	69,18%	58,91%	$3,12 \times 10^{-6}$

Tabela 5.7: Comparação entre os avaliadores para primeira, segunda e terceira dimensões.

Avaliadores	Acerto Percentual	Coefficiente Kappa	Variância
MATRNB	27,34%	3,13%	$7,18 \times 10^{-6}$
MATRENB	73,91%	65,21%	$2,84 \times 10^{-6}$
MATRFENB	59,42%	45,90%	$3,47 \times 10^{-6}$
MATRNBEPF	72,83%	63,78%	$2,91 \times 10^{-6}$

Tabela 5.8: Comparação entre os avaliadores para todas as quatro dimensões.

Avaliadores	Acerto Percentual	Coefficiente Kappa	Variância
MATRNB	27,15%	2,86%	$7,52 \times 10^{-6}$
MATRENB	76,01%	68,02%	$2,69 \times 10^{-6}$
MATRFENB	58,15%	44,20%	$3,50 \times 10^{-6}$
MATRNBEPF	72,99%	63,98%	$2,90 \times 10^{-6}$

Tabela 5.9: Tempo Computacional por métodos e dimensão, em segundos.

	MATRNB	MATRENB	MATRFENB	MATRNBEPF
Dimensão 1	7,47	6,84	6,88	32,69
Dimensão 2	10,86	9,51	10,18	34,86
Dimensão 3	14,46	12,37	12,68	37,73
Dimensão 4	17,90	14,96	15,76	39,16

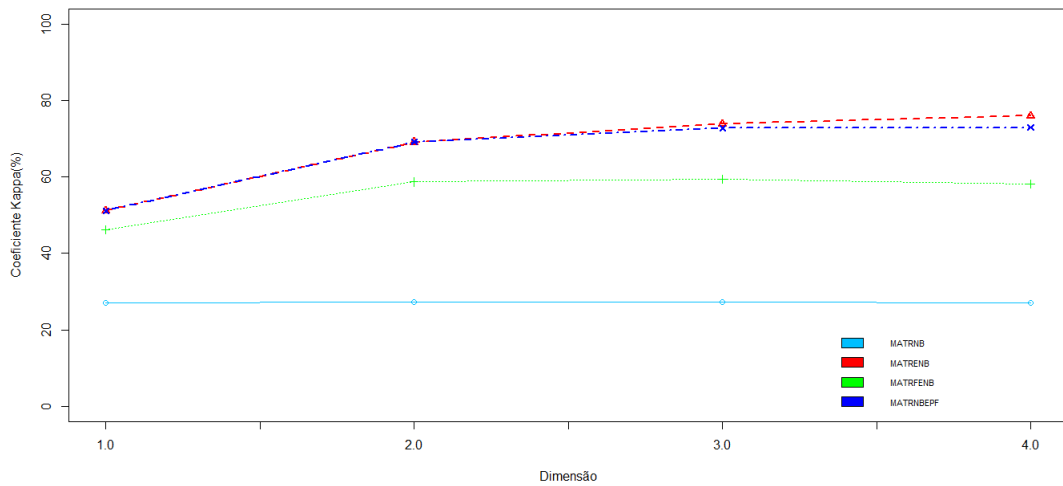


Figura 5.2: Comparação entre os avaliadores de acordo com o Acerto Percentual por dimensão.

A Figura 5.3 resume a precisão de cada método por meio de Coeficiente Kappa (lado esquerdo) e o desempenho computacional usando o Tempo da CPU (lado direito). É possível observar que o MATRNB apresentou a menor precisão e o maior tempo computacional quando comparado com os métodos MATRENB e MATRFENB. O método MATRNBEPF apresentou uma boa precisão de acordo com o coeficiente Kappa, porém seu tempo computacional foi o maior entre todos os avaliadores estudados. Por outro lado, o MATRENB forneceu a maior precisão com menor tempo computacional.

