
Seleção e controle do viés de aprendizado ativo

Davi Pereira dos Santos

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Davi Pereira dos Santos

Seleção e controle do viés de aprendizado ativo

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências – Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e Matemática Computacional

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

USP – São Carlos
Janeiro de 2017

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

S237s Santos, Davi Pereira dos
Seleção e controle do viés de aprendizado ativo /
Davi Pereira dos Santos; orientador André Carlos Ponce
de Leon Ferreira de Carvalho. -- São Carlos -- SP,
2017.
183 p.

Tese (Doutorado - Programa de Pós-Graduação em
Ciências de Computação e Matemática Computacional) --
Instituto de Ciências Matemáticas e de Computação,
Universidade de São Paulo, 2017.

1. Aprendizado de máquina. 2. Aprendizado ativo.
3. Meta-aprendizado. I. Carvalho, André Carlos Ponce
de Leon Ferreira de, orient. II. Título.

Davi Pereira dos Santos

Selection and control of the active learning bias

Doctoral dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of the Doctorate Program in Computer Science and Computational Mathematics. *FINAL VERSION*

Concentration Area: Computer Science and Computational Mathematics

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

USP – São Carlos
January 2017

*Dedicado à diversidade da vida, ao sorriso, às dimensões subatômicas e cósmicas, à
tristeza, à compaixão... às perguntas sem resposta.*

Agradecimentos

Algumas virtudes surgem sem anúncio. Imagino que a experiência quádrupla da paternidade/maternidade tenha surpreendido meus pais em muitos aspectos. Hoje compreendo os momentos possivelmente mais difíceis, aqueles em que a paciência e dedicação é resultado de puro esforço - quando é preciso manter o mundo girando, com ou sem instinto, romantismo ou qualquer outro impulso adicional. Expresso aqui meu agradecimento a Julio e Eliana, e meus irmãos. Similarmente, a orientação acadêmica também traz seus ensinamentos para a vida. Desde o prof. João da época do mestrado e seus sábios conselhos existenciais disponíveis até hoje, até o prof. André, exemplo de ânimo e humildade. Agradeço a ele pela sua permanente disposição em ajudar e sua paciência ao longo de cinco anos. Agradeço também às contribuições da banca, cujo interesse e dedicação na apreciação do trabalho ajudaram a dar significado a este momento.

Este trabalho mal teria um começo sem a ajuda de Luís P. Garcia e certamente teria sido mais difícil sem o apoio ou material de Angela Giampetro, Lhaís V., T. Covões, R. Mantovani, R. Cerri, P. Pisani, P. Jaskowiak, L. Colleta, D. Horta e F. Zuher. Outras pessoas, como R. Sadao, R. Rigolin, Rafaela D., Uziel, Ligia T., Danira e Fernando, Virgínia V., Camila T., Elaine K., Pâmela C., Vital F., T. Galante, M. Manzato, Herr Marcos V., Nayara, Renan e Larissa Campos merecem meu carinho por terem sido minha família sancarlense por tantos anos. Por fim, escrevo em memória das gigantes esverdeadas que me acompanharam desde o ingresso na universidade, em São Paulo, São Carlos e Araraquara, mas acabaram convertidas em serragem.

Pesquisa desenvolvida com apoio nºDS-3136101/D da CAPES¹ no período de 01/04/2011 a 28/02/2015 e utilização dos recursos computacionais do CeMEAI² financiados pela ag. FAPESP³. *I would like to thank J. Hammersley for sponsoring the hosting of my thesis at* <https://www.overleaf.com>.

Este texto foi escrito com ferramentas disponibilizadas pela comunidade de *software* livre: L^AT_EX, github⁴/git 2.1.4, Debian 8.2, Kile 2.1.3, Okular 4.14.2, PGFPlots 1.12 e outras a elas vinculadas.

¹ Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

² Centro de Ciências Matemáticas Aplicadas à Indústria

³ Fundo de Amparo à Pesquisa do Estado de São Paulo

⁴ <https://github.com>

*“We are not to tell nature what she’s gotta be...
She’s always got better imagination than we have.”
Richard Feynman (lecturing)*

RESUMO

SANTOS, D. P.. **Seleção e controle do viés de aprendizado ativo**. 2017. 183 f. Tese (Doutorado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

A área de aprendizado de máquina passa por uma grande expansão em seu universo de aplicações. Algoritmos de indução de modelos preditivos têm sido responsáveis pela realização de tarefas que eram inviáveis ou consideradas exclusividade do campo de ação humano até recentemente. Contudo, ainda é necessária a supervisão humana durante a construção de conjuntos de treinamento, como é o caso da tarefa de classificação. Tal construção se dá por meio da rotulação manual de cada exemplo, atribuindo a ele pelo menos uma classe. Esse processo, por ser manual, pode ter um custo elevado se for necessário muitas vezes.

Uma técnica sob investigação corrente, capaz de mitigar custos de rotulação, é o aprendizado ativo. Dado um orçamento limitado, o objetivo de uma estratégia de amostragem ativa é direcionar o esforço de treinamento para os exemplos essenciais. Existem diversas abordagens efetivas de selecionar ativamente os exemplos mais importantes para consulta ao supervisor. Entretanto, não é possível, sem incorrer em custos adicionais, testá-las de antemão quanto à sua efetividade numa dada aplicação. Ainda mais crítica é a necessidade de que seja escolhido um algoritmo de aprendizado para integrar a estratégia de aprendizado ativo antes que se disponha de um conjunto de treinamento completo.

Para lidar com esses desafios, esta tese apresenta como principais contribuições: uma estratégia baseada na inibição do algoritmo de aprendizado nos momentos menos propícios ao seu funcionamento; e, a experimentação da seleção de algoritmos de aprendizado, estratégias ativas de consulta ou pares estratégia-algoritmo baseada em meta-aprendizado, visando a experimentação de formas de escolha antes e durante o processo de rotulação. A estratégia de amostragem proposta é demonstrada competitiva empiricamente. Adicionalmente, experimentos iniciais com meta-aprendizado indicam a possibilidade de sua aplicação em aprendizado ativo, embora tenha sido identificado que investigações mais extensivas e aprofundadas sejam necessárias para apurar sua real efetividade prática.

Importantes contribuições metodológicas são descritas neste documento, incluindo uma análise frequentemente negligenciada pela literatura da área: o risco devido à variabilidade dos algoritmos. Por fim, são propostas as curvas e faixas de ranqueamento, capazes de sumarizar, num único gráfico, experimentos de uma grande coleção de conjuntos de dados.

Palavras-chave: Aprendizado de máquina, Aprendizado ativo, Meta-aprendizado.

ABSTRACT

SANTOS, D. P.. **Seleção e controle do viés de aprendizado ativo**. 2017. 183 f. Tese (Doutorado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

The machine learning area undergoes a major expansion in its universe of applications. Algorithms for the induction of predictive models have made it possible to carry out tasks that were once considered unfeasible or restricted to be solved by humans. However, human supervision is still needed to build training sets, for instance, in the classification task. Such building is usually performed by manual labeling of each instance, providing it, at least, one class. This process has a high cost due to its manual nature.

A current technique under research, able to mitigate labeling costs, is called active learning. The goal of an active learning strategy is to manage the training effort to focus on the most relevant instances, within a budget. Several effective sampling approaches having been proposed. However, when one needs to choose the proper strategy for a given problem, they are impossible to test beforehand without incurring into additional costs. Even more critical is the need to choose a learning algorithm to integrate the active learning strategy before the existence of a complete training set.

This thesis presents two major contributions to cope with such challenges: a strategy based on the learning algorithm inhibition when it is prone to inaccurate predictions; and, an attempt to automatically select the learning algorithms, active querying strategies or pairs strategy-algorithm, based on meta-learning. This attempt tries to verify the feasibility of such kind of decision making before and during the learning process.

The proposed sampling approach is empirically shown to be competitive. Additionally, meta-learning experiments show that it can be applied to active learning, although more a extensive investigation is still needed to assess its real practical effectivity.

Important methodological contributions are made in this document, including an often neglected analysis in the literature of active learning: the risk due to the algorithms variability. A major methodological contribution, called ranking curves, is presented.

Key-words: Machine learning, Active learning, Meta learning.

Lista de ilustrações

1	Esquema de aprendizado supervisionado.	30
2	Esquema de aprendizado ativo.	31
3	Frequência de uso de estratégias na competição de aprendizado ativo.	32
4	Curvas das funções-critério para os exemplos selecionados por TU.	58
5	Curvas das funções-critério para os exemplos selecionados por ATU.	58
6	Curvas das funções-critério para os exemplos selecionados por HTU.	59
7	Colocação média para diferentes limites de correlação.	60
8	Esquema do sistema de recomendação.	66
9	Correlação entre ranqueamentos de algoritmos de aprendizado.	76
10	Valores de correlação entre conjuntos de dados mais correlacionados.	77
11	Média da distância de Hamming entre predições de modelos induzidos	84
12	Curvas de aprendizado para o conjunto <i>abalone-3class</i>	91
13	Curvas de aprendizado médias de EERent, Mar, Rnd e TUEuc.	92
14	Curvas de ranqueamento de EERent, Mar, Rnd e TUEuc.	92
15	Exemplo de curvas de ranqueamento com faixas.	93
16	Curvas de ranqueamento (μ_K).	96
17	Curvas representando a variabilidade das estratégias.	100
18	Curvas de ranqueamento (σ_K).	100
19	Média de duração das consultas.	101
20	Possíveis nichos de inadequação conjunto-estratégia-algoritmo.	103
21	Curvas de ranqueamento (μ_K).	104
22	Curvas de ranqueamento por algoritmo de aprendizado (μ_K).	105
23	Possíveis nichos de adequação conjunto-estratégia-algoritmo.	106
24	Curvas de ranqueamento com faixas - incluindo meta-aprendiz.	107
25	Comparação de valores de ALC- μ_K obtidos com meta-aprendizes.	108
26	Comparação da correlação na predição de ranqueamento de algoritmos.	110
27	Comparação da acurácia na predição do melhor algoritmo de aprendizado.	111
28	Comparação de K na recomendação do melhor algoritmo de aprendizado.	112
29	Comparação da acurácia balanceada na recomendação do melhor algoritmo de aprendizado.	113

30	Árvore com os meta-atributos mais relevantes.	115
31	Comparação da correlação na predição de ranqueamento de estratégias.	116
32	Comparação de κ na recomendação da melhor estratégia.	117
33	Comparação da correlação, predição: ranqueamento estratégia-algoritmo.	118
34	Comparação de κ na recomendação do melhor par estratégia-algoritmo.	120
35	Desempenho na predição de ranqueamento de medidas de distância.	121
36	Comparação de κ na predição da melhor medida de distância.	122
37	Comparação de κ na recomendação de uso de aprendizado ativo.	124
38	Curvas de ranqueamento com meta-aprendiz e recomendação perfeita.	134
39	Curvas de ranqueamento com faixas RFW, RoF e metaPCT.	150
40	Comparação de valores de $ALC-\mu_{\kappa}$ com RFW e RoF como aprendizes.	150
41	Coeficiente de correlação de Spearman por conjunto de dados.	154

Lista de quadros

1	Características de cada estratégia.	53
2	Descrição dos 53 meta-atributos.	68
3	Características de cada estratégia, incluindo propostas.	69

Lista de algoritmos

Algoritmo 1 – Validação cruzada.	42
Algoritmo 2 – Amostragem ativa baseada em reserva de exemplos.	45
Algoritmo 3 – Estratégia híbrida ponderada por densidade.	61
Algoritmo 4 – Estratégia SGmulti.	63
Algoritmo 5 – Identificação do melhor algoritmo de aprendizado.	65
Algoritmo 6 – Recomendação automática de algoritmos de aprendizado.	70

Lista de tabelas

1	Pares de conjuntos com as maiores correlações.	79
2	Conjuntos de dados mais desbalanceados.	80
3	Conjuntos de dados com mais atributos.	80
4	Conjuntos de dados com mais classes.	81
5	Conjuntos de dados com mais exemplos.	81
6	Conjuntos de dados com menos exemplos.	82
7	Exemplo com legenda das diferenças estatisticamente significativas. . .	89
8	Exemplo de contagem de colocações.	89
9	Significância estatística das diferenças (ALC- μ_K).	97
10	Contagem de colocações (ALC- μ_K).	97
11	Contagem das três melhores e três piores colocações (ALC- μ_K).	99
12	Significância estatística das diferenças (ALC- σ_K).	101
13	Contagem de colocações (ALC- σ_K).	101
14	Relevância dos meta-atributos	114
15	Significância estatística das diferenças (RFw - ALC- μ_K).	151
16	Contagem de colocações (RFw - ALC- μ_K).	151
17	Significância estatística das diferenças (RoF - ALC- μ_K).	151
18	Contagem de colocações (RoF - ALC- μ_K).	151
19	Características dos conjuntos de dados (1-45).	162
20	Características dos conjuntos de dados (46-90).	163
21	Estratégias comparadas, conjuntos 1-45 com 5NN.	172
22	Estratégias comparadas, conjuntos 46-90 com 5NN.	173
23	Estratégias comparadas, conjuntos 1-45 com C4.5w	174
24	Estratégias comparadas, conjuntos 46-90 com C4.5w.	175
25	Estratégias comparadas, conjuntos 1-45 com NB.	176
26	Estratégias comparadas, conjuntos 46-90 com NB.	177
27	Estratégias comparadas, conjuntos 1-45 com SVM.	178
28	Estratégias comparadas, conjuntos 46-90 com SVM.	179
29	Estratégias comparadas, conjuntos 1-45 com RFw.	180
30	Estratégias comparadas, conjuntos 46-90 com RFw.	181

31	Estratégias comparadas, conjuntos 1-45 com RoF.	182
32	Estratégias comparadas, conjuntos 46-90 com RoF.	183

Lista de símbolos

θ — modelo induzido

$\theta_{\mathcal{L}'}$ — modelo induzido com exemplos do conjunto \mathcal{L}'

Θ — conjunto de modelos possíveis

$y \in Y$ — classe e conjunto de classes possíveis

$\mathbf{x} \in X$ — tupla/vetor de atributos e conjunto de tuplas válidas

ϕ — função de indução de modelos

$A \in \mathcal{A}$ — conjunto de valores válidos do atributo A e conjunto de atributos existentes

\mathcal{L} — conjunto de exemplos rotulados

$\tilde{\mathcal{L}}$ — conjunto de exemplos rotulados ponderados

\mathcal{U} — reserva de exemplos

V_i — sequência de valores ocorridos para $x_i, \mathbf{x} \in \mathcal{U}$

\mathcal{F} — espaço de atributos transformado

$P_\theta(\mathbf{y} \mid \mathbf{x})$ — probabilidade de ocorrência da classe \mathbf{y} dado \mathbf{x} - segundo o modelo θ

$\mathbf{P}_{\sim\theta}(\mathbf{x})$ — vetor com a distribuição de probabilidades para \mathbf{x} - segundo o modelo θ

$\hat{y}_\theta(\mathbf{x})$ — classe mais provável do exemplo \mathbf{x} de acordo com modelo θ

$\text{Inf}_\theta(\mathbf{x})$ — medida de informatividade de \mathbf{x}

$f_i(\mathbf{x})$ — função preditiva da classe y_i

g — função sigmoide logística

$R_\theta(\mathcal{N})$ — matriz de confusão para o conjunto de teste \mathcal{N} de acordo com o modelo θ

$r(R_\theta)$ — acurácia convencional para a matriz de confusão R_θ

$\mathbf{e}(R_\theta)$ — vetor de ocorrências esperadas na matriz de confusão R_θ

$\mathbf{p}(R_\theta)$ — vetor de ocorrências preditas na matriz de confusão R_θ

$o(\mathbf{x})$ — função/oráculo que revela o rótulo do exemplo \mathbf{x}

$\mathcal{P}_i(\mathcal{L})$ — partição i do conjunto \mathcal{L}

k — número de partições da validação cruzada

$\kappa(S)$ — índice kappa multiclasse para a matriz de confusão S

\mathcal{M} — conjunto de exemplos de treinamento ou reserva para aprendizado ativo

\mathcal{N} — conjunto de exemplos separados para teste

S — matriz de confusão cumulativa

$\hat{\mathbf{x}}$ — exemplo escolhido para consulta

\dot{c} — orçamento

q — função-critério de consulta

$P_{\theta_{max}}(\mathbf{x})$ — máxima probabilidade a posteriori para o exemplo \mathbf{x} segundo o modelo θ

$\hat{y}_{\theta}(\mathbf{x})$ — classe mais provável do exemplo \mathbf{x} de acordo com modelo θ

M_{θ} — medida de informatividade baseada na margem de incerteza

E — entropia normalizada

E_{θ} — medida de informatividade baseada em entropia

$\text{sim}(\mathbf{x}, \mathbf{u})$ — similaridade entre dois exemplos \mathbf{x} e \mathbf{u}

α — importância relativa da densidade dos exemplos da reserva

δ — importância relativa da densidade dos exemplos rotulados

$\text{ID}_{\theta}(\mathbf{x})$ — densidade de informação do exemplo \mathbf{x} relacionada aos exemplos não rotulados

$\text{ID}_{\text{TU}_{\theta}}(\mathbf{x})$ — $\text{ID}_{\theta}(\mathbf{x})$ inversamente ponderado pelos exemplos já rotulados

$d(\mathbf{u}, \mathbf{z})$ — distância entre dois exemplos \mathbf{u} e \mathbf{z}

S — conjunto de hipóteses mais específicas

G — conjunto de hipóteses mais gerais

$\mathcal{O}(\dots)$ — ordem de complexidade

C — comitê, conjunto de modelos

$\text{JS}(C, \mathbf{x})$ — divergência não-ponderada de Jensen-Shannon no comitê C para o exemplo \mathbf{x}

O — função objetivo na estratégia de redução do erro

ρ — correlação de Pearson entre medidas de Mar e TU

ρ_{limite} — valor limite para ρ

κ — índice kappa multiclasse

$\text{U}(i, j)$ — distribuição uniforme no intervalo $[i; j] \cap \mathbb{Z}$

c — classe negativa

$\bar{\mathcal{L}}$ — grupo de conjuntos de dados ponderados e artificialmente rotulados

Φ — conjunto de funções indutoras

\mathcal{L} — coleção de conjuntos de dados rotulados

γ — metaclasses

χ — metaexemplo

Λ — metaconjunto de dados

η — metamodelo

$\psi(\Lambda)$ — função indutora para o metaconjunto de dados Λ

$\#at, \#ex, \#nc, \#ea, \%no, lgex, lgea, \#no_{min}, \#no_{max}, \#no_{mea}, \#no_{min/max}, \mu_{min}, \mu_{max}, \mu_{mea}, \mu_{min/max}$ — meta-atributos descritos no Quadro 2 (parte 1/4)

$\sigma_{min}, \sigma_{max}, \sigma_{mea}, \sigma_{min/max}, en_{min}, en_{max}, en_{mea}, en_{min/max}, \rho_{min}, \rho_{max}, \rho_{mea}, \rho_{min/max}$ — meta-atributos descritos no Quadro 2 (parte 2/4)

$sk_{min}, sk_{max}, sk_{mea}, sk_{min/max}, ku_{min}, ku_{max}, ku_{mea}, ku_{min/max}, cn_{k1}, cn_{k1.5}, cn_{k2}, cn_{h1}, cn_{h1.5}, cn_{h2}$ — meta-atributos descritos no Quadro 2 (parte 3/4)

$du_{k1}, du_{k1.5}, du_{k2}, du_{h1}, du_{h1.5}, du_{h2}, si_{k1}, si_{k1.5}, si_{k2}, si_{h1}, si_{h1.5}, si_{h2}$ — meta-atributos descritos no Quadro 2 (parte 4/4)

μ_{κ} — valor médio de κ para 5 execuções de validação cruzada em 5 partes

σ_{κ} — desvio padrão de κ para 5 execuções de validação cruzada em 5 partes

G — índice de impureza de Gini

$d_m(\mathbf{u}, \mathbf{z})$ — distância de Mahalanobis entre dois exemplos \mathbf{u} e \mathbf{z}

U — matriz de exemplos da reserva

$cov(U)$ — matriz de covariância de U

GiB — giga *bytes*

Sumário

1	Introdução	29
1.1	Motivação	30
1.2	Hipóteses	33
1.3	Proposta	34
1.4	Contribuição	35
1.5	Estrutura do documento	37
2	Contexto	39
2.1	Classificação	39
2.2	Aprendizado ativo	42
2.3	Estratégias de consulta	44
2.4	Considerações	51
3	Propostas	55
3.1	Estratégias	56
3.2	Aprendizado meta-ativo	64
3.3	Considerações	67
4	Método experimental	71
4.1	Cenário escolhido	72
4.2	Conjuntos de dados	74
4.3	Algoritmos de aprendizado	78
4.4	Estratégias	85
4.5	Avaliação	86
4.6	Curvas de ranqueamento	91
4.7	Considerações	93
5	Resultados	95
5.1	Nível base	96
5.2	Nível meta - Recomendação de algoritmos	108
5.3	Nível meta - Outras modalidades	116
5.4	Considerações	123
6	Conclusão	127
6.1	Dificuldades	128
6.2	Metas atingidas	129
6.3	Hipóteses comprovadas	130
6.4	Limitações	131
6.5	Desdobramentos Futuros	133

Referências	137
APÊNDICE A Comitês como aprendizes ativos	149
APÊNDICE B Dependência entre conjuntos de dados	153
APÊNDICE C Meta-aprendizado	155
C.1 Generalização em pilha	156
C.2 Caracterização por modelos	156
C.3 Marcadores de referência	156
C.4 Caracterização direta	157
APÊNDICE D Ferramentas	159
APÊNDICE E Conjuntos de dados	161
ANEXO A Atividades complementares	165
ANEXO B Resultados detalhados	171

Introdução

A área de pesquisa conhecida como *aprendizado de máquina* permeia o cotidiano humano provendo auxílio em diversas áreas, como bioinformática, visão computacional e recomendação de conteúdo - entre outras (WANG *et al.*, 2014; KRIZHEVSKY; SUTSKEVER; HINTON, 2012; RICCI; ROKACH; SHAPIRA, 2011). Uma das principais tarefas de algoritmos de aprendizado de máquina é a de classificação de dados. Ela busca associar objetos de interesse, *exemplos*¹, a classes. Esse tipo de categorização é típica do grupo de tarefas denominado *aprendizado supervisionado*, pois requer, usualmente, a supervisão de um especialista - chamado *oráculo* no contexto deste trabalho. Após um esforço suficiente de supervisão na categorização de um *conjunto* de exemplos para treinamento, torna-se possível a predição da classe de novos exemplos. Esse é o objetivo final da tarefa de classificação. Nesse contexto, o desempenho preditivo de algoritmos de aprendizado depende do cumprimento de pelo menos três requisitos:

- a **amostragem** de dados para a construção do conjunto de treinamento deve ser representativa do problema em questão;
- o processo de **categorização** desses dados deve disponibilizar conhecimento suficiente para o aprendizado da tarefa desejada; e,
- o algoritmo de **aprendizado** deve ser adequado ao domínio do problema.

Esses requisitos podem ser cumpridos de acordo com o ambiente de aprendizado de máquina ilustrado na Figura 1. Uma amostra dos exemplos (ou todos) disponíveis no domínio do problema é classificada de acordo com as classes atribuídas por um supervisor. O resultado é o conjunto de dados de treinamento para a geração de um *modelo* de classificação.

¹ Os termos adotados para a representação inequívoca de conceitos importantes estão marcados em itálico em sua definição/primeira ocorrência. Em tempo, optou-se pela grafia das palavras em português quando existente (*ranqueamento*, por exemplo, introduzida na Seção 1.4) e pelas contrações facultativas mais comuns de preposição e artigo indefinido, prescritas na norma padrão da língua portuguesa (LIMA, 1973; LETRAS, 2004).

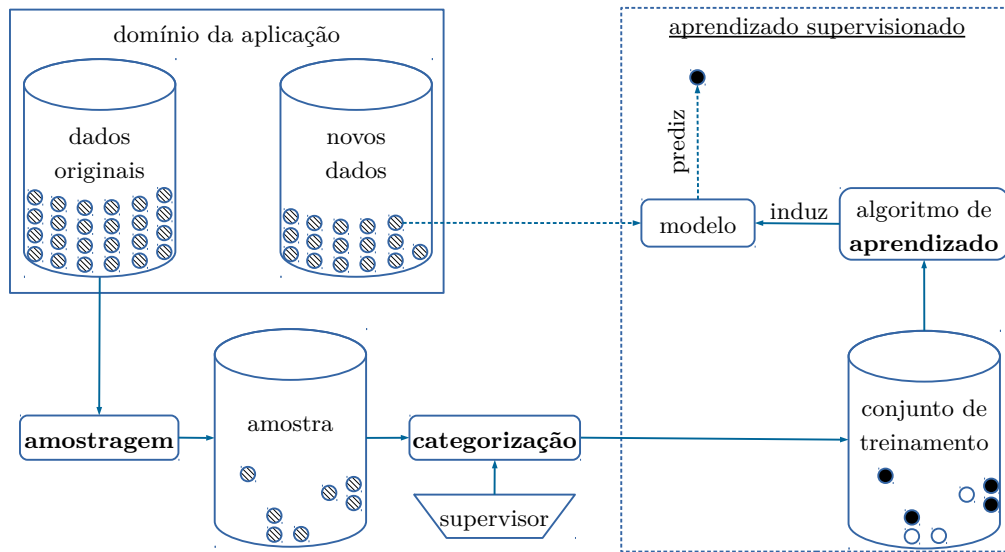


Figura 1 – Esquema de aprendizado supervisionado e componentes externos.

1.1 Motivação

Assumindo-se a disponibilidade de infraestrutura computacional, a necessidade de esforço humano pode se tornar o maior obstáculo para um projeto de aprendizado supervisionado ser bem sucedido. Esse esforço pode ser necessário em pelo menos duas perspectivas de atuação que abrangem os passos em destaque na Figura 1:

- no projeto do sistema, tanto na escolha da forma de amostragem quanto na determinação do algoritmo de aprendizado mais adequado ao problema; e,
- durante a categorização dos dados, normalmente realizada por um ou mais supervisores humanos.

No primeiro caso, um especialista em aprendizado de máquina toma decisões relacionadas à implementação do sistema. No segundo caso, especialistas no domínio do problema fazem a *rotulação*, ou seja, atribuem rótulos informando a classe de cada objeto de interesse. As atividades realizadas nos dois momentos podem ter um custo elevado.

O custo de rotulação pode ser facilmente contabilizado pela quantidade de rótulos necessários para a construção do conjunto de treinamento. Por outro lado, o custo das decisões do especialista é de estimativa mais incerta, pois envolve o esforço despendido no processo de escolha e suas possíveis consequências. Quanto menor a confiança desse especialista na escolha do processo de amostragem ou do algoritmo de aprendizado, maiores as chances da aplicação incorrer em custos adicionais ou prejuízo no desempenho preditivo futuro. Portanto, custo de rotulação e custo decorrente de inadequação do sistema à tarefa desejada são dois aspectos importantes - quando se deseja garantir a viabilidade financeira de uma aplicação.

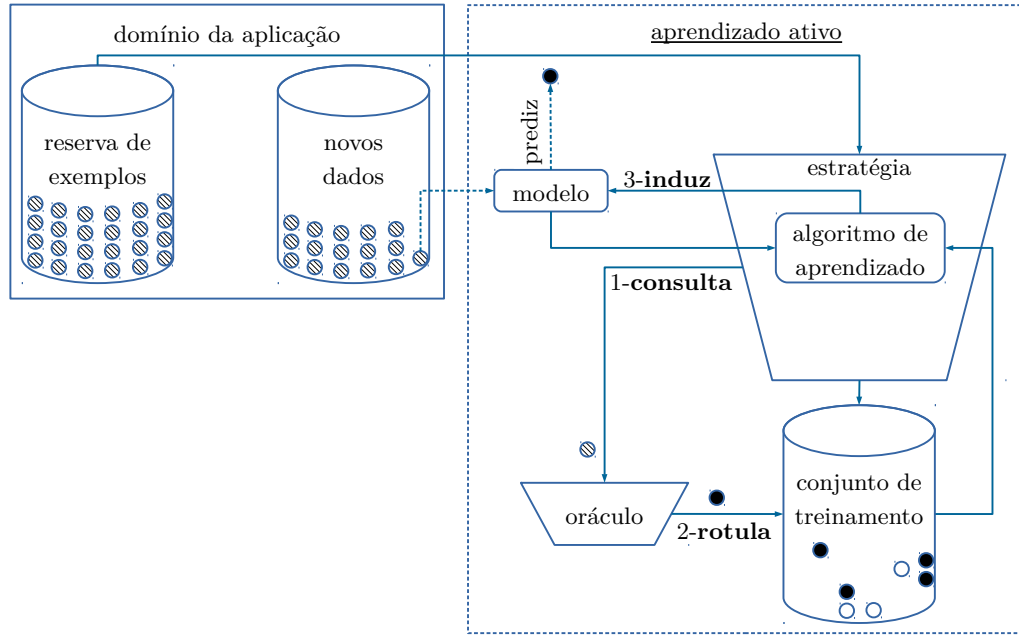


Figura 2 – Esquema de aprendizado ativo.

Além da questão econômica, o desempenho do sistema também depende das decisões tomadas na fase de projeto. Acurácia preditiva e tempo de processamento, por exemplo, são diretamente afetados pela escolha do algoritmo de aprendizado. Ademais, a qualidade dos dados amostrados e o próprio viés da *estratégia* de amostragem escolhida também influenciam diretamente no desempenho. Logo, a definição do par *estratégia-algoritmo* ideal para um dado problema tem um papel central no projeto de um sistema de aprendizado supervisionado como aquele exposto previamente na Figura 1. A necessidade de definição desse par motiva a realização de uma pesquisa que se aprofunde além da simples comparação de poucas estratégias de amostragem, com mais do que somente um algoritmo de aprendizado e mais do que uma dúzia de conjuntos de dados - como tem sido feito na literatura (EVANS; ADAMS; ANAGNOSTOPOULOS, 2014; SETTLES; CRAVEN, 2008; SCHEIN; UNGAR, 2007; KÖRNER; WROBEL, 2006).

O cenário de *aprendizado ativo* é explorado neste trabalho visando abordar o problema da redução do custo de rotulação (Seção 2.2). Apenas os exemplos mais relevantes do conjunto de dados são escolhidos. No cenário específico desta tese, que será detalhado na Seção 4.1, o conjunto de dados não rotulados é chamado *reserva de exemplos*². Nesse contexto, o conjunto de treinamento cresce enquanto houver *orçamento* - seguindo o ciclo *consulta* → *rotula* → *induz* apresentado na Figura 2. Consequentemente, a administração do custo de rotulação passa a ser atribuição de uma estratégia de amostragem ativa. Para fins experimentais, o orçamento indica o número de *consultas* permitidas.

² [pool]

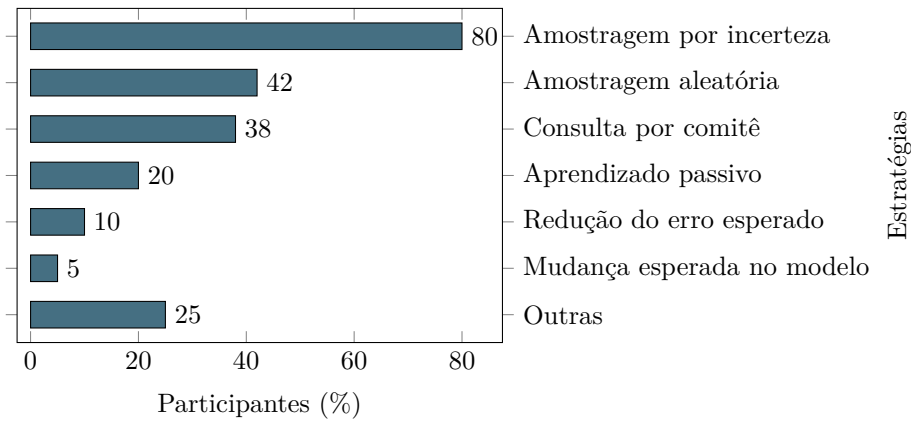


Figura 3 – Frequência de uso de estratégias na competição de aprendizado ativo - descritas na Seção 2.3. Alguns participantes adotaram mais de uma estratégia. *Adaptado de Guyon et al. (2011).*

Finalmente, a questão da possível inadequação do algoritmo de aprendizado ao domínio do problema permanece em aberto. Assim, a indução de um modelo adequado depende da escolha do par estratégia-algoritmo, ou apenas da estratégia, no caso excepcional em que tenha sido adotada uma estratégia sem *aprendiz* (Capítulo 2). Uma estratégia sem *aprendiz* é uma estratégia que não necessita de um algoritmo de aprendizado para realizar as consultas. A motivação para a pesquisa em ambas componentes do par estratégia-algoritmo é dada nas seções 1.1.1 e 1.1.2.

1.1.1 Escolha da estratégia

A existência de diversas estratégias na literatura de aprendizado ativo coloca o especialista em aprendizado de máquina diante de uma escolha para a qual não há um critério claro. A principal dificuldade advém da impossibilidade de comparação dos desempenhos dessas diferentes abordagens sem incorrer em custos de rotulação. Isso se deve ao fato de que a qualidade das consultas é dada, em última instância, pela acurácia do modelo induzido. Sendo a indução somente possível com a presença dos rótulos e esses são provenientes da interação estratégia-oráculo, tem-se um paradoxo: a escolha da estratégia depende dos rótulos que, por sua vez, dependem da estratégia.

Na literatura da área, a escolha da estratégia a ser utilizada tem sido quase arbitrária, tendendo a se concentrar no uso da mais simples - a *amostragem por incerteza*. Essa preferência foi reportada numa competição de aprendizado ativo (GUYON *et al.*, 2011) - conforme ilustrado na Figura 3. Em consonância com esse panorama está a ausência de estudos comparativos abrangentes que possam guiar o especialista na definição da estratégia de consulta a ser empregada - até onde o conhecimento do autor desta tese permite dizer. Consequentemente, a maneira ideal de redução do custo de rotulação ainda é um problema em aberto.

1.1.2 Escolha do algoritmo

A questão da escolha do algoritmo de aprendizado para a indução de modelos de classificação é análoga à questão da escolha da estratégia. A escolha do algoritmo é definida neste texto da seguinte forma: seleção *manual* - baseada em experiência pessoal de experimentos anteriores do especialista; ou, seleção *automática* - baseada em sistemas de recomendação, como o *meta-aprendizado* (Apêndice C).

Usualmente, ambas abordagens (manual e automática) empregam validação cruzada durante a comparação de algoritmos candidatos (BOUCKAERT; FRANK, 2004; BRAZDIL *et al.*, 2009). Isso as torna mais aplicáveis ao cenário em que o conjunto de treinamento já esteja construído. Entretanto, no cenário de aprendizado ativo, a rotulação é um processo em andamento, assim como seu conjunto de treinamento resultante. Tal característica impede que o especialista disponha da importante fonte de informação que um conjunto de treinamento completo poderia representar. Reduz-se, assim, a quantidade de informações disponíveis para decisões bem fundamentadas. Logo, além da escolha da estratégia ser incerta pela falta de critérios, conforme visto na Seção 1.1.1, a escolha do algoritmo mais adequado para o papel de aprendiz também carece de melhores fundamentos.

1.2 Hipóteses

O teorema conhecido como *No Free Lunch* (NFL) permite afirmar que nenhum algoritmo de aprendizado pode ser o mais adequado para todos os domínios (SCHAFFER, 1994). A mesma afirmação é válida para algoritmos de otimização (WOLPERT; MACREADY, 1997) e, conseqüentemente, válida para estratégias de amostragem ativa. Assim, é garantida a existência da questão da escolha no presente cenário do ponto de vista teórico. A razão de uma estratégia poder ser entendida como um procedimento de otimização é ela consistir na busca pela melhor sequência de exemplos dentro das possibilidades que a reserva de exemplos e o orçamento permitem. Dessa forma, é possível definir o contexto para as hipóteses desta tese: *cada estratégia tem um viés de amostragem, normalmente dependente de um viés de aprendizado, que favorece determinados domínios de problemas e prejudica outros*. Diante desse cenário, e considerando a motivação apresentada previamente na Seção 1.1, duas hipóteses foram formuladas. A Hipótese I corresponde à principal tese defendida neste trabalho:

É possível explorar relações entre conjuntos de dados e *algoritmos de aprendizado, estratégias de aprendizado ativo ou pares de ambos*, de maneira que seja possível escolher abordagens para essas duas componentes que resultem num melhor desempenho preditivo que uma escolha arbitrária.

A Hipótese II, secundária, supõe que *a presença ou tipo do viés de aprendizado podem ser controlados durante o processo de rotulação, aumentando o desempenho preditivo*.

A Seção 1.3 contém a proposta de investigação das hipóteses.

1.3 Proposta

A proposta desta tese é a investigação de alternativas para facilitar a escolha da estratégia de amostragem e/ou seu correspondente algoritmo de aprendizado mais adequados para um conjunto de dados. Uma particularidade do cenário de aprendizado ativo é que o subconjunto inicial de exemplos rotulados é vazio ou pequeno demais para escolhas apropriadas. Assim, o resultado esperado da investigação é que a escolha da estratégia, do algoritmo ou do par estratégia-algoritmo possa ser realizada de uma maneira mais objetiva que a convencional, induzindo modelos com melhor acurácia preditiva para um dado orçamento.

A investigação concentra-se nos seguintes objetivos, em ordem decrescente do grau de participação requerida do especialista que hipoteticamente usufruiria dos resultados:

- identificação de nichos de problemas onde a utilização de certas estratégias seja adequada - teste qualitativo da Hipótese I;
- desenvolvimento de uma estratégia capaz de suprimir a influência do algoritmo de aprendizado na fase de rotulação, quando seu uso for prejudicial - teste da Hipótese II; e,
- teste inicial de uma estratégia que poderia ser chamada *aprendizado meta-ativo*³ (meta-aprendizado aplicado a aprendizado ativo) capaz de escolher (e trocar) o algoritmo antes de (e durante) o aprendizado - teste quantitativo das Hipóteses I e II.

Os meios escolhidos para atingir essas metas são: revisão bibliográfica; implementação própria de *software* e reuso de bibliotecas de terceiros; aplicação dos métodos propostos em problemas relevantes; comparação com referências importantes da área; e, avaliação e validação dos resultados obtidos.

As principais contribuições e resultados decorrentes da implementação da proposta são apresentados na Seção 1.4.

³ Note-se a distinção entre *aprendizado meta-ativo* e *meta-aprendizado ativo* (SOUSA *et al.*, 2013) explicada na Seção 3.2.

1.4 Contribuição

As contribuições deste trabalho são originais até onde o conhecimento do autor permite dizer. Os resultados obtidos e publicações resultantes são listados nas seções seguintes.

1.4.1 Resultados

Este trabalho foi baseado em experimentos que simulassem situações representativas da realidade - naturalmente, dentro das limitações próprias de qualquer avaliação empírica, como a definição da coleção de conjuntos de dados, do conjunto de estratégias de aprendizado ativo e do conjunto de algoritmos de aprendizado. Os resultados obtidos, dentro do arcabouço experimental adotado, são apresentados a seguir.

- A estratégia HTU (Seção 3.1.2), proposta para controle da atuação do aprendiz e, similarmente, ATU (Seção 3.1.1), proposta para avaliar o efeito da ausência de aprendiz, mostraram-se, em vários aspectos, as de menor risco para o orçamento dentre as abordagens comparadas - sem incorrer em redução na acurácia preditiva.
- As possíveis adequações e inadequações entre estratégias, algoritmos de aprendizado e conjuntos de dados foram ilustradas por meio de árvores de decisão e curvas de aprendizado, de forma que foi possível identificar a influência preponderante do algoritmo no desempenho preditivo ao longo do aprendizado.
- A escolha automática do algoritmo de aprendizado requer investigações mais extensivas. Na presença, dentro da coleção, de mais de conjunto de dados de cada domínio, tal escolha automática superou a escolha arbitrária e a referência da área (Apêndice C).
- Outras modalidades de recomendação automática também foram experimentadas e se mostraram mais desafiadoras: recomendação de estratégias e recomendação de pares estratégia-algoritmo.
- Algumas modalidades de recomendação foram percebidas como desafios sem indícios de viabilidade: recomendação de métricas de distância para estratégias baseadas em densidade e recomendação sobre adotar ou não aprendizado ativo.
- As características mais relevantes dos conjuntos de dados para a recomendação automática foram organizadas visualmente numa estrutura de árvore de decisão e brevemente discutidas.

Em síntese, os experimentos mostraram que é possível obter modelos com melhor desempenho preditivo no cenário de aprendizado ativo, caso alguma das abordagens propostas, ou uma combinação delas, seja adotada: ATU, HTU e, num cenário mais restrito

e dependente de maiores investigações quanto à sua real viabilidade, aprendiz meta-ativo.

1.4.2 Publicações

As contribuições são diretamente ligadas às hipóteses previamente apresentadas e são enumeradas como segue, juntamente com as publicações decorrentes.

1. Demonstração da influência determinante do algoritmo de aprendizado adotado como aprendiz - submetido por Santos e Carvalho (2016b).
2. Experimentação inicial de uma nova abordagem para aprendizado ativo, chamada *aprendizado meta-ativo*, cujo objetivo é buscar o algoritmo com o melhor viés para um dado conjunto de dados - submetido por Santos e Carvalho (2016a).
3. Experimentação da mesma abordagem do Item 2 é em outras modalidades de recomendação como: estratégia e pares estratégia-algoritmo - artigo em processo de escrita.
4. Uma estratégia baseada na inibição do aprendiz enquanto ele for potencialmente prejudicial - publicado por Santos e Carvalho (2015b).

Publicações referentes a contribuições metodológicas ou visando contribuir com a organização da área de aprendizado ativo são enumeradas a seguir.

5. Comparação descritiva e experimental de estratégias da literatura utilizando 28 conjuntos de dados e algoritmos de aprendizado com vieses diversos - publicado por Santos e Carvalho (2014a).
6. Adaptação de estratégias para o cenário multiclasse - idem ao Item 5.
7. Comparação e adaptação de estratégias utilizando um algoritmo ainda pouco explorado no cenário de aprendizado ativo: as redes neurais com pesos aleatórios (SCHMIDT *et al.*, 1992) - publicado por Santos e Carvalho (2014b).
8. Proposta metodológica das *curvas de ranqueamento*⁴, que contornam a dificuldade de apresentação das curvas de aprendizado obtidas em experimentos com 39 conjuntos de dados - idem aos itens 1 e 4.

Adicionalmente, os seguintes sistemas de *software* foram desenvolvidos e disponibilizados publicamente.

⁴ Apesar do vocábulo estrangeiro *ranking* constar no léxico contemporâneo da língua portuguesa (LETRAS, 2004), optou-se aqui pela grafia *ranqueamento* já presente na literatura brasileira da área (COLARES, 2005), pois seu radical permite outras possibilidades não contempladas pelo léxico, como *ranqueado* e *ranqueável*.

9. Biblioteca de encapsulamento de algoritmos de classificação (SANTOS; CARVALHO, 2015c).
10. Biblioteca de estratégias de amostragem ativa e ambiente experimental (SANTOS; CARVALHO, 2015a).

Finalmente, dois artigos resultaram de colaboração com outros grupos de pesquisa, envolvendo extração de atributos de séries financeiras (BEDO *et al.*, 2013) e seleção de métodos de extração de atributos (BEDO *et al.*, 2015).

1.5 Estrutura do documento

O contexto da pesquisa e a notação matemática, incluindo a terminologia adotada, são apresentados no Capítulo 2. O capítulo também contém uma revisão da literatura fundamental para esta tese: aprendizado de máquina e aprendizado ativo. No Capítulo 3, são apresentadas as abordagens propostas ou adaptadas: ATU, HTU, SGmulti e a possibilidade experimentada de aprendizado meta-ativo. O Capítulo 4 contém a descrição do cenário escolhido; a enumeração das estratégias, algoritmos e conjuntos de dados adotados; e, a apresentação dos métodos de avaliação comuns a todos os experimentos realizados. A proposta metodológica chamada *curvas de ranqueamento* é introduzida nesse capítulo.

A avaliação empírica e seus resultados associados são apresentados no Capítulo 5. Por fim, no Capítulo 6, é feita uma análise geral das metas atingidas, contribuições, dificuldades e limitações da pesquisa realizada. Possíveis desdobramentos futuros são delineados ao final do capítulo.

Alguns conteúdos foram registrados em apêndices e anexos para brevidade do corpo principal do texto: revisão sobre meta-aprendizado (Apêndice C), ferramentas utilizadas (Apêndice D), descrição da coleção de conjuntos de dados adotada (Apêndice E), atividades complementares (Anexo A) e resultados detalhados (Anexo B). Os apêndices A e B contêm experimentos e resultados adicionais sobre aprendizes ativos baseados em comitês e o efeito da similaridade entre conjuntos de dados, respectivamente.

Contexto

*...no matter how many
instances of white swans we
may have observed, this does
not justify the conclusion that
all swans are white.*

Karl Popper¹

A pesquisa realizada exige a apresentação dos temas pertinentes e as definições adotadas no presente documento, incluindo notação e terminologia. A contextualização se inicia na Seção 2.1, onde a tarefa de classificação, no contexto de aprendizado de máquina, é apresentada em termos gerais. O assunto central desta tese, que engloba o aprendizado ativo e suas principais estratégias, é introduzido nas seções 2.2 e 2.3. Finalmente, na Seção 2.4, são feitas considerações a respeito dos temas abordados e uma comparação objetiva entre as estratégias.

2.1 Classificação

A atividade de identificação da categoria de um determinado objeto de interesse é chamada *classificação*. Ela é a tarefa que pode se beneficiar mais amplamente dos resultados da presente pesquisa. Seus principais conceitos são apresentados nas seções seguintes juntamente com a forma de representação adotada neste texto. Essa exposição tem o objetivo de facilitar consultas posteriores sobre terminologia durante a leitura dos demais capítulos. As escolhas de alguns símbolos são baseadas no livro de Settles (2012).

2.1.1 Modelo

O aprendizado de um determinado conceito pode ser visto como uma busca dentro do espaço de hipóteses representado pelo conjunto Θ de modelos de representação

¹ “...não importa quantos exemplos de cisne branco tenhamos observado, isso não justifica a conclusão de que *todos* os cisnes sejam brancos.” - Popper (1959), sobre o problema da indução na ciência e a necessidade do princípio da falseabilidade.

possíveis. Assim, a tarefa de classificação induz um modelo $\theta \in \Theta$ baseado num conjunto de dados de treinamento, cujos rótulos foram previamente fornecidos por um especialista no domínio do problema (Seção 1.1). Os algoritmos de aprendizado são considerados determinísticos, por simplicidade.

Antes da construção de cada modelo θ , é desejável saber qual parte dos dados é relevante. Esse é um problema em aberto que pode ser entendido como uma busca pela melhor maneira de se amostrarem os dados. A motivação para essa busca pode frequentemente advir da intratabilidade de um excesso de dados ou da escassez de recursos para manutenção de um especialista no papel de supervisor. Ambos os casos podem se beneficiar do tipo de aprendizado denominado *ativo*; que será apresentado na Seção 2.2.

2.1.2 Classe

Neste texto, o termo *classificador* refere-se ao modelo θ mais recente induzido pelo algoritmo de aprendizado, quando o objetivo é a predição de classes - diferentemente de quando o objetivo é a estimação de probabilidades, por exemplo, cujos valores de retorno são contínuos. A classe y de um exemplo é um dos elementos do limitado conjunto Y de classes de um dado problema.

2.1.3 Atributos

Um exemplo não rotulado é dado por uma tupla de *atributos*² \mathbf{x} pertencente ao conjunto X de tuplas válidas - ou *vetores* válidos, se o contexto exigir e assumindo uma binarização prévia dos atributos nominais, caso existam. Dessa forma, a representação de um exemplo rotulado é um par $\langle \mathbf{x}, y \rangle \in X \times Y$. Dado um subconjunto contido em $X \times Y$, cada algoritmo de aprendizado tem sua própria função de indução de modelos³ $\phi: 2^{X \times Y} \rightarrow \Theta$.

Cada atributo é representado por um conjunto de valores $A \in \mathcal{A}$, onde \mathcal{A} é o conjunto de atributos do problema. Este é dado pela Equação 2.1, onde $\dim \mathbf{x}'$ indica a dimensão de qualquer $\mathbf{x}' \in X$.

$$\mathcal{A} = \{ \{x_i \mid \mathbf{x} \in X\} \mid 1 \leq i \leq \dim \mathbf{x}' \} \quad (2.1)$$

Assim, cada conjunto de valores A contém os valores que cada atributo de \mathbf{x} pode assumir. Por exemplo, sendo o primeiro atributo $A = \{alto, medio, baixo\}$, é possível que $x_1 = medio$; ou, sendo o segundo atributo $A' = \mathbb{R}$, $x_2 = 3,7$. Diferentemente, V_i é a sequência - ou vetor, dependendo do contexto - de todos os valores ocorridos para o atributo correspondente à componente x_i de todos os exemplos da reserva $\mathcal{U} \subset X$ - conforme Equação 2.2.

$$V_i = \langle x_i \mid \mathbf{x} \in \mathcal{U} \rangle \quad (2.2)$$

² Neste texto, os *atributos* de um exemplo não incluem a classe do exemplo, exceto quando explicitado como *atributo alvo*.

³ A notação 2^B representa o conjunto de todos os subconjuntos de B - chamado *potência* ou *de partes* (DEVLIN, 2012).

A reserva de exemplos (\mathcal{U}) consiste no conjunto de exemplos não rotulados disponíveis numa dada aplicação. O tipo do atributo é dado pela função *isnom*: $\mathcal{A} \rightarrow \{0,1\}$ que retorna 1 para atributos nominais e 0 para atributos numéricos. O conjunto de exemplos rotulados é representado por \mathcal{L} . Eventualmente, nas situações em que seja necessário atribuir pesos aos exemplos, cada exemplo é representado por uma tripla $\langle \mathbf{x}, y, w \rangle, 0 \leq w \leq 1$. Nesse caso, o *conjunto rotulado de exemplos ponderados* é dado pelo símbolo $\tilde{\mathcal{L}}$.

O termo *espaço de atributos* representa um espaço em $\mathbb{R}^{|\mathcal{A}|}$ onde cada exemplo pode ser situado. Essa definição pressupõe a binarização de atributos nominais, quando necessário. Ela difere do *espaço de parâmetros* \mathcal{F} que é um espaço de atributos transformado, cujo exemplo típico de coordenadas é o conjunto de pesos de uma rede neural (HAYKIN; NETWORK, 2004).

2.1.4 Valores de retorno

Dado um exemplo \mathbf{x} , e supondo que o modelo θ permita a estimação da probabilidade P_θ de \mathbf{x} ter a classe y , a classe mais provável \hat{y}_θ é obtida pela Equação 2.3.

$$\hat{y}_\theta(\mathbf{x}) = \arg \max_y P_\theta(y | \mathbf{x}) \quad (2.3)$$

P_θ é a base de grande parte das medidas de *informatividade* $\text{Inf}_\theta(\mathbf{x})$ empregadas pelas estratégias. Entretanto, modelos não probabilísticos podem retornar valores numéricos que não representam uma distribuição de probabilidades. Esses valores são resultado de uma função preditiva $f_y(\mathbf{x})$: valores próximos de 1 indicam pertinência à classe y . Classificadores cujos modelos retornam valores fora do intervalo $[0;1]$ podem fazer uso da função sigmoide logística exibida na Equação 2.4 para cada classe y (SEGGERN, 2006).

$$g_y(\mathbf{x}) = \frac{1}{1 + e^{-f_y(\mathbf{x})}} \quad (2.4)$$

$\mathbf{P}_{\sim\theta}$ representa a sequência com a distribuição de probabilidade das classes para um dado exemplo \mathbf{x} . Por exemplo, $\mathbf{P}_{\sim\theta}(\mathbf{x}) = \langle 0; 0,5; 0,4; 0,1 \rangle, |Y| = 4$.

A matriz de confusão $\mathbf{R}_\theta: 2^{X \times Y} \rightarrow \mathbb{R}^{|Y| \times |Y|}$ possibilita o cálculo de importantes medidas de desempenho (STEHRMAN, 1997). Ela representa uma tabela de frequências de classes preditas para cada classe esperada. Supondo um conjunto de teste \mathcal{N} , a matriz de confusão seria denotada por $\mathbf{R}_\theta(\mathcal{N})$, porém, dependendo do contexto, o parâmetro θ e o conjunto de teste \mathcal{N} podem ser omitidos. A notação para a acurácia convencional é representada, neste texto, por $r(\mathbf{R}_\theta)$. Os vetores de ocorrências esperadas e preditas são dados, respectivamente, por $\mathbf{e}(\mathbf{R}_\theta)$ e $\mathbf{p}(\mathbf{R}_\theta)$. A verdadeira classe de um exemplo é revelada pela função oráculo $o: X \rightarrow Y$.

2.1.5 Validação

Para a aplicação de validação cruzada (Capítulo 4), \mathcal{L} é dividido em k partições. Assume-se, por simplicidade, que $|\mathcal{L}|$ é divisível por k . Assim, $\mathcal{P}_i(\mathcal{L})$ representa a partição de índice i conforme condições 2.5, 2.6, 2.7 e 2.8.

$$\bigcup_{1 \leq i \leq k} \mathcal{P}_i(\mathcal{L}) = \mathcal{L} \quad (2.5)$$

$$\mathcal{P}_i(\mathcal{L}) \neq \emptyset, 1 \leq i \leq k \quad (2.6)$$

$$|\mathcal{P}(\mathcal{L})| = \frac{|\mathcal{L}|}{k} \quad (2.7)$$

$$\mathcal{P}_i(\mathcal{L}) \cap \mathcal{P}_j(\mathcal{L}) \neq \emptyset \Rightarrow \mathcal{P}_i(\mathcal{L}) = \mathcal{P}_j(\mathcal{L}) \quad (2.8)$$

O Algoritmo 1 descreve o procedimento de validação cruzada para a medida de desempenho adotada nesta tese (Seção 3.1.2).

Algoritmo 1: Validação cruzada.

Entrada:

k - número de partições da validação cruzada

$\phi : 2^{X \times Y} \rightarrow \Theta$ - função indutora (algoritmo de aprendizado ou estratégia de amostragem ativa)

\mathcal{L} - conjunto de dados rotulados

Resultado:

κ - índice kappa (Seção 4.5.1)

```

1 função val.cruzada( $k, \phi, \mathcal{L}$ )
2    $S \leftarrow 0_{|Y|, |Y|}$   $\triangleleft$  matriz de confusão cumulativa
3   para cada  $i, 1 \leq i \leq k$  faça
4      $\mathcal{M} = \bigcup_{j \neq i} \mathcal{P}_j(\mathcal{L})$   $\triangleleft$  exemplos disponíveis
5      $\mathcal{N} = \mathcal{P}_i(\mathcal{L})$   $\triangleleft$  exemplos reservados para teste
6      $\theta = \phi(\mathcal{M})$ 
7      $S \leftarrow S + R_\theta(\mathcal{N})$ 
8   retorna  $\kappa(S)$ 

```

2.1.6 Algoritmos de aprendizado

Os algoritmos de aprendizado relevantes para esta tese serão mencionados no Capítulo 4 juntamente com os valores adotados para seus parâmetros.

2.2 Aprendizado ativo

Segundo Settles (2012), *aprendizado ativo* é o estudo de máquinas de aprendizado capazes de se aprimorar fazendo perguntas. A origem do termo remete à Pedagogia. Nela,

o termo se refere ao aproveitamento da individualidade de cada aluno e seu próprio ritmo de aprendizado. Guardadas as devidas proporções, pode-se traçar um paralelo com a área de aprendizado de máquina: um algoritmo pode usufruir de uma atenção seletiva que priorize os exemplos mais difíceis para ele num dado momento, ou seja, que possam trazer mais benefícios para a indução de modelos preditivos num dado instante.

Em sua forma mais geral, como esboçado por [Forman \(2012\)](#), o aprendizado de máquina ativo pode lançar mão de todo o conhecimento que um especialista (supervisor humano) seja capaz de transmitir dentro das limitações de configurabilidade do sistema. Alguns pontos de configuração seriam, por exemplo: reescrita do programa de extração de atributos visando maior separabilidade entre as classes; composição de expressões regulares para extrair termos adequados de textos técnicos; e, criação de regras de classificação. Outras informações, mais diretamente obteníveis do supervisor humano, incluiriam: valores de atributos, classes associadas com atributos, exemplos completos sob demanda, entre outros. Em linhas gerais, a aquisição de informação deve ser guiada por duas noções: preferir aquela para a qual o estado corrente do modelo é incerto; e, preferir aquela estimada como a mais relevante ([KRISHNAPURAM; YU; RAO, 2011](#)).

Dentre as perguntas que as máquinas de aprendizado ativo são capazes de elaborar, a mais direta seria “Qual é a classe do exemplo \mathbf{x} ?”. A obtenção de um rótulo confiável normalmente é um processo custoso, logo, trata-se de um ponto de fundamental importância, dada a necessidade de exemplos rotulados em muitas aplicações (Capítulo 1). Isso torna o problema de decidir qual seria o melhor \mathbf{x} a ser consultado a motivação basilar desta tese. Adicionalmente, a quantidade de exemplos disponíveis pode ser abundante e a capacidade de esforço humano disponível limitada. Dessa forma, apenas uma parcela criteriosamente escolhida dos exemplos deve ser rotulada - *consultada*, na terminologia de aprendizado ativo. Essa abordagem se dá em oposição ao *aprendizado por exemplos* convencional ([VALIANT, 1984](#)), também chamado *passivo*. No aprendizado passivo, procura-se pelo maior conjunto de treinamento possível ou realiza-se uma amostragem aleatória. No primeiro caso, o custo de rotulação pode se tornar proibitivo. No segundo caso, a decisão quanto à relevância dos exemplos é deixada ao acaso. Em ambos os casos, dependendo da aplicação, a construção do conjunto de treinamento pode ser crítica. Por exemplo, quando a consulta de um exemplo envolve reações químicas destrutivas, é desejável fazer o mínimo possível de consultas visando um reduzido custo material. Similarmente, se o mecanismo rotulador, o oráculo, for um especialista humano ou mesmo um robô ([BRYANT et al., 2001](#)), é desejável parcimônia nas consultas para não incorrer num esforço elevado de atenção humana ou movimento mecânico. Assim, um direcionamento adequado do esforço de aprendizado tem como resultado um processo de rotulação mais econômico. A variedade de abordagens existentes para esse direcionamento será assunto da Seção 2.3. As definições necessárias sobre aprendizado ativo são dadas a seguir - analogamente à Seção 2.1, que discorrera sobre classificação.

Na literatura, os termos *aprendiz* (ativo) e *estratégia* (de aprendizado ativo) são frequentemente usados de forma intercambiável (BALCAN; BEYGELZIMER; LANGFORD, 2009). Entretanto, neste texto, optou-se por chamar *aprendiz* apenas a parte da amostragem ativa referente ao aprendizado, ou seja, o algoritmo e seus modelos induzidos. Existem estratégias sem aprendiz, apesar da ausência de aprendiz contradizer o princípio motivador original do aprendizado ativo anteriormente explicitado.

Nesta tese, o aprendizado convencional é chamado *passivo*, pois é realizado com todos os exemplos disponíveis devidamente rotulados. A estratégia de amostragem de maior correspondência com o aprendizado passivo é a *aleatória*, pois é uma aproximação não enviesada de todo o conjunto de exemplos.

