

Universidade Federal do Rio Grande do Norte
Centro de Ciências Exatas e da Terra
Departamento de Informática e Matemática Aplicada
Programa de Pós-Graduação em Sistemas e Computação

**Analisando o desempenho do *ClassAge*: Um Sistema
Multiagentes para Classificação de Padrões**

Márjory Cristiany da Costa Abreu

Natal
Outubro de 2006

Universidade Federal do Rio Grande do Norte
Centro de Ciências Exatas e da Terra
Departamento de Informática e Matemática Aplicada
Programa de Pós-Graduação em Sistemas e Computação

Analisando o desempenho do *ClassAge*: Um Sistema Multiagentes para Classificação de Padrões

Dissertação de mestrado submetida ao Programa de Pós-Graduação em Sistemas e Computação do Departamento de Informática e Matemática Aplicada da Universidade Federal do Rio Grande do Norte como parte dos requisitos para a obtenção do grau de Mestre em Sistemas e Computação (MSc.).

Márjory Cristiany da Costa Abreu

Profa. Dra. Anne Magály de Paula Canuto
Orientador

Natal, Outubro de 2006

Divisão de Serviços Técnicos

Catálogo da Publicação na Fonte. UFRN / Biblioteca Central Zila Mamede

ABREU, Márjory Cristiany da

Analisando o desempenho do *ClassAge*: Um Sistema Multiagentes para Classificação de Padrões / Márjory Cristiany da Costa Abreu- Natal, RN, 2006.

133 f. : il.

Orientadora: Profa. Dra. Anne Magály de Paula Canuto.

Dissertação (Mestrado) – Universidade Federal do Rio Grande do Norte. Centro de Ciências Exatas e da Terra. Programa de Pós-graduação em Sistemas e Computação.

1. Sistemas Multiagentes – Dissertação. 2. Métodos de classificação – Dissertação. 3. Métodos de Negociação – Dissertação. 4. Sistemas multiclassificadores - Dissertação. 5. *Ensemble* – Dissertação. I. Canuto, Anne Magály de Paula. II. Título.

RN/UF/BCZM

CDU 004.78

Analizando o desempenho do *ClassAge*: Um Sistema Multiagentes para Classificação de Padrões

Márjory Cristiany da Costa Abreu

'Esta dissertação de mestrado foi avaliada e considerada aprovada pelo Programa de Pós-Graduação em Sistemas e Computação do Departamento de Informática e Matemática Aplicada da Universidade Federal do Rio Grande do Norte.'

Profa. Dra. Anne Magály de Paula Canuto
Orientador

Banca Examinadora:

Profa. Dra. Anne Magály de Paula Canuto
Presidente

Prof. Dr. André Carlos Ponce de Leon Ferreira de Carvalho

Prof. Dr. André Maurício Cunha Campos

Este trabalho é dedicado à *Francisco das Chagas Abreu* (painho) e *Maria Aparecida Costa Abreu* (mainha).

AGRADECIMENTOS

Eu não poderia começar os agradecimentos dessa dissertação de outra forma senão agradecendo a **Deus** e a **Nossa Senhora Aparecida**, que estão sempre presentes na minha vida e se manifestam da maneira mais maravilhosa, sempre me mostrando o caminho certo a seguir e sempre me dando forças, esperança e coragem para que eu realize meus sonhos.

Existem duas pessoas sem as quais a maioria das minhas realizações não seriam possíveis. Essas pessoas me ensinaram a ter responsabilidade e a jamais desistir dos meus sonhos e ideais, mesmo que o caminho para atingi-los parecesse extremamente difícil. Eles sempre estiveram ao meu lado, me dando força para que eu pudesse seguir a diante. Agradeço ao meu Pai, **Francisco das Chagas Abreu**, e à minha Mãe, **Maria Aparecida Costa Abreu**, por tudo o que vocês me fizeram, por tudo o que vocês me tornaram. OBRIGADA!

Agradeço a minha orientadora e amiga, professora **Anne Magaly de Paula Canuto**, por acreditar na minha capacidade de produção, embora, muitas vezes, nem eu mesmo acreditasse nisso. Agradeço pelo apoio, pelas broncas, pela cobrança, pela confiança, pela compreensão, enfim, por tudo. Sinto-me privilegiada por ter a honra de conhecer e trabalhar com uma profissional de tão grande competência e de tamanha grandeza. Sem dúvida esse trabalho não seria possível sem seu incentivo.

É possível realizar alguma coisa sem amor? É possível crescer sem uma pessoa para dividir as alegrias e as tristezas da vida? Quero agradecer ao meu amor, à pessoa que está ao meu lado desde de 2002 e que tem sido o apoio, que tem tido compreensão, que tem me aturado nos momentos mais estressantes, que tem estado comigo nos momentos mais felizes. Quero agradecer a **Carlos Eduardo da Silva** ou Carioca ou Kaduardo ou apenas Mô! Obrigado, meu amor! Por tudo que já vivemos juntos e por tudo que ainda vamos viver juntos, por que ainda temos um doutorado e o resto da vida pela frente.

Agradeço, também, aos meus amigos Ivanilson Ninho, Jorgiano, Fellipão pelos memoráveis almoços no MidWay e cafés no São Bras. Agradeço aos meus irmãos de orientadora Pow, Laura e Raul por terem tido a paciência de ler a minha dissertação. :D Agradeço aos meus amigos de DIMAp, além dos meus irmãos de orientadora: Fred, Girão, Bruno Cruz, Laurindo, Gilbran, Lucas, Isanio, Toba, Nazário, Jones pelas muitas horas de T.O. que me fizeram rir bastante depois de um dia cansativo de trabalho.

Agradeço a CAPES, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pelo apoio financeiro que foi indispensável para a conclusão deste trabalho.

Resumo

A utilização de sistemas baseados no paradigma dos agentes para resolução de problemas de reconhecimento de padrões vem sendo propostos com o intuito de resolver, ou atenuar, o problema de tomada de decisão centralizada dos sistemas multi-classificadores e, como consequência, melhorar sua capacidade de classificação. Com a intenção de solucionar este problema, o Sistema *NeurAge* foi proposto. Este sistema é composto por agentes neurais que podem se comunicar e negociar um resultado comum para padrões de teste. No Sistema *NeurAge*, os métodos de negociação são muito importantes para prover uma melhor precisão ao sistema, pois os agentes necessitam alcançar a melhor solução e resolver conflitos, quando estes existem, em relação a um problema. Esta dissertação apresenta uma extensão do Sistema *NeurAge* que pode utilizar qualquer tipo de classificador e agora será chamado de Sistema *ClassAge*. Aqui é feita uma análise do comportamento do Sistema *ClassAge* diante de várias modificações na topologia e nas configurações dos componentes deste sistema.

Área de Concentração: Inteligência Computacional

Palavras-chave: Sistemas Multiagentes; Métodos de Classificação; Métodos de Negociação; Sistemas Multi-Classificadores; *Ensemble*.

Abstract

The use of multi-agent systems for classification tasks has been proposed in order to overcome some drawbacks of multi-classifier systems and, as a consequence, to improve performance of such systems. As a result, the NeurAge system was proposed. This system is composed by several neural agents which communicate and negotiate a common result for the testing patterns. In the NeurAge system, a negotiation method is very important to the overall performance of the system since the agents need to reach and agreement about a problem when there is a conflict among the agents. This thesis presents an extensive analysis of the NeurAge System where it is used all kind of classifiers. This systems is now named ClassAge System. It is aimed to analyze the reaction of this system to some modifications in its topology and configuration.

Area of Concentration: Computational Intelligence

Key words: Multi-Agent Systems; Classification Methods; Negotiation Methods; Multi-Classifiers Systems; Ensemble.

Sumário

1	Introdução	1
1.1	Motivação	3
1.2	Objetivos	4
1.3	Organização do Trabalho	5
2	Estado da Arte	7
2.1	Sistemas Multiagente (SMA)	7
2.1.1	Sistema Multiagentes Reativos	9
2.1.2	Sistemas Multiagentes Cognitivos	10
2.2	Sistemas Multi-Classificador (SMC)	11
2.2.1	Sistemas Modulares	13
2.2.2	Sistemas Paralelos	14
2.3	Trabalhos Relacionados	15
2.3.1	Sistemas Multi-Classificadores	15
2.3.2	Sistemas Multiagentes	17
3	O Sistema <i>NeurAge</i> e Suas Características	23
3.1	Um Sistema Multiagente para Classificação	24
3.2	Agente Classificador do Sistema <i>NeurAge</i>	26
3.3	Métodos de Negociação	28

3.3.1	Método de Negociação Baseado na Teoria dos Jogos	29
3.3.2	Método de Negociação Baseado no Leilão	33
3.3.3	Método de Negociação Baseado na Sensibilidade	37
4	Detalhamento da Investigação	43
4.1	Primeiros Resultados do Sistema <i>NeurAge</i>	43
4.2	Metodologia do Trabalho	45
4.2.1	Validação Cruzada	45
4.2.2	Análise Estatística	46
4.2.3	Medidas de Diversidade	47
4.2.4	Métodos de Classificação	50
4.2.5	Bases de Dados	53
4.2.6	Métodos de Combinação	57
4.2.7	Métodos Baseados em Seleção	59
4.2.8	Topologias dos Sistemas	61
5	Experimentos e Análises de Resultados	63
5.1	Métodos Individuais	65
5.2	Sistemas com Três Componentes	65
5.2.1	Métodos de Combinação ou Fusão para Sistemas com Três Com- ponentes	66
5.2.2	Métodos Baseados em Seleção para Sistemas com Três Compo- nentes	68
5.2.3	Métodos de Negociação para Sistemas com Três Componentes . .	70
5.2.4	Negociação <i>versus</i> Combinação para Sistemas com Três Compo- nentes	72
5.2.5	Diversidade para Sistemas com Três Componentes	73
5.3	Sistemas com Cinco Componentes	74

5.3.1	Métodos de Combinação ou Fusão para Sistemas com Cinco Componentes	74
5.3.2	Métodos Baseados em Seleção para Sistemas com Cinco Componentes	76
5.3.3	Métodos de Negociação para Sistemas com Cinco Componentes	78
5.3.4	Negociação <i>versus</i> Combinação para Sistemas com Cinco Componentes	80
5.3.5	Diversidade para Sistemas com Cinco Componentes	81
5.4	Sistemas com Sete Componentes	82
5.4.1	Métodos de Combinação ou Fusão para Sistemas com Sete Componentes	83
5.4.2	Métodos Baseados em Seleção para Sistemas com Sete Componentes	85
5.4.3	Métodos de Negociação para Sistemas com Sete Componentes	87
5.4.4	Negociação <i>versus</i> Combinação para Sistemas com Sete Componentes	89
5.4.5	Diversidade para Sistemas com Sete Componentes	90
5.5	Sistemas com Nove Componentes	91
5.5.1	Métodos de Combinação ou Fusão para Sistemas com Nove Componentes	92
5.5.2	Métodos Baseados em Seleção para Sistemas com Nove Componentes	93
5.5.3	Métodos de Negociação para Sistemas com Nove Componentes	95
5.5.4	Negociação <i>versus</i> Combinação para Sistemas com Nove Componentes	97
5.5.5	Diversidade para Sistemas com Nove Componentes	98
5.6	Discussão dos Resultados	100
5.6.1	Tamanho dos Sistemas	100
5.6.2	Estrutura dos Sistemas	100

5.6.3	Métodos de Combinação	101
5.6.4	Métodos de Seleção	104
5.6.5	Métodos de Negociação	107
5.6.6	Sensibilidade nos Sistemas	110
5.6.7	Diversidade nos Sistemas	110
5.6.8	Negociação <i>versus</i> Combinação	112
6	Considerações Finais	115
6.1	Trabalhos Futuros	116
A		128

Lista de Figuras

2.1	Modelo de um Sistema Multi-Classificador Modular	14
2.2	Modelo de um Sistema Multi-Classificador Paralelo	14
3.1	Modelo de um Sistema Multi-Classificador Paralelo Tradicional	25
3.2	Eliminação do Módulo de Combinação do Sistema Multi-Classificador . .	25
3.3	Arquitetura Interna de Um Agente Classificador	27
5.1	Resultados dos Métodos de Fusão para base A (breast cancer)	101
5.2	Resultados dos Métodos de Fusão para base B (imagem)	101
5.3	Resultados dos Métodos de Fusão para base C (proteínas)	102
5.4	Resultados dos Métodos de Fusão para base D (splice)	102
5.5	Resultados dos Métodos de Fusão para base E (wisconsin)	103
5.6	Resultados dos Métodos de Seleção para base A (breast cancer)	104
5.7	Resultados dos Métodos de Seleção para base B (imagem)	104
5.8	Resultados dos Métodos de Seleção para base C (proteínas)	105
5.9	Resultados dos Métodos de Seleção para base D (splice)	105
5.10	Resultados dos Métodos de Seleção para base E (wisconsin)	106
5.11	Resultados dos Métodos de Negociação para base A (breast cancer) . . .	107
5.12	Resultados dos Métodos de Negociação para base B (imagem)	107
5.13	Resultados dos Métodos de Negociação para base C (proteínas)	108
5.14	Resultados dos Métodos de Negociação para base D (splice)	109

5.15	Resultados dos Métodos de Negociação para base E (wisconsin)	109
5.16	Medidas de Diversidade Dupla-Falsa para todas as bases	110
5.17	Medidas de Diversidade de Entropia para todas as bases	111
5.18	Medidas de Diversidade Q Estatístico para todas as bases	111
5.19	Comparação entre os Melhores Mérodo de Seleção, Negociação e Fusão para a base A (breast cancer)	112
5.20	Comparação entre os Melhores Mérodo de Seleção, Negociação e Fusão para a base B (imagem)	113
5.21	Comparação entre os Melhores Mérodo de Seleção, Negociação e Fusão para a base C (proteínas)	113
5.22	Comparação entre os Melhores Mérodo de Seleção, Negociação e Fusão para a base D (splice)	114
5.23	Comparação entre os Melhores Mérodo de Seleção, Negociação e Fusão para a base E (wisconsin)	114
A.1	Resumo dos Resultados para a base A (breast cancer)	129
A.2	Resumo dos Resultados para a base B (imagem)	130
A.3	Resumo dos Resultados para a base C (proteínas)	131
A.4	Resumo dos Resultados para a base D (splice)	132
A.5	Resumo dos Resultados para a base E (wisconsin)	133

Lista de Tabelas

3.1	Matriz de <i>Payoff</i>	30
3.2	Matriz de <i>Payoff</i> para o <i>NeurAge</i>	30
3.3	Confiabilidades dos dois Agentes.	32
3.4	Aplicação da Matriz de <i>Payoff</i> para o <i>NeurAge</i>	33
3.5	Matriz de Custo do Leilão.	36
3.6	Análise de Sensibilidade dos Agentes.	41
3.7	Média de Treinamento para Todos os Atributos.	41
3.8	Distância absoluta entre os atributos do padrão de teste e da média de treinamento.	41
3.9	Ordem decrescente da distância entre os atributos.	42
4.1	Resultados do Trabalho (Abreu, 2005).	44
4.2	Características das Bases de Dados deste trabalho.	54
4.3	Representação Numérica da Base de Dados de Imagem.	55
4.4	Distribuição de padrões da Base de Dados de Proteínas.	56
5.1	Porcentagem dos Acertos e Desvio Padrão dos Métodos Individuais. . . .	65
5.2	Método baseado em Fusão para Sistemas com Três Componentes.	67
5.3	Método baseado em Seleção para Sistemas com Três Componentes.	69
5.4	Método de Negociação para Sistemas com Três Componentes.	71
5.5	Valores de p para os testes estatísticos entre os melhores métodos de ne- gociação e os de fusão e seleção para sistemas com Três Componentes. .	72

5.6	Medidas de Diversidade para Sistemas com Três Componentes.	73
5.7	Método baseado em Fusão para Sistemas com Cinco Componentes. . . .	75
5.8	Método baseado em Seleção para Sistemas com Cinco Componentes. . .	77
5.9	Método de Negociação para Sistemas com Cinco Componentes.	79
5.10	Teste de Hipótese entre os melhores métodos de seleção e fusão com o melhor método de negociação para sistemas com Cinco Componentes. . .	80
5.11	Medidas de Diversidade para Sistemas com Cinco Componentes.	81
5.12	Método baseado em Fusão para Sistemas com Sete Componentes.	84
5.13	Método baseado em Seleção para Sistemas com Sete Componentes. . . .	86
5.14	Método de Negociação para Sistemas com Sete Componentes.	88
5.15	Teste de Hipótese entre os melhores métodos de seleção e fusão com o melhor método de negociação para sistemas com Sete Componentes. . . .	89
5.16	Medidas de Diversidade para Sistemas com Sete Componentes.	90
5.17	Método baseado em Fusão para Sistemas com Nove Componentes.	93
5.18	Método baseado em Seleção para Sistemas com Nove Componentes. . . .	94
5.19	Método de Negociação para Sistemas com Nove Componentes.	96
5.20	Teste de Hipótese entre os melhores métodos de seleção e fusão com o melhor método de negociação para sistemas com Nove Componentes. . .	98
5.21	Medidas de Diversidade para Sistemas com Nove Componentes.	99

Lista de Siglas

- ClassAge: *Classifier Agent System*.
- DCS: *Dynamic Classifier Selection*.
- DCS-DT: *Dynamic Classifier Selection - Decision-Theoretic*.
- DCS-MCB: *Dynamic Classifier Selection - Multiple Classifier Behaviour*.
- FuzzyMLP: *Fuzzy Multi-Layer Perceptron*.
- IREP: *Incremental Reduced Error Pruning*.
- JRip: *Optimized IREP*.
- K-nn: *K-nearest neighbors*.
- LCA: *Local Class Accuracy*.
- MLP: *Multi-Layer Perceptron*.
- NeurAge: *Neural Agent System*.
- RBF: *Radial Basis Function*.
- SMA: *Sistemas Multi-Agente*.
- SMC: *Sistemas Multi-Classificadores*.
- SVM: *Support Vector Machine*.

Capítulo 1

Introdução

A necessidade dos pesquisadores que trabalham com mineração de dados de terem um sistema computacional que trabalhe com o reconhecimento de padrões de forma eficiente, precisa e com alto desempenho motivou o interesse pelo estudo na área de métodos de reconhecimento de padrões (Cho, 1999). Apesar desses sistemas conseguirem resultados satisfatórios, eles não estavam, muitas vezes, resolvendo as tarefas de reconhecimento de padrões de maneira eficiente. Afim de obter um melhor desempenho de um sistema de reconhecimento de padrões, utilizam-se vários tipos de classificadores trabalhando conjuntamente, dando origem ao termo Sistema Multi-Classificador (SMC) (Ho et al., 1994).

Os Sistemas Multi-Classificadores podem ser de dois tipos: Modular e Paralelo (Mitchell, 1997). O primeiro acontece quando cada um dos classificadores ou um conjunto de classificadores é responsável por solucionar uma parte do problema e depois enviar suas respostas para um método de combinação comum. Já os Paralelos funcionam de maneira que todos os classificadores resolvem o mesmo problema e enviam suas respostas, mais uma vez, para um combinador comum. Neste trabalho, a abordagem paralela será analisada.

Dentre as aplicações mais recentes dos SMCs existentes na literatura podem-se citar como exemplo, o reconhecimento de fala (Mukhopadhyay et al., 2003), reconhecimento de caracteres (Heutte et al., 2004), reconhecimento de intenção (objetivo) (Zhou, 2002), reconhecimento de faces (Lemieux and Parizeau, 2003), reconhecimento de emoção (Petrushin, 2000), reconhecimento de sinais em ambiente de internet (Park et al., 2002), biométricas (Shou-Jue and Xu, 2003), classificação de regras (Sirlantzis et al., 2002a), reconhecimento de imagem (Giacinto and Roli, 2001), classificação de proteínas (Tan et al., 2003), dentre outros.

Na abordagem paralela, a escolha do método combinador, ou seja, o método que realiza a tomada de decisão do sistema, é muito complicada, pois pequenas alterações na

estrutura do sistema podem resultar em relevantes alterações no seu desempenho (Abreu et al., 2004). O fato de um SMC possuir um método de tomada de decisão centralizado faz com que esse tipo de sistema se torne dependente pois, independentemente dos classificadores que o sistema estará trabalhando, a tomada de decisão sempre dependerá de um único método de combinação (Abreu and Canuto, 2006).

Com o objetivo de solucionar o problema de tomada de decisão centralizada dos SMC, de maneira mais natural, tentou-se encontrar uma forma que tornasse o funcionamento e a arquitetura desses sistemas mais flexíveis. Diante de uma primeira análise, pode-se considerar um SMC como sendo um Sistema Multiagentes sem uma estrutura organizacional pré-definida, onde cada classificador é considerado um agente e o módulo combinador é, também, um agente que, sozinho, encontrava a solução do sistema a partir dos resultados de cada agente.

A idéia mais simples para descentralizar o processo de tomada de decisão centralizada dos SMCs vistos, agora, como SMAs foi retirar o método de combinação desse sistema. Para isso, os classificadores precisariam se comunicar. Diante disso, observou-se que a comunicação entre os classificadores poderia resolver este problema, uma vez que se eles se comunicam, eles podem concluir, mais precisamente, o resultado do sistema. A melhor maneira é transformar os Sistemas Multi-Classificadores em Sistemas Multiagentes (Abreu et al., 2006a).

Na nova visão dos SMC, os classificadores podem se comunicar e decidir a solução do sistema sem a necessidade de consultar o módulo de tomada de decisão. Sendo assim, agora eles são vistos como entidades autônomas que são capazes de tomar decisões, analisar as outras entidades e compartilhar informações. Logo, observou-se que a solução para isto era utilizar Agentes Inteligentes, pois estes já possuem todas estas características. Sendo assim, a utilização dos conceitos dos SMA para a resolução de problemas de reconhecimento de padrões ficou sendo viável e um novo termo foi criado, Sistemas Multiagente para Classificação (Abreu et al., 2004). O sistema agora é mais dinâmico, autônomo, distribuído e possui um desempenho mais eficaz.

O Sistema *NeurAge* (*Neural Agents System*), um Sistema Multiagente Neural para Classificação (Abreu et al., 2004), (Abreu, 2005) foi proposto para utilizar o conceito de agentes trabalhando com reconhecimento de padrões. Este sistema possui vários agentes com a capacidade de realização de reconhecimento de padrões. Inicialmente, eles recebem uma entrada do ambiente e cada um, individualmente, calcula sua saída. Após esta etapa, cada um dos agentes compara sua saída com os demais, tentando contestar as saídas dos outros, se essas saídas não forem iguais a dele. Quando acontece uma discordância, o agente que está negociando tenta persuadir os demais agentes a mudarem de opinião. A negociação termina quando os agentes conseguem chegar a uma opinião comum.

Diante do que foi explicado anteriormente, observa-se que existe a necessidade de comunicação entre os agentes do *NeurAge*. Isto faz com que o uso de protocolos de negociação seja necessário para o funcionamento do sistema (Wooldridge, 2002). Sendo assim, os mecanismos de negociação são parte importante para o bom funcionamento destes sistemas. É importante, também, que esses protocolos de negociação sejam eficientes e que possam trabalhar com todas as informações que os agentes fornecem, tornando o sistema mais seguro e confiável (Dunne et al., 2003).

Este trabalho propõe explorar melhor o Sistema *NeurAge*, abrangendo extensões tanto a nível de componentes do sistema quanto a nível de protocolos de negociação. Uma das extensões é o uso de outros tipos de classificadores além dos neurais. Com isso esse sistema passa a ser conhecido como o Sistema *ClassAge* (*Classifier Agents System*), um sistema multiagente para reconhecimento de padrões que pode conter qualquer tipo de classificador em sua arquitetura.

1.1 Motivação

Na proposta inicial do Sistema *NeurAge*, foram utilizadas as configurações de sistema mostradas a seguir:

- Todos os teste do Sistema *NeurAge* inicial foram realizados com apenas agentes neurais contendo o apenas o método de classificação MLP (*Multi-Layer Perceptron*) (Park et al., 2002);
- O tamanho do sistema era fixado em quatro agentes;
- Os três métodos de negociação que foram adaptados neste trabalho foram o Baseado na Teoria dos Jogos, o Baseado no Leilão e Baseado na Sensibilidade; e
- Foram utilizadas três bases de dados, duas delas balanceadas e com quantidade de padrões suficiente para um treinamento adequado e a terceira era desbalanceada com uma menor quantidade de padrões para um treinamento adequado.

Os resultados dos métodos de negociação foram comparados com os resultados de dois métodos de combinação centralizados: Soma e Voto. Além disso, foi utilizado o teste de hipótese para fazer uma comparação estatística entre os métodos. Para os métodos Soma e Voto obteve-se sempre um desempenho superior ao desempenho dos métodos individuais, mas inferior aos métodos de negociação. Dentre os métodos de negociação, o que obteve o melhor desempenho para todos os experimentos foi o método baseado na sensibilidade.

Apesar dos resultados serem satisfatórios a uma primeira vista, pode ser observado que a pesquisa não se aprofundou o suficiente nas diversas possibilidades de cada um dos parâmetros citados anteriormente. Isso serviu como motivação para este trabalho, de forma que está sendo proposto o uso desse sistema de forma mais aprofundada. Acredita-se que os resultados a serem obtidos serão ainda melhores que os atuais.

1.2 Objetivos

Como foi mostrado anteriormente, o principal objetivo desse trabalho é investigar o desempenho do Sistema *ClassAge* diante de várias modificações do seus principais parâmetros. Logo, este sistema deve ser submetido a diversas situações com o intuito de observar a sua reação diante das modificações propostas. A seguir serão citados os sub-objetivos propostos por este trabalho.

- Como primeiro sub-objetivo, analisar-se-á outros tipos de classificadores, pois os trabalhos realizados em (Abreu, 2005) e (Abreu et al., 2004) utilizaram apenas classificadores neurais, mais especificamente, Redes Neurais MLP (*Multi-Layer Perceptron*). Neste trabalho, serão utilizados as Árvores de Decisão, as Máquinas de Vetores Suportes, as Redes RBF (*Radial Basis Function*), os k-vizinhos mais próximos, o método JRip e a Rede Fuzzy MLP.
- Outro sub-objetivo proposto é trabalhar com sistemas de tamanho variado, porque em (Abreu, 2005) e (Abreu et al., 2004) foi utilizado apenas um Sistema Multiagentes com quatro agentes. Neste trabalho, serão feitos experimentos com três, cinco, sete e nove agentes para que se possa observar a reação do sistema a essa variação na sua quantidade de agentes. Além disso, será observado o comportamento do sistema quando se tem variações nos tipos de classificadores.
- Mais um sub-objetivo é comparar os métodos de negociação com outros métodos de combinação. Isso é importante, pois uma boa maneira de avaliar o desempenho dos métodos de negociação aplicados ao reconhecimento de padrões é fazer comparações com Sistemas Multi-Classificador. Em (Abreu, 2005) e (Abreu et al., 2004) os métodos de negociação só foram comparados a Soma e a Votação. Neste trabalho, além do Voto e da Soma, serão utilizados outros métodos combinadores, como, por exemplo, Redes MLP, Redes Fuzzy MLP, a Média, a Mediana e o *Naive Bayes*.
- Um outro sub-objetivo é comparar os métodos de negociação com métodos baseados em seleção. Da mesma maneira que é importante comparar-se os métodos de

negociação com os métodos de combinação; é importante comparar também com os métodos de seleção. No trabalho inicial não foi feita nenhuma comparação com métodos de seleção. Neste trabalho serão utilizados os métodos DCS's (*Dynamic Classifier Selection*) com três variações (DCS-MCB, DCS-LA e DCS-DT).

- E, como último sub-objetivo, serão utilizadas outras bases de dados. Em (Abreu, 2005) e (Abreu et al., 2004) foram utilizadas três bases de dados para avaliação. Uma delas era desbalanceada e as outras duas eram balanceadas, porém todas possuíam um grande número de exemplos. Aqui, nesta dissertação, serão trabalhadas cinco bases de dados, pois assim, pode-se observar se o desempenho deste sistema é dependente dos dados ou não.

Em suma, esta dissertação tem como principal objetivo fazer uma análise mais detalhada do Sistema *NeurAge*. Como consequência desta análise, este sistema passará a ser conhecido como *ClassAge*, pois o mesmo englobará agentes classificadores, podendo ser neurais ou não.

1.3 Organização do Trabalho

Nesta dissertação, existem seis capítulos os quais estão organizados da maneira exposta a seguir.

- No capítulo 1, uma breve introdução do que será mostrado, a motivação e os objetivos do trabalho são feitos;
- No capítulo 2, o embasamento teórico do trabalho é exposto, além de fazer uma mostra do estado da arte na área de pesquisa em que o trabalho se encontra, apresentando os principais trabalhos que foram feitos;
- No capítulo 3, uma explanação sobre o Sistema *NeurAge* e de suas principais características é feita. Além disso, é apresentada uma explicação de cada método de negociação utilizado, assim como as adaptações sofridas por estes métodos;
- No capítulo 4, a proposta principal do trabalho é mostrada, assim como também, as etapas de desenvolvimento, a metodologia dos experimentos e uma breve explicação de cada etapa;
- No capítulo 5, os resultados são apresentados, assim como são apresentadas análises destes para cada uma das bases de dados escolhidas, assim como as configurações que foram utilizadas;

- No capítulo 6, um comentário é feito sobre trabalho de forma geral e as suas considerações;
- E, por fim, é apresentada a bibliografia deste trabalho.

Capítulo 2

Estado da Arte

O estudo dos Sistemas Multiagentes (SMA) para classificação é considerado um tópico recente da Inteligência Artificial, pois mistura conceitos de Sistemas Multiagentes com conceitos de Sistemas Multi-Classificadores (SMC) para realizar tarefas de reconhecimento de padrões. É possível encontrar alguns trabalhos relacionados a esse tópico, mas que utilizam abordagens diferentes. Apesar disso, nenhum desses trabalhos utiliza conceitos de negociação como solução deste problema.

Neste capítulo serão introduzidos os principais conceitos relacionados a Sistemas Multiagentes e Sistemas Multi-classificadores. Além disso, serão citados e comentados alguns dos principais trabalhos sobre estes tópicos e os trabalhos mais recentes que utilizam SMA para reconhecimento de padrões.

As seções 2.1 e 2.2 apresentam o embasamento teórico necessário para explicar o modelo de Agente Classificador. Por fim, na Seção 2.3 é feita uma apresentação sobre os principais trabalhos que existem na literatura e utilizam os agentes para resolver problemas de reconhecimento de padrões.

2.1 Sistemas Multiagente (SMA)

Não existe uma definição para agente que seja aceita por toda a comunidade de Inteligência Artificial. Uma definição inicial proposta por (Ferber and Gasser, 1991), é que um agente é uma entidade real ou abstrata que é capaz de agir sobre ela mesma e sobre seu ambiente. Em geral, o agente dispõe de uma representação parcial deste ambiente em um universo multiagentes e ele pode se comunicar com os demais agentes. O comportamento desse agente é consequência de suas observações, de seu conhecimento e das interações com os demais agentes. Ou seja, um Agente Inteligente é um sistema computacional que

está em um ambiente definido, onde é capaz de realizar ações autônomas neste ambiente para que um determinado objetivo seja alcançado (Wooldridge, 2002).

As definições propostas anteriormente preocupam-se com os mecanismos internos para o tratamento de cada agente, não estabelecendo o tipo de comunicação possível nem a granularidade dos agentes. Em (Rezende, 2002), tem-se uma definição que ressalta o aspecto da identidade de cada agente onde um agente é uma entidade que pode ser associada a uma identidade única e que é capaz de realizar tarefas formalizadas. Um agente pode ser considerado como um meio que produz um certo número de ações a partir dos conhecimentos e mecanismos internos que lhe são próprios.

Denomina-se interação entre agentes ou entre agente/ambiente uma troca de informações, que pode ser realizada de forma direta (comunicação explícita) ou de modo indireto (emissão de sinais através do ambiente). Uma organização define todas as restrições aplicadas aos agentes pertencentes a uma determinada sociedade, ou seja, os meios dos quais o projetista do sistema pode garantir que cada agente desejará e realizará a resolução dos problemas propostos (Petrushin, 2000).

A partir disso, dado um determinado sistema, denomina-se agente cada uma de suas entidades ditas ativas. Este conjunto de agentes forma uma sociedade. As entidades passivas serão designadas pelo termo ambiente. Um agente recebe informações e raciocina sobre o ambiente, sobre outros agentes e decide quais ações deve realizar e quais objetivos deve seguir. Um agente é uma entidade ativa, ou seja, capaz de controlar suas ações, diferentemente das noções estáticas tais como módulos, conjunto de regras e bases de conhecimentos (Toda et al., 2001).

Um conjunto de ações pré-definidas é associado a cada um dos agentes e essas ações representam a sua capacidade de modificar o ambiente. O grande dilema é o agente decidir qual ação será mais satisfatória para que o seu objetivo seja alcançado (Vuurpijl and Schomaker, 1998a). De forma geral, as principais capacidades que um sistema deve possuir para ser considerado um Agente Inteligente são: reatividade, proatividade e habilidade social.

- Reatividade é a capacidade de perceber o ambiente e responder de forma oportuna às mudanças que ocorrem nele afim de satisfazer algum objetivo.
- Proatividade é a capacidade de possuir comportamento dirigido diretamente a um determinado objetivo. O agente faz um exame da iniciativa e pondera se essa ação ajudará de forma satisfatória o cumprimento dos seus objetivos.
- Habilidade Social é a capacidade de interagir com o ambiente e com outros agentes a fim de satisfazer seus objetivos.

Os Sistemas Multiagentes (SMA) são sistemas que possuem vários agentes, cada um com seu objetivo ou função. Eles trabalham, em geral, por um objetivo comum e esse objetivo comum não necessariamente é igual aos objetivos individuais de cada agente. Pelo fato de assumir-se que eles são autônomos, eles devem ser capazes de coordenar dinamicamente suas atividades e cooperar uns com os outros para que o objetivo do sistema seja alcançado (Raje et al., 1998).

Pode-se definir um sistema como sendo multiagente quando ele possui determinadas características: um ambiente, um conjunto de agentes, um conjunto de objetos, as relações que podem existir entre esses agentes e os objetos ou entre os próprios agentes e um conjunto de operações que podem ser realizadas (Li et al., 2005b).

Em relação ao tipo de agente contido no sistema, os SMA podem ser caracterizados didaticamente em duas classes, que serão apresentadas a seguir. A primeira denomina-se Sistemas Multiagentes Reativos e trabalha com o desenvolvimento de sistemas que utilizam um grande número de agentes simples para a resolução de um determinado problema. A segunda abordagem, denominada Sistemas Multiagentes Cognitivos trabalha com poucos agentes que realizam tarefas mais complexas que os primeiros.

2.1.1 Sistema Multiagentes Reativos

Em geral, SMAs reativos são baseados em modelos de organização biológica ou etológica (formigas, cupins, abelhas, etc.). O modelo de funcionamento de um agente reativo é formado pelo par Estímulo-Resposta, ou seja, Ação-Reação. As principais características deste tipo de agente e dos sistemas onde eles se encontram são destacadas a seguir e podem ser encontradas mais detalhadamente em (Ferber and Gasser, 1991).

- Não há representação explícita do conhecimento: o conhecimento dos agentes é implícito (as suas regras de comportamento) e sua manifestação se externa através do seu comportamento em interação com os demais agentes;
- Não há representação interna do ambiente: o comportamento (resposta) de cada agente é baseado no que ele percebe (estímulo) a cada instante. Não há uma representação interna explícita do ambiente;
- Não há memória das ações: os agentes reativos não mantêm nenhum tipo de histórico de suas ações, ou seja, o resultado de uma determinada ação passada não influencia diretamente na decisão de uma ação futura;
- Organização etológica: a forma de organização dos SMA reativos é similar à observada por animais que vivem em grandes comunidades;

- Grande número de membros: em geral, os SMA reativos possuem um grande número de agentes, com populações que podem chegar à ordem de milhares de membros.

Em suma, os agentes reativos são muito simples e não possuem representação do seu ambiente. Suas reações dependem unicamente de sua percepção deste ambiente. Na Inteligência Artificial clássica, por exemplo, define-se um problema de uma maneira global, criando métodos de resolução que se aplicam diretamente e exclusivamente sobre esta definição. Os modelos de SMAs Reativos, por outro lado, concebem o problema como sendo um conjunto de agentes interagindo entre si, onde cada um destes possui seus próprios objetivos individuais. Uma forma usual de representar os comportamentos dos agentes é por meio de um conjunto de regras.

2.1.2 Sistemas Multiagentes Cognitivos

Em geral, os SMAs cognitivos são baseados em modelos de organização social, de sociedades humanas: grupos, hierarquias, mercados etc. Esses Agentes Cognitivos, segundo (Ferber and Gasser, 1991), possuem uma representação explícita do ambiente e dos membros da comunidade e podem raciocinar sobre as ações tomadas no passado e planejar as futuras ações. Os Agentes Cognitivos podem ainda interagir com os demais membros da comunidade através de linguagens e protocolos de comunicação complexos, estratégias sofisticadas de negociação etc. As principais características dos Sistemas Multiagentes Cognitivos são citadas a seguir.

- Representação explícita do ambiente e dos outros agentes da sociedade;
- Podem manter um histórico das interações e ações passadas e, graças a esta memória, são capazes de planejar suas ações futuras;
- Seu sistema de percepção, que permite examinar o ambiente, e o de comunicação, que permite a troca de mensagens entre agentes, são distintos. A comunicação entre agentes pode ser realizada de modo direto, através do envio e recebimento de mensagens;
- Seu mecanismo de controle é deliberativo. Os agentes cognitivos raciocinam e decidem em conjunto sobre quais ações devem ser executadas, que planos seguir e que objetivos devem alcançar;
- Os modelos de organização dos SMA cognitivos, em geral, são modelos sociológicos, como, por exemplo, as organizações humanas;

- Um SMA cognitivo contém, usualmente, poucos agentes, na ordem de algumas dezenas, no máximo.

Esses agentes normalmente apresentam elevada complexidade computacional e são caracterizados por apresentar um comportamento inteligente tanto em uma Comunidade de Agentes quanto isoladamente. Estas comunidades geralmente são compostas por um pequeno número de participantes. Os agentes autônomos possuem um alto grau de determinação, eles podem decidir por motivações próprias, quando e sob que condições suas ações devem ser tomadas. Em muitos casos estes agentes precisam interagir com outros agentes autônomos para atingir seus objetivos, por não possuírem habilidades ou recursos suficientes para solucionar seus problemas sozinhos ou ainda pela interdependência em relação a outros agentes.

Os objetivos das interações entre agentes são para fazer outros agentes tomarem suas decisões conscientes do que podem causar aos demais agentes (como por exemplo executar um serviço em particular), modificar uma linha de ação planejada, ou ainda atingir um acordo sobre ações conjuntas. Uma vez que estes agentes não possuem um controle direto sobre os outros faz-se necessário utilizar uma estratégia de cooperação para aglutinar outros agentes autônomos na realização de uma dada tarefa formando assim um sistema multiagente para solução de problemas através de ação cooperativa (Mukhopadhyay et al., 2003).

De forma geral, pode-se dizer que os agentes cognitivos, diferentemente dos reativos, possuem uma representação explícita e detalhada, além de possuir um histórico no qual ele se baseia para tomar suas decisões. Por outro lado, os agentes reativos, diferentemente dos cognitivos, possuem uma comunicação indireta e um controle não deliberativo, fazendo com que seja necessário uma quantidade considerável de agentes no sistema para produzir, de forma emergente, um comportamento complexo.

2.2 Sistemas Multi-Classificador (SMC)

Com o crescimento da pesquisa na área de reconhecimento de padrões, o uso de classificadores individuais se tornou pouco eficiente na maioria dos problemas complexos (Ho et al., 1994). Diante disso, sistemas de classificação baseados na combinação de saídas de um conjunto de diferentes classificadores vêm sendo usados como solução e estão alcançando alta eficácia (Roli and Giacinto, 2002). Sistemas Multi-Classificador (SMC) são sistemas compostos de vários classificadores.

A combinação de classificadores explora a idéia que diferentes classificadores, especialistas ou módulos de reconhecimento podem oferecer informação complementar sobre

padrões a serem classificados (Canuto, 2001). Isso facilita o trabalho do módulo que faz o papel de combinador, pois ele tem algumas opções formadas através de diferentes ângulos de observação sobre o mesmo padrão de entrada.

Como já foi dito, um SMC é formado por diferentes classificadores, ou algoritmos de classificação, e uma função para a combinação das saídas desses classificadores. Estes sistemas funcionam recebendo um dado externo, processando-o e emitindo uma resposta. Esta resposta é direcionada a um classificador especial que executa a combinação destas respostas e toma a decisão de qual é a melhor solução para o sistema, considerando todas as informações fornecidas pelos demais classificadores (Cho, 1999).

A condição suficiente e necessária para um SMC ser mais eficaz que qualquer de seus membros, além da escolha adequada do método de combinação, é que seus classificadores sejam eficazes e diversificados (Tan et al., 2003). A diversidade entre algoritmos de classificação é somente a diferença entre os seus erros produzidos. Esta diversidade pode ser medida entre dois classificadores ou entre conjuntos de classificadores, por exemplo. Um classificador eficaz é aquele que possui uma taxa de acerto melhor que uma simples decisão aleatória. Existem três razões que justificam que é possível construir bons SMC, ou seja, classificadores com pequenas diversidades (Yang et al., 2004).

- A primeira delas é estatística. Um algoritmo de aprendizado pode ser visto como a busca, dentro de um espaço de hipóteses, para identificar a melhor resposta. O problema estatístico começa quando o conjunto de dados de treinamento é muito pequeno quando comparado com o tamanho do espaço hipotético. Sem dados suficientes, o algoritmo pode encontrar muitas hipóteses diferentes, mas que possuem a mesma exatidão (Smits, 2002). Construindo um SMC diferente dos classificadores exatos, o algoritmo pode fazer uma média das saídas e reduzir o risco da escolha da classificação errada.
- A segunda razão é computacional. Muitos algoritmos de aprendizado trabalham baseado em buscas locais que devem culminar em um ótimo local. Por exemplo, redes neurais empregam o gradiente descendente para minimizar uma função de erro sobre o conjunto de treinamento. Em todos os casos, se existe um conjunto de treinamento suficiente, isto faz com que seja difícil, computacionalmente, para o algoritmo encontrar a melhor hipótese (Sirlantzis et al., 2002b). Um SMC construído para fazer uma busca local iniciada por muitos diferentes pontos deve prover uma melhor aproximação da função real desconhecida que qualquer um dos classificadores individuais.
- A terceira razão é representacional. Em muitas aplicações de aprendizado de máquina, a função real do problema não pode ser representada por nenhuma das hi-

póteses. A representação é complicada porque existem muitos algoritmos no qual o espaço de hipóteses é possível para todos os classificadores ao mesmo tempo (Sirlantzis and Fairhurst, 2001). Sendo assim, dado um conjunto de treinamento suficiente, os algoritmos irão explorar o espaço de todos os possíveis classificadores.

Como foi citado anteriormente, um SMC possui um mecanismo de combinação que recebe as saídas dos classificadores individuais e decide qual será a saída do sistema. A escolha desse mecanismo nem sempre é trivial e influencia diretamente o desempenho final do sistema. As características presentes no sistema criado serão diretamente dependentes das características do problema a ser atacado (Kang and Lee, 2000). Esse mecanismo pode ser uma única função em arquiteturas paralelas ou múltiplas funções para uma arquitetura serial. A existência do módulo de combinação, que necessita de todas as saídas dos classificadores para poder gerar a saída final, pode tornar o sistema pouco flexível, pois torna a tomada de decisão um processo centralizado.

Existem alguns tipos de Sistemas Multi-Classificador que podem ser citados. De acordo com a sua arquitetura interna, no mais alto nível de abstração, eles podem ser Modular ou Paralelo (*Ensemble* ou Comitês) (Lemieux and Parizeau, 2003). Eles serão explicados nas seções 2.2.1 e 2.2.2.

2.2.1 Sistemas Modulares

O termo modular é usado quando uma tarefa é decomposta em um número de sub-tarefas e a solução da tarefa completa requer a contribuição de todos os módulos do sistema.

Neste caso, os SMCs Modulares são aqueles em que existem vários especialistas dentro do sistema e cada um deles executa uma tarefa em uma determinada área diferente. Eles trabalham dividindo a tarefa a ser realizada em partes e cada uma dessas partes é responsabilidade de um grupo de classificadores. A figura 2.1 mostra como seria a disposição dos classificadores nessa abordagem.