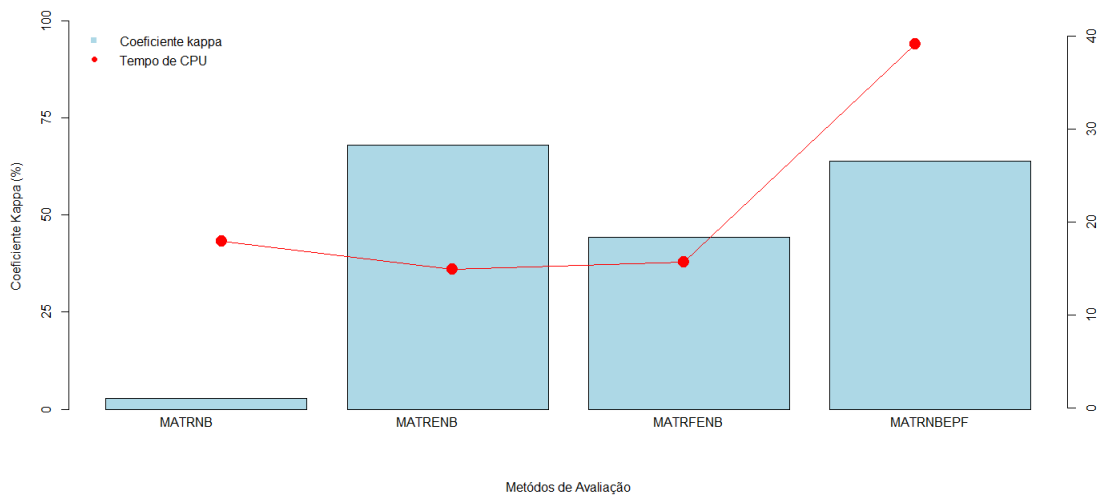


Figura 5.3: Comparação entre os métodos avaliação de acordo com o Coeficiente Kappa e o Tempo Computacional para quatro dimensões.

Capítulo 6

Conclusão

Como resultado desse trabalho foi apresentada uma comparação entre os avaliadores MATRNB, MATRENB, MATRFENB e MATRNBEPF, utilizando como parâmetros coeficientes de concordância e o tempo computacional. Para cada avaliador foram utilizados dados com quatro dimensões e avaliados segundo quatro classes de desempenho.

Para a primeira dimensão, todos os métodos obtiveram resultados espúrios. Possivelmente isso se deve à grande intersecção dos dados simulados entre as classes de desempenho. A partir de duas dimensões os resultados alcançam melhores graus de concordância de acordo com o coeficiente Kappa. Com três ou mais dimensões, a maioria dos avaliadores fornece um grau de concordância moderado, para os dados simulados da distribuição Exponencial. Com quatro dimensões o MATRENB obteve um resultado considerável de acordo com o coeficiente Kappa. Em relação ao tempo computacional o MATRENB conseguiu avaliar dados com baixa complexidade computacional para todas as dimensões analisadas.

O método MATRNBEPF possui um percentual de acerto semelhante ao MATRENB para as três primeiras dimensões sendo confirmado através do teste de hipótese para diferença dos coeficientes Kappa. Entretanto, esse o MATRNBEPF possui o maior tempo de CPU observado nesse estudo. O MATRNB obteve o menor percentual de acerto para todas as dimensões analisadas quando comparadas com o desempenho dos outros três avaliadores estudados.

Em relação ao simulador do exame ginecológico (Siteg), temos que de acordo com os resultados obtidos nesse trabalho o melhor método de avaliação para modelar os tempos de execução presentes nas diferentes fases do simulador é o MATRENB que apresentou a

melhor performance e precisão dentre todas as dimensões analisadas.

Como possíveis trabalhos futuros podemos sugerir:

- avaliar o desempenho desses métodos utilizando outras distribuições Estatísticas;
- incluir outras metodologias na comparação de desempenho utilizando esses dados;
- incluir um maior número de dimensões e avaliar performance dos métodos citados acima;
- a utilização de métodos de correção de viés para o parâmetro da distribuição Exponencial.

Publicações Realizadas e Registros Solicitados

- RODRIGUES, A.K.G.; BATISTA, T.V.V.; MORAES, R.M., & MACHADO, L.S. (2016, July). A new Exponential Naive Bayes Classifier with Fuzzy Parameters. In Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference, Vancouver. p. 1188-1194.
- RODRIGUES, A.K.G.; MORAES, R.M. Avaliação da Rede Exponencial Naive Bayes com Parâmetros Fuzzy Usando Diferentes Distribuições Estatísticas. In: Simpósio Nacional de Probabilidade e Estatística, 2016, Porto Alegre. Anais do 22º Simpósio Nacional de Probabilidade e Estatística, 2016. p.81.
- RODRIGUES, A.K.G.; BATISTA, T.V.V. ; MORAES, R.M. ; MACHADO, L.S. Rede Exponencial Naive Bayes com Parâmetros Fuzzy (Rede ExpNB-Parâmetros Fuzzy) 2016. Referências adicionais: Brasil/; Finalidade: Classificação de dados e Avaliação de treinamento; Plataforma: Plataforma: C/C++; Inst. promotora/financiadora: Universidade Federal da Paraíba. Registro solicitado em 11/04/2016, sob o protocolo nº 000032-0.