Cada consulta de uma dada estratégia ao oráculo visa a obtenção do rótulo que indica a classe verdadeira do exemplo \hat{x} escolhido. As consultas podem ter um viés mais *exploratório*, que busca maximizar a variedade na escolha de exemplos; ou, mais *prospec-tivo*, que se concentra apenas nos casos mais críticos, ou seja, com mais informatividade. O limite de duração para o processamento computacional entre consultas é aqui equiparado ao *tempo de espera tolerável* (NAH, 2003), pois é a parte do tempo de processamento que realmente afeta o oráculo. O símbolo ϵ representa o valor do orçamento, que é o número de consultas permitidas.

2.3 Estratégias de consulta

Na literatura de aprendizado ativo há diversos subcasos, cenários e estratégias. O subcaso e o cenário mais frequentes, onde se situa a presente tese, são a consulta de rótulos e o cenário baseado em reserva de exemplos, respectivamente. No caso das estratégias, elas são frequentemente baseadas em diferentes concepções de relevância de exemplos ou mesmo diferentes teorias do aprendizado. A amostragem por incerteza, por exemplo, assume que os exemplos - e seus rótulos - são gerados de acordo com uma distribuição de probabilidades. A amostragem por busca no espaço de hipóteses, por sua vez, assume a existência de hipóteses integrantes de um *espaço de versões* (MITCHELL, 1997). Estratégias sem aprendiz, por outro lado, são independentes de algoritmo de aprendizado. Essa diversidade de embasamentos configura-se, praticamente, como um conjunto de paradigmas de consulta. Suas principais representantes são apresentadas nas seções seguintes, com os respectivos princípios de funcionamento e características principais. A abreviação definida para cada estratégia é dada entre parênteses na seção correspondente. Na apresentação das ordens de complexidade, é assumido que todos os cálculos possíveis de realização antes do início do processo de rotulação já foram feitos e seus resultados estão disponíveis em memória. Essa premissa se baseia no cenário que será especificado na Seção 4.1. Ele garante a disponibilidade de todos os exemplos antes do início do processo de rotulação.

Dessa forma, apenas o custo computacional entre consultas se configura como potencial custo financeiro, pois consome o tempo do oráculo.

O Algoritmo 2 é o invólucro mais comum das estratégias apresentadas. Ele aceita diferentes critérios de consulta (*funções-critério*), descritos nas seções a seguir.

Algoritmo 2: Amostragem ativa baseada em reserva de exemplos.

Entrada:

\mathcal{U} - reserva de exemplos

\mathcal{L} - conjunto inicial de exemplos rotulados

ϕ - função indutora (algoritmo de aprendizado)

$q: X \times \dots \times \Theta \rightarrow \mathbb{R}$ - função-critério de consulta a se maximizar

ζ - orçamento (quantidade de exemplos a rotular)

Resultado:

\mathcal{L}' - conjunto final de exemplos rotulados

```

1 função amostragem( $\mathcal{U}, \mathcal{L}, \phi, q, \zeta$ )
2   se  $\zeta = 0$  então
3     retorna  $\mathcal{L}$ 
4   senão
5      $\theta = \phi(\mathcal{L})$ 
6      $\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} [q(\mathbf{x}, \dots, \theta)]$  ◀ argumentos requeridos por q dependem de cada estratégia
7      $\mathcal{L}' = \mathcal{L} \cup \{\langle \hat{\mathbf{x}}, o(\hat{\mathbf{x}}) \rangle\}$ 
8      $\mathcal{U}' = \mathcal{U} \setminus \{\hat{\mathbf{x}}\}$ 
9     retorna amostragem( $\mathcal{U}', \mathcal{L}', \phi, q, \zeta - 1$ )

```

2.3.1 Estratégias baseadas em incerteza

Provavelmente a medida de informatividade $Inf(\mathbf{x})$ mais simples para a decisão de quando se deva selecionar um exemplo \mathbf{x} da reserva \mathcal{U} , ou um grupo deles, é a máxima probabilidade a posteriori dada por um modelo probabilístico θ (LEWIS, 1995) - conforme Equação 2.9.

$$P_{\theta max}(\mathbf{x}) = \max_{y \in Y} P_{\theta}(\mathbf{y} | \mathbf{x}) \quad (2.9)$$

Classificadores não probabilísticos e com saídas numéricas f podem simular uma distribuição de probabilidades por meio da aplicação da função sigmoide logística g , conforme Equação 2.10.

$$P_{\theta}(y_o = 1 | \mathbf{x}) = \frac{g(f_o(\mathbf{x}))}{\sum_{1 \leq p \leq |Y|} g(f_p(\mathbf{x}))} \quad (2.10)$$

2.3.1.1 Amostragem por incerteza

A estratégia de *amostragem por incerteza* (*Unc*) consiste em consultar o exemplo mais informativo $\hat{\mathbf{x}}$, ou seja, aquele com o maior valor para $1 - P_{\theta max}(\mathbf{x})$ - conforme Equação

2.11. Seu objetivo é explorar a fronteira de decisão no espaço de exemplos. A complexidade dessa estratégia é $\mathcal{O}(1)$ - equivalente a apenas um exemplo aprendido por consulta.

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x} \in \mathcal{U}} [1 - P_{\theta_{\max}}(\mathbf{x})] \quad (2.11)$$

2.3.1.2 Amostragem por margem ou entropia

Em problemas multiclasse, o menor $P_{\theta_{\max}}(\mathbf{x})$ pode não indicar o exemplo mais controverso. Por exemplo, na reserva, um exemplo pode ter duas ou mais classes igualmente prováveis. Essa evidente incerteza inerente ao exemplo não é refletida pelo valor de $P_{\theta_{\max}}(\mathbf{x})$. A estratégia baseada na *margem de incerteza* (Mar), cuja medida M de informatividade é apresentada na Equação 2.12, evita esse problema utilizando o valor da diferença entre as duas maiores probabilidades - o subscrito θ é subentendido para P e a classe predita $\hat{y}(\mathbf{x})$.

$$M_{\theta}(\mathbf{x}) = 1 - \{P[\hat{y}(\mathbf{x}) | \mathbf{x}] - \max_{y \in Y \setminus \{\hat{y}(\mathbf{x})\}} P(y | \mathbf{x})\} \quad (2.12)$$

Outra possibilidade é a estratégia de entropia normalizada (Ent), em que $\text{Inf}_{\theta} = E_{\theta}$ (LEWIN *et al.*, 2004). A medida E_{θ} é dada pela Equação 2.14, que é baseada na medida de entropia $E(\mathbf{v})$ para um dado vetor \mathbf{v} , apresentada na Equação 2.13.

$$E(\mathbf{v}) = -\log^{-1}(\dim \mathbf{v}) \sum_{1 \leq i \leq \dim \mathbf{v}} v_i \log v_i \quad (2.13)$$

$$E_{\theta}(\mathbf{x}) = E[\mathbf{P}_{\sim \theta}(\mathbf{x})] \quad (2.14)$$

2.3.1.3 Margem simples

A estratégia de amostragem por margem pode ser estendida para um espaço de parâmetros \mathcal{F} , resultante da transformação do espaço de atributos por uma função núcleo. Essa foi a abordagem de Tong e Koller (2001) para *Support Vector Machines* (SVM) - (HEARST *et al.*, 1998); cuja principal variante é chamada *margem simples* (SVMsim). Ela seleciona o exemplo mais próximo do hiperplano (fronteira de decisão) que divide linearmente o espaço \mathcal{F} .

2.3.2 Estratégias (não-agnósticas) ponderadas por densidade

As estratégias baseadas em densidade ponderam cada exemplo de acordo com a densidade de sua vizinhança. Sua ideia principal é consultar exemplos representativos da distribuição dos dados e evitar pontos aberrantes. Isso é feito por meio do cálculo da medida de densidade de informação (ID). Ela atribui diferentes pesos à medida de informatividade. Cada exemplo tem um peso que corresponde a um valor de densidade que, de acordo com a Equação 2.15, é mais intenso quando a distância é curta para muitos vizinhos. Assim, os pesos variam conforme o nível de concentração de exemplos não

rotulados no entorno do exemplo \mathbf{x} sob análise (SETTLES, 2008), onde α é um parâmetro de ajuste da importância relativa da densidade. Qualquer medida de similaridade (sim) e de informatividade Inf_θ podem ser adotadas. Assim, a *estratégia ponderada pela densidade* (DW)⁴ consiste em consultar o exemplo com a maior ID_θ .

$$\text{ID}_\theta(\mathbf{x}) = \text{Inf}_\theta(\mathbf{x}) \left(\frac{1}{|\mathcal{U}|} \sum_{\mathbf{u} \in \mathcal{U}} \text{sim}(\mathbf{x}, \mathbf{u}) \right)^\alpha \quad (2.15)$$

Um desdobramento natural é a *estratégia baseada na utilidade de treinamento* (TU)⁵. Ela faz uso da densidade de informação inversamente ponderada pela concentração de exemplos rotulados $\text{ID}_{\text{TU}_\theta}$, resultando num afastamento das regiões mais consultadas anteriormente (SETTLES, 2010; FUJII *et al.*, 1998) - segundo a Equação 2.16, onde δ é um parâmetro de ajuste da importância relativa da densidade dos exemplos rotulados.

$$\text{ID}_{\text{TU}_\theta}(\mathbf{x}) = \text{ID}_\theta(\mathbf{x}) \left(\frac{1}{|\mathcal{L}|} \sum_{\mathbf{l} \in \mathcal{L}} \text{sim}(\mathbf{x}, \mathbf{l}) \right)^{-\delta} \quad (2.16)$$

Neste trabalho, $\text{Inf}_\theta = M$ e as principais medidas de distância $d(\mathbf{x}, \mathbf{u})$ adotadas são a euclidiana e a de Manhattan. Elas são transformadas em medidas de similaridade $\text{sim}(\mathbf{x}, \mathbf{u})$ pela fórmula 2.17. A ordem de complexidade é $\mathcal{O}(1)$, se os $|\mathcal{U}|^2$ resultados dos cálculos de similaridade forem, antes do início do processo de rotulação, devidamente armazenados em memória para acesso rápido posterior.

$$\text{sim}(\mathbf{x}, \mathbf{u}) = \frac{1}{1 + d(\mathbf{x}, \mathbf{u})} \quad (2.17)$$

2.3.3 Estratégias agnósticas

Neste documento, são consideradas agnósticas aquelas estratégias que ignoram a fronteira de decisão traçada pelo modelo de classificação (KEARNS; SCHAPIRE; SELLIE, 1994). Nas próximas seções são apresentadas algumas variações dessa abordagem.

2.3.3.1 Amostragem aleatória

A *amostragem aleatória* (Rnd) corresponde ao aprendizado passivo aplicado a apenas uma parte dos exemplos. Não há uma ordem de preferência ou critério para a realização das consultas - apenas a (pseudo)aleatoriedade. Suas principais características são seu viés totalmente exploratório, ou seja, sem focar nenhuma região especial do espaço de exemplos; e, a ausência de aprendiz. Seu custo computacional pode ser considerado nulo.

⁴ [Density Weighted]

⁵ [Training Utility]

2.3.3.2 Busca no espaço de hipóteses

É possível fazer uma amostragem ativa baseada na perspectiva do espaço de hipóteses. A intuição por trás dessa abordagem é o fato dos exemplos mais importantes residirem na região onde as hipóteses se contradizem. Isso equivale a consultar os exemplos que reduziriam o espaço de versões depois de inseridos no conjunto de treinamento. A busca no espaço de hipóteses é feita pelo acompanhamento das hipóteses mais específicas e das mais gerais pertencentes aos conjuntos S e G de todas as hipóteses possíveis, representadas pelos modelos $\theta_S \in S$ e $\theta_G \in G$, respectivamente. Uma característica distintiva de estratégias desse tipo é seu modelo de decisão binário: todos os exemplos controversos são considerados igualmente informativos, podendo ser consultados em qualquer ordem ou em lotes.

A *rede específica/geral (SG-network)*⁶ (COHN; ATLAS; LADNER, 1994), também chamada de CAL (DASGUPTA, 2011) em referência a seus proponentes, é baseada na busca no espaço de hipóteses. Ela foi um dos primeiros algoritmos de aprendizado ativo. Ela faz uma aproximação para ser capaz de induzir apenas o modelo mais específico θ_S e o modelo mais geral θ_G , pois a quantidade de hipóteses possíveis pode ser infinita. A aproximação é feita pela geração ou amostragem de *exemplos de fundo*. Eles são chamados assim, pois seus rótulos são artificialmente gerados de acordo com a meta desejada de treinamento: especificidade (tendência a prever a classe positiva) ou generalidade (tendência a prever a classe negativa). Depois de criados os modelos iniciais, eventuais exemplos que causem desacordo entre θ_S e θ_G são selecionados para consulta. Duas redes *perceptron multicamadas (MLP)*⁷, treinadas com o algoritmo *backpropagation* (HAYKIN; NETWORK, 2004), foram empregadas no trabalho original. Entretanto, outros algoritmos capazes de induzir modelos aptos a lidar com exemplos ponderados também poderiam ter sido usados. A ordem de complexidade é $\mathcal{O}(1)$ - ou $\mathcal{O}(|Y|)$ na adaptação multiclasse que será proposta neste trabalho.

2.3.3.3 Amostragem por agrupamento

O processo de amostragem pode explorar agrupamentos naturais na reserva de exemplos, pois são independentes da existência de rótulos. Essa abordagem é uma alternativa à realização de consultas que enfocam a fronteira de decisão ou o manejo de hipóteses em desacordo, citados anteriormente. Uma importante representante desse tipo de estratégia é *amostragem hierárquica (HS)*⁸ proposta por Dasgupta (2011), baseada em agrupamento hierárquico (MURTAGH, 1983).

O método de agrupamento hierárquico organiza os exemplos numa hierarquia que

⁶ [Specific/General network]

⁷ [MultiLayer Perceptron]

⁸ [Hierarchical Sampling]

pode ser representada por uma árvore. Cada nó folha simboliza um exemplo e cada nó pai representa uma relação de proximidade/similaridade entre seus filhos. Os filhos, por sua vez, podem ser exemplos isolados ou novas relações de parentesco. Qualquer nó tem o potencial de ser visto como um só grupo contenedor de todos os exemplos representados pelos nós folhas descendentes diretos ou indiretos dele na hierarquia. Logo, na árvore como um todo, as diferentes podas possíveis definem diferentes organizações em grupos. A amostragem hierárquica faz uso da árvore para definir a relevância dos exemplos. Eles têm maior probabilidade de serem consultados caso pertençam aos grupos mais impuros e representativos. O grau de impureza é dado pela proporção de exemplos da mesma classe; o grau de representatividade é dado pela quantidade de exemplos em cada grupo. A implementação original do autor, adotada neste trabalho, fez uso do algoritmo de agrupamento chamado *Ward's average linkage method* (JAIN; MURTY; FLYNN, 1999). A ordem de complexidade de cada consulta de HS não é especificada, mas sabe-se o custo da etapa de agrupamento, que é $\mathcal{O}(n^2 \log n)$. Apesar de alto, se comparado às demais estratégias, esse custo pode ser antecipado, analogamente ao processamento prévio de distâncias seguido de armazenamento em memória sugerido previamente na Seção 2.3.2 para DW e TU.

2.3.4 Outras estratégias

Outros tipos de estratégia são apresentados nesta seção.

2.3.4.1 Balanceamento exploração-prospecção

A natureza puramente prospectiva de SVMsim (Seção 2.3.1.3) possui o viés de consultar prioritariamente exemplos da fronteira de decisão, deixando de explorar as demais regiões de \mathcal{F} . Por esse motivo, Osugi, Deng e Scott (2005) propuseram um balanceamento entre prospecção e exploração. A exploração é realizada por um algoritmo que escolhe sempre o exemplo mais distante, chamado *Kernel Farthest First* (KFF). A heurística de adotar os mais distantes primeiro, sem a transformação no espaço de atributos, já havia sido usada previamente no cálculo de agrupamentos aproximadamente ótimos por Hochbaum e Shmoys (1985).

O algoritmo de balanceamento exploração-prospecção (SVMbal) inicia com a aplicação de KFF. Seu primeiro exemplo é aleatório e o segundo é seu par mais distante. Esse par consiste no conjunto inicial que vai ser estendido com exemplos, um a um, da mesma maneira que o segundo foi escolhido. Esse processo exploratório é alternado com SVMsim de acordo com uma probabilidade definida em função da distância entre as predições, para todos os exemplos, antes e depois da última consulta realizada até o momento.

2.3.4.2 Consulta por comitê

Um comitê $\mathcal{C} = \{\theta_1, \theta_2, \dots, \theta_n\}$ é um conjunto de modelos combinados com o objetivo de superar as predições que a princípio seriam feitas por um único modelo isoladamente (DIETTERICH, 2000). Quando é adotado um esquema de votação, o princípio do *teorema do júri de Condorcet* (VALENTINI; MASULLI, 2002) prova que o comitê é superior a um único indivíduo, desde que seus componentes sejam razoavelmente competentes ou, em outras palavras, sejam melhores que um decisor ao acaso. De fato, empiricamente, comitês geralmente aumentam a acurácia preditiva de algoritmos de aprendizado que a princípio seriam usados isoladamente (BAUER; KOHAVI, 1999). Entretanto, é necessário que exista divergência entre os membros (HANSEN; SALAMON, 1990). Portanto, a fonte da capacidade dos comitês de melhorar a acurácia preditiva de classificadores reside na diversidade de vieses de aprendizado. Essa diversidade pode ser induzida por técnicas de subamostragem - como as populares *boosting* (SCHAPIRE, 1990) e *bagging* (BREIMAN, 1996).

Query by bagging e *query by boosting* são dois exemplos de *consulta por comitê* (QBC)⁹ utilizados por Abe e Mamitsuka (1998). Dependendo do tipo de valor de retorno fornecido pelo modelo, diferentes medidas de desacordo podem ser utilizadas. A divergência de Jensen-Shannon (LIN, 1991), por exemplo, é uma medida de teoria da informação que compara distribuições de probabilidade, comumente usada em comitês para avaliar o grau de desacordo entre os membros (MELVILLE; MOONEY, 2004). A divergência não-ponderada de Jensen-Shannon é definida em termos da entropia das distribuições na Equação 2.18.

$$\text{JS}(\mathcal{C}, \mathbf{x}) = E_{\theta_{\mathcal{C}}}(\mathbf{x}) - \sum \langle E_{\theta}(\mathbf{x}) \mid \theta \in \mathcal{C} \rangle \quad (2.18)$$

O modelo geral $\theta_{\mathcal{C}}$ é a agregação de todos os modelos e representa o comitê como um todo. Quanto maior o valor de JS, mais distante os membros estão de um consenso. Assim, o exemplo com o maior valor deve ser consultado primeiro. A complexidade computacional é $\mathcal{O}(1)$, se o comitê for visto como um único algoritmo de aprendizado; ou $\mathcal{O}(|\mathcal{C}|)$, se o número de membros $|\mathcal{C}|$ for considerado.

2.3.4.3 Redução do erro esperado

A *estratégia de redução do erro esperado* (EER)¹⁰ adotada neste trabalho é baseada no *exemplo de redução da entropia* proposto por (GUO; GREINER, 2007). É um método que busca pelo exemplo que mais reduz a entropia na predição geral do modelo para todo o conjunto de dados. Ele considera, assim, implicitamente, a informação sobre eventuais agrupamentos subjacentes, evitando depender apenas dos escassos exemplos rotulados.

⁹ [Query By Committee]

¹⁰ [Expected Error Reduction]

O exemplo $\hat{\mathbf{x}}$ selecionado para consulta e sua classe mais provável, representada pelo vetor preditivo \hat{y} , são obtidos conforme a Equação 2.19, onde O é a função objetivo.

$$\langle \hat{\mathbf{x}}, \hat{y} \rangle = \arg \min_{\langle \mathbf{x}, y \rangle \in \mathcal{U} \times Y} \sum_{\mathbf{u} \in \mathcal{U}} O(\theta_{\mathcal{L} \cup \{\langle \mathbf{u}, y \rangle\}}, \mathbf{x}) \quad (2.19)$$

Após cada consulta, caso a classe real seja diferente da esperada, o método recorre à estratégia de amostragem por incerteza (Seção 2.3.1.1) como medida de contingência. No artigo original, optou-se pela entropia como função objetivo. Outras medidas também podem ser adequadas, dependendo da meta, como a acurácia balanceada ou kappa multiclasse (Seção 4.5.1). A complexidade computacional é $\mathcal{O}(|Y||\mathcal{U}|^2)$.

2.3.4.4 Impacto esperado no modelo

O futuro impacto de um exemplo sobre o modelo é a base de um tipo de estratégia chamado *mudança esperada no modelo*. Esse impacto é uma indicação de sua possível contribuição para o aprendizado. O procedimento proposto por (SETTLES; CRAVEN; RAY, 2007) é chamado *comprimento esperado do gradiente (EGL)*¹¹. Ele se aplica a modelos baseados na técnica de gradiente descendente no espaço \mathcal{F} (HAYKIN; NETWORK, 2004): deve ser escolhido para treinamento o exemplo capaz de contribuir com a descida de maior magnitude na curva de erro. Como o rótulo não é sabido de antemão, usa-se a soma das contribuições de cada rótulo ponderada pelas suas respectivas probabilidades a posteriori.

2.4 Considerações

Neste capítulo, os assuntos que contextualizam esta tese foram revisados: classificação, aprendizado ativo e suas estratégias. Outro assunto relevante por integrar a proposta de *aprendizado meta-ativo* é o *meta-aprendizado*, cuja revisão bibliográfica é resumida no Apêndice C.

A literatura de aprendizado ativo é vasta. Logo, a revisão bibliográfica das estratégias é invariavelmente incompleta. Porém, a variedade de abordagens contidas na presente revisão é suficiente para a finalidade de teste das hipóteses desta tese. Essa diversidade proporcionou diferentes níveis de adequação entre estratégias, algoritmos de aprendizado e conjuntos de dados (conforme será apresentado no Capítulo 5).

Por fim, o Quadro 1 contém uma síntese das estratégias e suas características, conforme descrito a seguir.

- Forma de **busca**: define se a consulta favorece abrangência ou eficiência; se ambas as metas são balanceadas de acordo com um critério simples ou combinadas de maneira alternada; e, se cada consulta se restringe apenas a parte dos dados.

¹¹ [Expected Gradient Length]

- **Aprendiz:** indica se a estratégia tem ou não aprendiz e, caso tenha, se o algoritmo de aprendizado é fixo, específico da estratégia.
- **Dependência** entre consultas: define se o oráculo é necessário durante a amostragem ou pode ser consultado quando todos os exemplos relevantes já tiverem sido definidos pela estratégia.
- Ordem de **complexidade:** custo computacional considerando a quantidade de exemplos a aprender por consulta.

As propostas desta tese são apresentadas no próximo capítulo.

Quadro 1 – Características de cada estratégia.

Estratégia	Busca	Aprendiz	Dependência	Complexidade
Rnd ^a	exploratória aleatória	ausente	nenhuma	dispensa treinamento
HS ^b	balanceada exploratória prospectiva	ausente	total	dispensa treinamento [*]
Mar/Ent ^a /QBC ^c	prospectiva	presente	total	$\mathcal{O}(1)$
DW ^d	prospectiva	presente	total	$\mathcal{O}(1)$
EER ^e	prospectiva	presente	total	$\mathcal{O}(Y \mathcal{U} ^2)$
TU ^f	balanceada exploratória prospectiva	presente	total	$\mathcal{O}(1)$
SGnetwork ^g	limitada explora- tória aleatória	presente	total	$\mathcal{O}(1)$
SVMsim ^h	prospectiva	presente específico	total	$\mathcal{O}(\mathcal{L} \mathcal{U})$
EGL ⁱ	prospectiva	presente específico	total	$\mathcal{O}(Y \mathcal{U})$
SVMbal ^j	combinada exploratória prospectiva	presente específico	total	$\mathcal{O}(\mathcal{L} \mathcal{U})$

^{*} Na ausência de aprendiz não há treinamento, porém HS tem sua própria complexidade a ser considerada.

^a Amostragem aleatória, por margem ou entropia (SETTLES, 2012).

^b Amostragem hierárquica (DASGUPTA, 2011).

^c Consulta por comitê (ABE; MAMITSUKA, 1998).

^d Amostragem ponderada por densidade (SETTLES, 2008).

^e Redução esperada do erro (GUO; GREINER, 2007).

^f Amostragem ponderada por densidade e utilidade de treinamento (SETTLES, 2010; FUJII *et al.*, 1998).

^g SGnetwork (COHN; ATLAS; LADNER, 1994).

^h Margem simples (TONG; KOLLER, 2001).

ⁱ Comprimento esperado do gradiente (SETTLES; CRAVEN; RAY, 2007).

^j Balanceamento exploração-prospecção (OSUGI; DENG; SCOTT, 2005).

Propostas

*Fast is busy, controlling,
aggressive...analytical...
Slow is...calm,
careful...intuitive...patient...
[The ideal] philosophy can be
summed up in a single word:
balance.*

Carl Honoré¹

O uso efetivo do aprendizado ativo depende de escolhas adequadas diante da diversidade existente de estratégias e de algoritmos de aprendizado de máquina. Nesta tese, são empreendidas investigações em três níveis sobre esse problema:

- análise comparativa de estratégias de amostragem ativa;
- controle da influência do aprendiz durante o processo de amostragem ativa; e,
- recomendação automática de estratégias.

O primeiro nível de investigação é essencialmente qualitativo e será apresentado por meio de experimentos no Capítulo 5. Os dois últimos níveis resultaram na proposta de duas abordagens automáticas, respectivamente: estratégia híbrida baseada na inibição do aprendiz - apresentada na Seção 3.1 juntamente com sua versão puramente agnóstica (Seção 2.3.3); e, aprendizado meta-ativo - apresentado na Seção 3.2. A Seção 3.1 também contém a proposta de adaptação da estratégia *SG-network* a problemas com mais de duas classes.

¹ “Rápido é ocupado, controlador, agressivo... analítico... Lento é... tranquilo, cuidadoso... intuitivo... paciente... A filosofia [ideal] pode ser resumida em uma única palavra: equilíbrio. ” - [Honoré \(2006\)](#), sobre o culto à velocidade e o benefício da lentidão para uma vida com mais significado.

3.1 Estratégias

Em certos casos, o problema da escolha do algoritmo de aprendizado mencionado previamente na Seção 1.1.2 pode ser contornado pela adoção de uma estratégia sem aprendiz. Uma aplicação hipotética seria o treinamento de um sistema de reconhecimento de conteúdo impróprio, onde o especialista em aprendizado de máquina estivesse interessado em disponibilizar ao público somente conteúdos que tenham sido aprovados previamente por modelos preditivos suficientemente treinados. Nesse caso, apenas o modelo final é de interesse. Logo, pode-se empregar uma estratégia sem aprendiz, como a baseada em densidade proposta a seguir, na Seção 3.1.1. No contexto desta tese, sua finalidade é facilitar a identificação do efeito da ausência de aprendiz.

Por outro lado, a presença do aprendiz durante o aprendizado pode ser útil mesmo em aplicações que o dispensem. Na Seção 3.1.2, uma estratégia híbrida, que alterna períodos de presença e ausência de aprendiz, é proposta.

3.1.1 Ponderada por densidade sem aprendiz

Resumidamente, a estratégia TU (baseada em densidade e *Training Utility* - Seção 2.3.2) pondera a medida de informatividade Inf_θ de cada exemplo, afastando a possibilidade de consulta daqueles próximos aos já rotulados e aumentando as chances daqueles situados nas regiões com maior concentração de exemplos não rotulados - conforme depreende-se das equações 2.15 e 2.16. Esse efeito é mais forte quando a densidade é muito alta, pois ela tende a ser a componente dominante da densidade de informação ID_θ . Isso faz com que Inf_θ tenda a perder relevância.

Dado que a medida Inf_θ já é explorada isoladamente pela estratégia Mar (baseada em incerteza - Seção 2.3.1), resta a possibilidade de exploração isolada da parte referente à densidade. Assim, a fórmula da função-critério de consulta da nova estratégia (SANTOS; CARVALHO, 2015b), chamada *amostragem agnóstica ponderada por densidade (ATU)*² é apresentada na Equação 3.1.

$$\text{ID}_{\text{ATU}}(\mathbf{x}) = \left(\frac{1}{|\mathcal{U}|} \sum_{\mathbf{u} \in \mathcal{U}} \text{sim}(\mathbf{x}, \mathbf{u}) \right)^\alpha \left(\frac{1}{|\mathcal{L}|} \sum_{\mathbf{l} \in \mathcal{L}} \text{sim}(\mathbf{x}, \mathbf{l}) \right)^{-\delta} \quad (3.1)$$

Onde os parâmetros α e δ mantêm os significados dados na definição de TU. Ela é chamada *agnóstica* com base no uso original do termo, que propõe chamar agnóstico o aprendizado que não faz suposições sobre a distribuição dos dados³ (KEARNS; SCHAPIRE; SELLIE, 1994). Seu funcionamento segue o Algoritmo 2, geral, apresentado previamente no Capítulo 2, onde a função-critério passa a ser definida da seguinte forma: $q = \text{ID}_{\text{ATU}}$.

² [density-weighted Agnostic sampling (Training Utility)]

³ Dasgupta, Hsu e Monteleoni (2008) apresentam outro ponto de vista, definindo como *aprendiz ativo agnóstico* aquele que não assume a existência de uma fronteira de decisão perfeita.

Conforme já mencionado, uma vantagem de ATU com relação à maioria das estratégias, é a possibilidade de adiamento da escolha do algoritmo de aprendizado em aplicações que dispensem a disponibilidade do modelo preditivo durante o aprendizado. Em determinadas situações, também pode ser útil a possibilidade de conhecer toda a sequência de consultas antes de qualquer interação com o oráculo. Como, frequentemente, o limite de consultas é uma quantidade conhecida de antemão (SETTLES, 2010), elas podem ser repartidas entre vários oráculos que podem trabalhar de forma independente entre si. Diferentemente, essa possibilidade de paralelização não existe plenamente, por exemplo, na estratégia HS (baseada em agrupamento hierárquico - Seção 2.3.3). Mesmo sem aprendiz, esta é dependente da revelação dos rótulos reais para ser capaz de determinar a pureza de cada grupo na hierarquia. Outra vantagem de ATU, partilhada por poucas estratégias, como Rnd (amostragem aleatória - Seção 2.3.3), é a inexistência de tempo de espera entre as consultas.

3.1.2 Híbrida ponderada por densidade

A abordagem sem aprendiz baseada em densidade (ATU - Seção 3.1.1) é de natureza exploratória, pois enfoca as regiões mais desconhecidas (sem rótulos) do espaço de exemplos. Devido à ausência de aprendiz, não há uma fronteira de decisão a considerar. Logo, não é possível realizar consultas de caráter prospectivo que ataquem diretamente os exemplos mais críticos de forma análoga a uma busca binária. Entretanto, ambas abordagens, prospectiva e exploratória, têm limitações. Se, por um lado, a premissa de fronteira única ou perfeitamente separável de uma busca binária não pode ser garantida; por outro lado, a amostragem puramente exploratória pode ser dispendiosa, pois se assemelha a uma busca exaustiva. Assim, é intuitivo supor que o equilíbrio entre exploração e prospecção possa vir a ser proveitoso. A estratégia TU, por exemplo, combina ambas as abordagens em sua fórmula, mas não permite que elas ajam separadamente.

A Figura 4 ilustra, para o conjunto de dados *Banana* (BACHE; LICHMAN, 2013), a evolução da função-critério de TU e suas respectivas componentes de exploração e prospecção, correspondentes às estratégias ATU e Mar, respectivamente. As curvas ATU e Mar são constituídas pelos valores que essas estratégias atribuiriam ao exemplo selecionado por TU. Os modelos foram gerados pelo algoritmo *Naive Bayes* (LEWIS, 1998). Nota-se uma alta correspondência entre as curvas Mar e TU, especialmente nas primeiras 50 consultas. Essa dominância da componente Mar na determinação da função-critério TU impede que a característica exploratória de ATU influencie significativamente na escolha dos exemplos. A curva da componente ATU é descendente, indicando a progressiva redução de áreas desconhecidas no espaço de exemplos, porém com oscilações devido à ordem imposta por TU às consultas.

Caso os exemplos fossem escolhidos de acordo com o critério de ATU, as curvas

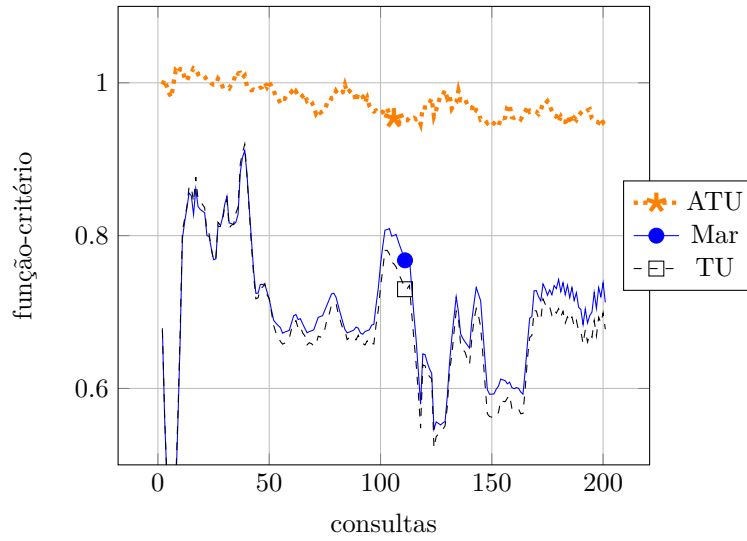


Figura 4 – Curvas de valores das funções-critério para os exemplos selecionados por TU.

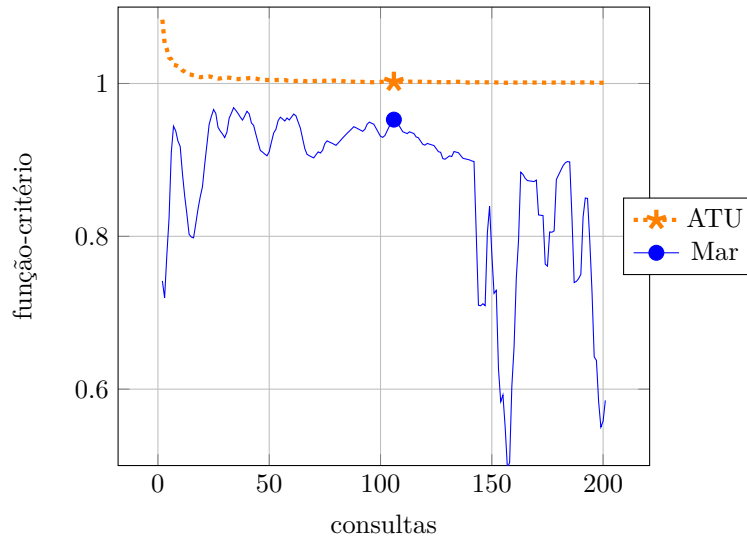


Figura 5 – Curvas de valores das funções-critério para os exemplos selecionados por ATU.

seriam como aquelas apresentadas na Figura 5. A curva ATU é monotônica decrescente conforme esperado, pois cada exemplo consultado (e retirado da reserva) tem o valor de densidade mais alto naquele momento. Em divergência com a curva da componente Mar na figura anterior, o aprendiz se manteve na região de incerteza por mais tempo e atingiu valores acima de 0,9 em quase todas as 140 primeiras consultas. Pode-se supor que essa maior incidência de exemplos incertos se deva à descoberta antecipada de novos segmentos da fronteira de decisão em regiões que permaneceriam inexploradas caso a incerteza do aprendiz (componente Mar) influenciasse as consultas. Assim, apesar de ser um resultado para um conjunto de dados específico, existe a possibilidade de que outros conjuntos possam se beneficiar desse potencial de antecipação da consulta de exemplos da parte desconhecida da fronteira.

Propõe-se, aqui, uma nova estratégia ([SANTOS; CARVALHO, 2015b](#)) - chamada

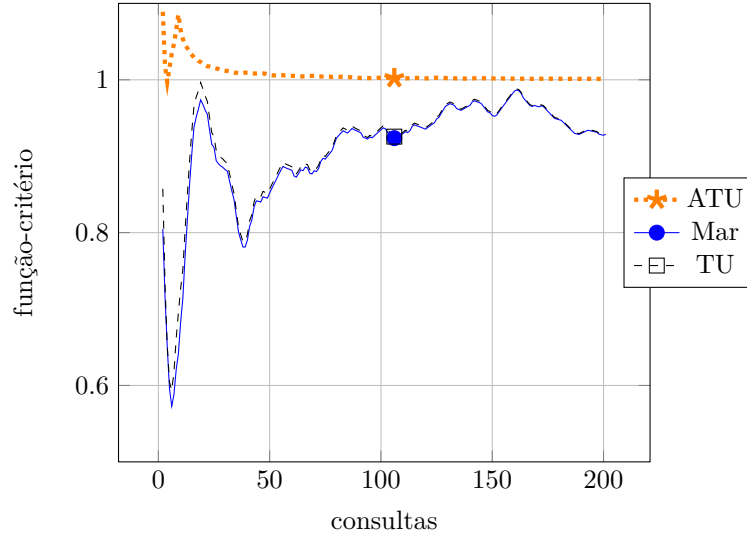


Figura 6 – Curvas de valores das funções-critério para os exemplos selecionados por HTU.

*amostragem híbrida ponderada por densidade (HTU)*⁴. Ela procura alternar as estratégias TU e ATU de acordo com o nível de contribuição corrente da componente exploratória em TU. O nível de contribuição é estimado por meio do cálculo da correlação de Pearson ρ (ROSS, 2000) entre as funções-critério de TU e sua componente não-agnóstica (Mar) para toda a reserva de exemplos. Um valor acima de 0,8, ou seja, representando uma correlação *muito forte* de acordo com Evans (1996), indica uma baixa contribuição da componente exploratória. Outra interpretação é que, havendo baixa correlação entre os níveis de relevância atribuídos aos exemplos, o aprendiz esteja interferindo excessivamente no critério de consulta - possivelmente por estar ainda no estágio inicial do aprendizado e ser incapaz de realizar previsões estáveis ou confiáveis. Ainda outra interpretação, mais intuitiva, é que a exploração (ATU) prossegue até que as consultas se aproximem tanto da fronteira a ponto de serem similares aos exemplos mais incertos, isto é, consultas similares às aquelas que seriam feitas por TU. A partir desse ponto, ou enquanto a correlação se mantivesse elevada, poder-se-ia considerar que a etapa de exploração não é mais necessária. Isso é observável na Figura 6, resultante da sequência de consultas gerada por HTU. Exceto pela mesma brusca oscilação inicial, provavelmente devido aos primeiros segmentos de fronteira descobertos, a situação é, de certa forma, invertida com relação à situação da Figura 4: Mar e TU passam a coincidir na segunda metade do gráfico e o nível de incerteza se mantém elevado. Como resultado, Mar apresenta a curva mais elevada dentre as três figuras. Isso indica que exemplos mais incertos e mais informativos⁵ foram consultados.

O valor limite de correlação ρ_{limite} escolhido para os experimentos deste documento,

⁴ [Density-weighted Hybrid Sampling (Training Utility)]

⁵ A única diferença entre HTU e TU é a possibilidade de ausência de aprendiz da primeira. Consequentemente, se todos os segmentos de fronteira já estivessem revelados, os valores de incerteza apresentados por TU seriam o limite teórico para os valores de incerteza apresentados por HTU (e ATU). Logo, a elevação que HTU causou nos valores de incerteza se dá devido ao descobrimento de novos segmentos de fronteira. Trata-se, portanto, de exemplos mais informativos.

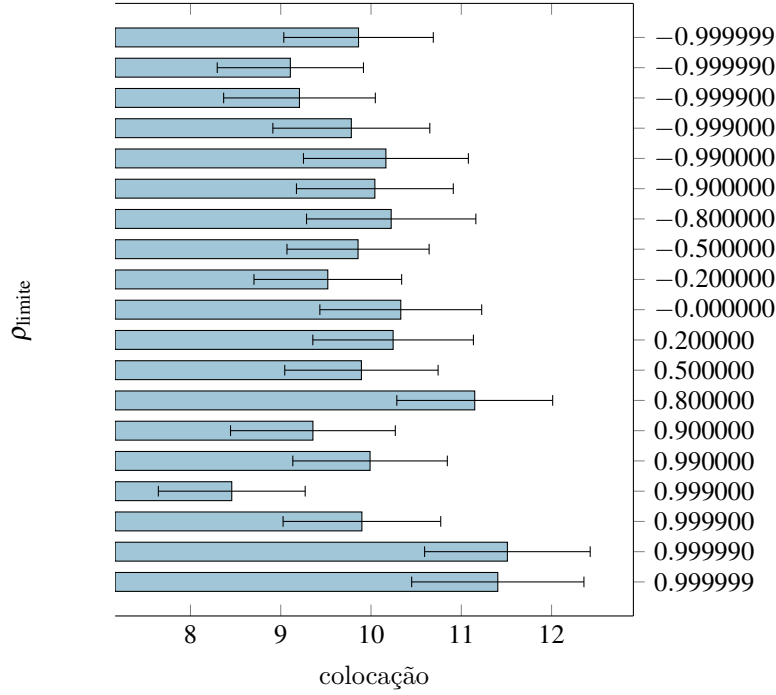


Figura 7 – Colocação média para diferentes limites de correlação. O intervalo de confiança é dado para $p = 0,05$ e $n = 120$.

inclusive o descrito anteriormente, foi 0,999. Esse valor foi definido conforme explicado a seguir. Um subconjunto de 1000 exemplos foi extraído aleatoriamente de cada conjunto de dados pertencente a um grupo de 12 conjuntos que não foram empregados na fase experimental da presente pesquisa: *multiple-features*, *appendicitis*, *micro-mass-pure-spectra*, *hayes-roth*, *cnae-9*, *breast-tissue-4class*, *qualitative-bankruptcy*, *fertility-diagnosis*, *acute-inflammations-urinary*, *digits2*, *lsvt-voice-rehabilitation* e *micro-mass-mixed-spectra* (BACHE; LICHMAN, 2013). O valor médio do índice kappa (para problemas multiclasse - (FLEISS; LEVIN; PAIK, 2013; SHAH, 2011)) foi adotado como indicador de desempenho numa validação cruzada em 10 partes. O intervalo de confiança, consequentemente, é dado para $p = 0,05$ e $n = 120$. Ele é dado pela Equação 3.2 em função da matriz de confusão R após 100 consultas - onde $e_i(R)$ e $p_i(R)$ são as componentes dos vetores de classes esperadas e preditas de acordo com R , respectivamente (Seção 2.1.4). O algoritmo empregado foi o *Random Forest* com 10 árvores (BREIMAN, 2001).

$$\kappa(R) = \left(r(R) - \frac{p_i(R) \cdot e_i(R)}{\sum_{1 \leq i \leq |Y|} e_i(R)} \right) \left(1 - \frac{p_i(R) \cdot e_i(R)}{\sum_{1 \leq i \leq |Y|} e_i(R)} \right)^{-1} \quad (3.2)$$

A Figura 7 contém a colocação média para valores de ρ_{limite} em diferentes valores e ordens de magnitude no intervalo $[-0,999999; 0,999999]$. Os valores de colocação se encontram entre 8 e 12, sendo 8,46 a melhor. Assim, optou-se nesta tese pela sua correlação correspondente, $\rho_{\text{limite}} = 0,999$. O autor desta pesquisa não considera esse valor crítico. Ele foi adotado como prova de conceito para o estudo do efeito do controle da presença do

aprendiz na curva de aprendizado e no desempenho preditivo.

A estratégia HTU é descrita pelo Algoritmo 3. O efeito esperado é que ela iniba o aprendiz nos períodos em que a influência dele possa ser negativa. Esse período de inibição pode se estender por todo o aprendizado, no caso de algoritmos que gerem modelos impróprios para prospecção ou mesmo completamente inadequados para o dado problema.

Algoritmo 3: Estratégia híbrida ponderada por densidade.

Entrada:

\mathcal{U} - reserva de exemplos

\mathcal{L} - conjunto inicial de exemplos rotulados

ζ - orçamento (quantidade de exemplos a rotular)

ρ_{limite} - limite de correlação entre medidas de Mar e TU

Resultado:

\mathcal{L}' - conjunto final de exemplos rotulados

```

1 função HTU( $\mathcal{U}$ ,  $\mathcal{L}$ ,  $\zeta$ )
2   se  $\zeta = 0$  então
3     retorna  $\mathcal{L}$ 
4   senão
5      $\theta = \phi(\mathcal{L})$ 
6      $\mu_{M_\theta} = \frac{1}{|\mathcal{U}|} \sum_{\mathbf{x} \in \mathcal{U}} M_\theta(\mathbf{x})$ 
7      $\sigma_{M_\theta} = \frac{1}{|\mathcal{U}|} \sum_{\mathbf{x} \in \mathcal{U}} (M_\theta(\mathbf{x}) - \mu_{M_\theta})^2$ 
8      $\mu_{\text{ID}_{\text{TU}_\theta}} = \frac{1}{|\mathcal{U}|} \sum_{\mathbf{x} \in \mathcal{U}} \text{ID}_{\text{TU}_\theta}(\mathbf{x})$ 
9      $\sigma_{\text{ID}_{\text{TU}_\theta}} = \frac{1}{|\mathcal{U}|} \sum_{\mathbf{x} \in \mathcal{U}} (\text{ID}_{\text{TU}_\theta}(\mathbf{x}) - \mu_{\text{ID}_{\text{TU}_\theta}})^2$ 
10     $\rho = (\sigma_{M_\theta}^2 \cdot \sigma_{\text{ID}_{\text{TU}_\theta}}^2)^{-\frac{1}{2}} \sum_{\mathbf{x} \in \mathcal{U}} (M_\theta(\mathbf{x}) - \mu_{M_\theta})(\text{ID}_{\text{TU}_\theta}(\mathbf{x}) - \mu_{\text{ID}_{\text{TU}_\theta}}) \triangleleft \text{correlação Mar-TU}$ 
11    se  $\rho < \rho_{\text{limite}}$  então
12       $\hat{\mathbf{x}} \leftarrow \arg \max_{\mathbf{x}} [\text{ID}_{\text{ATU}}(\mathbf{x})]$ 
13    senão
14       $\hat{\mathbf{x}} \leftarrow \arg \max_{\mathbf{x}} [\text{ID}_{\text{TU}_\theta}(\mathbf{x})]$ 
15     $\mathcal{L}' \leftarrow \mathcal{L} \cup \{\langle \hat{\mathbf{x}}, o(\hat{\mathbf{x}}) \rangle\}$ 
16     $\mathcal{U}' \leftarrow \mathcal{U} \setminus \{\hat{\mathbf{x}}\}$ 
17    retorna HTU( $\mathcal{U}'$ ,  $\mathcal{L}'$ ,  $\zeta - 1$ )

```

Além das duas estratégias propostas, algumas estratégias precisaram ser adaptadas para problemas multiclasse. Essas adaptações são apresentadas na Seção 3.1.3.

3.1.3 Adaptações multiclasse

Algumas estratégias citadas previamente no Capítulo 2 são originalmente voltadas a problemas binários. Por esse motivo, adaptações precisaram ser feitas para que o conjunto de estratégias pudesse ser aplicado a conjuntos de dados com mais de duas

classes.

3.1.3.1 Busca no espaço de hipóteses multiclasse

Conforme explicado na Seção 2.3.3.2, é possível fazer uma amostragem ativa dentro da perspectiva do espaço de hipóteses. Uma adaptação chamada *SGmulti* foi proposta para contornar a limitação de aplicabilidade da abordagem original apenas a problemas binários (SANTOS; CARVALHO, 2014a). A abordagem original (*SG-network*) depende da existência de uma única classe *positiva*. São definidos um modelo mais geral θ_G e um modelo mais específico θ_S . Cada modelo simula a especificidade ou a generalidade. A simulação se dá por meio de exemplos de fundo artificiais da classe oposta (SETTLES, 2012).

Na presente proposta de adaptação, a existência de mais de duas classes exige mais que dois modelos. Cada classe $c \in Y$ tem um modelo associado θ_c , onde c corresponde à classe negativa. O conjunto $Y \setminus \{c\}$ corresponde ao que poderia ser chamado conjunto de “classes positivas”. Cada modelo θ_c é induzido com um conjunto de exemplos ponderados $\tilde{\mathcal{L}}_c$, especificado na Equação 3.3.

$$\tilde{\mathcal{L}}_c = \tilde{\mathcal{L}} \cup \{\langle \mathbf{x}, c, w \rangle \mid \mathbf{x} \in \mathcal{U}, w = (|Y||\mathcal{U}|)^{-1}\} \quad (3.3)$$

Resumidamente, a equação define cada conjunto $\tilde{\mathcal{L}}_c$ como $\tilde{\mathcal{L}}$ acrescido dos exemplos de \mathcal{U} artificialmente rotulados com a classe c . Tais exemplos são ditos artificiais porque a classe é fictícia, ou seja, não é dada por um rótulo atribuído pelo oráculo. Tais exemplos são incorporados ao conjunto com um peso w diminuto ($w \ll 1$), conforme sugerido recentemente na literatura Settles (2012). O valor exato de w foi definido de forma que a contribuição de todos os exemplos artificiais nunca superasse a contribuição de um único exemplo real: $\sum_{c \in Y} |\tilde{\mathcal{L}}_c \setminus \tilde{\mathcal{L}}| = 1$. Essa medida evita que os exemplos de fundo se sobreponham aos exemplos reais no início do processo de rotulação. Essa não sobreposição é importante porque, no início, os exemplos reais rotulados são escassos.

O conjunto $\tilde{\mathcal{L}}$ de todos os conjuntos $\tilde{\mathcal{L}}_c$ é definido pela Equação 3.4.

$$\tilde{\mathcal{L}} = \{\tilde{\mathcal{L}}_c \mid c \in Y\} \quad (3.4)$$

A função de predição \hat{y}_θ retorna a classe mais provável de um dado exemplo \mathbf{x} de acordo com o modelo θ_c . É possível encontrar um exemplo $\hat{\mathbf{x}}$ para o qual não haja consenso comparando os retornos de todas as diferentes funções de predição. A sequência⁶ D de exemplos em desacordo é dada pela Equação 3.5.

$$D = \langle \mathbf{x} \in \mathcal{U} \mid c, d \in Y, \hat{y}_{\theta_c}(\mathbf{x}) \neq \hat{y}_{\theta_d}(\mathbf{x}) \rangle \quad (3.5)$$

⁶ Sequências são mais convenientes que conjuntos para a descrição do procedimento no Algoritmo 4.

Exemplos da região de desacordo ($\hat{\mathbf{x}} \in D$) são consultados um a um. Após cada consulta, seu exemplo de fundo correspondente em todos os conjuntos de treinamento recebe o rótulo verdadeiro e peso integral, gerando os novos conjuntos $\tilde{\mathcal{L}}'_c, c \in Y$ - conforme a Equação 3.6.

$$\tilde{\mathcal{L}}'_c = (\tilde{\mathcal{L}}_c \setminus \{\langle \hat{\mathbf{x}}, c, w \rangle\}) \cup \{\langle \hat{\mathbf{x}}, o(\hat{\mathbf{x}}), 1 \rangle\} \quad (3.6)$$

Depois de cada consulta, os modelos θ_c são atualizados com os novos conjuntos $\tilde{\mathcal{L}}'_c, c \in Y$. Dado o peso diminuto do exemplo artificial que foi substituído, essa atualização equivale a um treinamento incremental apenas com o exemplo $\langle \hat{\mathbf{x}}, o(\hat{\mathbf{x}}), 1 \rangle$ - resultando numa ordem de complexidade de $\mathcal{O}(|Y|)$, pois requer treinamento em $|Y|$ exemplos por consulta, um para cada modelo.

Por fim, a estratégia SGmulti é descrita pelo Algoritmo 4.

Algoritmo 4: Estratégia SGmulti.

Entrada:
 $\tilde{\mathcal{L}}$ - conjunto dos conjuntos rotulados (parcialmente) artificiais
 ζ - orçamento (quantidade de exemplos a rotular)
Resultado:
 \mathcal{L}' - conjunto final de exemplos rotulados

```

1 função SGmulti( $\mathcal{U}, \tilde{\mathcal{L}}, \zeta$ )
2   se  $\zeta = 0$  então
3     retorna  $\tilde{\mathcal{L}}$ 
4   senão
5      $\{\tilde{\mathcal{L}}_1, \tilde{\mathcal{L}}_2, \dots, \tilde{\mathcal{L}}_{|Y|}\} = \tilde{\mathcal{L}}$ 
6      $\theta_c = \phi(\tilde{\mathcal{L}}_c), 1 \leq c \leq |Y|$ 
7      $D = \{\mathbf{x} \in \mathcal{U} \mid i, j \in Y, \hat{y}_{\theta_i}(\mathbf{x}) \neq \hat{y}_{\theta_j}(\mathbf{x})\}_{1 \leq i, j \leq |Y|}$ 
8      $i \sim U(1, |D|)$  ◀ sorteia um índice da sequência D (U: distribuição uniforme)
9      $\hat{\mathbf{x}} = D_i$ 
10     $\hat{\mathcal{X}} = \{\langle \hat{\mathbf{x}}, o(\hat{\mathbf{x}}), 1 \rangle\}$  ◀ conjunto unitário contendo o exemplo real novo (rotulado)
11     $\tilde{\mathcal{L}}' = \{(\tilde{\mathcal{L}}_c \setminus \{\langle \hat{\mathbf{x}}, c, w \rangle\}) \cup \hat{\mathcal{X}} \mid c \in Y\}$ 
12     $\mathcal{U}' = \mathcal{U} \setminus \{\hat{\mathbf{x}}\}$ 
13    retorna SGmulti( $\mathcal{U}', \tilde{\mathcal{L}}', \zeta - 1$ )
```

3.1.3.2 Outras adaptações

As versões multiclasse de SVMsim e SVMbal foram implementadas baseadas nos programas originais dos autores (Seção 2.3.4). Optou-se pela forma *um-contra-muitos*: uma instância da estratégia de amostragem ativa independente para cada subproblema binário. Visando a distribuição equitativa das consultas entre as instâncias, cada nova consulta era feita por uma instância diferente, em forma de rodízio. Entretanto, o autor deste trabalho optou por descartar essas duas estratégias, dada a sua dependência de um algoritmo de aprendizado específico e seu desempenho excessivamente baixo em testes

iniciais. Supõe-se que a causa desse desempenho tenha sido a forma de adaptação adotada. Logo, ela provavelmente não corresponderia a uma opção realmente representativa ou relevante da literatura de estratégias de amostragem ativa. Trata-se de um tópico que requer maior aprofundamento para a definição da melhor maneira de adaptação e prosseguimento com a comparação com as demais estratégias. Por outro lado, as abordagens podem ser consideradas parcialmente representadas, pois SVMsim é análoga à estratégia Unc (e Mar), por exemplo.

3.2 Aprendizado meta-ativo

A proposta de estudo sobre recomendação automática desta tese visa, primariamente, a escolha do algoritmo de aprendizado mais adequado para um novo conjunto de dados, ainda não rotulado. O sistema baseia-se em conhecimento prévio, supostamente adquirido por meio de experimentos anteriores realizados pelo especialista em aprendizado de máquina. Assim, supõe-se que exista uma coleção de conjuntos já conhecidos e devidamente rotulados.

Outras variações também foram investigadas dentro das mesmas premissas, como a recomendação de estratégias e pares estratégia-algoritmo, entre outras. Todas essas possibilidades de recomendação são implementadas por meio da técnica de meta-aprendizado - apresentada mais detalhadamente no Apêndice C. A técnica de meta-aprendizado tem sido utilizada na literatura para sugerir algoritmos e/ou parâmetros mais apropriados, ou um ranqueamento das opções disponíveis, para um novo conjunto de dados (BRAZDIL *et al.*, 2009).

A abordagem geral da nova proposta é chamada *aprendizado meta-ativo*, pois a recomendação da forma de aprendizado ou consulta ocorre no nível meta; enquanto que os algoritmos de aprendizado e as estratégias de amostragem ativa situam-se no nível base. Nesse caso, idealmente, o meta-aprendizado poderia ser visto como uma ferramenta de suporte ao aprendizado ativo: as recomendações automáticas seriam fornecidas em busca da melhor forma de realizar consultas. É importante diferenciar esse novo esquema daquele presente no *meta-aprendizado ativo* (SOUSA *et al.*, 2013), em que, inversamente, o aprendizado ativo auxilia no processo de meta-aprendizado. A finalidade deste é distinta: procura-se reduzir a quantidade necessária de rótulos no nível meta.

O sistema de recomendação de algoritmos de aprendizado baseado em meta-aprendizado, ilustrado na Figura 8, é composto por sete passos principais:

1. organização de uma coleção variada de conjuntos de dados rotulados;
2. definição de uma estratégia;

3. determinação experimental do melhor algoritmo para cada conjunto via simulação⁷ de aplicação da estratégia;
4. caracterização de todos os conjuntos;
5. geração de um metaexemplo para cada conjunto - a classe é dada pelo nome do melhor algoritmo;
6. treinamento de um meta-aprendiz com os metaexemplos do passo anterior (optou-se inicialmente por um comitê de árvores de decisão nos experimentos - Capítulo 4); e,
7. consulta ao metamodelo induzido pelo meta-aprendiz para classificar o novo conjunto de dados e obter a recomendação.

No Passo 3, a determinação experimental do melhor algoritmo é feita conforme o procedimento apresentado no Algoritmo 5. A simulação de aplicação da estratégia deve, preferencialmente, manter as mesmas condições esperadas para o problema alvo, por exemplo, o orçamento ζ e a quantidade inicial de exemplos rotulados $|\mathcal{L}|$. Uma simulação com orçamento diferente daquele pretendido para o novo conjunto de dados pode gerar informações imprecisas, dado que algoritmos com bom desempenho em treinamentos com poucos dados podem ter baixo desempenho com muitos dados (ATTENBERG; PROVOST, 2010a; PERLICH; PROVOST; SIMONOFF, 2003).

Algoritmo 5: Identificação do melhor algoritmo de aprendizado.

Entrada:

Φ - conjunto de funções indutoras ϕ (algoritmos)

\mathcal{L} - conjunto de dados rotulados

ζ - orçamento (quantidade de exemplos a rotular)

q - conjunto de funções que representam estratégias (Algoritmo 2)

k - número de partições da validação cruzada

Resultado:

ϕ - função indutora (melhor algoritmo)

```

1 função induz( $\phi, \zeta, q, \mathcal{L}$ )
2    $\mathcal{L}' \subset \mathcal{L} \triangleleft$  conjunto de treinamento inicial contém um exemplo por classe:
3    $|\mathcal{L}'| = |\mathcal{Y}|, y \neq z \forall \langle \_, y \rangle, \langle \_, z \rangle \in \mathcal{L}'$ 
4    $\mathcal{U}' = \{\mathbf{x} \mid \langle \mathbf{x}, \_ \rangle \in \mathcal{L} \setminus \mathcal{L}'\}$ 
5    $\mathcal{Q} = \text{amostragem}(\mathcal{U}', \mathcal{L}', q, \zeta)$ 
6   retorna  $\phi(\mathcal{Q})$ 

7 função melhor( $\Phi, \mathcal{L}, \zeta, q, k$ )
8   retorna  $\arg \max_{\phi \in \Phi} [\text{val.cruzada}(k, \mathcal{L}' \mapsto \text{induz}(\phi, \zeta, q, \mathcal{L}'), \mathcal{L})]$ 
```

⁷ A simulação ignora que os conjuntos de dados da coleção já contêm todos os rótulos. Assim, apenas os exemplos consultados junto a um oráculo imaginário podem ser considerados rotulados durante a simulação.