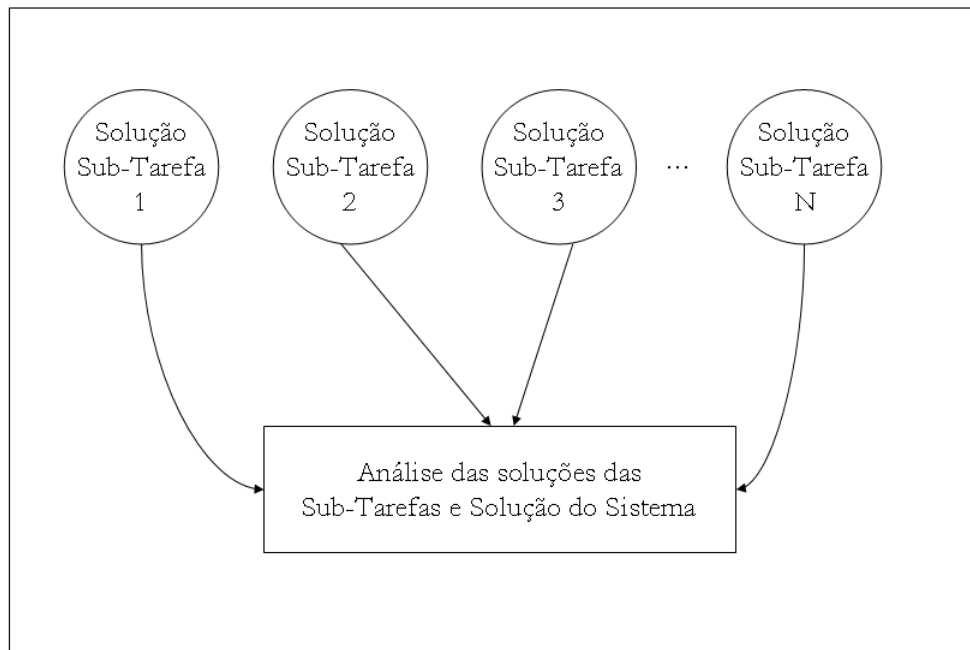


Figura 2.1: Modelo de um Sistema Multi-Classificador Modular

2.2.2 Sistemas Paralelos

O termo paralelo, também conhecido como *Ensemble* ou Comitês é usado para a combinação redundante de um conjunto de classificadores. Os componentes classificadores são redundantes de maneira que cada classificador provê a solução para a mesma tarefa.

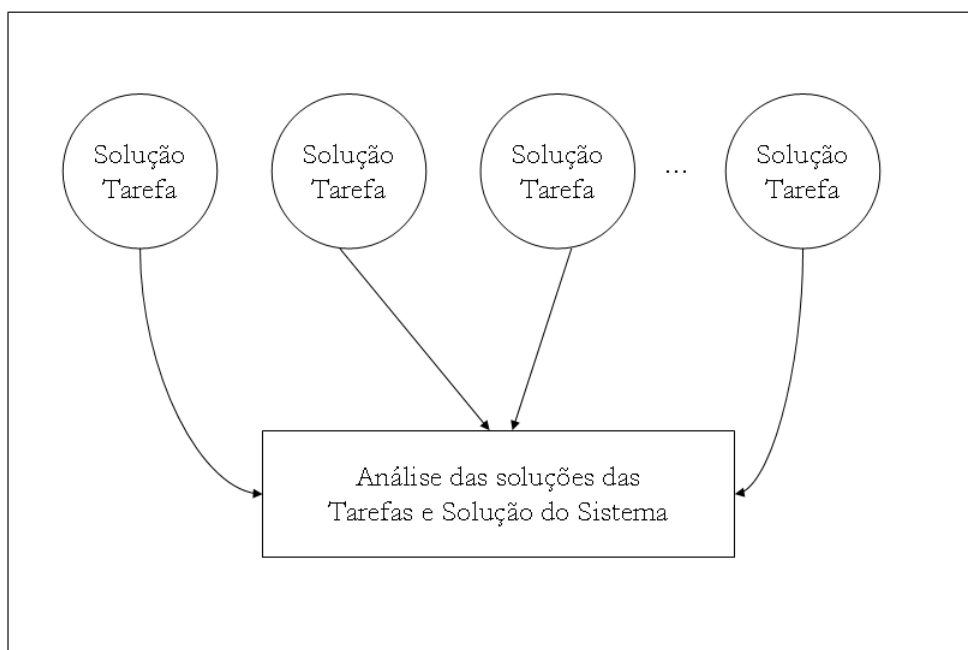


Figura 2.2: Modelo de um Sistema Multi-Classificador Paralelo

Neste caso, os SMC Paralelos são aqueles em que todos os especialistas podem, cada um de maneira diferente, responder ao mesmo estímulo externo. Não existe divisão de tarefas. Todos os classificadores possuem a mesma responsabilidade. A figura 2.2 mostra como seria a disposição dos classificadores nessa abordagem.

2.3 Trabalhos Relacionados

Existem muitos trabalhos relacionados com SMA, SMC e Sistemas Multiagentes para reconhecimento de padrões que podem ser citados.

Dentro da abordagem desta dissertação, ou seja os trabalhos que envolvem Sistemas Multiagentes para classificação, existem duas linhas básicas: aqueles que subdividem as tarefas entre os agentes, de forma que cada agente é responsável por uma tarefa diferente, similares a SMC Modular e aqueles em que todos os agentes possuem a mesma tarefa, similares a *Ensemble* ou Comitês.

Na Seção 2.3.1 são mostrados os trabalhos mais recentes sobre Sistemas de múltiplos classificadores. Na Seção 2.3.2 são mostrados os trabalhos mais recentes sobre Sistemas Multiagentes, onde serão também mostrados os principais trabalhos relacionados com Sistemas Multiagentes para Classificação.

2.3.1 Sistemas Multi-Classificadores

Como já foi mencionado anteriormente, os Sistemas Multi-Classificadores podem ser divididos em dois tipos: Sistema Modulares e Sistemas Paralelos. Nas seções 2.3.1.1 e 2.3.1.2 são apresentados os trabalhos mais recentes e relevantes para esta dissertação que estão relacionados com os dois tipos de SMC.

2.3.1.1 Sistemas Classificadores Modulares

Como já foi falado anteriormente, Sistemas Multi-classificadores Modulares são aqueles em que cada agente, ou conjunto de agentes, possui uma determinada tarefa diferente das demais. Dentre os principais trabalhos existentes na literatura relacionados com esta abordagem, pode-se citar, por exemplo, (Shou-Jue and Xu, 2003), (Brewer et al., 2003), (Heutte et al., 2004), (Vuurpijl and Schomaker, 1998b), (Mukhopadhyay et al., 2003), (Peng et al., 2001), (Fairhurst et al., 1989), (Sirlantzis et al., 2002a), dentre outros.

Em (Shou-Jue and Xu, 2003), por exemplo, um novo modelo teórico dos princípios do Reconhecimento de Padrões foi proposto. Este modelo é baseado na importância da cognição relacionada com a importância da classificação no Reconhecimento de Padrões estatístico tradicional. Além disso, ele é mais próximo das funções humanas que o Reconhecimento de Padrões tradicional que utiliza a separação ótima como princípio básico. Aqui ele utiliza os sistemas modulares para trabalhar com essa distinção de tarefas. Então esse novo modelo é chamado de *Biomimetic Pattern Recognition (BPR)*. Sua matemática básica é aplicada na análise topológica de um conjunto de exemplos em um alto espaço dimensional.

Já em (Brewer et al., 2003) um sistema de classificação é utilizado para analisar os bancos de dados que armazenam dados colhidos de florestas como vegetação, animais, dentre outras coisas. Os componentes desse sistema incluem uma estratificação ecogeográfica, a produção de objetos de imagens através da segmentação de imagens, a incorporação de dados de imagem multi-temporal e detecção de mudanças e a utilização de múltiplos classificadores para diferentes níveis de hierarquia de classificação. Neste sistema, é criado um classificador para reconhecer cada região ou característica deste banco de dados.

Ainda nesta mesma linha, em (Heutte et al., 2004) e (Vuurpijl and Schomaker, 1998b) o processo de classificação é dividido em estágios e cada estágio é associado a um tipo de classificador. Assim as tarefas são divididas e cada classificador fica responsável por solucionar uma determinada parte do problema.

Em SMC Modulares, é possível utilizar técnicas de otimização para melhorar o seu desempenho. Em (Sirlantzis et al., 2002a), por exemplo, é introduzido um Sistema Multi-Classificador que incorpora um algoritmo genético de forma a prover uma seleção simultânea e dinâmica não somente nos participantes da classificação, mas também na regra de combinação utilizada. É explorada a eficiência do algoritmo evolucionário em relação ao comportamento da configuração dos especialistas resultante.

2.3.1.2 Sistemas Classificadores Paralelos

Os Sistemas Multi-classificadores Paralelos, que também são conhecidos como *Ensembles*, como também já foi citado anteriormente, são aqueles em que todos os classificadores do sistema possuem a mesma tarefa. Dentre os trabalhos mais recentes encontrados na literatura, pode-se citar (Li et al., 2005a), (Prudent and Ennaji, 2005), (Chen et al., 2005), (Abreu et al., 2005a), (Abreu et al., 2005a) e (Abreu et al., 2005f), dentre outros.

E em *Sistemas Inteligentes* (Rezende, 2002) existe uma grande série de aplicações dentro de classificadores. Neles é explicado com funcionam os classificadores, que são

os algoritmos responsáveis por solucionar problemas e como funcionam o combinadores, que são responsáveis por combinar os resultados dos classificadores e produzir uma saída para o sistema.

Já em (Prudent and Ennaji, 2005) é introduzido um novo esquema para a resolução de tarefas de reconhecimento de padrões de forma geral. Esse esquema foi desenvolvido utilizando um Sistema Multi-Classificador. A distribuição do processo respeita a topologia dos dados e suas características no espaço, tudo isso para encontrar decisões confiáveis. Para isso, foi utilizada uma rede auto-organizável que provê um grafo que representa a topologia dos dados. Durante o processo de decisão este grafo é usado para ativar o classificador apropriado entre um conjunto de especialistas.

Em (Chen et al., 2005) foi feita uma investigação das edições de independência e diversidade entre os classificadores individuais participando de um esquema de multi-classificador de fusão. É apresentada uma definição formal da independência estatística dos classificadores. Neste trabalho, cada classificador tenta resolver o mesmo problema e o combinador final é um método de fusão que escolhe a melhor saída.

E em (Abreu et al., 2005a) foram feitos experimentos utilizando a ferramenta *Weka* (Witten and Frank, 2000). Foram utilizados vários métodos de classificação aplicados na base SCOP de dobra de proteínas (Blake and Merz, 1998). Nele foram utilizados os métodos de combinação: *Stacking*, *StackingC* e Votação. Já, como métodos de classificação, na tarefa de predição das classes de estrutura de proteínas, foram utilizados: Árvore de Decisão, k-Vizinhos mais Próximos, *Naive Bayes*, Máquinas de Vetores de Suporte e Redes Neurais. Neste trabalho foi observado que sistemas de multiclassificação tem um desempenho melhor que os classificadores individuais de forma geral.

Seguindo a mesma linha, em (Abreu et al., 2005a) e em (Abreu et al., 2005f) foram feitos trabalhos, mais uma vez utilizando a base de dados de proteínas (Blake and Merz, 1998). Neles são feitas alguns experimentos com SMC Híbridos e Não Híbridos, utilizando um conjunto de tipos diferentes de classificadores e um conjunto diferente de combinadores. Eles são trabalhos que analisam e avaliam resultados de forma diferente, mudando os tipos de experimentos que são realizados e os tipos de comparações que são feitas entre eles. Todos os experimentos foram realizados na plataforma *Weka* (Witten and Frank, 2000).

2.3.2 Sistemas Multiagentes

Como já foi mencionado anteriormente, os Sistemas Multiagentes podem ser divididos em dois tipos: Sistema Reativos e Sistemas Cognitivos. Nas seções 2.3.2.1 e 2.3.2.2

são apresentados os trabalhos mais recentes e relevantes para esta dissertação que estão relacionados com os dois tipos de SMA. Na Seção 2.3.2.3 são mostrados os trabalhos mais atuais que abordam o reconhecimento de padrões solucionado por SMA.

2.3.2.1 Sistemas Multiagentes Reativos

Como citado anteriormente, os Sistemas Multiagentes Reativos são aqueles no qual o modelo de funcionamento do agente é formado pelo par Estímulo-Resposta (Ação-Reação). Ou seja, ele possui reações pré-definidas que são ativadas quando ele reconhece que sofreu ações que também são pré-definidas. Como principais trabalhos presentes na literatura no momento, pode-se citar (Anja, 2003), (Lawley et al., 2003), (Waller and Dodds, 1999), (Panwai and Dia, 2005), (Neves and Oliveira, 1998), (Corchado et al., 1997), (Popovici et al., 2005), (Popovici et al., 2004), dentre outros.

Em (Anja, 2003), por exemplo, é mostrado que uma das tarefas dos agentes na Internet seria fechar contratos de interesse por conta própria. O fechamento de contratos é sujeito a responsabilidades, pois não existe diferença entre o mundo real e o virtual. Exatamente como no mundo real, existem duas fases no fechamento do contrato, a fase pré-contratual e a contratual. A responsabilidade de ambas é diferente. Assim, os métodos de negociação para agentes precisam saber distinguir essas duas fases e saber o que fazer em cada uma delas.

Já em (Lawley et al., 2003) é mostrado como o mecanismo de negociação permite usuários controlarem os níveis de Qualidade de Serviço que lhes são requisitados. Isso é feito através de um protocolo de negociação automático que identifica termos aceitáveis mutualmente em Grades de Serviço de Notificação. Esse protocolo resolve os problemas de informações destacadas entre a disseminação e as preferências do usuário.

Em (Waller and Dodds, 1999), por outro lado, apresenta uma aplicação dos agentes reativos na análise e construção de avatares humanos na utilização de ferramentas de realidade virtual. Normalmente as ferramentas usadas não possuem informações de fatores humanos, isto é solucionado adicionando um planejamento avançado que pode prover ao avatar se construir automaticamente.

As aplicações da robótica atualmente utilizam agentes reativos como base, como pode ser observado em (Neves and Oliveira, 1998). Neste trabalho, é feita uma proposta de desenvolvimento de uma plataforma móvel em um robô autônomo e móvel que é capaz de executar diversas tarefas em diferentes ambientes e, além disso, é capaz de agregar novas capacidades.

Outra área que pode-se encontrar grande utilização de agentes reativos é em sistemas

desenvolvidos para crianças, como pode ser visto em (Popovici et al., 2004). Neste trabalho apresenta um ambiente virtual educacional que é usado no ensino fundamental como uma ferramenta suplementar para ensinar as crianças. Neste ambiente, as crianças criam avatares e interagem entre si tomando decisões utilizando mecanismos cooperativos.

2.3.2.2 Sistemas Multiagentes Cognitivos

Diferente dos Sistemas Multiagentes Reativos, os Cognitivos, como já foi citado anteriormente, são sistemas muito mais complexos e apurados. Eles possuem uma representação do ambiente detalhada e possuem uma memória, o que possibilita a consulta a experiências anteriores na hora da tomada de decisão. Eles são mais adequados para representação humana exatamente por possuírem as características enumeradas anteriormente. Como alguns dos trabalhos existentes na literatura atualmente, pode-se citar (Dunne et al., 2003), (Dunne, 2004), (Farrugia, 2002), (Miao et al., 2002), (Ogata, 2002), (Fletcher, 2006), (Ayesh and Cowell, 2004), (van Otterlo et al., 2003), (Kawamura, 2002), (Davis, 2000), (Boella and van der Torre, 2004), (Leong and Chunyan, 2005), dentre outros.

Em (Dunne et al., 2003) foi feito um estudo sobre a complexidade do uso de SMA como o princípio de implementação de contrato de negociação no comércio de Internet. Além disso foi observado a complexidade desse uso na negociação de Internet como o modelo abstrato, considerando que o conjunto de agentes possuem uma coleção de recursos compartilhados entre eles. Eles tentam construir um benefício mútuo de relocação desses recursos. A simples procura de negociação é aquela em que cada agente transfere exatamente um recurso para outro. Isso é chamado de *O-contracts*.

Já em (Dunne, 2004) são examinadas as propriedades do modelo de alocação de recursos onde é considerado um número de alternativas de noções de racionalidade. Enquanto isso, é conhecido que impondo restrições de racionalidade e da estrutura a forma de mudança destes recursos deve ser suficiente para realizar toda a realocação durante a negociação.

E em (Farrugia, 2002) é mostrado que antes que os agentes baseados em Internet pudessem entender negociação, eles precisavam concordar com protocolos de alto nível baseados na lógica do sistema. A comunidade da *Websemantic* está desenvolvendo um conjunto de estados para a interoperabilidade entre *software* de agentes baseados em Internet e linguagens com semânticas bem definidas. Porém, não está claro como essas linguagens podem ser usadas para compreensão da negociação. Logo, nesse trabalho a lógica de sistema é mostrada como solução para este problema.

Um uso bastante comum de agentes cognitivos, também, é no ambiente de Internet. Existe uma crescente necessidade de vários *web-services*, comércio eletrônico e negócios

feitos pela Internet que provenham recomendações customizadas dos seus clientes. Em (Miao et al., 2002) é proposto um novo tipo de agente de recomendação personalizado utilizando agentes cognitivos baseados em fusão. Estes agentes são designados a sugerir baseados nas preferências dos usuários, outros usuários e especialistas no assunto abordado.

2.3.2.3 Sistemas Multiagentes para Reconhecimento de Padrões

O uso de agentes para realizar tarefas de reconhecimento de padrões vem crescendo cada vez mais na literatura. Basicamente eles podem ser de dois tipos: agentes que realizam a mesma tarefa e agentes que realizam tarefas diferentes. A seguir apresentam-se alguns trabalhos que usam esses conceitos. Na primeira subseção a seguir mostram-se alguns trabalhos que usam agentes para resolver tarefas diferentes, já na segunda subseção a seguir mostram-se alguns trabalhos que usam agentes para resolver tarefas iguais.

Agentes Classificadores com Tarefas Diferentes

A partir do que já foi mostrado nesta dissertação, Sistemas compostos por Agentes Classificadores que realizam tarefas diferentes são aqueles em que cada agente, ou conjunto de agente, deve resolver apenas uma parte do problema a ser trabalhado e quando cada agente, ou conjunto de agentes, fornece sua resposta para aquela parte do problema, um combinador é responsável por concluir, de acordo com as informações fornecidas, qual será a resposta mais adequada do sistema. Dentre os principais trabalhos que seguem essa abordagem, pode-se citar (Raje et al., 1998), (Heutte et al., 2004), (Heutte et al., 2004), (Hull, 1994), (Schomaker et al., 1998), (Vuurpijl and Schomaker, 1998a), (Morency et al., 2005), (Abreu et al., 2006a), (Abreu et al., 2006b).

Em (Raje et al., 1998), o principal foco da pesquisa é desenvolver um sistema baseado em agentes capaz de prover serviço de informação personalizada para um usuário com o mínimo de intervenção do mesmo. Eles utilizam uma biblioteca virtual onde é filtrada a informação e é feita a classificação de dois tipos: Ambientes de um único agente e Ambientes Multiagentes.

Já em (Heutte et al., 2004) é feita uma análise dos sistemas que fazem a leitura automática de textos escritos à mão. O trabalho mostra como a capacidade do sistema de leitura se adapta para o reconhecimento para cada escrita humana. Isso ocorre através da exploração do contexto gráfico definido entre as invariantes do escritor. Esta adaptação é garantida por ligações de interação ativada sobre todo o texto entre os procedimentos de reconhecimento da palavra e da letra. Além disso, é utilizado uma arquitetura de multiagentes para suportar os princípios da implementação de adaptação.

A essência do projeto apresentado em (Schomaker et al., 1998) é que em muitos campos do reconhecimento de padrões, várias técnicas são avaliadas para classificar instâncias de padrões. Cada aproximação é caracterizada pelas suas próprias virtudes e respostas. A idéia de combinar a saída de muitos classificadores tem sido estudada por muitos anos na expectativa de quando comparada ao melhor classificador, SMC, deve-se ter uma maior capacidade de reconhecimento. A combinação de muitos classificadores é um exemplo do mais geral e fundamental problema da integração de informação de muitas fontes. Como escolher uma combinação de algoritmos é ainda um problema e esse é o tópico de pesquisa desse projeto.

Em (Vuurpijl and Schomaker, 1998a) é descrito um novo *framework* que utiliza agentes inteligentes para reconhecimento de padrões. Com ele é implementado um reconhecedor de dígito para validação de senha. A modularidade intrínseca do uso de agentes provê uma resposta mais rápida, um desenvolvimento mais estratificado e um paralelismo para o sistema proposto.

Já em (Morency et al., 2005) é apresentado um modelo de reconhecimento visual que integra sugestões de um diálogo falado de um agente com a observação direta da reação mais aparente do usuário. Isso é motivado pelo fato que as reações e os gestos oferecem muitas chaves convencionais e são usadas essencialmente na interação entre pessoas. Foi investigado como a informação contextual pode gerar uma reação de ações durante interações com agentes.

Em (Abreu et al., 2006a) e em (Abreu et al., 2006b) é feito a implementação de um sistema multiagente para classificação de padrões utilizando-se da plataforma multiagente *JADE*, onde os agentes têm visões diferentes do padrão a ser classificado apesar de se proporem a resolver a mesma tarefa. Além disso, dois tipos de métodos de classificação são utilizados (redes neurais MLP e RBF).

Agentes Classificadores com Tarefas Iguais

Por outro lado, Sistemas compostos por Agentes Classificadores que realizam a mesma tarefa são aqueles em que cada agente deve resolver o mesmo problema. Dentre os principais trabalhos que seguem essa abordagem, pode-se citar (Al-Ani and Derich, 2002), (Petrushin, 2000), (Julia et al., 1997), (Sidorova, 1999), (Abreu et al., 2005e), (Abreu et al., 2004) e (Abreu, 2005).

Mas em (Al-Ani and Derich, 2002) é apresentada uma nova técnica de combinação de classificadores baseada na teoria de *Dempster-Shafer* de evidência. Essa teoria é um poderoso método para combinar medidas de evidência para diferentes classificadores. Entretanto, cada um dos métodos avaliados que estima a evidência dos classificadores é

analisado. Estes possuem suas próprias limitações. Além disso, é proposta aqui uma nova implementação no qual existe adaptação para os dados de treinamento.

Em (Petrushin, 2000) é feito uma descrição dos agentes para reconhecimento de emoções em apresentações públicas e suas aplicações em problemas do mundo real. Os agentes podem reconhecer cinco estados emocionais com a seguinte exatidão: Normal ou Não-Emocional (55-75%), Felicidade (60-70%), Raiva (70-80%), Tristeza (75-85%) e Medo (35-55%). A média total de exatidão é, em torno, de 70%. Os agentes podem ser adaptados para um determinado ambiente dependendo dos parâmetros da apresentação e do número de alvos de emoções. Para uma aplicação prática um agente tem sido criado para ser hábil na análise da qualidade do sinal de fala e diferenciar entre dois estados de emoção: Agitação (Raiva, Felicidade e Medo) e Calma (Estado Normal e Tristeza).

Em (Julia et al., 1997) é feita uma descrição de um protótipo para uma aplicação que combina tecnologia de identificação de fala e uma arquitetura de agentes para prover monitoração de mensagens de voz incomuns de usuários. Nele, agentes distribuídos processam o pedido, testando periodicamente o sistema de correio de voz do usuário para identificar a composição das mensagens a partir de um conjunto de vozes selecionadas. Quando a mensagem encontra a condição especificada pelo usuário, os agentes localizam a posição do pedido e notificam o usuário que chegou uma mensagem importante.

E em (Sidorova, 1999) uma estratégia de reconhecimento de padrões baseada na inferência de árvore gramatical e um método de aprendizado estatístico baseado na Árvore de Decisão - Algoritmo C4.5 (Quinlan, 1993) para uma solução é apresentado. O agente trabalha com linguagem Russa. A idéia principal do algoritmo é que o sinal da linguagem é muito rica em detalhes e que essa riqueza reflete na riqueza estrutural da linguística. Árvores são usadas para representar a organização estrutural. Primeiro o sistema aprende um conjunto de árvores automáticas até a técnica de correção de erro baseada em inferência gramatical. Então, um conjunto de distâncias de cada estrutura para cada árvore automática é obtida. Finalmente, a árvore de decisão aprende pelo último conjunto de distâncias usadas no algoritmo.

Já em (Abreu et al., 2005e) foi utilizada a mesma base de dados que em (Abreu et al., 2005a), mas os experimentos foram diferentes. Nesse trabalho é apresentada uma análise entre sistemas de classificação e métodos de negociação propostos no Sistema *NeurAge*. Foram utilizados os métodos de negociação adaptados propostos em (Abreu et al., 2004) e (Abreu, 2005) que são: Teoria dos Jogos, Leilão e Negociação Baseado na Sensibilidade de agente. Dentre os métodos de combinação consagrados na literatura, foram utilizados a Soma e o Voto. Por fim, o trabalho mostra que o método baseado na sensibilidade possuiu um melhor desempenho comparados com os demais.

Capítulo 3

O Sistema *NeurAge* e Suas Características

Os sistemas com classificadores combinados possuem potenciais vantagens sobre os métodos individuais mas, a implementação desses sistemas não é uma tarefa fácil e simples. O principal problema é a determinação da estratégia de combinação. Normalmente, a escolha do melhor método de combinação para um Sistema Multi-Classificador (SMC) necessita de execuções exaustivas. Essa limitação acontece pelo fato de existir um processo de tomada de decisão centralizado.

Os SMCs podem utilizar muitos classificadores em sua topologia, porém apenas o método de combinação é responsável pela escolha da saída final do sistema. Este fato caracteriza um processo de tomada de decisão centralizado e mostra uma grande dependência entre a escolha de um método de combinação e o desempenho do sistema (Kuncheva, 2004). Com o intuito de transformar a tomada de decisão dos SMCs mais distribuída, alguns métodos baseados em seleção foram propostos (Giacinto and Roli, 1999) e (Didaci and Giacinto, 2004). Estes métodos baseados em seleção são conhecidos como DCSs (*Dynamic Classifier Selection*) e eles escolhem o classificador que foi considerado mais apto para a classificação do padrão de entrada. A escolha do classificador que rotula o padrão é feita durante uma fase de operação. Esta escolha é normalmente baseada na escolha do método vencedor.

Nestes métodos de seleção de um SMC acontece uma escolha direta, conhecida também como *one-step process* ou processo de um passo. Neste processo, a escolha do classificador é feita somente uma vez por padrão, onde o classificador considerado o mais competente é escolhido. Podem acontecer situações onde todos os classificadores escolhem respostas erradas, desta maneira produzindo uma resposta errada. O ideal é que um conjunto de classificadores decida, de maneira interativa e incremental, qual dos classi-

ficadores é mais adequado para classificar aquele determinado padrão de entrada. Além disso, os classificadores também devem poder mudar de opinião, baseados em critérios ou informações sobre os padrões de entrada.

O Sistema *NeurAge* é um Sistema Multiagentes (SMA) que foi proposto como uma possível solução para o problema de tomada de decisão centralizado dos SMCs. O funcionamento geral do *NeurAge* pode ser descrito da seguinte maneira: quando uma entrada é mostrada para o sistema, todos os agentes neurais produzem suas próprias respostas. Ao contrário dos SMCs centralizados ou paralelos, onde existe um método de combinação que escolhe a saída do sistema, todos os agentes se comunicarão com os demais para encontrar uma saída comum para o sistema. Dessa maneira, tem-se um caminho mais flexível e ainda distribuído para realizar a tomada de decisão.

A Seção 3.1 mostra como foi feita a ligação entre a topologia bem como o funcionamento dos SMCs e os SMAs e a migração. Na Seção 3.2 explica-se como foi modelado um agente de classificação, ou seja, aquele que possui a capacidade de realizar reconhecimento de padrão. E, por fim, na Seção 3.3 faz-se uma introdução e uma explicação de como funcionam alguns dos métodos de negociação mais utilizados na literatura e como foram feitas as adaptações para que eles fossem utilizados no Sistema *NeurAge*.

3.1 Um Sistema Multiagente para Classificação

A Figura 3.1 ilustra um exemplo típico de um SMC paralelo, também conhecido como *Ensemble* ou Comitês. Como já foi mencionado anteriormente, neste tipo de sistema os classificadores analisam padrões de entrada mostrados a eles e enviam suas respostas para um método combinador. Este combinador é, também, um classificador. Este combinador decide qual a classe adequada para o padrão que foi apresentado ao sistema.

Como já mencionado anteriormente, a existência de um módulo de combinação nos SMCs, que necessita de todas as saídas dos classificadores para obter a saída final de um sistema, pode torná-los sistemas dependentes deste módulo. Além disso, se, por algum motivo, o combinador parar de funcionar o sistema parará. Isso ocorre porque a tomada de decisão é um processo centralizado. Neste contexto, iniciou-se o estudo de uma solução viável com o objetivo de atenuar esse problema. A partir do modelo da Figura 3.1 começou-se a pensar como poderia ser resolvido esse problema de centralização. Pensou-se, então, em criar estruturas que fizessem a classificação de uma maneira diferente, no qual os classificadores pudessem decidir, por si só, qual seria a saída do sistema.

A Figura 3.2 mostra o sistema proposto inicialmente em (Abreu et al., 2004). Na sua estrutura observa-se um sistema no qual os todos os seus componentes podem se

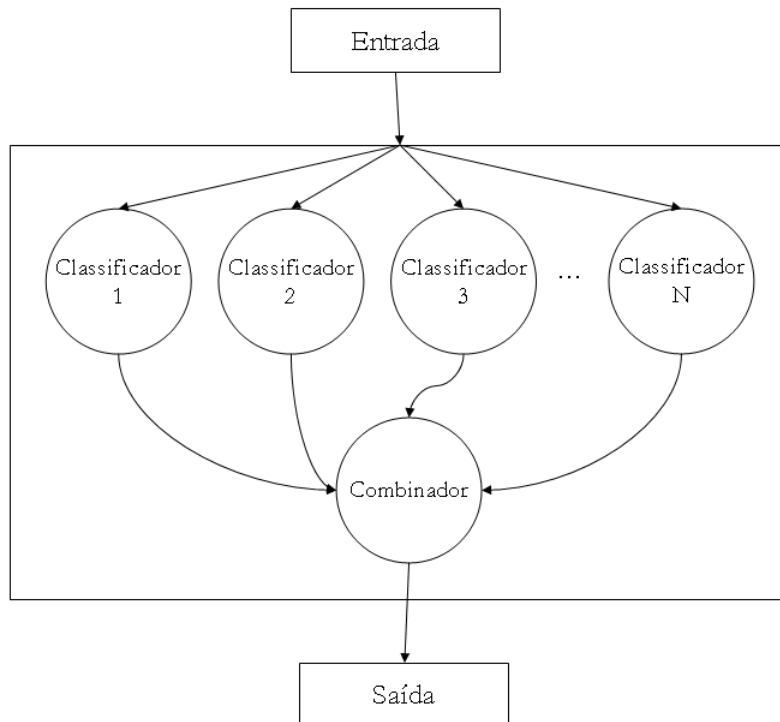


Figura 3.1: Modelo de um Sistema Multi-Classificador Paralelo Tradicional

comunicar. Todas as entidades seriam autônomas e uma entrada é mostrada para todas essas entidades.

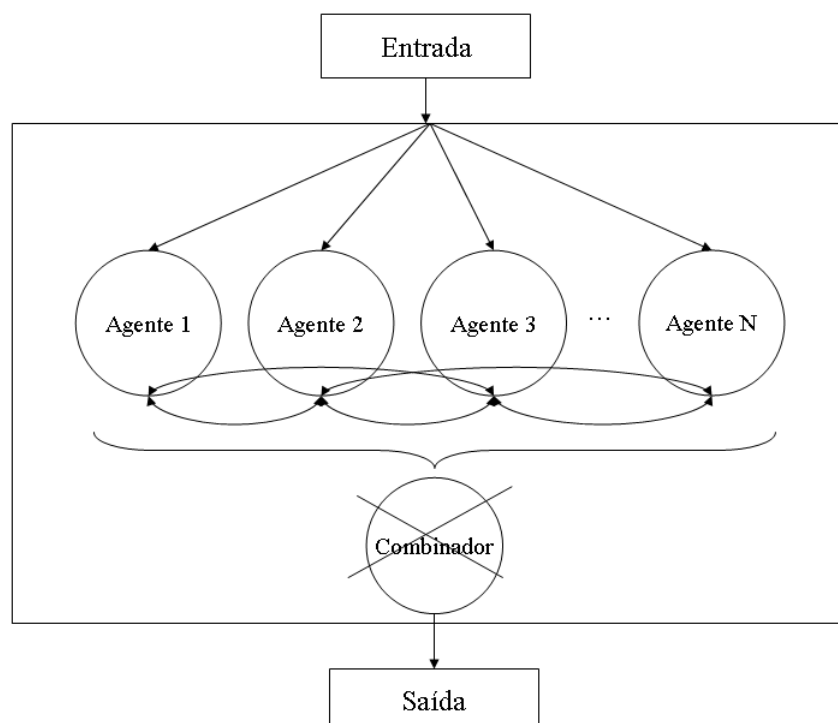


Figura 3.2: Eliminação do Módulo de Combinação do Sistema Multi-Classificador

As entidades supracitadas, por conseguinte, produziriam suas respectivas saídas e, diferentemente da Figura 3.1 no qual essas saídas eram enviadas para um módulo de combinação, as entidades se comunicariam entrando em acordo em relação à resposta para a entrada fornecida.

Percebeu-se que esse sistema proposto era bastante semelhante com um SMA. Porém, nele os agentes deveriam executar tarefas de reconhecimento de padrões. Para atender a essa nova funcionalidade foi proposta a idéia de modelar um agente de classificação que era composto pelos módulos básicos de um agente inteligente e, além disso, possuiria um módulo extra que executaria o reconhecimento de padrões. Essa nova entidade seria conhecida como Agente Classificador.

3.2 Agente Classificador do Sistema *NeurAge*

A arquitetura de um Agente Classificador pode ser representada pela Figura 3.3 que é composta pelos módulos citados a seguir.

- **Controlador:** Este módulo é responsável por receber as perguntas do usuário, assim como, por definir a ordem de ativação de seus processos internos.
- **Tomada de Decisão:** Este módulo é responsável por raciocinar sobre seu conhecimento a fim definir a melhor saída para um classificador.
- **Negociação:** Este módulo é responsável pela comunicação com outros agentes a fim de que um resultado comum seja obtido. Esta negociação é feita através de protocolos de negociações. Durante o processo da negociação, pode-se sugerir que um agente mude seu resultado. Entretanto, cada um deles possui a autonomia para decidir se muda ou confirma seu resultado atual.
- **Classificador:** Este módulo é responsável por executar o método de classificação do agente. Este método de classificação produz uma saída para cada classe possível do problema em questão, esta saída representa o grau de certeza que o agente tem de que o padrão de entrada pertença àquela classe. A classe escolhida pelo agente é aquela que possui o maior valor de saída do método de classificação, ou seja, o maior grau de certeza. O fato do método de classificação produzir esses graus de certezas possibilita o agente de analisar outras possibilidades de resposta durante a negociação.

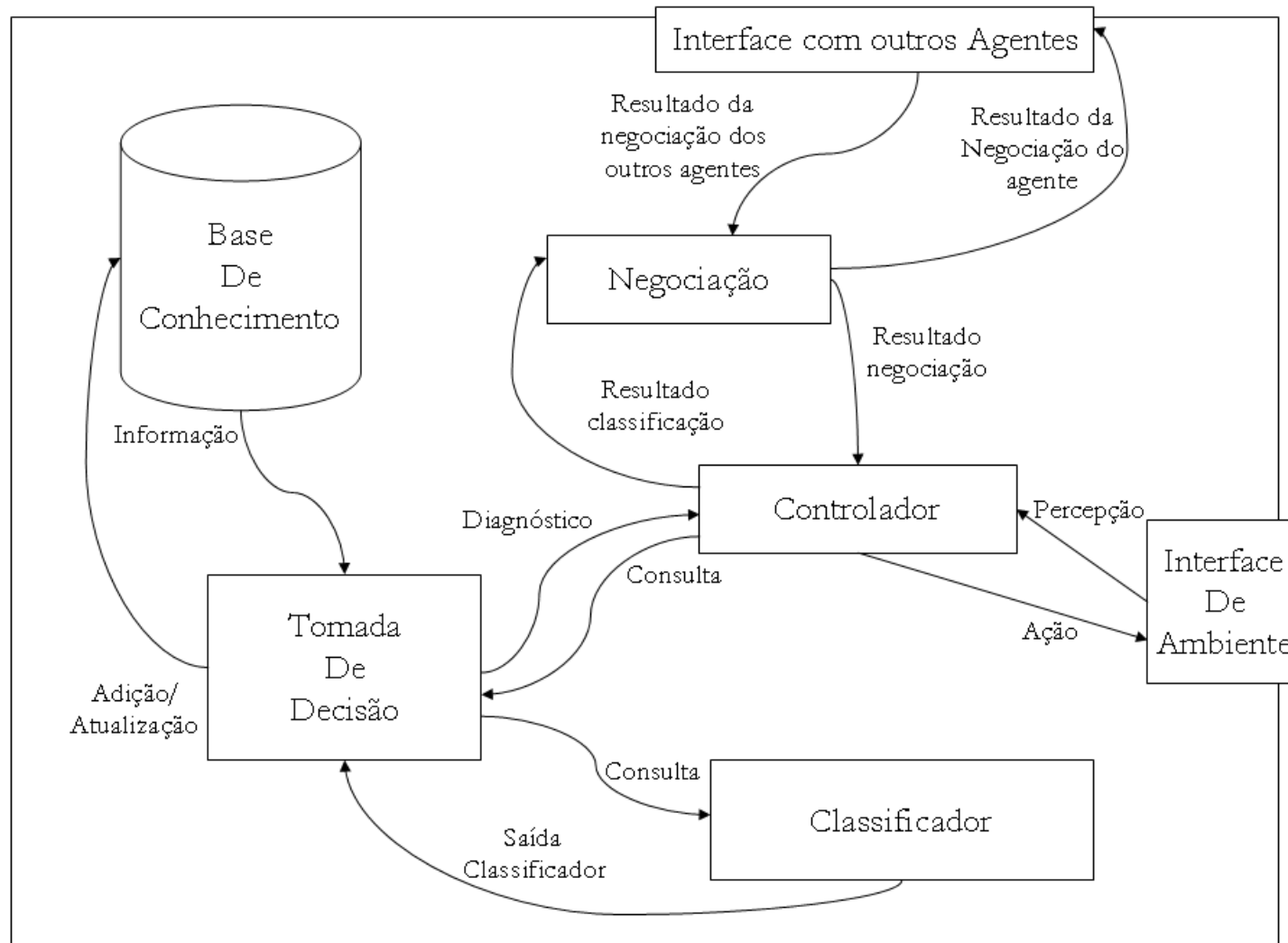


Figura 3.3: Arquitetura Interna de Um Agente Classificador

A idéia principal por trás do funcionamento do Agente Classificador mostrado na Figura 3.3 é que uma vez que a pergunta do usuário é fornecida, o controlador passa a informação ao módulo de tomada de decisão, que alcança o módulo classificador para produzir sua saída. Então, o controlador pode decidir se deve se comunicar com outros agentes, através do módulo de negociação, a fim de alcançar um resultado comum ou não.

Durante o processo da negociação, pode ser necessária uma mudança de opinião sobre a saída atual ou que se execute novamente alguma outra ação (Abreu et al., 2004). O processo de mudar um resultado é possível porque o módulo classificador fornecerá sempre uma lista de classes possíveis baseado em alguns critérios da avaliação. Além disso, um agente pode decidir executar a tomada de decisão mais uma vez, analisando outros critérios ou características do padrão de teste. Uma vez que um resultado comum é alcançado, fornece-se ao usuário.

3.3 Métodos de Negociação

A necessidade de que os agentes se comuniquem para que possam chegar a uma solução comum é bastante evidente no Sistema *NeurAge*. Com isso, observou-se que o processo de negociação é muito importante para o seu bom funcionamento (Abreu et al., 2005d). Os protocolos de negociação usados nos SMAs não podem ser aplicados adequadamente no Sistema *NeurAge*. Como exemplo, tem-se que durante a negociação em um sistema baseado em Leilão, dois agentes (comprador e vendedor) tentam encontrar um valor comum que seja bom para ambos (Tomohara, 2002). Por outro lado, dois agentes neurais do *NeurAge* devem decidir a qual classe pertence um determinado padrão de entrada. No *NeurAge*, ambos podem produzir diferentes classes para o mesmo padrão. A mudança do valor da saída de um agente do *NeurAge* não beneficiará o outro agente como acontece no sistema baseado no Leilão. Dessa maneira, um deles ou ambos terão que mudar seus resultados e aceitar a saída do outro (Toda et al., 2001). Daí, a necessidade de adaptá-los. Existem vários métodos de negociação na literatura, dentre eles foram utilizados os citados a seguir.

- Método baseado na Teoria dos Jogos, (De Vos and Vermeir, 2001), (Mathieu et al., 2006),
- Método baseado no Leilão, (Li et al., 2005b), (Ben-Ameur et al., 2002),
- Método baseado na Sensibilidade (Abreu, 2005), (Abreu et al., 2005d).

Quando usam-se métodos de classificação, os valores de confiabilidade podem ser associados com as amostras para mostrar a confiabilidade do método em questão na classificação de um padrão de entrada para uma classe particular. Valores de confiabilidade podem, também, ser associados a todas as outras classes e eles definem um certo grau de pertinência de um exemplo e todas as classes. Estes valores de confiabilidade, também conhecidos como grau de confiabilidade, são providos por todos os métodos de classificação e podem ser considerados como uma das mais valiosas informações que podem ser extraídas das saídas dos métodos de classificação (Zhou, 2002). No Sistema *NeurAge*, uma vez que uma classe ganhadora é definida pelo classificador de um agente, sua confiabilidade correspondente é considerada a confiabilidade do agente. Quando um agente muda o seu resultado (classe), sua confiabilidade também muda.

É importante enfatizar que o uso da confiabilidade foi aplicado nos três métodos de negociação. Entretanto, não é permitido que um agente possua confiabilidade negativa. Neste caso, ou o agente muda a classe vencedora ou será descartado do processo de negociação, pois não existe sentido em manter um agente com confiabilidade negativa no processo de negociação. Em todos os métodos que serão descritos a seguir, a decisão de manter ou descartar um agente da negociação é baseado em regras internas e no domínio da base de conhecimento de cada agente.

Outro detalhe importante é que nestes métodos de negociação o agente não pode optar por não negociar, ou seja, sempre que houver discordância de opiniões, haverá negociação. A seguir serão explicados como funcionam os métodos de negociação abordados neste trabalho e como são as adaptações feitas para a utilização do Sistema *ClassAge*.

3.3.1 Método de Negociação Baseado na Teoria dos Jogos

A Teoria dos Jogos vem sendo amplamente usada como uma ferramenta de cooperação em Sistemas Multiagentes (SMA) (Osborne, 2003). Na Teoria dos Jogos, uma descrição sistemática dos resultados pode ser feita através da utilização da estratégia dos jogos. Uma estratégia de jogo é um jogo no qual um jogador, neste caso um agente, escolhe o seu plano de ação apenas uma vez e ao mesmo tempo que o seu oponente. De maneira a ajudar os jogadores a tomar suas decisões, uma matriz de *payoff* (matriz de retorno) é usada. Nesta matriz, cada célula representa o valor de *payoff* que os jogadores terão no caso dessas ações serem escolhidas. Um modelo da matriz de *payoff* é mostrado na Tabela 3.1, onde existem dois jogadores e cada um deles só possui duas ações.

Como pode ser visto na Tabela 3.1, os jogadores devem analisar cada uma das possibilidades e escolher aquela onde ele mesmo possui a menor perda. Para o caso do jogador

Jogador 2	Jogador 1	
	AçãoA/Jogador1; AçãoB/Jogador2	AçãoA/Jogador1; AçãoD/Jogador2
	AçãoC/Jogador1; AçãoB/Jogador2	AçãoC/Jogador1; AçãoD/Jogador2

Tabela 3.1: Matriz de *Payoff*.

1, ele pode executar apenas as ações A e C, já o jogador 2, pode executar apenas as ações B e D. A partir disso, eles negociam até chegar em comum.

Baseado no que foi introduzido anteriormente, a estratégia da teoria dos jogos foi ajustada para que fosse possível ser utilizada no *NeurAge*. Como já mencionado, quando um padrão de entrada é apresentado ao Sistema *NeurAge*, todos os agentes escolhem sua classe de saída, que representa a classe ganhadora para aquele padrão, ou seja aquela com o maior valor de confiabilidade. Além disso, é fornecido por cada agente a lista de confiabilidades, ou seja, o grau de pertinência do padrão de entrada para cada classe. A partir disso, todos os agentes do *NeurAge* possuem sempre duas ações possíveis, que são: manter a classe vencedora ou mudar a classe vencedora para a classe vencedora do agente oponente. A matriz de *payoff* para esta situação seria a mostrada na Tabela 3.2.

Ag_j	Ag_i	
	muda/ Ag_i ; muda/ Ag_j	nãoMuda/ Ag_i ; muda/ Ag_j
	muda/ Ag_i ; nãoMuda/ Ag_j	nãoMuda/ Ag_i ; nãoMuda/ Ag_j

Tabela 3.2: Matriz de *Payoff* para o *NeurAge*.