Referências Bibliográficas

- BASDOGAN, C.; DE, S.; KIM, J.; MUNIYANDI, M.; KIM, H.; SRINIVASAN, M. A. Haptics in minimally invasive surgical simulation and training. *IEEE computer graphics and applications*, IEEE, v. 24, n. 2, p. 56–64, 2004.
- BISHOP, C. *Pattern Recognition and Machine Learning*. 1st. ed. Berlin: Springer, 2007.
- BUCKLEY, J. J. *Simulating fuzzy systems*. [S.l.]: Springer Science & Business Media, 2005. v. 171.
- BURDEA, G.; COIFFET, P. *Virtual Reality Technology*. New Jersey: John Wiley and Sons, 2003.
- CAETANO, S. L.; LOUZADA-NETO, F. Controle de qualidade via dados acelerados com distribuição exponencial e relação estresse-resposta lei de potência inversa. *Rev. Mat. Estat*, v. 25, n. 1, p. 85–98, 2007.
- CATALUNHA, M. J.; SEDIYAMA, G. C.; LEAL, B. G.; SOARES, C. P. B.; RIBEIRO, A. Aplicação de cinco funções densidade de probabilidade a séries de precipitação pluvial no estado de minas gerais. *Revista Brasileira de Agrometeorologia*, v. 10, n. 1, p. 153–162, 2002.
- COHEN, J. A coefficient of agreement for nominal scales. In: *Educat. Psyc. Measurement*. [S.l.: s.n.], 1960.
- CONGALTON, R. G.; GREEN, K. *Assessing the accuracy of remotely sensed data: principles and practices*. [S.l.]: CRC press, 2008.
- DINSMORE, M.; LANGRANA, N.; BURDEA, G.; LADEJI, J. Virtual reality training simulation for palpation of subsurface tumors. In: IEEE. *Virtual Reality Annual International Symposium, 1997., IEEE 1997*. [S.l.], 1997. p. 54–60.
- DUAN, J.; SELKER, J.; GRANT, G. E. *Evaluation Probability Density Functions in Precipitation Models for The Pacific Northwest1*. [S.l.]: Wiley Online Library, 1998.
- DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification*. 2nd. ed. [S.l.]: Wiley Interscience, 2000.
- FELLER, W. *An Introduction to Probability Theory and its Applications*. 2nd. ed. [S.l.]: Wiley, 1971.
- FERREIRA, J. A.; SOARES, E. A. M. G.; MORAES, R. M. Assessment of fuzzy gaussian naive bayes classifier using data with different statistical distributions. In: *Congresso Brasileiro de Sistemas Fuzzy (CBSF2014)*. [S.l.: s.n.], 2014.

- FRANK, E.; HALL, M.; PFAHRINGER, B. Locally weighted naive bayes. In: MORGAN KAUFMANN PUBLISHERS INC. *Proceedings of the Nineteenth conference on Uncertainty in Artificial Intelligence*. [S.l.], 2002. p. 249–256.
- GENTLE, J. E. *Elements of Computational Statistics*. [S.l.]: Springer, 2005.
- HARDERS, M. *Surgical scene generation for virtual reality-based training in medicine*. [S.l.]: Springer Science & Business Media, 2008.
- KLIR, G. J.; BO, Y. Fuzzy sets and fuzzy logic, theory and applications. 2008.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. In: *Biometrics*. [S.l.: s.n.], 1977. p. 159–174.
- MACHADO, L. S.; MORAES, R. M. Vr-based simulation for the learning of gynaecological examination. *Advances in Artificial Reality and Tele-Existence*, Springer, p. 97–104, 2006.
- MACHADO, L. S.; MORAES, R. M.; ZUFFO, M. K. Fuzzy rule-based evaluation for a haptic and stereo simulator for bone marrow harvest for transplant. In: CITESEER. *5th Phantom Users Group Workshop Proceedings*. [S.l.], 2000.
- MORAES, R. M. Performance analysis of evolving fuzzy neural networks for pattern recognition. In: *Mathware and Soft Computing*. [S.l.: s.n.], 2013. p. 63–69.
- MORAES, R. M. A new generalization for naive bayes style fuzzy probabilistic classifier. In: *XV Safety, Health and Environment World Congress (SHEWC'2015)*. Portugal: [s.n.], 2015.
- MORAES, R. M.; MACHADO, L. S. Using fuzzy hidden markov models for online training evaluation and classification in virtual reality simulators. *International Journal of General Systems*, Taylor & Francis, v. 33, n. 2-3, p. 281–288, 2004.
- MORAES, R. M.; MACHADO, L. S. Maximum likelihood for on-line evaluation of training based on virtual reality. In: *Proceedings of Global Congress on Engineering and Technology Education (GCETE'2005)*. Março, Santos, Brasil. [S.l.: s.n.], 2005. p. 299–302.
- MORAES, R. M.; MACHADO, L. S. Assessment based on naive bayes for training based on virtual reality. 2007.
- MORAES, R. M.; MACHADO, L. S. Using embedded systems to improve performance of assessment in virtual reality training environments. In: *International Conference on Engineering and Technology Education*. [S.l.: s.n.], 2008.
- MORAES, R. M.; MACHADO, L. S. Another approach for fuzzy naive bayes applied on online training assessment in virtual reality simulators. In: *Safety, Health and Environmental World Congress*. [S.l.: s.n.], 2009.
- MORAES, R. M.; MACHADO, L. S. Online training evaluation in virtual reality simulators using possibilistic networks. In: *Proceedings of Safety Health and Environmental World Congress 2009 (SHEWC'2009)*. [S.l.: s.n.], 2009. p. 67–71.