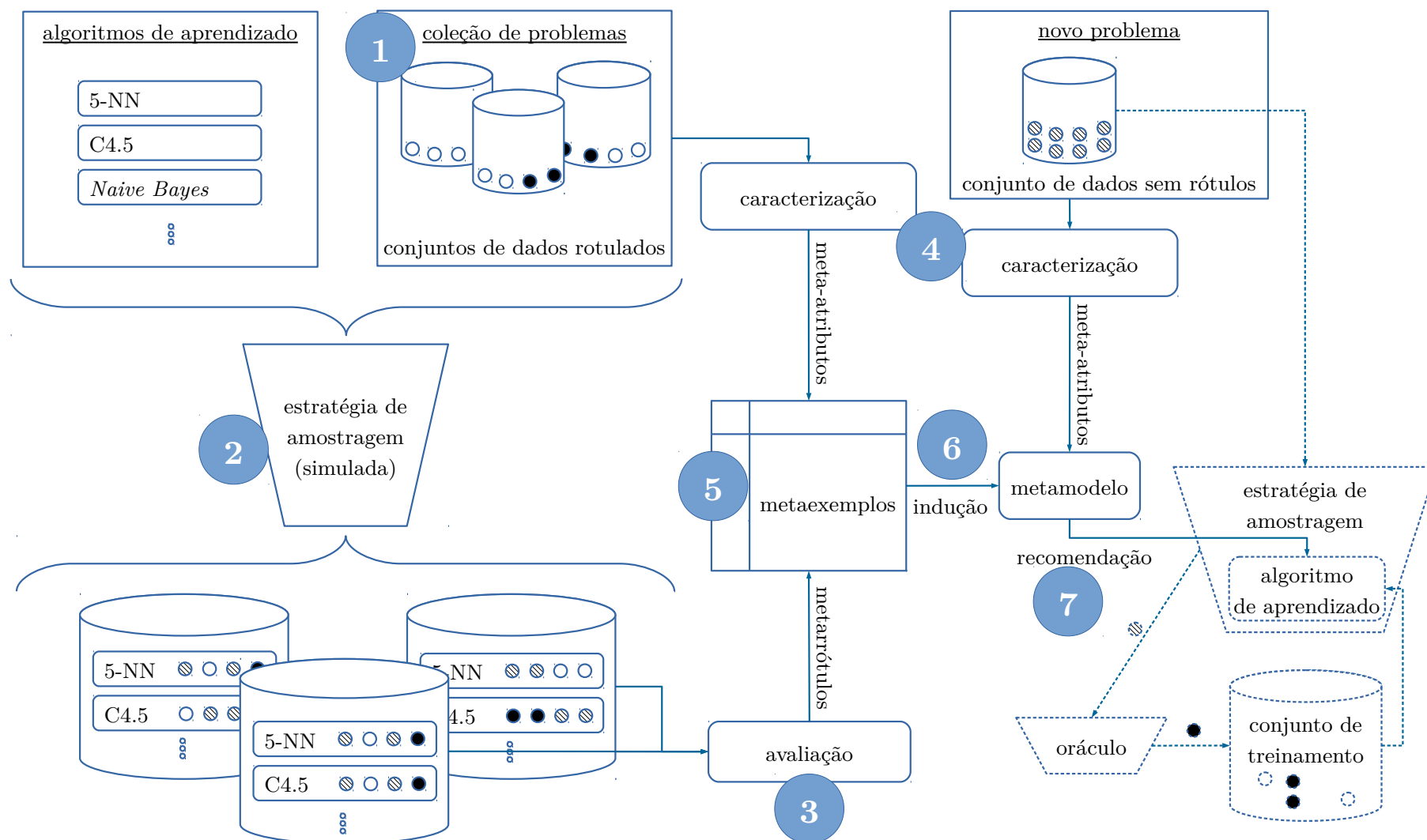


Figura 8 – Esquema do sistema de recomendação. Os elementos tracejados representam o ciclo usual de aprendizado ativo, que aqui usufrui do meta-aprendizado. Nessa proposta inicial, a estratégia é sempre a mesma.

A obtenção de informações sobre o desempenho de estratégias em conjuntos já rotulados requer a caracterização destes (Passo 4). A caracterização ocorre por meio da extração de medidas que descrevam os dados, chamadas *meta-atributos*. Uma vez caracterizados, os conjuntos tornam-se comparáveis entre si e viáveis para a construção de um metaconjunto de treinamento, pois passaram a ser representados num mesmo espaço de meta-atributos, como metaexemplos (Passo 5). O metamodelo induzido com os metaexemplos (Passo 6) é, então, capaz de fazer recomendações de algoritmos (Passo 7) para a estratégia de amostragem ativa (não mais simulada).

Usualmente, medidas estatísticas simples, como *assimetria* e *curtose* são calculadas para cada atributo numérico do conjunto de dados. Essas medidas estão presentes no STATLOG (Seção C.4), um dos primeiros sistemas de meta-aprendizado. Ele contempla conjuntos de dados com diferentes quantidades de atributos numéricos, criando um meta-atributo para cada medida. Assim, cada meta-atributo consiste na média dos valores da medida ao longo de todos os atributos numéricos. Segundo Kalousis (2002), essa abordagem por médias incorre em perda de poder discriminatório. Sua alternativa foi a adoção de histogramas com a contagem de ocorrências em determinadas faixas de valores. Nesta tese, devido à impossibilidade de prever o intervalo total de ocorrência dos valores e definir seu particionamento, optou-se por substituir o esquema de histogramas por um esquema limitado a alguns valores relevantes. Esses valores seriam o mínimo, o máximo, o médio e a razão entre o mínimo e o máximo. Na presente proposta, as medidas escolhidas para caracterização dos conjuntos foram obtidas de alguns trabalhos da literatura de meta-aprendizado. As medidas são descritas no Quadro 2.

Finalmente, o esquema proposto segue o Algoritmo 6. O esquema é o mesmo nas demais formas de recomendação, com pequenas alterações triviais. Na recomendação de estratégias, escolhe-se um algoritmo de antemão e, na recomendação de pares estratégia-algoritmo, ambos compõem a informação de metaclasses.

3.3 Considerações

Neste capítulo, as estratégias ATU e HTU foram propostas juntamente com a adaptação SGmulti e o sistema de recomendação baseado em meta-aprendizado. Todas as propostas foram textualmente e algoritmicamente descritas. No Quadro 3, as novas estratégias são incorporadas à lista comparativa de estratégias previamente descritas no Capítulo 2.

Resumidamente, a estratégia ATU consiste na remoção total do aprendiz de TU. HTU consiste na ativação do aprendiz somente nos momentos em que o modelo preditivo esteja suficientemente estável, a ponto de manter a correlação entre a medida de incerteza e a função-critério de TU acima de um elevado limiar. SGmulti apenas estende o conceito

Quadro 2 – Descrição dos 53 meta-atributos.

Meta-atributo	Descrição	Fórmula
#at	número de atributos ^a	$ \mathcal{A} $
#ex	número de exemplos ^a	$ \mathcal{U} $
#nc	número de classes ^b	$ Y $
#ea	proporção de exemplos para atributos ^c	$\frac{ \mathcal{U} }{ \mathcal{A} }$
%no	proporção de atributos nominais ^c	$\frac{1}{ \mathcal{A} } \{A \in \mathcal{A} \mid \text{isnom}(A) = 1\} $
lgex	logaritmo do número de exemplos ^d	$\log \mathcal{U} $
lgea	logaritmo da proporção exemplos/atributos ^d	$\log \frac{ \mathcal{U} }{ \mathcal{A} }$
#no _{min} , #no _{max} , #no _{mea} , #no _{min/max}	qtd. de valores nominais: ^c <i>mínima, máxima, média e mín./máx. por atributo^c</i>	$\#no_A = A , A \in \mathcal{A}$
μ_{min} , μ_{max} , μ_{mea} , $\mu_{min/max}$	média ^c (<i>idem</i>)	$\mu_j = \frac{1}{ \mathcal{U} } \sum_{x \in \mathcal{U}} x_j \quad 1 \leq j \leq \mathcal{A} $
σ_{min} , σ_{max} , σ_{mea} , $\sigma_{min/max}$	desvio padrão ^b (<i>idem</i>)	$\sigma_j = \frac{1}{ \mathcal{U} } \sum_{x \in \mathcal{U}} (x_j - \mu_j)^2$
en _{min} , en _{max} , en _{mea} , en _{min/max}	entropia normalizada ^b (<i>idem</i>)	$en_j = \frac{-1}{\log \mathcal{U} } \sum_{x \in \mathcal{U}} x_j \log x_j$
ρ_{min} , ρ_{max} , ρ_{mea} , $\rho_{min/max}$	correlação entre pares de atributos ^b (<i>idem</i>)	$\rho_{jk} = (\sigma_j^2 \cdot \sigma_k^2)^{-\frac{1}{2}} \sum_{x \in \mathcal{U}} (x_j - \mu_j)(x_k - \mu_k)$
sk _{min} , sk _{max} , sk _{mea} , sk _{min/max}	assimetria ^b (<i>idem</i>)	$sk_j = \frac{n}{(n-1)(n-2)} \sum_{x \in \mathcal{U}} \frac{(x_j - \mu_j)^3}{\sigma_j^3}$
ku _{min} , ku _{max} , ku _{mea} , ku _{min/max}	curtose ^b (<i>idem</i>)	$ku_j = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{x \in \mathcal{U}} \frac{(x_j - \mu_j)^4}{\sigma_j^4} - 3(n-1)^2 [(n-2)(n-3)]^{-1}$
cn _{k1} , cn _{k1.5} , cn _{k2} , cn _{h1} , cn _{h1.5} , cn _{h2}	conectiv. ^e <i>k</i> -médias conectiv agrup. hierárq.	agrupam. (XU; WUNSCH, 2008) qtd. de grupos: $ Y $; $1,5 Y $ e $2 Y $
du _{k1} , du _{k1.5} , du _{k2} , du _{h1} , du _{h1.5} , du _{h2}	índ. Dunn ^e <i>k</i> -médias índ. Dunn agr. hierárq.	agrupam. (DUNN, 1973) (<i>idem</i>)
si _{k1} , si _{k1.5} , si _{k2} , si _{h1} , si _{h1.5} , si _{h2}	silhueta ^e <i>k</i> -médias silhueta agrup. hierárq.	agrupam. (ROUSSEEUW, 1987) (<i>idem</i>)

^a Caracterização sugerida por Rendell, Sheshu e Tcheng (1987).

^b Baseado no projeto STATLOG (BRAZDIL; HENERY, 1994).

^c Baseado no conjunto de Kalousis (2002).

^{c'} Adaptação da sumarização proposta por Kalousis (2002).

^d Meta-atributos para recomendar algoritmos de agrupamento (SOUTO *et al.*, 2008).

^e Meta-atributos para recomendação de algoritmos de classificação (SOUZA, 2010) e de algoritmos de agrupamento (FERRARI; CASTRO, 2015).

Quadro 3 – Características de cada estratégia, incluindo estratégias propostas (em negrito).

Estratégia	Busca	Aprendiz	Dependência	Complexidade
Rnd ^a	exploratória aleatória	ausente	nenhuma	dispensa treinamento
ATU ^k	exploratória	sem aprendiz	nenhuma	dispensa treinamento
HS ^b	balanceada exploratória prospectiva	ausente	total	dispensa treinamento [*]
Mar/Ent ^a /QBC ^c	prospectiva	presente	total	$\mathcal{O}(1)$
DW ^d	prospectiva	presente	total	$\mathcal{O}(1)$
EER ^e	prospectiva	presente	total	$\mathcal{O}(Y \mathcal{U} ^2)$
TU ^f	balanceada exploratória prospectiva	presente	total	$\mathcal{O}(1)$
SGnetwork ^g	limitada exploratória aleatória	presente	total	$\mathcal{O}(1)$
SGmulti ^g	limitada exploratória aleatória	presente	total	$\mathcal{O}(Y)$
SVMsim ^h	prospectiva	presente específico	total	$\mathcal{O}(\mathcal{L} \mathcal{U})$
EGL ⁱ	prospectiva	presente específico	total	$\mathcal{O}(Y \mathcal{U})$
SVMbal ^j	combinada exploratória prospectiva	presente específico	total	$\mathcal{O}(\mathcal{L} \mathcal{U})$
HTU ^k	combinada exploratória prospectiva	combinação ausente presente	total	$\mathcal{O}(1)$

^{*} Na ausência de aprendiz não há treinamento, porém HS tem sua própria complexidade a ser considerada.

^a Amostragem aleatória, por margem ou entropia (SETTLES, 2012).

^b Amostragem hierárquica (DASGUPTA, 2011).

^c Consulta por comitê (ABE; MAMITSUKA, 1998).

^d Amostragem ponderada por densidade (SETTLES, 2008).

^e Redução esperada do erro (GUO; GREINER, 2007).

^f Amostragem ponderada por densidade e utilidade de treinamento (SETTLES, 2010; FUJII *et al.*, 1998).

^g SGmulti (SANTOS; CARVALHO, 2014a).

^h Margem simples (TONG; KOLLER, 2001).

ⁱ Comprimento esperado do gradiente (SETTLES; CRAVEN; RAY, 2007).

^j Balanceamento exploração-prospecção (OSUGI; DENG; SCOTT, 2005).

^k Amostragem ponderada por densidade sem aprendiz e híbrida (SANTOS; CARVALHO, 2015b).

Algoritmo 6: Recomendação automática de algoritmos de aprendizado.**Entrada:**

Φ - conjunto de funções indutoras ϕ (algoritmos)
 \mathcal{L} - coleção de conjuntos de dados rotulados
 ψ - função de indução (algoritmo do meta-aprendiz)
 ζ - orçamento (50 ou 100 exemplos a rotular)
 q - função-critério que representa a estratégia (Algoritmo 2)
 k - número de partições da validação cruzada
 \mathcal{U} - novo conjunto de dados sem rótulos

Resultado:

ϕ - função indutora recomendada (algoritmo do aprendiz ativo)

```

1 função treina( $\Phi, \mathcal{L}, \psi, \zeta, q, k, \mathcal{U}$ )
2    $\Lambda \leftarrow \emptyset$ 
3   para cada  $\mathcal{L} \in \mathcal{L}$  faça
4      $\gamma = \text{melhor}(\Phi, \mathcal{L}, \zeta, q, k)$  ◁ metaclasse
5      $\chi = (\#at, \#ex, \#nc, \#ea, \%no, lgex, lgea, \#no_{min}, \#no_{max}, \#no_{mea}, \#no_{min/max},$ 
         $\mu_{min}, \mu_{max}, \mu_{mea}, \mu_{min/max}, \sigma_{min}, \sigma_{max}, \sigma_{mea}, \sigma_{min/max},$ 
         $en_{min}, en_{max}, en_{mea}, en_{min/max}, \rho_{min}, \rho_{max}, \rho_{mea}, \rho_{min/max},$ 
         $sk_{min}, sk_{max}, sk_{mea}, sk_{min/max}, ku_{min}, ku_{max}, ku_{mea}, ku_{min/max},$ 
         $cn_{k1}, cn_{k1.5}, cn_{k2}, cn_{h1}, cn_{h1.5}, cn_{h2}, du_{k1}, du_{k1.5}, du_{k2}, du_{h1}, du_{h1.5}, du_{h2},$ 
         $si_{k1}, si_{k1.5}, si_{k2}, si_{h1}, si_{h1.5}, si_{h2})$  ◁ meta-atributos extraídos de  $\mathcal{L}$ 
6      $\Lambda \leftarrow \Lambda \cup \{\langle \chi, \gamma \rangle\}$ 
7    $\eta \leftarrow \psi(\Lambda)$  ◁ induz metamodelo
8   retorna  $\eta$ 

9 função recomenda( $\eta, \zeta, \mathcal{U}$ )
10   $\chi = (\#at, \dots, si_{h2})$  ◁ meta-atributos extraídos de  $\mathcal{U}$ 
11   $\phi \leftarrow \hat{y}_\eta(\chi)$ 
12  retorna  $\phi$ 

```

da abordagem original, mantendo uma reserva de exemplos artificiais para cada classe.

A tentativa de aprendizado meta-ativo é uma nova abordagem na área de aprendizado ativo - até onde alcança o conhecimento do autor. Ela é baseada na recomendação automática de algoritmos de aprendizado. Inicialmente, meta-atributos são extraídos do conjunto de dados, caracterizando-o. Uma base de conhecimento previamente construída por meio da caracterização de outros conjuntos, cujos melhores algoritmos de aprendizado são conhecidos, é utilizada como metaconjunto. Idealmente, um metamodelo induzido com esse metaconjunto permitiria prever qual o melhor algoritmo e, possivelmente, a melhor estratégia ou o melhor par estratégia-algoritmo.

As propostas aqui apresentadas foram avaliadas empiricamente de acordo com o método descrito no Capítulo 4.

Método experimental

Bien loin que l'objet précède le point de vue, on dirait que c'est le point de vue qui crée l'objet...

Ferdinand de Saussure¹

As escolhas metodológicas expostas neste capítulo fundamentam a avaliação experimental das propostas deste trabalho. Na Seção 4.1, são apresentados detalhes sobre o cenário adotado. Os conjuntos de dados, utilizados visando a simulação de aplicações reais, são descritos e analisados sob diferentes aspectos na Seção 4.2. Nas seções 4.3 e 4.4, os parâmetros são definidos e considerações são feitas sobre as estratégias de amostragem ativa e os algoritmos de aprendizado empregados - respectivamente. As medidas de desempenho, a forma de validação e o tipo de teste empregado para a avaliação da significância estatística dos resultados experimentais são apresentados na Seção 4.5. Na Seção 4.6, são introduzidas as curvas de ranqueamento - importante contribuição metodológica deste trabalho. Por fim, a Seção 4.7 contém as considerações gerais.

Parte das decisões tomadas foram baseadas num experimento preliminar. Ele permitiu a definição de experimentos definitivos adequados aos recursos computacionais disponíveis. Aproximadamente um mês foi necessário para a conclusão do experimento nos 28 menores conjuntos de dados. Assim, considerando-se o crescimento exponencial do custo computacional em alguns pares estratégia-algoritmo com o aumento do número de classes, atributos e/ou exemplos, foi necessário reduzir a quantidade de iterações do procedimento de validação e antecipar o término das consultas. Os resultados desse experimento preliminar com relação a desempenho não são reportados nesta tese por serem menos abrangentes que o experimento definitivo. Mais detalhes podem ser encontrados em Santos e Carvalho (2014a).

¹ “Bem longe de dizer que o objeto precede o ponto de vista, diríamos que é o ponto de vista que cria o objeto...” - Saussure (1972), sobre as peculiaridades metodológicas em sua área de estudo.

4.1 Cenário escolhido

É necessário situar o escopo do presente trabalho, dado o grande número de possibilidades dentro dos cenários e restrições que possam existir num dado problema. Existem três principais cenários na literatura de aprendizado ativo (SETTLES, 2010): *síntese de consulta por associação* ou *consulta de exemplos sintetizados*²; *amostragem baseada em reserva de exemplos*; e, *amostragem seletiva baseada em fluxo*. O cenário adotado neste trabalho é baseado em reserva de exemplos. Especificamente, há as seguintes restrições de escopo:

- problemas monorrótulo multiclasse;
- consulta pela classe (não por valores de atributos, por exemplo);
- conjunto ou quantidade de classes possíveis previamente conhecida;
- distribuição das classes não necessariamente balanceada;
- distribuição estacionária;
- custo por erro de classificação uniforme;
- custo por consulta uniforme;
- oráculo único e sujeito a ruído consistente (sempre comete os mesmos erros);
- os domínios dos atributos nominais são previamente conhecidos;
- atributos sem valores faltantes; e,
- consulta *on-line*, ou seja, um a um.

As seguintes condições foram criadas visando maior rigor experimental, replicabilidade e verossimilhança com aplicações reais:

- um rótulo inicial por classe;
- todo o conjunto de dados original é utilizado no processo de validação cruzada - exceto exemplos duplicados e no caso de EER (Seção 4.4.1); e,
- critério de parada determinado pelo orçamento de cem consultas ($\zeta = 100$).

Uma condição mínima necessária a diversas estratégias é existir, no modelo de classificação, a capacidade de estimar pelo menos um esboço da fronteira de decisão - vide critérios

² [*membership query synthesis*]

de consulta no Capítulo 2. Consequentemente, optou-se pela adoção de um exemplo rotulado inicial por classe. Além dessa ser a quantidade mínima para que um modelo seja induzido e possa estimar uma fronteira de decisão, ela também reflete a cautela necessária em aplicações reais, pois o aprendizado ativo não garante que exemplos de todas as classes sejam consultados. Por exemplo, quanto maior o grau de desbalanceamento entre as classes, maior o custo com consultas adicionais necessárias até que todas as classes sejam contempladas. Portanto, é importante obter criteriosamente alguns rótulos antes de ser iniciado o processo de amostragem ativa.

Uma maneira de obter um rótulo por classe é por meio do *aprendizado guiado* (ATTENBERG; PROVOST, 2010b). Esse tipo de aprendizado realiza uma amostragem em que o usuário escolhe os exemplos a rotular. Normalmente, ele sabe de antemão as classes menos frequentes e é capaz de encontrar exemplos correspondentes, seja por meio de consulta a motores de busca na *internet*, da própria memória pessoal ou de outras formas. Essa modalidade de amostragem faz maior usufruto da capacidade da inteligência humana enquanto agente supervisor do que a simples rotulação. Embora custoso, esse potencial humano pode ser aproveitado no estágio inicial da rotulação e favorecer a posterior amostragem ativa, minimizando as futuras intervenções humanas.

A prática de rotulação prévia em experimentos também tem sido parte do método experimental adotado em outros trabalhos (GUO; SCHUURMANS, 2007; BISWAS; PARIKH, 2013; GU; JIN; CHIU, 2014), com variações mais permissivas com relação à quantidade de exemplos (PATRA; BRUZZONE, 2012) ou substituída por uma amostragem aleatória aplicada até que se encontre um exemplo de cada classe (CHERMAN, 2013).

A limitação a um número fixo de consultas segue a observação de Settles (2012) sobre o critério desejável de parada do processo de aprendizado ativo. Ele relata que, em aplicações reais, normalmente trata-se de uma restrição financeira. Consequentemente, a definição do orçamento para fins experimentais tem sido arbitrária na literatura, sendo 100 consultas um valor recorrente (CRAWFORD; TUIA; YANG, 2013; CHERMAN, 2013; SETTLES; CRAVEN; RAY, 2007; ROY; MCCALLUM, 2001; SETTLES; CRAVEN, 2008).

Outra característica presente nos experimentos está no não aproveitamento da informação de orçamento disponível - decisão usual na literatura (ROY; MCCALLUM, 2001). Isso implica na geração de sequências de consultas provavelmente não ótimas, ou, em outras palavras, cada consulta é realizada como se fosse a última. Uma consulta ótima dependeria de quantas ainda poderiam ser feitas. Apesar de ser uma possibilidade teórica, seu não aproveitamento não é uma limitação real, pois não é esperado que existam estratégias capazes de tirar proveito dessa informação de forma significativa.

4.2 Conjuntos de dados

Os conjuntos de dados do projeto de Garcia (2015), baseado no repositório da Universidade da Califórnia, Irvine (UCI) - (BACHE; LICHMAN, 2013), foram utilizados com o objetivo de representar os variados domínios existentes em aplicações reais. Parte dos conjuntos foi descartada conforme descrito nas seções 4.2.1 e 4.2.2. Os critérios de descarte objetivaram selecionar os conjuntos mais adequados³ para a coleção utilizada nos experimentos desta tese. Contudo, a análise dos experimentos leva em conta que o propósito da construção de uma coleção representativa de conjuntos de dados não é a demonstração da superioridade de algum algoritmo em todos os casos, mas identificar os pontos fortes de cada algoritmo (JAPKOWICZ; SHAH, 2011).

4.2.1 Custo computacional

Dada a quantidade envolvida de estratégias de amostragem ativa, algoritmos de aprendizado e execuções de validação por conjunto de dados (Seção 4.5.2), a limitação do tempo de processamento foi essencial para que os experimentos previstos fossem realizados. Essa limitação não afetou os experimentos, pois o foco é a acurácia preditiva. Logo, os conjuntos de dados que demandam excessivo tempo para a inicialização de estratégias ou para a realização de consultas foram descartados, permitindo a viabilidade computacional dos experimentos. Esse tempo foi estimado por meio de um experimento reduzido. Ele consistiu na execução de validação cruzada em duas partes aplicada apenas ao algoritmo com maior custo computacional (RoF - Seção 4.3) e três estratégias, cada uma com sua respectiva particularidade referente ao custo computacional:

- ATU - inicialização requer quantidade quadrática de cálculos de distância em relação ao tamanho da reserva de exemplos;
- SGmulti - consulta requer quantidade de treinamentos proporcional ao número de classes; e,
- EER - consulta requer quantidade de treinamentos proporcional ao número de classes e quadrática com relação ao número de exemplos.

Os seguintes conjuntos de dados foram descartados, uma vez que atingiram um tempo de processamento entre 4 e 16 horas para as primeiras 50 consultas no experimento reduzido: *arcene*, *micro mass pure spectra*, *micro mass mixed spectra*, *multiple features*, *digits2*, *lsvt voice rehabilitation*, *semeion*, *cnae 9*, *gas drift*, *gas drift different concentrations* e *hill valley with noise*. Estimou-se um tempo entre 700 e 2800 dias para cada um desses conjuntos no experimento definitivo, tomando por base as seguintes considerações:

³ Análises posteriores à primeira versão deste documento demonstraram uma importante deficiência na coleção adotada (Apêndice B).

- 100 consultas contêm 3 vezes mais unidades de treinamento⁴ que 50 consultas;
- seriam avaliados, no pior caso, 8 algoritmos e 14 estratégias; e,
- o processo de validação consistiria de 25 execuções.

Mesmo dividindo esses valores estimados de duração do experimento definitivo pelo número de núcleos de processamento inicialmente disponíveis, 100, não seria viável incluir tais conjuntos de dados.

4.2.2 Similaridade

Conjuntos de dados equivalentes ou muito similares foram excluídos dos experimentos desta tese com a finalidade de aumentar a *independência entre amostras* para o teste estatístico (seções 4.5.3 e 4.5.4). A ausência de redundância entre os conjuntos também é importante para a avaliação adequada da capacidade de generalização do meta-aprendizado: tal como no nível base, metaexemplos não devem aparecer simultaneamente nos conjuntos de teste e treinamento. A similaridade entre conjuntos foi indiretamente estimada do ponto de vista do viés de aprendizado. Trata-se de uma escolha conservadora, pois os mesmos vieses podem ser adequados a problemas distintos. Essa abordagem é baseada na tabela de distâncias entre taxas de erro empregada por Brazdil e Henery (1994). Neste trabalho, entretanto, optou-se pela comparação de desempenhos relativos, feita por uma medida de correlação, pois ela proporciona a análise de um ponto de vista mais amplo, independente de valores absolutos de acurácia, por exemplo.

A correlação de Spearman (GIBBONS; CHAKRABORTI, 2011) foi calculada entre os ranqueamentos médios de desempenho de 10 algoritmos de aprendizado (Seção 4.3) com o objetivo de estimar indiretamente a similaridade entre cada par de conjuntos. A medida de desempenho adotada, κ (Seção 3.1.2), foi obtida em 100 execuções de validação cruzada em dez partes. Valores elevados de correlação foram obtidos devido à frequente superioridade de alguns algoritmos sobre outros (DELGADO *et al.*, 2014). Valores acima de 0,95 foram considerados indicativos de similaridade excessiva. Esse foi o maior valor atingido por conjuntos de domínios conhecidamente distintos.

Apenas um conjunto dentre cada grupo de conjuntos similares permaneceu e os demais foram descartados. Os conjuntos descartados são listados a seguir: *thyroid allrep*, *thyroid allhyper*, *thyroid allhypo*, *thyroid allbp*, *thyroid dis*, *robot failure lp4*, *leukemia haslinger*, *robot nav sensor readings 4*, *cardiotocography 10class*, *nursery 4class*, *movement libras 10*, *mushroom expanded*, *waveform v1*, *connectionist vowel reduced*, *breast tissue 6class*, *yeast*, *wine quality 5class*, *abalone 11class*, *robot nav sensor readings 24* e variações dos conjuntos *volcanoes a*, *b*, *d* e *e*. Conjuntos em que todos os algoritmos obtiveram um

⁴ Quantidade de exemplos de treinamento. Todos são reaprendidos a cada nova consulta.

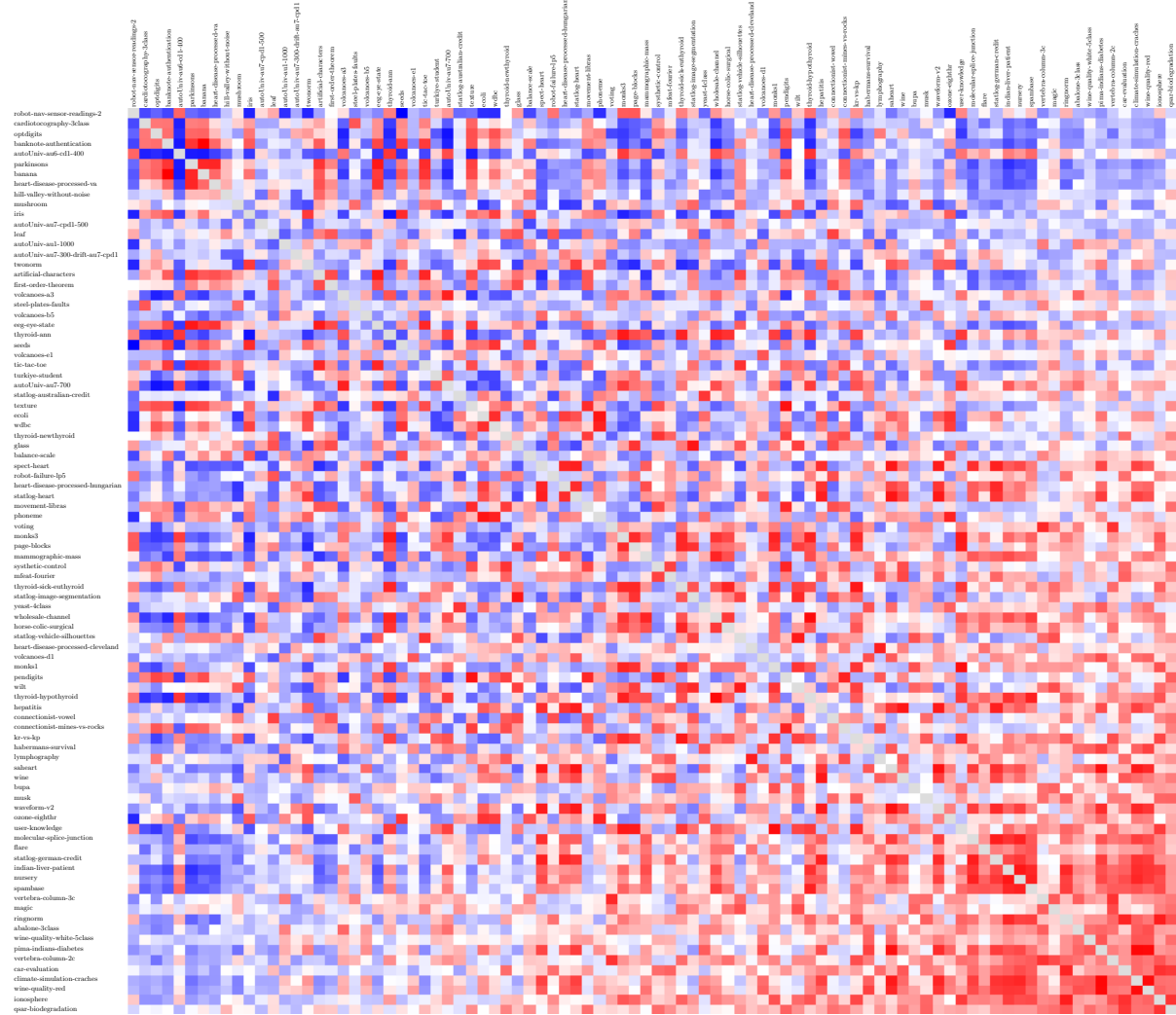


Figura 9 – Visão geral da intensidade de correlação entre ranqueamentos de algoritmos de aprendizado para cada par possível na coleção. Azul: mais adequado. Vermelho: menos adequado.

desempenho próximo ao acaso ($\kappa < 0,02$) poderiam gerar colocações de ranqueamento de pares estratégia-algoritmo espúrias, afetando os resultados sem trazer qualquer benefício para os experimentos. Consequentemente, eles também foram descartados: *autoUniv-au6-1000*, *autoUniv-au6-1000*, *autoUniv-au6-250-drift-au6-cd1-500*, *meta-data*, *planning-relax* e *monks2*.

A Figura 9 contém o mapa de calor (WILKINSON; FRIENDLY, 2009) que atribui uma intensidade de correlação a cada par dentre os 90 conjuntos de dados que restaram ao fim do processo de descarte. Este e os demais mapas de calor neste texto estão com as linhas ordenadas pela soma dos valores nas colunas. O azul e o vermelho mais intensos indicam, respectivamente, o melhor e o pior valor. A cor branca indica o valor central. No presente caso, o melhor valor é o mais baixo; em valores de acurácia, distância e correlação na predição de ranqueamentos, o melhor valor é o mais elevado. O mapa de calor com os valores de correlação dos conjuntos que apresentaram valores acima de 0,8 é apresentado com maior clareza na Figura 10.

		banknote-authentication				eeg-eye-state				statlog-image-segmentation				monks3				texture				monks1				user-knowledge				thyroid-hypothyroid				wholesale-channel				pendigits				spect-heart				ecoli				statlog-heart				phoneme				indian-liver-patient				saheart				hepatitis				wine				nursery				waveform-v2				pima-indians-diabetes				climate-simulation-craches				wine-quality-red			
		parkinsons	banknote-authentication	banana	eeg-eye-state	glass	statlog-image-segmentation	monks3	texture	monks1	user-knowledge	thyroid-hypothyroid	wholesale-channel	pendigits	spect-heart	ecoli	statlog-heart	phoneme	indian-liver-patient	saheart	hepatitis	wine	nursery	waveform-v2	pima-indians-diabetes	climate-simulation-craches	wine-quality-red																																																																		
parkinsons			0.84	0.87	0.88	0.36	0.36	-0.49	0.78	-0.20	-0.31	-0.62	-0.50	0.65	-0.58	0.22	-0.28	0.46	-0.69	-0.39	-0.27	-0.06	-0.56	-0.29	-0.32	-0.41	-0.34																																																																		
banknote-authentication	0.84			0.95	0.76	0.12	0.08	-0.61	0.73	-0.23	-0.35	-0.73	-0.71	0.71	-0.46	0.29	-0.16	0.55	-0.57	-0.21	-0.16	0.03	-0.42	-0.13	-0.12	-0.25	-0.24																																																																		
banana	0.87	0.95			0.79	0.17	0.12	-0.58	0.75	-0.22	-0.33	-0.69	-0.67	0.74	-0.49	0.30	-0.16	0.56	-0.57	-0.25	-0.17	0.06	-0.42	-0.13	-0.14	-0.27	-0.22																																																																		
eeg-eye-state	0.88	0.76	0.79			0.56	0.45	-0.37	0.81	-0.03	-0.12	-0.46	-0.43	0.66	-0.53	0.12	-0.30	0.42	-0.54	-0.29	-0.24	-0.11	-0.49	-0.24	-0.17	-0.27	-0.16																																																																		
glass	0.36	0.12	0.17	0.56			0.92	0.44	0.31	0.59	0.57	0.32	0.24	0.25	-0.52	-0.40	-0.59	-0.15	-0.04	-0.46	-0.34	-0.21	-0.14	-0.39	-0.24	-0.22	-0.07																																																																		
statlog-image-segmentation	0.36	0.08	0.12	0.45	0.92			0.55	0.28	0.66	0.63	0.34	0.31	0.19	-0.47	-0.42	-0.57	-0.21	0.00	-0.47	-0.39	-0.19	-0.06	-0.41	-0.27	-0.22	-0.08																																																																		
monks3	-0.49	-0.61	-0.58	-0.37	0.44	0.55			-0.48	0.85	0.88	0.87	0.79	-0.39	0.14	-0.35	-0.15	-0.38	0.65	0.00	-0.07	-0.04	0.56	0.00	0.15	0.26	0.38																																																																		
texture	0.78	0.73	0.75	0.81	0.31	0.28	-0.48			-0.27	-0.26	-0.45	-0.47	0.92	-0.26	0.26	0.02	0.45	-0.34	-0.06	0.15	0.35	-0.21	0.07	-0.02	-0.05	-0.06																																																																		
monks1	-0.20	-0.23	-0.22	-0.03	0.59	0.66	0.85	-0.27			0.92	0.69	0.60	-0.20	-0.07	-0.25	-0.26	-0.12	0.50	-0.07	-0.17	-0.12	0.40	-0.10	0.21	0.25	0.38																																																																		
user-knowledge	-0.31	-0.35	-0.33	-0.12	0.57	0.63	0.88	-0.26	0.92			0.80	0.69	-0.15	0.08	-0.27	-0.12	-0.16	0.66	0.09	-0.04	0.02	0.53	0.05	0.34	0.39	0.52																																																																		
thyroid-hypothyroid	-0.62	-0.73	-0.69	-0.46	0.32	0.34	0.87	-0.45	0.69	0.80			0.93	-0.33	0.44	-0.11	0.16	-0.21	0.83	0.30	0.31	0.22	0.72	0.29	0.41	0.52	0.57																																																																		
wholesale-channel	-0.50	-0.71	-0.67	-0.43	0.24	0.31	0.79	-0.47	0.60	0.69	0.93			-0.37	0.50	0.06	0.25	-0.07	0.74	0.34	0.32	0.22	0.64	0.32	0.40	0.49	0.56																																																																		
pendigits	0.65	0.71	0.74	0.66	0.25	0.19	-0.39	0.92	-0.20	-0.15	-0.33	-0.37			-0.09	0.35	0.20	0.54	-0.10	0.10	0.32	0.56	0.02	0.28	0.17	0.13	0.14																																																																		
spect-heart	-0.58	-0.46	-0.49	-0.53	-0.52	-0.47	0.14	-0.26	-0.07	0.08	0.44	0.50	-0.09			0.51	0.88	0.26	0.75	0.92	0.79	0.59	0.76	0.86	0.78	0.82	0.72																																																																		
ecoli	0.22	0.29	0.30	0.12	-0.40	-0.42	-0.35	0.26	-0.25	-0.27	-0.11	0.06	0.35	0.51			0.74	0.92	0.12	0.67	0.68	0.60	0.34	0.68	0.63	0.54	0.52																																																																		
statlog-heart	-0.28	-0.16	-0.16	-0.30	-0.59	-0.57	-0.15	0.02	-0.26	-0.12	0.16	0.25	0.20	0.88	0.74			0.56	0.56	0.91	0.87	0.77	0.62	0.90	0.80	0.80	0.73																																																																		
phoneme	0.46	0.55	0.56	0.42	-0.15	-0.21	-0.38	0.45	-0.12	-0.16	-0.21	-0.07	0.54	0.26	0.92	0.56			0.00	0.51	0.53	0.53	0.19	0.55	0.57	0.44	0.48																																																																		
indian-liver-patient	-0.69	-0.57	-0.57	-0.54	-0.04	0.00	0.65	-0.34	0.50	0.66	0.83	0.74	-0.10	0.75	0.12	0.56	0.00			0.67	0.59	0.52	0.91	0.64	0.73	0.83	0.81																																																																		
saheart	-0.39	-0.21	-0.25	-0.29	-0.46	-0.47	0.00	-0.06	-0.07	0.09	0.30	0.34	0.10	0.92	0.67	0.91	0.51	0.67			0.82	0.63	0.72	0.89	0.93	0.92	0.83																																																																		
hepatitis	-0.27	-0.16	-0.17	-0.24	-0.34	-0.39	-0.07	0.15	-0.17	-0.04	0.31	0.32	0.32	0.79	0.68	0.87	0.53	0.59	0.82			0.90	0.69	0.90	0.74	0.78	0.66																																																																		
wine	-0.06	0.03	0.06	-0.11	-0.21	-0.19	-0.04	0.35	-0.12	0.02	0.22	0.22	0.56	0.59	0.60	0.77	0.53	0.52	0.63	0.90			0.65	0.84	0.59	0.65	0.59																																																																		
nursery	-0.56	-0.42	-0.42	-0.49	-0.14	-0.06	0.56	-0.21	0.40	0.53	0.72	0.64	0.02	0.76	0.34	0.62	0.19	0.91	0.72	0.69	0.65			0.77	0.76	0.85	0.81																																																																		
waveform-v2	-0.29	-0.13	-0.13	-0.24	-0.39	-0.41	0.00	0.07	-0.10	0.05	0.29	0.32	0.28	0.86	0.68	0.90	0.55	0.64	0.89	0.90	0.84	0.77			0.81	0.84	0.78																																																																		
pima-indians-diabetes	-0.32	-0.12	-0.14	-0.17	-0.24	-0.27	0.15	-0.02	0.21	0.34	0.41	0.40	0.17	0.78	0.63	0.80	0.57	0.73	0.93	0.74	0.59	0.76	0.81			0.95	0.93																																																																		
climate-simulation-craches	-0.41	-0.25	-0.27	-0.27	-0.22	-0.22	0.26	-0.05	0.25	0.39	0.52	0.49	0.13	0.82	0.54	0.80	0.44	0.83	0.92	0.78	0.65	0.85	0.84	0.95			0.93																																																																		
wine-quality-red	-0.34	-0.24	-0.22	-0.16	-0.07	-0.08	0.38	-0.06	0.38	0.52	0.57	0.56	0.14	0.72	0.52	0.73	0.48	0.81	0.83	0.66	0.59	0.81	0.78	0.93	0.93																																																																				

Figura 10 – Valores de correlação entre os conjuntos de dados mais correlacionados.

Variações como *thyroid-sick-euthyroid* e *thyroid-hypothyroid* ou *volcanoes-d1* e *volcanoes-e1* foram consideradas do mesmo domínio, porém como problemas distintos. Na Tabela 1, é possível observar que os domínios começam a coincidir somente com correlações abaixo de 0,830, havendo 18 pares de domínios distintos acima desse valor. Logo, foi suposto que abaixo desse valor a coincidência de domínios fosse pouco preocupante. Entretanto, análises posteriores ao término da presente pesquisa (Apêndice B) indicam que seria fundamental manter apenas um conjunto de dados por domínio, para que a generalidade do aprendiz meta-ativo pudesse ser verificada.

4.2.3 Descrição

Todos os conjuntos selecionados, após particionamento dos dados durante o processo de validação cruzada, oferecem uma reserva com pelo menos 100 exemplos. Essa quantidade corresponde a um padrão de quantidade de consultas permitidas, de modo que as curvas de aprendizado de todos os conjuntos de dados sejam comparáveis entre si. Os conjuntos têm suas características expostas no Apêndice E - tabelas 19 e 20. Para ilustrar a diversidade dos conjuntos, alguns grupos especiais foram criados da seguinte forma:

- classe majoritária 20 vezes mais numerosa que a classe minoritária (Tabela 2);
- 60 ou mais atributos (Tabela 3);
- mais de 6 classes (Tabela 4);
- mais de 4000 exemplos (Tabela 5); e,
- menos de 200 exemplos (Tabela 6).

Devido à presença de atributos nominais e numéricos, os algoritmos providos pela ferramenta Weka (HALL *et al.*, 2009) fizeram eventual uso de seus próprios filtros internos para lidar com cada tipo de atributo de acordo com a natureza do algoritmo: binarização, discretização e *padronização*⁵ (KREYSZIG, 2007).

Exemplos duplicados foram removidos visando simular um oráculo consistente. A cada grupo de exemplos repetidos, um único representante foi mantido cuja classe atribuída foi a moda das classes contempladas pelo grupo.

4.3 Algoritmos de aprendizado

Cada algoritmo de aprendizado tem, intrinsecamente, um viés próprio (Apêndice C) que, quando integrado a uma estratégia de amostragem ativa, pode influenciar no

⁵ [z-score]

Tabela 1 – Pares de conjuntos com as maiores correlações. *Pares em negrito relacionam conjuntos de domínios possivelmente próximos. Linha pontilhada simples indica omissão de pares de conjuntos de domínios distintos. Linha pontilhada dupla indica omissão de pares em geral.*

Primeiro conjunto	Segundo conjunto	Correlação
banana	banknote-authentication	0,950
climate-simulation-craches	pima-indians-diabetes	0,949
pima-indians-diabetes	wine-quality-red	0,944
thyroid-hypothyroid	wholesale-channel	0,941
climate-simulation-craches	wine-quality-red	0,940
monks1	user-knowledge	0,932
pima-indians-diabetes	saheart	0,926
saheart	waveform-v2	0,924
pendigits	texture	0,921
saheart	statlog-heart	0,912
ecoli	phoneme	0,912
glass	statlog-image-segmentation	0,909
indian-liver-patient	nursery	0,906
saheart	spect-heart	0,903
eeg-eye-state	parkinsons	0,903
hepatitis	wine	0,902
monks3	user-knowledge	0,900
statlog-heart	waveform-v2	0,900
thyroid-ann	thyroid-hypothyroid	0,829
vertebra-column-2c	vertebra-column-3c	0,821
statlog-german-credit	statlog-heart	0,821
statlog-image-segmentation	statlog-vehicle-silhouettes	0,776
autoUniv-au6-cd1	autoUniv-au7-700	0,744
thyroid-hypothyroid	thyroid-sick-euthyroid	0,740
wine-quality-red	wine-quality-white-5class	0,712
volcanoes-d1	volcanoes-e1	0,712
robot-failure-lp5	robot-nav-sensor-readings-2	0,618
volcanoes-b5	volcanoes-d1	0,606
wine	wine-quality-red	0,590
connectionist-mines-vs-rocks	connectionist-vowel	0,503
volcanoes-b5	volcanoes-e1	0,500
autoUniv-au6-cd1-400	autoUniv-au7-300-drift-au7,,,	0,000
statlog-heart	statlog-image-segmentation	-0,524
robot-nav-sensor-readings-2	seeds	-0,888

desempenho da estratégia. Dessa forma, diferentes algoritmos foram adotados na comparação de estratégias.

As implementações adotadas são aquelas presentes na biblioteca Weka: k -vizinhos mais próximos (k -NN)⁶; árvore de decisão C4.5 (C4.5w)⁷; NB (*Naive Bayes*); e, *Support Vector Machines* (SVM)⁸ (HART, 1968; QUINLAN, 1993; LEWIS, 1998; HEARST *et al.*, 1998). O sufixo *w* foi acrescentado como indicativo de que a implementação Weka não corresponde diretamente ao algoritmo original.

⁶ Chamado IBk na implementação Weka.

⁷ Chamado J48 na implementação Weka.

⁸ Invólucro LibSVM para Weka.

Tabela 2 – Conjuntos de dados mais desbalanceados.

Conjunto de dados	$ \mathcal{U} $	$ Y $	Atributos	Nominais	Majoritária (%)	Minoritária (%)
44-nursery	10368	5	8	8	33,3	1 exemplo
79-volcanoes b5	7986	5	3	0	96,2	0,3
47-page blocks	4314	5	10	0	90,5	0,5
80-volcanoes d1	7002	5	3	0	94,4	0,6
81-volcanoes e1	946	5	3	0	91,6	0,8
17-ecoli	269	8	7	0	42,6	0,6
88-wine quality red	1087	6	11	0	42,5	0,7
78-volcanoes a3	1217	5	3	0	90,0	1,9
34-lymphography	118	4	18	15	54,7	1,4
68-thyroid ann	2967	3	21	0	92,3	2,5
73-turkiye student	2667	13	32	0	13,6	0,5
69-thyroid hypothyroi...	2468	2	25	18	95,5	4,5

Tabela 3 – Conjuntos de dados com mais atributos.

Conjunto de dados	$ \mathcal{U} $	$ Y $	Atributos	Nominais	Majoritária (%)	Minoritária (%)
43-musk	5265	2	167	1	84,5	15,5
27-hill valley withou...	970	2	100	0	50,5	49,5
54-robot failure lp5	130	5	90	0	27,8	13,0
41-movement libras	264	15	90	0	7,3	6,4
37-mfeat fourier	1595	10	76	0	10,0	9,9
46-ozone eighthr	2021	2	72	0	93,7	6,3
45-optdigits	4496	10	62	0	10,2	9,9
66-systhetic control	480	6	60	0	16,7	16,7
38-molecular splice j...	2404	3	60	60	55,0	22,4
15-connect. mines vs...	166	2	60	0	53,4	46,6

Algoritmos que requeiram ajuste externo de parâmetros frequentemente dependem de métodos de seleção de modelo (ARLOT; CELISSE *et al.*, 2010). Entretanto, no cenário de aprendizado ativo, o conjunto de treinamento é pequeno durante a maior parte da curva de aprendizado, tornando a seleção de modelos pouco confiável. Adicionalmente, a complexidade dos métodos de seleção de modelo é computacionalmente incompatível com a necessidade de induzir um novo modelo a cada nova consulta. Por fim, tendo em vista que o objetivo da seleção de algoritmos descrita nesta seção não foi a maximização da acurácia de algum algoritmo em particular, mas fornecer uma diversidade de vieses de aprendizado, não era mandatório que os parâmetros fossem otimamente ajustados. Consequentemente, optou-se por valores pré-definidos. O ajuste de parâmetros foi feito de acordo os valores padrão das implementações e as necessidades do aprendizado ativo, conforme listado a seguir.

- k -NN (denominado 5NN daqui em diante) - A quantidade adotada de 5 vizinhos

Tabela 4 – Conjuntos de dados com mais classes.

Conjunto de dados	$ \mathcal{U} $	$ Y $	Atributos	Nominais	Majoritária (%)	Minoritária (%)
33-leaf	272	30	15	0	4,7	2,4
41-movement libras	264	15	90	0	7,3	6,4
73-turkiye student	2667	13	32	0	13,6	0,5
67-texture	4378	11	40	0	9,1	9,1
16-connect. vowel...	792	11	13	0	9,1	9,1
49-pendigits	8794	10	16	0	10,4	9,6
45-optdigits	4496	10	62	0	10,2	9,9
37-mfeat fourier	1595	10	76	0	10,0	9,9
2-artificial charact...	3890	10	7	0	14,4	6,0
17-ecoli	269	8	7	0	42,6	0,6
4-autoUniv au6 cd1 4...	320	8	40	3	27,8	6,3
63-statlog image segm...	1669	7	18	0	14,4	14,0

Tabela 5 – Conjuntos de dados com mais exemplos.

Conjunto de dados	$ \mathcal{U} $	$ Y $	Atributos	Nominais	Majoritária (%)	Minoritária (%)
35-magic	15124	2	10	0	65,2	34,8
18-eeg eye state	11984	2	14	0	55,1	44,9
44-nursery	10368	5	8	8	33,3	0,0
49-pendigits	8794	10	16	0	10,4	9,6
79-volcanoes b5	7986	5	3	0	96,2	0,3
80-volcanoes d1	7002	5	3	0	94,4	0,6
42-mushroom	6499	2	21	21	51,8	48,2
74-twonorm	5920	2	20	0	50,0	50,0
53-ringnorm	5920	2	20	0	50,5	49,5
43-musk	5265	2	167	1	84,5	15,5
45-optdigits	4496	10	62	0	10,2	9,9
19-first order theore...	4402	6	51	0	42,2	8,0
67-texture	4378	11	40	0	9,1	9,1
50-phoneme	4316	2	5	0	70,8	29,2
47-page blocks	4314	5	10	0	90,5	0,5
9-banana	4233	2	2	0	55,2	44,8
55-robot nav sensor r...	4142	4	2	0	42,0	6,3

possibilitou uma estimação de probabilidades com uma resolução que permite 6 valores quando todos os exemplos estão à mesma distância:

$$P_{\theta}(y|\mathbf{x}) \in \{0,0;0,2;0,4;0,6;0,8;1,0\} \forall \mathbf{x} \in X, y \in Y$$

Quanto maior a resolução, mais detalhada é a comparação de medidas de informatividade. O voto de cada vizinho foi ponderado por $1 - d$ para atenuar as estimativas de probabilidade, onde d é a distância euclidiana padronizada (KREYSZIG, 2007).

- C4.5w - A ramificação dos nós de decisão foi múltipla (não apenas binária), com fator de confiança na poda 0,25 e um valor mínimo de 2 exemplos por folha. A

Tabela 6 – Conjuntos de dados com menos exemplos.

Conjunto de dados	$ \mathcal{U} $	$ \mathcal{Y} $	Atributos	Nominais	Majoritária (%)	Minoritária (%)
31-iris	118	3	4	0	34,0	32,7
34-lymphography	118	4	18	15	54,7	1,4
26-hepatitis	124	2	19	13	79,4	20,6
54-robot failure lp5	130	5	90	0	27,8	13,0
87-wine	142	3	13	0	39,9	27,0
48-parkinsons	156	2	22	0	75,4	24,6
25-heart disease va...	159	5	13	0	28,1	5,0
15-connect. mines vs...	166	2	60	0	53,4	46,6
57-seeds	168	3	7	0	33,3	33,3
21-glass	170	6	9	0	35,7	4,2
70-thyroid newthyroid	172	3	5	0	69,8	14,0
59-spect heart	178	2	22	22	55,9	44,1

poda se deu com a *substituição de nós por folhas*⁹. A estimação de probabilidades foi realizada com suavização aditiva (MANNING; RAGHAVAN; SCHÜTZE, 2008), com o objetivo de atenuar as medidas de informatividade.

- NB - Uma discretização supervisionada de atributos numéricos foi realizada antes de cada treinamento (FAYYAD; IRANI, 1993).
- SVM - Tipo C-SVC (CORTES; VAPNIK, 1995) com núcleo função de base radial. Os parâmetros, de acordo com a notação usual na literatura de SVM, seguiram os seguintes valores: $\gamma = 0,5$; $C = 1$; e, $\epsilon = 0,001$ (tolerância do critério de parada).

Um experimento envolvendo comitês também foi realizado. Eles são apresentados na Seção 4.3.1.

4.3.1 Comitês

Um comitê de classificadores é composto por diversos modelos visando um aumento na acurácia preditiva (Seção 2.3.4.2). Um efeito colateral está na geração de estimativas de probabilidade mais tênues por serem baseadas na sumarização de predições feitas por modelos distintos e, em algum grau, independentes. Estratégias de amostragem ativa baseadas na incerteza do modelo podem se beneficiar de predições atenuadas, pois elas podem aumentar seu poder discriminatório durante a comparação de exemplos para consulta. Esse benefício é mais notável quando o algoritmo tende a gerar modelos excessivamente confiantes. Segundo Roy e McCallum (2001), o algoritmo NB, por exemplo, produz estimativas de probabilidade erroneamente elevadas para exemplos de fronteira quando sua premissa de independência entre os atributos é quebrada. Uma solução proposta por esses

⁹ [subtree replacement]

autores é a adoção de um comitê do tipo *bagging* (Seção 2.3.4.2).

Visando aumentar a abrangência dos resultados, optou-se, aqui, pela inclusão de um experimento com comitês. Dois comitês equiparáveis e com desempenho geral elevado (DELGADO *et al.*, 2014) foram escolhidos: *Rotation Forest* (RoF) e *Random Forest* (RFw) (RODRÍGUEZ; KUNCHEVA; ALONSO, 2006; BREIMAN, 2001). A escolha desses dois algoritmos permitiu a realização de um experimento com a acurácia preditiva num patamar superior à que foi obtida no experimento de modelos únicos. Os comitês foram configurados com apenas 10 membros, devido às limitações de recursos computacionais anteriormente mencionadas (Seção 4.2.1).

Nesta tese, comitês também foram aplicados a uma tarefa de regressão, que consistiu na predição de ranqueamentos pelo algoritmo *Predictive Clustering Trees* (PCT) para o sistema de recomendação automática - introduzido na Seção 4.3.3.

4.3.2 Diversidade de algoritmos

O mesmo procedimento experimental da Seção 4.2.2 foi adotado para avaliar a similaridade entre os algoritmos de aprendizado, com exceção da medida de correlação, que foi substituída pela distância de Hamming (HAMMING, 1950). Essa métrica quantifica as predições distintas entre cada par de modelos para uma dada coleção de conjuntos de dados. Ela representa diretamente o grau de equivalência entre algoritmos do ponto de vista da predição de classes. Logo, é mais precisa que a medida anterior (Seção 4.2.2).

O algoritmo 5NN foi testado com duas métricas de distância e com a opção de ponderação. A versão com ponderação (indicada nesta seção pelo sufixo *p*) foi preferida por permitir estimativas de probabilidade atenuadas, necessárias para parte das estratégias de amostragem ativa. As versões com a distância de Manhattan (indicada pelo sufixo *m*) foram descartadas devido à excessiva semelhança com as versões baseadas na distância euclidiana, pois obteve valores 0 e 13 no cálculo da distância de Hamming - conforme Figura 11.

4.3.3 Algoritmos de (meta) aprendizado

É esperado que apenas parte dos 53 meta-atributos definidos na Seção 3.2 seja relevante na tarefa de predição de ranqueamento ou recomendação de algoritmos. Logo, os algoritmos de aprendizado mais apropriados são aqueles capazes de lidar com atributos irrelevantes. Um tipo de algoritmo com essa capacidade frequentemente empregado é o comitê de árvores (STROBL; MALLEY; TUTZ, 2009). Quatro algoritmos baseados em comitê de árvores com relatos de alto desempenho em geral (DELGADO *et al.*, 2014) frente a outros algoritmos são RFw, RoF (Seção 4.3.1), *Predictive Clustering Trees* (PCT) e, não necessariamente baseado em árvores, AdaBoost (ABoo) - (DELGADO *et al.*, 2014; FREUND; SCHAPIRE, 1996; TODOROVSKI; BLOCKEEL; DZEROSKI, 2002). Eles foram adotados para

	SVM	NB	C4.5w	BagNB	RFw	RoF	5NN	5NNm	5NNp	5NNmp
SVM	-	53	56	55	55	54	55	55	56	56
NB	53	-	36	24	36	37	38	37	41	41
C4.5w	56	36	-	33	30	33	36	35	37	37
BagNB	55	24	33	-	30	33	34	33	38	38
RFw	55	36	30	30	-	28	32	31	32	32
RoF	54	37	33	33	28	-	30	30	28	28
5NN	55	38	36	34	32	30	-	13	29	29
5NNm	55	37	35	33	31	30	13	-	30	30
5NNp	56	41	37	38	32	28	29	30	-	0
5NNmp	56	41	37	38	32	28	29	30	0	-

Figura 11 – Média da distância de Hamming entre as predições de modelos induzidos por diferentes algoritmos.

gerar os modelos de recomendação do sistema de meta-aprendizado. O prefixo *meta* é acrescentado para desfazer qualquer ambiguidade entre o nível base e o nível meta - exemplos de uso: metaPCT, EERent-metaRoF, meta-aprendiz, metaclasse, metaexemplos, metaconjunto de dados etc.

Optou-se pela quantidade padrão de 500 membros definido na implementação do algoritmo *Random Forest* para a linguagem R (TEAM, 2014). Esse número foi considerado válido pelos seguintes motivos:

- é importante garantir que a quantidade de membros seja grande o bastante para não haver perda de acurácia devido a falta de membros - 128 árvores foram suficientes para atingir a acurácia máxima numa coleção de 29 conjuntos de dados, segundo Oshiro, Perez e Baranauskas (2012);
- o fator determinante na escolha do tamanho do comitê é o custo computacional (OSHIRO; PEREZ; BARANAUSKAS, 2012);
- o risco de sobreajuste nesse parâmetro é baixo ou inexistente devido ao *paradoxo da generalização do comitê* (ELDER-IV, 2003; SENI; ELDER-IV, 2010);
- é uma quantidade computacionalmente viável para o experimento do nível meta; e,
- sendo um número padrão, facilita a replicabilidade ou comparação com resultados de terceiros.