Na Tabela 3.2 observa-se que existem quatro configurações possíveis que são: ambos mudam de classe, ambos não mudam de classe ou um deles muda e o outro não muda. No *NeurAge*, o agente escolhe o que for melhor para ele. É importante, ainda, definir a medida de *payoff* que será utilizada na matriz. No caso usado no *NeurAge*, essas medidas devem ser baseadas na confiabilidade da classe vencedora do agente e na confiabilidade da classe correspondente a classe vencedora do agente oponente. As equações 3.1 e 3.2 mostram como seria o cálculo da matriz para as possibilidades do agente mudar ou não mudar de classe respectivamente para o Ag_i .

$$muda/Ag_i = \frac{conf Ag_i[VencAg_i] + conf Ag_i[VencAg_j]}{2} \quad (3.1)$$

$$naoMuda/Ag_i = conf Ag_i[VencAg_i] - conf Ag_i[VencAg_j] \quad (3.2)$$

onde $conf Ag_i[VencAg_i]$ é a confiabilidade do agente i para a classe atualmente vencedora do agente i e $conf Ag_i[VencAg_j]$ é a confiabilidade do agente i para a classe atualmente vencedora do agente j . Esses valores de *payoff* serão correspondentes as novas

confiabilidades dos agentes. Por exemplo, se um agente muda de classe, a confiabilidade da classe que era vencedora vai ser igual ao valor calculado na Fórmula 3.1. Já se ele não muda de classe, a confiabilidade da classe que era vencedora vai ser igual ao valor calculado na Fórmula 3.2. Com isso, tem-se que os valores de *payoff* para as duas ações são, como mostrados anteriormente:

- A ação muda: o valor de *payoff* é a média da soma entre as confiabilidades dos agentes;
- A ação não muda: o valor de *payoff* é a diferença entre as confiabilidades dos agentes.

Essa negociação só acontece entre dois agentes de cada vez. Caso existam mais de dois agentes na negociação, é feita uma ordenação pelo valor da confiabilidade da classe vencedora de cada agente. A partir disso, é criada uma lista de ordenação que contém as referências para os agentes. Sempre que um agente é excluído do processo de negociação, o que está imediatamente após ele na lista de ordenação entra nesta negociação.

Quando um agente decide mudar de classe e o outro decide não mudar, aquele que mudou é considerado o perdedor e é eliminado da negociação. Então um novo agente entra no seu lugar. Quando os dois mudam ou não mudam, ambos continuam na negociação. A negociação termina quando não existe mais nenhum agente na lista de ordenação para substituir aquele que é retirado da negociação. Um plano de ação que representa a Teoria dos Jogos aplicado ao *NeurAge* é mostrado a seguir.

1. Apresente o padrão para cada um dos agentes e estes escolhem suas classes vencedoras;
2. Se todos os agentes possuírem a mesma classe vencedora, vá para o Passo 9;
3. Ordene os agentes decrescentemente de acordo com suas confiabilidades e coloque-os em uma lista de negociação;
4. Se ainda existir mais de dois agentes na lista de negociação, escolha os dois com as maiores confiabilidades e com classes vencedoras diferentes para negociar. Senão vá para o Passo 9;
5. Para cada um deles calcule os novos valores de *payoff* para os casos deles mudarem e não mudarem de classe;
6. Escolha a ação onde acontece menor perda para cada um dos agentes;

7. Se apenas um dos dois agentes escolheu mudar de classe, modifica-se a sua confiabilidade e ele está excluído da negociação. Vá para o Passo 3;
8. Se ambos escolheram não mudar ou mudar de classe, modificam-se as suas confiabilidades e ambos continuam na negociação. Vá para o Passo 3;
9. Se ainda existir padrão, vá para o próximo padrão e para o Passo 1, senão vá para o Passo 10;
10. Termina a negociação.

3.3.1.1 Exemplo do Método de Negociação Baseado na Teoria dos Jogos

Como exemplo para ilustrar a operação desse método, será analisada uma tarefa de reconhecimento de padrões. Para tal serão consideradas de três classes (A, B e C) no qual cada padrão contém cinco atributos (at1, at2, at3, at4 e at5). Supondo um sistema composto de dois agentes (Ag1, Ag2). Após o processo de treinamento o seguinte padrão de teste (0,7; 0,4; 0,34; 0,9; 0,22) é apresentado aos agentes. O módulo de classificação de cada agente produz suas respectivas saídas. A Tabela 3.3 mostra a saída de cada um dos agentes.

Agente 1 (Ag1)		Agente 2 (Ag2)	
Classe	Confiabilidade	Classe	Confiabilidade
A	0,90	B	0,87
B	0,30	A	0,56
C	0,25	C	0,34

Tabela 3.3: Confiabilidades dos dois Agentes.

A Tabela 3.3 representa o Passo 3 do plano de ação mostrado anteriormente. Nesse caso, a classe escolhida pelo Ag1 foi a A, por ter tido a maior confiabilidade, seguido pela B e C. Já, a classe escolhida pelo Ag2 foi a B, seguida pela A e C. O Passo 4 do plano de ação é satisfeito, pois existem dois agentes para negociar. Como já foi dito, uma matriz de *payoff* é calculada para cada ação dos agentes, isto corresponde ao Passo 5 do plano de ação. Logo, para o Ag1 e o Ag2 tem-se que, aplicando as Funções 3.1 e 3.2, os resultados mostrados a seguir.

- $muda/Ag1 = \frac{0,90+0,30}{2} = 0,60$, referente a Função 3.1;
- $naoMuda/Ag1 = 0,90 - 0,30 = 0,60$, referente a Função 3.2;
- $muda/Ag2 = \frac{0,87+0,56}{2} = 0,715$, referente a Função 3.1;

- $naoMuda/Ag2 = 0,87 - 0,56 = 0,31$, referente a Função 3.2.

Para esta situação a matriz de *payoff* seria como a mostrada na Tabela 3.4. Nela, pode-se observar que para o agente 1 não existe diferença entre as ações mudar ou não, pois em ambos os casos sua confiabilidade para a classe A teria uma diminuição de 0,30. Já para o agente 2 é mais vantajoso escolher a ação muda de classe, pois a sua confiabilidade para a classe B diminuiria 0,155 apenas, já se ele escolhesse a ação não muda a sua confiabilidade diminuiria 0,56. Sendo assim, a célula escolhida da matriz de *payoff* é a célula $C_{1,2}$, ou seja, a célula correspondente a linha 1 e coluna 2 ou 1,2.

	Agente 1	
	Mudar	Não Mudar
Agente 2	0,60; 0,715	0,60; 0,715
	0,60; 0,31	0,60; 0,31

Tabela 3.4: Aplicação da Matriz de *Payoff* para o *NeurAge*.

Como já foi mencionado anteriormente, se apenas um dos dois agentes escolhe a ação mudar, este agente está descartado da negociação. Seguindo o plano de ação mostrado anteriormente, no Passo 6 tem-se que o agente 1 escolhe a ação não mudar e o agente 2 escolhe a ação mudar. Pelo Passo 7, o agente 2 está fora da negociação e volta-se para o Passo 3. Fazendo uma nova ordenação das confiabilidades do agente 1, tem-se que a sua classe vencedora continuaria sendo a classe A. Indo para o Passo 4 observa-se que não existem mais agentes para negociar, sendo assim, vai-se para o Passo 10 que finaliza a negociação para aquele padrão de entrada.

Na arquitetura interna dos agentes que usam este método de negociação, a base de conhecimento compartilhada contém toda a informação sobre os valores de mudar ou não mudar de todos os agentes que estão na negociação. Por outro lado, o domínio da base de conhecimento privada possui informações sobre as confiabilidades dos agentes, as novas confiabilidades e um limiar, que é individual para cada agente. Existindo este limiar, um agente pode decidir mudar, por ele mesmo, uma classe vencedora. Como já mencionado, estas decisões podem ser baseadas em regras internas e individuais de cada agente, dependendo de experiências passadas.

3.3.2 Método de Negociação Baseado no Leilão

Leilão é outro protocolo de negociação que foi ajustado para ser usado no Sistema *NeurAge*. O leilão é um método no qual uma série de regras devem ser seguidas. Isto pode determinar recursos e valores para uma boa base em uma proposta inicial feita por um dos seus participantes. Este método é bastante popular no ambiente da Internet (Farrugia,

2002). O leilão pode ser visto como uma transação comercial onde existe um vendedor e vários compradores. Um comprador é considerado vencedor de um leilão quando ele consegue dar o melhor lance, ou seja, fazer a proposta mais vantajosa para o vendedor. Uma seção de leilão pode durar uma considerável porção de tempo. Por exemplo, um leilão de internet pode durar dias até que se defina qual é o comprador vencedor, ou seja, aquele com a melhor proposta.

Diferentemente do leilão original, neste método, todos os agentes são considerados como de um tipo, ou seja ou todos vendedores ou todos compradores. Eles tentam alcançar um acordo em relação a um padrão de entrada. Em analogia com o leilão original, esta abordagem é similar a uma situação onde todos os compradores ou vendedores decidem encontrar uma maneira no qual cada comprador tenha o melhor preço em relação ao início da Seção do leilão.

Da mesma forma que mostrado no método anterior, este método também utiliza a medida de confiabilidade como base do seu funcionamento. Mas, diferentemente do que foi mostrado na Seção anterior, todos os agentes podem fazer parte da mesma iteração ao mesmo tempo. Uma vez que uma iteração do leilão é iniciada, um custo para cada agente é calculado. Este custo é baseado na soma das diferenças entre a confiabilidade da classe vencedora e a confiabilidade correspondente à classe vencedora dos demais agentes. Esta soma é dividida por uma constante que varia de acordo com o problema atacado.

Neste método, o agente com o maior custo é considerado o perdedor e a iteração termina. As confiabilidades das classes vencedoras de todos os agentes são mudadas para as diferenças entre a confiabilidade da classe vencedora atual e o seu custo. Quando um agente perde duas vezes seguidas, ele é descartado da negociação. O agente que permanecer até o final é considerado o vencedor e a classe vencedora deste agente é vista como a classe vencedora do sistema. A tabela de custos possui dois valores possíveis para cada célula que podem ser vistos nas Fórmulas 3.3 e 3.4.

$$parcela = conf[Venc] - conf[VencOpo] \quad (3.3)$$

$$custo = \frac{parcela1 + parcela2 + \dots + parcelaN}{constante} \quad (3.4)$$

Na Fórmula 3.3, o valor de $conf[Venc]$ é o valor da confiabilidade da classe vencedora do agente que está negociando, já $conf[VencOpo]$ é o valor da confiabilidade do agente que está negociando para a classe que o agente oponente acha que é a vencedora. Nesta equação é calculada cada parcela que vai ser utilizada no cálculo do custo que é mostrado na Fórmula 3.4. Nesta Fórmula a quantidade de parcelas vai ser igual a quantidade de

agentes menos um, ou seja, o valor de N será a quantidade de agentes menos um. Um plano de ação que representa o Leilão aplicado ao *NeurAge* é mostrado a seguir.

1. Apresente o padrão para cada um dos agentes e estes escolhem suas classes vencedoras;
2. Se ainda existir mais de dois agentes continue, senão vá para o Passo 8;
3. Para cada um deles calcule os valores das parcelas correspondentes aos seus oponentes, utilizando a Fórmula 3.3;
4. Calcule o valor do custo de cada agente, utilizando a Fórmula 3.4;
5. O agente que teve o maior custo é marcado;
6. Se este agente marcado já tinha sido marcado na iteração anterior, então ele está fora da negociação.
7. Vá para o Passo 2;
8. Se ainda existir padrões, vá para o Passo 1, senão vá para o Passo 9;
9. Termina a negociação.

3.3.2.1 Exemplo do Método de Negociação Baseado no Leilão

Usando o mesmo exemplo da Seção 3.3.1, mostrado na Tabela 3.3, um terceiro agente pode ser acrescentado no sistema. Baseado na Tabela 3.3 e tendo como confiabilidade para o terceiro agente ($C = 0,85$; $B = 0,4$; $A = 0,37$). Esta etapa corresponde ao Passo 1 do plano de ação. Como ainda existem três agentes para negociar, o Passo 2 é satisfeito e é executado o Passo 3 da maneira mostrada a seguir e aplicando sempre a Fórmula 3.3.

Parcelas para o Agente 1:

- Em relação ao agente 2, tem-se que $parcela_{12} = 0,90 - 0,30 = 0,60$;
- Em relação ao agente 3, tem-se que $parcela_{13} = 0,90 - 0,25 = 0,65$;

Parcelas para o Agente 2:

- Em relação ao agente 1, tem-se que $parcela_{21} = 0,87 - 0,56 = 0,31$;
- Em relação ao agente 3, tem-se que $parcela_{23} = 0,87 - 0,34 = 0,53$;

Parcelas para o Agente 3:

- Em relação ao agente 1, tem-se que $parcela_{31} = 0,85 - 0,37 = 0,48$;
- Em relação ao agente 2, tem-se que $parcela_{32} = 0,85 - 0,40 = 0,45$;

A costanta serve para ponderar o valor do custo que cada agente vai possuir no momento de cada lance. Após o cálculo de cada parcela, é executado o Passo 4, com a aplicação da Fórmula 3.4. Considerando que o valor da constante é de 10 (dez). O cálculo do custo para cada agente pode ser visto a seguir.

Custo para o Agente 1:

$$custo_{Ag1} = \frac{parcela_{12} + parcela_{13}}{10} = 0,125$$

Custo para o Agente 2:

$$custo_{Ag2} = \frac{parcela_{21} + parcela_{23}}{10} = 0,084$$

Custo para o Agente 3:

$$custo_{Ag3} = \frac{parcela_{31} + parcela_{32}}{10} = 0,093$$

Para resumir, a matriz completa com o cálculo dos custos e das parcelas de cada custo é mostrada na Tabela 3.5.

-	Ag1	Ag2	Ag3
Ag1	0,125	0,6	0,65
Ag2	0,31	0,084	0,53
Ag3	0,48	0,45	0,093

Tabela 3.5: Matriz de Custo do Leilão.

Como já foi explicado anteriormente, o valor do custo para o Ag1, por exemplo, foi calculado pela soma das células C_{12} (0,6) e C_{13} (0,65) dividido por 10 (constante) ($\frac{1,25}{10} = 0,125$). As células C_{12} e C_{13} são calculadas pela diferença entre a classe escolhida pelo Ag1 e a classe escolhida correspondente aos demais agentes, como foi mostrado anteriormente. Nesse exemplo, Ag1 é considerado como o perdedor para esta iteração, pois o seu custo foi o maior. A partir do valor de custo que foi calculado, subtrai-se esse valor a confiabilidade atual de cada agente para se obter as suas novas confiabilidades. As

novas confiabilidades de cada agente serão $Ag1 = 0,775$; $Ag2 = 0,786$ e $Ag3 = 0,757$. A partir deste ponto, considerando o Passo 6, como é apenas a primeira iteração, pelo plano de ação a negociação continua.

A arquitetura interna dos agentes que utilizam o método de negociação é similar ao anterior, onde a base de conhecimento compartilhada contém todas as informações sobre os valores de custo de todos os agentes para a iteração atual e a anterior. Por outro lado, o domínio da base de conhecimento privada possui informações sobre as confiabilidades dos agentes, as novas confiabilidades e as classes vencedoras. Além disso, o agente pode, mais uma vez, decidir sair do processo de negociação a qualquer momento baseado em regras internas.

3.3.3 Método de Negociação Baseado na Sensibilidade

Este método de negociação foi proposto inicialmente em (Abreu et al., 2004). Ele é uma aplicação baseada no método Teoria dos Jogos, mas utiliza mais informações que a aplicação tradicional da Teoria dos Jogos que foi apresentada na Seção 3.3.1. Neste método é introduzido o conceito de sensibilidade de um agente.

A sensibilidade de um agente é a capacidade deste sofrer alterações em sua classificação diante de alterações nos atributos de entrada para este método. Se um agente tem uma grande sensibilidade para um determinado atributo, significa que ele sofre grandes variações de classificação a partir de pequenas variações nos atributos. Por outro lado, se um agente tem uma pequena sensibilidade para um determinado atributo, significa que ele sofre pequenas variações de classificação, mesmo que aconteçam grandes alterações nos atributos de entrada. Esta sensibilidade é usada durante o processo de negociação entre os agentes.

Este método é também composto por muitas iterações e o agente com a maior confiabilidade na última iteração é escolhida para ser o mais adequado para classificar o padrão de entrada em execução. Em cada iteração, existe um decremento na confiabilidade de todos os agentes que é baseada na análise de sensibilidade. O principal detalhe desta análise é investigar a sensibilidade de um método de classificação para um determinado atributo e usar esta informação no processo de negociação. Além do uso da sensibilidade, também são usadas informações como os atributos de entrada e a média dos atributos de treinamento. Essas informações são usadas para definir a magnitude de decréscimo da confiabilidade dos agentes durante a negociação.

Uma maneira possível de criar um plano de ação para este método é selecionar os agentes, dois a dois, para calcular o risco de conflito em uma específica iteração. Normalmente, dois agentes são escolhidos para fazer parte da negociação. Sendo assim, o risco

de conflito é calculado para ambos e, assim, um deles é escolhido para ceder. Quando um agente é considerado como aquele que cede, ele não fará mais parte do processo de negociação para esta iteração. Em outras palavras, um agente só pode ceder uma vez por iteração. O plano de ação pode ser descrito como se segue.

1. Após a fase de treinamento do módulo de classificação do agente, calcula-se a análise de sensibilidade para todos os atributos de cada classe.
2. Calcula-se a média de treinamento para cada atributo de cada classe em relação àquele método de classificação;
3. Para cada padrão de teste, faça:
 - (a) Calcule a diferença entre os atributos do padrão de entrada e a média de treinamento dos atributos para todas as classes;
 - (b) Ordene as diferenças entre os atributos em ordem decrescente, ou seja, do menos similar para o mais similar;
 - (c) Começa o processo de negociação;
 - (d) Após o processo de negociação, o agente do *NeurAge* com o maior grau de confiabilidade é dito o mais adequado para classificar o padrão de teste e sua saída é considerada a saída do sistema.

O processo de negociação, ou seja, o Passo (c) citado anteriormente, pode ser descrito como a seguir.

1. Escolha o primeiro atributo da lista das diferenças que foi ordenada decrescentemente, ou seja, o atributo menos similar e faça:
 - (a) Escolha dois agentes para fazer parte da negociação;
 - (b) Calcule a possibilidade do risco de conflito de cada agente;
 - (c) Escolha o agente que vai ceder. O risco é usado de maneira similar que a estratégia de Zeuthen para uma negociação PCM (protocolo de concessão monotônica) (Wooldridge, 2002). Esta escolha é baseada no risco do agente, onde o agente a ceder deve ser aquele com o menor valor de risco;
 - (d) A concessão de um agente significa o decréscimo do grau da confiabilidade deste agente para a classe escolhida. Este agente é considerado como um agente que já cedeu para esta iteração;

- (e) Se todos os agente que escolheram classe vencedoras diferentes são considerados como já cedidos, a iteração termina, então retira-se o atributo da lista de diferenças e vai para o Passo 1. Caso contrário, vai para o Passo (a);
2. Este processo acontece até que o máximo número de iterações seja alcançado ou até que só reste apenas um agente na negociação. É importante lembrar que em caso de confiabilidades não negativas, o agente pode decidir mudar a classe vencedora ou sair do processo de negociação.

Como já foi mencionado, um agente deve calcular a análise de sensibilidade para todos os atributos de entrada. Esta análise pode ser feita excluindo e/ou variando o valor do atributo de entrada e analisando a variação do desempenho do método de classificação. Esta análise é feita para todos os atributos de todas as classes do problema de classificação.

A principal idéia por trás do processo mencionado anteriormente é que quanto maior a distância entre o atributo de entrada e o atributo da média do treinamento, maior a probabilidade que a análise da sensibilidade estar errada. Desta maneira, este fato é utilizado para calcular o conflito do risco dos agentes e a concessão, ou seja o decréscimo do grau de confiabilidade. O risco e a medida de concessão são mostrados nas Fórmulas 3.5 e 3.6.

$$risco_i = conf_i + \frac{1}{S_i * (1 - D)} \quad (3.5)$$

$$concessao_i = \frac{D * S_i}{R * C} \quad (3.6)$$

onde:

- $conf_i$ é a confiabilidade do agente i para sua classe vencedora;
- D é a diferença normalizada entre os atributos do padrão de teste atual e os padrões da média de treinamento;
- S_i é a sensibilidade do método de classificação para o atributo correspondente da classe ganhadora;
- R é a posição do atributo na lista da diferença calculada entre o padrão de entrada e o da média de treinamento;
- C é uma constante que define a intensidade da punição.

O detalhe relacionado a medida de risco é que quanto maior a certeza sobre uma classe vencedora, ou seja, a $conf_i$, e quanto menor a sensibilidade para um determinado atributo, ou seja, a S_i , maior este risco é. Enquanto o agente com a menor valor de risco é escolhido, o agente com maior valor de risco, provavelmente, não será escolhido. Além disso, a inversão da diferença normalizada entre o valor do atributo do padrão de teste corrente e o da média de treinamento é também levado em consideração. Se um agente não é muito confiante sobre um padrão de teste, ou seja, ele possui um valor baixo de confiabilidade, e este agente é sensível para o atributo atual, ou seja, possui um alto valor do S_i , mas o atributo em questão para o padrão de treinamento é distante da média de treinamento para este atributo, ou seja, o valor de D próximo de um, este atributo do padrão de teste será considerado como não muito confiável e todos os problemas com a confiabilidade e sensibilidade deste agente são atenuado pela inversão da diferença normalizada.

Como já mencionado, um agente pode mudar a sua classe vencedora. Neste caso, quando a confiabilidade de uma classe vencedora se torna menor que a confiabilidade de alguma outra classe, o agente pode decidir, baseado em experiências passadas e em regras internas, escolher uma classe com confiabilidade maior ou manter a classe vencedora. Além disso, um agente pode decidir executar mais uma vez o processo de tomada de decisão ou treinar novamente o seu método classificador. Isto é baseado na análise do desempenho do agente, se ele teve um desempenho satisfatório ou não na realização da tarefa de classificação.

O final da execução do processo de negociação pode ser feito de duas maneiras. A primeira é definir um número máximo de iterações, ou seja, independente de quantos agentes ainda estejam no processo, quando o número de iterações é atingido a negociação para. Já a segunda maneira é quando só existem agentes com confiabilidades de valor não negativo. Quando o usuário não define um número de iterações, o número de atributos do padrão de entrada é considerado. Além disso, existe sempre um ganhador na negociação. O único problema que pode surgir na utilização deste método é que todos os agentes podem alcançar confiabilidades negativas ao mesmo tempo. Neste caso, eles devem ser descartados da negociação e a confiabilidade do sistema ser a da iteração anterior. Então, o agente com a maior confiabilidade é definido como ganhador.

Na arquitetura interna dos agentes, todas as informações relacionadas com a análise de sensibilidade, a confiabilidade e a média de treinamento de todos os agentes compõem a base de conhecimento. Por outro lado, informações sobre métodos de negociações anteriores, limiares, regras internas e tudo que é individual para cada agente compõem o domínio da base de conhecimento.

3.3.3.1 Exemplo do Método de Negociação Baseado na Sensibilidade

Como um exemplo para ilustrar a operação deste método de negociação, o mesmo exemplo usado na Seção 3.3.1 com o acréscimo do Ag3 que foi feito na Seção 3.3.2. A confiabilidade dos dois primeiros agentes pode ser vista na Tabela 3.3 e a do terceiro agente é dada por ($C = 0,85$; $B = 0,4$; $A = 0,37$). De acordo com o Passo 1 do plano de ação deste método é necessário calcular a análise de sensibilidade e a média de treinamento para todos os atributos de todas as classes. Estes valores podem ser vistos nas Tabelas 3.6 e 3.7.

-	At1	At2	At3	At4	At5
Ag1	1.7%	0.9%	6.2%	5.3%	6.9%
Ag2	8.0%	3.4%	10.7%	4.6%	9.0%
Ag3	7.5%	6.4%	6.8%	3.9%	4.8%

Tabela 3.6: Análise de Sensibilidade dos Agentes.

-	At1	At2	At3	At4	At5
A	0,40	0,50	0,25	0,84	0,73
B	0,20	0,70	0,35	0,88	0,61
C	0,10	0,47	0,46	0,91	0,75

Tabela 3.7: Média de Treinamento para Todos os Atributos.

Seguindo o plano de ação, executa-se o Passo (a), que é exatamente calcular a diferença absoluta entre os atributos do padrão de entrada e os atributos calculados na média do treinamento. Esses valores são normalizados e podem ser vistos na Tabela 3.8. De acordo com o Passo (b) do plano de ação, essas diferenças devem ser ordenadas decrescentemente, e elas podem ser vistas na Tabela 3.9.

-	At1	At2	At3	At4	At5
A	0,30	0,10	0,09	0,06	0,51
B	0,50	0,30	0,01	0,02	0,39
C	0,60	0,07	0,12	0,01	0,53

Tabela 3.8: Distância absoluta entre os atributos do padrão de teste e da média de treinamento.

Na primeira iteração do processo de negociação, sinalizado no Passo (c) do plano de ação, supõe-se que os agentes Ag1 e Ag2 são escolhidos para fazer a negociação. A classe ganhadora para o Ag1 é a classe A e o At5 é o atributo com menor similaridade para esta classe. Além disso, a sensibilidade do Ag1 para o At5 é de 6.9% e a diferença

A	At5, At1, At2, At3, At4
B	At1, At5, At2, At4, At3
C	At1, At5, At3, At2, At4

Tabela 3.9: Ordem decrescente da distância entre os atributos.

para a média de treinamento é 0.51. Por outro lado, a classe ganhadora do Ag2 é B e o At1 é o atributo menos similar desta classe. A sensibilidade do Ag2 para o At1 é de 8.0% e a sua diferença para a média de treinamento da a classe B é 0.5. Baseado nisto, a medida de risco, dada pela Fórmula 3.5 é calculada para ambos os agentes, como mostrado a seguir.

Medida de Custo para o Agente 1:

$$risco_{Ag1/ClasseA/At5} = 0.90 + \frac{1}{6.9*(1-0.51)} = 0.90 + \frac{1}{3.381} = 0.90 + 0.2957 = 1.1957$$

Medida de Custo para o Agente 2:

$$risco_{Ag2/ClasseB/At1} = 0.87 + \frac{1}{8.0*(1-0.50)} = 0.87 + \frac{1}{4.0} = 0.87 + 0.25 = 1.12$$

Como pode ser visto no cálculo feito anteriormente, os valores dos riscos para os Ag1 e Ag2 são, respectivamente, 1.1957 e 1.12. Pelo Passo (c) da negociação, o agente com o menor valor de risco deve ceder, neste caso é o Ag2, que cede e o seu valor de concessão é dado pela Fórmula 3.6, é calculado usando a constante igual a vinte (20), como pode ser visto a seguir.

$$concessao_{Ag2} = \frac{0.5*8.0}{1*20} = \frac{4.0}{20} = 0.20$$

Mais uma vez, como pode ser visto no cálculo feito anteriormente, o valor da concessão para o Ag2 é de 0.2, que é subtraído da confiabilidade do Ag2 sobre a classe B, que agora possui o valor 0.67, pois $0.87 - 0.20 = 0.67$. O agente Ag2 é considerado como o agente que cedeu e não mais participará desta iteração da negociação. Como a classe ganhadora do agente Ag1 é A e a classe ganhadora do Ag3 é C, eles farão parte da negociação. A medida de risco do Ag1 é 1.1957 e a do Ag3, como o At1 com menor similaridade, pode ser vista a seguir.

$$risco_{Ag3/ClasseC/At1} = 0.85 + \frac{1}{7.5*(1-0.60)} = 0.85 + \frac{1}{3} = 0.85 + 0.3333 = 1.1833$$

Mais uma vez, como pode ser visto no cálculo feito anteriormente, o risco para o Ag3 é de 1.1833, ou seja, agora o Ag3 é aquele que cede. Como só existem um agente que não cedeu ainda, que é o Ag1, a primeira iteração é finalizada. Como todos os agentes possuem confiabilidades positivas e o número total de atributos não foi alcançado, uma segunda iteração deve começar.

Capítulo 4

Detalhamento da Investigação

Como foi mostrado no Capítulo 3, o trabalho original do Sistema *NeurAge* (Abreu et al., 2004) apresentou uma nova aplicação para os Sistemas Multiagentes (SMA). No trabalho original, os agentes são capazes de resolver problemas de reconhecimento de padrões. Para que isto fosse possível, foram utilizadas algumas características dos Sistemas Multi-Classificador (SMC). Porém, esse trabalho fez uma exploração superficial dos diversos parâmetros que poderiam ser analisados.

A partir disso, esta dissertação propõe estender a investigação iniciada no primeiro trabalho. Ou seja, será feita uma expansão dos conceitos aplicados inicialmente no Sistema *NeurAge* sendo chamado, agora, de Sistema *ClassAge*. Para isso, fez-se uma análise sobre os parâmetros que poderiam ser mudados e, analogamente a isso, fez-se uma análise comparativa de como esses sistemas se comportariam diante das variações propostas.

Na Seção 4.1 é feita uma explanação de como os primeiros experimentos, utilizando os conceitos do *NeurAge*, foram planejados e executados, além de mostrar os seus resultados. Na Seção 4.2 é feita uma explicação da metodologia que será utilizada neste trabalho, bem como os parâmetros que serão explorados e seus possíveis valores.

4.1 Primeiros Resultados do Sistema *NeurAge*

O Sistema *NeurAge* foi proposto inicialmente em (Abreu et al., 2004) e (Abreu, 2005). Nele foram feitos alguns experimentos para que fosse possível a validação do mesmo. Nesse sistema inicial foram utilizadas algumas configurações padrões que podem ser vistas a seguir.

- Agentes Neurais contendo a rede MLP (*Multi-Layer Perceptron*) (Xiang et al., 2001).

- Sistema contendo quatro agentes.
- Os métodos de negociação:
 - Baseado na Teoria dos Jogos (TeoJog).
 - Baseado no Leilão (Lei).
 - Baseado na Sensibilidade (Sensi).
- Três bases de dados:
 - Proteínas (Tan et al., 2003).
 - Caracteres Numéricos (Hull, 1994).
 - Imagens (Blake and Merz, 1998).
- Dois métodos de Combinação:
 - Voto.
 - Soma.

Os resultados dos experimentos estão na Tabela 4.1. Observando as porcentagens de acerto e os desvios padrões pode-se concluir que foram obtidos resultados satisfatórios. Os trabalhos (Abreu et al., 2005c) e (Abreu et al., 2005b) também se basearam nesses experimentos para sua elaboração, assim como, os trabalhos (Abreu et al., 2005d) e (Abreu and Canuto, 2004).

Base de Dados de Caracteres					
-	Soma	Voto	Sensi	Lei	TeoJog
Acerto	93,81%	91,61%	96,05%	95,17%	95,21%
Desvio Padrão	7,59	10,16	11,93	13,64	13,20
Base de Dados de Imagens					
-	Soma	Voto	Sensi	Lei	TeoJog
Acerto	89,25%	81,58%	93,33%	91,33%	93,33%
Desvio Padrão	2,56	3,03	2,64	2,61	2,64
Base de Dados de Proteínas					
-	Soma	Voto	Sensi	Lei	TeoJog
Acerto	83,00%	74,67%	94,00%	87,00%	94,00%
Desvio Padrão	4,83	5,33	2,72	3,33	3,88

Tabela 4.1: Resultados do Trabalho (Abreu, 2005).

O trabalho apresentado aqui estende a investigação iniciada em (Abreu et al., 2004), explorando tantos os pontos que foram citados anteriormente no início da Seção, como

outros pontos relevantes para o refinamento da análise do novo sistema. Como o trabalho realizado, conseguiu-se alcançar um entendimento melhor dos potenciais e das limitações deste sistema.

4.2 Metodologia do Trabalho

O Sistema *ClassAge* possui muitas possibilidades de parâmetros que podem ser explorados. A partir das configurações que foram mostradas na Seção 4.1 pode-se pensar em vários tipos de variações possíveis, mas as que serão utilizadas nessa proposta de trabalho serão as seguintes:

- Quantidade de componentes do Sistema, ou seja, de classificadores ou agentes.
- Variações nos tipos dos componentes do Sistema, ou seja, sistemas homogêneos ou heterogêneos.
- Tipos de Bases de Dados.

Antes de começar a descrever os pontos supracitados, é importante descrever alguns procedimentos a serem analisados, que serão mostrados nas Subseções 4.2.1, 4.2.2 e 4.2.3. Na Subseção 4.2.4 será mostrado e explicado como funcionam os outros tipo de classificadores que serão utilizados no trabalho. Na Subseção 4.2.5 serão citadas as bases de dados escolhidas. Na Subseção 4.2.6 será explicado como funcionam os outros tipos de classificadores que serão utilizados nesse trabalho. E, por fim, na Seção 4.2.8 será explicado como será administrada a questão do tamanho do sistema.

4.2.1 Validação Cruzada

Para a avaliação da robustez de um classificador, nesse caso um Agente Classificador, pode-se utilizar uma metodologia que executa validação cruzada. O método de amostragem *10-fold-cross-validation* vem sendo uma boa escolha na avaliação da performance de classificadores (Mitchell, 1997).

Esse método divide o conjunto de dados em k partes iguais onde cada treinamento é feito utilizando $k-1$ partes e aquele *fold* que sobrou é usado para o teste. De preferência esses *folds* devem possuir a mesma quantidade de padrões para cada classe. No caso desta dissertação, usar-se o valor de k igual a dez (10), ou seja, a cada execução do treinamento e teste, serão utilizados nove partes para treinamento e uma parte para teste.

4.2.2 Análise Estatística

Um estudo acerca dos métodos de análise estatística foi realizado abordando principalmente três métodos: o teste de McNemar (Mitchell, 1997), o teste para a diferença de duas proporções (Mitchell, 1997) e o t-teste de variância combinada (Levine and Berenson, 2000). Neste estudo, o t-teste de variância combinada foi escolhido para ser utilizado na análise estatística dos resultados obtidos pelos experimentos. Para todas as bases de dados utilizadas nesta investigação, o t-teste foi aplicado usando o método de validação cruzada com dez grupos, como foi citado na Seção 4.2.1, e grau de confiança de acerto de 95%, para definir se as diferenças obtidas são significantemente importantes do ponto de vista estatístico.

O t-teste é um método para comparar dois métodos de aprendizagem e é amplamente utilizado com algoritmos de aprendizagem de máquina (Mitchell, 1997). Este método é utilizado para comparar duas amostras (conjuntos de resultados), e, para melhor entendê-lo, é necessário explicar alguns conceitos estatísticos em que o teste se baseia. Entre eles, o teste utiliza o conceito de inferência estatística para obter estimativas sobre as características de uma população (uma amostra) (Levine and Berenson, 2000).

Considerando uma amostra X com n valores selecionados de uma população de valores, define-se a média aritmética da amostra, como mostrado na Fórmula 4.1.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (4.1)$$

Desta forma, \bar{X} é uma estimativa da média aritmética populacional que normalmente é desconhecida e denominada de μ_x . Também define-se a variância da amostra como mostrado na Fórmula 4.2.

$$S_x^2 = \frac{n * \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2}{n * (n - 1)} \quad (4.2)$$

Sendo assim, S_x^2 é uma estimativa da variância populacional σ_x^2 . Nesse contexto, denomina-se de nível de significância de erro da estimativa, ou α , a probabilidade de erro que a estimativa amostral esteja incorreta. Por exemplo, $\alpha=5\%$ ou 0,05 significa que, se todas as amostras possíveis de um mesmo tamanho n fossem retiradas da população, 5% delas não iriam conter a verdadeira média aritmética da população no intervalo de confiança de $(1 - \alpha)$. O nível de confiança de acerto da estimativa é, então, definido como o complemento do nível de significância de erro da estimativa, ou seja, $(1 - \alpha)$ (Levine and Berenson, 2000).

Todos esses conceitos descritos anteriormente servem para inferir a média aritmética populacional μ , a partir de uma determinada amostra de valores com certo intervalo de confiança $(1 - \alpha)$. Com o objetivo de comparar duas amostras independentes, esses conceitos básicos podem ser entendidos. Neste trabalho, todas as amostras possuem o mesmo tamanho n . Desse modo, a Fórmula 4.3 pode ser utilizada para calcular a diferença de seus intervalos de confiança.

$$-t_{(1-\alpha)/2, 2n-1} \leq \frac{(\bar{X} - \bar{Y}) - (x - y)}{\sqrt{\frac{S_x^2 + S_y^2}{n}}} \leq t_{(1-\alpha)/2, 2n-1} \quad (4.3)$$

Ao se comparar os dois intervalos de confiança, caso o módulo da diferença entre as duas amostras seja maior que $t_{(1-\alpha)/2, 2n-1}$, conclui-se que eles são significantemente diferentes em nível de significância, ou seja α escolhido. Todavia, caso o módulo da diferença entre as duas amostras seja menor ou igual a $t_{(1-\alpha)/2, 2n-1}$, não é possível afirmar relação de equivalência das amostras, mas, apenas, que não há diferenças significativas, detectadas em nível de significância, ou seja, α escolhido.

Esse teste que utiliza a distribuição t de *Student*, para comparar duas amostras independentes, é conhecido como teste t de variância combinada. Esse teste estatístico é robusto, pois não é sensível a distanciamientos moderados do pressuposto da normalidade das amostras (Levine and Berenson, 2000). Inicialmente, assume-se uma hipótese nula, denominada de H_0 . Sempre que se determina a hipótese nula, deve-se especificar a hipótese alternativa, oposta à nula, denominada de H_1 , que será verdadeira, caso H_0 seja falsa (Levine and Berenson, 2000).

O valor p é o melhor modo de analisar os resultados de um teste de hipótese. Um valor de p é o menor nível de significância à qual H_0 é rejeitado quando um procedimento de teste específico é aplicado em um determinado grupo de dados. A simples declaração de que a hipótese H_0 está ou não na área de rejeição não informa o grau em que o valor computado no teste estatístico está na área de rejeição ou de não-rejeição. Uma vez que o valor p tenha sido calculado, pode-se tomar a decisão se as diferenças nos resultados dos dois grupos de teste são significantes ou não do ponto de vista estatístico. Neste caso, se o valor de p for menor que 0,05, os resultados são ditos significantemente diferentes. Caso contrário, as diferenças não são estatisticamente significantes.

4.2.3 Medidas de Diversidade

Os sistemas que combinam classificadores para um melhor desempenho na resolução de problemas de reconhecimento de padrões vem ganhando crescente atenção na literatura.

As saídas dos classificadores são combinadas com a intenção de encontrar uma decisão mais correta que as encontrada a partir de classificadores individuais. Estes classificadores combinados são diferentes entre si, senão esta decisão não seria melhor que as decisões dos classificadores individuais. Esta diferença entre eles, que é também conhecida como diversidade, deve alcançar uma melhor ou pior decisão, sendo assim, conclui-se que tanto existem boas diversidades, quanto más diversidades (Kuncheva and Whitaker, 2001).

Não existe uma escolha única de medida de diversidade ou dependência. Existem diferentes medidas de avaliação da diversidade em diferentes campos de pesquisa (Kuncheva, 2004). Existem medidas que trabalham com paridade (*pairwise*), ou seja, dois a dois, e medidas que não trabalham com paridade (*no-pairwise*), ou seja, a análise é feita individualmente. As do primeiro grupo são calculadas para cada par de classificadores e as do segundo grupo utilizam a idéia de entropia e correlação de saídas individuais ou são baseadas na dificuldade de distribuição dos pontos dos dados trabalhados.

Dentre as medidas que utilizam paridade, pode-se citar quatro principais: o Q estatístico, o coeficiente de correlação, a medida de desacordo e a medida do duplo-falso. Já dentre as medidas que não utilizam paridade, pode-se citar seis principais: a medida de entropia E , a medida de dificuldade θ , a variância de Kohavi-Wolpert, a medida de acordo κ , a diversidade generalizada e a diversidade coincidente da falha (Kuncheva and Whitaker, 2001).

Dentre as medidas citadas anteriormente, foram utilizadas neste trabalho duas medidas que utilizam paridade e uma que não utiliza paridade. Elas serão explicadas a seguir nas Seções 4.2.3.1 e 4.2.3.2.

4.2.3.1 Medidas que utilizam paridade

Como já foi mencionado anteriormente, existem várias medidas de similaridade que trabalham com o conceito de paridade. Dentre as principais, foram escolhidas para serem aplicadas neste trabalho o Q estatístico e a medida do duplo-falso.

O Q Estatístico: O Q estatístico de Yule (Kuncheva, 2004) para dois classificadores, D_i e D_j é dados pela Fórmula 4.4.

$$Q_{i,j} = \frac{N^{11}N^{00} - N^{01}N^{10}}{N^{11}N^{00} + N^{01}N^{10}} \quad (4.4)$$

onde N^{11} é o número de padrões onde ambos os classificadores erram; N^{00} é o número de padrões onde ambos os classificadores acertam; N^{01} é o número de padrões onde o

primeiro classificador erra e o segunda acerta; N^{10} é o número de padrões onde o primeiro classificador acerta e o segunda erra;

Para classificadores estatisticamente independentes, espera-se que o valor de Q seja 0. Caso contrário o valor de Q pode variar de -1 a 1. Essa variação se comporta da maneira mostrada a seguir.

- Quando os classificadores tendem a acertar ou errar os mesmos padrões, eles tendem a possuir valores de Q positivos.
- Quando os classificadores tendem a acertar ou errar padrões diferentes, eles tendem a possuir valores de Q negativos.

Sendo assim, pode-se afirmar que os classificadores são considerados mais diversos quando possuem valores de Q negativos. Por outro lado, pode-se afirmar que os classificadores são menos diversos quando possuem valores de Q positivos.

A **medida de duplo-falso** (Giacinto and Roli, 2001): Essa medida utiliza a proporção de casos que em que existem erros de classificação para ambos os classificadores a define conjuntos. A Fórmula 4.5 mostra como isso é feito.

$$df = \frac{N^{00}}{N^{00} + N^{11} + N^{01} + N^{10}} \quad (4.5)$$

É importante observar que isso é uma medida de não similaridade, ou seja, quanto maior o valor df , menor a similaridade entre os classificadores e quanto menor esse valor, maior é essa similaridade.

4.2.3.2 Medidas que não utilizam paridade

Como já foi mencionado anteriormente, existem várias medidas de similaridade que não trabalham com o conceito de paridade. Dentre as principais, foi escolhida para ser aplicada neste trabalho a Medida de Entropia.

Medida de Entropia (Kuncheva, 2004): Essa é uma medida baseada na suposição que a maior diversidade entre classificadores é mostrada a por $\frac{L}{2}$ dos votos em y , onde L é a quantidade de classificadores. Se todos eles são 0's ou 1's, não existe similaridade e eles são diversos. Uma possível maneira de retratar isso é mostrada na Fórmula 4.6.

$$e = \frac{1}{N} \sum_{m=1}^N \frac{1}{L - \frac{L}{2}} \min\{l(z_m), L - l(z_m)\} \quad (4.6)$$

$l(z_m)$ é o número de classificadores que reconhecem corretamente z_m . O e varia entre 0 e 1, onde 0 indica nenhuma diferença e 1 indica a mais alta diversidade possível.

4.2.4 Métodos de Classificação

A pesquisa inicial do *NeurAge* envolvia apenas agentes com módulo de reconhecimento de padrões contendo Redes MLP (Abreu et al., 2004). De forma a englobar uma maior quantidade de problemas foram sugeridos outros tipos de classificadores, não excluindo a MLP, os quais são apresentados a seguir.

4.2.4.1 Redes Neurais

Neste trabalho foram utilizados três tipos de redes neurais: Redes do Tipo MLP (*Multi-Layer Perceptron*) (Xiang et al., 2001), FuzzyMLP (*Fuzzy Multi-Layer Perceptron*) (Canuto, 2001) e RBF (*Radial Basis Function*) (Park et al., 2002). A seguir elas são descritas sucintamente.

As Redes Neurais MLP são redes de múltiplas camadas onde cada camada tem uma função específica. A camada de saída recebe os estímulos da camada intermediária e constrói o padrão que será a resposta. As camadas intermediárias funcionam como extratoras de características, seus pesos são uma codificação de características apresentadas nos padrões de entrada e permitem que a rede crie sua própria representação, mais rica e complexa, do problema.

Durante o treinamento com o algoritmo *backpropagation*, a rede opera em uma sequência de dois passos. Primeiro, um padrão é apresentado à camada de entrada da rede. A atividade resultante flui através da rede, camada por camada, até que a resposta seja produzida pela camada de saída. No segundo passo, a saída obtida é comparada à saída desejada para esse padrão particular. Se esta não estiver correta, o erro é calculado. O erro é propagado a partir da camada de saída até a camada de entrada, e os pesos das conexões das unidades das camadas internas vão sendo modificados conforme o erro é retropropagado.

Existem várias extensões do algoritmo *backpropagation*, tais como RProp, Quick-Prop, FuzzyMLP, dentre outras. A extensão FuzzyMLP (*Fuzzy Multi-Layer Perceptron*) (Canuto, 2001) é uma implementação da teoria dos conjuntos fuzzy em uma rede neural perceptron de múltiplas camadas. Ou seja, é o resultado da *fuzzificação* direta no nível de rede, no nível de aprendizado ou em ambos na rede MLP. Além disso, a saída desejada é calculada de forma diferente que nas Redes MLP, no qual utilizam o método do

winner-takes-all. Sendo assim, os nodos correspondentes a saída desejada são modificados durante a fase de treinamento dando origem a saída desejada fuzzy.

Uma rede neural RBF (*Radial Basis Function*) (Park et al., 2002), ou seja, com função de ativação de base radial pode ser vista como uma rede *feedforward* de múltiplas camadas, contendo três camadas com papéis completamente distintos. A camada de entrada é composta por unidades sensoriais que conectam a rede a seu ambiente. A segunda camada (única camada intermediária) aplica uma transformação não-linear do espaço de entrada para o espaço intemediário. As unidades intermediárias não-lineares têm ativação local e suas respostas correspondem às ativações das funções de base radial.