- MORAES, R. M.; MACHADO, L. S. Development of a medical training system with integration of users' assessment. In: *Virtual Reality*. [S.l.]: InTech, 2010.
- MORAES, R. M.; MACHADO, L. S. *Intelligent Decision Making in Training Based on Virtual Reality*. [S.l.]: Atlantis Press, 2010. 85-126 p.
- MORAES, R. M.; MACHADO, L. S. Assessment systems for training based on virtual reality: A comparison study. In: . [S.l.]: SBC Journal on 3D Interactive Systems, 2012.
- MORAES, R. M.; MACHADO, L. S. An online training assessment based on naive bayes network modeled by gaussian mixture models for medical simulators. In: *Proceedings of International Conference on Engineering and Computer Education*. [S.l.: s.n.], 2013. v. 7, p. 120–124.
- MORAES, R. M.; MACHADO, L. S. Psychomotor skills assessment in medical training based on virtual reality using a weighted possibilistic approach. In: *Knowledge Based Systems*. [S.l.: s.n.], 2014. p. 97–62.
- MORAES, R. M.; MACHADO, L. S. A fuzzy exponential naive bayes classifier. In: WORLD SCIENTIFIC. *Uncertainty Modelling in Knowledge Engineering and Decision Making: Proceedings of the 12th International FLINS Conference (FLINS 2016)*. [S.l.], 2016. v. 10, p. 207.
- NETTO, A.; MACHADO, L.; OLIVEIRA, M. C. F. Realidade virtual-definições, dispositivos e aplicações. *Revista Eletrônica de Iniciação Científica-REIC. Ano II*, v. 2, 2002.
- RAMONIA, M.; SEBASTIANI, P. Robust bayes classifiers. In: *Artificial Intelligence*. [S.l.: s.n.], 2001. p. 209–226.
- RHEINGOLD, H. *Virtual Reality: Exploring the Brave New Technologies*. [S.l.]: Simon & Schuster Adult Publishing Group, 1991.
- ROBBINS, S. L.; COTRAN, R. S.; KUMAR, V.; COLLINS, T. Fundamentos de robbins: patologia estrutural e funcional. In: *Fundamentos de Robbins: patologia estrutural e funcional*. [S.l.]: Guanabara Koogan, 2001.
- RODRIGUES, A. K. G.; BATISTA, T. V. V.; MORAES, R. M.; MACHADO, L. S. A new exponential naive bayes classifier with fuzzy parameters. In: IEEE. *Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on*. [S.l.], 2016. p. 1188–1194.
- SANTOS, A. D. Simulação médica baseada em realidade virtual para ensino e treinamento em ginecologia. Universidade Federal da Paraíba, 2010.
- SANTOS, A. D.; MACHADO, L. S.; MORAES, R. M.; GOMES, R. G. Avaliação baseada em lógica fuzzy para um framework voltado a construção de simuladores baseados em rv. In: *Anais do XII Symposium on Virtual and Augmented Reality. Natal: Sociedade Brasileira de Computação*. [S.l.: s.n.], 2010. p. 194–202.
- TEAM, R. C. *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013*. 2014.

WEBSTER, R. W.; ZIMMERMAN, D. I.; MOHLER, B. J.; MELKONIAN, M. G.; HALUCK, R. S. A prototype haptic suturing simulator. *Studies in health technology and informatics*, IOS Press; 1999, p. 567–569, 2001.

WEISS, P. L.; RAND, D.; KATZ, N.; KIZONY, R. Video capture virtual reality as a flexible and effective rehabilitation tool. *Journal of neuroengineering and rehabilitation*, BioMed Central, v. 1, n. 1, p. 12, 2004.

ZADEH, L. A. Fuzzy sets. *Information and control*, Elsevier, v. 8, n. 3, p. 338–353, 1965.

ZADEH, L. A. *Probability measures of fuzzy events*. [S.l.]: Elsevier, 1968. v. 23. 421–427 p.