4.4 Estratégias

Nos experimentos realizados, 9 estratégias e 5 variantes foram avaliadas. O conjunto de estratégias selecionadas foi considerado representativo da diversidade de paradigmas relevantes de aprendizado ativo, de acordo com os artigos referenciados por [Settles \(2012\)](#).

Foram acrescentados os sufixos *euc* e *man* nas siglas de estratégias que façam uso das distâncias euclidiana e de Manhattan, respectivamente. Os sufixos *ent* e *acu* indicam, respectivamente, os critérios de entropia e de acurácia balanceada para a estratégia EER (baseada na *redução do erro esperado*). Dentre as possibilidades iniciais, três estratégias foram descartadas. SVMsim e SVMbal foram implementadas (Seção 2.3.4), mas descartadas dada a sua dependência de um algoritmo específico e seu desempenho excessivamente baixo em testes iniciais - possivelmente devido a um insucesso da adaptação multiclasse escolhida. QBC (*Query By Committee*) foi implementada para o comitê RFW com a medida JS (Seção 2.3.4), mas também descartada pelos mesmos motivos.

Um conjunto reduzido de estratégias foi utilizado na parte inicial da análise comparativa (Seção 5.1.1). Assim, somente uma variante por estratégia foi selecionada para esse primeiro momento. As variantes não incluídas foram as seguintes:

- ATUman, DWman, TUman e HTUman (todas as estratégias baseadas em densidade) - a distância euclidiana foi considerada mais usual que a distância de Manhattan, cujo desempenho, em geral, é similar ([SANTOS; CARVALHO, 2015b](#)); e,
- EERacu - a variante EERent foi preferida porque *E* (entropia) é a medida proposta no trabalho original.

Algumas das estratégias exigiram configurações específicas. As escolhas adotadas nessas configurações são detalhadas nas seções seguintes.

4.4.1 EERent e EERacu

A utilização da acurácia balanceada ([MASSO; VAISMAN, 2010](#)) como função objetivo alternativa configurou-se como uma variação do método (EERacu), visando agir diretamente na medida de interesse nos experimentos (Seção 4.5.1). A acurácia balanceada foi preferida à κ por ser uma medida multiclasse com menos propensão a ter valores próximos ou iguais a zero, evitando, assim, o anulamento da medida de informatividade. Apenas 100 exemplos foram aleatoriamente amostrados de \mathcal{U} antes de cada consulta para reduzir o alto custo computacional de EER.

4.4.2 HS

A implementação original do autor foi empregada neste trabalho com o mesmo algoritmo de agrupamento, *Ward's average linkage method*¹⁰.

4.4.3 DW, TU, ATU e HTU

Os parâmetros de ponderação foram fixados no valor 1 ($\alpha = 1$, $\delta = 1$).

4.5 Avaliação

A avaliação dos resultados experimentais foi realizada qualitativamente por meio de curvas de aprendizado e quantitativamente por meio de medidas de desempenho (Seção 4.5.1). As formas de validação são apresentadas na Seção 4.5.2. A parte quantitativa foi confirmada por testes estatísticos (seções 4.5.3 e 4.5.4).

4.5.1 Medidas

Nos experimentos desta tese, foram necessárias medidas de avaliação de desempenho preditivo específicas para o nível base e para o nível meta. Essas medidas são apresentadas nas seções seguintes.

4.5.1.1 Nível base

Uma medida de desempenho preditivo comumente utilizada na área de aprendizado ativo é a *área abaixo da curva de aprendizado (ALC)*¹¹. Uma das primeiras menções a esse tipo de medida em aprendizado ativo, até onde o conhecimento do autor permite dizer, foi feita por [Raghavan, Madani e Jones \(2007\)](#). A medida ALC foi empregada em diversos artigos, como no trabalho de [Settles e Craven \(2008\)](#) e aplicado em momentos relevantes, como na competição de aprendizado ativo mencionada no Capítulo 1.

A ALC é o valor resultante do somatório de alguma medida de desempenho ao longo das consultas. Seu intuito é avaliar o desempenho geral de uma estratégia, ou seja, em uma ampla faixa de orçamentos. Logo, a medida se mostra adequada para os objetivos desta tese, relacionados à comparação de estratégias. Neste trabalho, ela foi adotada em conjunto com a extensão do índice kappa de Cohen, κ - detalhada na Seção 3.1.2. Na descrição dos resultados, os valores de ALC apresentados são o resultado da divisão da ALC pela quantidade de consultas, visando manter a interpretabilidade da medida.

O índice κ é utilizado para comparar o grau de consenso entre avaliadores ([EUGENIO; GLASS, 2004](#)). Seu poder discriminatório tem sido satisfatório frente a outras medidas

¹⁰ Implementação disponível no Weka ([HALL et al., 2009](#)).

¹¹ [Area under the Learning Curve]

de sumarização da matriz de confusão (DEMIRKESEN; CHERIFI, 2008). Quando aplicado como medida de desempenho de classificação, o valor limite 1 indica acerto total, 0 desempenho equivalente ao aleatório e -1 erro total. Entretanto, nem sempre o limite inferior coincide com -1 (EMAM, 1999). Esse índice permite realizar a comparação sob dois pontos de vista: entre estratégias (valor relativo) e em relação ao acaso (valor absoluto).

Dada a presença de um alto grau de desbalanceamento entre as classes nos conjuntos de dados da coleção, a medida de acurácia convencional seria uma medida de desempenho inadequada, pois permitiria que a classe majoritária dominasse a composição de seu valor final. De fato, as proporções das classes majoritária e minoritária chegam a extremos de, respectivamente, 96,5% e 0,0% (1 exemplo) do tamanho da reserva - conforme Tabela 2.

As curvas de aprendizado e valores de ALC são baseadas nas médias das curvas resultantes de um procedimento de validação cruzada (SETTLES, 2010). Cada curva exibe a medida de interesse κ em função da quantidade de consultas. Assim, neste texto, o valor médio μ_κ de κ para as 5 execuções de validação cruzada em 5 partes (Seção 4.5.2) é a medida base de todos os resultados sobre desempenho relacionado à capacidade preditiva. Como consequência, cada conjunto de dados tem um valor de μ_κ e seu correspondente desvio padrão σ_κ para cada consulta de uma dada estratégia. É importante diferenciar μ_κ (e σ_κ), que é uma média de valores κ interna a cada conjunto de dados, da *média de μ_κ* (e *média de σ_κ*), que é calculada para a coleção como um todo.

4.5.1.2 Nível meta

A avaliação do sistema de recomendação foi realizada conforme dois cenários: classificação e predição de ranqueamento de algoritmos. Seguindo a escolha de Souto *et al.* (2008), a predição pelo ranqueamento médio (Def) foi uma das referências utilizadas na avaliação do sistema de recomendação. O ranqueamento aleatório (Alea) e a escolha fixada na classe majoritária (Maj) também serviram de referência. A comparação entre as predições de ranqueamento foi feita por meio do cálculo do coeficiente de correlação de Spearman (Seção 4.2.2). O desempenho de classificação foi avaliado pela comparação da medida κ e, paralelamente, pelas acurácias ordinária e balanceada na predição de melhor algoritmo de aprendizado. O mesmo método experimental foi adotado na investigação inicial de outras possibilidades de recomendação: de estratégias de amostragem ativa, de pares estratégia-algoritmo, de métricas de distância para estratégias baseadas em densidade e da própria utilização ou não de aprendizado ativo.

Para verificação da efetividade do sistema de recomendação, o efeito do meta-aprendiz também foi avaliado no nível base. O algoritmo para o papel de aprendiz para cada dada estratégia foi escolhido automaticamente e seu desempenho foi comparado diretamente com os demais algoritmos no nível base por meio dos valores de ALC e

curvas de aprendizado.

4.5.2 Validação

Não há consenso na literatura de aprendizado ativo sobre a melhor maneira de validação. Dados os recursos disponíveis, optou-se por 5 execuções de validação cruzada em 5 partes. Ela tem uma menor variabilidade quando comparada à *Leave-One-Out* (LOO) e à validação em 10 partes (KEARNS; RON, 1999) e um menor enviesamento quando comparada, por exemplo, com a validação em 2 partes ou *hold-out* (ARLOT; CELISSE *et al.*, 2010). Adicionalmente, a validação em 5 partes é uma configuração comum na literatura. Algumas das possibilidades encontradas em trabalhos relevantes são as seguintes:

- validação em conjunto à parte pré-definido pelos autores dos conjuntos de dados (ERTEKIN *et al.*, 2007);
- 100 execuções de validação em conjunto com 40% dos exemplos à parte (BARAM; EL-YANIV; LUZ, 2004);
- validação cruzada em 5 partes (CHEN *et al.*, 2015; SETTLES; CRAVEN; FRIEDLAND, 2008; LOMASKY *et al.*, 2007; XU; AKELLA; 0001, 2007);
- 4 execuções de validação cruzada em 5 partes (BECKER; OSBORNE, 2005; MUSLEA; MINTON; KNOBLOCK, 2002);
- validação cruzada em 9 partes (STIKIC; LAERHOVEN; SCHIELE, 2008);
- validação cruzada em 10 partes (SETTLES; CRAVEN; FRIEDLAND, 2008);
- 2 execuções de validação cruzada em 10 partes (KÖRNER; WROBEL, 2006; MELVILLE; MOONEY, 2004); e,
- validação cruzada em 20 partes (MUSLEA; MINTON; KNOBLOCK, 2003).

No nível meta, LOO foi utilizada na avaliação do desempenho na predição de ranqueamentos conforme sugerido por Brazdil, Soares e Costa (2003). Essa abordagem maximizou a utilidade dos dados disponíveis, que correspondem a apenas 90 metaexemplos, um para cada conjunto de dados. Além disso, LOO fornece a estimativa menos enviesada da acurácia quando comparada às demais formas de validação cruzada (JOACHIMS, 2000) e, nesse contexto, permite a aplicação de teste estatístico, conforme discutido na Seção 4.5.4.

Por outro lado, no caso da avaliação do desempenho de classificação no nível meta, o teste estatístico não pôde ser aplicado com confiabilidade pelos motivos descritos na Seção 4.5.4. Optou-se, então, por dez execuções de validação cruzada em dez partes (BOUCKAERT; FRANK, 2004),

Tabela 7 – Exemplo de tabela de sumarização de diferenças estatisticamente significativas. Cada símbolo representa um nível de significância estatística α : * ($\alpha = 0,01$), + ($\alpha = 0,05$) e . ($\alpha = 0,10$).

	1	2	3	4	5
1 - algoritmo 1	-	*			
2 - algoritmo 2		-	+		
3 - algoritmo 3			-	.	
4 - algoritmo 4				-	
5 - algoritmo 5					-

Tabela 8 – Exemplo de contagem de colocações. Medida: $ALC-\sigma_k$. O melhor e o pior valor de cada coluna estão em **negrito azul** e **negrito vermelho** respectivamente. Apenas negrito indica segundo melhor valor.

	Primeiras colocações	Últimas colocações
algoritmo 1	78	13
algoritmo 2	67	91
algoritmo 3	55	24
algoritmo 4	43	44
algoritmo 5	35	44

4.5.3 Teste estatístico no nível base

As diferenças de desempenho preditivo foram atestadas pelo teste não paramétrico de Friedman com teste post-hoc de Nemenyi seguindo a abordagem proposta por Demšar (2006) para comparações de classificadores. Ainda conforme a abordagem do autor daquele estudo, os valores foram arredondados na terceira casa decimal, para forçar o empate nas diferenças irrelevantes. Os resultados do teste são sumarizados numa tabela cujas células contêm símbolos que indicam com que nível de significância estatística uma estratégia na linha vence a outra na coluna. No exemplo com algoritmos hipotéticos dado na Tabela 7, o algoritmo 1 vence o algoritmo 2 com $p\text{-valor} < 0,01$; o algoritmo 2 vence o algoritmo 3 com $p\text{-valor} < 0,05$ e o algoritmo 3 vence o algoritmo 4 com $p\text{-valor} < 0,10$. A contagem de ocorrências de primeiros e últimos lugares é apresentada na Tabela 8. O critério de vitória é baseado na comparação dos valores de ALC de μ_k ($ALC-\mu_k$) para os dois algoritmos em questão em cada conjunto de dados.

4.5.4 Teste estatístico no nível meta

As diferenças estatisticamente significativas entre metaPCT e Def foram reveladas pelo teste de Wilcoxon pareado (DEMŠAR, 2006) aplicado aos valores de correlação entre os ranqueamentos esperados e seus correspondentes ranqueamentos preditos (Seção 4.5.1.2). Cada metaexemplo representa um conjunto de dados. Logo, houve uma aproximação da premissa de *independência entre amostras* devido ao isolamento de cada metaexemplo sob teste - situação intrínseca ao procedimento LOO. Assim, a avaliação da predição de ranqueamento está estatisticamente fundamentada.

Por outro lado, a comparação de acurácias entre múltiplos metaclassificadores não permite aplicar um teste estatístico baseado em ranqueamento, quando há um único metaexemplo no conjunto de teste. O motivo dessa impossibilidade é o valor da predição de classe ser discreto e, consequentemente, só poder ser avaliado como correto ou incorreto.

Outras coleções de conjuntos de dados precisariam ser incorporadas ao experimento, mas isso não é viável com a presente disponibilidade de conjuntos pré-processados. Entretanto, é possível obter valores informativos do metaconjunto que representa a única coleção disponível, caso seja alterada a forma de validação cruzada. A média e o desvio padrão do desempenho preditivo podem ser obtidos por meio de dez execuções de validação cruzada em dez partes.

A troca da forma de validação é preferível, pois o desvio padrão retornado por LOO seria desprovido de significado prático. Ele teria um valor elevado que refletiria o fato dele ser fruto do caso extremo em que o conjunto de teste tem apenas um elemento. Além disso, quando há apenas um conjunto de dados e uma única execução, o desvio padrão obtido por LOO é mera função direta da acurácia. Isso pode ser verificado no desenvolvimento das equações 4.1 e 4.2, onde μ , σ , e $h_i \in \{0;1\}$ são, respectivamente, a acurácia média, seu desvio padrão e o valor correspondente a acerto (1) ou erro (0) para um dado exemplo de índice i . A quantidade de acertos e o número de exemplos são dados por a e n , respectivamente.

$$\mu = an^{-1} \quad (4.1)$$

$$\begin{aligned} \sigma &= \sqrt{\sum_{1 \leq i \leq n} (h_i - \mu)^2 n^{-1}} \\ &= \sqrt{[a(1 - \mu)^2 + (n - a)(0 - \mu)^2]n^{-1}} \\ &= \sqrt{[a - 2a\mu + \mu^2 + (n - a)\mu^2]n^{-1}} \\ &= \sqrt{[a - 2a\mu + a\mu^2 + n\mu^2 - a\mu^2]n^{-1}} \\ &= \sqrt{[a - 2a\mu + n\mu^2]n^{-1}} \\ &= \sqrt{an^{-1} - 2an^{-1}\mu + \mu^2} \\ &= \sqrt{\mu - 2\mu^2 + \mu^2} \\ &= \sqrt{\mu - \mu^2} \end{aligned} \quad (4.2)$$

Adicionalmente, não seria possível contornar o problema por meio de múltiplas execuções, pois o procedimento LOO não dá margem para o acréscimo de perturbações na composição do conjunto de teste.

Note-se que a não aplicabilidade do teste estatístico permanece, mesmo com a mudança na forma de validação cruzada ou seu consequente surgimento da possibilidade de ranqueamento da nova medida gerada. As repetições do processo de validação quebram diretamente a premissa de independência entre amostras em virtude da coleção de conjuntos ser sempre a mesma.

Por fim, até onde o conhecimento do autor permite dizer, testes de significância estatística na comparação de múltiplos classificadores num único (meta)conjunto de dados são um caso omissso na literatura da área (SANTAFE; INZA; LOZANO, 2015; JAPKOWICZ; SHAH, 2011).

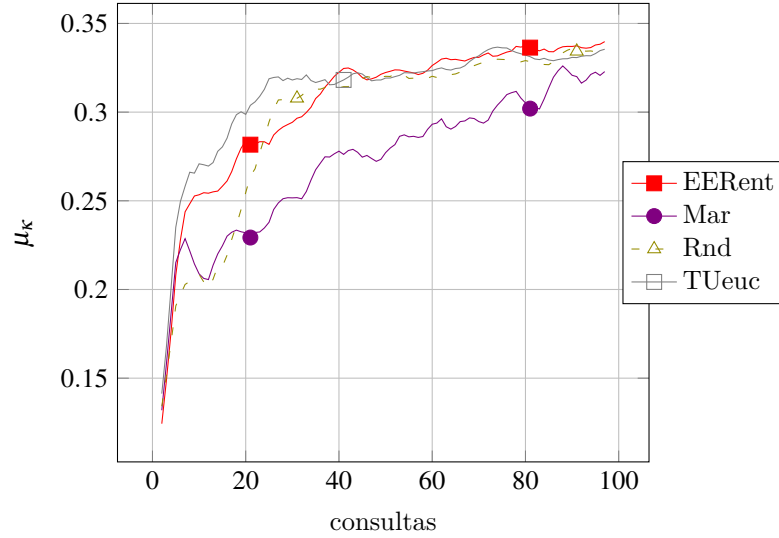


Figura 12 – Curvas de aprendizado para o conjunto *abalone-3class*.

4.6 Curvas de ranqueamento

As curvas de ranqueamento são uma importante contribuição metodológica desta tese. Apesar de serem aqui aplicadas ao aprendizado ativo, elas são também aplicáveis a outros domínios, como a classificação em fluxos de dados. Tradicionalmente, estratégias de amostragem ativa são avaliadas por meio de curvas de aprendizado convencionais (Seção 4.5.1.1). O comportamento típico da curva de aprendizado ativo pode ser visto nas curvas das estratégias EERent, Mar (baseada na margem de incerteza), Rnd (amostragem aleatória) e TUeuc (baseada em densidade que considera exemplos rotulados) com o algoritmo NB para o conjunto de dados *abalone-3class* na Figura 12.

É possível identificar, na figura, os trechos onde uma estratégia supera outra. Entretanto, se mais de um conjunto for considerado, os valores da medida de desempenho podem se tornar incomensuráveis (DEMŠAR, 2006) ou de difícil interpretação, como é o caso da comparação das curvas médias de Mar e TUeuc para toda a coleção, conforme Figura 13.

Comparações de curvas neste gráfico são imprecisas devido aos diferentes pesos que os conjuntos podem ter: conjuntos mais difíceis, por exemplo, tendem a ter valores menores para μ_K e acabam sub-representados. Uma forma de neutralização desse tipo de desigualdade entre os conjuntos é a adoção de um ranqueamento das estratégias de acordo com seus valores μ_K a cada consulta. A média dos ranqueamentos para todos os conjuntos de dados resulta nas curvas da Figura 14. A curva ranqueada exhibe a colocação média da estratégia no total de conjuntos de dados em função da quantidade de consultas realizadas. As curvas foram suavizadas por meio de médias móveis sobre uma janela deslizante de 5 exemplos para melhor visualização.

No caso de um conjunto com mais de um algoritmo de aprendizado a ser adotado

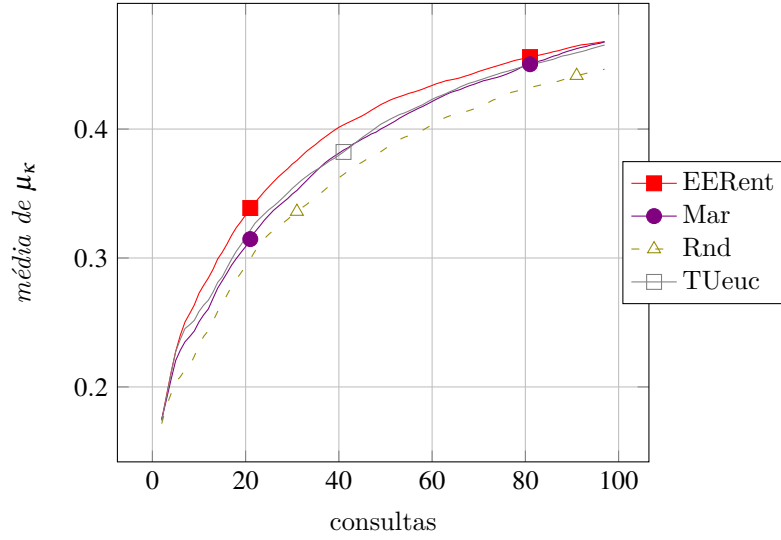


Figura 13 – Curvas de aprendizado médias de EERent, Mar, Rnd e TUEuc.

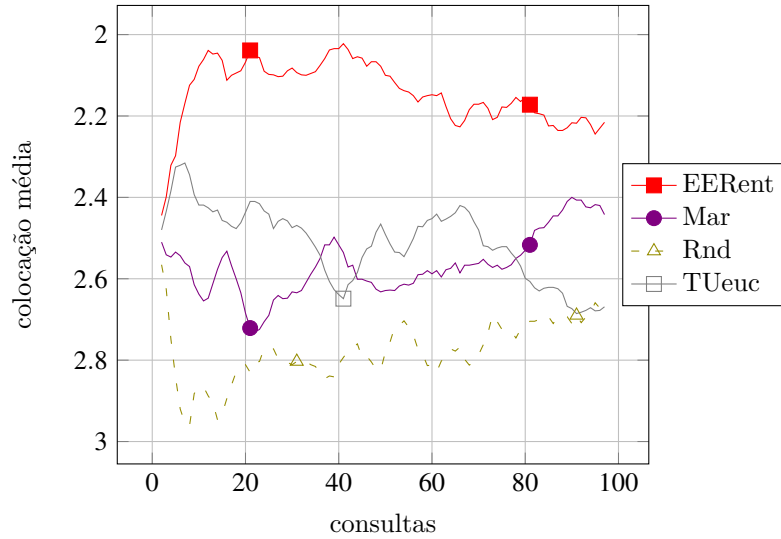


Figura 14 – Curvas de ranqueamento de EERent, Mar, Rnd e TUEuc.

como aprendiz, é possível computar a colocação média para todos os pares conjunto-algoritmo. Outra possibilidade é comparar o desempenho geral das estratégias num mesmo gráfico, omitindo as menos relevantes e indicando com uma faixa por algoritmo as suas colocações mínimas e máximas - conforme Figura 15.

Finalmente, considerou-se mais adequado confirmar eventuais achados durante a comparação de curvas por meio dos testes estatísticos citados na Seção 4.5.3, do que acrescentar às curvas marcadores visuais com intervalos de confiança ou desvio padrão. No entanto, uma consequência dessa escolha é que os testes adotados são mais conservadores do que testes paramétricos.

Duas características intrínsecas ao método podem ser consideradas limitações, se comparadas às curvas de aprendizado convencionais:

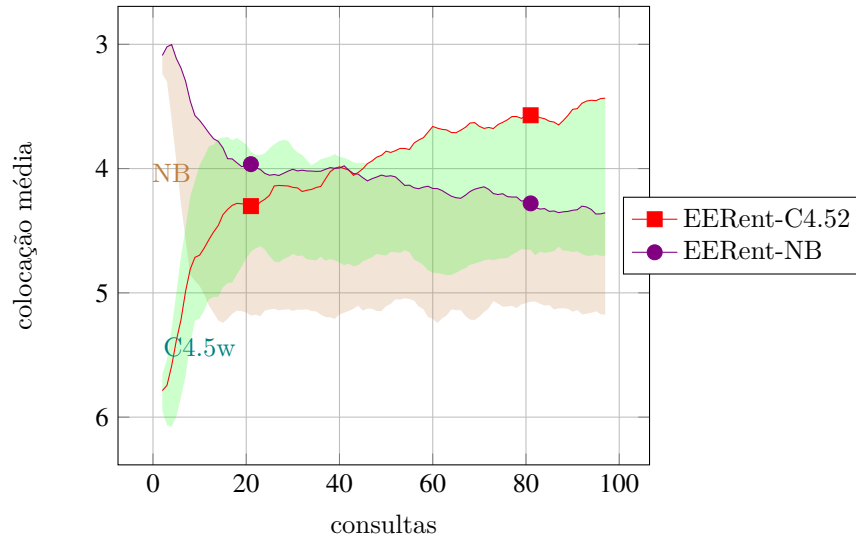


Figura 15 – Exemplo de curvas de ranqueamento com faixas. Medida comparada: μ_k . Cada faixa corresponde a um algoritmo de aprendizado e representa os limites atingidos pelas estratégias. A curva da melhor estratégia de cada algoritmo é explícita.

curvas relativas, cada curva resulta da composição do comportamento de todas as estratégias, logo, uma curva descendente não necessariamente indica uma perda de acurácia preditiva, mas tão somente um desempenho inferior *relativo* em um número crescente de conjuntos de dados - de fato, curvas de aprendizado convencionais raramente apresentam queda, como pode ser exemplificado pela Figura 13; e,

coleções grandes, a interpretabilidade das curvas depende do tamanho da coleção - cada consulta pode resultar numa oscilação abrupta na colocação média, caso haja poucos conjuntos de dados.

4.7 Considerações

Neste capítulo, foi possível delimitar o alcance dos experimentos, lidando com dificuldades como a escassez de conjuntos de dados pré-processados, a inviabilidade prática de adotar muitos algoritmos de aprendizado e as restrições de tempo. Por outro lado, foi possível reduzir a possibilidade de redundância entre os conjuntos selecionados e proporcionar a presença de vieses de aprendizado distintos. Os algoritmos selecionados correspondem a diferentes vieses de busca e representação: baseado em abordagens gulosas capazes de induzir árvores de decisão (C4.5w), baseado em exemplos/distância (k -NN), baseado em probabilidades na indução de um modelo probabilístico (NB) e baseado na teoria do aprendizado estatístico (SVM).

A escolha do cenário, da maneira de validação e dos parâmetros de algoritmos foi apresentada e discutida de acordo com a literatura da área, fornecendo condições para a replicabilidade experimental. As principais características da coleção de conjuntos de

dados adotados foram descritas.

Por fim, uma proposta de técnica de avaliação, chamada *curvas de ranqueamento*, foi apresentada e conceitualmente justificada. Ela foi necessária devido à ausência de outras alternativas viáveis na literatura consultada. Outros aspectos experimentais, como as ferramentas de *software* e recursos computacionais utilizados, estão listados no Apêndice [D](#).

Resultados

*Kleinere
Laboratoriums-Explosionen
werden bei der Natur des
Stoffes, mit dem wir arbeiten,
nie zu vermeiden sein.*

Sigmund Freud¹

As propostas deste trabalho foram avaliadas empiricamente por meio de experimentos comparativos baseados em estratégias de aprendizado ativo e valores de referência da literatura. A coleção consistiu de noventa conjuntos de dados (Capítulo 4). A apresentação dos resultados é dividida em duas partes principais que correspondem às seções 5.1, 5.2, respectivamente:

nível base, em que as estratégias propostas e o efeito do meta-aprendiz utilizando PCT (*Predictive Clustering Trees*) são avaliados - incluindo a investigação de algumas relações entre estratégias, algoritmos de aprendizado e conjuntos de dados; e,

nível meta, que contém evidências da ocorrência de aprendizado na proposta de recomendação automática, tanto com o algoritmo PCT adotado inicialmente, quanto com outros algoritmos de aprendizado utilizados posteriormente - incluindo uma breve análise dos meta-atributos mais relevantes.

Adicionalmente, a Seção 5.3 reporta os resultados da extensão da análise no nível meta a outras modalidades de recomendação.

Considerações gerais são feitas na Seção 5.4. Detalhes sobre os aspectos básicos das técnicas de avaliação envolvidas foram previamente descritos no Capítulo 4. Um breve experimento com um novo conjunto de algoritmos de aprendizado, baseados em comitês, será apresentado, separadamente, no Apêndice A.

¹ “Pela natureza da matéria com que trabalhamos, nunca será possível evitar pequenas explosões de laboratório.” - Freud numa carta a Jung, dando conselhos profissionais e amorosos.

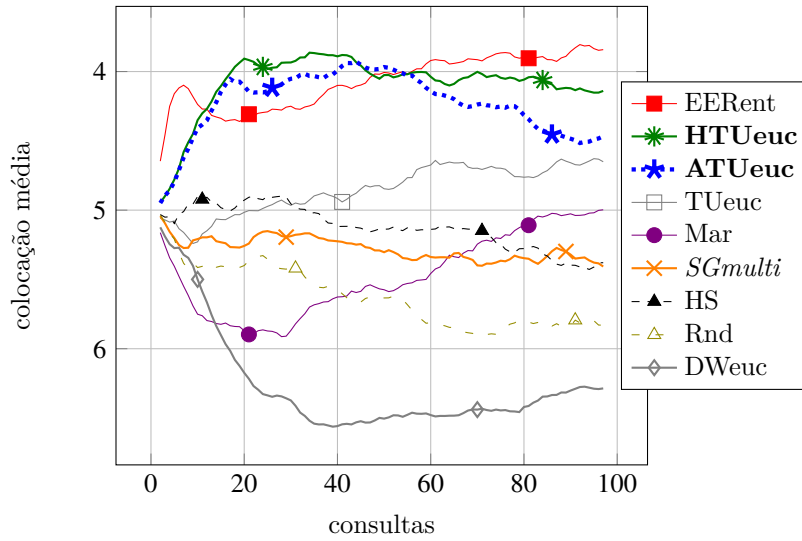


Figura 16 – Curvas de ranqueamento. Medida comparada: μ_{κ} . As siglas das estratégias propostas estão grafadas em **negrito** e a estratégia adaptada para multiclasse em *itálico*. A média considera quatro algoritmos e noventa conjuntos de dados.

5.1 Nível base

As estratégias propostas, ATU (agnóstica baseada em densidade) e HTU (híbrida baseada em densidade), a adaptação para multiclasse SGmulti e aquelas implementadas conforme a literatura, tiveram seu desempenho geral comparado visualmente por meio de curvas de ranqueamento. As diferenças estatisticamente significativas foram reveladas pelo teste de Friedman com teste post-hoc de Nemenyi (Capítulo 4). A avaliação no nível base é dividida nas seguintes seções: desempenho preditivo (Seção 5.1.1), variabilidade (Seção 5.1.2) e custo computacional (Seção 5.1.3). Adicionalmente, na Seção 5.1.4, uma análise separada por algoritmos de aprendizado é realizada. Por fim, na Seção 5.1.5, o meta-aprendiz é situado no contexto do nível base de forma a conduzir aos experimentos no nível meta, das macrosseções posteriores 5.2 e 5.3.

5.1.1 Acurácia preditiva

As curvas de ranqueamento das estratégias EERent², HTUeuc, ATUeuc, TUEuc³, Mar⁴, SGmulti⁵, HS⁶, Rnd⁷ e DWeuc⁸ são exibidas na Figura 16 - toda a coleção de conjuntos de dados foi utilizada. As propostas ATUeuc e HTUeuc compartilham, juntamente com EERent, o topo do gráfico com colocações acima de 4,5 em praticamente todo o período experimental. Exceto por TUEuc, as demais estratégias, inclusive SGmulti,

² baseada no erro esperado

³ baseada em densidade que considera exemplos rotulados

⁴ baseada na margem de incerteza

⁵ proposta multiclasse de busca no espaço de hipóteses

⁶ baseada em agrupamento hierárquico

⁷ amostragem aleatória

⁸ baseada em densidade sem considerar exemplos rotulados

Tabela 9 – Diferenças estatisticamente significativas. Medida: $ALC-\mu_K$. *Legenda na Tabela 7.*

Estratégia	1	2	3	4	5	6	7	8	9
1 - ATUeuc	-	*	*			*	*	*	*
2 - HS		-	*			*	*		
3 - DWeuc			-						
4 - EERent		*	*	-		*	*	*	*
5 - HTUeuc		*	*		-	*	*	*	*
6 - Mar			*			-			
7 - Rnd			*				-		
8 - <i>SGmulti</i>			*					-	
9 - TUeuc			*			*	*		-

Tabela 10 – Contagem de colocações (de 360). Medida: $ALC-\mu_K$. *Detalhes na Tabela 8.*

Estratégia	Primeiras colocações	Derrotas para Rnd	Últimas colocações
EERent	97	92	15
ATUeuc	70	96	21
HTUeuc	64	109	8
DWeuc	37	242	166
TUeuc	31	148	24
Mar	30	187	68
HS	18	129	19
<i>SGmulti</i>	13	88	19
Rnd	11	-	36

se mantiveram abaixo do ranqueamento médio (5ª colocação). Logo, as duas propostas se mostraram competitivas, enquanto que a estratégia adaptada para multiclasse obteve um desempenho abaixo da média. Por outro lado, o bom desempenho em geral, e especialmente a predominância de EERent nas primeiras 10 consultas, pode ser resultado dela ser a única estratégia que requereu, antes de cada consulta, a subamostragem de 100 exemplos explicada previamente na Seção 4.4.1. Essa subamostragem tem, incidentalmente, um efeito exploratório, pois introduz (pseudo)aleatoriedade numa estratégia que, de outra forma, seria determinística.

HTUeuc e ATUeuc superam EERent visualmente durante a maior parte da primeira metade das consultas. Entretanto, após aproximadamente 55 consultas, elas apresentam uma baixa gradual em sua posição no gráfico. Esse efeito é mais intenso para a estratégia ATUeuc - conforme esperado. Tal constatação é compatível com a expectativa de que um viés de amostragem exploratório deixe de ser vantajoso após o trecho inicial do aprendizado. O mesmo efeito pode ser notado nas outras estratégias agnósticas, HS, SGmulti e Rnd. Elas cedem posições para as estratégias prospectivas com o decorrer do aprendizado. De forma correspondente, as curvas de estratégias prospectivas, como TUeuc e Mar, apresentam uma trajetória ascendente conforme aumenta a quantidade de consultas realizadas. Por fim, o destaque negativo foi DWeuc, demonstrando que a ponderação pela densidade pode ser bastante prejudicial em alguns casos. A diferença entre TU (incluindo sucessoras, ATU e HTU) e DW é que a primeira assegura o afastamento do grupo de exemplos rotulados. O afastamento ocorre por meio da inversa ponderação pela densidade desse grupo de exemplos.

Quando considerada a ALC (*Area Under the Learning Curve* - Capítulo 4), não foram detectadas diferenças estatisticamente significativas entre EERent, HTUeuc e ATUeuc - conforme ilustra a Tabela 9. Entretanto, alguns pontos podem ser destacados acerca das propostas com diferenças estatisticamente significativas ($\alpha = 0,01$) - listados a seguir.

- ATUeuc superou todas as outras estratégias agnósticas.
- ATUeuc e HTUeuc superaram sua antecessora comum, TUeuc.
- Ambas também superaram Mar, SGmulti e DWeuc.
- Embora SGmulti tenha superado significativamente apenas a pior estratégia (DWeuc), ela obteve o menor número de derrotas para Rnd, provavelmente por selecionar aleatoriamente exemplos da região de desacordo. Isso está de acordo com a expectativa de que as vantagens de um maior enviesamento em alguns domínios está associada a desvantagens em outros domínios.
- Enquanto nenhuma das outras estratégias empataram com Rnd (Anexo B), SGmulti empatou 112 vezes, devido à sua semelhança com Rnd - conforme contagem na Tabela 10.
- HTUeuc obteve 8 últimas colocações - a menor quantidade de ocorrências dentre as estratégias sob comparação.

Com relação ao patente mau desempenho apresentado na curva de DWeuc, pode-se atribuí-lo à grande quantidade de derrotas para Rnd (242). Entretanto, DWeuc obteve o quarto maior número de primeiras colocações, indicando que uma estratégia de amostragem pode ser positivamente enviesada para alguns conjuntos de dados/algoritmos e negativamente enviesada para muitos outros, em conformidade com o teorema NFL aplicado ao aprendizado ativo (*No Free Lunch* - Seção 1.2).

Um indicativo da viabilidade do aprendizado ativo em geral é que a amostragem aleatória (Rnd) teve o menor número de primeiros lugares e venceu as demais estratégias (exceto DWeuc e Mar) menos vezes do que perdeu. Isto é, a maioria dos valores da coluna *derrotas para Rnd* está abaixo de 180, que é a metade do total de pares conjunto-algoritmo testados. Ainda confirmando a efetividade do aprendizado ativo, HS superou a amostragem aleatória, respeitando, assim, a garantia teórica fornecida pelo seu autor.

Finalmente, há uma limitação nessa avaliação baseada apenas na contagem dos primeiros colocados. Ela ignora estratégias bem posicionadas que não tenham atingido o maior valor da medida de desempenho, mesmo que a diferença para a primeira colocação seja diminuta. Na prática, uma estratégia frequentemente bem colocada é preferível a uma estratégia frequentemente mal colocada que tenha fortuitamente o maior número de primeiras colocações. Portanto, optou-se por apresentar, adicionalmente, uma contagem de vitórias que considerasse os três melhores colocados. É possível verificar na Tabela 11 que HTUeuc, seguida de ATUeuc, se torna preferível a EERent segundo a nova forma de contagem. Além disso, HTUeuc mantém o melhor número de ocorrências entre as três últimas colocações. Consequentemente, pode-se afirmar que, para a coleção, estratégias e

Tabela 11 – Contagem das três melhores e três piores colocações. Medida: $ALC-\mu_K$. Detalhes na Tabela 7.

Estratégia	Primeiras colocações	Derrotas para Rnd	Últimas colocações
HTUeuc	197	109	61
ATUeuc	190	96	74
EERent	189	92	65
TUeuc	131	148	113
HS	95	129	95
Mar	80	187	172
SGmulti	75	88	127
DWeuc	73	242	227
Rnd	66	-	165

algoritmos empregados *HTUeuc é a estratégia mais segura em termos de consistência/regularidade, pois frequentemente está entre as melhores e raramente está entre as piores.*

Outro aspecto da qualidade de uma estratégia de amostragem ativa é a estabilidade de seu desempenho. Os resultados relativos à estabilidade são reportados a seguir, na Seção 5.1.2, por meio do desvio padrão do valor da medida de desempenho preditivo (σ_K).

5.1.2 Variabilidade

Uma estratégia estável deve ser pouco sensível a variações na constituição da reserva de exemplos. Uma grande variabilidade coloca o orçamento sob o risco de ser gasto sem que seja atingido um desempenho adequado e pode levar a custos excedentes de rotulação imprevisíveis. Assim, analisar a estabilidade das estratégias tem fundamental importância na escolha de estratégias. A medida aqui adotada para verificar a variabilidade de desempenho das estratégias é baseada em σ_K , analogamente aos resultados anteriores, baseados em μ_K . A Figura 17 contém as curvas de valores σ_K .

Cada estratégia começou com σ_K abaixo do valor de pico da curva, possivelmente devido à constante presença inicial de um rótulo por classe em todas as 25 reservas da validação cruzada. A partir da vigésima consulta, todas as estratégias passaram a apresentar valores descendentes. Elas se tornavam mais estáveis com o crescimento do conjunto de treinamento. Portanto, pode-se supor que a escolha de estratégias mais estáveis é prioritária em orçamentos mais restritos. Os mesmos resultados são apresentados em termos relativos na Figura 18, por meio das curvas de ranqueamento baseadas nos valores de σ_K .

As curvas ranqueadas de ATUeuc e HTUeuc atingiram as melhores colocações na primeira metade das consultas e foram similares a EERent na segunda metade. Conforme ilustra a Tabela 12, é possível atestar com diferenças estatisticamente significativas ($\alpha = 0,01$) que, para a coleção de conjuntos de dados, estratégias e algoritmos de aprendizado

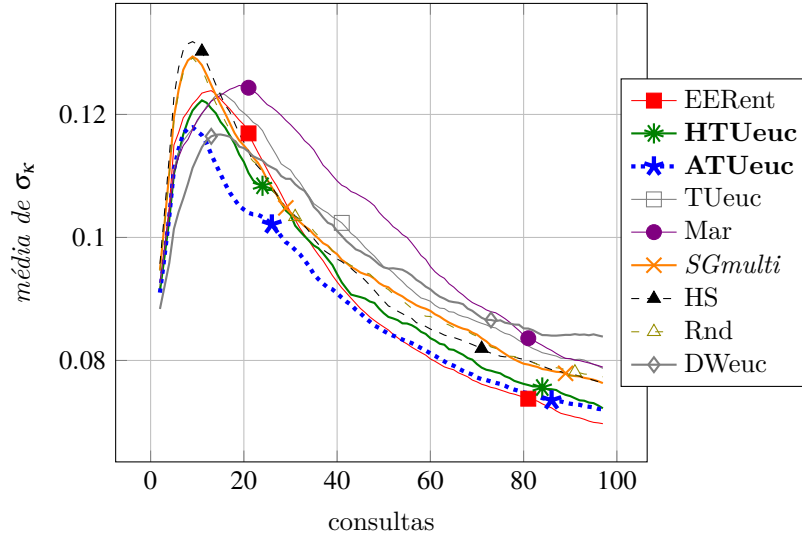


Figura 17 – Curvas representando a variabilidade das estratégias.

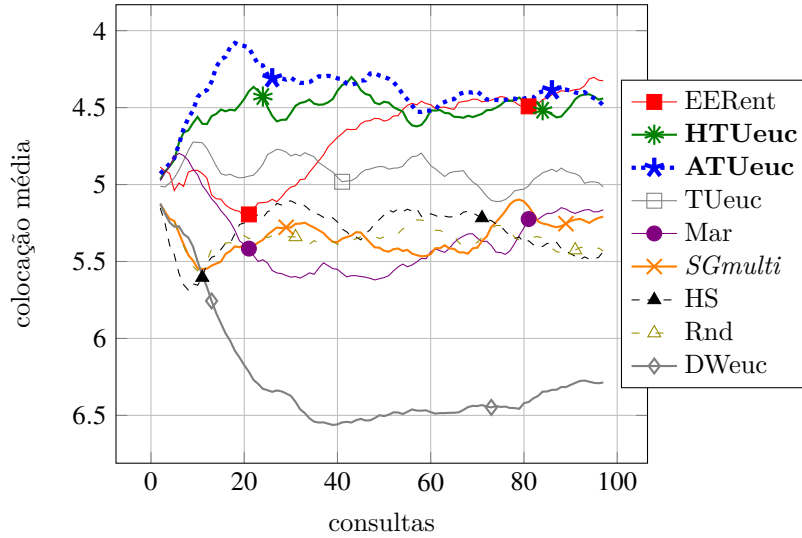


Figura 18 – Curvas de ranqueamento. Medida comparada: σ_k . A colocação média é calculada ao longo de toda a coleção e para os quatro algoritmos de aprendizado.

empregados, *ATUeuc* e *HTUeuc* têm a menor variabilidade.

DWeuc obteve o segundo maior número de primeiras colocações, porém também obteve a maior quantidade de últimas colocações - conforme mostrado na Tabela 13. Uma desvantagem adicional de EERent é seu elevado custo computacional entre consultas. Os custos computacional e de esforço humano são discutidos e avaliados na Seção 5.1.3.

5.1.3 Custo

Idealmente, o tempo requerido por uma estratégia para escolher o próximo exemplo a ser consultado nunca deveria ultrapassar o *tempo de espera tolerável* (NAH, 2003) por um oráculo humano. As consequências vão desde o aumento no consumo do tempo de disponibilidade humana até a redução na confiabilidade do processo de supervisão.

Tabela 12 – Diferenças estatisticamente significativas. Medida: $ALC-\sigma_K$. Legenda na Tabela 7.

Estratégia	1	2	3	4	5	6	7	8	9
1 - ATUeuc	-	*	*	*		*	*	*	*
2 - HS		-							
3 - DWeuc			-						
4 - EERent				-	+				
5 - HTUeuc		*	*	*	-	*	*	*	*
6 - Mar						-			
7 - Rnd							-		
8 - <i>SGmulti</i>								-	
9 - TUeuc						*			-

Tabela 13 – Contagem de colocações. Medida: $ALC-\sigma_K$. Legenda na Tabela 7.

Estratégia	Primeiras colocações	Derrotas para Rnd	Últimas colocações
ATUeuc	78	118	13
DWeuc	67	181	91
HTUeuc	55	130	24
EERent	43	172	44
TUeuc	35	167	44
Mar	29	205	60
HS	29	187	34
Rnd	24	-	34
<i>SGmulti</i>	15	130	33

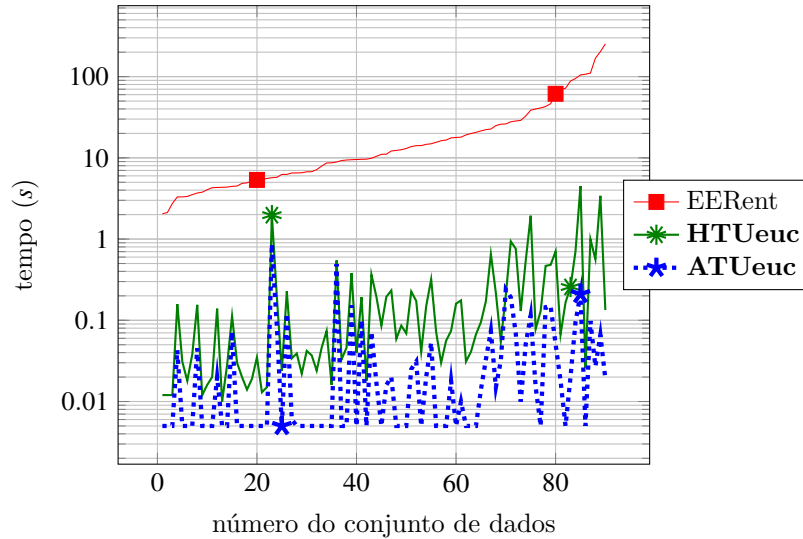


Figura 19 – Média de duração das consultas com os maiores períodos de espera de cada execução do procedimento de validação.

A capacidade humana de processamento de informação pode ser afetada por interferências na memória de curto prazo. De acordo com Miller (1968), um exemplo de interferência é a imposição de um período de espera superior a 2s. Um período de espera inferior a 1s permite que o usuário mantenha o fluxo de pensamento ininterrupto (NAH, 2003). Na Figura 19, a duração média dos períodos de espera mais longos é apresentada para ATUeuc, HTUeuc e EERent. Note-se que EERent foi diretamente beneficiada pela realização da subamostragem que reduziu a reserva para 100 exemplos (Seção 4.4.1) - os dez maiores conjuntos de dados da coleção, por exemplo, têm mais de 5000 exemplos e tornariam o experimento computacionalmente inviável para EER.

Naturalmente, a curva de ATUeuc situa-se sempre abaixo da curva de HTUeuc, devido à dispensa do aprendiz por parte daquela. Contudo, HTU ultrapassou o limite menos crítico (1s) em apenas seis conjuntos de dados, mantendo-se abaixo de 5s em toda

a coleção. Logo, ambas mostraram-se de baixo impacto do ponto de vista de esforço humano. As esperas geradas por EERent, por sua vez, atingiram durações superiores ao limite crítico (2s) em todos os conjuntos de dados e ultrapassaram 100s em seis deles.

Assim, apesar do bom desempenho preditivo, EERent pode aumentar o custo com recursos humanos ou mesmo se tornar inviável com certos conjuntos de dados e algoritmos. Em oposição, estratégias com menor ordem de complexidade reduzem o tempo entre consultas em várias ordens de magnitude. Logo, dentre as três estratégias com melhor desempenho preditivo, *ATUeuc e HTUeuc são consideravelmente mais apropriadas com relação ao custo computacional entre consultas e, conseqüentemente, de esforço humano.*

5.1.4 Acurácia preditiva por algoritmo

Os resultados apresentados nas seções anteriores são provenientes de um cenário que permite avaliar o desempenho das estratégias de amostragem ativa sem fazer distinção entre os algoritmos de aprendizado. Entretanto, é possível que alguns algoritmos sejam intrinsecamente inadequados para certas estratégias. O especialista em aprendizado de máquina, ciente dessa possibilidade, poderia evitar determinadas combinações.

Do ponto de vista experimental, combinações estratégia-algoritmo inadequadas prejudicam a colocação média no ranqueamento, rebaixando estratégias que, de outra forma, poderiam se posicionar entre as melhores. De fato, quando é construída uma árvore de decisão para organizar as estratégias de acordo com seu desempenho em todos os pares conjunto-algoritmo, o algoritmo de aprendizado surge como condição primária; ele é a condição do nó raiz - conforme Figura 20. A árvore foi induzida pelo algoritmo C4.5w com poda de pelo menos 150 exemplos por folha.

Cada nó folha contém a contagem de derrotas de cada algoritmo de acordo com a ALC- μ_K . Os atributos considerados foram: *quantidade de classes, quantidade de atributos, quantidade de exemplos, razão entre quantidade de exemplos e de atributos, percentual de atributos nominais, percentual de exemplos correspondente à classe majoritária, percentual de exemplos correspondente à classe minoritária, razão entre percentual correspondente à classe majoritária e correspondente à classe minoritária e algoritmo de aprendizado.*

Apesar das variantes de DW serem destaque negativo em praticamente todas as folhas, é possível notar indícios de inadequação entre algoritmos e estratégias. Por exemplo, EERent poderia ser, de acordo com a árvore, inadequada para C4.5w, 5NNw ou NB com atributos nominais, enquanto que poderia ser adequada nos demais caminhos, dado que não aparece entre as menos adequadas em nenhuma das outras folhas.

Observação similar poderia ser feita para as propostas baseadas em densidade ou para HS, por exemplo. Logo, uma análise separada por algoritmo mostra-se relevante

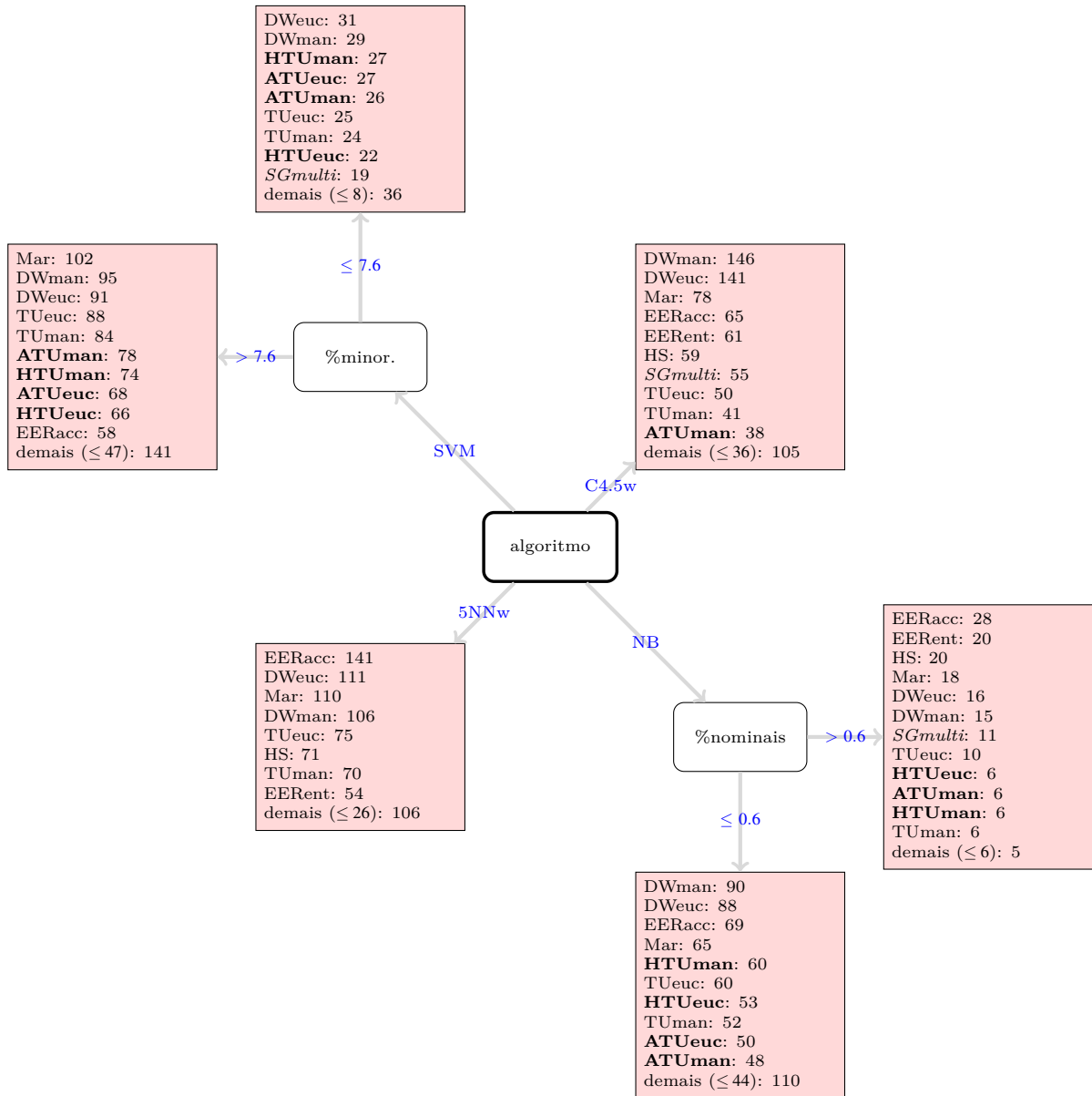


Figura 20 – Possíveis nichos de inadequação conjunto-estratégia-algoritmo. Contagem de ocorrências na última colocação.

para limitar a influência das inadequações estratégia-algoritmo na comparação de desempenhos.

Na Figura 21, é exibida a curva ranqueada da melhor estratégia para cada algoritmo de aprendizado. No início do aprendizado, o comportamento de cada algoritmo foi distinto - conforme evidenciado pelas faixas coloridas que indicam os limites atingidos pelas estratégias. Próximo ao final da curva, após 60 consultas, a faixa de estratégias correspondente ao algoritmo SVM começa a se destacar, indicando a existência de uma relação entre o desempenho do algoritmo e o instante do aprendizado.

A melhor estratégia variou conforme o algoritmo: EERent foi a melhor para SVM e NB, enquanto que HTUeuc e ATUeuc foram as melhores para C4.5w e 5NNw, res-

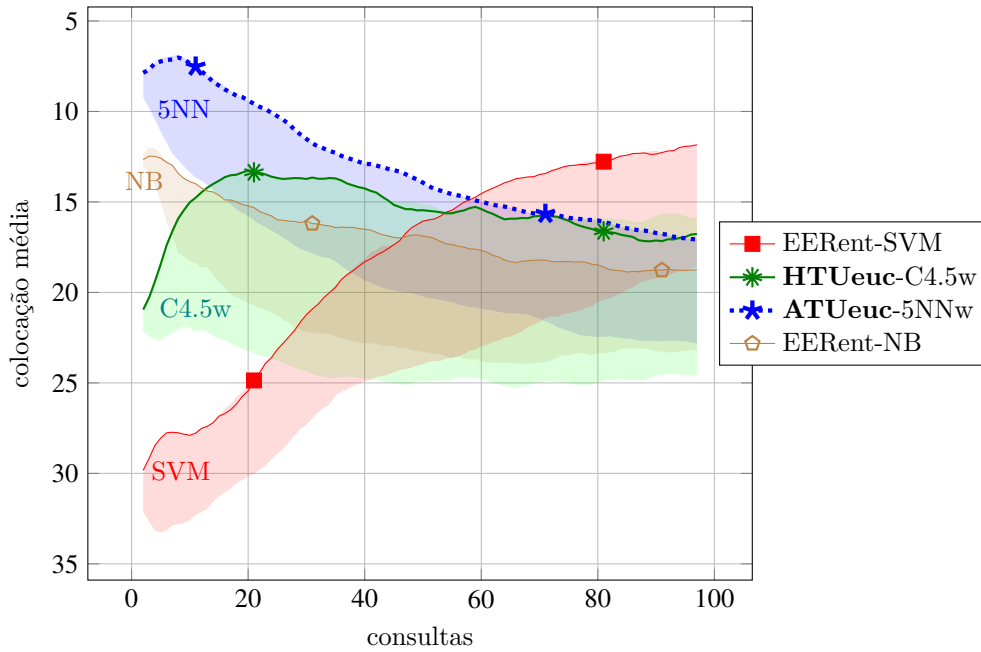


Figura 21 – Curvas de ranqueamento. Medida comparada: μ_{κ} . Detalhes na Figura 15. Cada faixa contém, implicitamente, as curvas de todas as estratégias.

pectivamente. A Figura 22 apresenta as faixas 5NN, C4.5w, NB e SVM separadamente, explicitando as curvas de todas as estratégias. Em geral, ATUeuc, HTUeuc e EERent permanecem como as melhores opções com exceção de alguns casos. SGMulti, por exemplo, se desvinculou positivamente de Rnd com os algoritmos C4.5w e NB e chegou a figurar entre as duas melhores estratégias na segunda metade das consultas com NB.

Dada a diversidade de estratégias, seria esperada uma grande intersecção entre todas as faixas da Figura 21, apresentada previamente. Entretanto, a existência de amplas regiões sem intersecção sugere uma predominância do algoritmo de aprendizado na determinação do desempenho do par estratégia-algoritmo. Consequentemente, seria proveitoso o especialista em aprendizado de máquina dispor de um guia análogo à árvore de decisão da Figura 20 que auxilie na escolha do algoritmo, dada a estratégia; ou, na escolha da estratégia, dado o algoritmo. Assim, uma possível árvore guia para a escolha manual do par estratégia-algoritmo é apresentada na Figura 23.

É possível notar que a distribuição de vitórias muda consideravelmente em cada caminho. Essa característica é a base do sistema automático de recomendação de algoritmos proposto, cujo meta-aprendiz resultante é comparado no cenário de aprendizado ativo com os demais algoritmos, a seguir, na Seção 5.1.5.

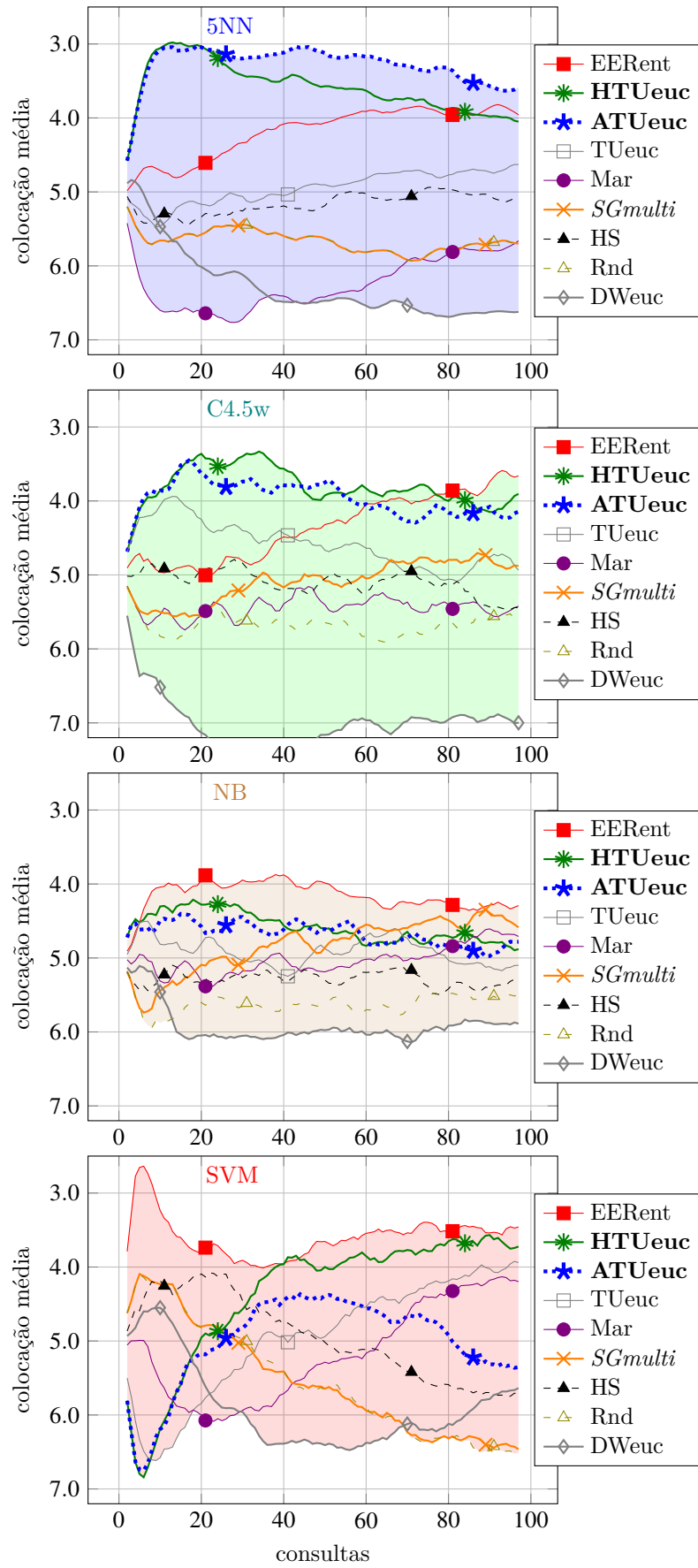


Figura 22 – Curvas de ranqueamento separadas por algoritmo de aprendizado. Medida comparada: μ_K . Detalhes na Figura 15.