4.2.4.2 Máquinas de Vetores Suporte

As SVM (*Support Vector Machines*) (Meir, 2002), ou Máquinas de Vetores Suportes foram primeiramente aplicadas em Classificação de Textos no final dos anos 90. De acordo com a teoria deste método, enquanto técnicas tradicionais para reconhecimento de padrões são baseadas na minimização do risco empírico, ou seja, tenta-se otimizar o desempenho sobre o conjunto de treinamento, as SVM minimizam o risco estrutural, ou seja, tenta-se minimizar a probabilidade de classificar de forma errada padrões ainda não vistos por uma distribuição de probabilidade dos dados fixa e desconhecida.

De forma direta, este método é um novo princípio de indução que é equivalente a minimizar um limite superior do erro de generalização, dependendo da teoria de convergência uniforme de probabilidade. SVM divide o espaço de termos em hiperplanos ou superfícies, separando as amostras de treinamento positivas das negativas. Algumas vezes estas superfícies são referidas como superfícies de decisão. Então a superfície que provê a maior separação (a maior margem possível entre as amostras positivas e negativas) é selecionada.

4.2.4.3 *K*-Vizinhos Mais Próximos

O K-nn (*K-nearest neighbors*) (Jiangsheng, 2002) é um dos métodos de aprendizado mais simples. Ele é um classificador onde o aprendizado é baseado em analogia. O conjunto de treinamento é formado por vetores de n dimensões e cada elemento deste conjunto representa um ponto no espaço n -dimensional. Ele funciona de maneira que é calculado os k 's vizinhos mais próximos de um determinado conjunto de dados. Isso é feito usando métricas específicas, como por exemplo a distância euclidiana.

Para determinar a classe de um elemento que não pertença ao conjunto de treinamento, o classificador procura k elementos do conjunto de treinamento que estejam mais

próximos deste elemento, ou seja, que tenha, a menor distância. Estes k elementos são chamados de k -vizinhos mais próximos. Verifica-se quais são as classes desses k vizinhos e a classe mais freqüente será à atribuída ao padrão de entrada.

O K -nn é um classificador que possui apenas um parâmetro livre, no caso o número de K -vizinhos. Este número é controlado pelo usuário com o objetivo de obter uma melhor classificação. Este processo de classificação pode ser computacionalmente exaustivo se considerado um conjunto com muitos dados. Para determinadas aplicações, no entanto, o processo é bem aceitável.

4.2.4.4 Árvores de Decisão

Uma árvore de decisão (*Decision Trees*) (Quinhan, 1993) utiliza a estratégia de dividir para conquistar. Nelas, um problema complexo é decomposto em sub-problemas mais simples e recursivamente, a mesma estratégia é aplicada a cada sub-problema. A capacidade de discriminação de uma árvore vem da divisão do espaço definido pelos atributos em sub-espacos e cada sub-espaco é associado uma classe.

Este método consiste de uma hierarquia de nós internos e externos que são conectados por ramos. O nó interno, também conhecido como nó decisório ou nó intermediário, é a unidade de tomada de decisão que avalia através de teste lógico qual será o próximo nó descendente ou filho. Em contraste, um nó externo, também conhecido como folha ou nó terminal, está associado a um rótulo ou um valor. Em geral, o procedimento de uma árvore de decisão é dado a seguir.

1. Apresenta-se um conjunto de dados ao nó inicial (ou nó raiz que também é um nó interno) da árvore;
2. Dependendo do resultado do teste lógico usado pelo nó, a árvore ramifica-se para um dos nós filhos;
3. Este procedimento é repetido até que um nó terminal seja alcançado.

De forma resumida, em uma árvore de decisão, a classificação de um caso se inicia pela raiz da árvore, e esta árvore é percorrida até que se chegue a uma folha. Em cada nó de decisão será feito um teste que irá direcionar o padrão de teste para uma sub-árvore. Este processo irá guiar-se para uma folha. A classe do padrão se presupõe que seja a mesma que está armazenada nesta folha. O método utilizado neste trabalho foi o J4.8.

4.2.4.5 JRip

Muitas das técnicas utilizadas pelo aprendizado moderno vêm sendo adaptadas para o aprendizado e árvores de decisão. A maioria das árvores de decisão utiliza uma estratégia de sobrecarregar para simplificar para tratar dados com ruído. Nesta estratégia, uma hipótese é inicialmente formada gerando uma árvore complexa que super-utiliza os dados, e assim, simplifica a árvores através de técnicas de poda.

Normalmente as técnicas de poda melhoram taxas de erros de dados não vistos quando o conjunto de dados possui ruído. Uma grande variedade de métodos vem sendo proposto para podar árvores e uma das técnicas mais eficientes é a poda reduzindo o erro (*REP - Reduced Error Pruning*) (Cohen, 1995). O REP pode facilmente adaptar sistemas de regras de aprendizado. Existem algumas variações deste método e uma delas é o *IREP - Incremental Reduced Error Pruning* (Cohen, 1995).

O (*IREP - Incremental Reduced Error Pruning*) integra firmemente a poda pela redução do erro com um algoritmo da regra de aprendizado que trabalha com a regra dividir para conquistar. Este algoritmo possui um conjunto de regras e testa todas as regras, uma por vez. Depois que uma regra é encontrada, todos os exemplos que são cobertos pela regra são deletados. Este processo é repetido até que não exista exemplos corretamente classificados, ou até que a regra encontrada pelo *IREP* possua um erro inaceitável.

A abordagem que repete o crescimento e a simplificação usada no *IREP* pode produzir resultados diferentes da poda para redução de erro convencional. Um caminho para melhorar a abordagem incremental do *IREP* é adiar o processo de produção de regras deste método. Dessa maneira, esse método se aproxima ao máximo do método de poda pelo erro. Esta otimização é conhecida como JRip (*Optimized IREP - Incremental Reduced Error Pruning*) (Cohen, 1995).

4.2.5 Bases de Dados

As bases de dados utilizadas nos primeiros experimentos feitos na utilização dos conceitos do Sistema *NeurAge* foram citadas na Seção 4.1. São elas a base de Caracteres Numéricos, Imagens e Proteínas. Cada uma delas possui suas características próprias, porém todas são bases contendo um número suficiente de padrões para treinamento e um número reduzido de atributos em relação a quantidade de seus padrões.

Neste trabalho foi pensado em usar outras bases de dados para uma nova avaliação dos métodos propostos. Logo, serão utilizadas as bases de Splice, Breast Cancer e Wisconsin, além da Imagem e Proteínas. As configurações básicas de cada base podem ser vistas

Base de Dados de Imagens		
Qtd Padrões	Qtd Atributos	Qtd Classes
2100	18	7
Base de Dados de Proteínas		
Qtd Padrões	Qtd Atributos	Qtd Classes
530	126	5
Base de Dados de Splice		
Qtd Padrões	Qtd Atributos	Qtd Classes
2890	60	3
Base de Dados de Breast Cancer		
Qtd Padrões	Qtd Atributos	Qtd Classes
620	9	2
Base de Dados de Wisconsin		
Qtd Padrões	Qtd Atributos	Qtd Classes
510	30	2

Tabela 4.2: Características das Bases de Dados deste trabalho.

na Tabela 4.2. Cada uma dessas bases possui suas próprias características. Nas Seções a seguir estão descritas mais detalhadamente cada uma das bases.

4.2.5.1 Base Breast Cancer

A Base Breast Cancer (Mangasarian and Wolberg, 1992), que aqui será sempre referenciada como Base A, trata do câncer de mama e foi desenvolvida pela Universidade de Wisconsin. As instâncias foram extraídas de imagens digitalizadas da massa do seio e descrevem as características do núcleo celular presente na imagem.

Os nove atributos desta base são denominados *Clump Thickness*, *Uniformity of Cell Size*, *Uniformity of Cell Shape*, *Marginal Adhesion*, *Single Epithelial Cell Size*, *Bare Nuclei*, *Bland Chromatin*, *Normal Nucleoli* e *Mitoses*. Esses atributos são nominais e representam os possíveis deoxinucleotídeos do DNA que são adenina, citosina, guanina e timina. Neste caso esses valores foram transformados para valores numéricos.

A base possui duas classes no qual a primeira representa ter câncer e a segunda representa não ter câncer. Nesta base existem 682 instâncias, onde 620 são de treinamento e teste e 62 são de validação. No conjunto de treinamento, 210 padrões são da classe do tipo 1, que representa ter câncer, e 410 são da classe do tipo 2, que representa não ter câncer. O conjunto de validação possui 21 padrões do tipo 1, que representa ter câncer, e 41 do tipo 2, que representa não ter câncer. É importante destacar que esta base de dados

é uma base desbalancedada.

4.2.5.2 Base de Imagem

A Base de Imagem (Blake and Merz, 1998), que aqui será referenciada como Base B, foi criada pelo *Vision Group* da Universidade de Massachusetts. Nesta base, as instâncias (cada instância é uma região de 3x3) foram aleatoriamente extraídas de uma base de dados de 7 imagens segmentadas para criar uma classificação para cada pixel.

Esta base de dados possui dezenove atributos reais que foram extraídos das regiões das imagens, mas só foram utilizados 18 atributos, pois um deles era um atributo contínuo. No total, existem 7 diferentes classes de regiões, *brickface*, *sky*, *foliage*, *cement*, *window*, *path* e *grass*. Na Tabela 4.3 pode ser observada representação numérica de cada uma das classes.

Classe	Representação
<i>brickface</i>	0.000000
<i>sky</i>	0.1667
<i>foliage</i>	0.3333
<i>cement</i>	0.5
<i>window</i>	0.6667
<i>path</i>	0.8333
<i>grass</i>	1.000000

Tabela 4.3: Representação Numérica da Base de Dados de Imagem.

Esta base contém 2310 instâncias, onde cada uma das sete classes possui 330 instâncias. Foram utilizadas 2100 instâncias para treinamento e teste, onde cada uma das sete classes possui 300 representantes. Já para validação, foram utilizadas 210 instâncias, onde cada uma das sete classes possui 30 representantes.

4.2.5.3 Base de Proteínas

A Base de Proteínas (Tan et al., 2003), que aqui será referenciada como Base C, é uma classificação hierárquica, detalhada manualmente, derivada de estruturas conhecidas de proteínas e organizada de acordo com seus relacionamentos evolucionários e estruturais. Ela é dividida em quatro níveis hierárquicos: classe, dobra, superfamília e família.

Aqui neste trabalho é utilizada a parte dessa base que contém a classificação das classes de proteínas. As principais classes de proteínas são $\text{all-}\alpha$, $\text{all-}\beta$, $\alpha + \beta$, $\frac{\alpha}{\beta}$ e *small*.

Nesta base existem 125 seqüências que caracterizam as classes de proteínas. As propriedades de cada seqüência contém 21 características contínuas, a menos da última que contém 20. A partir disso, conclui-se que esta base possui 126 atributos, pois são 125 seqüências e mais um atributo de rótulo. Estes atributos são nominais e representam os possíveis deoxinucleotídeos do DNA que são adenina, citosina, guanina e timina. Neste caso esses valores foram transformados para valores numéricos.

Classe	Quantidade de Padrões
all- α	111
all- β	177
$\alpha + \beta$	203
$\frac{\alpha}{\beta}$	46
<i>small</i>	45

Tabela 4.4: Distribuição de padrões da Base de Dados de Proteínas.

Esta base possui 584 padrões distribuídos de acordo com a Tabela 4.4. Foram utilizados 531 padrões para treinamento e teste e 53 padrões para validação. No conjunto de treinamento e validação existem, respectivamente, 101 e 10 padrões do tipo all- α , 159 e 18 padrões do tipo $\frac{\alpha}{\beta}$, 186 e 17 padrões do tipo all- β , 42 e 4 padrões do tipo $\alpha + \beta$ e 41 e 4 padrões do tipo *small*. É muito importante destacar que esta base de dados é desbalanceada.

4.2.5.4 Base de Splice

A Base de Splice (Towell and Shavlik, 1992), que aqui será referenciada como Base D, é de junções de *splice* de seqüências de DNA, com teoria imperfeita associada do domínio. Junção de *Splice* são trechos na seqüência de DNA que são removidos dele durante o processo de criação de proteínas nos mais altos organismos. Assim como a base A, esta base foi desenvolvida pela Universidade de Wisconsin.

Esta base possui 60 atributos que descrevem as seqüências de DNA usadas no processo de criação da proteína. Esses atributos, como na Base de Proteínas, são nominais e representam os possíveis deoxinucleotídeos do DNA. Neste caso também, esses valores foram transformados para valores numéricos.

As possíveis classes desta base são de três tipos que representam o tipo de junção de seqüência de DNA. Podem ser *intron*, representada pela classe 2, *exon*, representada pela classe 3 ou nenhuma das duas, representada pela classe 1. Ao total esta base de dados possui 3179 instâncias. A distribuição para os arquivos de treinamento e validação são,

respectivamente, de 1500 e 150 para a classe 1, de 700 e 70 para classe 2 e de 690 e 69 para a classe 2.

4.2.5.5 Base de Wisconsin

A Base de Wisconsin (Street et al., 1995), que aqui será referenciada como Base E, foi obtida de hospitais da Universidade de Wisconsin, esta base contém dados sobre câncer de mama assim como na Base de Breast-Cancer. Suas instâncias estão separadas em duas classes, benigno e maligno de acordo com o tipo do câncer. Esta base possui 30 atributos que representam amostras de informação sobre câncer.

Ela possui 561 instâncias e a representação numérica para as classes é maligno 1 e benigno 0. A distribuição delas para os arquivos de treinamento e validação são, respectivamente, 330 e 33 para a classe 0 e 190 e 19 para a classe 1.

4.2.6 Métodos de Combinação

Como foi dito anteriormente, para analisar se as modificações propostas no Sistema *ClassAge* possuem maior eficácia que os correspondentes SMCs foram feitas comparações do *ClassAge* com alguns métodos de combinação que também podem ser chamados de métodos baseados em fusão.

Dentre os método de combinação, que visam aproveitar as várias vantagens que os métodos individuais possuem (Mitchell, 1997), foram escolhidos alguns métodos para a análise. Existem vários métodos baseados em fusão na literatura. Eles pertencem a esse grupo pois, fazem a classificação através da fusão de várias informações. Vários grupos podem ser citados, dentre os a seguir.

- Métodos de Combinação Linear: São os métodos mais simples de combinar múltiplos classificadores, eles executam combinação linear das saídas dos classificadores. Os métodos deste tipo que foram utilizados neste trabalho são Soma (*Sum*) (Kittler and Alkoot, 2003), Média (*Average*) (Kuncheva, 2002b) e Mediana (*Median*) (Kuncheva, 2002b).
- Métodos Não Lineares: São métodos que realizam combinação não linear das saídas dos classificadores. Em geral, são mais complexos que os métodos de combinação linear. Nesta classe de métodos, incluem-se combinadores baseados em ordenação. Dentre eles o utilizado aqui é o Voto (*Majority Vote*) (Stefano et al., 2002).

- Métodos baseados em Estatística: São métodos que utilizam probabilidade ou combinações estatísticas para combinar as saídas dos classificadores. O método deste tipo que foi utilizado neste trabalho foi o *Naive Bayes* (Davis et al., 2004).
- Métodos baseados em Inteligência Computacional: Aqui os métodos fazem a combinação das saídas dos classificadores a partir de técnicas de inteligência computacional. Os utilizados neste trabalho são o FuzzyMLP (*Fuzzy Multi-Layer Perceptron*) (Canuto, 2001) e o MLP (*Multi-Layer Perceptron*) (Xiang et al., 2001).

Como citado anteriormente, serão utilizados neste trabalho a Soma, a Média, a Mediana, o Voto, o *Naive Bayes*, o FuzzyMLP e o MLP. Como os dois últimos já foram descritos na Seção 4.2.4.1, pois também são utilizados como classificadores em alguns dos experimentos, não serão explicados novamente. A seguir serão descritos os demais métodos que serão usados.

O *Naive Bayes* (Davis et al., 2004) implementa o classificador Bayes, no qual para cada classe de decisão é calculada a probabilidade condicional de que essa classe de decisão é correta. Nesse método de classificação assume-se que os atributos dos exemplos são independentes dada a classe. As probabilidades envolvidas são calculadas como frequência obtida da tabela de decisão.

Um classificador *Naive Bayes* (Davis et al., 2004) é um classificador probabilístico simples baseado na aplicação do teorema de Bayes com suposições de forte independência. Dependendo da natureza precisa do modelo probabilístico, estes classificadores podem ser treinados com bastante eficiência em um aprendizado supervisionado. Em algumas aplicações práticas, o parâmetro estimação para os modelos do *Naive Bayes* utiliza o método da máxima probabilidade; em outras palavras, o classificador pode trabalhar sem acreditar na probabilidade bayesiana ou utilizando qualquer método bayesiano. De acordo com o modelo do *naive* e a sua aparente simplificação de suposições, este classificador, muitas vezes, trabalha melhor com situações complexas do mundo real do que se é esperado dele.

O método Soma, funciona de maneira que, quando apresentado um padrão de teste para os classificadores, todas as saídas correspondentes de cada classificador são somadas e a classe vencedora é aquela que possuir o maior valor absoluto. Já o Voto, funciona de maneira que, quando apresentado um padrão de teste para os classificadores, cada um deles vota na classe que ele achar que é a correta. A classe vencedora é aquela que possuir o maior número de votos. O método Média calcula a média aritmética dos valores das saídas dos classificadores, já a mediana calcula o valor mediano entre as saídas destes classificadores.

4.2.7 Métodos Baseados em Seleção

Como já mencionado anteriormente, diferentemente dos combinadores baseados em fusão, nos métodos baseados em seleção, somente um classificador é necessário para classificar um padrão de entrada. Para que isto ocorra, é importante definir um processo de escolha do membro do sistema para tomar a decisão. O fato de apenas um classificador ser escolhido para realizar a classificação faz com que estes métodos sejam denominados de seleções. A escolha deste classificador que fará a rotulação do padrão de entrada acontece durante a fase de operação ou teste. Esta escolha é, normalmente, baseada na certeza da decisão atual. A preferência é dada para os classificadores que possuem maior certeza dos seus resultados (Giacinto and Roli, 1999).

Existem vários métodos baseados em seleção na literatura. Alguns deles são considerados híbridos, pois só efetuam a seleção se o classificador tiver uma certeza estatística de que o seu resultado está correto. Dentre os métodos existentes foram escolhidos os citados a seguir, onde os dois últimos são considerados híbridos.

- DCS-LA (*Dynamic Classifier Selection based on Local Accuracy*) (Giacinto and Roli, 1999).
- DCS-DT (*Dynamic Classifier Selection using Decision-Theoretic*) (Kuncheva, 2002a).
- DCS-MCB (*Dynamic Classifier Selection based on Multiple Classifier Behaviour*) (Giacinto and Roli, 2000).

A seguir, nas seções 4.2.7.1, 4.2.7.2 e 4.2.7.3, uma descrição de cada um desses métodos é feita.

4.2.7.1 DCS-LA

O Classificador de Seleção Dinâmica baseado em exatidão local, também conhecida como LCA (*local class accuracy*), utiliza análise de competência para o classificador que denomina um padrão de entrada com um determinado rótulo ou classe. O principal passo para calcular a exatidão local de uma classe (LCA) para um padrão de teste x pode ser definido como:

1. Pega-se os rótulos fornecidos por todos os classificadores (D).
2. Para cada classificador ($D_i, i = 1, \dots, L$) encontre k pontos mais próximos de x para cada D_i que provê o mesmo rótulo.

3. Calcule a proporção de pontos no qual D_i provê o rótulo correto e faça com que essa proporção seja a exatidão do classificador (LCA).
4. Escolha o classificador com o maior LCA. Nesse caso podem acontecer três coisas:
 - (a) Se existe somente um vencedor, ele rotula o padrão x .
 - (b) Se dois classificadores empatam, escolha o terceiro.
 - (c) Se todos os classificadores empatam, escolha uma classe randomicamente entre os rótulos empatados.

A seleção do classificador mais apropriado é feita durante a fase de teste e acontece em caso de haver um desacordo entre os classificadores. Os métodos baseados em seleção são muito sensíveis a variações no tipo dos membros do *ensemble* e na diversidade destes. Isso significa que a escolha do tipo de classificador que vai fazer parte do sistema é muito importante para o bom desempenho de um DCS.

4.2.7.2 DCS-MCB

Como dito anteriormente, o Classificador de Seleção Dinâmica baseado em classificador de comportamento múltiplo é um método híbrido, pois este método pode optar utilizar ou não o classificador selecionado, quando este não tiver certeza estatística que está correto. Este método possui duas principais diferenças entre o anterior. Primeiramente, k é variável e, segundo, um classificador é selecionado se e somente se o maior LCA é substancialmente maior que os valores dos LCA's dos demais classificadores. Por outro lado, o padrão de teste é classificado através da aplicação da técnica do Voto. Os principais Passos do DCS-MCB são descritos a seguir.

1. Para cada padrão de teste, selecione os k 's vizinhos mais próximos.
2. Selecione somente os vizinhos que possuem maior similaridade que um limiar que é definido previamente.
3. Calcule a competência de cada classificador para os vizinhos selecionados.
4. Se o melhor classificador é substancialmente maior que os outros, selecione ele.
5. Caso contrário, utilize uma técnica de votação.

A principal diferença entre este método e o DCS-LA é o fato do DCS-MCB ser um método híbrido, pois pode optar, quando não tiver uma certeza estatística, entre escolher o classificador com o maior LCA ou executar uma técnica de Votação.

4.2.7.3 DCS-DT

Como dito anteriormente, o Classificador de Seleção Dinâmica é um método híbrido e pode ser usado com modelos de decisão. Como o método mostrado anteriormente, se o classificador que foi selecionado, ou seja, o que foi escolhido como sendo o mais competente para rotular um padrão de entrada, não tiver uma forte certeza, neste caso uma certeza estatística, o método pode escolher entre usar este classificador ou técnicas de fusão. Os principais Passos para esse método são descritos a seguir.

1. Realiza o processo de treinamento de todos os classificadores;
2. Agrupa-se os padrões de treinamento em conjuntos, usando o procedimento de agrupamento K-médias;
3. A exatidão da classificação de todos os classificadores é estimada;
4. O classificador com o maior LCA, ou seja, com a melhor classificação é definida para cada conjunto;
5. Para cada padrão de teste, faça:
 - (a) A partir do padrão de entrada, encontra-se o conjunto com o centro mais próximo deste;
 - (b) Se o melhor classificador do conjunto mais próximo é significativamente melhor que os outros, ele rotula o padrão de entrada;
 - (c) Caso contrário, uma técnica de fusão é usada com todos os classificadores.

A principal diferença entre este método e o DCS-LA é, igualmente ao DCS-MCB, o DCS-DT é um método híbrido. Isto possibilita que um teste estatístico seja feito para definir se o melhor classificador, no caso o que possui o maior LCA, é significativamente diferente dos outros. Já a principal diferença entre o DCS-DT e o DCS-MCB é que o primeiro utiliza, como segunda opção, uma técnica de votação, já o segundo utiliza uma técnica de combinação baseada em fusão.

4.2.8 Topologias dos Sistemas

Dentre as alterações propostas neste trabalho está a variação na topologia dos sistemas. Variar a topologia de um sistema significa mudar a quantidade de membros deste sistema, ou mudar o tipo destes membros, ou, ainda, realizar as duas modificações anteriores ao mesmo tempo. Os trabalhos encontrados na literatura apontam que quando analisam-se as

reações dos sistemas, a variação da diversidade destes só é relevante quando o tamanho do sistema é igual ou inferior a dez (Abreu et al., 2006b). Por este motivo, foram escolhidos os sistemas com: Três Componentes, Cinco Componentes, Sete Componentes e Nove Componentes.

Como foi dito anteriormente, a variação da topologia pode significar variação na quantidade de componentes ou no tipo dos componentes. Neste trabalho, além da variação na quantidade de componentes, serão feitas variações nos tipos dos componentes. Neste caso, as variações serão com três, cinco e sete tipos de componentes. Por exemplo, no caso de experimentos com três classificadores serão feitos experimentos com três classificadores de tipos diferentes ou dois de tipos iguais possuindo configurações diferentes e um diferente destes dois.

A partir do que foi visto, pode-se concluir que serão utilizados sistemas híbridos ou heterogêneos e não híbridos ou homogêneos. No trabalho (Abreu et al., 2004) foram utilizados apenas sistemas não-híbridos. Um sistema Não Híbrido é aquele que possui apenas classificadores do mesmo tipo, por exemplo, um sistema formado apenas por SVM's ou por K-nn's. Ao contrário dos sistemas não híbridos, os sistemas híbridos são aqueles onde existem pelo menos dois métodos de classificação diferentes. Por exemplo, um sistema com quatro componentes formados por dois classificadores SVM's e dois classificadores de Árvore de Decisão.

Capítulo 5

Experimentos e Análises de Resultados

O Sistema *NeurAge* propôs, como mencionado anteriormente, uma nova forma de trabalhar com agentes para reconhecimento de padrões. Inicialmente foram feitos alguns testes que comprovaram a sua eficiência. Porém, ainda é possível aprofundar a investigação desse sistema, pois muitos dos parâmetros dele ainda foram pouco explorados.

O trabalho proposto aqui é exatamente investigar mais aprofundadamente os conceitos utilizados no *NeurAge*, porém com uma abordagem mais abrangente criando o *ClassAge* (*Classifier Agents System*), um Sistema Multiagentes Classificador.

Com o intuito de identificar mais facilmente as topologias das estruturas utilizadas aqui foi adotada uma nomenclatura especial. Como já foi dito anteriormente, serão trabalhados dois tipos principais de estruturas: as não híbridas, com apenas um tipo de componente e as híbridas com tipos diferentes de componentes. Neste trabalho, os sistemas com apenas um tipo de componente, ou seja, um sistema não híbrido, serão chamados de *NH*, sistemas com três tipos de componentes serão chamados de *H3*, os sistemas com cinco tipos de componentes serão chamados de *H5* e, por fim, os sistemas com sete tipos de componentes serão chamados de *H7*.

Como existem muitas possibilidades para cada uma das estruturas citadas anteriormente, este trabalho apresenta a média de todas as possibilidades para as correspondentes estruturas. Por exemplo, quando se trabalha com sistemas com três componentes e usa-se *H3*, levando em consideração que estão sendo usados sete tipos de classificadores diferentes, tem-se 35 possibilidades de combinação. Dessa maneira, sistemas com três tipos de componentes (*H3*) representam a média das 35 possibilidades. Essa metodologia foi adotada para todos os tamanhos de sistemas. Desta maneira fica mais fácil fazer uma melhor análise dos resultados.

Nas tabelas 5.2, 5.7, 5.12 e 5.17 são mostrados a média de acurácia e o desvio padrão de todos os sistemas de classificação para 3, 5, 7 e 9 componentes, respectivamente, para

todas as cinco bases de dados para sistemas baseados em fusão. Já nas tabelas 5.3, 5.8, 5.13 e 5.18 são mostrados a média de acurácia e o desvio padrão de todos os sistemas de classificação para 3, 5, 7 e 9 componentes, respectivamente, para todas as cinco bases de dados para sistemas baseados em seleção. E nas tabelas 5.4, 5.9, 5.14 e 5.19 são mostrados a média de acurácia e o desvio padrão de todos os sistemas de classificação para 3, 5, 7 e 9 componentes, respectivamente, para todas as cinco bases de dados os métodos de negociação.

Como foi citado anteriormente, neste trabalho foram analisadas as diversidades dos sistemas. Foram utilizadas três medidas de diversidade: o Q estatístico, o dupla-falsa e a medida de entropia. Cada uma delas está representada nas Tabelas 5.6, 5.11, 5.16 e 5.21 como sendo, q , df , e , respectivamente. Essas diferenças foram calculadas em relação as diferentes topologias dos sistemas, ou seja, entre os sistemas não híbridos e os híbridos.

Outra informação importante que será analisada é a quantidade de vezes que os métodos híbridos baseados em seleção, ou seja, o DCS-DT e o DCS-MCB, escolhem não usar a seleção. Esta informação está exposta nas Tabelas 5.3, 5.8, 5.13 e 5.18. No caso do DCS-DT será analisada a porcentagem de vezes que este método utilizou um método de fusão ao invés do melhor classificador. Já no caso do DCS-MCB, será analisada a porcentagem de vezes que este método utilizou o método de votação.

A maioria dos métodos de negociação executa algumas iterações para que se chegue a uma decisão em comum. No caso do método do Leilão, isso não acontece, mas na Teoria dos Jogos e no método de sensibilidade, isto ocorre. Esses valores estão representados nas Tabelas 5.4, 5.9, 5.14 e 5.19 como *iteracaoSensibi* e *iteracaoTJ*.

Mais uma informação importante que deve ser analisada é a diferença entre os melhores resultados e os piores resultados dentro das diferentes topologias dos sistemas. Por exemplo, observam-se os resultados dos sistemas não híbridos e híbridos para sistemas com nove componentes, se o menor desempenho foi o do sistema *NH* e o maior desempenho foi do sistema *H5*, então essa diferença vai ser entre os valores das confiabilidades deste sistemas. Essas diferenças estão representados na coluna *Dif*. Essa medida mostra como o resultado do sistema é afetado pela diversidade dos componentes.

E, por fim, outra informação relevante é o valor do teste estatístico que é feito para analisar se o melhor resultado é estatisticamente melhor que o pior resultado. Essa informação está representada na coluna p . Como já foi dito anteriormente, foi utilizado uma taxa de 5% de aceitação. Isto significa que quando o valor do p é menor que 0,05, pode-se afirmar que o sistema com confiabilidade melhor é estatisticamente melhor que o de menor desempenho.

5.1 Métodos Individuais

Antes de começar a análise do desempenho do Sistema *ClassAge*, é importante observar o desempenho dos classificadores individuais. Foram feitas nove topologias, ou configurações, diferentes de cada um dos métodos de classificação utilizados. Além disso, foi utilizado o método de amostragem *10fold-cross-validation*. Por motivo de simplicidade, nesta dissertação, será utilizada, como parâmetro de comparação, a média destes nove experimentos. A Tabela 5.1 mostra este desempenho dos classificadores individuais utilizados nos sistemas. Os valores dos parâmetros de todas as nove configurações dos classificadores foram escolhidas de acordo com as melhores porcentagens de corretos e os menores desvios padrões. Os valores presentes na Tabela 5.1 são a *Media ± DesvioPadrao*.

Classificadores Individuais					
-	Base A	Base B	Base C	Base D	Base E
K-nn	75, 24 ± 3, 26	77, 67 ± 2, 33	70, 83 ± 5, 29	76, 59 ± 2, 59	73, 69 ± 5, 99
SVM	81, 74 ± 4, 21	83, 84 ± 2, 11	75, 81 ± 4, 87	82, 66 ± 3, 84	78, 26 ± 5, 21
MLP	88, 21 ± 2, 71	90, 26 ± 2, 03	79, 51 ± 3, 26	89, 22 ± 3, 99	83, 64 ± 3, 66
FMLP	89, 22 ± 2, 94	91, 34 ± 2, 36	84, 37 ± 3, 12	90, 26 ± 2, 64	86, 81 ± 3, 22
RBF	86, 23 ± 2, 99	89, 63 ± 2, 91	82, 34 ± 3, 59	85, 29 ± 2, 66	83, 67 ± 4, 52
Árvore de Decisão	80, 21 ± 3, 71	84, 26 ± 3, 26	78, 84 ± 4, 26	82, 54 ± 3, 59	85, 84 ± 4, 91
JRip	82, 33 ± 3, 52	84, 91 ± 3, 89	75, 99 ± 4, 88	83, 97 ± 3, 67	86, 22 ± 4, 24

Tabela 5.1: Porcentagem dos Acertos e Desvio Padrão dos Métodos Individuais.

Como pode ser observados na Tabela 5.1 as médias de acertos para o classificador FuzzyMLP sempre são maiores que os demais métodos. Isto pode ser explicado pelo fato deste método utilizar conceitos fuzzy associados a redes neurais, pois desta maneira este método tem uma melhor capacidade de captar o conhecimento intrínseco das aplicações resultando, assim, em um melhor desempenho. Por outro lado, o classificador K-nn se mantém com as menores médias de acertos. Isto pode ser explicado, ao contrário do FuzzyMLP, pelo fato do K-nn ser o classificador mais simples utilizado. De maneira geral os classificadores SVM, Árvore e JRip possuem desempenhos semelhantes.

Observando os desvios padrões, conclui-se que as Bases C e E possuem os maiores desvios padrões, o que mostra que essas são as bases que sofrem maiores variações de estabilidade por classificador. Isto pode ser explicado pelo fato destas bases serem, como já foi mencionado, desbalanceadas.

5.2 Sistemas com Três Componentes

Os resultados dos experimentos utilizando sistemas com três componentes podem ser observados nas Tabelas 5.2, 5.3 e 5.4. Nesse tamanho de sistema, para cada classificação, existem duas diferentes estruturas:

- Um tipo de componente, ou sistema não híbrido (*NH*)
- Três tipos de componentes, ou sistema híbrido (*H3*).

Para os sistemas não híbridos, foram possíveis construir sete variações, pois só existem sete tipos diferentes de classificadores. Para os experimentos dos sistemas híbridos, que neste tamanho de sistemas é somente o *H3*, foram construídos onze variações. Como já foi mencionado anteriormente, os valores apresentados nas tabelas citadas anteriormente são a média dos melhores desvios padrões de todas as possibilidades. Os valores em negrito representam o melhor desempenho entre as duas variações do sistema, *NH* e *H3*.

Neste tamanho de sistemas, os valores de *Diff* são sempre calculados entre os sistemas *NH* e *H3*. E, por conseguinte, os valores dos testes estatísticos, neste caso os valores de *p*, foram calculados para verificar se essas diferenças são estatisticamente significantes ou não.

Nestas tabelas, sete diferentes métodos de combinação foram analisados, que são: Voto, Naive Bayes (*NB*), Soma, Média, Mediana, *MLP*, FuzzyMLP (*FMLP*); três diferentes métodos de seleção foram analisados, que são: DCS-LA, DCS-DT, DCS-MCB; e três diferentes métodos de negociação foram analisados, que são: Sensibilidade (*Sensibi*), *Lei-lão* e Teoria dos Jogos (*TJ*). Para definir o conjunto de treinamento para os combinadores treináveis, aproximadamente 10% das instâncias de cada base de dados foram separada para criar um conjunto de validação.

5.2.1 Métodos de Combinação ou Fusão para Sistemas com Três Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de combinação para sistemas com três componentes.

Como pode ser visto nas Tabelas 5.1 e 5.2, pode-se concluir que os sistemas baseados em fusão com três componentes tiveram um melhor desempenho em relação aos métodos individuais. Este comportamento já era esperado, pois, como já foi dito anteriormente, os *Ensembles* conseguem captar o conhecimento mais facilmente porque gera sua resposta a partir de um conjunto de classificadores.

Na Tabela 5.2 pode ser observado os resultados dos sete métodos baseados em fusão que foram usados neste trabalho. De maneira geral, os melhores desempenhos acontecem quando o sistema possui três componentes diferentes, ou seja, os melhores desempenhos

Sistemas Com Três Componentes								
BD	Métodos Baseados em Fusão							
A	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
	NH	72,02±7,01	78,76±4,71	77,54±6,03	77,78±4,69	73,60±3,71	91,18±3,82	91,95±3,23
	H3	75,58±2,87	78,58±2,34	78,21±2,22	78,67±3,31	76,94±4,24	92,37±2,17	94,75±2,16
	Dif	3,56	0,18	0,67	0,89	3,34	1,19	2,80
	p	0,00000214	0,366	0,145	0,0674	0,000000102	0,00416	3,20E-11
B	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
	NH	80,97±6,06	85,56±6,58	82,51±5,55	83,96±6,40	81,57±6,62	95,25±2,97	95,95±2,40
	H3	82,56±6,61	85,11±7,13	83,02±5,40	84,36±7,29	82,72±7,11	94,49±2,50	96,35±2,33
	Dif	1,59	0,45	0,51	0,40	1,15	0,76	0,40
	p	0,0736	0,143	0,00000565	0,176	0,215	0,00845	0,177
C	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
	NH	69,21±11,0	73,07±9,56	70,71±10,2	72,31±10,3	70,05±11,5	84,33±7,31	89,69±6,14
	H3	71,71±4,58	76,51±3,75	76,02±3,85	76,00±3,90	74,60±4,88	90,25±2,82	93,16±3,53
	Dif	2,50	3,44	5,31	3,69	4,55	5,92	3,47
	p	0,00000978	0,001	0,00000122	0,00065	0,0000531	8,94E-16	3,78E-09
D	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
	NH	76,15±4,68	81,14±4,46	78,65±3,48	79,39±4,40	77,07±5,36	89,92±2,12	93,02±3,01
	H3	78,30±4,61	82,93±3,93	80,67±3,83	81,85±3,92	79,59±4,75	92,73±2,77	95,34±2,28
	Dif	2,15	1,79	2,02	2,46	2,52	2,81	2,32
	p	0,0096	0,0182	0,00183	0,0449	0,0793	0,00015	0,00013
E	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
	NH	73,62±6,39	78,95±4,86	76,03±4,47	77,72±5,02	75,20±6,35	89,10±2,48	91,09±2,95
	H3	77,26±5,87	80,21±5,38	78,58±4,58	79,28±5,75	77,80±6,65	91,45±2,50	93,27±3,06
	Dif	3,64	1,26	2,55	1,56	2,60	2,35	2,18
	p	0,00000147	0,00061	0,0000123	0,0002	0,0000255	3,23E-11	2,71E-11

Tabela 5.2: Método baseado em Fusão para Sistemas com Três Componentes.

são alcançados nos sistemas híbridos. Apenas nas bases A e B com o método de combinação Naive Bayes (*NB*) o melhor resultado acontece com sistemas *NH* e ainda na base B isso acontece com o método *MLP*. Isto mostra que aumentando a diversidade do sistema, consegue-se um melhor desempenho.

Ainda na Tabela 5.2, em relação aos métodos baseados em fusão, observa-se que para todas as cinco bases de dados, de maneira geral, os melhores desempenhos foram alcançados com o método FuzzyMLP (*FMLP*). Apenas para as bases A e B, o método com melhor desempenho foi o *MLP*. Além disso, observa-se que os métodos *MLP* e *FMLP* possuem resultados próximos e que eles se distanciam consideravelmente dos demais métodos. Isto pode ser explicado pelo fato destas redes neurais serem métodos mais complexos que os demais, por isso, conseguem realizar a classificação com maior eficiência.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Observa-se que, para as bases A, B e E, o método Voto possui os maiores valores de *Dif*. Já para as bases C e D isto acontece com o método *MLP*. Por outro lado, observando o método que possui os menores valores de *Dif*, conclui-se que para as bases A, D e E, este método é o Naive Bayes (*NB*). Já na base C, o método Voto é o que possui o menor valor de *Dif*, em contraste com

as bases A, B e E no qual ele possui o maior. Na base B acontece um empate, os métodos Média e FuzzyMLP (*FMLP*) possuem o menor valor de *Dif*. Isto mostra que, de forma geral, os sistemas deste tipo mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais complexos. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais simples.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de p deste teste pode ser observado na Tabela 5.2. Dentre os testes feitos entre os métodos que tiveram os maiores *Dif*, a única base no qual essa diferença não é estatística é a base B, pois obteve um valor de p de 0,0736 que é maior que o limite de 0,05. Já entre os testes feitos entre os métodos que tiveram os menores *Dif*, conclui-se que nas bases C, D e E essa diferença é estatisticamente comprovada, mas nas bases A e B isto não acontece.

5.2.2 Métodos Baseados em Seleção para Sistemas com Três Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de seleção para sistemas com três componentes.

Da mesma forma que aconteceu com os métodos baseados em fusão, como pode ser visto nas Tabelas 5.1 e 5.3, pode-se concluir que os sistemas baseados em seleção com três componentes tiveram um melhor desempenho em relação aos métodos individuais. Este comportamento também já era esperado, pois, como já foi dito anteriormente, os *Ensembles* conseguem captar o conhecimento mais facilmente porque geram sua resposta a partir de um conjunto de classificadores.

Na Tabela 5.3 podem ser observados os resultados dos três métodos baseados em seleção que foram usados neste trabalho. Nela, observa-se que, para todas as bases, os melhores desempenhos acontecem quando o sistema possui três componentes diferentes, ou seja, quando o sistema é híbrido. Este é um comportamento esperado, pois há uma tendência a se ter uma maior diversidade em sistemas com diferentes tipos de classificadores.

De maneira geral, os melhores desempenhos são alcançados com o método DCS-MCB. Porém, na base A isto acontece com o método DCS-DT. Dentre esses métodos de seleção não é observada uma grande variação entre as porcentagens de corretude, porém, em todas as bases o DCS-LA possui o menor desempenho dos três. Isto pode ser justificado pelo fato dos DCS-DT e DCS-MCB serem métodos híbridos, ou seja, eles podem realizar seleção ou fusão durante a sua classificação.

Sistemas Com Três Componentes						
BD		Métodos Baseados em Seleção				
A	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	90,80±3,30	90,79±3,28	18,94	92,41±2,77	12,99
	H3	92,21±1,74	94,75±2,31	20,82	94,45±1,67	15,24
	Dif	1,41	3,96	-	2,04	-
	p	0,000122	6,96E-18	-	2,41E-09	-
B	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	93,07±2,95	94,80±3,04	22,13	95,60±2,04	16,81
	H3	94,79±2,53	95,63±2,73	20,95	95,72±2,13	15,77
	Dif	1,72	0,83	-	0,12	-
	p	0,0000352	0,00231	-	0,215	-
C	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	81,04±6,64	86,01±6,91	22,99	88,39±6,26	17,25
	H3	89,43±2,46	91,90±2,86	18,10	93,55±2,83	12,79
	Dif	8,39	5,89	-	5,16	-
	p	3,62E-15	1,46E-12	-	8,91E-11	-
D	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	90,35±2,77	91,69±2,51	24,34	93,88±2,80	17,27
	H3	92,77±2,01	94,15±2,71	23,95	95,87±2,77	15,80
	Dif	2,42	2,46	-	1,99	-
	p	0,00000706	0,00507	-	0,0681	-
E	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	88,43±2,53	89,83±2,75	20,99	91,82±2,15	15,24
	H3	90,82±2,12	92,29±2,97	20,91	94,28±2,18	13,97
	Dif	2,39	2,46	-	2,46	-
	p	1,12E-11	4,05E-14	-	1,34E-14	-

Tabela 5.3: Método baseado em Seleção para Sistemas com Três Componentes.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Os valores de *Dif* são os maiores, para as bases A, D e E, quando o método DCS-DT é aplicado. Já as bases B e C isto acontece quando o método DCS-LA é aplicado. Por outro lado, os valores de *Dif* são os menores, para as bases B, C e D, quando o DCS-MCB é utilizado. Já para as bases A e E isto ocorre com o método DCS-LA. Isto mostra que, de forma geral, os sistemas deste tipo são mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais simples. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais complexos.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de *p* deste teste pode ser observado na Tabela 5.3. Dentre os testes feitos entre os métodos que tiveram os maiores *Dif*, pode-se concluir que todos são estatisticamente comprováveis, pois não atingiram o limiar que é de 0.05. Já dentre os testes feitos entre os métodos que tiveram os menores *Dif*, as únicas bases que não se pode afirmar nada são as B e D, pois possuem valores de *p* menores que o limiar.

Na maioria dos casos, o método DCS-MCB possui um desempenho melhor que o método DCS-DT. Nas colunas 4 e 6 da Tabela 5.3 são apresentados as porcentagens de

fusão para os métodos híbridos DCS-DT e DCS-MCB. Do ponto de vista do método DCS-DT observa-se que, de forma geral, a porcentagem da utilização do método de fusão vai diminuindo quando se faz a variação dos tipos de componentes. Apenas para a base de dados A isto não acontece. Da mesma maneira, do ponto de vista do método DCS-MCB, observa-se o mesmo comportamento, uma diminuição do uso do Voto na medida que se faz a variação do tipo de combinador. Mais uma vez, para a base A isto não acontece. Isto pode ser explicado pelo fato de quanto mais tipos diferentes de classificadores são usados, maior a chance do classificador selecionado ser melhor, estatisticamente, que os demais.

Em geral, os métodos de seleção possuem um melhor desempenho que os métodos de fusão. Porém, na base A acontece um empate e na base B o melhor método de fusão é melhor que o melhor método de seleção. Outro detalhe importante, que já foi dito, é que os melhores métodos de seleção são os híbridos, ou seja, eles trabalham com seleção ou com fusão. Isto pode ser explicado pelo fato métodos de seleção trabalharem com mais informação que os métodos de fusão.

5.2.3 Métodos de Negociação para Sistemas com Três Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de negociação para sistemas com três componentes.

Os resultados dos testes feitos com os métodos de negociação podem ser vistos na Tabela 5.4. Diante disto, observa-se que, assim como nos *Ensembles*, para todas as bases, de maneira geral, os melhores desempenhos acontecem quando o sistema possui três agentes com classificadores diferentes, ou seja, quando utiliza sistema híbrido. Isso somente não acontece no método *Leilão* da base D.