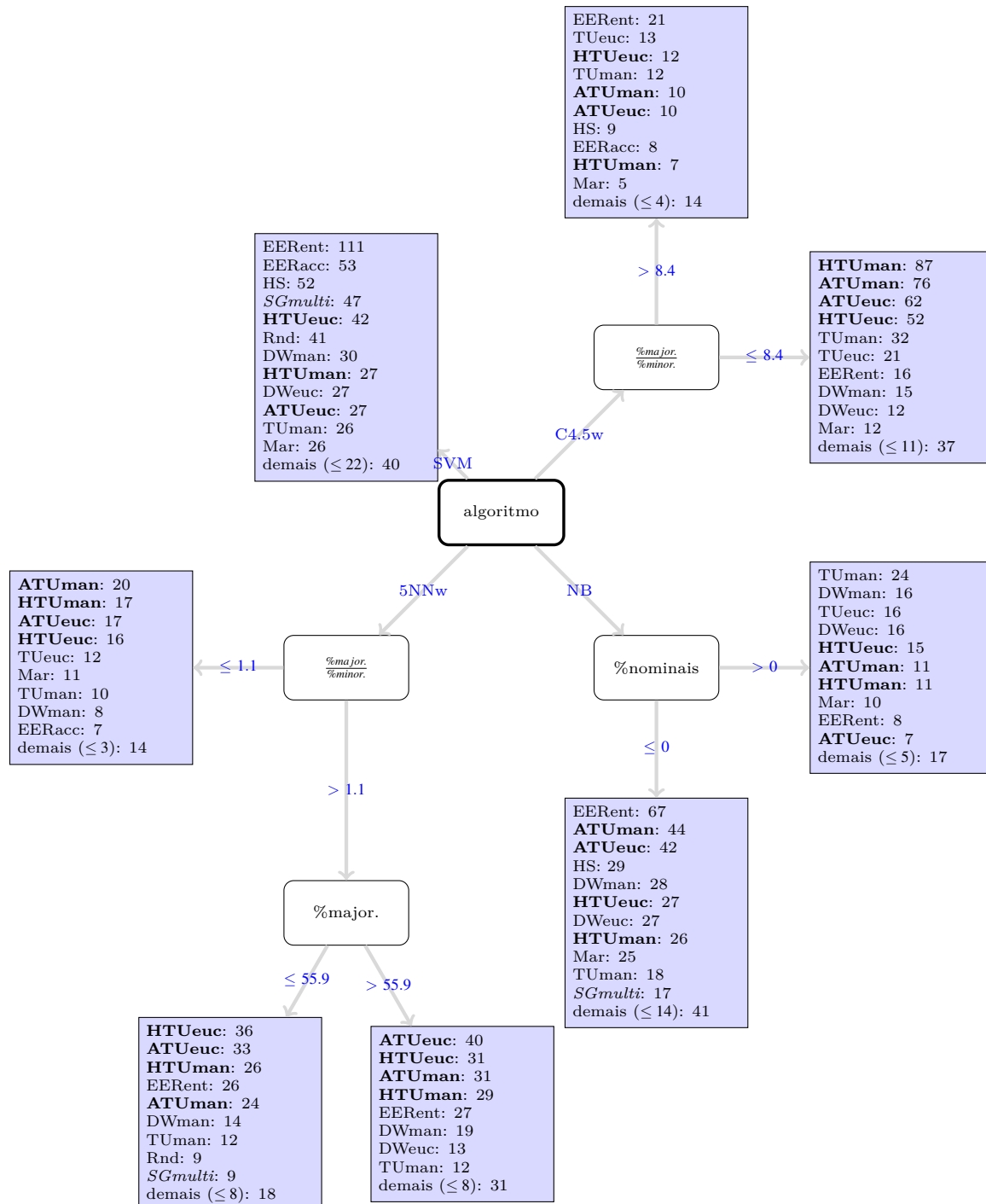


Figura 23 – Possíveis nichos de adequação conjunto-estratégia-algoritmo. Contagem de ocorrências entre as três melhores colocações.

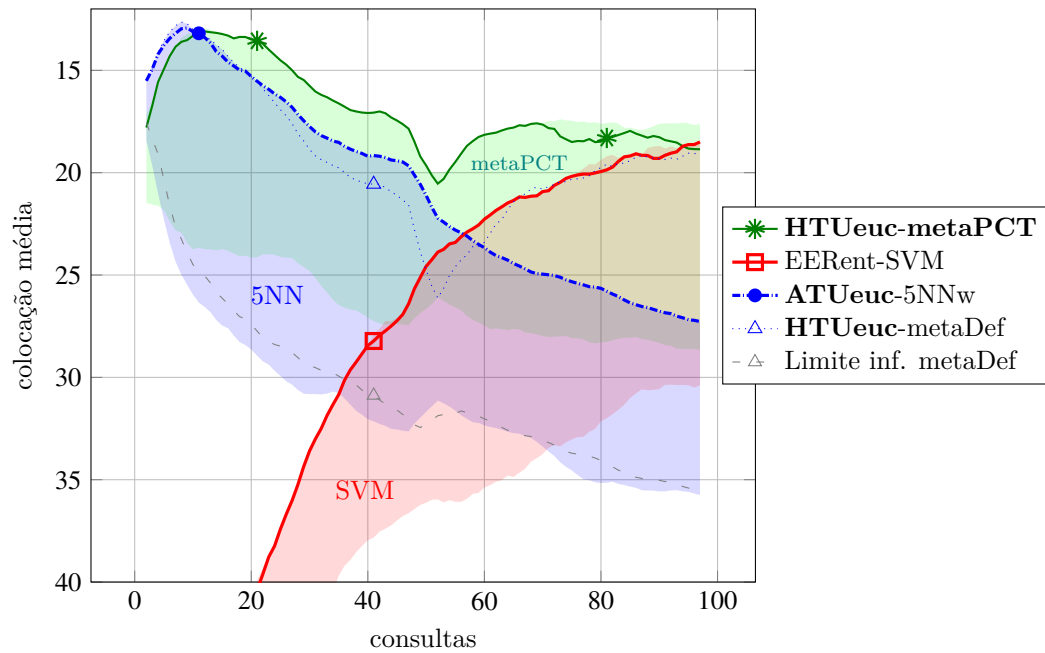


Figura 24 – Curvas de ranqueamento com faixas - incluindo meta-aprendiz. Medida comparada: μ_κ . Referência: meta-aprendiz baseado na predição do ranqueamento médio - *MetaDef*. Detalhes na Figura 21.

5.1.5 Acurácia preditiva com meta-aprendiz

Conforme discutido previamente no Capítulo 1, o especialista frequentemente não dispõe dos recursos necessários para determinar o melhor algoritmo de aprendizado no cenário de aprendizado ativo. Por exemplo, a ausência de um conjunto de treinamento suficientemente grande e o desconhecimento dos efeitos da combinação de uma dada estratégia com um dado conjunto de dados aumentam o risco de uma escolha inadequada. O aprendizado meta-ativo proposto visa reduzir esse risco e tornar menos subjetiva a definição das partes de um sistema de aprendizado ativo. **No presente experimento e nos seguintes, o processo de rotulação foi considerado como dois períodos de consultas que correspondem às suas metades. Isso permitiu que o meta-aprendiz pudesse reconsiderar, após 50 consultas, sua escolha de algoritmo.**

A Figura 24 contém a faixa de curvas de ranqueamento do meta-aprendiz (metaPCT) comparada às duas faixas de maior destaque no experimento anterior, 5NN e SVM. O experimento é análogo ao anterior: cada faixa contém as curvas de todas as estratégias. Consequentemente, o meta-aprendiz fez uma recomendação para cada estratégia e cada período de consultas.

A faixa metaPCT está visualmente melhor posicionada que as demais em praticamente todo o intervalo de consultas. A estratégia que tirou maior proveito do meta-aprendiz foi HTUeuc. Ela superou os pares estratégia-algoritmo ATUeuc-5NN e EERent-SVM, outrora vencedores. Por outro lado, o melhor par contendo o meta-aprendiz referên-

Estratégia	PCT	Def	α	p-valor
HTUeuc	0,45	0,44	*	0,005
HTUman	0,45	0,43	+	0,022
TUman	0,45	0,43	*	0,001
EERent	0,45	0,43	*	0,001
ATUeuc	0,44	0,43	+	0,020
TUeuc	0,44	0,42	+	0,027
ATUman	0,44	0,43	+	0,029
HS	0,43	0,41	+	0,016
Mar	0,43	0,41	*	0,004
Rnd	0,42	0,41	+	0,012
EERacc	0,42	0,40	*	0,004
SGmulti	0,42	0,40		0,104
DWeuc	0,43	0,39	*	0,000
DWman	0,41	0,38	*	0,002
Vitórias	14	0		

Figura 25 – Comparação de valores de $ALC-\mu_k$ obtidos com meta-aprendizes. *Legenda na Tabela 7. Obs.: uma quantidade expressiva de vitórias não significa que os resultados possam ser generalizados a outras coleções, pois todas as linhas são altamente correlacionadas entre si.*

cia (HTUeuc-metaDef) foi equivalente ou inferior a eles. Essa superioridade de metaPCT em relação à referência indica que o sistema de meta-aprendizado foi efetivamente capaz de extrair conhecimento da coleção. A significância estatística α das diferenças observadas é confirmada na Figura 25 para todas as estratégias, exceto SGmulti.

Assim, é possível afirmar que, para a coleção, estratégias de amostragem ativa e algoritmos de aprendizado empregados *a curva de aprendizado resultante de algoritmos de aprendizado recomendados automaticamente corresponde a um melhor desempenho preditivo do que a curva produzida por um algoritmo fixo, mesmo que o desempenho deste seja predominante na maioria dos conjuntos de dados da coleção - desde que a coleção contenha mais do que um conjunto de um mesmo domínio* (conforme demonstrado posteriormente no Apêndice B).

Outros algoritmos de aprendizado também podem ser empregados como meta-aprendizes. Essa possibilidade é avaliada juntamente com os experimentos no nível meta, na Seção 5.2.

5.2 Nível meta - Recomendação de algoritmos

Previamente, no nível base (Seção 5.1), a efetividade prática das estratégias propostas foi experimentalmente verificada pela avaliação direta de suas curvas de aprendizado, incluindo a adoção do meta-aprendiz como alternativa ao uso de um algoritmo de aprendizado específico. No nível meta, entretanto, a avaliação se concentra na verificação da ocorrência de aprendizado, isto é, se a capacidade preditiva do meta-aprendiz supera os métodos de referência: Def (ranqueamento médio), Maj (predição da classe majoritária) e

Alea (predição aleatória). O algoritmo de recomendação foi avaliado com foco nas PCT, pelo fato de sua implementação disponível viabilizar a predição de ranqueamento (Seção 5.2.1). No entanto, outros algoritmos também foram adotados como meta-aprendizes na tarefa de predição do melhor algoritmo de aprendizado para uma dada estratégia de amostragem ativa, na Seção 5.2.2. Adicionalmente, na Seção 5.2.3, a relevância dos meta-atributos é investigada.

O mesmo esquema de reconsideração de escolhas do sistema de recomendação da Seção anterior foi adotado: o processo de rotulação foi considerado como dois períodos de consultas que correspondem às suas metades do processo de rotulação. Logo, para cada estratégia, duas metapredições foram feitas.

5.2.1 Ranqueamento de algoritmos de aprendizado

As diferenças estatisticamente significativas entre PCT e Def foram reveladas pelo teste de Wilcoxon aplicado a 90 valores de correlação entre ranqueamentos esperados e preditos. Esses valores correspondem aos testes do procedimento LOO (*Leave-One-Out* - Seção 4.5.4) aplicado a toda a coleção de conjuntos de dados. O nível de significância estatística é dado na Figura 26, onde cada linha corresponde a um experimento completo, sendo que, em cada experimento, uma estratégia diferente foi adotada. O número sobrescrito indica a faixa de orçamentos: “HTUeuc¹”, por exemplo, indica que a predição do ranqueamento é referente às primeiras 50 consultas; “HTUeuc²” remete às 50 últimas consultas da mesma estratégia.

O meta-aprendiz PCT obteve valores superiores à referência Def com todas as 28 estratégias, a maioria com diferenças estatisticamente significativas, reforçando assim a constatação de superioridade do sistema de recomendação.

Na metade inferior da figura, os valores de correlação são sistematicamente mais baixos para ambos metaclassificadores e muito próximos de zero para Def. Essa metade corresponde às últimas 50 consultas do período de aprendizado (sobrescrito 2). Supõe-se que, à medida que o aprendizado avança, a colocação média de cada algoritmo de aprendizado (aprendiz ativo) se torna mais próxima das demais, levando a um ranqueamento médio pouco representativo - conforme previamente ilustrado pela crescente intersecção entre as faixas na Figura 21. Apesar dessa crescente dificuldade, nota-se também que, na segunda metade, o meta-aprendiz proposto se distingue ainda mais de Def, pois os p -valores são mais baixos, ocasionando uma maior frequência de $\alpha = 0,01$ (símbolo *) que na metade superior. Sucintamente, no panorama geral, PCT superou Def.

Consequentemente, é possível afirmar que, considerando a medida de desempenho adotada no nível base (μ_K) e dada a coleção de conjuntos de dados e os algoritmos adotados, para a maioria das estratégias ***o meta-aprendiz foi capaz de induzir modelos que efetivamente representam conhecimento presente na relação entre os***

Estratégia	PCT	Def	α	p-valor
ATUman ¹	0,55	0,52		0,158
HTUeuc ¹	0,54	0,48	+	0,012
ATUeuc ¹	0,56	0,49	*	0,003
TUeuc ¹	0,52	0,46	*	0,002
HTUman ¹	0,53	0,50		0,284
TUman ¹	0,53	0,41	*	0,000
SGmulti ¹	0,51	0,47	.	0,053
Rnd ¹	0,50	0,46		0,109
DWeuc ¹	0,53	0,47	*	0,007
Mar ¹	0,48	0,42	*	0,002
EERacc ¹	0,38	0,33	.	0,060
HS ¹	0,48	0,35	*	0,000
DWman ¹	0,51	0,48	.	0,077
EERent ¹	0,42	0,38		0,326
EERacc ²	0,31	0,19	*	0,004
DWman ²	0,31	0,17	*	0,001
DWeuc ²	0,32	0,11	*	0,000
HTUeuc ²	0,25	0,13	*	0,008
HTUman ²	0,25	0,11	*	0,007
EERent ²	0,24	0,14	+	0,019
TUeuc ²	0,25	0,12	+	0,020
HS ²	0,24	0,05	*	0,000
ATUeuc ²	0,21	0,09	+	0,013
SGmulti ²	0,21	-0,01	*	0,000
Rnd ²	0,22	-0,03	*	0,000
TUman ²	0,23	0,02	*	0,000
ATUman ²	0,20	-0,03	*	0,000
Mar ²	0,23	0,02	*	0,000
Vitórias	28	0		

Figura 26 – Comparação de valores do coeficiente de correlação de Spearman na predição de ranqueamento de algoritmos de aprendizado. *Detalhes na Figura 25.*

meta-atributos propostos e as metaclasses⁹; em oposição ao que seria esperado se os modelos houvessem sido gerados ao acaso.

A avaliação experimental prossegue, na Seção 5.2.2, com a inclusão dos algoritmos RoF (*Rotation Forest*), RFW (*Random Forest*) e ABoo (*AdaBoost*) enquanto meta-aprendizes, visando a comparação de desempenho no nível meta na tarefa de predição de classes.

5.2.2 Recomendação de algoritmos de aprendizado

A recomendação do melhor algoritmo distancia-se da predição de ranqueamentos e aproxima-se da tarefa de classificação. Ela visa a recomendação do melhor algoritmo apenas, descartando informações sobre as demais colocações. Diferentemente da abordagem metodológica da Seção 5.2.1, o método de validação LOO, e o subsequente teste de

⁹ Desde que a coleção contenha mais do que um conjunto de um mesmo domínio (conforme demonstrado posteriormente no Apêndice B)

Estratégia	RoF		RFw		PCT		ABoo		Maj		Alea	
DWeuc ¹	0,67	/ 0,00	0,66	/ 0,01	0,64	/ 0,02	0,55	/ 0,05	0,59	/ 0,00	0,30	/ 0,05
SGmulti ¹	0,65	/ 0,01	0,67	/ 0,02	0,66	/ 0,01	0,54	/ 0,03	0,57	/ 0,00	0,29	/ 0,03
DWman ¹	0,65	/ 0,02	0,64	/ 0,01	0,61	/ 0,02	0,56	/ 0,04	0,61	/ 0,00	0,29	/ 0,05
Rnd ¹	0,61	/ 0,02	0,60	/ 0,01	0,62	/ 0,01	0,53	/ 0,02	0,59	/ 0,00	0,29	/ 0,04
HS ¹	0,61	/ 0,01	0,62	/ 0,02	0,60	/ 0,01	0,55	/ 0,04	0,56	/ 0,00	0,28	/ 0,03
ATUman ¹	0,61	/ 0,02	0,62	/ 0,02	0,61	/ 0,02	0,51	/ 0,02	0,58	/ 0,00	0,28	/ 0,03
ATUeuc ¹	0,62	/ 0,01	0,60	/ 0,01	0,62	/ 0,01	0,53	/ 0,06	0,56	/ 0,00	0,29	/ 0,03
HTUeuc ¹	0,62	/ 0,01	0,64	/ 0,01	0,62	/ 0,03	0,53	/ 0,07	0,51	/ 0,00	0,27	/ 0,04
HTUman ¹	0,60	/ 0,02	0,62	/ 0,04	0,56	/ 0,03	0,56	/ 0,04	0,51	/ 0,00	0,27	/ 0,03
TUeuc ¹	0,62	/ 0,02	0,60	/ 0,02	0,61	/ 0,01	0,48	/ 0,05	0,48	/ 0,00	0,28	/ 0,05
Mar ¹	0,62	/ 0,01	0,59	/ 0,01	0,62	/ 0,02	0,45	/ 0,08	0,48	/ 0,00	0,28	/ 0,05
TUman ¹	0,54	/ 0,00	0,54	/ 0,02	0,56	/ 0,02	0,48	/ 0,04	0,49	/ 0,00	0,29	/ 0,05
EERent ¹	0,57	/ 0,02	0,53	/ 0,01	0,52	/ 0,02	0,46	/ 0,04	0,49	/ 0,00	0,28	/ 0,03
EERacc ¹	0,53	/ 0,01	0,49	/ 0,01	0,47	/ 0,00	0,38	/ 0,04	0,43	/ 0,00	0,27	/ 0,05
DWeuc ²	0,50	/ 0,02	0,49	/ 0,03	0,48	/ 0,03	0,36	/ 0,06	0,39	/ 0,00	0,25	/ 0,08
DWman ²	0,49	/ 0,02	0,44	/ 0,04	0,48	/ 0,01	0,38	/ 0,04	0,39	/ 0,00	0,25	/ 0,09
Mar ²	0,48	/ 0,03	0,45	/ 0,03	0,47	/ 0,02	0,38	/ 0,03	0,40	/ 0,00	0,25	/ 0,06
EERacc ²	0,44	/ 0,03	0,42	/ 0,03	0,44	/ 0,03	0,39	/ 0,04	0,42	/ 0,00	0,25	/ 0,04
TUeuc ²	0,43	/ 0,01	0,41	/ 0,02	0,41	/ 0,02	0,35	/ 0,04	0,39	/ 0,00	0,24	/ 0,06
TUman ²	0,40	/ 0,02	0,44	/ 0,01	0,42	/ 0,04	0,36	/ 0,02	0,34	/ 0,00	0,25	/ 0,06
HTUeuc ²	0,40	/ 0,03	0,38	/ 0,01	0,35	/ 0,03	0,36	/ 0,03	0,39	/ 0,00	0,24	/ 0,07
HTUman ²	0,37	/ 0,03	0,36	/ 0,02	0,36	/ 0,01	0,33	/ 0,06	0,32	/ 0,00	0,26	/ 0,09
EERent ²	0,37	/ 0,02	0,37	/ 0,03	0,33	/ 0,03	0,34	/ 0,03	0,33	/ 0,00	0,25	/ 0,07
SGmulti ²	0,40	/ 0,02	0,37	/ 0,02	0,38	/ 0,03	0,32	/ 0,04	0,29	/ 0,00	0,24	/ 0,07
HS ²	0,39	/ 0,03	0,37	/ 0,03	0,35	/ 0,04	0,29	/ 0,02	0,32	/ 0,00	0,26	/ 0,08
ATUman ²	0,33	/ 0,03	0,36	/ 0,02	0,36	/ 0,01	0,34	/ 0,03	0,29	/ 0,00	0,25	/ 0,09
ATUeuc ²	0,36	/ 0,02	0,35	/ 0,02	0,35	/ 0,02	0,31	/ 0,06	0,30	/ 0,00	0,25	/ 0,08
Rnd ²	0,37	/ 0,02	0,36	/ 0,01	0,33	/ 0,02	0,35	/ 0,07	0,29	/ 0,00	0,23	/ 0,09
Ranq. médio	1,5		2,0		2,3		4,5		4,1		6,0	

Figura 27 – Comparação da acurácia (média/desvio padrão) na predição do melhor algoritmo de aprendizado. *Obs.: uma boa colocação no ranqueamento médio não implica em generalidade dos resultados, pois todas as linhas são altamente correlacionadas.*

significância estatística, não puderam ser utilizados na comparação entre os algoritmos PCT, RoF, RFw e ABoo e as referências Maj e Alea - dada a natureza do experimento. Mais detalhes foram expostos previamente na Seção 4.5.4.

A acurácia média e o desvio padrão de cada algoritmo utilizado como meta-aprendiz, resultantes de dez execuções de validação cruzada em dez partes, são apresentados na Figura 27. As linhas estão ordenadas a começar pelas acurácias mais altas. Essa ordenação praticamente corresponde ao decréscimo dos valores para Maj, que representa a proporção de exemplos da metaclassa majoritária. Essa proporção influencia diretamente o patamar das demais acurácias.

Tendo em vista os ranqueamentos médios, na última linha da figura, com exceção de ABoo, os meta-aprendizes alternativos obtiveram desempenho superior ao obtido pelo algoritmo PCT. Isso favorece uma avaliação positiva do sistema de recomendação, pois o desempenho preditivo de PCT já havia sido reportado, na Seção 5.2.1, com superioridade

Estratégia	RoF		RFw		PCT		ABoo		Maj		Alea	
DWeuc ¹	0,36	/ 0,02	0,38	/ 0,01	0,35	/ 0,05	0,25	/ 0,09	0,00	/ 0,00	0,04	/ 0,03
SGmulti ¹	0,38	/ 0,03	0,36	/ 0,03	0,35	/ 0,02	0,20	/ 0,08	0,00	/ 0,00	0,06	/ 0,04
DWman ¹	0,34	/ 0,03	0,37	/ 0,04	0,34	/ 0,03	0,20	/ 0,07	0,00	/ 0,00	0,07	/ 0,02
Rnd ¹	0,38	/ 0,02	0,33	/ 0,02	0,37	/ 0,03	0,15	/ 0,12	0,00	/ 0,00	0,06	/ 0,04
HS ¹	0,33	/ 0,03	0,35	/ 0,06	0,23	/ 0,05	0,29	/ 0,07	0,00	/ 0,00	0,05	/ 0,03
ATUman ¹	0,33	/ 0,00	0,31	/ 0,03	0,26	/ 0,03	0,21	/ 0,07	0,00	/ 0,00	0,08	/ 0,04
ATUeuc ¹	0,28	/ 0,02	0,30	/ 0,03	0,28	/ 0,02	0,22	/ 0,07	0,00	/ 0,00	0,06	/ 0,02
HTUeuc ¹	0,29	/ 0,02	0,25	/ 0,02	0,27	/ 0,03	0,18	/ 0,09	0,00	/ 0,00	0,07	/ 0,02
HTUman ¹	0,25	/ 0,01	0,24	/ 0,03	0,28	/ 0,03	0,20	/ 0,08	0,00	/ 0,00	0,07	/ 0,05
TUeuc ¹	0,26	/ 0,04	0,27	/ 0,04	0,24	/ 0,04	0,15	/ 0,03	0,00	/ 0,00	0,06	/ 0,03
Mar ¹	0,29	/ 0,02	0,26	/ 0,04	0,25	/ 0,04	0,12	/ 0,07	0,00	/ 0,00	0,00	/ 0,10
TUman ¹	0,23	/ 0,04	0,20	/ 0,02	0,23	/ 0,03	0,17	/ 0,04	0,00	/ 0,00	0,07	/ 0,04
EERent ¹	0,28	/ 0,05	0,20	/ 0,03	0,19	/ 0,03	0,14	/ 0,06	0,00	/ 0,00	0,06	/ 0,02
EERacc ¹	0,27	/ 0,03	0,19	/ 0,06	0,23	/ 0,02	0,13	/ 0,06	0,00	/ 0,00	0,00	/ 0,12
DWeuc ²	0,25	/ 0,04	0,19	/ 0,04	0,24	/ 0,03	0,14	/ 0,05	0,00	/ 0,00	-0,01	/ 0,07
DWman ²	0,23	/ 0,06	0,20	/ 0,04	0,13	/ 0,04	0,14	/ 0,07	0,00	/ 0,00	0,06	/ 0,03
Mar ²	0,27	/ 0,02	0,20	/ 0,02	0,15	/ 0,01	0,05	/ 0,05	0,00	/ 0,00	0,06	/ 0,06
EERacc ²	0,17	/ 0,03	0,22	/ 0,02	0,21	/ 0,05	0,12	/ 0,03	0,00	/ 0,00	-0,01	/ 0,08
TUeuc ²	0,19	/ 0,03	0,14	/ 0,02	0,16	/ 0,03	0,09	/ 0,06	0,00	/ 0,00	-0,01	/ 0,08
TUman ²	0,16	/ 0,05	0,12	/ 0,05	0,16	/ 0,04	0,11	/ 0,07	0,00	/ 0,00	0,00	/ 0,05
HTUeuc ²	0,17	/ 0,02	0,14	/ 0,03	0,14	/ 0,04	0,08	/ 0,05	0,00	/ 0,00	-0,02	/ 0,08
HTUman ²	0,14	/ 0,03	0,12	/ 0,02	0,09	/ 0,03	0,12	/ 0,10	0,00	/ 0,00	-0,01	/ 0,11
EERent ²	0,13	/ 0,02	0,13	/ 0,04	0,07	/ 0,04	0,10	/ 0,04	0,00	/ 0,00	0,00	/ 0,09
SGmulti ²	0,15	/ 0,04	0,13	/ 0,04	0,10	/ 0,06	0,03	/ 0,02	0,00	/ 0,00	0,01	/ 0,10
HS ²	0,12	/ 0,04	0,10	/ 0,02	0,09	/ 0,02	0,07	/ 0,07	0,00	/ 0,00	0,01	/ 0,12
ATUman ²	0,08	/ 0,04	0,11	/ 0,03	0,11	/ 0,02	0,10	/ 0,03	0,00	/ 0,00	0,00	/ 0,12
ATUeuc ²	0,11	/ 0,03	0,08	/ 0,02	0,09	/ 0,03	0,06	/ 0,09	0,00	/ 0,00	0,00	/ 0,11
Rnd ²	0,11	/ 0,05	0,07	/ 0,01	0,04	/ 0,05	0,07	/ 0,04	0,00	/ 0,00	-0,02	/ 0,09
Ranq. médio	1,4		2,0		2,5		3,8		5,6		5,2	

Figura 28 – Comparação de κ (média/desvio padrão) na recomendação do melhor algoritmo de aprendizado. Detalhes na Figura 27.

estatisticamente significativa com relação à referência Def. Logo, é razoável estender essa constatação a RoF e RFw. A validade dos valores de acurácia é reforçada pelos valores relativamente baixos de desvio padrão, quando consideradas as diferenças entre as médias atingidas pelos algoritmos alternativos e as médias atingidas pelos algoritmos de referência. De fato, a superioridade dos meta-aprendizes é confirmada, de maneira mais clara, e incluindo ABoo, por meio de medidas que equilibrem o peso de todas as classes, como a acurácia balanceada e κ .

A relativa manutenção do comportamento dos valores de ranqueamento médio nas três figuras 27, 28 e 29 aponta para a existência de redundância pelo menos parcial entre as suas respectivas medidas: acurácia, acurácia balanceada e κ . Por simplicidade, apenas a medida κ é reportada nas seções subsequentes, pois ela tem a vantagem de considerar a possibilidade do algoritmo acertar ao acaso, ou seja, ela já incorpora em seu valor a comparação com os métodos de referência. Isso pode ser comprovado pelo desempenho praticamente nulo nas duas últimas colunas, referentes a esses métodos, na Figura 28. Na

Estratégia	RoF		RFw		PCT		ABoo		Maj		Alea	
DWeuc ¹	0,42	/ 0,01	0,42	/ 0,01	0,41	/ 0,03	0,38	/ 0,06	0,25	/ 0,00	0,30	/ 0,12
SGmulti ¹	0,43	/ 0,02	0,41	/ 0,01	0,41	/ 0,01	0,34	/ 0,04	0,25	/ 0,00	0,32	/ 0,11
DWman ¹	0,44	/ 0,01	0,43	/ 0,04	0,44	/ 0,03	0,34	/ 0,05	0,25	/ 0,00	0,25	/ 0,08
Rnd ¹	0,41	/ 0,02	0,42	/ 0,02	0,40	/ 0,02	0,35	/ 0,04	0,25	/ 0,00	0,32	/ 0,12
HS ¹	0,43	/ 0,01	0,40	/ 0,01	0,42	/ 0,02	0,33	/ 0,06	0,25	/ 0,00	0,31	/ 0,11
ATUman ¹	0,40	/ 0,01	0,41	/ 0,02	0,37	/ 0,02	0,39	/ 0,04	0,25	/ 0,00	0,31	/ 0,11
ATUeuc ¹	0,41	/ 0,03	0,38	/ 0,03	0,41	/ 0,02	0,36	/ 0,04	0,25	/ 0,00	0,24	/ 0,05
HTUeuc ¹	0,38	/ 0,01	0,37	/ 0,01	0,37	/ 0,01	0,34	/ 0,04	0,25	/ 0,00	0,33	/ 0,11
HTUman ¹	0,37	/ 0,01	0,38	/ 0,01	0,37	/ 0,01	0,35	/ 0,03	0,25	/ 0,00	0,32	/ 0,11
TUeuc ¹	0,41	/ 0,02	0,38	/ 0,04	0,40	/ 0,01	0,33	/ 0,05	0,25	/ 0,00	0,25	/ 0,10
Mar ¹	0,36	/ 0,01	0,35	/ 0,01	0,38	/ 0,02	0,34	/ 0,04	0,25	/ 0,00	0,33	/ 0,10
TUman ¹	0,36	/ 0,00	0,36	/ 0,01	0,34	/ 0,01	0,35	/ 0,07	0,25	/ 0,00	0,33	/ 0,13
EERent ¹	0,36	/ 0,02	0,35	/ 0,01	0,36	/ 0,02	0,33	/ 0,02	0,25	/ 0,00	0,33	/ 0,13
EERacc ¹	0,36	/ 0,02	0,36	/ 0,01	0,35	/ 0,02	0,32	/ 0,02	0,25	/ 0,00	0,32	/ 0,12
DWeuc ²	0,36	/ 0,02	0,39	/ 0,01	0,39	/ 0,03	0,33	/ 0,03	0,25	/ 0,00	0,24	/ 0,05
DWman ²	0,38	/ 0,02	0,34	/ 0,01	0,34	/ 0,02	0,32	/ 0,03	0,25	/ 0,00	0,32	/ 0,11
Mar ²	0,40	/ 0,02	0,37	/ 0,01	0,37	/ 0,02	0,32	/ 0,04	0,25	/ 0,00	0,24	/ 0,07
EERacc ²	0,36	/ 0,02	0,36	/ 0,02	0,33	/ 0,02	0,35	/ 0,07	0,25	/ 0,00	0,23	/ 0,09
TUeuc ²	0,37	/ 0,01	0,34	/ 0,01	0,32	/ 0,01	0,27	/ 0,02	0,25	/ 0,00	0,31	/ 0,13
TUman ²	0,36	/ 0,02	0,35	/ 0,01	0,33	/ 0,01	0,31	/ 0,04	0,24	/ 0,00	0,26	/ 0,09
HTUeuc ²	0,35	/ 0,01	0,33	/ 0,02	0,34	/ 0,02	0,31	/ 0,03	0,25	/ 0,00	0,24	/ 0,05
HTUman ²	0,33	/ 0,03	0,31	/ 0,03	0,34	/ 0,02	0,32	/ 0,05	0,25	/ 0,00	0,25	/ 0,04
EERent ²	0,36	/ 0,03	0,34	/ 0,03	0,32	/ 0,04	0,27	/ 0,02	0,25	/ 0,00	0,26	/ 0,08
SGmulti ²	0,34	/ 0,02	0,33	/ 0,03	0,31	/ 0,02	0,32	/ 0,02	0,25	/ 0,00	0,25	/ 0,05
HS ²	0,31	/ 0,04	0,34	/ 0,02	0,34	/ 0,02	0,33	/ 0,02	0,23	/ 0,00	0,25	/ 0,08
ATUman ²	0,31	/ 0,01	0,30	/ 0,01	0,28	/ 0,01	0,28	/ 0,02	0,25	/ 0,00	0,32	/ 0,13
ATUeuc ²	0,32	/ 0,02	0,30	/ 0,01	0,30	/ 0,02	0,29	/ 0,06	0,23	/ 0,00	0,25	/ 0,08
Rnd ²	0,30	/ 0,03	0,29	/ 0,01	0,28	/ 0,02	0,29	/ 0,04	0,25	/ 0,00	0,24	/ 0,06
Ranq. médio	1,4		2,0		2,4		3,8		5,6		4,9	

Figura 29 – Comparação da acurácia balanceada (média/desvio padrão) na recomendação do melhor algoritmo de aprendizado. *Detalhes na Figura 27.*

mesma figura, é possível notar novamente a superioridade das colocações médias de RoF e RFw sobre PCT.

Com relação à acurácia balanceada, apenas dois dos meta-aprendizes foram superados em algum momento por uma das referências, conforme pode ser visto na Figura 29: PCT com a estratégia DWman obteve menor acurácia balanceada do que Alea; e, ABoo com as estratégias EERacc e DWman obteve menor acurácia balanceada do que Alea.

Finalmente, pode-se concluir que *o desempenho do sistema de recomendação proposto pode ser melhorado com a troca do meta-aprendiz, pois não se restringe ao uso do algoritmo PCT nesse papel*, embora ABoo, especificamente, tenha se mostrado pouco recomendável quando comparado aos demais, dentro das limitações impostas pelo arcabouço experimental.

Experimentos com outras modalidades de recomendação serão reportados na Seção 5.3. Os meta-atributos são analisados a seguir, na Seção 5.2.3.

Tabela 14 – Relevância dos meta-atributos na recomendação de algoritmos de aprendizado. *Obtida para a primeira metade das consultas de HTUeuc com o meta-aprendiz RFW.*

Meta-atributo	Relevância	Meta-atributo	Relevância	Meta-atributo	Relevância
%no	0,0389	du _{h1.5}	0,0207	du _{k2}	0,0150
#no _{max}	0,0344	$\sigma_{min/max}$	0,0196	σ_{max}	0,0150
en _{min}	0,0307	si _{h2}	0,0195	μ_{mea}	0,0149
ku _{min}	0,0294	ρ_h	0,0193	si _h	0,0147
du _{k1}	0,0279	sk _{min}	0,0189	si _{k1.5}	0,0146
ρ_{min}	0,0278	sk _{min/max}	0,0187	#at	0,0145
en _{mea}	0,0273	$\rho_{h1.5}$	0,0182	en _{max}	0,0144
#no _{mea}	0,0271	du _{h2}	0,0179	lgex	0,0139
$\rho_{min/max}$	0,0257	si _k	0,0176	$\sigma_{min/max}$	0,0132
#no _{min}	0,0245	$\rho_{k1.5}$	0,0172	ku _{mea}	0,0129
cn _{h2}	0,0235	#ea	0,0169	$\mu_{min/max}$	0,0123
ρ_{mea}	0,0235	ρ_{k2}	0,0169	ku _{max}	0,0119
μ_{min}	0,0228	lgea	0,0169	μ_{max}	0,0119
du _{k1.5}	0,0224	si _{h1.5}	0,0165	#nc	0,0113
cn _{k1}	0,0222	σ_{mea}	0,0163	sk _{max}	0,0112
du _{h1}	0,0215	sk _{mea}	0,0159	#no _{min/max}	0,0078
si _{k2}	0,0212	#ex	0,0157	ρ_{max}	0,0005
σ_{min}	0,0212	ku _{min/max}	0,0154		

5.2.3 Análise dos meta-atributos

É possível analisar quais aspectos são mais importantes na caracterização dos conjuntos de dados, por meio da quantificação da relevância de cada atributo para o meta-aprendiz. Em comitês de árvores de decisão, a relevância¹⁰ de cada meta-atributo (BREIMAN *et al.*, 1984) é calculada por meio da média das reduções que ele causa na impureza, dadas pelas diferenças no *índice Gini* entre nós pais e nós filhos durante a indução do meta-aprendiz RFW. A impureza de cada nó é dada pelo índice Gini - conforme Equação 5.1.

$$G_n = \sum_{y \in Y} P_n(y)[1 - P_n(y)] \quad (5.1)$$

A probabilidade $P_n(y)$ indica a proporção de exemplos da classe y na sub-árvore representada pelo nó n .

A Tabela 14 contém os valores de relevância dos meta-atributos em ordem decrescente. Os dois meta-atributos com os maiores valores de relevância são humanamente interpretáveis: percentual de atributos nominais (%no) e quantidade máxima de atributos nominais (#no_{max}). Os próximos cinco são medidas estatísticas a respeito da distribuição dos dados (entropia, curtose, correlação) e uma medida de agrupamento.

A análise de uma árvore de decisão isolada permite verificar diretamente a dis-

¹⁰ A implementação do algoritmo RF presente na biblioteca *scikit learn* foi utilizada no cálculo da relevância de atributos (PEDREGOSA *et al.*, 2011).

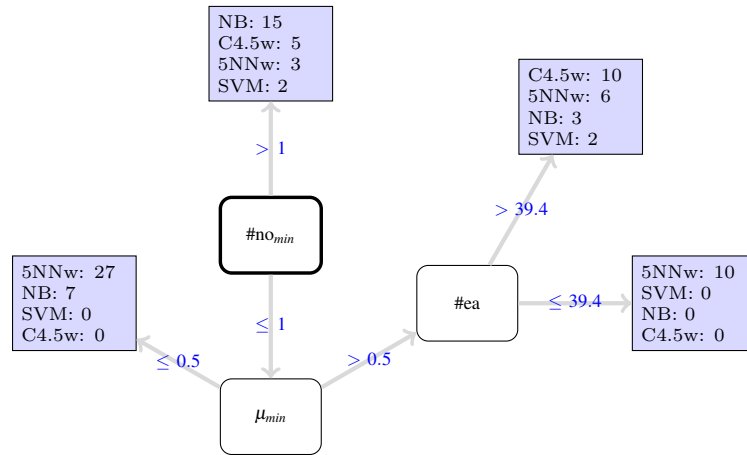


Figura 30 – Árvore com os meta-atributos mais relevantes de acordo com o meta-aprendiz C4.5w.

tribuição das predições de acordo com os meta-atributos mais relevantes. Uma árvore induzida pelo algoritmo (meta)C4.5w com no mínimo 10 exemplos por folha é apresentada na Figura 30. A árvore contém os meta-atributos mais relevantes. Ela foi induzida utilizando todos os metaexemplos.

Os meta-atributos, naturalmente, não coincidem totalmente com os mais relevantes segundo as estimativas do meta-aprendiz RFW (com 500 árvores), mas são igualmente representativos, dado o mesmo valor de acurácia (0,62) de (meta)C4.5w e RFW obtido na validação cruzada em dez partes.

Algumas observações podem ser feitas. O algoritmo NB, seguido de C4.5w, obteve mais vitórias (15) do que todos os demais juntos (10) quando *havia mais de um valor em atributos nominais* ($\#no_{min} > 1$). Essa constatação é compatível com a natureza de NB e C4.5w, pois são capazes de processar atributos nominais diretamente - diferentemente de SVM e 5NNw. O outro ramo da raiz agrupa os conjuntos que contêm apenas atributos numéricos, sendo que o algoritmo 5NNw obteve as 10 vitórias do caminho $\#no_{min} \rightarrow \mu_{min} \rightarrow \#ea \leq 39,4$. Outro caminho com clara capacidade discriminatória é $\#no_{min} \rightarrow \mu_{min} \leq 0,5$, em que apenas 5NNw (27) e NB (7) obtêm vitórias. Em ambos os casos, não foi encontrada uma relação evidente entre o caminho e a folha. Por fim, nota-se que o comitê RFW com 500 membros poderia, nesse caso, ser substituído por uma árvore isolada, como o (meta)C4.5w, sem perda de acurácia. Apenas três meta-atributos foram suficientes para caracterizar os conjuntos de dados: $\#no_{min}$, μ_{min} e $\#ea$. Uma característica comum entre eles é não serem medidas específicas de agrupamento de dados.

A presente análise não foi estendida à recomendação de pares devido à sua semelhança com a recomendação de algoritmos de aprendizado, causada pela forte influência do algoritmo no par. Nas demais modalidades de recomendação, as árvores isoladas geradas pelo algoritmo (meta)C4.5w não superaram a acurácia de referência.

Algoritmo	PCT	Def	α	p -valor
5NN ¹	0,22	0,21		0,980
C4.5w ¹	0,19	0,18		0,400
NB ¹	0,17	0,14		0,235
SVM ¹	0,25	0,22		0,299
5NN ²	0,15	0,13		0,591
C4.5w ²	0,16	0,12		0,197
NB ²	0,19	0,12	+	0,039
SVM ²	0,22	0,24		0,909
Vitórias	7	1		

Figura 31 – Comparação do coeficiente de correlação de Spearman na predição de ranqueamento de estratégias. *Detalhes na Figura 25.*

5.3 Nível meta - Outras modalidades

O mesmo aparato de caracterização de conjuntos de dados foi testado em outras modalidades de recomendação automática, como segue:

- predição da melhor estratégia para um dado algoritmo de aprendizado (Seção 5.3.1);
- predição do melhor par estratégia-algoritmo (Seção 5.3.2);
- predição da melhor métrica para estratégias baseadas em distância (Seção 5.3.3); e,
- recomendação sobre a aplicação ou não de aprendizado ativo (Seção 5.3.4).

De todas essas possibilidades testadas, a recomendação de estratégias e de pares mostraram-se as modalidades mais promissoras.

5.3.1 Recomendação de estratégias

Uma investigação inicial sobre a possibilidade da recomendação automática de estratégias de amostragem ativa, dado um algoritmo de aprendizado, foi realizada. Trata-se de uma abordagem, de certa forma, complementar à contemplada na Seção 5.2.2.

O conjunto de estratégias foi arbitrariamente reduzido para o seguinte, visando uma quantidade reduzida de metaclasses, mas representando as diferentes abordagens estudadas: EERent, HS e HTUeuc. Considerou-se que as estratégias baseadas em densidade ou incerteza foram representadas por HTUeuc e as estratégias agnósticas por HS. A ausência de aprendiz da estratégia HS e seu peculiar viés de exploração colaboram para a diversidade do conjunto.

Os valores de correlação média entre os ranqueamentos preditos e os esperados são apresentados na Figura 31. De acordo com esses resultados, PCT superou Def com significância estatística na predição de ranqueamentos somente para o algoritmo NB, durante a segunda metade do período de aprendizado (2 sobrescrito). Entretanto, por situar-se entre

Algoritmo	RoF		RFw		PCT		ABoo		Maj		Alea	
C4.5w ²	0,34	/ 0,03	0,30	/ 0,05	0,31	/ 0,07	0,18	/ 0,10	0,00	/ 0,00	-0,02	/ 0,09
C4.5w ¹	0,32	/ 0,05	0,29	/ 0,03	0,32	/ 0,03	0,17	/ 0,07	0,00	/ 0,00	-0,01	/ 0,06
NB ¹	0,19	/ 0,02	0,22	/ 0,05	0,18	/ 0,05	0,08	/ 0,07	0,00	/ 0,00	-0,03	/ 0,09
NB ²	0,19	/ 0,03	0,15	/ 0,04	0,20	/ 0,04	0,11	/ 0,04	0,00	/ 0,00	-0,01	/ 0,09
5NN ²	0,17	/ 0,05	0,09	/ 0,05	0,17	/ 0,04	0,06	/ 0,11	0,00	/ 0,00	-0,03	/ 0,06
SVM ¹	0,06	/ 0,05	0,12	/ 0,04	0,10	/ 0,04	0,04	/ 0,08	0,00	/ 0,00	-0,01	/ 0,06
SVM ²	0,01	/ 0,08	0,10	/ 0,03	0,06	/ 0,02	0,07	/ 0,12	0,00	/ 0,00	-0,01	/ 0,08
5NN ¹	-0,06	/ 0,03	0,07	/ 0,03	0,04	/ 0,02	0,01	/ 0,14	0,00	/ 0,00	0,00	/ 0,09
Ranq. médio	2,5		2,0		1,9		3,6		4,9		5,8	

Figura 32 – Comparação de κ (média/desvio padrão) na recomendação da melhor estratégia. *Detalhes na Figura 27.*

múltiplas comparações, essa vantagem isolada pode ser resultado do acaso devido à possibilidade de ocorrência do erro de tipo I global (DEMŠAR, 2006). Logo, embora superior numericamente em sete casos, não é possível afirmar com significância estatística, diante desses resultados, que o meta-aprendizado tenha sido vantajoso. O sucesso da proposta, assim, é menos evidente do que na Figura 26, apresentada previamente, correspondente à recomendação de algoritmos.

A viabilidade da recomendação de estratégias é mais clara na tarefa de predição de metaclasses - conforme ilustrado na Figura 32. Os meta-aprendizes obtiveram desempenho inferior à referência ($\kappa = 0$) em apenas um caso. PCT atingiu a melhor colocação média dentre os quatro meta-aprendizes considerados. Porém, um aspecto negativo nesses resultados é que os piores desempenhos preditivos no nível meta ocorreram com os algoritmos de maior destaque no nível base: SVM e 5NN. Isso significa que o nível meta pode ter um bom desempenho preditivo para algoritmos possivelmente inadequados - reduzindo a utilidade prática da metaestratégia. Apesar disso, de acordo com os resultados apresentados e consideradas as limitações do aparato experimental adotado, pode-se dizer que **a recomendação de estratégias de amostragem ativa é promissora**¹¹, embora represente um maior desafio que a recomendação de algoritmos.

5.3.2 Recomendação de pares estratégia-algoritmo

Um desdobramento natural da investigação realizada nas seções anteriores é a recomendação automática de pares estratégia-algoritmo. Considerando os quatro algoritmos do nível base e as três estratégias adotadas como opções na recomendação de estratégias, tem-se um total de doze pares possíveis. Apesar dessa grande quantidade de metaclasses, o meta-aprendiz foi capaz de superar Def na predição de ranqueamento de pares nas duas metades da curva de aprendizado com $\alpha = 0,01$. Isso pode ser visto nas duas linhas contendo os termos *Todos os pares* na Figura 33.

¹¹ Desde que a coleção contenha mais do que um conjunto de um mesmo domínio (conforme demonstrado posteriormente no Apêndice B)

Pares disponíveis

	PCT	Def	α	p-valor
HTUeuc-5NN,HTUeuc-SVM,HS-5NN,HS-SVM ¹	0,64	0,63		0,104
EERent-5NN,EERent-SVM,HTUeuc-5NN,HTUeuc-SVM ¹	0,60	0,60		0,576
EERent-5NN,EERent-SVM,HTUeuc-5NN,HTUeuc-SVM,HS-5NN,HS-SVM ¹	0,60	0,59	+	0,032
HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM,HS-5NN,HS-NB,HS-SVM ¹	0,57	0,54	*	0,002
EERent-5NN,EERent-SVM,HS-5NN,HS-SVM ¹	0,55	0,55		0,317
<i>HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-C4.5w,HS-SVM¹</i>	<i>0,52</i>	<i>0,54</i>	<i>+</i>	<i>0,030</i>
EERent-5NN,EERent-NB,EERent-SVM,HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM,HS-5NN,HS-NB,HS-SVM ¹	0,52	0,50		0,115
EERent-5NN,EERent-NB,EERent-SVM,HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM ¹	0,52	0,50		0,276
EERent-5NN,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM ¹	0,51	0,51		0,750
EERent-5NN,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-C4.5w,HS-SVM ¹	0,50	0,51		0,926
EERent-5NN,EERent-NB,EERent-SVM,HS-5NN,HS-NB,HS-SVM ¹	0,47	0,47		0,973
EERent-5NN,EERent-C4.5w,EERent-SVM,HS-5NN,HS-C4.5w,HS-SVM ¹	0,47	0,47		0,783
HTUeuc-C4.5w,HTUeuc-SVM,HS-C4.5w,HS-SVM ¹	0,45	0,48		0,360
EERent-C4.5w,EERent-SVM,HTUeuc-C4.5w,HTUeuc-SVM ¹	0,46	0,47		0,353
HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-NB,HS-C4.5w,HS-SVM ¹	0,49	0,44	*	0,001
HTUeuc-NB,HTUeuc-SVM,HS-NB,HS-SVM ¹	0,45	0,46		0,811
EERent-NB,EERent-SVM,HTUeuc-NB,HTUeuc-SVM ¹	0,44	0,45		0,774
Todos os pares¹	0,46	0,42	*	0,002
EERent-C4.5w,EERent-SVM,HTUeuc-C4.5w,HTUeuc-SVM,HS-C4.5w,HS-SVM ¹	0,43	0,44		0,861
EERent-5NN,EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM ¹	0,46	0,41	*	0,002
EERent-NB,EERent-SVM,HTUeuc-NB,HTUeuc-SVM,HS-NB,HS-SVM ¹	0,43	0,44		0,606
HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-NB,HS-C4.5w,HS-SVM ¹	0,45	0,39	*	0,005
EERent-5NN,EERent-NB,EERent-C4.5w,EERent-SVM,HS-5NN,HS-NB,HS-C4.5w,HS-SVM ¹	0,43	0,41		0,153
<i>EERent-NB,EERent-SVM,HS-NB,HS-SVM¹</i>	<i>0,40</i>	<i>0,42</i>	<i>+</i>	<i>0,044</i>
EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM ¹	0,44	0,35	*	0,000
EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-NB,HS-C4.5w,HS-SVM ¹	0,42	0,36	*	0,001
EERent-C4.5w,EERent-SVM,HS-C4.5w,HS-SVM ¹	0,38	0,39		0,793
EERent-NB,EERent-C4.5w,EERent-SVM,HS-NB,HS-C4.5w,HS-SVM ¹	0,39	0,36		0,205
HTUeuc-5NN,HTUeuc-NB,HS-5NN,HS-NB ¹	0,33	0,26		0,134
___HTUeuc-NB,HTUeuc-SVM,HS-NB,HS-SVM ²	0,32	0,25		0,178
___EERent-NB,EERent-SVM,HTUeuc-NB,HTUeuc-SVM,HS-NB,HS-SVM ²	0,29	0,24		0,899
___EERent-NB,EERent-SVM,HS-NB,HS-SVM ²	0,26	0,26		0,574
EERent-5NN,EERent-NB,HTUeuc-5NN,HTUeuc-NB,HS-5NN,HS-NB ¹	0,29	0,23		0,086
___EERent-NB,EERent-SVM,HTUeuc-NB,HTUeuc-SVM ²	0,27	0,21		0,490
___EERent-5NN,EERent-SVM,HS-5NN,HS-SVM ²	0,31	0,17	*	0,010
EERent-5NN,EERent-C4.5w,HTUeuc-5NN,HTUeuc-C4.5w,HS-5NN,HS-C4.5w ¹	0,23	0,25		0,654
HTUeuc-5NN,HTUeuc-C4.5w,HS-5NN,HS-C4.5w ¹	0,21	0,27		0,166
___EERent-5NN,EERent-SVM,HTUeuc-5NN,HTUeuc-SVM,HS-5NN,HS-SVM ²	0,30	0,18	+	0,019
___HTUeuc-5NN,HTUeuc-SVM,HS-5NN,HS-SVM ²	0,29	0,18	+	0,041
EERent-5NN,EERent-NB,EERent-C4.5w,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HS-5NN,HS-NB,HS-C4.5w ¹	0,27	0,20	+	0,016
EERent-5NN,EERent-C4.5w,HTUeuc-5NN,HTUeuc-C4.5w ¹	0,24	0,24		0,956
EERent-5NN,EERent-NB,HS-5NN,HS-NB ¹	0,24	0,23		0,727
___HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-NB,HS-C4.5w,HS-SVM ²	0,28	0,18	+	0,021
HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HS-5NN,HS-NB,HS-C4.5w ¹	0,27	0,19	*	0,006
EERent-5NN,EERent-NB,HTUeuc-5NN,HTUeuc-NB ¹	0,25	0,20		0,220
___HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM,HS-5NN,HS-NB,HS-SVM ²	0,30	0,15	*	0,000
___EERent-5NN,EERent-NB,EERent-SVM,HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM ²	0,27	0,17	+	0,018
___EERent-NB,EERent-C4.5w,EERent-SVM,HS-NB,HS-C4.5w,HS-SVM ²	0,26	0,18	+	0,021
___EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-NB,HS-C4.5w,HS-SVM ²	0,26	0,18	*	0,010
EERent-5NN,EERent-C4.5w,HS-5NN,HS-C4.5w ¹	0,21	0,22		0,936
___EERent-5NN,EERent-NB,EERent-SVM,HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM,HS-5NN,HS-NB,HS-SVM ²	0,27	0,15	*	0,001
EERent-5NN,EERent-NB,EERent-C4.5w,HS-5NN,HS-NB,HS-C4.5w ¹	0,25	0,17	+	0,016
EERent-5NN,EERent-NB,EERent-C4.5w,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w ¹	0,27	0,15	*	0,001
___EERent-C4.5w,EERent-SVM,HS-C4.5w,HS-SVM ²	0,25	0,16	+	0,030
___EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM ²	0,24	0,16	+	0,030
___EERent-5NN,EERent-SVM,HTUeuc-5NN,HTUeuc-SVM ²	0,26	0,14		0,137
___EERent-C4.5w,EERent-SVM,HTUeuc-C4.5w,HTUeuc-SVM,HS-C4.5w,HS-SVM ²	0,24	0,15	+	0,019
___HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-NB,HS-C4.5w,HS-SVM ²	0,28	0,10	*	0,000
___HTUeuc-C4.5w,HTUeuc-SVM,HS-C4.5w,HS-SVM ²	0,22	0,16		0,223
Todos os pares²	0,27	0,09	*	0,000
___EERent-5NN,EERent-NB,EERent-SVM,HS-5NN,HS-NB,HS-SVM ²	0,26	0,09	*	0,000
___HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-C4.5w,HS-SVM ²	0,26	0,09	*	0,000
___EERent-5NN,EERent-C4.5w,EERent-SVM,HS-5NN,HS-C4.5w,HS-SVM ²	0,26	0,08	*	0,000
___EERent-5NN,EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM ²	0,25	0,09	*	0,000
___EERent-5NN,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-C4.5w,HS-SVM ²	0,26	0,07	*	0,000
___EERent-5NN,EERent-NB,EERent-C4.5w,EERent-SVM,HS-5NN,HS-NB,HS-C4.5w,HS-SVM ²	0,25	0,08	*	0,000
___EERent-NB,EERent-C4.5w,HS-NB,HS-C4.5w ²	0,18	0,14		0,636
___HTUeuc-5NN,HTUeuc-C4.5w,HS-5NN,HS-C4.5w ²	0,21	0,09	*	0,009
___EERent-C4.5w,EERent-SVM,HTUeuc-C4.5w,HTUeuc-SVM ²	0,21	0,09	*	0,007
EERent-NB,EERent-C4.5w,HTUeuc-NB,HTUeuc-C4.5w,HS-NB,HS-C4.5w ¹	0,22	0,08	*	0,001
___EERent-5NN,EERent-C4.5w,HS-5NN,HS-C4.5w ²	0,22	0,08	*	0,005
___EERent-5NN,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM ²	0,27	0,01	*	0,000
___HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HS-5NN,HS-NB,HS-C4.5w ²	0,22	0,05	*	0,001
___EERent-5NN,EERent-NB,EERent-C4.5w,HS-5NN,HS-NB,HS-C4.5w ²	0,17	0,10	+	0,049
___EERent-NB,EERent-C4.5w,HTUeuc-NB,HTUeuc-C4.5w,HS-NB,HS-C4.5w ²	0,19	0,08	+	0,035
___EERent-5NN,EERent-C4.5w,HTUeuc-5NN,HTUeuc-C4.5w,HS-5NN,HS-C4.5w ²	0,21	0,06	*	0,003
___EERent-5NN,EERent-NB,HS-5NN,HS-NB ²	0,15	0,11		0,343
___HTUeuc-NB,HTUeuc-C4.5w,HS-NB,HS-C4.5w ²	0,22	0,03	*	0,005
HTUeuc-NB,HTUeuc-C4.5w,HS-NB,HS-C4.5w ¹	0,23	0,02	*	0,000
___EERent-5NN,EERent-NB,EERent-C4.5w,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HS-5NN,HS-NB,HS-C4.5w ²	0,20	0,04	*	0,000
___HTUeuc-5NN,HTUeuc-NB,HS-5NN,HS-NB ²	0,19	0,01	*	0,000
___EERent-5NN,EERent-NB,HTUeuc-5NN,HTUeuc-NB,HS-5NN,HS-NB ²	0,15	0,05	+	0,015
___EERent-NB,EERent-C4.5w,HTUeuc-NB,HTUeuc-C4.5w ²	0,16	0,03	+	0,019
EERent-NB,EERent-C4.5w,HS-NB,HS-C4.5w ¹	0,18	0,00	*	0,000
___EERent-5NN,EERent-NB,EERent-C4.5w,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w ²	0,19	-0,02	*	0,000
___EERent-5NN,EERent-NB,HTUeuc-5NN,HTUeuc-NB ²	0,11	0,06		0,381
___EERent-5NN,EERent-C4.5w,HTUeuc-5NN,HTUeuc-C4.5w ²	0,19	-0,05	*	0,001
EERent-NB,EERent-C4.5w,HTUeuc-NB,HTUeuc-C4.5w ¹	0,18	-0,05	*	0,000

Vitórias

71 15

Figura 33 – Comparação do coeficiente de correlação de Spearman na predição de ranqueamento de pares estratégia-algoritmo. *Linhas grafadas em itálico indicam diferença com significância estatística a favor do método de referência. Detalhes na Figura 27.*

As demais linhas representam experimentos em que todos os outros subconjuntos possíveis contendo pelo menos dois algoritmos e duas estratégias foram disponibilizados para o sistema de recomendação. As linhas correspondentes à segunda metade do período de consultas estão deslocadas para a direita visando compensar a baixa legibilidade dos números sobrescritos.

O meta-aprendiz PCT obteve um bom desempenho geral, sendo superior à referência 71 vezes contra 15; a maioria delas com diferenças estatisticamente significativas. Portanto, no contexto do experimento, é possível afirmar que **a recomendação automática de pares estratégia-algoritmo é, no mínimo, promissora**¹². Tal constatação é reforçada na tarefa de classificação, em que as colocações médias dos meta-aprendizes (2,1; 2,1; 2,4 e 3,3) superaram aquelas obtidas pelas referências (5,0 e 5,2) - conforme mostrado na Figura 34. Nessa modalidade de meta-aprendizado, RFW e PCT compartilharam a segunda colocação, seguidos por RoF.

5.3.3 Recomendação de métrica de distância

Um dos parâmetros que poderiam ser ajustados via meta-aprendizado é a métrica de distância das estratégias baseadas em densidade. A distância de Mahalanobis foi acrescentada como opção (MAHALANOBIS, 1936), apesar de seu alto custo computacional ter levado à exclusão dos seguintes conjuntos: *musk*, *texture*, *volcanoes-d1* e *volcanoes-b5*.

Considerando cada exemplo da reserva $\mathbf{x}_i, 1 \leq i \leq |\mathcal{U}|$ como um vetor linha da matriz U , a distância de Mahalanobis d_m entre dois exemplos \mathbf{u} e \mathbf{z} é baseada na matriz de covariância $\text{cov}(U)$ - conforme as equações 5.2 e 5.3.

$$U = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_{|\mathcal{U}|} \end{pmatrix} \in \mathbb{R}^{|\mathcal{U}| \times |\mathcal{A}|} \quad (5.2)$$

$$d_m(\mathbf{u}, \mathbf{z}) = \sqrt{(\mathbf{u} - \mathbf{z})^\top \text{cov}(U)^{-1} (\mathbf{u} - \mathbf{z})} \quad (5.3)$$

Cada par [estratégia baseada em densidade]-algoritmo foi representado por um experimento que consistia na recomendação automática de uma das três métricas de distância: euclidiana, Manhattan e Mahalanobis. Os resultados contrariam a tendência promissora observada anteriormente nas outras modalidades de recomendação. Na Figura 35, Def obteve mais vitórias que PCT, algumas com significância estatística (sujeita ao erro de tipo I global). Na predição de metaclasses, cujos resultados são apresentados na Figura 36, a colocação média de todos os meta-aprendizes se situou próxima da colocação

¹² Desde que a coleção contenha mais do que um conjunto de um mesmo domínio (conforme demonstrado posteriormente no Apêndice B)

Pares disponíveis

Pares disponíveis		RoF		RFw		PCT		ABoo		Maj		Alea	
HTUeuc-NB,HTUeuc-SVM,HS-NB,HS-SVM ¹		0,33	/	0,04	0,32	/	0,04	0,42	/	0,03	0,24	/	0,03
EERent-5NN,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM ¹		0,23	/	0,01	0,27	/	0,06	0,33	/	0,06	0,23	/	0,04
HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM,HS-5NN,HS-NB,HS-SVM ¹		0,28	/	0,03	0,29	/	0,03	0,29	/	0,05	0,10	/	0,06
HTUeuc-5NN,HTUeuc-NB,HS-5NN,HS-NB ¹		0,26	/	0,04	0,28	/	0,02	0,28	/	0,03	0,17	/	0,06
EERent-5NN,EERent-C4.5w,HS-5NN,HS-C4.5w ²		0,24	/	0,05	0,23	/	0,04	0,23	/	0,04	0,13	/	0,12
EERent-5NN,EERent-C4.5w,HTUeuc-5NN,HTUeuc-C4.5w ¹		0,19	/	0,03	0,23	/	0,06	0,19	/	0,03	0,21	/	0,09
EERent-5NN,EERent-C4.5w,HTUeuc-5NN,HTUeuc-C4.5w,HS-5NN,HS-C4.5w ¹		0,19	/	0,02	0,22	/	0,03	0,20	/	0,05	0,19	/	0,03
EERent-5NN,EERent-NB,EERent-SVM,HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM ¹		0,15	/	0,03	0,19	/	0,03	0,27	/	0,06	0,12	/	0,04
EERent-5NN,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-C4.5w,HS-SVM ¹		0,15	/	0,03	0,22	/	0,03	0,18	/	0,06	0,17	/	0,06
EERent-C4.5w,EERent-SVM,HTUeuc-C4.5w,HTUeuc-SVM ¹		0,24	/	0,00	0,13	/	0,04	0,18	/	0,01	0,12	/	0,05
HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HS-5NN,HS-NB,HS-C4.5w ¹		0,18	/	0,03	0,19	/	0,02	0,21	/	0,03	0,11	/	0,06
EERent-5NN,EERent-SVM,HS-5NN,HS-SVM ¹		0,23	/	0,04	0,15	/	0,02	0,22	/	0,04	0,08	/	0,05
EERent-5NN,EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM ¹		0,15	/	0,02	0,17	/	0,02	0,17	/	0,05	0,12	/	0,05
HTUeuc-NB,HTUeuc-C4.5w,HS-NB,HS-C4.5w ²		0,19	/	0,01	0,22	/	0,04	0,19	/	0,02	0,07	/	0,05
EERent-5NN,EERent-NB,HTUeuc-5NN,HTUeuc-NB ¹		0,13	/	0,03	0,21	/	0,06	0,24	/	0,06	0,06	/	0,05
HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-NB,HS-C4.5w,HS-SVM ¹		0,17	/	0,03	0,17	/	0,02	0,17	/	0,04	0,09	/	0,04
EERent-5NN,EERent-NB,HTUeuc-5NN,HTUeuc-NB,HS-5NN,HS-NB ¹		0,12	/	0,05	0,18	/	0,03	0,23	/	0,04	0,09	/	0,09
EERent-5NN,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM ²		0,16	/	0,01	0,17	/	0,03	0,15	/	0,03	0,16	/	0,04
EERent-C4.5w,EERent-SVM,HTUeuc-C4.5w,HTUeuc-SVM ¹		0,19	/	0,02	0,20	/	0,04	0,17	/	0,04	0,07	/	0,08
HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-NB,HS-C4.5w,HS-SVM ²		0,16	/	0,02	0,17	/	0,02	0,17	/	0,02	0,16	/	0,05
EERent-C4.5w,EERent-SVM,HTUeuc-C4.5w,HTUeuc-SVM,HS-C4.5w,HS-SVM ²		0,14	/	0,01	0,16	/	0,03	0,21	/	0,04	0,13	/	0,05
HTUeuc-5NN,HTUeuc-C4.5w,HS-5NN,HS-C4.5w ²		0,23	/	0,02	0,17	/	0,02	0,16	/	0,03	0,07	/	0,07
EERent-C4.5w,EERent-SVM,HTUeuc-C4.5w,HTUeuc-SVM,HS-C4.5w,HS-SVM ¹		0,19	/	0,03	0,13	/	0,04	0,17	/	0,06	0,10	/	0,08
EERent-NB,EERent-C4.5w,HTUeuc-NB,HTUeuc-C4.5w ²		0,16	/	0,02	0,16	/	0,03	0,13	/	0,01	0,11	/	0,07
EERent-5NN,EERent-C4.5w,EERent-SVM,HS-5NN,HS-C4.5w,HS-SVM ¹		0,15	/	0,02	0,19	/	0,01	0,11	/	0,04	0,11	/	0,06
EERent-5NN,EERent-NB,EERent-SVM,HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM,HS-5NN,HS-NB,HS-SVM ¹		0,12	/	0,03	0,14	/	0,04	0,24	/	0,02	0,07	/	0,03
EERent-5NN,EERent-C4.5w,HTUeuc-5NN,HTUeuc-C4.5w ²		0,18	/	0,04	0,17	/	0,02	0,16	/	0,02	0,12	/	0,03
EERent-NB,EERent-C4.5w,HTUeuc-NB,HTUeuc-C4.5w,HS-NB,HS-C4.5w ¹		0,16	/	0,03	0,15	/	0,05	0,13	/	0,02	0,11	/	0,02
HTUeuc-C4.5w,HTUeuc-SVM,HS-C4.5w,HS-SVM ¹		0,12	/	0,03	0,15	/	0,05	0,15	/	0,03	0,14	/	0,05
EERent-NB,EERent-SVM,HS-NB,HS-SVM ²		0,15	/	0,02	0,17	/	0,03	0,16	/	0,05	0,11	/	0,08
EERent-5NN,EERent-SVM,HS-5NN,HS-SVM ¹		0,13	/	0,02	0,14	/	0,03	0,14	/	0,06	0,16	/	0,06
HTUeuc-C4.5w,HTUeuc-SVM,HS-C4.5w,HS-SVM ²		0,11	/	0,04	0,14	/	0,04	0,15	/	0,04	0,18	/	0,06
EERent-5NN,EERent-NB,HTUeuc-5NN,HTUeuc-NB ²		0,15	/	0,05	0,15	/	0,04	0,17	/	0,04	0,08	/	0,08
EERent-C4.5w,EERent-SVM,HS-C4.5w,HS-SVM ²		0,13	/	0,03	0,19	/	0,03	0,15	/	0,02	0,12	/	0,05
HTUeuc-5NN,HTUeuc-NB,HS-5NN,HS-NB ¹		0,16	/	0,03	0,17	/	0,03	0,16	/	0,04	0,07	/	0,07
EERent-5NN,EERent-C4.5w,EERent-SVM,HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-C4.5w,HS-SVM ²		0,16	/	0,03	0,16	/	0,03	0,13	/	0,03	0,10	/	0,04
		0,09	/	0,03	0,14	/	0,02	0,21	/	0,03	0,09	/	0,02
HTUeuc-NB,HTUeuc-C4.5w,HS-NB,HS-C4.5w ¹		0,12	/	0,03	0,18	/	0,06	0,16	/	0,05	0,08	/	0,05
EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM ¹		0,13	/	0,04	0,13	/	0,03	0,14	/	0,02	0,11	/	0,02
HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM,HS-5NN,HS-NB,HS-SVM ²		0,16	/	0,03	0,16	/	0,03	0,13	/	0,02	0,11	/	0,06
EERent-5NN,EERent-NB,EERent-SVM,HS-5NN,HS-NB,HS-SVM ²		0,14	/	0,04	0,17	/	0,02	0,15	/	0,04	0,09	/	0,05
EERent-5NN,EERent-NB,EERent-C4.5w,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w ¹		0,11	/	0,03	0,15	/	0,03	0,17	/	0,03	0,09	/	0,04
EERent-5NN,EERent-NB,EERent-C4.5w,EERent-SVM,HS-5NN,HS-NB,HS-C4.5w,HS-SVM ¹		0,10	/	0,04	0,12	/	0,03	0,18	/	0,04	0,09	/	0,08
EERent-5NN,EERent-C4.5w,HTUeuc-5NN,HTUeuc-C4.5w,HS-5NN,HS-C4.5w ²		0,18	/	0,03	0,14	/	0,02	0,15	/	0,03	0,07	/	0,05
HTUeuc-5NN,HTUeuc-C4.5w,HS-5NN,HS-C4.5w ¹		0,16	/	0,03	0,11	/	0,03	0,12	/	0,03	0,10	/	0,06
HTUeuc-5NN,HTUeuc-SVM,HS-5NN,HS-SVM ²		0,14	/	0,04	0,07	/	0,06	0,17	/	0,05	0,10	/	0,08
EERent-NB,EERent-C4.5w,HTUeuc-NB,HTUeuc-C4.5w ¹		0,13	/	0,03	0,12	/	0,03	0,14	/	0,03	0,07	/	0,05
EERent-5NN,EERent-SVM,HTUeuc-5NN,HTUeuc-SVM ¹		0,01	/	0,06	0,14	/	0,04	0,12	/	0,05	0,14	/	0,06
EERent-5NN,EERent-SVM,HTUeuc-5NN,HTUeuc-SVM,HS-5NN,HS-SVM ²		0,15	/	0,02	0,13	/	0,01	0,11	/	0,03	0,08	/	0,02
HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-NB,HS-C4.5w,HS-SVM ²		0,11	/	0,02	0,12	/	0,03	0,10	/	0,05	0,11	/	0,09
EERent-5NN,EERent-NB,EERent-C4.5w,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HS-5NN,HS-NB,HS-C4.5w ¹		0,06	/	0,02	0,13	/	0,04	0,15	/	0,01	0,08	/	0,04
HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-NB,HS-C4.5w,HS-SVM ¹		0,09	/	0,01	0,11	/	0,03	0,10	/	0,03	0,09	/	0,05
EERent-5NN,EERent-SVM,HTUeuc-5NN,HTUeuc-SVM ²		0,17	/	0,04	0,11	/	0,03	0,09	/	0,04	0,03	/	0,03
HTUeuc-5NN,HTUeuc-C4.5w,HTUeuc-SVM,HS-5NN,HS-C4.5w,HS-SVM ¹		0,14	/	0,03	0,12	/	0,03	0,09	/	0,05	0,04	/	0,01
EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-NB,HTUeuc-SVM,HS-NB,HS-C4.5w,HS-SVM ¹		0,12	/	0,03	0,07	/	0,03	0,07	/	0,04	0,09	/	0,04
EERent-NB,EERent-C4.5w,HTUeuc-NB,HTUeuc-C4.5w,HS-NB,HS-C4.5w ²		0,11	/	0,02	0,11	/	0,04	0,08	/	0,03	0,08	/	0,02
EERent-5NN,EERent-C4.5w,EERent-SVM,HS-5NN,HS-C4.5w,HS-SVM ¹		0,18	/	0,04	0,07	/	0,03	0,10	/	0,03	0,04	/	0,05
EERent-5NN,EERent-NB,EERent-C4.5w,HS-5NN,HS-NB,HS-C4.5w ¹		0,08	/	0,02	0,10	/	0,02	0,11	/	0,05	0,03	/	0,02
EERent-5NN,EERent-NB,EERent-C4.5w,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w ²		0,07	/	0,04	0,10	/	0,04	0,08	/	0,04	0,05	/	0,05
EERent-5NN,EERent-SVM,HTUeuc-5NN,HTUeuc-SVM,HS-5NN,HS-SVM ¹		0,04	/	0,05	0,11	/	0,04	0,15	/	0,03	-0,01	/	0,07
HTUeuc-NB,HTUeuc-SVM,HS-NB,HS-SVM ²		0,02	/	0,02	0,09	/	0,02	0,09	/	0,02	0,05	/	0,05
EERent-5NN,EERent-NB,EERent-C4.5w,EERent-SVM,HS-5NN,HS-NB,HS-C4.5w,HS-SVM ²		0,08	/	0,01	0,07	/	0,04	0,06	/	0,03	0,06	/	0,05
EERent-5NN,EERent-NB,EERent-SVM,HS-5NN,HS-NB,HS-SVM ¹		0,02	/	0,02	0,05	/	0,06	0,13	/	0,02	0,02	/	0,06
EERent-5NN,EERent-C4.5w,HS-5NN,HS-C4.5w ¹		0,07	/	0,03	0,08	/	0,04	0,06	/	0,02	0,05	/	0,10
EERent-5NN,EERent-NB,EERent-C4.5w,HTUeuc-5NN,HTUeuc-NB,HTUeuc-C4.5w,HS-5NN,HS-NB,HS-C4.5w ²		0,08	/	0,02	0,06	/	0,00	0,03	/	0,02	0,07	/	0,02
EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM ²		0,05	/	0,03	0,08	/	0,00	0,06	/	0,02	0,04	/	0,04
EERent-5NN,EERent-NB,EERent-SVM,HTUeuc-5NN,HTUeuc-NB,HTUeuc-SVM,HS-5NN,HS-NB,HS-SVM ²		0,08	/	0,03	0,05	/	0,03	0,06	/	0,04	0,03	/	0,03
EERent-NB,EERent-C4.5w,EERent-SVM,HTUeuc-NB,HTUeuc-C4.5w,HTUeuc-SVM,HS-NB,HS-C4.5w,HS-SVM ²		0,05	/	0,01	0,04	/	0,02	0,04	/	0,01	0,04	/	0,04
EERent-NB,EERent-C4.5w,HS-NB,HS-C4.5w ¹		0,06	/	0,02	0,05	/	0,02	0,01	/	0,01	0,08	/	0,02
EERent-NB,EERent-C4.5w,EERent-SVM,HS-NB,HS-C4.5w,HS-SVM ²		0,03	/	0,03	0,07	/	0,02	0,05	/	0,02	0,05	/	0,04
EERent-5NN,EERent-NB,HTUeuc-5NN,HTUeuc-NB,HS-5NN,HS-NB ²		0,08	/	0,02	0,07	/	0,02	0,04	/	0,07	-0,01	/	0,06
EERent-5NN,EERent-NB,HS-5NN,HS-NB ¹		0,00	/	0,02	0,03	/	0,03	0,04	/	0,04	0,09		

Estratégia-Algoritmo	PCT	Def	α	p-valor
____DW-NB ²	0,12	0,13		0,356
DW-NB ¹	0,12	0,12		0,834
TU-5NN ¹	0,13	0,09		0,110
____TU-5NN ²	0,11	0,10		0,642
HTU-C4.5w ¹	0,10	0,09		0,511
DW-5NN ¹	0,11	0,06		0,391
____DW-SVM ²	0,11	0,05		0,121
DW-C4.5w ¹	0,06	0,09		0,832
ATU-5NN ¹	0,08	0,05		0,167
TU-C4.5w ¹	0,06	0,07		0,642
ATU-C4.5w ¹	0,06	0,07		0,295
HTU-5NN ¹	0,07	0,05		0,300
____HTU-C4.5w ²	0,06	0,06		0,925
____TU-C4.5w ²	0,04	0,06		0,299
____ATU-SVM ²	0,04	0,06		0,054
DW-SVM ¹	0,04	0,05		0,771
____DW-C4.5w ²	0,01	0,06		0,064
____ATU-NB ²	0,06	0,01	+	0,038
TU-NB ¹	0,05	0,02		0,191
____HTU-SVM ²	0,02	0,04		0,409
HTU-SVM ¹	0,02	0,04		0,261
____HTU-5NN ²	0,03	0,03		0,928
____DW-5NN ²	0,12	-0,06	*	0,000
____TU-NB ²	0,02	0,04		0,170
____HTU-NB ²	0,03	0,02		0,586
ATU-SVM ¹	0,02	0,03		0,389
TU-SVM ¹	0,02	0,03		0,558
HTU-NB ¹	0,02	0,02		0,751
ATU-NB ¹	0,01	0,03		0,972
____ATU-5NN ²	0,01	0,04		0,162
____ATU-C4.5w ²	0,00	0,04	+	0,012
____TU-SVM ²	-0,01	0,03	+	0,036
Vitórias	12	19		

Figura 35 – Desempenho na predição de ranqueamento de medidas de distância para estratégias baseadas em densidade. *Detalhes na Figura 33.*

central (3,5), indicando que a escolha do algoritmo de aprendizado praticamente não influencia o resultado.

Apesar de existir uma pequena vantagem na colocação média de PCT (2,8) em relação à melhor referência (Maj na colocação 3,4), não é possível afirmar que a diferença seja significativa, pois, predominam valores negativos na metade inferior da figura, e metade dos valores positivos são muito próximos de zero quando comparados à magnitude do desvio padrão.

A recomendação automática da métrica de distância mostra-se, assim, um maior desafio do que aquele apresentado pelas outras modalidades, pois não há indícios suficientes de que ela seja viável - consideradas as limitações experimentais e especificidades da abordagem proposta. Logo, a princípio, **a recomendação automática da métrica de distância não mostrou-se suficientemente promissora** para investigações futuras prioritárias. No entanto, também não é possível descartá-la completamente, pois outros

Estratégia-Algoritmo	RoF		RFw		PCT		ABoo		Maj		Alea	
DW-C4.5w ¹	0,39	/ 0,07	0,32	/ 0,02	0,32	/ 0,04	0,20	/ 0,07	0,00	/ 0,00	0,00	/ 0,07
____DW-5NN ²	0,19	/ 0,06	0,21	/ 0,05	0,15	/ 0,05	0,22	/ 0,07	0,00	/ 0,00	0,00	/ 0,02
TU-5NN ¹	0,21	/ 0,04	0,14	/ 0,04	0,13	/ 0,01	0,07	/ 0,05	0,00	/ 0,00	-0,01	/ 0,11
ATU-5NN ¹	0,14	/ 0,02	0,19	/ 0,03	0,13	/ 0,04	0,09	/ 0,06	0,00	/ 0,00	-0,03	/ 0,02
____DW-NB ²	0,13	/ 0,02	0,16	/ 0,02	0,09	/ 0,01	0,10	/ 0,05	0,00	/ 0,00	-0,02	/ 0,06
DW-5NN ¹	0,14	/ 0,03	0,15	/ 0,06	0,19	/ 0,02	0,06	/ 0,05	0,00	/ 0,00	-0,03	/ 0,04
DW-NB ¹	0,06	/ 0,02	0,07	/ 0,03	0,18	/ 0,05	0,04	/ 0,03	0,00	/ 0,00	0,01	/ 0,06
HTU-C4.5w ¹	0,08	/ 0,03	0,10	/ 0,04	0,12	/ 0,05	0,08	/ 0,08	0,00	/ 0,00	-0,02	/ 0,06
HTU-NB ¹	0,06	/ 0,04	0,09	/ 0,03	0,09	/ 0,06	0,08	/ 0,05	0,00	/ 0,00	0,00	/ 0,05
____ATU-C4.5w ²	0,09	/ 0,05	0,07	/ 0,04	0,13	/ 0,03	0,02	/ 0,07	0,00	/ 0,00	-0,02	/ 0,07
TU-NB ¹	0,16	/ 0,05	0,03	/ 0,03	0,04	/ 0,06	0,03	/ 0,11	0,00	/ 0,00	0,01	/ 0,07
HTU-5NN ¹	0,11	/ 0,04	0,09	/ 0,05	0,08	/ 0,04	0,08	/ 0,07	0,00	/ 0,00	-0,03	/ 0,04
HTU-SVM ¹	0,01	/ 0,04	0,07	/ 0,04	0,07	/ 0,04	0,07	/ 0,06	0,00	/ 0,00	-0,03	/ 0,05
____DW-SVM ²	0,02	/ 0,05	0,08	/ 0,03	0,11	/ 0,05	-0,04	/ 0,04	0,00	/ 0,00	-0,01	/ 0,06
____TU-5NN ²	0,07	/ 0,04	0,00	/ 0,06	0,00	/ 0,05	0,00	/ 0,05	0,00	/ 0,00	-0,01	/ 0,07
____ATU-SVM ²	0,01	/ 0,03	0,00	/ 0,03	0,01	/ 0,04	0,03	/ 0,08	0,00	/ 0,00	0,00	/ 0,05
____ATU-NB ²	0,03	/ 0,02	0,00	/ 0,03	-0,03	/ 0,05	-0,02	/ 0,05	0,00	/ 0,00	0,01	/ 0,07
____ATU-5NN ²	0,03	/ 0,05	-0,01	/ 0,07	0,02	/ 0,07	-0,03	/ 0,09	0,00	/ 0,00	-0,03	/ 0,02
____DW-C4.5w ²	0,06	/ 0,04	-0,02	/ 0,02	-0,02	/ 0,05	-0,01	/ 0,06	0,00	/ 0,00	-0,03	/ 0,03
TU-SVM ¹	-0,03	/ 0,07	-0,02	/ 0,07	0,03	/ 0,07	0,01	/ 0,09	0,00	/ 0,00	-0,01	/ 0,07
ATU-C4.5w ¹	-0,09	/ 0,05	-0,04	/ 0,12	0,05	/ 0,06	-0,05	/ 0,09	0,00	/ 0,00	0,01	/ 0,06
DW-SVM ¹	-0,03	/ 0,04	-0,04	/ 0,03	0,00	/ 0,05	-0,03	/ 0,06	0,00	/ 0,00	-0,03	/ 0,05
____HTU-NB ²	-0,06	/ 0,05	-0,05	/ 0,05	0,00	/ 0,03	-0,02	/ 0,05	0,00	/ 0,00	-0,02	/ 0,07
____TU-NB ²	-0,09	/ 0,04	-0,05	/ 0,06	-0,05	/ 0,02	0,03	/ 0,07	0,00	/ 0,00	0,01	/ 0,06
____TU-C4.5w ²	0,00	/ 0,03	-0,05	/ 0,05	-0,04	/ 0,04	-0,07	/ 0,05	0,00	/ 0,00	0,01	/ 0,12
ATU-NB ¹	-0,04	/ 0,01	-0,01	/ 0,06	-0,08	/ 0,06	-0,05	/ 0,13	0,00	/ 0,00	0,01	/ 0,05
____HTU-5NN ²	-0,06	/ 0,04	-0,04	/ 0,04	-0,06	/ 0,02	-0,03	/ 0,04	0,00	/ 0,00	0,00	/ 0,05
____TU-SVM ²	-0,04	/ 0,04	-0,08	/ 0,03	-0,09	/ 0,03	-0,06	/ 0,04	0,00	/ 0,00	0,00	/ 0,07
ATU-SVM ¹	-0,05	/ 0,04	-0,09	/ 0,03	-0,05	/ 0,05	-0,05	/ 0,07	0,00	/ 0,00	-0,03	/ 0,06
TU-C4.5w ¹	-0,10	/ 0,05	-0,09	/ 0,06	-0,10	/ 0,06	-0,03	/ 0,09	0,00	/ 0,00	0,00	/ 0,08
____HTU-SVM ²	-0,11	/ 0,03	-0,08	/ 0,04	-0,09	/ 0,03	-0,07	/ 0,03	0,00	/ 0,00	-0,02	/ 0,07
____HTU-C4.5w ²	-0,16	/ 0,04	-0,17	/ 0,04	-0,16	/ 0,06	-0,14	/ 0,10	0,00	/ 0,00	-0,01	/ 0,07
Ranq. médio	3,1		3,3		2,8		3,3		3,3		3,9	

Figura 36 – Comparação de κ (média/desvio padrão) na predição da melhor medida de distância para estratégias baseadas em densidade. Detalhes na Figura 27.

conjuntos de meta-atributos ou meta-aprendizes podem se mostrar mais apropriados para a tarefa.

5.3.4 Recomendação de emprego de aprendizado ativo

A última modalidade de recomendação avaliada foi a escolha entre um dado par estratégia-algoritmo e o par composto por esse mesmo algoritmo e Rnd (amostragem aleatória). A utilização de Rnd é equivalente a não utilizar uma estratégia de amostragem ativa. Logo, a finalidade do experimento é simular a decisão de emprego ou não de aprendizado ativo.

São apenas duas metaclasses por experimento, logo, não há predição de ranqueamento. Na Figura 37, PCT obteve a melhor colocação média (2,5). Comparativamente, esse valor é melhor posicionado do que a colocação média obtida na modalidade de recomendação de métricas de distância (2,8). De fato, em aproximadamente dois terços

dos experimentos, PCT obteve médias de κ acima do desvio padrão. Cada linha em que a média supera o desvio padrão sugere que os valores sejam majoritariamente positivos. Apesar desse resultado mostrar que o uso do meta-aprendiz levou a recomendações melhores que a referência Maj, a referência Alea mostra-se mais segura, por ter apenas valores positivos. Consequentemente, ***há indícios de que a predição automática da efetividade do aprendizado ativo, para uma dada estratégia e um dado algoritmo de aprendizado arbitrariamente definidos, não seja recomendável.*** No entanto, HTU constou 5 vezes entre os 30 experimentos na parte superior da figura e apenas 1 vez entre os 30 experimentos da parte inferior. Assim, ***há indícios de que a estratégia HTU ofereça maior previsibilidade quanto à possibilidade de predição da efetividade do aprendizado ativo.***

Um ponto importante a ser considerado é que, numa situação real, os pares mais adequados - ou os pares menos adequados - podem ser conhecidos de antemão pelo especialista, determinando quais linhas da Figura 37 correspondem às necessidades da aplicação. Essa quantidade reduzida de linhas permitiria saber mais precisamente se a recomendação de uso de aprendizado ativo supera ou não as referências.

5.4 Considerações

Em geral, o emprego de estratégias de amostragem ativa mostrou-se efetivo no aumento da acurácia preditiva, apesar das dúvidas levantadas em parte da literatura (EVANS; ADAMS; ANAGNOSTOPOULOS, 2014; ATTENBERG; PROVOST, 2010a). Especificamente, os experimentos permitiram: verificar a existência de relações entre conjuntos de dados, estratégias e algoritmos; testar, no nível base, a hipótese da possibilidade de inibição e controle do tipo de algoritmo empregado como aprendiz; e, comprovar que os metamodelos induzidos representam o conceito do problema de recomendação automática de forma a superar a acurácia preditiva das referências, tanto no nível meta quanto no nível base. O algoritmo adotado como aprendiz mostrou-se de fundamental importância para o desempenho de cada par estratégia-algoritmo (Seção 5.1.4).

Em termos de custo computacional/de esforço humano, da regularidade no desempenho e de acurácia preditiva e sua variabilidade, as estratégias propostas (ATU e HTU) mostraram-se competitivas com relação a seus pares. A proposta de adaptação (SGmulti) não obteve um bom desempenho geral, mas obteve o menor número de derrotas para a amostragem aleatória. Segundo o experimento adicional, registrado apenas no Apêndice A, SGmulti obteve um desempenho acima da média quando utilizada com RFW enquanto aprendiz. Trata-se de um resultado relevante, pois esse algoritmo tem sido reportado entre aqueles com o melhor desempenho preditivo (DELGADO *et al.*, 2014).

Além das abordagens propostas, uma das estratégias que obtiveram melhor desem-

Estratégia-Algoritmo	RoF		RFw		PCT		ABoo		Maj		Alea	
SGmulti-5NN ²	-	/ 0,00	-	/ 0,00	-	/ 0,00	-	/ 0,00	-	/ 0,00	-	/ 0,00
SGmulti-5NN ¹	-	/ 0,00	-	/ 0,00	-	/ 0,00	-	/ 0,00	-	/ 0,00	-	/ 0,00
TUeuc-SVM ¹	0,50	/ 0,03	0,49	/ 0,03	0,41	/ 0,06	0,37	/ 0,05	0,00	/ 0,00	0,14	/ 0,04
HTUman-SVM ¹	0,31	/ 0,03	0,33	/ 0,04	0,38	/ 0,03	0,29	/ 0,03	0,00	/ 0,00	0,06	/ 0,03
HTUman-NB ¹	0,25	/ 0,07	0,29	/ 0,02	0,40	/ 0,03	0,23	/ 0,11	0,00	/ 0,00	0,10	/ 0,04
HTUeuc-NB ¹	0,19	/ 0,06	0,37	/ 0,06	0,40	/ 0,05	0,20	/ 0,07	0,00	/ 0,00	0,10	/ 0,02
HS-5NN ²	0,22	/ 0,04	0,31	/ 0,06	0,29	/ 0,07	0,19	/ 0,12	0,00	/ 0,00	0,12	/ 0,03
ATUeuc-SVM ¹	0,35	/ 0,06	0,29	/ 0,07	0,24	/ 0,08	0,18	/ 0,04	0,00	/ 0,00	0,06	/ 0,03
HTUeuc-5NN ¹	0,27	/ 0,04	0,27	/ 0,04	0,29	/ 0,07	0,13	/ 0,07	0,00	/ 0,00	0,14	/ 0,05
DWeuc-NB ²	0,31	/ 0,02	0,33	/ 0,05	0,26	/ 0,05	0,13	/ 0,07	0,00	/ 0,00	0,07	/ 0,04
DWman-5NN ¹	0,31	/ 0,05	0,26	/ 0,05	0,32	/ 0,05	0,12	/ 0,09	0,00	/ 0,00	0,08	/ 0,04
TUman-SVM ¹	0,22	/ 0,07	0,25	/ 0,04	0,29	/ 0,03	0,18	/ 0,13	0,00	/ 0,00	0,12	/ 0,03
EERacc-5NN ¹	0,30	/ 0,02	0,23	/ 0,05	0,34	/ 0,05	0,16	/ 0,08	0,00	/ 0,00	0,03	/ 0,05
TUeuc-NB ²	0,22	/ 0,04	0,29	/ 0,07	0,32	/ 0,10	0,16	/ 0,11	0,00	/ 0,00	0,05	/ 0,04
Mar-C4.5w ²	0,26	/ 0,05	0,21	/ 0,05	0,32	/ 0,04	0,15	/ 0,08	0,00	/ 0,00	0,10	/ 0,02
HTUeuc-C4.5w ²	0,22	/ 0,02	0,23	/ 0,01	0,27	/ 0,06	0,15	/ 0,04	0,00	/ 0,00	0,14	/ 0,04
DWman-5NN ²	0,29	/ 0,07	0,25	/ 0,04	0,21	/ 0,05	0,13	/ 0,10	0,00	/ 0,00	0,06	/ 0,03
HTUeuc-NB ²	0,22	/ 0,05	0,29	/ 0,06	0,25	/ 0,03	0,09	/ 0,09	0,00	/ 0,00	0,08	/ 0,02
Mar-NB ²	0,30	/ 0,02	0,26	/ 0,04	0,21	/ 0,06	0,06	/ 0,14	0,00	/ 0,00	0,09	/ 0,01
EERent-5NN ²	0,20	/ 0,03	0,24	/ 0,03	0,24	/ 0,04	0,09	/ 0,13	0,00	/ 0,00	0,14	/ 0,04
EERacc-5NN ²	0,22	/ 0,07	0,25	/ 0,05	0,20	/ 0,09	0,17	/ 0,09	0,00	/ 0,00	0,07	/ 0,03
DWman-C4.5w ²	0,28	/ 0,02	0,20	/ 0,05	0,29	/ 0,08	0,09	/ 0,05	0,00	/ 0,00	0,04	/ 0,04
Mar-SVM ¹	0,22	/ 0,05	0,28	/ 0,05	0,19	/ 0,04	0,14	/ 0,07	0,00	/ 0,00	0,05	/ 0,05
TUman-C4.5w ²	0,15	/ 0,07	0,16	/ 0,02	0,29	/ 0,04	0,21	/ 0,06	0,00	/ 0,00	0,06	/ 0,02
DWeuc-C4.5w ²	0,28	/ 0,04	0,16	/ 0,05	0,32	/ 0,04	0,07	/ 0,12	0,00	/ 0,00	0,03	/ 0,02
HTUeuc-SVM ¹	0,22	/ 0,05	0,20	/ 0,08	0,19	/ 0,06	0,19	/ 0,07	0,00	/ 0,00	0,06	/ 0,03
HS-C4.5w ²	0,22	/ 0,06	0,23	/ 0,05	0,18	/ 0,02	0,15	/ 0,08	0,00	/ 0,00	0,08	/ 0,03
ATUeuc-5NN ¹	0,12	/ 0,09	0,15	/ 0,08	0,30	/ 0,09	0,13	/ 0,10	0,00	/ 0,00	0,16	/ 0,05
Mar-C4.5w ¹	0,18	/ 0,05	0,18	/ 0,03	0,23	/ 0,04	0,17	/ 0,10	0,00	/ 0,00	0,09	/ 0,01
HTUman-C4.5w ²	0,15	/ 0,06	0,16	/ 0,02	0,23	/ 0,03	0,13	/ 0,09	0,00	/ 0,00	0,15	/ 0,04
TUeuc-NB ¹	0,11	/ 0,06	0,25	/ 0,05	0,26	/ 0,06	0,14	/ 0,03	0,00	/ 0,00	0,06	/ 0,03
DWman-SVM ¹	0,18	/ 0,04	0,24	/ 0,02	0,21	/ 0,04	0,13	/ 0,13	0,00	/ 0,00	0,05	/ 0,06
Mar-SNN ²	0,22	/ 0,04	0,20	/ 0,04	0,15	/ 0,08	0,18	/ 0,10	0,00	/ 0,00	0,05	/ 0,04
EERacc-SVM ¹	0,22	/ 0,06	0,17	/ 0,12	0,21	/ 0,10	0,10	/ 0,11	0,00	/ 0,00	0,10	/ 0,04
DWman-C4.5w ¹	0,17	/ 0,03	0,17	/ 0,05	0,21	/ 0,05	0,17	/ 0,09	0,00	/ 0,00	0,06	/ 0,04
DWeuc-SVM ¹	0,16	/ 0,05	0,25	/ 0,02	0,23	/ 0,04	0,09	/ 0,05	0,00	/ 0,00	0,03	/ 0,05
DWeuc-NB ¹	0,19	/ 0,05	0,20	/ 0,07	0,16	/ 0,05	0,15	/ 0,08	0,00	/ 0,00	0,07	/ 0,04
ATUman-C4.5w ¹	0,19	/ 0,07	0,20	/ 0,08	0,19	/ 0,05	0,02	/ 0,09	0,00	/ 0,00	0,16	/ 0,03
HTUman-SVM ¹	0,20	/ 0,05	0,11	/ 0,03	0,21	/ 0,08	0,11	/ 0,08	0,00	/ 0,00	0,13	/ 0,03
DWman-SVM ¹	0,24	/ 0,08	0,16	/ 0,06	0,11	/ 0,05	0,09	/ 0,07	0,00	/ 0,00	0,14	/ 0,03
DWman-NB ¹	0,20	/ 0,07	0,17	/ 0,09	0,21	/ 0,08	0,08	/ 0,05	0,00	/ 0,00	0,07	/ 0,05
TUeuc-SNN ²	0,12	/ 0,01	0,20	/ 0,10	0,20	/ 0,07	0,17	/ 0,16	0,00	/ 0,00	0,05	/ 0,03
Mar-SNN ²	0,20	/ 0,03	0,20	/ 0,05	0,16	/ 0,04	0,09	/ 0,07	0,00	/ 0,00	0,06	/ 0,03
ATUeuc-C4.5w ¹	0,00	/ 0,08	0,24	/ 0,05	0,27	/ 0,06	0,06	/ 0,14	0,00	/ 0,00	0,14	/ 0,05
DWman-NB ²	0,28	/ 0,04	0,16	/ 0,02	0,12	/ 0,06	0,15	/ 0,16	0,00	/ 0,00	0,07	/ 0,04
ATUman-SVM ¹	0,18	/ 0,05	0,17	/ 0,07	0,13	/ 0,06	0,08	/ 0,10	0,00	/ 0,00	0,13	/ 0,03
TUeuc-SNN ¹	0,12	/ 0,02	0,22	/ 0,08	0,17	/ 0,06	0,11	/ 0,08	0,00	/ 0,00	0,07	/ 0,04
HTUman-NB ²	0,11	/ 0,07	0,18	/ 0,03	0,15	/ 0,04	0,18	/ 0,10	0,00	/ 0,00	0,10	/ 0,02
EERent-NB ¹	0,10	/ 0,03	0,12	/ 0,05	0,21	/ 0,04	0,09	/ 0,08	0,00	/ 0,00	0,10	/ 0,04
DWman-SVM ²	0,06	/ 0,09	0,19	/ 0,06	0,21	/ 0,01	0,08	/ 0,12	0,00	/ 0,00	0,08	/ 0,02
DWman-SVM ²	0,10	/ 0,06	0,18	/ 0,09	0,17	/ 0,08	0,04	/ 0,09	0,00	/ 0,00	0,09	/ 0,03
HTUeuc-SNN ²	0,04	/ 0,04	0,14	/ 0,04	0,10	/ 0,06	0,18	/ 0,08	0,00	/ 0,00	0,13	/ 0,03
ATUman-SVM ²	0,04	/ 0,10	0,14	/ 0,06	0,12	/ 0,10	0,10	/ 0,10	0,00	/ 0,00	0,05	/ 0,03
ATUman-NB ¹	0,05	/ 0,04	0,07	/ 0,08	0,15	/ 0,05	0,19	/ 0,14	0,00	/ 0,00	0,10	/ 0,02
SGmulti-NB ¹	0,10	/ 0,07	0,12	/ 0,08	0,10	/ 0,10	0,13	/ 0,08	0,00	/ 0,00	0,07	/ 0,05
EERent-SNN ¹	0,07	/ 0,07	0,11	/ 0,06	0,10	/ 0,04	0,12	/ 0,10	0,00	/ 0,00	0,12	/ 0,03
DWeuc-C4.5w ¹	0,05	/ 0,11	0,05	/ 0,05	0,30	/ 0,04	0,03	/ 0,09	0,00	/ 0,00	0,07	/ 0,03
HS-C4.5w ¹	0,14	/ 0,03	0,14	/ 0,04	0,16	/ 0,04	-0,03	/ 0,08	0,00	/ 0,00	0,08	/ 0,05
TUman-NB ¹	0,11	/ 0,04	0,10	/ 0,04	0,21	/ 0,07	0,02	/ 0,06	0,00	/ 0,00	0,06	/ 0,03
TUman-SVM ¹	0,21	/ 0,04	0,13	/ 0,06	0,05	/ 0,05	0,03	/ 0,09	0,00	/ 0,00	0,07	/ 0,05
TUman-SNN ¹	0,15	/ 0,08	0,18	/ 0,10	0,07	/ 0,05	0,09	/ 0,08	0,00	/ 0,00	0,06	/ 0,04
SGmulti-SVM ¹	-0,01	/ 0,06	0,06	/ 0,04	0,21	/ 0,08	0,06	/ 0,08	0,00	/ 0,00	0,12	/ 0,03
ATUman-SVM ¹	0,16	/ 0,06	0,07	/ 0,03	0,00	/ 0,03	0,07	/ 0,08	0,00	/ 0,00	0,15	/ 0,02
SGmulti-C4.5w ¹	0,08	/ 0,04	0,05	/ 0,05	0,15	/ 0,05	0,11	/ 0,06	0,00	/ 0,00	0,08	/ 0,04
HTUman-SNN ²	0,12	/ 0,08	0,13	/ 0,04	0,04	/ 0,04	-0,01	/ 0,07	0,00	/ 0,00	0,14	/ 0,04
DWeuc-SNN ¹	0,06	/ 0,07	0,12	/ 0,04	0,08	/ 0,10	0,07	/ 0,07	0,00	/ 0,00	0,08	/ 0,04
SGmulti-NB ¹	-0,01	/ 0,08	0,15	/ 0,08	0,15	/ 0,03	0,05	/ 0,14	0,00	/ 0,00	0,06	/ 0,03
HTUman-C4.5w ¹	0,01	/ 0,08	0,03	/ 0,08	0,06	/ 0,05	0,12	/ 0,18	0,00	/ 0,00	0,17	/ 0,02
EERacc-SVM ¹	0,02	/ 0,03	0,14	/ 0,06	0,13	/ 0,11	-0,02	/ 0,09	0,00	/ 0,00	0,09	/ 0,01
SGmulti-C4.5w ¹	0,11	/ 0,05	0,08	/ 0,07	0,11	/ 0,08	-0,04	/ 0,03	0,00	/ 0,00	0,07	/ 0,04
HTUeuc-SVM ¹	0,00	/ 0,04	0,07	/ 0,07	0,06	/ 0,05	0,05	/ 0,04	0,00	/ 0,00	0,15	/ 0,05
SGmulti-SNN ¹	0,02	/ 0,01	0,04	/ 0,03	0,01	/ 0,04	0,07	/ 0,02	0,00	/ 0,00	0,06	/ 0,03
ATUman-C4.5w ²	0,01	/ 0,06	0,06	/ 0,02	0,07	/ 0,06	0,03	/ 0,09	0,00	/ 0,00	0,14	/ 0,04
EERent-SVM ¹	0,00	/ 0,02	0,04	/ 0,04	0,14	/ 0,03	-0,02	/ 0,06	0,00	/ 0,00	0,14	/ 0,03
HS-SVM ²	0,14	/ 0,05	0,08	/ 0,04	-0,02	/ 0,03	-0,05	/ 0,09	0,00	/ 0,00	0,10	/ 0,04
ATUeuc-SVM ²	-0,03	/ 0,04	0,07	/ 0,06	0,08	/ 0,05	-0,02	/ 0,06	0,00	/ 0,00	0,14	/ 0,04
TUeuc-C4.5w ²	0,07	/ 0,06	-0,02	/ 0,05	0,11	/ 0,04	0,03	/ 0,12	0,00	/ 0,00	0,05	/ 0,05
ATUeuc-NB ¹	-0,01	/ 0,06	0,03	/ 0,05	0,13	/ 0,07	-0,03	/ 0,07	0,00	/ 0,00	0,10	/ 0,04
HS-SVM ¹	0,00	/ 0,06	0,03	/ 0,05	0,06	/ 0,04	-0,03	/ 0,10	0,00	/ 0,00	0,12	/ 0,03
ATUeuc-5NN ²	0,13	/ 0,06	-0,06	/ 0,01	-0,07	/ 0,01	0,03	/ 0,21	0,00	/ 0,00	0,12	/ 0,05
DWeuc-5NN ²	0,00	/ 0,07	0,07	/ 0,04	-0,03	/ 0,08	0,03	/ 0,09	0,00	/ 0,00	0,06	/ 0,03
Mar-NB ¹	0,02	/ 0,06	-0,04	/ 0,04	0,03	/ 0,07	-0,01	/ 0,15	0,00	/ 0,00	0,09	/ 0,02
HTUman-5NN ¹	-0,04	/ 0,03	-0,01	/ 0,01	-0,04	/ 0,01	-0,01	/ 0,13	0,00	/ 0,00	0,17	/ 0,02
Mar-SVM ²	0,02	/ 0,02	-0,03	/ 0,06	-0,07	/ 0,04	0,01	/ 0,12	0,00	/ 0,00	0,15	/ 0,05
EERent-C4.5w ²	0,05	/ 0,07	-0,07	/ 0,03	-0,06	/ 0,05	-0,01	/ 0,09	0,00	/ 0,00	0,15	/ 0,04
TUeuc-C4.5w ¹	-0,05	/ 0,03	-0,01	/ 0,04	0,07	/ 0,05	-0,02	/ 0,05	0,00	/ 0,00	0,05	/ 0,03
EERacc-C4.5w ¹	-0,07	/ 0,04	-0,04	/ 0,02	0,02	/ 0,03	0,01	/ 0,05	0,00	/ 0,00	0,10	/ 0,04
HTUeuc-C4.5w ¹	-0,07	/ 0,07	-0,04	/ 0,06	-0,02	/ 0,06	-0,02	/ 0,10	0,00	/ 0,00	0,16	/ 0,01
EERent-SVM ²	-0,06	/ 0,01	-0,05	/ 0,02	0,03	/ 0,03	-0,09	/ 0,08	0,00	/ 0,00	0,15	/ 0,05
ATUman-5NN ¹	-0,02	/ 0,05</										

penho foi EER. No entanto, sua complexidade computacional é a mais elevada, tornando-a inadequada em aplicações que tenham restrição de tempo de processamento, como a dependência da interação com um supervisor humano.

Com relação a meta-aprendizado, sua aplicação em aprendizado ativo mostrou-se efetiva na recomendação de algoritmos e pares estratégia-algoritmo. Outras modalidades também foram investigadas, sendo que a recomendação de estratégias mostrou-se a mais promissora.

Finalmente, conclui-se que, possivelmente, cada algoritmo de aprendizado seja uma questão à parte em aprendizado ativo, devendo ter uma análise prioritária no momento da definição do sistema de aprendizado. Por sua vez, a escolha da estratégia depende, inicialmente, dos detalhes do cenário da aplicação, principalmente: se ela requer a disponibilidade do oráculo; se ele é sujeito à fadiga; e, se existe a necessidade de um modelo capaz de realizar previsões antes do término do processo de rotulação. A delimitação adequada dos possíveis candidatos aos papéis dessas duas importantes componentes, estratégia e algoritmo, é fundamental para o sucesso do aprendizado de máquina ativo e do sistema de recomendação automática, caso seja adotado.

Conclusão

```
um dia os sóis  
    acabam  
e nada mais  
existe  
  
um dia o Sol  
    não nasce  
um dia o Sol  
    se apaga  
um dia ou não  
    acordo  
um dia virá  
    antes  
e antes desse  
um outro  
o outro e por  
diante  
e outro e outro  
e hoje1
```

Na literatura de aprendizado ativo, muitas abordagens têm sido propostas para a realização de consultas relevantes junto ao oráculo. O principal aspecto dessa área, explorado neste trabalho, é a pouca atenção dada à relevância dos algoritmos de aprendizado enquanto aprendizes, especialmente se consideradas as especificidades de cada estratégia de amostragem e cada conjunto de dados. Foi necessária, portanto, uma investigação dos fatores que influenciam o *viés de aprendizado ativo*, tais como a presença e tipo do *viés de aprendizado* do algoritmo e o *viés de amostragem* da estratégia.

A investigação conduzida resultou nesta tese, que mostra ser possível atingir melhores desempenhos preditivos de acordo com as propriedades do conjunto de dados e o momento em que se situe o processo de aprendizado. O meio proposto é uma escolha (ou inibição) mais criteriosa, preferencialmente automática, do tipo do viés de aprendizado, da estratégia ou de ambos.

O desenvolvimento da pesquisa requereu trabalho conceitual, de criação, de implementação e metodológico. O autor deixa neste parágrafo registrada a extensão do esforço

¹ 17/04/2014 - Baseado no conto *The last question* de Isaac Asimov ([ASIMOV et al., 2007](#)).

empreendido no desenvolvimento e a carga computacional resultante dos experimentos. Embora os experimentos realizados nesta pesquisa tenham tido um elevado custo computacional, sua real finalidade é a economia em termos de esforço humano, provavelmente mais custoso que a energia consumida. Nesse aspecto, este trabalho pode ser considerado um projeto bem sucedido, no mínimo, pela confirmação da efetividade do aprendizado ativo em geral. Além disso, a pesquisa também resulta em benefícios intangíveis, que estão além da mera aplicabilidade prática de seus resultados. Durante a pesquisa, 13 estratégias e variações foram estudadas, implementadas e propostas, totalizando 22 alternativas, se considerados os parâmetros explorados. Uma coleção com 90 conjuntos de dados foi elaborada visando, tanto quanto possível, um embasamento estatístico confiável para as observações feitas. Foram realizados experimentos no nível base, aprendizado de máquina convencional, e no nível meta, para o meta-aprendizado. Quatro algoritmos de aprendizado, com vieses e desempenhos preditivos variados, e dois comitês foram empregados no nível base. No nível meta, duas medidas de referência e quatro comitês foram utilizados. O esforço de implementação e custo computacional podem ser resumidos pelo número de 40 milhões de consultas registradas; rotuladas por um incansável oráculo, que felizmente era simulado - mais detalhes são dados no Apêndice D. Numa aplicação real, essa tarefa seria atribuída a um especialista no domínio correspondente.

Por fim, esta pesquisa enfrentou algumas dificuldades, listadas na Seção 6.1, que não impediram o atingimento das metas. As contribuições, cristalizadas pela comprovação das hipóteses formuladas, são apresentadas nas seções 6.2 e 6.3, respectivamente. Adicionalmente, algumas limitações e possibilidades de trabalhos futuros foram identificadas e apresentadas nas seções 6.4 e 6.5, respectivamente.

6.1 Dificuldades

Embora essencial para aumentar a generalidade dos resultados, o tamanho da coleção organizada criou uma dificuldade sem relatos prévios na pesquisa bibliográfica realizada. Normalmente, as estratégias são testadas em poucos conjuntos (Seção 1.1) frequentemente acompanhados da ilustração individual de suas respectivas curvas de aprendizado. Durante a pesquisa que resultou no presente trabalho, houve diversas tentativas de sumarização de todas as curvas no mesmo gráfico, como: normalização de medidas ou mesmo de compatibilização entre diferentes limites de orçamento; contagem de vitórias; e, neutralização das especificidades de cada conjunto de dados por meio da subtração do desempenho de estratégias de referência (Rnd, limites teóricos superior e inferior etc.) - entre outras. Finalmente, as curvas de ranqueamento, propostas na Seção 4.6, se mostraram a solução mais direta do ponto de vista conceitual, desde que todos os conjuntos de dados fossem consultados com o mesmo limite de orçamento. Uma limitação desse tipo de curva, de certa forma positiva, é requerer que muitos conjuntos sejam adotados - o oposto

da limitação das curvas convencionais, para as quais apenas poucos conjuntos podem ser viavelmente representados.

Há uma tendência ao emprego da ALC na literatura (*Area Under the Learning Curve* - Seção 4.5.1). Contudo, não há consenso sobre a medida que a deva compor. Esse problema também ocorre em outras áreas que envolvam a tarefa de classificação. Optou-se pelo índice κ (kappa multiclasse - Seção 3.1.2) devido à sua interpretabilidade e à sua dependência de poucos dados, permitindo que apenas a matriz de confusão fosse armazenada. Os dados necessários para o cálculo de medidas como a *Area Under the ROC Curve* (AUC) (LOBO; JIMÉNEZ-VALVERDE; REAL, 2008) fariam com que a base de dados excedesse o espaço disponível em disco. A alternativa, que seria registrar o valor da AUC e descartar dos demais dados, aumentaria o custo computacional, não seria adequada para todos os algoritmos adotados e não permitiria a troca para outras medidas de desempenho posteriormente.

O cálculo da distância de Mahalanobis mostrou-se proibitivamente custoso, fazendo com que essa métrica fosse adotada apenas no experimento de recomendação de métricas de distância (Seção 5.3.3). Provavelmente, o uso de toda a reserva de exemplos para seu cálculo não tenha sido uma boa escolha.

Por fim, as bibliotecas de estratégias implementadas por terceiros se mostraram muito modestas, com poucas, ou pouco relevantes, abordagens implementadas. Essas bibliotecas compreendem um grupo heterogêneo de linguagens de programação e de cenários de aprendizado ativo. Alguns exemplos encontram-se na página de Burr Settles na *internet*². A única implementação de terceiros efetivamente utilizada nos experimentos foi a referente à estratégia HS. Apesar disso, seu uso também requereu uma implementação, parcial, que fornecesse o pré-processamento da reserva de exemplos na forma de agrupamento hierárquico (Seção 4.4.2).

6.2 Metas atingidas

Um sumário dos objetivos desta tese de doutorado, delineados previamente na Seção 1.3, e seus respectivos resultados obtidos nos experimentos realizados são descritos a seguir:

- **Identificação da existência de relações de adequação entre nichos de problemas e estratégias:** a presença dessas relações foi identificada por meio da análise qualitativa de árvores de decisão baseadas em meta-atributos humanamente interpretáveis; adicionalmente, foi constatado que a escolha do algoritmo de aprendizado precede as demais variáveis na determinação do sucesso da estratégia (Seção

² <<http://active-learning.net>> - Acessado em 07/01/2016.

5.1.4).

- **Desenvolvimento de uma estratégia capaz de suprimir, sob demanda, a influência do aprendiz:** a estratégia HTU foi proposta como uma alternativa para controlar a atuação do aprendiz na estratégia baseada em densidade, TU; ela se mostrou competitiva frente às demais estratégias consideradas e apresentou propriedades relevantes, como estabilidade, segurança e custo computacional interconsultas *humanamente tolerável* (Seção 5.1.3) - características que permitem redução nos custos de esforço humano.
- **Desenvolvimento de um aprendiz meta-ativo:** a abordagem de recomendação automática proposta, chamada *aprendizado meta-ativo*, *superou*³, no nível base, o uso de um único algoritmo de aprendizado específico e mostrou-se viável, no nível meta, com diversos algoritmos no papel de meta-aprendizes (principalmente PCT, RoF e RFw). Além da recomendação automática de algoritmos de aprendizado no contexto de aprendizado ativo, outras modalidades de recomendação também se mostraram promissoras (dentro das limitações metodológicas constatadas posteriormente), como a recomendação de estratégias e pares estratégia-algoritmo.

Tais resultados incluíram a comprovação das hipóteses, conforme descrito a seguir, na Seção 6.3.

6.3 Hipóteses comprovadas

A hipótese principal, de que, em tarefas de classificação, relações entre conjuntos de dados, algoritmos de aprendizado ativo e estratégias de amostragem ativa podem ser exploradas visando um maior desempenho preditivo frente à escolha arbitrária é válida, dentro das limitações experimentais. Essa conclusão se baseia na comparação de medidas relevantes de acurácia como: a ALC da medida μ_K ; a correlação entre ranqueamentos preditos e esperados; e, a acurácia ordinária, a acurácia balanceada e a medida κ no nível meta.

Da mesma forma, a hipótese secundária, de que o aprendiz pode ser automaticamente inibido ou substituído com vantagem durante o aprendizado, também foi demonstrada válida. Essa conclusão se baseia na comparação da estratégia HTU com sua antecessora na literatura, cuja presença do aprendiz é constante, e também com estratégias representantes de outros paradigmas. Também contribui para a validade da hipótese os resultados favoráveis da comparação do meta-aprendiz, ativado antes da 1ª e depois

³ A validade da avaliação do sucesso do meta-aprendizado é limitada pelo método experimental empregado. Algumas análises posteriores (Apêndice B) apontam no sentido de ser necessário maior rigor experimental.

da 50ª consulta, com as outras estratégias no nível base e com valores de referência no nível meta.

As seguintes afirmações foram empiricamente verificadas, segundo o método experimental empregado e considerados os conjuntos de dados, o conjunto específico de estratégias e os algoritmos adotados.

1. HTUeuc tem o desempenho mais consistente, ou seja, é a mais segura em termos financeiros/de esforço humano.
2. ATUeuc e HTUeuc apresentam a menor variabilidade de desempenho preditivo (σ_{κ}).
3. Dentre as estratégias com melhor desempenho, ATUeuc e HTUeuc possuem o menor custo computacional entre consultas e, conseqüentemente, de esforço humano.
4. A recomendação automática de algoritmos aumenta o desempenho da estratégia.
5. Diferentes algoritmos podem gerar metaclassificadores com desempenhos preditivos similares, seguindo a abordagem de recomendação automática proposta.
6. Embora requeira um método experimental mais rigoroso (Apêndice B), a recomendação automática de algoritmos, estratégias ou pares estratégia-algoritmo inicialmente mostrou-se viável.

Quanto às outras modalidades de recomendação automática, foi observado que pesquisas adicionais ainda são necessárias para que se possa chegar a conclusões melhor fundamentadas.

6.4 Limitações

Durante e após a pesquisa realizada nesta tese, algumas limitações foram identificadas. Elas não puderam ser superadas devido a diversos fatores, como baixa prioridade, percepção ou aparecimento tardio, impossibilidade prática, ausência de consenso, limitação intrínseca, entre outros.

Uma limitação comum de trabalhos na literatura de classificação também está presente neste. Cada domínio de aplicação tem um modelo de custos de erro de classificação mais adequado e, conseqüentemente, requer uma medida de avaliação de desempenho apropriada. Logo, a adoção uniforme da medida de desempenho no nível base (κ) para todos os conjuntos de dados configura-se como uma limitação metodológica. Uma investigação aprofundada do domínio de cada conjunto de dados seria capaz de identificar a medida mais apropriada para cada conjunto, de acordo com as melhores práticas na literatura para o domínio do problema em questão. A solução ideal seria que cada conjunto de

dados pré-processado fosse disponibilizado com essa informação. Entretanto, em muitos casos não há um critério para isso.