Além disso, para todas as bases de dados, os melhores desempenhos são alcançados com o método baseado em Sensibilidade (*Sensibi*). Isto pode ser explicado pelo fato deste método utilizar mais informações que os demais para executar a negociação. Nos métodos baseados em negociação não é observada uma grande variação entre as porcentagens de correteude, porém, em todas as bases o *Leilão* possui o menor desempenho dentre três.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Dentre os métodos de negociação, observa-se, ainda, que, para as bases A, C, D e E, o método *Sensibi* possui os maiores valores de *Dif*. Porém, para a base B isto acontece com o método Teoria dos Jogos (*TJ*). Observando o fato contrário, ou seja, o método que possui os menores valores de *Dif*, pode-se afirmar que para as bases A e E isto ocorre com o método Teoria dos Jogos (*TJ*)

e para as demais isto ocorre com o método *Leilão*. Isto mostra que, de forma geral, os sistemas deste tipo são mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais complexos. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais simples.

Sistemas Com Três Componentes						
BD	Métodos de Negociação					
	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
A	NH	88,78±2,77	91,41±2,38	10,28	90,96±2,45	9,25
	H3	90,71±2,13	95,17±1,77	10,29	92,65±2,08	10,12
	Dif	1,93	3,76	-	1,69	-
	p	0,00000019	1,71E-25	-	0,000000818	-
B	NH	91,70±3,15	95,78±2,14	11,38	93,38±3,12	10,52
	H3	91,74±3,14	96,69±2,24	12,36	94,70±3,12	10,85
	Dif	0,04	0,91	-	1,32	-
	p	0,00056	0,0541	-	0,35	-
C	NH	87,25±5,02	90,85±5,37	16,86	88,81±5,61	12,61
	H3	89,82±3,36	94,49±3,40	14,47	92,03±3,76	11,61
	Dif	2,57	3,64	-	3,22	-
	p	6,15E-10	1,57E-13	-	2,13E-11	-
D	NH	92,05±2,87	94,99±2,64	16,10	93,85±2,31	12,24
	H3	91,76±3,75	96,41±3,00	13,58	94,80±3,18	11,78
	Dif	0,29	1,42	-	0,95	-
	p	0,0224	0,0657	-	0,285	-
E	NH	89,59±3,03	92,48±1,93	15,38	91,25±2,14	11,36
	H3	92,07±2,81	95,54±2,50	16,25	93,24±2,26	11,75
	Dif	2,48	3,06	-	1,99	-
	p	0,00395	6,79E-17	-	0,000000682	-

Tabela 5.4: Método de Negociação para Sistemas com Três Componentes.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de p deste teste pode ser observado na Tabela 5.4. Dentre os testes feitos entre os métodos que tiveram os maiores *Dif*, essa diferença é estatisticamente comprovada apenas para as bases A, C e E. Para as bases B e D, o valor do p estatístico é, respectivamente, 0,35 e 0,065, o que não é menor que 0,05 e prova que não se pode afirmar nada sobre esses dois métodos. Por outro lado, quando se observa os métodos com os menores valores de *Dif*, é interessante observar que, apesar das diferenças serem relativamente pequenas, todos os métodos podem-se dizer melhor estatisticamente que suas variações.

Em relação ao número de iterações (*rounds*) do método *Sensibi* observa-se que, de forma geral, a porcentagem das iterações da negociação para as bases A, B e E diminuem com o a utilização de um sistema híbrido. Por outro lado, para as bases C e D isso ocorre de forma contrária. Já analisando o método Teoria dos Jogos (*TJ*), observa-se um comportamento semelhante, uma diminuição na porcentagem das iterações da negociação

para as bases C e D e um aumento para as bases A, B e E. Logo, conclui-se que de forma geral, o número de *rounds* aumenta quando trabalha-se com estruturas híbridas. Isto pode ser explicado pelo fato de que quando se trabalha com componentes diferentes estes demoram mais a chegar a uma opinião comum.

5.2.4 Negociação *versus* Combinação para Sistemas com Três Componentes

Fazendo uma comparação entre os métodos de negociação e os de combinação, observa-se que, para todas as bases de dados, o método de negociação baseado em Sensibilidade (*Sensibi*) obteve o melhor desempenho. Para saber se um método é estatisticamente melhor que o outro, realiza-se um teste de hipótese entre estes dois métodos. Os resultados dos testes estatísticos entre o melhor método de negociação dos sistemas *NH* e *H3* e os melhores métodos baseados em fusão e seleção podem ser observados na Tabela 5.5.

Na Tabela 5.5, quando aparece 'Sim' significa que o melhor método de negociação é estatisticamente melhor que o melhor método do tipo que está sendo comparado e 'Não' significa que não se pode afirmar nada sobre os sistemas comparados. Para sistemas com três componentes e considerando que existe apenas um tipo de sistema híbrido, foram feitos vinte testes estatísticos. Como resultado, tem-se que em 8 das 20 comparações o método de negociação é estatisticamente melhor. Isso significa que em 40% das comparações conseguiu-se uma superioridade estatística.

Base	NH		H3	
	Fusão	Seleção	Fusão	Seleção
A	Não - 0,12697517	Sim - 0,01154840	Sim - 0,01627784	Não - 0,21365802
B	Não - 0,33515785	Não - 0,30154605	Não - 0,29236770	Não - 0,37034379
C	Não - 0,11784616	Sim - 0,00691982	Não - 0,38612846	Não - 0,26218269
D	Não - 0,00003050	Sim - 0,00834341	Sim - 0,00050816	Não - 0,46856358
E	Sim - 0,00063044	Não - 0,02955606	Sim - 8,27E-14	Sim - 4,38E-10

Tabela 5.5: Valores de p para os testes estatísticos entre os melhores métodos de negociação e os de fusão e seleção para sistemas com Três Componentes.

Observando Tabela 5.5, conclui-se que, na base de dados A para sistemas *NH*, o melhor método de negociação é estatisticamente melhor que o melhor método de seleção, pois o valor do p ficou abaixo do limiar que é 0.05, mas não se pode dizer nada em relação ao melhor método de negociação e o melhor de seleção. Já para sistemas *H3* acontece exatamente o contrário, ou seja, o método de negociação é melhor que o de fusão, mas não se pode afirmar nada em relação ou método de seleção.

Em relação a base B, não se pode afirmar nada, pois os valores de p para os sistemas *NH* e *H3* são maiores que o limiar. De maneira contrária, na base de dados E, o método

de negociação é melhor que todos os outros.

Na base C, observa-se que apenas quando o método de negociação é comparado com o método de seleção para sistemas *H3*, pode-se dizer que a negociação é melhor. Por outro lado, em relação a base D, apenas quando o método de negociação é comparado com o de seleção para sistemas *H3* é que não se pode afirmar nada.

5.2.5 Diversidade para Sistemas com Três Componentes

Com o intuito de analisar o nível de diversidade dos sistemas, a Tabela 5.6 mostra os valores das três diversidades que foram aplicadas a todas as estruturas dos sistemas com três componentes. Essas medidas de diversidade foram discutidas no capítulo anterior. Como já mencionado, o Q estatístico (q) e o dupla-falsa (df) são medidas que utilizam paridade e a entropia (e) é uma medida que não utiliza paridade.

Na Tabela 5.6, pode-se observar que, quase sempre, variações nos tipos de componentes dos sistemas influenciam as diversidades. De acordo com o comportamento das medidas de diversidade, observa-se que as medidas de q e e tendem a aumentar o seu valor quando são calculadas para sistemas híbridos, neste caso o sistema *H3*. Por outro lado, as medidas de df tendem a diminuir com o aumento da diversidade.

BD		Medidas de Diversidade		
A	Var	q	df	e
	NH	0,79	0,31	0,80
	H3	0,84	0,28	0,82
B	Var	q	df	e
	NH	0,88	0,28	0,84
	H3	0,86	0,27	0,83
C	Var	q	df	e
	NH	0,83	0,31	0,82
	H3	0,82	0,29	0,88
D	Var	q	df	e
	NH	0,88	0,25	0,83
	H3	0,88	0,22	0,85
E	Var	q	df	e
	NH	0,83	0,24	0,82
	H3	0,87	0,25	0,85

Tabela 5.6: Medidas de Diversidade para Sistemas com Três Componentes.

De acordo com os resultados mostrados na Tabela 5.6, para as três medidas conclui-se que sistemas híbridos são capazes de prover maiores níveis de diversidade que sistemas não híbridos. Outra conclusão importante é que este fato também reflete na acurácia dos sistemas, já que os melhores resultados para sistemas com três componentes foram alcançados quando foram utilizados os sistemas *H3*.

5.3 Sistemas com Cinco Componentes

Os resultados dos experimentos utilizando sistemas com cinco componentes podem ser observados nas Tabelas 5.7, 5.8 e 5.9. Nesse tamanho de sistema, para cada classificação, existem três diferentes estruturas:

- Um tipo de componente (*NH*)
- Três tipos de componentes (*H3*)
- Cinco tipos de componentes (*H5*).

Para os sistemas não híbridos, foram possíveis construir sete variações, pois só existem sete tipos diferentes de classificadores. Para os experimentos dos sistemas híbridos, que neste tamanho de sistemas podem ser de dois tipos, *H3* e *H5*, foram construídos dez e quatro variações, respectivamente. Como já foi mencionado anteriormente, os valores apresentados nas tabelas citadas anteriormente são a média dos melhores desvios padrões de todas as possibilidades. Os valores em negrito representam o melhor desempenho entre as três variações do sistema, *NH*, *H3* e *H5*.

Neste tamanho de sistemas, os valores de *Diff* são calculados os maiores e menores valores dentre os três tipos de sistemas, *NH*, *H3* e *H5*. E, por conseguinte, os valores dos testes estatísticos, neste caso os valores de *p*, foram calculados para verificar se essas diferenças são estatisticamente significantes ou não.

Nestas tabelas, sete diferentes métodos de combinação foram analisados, que são: Voto, Naive Bayes (*NB*), Soma, Média, Mediana, *MLP*, FuzzyMLP (*FMLP*); três diferentes métodos de seleção foram analisados, que são: DCS-LA, DCS-DT, DCS-MCB; e três diferentes métodos de negociação foram analisados, que são: Sensibilidade (*Sensibi*), *Lei-lão* e Teoria dos Jogos (*TJ*). Para definir o conjunto de treinamento para os combinadores treináveis, aproximadamente 10% das instâncias de cada base de dados foram separadas para criar um conjunto de validação.

5.3.1 Métodos de Combinação ou Fusão para Sistemas com Cinco Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de combinação para sistemas com cinco componentes.

Na Tabela 5.7 pode ser observado os resultados dos sete métodos baseados em fusão que foram usados neste trabalho. Como pode ser observado, de maneira geral, os melhores desempenhos acontecem quando o sistema possui cinco agentes com classificadores diferentes. Apesar disso, cada base de dados apresentou um comportamento diferente para esses métodos. Por exemplo, para a base A, apenas um dos métodos, (*NB*) apresentou um melhor desempenho com o sistema *H3*. Já a base B, três dos classificadores (*Voto*, *NB* e *Média*) apresentou um melhor desempenho com os sistemas *NH*. A base D, apresentou dois classificadores (*NB* e *Média*) com melhor desempenho para *H3*. As bases C e E apresentaram 100% de melhor desempenho para sistemas *H5*. Observa-se, ainda, que para todas as cinco bases de dados os melhores desempenhos foram alcançados com o método FuzzyMLP (*FMLP*). Mais uma vez, isto pode ser explicado pelo fato deste método ser mais complexo que os demais, por isso, consegue obter melhores resultados.

Sistemas Com Cinco Componentes								
BD	Métodos Baseados em Fusão							
	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
A	NH	73,85±6,26	78,44±5,12	77,56±5,90	75,98±4,90	76,02±5,61	93,47±4,85	94,27±2,85
	H3	80,75±6,11	81,58±4,66	81,12±3,78	81,63±3,88	80,72±5,45	94,42±2,41	96,43±2,01
	H5	82,61±5,72	81,23±5,57	84,24±3,99	83,03±4,89	83,50±8,31	95,45±1,42	97,27±2,26
	Dif	8,76	3,14	6,68	7,05	7,48	1,98	3,00
	p	2,88E-11	0,0000275	2,43E-09	3,17E-11	7,40E-08	0,00678	5,03E-08
B	NH	85,26±3,48	89,74±4,68	86,24±4,09	88,15±4,86	85,73±5,30	95,49±3,00	96,86±2,08
	H3	84,43±3,73	88,97±4,50	85,45±3,07	87,40±5,25	84,84±6,62	94,59±2,21	96,58±1,90
	H5	85,00±2,65	89,35±4,61	87,86±2,05	88,01±6,58	85,82±6,57	95,83±1,64	96,92±1,92
	Dif	0,83	0,77	2,41	0,75	0,98	1,24	0,34
	p	0,0736	0,143	0,00000565	0,176	0,215	0,00085	0,177
C	NH	72,84±6,79	77,88±8,22	74,58±6,99	76,31±8,47	73,65±7,80	84,47±4,88	89,27±4,48
	H3	75,95±4,18	79,39±3,97	78,47±4,02	78,93±5,06	77,34±5,47	89,19±2,56	93,61±3,90
	H5	78,45±5,43	82,39±4,77	80,81±4,87	81,15±4,95	79,60±6,81	92,28±2,75	94,38±3,37
	Dif	5,61	4,51	6,23	4,84	5,95	7,81	5,11
	p	0,00000978	0,001	0,00000122	0,00065	0,0000531	8,94E-16	3,78E-09
D	NH	78,86±4,95	82,16±3,57	80,58±2,80	81,48±3,47	79,86±5,55	92,99±2,11	94,97±2,23
	H3	79,56±4,62	83,42±4,02	81,49±3,40	82,57±4,49	80,62±5,71	93,45±2,35	95,33±2,19
	H5	81,31±5,62	82,57±4,92	82,50±3,97	82,47±4,15	81,52±6,49	94,65±2,44	96,64±2,20
	Dif	2,45	1,26	1,92	1,09	1,66	1,66	1,67
	p	0,0096	0,0182	0,00183	0,0449	0,0793	0,00015	0,00013
E	NH	74,73±6,40	78,15±4,84	77,01±4,90	77,47±5,68	76,10±5,99	89,85±2,88	91,79±3,41
	H3	77,19±4,90	80,75±4,73	77,74±5,11	79,54±5,45	77,06±6,14	90,96±2,57	93,20±3,39
	H5	80,89±6,11	81,37±4,95	81,10±4,26	81,36±4,74	81,03±5,73	94,01±2,90	96,33±2,61
	Dif	6,16	3,22	4,09	3,89	4,93	4,16	4,54
	p	0,00000147	0,00061	0,0000123	0,0002	0,0000255	3,23E-11	2,71E-11

Tabela 5.7: Método baseado em Fusão para Sistemas com Cinco Componentes.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Em relação as maiores diferenças, ou seja, os maiores valores de *Dif*, observa-se que, para as bases A, D e E, o

método Voto possui os maiores valores de *Dif*. Na base B isto acontece com o método Soma, e na base C com o método *MLP*. Observando o caso oposto, os menores valores de *Dif* podem ser vistos, para as bases C e E no método Naive Bayes (*NB*). Para a base A isto ocorre com a *MLP*, na base B com a FuzzyMLP (*FMLP*) e na base D com a Média. Ao contrário do que aconteceu nos métodos de fusão com três componentes, de forma geral, os sistemas de cinco componentes mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais simples. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais complexos.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de p deste teste pode ser observado na Tabela 5.7. Em todas essas ocorrências é provado pelo teste de hipótese que os métodos são significantemente melhores que a sua variação, pois todos os valores de p para essas ocorrências são menores que o limiar, que é de 0.05.

5.3.2 Métodos Baseados em Seleção para Sistemas com Cinco Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de seleção para sistemas com cinco componentes.

Na Tabela 5.8, pode-se observar o desempenho dos métodos baseados em seleção. Assim como aconteceu com os métodos de fusão, em geral, os sistemas que apresentaram melhor desempenho foram os *H5*. Apenas na base B é que um deles ficou melhor para os sistemas *H3*. Em geral, o método que apresenta o melhor desempenho é o DCS-MBC. Apenas para a base B o DCS-DT apresenta um desempenho melhor. Outro detalhe desses resultados é que as variações da porcentagem dos acertos entre os métodos DCS-LA, DCS-DT e DCS-MCB são pequenas.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Diante da diferença entre os resultados dos sistemas *NH*, *H3* e *H5*, ou seja, os valores de *Dif*, observa-se que as maiores diferenças ocorrem na utilização do método DCS-LA. Apenas na base E este fato ocorre com o método DCS-DT. Por outro lado, todas as menores diferenças acontecem quando o método DCS-MCB é aplicado. Da mesma forma que nos métodos de seleção com três componentes, Isto mostra que, de forma geral, os sistemas mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais simples. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais complexos.

Sistemas Com Cinco Componentes						
BD		Métodos Baseados em Seleção				
A	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	91,2±3,58	91,83±2,99	19,98	94,13±3,34	13,81
	H3	93,99±2,22	94,33±2,91	18,57	95,32±1,97	15,37
	H5	95,53±3,02	95,47±2,05	19,44	96,37±2,29	17,37
	Dif	4,33	3,64	-	2,24	-
	p	1,64E-09	2,59E-10	-	0,00014	-
B	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	94,51±2,78	96,10±1,94	22,61	96,54±1,87	16,94
	H3	95,51±2,30	97,06±2,27	22,50	96,67±2,12	15,26
	H5	96,45±1,33	96,58±1,54	21,53	96,82±1,68	13,19
	Dif	1,94	0,96	-	0,28	-
	p	0,0000352	0,00231	-	0,215	-
C	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	80,41±7,05	85,33±5,68	22,61	87,94±5,91	17,62
	H3	86,65±3,10	89,44±2,07	20,97	92,05±1,87	16,62
	H5	90,80±2,32	92,88±2,77	21,70	94,89±2,60	14,87
	Dif	10,39	7,55	-	6,95	-
	p	3,62E-15	1,46E-12	-	8,91E-11	-
D	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	93,16±2,44	93,76±3,00	22,68	95,38±2,17	15,49
	H3	94,38±1,98	94,93±3,03	23,25	95,57±2,05	15,38
	H5	95,34±2,37	95,29±2,85	20,11	96,06±2,40	15,83
	Dif	2,18	1,53	-	0,68	-
	p	0,00000706	0,00507	-	0,0681	-
E	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	89,78±2,72	90,45±2,19	21,67	92,52±1,76	16,02
	H3	90,74±2,79	93,02±2,51	21,79	95,17±2,48	16,06
	H5	93,64±2,41	94,55±2,76	19,19	95,64±1,85	15,99
	Dif	3,86	4,10	-	3,12	-
	p	1,12E-11	4,05E-14	-	1,34E-14	-

Tabela 5.8: Método baseado em Seleção para Sistemas com Cinco Componentes.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de p deste teste pode ser observado na Tabela 5.8. Sendo assim, todos os métodos que tiveram os maiores valores de Dif são estatisticamente melhores que a sua variação. Mas, de acordo com os métodos que apresentaram os menores valores de Dif , pode-se dizer que para as bases A, C e E eles são estatisticamente melhores que suas variações, mas para as base B e D não se pode afirmar nada.

Na maioria dos casos mais uma vez, o método DCS-MCB possui um desempenho melhor que o método DCS-DT. Ainda observando os métodos de seleção, agora somente o método DCS-DT, conclui-se que para as bases A, B e C, as porcentagens do uso do método de fusão diminui para sistemas $H3$ e aumenta para sistemas $H5$. Já no método DCS-MCB para as bases B e C a porcentagem de execução do Voto diminui de acordo com o acréscimo de novos classificadores. Por outro lado, para a base A acontece exatamente o oposto. Na base D acontece uma diminuição da porcentagem do uso do Voto nos sistemas $H3$ seguido pelo aumento para sistemas $H5$. Já na base E acontece o processo

oposto.

Para este tamanho de sistema, comparando-se os métodos de seleção e fusão, observa-se que os de fusão apresentam um desempenho melhor que os de seleção. Como pode ser visto nas Tabelas 5.7 e 5.8, nas bases de dados A, D e E o desempenho do melhor método de fusão, que no caso foi sempre o FuzzyMLP, é melhor que o melhor método de seleção, que no caso foi sempre o DCS-MCB. Porém, como nos sistemas com três componentes, os métodos de seleção apresentam um melhor desempenho que os de fusão para as bases B e C.

5.3.3 Métodos de Negociação para Sistemas com Cinco Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de negociação para sistemas com cinco componentes.

Os resultados dos testes feitos com os métodos de negociação podem ser vistos na Tabela 5.9. Analisando esta tabela, pode-se afirmar que, em geral assim como nos *Ensembles*, eles apresentam melhor desempenho quando são utilizados sistemas *H5*. Porém, para as bases D e E, o *Leilão* apresentou melhor desempenho com sistemas do tipo *NH* e *H3*, respectivamente. E, na base B, a Teoria dos Jogos (*TJ*) apresenta um melhor desempenho com o sistema do tipo *H3*. O método que obteve o melhor desempenho para todas as bases foi o métodos de negociação baseado em Sensibilidade (*Sensibi*). Além disso, observa-se uma diferença considerável entre o desempenho deste método em relação aos demais. Este é um comportamento esperado, já que este método utiliza mais informações durante o processo de negociação.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Avaliando, agora, as diferenças entre os maiores e menores sistemas entre *NH*, *H3* e *H5*, ou seja, os maiores e menores valores de *Dif*, tem-se que as maiores diferenças nas bases A, B e D acontecem com o método *Leilão*. Para as demais bases, isto acontece com o método baseado em Sensibilidade (*Sensibi*). Por outro lado, ainda observando as diferenças, o método que possui as menores diferenças, para as bases B e D é a Teoria dos Jogos (*TJ*), já para as bases C e E, isto ocorre com o método *Leilão*. E, por fim, a base A possui uma menor diferença quando trabalha com o método Sensibilidade (*Sensibi*). Da mesma forma que em sistema com três agentes, isto mostra que, de forma geral, os sistemas deste tipo são mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais complexos. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais simples.

Sistemas Com Cinco Componentes						
BD	Métodos de Negociação					
A	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
	NH	87,64±2,43	94,68±2,99	12,74	90,85±2,95	10,18
	H3	88,88±2,26	97,14±1,28	12,56	94,20±2,12	10,53
	H5	92,30±2,23	98,07±1,82	13,21	95,66±2,04	10,60
	Dif	5,02	3,39	-	4,81	-
	p	2,54E-17	1,32E-09	-	2,09E-15	-
B	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
	NH	93,57±2,53	97,21±2,39	13,97	95,72±1,91	10,16
	H3	93,13±2,63	97,90±2,03	13,14	96,09±2,10	10,43
	H5	94,82±2,88	97,93±2,00	15,62	95,94±1,91	10,38
	Dif	1,69	0,72	-	0,15	-
	p	0,00056	0,0541	-	0,35	-
C	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
	NH	87,10±4,34	90,60±4,30	17,73	88,99±4,36	13,11
	H3	89,40±2,84	94,93±2,78	17,09	93,91±3,42	11,66
	H5	92,29±3,08	96,74±2,40	16,80	94,68±2,96	12,25
	Dif	5,19	6,14	-	5,69	-
	p	6,15E-10	1,57E-13	-	2,13E-11	-
D	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
	NH	92,27±3,34	97,41±2,10	14,67	95,22±2,15	12,59
	H3	92,24±3,38	97,09±2,26	14,75	95,13±2,96	12,47
	H5	90,83±3,99	97,71±2,01	14,01	95,44±2,63	11,52
	Dif	1,44	0,62	-	0,31	-
	p	0,0224	0,0657	-	0,285	-
E	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
	NH	90,23±3,29	93,14±2,36	15,32	91,98±2,89	10,63
	H3	91,51±2,88	94,61±3,13	16,66	93,55±2,96	11,13
	H5	90,51±2,94	97,32±1,73	13,77	94,78±2,51	10,65
	Dif	1,28	4,18	-	2,80	-
	p	0,00395	6,79E-17	-	0,000000682	-

Tabela 5.9: Método de Negociação para Sistemas com Cinco Componentes.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de p deste teste pode ser observado na Tabela 5.9. Com isto, conclui-se que para os métodos que apresentaram os maiores valores de Dif , todos são estatisticamente que a sua variação. Já para aqueles que apresentaram os menores valores de Dif , apenas nas bases B e D não se pode afirmar nada.

Analisando o método *Sensibi* observa-se que, de forma geral, a porcentagem das iterações da negociação sofrem picos. Por exemplo, para as bases A e B a média das iterações sofrem uma pequena queda entre os sistemas *NH* e *H3*, mas sobe para sistemas *H5*. Já a base C para as bases D e E aumentam com o a utilização do sistema híbrido *H3* e diminui com o uso do *H5*. E, por fim, a base C que sofre diminuição quando aumenta a quantidade de agentes diferentes. Agora, analisando o método Teoria dos Jogos (*TJ*), nas bases B e E acontece um aumento da porcentagem de iterações para sistemas *H3* e diminuição para sistemas *H5*. Já a base C tem a reação contrária. A base A só sofre aumento com o acréscimo de agentes diferentes e a base D sofre o efeito contrário. Logo, conclui-se que

de forma geral, o número de *rounds* aumenta quando trabalha-se com estruturas híbridas. Isto pode ser explicado pelo fato de que quando se trabalha com componentes diferentes estes demoram mais a chegar a uma opinião comum.

5.3.4 Negociação *versus* Combinação para Sistemas com Cinco Componentes

Fazendo uma comparação entre os métodos de negociação e os de combinação, que englobam os métodos de seleção e fusão, observa-se que, para todas as bases de dados, mais uma vez, o método de negociação baseado em Sensibilidade (*Sensibi*) obteve o melhor desempenho. Para saber se um método é estatisticamente melhor que o outro, realiza-se um teste de hipótese entre esse método e os demais. Os resultados dos testes estatísticos entre o melhor método de negociação dos sistemas *NH*, *H3* e *H5* e os melhores métodos baseados em fusão e seleção podem ser observados na Tabela 5.10.

Na Tabela 5.10, quando aparece 'Sim' significa que o melhor método de negociação é estatisticamente melhor que o melhor método do tipo que está sendo comparado e 'Não' significa que não se pode afirmar nada sobre os sistemas comparados. Para sistemas com cinco componentes e considerando que existem dois tipos de sistemas híbridos, foram feitos trinta testes estatísticos. Como resultado, tem-se que em 26 das 30 comparações o método de negociação é estatisticamente melhor. Isso significa que em 86,67% das comparações conseguiu-se uma superioridade estatística. Observa-se uma melhora quando compara-se o desempenho dos sistemas com três componentes, onde somente 40% das comparações foram estatisticamente melhores.

	NH		H3		H5	
Base	Fusão	Seleção	Fusão	Seleção	Fusão	Seleção
A	Não - 0,20378700	Não - 0,15274868	Sim - 0,00162633	Sim - 3,02E-13	Sim - 0,04425290	Sim - 0,00022066
B	Não - 0,18325493	Sim - 0,03363419	Sim - 2,08E-06	Sim - 0,00309751	Sim - 0,01168165	Sim - 0,00442876
C	Sim - 0,03798267	Sim - 0,00141797	Sim - 0,00307054	Sim - 1,27E-15	Sim - 0,00028575	Sim - 0,00072196
D	Sim - 2,97E-10	Sim - 5,03E-08	Sim - 3,94E-08	Sim - 6,65E-07	Sim - 0,01284674	Sim - 0,00063113
E	Sim - 0,00359252	Sim - 0,03886376	Sim - 0,00126007	Não - 0,07913151	Sim - 0,02465112	Sim - 0,00003479

Tabela 5.10: Teste de Hipótese entre os melhores métodos de seleção e fusão com o melhor método de negociação para sistemas com Cinco Componentes.

Na primeira linha da Tabela 5.10, observa-se os testes entre o melhor método de negociação e os melhores método de fusão e seleção. Logo, conclui-se que para a base de dados A, apenas quando os sistemas *NH* são utilizados é que não se pode afirmar nada, pois estes ultrapassaram o limiar, que é de 0.05. Na base de dados B, não se pode afirmar nada quando se compara o melhor método de fusão com o melhor método de negociação, nas demais comparações, a negociação é melhor estatisticamente. Nas bases C, D e E

todos as comparações concluíram que o melhor método de negociação é melhor que os melhores métodos de fusão e seleção.

5.3.5 Diversidade para Sistemas com Cinco Componentes

Com o intuito de analisar o nível de diversidade dos sistemas, a Tabela 5.11 mostra os valores das três diversidades que foram aplicadas a todas as estruturas dos sistemas com cinco componentes. Essas medidas de diversidade foram discutidas no capítulo anterior. Como já mencionado, o Q estatístico (q) e o dupla-falsa (df) são medidas que utilizam paridade e a entropia (e) é uma medida que não utiliza paridade.

Na Tabela 5.11, pode-se observar que, quase sempre, variações nos tipos de componentes dos sistemas influenciam as diversidades. De acordo com o comportamento das medidas de diversidade, observa-se que as medidas de q e e tendem a aumentar o seu valor quando são calculadas para sistemas híbridos, neste caso os sistemas $H3$ e $H5$. Por outro lado, as medidas de df tendem a diminuir com o aumento da diversidade.

Neste tamanho de sistema, observa-se que, em geral, as diversidades para as bases D e E tendem a manter uma estabilidade, mas nas bases A, B e C acontece exatamente o que é esperado. Este comportamento apenas não é observado nas bases B e C.

BD		Medidas de Diversidade		
A	Var	q	df	e
	NH	0,80	0,29	0,82
	H3	0,82	0,26	0,81
	H5	0,86	0,25	0,83
B	Var	q	df	e
	NH	0,91	0,23	0,86
	H3	0,92	0,23	0,87
	H5	0,84	0,29	0,84
C	Var	q	df	e
	NH	0,83	0,29	0,84
	H3	0,85	0,31	0,84
	H5	0,88	0,23	0,87
D	Var	q	df	e
	NH	0,85	0,22	0,79
	H3	0,87	0,34	0,79
	H5	0,85	0,22	0,82
E	Var	q	df	e
	NH	0,87	0,24	0,84
	H3	0,86	0,26	0,83
	H5	0,85	0,24	0,83

Tabela 5.11: Medidas de Diversidade para Sistemas com Cinco Componentes.

De acordo com os resultados mostrados na Tabela 5.11 e mantendo o padrão do que foi observado nos sistemas com três componentes, para as três medidas conclui-se que

sistemas híbridos são capazes de prover maiores níveis de diversidade que sistemas não híbridos. Outra conclusão importante é que este fato também reflete na acurácia dos sistemas, já que os melhores resultados para sistemas com cinco componentes foram alcançados quando foram utilizados os sistemas *H5*.

5.4 Sistemas com Sete Componentes

Os resultados dos experimentos utilizando sistemas com sete componentes podem ser observados nas Tabelas 5.12, 5.13 e 5.14. Nesse tamanho de sistema, para cada classificação, existem quatro diferentes estruturas:

- Um tipo de componente (*NH*)
- Três tipos de componentes (*H3*)
- Cinco tipos de componentes (*H5*)
- Sete tipos de componentes (*H7*).

Para os sistemas não híbridos, foram possíveis construir sete variações, pois só existem sete tipos diferentes de classificadores. Para os experimentos dos sistemas híbridos, que neste tamanho de sistemas podem ser de três tipos, *H3*, *H5* e *H7*, foram construídos dez, quatro e uma variação, respectivamente. Como já foi mencionado anteriormente, os valores apresentados na tabela citadas anteriormente são a média dos melhores desvios padrões de todas as possibilidades. Os valores em negrito representam o melhor desempenho entre as quatro variações do sistema, *NH*, *H3*, *H5* e *H7*.

Neste tamanho de sistemas, os valores de *Dif* são calculados os maiores e menores valores dentre os quatro tipos de sistemas, *NH*, *H3*, *H5* e *H7*. E, por conseguinte, os valores dos testes estatísticos, neste caso os valores de *p*, foram calculados para verificar se essas diferenças são estatisticamente significantes ou não.

Nestas tabelas, sete diferentes métodos de combinação foram analisados, que são: Voto, Naive Bayes (*NB*), Soma, Média, Mediana, *MLP*, FuzzyMLP (*FMLP*); três diferentes métodos de seleção foram analisados, que são: DCS-LA, DCS-DT, DCS-MCB; e três diferentes métodos de negociação foram analisados, que são: Sensibilidade (*Sensibi*), *Leilão* e Teoria dos Jogos (*TJ*). Para definir o conjunto de treinamento para os combinadores treináveis, aproximadamente 10% das instâncias de cada base de dados foram separada para criar um conjunto de validação.

5.4.1 Métodos de Combinação ou Fusão para Sistemas com Sete Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de combinação para sistemas com sete componentes.

Na Tabela 5.12 pode ser observado os resultados dos sete métodos baseados em fusão que foram usados neste trabalho. Analisando esta tabela, tem-se que, em geral, o sistema que obteve as melhores porcentagens de acerto foi o com sete tipos de classificadores diferentes, ou seja, o sistemas híbrido *H7*. Isto é sensivelmente percebido quando se observa as bases de dados C e D. Nelas os melhores desempenhos são 100% pertencentes aos sistemas *H7*. Apesar de possuir apenas o método FuzzyMLP (*FMLP*) com melhor desempenho para sistemas *H5*, a base B também tem uma grande concentração de melhores resultados com os sistemas *H7*. Por outro lado, a base de dados A mostra um desempenho particular. Nela, os métodos Voto, Soma e Mediana possuem um melhor desempenho com sistemas *H3*. Ainda assim, os demais métodos possuem melhor desempenho com sistemas *H7*. Já a base de dados E apresenta uma melhor desempenho sempre com sistemas *H5*.

Para esse tamanho de sistema, os métodos que apresentam os melhores desempenhos para cada base de dados estão concentrados nas redes neurais. Para as bases A e B os melhores desempenhos foram alcançados com os classificadores *MLP* e para as demais bases, isto ocorreu com os classificadores FuzzyMLP (*FMLP*). E, mais uma vez, os métodos neurais apresentam uma diferença de desempenho acentuada em relação aos demais. E, mais uma vez, isto pode ser explicado pelo fato do *MLP* e do FuzzyMLP serem métodos mais complexos que os demais, conseguindo assim realizar mais classificações corretamente.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Em relação a diferença de desempenho entre os sistemas híbridos e não híbridos, ou seja, os valores de *Dif*, observa-se que os maiores valores de *Dif* ocorreram, em sua maioria, com o classificador Naive Bayes (*NB*). Apenas a base de dados D apresenta essa maior diferença quando é utilizado o classificador Voto. Em relação aos menores valores de *Dif*, observa-se que eles ocorrem, para as bases de dados A e C, quando o classificador FuzzyMLP (*FMLP*) é utilizado. Para a base B, isto ocorre para o método Voto, para a base D, para o método *MLP* e para a base E, para o método Soma. Da mesma forma que nos sistemas com três e cinco componentes, isto mostra que, de forma geral, os sistemas deste tipo são mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais complexos. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais simples.

Sistemas Com Sete Componentes								
BD		Métodos Baseados em Fusão						
A	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
	NH	75,32±5,87	81,21±6,12	79,23±5,56	80,11±7,03	78,25±7,95	91,68±3,61	93,96±2,88
	H3	78,57±2,98	83,98±2,51	82,98±3,35	83,23±4,97	81,86±6,39	93,46±2,11	95,86±2,37
	H5	84,31±3,26	88,78±3,67	87,20±3,51	87,08±6,42	85,36±6,64	95,48±2,61	96,71±2,12
	H7	85,69±2,16	90,51±3,22	88,84±2,24	87,95±6,50	86,88±6,82	96,72±1,90	96,73±2,00
	Dif	10,37	9,30	9,61	7,84	8,63	5,04	2,77
	p	3,12E-28	1,61E-21	4,51E-27	1,15E-10	9,13E-11	3,32E-19	3,90E-10
B	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
	NH	85,21±4,60	88,89±6,66	85,01±3,38	88,41±6,03	84,80±4,70	95,49±1,81	96,84±2,61
	H3	82,86±4,63	87,90±3,87	85,36±3,42	86,91±5,21	84,13±5,48	95,43±2,58	96,44±2,51
	H5	86,07±4,11	90,70±4,60	87,31±4,07	89,75±5,99	86,39±4,65	95,70±2,40	97,20±2,40
	H7	85,69±2,16	90,51±3,22	88,84±2,24	87,95±6,50	86,88±6,82	96,72±1,90	96,73±2,00
	Dif	3,21	2,80	3,83	2,84	2,75	1,29	0,76
	p	0,00000315	0,0000147	1,08E-13	0,00061	0,00925	0,000597	0,0247
C	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
	NH	74,80±3,93	81,50±2,64	75,92±3,65	79,94±4,08	75,06±6,22	89,76±2,84	91,71±2,92
	H3	78,36±3,00	82,96±2,17	79,64±2,25	81,79±3,63	79,09±5,48	91,83±1,47	92,59±2,56
	H5	82,54±4,84	86,92±6,40	85,45±5,34	86,49±6,53	83,76±7,49	93,03±2,39	94,60±2,34
	H7	85,43±2,19	90,09±3,91	88,17±2,37	87,81±6,05	86,49±6,40	96,19±1,78	96,78±1,88
	Dif	10,63	8,59	12,25	7,87	11,43	6,43	5,07
	p	1,92E-42	1,21E-31	1,71E-50	6,81E-16	3,78E-20	1,28E-33	4,87E-24
D	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
	NH	77,45±3,88	82,42±3,47	80,38±3,03	81,27±4,86	79,44±6,67	91,23±2,38	94,42±1,97
	H3	79,45±4,57	83,60±3,93	82,06±4,01	82,26±5,03	80,15±5,94	93,00±3,59	95,34±2,35
	H5	81,61±4,72	85,08±3,97	83,01±3,48	84,56±5,38	81,94±5,55	93,33±2,53	96,50±2,36
	H7	82,56±4,87	88,80±3,57	85,90±5,04	86,46±6,12	84,71±6,86	96,20±1,78	96,38±1,88
	Dif	5,11	6,38	5,52	5,19	5,27	4,97	2,08
	p	9,95E-11	3,27E-20	5,02E-13	6,79E-08	0,00000462	1,72E-28	4,35E-08
E	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
	NH	78,98±5,06	83,02±4,92	80,49±4,71	81,39±4,89	79,57±5,31	92,16±3,38	93,25±2,75
	H3	80,13±5,07	84,37±4,80	82,49±4,56	83,27±4,91	81,50±5,52	92,28±3,23	94,31±2,64
	H5	82,77±5,52	86,64±6,09	84,86±6,04	85,46±5,60	84,15±4,98	94,41±4,05	96,60±2,64
	H7	84,65±3,27	88,71±4,37	87,77±3,47	87,00±6,56	85,56±6,81	96,27±1,97	96,73±1,96
	Dif	5,67	5,69	7,28	5,61	5,99	4,11	3,48
	p	4,47E-13	0,00061	2,02E-19	2,88E-08	2,15E-08	2,73E-15	7,32E-15

Tabela 5.12: Método baseado em Fusão para Sistemas com Sete Componentes.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de p deste teste pode ser observado na Tabela 5.12. Diante disso, pode-se afirmar que todos os sistemas que possuíram os maiores valores de Dif são estatisticamente melhores que suas variações. Por outro lado, conclui-se que, quando analisando os menores valores de Dif , o único método que não se pode afirmar nada é o MLP da base D, os demais são estatisticamente melhores que suas variações.

5.4.2 Métodos Baseados em Seleção para Sistemas com Sete Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de seleção para sistemas com sete componentes.

Na Tabela 5.13, pode-se observar o desempenho dos métodos baseados em seleção. Nesta tabela, observa-se que, em geral, os sistemas que apresentam os melhores desempenhos são os *H7*. Apenas para a base de dados D isto não ocorre com o método DCS-DT. De acordo com os melhores desempenhos do sistema, observa-se que para as bases A e D isto acontece com o método DCS-LA. Por outro lado, para as bases B, C e E isto ocorre com o método DCS-MCB. Além disso, observa-se que as diferenças de desempenho entre os métodos é relativamente pequena, ou seja, os desempenhos são bastante próximas.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Em relação a diferença de desempenho entre os sistemas híbridos e não híbridos, ou seja, os valores de *Dif*, pode-se concluir que, em geral, os maiores valores de *Dif* ocorrem quando o método DCS-LA é utilizado. Apenas para a base de dados A, isto ocorre com o método DCS-DT. Já os menores valores de *Dif* acontecem, para as bases A, C, D e E quando o método DCS-MCB é utilizado. Apenas com a base B isto ocorre para o método DCS-DT. Mais uma vez, assim como nos sistemas com três e cinco componentes, isto mostra que, de forma geral, os sistemas deste tipo são mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais simples. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais complexos.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de *p* deste teste pode ser observado na Tabela 5.13. Diante do exposto e dos valores de *p*, pode-se concluir que todos os sistemas com os combinadores com maiores e menores valores de *Dif* são estatisticamente melhores.

Na maioria dos casos mais uma vez, o método DCS-MCB possui um desempenho melhor que o método DCS-DT. Avaliando, agora, o método DCS-DT, ainda na Tabela 5.13, observa-se que nas bases B, C, D e E a média de utilização do método de fusão diminui com o acréscimo de classificadores de tipos diferentes. Apenas a base A possui um aumento progressivo nessa variação. Já para o método DCS-MCB, a média de utilização do método Voto para resolver o sistema, para as bases A e B diminui com a utilização do sistema *H3*, aumenta com a utilização do sistema *H5* e volta a diminuir com a utilização do sistema *H7*. A base C funciona de forma que essa média diminui quando aumenta a quantidade de tipos dos classificadores e quando usa-se *H7* ele diminui. Por outro lado, a base D se comporta exatamente de forma contrária. E, por fim, a base E apresenta

Sistemas Com Sete Componentes						
BD	Métodos Baseados em Seleção					
A	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	93,09±2,41	93,34±2,77	21,13	95,09±3,10	13,13
	H3	96,23±1,69	95,81±1,49	21,91	95,52±1,81	13,49
	H5	96,63±1,13	97,04±1,34	21,78	97,10±2,15	12,36
	H7	96,89±1,20	96,86±1,35	22,42	97,38±1,78	12,28
	Dif	3,80	3,70	-	2,29	-
	p	5,79E-23	1,52E-18	-	0,000000173	-
B	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	95,95±2,17	96,40±2,38	20,19	96,54±1,80	16,47
	H3	95,55±1,98	96,43±2,15	21,64	96,47±2,62	16,04
	H5	96,18±1,83	97,39±1,87	19,96	97,73±1,87	15,09
	H7	96,62±1,20	97,18±1,41	22,26	97,38±1,78	12,22
	Dif	1,07	1,33	-	1,26	-
	p	0,0000379	0,00354	-	0,000349	-
C	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	91,97±2,70	93,59±3,02	20,06	93,80±3,38	13,67
	H3	92,83±2,31	95,15±2,67	20,82	95,87±2,28	13,95
	H5	95,10±2,81	96,58±1,83	19,08	96,29±2,63	13,50
	H7	96,30±1,55	96,93±1,46	22,56	96,99±1,70	13,59
	Dif	4,33	3,34	-	3,19	-
	p	1,51E-22	3,52E-14	-	4,08E-11	-
D	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	91,09±2,74	94,07±3,08	21,48	95,58±2,82	14,74
	H3	92,89±2,68	95,38±3,15	19,82	95,96±2,61	12,96
	H5	95,53±2,61	96,63±3,32	20,15	96,43±2,84	13,47
	H7	96,46±1,16	96,77±1,62	22,17	96,67±1,96	13,02
	Dif	5,37	2,70	-	1,09	-
	p	2,87E-31	7,15E-10	-	0,00444	-
E	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	91,31±2,78	92,63±1,73	22,10	94,01±2,27	16,30
	H3	92,00±2,73	93,22±1,98	22,77	95,10±2,18	16,26
	H5	92,84±3,04	94,13±2,03	21,31	97,32±1,67	14,85
	H7	96,37±1,49	96,59±1,99	21,49	96,73±1,97	12,73
	Dif	5,06	3,96	-	3,31	-
	p	4,52E-27	6,48E-25	-	7,74E-18	-

Tabela 5.13: Método baseado em Seleção para Sistemas com Sete Componentes.

uma diminuição dessa média na medida que a quantidade de tipos de classificadores é aumentada.

Para este tamanho de sistema, comparando-se os métodos de seleção e fusão, observa-se que os de seleção apresentam um desempenho melhor que os de fusão. Como pode ser visto nas Tabelas 5.12 e 5.13, para todas as bases de dados o melhor método de fusão foi sempre o FuzzyMLP. Já o melhor método de seleção, foi em geral, o método DCS-MCB, mas, na base D isto ocorre com o método DCS-DT. Mas sempre os métodos FuzzyMLP possuíram desempenho inferior que os DCS.

5.4.3 Métodos de Negociação para Sistemas com Sete Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de negociação para sistemas com sete componentes.

Analisando a Tabela 5.14 e observando os métodos de negociação, pode-se afirmar que, assim como nos *Ensembles*, em geral, eles apresentam melhor desempenho quando são utilizados sistemas *H7*. Porém, para a base B os métodos *Leilão* e Sensibilidade (*Sensibi*) apresentam melhor desempenho com sistemas do tipo *H5*. E, na base D, a Teoria dos Jogos (*TJ*) e a Sensibilidade (*Sensibi*) apresentam um melhor desempenho com o sistema do tipo *H5* também. Observando o desempenho dos métodos de negociação, conclui-se que o método baseado em Sensibilidade (*Sensibi*) apresenta, para todas as bases, o melhor desempenho. Para as bases B e D, como foi mencionado anteriormente, apresentam esse resultado para sistemas *H5*, mas as demais são para sistemas *H7*.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Em relação a diferença de desempenho entre os sistemas híbridos e não híbridos, ou seja, os valores de *Dif*, observa-se que as maiores diferenças ocorreram, para as bases A e D, com o método *Leilão*. Para as bases B e C isto ocorre quando o método de Sensibilidade (*Sensibi*) é aplicado. E, por fim, para a base E isto ocorre para o método Teoria dos Jogos (*TJ*). Observando os métodos que alcançaram os menores valores de *Dif*, conclui-se que para as bases B, C e E isto ocorre quando o método *Leilão* é aplicado. Já nas bases A e D, isto acontece para os métodos de Teoria dos Jogos (*TJ*). Mais uma vez, assim como nos sistemas com três e cinco agents, isto mostra que, de forma geral, os sistemas deste tipo são mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais complexos. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais simples.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de p deste teste pode ser observado na Tabela 5.14. Diante do exposto e dos testes de hipótese, pode-se afirmar que todos os sistemas que possuíram os maiores valores de *Dif* são estatisticamente melhores que suas variações. Já em relação aos menores valores de *Dif*, o único método que não se pode afirmar que ele é estatisticamente melhor que sua variação é o método Teoria dos Jogos (*TJ*) da base D.

Observando, na Tabela 5.14, o método Teoria dos Jogos (*TJ*) conclui-se que para as bases A, C e D, a média da aplicação do método de fusão diminui com o aumento do tipo dos agentes presentes no sistema. Já a base E apresenta uma diminuição para o sistema *H3*, mas nos sistemas *H5* e *H7* ele sofre um aumento. E, por fim, a base B sofre sempre aumento desta média. Analisando, agora, o método baseado em Sensibilidade (*Sensibi*),

Sistemas Com Sete Componentes						
BD	Métodos de Negociação					
A	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
	NH	88,86±2,43	94,54±3,04	17,88	93,01±3,40	12,77
	H3	92,25±2,34	96,38±2,04	16,72	95,24±2,08	12,60
	H5	93,64±2,54	97,17±2,46	15,21	95,74±1,95	11,90
	H7	95,36±3,59	98,36±2,48	15,26	96,59±2,94	11,26
	Dif	6,50	3,82	-	3,58	-
	p	6,38E-11	0,00015	-	0,00112	-
B	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
	NH	94,41±2,73	97,74±2,15	13,03	96,14±2,40	10,12
	H3	94,80±2,21	97,97±2,65	14,17	96,28±2,49	10,35
	H5	95,35±2,94	98,25±2,13	13,52	96,87±2,23	10,98
	H7	94,46±2,01	96,19±2,19	16,37	97,56±1,88	10,99
	Dif	0,94	2,06	-	1,42	-
	p	0,0253	0,00275	-	0,0384	-
C	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
	NH	88,01±3,76	92,04±3,47	14,14	90,17±3,44	18,10
	H3	89,78±2,52	92,95±2,16	14,61	90,88±2,23	17,47
	H5	91,10±2,80	94,47±2,48	14,44	93,48±3,14	16,44
	H7	93,55±2,65	96,81±3,61	17,98	95,88±2,81	10,84
	Dif	5,54	6,77	-	5,71	-
	p	0,0000117	0,0000607	-	0,00000165	-
D	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
	NH	94,13±3,14	97,09±2,67	15,34	95,89±2,25	11,22
	H3	93,30±3,44	97,49±1,68	14,75	95,71±2,20	11,14
	H5	94,54±3,71	98,12±1,59	15,46	96,07±2,18	11,06
	H7	96,94±1,67	97,40±2,28	16,09	95,70±1,67	10,29
	Dif	3,64	1,03	-	0,36	-
	p	0,00066	0,0031	-	0,144	-
E	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
	NH	91,25±2,45	94,72±2,79	15,92	92,70±3,21	12,48
	H3	91,56±3,28	95,27±2,77	15,72	93,16±3,25	10,79
	H5	92,99±2,74	96,83±2,62	16,14	94,72±2,62	11,35
	H7	93,84±3,25	97,84±2,64	16,23	96,59±2,51	13,54
	Dif	2,59	3,12	-	3,89	-
	p	0,0018	0,00066	-	0,00022	-

Tabela 5.14: Método de Negociação para Sistemas com Sete Componentes.

observa-se a média da utilização do método Voto para resolução do sistema diminui, mas com o aumento da quantidade de tipos de agentes ele sofre um aumento também. Já as bases B e C apresentam uma oscilação quando se modifica a quantidade dos tipos, primeiramente ele aumenta, depois diminui para aumentar novamente. Logo, conclui-se que de forma geral, o número de *rounds* aumenta quando trabalha-se com estruturas híbridas. Isto pode ser explicado pelo fato de que quando se trabalha com componentes diferentes estes demoram mais a chegar a uma opinião comum.

5.4.4 Negociação *versus* Combinação para Sistemas com Sete Componentes

Comparando-se, agora, os métodos de negociação e os de combinação, que englobam os métodos de seleção e fusão, observa-se que, para todas as bases de dados, mais uma vez, o método de negociação baseado em Sensibilidade (*Sensibi*) obteve o melhor desempenho. Para saber se um método é estatisticamente melhor que o outro, realiza-se um teste de hipótese entre esse método e os demais. Os resultados dos testes estatísticos entre o melhor método de negociação dos sistemas *NH*, *H3*, *H5* e *H7* e os melhores métodos baseados em fusão e seleção podem ser observados na Tabela 5.15.

Na Tabela 5.15, quando aparece 'Sim' significa que o melhor método de negociação é estatisticamente melhor que o melhor método do tipo que está sendo comparado e 'Não' significa que não se pode afirmar nada sobre os sistemas comparados. Para sistemas com sete componentes e considerando que existem três tipos de sistemas híbridos, foram feitos quarenta testes estatísticos. Como resultado, tem-se que em 28 das 40 comparações o método de negociação é estatisticamente melhor. Isso significa que em 70% das comparações conseguiu-se uma superioridade estatística. Observa-se uma ligeira queda quando compara-se o desempenho dos sistemas com cinco componentes, onde 86,67% das comparações foram estatisticamente melhores, mas ainda é melhor que os sistemas com três componentes, onde obteve-se apenas 40% de certeza de melhor desempenho estatístico.

	NH		H3	
Base	Fusão	Seleção	Fusão	Seleção
A	Não - 0,07383105	Não - 0,47505606	Sim - 0,04272615	Não - 0,35760382
B	Sim - 0,04659815	Sim - 0,00008614	Sim - 0,00878909	Sim - 0,00378668
C	Sim - 0,00335861	Sim - 0,00002751	Sim - 1,87E-13	Sim - 4,18E-08
D	Sim - 7,52E-08	Sim - 0,00178451	Sim - 1,09E-11	Sim - 5,80E-11
E	Sim - 0,00003751	Sim - 0,01942831	Sim - 0,00002458	Sim - 0,00626104
	H5		H7	
Base	Fusão	Seleção	Fusão	Seleção
A	Sim - 0,00001909	Sim - 0,00087235	Não - 0,29011638	Não - 0,13053520
B	Sim - 0,01696520	Sim - 0,00025943	Não - 0,35376177	Não - 0,31715769
C	Sim - 6,64E-10	Sim - 0,00052588	Não - 0,07272109	Não - 0,15310682
D	Sim - 1,66E-11	Sim - 1,30E-06	Não - 0,40893985	Não - 0,34466944
E	Sim - 8,11E-06	Sim - 0,00010503	Sim - 0,00535126	Não - 0,06732302

Tabela 5.15: Teste de Hipótese entre os melhores métodos de seleção e fusão com o melhor método de negociação para sistemas com Sete Componentes.

Analisando a primeira linha da Tabela 5.15, tem-se os valores de p para os teste com a base de dados A. Para esta base, apenas quando compara-se o melhor método de negociação com os melhores de fusão e seleção dos sistemas *H5* e a fusão dos sistemas *H3* pode-se dizer que o método de negociação é estatisticamente melhor. Nas bases de dados B, C e D, apenas para as comparações feitas com sistemas *H7* não se pode afirmar

nada, nos demais a negociação é melhor. E, por fim, para a base de dados E apenas a comparação do método de seleção com o de negociação é que não se pode afirmar nada.

5.4.5 Diversidade para Sistemas com Sete Componentes

Com o intuito de analisar o nível de diversidade dos sistemas, a Tabela 5.16 mostra os valores das três diversidades que foram aplicadas a todas as estruturas dos sistemas com sete componentes. Essas medidas de diversidade foram discutidas no capítulo anterior. Como já mencionado, o Q estatístico (q) e o dupla-falsa (df) são medidas que utilizam paridade e a entropia (e) é uma medida que não utiliza paridade.

BD		Medidas de Diversidade		
A	Var	q	df	e
	NH	0,80	0,28	0,82
	H3	0,83	0,24	0,85
	H5	0,88	0,24	0,87
	H7	0,91	0,20	0,92
B	Var	q	df	e
	NH	0,84	0,26	0,81
	H3	0,83	0,25	0,79
	H5	0,86	0,29	0,85
	H7	0,83	0,28	0,84
C	Var	q	df	e
	NH	0,85	0,27	0,85
	H3	0,84	0,28	0,82
	H5	0,90	0,21	0,87
	H7	0,88	0,19	0,90
D	Var	q	df	e
	NH	0,81	0,24	0,75
	H3	0,83	0,24	0,78
	H5	0,86	0,22	0,80
	H7	0,79	0,28	0,81
E	Var	q	df	e
	NH	0,87	0,20	0,83
	H3	0,84	0,25	0,84
	H5	0,88	0,20	0,82
	H7	0,83	0,19	0,79

Tabela 5.16: Medidas de Diversidade para Sistemas com Sete Componentes.

Na Tabela 5.16, pode-se observar que, quase sempre, variações nos tipos de componentes dos sistemas influenciam as diversidades. De acordo com o comportamento das medidas de diversidade, observa-se que as medidas baseadas em paridade tendem a aumentar o seu valor quando são calculadas para sistemas híbridos, neste caso os sistemas H3, H5 e H7. Por outro lado, as medidas que não são baseadas em paridade tendem a diminuir com o aumento da diversidade.

Neste tamanho de sistema, observa-se que, em geral, as diversidades tendem a manter

uma estabilidade. Este comportamento é esperado quando aumenta-se a quantidade de componentes dos sistemas.

De acordo com os resultados mostrados na Tabela 5.16 e mantendo o padrão do que foi observado nos sistemas com três e cinco componentes, para as três medidas conclui-se que sistemas híbridos são capazes de prover maiores níveis de diversidade que sistemas não híbridos. Outra conclusão importante, que mais uma vez foi observada, é que este fato também reflete na acurácia dos sistemas, já que os melhores resultados para sistemas com cinco componentes foram alcançados quando foram utilizados os sistemas *H7*.

5.5 Sistemas com Nove Componentes

Os resultados dos experimentos utilizando sistemas com nove componentes podem ser observados nas Tabelas 5.17, 5.18 e 5.19. Nesse tamanho de sistema, da mesma forma que para os sistemas com sete componentes, para cada classificação, existem quatro diferentes estruturas:

- Um tipo de componente (*NH*)
- Três tipos de componentes (*H3*)
- Cinco tipos de componentes (*H5*)
- Sete tipos de componentes (*H7*).

Para os sistemas não híbridos, foram possíveis construir sete variações, pois só existem sete tipos diferentes de classificadores. Para os experimentos dos sistemas híbridos, que neste tamanho de sistemas podem ser de três tipos, *H3*, *H5* e *H7*, foram construídos dez, sete e sete variações, respectivamente. Como já foi mencionado anteriormente, os valores apresentados nas tabelas citadas anteriormente são a média dos melhores desvios padrões de todas as possibilidades. Os valores em negrito representam o melhor desempenho entre as quatro variações do sistema, *NH*, *H3*, *H5* e *H7*.

Neste tamanho de sistemas, os valores de *Dif* são calculados os maiores e menores valores dentre os quatro tipos de sistemas, *NH*, *H3*, *H5* e *H7*. E, por conseguinte, os valores dos testes estatísticos, neste caso os valores de *p*, foram calculados para verificar se essas diferenças são estatisticamente significantes ou não.

Nestas tabelas, sete diferentes métodos de combinação foram analisados, que são: Voto, Naive Bayes (*NB*), Soma, Média, Mediana, *MLP*, FuzzyMLP (*FMLP*); três diferentes métodos de seleção foram analisados, que são: DCS-LA, DCS-DT, DCS-MCB; e três

diferetes métodos de negociação foram analisados, que são: Sensibilidade (*Sensibi*), *Lei-lão* e Teoria dos Jogos (*TJ*). Para definir o conjunto de treinamento para os combinadores treináveis, aproximadamente 10% das instâncias de cada base de dados foram separada para criar um conjunto de validação.

5.5.1 Métodos de Combinação ou Fusão para Sistemas com Nove Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de combinação para sistemas com nove componentes.

Na Tabela 5.17 pode ser observado os resultados dos sete métodos baseados em fusão que foram usados neste trabalho. Nesta tabela tem-se que, em geral, o sistema que obteve as melhores porcentagens de acerto foi o com sete tipos de classificadores diferentes. Isto é percebido quando se observa as bases de dados A, C e E. Nelas os melhores desempenhos são 100% pertencentes aos sistemas *H7*. Por outro lado, a base B apresenta mais da metade (Voto, *NB*, Média, *FMLP*) dos seus métodos com melhor desempenho com sistemas *H5*. E a base D, apenas o FuzzyMLP (*FMLP*) apresenta o melhor desempenho com sistemas *H5*. Os melhores desempenhos entre os sistemas foram, unanimamente alcançados pelo método FuzzyMLP (*FMLP*). Apesar disso, como foi mencionado anteriormente, para as bases B e D isto ocorre quando ele é executado para sistemas *H5*.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Em relação a diferença de desempenho entre os sistemas híbridos e não híbridos, ou seja, os valores de *Dif*, observa-se que, os maiores valores de *Dif*, para as bases B e C, acontecem quando é aplicado o combinador Soma. Já na base A isto acontece para o Voto, na base D, para o Naive Bayes (*NB*) e, para a base E, isto acontece para o método media. Analisando os menores valores de *Dif*, observa-se para todas as bases, as menores diferenças aconteceram com o método FuzzyMLP (*FMLP*). Mais uma vez, isto mostra que, de forma geral, os sistemas deste tipo são mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais complexos. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais simples.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de *p* deste teste pode ser observado na Tabela 5.17. Em relação aos maiores valores de *Dif*, os métodos avaliados são estatisticamente melhores que as suas variações. Em relação aos menores valores de *Dif*, conclui-se, também, que todos os métodos são estatisticamente melhores que suas variações.

Sistemas Com Nove Componentes								
BD	Métodos Baseados em Fusão							
	Var	Voto	NB	Soma	Média	Mediana	MLP	FMLP
A	NH	75,32±5,87	81,21±6,12	79,23±5,56	80,11±7,03	78,25±7,95	91,68±3,61	93,96±2,88
	H3	78,57±2,98	83,98±2,51	82,98±3,35	83,23±4,97	81,86±6,39	93,46±2,11	95,86±2,37
	H5	84,31±3,26	88,78±3,67	87,20±3,51	87,08±6,42	85,36±6,64	95,48±2,61	96,71±2,12
	H7	85,69±2,16	90,51±3,22	88,84±2,24	87,95±6,50	86,88±6,82	96,72±1,90	96,73±2,00
	Dif	10,37	9,30	9,61	7,84	8,63	5,04	2,77
	p	3,12E-28	1,61E-21	4,51E-27	1,15E-10	9,13E-11	3,32E-19	3,90E-10
B	NH	85,21±4,60	88,89±6,66	85,01±3,38	88,41±6,03	84,80±4,70	95,49±1,81	96,84±2,61
	H3	82,86±4,63	87,90±3,87	85,36±3,42	86,91±5,21	84,13±5,48	95,43±2,58	96,44±2,51
	H5	86,07±4,11	90,70±4,60	87,31±4,07	89,75±5,99	86,39±4,65	95,70±2,40	97,20±2,40
	H7	85,69±2,16	90,51±3,22	88,84±2,24	87,95±6,50	86,88±6,82	96,72±1,90	96,73±2,00
	Dif	3,21	2,80	3,83	2,84	2,75	1,29	0,76
	p	0,00000315	0,0000147	1,08E-13	0,00061	0,00925	0,000597	0,0247
C	NH	74,80±3,93	81,50±2,64	75,92±3,65	79,94±4,08	75,06±6,22	89,76±2,84	91,71±2,92
	H3	78,36±3,00	82,96±2,17	79,64±2,25	81,79±3,63	79,09±5,48	91,83±1,47	92,59±2,56
	H5	82,54±4,84	86,92±6,40	85,45±5,34	86,49±6,53	83,76±7,49	93,03±2,39	94,60±2,34
	H7	85,43±2,19	90,09±3,91	88,17±2,37	87,81±6,05	86,49±6,40	96,19±1,78	96,78±1,88
	Dif	10,63	8,59	12,25	7,87	11,43	6,43	5,07
	p	1,92E-42	1,21E-31	1,71E-50	6,81E-16	3,78E-20	1,28E-33	4,87E-24
D	NH	77,45±3,88	82,42±3,47	80,38±3,03	81,27±4,86	79,44±6,67	91,23±2,38	94,42±1,97
	H3	79,45±4,57	83,60±3,93	82,06±4,01	82,26±5,03	80,15±5,94	93,00±3,59	95,34±2,35
	H5	81,61±4,72	85,08±3,97	83,01±3,48	84,56±5,38	81,94±5,55	93,33±2,53	96,50±2,36
	H7	82,56±4,87	88,80±3,57	85,90±5,04	86,46±6,12	84,71±6,86	96,20±1,78	96,38±1,88
	Dif	5,11	6,38	5,52	5,19	5,27	4,97	2,08
	p	9,95E-11	3,27E-20	5,02E-13	6,79E-08	0,00000462	1,72E-28	4,35E-08
E	NH	78,98±5,06	83,02±4,92	80,49±4,71	81,39±4,89	79,57±5,31	92,16±3,38	93,25±2,75
	H3	80,13±5,07	84,37±4,80	82,49±4,56	83,27±4,91	81,50±5,52	92,28±3,23	94,31±2,64
	H5	82,77±5,52	86,64±6,09	84,86±6,04	85,46±5,60	84,15±4,98	94,41±4,05	96,60±2,64
	H7	84,65±3,27	88,71±4,37	87,77±3,47	87,00±6,56	85,56±6,81	96,27±1,97	96,73±1,96
	Dif	5,67	5,69	7,28	5,61	5,99	4,11	3,48
	p	4,47E-13	0,00061	2,02E-19	2,88E-08	2,15E-08	2,73E-15	7,32E-15

Tabela 5.17: Método baseado em Fusão para Sistemas com Nove Componentes.

5.5.2 Métodos Baseados em Seleção para Sistemas com Nove Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de seleção para sistemas com nove componentes.

Diante dos resultados apresentados sobre os métodos de seleção na Tabela 5.18, conclui-se que, de maneira geral, os melhores desempenhos são alcançados utilizando os sistemas *H7*. Porém, para a base B, os métodos DCS-DT e DCS-MCB apresentam melhor desempenho com sistemas *H5*. Na base A apenas o método DCS-DT apresenta melhores resultados com *H5* e na base E isto ocorre com o método DCS-MCB. Avaliando os melhores desempenhos para cada base, conclui-se que, de forma geral, o método DCS-MCB

apresenta os melhores resultados. Apenas na base D isto ocorre com o método DCS-DT. Outro detalhe interessante é que nas bases B e E os métodos com melhor desempenho do conjunto são para sistemas *H5*.

Sistemas Com Nove Componentes						
BD	Métodos Baseados em Seleção					
A	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	93,09±2,41	93,34±2,77	21,13	95,09±3,10	13,13
	H3	96,23±1,69	95,81±1,49	21,91	95,52±1,81	13,49
	H5	96,63±1,13	97,04±1,34	21,78	97,10±2,15	12,36
	H7	96,89±1,20	96,86±1,35	22,42	97,38±1,78	12,28
	Dif	3,80	3,70	-	2,29	-
	p	5,79E-23	1,52E-18	-	0,000000173	-
B	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	95,95±2,17	96,40±2,38	20,19	96,54±1,80	16,47
	H3	95,55±1,98	96,43±2,15	21,64	96,47±2,62	16,04
	H5	96,18±1,83	97,39±1,87	19,96	97,73±1,87	15,09
	H7	96,62±1,20	97,18±1,41	22,26	97,38±1,78	12,22
	Dif	1,07	1,33	-	1,26	-
	p	0,0000379	0,00354	-	0,000349	-
C	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	91,97±2,70	93,59±3,02	20,06	93,80±3,38	13,67
	H3	92,83±2,31	95,15±2,67	20,82	95,87±2,28	13,95
	H5	95,10±2,81	96,58±1,83	19,08	96,29±2,63	13,50
	H7	96,30±1,55	96,93±1,46	22,56	96,99±1,70	13,59
	Dif	4,33	3,34	-	3,19	-
	p	1,51E-22	3,52E-14	-	4,08E-11	-
D	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	91,09±2,74	94,07±3,08	21,48	95,58±2,82	14,74
	H3	92,89±2,68	95,38±3,15	19,82	95,96±2,61	12,96
	H5	95,53±2,61	96,63±3,32	20,15	96,43±2,84	13,47
	H7	96,46±1,16	96,77±1,62	22,17	96,67±1,96	13,02
	Dif	5,37	2,70	-	1,09	-
	p	2,87E-31	7,15E-10	-	0,00444	-
E	Var	DCS-LA	DCS-DT	Uso Fusão	DCS-MCB	Uso Voto
	NH	91,31±2,78	92,63±1,73	22,10	94,01±2,27	16,30
	H3	92,00±2,73	93,22±1,98	22,77	95,10±2,18	16,26
	H5	92,84±3,04	94,13±2,03	21,31	97,32±1,67	14,85
	H7	96,37±1,49	96,59±1,99	21,49	96,73±1,97	12,73
	Dif	5,06	3,96	-	3,31	-
	p	4,52E-27	6,48E-25	-	7,74E-18	-

Tabela 5.18: Método baseado em Seleção para Sistemas com Nove Componentes.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Em relação a diferença de desempenho entre os sistemas híbridos e não híbridos, ou seja, os valores de *Dif* da Tabela 5.18, observa-se que os maiores valores de *Dif* ocorrem quando é utilizado o método DCS-LA. Apenas na base B, esse fato ocorre com o método DCS-DT. Avaliando os menores valores de *Dif* agora, tem-se que, de maneira geral, o método que apresenta as menores diferenças é o DCS-MCB. Apenas na base B isto ocorre com o método DCS-LA. Da mesma forma que nos sistemas com três, cinco e sete componentes, isto mostra que, de forma geral, os sistemas deste tipo são mais afetados pelo uso de componentes

diferentes são aqueles que são combinados por métodos mais simples. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais complexos.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de p deste teste pode ser observado na Tabela 5.18. Em relação aos maiores valores de Dif , conclui-se que, da mesma maneira que nos métodos de fusão, todos os métodos são estatisticamente melhores que suas variações. Isto se repete quando os métodos com menores valores de Dif são analisados.

Na maioria dos casos mais uma vez, o método DCS-MCB possui um desempenho melhor que o método DCS-DT. Observando o método DCS-DT, tem-se que a média do uso do método de fusão aumenta, diminui e aumenta para cada uma das variações dos sistemas $H3$, $H5$ e $H6$, para as bases A, B, C e D. Já para a base E, essa média sofre um aumento seguido de uma diminuição. Agora, observando o DCS-MCB, tem-se que a média do uso do método Voto para resolver o sistema possui várias variações dependendo da base utilizada. Para a base A essa média sofre um aumento, seguido de diminuição e, por fim, um aumento para os sistemas $H3$, $H5$ e $H7$. Já a base D sofre a reação contrária. Para as bases B e E, essa média sofre apenas redução de acordo com o aumento da quantidade dos tipos dos classificadores. E, por fim, a base A sofre aumento seguido de redução desta média.

Para este tamanho de sistema, comparando-se os métodos de seleção e fusão, observa-se que os de seleção apresentam um desempenho melhor que os de fusão. Como pode ser visto nas Tabelas 5.17 e 5.18, para todas as bases de dados o melhor método de fusão foi sempre o FuzzyMLP. Já o melhor método de seleção, foi em geral, o método DCS-MCB, mas, na base D isto ocorre com o método DCS-DT. Mas sempre os métodos FuzzyMLP possuíram desempenho inferior que os métodos baseados em seleção. Esse comportamento foi igual ao encontrado nos sistemas com sete componentes.

5.5.3 Métodos de Negociação para Sistemas com Nove Componentes

Nesta seção serão mostrados e analisados os resultados dos métodos de negociação para sistemas com nove componentes.

Analisando a Tabela 5.19 e observando os métodos de negociação, pode-se afirmar que, em geral assim como nos *Ensembles*, eles apresentam melhor desempenho quando são utilizados sistemas $H7$. Porém, para as bases A e C o método Sensibilidade (*Sensibi*) apresenta melhor desempenho com sistemas do tipo $H5$. E, na base B, a Teoria dos Jogos (*TJ*) e, mais uma vez, a Sensibilidade (*Sensibi*) apresentam um melhor desempenho com

o sistema do tipo *H5* também. Conclui-se, ainda, que o método baseado em Sensibilidade (*Sensibi*) apresenta, para todas as bases, o melhor desempenho. Para as bases A, B e D, como foi mencionado anteriormente, apresentam esse resultado para sistemas *H5*, mas as demais são para sistemas *H7*.

Sistemas Com Nove Componentes						
BD	Métodos de Negociação					
	Var	Leilao	Sensibi	iteracao Sensibi	TJ	iteracao TJ
A	NH	89,72±2,61	93,32±3,07	15,61	94,05±2,68	10,85
	H3	92,60±2,53	96,06±2,01	16,20	94,91±2,16	11,66
	H5	95,38±2,48	98,01±1,97	15,71	96,60±2,02	10,50
	H7	95,79±2,02	97,65±2,07	16,47	96,89±1,89	10,47
	Dif	6,07	4,69	-	2,84	-
	p	4,88E-32	3,11E-20	-	1,28E-11	-
B	NH	93,44±4,44	96,90±2,17	12,70	95,28±2,96	10,81
	H3	95,56±2,77	97,51±1,96	13,14	96,34±2,52	10,85
	H5	95,68±2,78	98,58±2,17	11,69	97,27±2,12	9,83
	H7	95,79±2,02	97,65±2,07	16,60	96,89±1,89	10,47
	Dif	2,35	1,07	-	1,99	-
	p	0,0000907	0,00052	-	0,00000533	-
C	NH	89,93±3,37	93,73±3,18	15,44	92,41±3,61	9,89
	H3	92,82±2,46	96,04±2,03	15,65	95,32±2,14	10,51
	H5	94,44±2,82	96,93±1,92	14,15	95,57±2,57	9,32
	H7	94,80±2,62	97,61±2,04	15,45	96,35±1,96	10,40
	Dif	4,87	3,88	-	3,94	-
	p	3,03E-17	7,29E-15	-	2,09E-13	-
D	NH	91,96±3,00	95,30±2,23	14,17	93,72±2,60	11,28
	H3	92,02±3,73	96,96±1,79	14,12	94,42±3,08	11,17
	H5	92,48±4,00	98,36±1,69	13,41	95,86±2,46	10,70
	H7	95,16±2,51	97,68±2,34	15,95	96,59±1,75	11,05
	Dif	3,20	3,06	-	2,87	-
	p	1,17E-10	3,79E-16	-	1,49E-12	-
E	NH	91,22±2,29	94,82±2,70	15,49	92,50±2,33	10,60
	H3	92,58±3,35	96,27±2,91	16,43	94,73±2,86	11,30
	H5	93,32±4,21	97,74±2,67	16,11	95,39±3,00	10,25
	H7	94,81±3,29	97,85±1,93	15,97	96,15±2,10	10,40
	Dif	3,59	3,03	-	3,65	-
	p	3,60E-12	1,68E-12	-	1,16E-17	-

Tabela 5.19: Método de Negociação para Sistemas com Nove Componentes.

Como já foi explicado anteriormente, a medida *Dif* serve para mostrar como a diversidade dos componentes influencia o resultado do sistema. Em relação a diferença de desempenho entre os sistemas híbridos e não híbridos, ou seja, os valores de *Dif* da Tabela 5.19, observa-se que as maiores diferenças ocorreram, em geral, com o método *Leilão*. Apenas para a E isto ocorre quando o método de Teoria dos Jogos (*TJ*). Ainda observando essas diferenças, mas agora vendo as menores delas, conclui-se que para as bases A e D isto ocorre para o método Teoria dos Jogos (*TJ*). Já para a base B, C e E isto ocorre quando o método baseado em Sensibilidade (*Sensibi*) é aplicado. Aqui, destoando

do quem vem acontecendo, observa-se que, de forma geral, os sistemas mais afetados pelo uso de componentes diferentes são aqueles que são combinados por métodos mais simples. Por outro lado, de forma geral, os sistemas menos afetados são aqueles que são combinados por métodos mais complexos.

Para verificar se essas diferenças são significantemente estatísticas, foi feito um teste estatístico e o valor de p deste teste pode ser observado na Tabela 5.19. Em relação aos sistemas que alcançaram os maiores valores de Dif , pode-se afirmar que todos esses sistemas são estatisticamente melhores que suas variações. E, da mesma maneira, os sistemas que apresentaram os menores valores de Dif também são estatisticamente melhores que suas variações.

Avaliando o método Teoria dos Jogos (TJ), observa-se que para as bases A e E acontece um aumento quando aumenta-se a quantidade de tipos de agentes e diminui. Já a base D apresenta o comportamento contrário. Para as bases B e C acontece um aumento para o sistema $H3$, diminuição para sistema $H5$ e aumento para $H7$. Para o método baseado em Sensibilidade ($Sensibi$), as bases A, B e C apresentam um aumento para sistema $H3$, uma diminuição para sistema $H5$ e mais uma vez um aumento para sistema $H7$. Para a base D, tem-se que o sistema sofre uma diminuição seguido de um aumento. Já a base E apresenta uma desempenho contrária. Logo, conclui-se que de forma geral, o número de *rounds* aumenta quando trabalha-se com estruturas híbridas. Isto pode ser explicado pelo fato de que quando se trabalha com componentes diferentes estes demoram mais a chegar a uma opinião comum.

5.5.4 Negociação versus Combinação para Sistemas com Nove Componentes

Comparando-se, agora, os métodos de negociação e os de combinação, que englobam os métodos de seleção e fusão, observa-se que, para todas as bases de dados, mais uma vez, o método de negociação baseado em Sensibilidade ($Sensibi$) obteve o melhor desempenho. Para saber se um método é estatisticamente melhor que o outro, realiza-se um teste de hipótese entre esse método e os demais. Os resultados dos testes estatísticos entre o melhor método de negociação dos sistemas NH , $H3$, $H5$ e $H7$ e os melhores métodos baseados em fusão e seleção podem ser observados na Tabela 5.20.

Na Tabela 5.20, quando aparece 'Sim' significa que o melhor método de negociação é estatisticamente melhor que o melhor método do tipo que está sendo comparado e 'Não' significa que não se pode afirmar nada sobre os sistemas comparados. Para sistemas com nove componentes e considerando que existem três tipos de sistemas híbridos, assim como nos sistemas com sete componentes, foram feitos quarenta testes estatísticos. Como

resultado, tem-se que, mais uma vez, em 28 das 40 comparações o método de negociação é estatisticamente melhor. Isso significa que em 70% das comparações conseguiu-se uma superioridade estatística. Observa-se uma estabilização do desempenho quando compara-se o desempenho dos sistemas com sete componentes que obteve 70% também.

Base	NH		H3	
	Fusão	Seleção	Fusão	Seleção
A	Não - 0,42539670	Sim - 0,01754553	Não - 0,26506068	Não - 0,25523294
B	Não - 0,44548408	Não - 0,14375855	Sim - 0,00045368	Sim - 0,00085046
C	Sim - 0,00007448	Não - 0,44583649	Sim - 2,85E-21	Não - 0,29089447
D	Sim - 0,00707814	Não - 0,26433799	Sim - 6,40E-08	Sim - 0,00091522
E	Sim - 0,00042150	Sim - 0,02887973	Sim - 6,92E-07	Sim - 0,00073519
Base	H5		H7	
	Fusão	Seleção	Fusão	Seleção
A	Sim - 0,00012606	Sim - 0,00514564	Sim - 0,00439235	Não - 0,20971863
B	Sim - 0,00024623	Sim - 0,00712848	Sim - 0,00764592	Não - 0,31759931
C	Sim - 8,67E-10	Não - 0,13655142	Sim - 0,00671829	Sim - 0,02587523
D	Sim - 1,77E-07	Sim - 0,00007658	Sim - 0,00019911	Sim - 0,00430161
E	Sim - 0,00610086	Não - 0,13311312	Sim - 0,00293487	Sim - 0,00045437

Tabela 5.20: Teste de Hipótese entre os melhores métodos de seleção e fusão com o melhor método de negociação para sistemas com Nove Componentes.

A primeira linha da Tabela 5.20 mostra os resultados das comparações dos melhores métodos de negociação com os de seleção e os de fusão. A partir desta linha, pode-se concluir que para sistemas *H3*, para a fusão dos sistemas *NH* e para a seleção dos sistemas *H7* não se pode afirmar nada, as demais comparações são estatisticamente melhores. Da mesma forma, a segunda linha representa os resultados dos teste com a base B. Nesta linha, os experimentos feitos com os sistemas *H3*, *H5* e a fusão do sistema *H7* são estatisticamente melhores que suas variações.

A terceira linha da Tabela 5.20 mostra os resultados das comparações feitas com a base C. Para todas as seleções desta base, exceto a feita para o sistema *H7*, não se pode afirmar nada. Já para a base D, representada na quarta linha, apenas para seleção do sistema *NH* não se pode afirmar nada. E, por fim, para a base E, apenas a seleção do sistema *H5* não se pode afirmar nada.

5.5.5 Diversidade para Sistemas com Nove Componentes

Com o intuito de analisar o nível de diversidade dos sistemas, a Tabela 5.21 mostra os valores das três diversidades que foram aplicadas a todas as estruturas dos sistemas com sete componentes. Essas medidas de diversidade foram discutidas no capítulo anterior. Como já mencionado, o Q estatístico (q) e o dupla-falsa (df) são medidas que utilizam paridade e a entropia (e) é uma medida que não utiliza paridade.

BD		Medidas de Diversidade		
A	Var	<i>q</i>	<i>df</i>	<i>e</i>
	NH	0,80	0,34	0,82
	H3	0,83	0,33	0,83
	H5	0,83	0,28	0,85
	H7	0,84	0,29	0,85
B	Var	<i>q</i>	<i>df</i>	<i>e</i>
	NH	0,84	0,26	0,81
	H3	0,83	0,25	0,79
	H5	0,86	0,29	0,85
	H7	0,83	0,28	0,84
C	Var	<i>q</i>	<i>df</i>	<i>e</i>
	NH	0,81	0,23	0,82
	H3	0,83	0,24	0,81
	H5	0,84	0,26	0,84
	H7	0,85	0,28	0,85
D	Var	<i>q</i>	<i>df</i>	<i>e</i>
	NH	0,84	0,28	0,82
	H3	0,84	0,25	0,83
	H5	0,90	0,24	0,86
	H7	0,84	0,28	0,86
E	Var	<i>q</i>	<i>df</i>	<i>e</i>
	NH	0,88	0,23	0,84
	H3	0,88	0,20	0,85
	H5	0,87	0,20	0,84
	H7	0,82	0,28	0,85

Tabela 5.21: Medidas de Diversidade para Sistemas com Nove Componentes.

Na Tabela 5.21, pode-se observar que, quase sempre, variações nos tipos de componentes dos sistemas influenciam as diversidades. De acordo com o comportamento das medidas de diversidade, observa-se que as medidas baseadas em paridade tendem a aumentar o seu valor quando são calculadas para sistemas híbridos, neste caso os sistemas *H3*, *H5* e *H7*. Por outro lado, as medidas que não são baseadas em paridade tendem a diminuir com o aumento da diversidade.

Assim como nos sistemas com sete componentes, neste tamanho de sistema, observa-se que, em geral, as diversidades tendem a manter uma estabilidade. Este comportamento é esperado quando aumenta-se a quantidade de componentes dos sistemas.

De acordo com os resultados mostrados na Tabela 5.21 e mantendo o padrão do que foi observado nos sistemas com três, cinco e com sete componentes, para as três medidas conclui-se que sistemas híbridos são capazes de prover maiores níveis de diversidade que sistemas não híbridos. Outra conclusão importante, que mais uma vez foi observada, é que este fato também reflete na acurácia dos sistemas, já que os melhores resultados para sistemas com cinco componentes foram alcançados quando foram utilizados os sistemas *H7*.

5.6 Discussão dos Resultados

O principal objetivo deste trabalho era analisar o comportamento do Sistema *ClassAge* diante da variação dos seus principais parâmetros. Para isto, diversas estruturas e tamanhos de sistemas foram analisados.

Como um resultado desta investigação, pode-se concluir que, como esperado, o padrão dos desempenhos é dependente dos dados utilizados, ou seja, os resultados dependem da base de dados utilizada. Entretanto, algumas indicações sobre esses experimentos podem ser percebidas. Estas indicações serão mostradas a seguir.

Como, nesta parte do trabalho, avalia-se os resultados de forma geral, utilizou-se uma nomenclatura diferenciada. Todas as vezes que observa-se os sistemas com três componentes, coloca-se um número 3 antes da configuração do sistema, como por exemplo, sistemas híbridos com três componentes 3-NH. Isso foi feito para todos os tamanhos de sistemas.

5.6.1 Tamanho dos Sistemas

Neste trabalho, foram analisados quatro tamanhos diferentes de sistema: sistemas com três, cinco, sete e nove componentes. Diante do que foi exposto neste capítulo conclui-se que a acurácia dos sistemas, em geral, aumenta com o aumento do número de componentes.

Isto já era um comportamento esperado, pois quanto mais especialistas tentando classificar um padrão de entrada maior a chance de acerto, pois eles complementam informação e conhecimento. De forma geral, para cada uma das bases de dados, esse comportamento pode ser melhor observado no Apêndice A.

5.6.2 Estrutura dos Sistemas

Os melhores resultados foram sempre alcançados pelas estruturas híbridas, para todos os quatro tamanhos de sistemas analisados. Esse comportamento pode ser melhor observado no Apêndice A, onde todos os resultados para todas as configurações para todas as bases de dados estão expostos.

De forma geral, o comportamento padrão percebido é que sempre os melhores desempenhos são alcançados pelos sistemas completamente híbridos, como já foi dito, seguido pelos parcialmente híbridos, respectivamente. Sempre os sistemas não híbridos, ou seja, sistemas *NH* possuem as acurácias inferiores.

5.6.3 Métodos de Combinação

Neste trabalho foram investigados sete métodos de combinação que também são conhecidos como métodos baseados em fusão. Esses métodos foram a MLP, a Fuzzy MLP, o Voto, a Soma, a Média, a Mediana e o Naive Bayes. Para esses métodos, observou-se que os sistemas são combinados pelo método Fuzzy MLP apresentaram, de forma geral, os melhores desempenhos para todas as bases de dados. Neste caso, o único padrão percebido foi que, em geral, os desempenhos do método *MLP* estão próximos dos do Fuzzy MLP, se distanciando consideravelmente dos demais métodos.

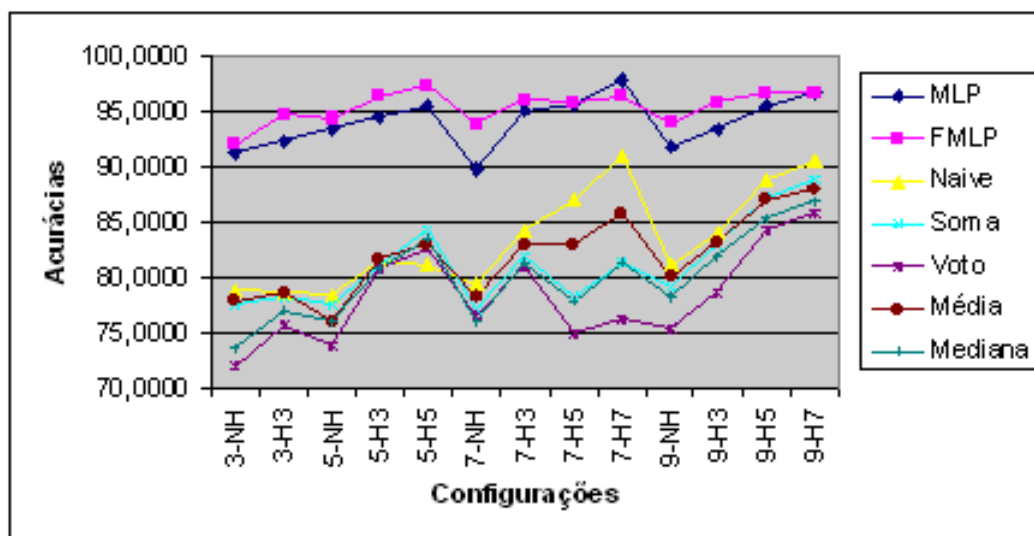


Figura 5.1: Resultados dos Métodos de Fusão para base A (breast cancer)

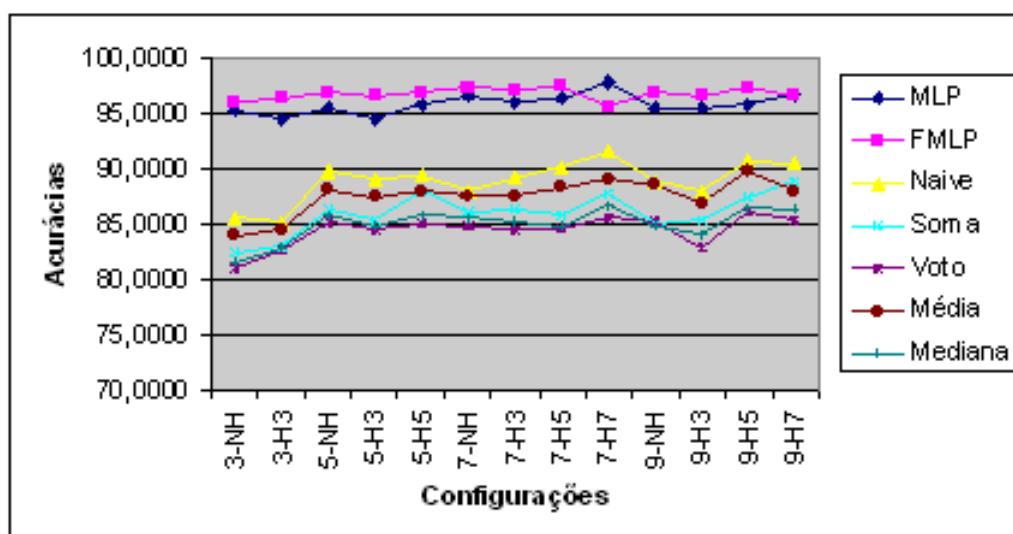


Figura 5.2: Resultados dos Métodos de Fusão para base B (imagem)

Como já foi dito anteriormente, este comportamento é esperado, já que os métodos *MLP* e Fuzzy MLP são os métodos mais complexos em relação aos outros métodos base-

ados em fusão, por este motivo, eles conseguem realizar a classificação com mais eficiência. Esse comportamento pode ser observado nas Figuras 5.1, 5.2, 5.3, 5.4 e 5.5.