A subamostragem mencionada na Seção 5.1.1 reduziu a reserva para 100 exemplos, visando maior tratabilidade computacional de EER (estratégia baseada na redução do erro esperado). Tal decisão pode ter beneficiado essa estratégia. Era esperada uma maior similaridade entre o comportamento de EER e o das outras estratégias não agnósticas. Assim, idealmente, todas as estratégias deveriam compartilhar dessa manipulação na reserva para garantir uma comparação em iguais condições. Embora essa parte do método experimental pudesse ser melhorado, seu provável impacto foi aumentar a competitividade de EER, limitando-se a reduzir a visibilidade do sucesso das propostas.

Lughofer (2012) publicou uma estratégia híbrida baseada em agrupamento da qual o autor deste documento não tomou conhecimento em tempo hábil para uma devida comparação com HTU. Embora a comparação com aquela e as diversas outras estratégias não contempladas pelos experimentos pudesse enriquecer este trabalho, a sua ausência não impactou seriamente os resultados, pois o objetivo deste trabalho se concentrou mais nos problemas de escolha no cenário de aprendizado ativo, como a estratégia e o tipo ou o momento de inibição de algoritmos de aprendizado, do que no desempenho de estratégias específicas.

Uma limitação importante de HTU é seu parâmetro ρ_{limite} . Apesar da superioridade de seu desempenho ter sido demonstrada experimentalmente, a validade do princípio de funcionamento motivador de sua proposta não foi comprovada. Seria necessário comparar HTU com uma estratégia híbrida aleatória ou conforme alguma heurística pré-definida, ou seja, com outros critérios de alternância entre ATU e TU que servissem de referência. Um indício de que a medida de correlação de Pearson talvez não quantifique a grandeza desejada, como o nível de contribuição da componente exploratória, é que, na maioria dos casos testados, $\rho_{\text{limite}} < 0$ levou a desempenhos melhores do que $\rho_{\text{limite}} > 0$ - conforme apresentado previamente na Figura 7. Isso precisaria ser investigado em mais detalhes, pois contradiz a motivação da proposta.

O cenário adotado é artificial, caso sejam considerados alguns aspectos de aplicações reais. Um desses aspectos é sobre os exemplos duplicados, ou seja, com os valores dos atributos coincidentes, mas com rótulos conflitantes. Eles foram evitados nos experimentos para melhor isolamento do objeto de estudo (Capítulo 4), apesar de serem um tópico ativo de pesquisa (IPEIROTIS *et al.*, 2014; SHENG; PROVOST; IPEIROTIS, 2008).

As estratégias DW, TU, ATU e HTU poderiam ter tido um melhor desempenho se os coeficientes de ponderação houvessem sido ajustados de acordo com alguma heurística, pois eles balanceiam a importância da densidade e do aprendiz. Nos experimentos, esses parâmetros foram arbitrariamente mantidos com o valor 1.

Souto *et al.* (2008) reportaram resultados sobre recomendação de algoritmos de agrupamento possivelmente mais efetivos, embora trate-se de outro domínio e outro aparato experimental. Segundo esses autores, as predições do meta-aprendiz atingiram um valor médio de correlação 0,75 contra o valor 0,59, de referência. O conjunto de meta-atributos proposto por eles poderia ter sido adotado.

Um dos pressupostos do sistema de recomendação automática proposto é que o conjunto inicial de treinamento seria pequeno demais para que fosse possível realizar uma seleção de modelos que visasse a escolha do melhor algoritmo de aprendizado. Realmente, a quantidade de apenas $|Y|$ exemplos antes do início do processo de rotulação inviabiliza qualquer tentativa de seleção via validação cruzada. Entretanto, após 50 consultas, a possibilidade de seleção do melhor algoritmo por meio de validação cruzada poderia ter sido verificada e comparada com a seleção automática. Ainda que 50 exemplos venham a ser insuficientes, espera-se que, em algum ponto da curva de aprendizado, dado o crescimento do conjunto de treinamento, a convencional seleção por meio de validação cruzada se torne preferível à seleção automática. Trata-se, assim, de uma referência a ser adotada em futuras comparações. Quanto maior o orçamento, mais importante se torna essa referência.

Por fim, uma limitação metodológica, presente também em outros trabalhos, diz respeito à escolha ideal da coleção de conjuntos de dados. Mesmo com o procedimento de eliminação de conjuntos muito similares, a dependência remanescente entre conjuntos teve uma influência quantificável nos resultados. Essa influência é analisada no Apêndice B. Resumidamente, as medidas de desempenho foram elevadas, em parte, por conta dessa parcial *dependência entre amostras* (conjuntos de dados similares). Apesar disso, a recomendação automática na parcela mais independente da coleção também superou a referência, mas em muito menor grau. Embora a presença de domínios similares tenha facilitado a tarefa de recomendação, isso não reduz a necessidade prática do sistema de recomendação - apenas sugere que ele seja, naturalmente, mais efetivo em coleções que contenham conjuntos cujos domínios sejam próximos ao do conjunto que se pretenda rotular. Diante dessa redução na generalidade dos resultados obtidos, ***ainda não é possível afirmar com segurança que meta-aprendizado seja a melhor forma de resolver os problemas de escolha envolvidos no aprendizado ativo.***

6.5 Desdobramentos Futuros

Um aprofundamento nas análises dos experimentos ainda é necessário. Trata-se de uma abordagem cujo princípio de funcionamento e efetividade requerem análises ainda mais rigorosas do que aquelas empregadas neste documento. Por exemplo, na Figura 41 do Apêndice B, é possível identificar os conjuntos de dados mais difíceis para a tarefa

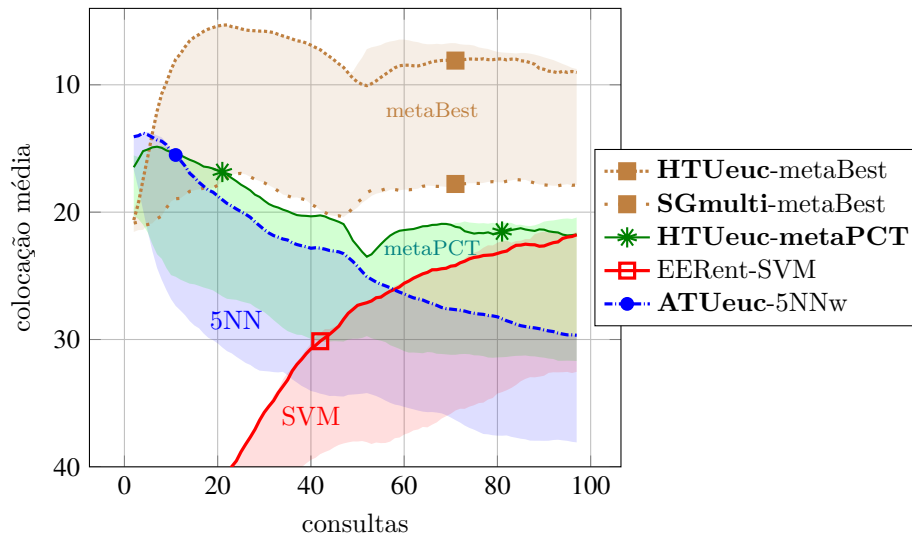


Figura 38 – Curvas de ranqueamento - incluindo recomendação perfeita. Medida comparada: μ_K . Detalhes na Figura 21.

de recomendação de algoritmos de aprendizado. Os motivos de cada dificuldade podem abrir possibilidades de melhoria nas propostas de recomendação automática, caso sejam identificados; ou, podem indicar não tratar-se de um assunto prioritário de pesquisa. Uma análise inicial indica que foram favorecidos justamente os conjuntos que dispuseram da presença de outros conjuntos do mesmo domínio no metaconjunto de treinamento. Logo, a coleção, mesmo sendo diversa, na realidade não continha conjuntos cuja extração de meta-atributos pudesse ser útil no aproveitamento de conhecimento entre domínios.

O autor desta tese considera a exploração das potenciais abordagens situadas na intersecção entre as áreas de aprendizado ativo e meta-aprendizado ainda em seus primórdios, pois abrange uma lacuna literária a ser preenchida. De fato, há margem para melhorias significativas conforme sugere a Figura 38. Nessa figura, é possível observar a faixa de colocações obtidas pelas estratégias quando dispõem do meta-aprendiz perfeito (metaBest), isto é, aquele que é capaz de prever, com acesso desleal às metaclasses, sempre o melhor algoritmo nas duas metades do intervalo de consultas. Até mesmo o limite inferior da faixa, que normalmente corresponde à pior estratégia, foi capaz de superar as demais abordagens por praticamente todo o período. Embora a faixa de metaBest certamente seja fruto de sobreajuste aos dados, ela permite verificar que a curva de metaPCT não é pressionada pelo limite do que é teoricamente possível.

A opção por dois momentos de recomendação automática, um para cada metade do período de consultas, foi arbitrária. A existência de um vale precisamente em torno do momento de transição (metade do orçamento, $\phi = 50$) sugere que o momento ideal para a reconsideração do melhor algoritmo varie de um conjunto de dados para outro - esse comportamento já havia se manifestado nos experimentos da Seção 5.1.5. Por exemplo,

na vigésima consulta, a curva de metaBest começa a ceder colocações; conforme a quantidade de consultas aumenta, maior a quantidade de conjuntos de dados em que algoritmos distintos do inicialmente escolhido tornam-se os mais adequados. Dado que a troca de algoritmo foi definida para ocorrer apenas naquele ponto (50^a consulta), é somente a partir dali que as estratégias puderam usufruir de um algoritmo mais adequado para o intervalo de consultas corrente. Essa constatação poderia ser explorada adicionando-se mais momentos em que o algoritmo de recomendação pudesse reconsiderar a escolha do algoritmo do aprendiz. No caso limite, o algoritmo do aprendiz poderia ser reconsiderado dinamicamente a cada consulta, tal como a abordagem de Rossi (2014) em fluxos de dados; porém, com um custo computacional \propto vezes maior na etapa de treinamento do algoritmo de recomendação. Dessa forma, uma extensão natural do trabalho seria a recomendação automática constante, isto é, a cada nova consulta. O limite teórico aumentaria consideravelmente para as curvas que fossem obtidas com essa eventual extensão do trabalho.

Um conjunto reduzido com apenas os meta-atributos mais relevantes pode ser investigado, de forma a aumentar a acurácia preditiva. Seria interessante também, conforme sugestão da Banca, investigar qual o padrão dos conjuntos de dados que fazem com que seja preferível a amostragem aleatória.

Outro desdobramento, mais imediato, seria a implementação da *metaestratégia* e, possivelmente, do *metapar estratégia-algoritmo*, para que a viabilidade dessas modalidades pudesse ser confirmada também no nível base, por meio de comparações das ALCs e das curvas de ranqueamento.

Por fim, outra possibilidade de meta-aprendizado aplicado a aprendizado ativo é a recomendação de parâmetros de estratégias.

Nota do autor

Espero que esta honesta jornada escrita, às vezes desenhada, tenha sido tão (ou mais) estimulante para o leitor quanto foi para mim em alguns momentos. Registro aqui meu desejo de ver mais instrução epistemológica em nossas universidades, de história da ciência e tudo o mais que se espera de conhecimento que colabore na formação de futuros portadores do título de Doutor em Ciências.

Referências

- ABE, N.; MAMITSUKA, H. Query learning strategies using boosting and bagging. In: **International Conference on Machine Learning**. [S.l.]: Morgan Kaufmann, 1998. p. 1–9. ISBN 1-55860-556-8. Citado 3 vezes nas páginas 50, 53 e 69.
- ARLOT, S.; CELISSE, A. *et al.* A survey of cross-validation procedures for model selection. **Statistics surveys**, Publicado pelo autor, Creative Commons Attribution License, v. 4, p. 40–79, 2010. Citado 2 vezes nas páginas 80 e 88.
- ASIMOV, I.; FLICK, B. E.; GALLANT, J.; THEATRE, T. E. R. **The last question**. [S.l.]: Ziggurat Productions, 2007. Citado na página 127.
- ATTENBERG, J.; PROVOST, F. J. Inactive learning?: difficulties employing active learning in practice. **Special Interest Group on Knowledge Discovery and Data Mining Explorations**, v. 12, n. 2, p. 36–41, 2010. Citado 2 vezes nas páginas 65 e 123.
- _____. Why label when you can search?: alternatives to active learning for applying human resources to build classification models under extreme class imbalance. In: **Proceedings of the 16th International Conference on Knowledge Discovery and Data Mining**. [S.l.]: ACM, 2010. p. 423–432. ISBN 978-1-4503-0055-1. Citado na página 73.
- BACHE, K.; LICHMAN, M. **UCI repository of machine learning databases**. Irvine, CA, 2013. Citado 3 vezes nas páginas 57, 60 e 74.
- BALCAN, M.-F.; BEYGELZIMER, A.; LANGFORD, J. Agnostic active learning. **Journal of Computing Systems and Sciences**, v. 75, n. 1, p. 78–89, 2009. Citado na página 44.
- BARAM, Y.; EL-YANIV, R.; LUZ, K. Online choice of active learning algorithms. **Journal of Machine Learning Research**, v. 5, p. 255–291, 2004. Citado na página 88.
- BAUER, E.; KOHAVI, R. An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. **Machine learning**, Springer, v. 36, n. 1-2, p. 105–139, 1999. Citado na página 50.
- BECKER, M.; OSBORNE, M. A two-stage method for active learning of statistical grammars. In: **International Joint Conference on Artificial Intelligence**. [S.l.]: Professional Book Center, 2005. p. 991–996. ISBN 0938075934. Citado na página 88.

BEDO, M. V. N.; SANTOS, D. P.; PONCIANO-SILVA, M.; AZEVEDO-MARQUES, P. M.; CARVALHO, A. P. L. F.; TRAINA JR., C. Endowing a Content-Based Medical Image Retrieval system with perceptual similarity using ensemble strategy. **Journal of Digital Imaging**, v. 1, n. 1, p. 1–18, 2015. Citado na página 37.

BEDO, M. V. N.; SANTOS, D. P. dos; KASTER, D. S.; JR, C. T. A similarity-based approach for financial time series analysis and forecasting. In: **Database and Expert Systems Applications**. [S.l.]: Springer, 2013. (Lecture Notes in Computer Science, v. 8056), p. 94–108. ISBN 978-3-642-40172-5. Citado na página 37.

BENSUSAN, H.; GIRAUD-CARRIER, C. G.; KENNEDY, C. J. A higher-order approach to meta-learning. In: **Inductive Logic Programming Work-in-progress reports**. [S.l.]: CEUR-WS.org, 2000. (CEUR Workshop Proceedings, v. 35). Citado na página 156.

BISWAS, A.; PARIKH, D. Simultaneous active learning of classifiers & attributes via relative feedback. In: **Conference on Computer Vision and Pattern Recognition**. [S.l.]: IEEE, 2013. p. 644–651. Citado na página 73.

BOUCKAERT, R. R.; FRANK, E. Evaluating the replicability of significance tests for comparing learning algorithms. In: **Advances in Knowledge Discovery and Data Mining**. [S.l.]: Springer, 2004. (Lecture Notes in Computer Science, v. 3056), p. 3–12. ISBN 3-540-22064-X. Citado 2 vezes nas páginas 33 e 88.

BRAZDIL, P.; GIRAUD-CARRIER, C. G.; SOARES, C.; VILALTA, R. **Metalearning - Applications to Data Mining**. [S.l.]: Springer, 2009. I–X, 1–176 p. (Cognitive Technologies). ISBN 978-3-540-73262-4. Citado 4 vezes nas páginas 33, 64, 155 e 157.

BRAZDIL, P.; HENERY, R. Analysis of results. **Machine learning, neural and statistical classification**, Ellis Horwood, New York, p. 175–212, 1994. Citado 3 vezes nas páginas 68, 75 e 157.

BRAZDIL, P.; SOARES, C.; COSTA, J. P. da. Ranking learning algorithms: Using IBL and meta-learning on accuracy and time results. **Machine Learning**, v. 50, n. 3, p. 251–277, 2003. Citado na página 88.

BREIMAN, L. Bagging predictors. **Machine learning**, Springer, v. 24, n. 2, p. 123–140, 1996. Citado na página 50.

_____. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001. Citado 2 vezes nas páginas 60 e 83.

BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A.; STONE, C. J. **Classification and Regression Trees**. [S.l.]: Wadsworth, 1984. ISBN 0-534-98053-8. Citado na página 114.

BRYANT, C. H.; MUGGLETON, S.; OLIVER, S. G.; KELL, D. B.; REISER, P. G. K.; KING, R. D. Combining inductive logic programming, active learning and robotics to discover the function of genes. **Electronic Transactions on Artificial Intelligence**, v. 5, n. B, p. 1–36, 2001. Citado na página 43.

CHEN, Y.; LASKO, T. A.; MEI, Q.; DENNY, J. C.; XU, H. A study of active learning methods for named entity recognition in clinical text. **Journal of biomedical informatics**, Elsevier, v. 58, p. 11–18, 2015. Citado na página 88.

CHERMAN, E. A. **Aprendizado de máquina multirrótulo: explorando a dependência de rótulos e o aprendizado ativo**. Tese (Doutorado) — Universidade de São Paulo, 2013. Citado na página 73.

COHN, D. A.; ATLAS, L. E.; LADNER, R. E. Improving generalization with active learning. **Machine Learning**, v. 15, n. 2, p. 201–221, 1994. Citado 2 vezes nas páginas 48 e 53.

COLARES, P. F. **Processo de Indução e Ranqueamento de Árvores de Decisão Sobre Modelos Olap. 2011. 109 f.** Dissertação (Mestrado) — Pontifícia Universidade Católica do Rio Grande do Sul, 2005. Citado na página 36.

CORTES, C.; VAPNIK, V. Support-vector networks. **Machine Learning**, v. 20, n. 3, p. 273–297, 1995. Citado na página 82.

CRAWFORD, M. M.; TUIA, D.; YANG, H. L. Active learning: Any value for classification of remotely sensed data? **Proceedings of the IEEE**, v. 101, n. 3, p. 593–608, 2013. Citado na página 73.

DASGUPTA, S. Two faces of active learning. **Theoretical Computer Science**, v. 412, n. 19, p. 1767–1781, 2011. Citado 3 vezes nas páginas 48, 53 e 69.

DASGUPTA, S.; HSU, D.; MONTELEONI, C. A general agnostic active learning algorithm. In: **International Symposium on Artificial Intelligence and Mathematics**. [S.l.: s.n.], 2008. Citado na página 56.

DELGADO, M. F.; CERNADAS, E.; BARRO, S.; AMORIM, D. G. Do we need hundreds of classifiers to solve real world classification problems? **Journal of Machine Learning Research**, v. 15, n. 1, p. 3133–3181, 2014. Citado 3 vezes nas páginas 75, 83 e 123.

DEMIRKESEN, C.; CHERIFI, H. A comparison of multiclass SVM methods for real world natural scenes. In: **Advanced Concepts for Intelligent Vision systems**. [S.l.]: Springer, 2008. (Lecture Notes in Computer Science, v. 5259), p. 752–763. ISBN 978-3-540-88457-6. Citado na página 87.

DEMŠAR, J. Statistical comparisons of classifiers over multiple data sets. **Journal of Machine Learning Research**, v. 7, p. 1–30, 2006. Citado 3 vezes nas páginas 89, 91 e 117.

DEVLIN, K. **The joy of sets: fundamentals of contemporary set theory**. [S.l.]: Springer Science & Business Media, 2012. Citado na página 40.

DIETTERICH, T. G. Ensemble methods in machine learning. In: **Multiple classifier systems**. [S.l.]: Springer, 2000. p. 1–15. Citado na página 50.

DUNN, J. C. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. Taylor & Francis, 1973. Citado na página 68.

ELDER-IV, J. F. The generalization paradox of ensembles. **Journal of Computational and Graphical Statistics**, v. 12, n. 4, 2003. Citado na página 84.

EMAM, K. E. Benchmarking kappa: Interrater agreement in software process assessments. **Empirical Software Engineering**, v. 4, n. 2, p. 113–133, 1999. Citado na página 87.

ERTEKIN, S.; 0002, J. H.; BOTTOU, L.; GILES, C. L. Learning on the border: active learning in imbalanced data classification. In: **Proceedings of the Sixteenth Conference on Information and Knowledge Management**. [S.l.]: ACM, 2007. p. 127–136. ISBN 978-1-59593-803-9. Citado na página 88.

EUGENIO, B. D.; GLASS, M. The kappa statistic: A second look. **Computational Linguistics**, v. 30, n. 1, p. 95–101, 2004. Citado na página 86.

EVANS, J. D. **Straightforward statistics for the behavioral sciences**. [S.l.]: Brooks/Cole, 1996. Citado na página 59.

EVANS, L. P. G.; ADAMS, N. M.; ANAGNOSTOPOULOS, C. When does active learning work? **Computing Research Repository**, abs/1408.1319, 2014. Citado 2 vezes nas páginas 31 e 123.

FAYYAD, U. M.; IRANI, K. B. Multi-interval discretization of continuous-valued attributes for classification learning. In: **International Joint Conference on Artificial Intelligence**. [S.l.]: Morgan Kaufmann, 1993. p. 1022–1029. ISBN 1-55860-300-X. Citado na página 82.

FERRARI, D. G.; CASTRO, L. N. de. Clustering algorithm selection by meta-learning systems: A new distance-based problem characterization and ranking combination methods. **Information Sciences**, v. 301, n. 0, p. 181 – 194, 2015. ISSN 0020-0255. Citado 2 vezes nas páginas 68 e 157.

FLEISS, J. L.; LEVIN, B.; PAIK, M. C. **Statistical methods for rates and proportions**. [S.l.]: John Wiley & Sons, 2013. Citado na página 60.

FORMAN, G. Programmer’s active learning: A broader perspective of choices for real-world classification tasks that matter. In: **Proceedings of the Active Learning in Real-world Applications**. [S.l.]: Springer, 2012. Citado na página 43.

FREUND, Y.; SCHAPIRE, R. E. Experiments with a new boosting algorithm. In: **International Conference on Machine Learning**. [S.l.]: Morgan Kaufmann, 1996. p. 148–156. ISBN 1-55860-419-7. Citado na página 83.

FUJII, A.; INUI, K.; TOKUNAGA, T.; TANAKA, H. Selective sampling for example-based word sense disambiguation. **Computational Linguistics**, v. 24, n. 4, p. 573–597, 1998. Citado 3 vezes nas páginas 47, 53 e 69.

GARCIA, L. P. F. A huge collection of preprocessed ARFF datasets for supervised classification problems. **GitHub Software Repository**, Jan 2015. DOI:10.5281/zenodo.13748. URL:<http://dx.doi.org/10.5281/zenodo.13748>. Citado na página 74.

GARCIA, L. P. F.; CARVALHO, A. C. P. F. de; LORENA, A. C. Noise detection in the meta-learning level. **Neurocomputing**, 2015. In press 2015. Citado na página 157.

GIBBONS, J. D.; CHAKRABORTI, S. Nonparametric statistical inference. In: **International Encyclopedia of Statistical Science**. [S.l.]: Springer, 2011. p. 977–979. ISBN 978-3-642-04897-5; 978-3-642-04898-2. Citado na página 75.

- GOMES, T. A. F.; PRUDÊNCIO, R. B. C.; SOARES, C.; ROSSI, A. L. D.; CARVALHO, A. C. P. L. F. Combining meta-learning and search techniques to select parameters for support vector machines. **Neurocomputing**, v. 75, n. 1, p. 3–13, 2012. Citado na página 157.
- GU, Y.; JIN, Z.; CHIU, S. C. Combining active learning and semi-supervised learning using local and global consistency. In: **International Conference on Neural Information Processing**. [S.l.]: Springer, 2014. (Lecture Notes in Computer Science, v. 8834), p. 215–222. ISBN 978-3-319-12636-4. Citado na página 73.
- GUO, Y.; GREINER, R. Optimistic active-learning using mutual information. In: **International Joint Conference on Artificial Intelligence**. [S.l.: s.n.], 2007. p. 823–829. Citado 3 vezes nas páginas 50, 53 e 69.
- GUO, Y.; SCHUURMANS, D. Discriminative batch mode active learning. In: **Neural Information Processing Systems**. [S.l.]: Curran Associates, Inc, 2007. Citado na página 73.
- GUYON, I.; CAWLEY, G. C.; DROR, G.; LEMAIRE, V. Results of the active learning challenge. In: **Active Learning and Experimental Design**. [S.l.]: Journal of Machine Learning Research.org, 2011. v. 16, p. 19–45. Citado na página 32.
- HALL, M. A.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The WEKA data mining software: an update. **Special Interest Group on Knowledge Discovery and Data Mining Explorations**, v. 11, n. 1, p. 10–18, 2009. Citado 2 vezes nas páginas 78 e 86.
- HAMMING, R. W. Error detecting and error correcting codes. **Bell System technical journal**, Wiley Online Library, v. 29, n. 2, p. 147–160, 1950. Citado na página 83.
- HANSEN, L. K.; SALAMON, P. Neural network ensembles. **IEEE Transactions on Pattern Analysis & Machine Intelligence**, IEEE, n. 10, p. 993–1001, 1990. Citado na página 50.
- HART, P. E. The condensed nearest neighbor rule (corresp.). **IEEE Transactions on Information Theory**, v. 14, n. 3, p. 515–516, 1968. Citado na página 79.
- HAYKIN, S.; NETWORK, N. A comprehensive foundation. **Neural Networks**, v. 2, n. 2004, 2004. Citado 4 vezes nas páginas 41, 48, 51 e 155.
- HEARST, M. A.; DUMAIS, S.; OSMAN, E.; PLATT, J.; SCHOLKOPF, B. Support vector machines. **Intelligent Systems and their Applications, IEEE**, IEEE, v. 13, n. 4, p. 18–28, 1998. Citado 2 vezes nas páginas 46 e 79.
- HOCHBAUM; SHMOYS. A best possible heuristic for the k-center problem. **Mathematics of Operations Research**, v. 10, n. 2, p. 180–184, 1985. Citado na página 49.
- HONORE, C. In praise of slowness. **Law Prac.**, HeinOnline, v. 32, p. 34, 2006. Citado na página 55.
- IPEIROTIS, P. G.; PROVOST, F. J.; SHENG, V. S.; WANG, J. Repeated labeling using multiple noisy labelers. **Data Mining Knowledge Discovery**, v. 28, n. 2, p. 402–441, 2014. Citado na página 132.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: A review. **Computing Surveys**, ACM, v. 31, n. 3, p. 264–323, 1999. Citado na página 49.

JAPKOWICZ, N.; SHAH, M. **Evaluating Learning Algorithms: A Classification Perspective**. [S.l.]: Cambridge University Press, 2011. ISBN 9780521196000. Citado 2 vezes nas páginas 74 e 90.

JOACHIMS, T. Estimating the generalization performance of an SVM efficiently. In: **International Conference on Machine Learning**. [S.l.]: Morgan Kaufmann, 2000. p. 431–438. ISBN 1-55860-707-2. Citado na página 88.

JONES, S. L. P. **Haskell 98 language and libraries: the revised report**. [S.l.]: Cambridge University Press, 2003. Citado na página 160.

KALOUSIS, A. **Algorithm selection via meta-learning**. Tese (Doutorado) — Université de Geneve, 2002. Citado 3 vezes nas páginas 67, 68 e 157.

KANDA, J.; CARVALHO, A. C. P. L. F. de; HRUSCHKA, E. R.; SOARES, C. Selection of algorithms to solve traveling salesman problems using meta-learning. **International Journal of Hybrid Intelligent Systems**, v. 8, n. 3, p. 117–128, 2011. Citado na página 157.

KEARNS, M. J.; RON, D. Algorithmic stability and sanity-check bounds for leave-one-out cross-validation. **Neural Computation**, v. 11, n. 6, p. 1427–1453, 1999. Citado na página 88.

KEARNS, M. J.; SCHAPIRE, R. E.; SELLIE, L. Toward efficient agnostic learning. **Machine Learning**, v. 17, n. 2-3, p. 115–141, 1994. Citado 2 vezes nas páginas 47 e 56.

KÖRNER, C.; WROBEL, S. Multi-class ensemble-based active learning. In: **17th European Conference on Machine Learning**. [S.l.]: Springer, 2006. (Lecture Notes in Computer Science, v. 4212), p. 687–694. ISBN 3-540-45375-X. Citado 2 vezes nas páginas 31 e 88.

KREYSZIG, E. **Advanced engineering mathematics**. [S.l.]: John Wiley & Sons, 2007. Citado 2 vezes nas páginas 78 e 81.

KRISHNAPURAM, B.; YU, S.; RAO, R. B. **Cost-sensitive machine learning**. [S.l.]: CRC Press, Inc., 2011. Citado na página 43.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: **Neural Information Processing Systems**. [S.l.: s.n.], 2012. p. 1106–1114. Citado na página 29.

LETRAS, A. B. de. **Vocabulário ortográfico da língua portuguesa**. Academia Brasileira de Letras, 2004. Disponível em: <<http://www.academia.org.br/nossa-lingua/busca-no-vocabulario>>. Citado 2 vezes nas páginas 29 e 36.

LEWIN, J.; SCHMITT, A. O.; ADORJÁN, P.; 0002, T. H.; PIEPENBROCK, C. Quantitative DNA methylation analysis based on four-dye trace data from direct sequencing of PCR amplicates. **Bioinformatics**, v. 20, n. 17, p. 3005–3012, 2004. Citado na página 46.

LEWIS, D. D. A sequential algorithm for training text classifiers: Corrigendum and additional data. **Special Interest Group on Information Retrieval Forum**, v. 29, n. 2, p. 13–19, 1995. Citado na página 45.

_____. Naive (bayes) at forty: The independence assumption in information retrieval. In: **10th European Conference on Machine Learning**. [S.l.]: Springer, 1998. (Lecture Notes in Computer Science, v. 1398), p. 4–15. ISBN 3-540-64417-2. Citado 2 vezes nas páginas 57 e 79.

LIMA, R. **Gramática normativa da língua portuguesa**. [S.l.]: J. Olympio, 1973. Citado na página 29.

LIN, J. Divergence measures based on the shannon entropy. **IEEE Transactions on Information Theory**, v. 37, n. 1, p. 145–151, 1991. Citado na página 50.

LOBO, J. M.; JIMÉNEZ-VALVERDE, A.; REAL, R. Auc: a misleading measure of the performance of predictive distribution models. **Global ecology and Biogeography**, Wiley Online Library, v. 17, n. 2, p. 145–151, 2008. Citado na página 129.

LOMASKY, R.; BRODLEY, C. E.; AERNECKE, M.; WALT, D.; FRIEDL, M. A. Active class selection. In: **18th European Conference on Machine Learning**. [S.l.]: Springer, 2007. (Lecture Notes in Computer Science, v. 4701), p. 640–647. ISBN 978-3-540-74957-8. Citado na página 88.

LUGHOFFER, E. Hybrid active learning for reducing the annotation effort of operators in classification systems. **Pattern Recognition**, v. 45, n. 2, p. 884–896, 2012. Citado na página 132.

MAHALANOBIS, P. C. On the generalized distance in statistics. **Proceedings of the National Institute of Sciences (Calcutta)**, v. 2, p. 49–55, 1936. Citado na página 119.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. **Introduction to information retrieval**. [S.l.]: Cambridge University Press, 2008. I–XXI, 1–482 p. ISBN 978-0-521-86571-5. Citado na página 82.

MASSO, M.; VAISMAN, I. I. Accurate and efficient gp120 V3 loop structure based models for the determination of HIV-1 co-receptor usage. **BioMed Central Bioinformatics**, v. 11, p. 494, 2010. Citado na página 85.

MELVILLE, P.; MOONEY, R. J. Diverse ensembles for active learning. In: **International Conference on Machine Learning**. [S.l.]: ACM, 2004. (ACM International Conference Proceeding Series, v. 69). Citado 2 vezes nas páginas 50 e 88.

MILLER, R. B. Response time in man-computer conversational transactions. In: **Fall Joint Computing Conference**. [S.l.]: Thomson Book Company, 1968. (American Federation of Information Processing Societies Conference Proceedings, v. 33), p. 267–277. Citado na página 101.

MITCHELL, T. M. The need for biases in learning generalizations. In: **Readings in Machine Learning**. [S.l.]: Morgan Kauffman, 1980. p. 184–191. Book published in 1990. Citado na página 155.

_____. **Machine learning**. [S.l.]: McGraw-Hill, 1997. I–XVII, 1–414 p. (McGraw Hill Series in Computer Science). ISBN 978-0-07-042807-2. Citado na página 44.

MURTAGH, F. A survey of recent advances in hierarchical clustering algorithms. **Computing Journal**, v. 26, n. 4, p. 354–359, 1983. Citado na página 48.

MUSLEA, I.; MINTON, S.; KNOBLOCK, C. A. Active + semi-supervised learning = robust multi-view learning. In: **International Conference on Machine Learning**. [S.l.]: Morgan Kaufmann, 2002. p. 435–442. ISBN 1-55860-873-7. Citado na página 88.

_____. Active learning with strong and weak views: A case study on wrapper induction. In: **International Joint Conference on Artificial Intelligence**. [S.l.]: Morgan Kaufmann, 2003. p. 415–420. Citado na página 88.

NAH, F. F.-H. A study on tolerable waiting time: How long are web users willing to wait? In: **Americas Conference on Information Systems**. [S.l.]: Association for Information Systems, 2003. p. 285. Citado 3 vezes nas páginas 44, 100 e 101.

OSHIRO, T. M.; PEREZ, P. S.; BARANAUSKAS, J. A. How many trees in a random forest? In: **Machine Learning and Data Mining in Pattern Recognition**. [S.l.]: Springer, 2012. (Lecture Notes in Computer Science, v. 7376), p. 154–168. ISBN 978-3-642-31536-7. Citado na página 84.

OSUGI, T. T.; DENG, K.; SCOTT, S. D. Balancing exploration and exploitation: A new algorithm for active machine learning. In: **International Conference on Data Mining**. [S.l.]: IEEE Computer Society, 2005. p. 330–337. ISBN 0-7695-2278-5. Citado 3 vezes nas páginas 49, 53 e 69.

PATRA, S.; BRUZZONE, L. A cluster-assumption based batch mode active learning technique. **Pattern Recognition Letters**, v. 33, n. 9, p. 1042–1048, 2012. Citado na página 73.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011. Citado 2 vezes nas páginas 114 e 160.

PERLICH, C.; PROVOST, F. J.; SIMONOFF, J. S. Tree induction vs. logistic regression: A learning-curve analysis. **Journal of Machine Learning Research**, v. 4, p. 211–255, 2003. Citado na página 65.

PFAHRINGER, B.; BENSUSAN, H.; GIRAUD-CARRIER, C. Tell me who can learn you and i can tell you who you are: Landmarking various learning algorithms. In: **Proceedings of the Seventeenth International Conference on Machine Learning**. [S.l.: s.n.], 2000. p. 743–750. Citado na página 156.

POPPER, K. A survey of some fundamental problems. **The logic of scientific discovery**, p. 27–47, 1959. Citado na página 39.

QUINLAN, J. R. **C4.5: Programs for Machine Learning**. [S.l.]: Morgan Kaufmann, 1993. ISBN 1-55860-238-0. Citado na página 79.

- RAGHAVAN, H.; MADANI, O.; JONES, R. When will feature feedback help? quantifying the complexity of classification problems. In: **International Joint Conference on Artificial Intelligence Workshop on Human in the Loop Computing**. [S.l.: s.n.], 2007. Citado na página 86.
- RENDELL, L. A.; SHESHU, R.; TCHENG, D. K. Layered concept-learning and dynamically variable bias management. In: **Proceedings of the 10th International Joint Conference on Artificial Intelligence**. [S.l.]: Morgan Kaufmann, 1987. p. 308–314. Citado 2 vezes nas páginas 68 e 157.
- RICCI, F.; ROKACH, L.; SHAPIRA, B. Introduction to recommender systems handbook. In: **Recommender Systems Handbook**. [S.l.]: Springer, 2011. p. 1–35. ISBN 978-0-387-85819-7. Citado na página 29.
- RODRÍGUEZ, J. J.; KUNCHEVA, L. I.; ALONSO, C. J. Rotation forest: A new classifier ensemble method. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 28, n. 10, p. 1619–1630, 2006. Citado na página 83.
- ROSS, S. M. **Introduction to probability and statistics for engineers and scientists (2. ed.)**. [S.l.]: Academic Press, 2000. I–XIV, 1–578 p. ISBN 978-0-12-598472-0. Citado na página 59.
- ROSSI, A. L. D. **Meta-aprendizado aplicado a fluxos contínuos de dados**. Tese (Doutorado) — Universidade de São Paulo, 2014. Citado na página 135.
- ROSSI, A. L. D.; CARVALHO, A. C. P. de Leon Ferreira de; SOARES, C.; SOUZA, B. F. de. Metastream: A meta-learning based method for periodic algorithm selection in time-changing data. **Neurocomputing**, v. 127, p. 52–64, 2014. Citado na página 157.
- ROUSSEEUW, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. **Journal of computational and applied mathematics**, Elsevier, v. 20, p. 53–65, 1987. Citado na página 68.
- ROY, N.; MCCALLUM, A. Toward optimal active learning through sampling estimation of error reduction. In: **International Conference on Machine Learning**. [S.l.]: Morgan Kaufmann, 2001. p. 441–448. ISBN 1-55860-778-1. Citado 2 vezes nas páginas 73 e 82.
- SANTAFE, G.; INZA, I.; LOZANO, J. A. Dealing with the evaluation of supervised classification algorithms. **Artificial Intelligence Review**, Springer, v. 44, n. 4, p. 467–508, 2015. Citado na página 90.
- SANTOS, D. P.; CARVALHO, A. C. P. L. F. Comparison of active learning strategies and proposal of a multiclass hypothesis space search. In: **Proceedings of the 9th International Conference on Hybrid Artificial Intelligence Systems**. [S.l.]: Springer, 2014. (Lecture Notes in Computer Science, v. 8480), p. 618–629. ISBN 978-3-319-07616-4. Citado 4 vezes nas páginas 36, 62, 69 e 71.
- _____. Viabilidade do aprendizado ativo em máquinas extremas. In: **Brazilian Conference on Intelligent Systems and Encontro Nacional de Inteligência Artificial e Computacional**. [S.l.: s.n.], 2014. Citado na página 36.
- _____. An Active Learning Library for Scala. **GitHub Software Repository**, Jan 2015. DOI:10.5281/zenodo.13733. URL:http://dx.doi.org/10.5281/zenodo.13733. Citado na página 37.

_____. Selectively inhibiting learning bias for active sampling. In: **Proceedings of the IV Brazilian Conference on Intelligent Systems**. Natal – Brasil: IEEE, 2015. To be published. Citado 5 vezes nas páginas 36, 56, 58, 69 e 85.

_____. Wrappers for Weka for Scala. **GitHub Software Repository**, Jan 2015. DOI:10.5281/zenodo.13735. URL: <http://dx.doi.org/10.5281/zenodo.13735>. Citado na página 37.

_____. Automatic selection of learning bias for active sampling. **Artificial Intelligence Review**, Springer, 2016. To be published. Citado na página 36.

_____. Investigation of different approaches for active learning. **Neurocomputing**, Elsevier, 2016. To be published. Citado na página 36.

SAUSSURE, F. d. **Curso de lingüística geral. Trad. Antônio Chelini, José Paulo Paes e Izidoro Blikstein**. [S.l.]: São Paulo: Cultrix/USP, 1972. Citado na página 71.

SCHAFFER, C. A conservation law for generalization performance. In: **International Conference on Machine Learning**. [S.l.]: Morgan Kaufmann, 1994. p. 259–265. ISBN 1-55860-335-2. Citado 2 vezes nas páginas 33 e 155.

SCHAPIRE, R. E. The strength of weak learnability. **Machine learning**, Springer, v. 5, n. 2, p. 197–227, 1990. Citado na página 50.

SCHEIN, A. I.; UNGAR, L. H. Active learning for logistic regression: an evaluation. **Machine Learning**, v. 68, n. 3, p. 235–265, 2007. Citado na página 31.

SCHMIDT, W. F.; KRAAIJVELD, M.; DUIN, R. P. *et al.* Feedforward neural networks with random weights. In: IEEE. **Proceedings of the 11th International Conference on Pattern Recognition Methodology and Systems**. [S.l.], 1992. v. 2, p. 1–4. Citado na página 36.

SEGGERN, D. H. V. **CRC standard curves and surfaces with Mathematica**. [S.l.]: CRC Press, 2006. Citado na página 41.

SENI, G.; ELDER-IV, J. F. **Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions**. [S.l.]: Morgan & Claypool Publishers, 2010. (Synthesis Lectures on Data Mining and Knowledge Discovery). Citado na página 84.

SETTLES, B. **Curious machines: active learning with structured instances**. Tese (Doutorado) — University of Madison Wisconsin, 2008. Citado 3 vezes nas páginas 47, 53 e 69.

SETTLES, B. Active learning literature survey. **University of Wisconsin, Madison**, v. 52, p. 55–66, 2010. Citado 6 vezes nas páginas 47, 53, 57, 69, 72 e 87.

_____. **Active Learning**. [S.l.]: Morgan & Claypool, 2012. (Synthesis Lectures on Artificial Intelligence and Machine Learning). Citado 7 vezes nas páginas 39, 42, 53, 62, 69, 73 e 85.

SETTLES, B.; CRAVEN, M. An analysis of active learning strategies for sequence labeling tasks. In: **Proceedings of the Conference on Empirical Methods on Natural Language Processing**. [S.l.]: ACL, 2008. p. 1070–1079. Citado 3 vezes nas páginas 31, 73 e 86.

SETTLES, B.; CRAVEN, M.; FRIEDLAND, L. Active learning with real annotation costs. In: **Proceedings of the Neural Information Processing Systems workshop on cost-sensitive learning**. [S.l.: s.n.], 2008. p. 1–10. Citado na página 88.

SETTLES, B.; CRAVEN, M.; RAY, S. Multiple-instance active learning. In: **Neural Information Processing Systems**. [S.l.]: Curran Associates, Inc, 2007. Citado 4 vezes nas páginas 51, 53, 69 e 73.

SHAH, M. Generalized agreement statistics over fixed group of experts. In: **Machine Learning and Knowledge Discovery in Databases**. [S.l.]: Springer, 2011. (Lecture Notes in Computer Science, v. 6913), p. 191–206. ISBN 978-3-642-23807-9. Citado na página 60.

SHENG, V. S.; PROVOST, F. J.; IPEIROTIS, P. G. Get another label? improving data quality and data mining using multiple, noisy labelers. In: LI, Y.; 0001, B. L.; SARAWAGI, S. (Ed.). **Proceedings of the 14th International Conference on Knowledge Discovery and Data Mining**. [S.l.]: ACM, 2008. p. 614–622. ISBN 978-1-60558-193-4. Citado na página 132.

SOUSA, A. F. M.; PRUDÊNCIO, R. B. C.; SOARES, C.; LUDERMIR, T. B. Active selection of training instances for a random forest meta-learner. In: **International Joint Conference on Neural Networks**. [S.l.]: IEEE, 2013. p. 1–7. ISBN 978-1-4673-6128-6. Citado 2 vezes nas páginas 34 e 64.

SOUTO, M. C. P.; PRUDÊNCIO, R. B. C.; SOARES, R. G. F.; ARAUJO, D. S. A. de; COSTA, I. G.; LUDERMIR, T. B.; SCHLIEP, A. Ranking and selecting clustering algorithms using a meta-learning approach. In: **International Joint Conference on Neural Networks**. [S.l.]: IEEE, 2008. p. 3729–3735. Citado 4 vezes nas páginas 68, 87, 133 e 157.

SOUZA, B. F. d. **Meta-aprendizagem aplicada à classificação de dados de expressão gênica**. Tese (Doutorado) — Universidade de São Paulo, 2010. Citado na página 68.

SOUZA, B. F. de; CARVALHO, A. C. P. L. F. de; SOARES, C. Empirical evaluation of ranking prediction methods for gene expression data classification. In: **Ibero-American Conference on Artificial Intelligence**. [S.l.]: Springer, 2010. (Lecture Notes in Computer Science, v. 6433), p. 194–203. ISBN 978-3-642-16951-9. Citado na página 157.

STEHRMAN, S. V. Selecting and interpreting measures of thematic classification accuracy. **Remote sensing of Environment**, Elsevier, v. 62, n. 1, p. 77–89, 1997. Citado na página 41.

STIKIC, M.; LAERHOVEN, K. V.; SCHIELE, B. Exploring semi-supervised and active learning for activity recognition. In: **International Semantic Web Conference**. [S.l.]: IEEE Computer Society, 2008. p. 81–88. ISBN 978-1-4244-2637-9. Citado na página 88.

STROBL, C.; MALLEY, J.; TUTZ, G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. **Psychological methods**, American Psychological Association, v. 14, n. 4, p. 323, 2009. Citado na página 83.

TEAM, R. C. **R: A language and environment for statistical computing**. **R Foundation for Statistical Computing, Vienna, Austria, 2012**. [S.l.]: ISBN 3-900051-07-0, 2014. Citado 2 vezes nas páginas 84 e 160.

TODOROVSKI, L.; BLOCKEEL, H.; DZEROSKI, S. Ranking with predictive clustering trees. In: **13th European Conference on Machine Learning**. [S.l.]: Springer, 2002. (Lecture Notes in Computer Science, v. 2430), p. 444–455. ISBN 3-540-44036-4. Citado na página 83.

TONG, S.; KOLLER, D. Support vector machine active learning with applications to text classification. **Journal of Machine Learning Research**, v. 2, p. 45–66, 2001. Citado 3 vezes nas páginas 46, 53 e 69.

VALENTINI, G.; MASULLI, F. Ensembles of learning machines. In: **Neural Nets**. [S.l.]: Springer, 2002. p. 3–20. Citado na página 50.

VALIANT, L. G. A theory of the learnable. **Communications of the ACM**, v. 27, n. 11, p. 1134–1142, 1984. Citado na página 43.

VILALTA, R.; DRISSI, Y. A perspective view and survey of meta-learning. **Artificial Intelligence Review**, v. 18, n. 2, p. 77–95, 2002. Citado na página 155.

WANG, K.; LIANG, C.; LIU, J.; XIAO, H.; HUANG, S.; XU, J.-H.; LI, F. Prediction of piRNAs using transposon interaction and a support vector machine. **BioMed Central Bioinformatics**, v. 15, p. 6593, 2014. Citado na página 29.

WILKINSON, L.; FRIENDLY, M. The history of the cluster heat map. **The American Statistician**, v. 63, n. 2, 2009. Citado na página 76.

WOLPERT, D.; MACREADY, W. G. No free lunch theorems for optimization. **IEEE Transactions on Evolutionary Computation**, v. 1, n. 1, p. 67–82, 1997. Citado 2 vezes nas páginas 33 e 155.

WOLPERT, D. H. Stacked generalization. **Neural Networks**, v. 5, n. 2, p. 241–259, 1992. Citado na página 156.

XU, R.; WUNSCH, D. **Clustering**. [S.l.]: John Wiley & Sons, 2008. Citado na página 68.

XU, Z.; AKELLA, R.; 0001, Y. Z. Incorporating diversity and density in active learning for relevance feedback. In: **29th European Conference on Advances in Information Retrieval**. [S.l.]: Springer, 2007. (Lecture Notes in Computer Science, v. 4425), p. 246–257. ISBN 978-3-540-71494-1. Citado na página 88.

Comitês como aprendizes ativos

Um novo experimento foi realizado, avaliando HTU e também a recomendação de algoritmos de aprendizado de máquina em outras condições. Por conveniência, apenas os aspectos mais relevantes do método experimental adotado nas seções anteriores foram considerados. Todos os experimentos reportados até este ponto no texto foram baseados no mesmo grupo de algoritmos de aprendizado enquanto aprendizes ativos: 5NNw, C4.5w, NB e SVM.

O novo experimento procura verificar o efeito da adoção, como aprendizes ativos, de um grupo de dois algoritmos baseados em comitê, (aprendiz ativo)RFw e (aprendiz ativo)RoF, possivelmente com maior poder preditivo do que o grupo anterior - conforme discutido no Capítulo 4. O uso do prefixo (*aprendiz ativo*) visa evitar ambiguidades com a utilização, em experimentos prévios, desses dois algoritmos em outra tarefa completamente distinta. Eles são agora adotados como aprendizes contidos nas estratégias, com apenas 10 árvores cada; diferentemente do papel de meta-aprendizes que desempenharam nos experimentos anteriores - em que puderam ser utilizadas 500 árvores.

A Figura 39 contém as curvas de ranqueamento geradas com os dois algoritmos pertencentes ao novo grupo e também a faixa de metaPCT, que alterna entre (aprendiz ativo)RFw e (aprendiz ativo)RoF conforme predito pelo sistema de recomendação automática. As únicas curvas explícitas são da melhor estratégia de cada faixa, de todas as estratégias da faixa metaPCT e de todos os pares contendo HTUeuc ou SGmulti.

Como pode ser visto, há uma intersecção expressiva entre as faixas correspondentes aos aprendizes base: (aprendiz ativo)RFw e (aprendiz ativo)RoF. Consequentemente, em geral, houve pouca margem de exploração do período de consultas (tanto na primeira quanto na segunda metade) por parte do meta-aprendiz PCT e da referência metaDef. De fato, a região hachurada, que corresponde à faixa de metaPCT, se estende praticamente do topo do gráfico até sua base, com exceção das estreitas áreas de RoF e RFw acima à esquerda e direita, respectivamente; e, RFw abaixo. Essa observação é confirmada quantitativamente pelo desempenho no nível meta, de acordo com os valores de $ALC-\mu\kappa$ - conforme apresentado na Figura 40.

Nessa figura, as três linhas em itálico contêm diferenças estatisticamente significa-

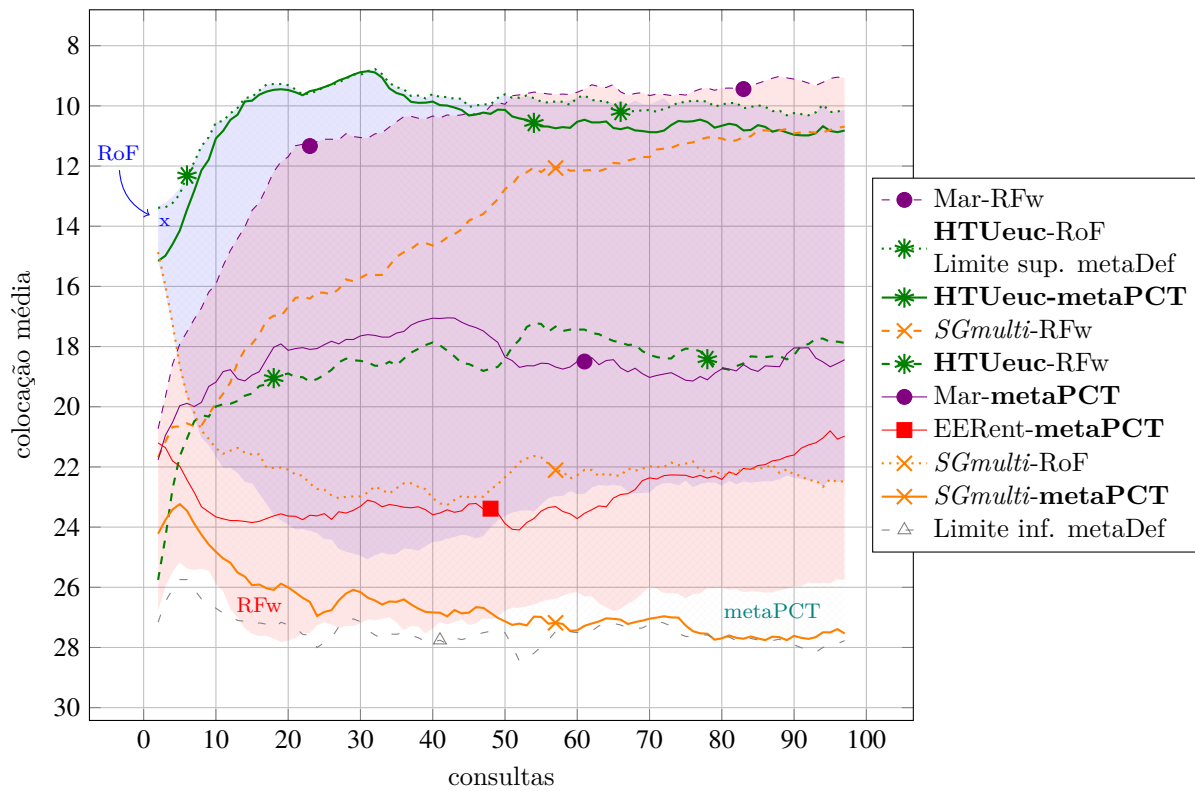


Figura 39 – Curvas de ranqueamento com faixas RFW, RoF e metaPCT. Medida comparada: μ_κ . O limite superior de metaDef coincide com a melhor estratégia de RoF, HTUeuc. Detalhes na Figura 21.

Estratégia	PCT	Def	α	p-valor
TUman	0,492	0,495		0,944
HTUeuc	0,491	0,494	+	0,014
HTUman	0,489	0,491		0,684
TUeuc	0,489	0,492		0,541
ATUman	0,468	0,469		1,000
ATUeuc	0,468	0,468		0,684
Mar	0,451	0,444		0,120
EERacc	0,442	0,444		0,616
HS	0,425	0,431	•	0,078
Rnd	0,417	0,420		0,780
EERent	0,427	0,402	+	0,042
DWeuc	0,401	0,409	+	0,041
SGmulti	0,406	0,403		0,955
DWman	0,398	0,405		0,166
Vitórias	3	11		

Figura 40 – Comparação de valores de ALC- μ_κ com RFW e RoF como aprendizes. Linhas em itálico indicam diferença com significância estatística a favor da referência. Legenda na Tabela 7.

Tabela 15 – Significância estatística das diferenças (RFw). Medida: $ALC-\mu_K$. *Legenda na Tabela 7.*

Estratégia	1	2	3	4	5	6	7	8	9
1 - ATUeuc	-		.						
2 - HS		-	*						
3 - DWeuc			-						
4 - EERent	*	*	*	-	*		*		*
5 - HTUeuc	.		*		-				
6 - Mar	*	*	*		*	-	*		*
7 - Rnd			+				-		
8 - SGmulti	*	*	*		*		*	-	+
9 - TUEuc	+		*				.		-

Tabela 16 – Contagem de colocações (RFw). Medida: $ALC-\mu_K$. *Detalhes na Tabela 8.*

Estratégia	Primeiras colocações	Derrotas para Rnd	Últimas colocações
Mar	49	14	1
SGmulti	10	7	1
EERent	8	17	2
ATUeuc	7	47	13
HTUeuc	6	38	2
Rnd	4	-	15
TUEuc	3	37	4
HS	3	25	0
DWeuc	1	65	52

Tabela 17 – Significância estatística das diferenças (RoF). Medida: $ALC-\mu_K$. *Legenda na Tabela 7.*

Estratégia	1	2	3	4	5	6	7	8	9
1 - ATUeuc	-	*	*				*	*	
2 - HS		-							
3 - DWeuc			-						
4 - EERent		*	*	-			*	*	
5 - HTUeuc	*	*	*	*	-	+	*	*	
6 - Mar		*	*			-	*	*	
7 - Rnd							-		
8 - SGmulti								-	
9 - TUEuc	*	*	*	+			*	*	-

Tabela 18 – Contagem de colocações (RoF). Medida: $ALC-\mu_K$. *Detalhes na Tabela 8.*

Estratégia	Primeiras colocações	Derrotas para Rnd	Últimas colocações
HTUeuc	33	3	0
Mar	18	7	1
TUEuc	16	6	0
ATUeuc	12	11	4
EERent	6	6	2
DWeuc	6	50	45
SGmulti	0	33	13
Rnd	0	-	17
HS	0	28	11

tivas a favor da referência, enquanto que em apenas um caso a diferença com significância foi a favor de PCT. Logo, os resultados indicam um desempenho similar para PCT e Def, ou mesmo a favor de Def, dada a maior quantidade de vitórias: 11 contra 3. Assim, **com o emprego dos comitês RFw e RoF como aprendizes, a recomendação automática não foi capaz de superar a referência.**

Por outro lado, a estratégia proposta, HTUeuc, obteve a melhor curva na primeira metade, utilizando RoF como aprendiz. Na segunda metade, contrariando todos resultados na Seção 5.1, a estratégia Mar obteve a melhor curva, utilizando RFw como aprendiz. O par HTUeuc-RoF obteve uma curva próxima, logo abaixo. Outra diferença com relação aos experimentos anteriores é o desempenho relativamente superior de SGmulti, utilizando RFw, especialmente na segunda metade.

O desempenho quantitativo das estratégias em todo o intervalo de consultas, separado por algoritmo, é apresentado com significância estatística nas tabelas 15 e 16 para RFw; e, 17 e 18 para RoF. Com o algoritmo RFw como aprendiz, a estratégia Mar se destacou com 49 primeiras colocações e apenas uma última colocação. Não foi detectada

vantagem com significância estatística de Mar contra EERent e SGmulti. As três superaram as demais com diferenças estatisticamente significativas. Nesse caso específico, ***com RFW como aprendiz, a adaptação proposta, SGmulti, mostrou-se competitiva.***

Com o algoritmo RoF, a situação se altera. HTUeuc passa a superar todas as demais, exceto TUEuc, com diferenças estatisticamente significativas; obtendo 33 primeiras colocações e nenhuma última colocação. Mar passa a obter o segundo melhor número de primeiras colocações, e mantém apenas uma última colocação.

Diante desses resultados heterogêneos, é possível concluir que, apesar de serem dois comitês baseados em árvores e com desempenhos similares de acordo com a Figura 40, as estratégias responderam à utilização deles de maneira distinta. Assim, ***a estratégia com melhor desempenho pode variar de um algoritmo para outro, mesmo que eles sejam competitivos entre si ou baseados em princípios similares.***

Finalmente, esse experimento adicional com comitês no papel de aprendizes ativos confirmou que as estratégias são dependentes do algoritmo adotado. Observou-se, também, que a diferenciação automática entre RFW e RoF não foi possível, supostamente devido a ambos apresentarem desempenhos comparáveis durante todo o aprendizado, como pode ser suposto a partir da grande intersecção entre suas faixas de curvas de ranqueamento.

Dependência entre conjuntos de dados

Mesmo com o procedimento de eliminação de conjuntos muito similares, a dependência remanescente entre conjuntos teve uma influência quantificável nos resultados.

A dificuldade imposta pelos conjuntos de dados pode ser estimada pela diferença entre os coeficientes de correlação de Spearman de PCT (*Predictive Clustering Trees*) e Def (ranqueamento médio) para cada metaexemplo. Assim, é possível comparar o nível de dificuldade oferecido pelos metaexemplos correspondentes a domínios similares e aquele oferecido pelos demais metaexemplos. Por simplicidade, foi escolhida apenas uma das melhores estratégias, HTUeuc, neste ensaio.

Na Figura 41, a coluna *Valor* indica a diferença de correlações por metaexemplo cujo conjunto de dados correspondente tem seu nome dado na primeira coluna. A segunda coluna (*#sim*) contém a quantidade de metaexemplos utilizados para treinamento cujos domínios são similares ao domínio do conjunto de dados indicado na primeira coluna. A coluna *Média* contém apenas dois valores: a média das diferenças de correlação para os metaexemplos com mais de zero similares no conjunto de treinamento e a média das diferenças de correlações para os demais metaexemplos.

Nas duas metades das consultas, o maior valor médio das diferenças de correlações ocorreu para os metaexemplos com maior quantidade de similares: na primeira metade, 0,08 contra 0,04; e, na segunda metade, 0,28 contra 0,04. Isso sugere que a coleção de conjuntos adotada possa ter elevado as medidas de desempenho por conta da existência de dependência entre as amostras (conjuntos de dados). Apesar disso, a média das diferenças entre PCT e Def ainda é positiva nas duas metades, indicando um desempenho acima da referência, mesmo em metaexemplos correspondentes a domínios independentes, isto é, sem similar no conjunto de treinamento. Embora a presença de domínios similares facilite a tarefa de recomendação, ela não reduz a necessidade do sistema de recomendação, apenas sugere que ele seja, naturalmente, mais efetivo em coleções que contenham conjuntos de domínios similares ao do conjunto que se pretenda rotular.