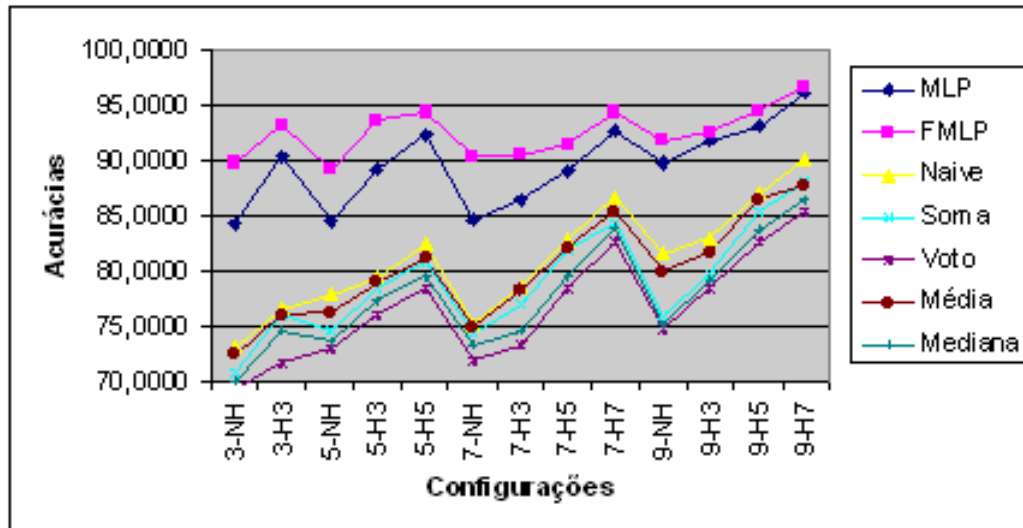


Figura 5.3: Resultados dos Métodos de Fusão para base C (proteínas)

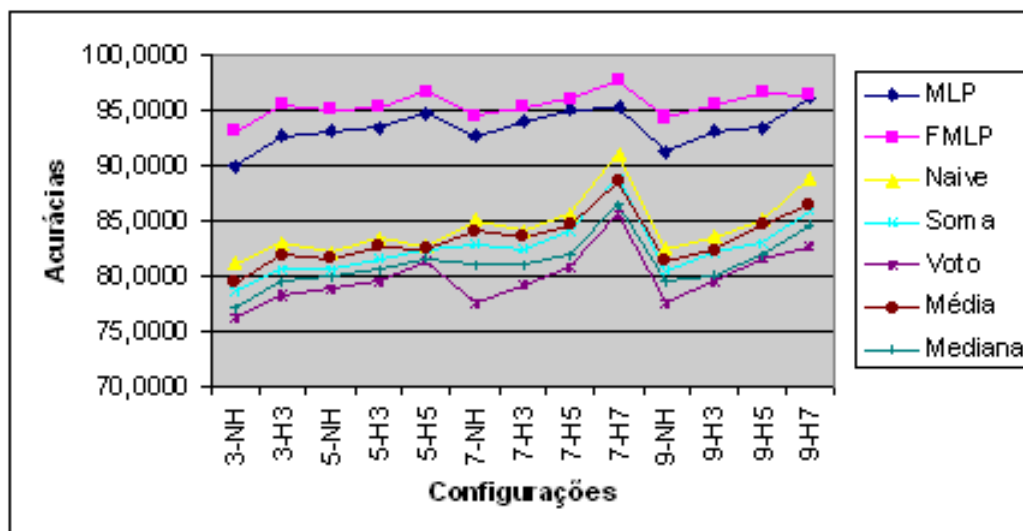


Figura 5.4: Resultados dos Métodos de Fusão para base D (splice)

É importante observar que o comportamento dos resultados é bastante semelhante, ou seja, quando um método possui uma melhora de resultado para uma determinada configuração, os demais também apresentam o mesmo comportamento. Os decréscimos das acurácias nos gráficos é dada pela mudança na quantidade de componentes da configuração que está sendo representada.

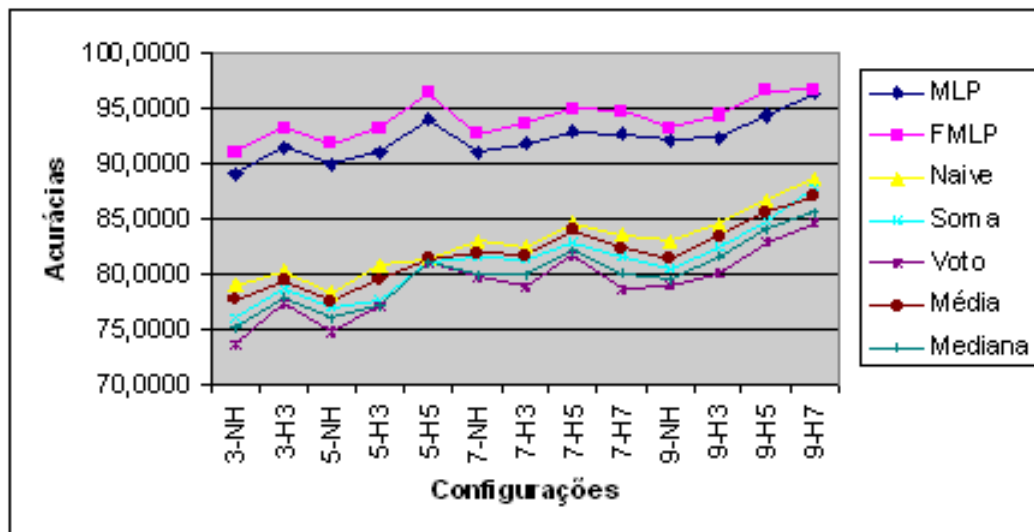


Figura 5.5: Resultados dos Métodos de Fusão para base E (wisconsin)

Outro comportamento interessante é que os valores dos resultados sempre aumentam de acordo com o aumento do número de componentes e com o aumento da diversidade do sistema, ou seja, quanto mais tipos de classificadores existir no sistema, maior é a acurácia desse sistema.

De forma geral, a base de dados que possui as melhores acurácias é a base B, base de imagens. Por outro lado, a base de dados que possui as menores acurácias é a base C, base de proteínas. As bases que possuem as menores variações de acurácias são as bases B e E. Já as bases que possuem as maiores variações de acurácias são as bases A, C e D.

5.6.4 Métodos de Seleção

Neste trabalho foram investigados três métodos baseados em seleção. Esses métodos foram os classificadores de seleção dinâmica ou DCSs: DCS-LA, DCS-DT e DCS-MCB. Para esses métodos, observou-se que os sistemas que trabalharam com o método DCS-MCB apresentaram, de forma geral, os melhores desempenhos para todas as bases. As Figuras 5.6, 5.7, 5.8, 5.9 e 5.10 mostram esse comportamento.

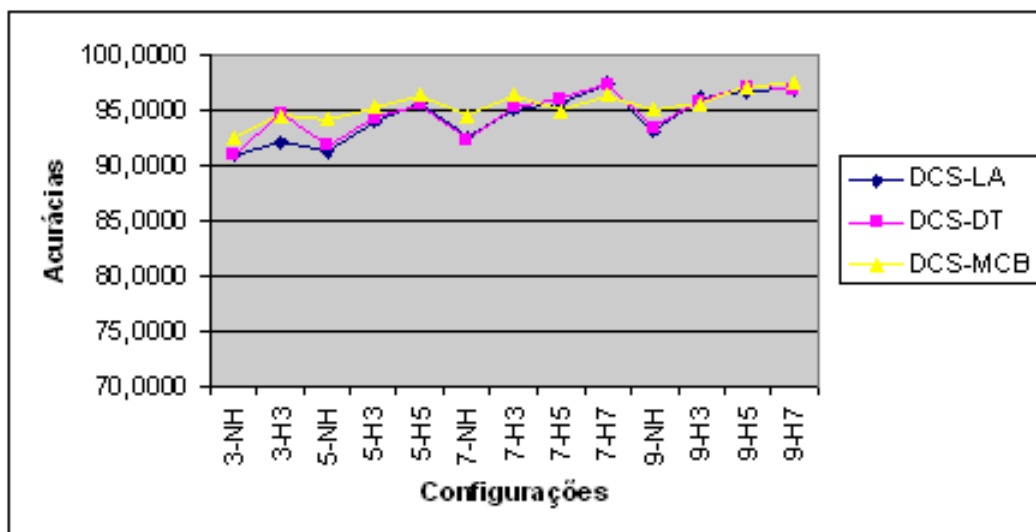


Figura 5.6: Resultados dos Métodos de Seleção para base A (breast cancer)

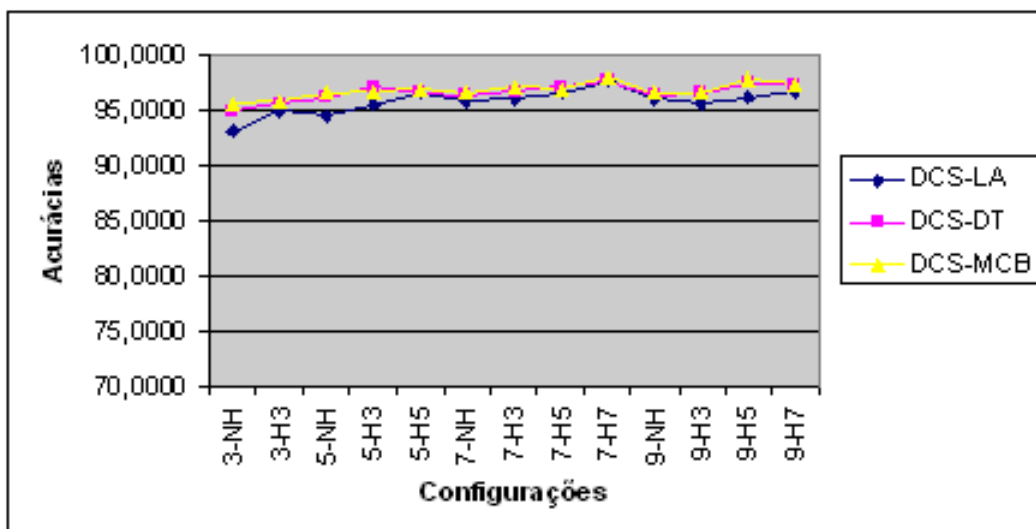


Figura 5.7: Resultados dos Métodos de Seleção para base B (imagem)

No caso comentado anteriormente, o comportamento padrão percebido é que os melhores desempenhos são alcançados pelo método DCS-MCB, depois pelo DCS-DT e, por fim, com pior desempenho, pelo DCS-LA. Mais uma vez, como já foi mencionado anteriormente, este comportamento é esperado pois os métodos DCS-MCB e DCS-LA são métodos híbridos, por isso podem ter mais chance de acertar na classificação.

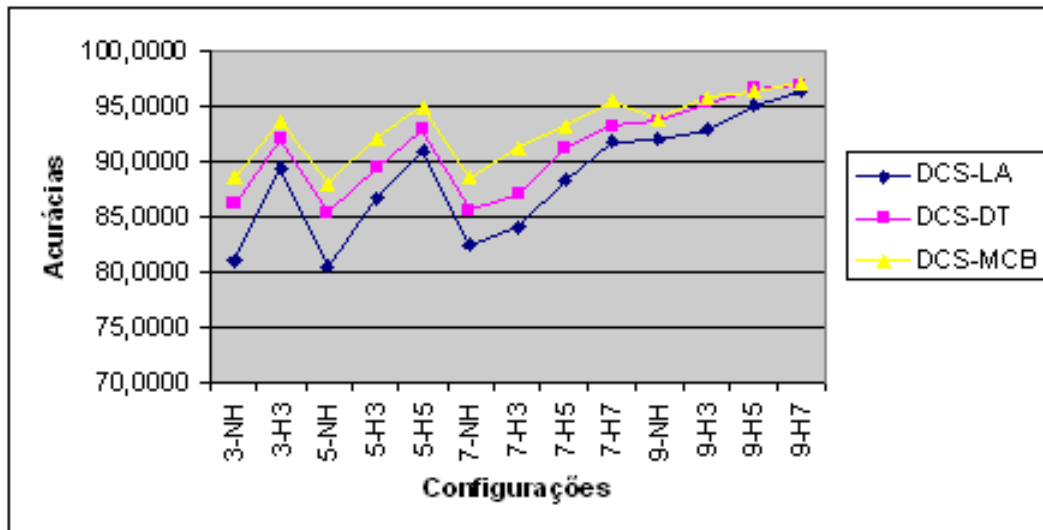


Figura 5.8: Resultados dos Métodos de Seleção para base C (proteínas)

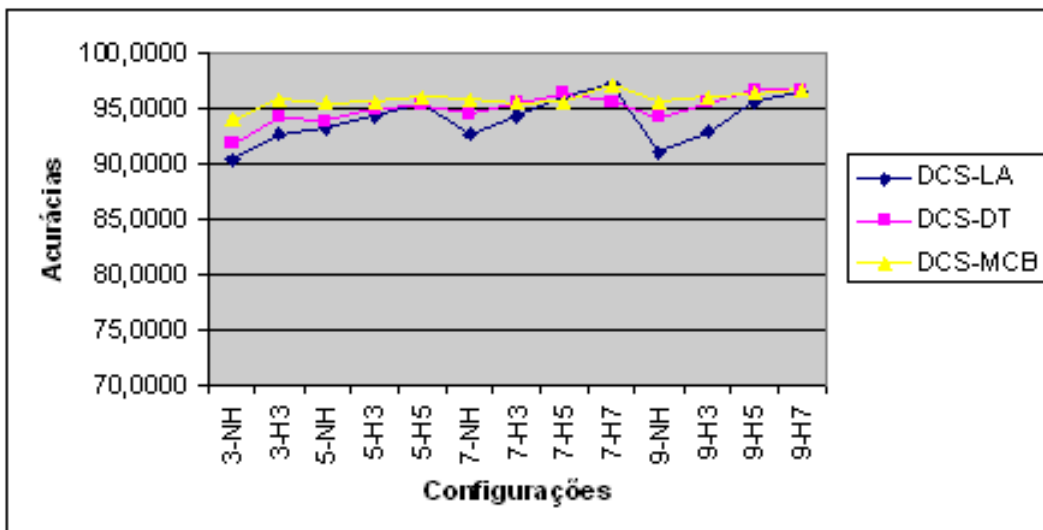


Figura 5.9: Resultados dos Métodos de Seleção para base D (splice)

Assim como foi observado nos métodos baseados em fusão, outro comportamento interessante é que os valores dos resultados sempre aumentam de acordo com o aumento do número de componentes e com o aumento da diversidade do sistema, ou seja, quanto mais tipos de classificadores existir no sistema, maior é a acurácia desse sistema.

Mais uma vez, da mesma forma que foi observado nos métodos baseados em seleção,

de forma geral, a base de dados que possui as melhores acurácias é a base B, base de imagens. Por outro lado, a base de dados que possui as menores acurácias é a base C, base de proteínas. As bases que possuem as menores variações de acurácias são as bases B, D e E. Já as bases que possuem as maiores variações de acurácias são as bases A e C.

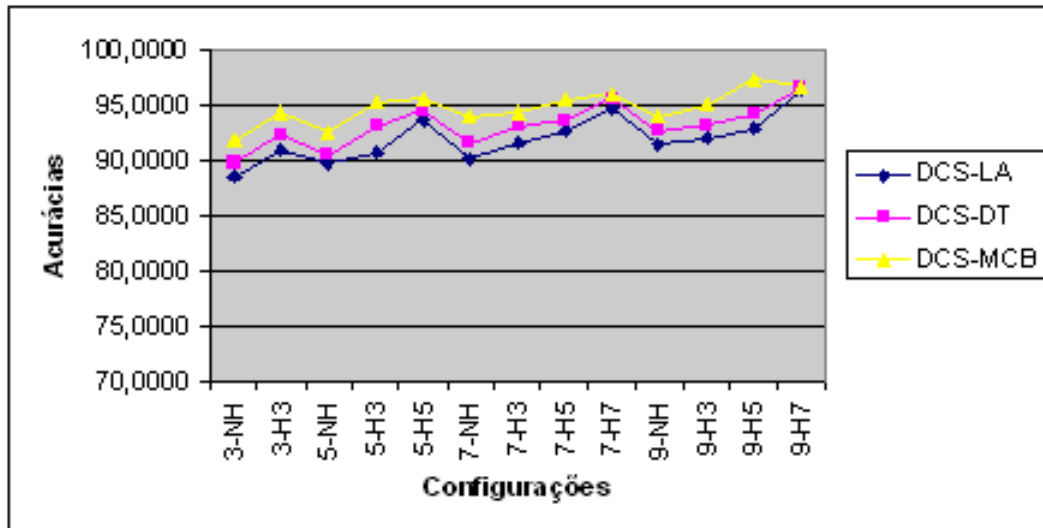


Figura 5.10: Resultados dos Métodos de Seleção para base E (wisconsin)

Ainda comparandos os métodos de seleção com os métodos baseados em fusão, observa-se que os valores mínimos das acurácias dos métodos de seleção são bem superiores que os métodos baseados em fusão. Esse comportamento é esperado, já que os métodos baseados em seleção foram propostos como solução do problema de tomada de decisão centralizada.

5.6.5 Métodos de Negociação

Dentre os métodos de negociação, o método baseado em Sensibilidade sempre apresentou a melhor desempenho. Isto acontece porque ele trabalha com mais informações que os demais métodos, conseguindo, assim, fazer uma análise mais detalhada do problema. Porém, ele necessita de mais tempo, ou seja, mais iterações do algoritmo para chegar a uma conclusão.

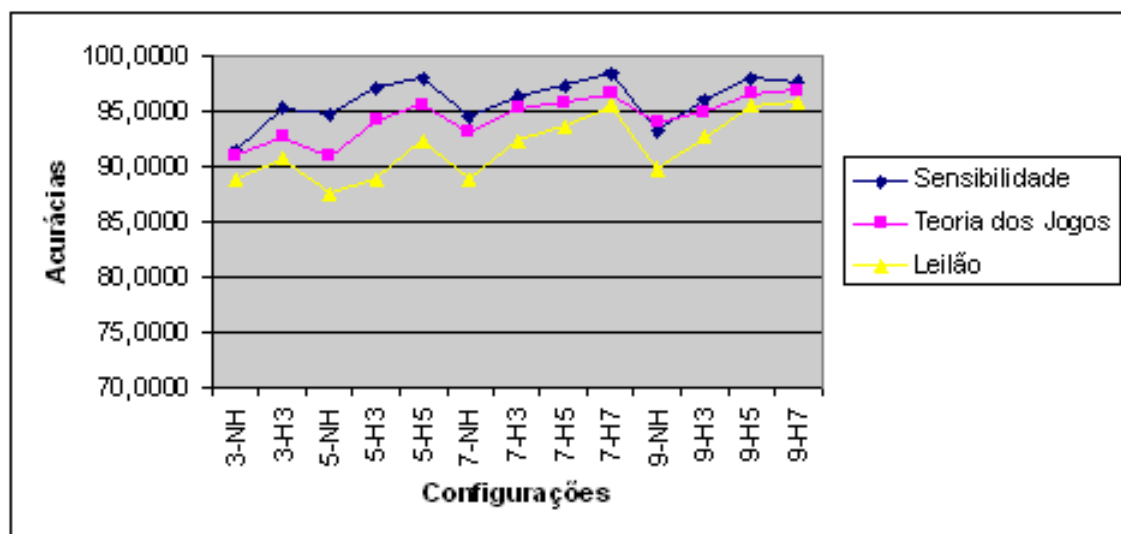


Figura 5.11: Resultados dos Métodos de Negociação para base A (breast cancer)

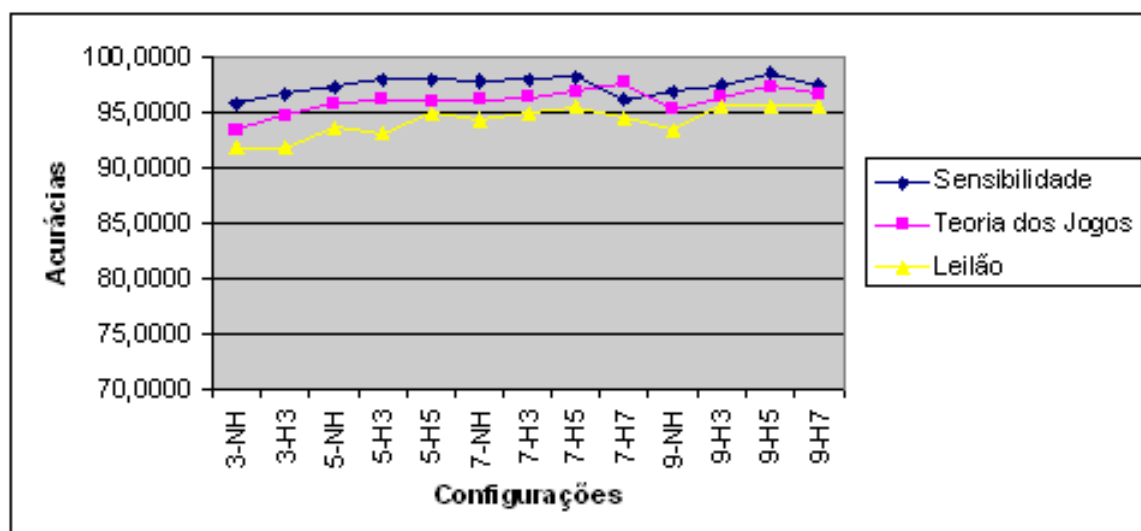


Figura 5.12: Resultados dos Métodos de Negociação para base B (imagem)

Para este conjunto de métodos, o padrão que se percebe é que, de maneira geral, os melhores desempenhos são alcançadas pelo método baseado em Sensibilidade, seguido pelo método baseado na Teoria dos Jogos, e, por fim, com os piores desempenhos, o

método baseado em *Leilão*. Esse comportamento pode ser observado nas Figuras 5.11, 5.12, 5.13, 5.14 e 5.15.

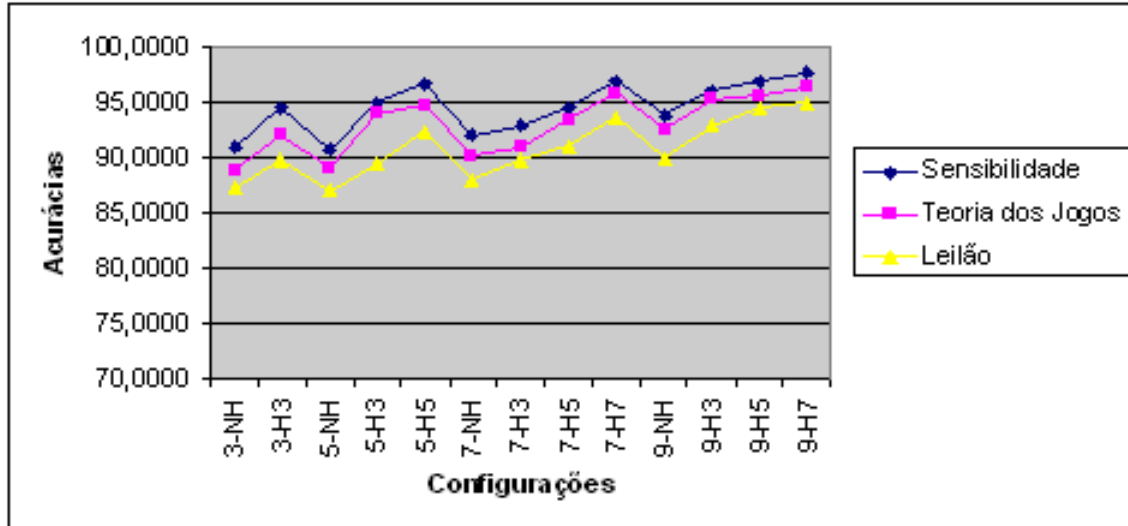


Figura 5.13: Resultados dos Métodos de Negociação para base C (proteínas)

Um detalhe interessante que deve ser observado é que existe uma relação entre a melhora do desempenho, para o método baseado em Sensibilidade, e o número de iterações que o algoritmo necessita para chegar a uma conclusão satisfatória. Por exemplo, o desempenho é menor quando é utilizado um sistema com três agentes, cresce e, então, se estabiliza quando cresce o número de agentes. Isso significa que em determinado momento acontece uma estabilização no desempenho do sistema em relação ao aumento da quantidade de agentes.

Por outro lado, a diferença no número de iterações entre o método baseado em Sensibilidade e os outros dois métodos são menores quando são usados sistemas com três agentes, então eles aumentam, diminuem e tendem a aumentar em relação ao aumento do número de agentes. Isto significa que em determinado momento, quanto maior o número de agentes, maior vai ser o número de iterações e mais demorará o processo de negociação.

Assim como foi observado nos métodos baseados em fusão e em seleção, outro comportamento interessante é que os valores dos resultados sempre aumentam de acordo com o aumento do número de componentes e com o aumento da diversidade do sistema, ou seja, quanto mais tipos de classificadores existir no sistema, maior é a acurácia desse sistema.

Mais uma vez, da mesma forma que foi observado nos métodos baseados em seleção, de forma geral, a base de dados que possui as melhores acurácias é a base B, base de imagens. Por outro lado, a base de dados que possui as menores acurácias é a base C,

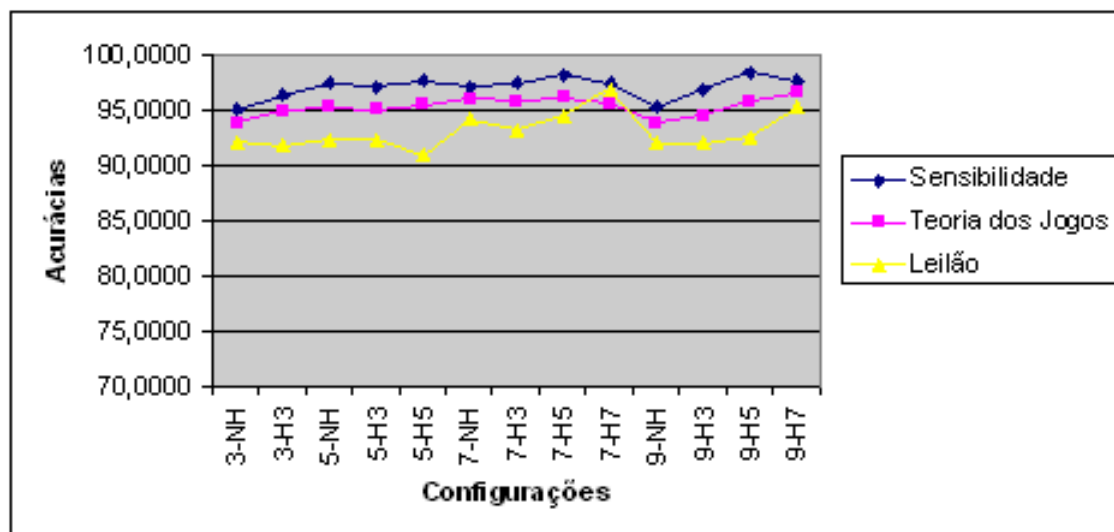


Figura 5.14: Resultados dos Métodos de Negociação para base D (splice)

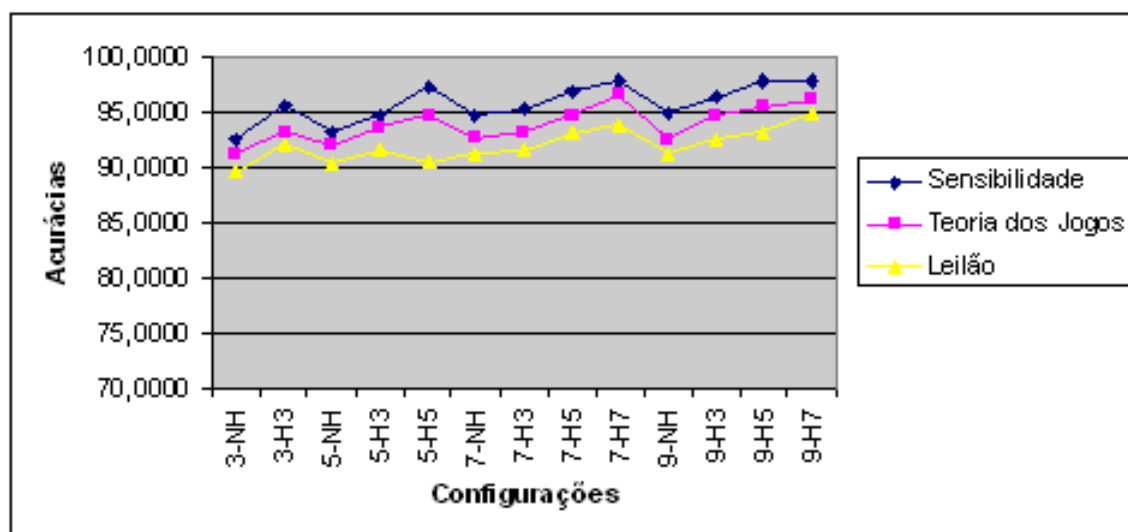


Figura 5.15: Resultados dos Métodos de Negociação para base E (wisconsin)

base de proteínas. As bases que possuem as menores variações de acurácias são as bases B e D. Já as bases que possuem as maiores variações de acurácias são as bases A, C e E.

Da mesma maneira que os métodos baseados em seleção, ainda comparandos com os métodos baseados em fusão, observa-se que os valores mínimos das acurácias dos métodos de negociação são bem superiores que os métodos baseados em fusão. E, comparando agora com os métodos baseados em seleção, observa-se que os métodos de negociação apresentam um desempenho superior, em geral, quando comparados com os métodos baseados em seleção. Esse comportamento é esperado, já que esses métodos de negociação foram propostos como solução do problema de tomada de decisão centralizada.

5.6.6 Sensibilidade nos Sistemas

Como já foi dito anteriormente, a sensibilidade a variações dos tipos de componentes que um sistema apresenta influencia no desempenho deste sistema. Neste trabalho foi observado que quanto maior a quantidade de componentes do sistema, maior esta diferença entre os sistemas híbridos e não híbridos.

Isto pode ser explicado pelo fato de que quanto maior o número de componentes, mais tipos diferentes de componentes podem fazer parte dos sistemas híbridos, logo mais informação será analisada e por conseguinte mais acurada será a classificação em relação a um sistema não híbrido. Esse comportamento pode ser melhor observado no Apêndice A que mostra todos os resultados compilados em um único gráfico para cada uma das bases.

5.6.7 Diversidade nos Sistemas

As três medidas de diversidade utilizadas neste trabalho tiveram o comportamento esperado. As Figuras 5.16, 5.17 e 5.18 mostram os resultados das medidas de diversidade Dupla-Falsa, Entropia e Q Estatístico, respectivamente para cada uma das cinco bases de dados A (breast cancer), B (imagem), C (proteínas), D (splice) e E (wisconsin).

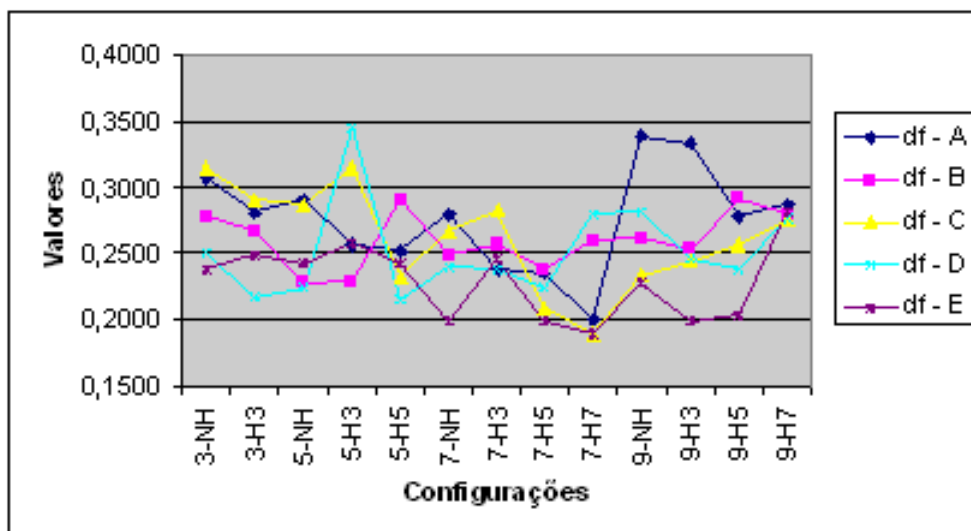


Figura 5.16: Medidas de Diversidade Dupla-Falsa para todas as bases

Através do resultados das medidas de diversidade, observou-se que a diversidade, em geral, aumentava quando aumentava-se a quantidade de componentes dos sistemas, no caso, sistemas com três, cinco, sete e nove componentes. Nas Figuras 5.16, 5.17 e 5.18, observou-se que quando eram utilizados componentes diferentes, ou seja, sistemas *H3*, *H5* e *H7*, as medidas também apontavam um aumento na diversidade dos sistemas.

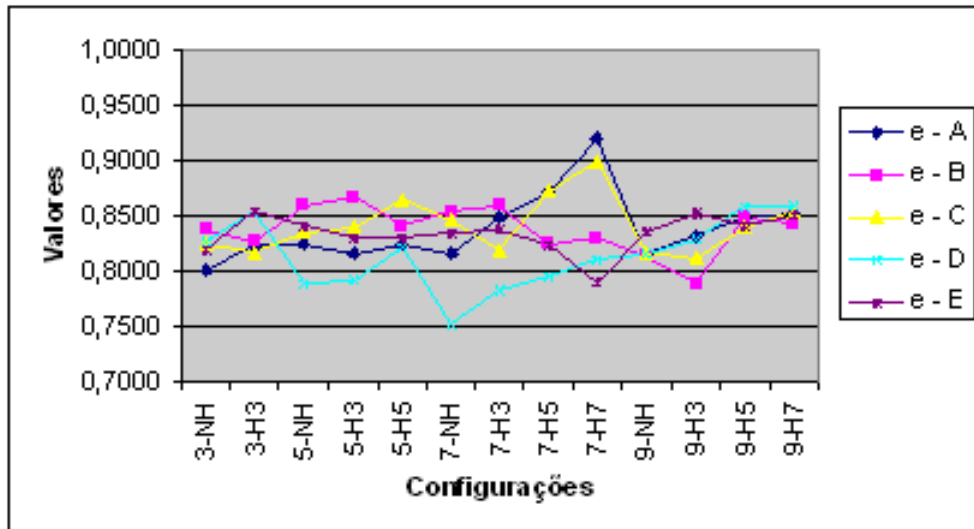


Figura 5.17: Medidas de Diversidade de Entropia para todas as bases

Como já foi dito anteriormente, as acurácias dos sistemas aumentaram da mesma forma que a diversidade, ou seja, quando se aumentou o número de componentes e quando variou-se estes componentes, os resultados sempre melhoraram. É importante observar que o comportamento das medidas não é semelhante. Isso ocorre pelo fato da diversidade variar de acordo com os membros do sistema e, principalmente, de acordo com a base de dados.

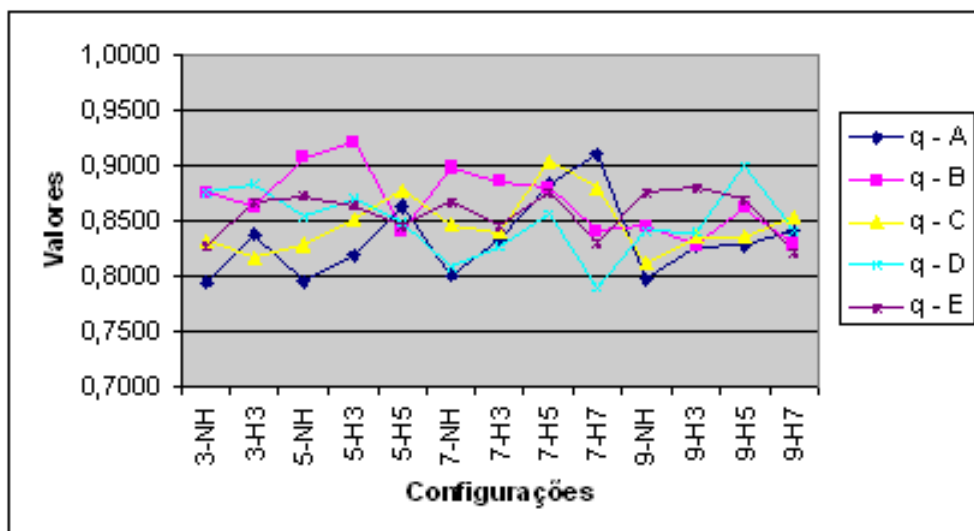


Figura 5.18: Medidas de Diversidade Q Estatístico para todas as bases

Baseado nos experimentos desta dissertação, conclui-se que deve existir uma relação entre o bom desempenho, ou acurácia, dos sistemas e a sua diversidade na escolha dos componentes do sistema. Entretanto, é importante enfatizar que este resultado é dependente dos dados utilizados, ou seja, isto deve ser verdade na utilização de outras bases.

5.6.8 Negociação *versus* Combinação

Quando se compara os melhores métodos baseado em fusão e baseado em seleção com o melhor método de negociação, o método de negociação é sempre melhor em todos os casos, como é explicitamente mostrado neste capítulo. Os comportamentos dos sistemas, em resumo, podem ser observados nas Figuras 5.19, 5.20, 5.21, 5.22 e 5.23.

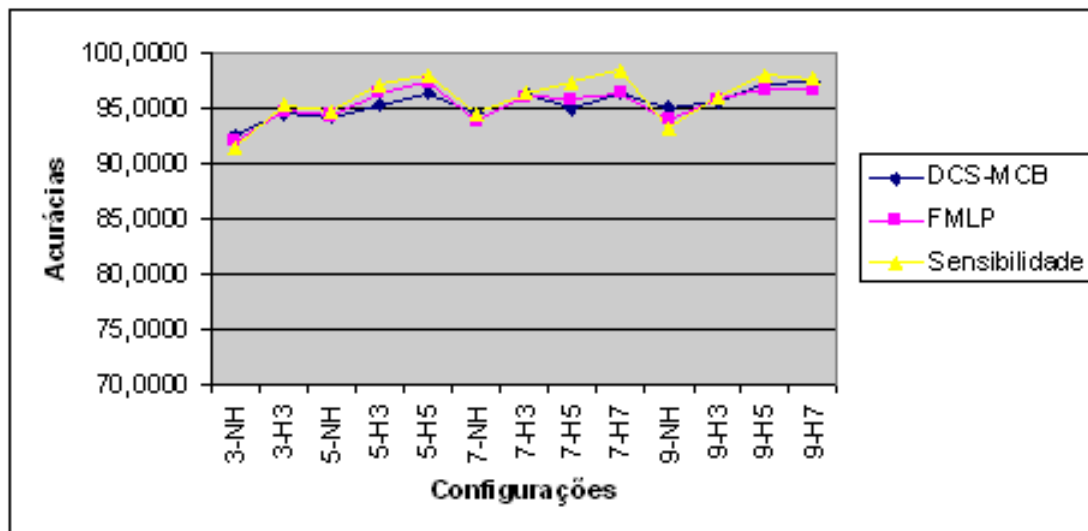


Figura 5.19: Comparação entre os Melhores Méodos de Seleção, Negociação e Fusão para a base A (breast cancer)

O padrão de desempenho dos sistemas de classificação foi confirmado pelo teste estatístico de hipótese. Nesses testes, que podem ser vistos nas Tabelas 5.5, 5.10, 5.15 e 5.20, é mostrado que, na maioria dos casos, ou seja em aproximadamente 70% dos casos, a acurácia do melhor método de negociação é estatisticamente significativa quando comparado com a acurácia dos melhores métodos de negociação.

Neste caso, observou-se que quando o sistema possui três componentes, a acurácia do melhor método de negociação é estatisticamente melhor que os demais em 50% dos casos. Quando o sistema possui cinco componentes essa acurácia é melhor em aproximadamente 80% dos casos e quando o sistema possui sete e nove componentes essa acurácia é melhor em 70% dos casos.

A melhora do desempenho do método de negociação começa baixa, com sistemas contendo três agentes. Então, ela aumenta, decresce e tende a estabilizar de acordo com o número de agentes. Por exemplo, quando usados sistemas com cinco e sete agentes, o método de negociação alcança o maior crescimento da acurácia, quando comparado com os métodos de combinação.

Assim como foi observado nos métodos baseados em fusão e em seleção, outro com-

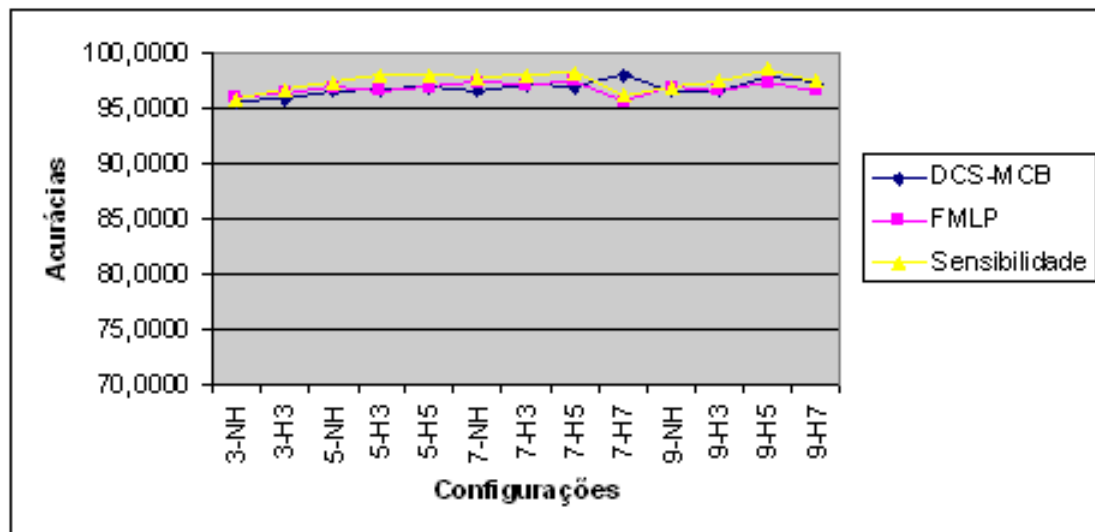


Figura 5.20: Comparação entre os Melhores Mérodo de Seleção, Negociação e Fusão para a base B (imagem)

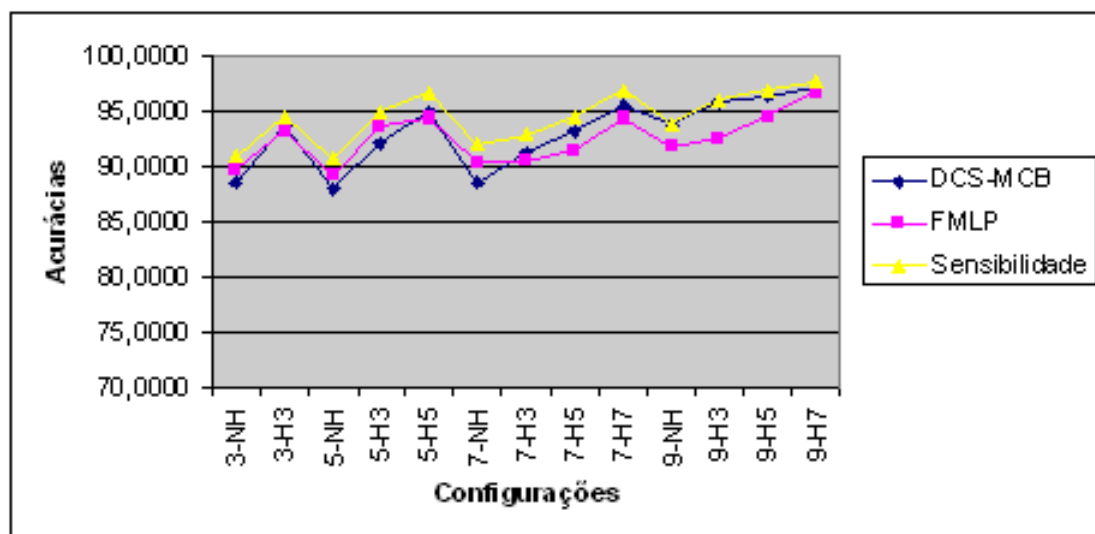


Figura 5.21: Comparação entre os Melhores Mérodo de Seleção, Negociação e Fusão para a base C (proteínas)

portamento interessante é que os valores dos resultados sempre aumentam de acordo com o aumento do número de componentes e com o aumento da diversidade do sistema, ou seja, quanto mais tipos de classificadores existir no sistema, maior é a acurácia desse sistema.

Mais uma vez, da mesma forma que foi observado nos métodos baseados em seleção, de forma geral, a base de dados que possui as melhores acurácias é a base B, base de imagens. Por outro lado, a base de dados que possui as menores acurácias é a base C, base de proteínas. As bases que possuem as menores variações de acurácias são as bases

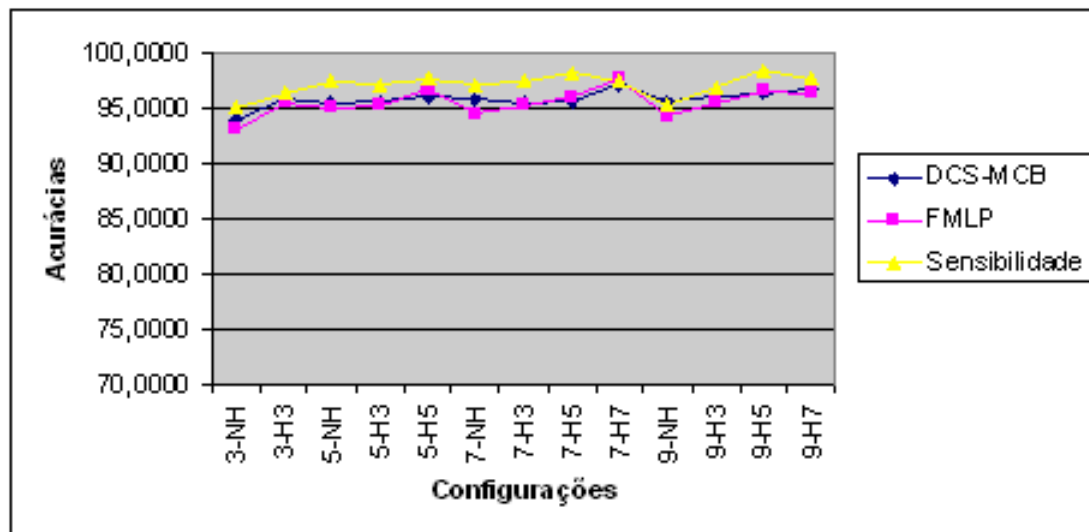


Figura 5.22: Comparação entre os Melhores Mérodo de Seleção, Negociação e Fusão para a base D (splice)

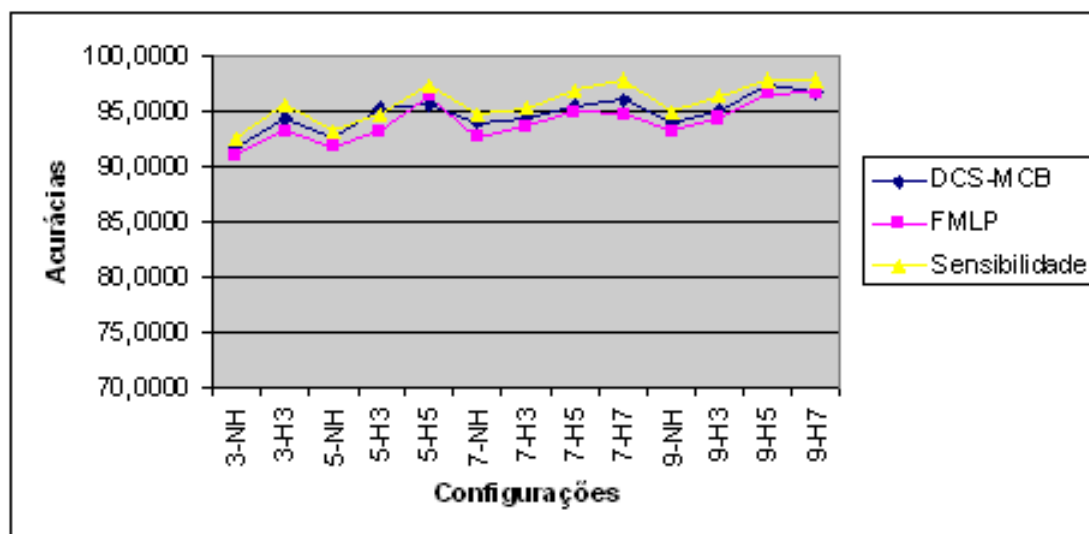


Figura 5.23: Comparação entre os Melhores Mérodo de Seleção, Negociação e Fusão para a base E (wisconsin)

A, B e D. Já as bases que possuem as maiores variações de aurácias são as bases C e E.

É importante enfatizar que o comportamento dos melhores métodos de cada um dos tipos possui uma grande similaridade. Porém, o melhor método de negociação, na maioria das vezes, apresenta resultados melhores que os demais métodos. Esse comportamento pode ser explicado pelo fato deste método, no caso o método baseado na sensibilidade dos agentes, analisar uma maior quantidade de informações antes de realizar a classificação.

Capítulo 6

Considerações Finais

Neste trabalho foi apresentado o Sistema *ClassAge*, um Sistema Multiagente para Classificação que foi baseado no Sistema *NeurAge*. Além disso, os três métodos de negociação para um sistema baseado em agentes que foram utilizados no Sistema *NeurAge* foram apresentados. Eles são o método baseado em Leilão, o método baseado na Teoria dos Jogos e o método baseado na Sensibilidade. A partir disso, um trabalho experimental foi feito usando cinco bases de dados diferentes com tamanhos e estruturas de sistemas diferentes. Foram observados sistemas híbridos e não híbridos com três, cinco, sete e nove componentes.

O desempenho do *ClassAge* foi comparado com o desempenho de classificadores individuais, com sistemas baseados em fusão e sistemas baseados em seleção. Os classificadores utilizados foram o K-nn, o SVM, as redes neurais MLP, FuzzyMLP e RBf, a Árvore de Decisão e o Jrip. Os métodos baseados em fusão utilizados foram o Voto, o Naive Bayes, a Soma, a Média, a Mediana e as redes neurais MLP e FuzzyMLP. E, por fim, os métodos de seleção utilizados foram o DCS-LA, o DCS-DT e o DCS-MCB. Desta maneira, através desta análise, pôde ser observado que as melhores performances foram quase sempre alcançadas pelos métodos de negociação em estruturas híbridas, para todos os quatro tamanhos de sistemas.

Para os métodos baseados em fusão, a média da maior performance foi alcançada por sistemas combinados pelo método FuzzyMLP, enquanto que para métodos baseados em seleção, isto ocorre para o método DCS-MCB. Já para os métodos de negociação, o método baseado em Sensibilidade sempre apresentou a melhor performance, mas isto acontece pelo custo de um maior número de iterações, quando comparado com os demais métodos de negociação.

Por fim, quando compara-se os melhores métodos de fusão, seleção e negociação; o último apresenta sempre a melhor performance em todos os tamanhos, estruturas e tipos

de sistemas. O padrão de performance dos sistemas de classificação foi confirmado por um teste estatístico que mostra que a acurácia do melhor método de negociação é estatisticamente significativa para a maioria dos casos quando comparada com a acurácia dos métodos de combinação (métodos baseados em fusão e seleção). Além disso, quando usa-se sistemas com cinco e sete agentes, o método de negociação alcança o melhor crescimento na acurácia, quando comparado com os métodos de fusão e seleção.

Estes resultados obtidos são muito importantes para as áreas de sistemas multiagentes e reconhecimento de padrões. Isto mostra que o uso de um processo incremental e distribuído de tomada de decisão é bom para a performance de um sistema de classificação.

O processo de tomada de decisão de um método de negociação pode ser considerado como um processo de seleção. Nesta investigação é mostrado que pelo menos um dos métodos de negociação possui uma performance melhor que todos os três métodos baseados em seleção que foram utilizados aqui.

6.1 Trabalhos Futuros

Esta análise que foi feita ainda pode ser bastante expandida. Desta maneira, alguns dos trabalhos futuros possíveis podem ser vistos a seguir.

- Utilizar de outros tipos de bases de dados:
 - Bases mais complexas;
 - Bases desbalanceadas;
 - Bases com alta dimensionalidade de atributos;
 - Bases com poucos atributos;
- Trabalhar com uma maior quantidade de agentes;
- Fazer a comparação do desempenho do *ClassAge* com o desempenho de sistemas combinados com outros tipos de classificadores;
- Utilizar outros tipos de classificadores nos agentes;
- Utilizar outras medidas de diversidade;
- Utilizar outros métodos de negociação:
 - Métodos de negociação fuzzy;
 - Métodos que trabalhem com negociação de vários agentes ao mesmo tempo;

- Colocar o *ClassAge* em alguma plataforma; e
- Desenvolver o *ClassAge* em alguma linguagem de programação baseada no paradigma de agentes.

Essas são algumas das variações que ainda podem ser feitas no Sistema *ClassAge*.

Referências Bibliográficas

- Abreu, M. C. C. (2005). Análise do desempenho de alguns métodos de negociação para agentes inteligentes classificadores. Relatório de Graduação.
- Abreu, M. C. C., Bettencourt, V. G., Canuto, A. M. P., and Souto, M. C. P. (2005a). Aplicação de multiclassificadores heterogêneos no reconhecimento de classes estruturais do proteínas. In *Proceedings of Congresso Brasileiro de Redes Neurais (CBRN'05)*. Universidade Federal do Rio Grande do Norte, Anais do CBRN.
- Abreu, M. C. C. and Canuto, A. M. d. P. (2006). Analyzing the benefits of using a fuzzy-neuro model in the accuracy of the neurage system. In *Proceedings of International Joint Conference on Neural Networks (IJCNN'06)*, volume 1, pages 2959–2966, Vancouver, Canadá.
- Abreu, M. C. C. and Canuto, A. M. P. (2004). Análise e desenvolvimento de um sistema multi-agentes classificador. volume 1, pages 207–215, Natal, RN. Workshop Técnico e Científico do DIMAp (WTCDimap'04).
- Abreu, M. C. C., Canuto, A. M. P., Medeiros, A., Oliveira, L. M., and Xavier Junior, J. C. (2005b). An analysis of performance in hybrid structures of ensembles. In *The 12th International Conference on Neural Information Processing (ICONIP'05)*, volume 1, pages 364–369, Taipei, Taiwan.
- Abreu, M. C. C., Canuto, A. M. P., Medeiros, A., Oliveira, L. M., and Xavier Júnior, J. C. (2005c). Performance and diversity evaluation in hybrid and non-hybrid structures of ensembles. In Society, I. C., editor, *Proceedings of Fifth International Conference on Hybrid Intelligent Systems (HIS'05)*, pages 285–290. Pontifícia Universidade Católica do Rio de Janeiro.
- Abreu, M. C. C., Canuto, A. M. P., Medeiros, A., Souza, F., Gomes Junior, M. F., and Bezerra, V. S. (2004). Investigating the use of an agent-based multi-classifier system for classification tasks. In on Computer Science, L. N., editor, *11^o International Conference on Neural Information Processing (ICONIP'04)*, volume 3316, pages 854–859, Calcutá. Springer Berlin / Heidelberg.

- Abreu, M. C. C., Canuto, A. M. P., and Santana, L. E. A. S. (2005d). A comparative analysis of negotiation methods for a multi-neural. In Society, I. C., editor, *Proceedings of Fifth International Conference on Hybrid Intelligent Systems (HIS'05)*, pages 451–456. Pontifícia Universidade Católica do Rio de Janeiro.
- Abreu, M. C. C., Canuto, A. M. P., and Souto, M. C. P. (2005e). Uma análise comparativa de sistemas de classificação usando uma base de estrutura de proteínas. In *Proceedings of Congresso Brasileiro de Redes Neurais (CBRN'05)*. Universidade Federal do Rio Grande do Norte, Anais do CBRN.
- Abreu, M. C. C., Santana, L. E. A. d. S., and Canuto, A. M. d. P. (2006a). Analyzing the performance of an agent-based neural system for classification tasks using data distribution among the agents. In *Proceedings of International Joint Conference on Neural Networks (IJCNN'06)*, volume 1, pages 2951–2958, Vancouver, Canadá.
- Abreu, M. C. C., Santana, L. E. A. d. S., and Canuto, A. M. d. P. (2006b). A neuro-fuzzy-based agent system with data distribution among the agents for classification tasks. volume 0, page 27, Los Alamitos, CA, USA. IEEE Computer Society.
- Abreu, M. C. C., Souto, M. C. P., Canuto, A. M. P., and Bettencourt, V. G. (2005f). Reconhecimento de classes estruturais de proteínas usando técnicas de aprendizagem de máquina individuais e de multiclassificação. In *Encontro Nacional de Inteligência Artificial (ENIA'05)*, volume 1, pages 1162–1165, São Leopoldo. Anais do XXV Congresso da Sociedade Brasileira de Computação (CSBC'05). Poster.
- Al-Ani, A. and Derich, M. (2002). A new technique for combining multiple classifiers using the dempster-shafer theory of evidence. In *Journal of Artificial Intelligence Research.*, volume 17, pages 333–361, Austrália.
- Anja, F. B. (2003). Can agents close contracts? In *Proceedings of LEA 2003: The Law and Electronic Agents.*, pages 9–20.
- Ayesh, A. and Cowell, J. (2004). Emotional analysis of facial expressions. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 1, pages 672–676.
- Ben-Ameur, H., Chaib-draa, B., and Kropf, P. (2002). Multi-item auctions for automatic negotiation. CIRANO Working Papers 2002s-68, CIRANO. available at <http://ideas.repec.org/p/cir/cirwor/2002s-68.html>.
- Blake, C. and Merz, C. (1998). Uci repository of machine learning databases. <http://www.ics.uci.edu/mlearn/MLRepository.html>.

- Boella, G. and van der Torre, L. (2004). Fulfilling or violating obligations in normative multiagent systems. In *Proceedings. IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'04)*, pages 483–486.
- Brewer, C., Barber, J., Willhauck, G., and Benz, U. (2003). Multi-source and multi-classifier system for regional landcover mapping. In *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pages 143–149.
- Canuto, A. M. P. (2001). *Combining Neural Networks and Fuzzy Logic for Applications in Character Recognition*. PhD thesis, Department of Electronics, University of Kent, Canterbury, UK.
- Chen, D., Sirlantzis, K., Hua, D., and Ma, X. (2005). On the relation between dependence and diversity in multiple classifier systems. In *International Conference on Information Technology: Coding and Computing (ITCC'05)*, volume 1, pages 134–139.
- Cho, S.-B. (1999). Pattern recognition with neural networks combined by genetic algorithm. In *Fuzzy Sets and Systems*, volume 2, pages 339–347. Elsevier Science.
- Cohen, W. W. (1995). Fast effective rule induction. In Friediris, A. and Russell, S., editors, *Proc. of the 12th International Conference on Machine Learning*, pages 115–123, Tahoe City, CA. Morgan Kaufmann.
- Corchado, F. F. R., Chen, L., Bui, M., and Faudemay, P. (1997). Evaluation of dynamic task allocation strategies for distributed information systems. In *Proceedings., Eighth International Workshop on Database and Expert Systems Applications.*, pages 120–125.
- Davis, D. N. (2000). Agents, emergence, emotion and representation. In *26th Annual Conference of the IEEE Industrial Electronics Society*, volume 4, pages 2577–2582.
- Davis, J., Santos Costa, V., Ong, I., Page, D., and Dutra, I. (2004). Using bayesian classifiers to combine rules. <http://www.cs.wisc.edu/~dpage/>.
- De Vos, M. and Vermeir, D. (2001). Logic programming agents and game theory. <http://tinf2.vub.ac.be/~dvermeir/papers/>.
- Didaci, L. and Giacinto, G. (2004). Dynamic classifier selection by adaptive k-nearest-neighbourhood rule. In *Lecture Notes in Computer Science*, volume 3077, pages 174–183.
- Dunne, P. (2004). Extremal behaviour in multiagent contract negotiation. Technical Report 04-001, Dept. of Computer Science, University of Liverpool.

- Dunne, P., Wooldridge, M., and Laurence, M. (2003). The complexity of contract negotiation. Technical Report 03-002, Dept. of Computer Science, University of Liverpool.
- Fairhurst, M., Abdel Wahab, H., and Brittan, P. (1989). Parallel implementation of image classifier architectures using transputer arrays. In *Third International Conference on Image Processing and its Applications*, volume 18-20, pages 136–140.
- Farrugia, J. (2002). Logical systems: Towards protocols for web-based meaning negotiation. In *Proceedings of AAAI 2002 Workshop on Meaning Negotiation*, pages 56–59.
- Ferber, J. and Gasser, L. (1991). Intelligence artificielle distribuée. In *International Workshop on Expert Systems and their Applications.*, volume 10, Avignon. Cours n. 9. France.
- Fletcher, M. (2006). A cognitive agent-based approach to varying behaviours in computer generated forces systems to model scenarios like coalitions. In *IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications*, pages 297–284.
- Giacinto, G. and Roli, F. (1999). Methods for dynamic classifier selection. In *Proceedings of the International Conference on Image Analysis and Processing*, volume 1, pages 659–664.
- Giacinto, G. and Roli, F. (2000). Dynamic classifier selection based on multiple classifier behaviour. In *The journal of The Pattern Recognition Society.*, volume 34, pages 1879–1881.
- Giacinto, G. and Roli, F. (2001). Design of effective neural network ensembles for image classification. In *Image and Vision Computing Journal*, volume 9-10, pages 686–693.
- Heutte, L., Nosary, A., and Paquet, T. (2004). A multiple agent architecture for handwritten text recognition. In *Pattern Recognition*, volume 37, pages 665–674.
- Ho, T. K., Hull, J., and Srihari, S. (1994). Decision combination in multiple classifier systems. In *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, volume 16, pages 66–75.
- Hull, J. J. (1994). A database for handwritten text recognition. In *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, volume 16, pages 550–554.
- Jiangsheng, Y. (2002). Method os k-nearest neighbors.
- Julia, L. E., Heck, L. P., and Cheyer, A. J. (1997). A speaker identification agent. In *Proceedings of the AVBPA'97.*, Crans Montana, Switzerland.

- Kang, H.-J. and Lee, S.-W. (2000). An information-theoretic strategy for constructing multiple classifier systems. In *Proceedings. 15th International Conference on Pattern Recognition*, volume 2, pages 483–486.
- Kawamura, K. (2002). The role of cognitive agent models in a multi-agent framework for human-humanoid interaction. In *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication*, pages 81–86.
- Kittler, J. and Alkoot, F. M. (2003). Sum versus vote fusion in multiple classifier systems. In *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, volume 25, pages 110–115.
- Kuncheva, L. I. (2002a). Switching between selection and fusion in combining classifiers: An experiment. In *IEEE Transactions on Systems, Man and Cybernetics.*, volume 32, pages 146–156.
- Kuncheva, L. I. (2002b). A theoretical study on six classifier fusion strategies. In *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, volume 24, pages 281–286.
- Kuncheva, L. I. (2004). Combining pattern classifier. In *Methods and Algorithms*. Wiley.
- Kuncheva, L. I. and Whitaker, C. J. (2001). Ten measures of diversity in classifier ensembles: limits for two classifiers. In *A DERA/IEE Workshop on Intelligent Sensor Processing*, pages 10/1–1010.
- Lawley, R., Decker, K., Luck, M., Payne, T., and Moreau, L. (2003). Automated negotiation for grid notification services. In Springer-Verlag, editor, *Ninth International Europar Conference (EURO-PAR'03)*, volume 2790, pages 384–393, Klagenfurt, Austria. Lecture Notes in Computer Science.
- Lemieux, A. and Parizeau, M. (2003). Flexible multi-classifier architecture for face recognition systems. In *Vision Interface*.
- Leong, P. and Chunyan, M. (2005). Fuzzy cognitive agents in shared virtual worlds. In *International Conference on Cyberworlds*, pages 23–25.
- Levine, D. M. and Berenson, M. L. (2000). *Estatística: teoria e aplicações usando o Microsoft excel*.
- Li, J., Barrett, N., Zhang, Y.-Q., and Washburn, D. (2005a). Genetic granular cognitive fuzzy neural networks and human brains for pattern recognition. In *IEEE International Conference on Granular Computing*, volume 1, pages 172–175.

- Li, Y., Chao, K.-M., Younas, M., Huang, Y., and Xiaohua, L. (2005b). Modeling e-marketplaces with multi-agents web services. In *Proceedings of the 11th International Conference on Parallel and Distributed Systems.*, volume 1, pages 175–181.
- Mangasarian, O. L. and Wolberg, W. H. (1992). Repository of machine learning databases.
- Mathieu, P., Beaufls, B., and Brandouy, O. (2006). *Agent-Based Methods in Finance, Game Theory and Their Applications.*, volume 564.
- Meir, R. (2002). A tutorial on support vector machines for pattern recognition.
- Miao, C., Yang, Q., Fang, H., and Goh, A. (2002). Fuzzy cognitive agents for personalized recommendation. In *Proceedings of the Third International Conference on Web Information Systems Engineering.*, pages 362–371.
- Mitchell, T. (1997). *Machine Learning*. McGraw Hill.
- Morency, L.-P., Sidner, C., and Darrell, T. (2005). Towards context-based visual feedback recognition for embodied agents. In *Proceedings of the Symposium on Conversational Informatics for Supporting Social Intelligence and Interaction (AISB'05).*, pages 69–72.
- Mukhopadhyay, S., Peng, S., Raje, R., Palakal, M., and Mostafa, J. (2003). Multi-agent information classification using dynamic acquaintance lists. In *Journal of the American Society for Information Science and Technology*, volume 54, pages 966–976.
- Neves, M. C. and Oliveira, E. (1998). A connectionist approach to the acquisition of elementary navigation competencies by a mobile robot. In *5th International Workshop on Advanced Motion Control (AMC'98).*, pages 530–535.
- Ogata, N. (2002). Dynamic semantics of cognitive agent language. In *Proceedings. IEEE Workshop on Knowledge Media Networking.*, pages 171–176.
- Osborne, M. (2003). *An Introduction to Game Theory*. Press.
- Panwai, S. and Dia, H. (2005). A reactive agent-based neural network car following model. In *Proceedings. 2005 IEEE Intelligent Transportation Systems.*, pages 375–380.
- Park, J.-W., Harley, R., and Venayagamoorthy, G. (2002). Comparison of mlp and rbf neural networks using deviation signals for on-line identification of a synchronous generator. In *Proceedings of International Joint Conference on Neural Networks (IJCNN'02)*, volume 1, pages 919–924.

- Peng, S., Mukhopadhyay, S., Raje, R. R., Palakal, M., and Mostafa, J. (2001). A comparison of single-agent and multi-agent architectures for the classification of handwritten text. In *Proceedings of 10th IEEE Heterogeneous Computing Workshop (HCW'01)*, page 10, San Francisco, California.
- Petrushin, V. A. (2000). Emotion recognition agents in real world. In *AAAI Fall Symposium on Socially Intelligent Agents: Human in the Loop*.
- Popovici, D. M., Buche, C., Querrec, R., and Harrouet, F. (2004). An interactive agent-based learning environment for children. In *International Conference on Cyberworlds.*, pages 233–240.
- Popovici, D. M., Querrec, R., Harrouet, F., Le Gal, C., Serbanati, L. D., and Morvan, S. (2005). Virtualdive - a vr-based educational virtual environment. In *Seventh International Symbolic and Numeric Algorithms for Scientific Computing (SYNASC'05).*, pages 0–8.
- Prudent, Y. and Ennaji, A. (2005). A topology based multi-classifier system. In *Proceedings Eighth International Conference on Document Analysis and Recognition*, volume 2, pages 670–674.
- Quinlan, J. R. (1993). C4.5: programs for machine learning.
- Raje, R. R., Mukhopadhyay, S., Boyles, M., Papiez, A., Patel, N., Palakal, M., and Mostafa, J. (1998). A bidding mechanism for web-based agents involved in information classification. In *World Wide Web*, volume 1, pages 155–165, Hingham, MA, USA. Kluwer Academic Publishers.
- Rezende, S. O. (2002). *Sistemas Inteligentes*, volume 1. Capítulo 10.
- Roli, F. and Giacinto, G. (2002). *Design of Multiple Classifier Systems*. Hybrid Methods in Pattern Recognition.
- Schomaker, L., Weert, C. d., and Erp, M. v. (1998). Heterogeneous information integration: theory and application of the multiple-agent paradigm in pattern recognition and user-interfacing. <http://www.onderzoekinformatie.nl/nl/oi/nod/onderzoek/OND1267138/>. Projeto de Pesquisa.
- Shou-Jue, W. and Xu, C. (2003). Biomimetic (topological) pattern recognition - a new model of pattern recognition theory and its application. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN'03)*, volume 3, pages 2258–2262.

- Sidorova, J. A. (1999). A combined approach of syntatic pattern recognition and machine learning to build a sensible to speech emotion agent.
- Sirlantzis, K., Fairhurst, M., and Guest, R. (2002a). An evolutionary algorithm for classifier and combination rule selection in multiple classifier systems. In *Proceedings 16th International Conference on Pattern Recognition*, volume 2, pages 771–774.
- Sirlantzis, K. and Fairhurst, M. C. (2001). Optimisation of multiple classifier systems using genetic algorithms. In *Proceedings. 2001 International Conference on Image Processing.*, volume 1, pages 1094–1097.
- Sirlantzis, K., Fairhurst, M. C., and Guest, R. M. (2002b). An evolutionary algorithm for classifier and combination rule selection in multiple classifier systems. In *Proceedings. 16th International Conference on Pattern Recognition.*, volume 2, pages 771–774.
- Smits, P. C. (2002). Multiple classifier systems for supervised remote sensing image classification based on dynamic classifier selection. In *IEEE Transactions on Geoscience and Remote Sensing.*, volume 40, pages 801–813.
- Stefano, C., Della Cioppa, A., and Marcelli, A. (2002). An adaptive weighted majority vote rule for combining multiple classifiers. In *16th International Conference on Pattern Recognition (ICPR'02).*, volume 2, page 20192.
- Street, W. N., Mangasarian, O. L., and Wolberg, W. H. (1995). Repository of machine learning databases.
- Tan, A. C., Gilbert, D., and Deville, Y. (2003). Multi-class protein fold classification using a new ensemble machine learning approach. In *Genome Informatics*, volume 14, pages 206–217. Wiley.
- Toda, Y., Yamashita, M., and Sawamura, H. (2001). An argument-based agent system with kqml as an agent communication language. In *PRIMA 2001: Proceedings of the 4th Pacific Rim International Workshop on Multi-Agents, Intelligent Agents*, pages 48–62, London, UK. Springer-Verlag.
- Tomohara, A. (2002). Externalities of non-cooperative tax policy in the globally integrated market. www.econ.jhu.edu/pdf/papers/WP476.pdf.
- Towell, G. and Shavlik, J. (1992). Interpretation of artificial neural networks: Mapping knowledge-based neural networks into rules. In *Advances in Neural Information Processing Systems*, volume 4. Morgan Kaufmann.

- van Otterlo, M., Wiering, M., Dastani, M., and Meyer, J.-J. (2003). A characterization of sapient agents. In *International Conference on Integration of Knowledge Intensive Multi-Agent Systems*, pages 172–177.
- Vuurpijl, L. and Schomaker, L. (1998a). A framework for using multiple classifiers in a multiple-agent architecture.
- Vuurpijl, L. and Schomaker, L. (1998b). Multiple-agent architectures for the classification of handwritten text.
- Waller, O. and Dodds, G. (1999). Reactive agent based planning for an avatar. In *Proceedings. 1999 IEEE/ASME International Conference on Advanced Intelligent Mechatronics.*, pages 713–718.
- Witten, L. H. and Frank, E. (2000). *Data Mining: practical machine learning tools and techniques with Java implementation*. USA.
- Wooldridge, M. (2002). *An Introduction to Multiagent Systems*. JOHN WILEY and SONS, LTD, University of Liverpool, UK.
- Xiang, C., Ding, S. Q., and Lee, T. H. (2001). Geometrical interpretation and architecture selection of mlp. In *IEEE Transactions on Neural Networks*, volume 16, pages 84–86.
- Yang, L.-Y., Qin, Z., and Huang, R. (2004). Design of a multiple classifier system. In *Proceedings of 2004 International Conference on Machine Learning and Cybernetics*, volume 5, pages 3272–3276.
- Zhou, L. (2002). Face recognition by combining several algorithms. In *Proceedings of the 16th International Conference on Pattern Recognition*, volume 3, pages 497–500, Beijing - China.

Apêndice A

Nas Figuras A.1, A.2, A.3, A.4 e A.5 pode ser observado o desempenho de todos os métodos baseados em fusão, que são MLP, FMLP, Voto, Soma, Naive Bayes, Média e Mediana. Além desses métodos observa-se ainda o desempenho dos métodos baseados em seleção, que são o DCS-LA, o DCS-DT e o DCS-MCB. E, por fim também são observados os desempenhos dos métodos de negociação, que são o método baseado na Sensibilidade, baseado na Teoria dos Jogos e baseado no Leilão.

Nessas figuras, a nomenclatura utilizada foi feita para identificar a qual configuração de sistema cada resultado está relacionado. Nessa nomenclatura, o primeiro número diz respeito a quantidade de componentes do sistema. Por exemplo, os modelos que começam com o número 3 estão associadas com os sistemas com 3 componentes.

Já as letras ou números que estão dispostas depois esse primeiro número dizem respeito a quantidade de componentes diferentes o sistema em questão possui. Por exemplo, a configuração 3-NH diz respeito aos sistema não híbridos ou não homogêneos com 3 componentes.

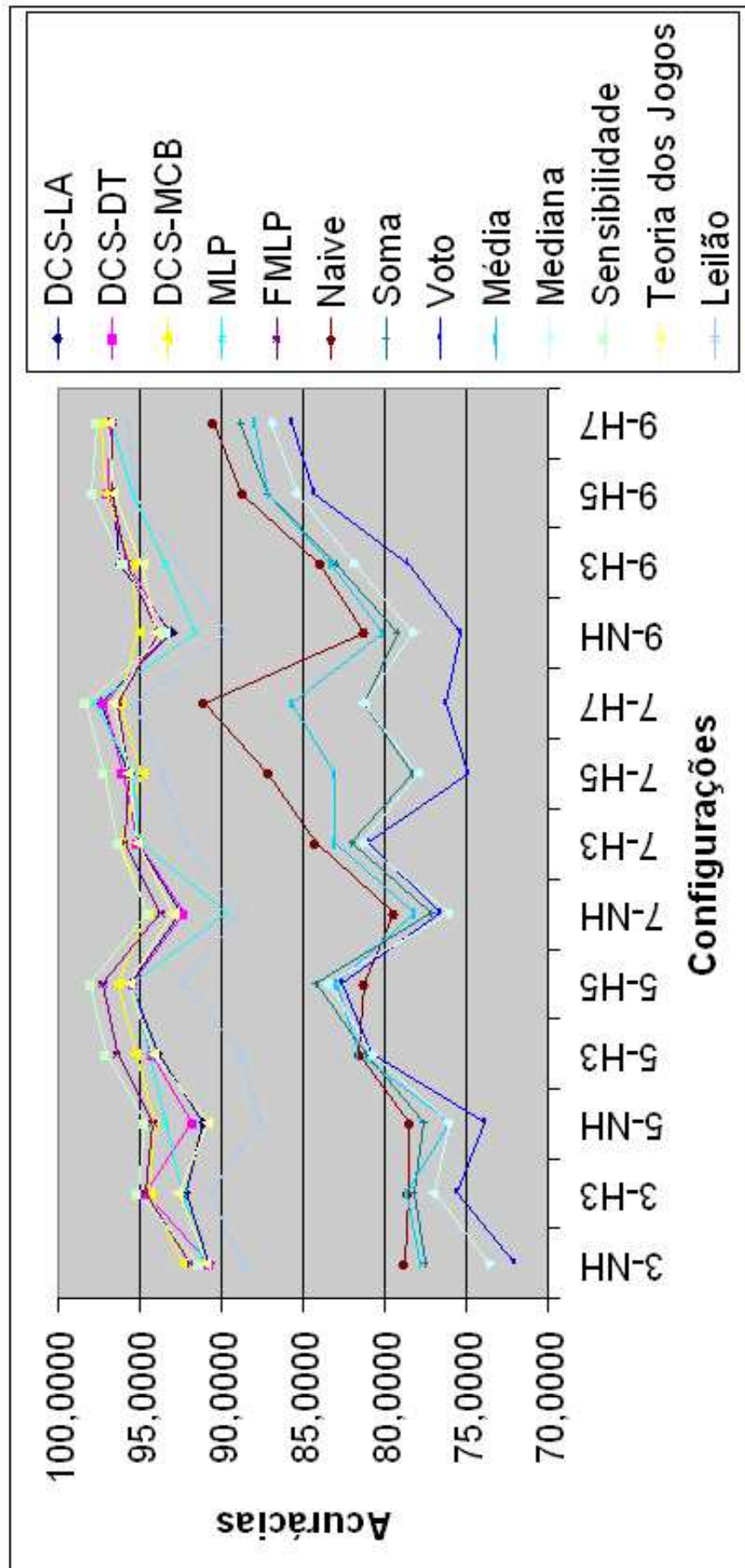


Figura A.1: Resumo dos Resultados para a base A (breast cancer)

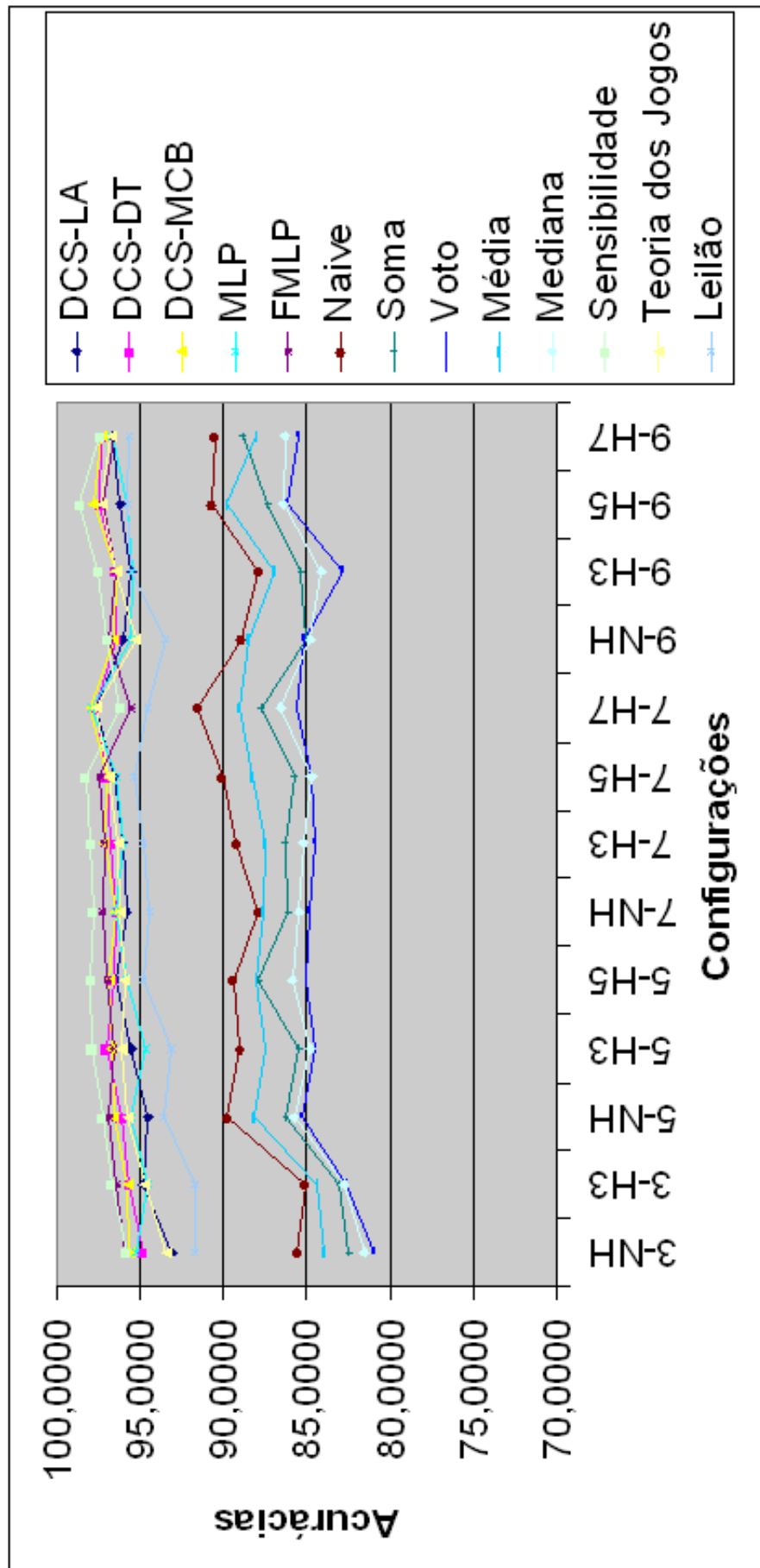


Figura A.2: Resumo dos Resultados para a base B (imagem)

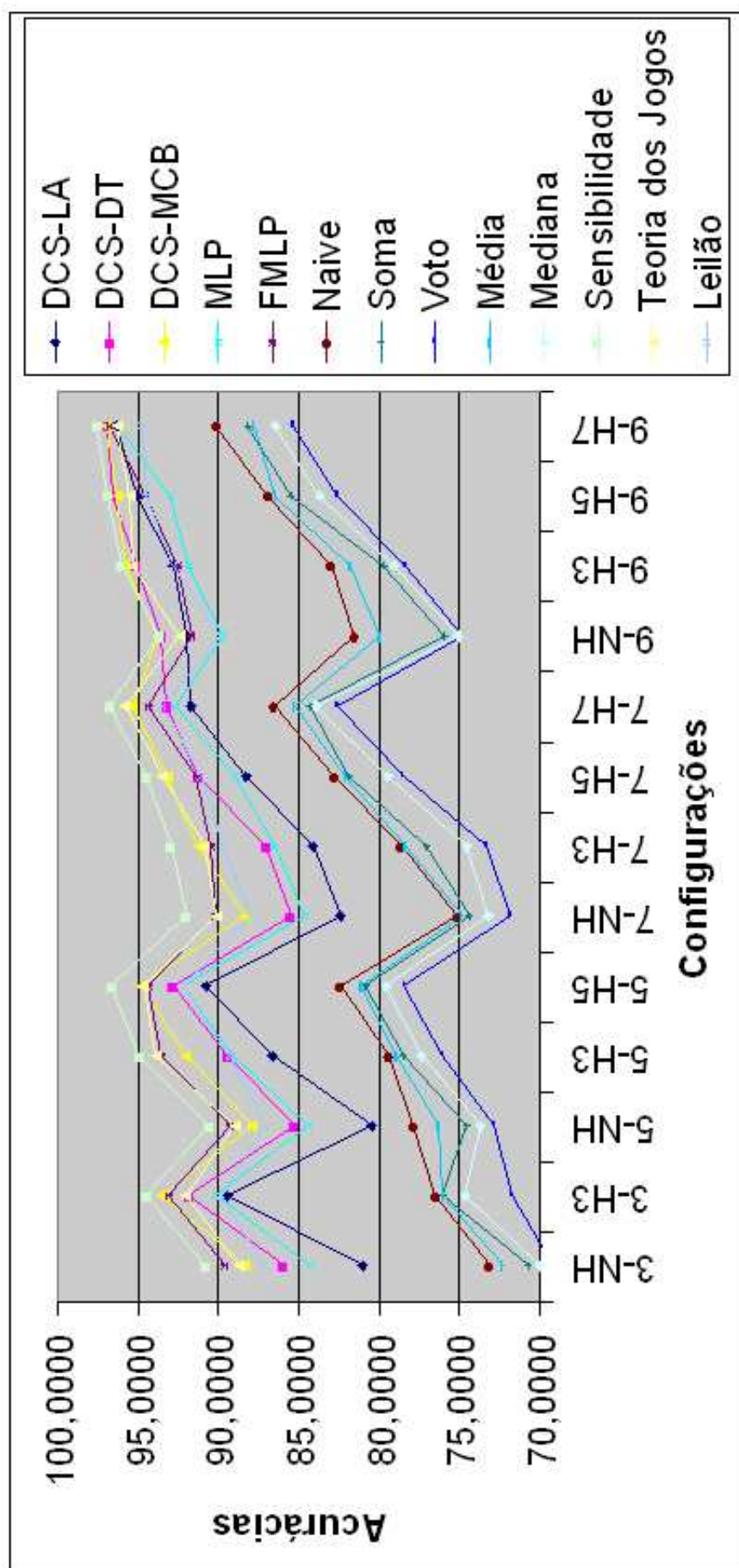


Figura A.3: Resumo dos Resultados para a base C (proteínas)

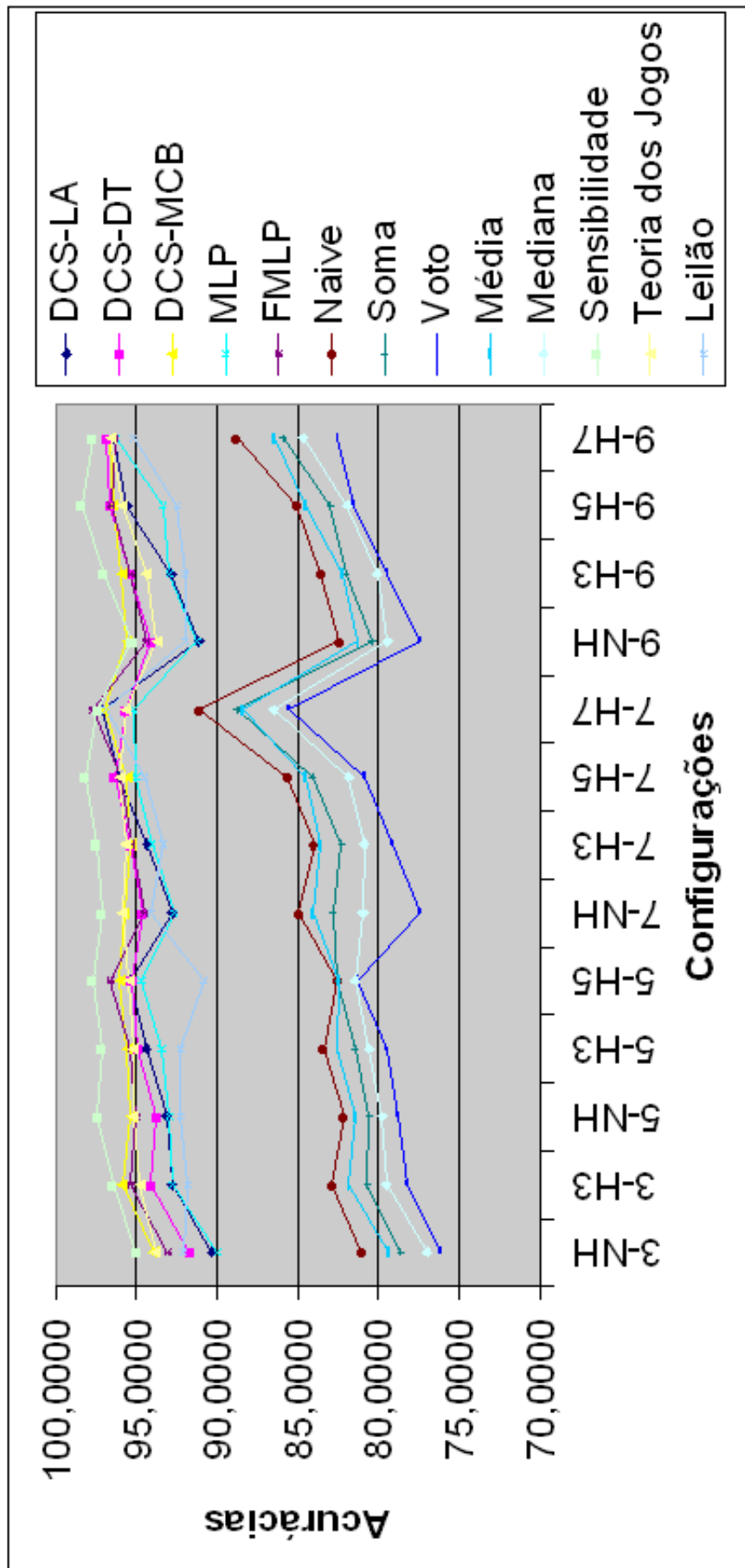


Figura A.4: Resumo dos Resultados para a base D (splice)

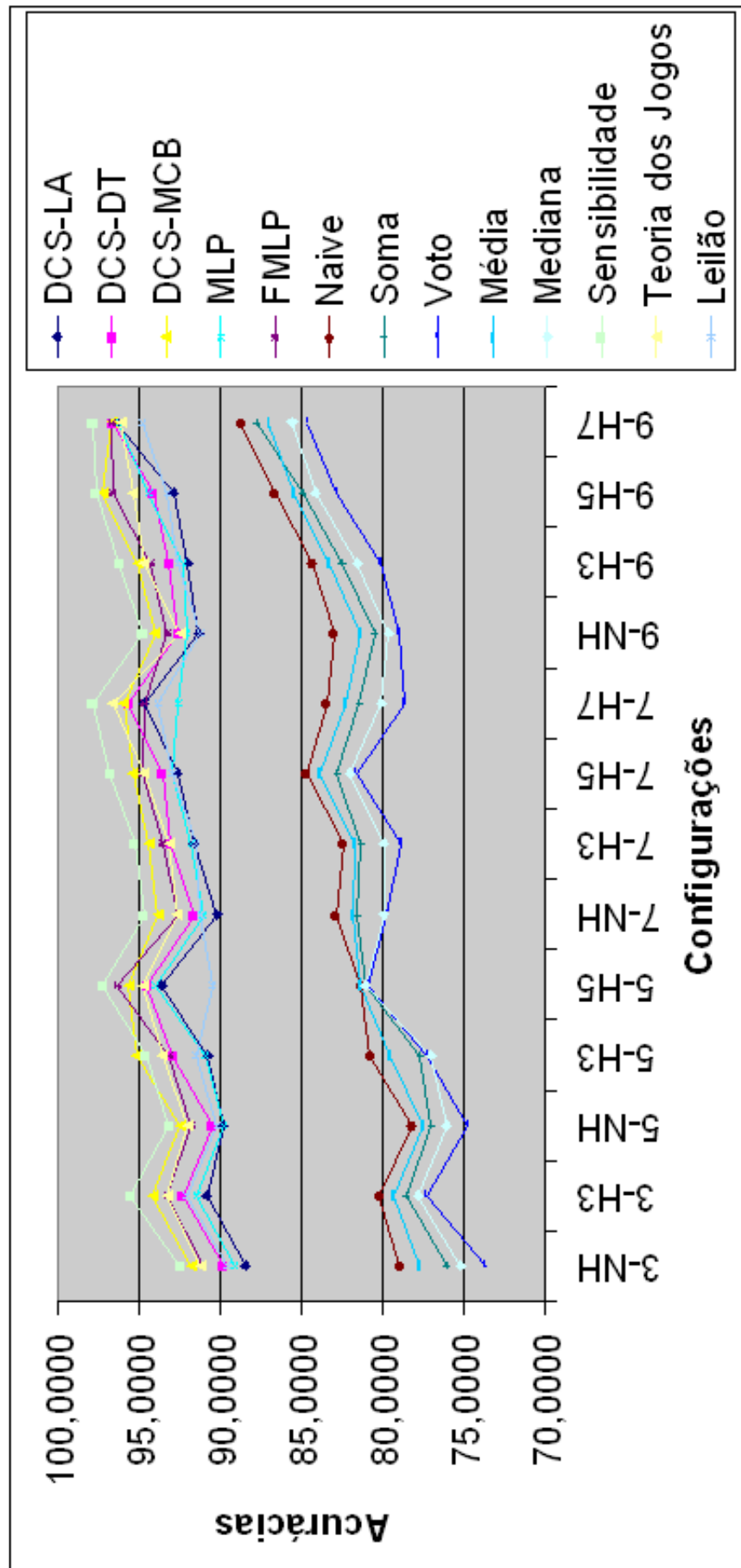


Figura A.5: Resumo dos Resultados para a base E (wisconsin)