Primeira metade das consultas					Segunda metade das consultas				
Conjunto de dados	#sim	Corr. Spearman			Conjunto de dados	#sim	Corr. Spearman		
		Valor	Média	Desvio padrão			Valor	Média	Desvio padrão
1-statlog-vehicle-silhouettes	4,0	0,22			1-statlog-german-credit	4,0	0,78		
2-statlog-australian-credit	4,0	0,19			2-autoUniv-au7-300-drift-au7-cpd1-800	4,0	0,63		
3-autoUniv-au7-700	4,0	0,03			3-statlog-australian-credit	4,0	0,56		
4-autoUniv-au7-300-drift-au7-cpd1-800	4,0	0,02			4-autoUniv-au6-cd1-400	4,0	0,49		
5-autoUniv-au1-1000	4,0	0,00			5-autoUniv-au7-cpd1-500	4,0	0,46		
6-autoUniv-au6-cd1-400	4,0	0,00			6-statlog-image-segmentation	4,0	0,33		
7-statlog-heart	4,0	0,00			7-statlog-heart	4,0	0,18		
8-statlog-german-credit	4,0	-0,09			8-autoUniv-au7-700	4,0	0,12		
9-autoUniv-au7-cpd1-500	4,0	-0,15			9-statlog-vehicle-silhouettes	4,0	0,00		
10-statlog-image-segmentation	4,0	-0,20			10-autoUniv-au1-1000	4,0	-0,59		
11-volcanoes-d1	3,0	0,36			11-thyroid-sick-euthyroid	3,0	1,01		
12-volcanoes-a3	3,0	0,34			12-thyroid-hypothyroid	3,0	0,83		
13-volcanoes-b5	3,0	0,32			13-spect-heart	3,0	0,74		
14-volcanoes-e1	3,0	0,23			14-volcanoes-d1	3,0	0,65		
15-spect-heart	3,0	0,20			15-volcanoes-e1	3,0	0,58		
16-thyroid-sick-euthyroid	3,0	0,10			16-volcanoes-a3	3,0	0,49		
17-heart-disease-processed-cleveland	3,0	0,05	0,08	0,24	17-volcanoes-b5	3,0	0,22	0,28	0,41
18-heart-disease-processed-va	3,0	0,00			18-thyroid-ann	3,0	0,11		
19-heart-disease-processed-hungarian	3,0	-0,05			19-heart-disease-processed-cleveland	3,0	0,06		
20-thyroid-ann	3,0	-0,09			20-heart-disease-processed-hungarian	3,0	0,02		
21-thyroid-newthyroid	3,0	-0,24			21-thyroid-newthyroid	3,0	-0,36		
22-thyroid-hypothyroid	3,0	-0,26			22-heart-disease-processed-va	3,0	-0,63		
23-wine-quality-red	2,0	0,31			23-wine	2,0	0,39		
24-wine-quality-white-5class	2,0	0,25			24-wine-quality-white-5class	2,0	-0,03		
25-wine	2,0	0,00			25-wine-quality-red	2,0	-0,19		
26-monks3	1,0	0,85			26-monks3	1,0	1,00		
27-monks1	1,0	0,59			27-connectionist-vowel	1,0	0,43		
28-vertebra-column-2c	1,0	0,14			28-monks1	1,0	0,38		
29-connectionist-mines-vs-rocks	1,0	0,00			29-vertebra-column-3c	1,0	0,33		
30-connectionist-vowel	1,0	0,00			30-vertebra-column-2c	1,0	0,31		
31-robot-nav-sensor-readings-2	1,0	-0,02			31-robot-failure-lp5	1,0	0,24		
32-vertebra-column-3c	1,0	-0,08			32-robot-nav-sensor-readings-2	1,0	0,00		
33-robot-failure-lp5	1,0	-0,22			33-connectionist-mines-vs-rocks	1,0	-0,35		
34-nursery	0,0	0,39			34-horse-colic-surgical	0,0	1,29		
35-flare	0,0	0,34			35-molecular-splice-junction	0,0	0,80		
36-voting	0,0	0,33			36-kr-vs-kp	0,0	0,75		
37-molecular-splice-junction	0,0	0,30			37-optdigits	0,0	0,75		
38-bupa	0,0	0,28			38-hepatitis	0,0	0,61		
39-horse-colic-surgical	0,0	0,27			39-leaf	0,0	0,55		
40-page-blocks	0,0	0,26			40-texture	0,0	0,50		
41-habermans-survival	0,0	0,24			41-mfeat-fourier	0,0	0,50		
42-cardiotocography-3class	0,0	0,23			42-mammographic-mass	0,0	0,50		
43-first-order-theorem	0,0	0,21			43-voting	0,0	0,43		
44-mammographic-mass	0,0	0,21			44-first-order-theorem	0,0	0,42		
45-artificial-characters	0,0	0,20			45-waveform-v2	0,0	0,36		
46-qsar-biodegradation	0,0	0,14			46-lymphography	0,0	0,30		
47-ecg-eye-state	0,0	0,14			47-indian-liver-patient	0,0	0,30		
48-ringnorm	0,0	0,14			48-wholesale-channel	0,0	0,29		
49-turkiye-student	0,0	0,13			49-pima-indians-diabetes	0,0	0,22		
50-kr-vs-kp	0,0	0,13			50-spambase	0,0	0,22		
51-magic	0,0	0,13			51-abalone-3class	0,0	0,22		
52-hepatitis	0,0	0,12			52-saheart	0,0	0,18		
53-indian-liver-patient	0,0	0,11			53-musk	0,0	0,16		
54-climate-simulation-crashes	0,0	0,11			54-yeast-4class	0,0	0,12		
55-banknote-authentication	0,0	0,10			55-ionosphere	0,0	0,10		
56-pima-indians-diabetes	0,0	0,10			56-habermans-survival	0,0	0,07		
57-car-evaluation	0,0	0,09			57-flare	0,0	0,06		
58-phoneme	0,0	0,06			58-car-evaluation	0,0	0,02		
59-pendigits	0,0	0,06			59-nursery	0,0	0,02		
60-lymphography	0,0	0,03			60-artificial-characters	0,0	0,00		
61-wilt	0,0	0,03			61-phoneme	0,0	0,00		
62-steel-plates-faults	0,0	0,02	0,04	0,18	62-ringnorm	0,0	0,00	0,04	0,41
63-banana	0,0	0,02			63-turkiye-student	0,0	-0,01		
64-ecoli	0,0	0,00			64-page-blocks	0,0	-0,01		
65-glass	0,0	0,00			65-movement-libras	0,0	-0,02		
66-hill-valley-without-noise	0,0	0,00			66-glass	0,0	-0,04		
67-ionosphere	0,0	0,00			67-parkinsons	0,0	-0,04		
68-iris	0,0	0,00			68-iris	0,0	-0,05		
69-leaf	0,0	0,00			69-climate-simulation-crashes	0,0	-0,05		
70-mfeat-fourier	0,0	0,00			70-magic	0,0	-0,06		
71-movement-libras	0,0	0,00			71-qsar-biodegradation	0,0	-0,06		
72-ozone-eighthr	0,0	0,00			72-hill-valley-without-noise	0,0	-0,07		
73-parkinsons	0,0	0,00			73-ecg-eye-state	0,0	-0,07		
74-saheart	0,0	0,00			74-seeds	0,0	-0,15		
75-seeds	0,0	0,00			75-banknote-authentication	0,0	-0,20		
76-synthetic-control	0,0	0,00			76-ozone-eighthr	0,0	-0,21		
77-waveform-v2	0,0	0,00			77-bupa	0,0	-0,25		
78-user-knowledge	0,0	-0,01			78-ecoli	0,0	-0,26		
79-abalone-3class	0,0	-0,03			79-wdbc	0,0	-0,26		
80-spambase	0,0	-0,10			80-cardiotocography-3class	0,0	-0,34		
81-musk	0,0	-0,11			81-synthetic-control	0,0	-0,35		
82-optdigits	0,0	-0,15			82-banana	0,0	-0,37		
83-twonorm	0,0	-0,17			83-balance-scale	0,0	-0,38		
84-texture	0,0	-0,18			84-steel-plates-faults	0,0	-0,39		
85-wholesale-channel	0,0	-0,24			85-wilt	0,0	-0,49		
86-tic-tac-toe	0,0	-0,26			86-pendigits	0,0	-0,50		
87-balance-scale	0,0	-0,27			87-user-knowledge	0,0	-0,52		
88-wdbc	0,0	-0,38			88-twonorm	0,0	-0,60		
89-mushroom	0,0	-0,41			89-mushroom	0,0	-0,83		
90-yeast-4class	0,0	-0,49			90-tic-tac-toe	0,0	-1,03		

#sim: quantidade de domínios similares no conjunto de treinamento

Figura 41 – Coeficiente de correlação de Spearman por conjunto de dados.

Meta-aprendizado

Um sistema de classificação baseado em aprendizado de máquina depende de um modelo induzido por um algoritmo (Seção 2). Diante da infinidade de vieses de aprendizado possíveis, muitos algoritmos de aprendizado têm sido propostos e alguns são frequentemente empregados de forma generalizada na solução dos mais diversos problemas, como é o caso das redes neurais artificiais (HAYKIN; NETWORK, 2004). Entretanto, nenhum algoritmo pode ser adequado a todos os domínios. Equivalentemente, um desempenho positivo em algumas situações de aprendizado precisa ser compensado por um igual grau de desempenho negativo em outras (WOLPERT; MACREADY, 1997; SCHAFFER, 1994). Isso decorre da existência de um viés necessário na forma de representação (árvores de decisão e redes neurais artificiais, entre outras) e de uma busca de hipóteses sobre um dado problema (busca gulosa e otimização de funções, entre outras). A existência do viés de aprendizado é essencial para a capacidade de generalização do algoritmo (MITCHELL, 1980).

Dessa forma, um sistema de aprendizado de máquina requer uma escolha criteriosa de qual algoritmo deva ser empregado. Normalmente, o problema da escolha do algoritmo é resolvido por um especialista em aprendizado de máquina. Ele utiliza conhecimentos sobre os dados e sobre os algoritmos disponíveis para escolher manualmente o melhor. Essa escolha é feita segundo alguma métrica de desempenho e as relações que ela estabelece entre algoritmos e conjuntos de dados (BRAZDIL *et al.*, 2009). Uma maneira de evitar a escolha manual é a adoção de um sistema de recomendação automática. Nos últimos anos, esses sistemas de recomendação têm sido gerados por meio de uma técnica denominada *meta-aprendizado*. Segundo Vilalta e Drissi (2002), meta-aprendizado é o estudo do aperfeiçoamento dos algoritmos de aprendizado por meio da experiência. Trata-se da investigação do desenvolvimento de sistemas de recomendação por meio de experiências passadas. Para isso, geralmente é utilizado um algoritmo de aprendizado no nível *meta*, que é um nível acima do aprendizado convencional, chamado de nível *base*.

Tanto no nível base quanto no nível meta, o aprendizado tem um viés. No nível meta, o modelo induzido seleciona o algoritmo do nível base cujo viés é mais adequado para o dado conjunto de dados.

Existem diferentes formas de meta-aprendizado. As utilizadas com mais frequência

e mais relevantes para esta tese são apresentadas nas seções seguintes.

A Seção C.4, em especial, descreve a abordagem mais aplicável ao problema de recomendação de estratégias de amostragem ativa. Dependendo do conjunto de meta-atributos escolhidos e do objetivo pretendido, ela permite caracterizar adequadamente os conjuntos de dados com poucos rótulos ou na ausência deles.

C.1 Generalização em pilha

Na *generalização em pilha* (WOLPERT, 1992), o meta-aprendiz lida com um meta-conjunto de dados que consiste de um conjunto de treinamento transformado por aprendizes no nível base. O resultado dessa transformação são metaexemplos cujos atributos são as predições de cada modelo base. Uma particularidade dessa abordagem é seu viés estático, pois ocorre uma combinação de algoritmos ao invés de uma seleção.

C.2 Caracterização por modelos

A própria estrutura dos modelos do nível base pode ser explorada na construção dos metaexemplos. Uma representante da *caracterização por modelos* é a indução de modelos tipados de ordem maior. Ela gera - de acordo com exemplo dado no trabalho de Bensusan, Giraud-Carrier e Kennedy (2000) - uma árvore de decisão para cada conjunto de dados. As árvores são completamente representadas por estruturas complexas que fazem o papel de metaexemplos que podem ser comparados entre si e são aprendidos por algoritmos especialmente desenvolvidos para esse tipo de tarefa.

C.3 Marcadores de referência

Os *marcadores de referência*¹ (PFAHRINGER; BENSUSAN; GIRAUD-CARRIER, 2000) são um conjunto de diversos algoritmos simples, de baixo custo computacional, cujos desempenhos são usados como referência para a caracterização de conjuntos de dados. A acurácia de cada um dos modelos marcadores de referência utilizados fornece o valor de um meta-atributo. Essa geração de meta-atributos ocorre por meio de processamentos que representem uma simplificação da tarefa base. Ela é desejável em cenários de recomendação de algoritmos cuja finalidade é evitar que todos os algoritmos candidatos, normalmente computacionalmente custosos, sejam experimentados. Logo, não é diretamente aplicável a aprendizado ativo, pois, sem rótulos, não é possível testar os algoritmos marcadores de referência.

¹ [landmarkers]

C.4 Caracterização direta

A caracterização direta consiste na extração de medidas simples e de baixo custo computacional diretamente dos exemplos que compõem um conjunto de dados, ou seja, sem o intermédio de um algoritmo de aprendizado. A primeira caracterização de conjuntos de dados foi feita por [Rendell, Sheshu e Tcheng \(1987\)](#) com o intuito de prever acurácia e tempo de processamento. Ela era baseada no número de exemplos e de atributos. O próximo conjunto de meta-atributos, proposto no projeto STATLOG ([BRAZDIL; HENERY, 1994](#)), era composto de medidas usuais na literatura atual:

- número de exemplos, atributos binários e não binários e classes;
- entropia das classes, informação mútua entre classe e atributos e razão sinal-ruído;
- entropia, curtose, assimetria, correlação e razão entre os desvios padrão entre atributos;
- primeira correlação canônica e variância pelo primeiro discriminante canônico.

Variações desse conjunto são propostas em trabalhos posteriores ([BRAZDIL *et al.*, 2009](#)), como a adoção de histogramas para evitar a perda de informações que ocorre quando se adota a média das medidas nos diferentes atributos base ([KALOUSIS, 2002](#)); ou a binarização de medidas, como o grau de dispersão do atributo alvo em tarefas de regressão ([GOMES *et al.*, 2012](#)). Há também trabalhos direcionados a: otimização ([KANDA *et al.*, 2011](#)), fluxos de dados ([ROSSI *et al.*, 2014](#)), predição de ranqueamentos ([SOUZA; CARVALHO; SOARES, 2010](#)) e detecção de ruído ([GARCIA; CARVALHO; LORENA, 2015](#)).

Finalmente, há trabalhos que visam a recomendação automática de algoritmos não supervisionados ([SOUTO *et al.*, 2008](#); [FERRARI; CASTRO, 2015](#)). Essa tarefa é mais próxima da recomendação de estratégias de amostragem ativa pela ausência de rótulos, diferentemente de muitas medidas dos conjuntos citados anteriormente, que dependem da presença do atributo alvo, ou seja, da existência de rótulos. Assim, dentro do contexto desta tese, a caracterização não supervisionada de conjuntos de dados, apesar de não voltada originalmente ao problema de seleção de estratégias, se mostra compatível.

As medidas adotadas como meta-atributos na presente pesquisa foram apresentadas em maior detalhe na Seção [3.2](#).

Ferramentas

Com relação ao desenvolvimento de código e infraestrutura, mais de 25 mil linhas de código em quase 300 arquivos foram escritas ou adaptadas entre as sucessivas versões dos programas implementados. Uma forma de quantificar a escalabilidade desse sistema é a base de dados resultante, com mais de 40 milhões de consultas registradas. No momento de maior volume experimental, o sistema gerenciador do banco de dados precisou suportar conexões vindas de quase 2000 processos de rotulação ocorrendo simultaneamente em aproximadamente 200 computadores situados em localizações distintas, terminando com um espaço ocupado de 34GiB em disco.

Diversos recursos computacionais foram utilizados para a realização dos experimentos. A maioria foi disponibilizada apenas na fase final. Os seguintes recursos foram empregados, em ordem de tempo de contribuição, do maior para o menor:

- *notebook* de trabalho com 8 núcleos de 2,4GHz;
- *cluster* Biocom com 120 núcleos de 3,5GHz;
- servidor GPU do laboratório Biocom com 24 núcleos de 2,5GHz;
- estação de trabalho com 8 núcleos de 4GHz;
- *nuvem* USP com 24 núcleos de 2,4GHz; e,
- *cluster* Euler, em parte da fase final dos experimentos, com 2080 núcleos de 2,8GHz.

Os seguintes programas foram utilizados (o número de versão é dado entre parênteses), em ordem de relevância nos experimentos, da maior para a menor:

- sistema operacional GNU-Linux Debian (7.0 e 8.2);
- banco de dados MySQL/*connector* (5.5/5.1);
- compilador da linguagem Java/JVM HotSpot(TM) (1.7/24.80);
- compilador da linguagem Scala/ScalaTest (2.11/2.2);

- biblioteca de aprendizado de máquina e interface gráfica Weka (3.7.11);
- biblioteca de álgebra linear *Matrix Toolkits Java* (1.0);
- biblioteca de álgebra linear LAPACK (implementação de referência 3);
- biblioteca de matemática Apache Commons io/math (1.3/3.3);
- biblioteca de parseamento para Scala (1.0);
- interpretador da linguagem R (3.1.1) - ([TEAM](#), 2014);
- biblioteca Python *scikit learn* ([PEDREGOSA et al.](#), 2011);
- SQLite-JDBC (3.7); e,
- compilador GHC da linguagem Haskell (7.8) - ([JONES](#), 2003).

Conjuntos de dados

As tabelas 19 e 20 contêm a lista de conjuntos de dados adotados nos experimentos definitivos. Suas principais características são apresentadas: tamanho médio da reserva de exemplos durante a validação cruzada ($|\mathcal{U}|$), quantidade de classes ($|Y|$), quantidade de atributos, quantidade de atributos nominais, proporção da classe majoritária e proporção da classe minoritária.

Tabela 19 – Características dos conjuntos de dados (1-45).

Conjunto de dados	$ U $	$ Y $	Atributos	Nominais	Majoritária (%)	Minoritária (%)
1-abalone 3class	3342	3	8	1	34,6	31,7
2-artificial charact...	3890	10	7	0	14,4	6,0
3-autoUniv au1 1000	798	2	20	0	74,1	25,9
4-autoUniv au6 cd1 4...	320	8	40	3	27,8	6,3
5-autoUniv au7 300 d...	880	5	12	4	27,7	13,9
6-autoUniv au7 700	560	3	12	4	35,0	30,6
7-autoUniv au7 cpd1...	400	5	12	4	38,4	8,6
8-balance scale	500	3	4	0	46,1	7,8
9-banana	4233	2	2	0	55,2	44,8
10-banknote authentic...	1078	2	4	0	54,7	45,3
11-bupa	273	2	6	0	58,4	41,6
12-car evaluation	1382	4	6	6	70,0	3,8
13-cardiotocography 3...	1692	3	35	0	77,9	8,3
14-climate simulation...	432	2	20	0	91,5	8,5
15-connect. mines vs...	166	2	60	0	53,4	46,6
16-connect. vowel...	792	11	13	0	9,1	9,1
17-ecoli	269	8	7	0	42,6	0,6
18-eeg eye state	11984	2	14	0	55,1	44,9
19-first order theore...	4402	6	51	0	42,2	8,0
20-flare	337	6	12	2	31,4	7,6
21-glass	170	6	9	0	35,7	4,2
22-habermans survival	226	2	3	0	72,1	27,9
23-heart disease clev...	242	5	13	2	54,1	4,3
24-heart disease hung...	234	2	13	0	63,8	36,2
25-heart disease va...	159	5	13	0	28,1	5,0
26-hepatitis	124	2	19	13	79,4	20,6
27-hill valley withou...	970	2	100	0	50,5	49,5
28-horse colic surgic...	240	2	27	14	63,7	36,3
29-indian liver patie...	456	2	10	1	71,2	28,8
30-ionosphere	280	2	33	0	64,3	35,7
31-iris	118	3	4	0	34,0	32,7
32-kr vs kp	2557	2	36	36	52,2	47,8
33-leaf	272	30	15	0	4,7	2,4
34-lymphography	118	4	18	15	54,7	1,4
35-magic	15124	2	10	0	65,2	34,8
36-mammographic mass	514	2	5	0	52,5	47,5
37-mfeat fourier	1595	10	76	0	10,0	9,9
38-molecular splice j...	2404	3	60	60	55,0	22,4
39-monks1	346	2	6	0	50,0	50,0
40-monks3	346	2	6	0	53,0	47,0
41-movement libras	264	15	90	0	7,3	6,4
42-mushroom	6499	2	21	21	51,8	48,2
43-musk	5265	2	167	1	84,5	15,5
44-nursery	10368	5	8	8	33,3	0,0
45-optdigits	4496	10	62	0	10,2	9,9

Tabela 20 – Características dos conjuntos de dados (46-90).

Conjunto de dados	$ \mathcal{U} $	$ \mathcal{Y} $	Atributos	Nominais	Majoritária (%)	Minoritária (%)
46-ozone eighthr	2021	2	72	0	93,7	6,3
47-page blocks	4314	5	10	0	90,5	0,5
48-parkinsons	156	2	22	0	75,4	24,6
49-pendigits	8794	10	16	0	10,4	9,6
50-phoneme	4316	2	5	0	70,8	29,2
51-pima indians diabe...	614	2	8	0	65,1	34,9
52-qsar biodegradatio...	842	2	41	0	66,3	33,7
53-ringnorm	5920	2	20	0	50,5	49,5
54-robot failure lp5	130	5	90	0	27,8	13,0
55-robot nav sensor r...	4142	4	2	0	42,0	6,3
56-saheart	370	2	9	1	65,4	34,6
57-seeds	168	3	7	0	33,3	33,3
58-spambase	3366	2	57	0	60,1	39,9
59-spect heart	178	2	22	22	55,9	44,1
60-statlog australian...	552	2	14	6	55,5	44,5
61-statlog german cre...	800	2	20	13	70,0	30,0
62-statlog heart	216	2	13	0	55,6	44,4
63-statlog image segm...	1669	7	18	0	14,4	14,0
64-statlog vehicle si...	677	4	18	0	25,8	23,5
65-steel plates fault...	1553	2	33	0	65,3	34,7
66-systhetic control	480	6	60	0	16,7	16,7
67-texture	4378	11	40	0	9,1	9,1
68-thyroid ann	2967	3	21	0	92,3	2,5
69-thyroid hypothyroi...	2468	2	25	18	95,5	4,5
70-thyroid newthyroid	172	3	5	0	69,8	14,0
71-thyroid sick euthy...	2468	2	25	18	90,9	9,1
72-tic tac toe	766	2	9	9	65,3	34,7
73-turkiye student	2667	13	32	0	13,6	0,5
74-twonorm	5920	2	20	0	50,0	50,0
75-user knowledge	322	5	5	0	32,0	6,0
76-vertebra column 2c	248	2	6	0	67,7	32,3
77-vertebra column 3c	248	3	6	0	48,4	19,4
78-volcanoes a3	1217	5	3	0	90,0	1,9
79-volcanoes b5	7986	5	3	0	96,2	0,3
80-volcanoes d1	7002	5	3	0	94,4	0,6
81-volcanoes e1	946	5	3	0	91,6	0,8
82-voting	223	2	16	16	67,0	33,0
83-waveform v2	4000	3	40	0	33,8	33,1
84-wdbc	455	2	30	0	62,7	37,3
85-wholesale channel	352	2	7	0	67,7	32,3
86-wilt	3855	2	5	0	94,7	5,3
87-wine	142	3	13	0	39,9	27,0
88-wine quality red	1087	6	11	0	42,5	0,7
89-wine quality white...	3149	5	11	0	45,4	3,3
90-yeast 4class	1015	4	8	0	34,5	12,8

Atividades complementares

Algumas atividades foram realizadas paralelamente ao doutorado. A mais relevante é apresentada a seguir.

Acompanhamento do processo de corte de árvores no campus

Diante do frenesi sobre o *aquecimento global*, pouco nos importamos com assuntos mais prioritários, como o *aquecimento local*, a poluição (do ar, sonora e visual) local, entre outros. Apesar de ser morada da maior concentração de doutores do Brasil, a cidade de São Carlos prossegue impermeabilizando ou desprotegendo o solo; removendo os resquícios de mata nativa culminando numa urbanização baseada em asfalto e cimento. Às vezes, bloquetes, grama e palmeiras são um paliativo de baixo custo de manutenção que visa apenas a permeabilidade. Nos remanescentes das fazendas, árvores não nativas também são removidas.

No campus I da USP, por exemplo, com o objetivo de evitar danos por queda de galhos, está previsto o corte inicial de 111 árvores consideradas de risco - de um total de aproximadamente 600 analisadas segundo laudo do IPT¹. Dependendo do porte da árvore cortada, devem ser plantadas quatro, seis ou oito em compensação. O replantio obrigatório estava previsto até fev/2016, totalizando 610 árvores preferencialmente nativas em qualquer dos campi. Até onde os frequentadores do campus I podem perceber, se houve um replantio significativo, ele deve ter ocorrido apenas no campus II.

Deixo registrado nas páginas seguintes, algumas fotos recentes e uma tentativa de poesia.

¹ O processo de licenciamento ambiental (número 29578/2013) encontra-se na Secretaria de Meio Ambiente da Prefeitura de São Carlos e a lista de árvores a remover também encontra-se na prefeitura da USP.



Foto panorâmica em metades: a seringueira-falsa à esquerda da lanchonete foi cortada ao final de 2015, seguindo o destino de sua semelhante que situava-se à direita e já não aparece na foto. Como resultado, os laboratórios do prédio à esquerda agora precisam de cortinas e seus frequentadores perderam a oportunidade de observar os pássaros que se abrigavam nos galhos da árvore.



Foto panorâmica em metades: fileira remanescente na praça e outras árvores.



A figueira estável e mais desenvolvida foi removida, enquanto que a figueira mais inclinada (à direita da foto) foi mantida. Dois dias antes do depósito deste documento, havia profissionais realizando levantamento topográfico na área. Faz pensar em qual seria o real critério do corte.

Compro Libras Arbóreas

compro libras arbóreas
2500lb de madeira viva

primeiro
a do figo
bem verdadeira

restam na praça
seringueiras-falsas uma fila
e a viúva, abstrata,
matemática,
à dentilhada valsa
bem resistia

F. elastica
ferrea fosse
menos brilharia
a *cementicia* viga

latina ironia
viga: *Arbos*, *concreta*

Resultados detalhados

Os valores de $ALC-\mu_k$ para todas as estratégias e conjuntos de dados são apresentados para cada algoritmo de aprendizado de máquina:

- 5NN - tabelas 21 e 22;
- C4.5w - tabelas 23 e 24;
- NB - tabelas 25 e 26;
- SVM - tabelas 27 e 28;
- RFw - tabelas 29 e 30; e,
- RoF - tabelas 31 e 32.

Tabela 21 – Estratégias comparadas nos conjuntos de dados 1-45 com aprendiz 5NN. Medida: $ALC-\mu_{\kappa}/\sigma_{\kappa}$. *Maior e menor média de cada linha estão em **negrito azul** e **negrito vermelho** respectivamente. Maiores médias isoladas estão sublinhadas. Os melhores valores de desvio padrão estão em verde. Apenas negrito indica segundo melhor valor.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
1-connect. mines vs...	0,54 /0,10	0,53/0,10	0,42 /0,09	0,51/ 0,07	0,50/0,10	0,55 /0,10	0,53/0,10	0,44/0,08	0,47/0,10	0,47/0,10	0,50/ 0,07	0,52/ 0,07	0,49/0,10	0,48/0,11
2-wine quality red	0,14/0,05	0,14/0,04	0,12/0,04	0,12/0,05	0,16 / 0,03	0,14/0,04	0,14/0,04	0,09 /0,04	0,15 /0,04	0,15 /0,04	0,11/ 0,03	0,12/ 0,03	0,13/ 0,03	0,10/0,04
3-qsar biodegradatio...	0,42/0,07	0,43 / 0,06	0,31/0,08	0,37/0,08	0,42/0,09	0,42/0,07	0,43 / 0,06	0,18 /0,15	0,41/0,08	0,41/0,08	0,34/0,08	0,38/ 0,06	0,36/0,07	0,34/0,07
4-waveform v2	0,48/0,03	0,47/ 0,02	0,45 /0,03	0,49/0,05	0,46/0,04	0,48/0,03	0,47/ 0,02	0,46/0,04	0,47/0,03	0,47/0,03	0,47/0,04	0,49/0,04	0,58 /0,03	0,59 / 0,02
5-steel plates fault...	0,78 / 0,02	0,76/ 0,02	0,56 /0,09	0,80 / 0,02	0,67/0,05	0,78 / 0,02	0,76/ 0,02	0,64/0,04	0,67/0,04	0,67/0,04	0,70/0,03	0,71/0,03	0,76/ 0,02	0,64/0,05
6-wine	0,91 /0,04	0,91 / 0,03	0,89 /0,05	0,92 /0,04	0,89 /0,04	0,91 /0,04	0,91 / 0,03	0,90/0,04	0,89 /0,05	0,89 /0,05	0,91 /0,04	0,91 /0,04	0,90/0,05	0,89 /0,04
7-volcanoes d1	0,08 /0,03	0,08 /0,03	0,06/ 0,02	0,07/ 0,02	0,07/ 0,02	0,08 /0,03	0,08 /0,03	0,05 / 0,02	0,07/0,03	0,07/0,03	0,07/0,03	0,07/0,03	0,05 /0,03	0,08 /0,04
8-mushroom	0,91/ 0,01	0,91/ 0,01	0,65 /0,18	0,90/0,03	0,87/0,02	0,95/ 0,01	0,94/ 0,01	0,95/ 0,01	0,87/0,03	0,87/0,03	0,96 / 0,01	0,96 / 0,01	0,95/ 0,01	0,95/ 0,01
9-climate simulation...	0,10/0,09	0,13/0,11	0,15 /0,10	0,09 /0,09	0,12/0,11	0,10/0,09	0,13/0,11	0,13/0,10	0,09 /0,10	0,09 /0,10	0,13/0,10	0,14 /0,10	0,10/ 0,07	0,10/ 0,08
10-eeg eye state	0,20 /0,05	0,21 /0,05	0,11 / 0,03	0,11 / 0,03	0,19/0,05	0,11 /0,04	0,11 / 0,03	0,11 /0,04	0,20 /0,04	0,20 /0,04	0,13/ 0,03	0,13/ 0,03	0,13/0,05	0,15/0,05
11-seeds	0,86 / 0,03	0,86 /0,04	0,83 /0,05	0,86 /0,04	0,86 /0,04	0,86 / 0,03	0,86 /0,04	0,85/0,05	0,84/0,04	0,84/0,04	0,85/0,04	0,86 /0,04	0,84/0,05	0,85/0,04
12-user knowledge	0,55 /0,05	0,54/0,05	0,46 /0,05	0,55 / 0,03	0,51/0,05	0,54/0,05	0,54/0,05	0,51/ 0,03	0,51/0,06	0,51/0,06	0,54/0,04	0,55 /0,05	0,53/0,06	0,55 / 0,03
13-wine quality white...	0,11 / 0,03	0,11 / 0,03	0,09/0,04	0,11 / 0,03	0,10/ 0,03	0,11 / 0,03	0,11 / 0,03	0,07 /0,04	0,10/ 0,03	0,10/ 0,03	0,08/0,04	0,09/0,04	0,07 / 0,03	0,08/ 0,03
14-phoneme	0,40 / 0,03	0,39/0,04	0,33 /0,06	0,40 /0,06	0,37/0,05	0,40 / 0,03	0,39/0,04	0,35/0,05	0,35/0,05	0,35/0,05	0,37/0,05	0,38/0,04	0,38/0,07	0,40 /0,04
15-mfeat fourier	0,59 /0,03	0,58 /0,03	0,53/0,04	0,56/0,03	0,53/ 0,02	0,58 /0,03	0,58 /0,03	0,51/0,04	0,53/0,04	0,53/0,04	0,56/0,04	0,57/0,04	0,49/0,03	0,48 / 0,02
16-ionosphere	0,57 / 0,09	0,56 / 0,09	0,35/0,18	0,55/0,16	0,51/0,11	0,54 / 0,09	0,56 / 0,09	0,45/0,16	0,52/0,12	0,52/0,12	0,46/0,14	0,55/0,15	0,18 /0,19	0,38/0,10
17-heart disease va...	0,07 /0,06	0,07 /0,04	0,07 /0,06	0,04/0,08	0,06/0,07	0,07 /0,05	0,07 /0,04	0,05/0,05	0,07 /0,06	0,07 /0,06	0,05/0,06	0,06/0,06	0,04/0,06	0,03 /0,07
18-thyroid ann	0,18 /0,07	0,18 /0,06	0,13/ 0,04	0,13/0,07	0,19 /0,06	0,18 /0,06	0,18 /0,06	0,05 / 0,05	0,17/0,07	0,17/0,07	0,09/0,06	0,09/0,06	0,16/0,10	0,14/0,09
19-connect. vowel...	0,35/0,03	0,35/0,03	0,35/0,03	0,36/ 0,02	0,32/0,03	0,35/0,03	0,35/0,03	0,38 /0,03	0,31/0,04	0,31/0,04	0,38 /0,03	0,38 / 0,02	0,33/0,03	0,27 / 0,02
20-statlog german cre...	0,14/0,07	0,15/0,07	0,11/0,06	0,17 / 0,05	0,14 / 0,05	0,15/0,07	0,15/0,07	0,10 /0,06	0,13/0,07	0,13/0,07	0,10 / 0,05	0,12/ 0,05	0,11/0,07	0,17 /0,07
21-turkiye student	0,10 /0,02	0,10 / 0,01	0,08 /0,02	0,10 /0,02	0,08 / 0,01	0,09/0,02	0,09/0,02	0,08 /0,02	0,09/0,02	0,09/0,02	0,09/ 0,01	0,09/0,02	0,09/0,02	0,10 /0,02
22-ringnorm	0,17/0,03	0,18/0,02	0,49 /0,12	0,11/0,03	0,16/0,04	0,13/0,03	0,18/0,02	0,33 /0,10	0,18/0,04	0,18/0,04	0,28/0,06	0,20/0,04	0,01 / 0,01	0,01 / 0,01
23-ozone eighthr	0,17/ 0,06	0,18 /0,07	0,15/0,09	0,19 /0,10	0,16/0,08	0,17/ 0,06	0,18 /0,07	0,12 /0,07	0,13/0,09	0,13/0,09	0,17/ 0,06	0,17/0,08	0,15/0,10	0,12 /0,11
24-yeast 4class	0,29/ 0,04	0,30 / 0,04	0,27/0,06	0,30 /0,05	0,28/0,05	0,30 /0,05	0,30 /0,05	0,27/0,08	0,27/ 0,04	0,27/ 0,04	0,26/0,06	0,26/0,05	0,21 /0,06	0,21 /0,06
25-parkinsons	0,67/0,08	0,67/0,08	0,53 /0,10	0,71 / 0,07	0,63/0,11	0,67/0,08	0,67/0,08	0,64/0,10	0,64/0,10	0,64/0,10	0,67/0,08	0,68 / 0,07	0,56/0,11	0,57/0,13
26-flare	0,35/ 0,04	0,34/0,05	0,32 /0,05	0,35/0,05	0,33/0,05	0,34/0,05	0,34/ 0,04	0,35/0,06	0,33/0,05	0,33/0,05	0,34/ 0,04	0,35/0,05	0,37 /0,05	0,37 /0,05
27-autoUniv au6 cd1 4...	-0,01/ 0,03	0,00 /0,04	0,00 /0,04	-0,02 /0,04	0,00 / 0,03	-0,01/ 0,03	0,00 /0,04	-0,01/ 0,03	-0,01/ 0,03	-0,01/ 0,03	-0,01/ 0,03	-0,01/ 0,03	-0,01/ 0,03	0,01 /0,04
28-glass	0,49/0,07	0,50 /0,07	0,42 /0,08	0,48/ 0,06	0,47/0,08	0,49/0,08	0,50 /0,07	0,42 /0,08	0,46/0,08	0,46/0,08	0,46/ 0,06	0,48/0,08	0,45/0,08	0,44/0,07
29-ecoli	0,71 /0,05	0,68 / 0,04	0,68/0,09	0,70/0,07	0,69/0,07	0,71 /0,06	0,68 / 0,04	0,66 /0,07	0,69/0,05	0,69/0,05	0,69/0,07	0,68/0,06	0,68/0,06	0,67/0,06
30-thyroid newthyroid	0,84/0,07	0,84/0,07	0,77/0,07	0,86 / 0,05	0,84/0,07	0,84/0,08	0,84/0,07	0,84/0,07	0,83/0,07	0,83/0,07	0,86 / 0,06	0,85/0,07	0,72/0,10	0,66 /0,11
31-wilt	0,10/0,06	0,12 /0,05	0,08/ 0,03	0,08/0,05	0,12 /0,05	0,08/ 0,04	0,08/ 0,04	0,07 / 0,04	0,11/0,05	0,11/0,05	0,07 /0,05	0,08/0,05	0,13 /0,06	0,10/0,06
32-optdigits	0,86 / 0,01	0,86 / 0,01	0,77 /0,03	0,85/ 0,01	0,81/0,03	0,86 / 0,01	0,86 / 0,01	0,79/0,02	0,80/0,02	0,80/0,02	0,82/0,02	0,82/0,02	0,81/0,02	0,81/0,02
33-pima indians diabe...	0,24 /0,07	0,24 /0,08	0,25 /0,08	0,22/0,07	0,21/0,07	0,23/0,08	0,24 /0,08	0,24 /0,09	0,24 / 0,06	0,24 / 0,06	0,23/0,08	0,22/0,07	0,19/0,07	0,17 /0,07
34-twonorm	0,82 / 0,03	0,81/0,04	0,82 /0,04	0,81/0,07	0,79/0,04	0,82 / 0,03	0,81/0,04	0,78/0,05	0,80/0,04	0,80/0,04	0,77/0,04	0,76/0,05	0,63/0,05	0,62 /0,04
35-volcanoes e1	0,15 / 0,04	0,15 /0,07	0,13 /0,06	0,14/0,06	0,16 /0,06	0,14/ 0,05	0,15 /0,07	0,13 /0,06	0,14/0,07	0,14/0,07	0,14/0,07	0,13 / 0,05	0,13 /0,06	0,14/0,07
36-nursery	0,44/ 0,01	0,44/ 0,01	0,39 /0,03	0,43/0,02	0,41/0,02	0,45 / 0,01	0,45 / 0,01	0,43/0,02	0,42/0,02	0,42/0,02	0,45 / 0,01	0,45 / 0,01	0,43/0,04	0,43/0,03
37-statlog heart	0,50/ 0,09	0,49/ 0,09	0,46/0,10	0,54 /0,11	0,51 /0,10	0,50/ 0,09	0,49/ 0,09	0,47/0,10	0,47/ 0,09	0,47/ 0,09	0,48/0,11	0,48/0,11	0,44 /0,13	0,47/0,10
38-statlog australian...	0,55/ 0,06	0,58 / 0,06	0,48 /0,08	0,58 /0,09	0,56/0,09	0,55/0,07	0,58 / 0,06	0,56 / 0,06	0,54/0,07	0,54/0,07	0,52/0,07	0,54/0,09	0,55/0,09	0,59 / 0,06
39-molecular splice j...	0,37/0,03	0,38 /0,02	0,35/0,03	0,34/0,02	0,35/0,02	0,37/0,03	0,38 /0,02	0,36/0,02	0,35/0,02	0,35/0,02	0,37/0,03	0,37/0,03	0,16 / 0,01	0,17/ 0,01
40-first order theore...	0,16 / 0,02	0,16 /0,03	0,11/0,03	0,15/0,04	0,12/0,03	0,16 / 0,02	0,16 /0,03	0,06 /0,04	0,12 / 0,02	0,12 / 0,02	0,13/0,03	0,14/0,03	0,09/0,03	0,06 / 0,02
41-page blocks	0,47 / 0,05	0,46 /0,05	0,41/0,07	0,46 /0,07	0,45/0,06	0,44/ 0,05	0,46 / 0,05	0,28 /0,12	0,43/0,07	0,43/0,07	0,41/0,07	0,42/0,07	0,35/0,07	0,32/0,07
42-artificial charact...	0,28 /0,02	0,28 /0,02	0,23/0,02	0,25/0,02	0,23/0,02	0,28 /0,02	0,28 /0,02	0,25/0,02	0,24/0,02	0,24/0,02	0,27/0,02	0,28 /0,02	0,22/0,02	0,16 /0,02
43-vertebra column 2c	0,46 /0,08	0,46 / 0,07	0,38/0,13	0,41/0,09	0,44/0,10	0,43/0,09	0,44/ 0,07	0,39/0,09	0,45/0,09	0,45/0,09	0,37 /0,11	0,38/0,08	0,41/ 0,07	0,47 / 0,06
44-pendigits	0,84 / 0,01	0,84 / 0,01	0,74 /0,03	0,84 /0,02	0,79/0,02	0,84 / 0,01	0,84 / 0,01	0,77/0,04	0,78/0,03	0,78/0,03	0,80/0,03	0,81/0,03	0,80/0,03	0,77/0,02
45-mammographic mass	0,47 /0,07	0,47 /0,07	0,38 /0,10	0,44/0,07	0,43/0,09	0,42/ 0,05	0,44/0,07	0,39/0,08	0,47 / 0,06	0,47 / 0,06	0,39/0,12	0,41/0,09	0,43/0,08	0,43/0,09

Tabela 22 – Estratégias comparadas nos conjuntos de dados 46-90 com aprendiz 5NN. Medida: $ALC-\mu_K/\sigma_K$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
46-spambase	0,47 /0,07	0,45/ 0,06	0,41/0,10	0,38/0,11	0,45/0,07	0,47 /0,07	0,45/ 0,06	0,19/0,14	0,45/0,08	0,45/0,08	0,43/ 0,06	0,36/0,08	0,38/ 0,06	0,14 /0,10
47-horse colic surgic...	0,43/ 0,07	0,42/0,09	0,38/0,12	0,50 /0,10	0,42/0,09	0,42/0,08	0,42/0,09	0,37/0,11	0,40/ 0,07	0,40/ 0,07	0,36 /0,11	0,38/0,12	0,38/0,09	0,51 /0,10
48-iris	0,91 / 0,04	0,91 / 0,04	0,85 /0,07	0,92 / 0,04	0,91 / 0,04	0,91 / 0,04	0,91 / 0,04	0,90/ 0,04	0,90/0,05	0,90/0,05	0,90/ 0,04	0,91 / 0,04	0,91 / 0,04	0,91 / 0,04
49-volcanoes b5	0,11 /0,04	0,11 /0,04	0,10/ 0,03	0,13 /0,04	0,11 /0,04	0,11 /0,04	0,11 /0,04	0,10/ 0,03	0,10/0,05	0,10/0,05	0,11 / 0,03	0,11 / 0,03	0,08 /0,04	0,09/0,04
50-statlog image segm...	0,78 / 0,02	0,78 / 0,02	0,69/0,05	0,76/0,03	0,76/ 0,02	0,78 / 0,02	0,78 / 0,02	0,70/0,04	0,73/ 0,02	0,73/ 0,02	0,75/ 0,02	0,77/ 0,02	0,70/0,04	0,66 /0,04
51-voting	0,71/ 0,07	0,71/ 0,07	0,64 /0,11	0,73/0,08	0,71/0,09	0,72/0,08	0,74/0,09	0,73/0,09	0,70/0,09	0,70/0,09	0,74/ 0,07	0,72/0,08	0,75 / 0,07	0,76 /0,08
52-wholesale channel	0,63 / 0,08	0,59/ 0,08	0,51 /0,13	0,63 / 0,07	0,59/0,09	0,63 / 0,08	0,62/ 0,08	0,53/0,10	0,57/0,09	0,57/0,09	0,61/0,09	0,61/0,10	0,54/0,11	0,55/0,11
53-cardiotocography 3...	0,92 / 0,02	0,91/ 0,02	0,71 /0,09	0,92 / 0,01	0,88/0,03	0,92 / 0,02	0,91/ 0,02	0,86/0,03	0,83/0,06	0,83/0,06	0,91/ 0,02	0,91/ 0,02	0,92 / 0,02	0,90/0,03
54-hill valley withou...	0,02/0,05	0,04 /0,05	0,01 /0,05	0,03/0,05	0,02/0,05	0,02/0,05	0,04 /0,05	0,01 / 0,04	0,02/0,05	0,02/0,05	0,03/ 0,04	0,03/ 0,04	0,01 /0,05	0,02/ 0,04
55-autoUniv au7 300 d...	0,08 / 0,03	0,08 / 0,03	0,06 / 0,03	0,08 / 0,03	0,07/ 0,03	0,08 / 0,03	0,08 / 0,03	0,07/ 0,03	0,08 / 0,03	0,08 / 0,03	0,06 / 0,03	0,08 / 0,02	0,08 / 0,03	0,09 / 0,03
56-texture	0,81 / 0,01	0,81 / 0,01	0,70/0,03	0,77/0,02	0,77/0,02	0,81 / 0,01	0,81 / 0,01	0,71/0,03	0,75/0,03	0,75/0,03	0,79/0,02	0,81 /0,02	0,73/0,03	0,68 /0,04
57-autoUniv au7 700	0,03/0,05	0,04 /0,05	0,03/0,05	0,04 / 0,04	0,03/0,05	0,02/ 0,04	0,04 /0,05	0,03/ 0,04	0,03/ 0,04	0,03/ 0,04	0,02/ 0,03	0,02/ 0,04	0,01 /0,05	0,06 / 0,04
58-musk	0,44 / 0,03	0,32/0,05	0,36/0,07	0,35/0,06	0,41/0,04	0,44 / 0,03	0,32/0,05	0,34/0,08	0,40/0,06	0,40/0,06	0,41/0,06	0,36/0,07	0,38/ 0,03	0,17 /0,04
59-kr vs kp	0,43/ 0,03	0,42/ 0,03	0,38 /0,05	0,42/0,04	0,39/0,04	0,43/0,04	0,43/ 0,03	0,42/0,04	0,41/0,04	0,41/0,04	0,44/ 0,02	0,43/ 0,03	0,45 /0,05	0,46 /0,06
60-autoUniv au1 1000	0,06 / 0,05	0,08 / 0,05	0,06 /0,07	0,07/ 0,04	0,06 / 0,05	0,08 /0,07	0,08 /0,06	0,08 /0,06	0,07/0,07	0,07/0,07	0,07/ 0,05	0,07/ 0,05	0,07/ 0,05	0,08 / 0,05
61-monks1	0,23/0,08	0,24/ 0,07	0,26 /0,08	0,20/0,08	0,25/ 0,06	0,22/0,09	0,21/ 0,07	0,26 /0,08	0,20/ 0,07	0,20/ 0,07	0,26 /0,08	0,24/0,08	0,22/0,09	0,16 / 0,07
62-car evaluation	0,37/0,05	0,37/0,04	0,36/0,05	0,40 /0,04	0,38/0,04	0,37/0,05	0,36/0,04	0,40 /0,04	0,36/0,05	0,36/0,05	0,38/0,04	0,39/ 0,03	0,33 /0,05	0,34/ 0,03
63-tic tac toe	0,39/0,05	0,38/0,05	0,34 /0,05	0,39/0,05	0,38/0,06	0,41 / 0,04	0,41 / 0,04	0,39/0,05	0,37/0,06	0,37/0,06	0,40/0,05	0,42 /0,05	0,38/0,06	0,39/0,05
64-thyroid sick euthy...	0,26 /0,10	0,26 / 0,07	0,20/0,10	0,16/0,10	0,25/0,09	0,26 /0,10	0,26 / 0,07	0,13 /0,10	0,22/0,09	0,22/0,09	0,13 / 0,07	0,15/ 0,07	0,24/0,09	0,17/0,12
65-bupa	0,11/ 0,07	0,10/0,09	0,08/0,09	0,07/ 0,07	0,09/0,09	0,10/0,08	0,10/0,09	0,05 /0,09	0,10/0,09	0,10/0,09	0,10/0,08	0,09/ 0,07	0,13 / 0,06	0,15 /0,08
66-abalone 3class	0,29/ 0,03	0,30 /0,04	0,26/0,04	0,28/0,06	0,28/0,04	0,29/ 0,03	0,30 /0,04	0,28/0,05	0,27/0,04	0,27/0,04	0,30 /0,04	0,29/0,06	0,26/0,07	0,22 /0,09
67-banana	0,63 / 0,04	0,60/ 0,04	0,51 /0,07	0,61/ 0,04	0,56/0,05	0,63 / 0,04	0,60/0,04	0,59/ 0,04	0,56 / 0,04	0,56 / 0,04	0,60/ 0,04	0,59/0,05	0,55/0,07	0,51 /0,05
68-lymphography	0,58 / 0,09	0,56/0,11	0,54/ 0,09	0,55/ 0,09	0,53/0,10	0,58 /0,10	0,55/0,10	0,54/0,10	0,52 /0,12	0,52 /0,12	0,54/0,10	0,57/0,11	0,54/0,10	0,56/ 0,09
69-volcanoes a3	0,27 /0,06	0,27 / 0,05	0,26/ 0,05	0,26/0,06	0,25/0,06	0,28 / 0,05	0,27 / 0,05	0,25/ 0,05	0,24 /0,06	0,24 /0,06	0,26/ 0,05	0,26/ 0,05	0,24 /0,07	0,25/0,07
70-leaf	0,44/0,05	0,43/0,05	0,44/0,06	0,45/ 0,04	0,42/0,06	0,44/0,05	0,43/0,05	0,45/0,05	0,42/0,05	0,42/0,05	0,46 /0,05	0,46 / 0,04	0,42/0,05	0,40 /0,05
71-heart disease hung...	0,51/0,12	0,52 /0,10	0,43/0,11	0,52 /0,09	0,49/ 0,08	0,51/0,12	0,52 /0,10	0,42 /0,12	0,49/ 0,08	0,49/ 0,08	0,46/0,14	0,48/0,11	0,48/0,11	0,53 /0,09
72-saheart	0,19 / 0,08	0,13/0,10	0,18/ 0,08	0,19 /0,10	0,14/0,09	0,19 / 0,08	0,13/0,10	0,22 / 0,08	0,14/0,09	0,14/0,09	0,19 / 0,08	0,18/ 0,08	0,11/0,09	0,10 /0,10
73-statlog vehicle si...	0,42/ 0,03	0,43 /0,04	0,36/0,04	0,38/ 0,03	0,38/0,04	0,42/ 0,03	0,43 /0,04	0,39/ 0,03	0,38/0,05	0,38/0,05	0,41/ 0,03	0,41/0,04	0,34/0,05	0,33 /0,04
74-spect heart	0,22/ 0,08	0,17 /0,10	0,19/0,11	0,24 /0,11	0,22/0,11	0,21/ 0,08	0,21/0,10	0,21/ 0,08	0,23 /0,10	0,23 /0,10	0,18/0,09	0,22/ 0,08	0,21/ 0,08	0,22/0,10
75-indian liver patie...	0,08 / 0,06	0,11 / 0,06	0,10/0,07	0,10/0,09	0,11 / 0,06	0,09/ 0,06	0,11 / 0,06	0,11 /0,08	0,11 / 0,06	0,11 / 0,06	0,12 /0,07	0,11 / 0,06	0,10/0,08	0,08 /0,07
76-banknote authentic...	0,93 / 0,02	0,93 / 0,02	0,78 /0,07	0,93 / 0,02	0,89/0,04	0,93 / 0,02	0,93 / 0,02	0,92/0,03	0,88/0,03	0,88/0,03	0,93 / 0,02	0,93 / 0,02	0,87/0,05	0,82/0,08
77-magic	0,37/ 0,04	0,36/0,05	0,32/0,07	0,39 /0,05	0,36/ 0,04	0,37/ 0,04	0,36/0,05	0,33/0,13	0,36/0,05	0,36/0,05	0,35/0,06	0,34/0,06	0,38 /0,06	0,30 /0,11
78-thyroid hypothyroi...	0,33 /0,13	0,33 / 0,11	0,24/ 0,11	0,25/0,16	0,28/0,13	0,33 /0,13	0,33 / 0,11	0,12 / 0,10	0,27/0,14	0,27/0,14	0,15/0,12	0,18/0,15	0,20/0,21	0,22/0,23
79-vertebra column 3c	0,54/0,07	0,54/0,06	0,46 /0,07	0,49/0,06	0,53/0,06	0,51/0,06	0,53 / 0,05	0,48/0,08	0,56 /0,07	0,56 /0,07	0,48/0,07	0,48 / 0,05	0,53 / 0,05	0,57 / 0,05
80-hepatitis	0,39 /0,16	0,34/0,14	0,34/0,13	0,37/0,13	0,38 /0,16	0,38 /0,15	0,34/0,14	0,32/ 0,10	0,30 /0,16	0,30 /0,16	0,34/ 0,12	0,33/ 0,12	0,32/ 0,12	0,34/0,16
81-balance scale	0,57 /0,05	0,56/0,06	0,52/0,08	0,51/ 0,04	0,53/0,06	0,57 /0,05	0,55/0,05	0,52/0,06	0,56/ 0,04	0,56/ 0,04	0,56/0,05	0,54/0,05	0,47 / 0,04	0,47 / 0,04
82-robot failure lp5	0,40 / 0,06	0,40 /0,07	0,35/0,08	0,32/0,08	0,36/0,08	0,40 / 0,06	0,40 /0,07	0,29 /0,08	0,37/0,07	0,37/0,07	0,41 /0,08	0,40 /0,07	0,32/0,08	0,31/0,08
83-autoUniv au7 cpd1...	0,12 /0,05	0,10/ 0,04	0,07 / 0,04	0,11 /0,05	0,10/ 0,04	0,11 /0,05	0,10/ 0,04	0,10/0,05	0,09/0,05	0,09/0,05	0,10/0,05	0,09/ 0,04	0,11/0,05	0,11/0,05
84-wdbc	0,85 /0,04	0,84/ 0,03	0,76/0,07	0,85 /0,04	0,84/0,04	0,84/0,04	0,84/ 0,03	0,79/0,05	0,81/0,05	0,81/0,05	0,83/0,04	0,85 /0,04	0,74/0,06	0,58 /0,10
85-robot nav sensor r...	0,81 / 0,02	0,78/ 0,02	0,68 /0,07	0,80/0,03	0,79/0,03	0,81 / 0,02	0,77/0,03	0,72/0,04	0,77/0,03	0,77/0,03	0,77/0,03	0,78/0,03	0,69/0,07	0,68 /0,07
86-habermans survival	0,12 /0,10	0,08/0,09	0,03 /0,08	0,06/0,09	0,10 /0,09	0,07/0,10	0,08/ 0,06	0,05/0,11	0,09/ 0,07	0,09/ 0,07	0,06/0,11	0,08/0,09	0,07/0,10	0,10 /0,08
87-synthetic control	0,88/0,03	0,87/ 0,02	0,86/0,03	0,86/0,03	0,86/0,04	0,88/0,03	0,87/ 0,02	0,89/ 0,02	0,85/0,03	0,85/0,03	0,90 / 0,02	0,90 / 0,02	0,71/0,04	0,69 /0,04
88-monks3	0,42/0,08	0,44/0,07	0,44/0,07	0,43/0,08	0,40 / 0,06	0,45/ 0,06	0,45/ 0,06	0,50 / 0,04	0,43/ 0,06	0,43/ 0,06	0,47/0,07	0,48/0,07	0,49 /0,07	0,48/ 0,06
89-movement libras	0,60/ 0,04	0,61 /0,05	0,54/0,06	0,59/0,05	0,54/0,06	0,60/ 0,04	0,61 /0,05	0,57/0,05	0,53/0,05	0,53/0,05	0,58/0,05	0,60/ 0,04	0,51/ 0,04	0,50 /0,05
90-heart disease clev...	0,26/0,07	0,27 /0,08	0,24 / 0,05	0,30 /0,06	0,27 /0,06	0,26/0,07	0,27 /0,08	0,24 /0,06	0,27 / 0,05	0,27 / 0,05	0,26/0,06	0,25/0,07	0,26/0,07	0,25/0,06

Tabela 23 – Estratégias comparadas nos conjuntos de dados 1-45 com aprendiz C4.5w. Medida: $ALC-\mu_{\kappa}/\sigma_{\kappa}$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
1-connect. mines vs...	0,33 /0,09	0,31/ 0,06	0,29/0,07	0,29/0,09	0,29/0,09	0,32/0,10	0,33 /0,07	0,28 / 0,06	0,28 / 0,05	0,28 /0,10	0,31/0,08	0,31/0,09	0,32/0,09	0,33 /0,07
2-wine quality red	0,14/0,04	0,15/0,04	0,15/0,04	0,14/ 0,03	0,15/0,04	0,16 / 0,03	0,16 / 0,03	0,16 /0,04	0,15/ 0,03	0,15/ 0,03	0,16 /0,04	0,15/0,04	0,06 /0,05	0,07/0,05
3-qsar biodegradatio...	0,41/0,06	0,42/0,06	0,42/ 0,05	0,39/ 0,05	0,39/ 0,05	0,40/ 0,05	0,43 / 0,05	0,42/0,06	0,40/0,06	0,42/0,06	0,40/ 0,05	0,43 / 0,05	0,25/0,09	0,20 /0,11
4-waveform v2	0,43 / 0,02	0,43 /0,03	0,43 /0,03	0,41 / 0,02	0,41 /0,03	0,43 /0,03	0,43 / 0,02	0,41 /0,03	0,42/ 0,02	0,42/0,03	0,42/0,04	0,43 /0,04	0,43 / 0,02	0,43 /0,03
5-steel plates fault...	0,78 / 0,04	0,76 / 0,04	0,67/0,09	0,68/0,07	0,73/0,08	0,75/0,07	0,75/0,05	0,62/0,08	0,69/0,07	0,70/0,07	0,72/0,07	0,64/0,15	0,18/ 0,04	0,14 /0,07
6-wine	0,75 / 0,05	0,75 /0,06	0,75 /0,06	0,74/0,07	0,73/ 0,05	0,75 / 0,05	0,76 / 0,05	0,73/0,06	0,73/0,07	0,75 / 0,05	0,73/0,07	0,75 / 0,05	0,68 /0,08	0,68 /0,06
7-volcanoes d1	0,09 /0,06	0,13 /0,07	0,02 / 0,03	0,07/0,04	0,08/0,06	0,07/ 0,03	0,06/0,04	0,02 / 0,02	0,08/0,07	0,07/0,06	0,08/0,04	0,04/ 0,03	0,04/ 0,03	0,05/ 0,03
8-mushroom	0,88 / 0,01	0,88 / 0,01	0,82/0,06	0,83/0,09	0,84/0,03	0,88 /0,03	0,89 / 0,01	0,83/0,04	0,82/0,04	0,86/0,04	0,85/0,05	0,86/0,03	0,01 /0,02	0,01 /0,02
9-climate simulation...	0,11 /0,09	0,09/0,07	0,09/0,07	0,06/ 0,04	0,09/0,07	0,08/0,06	0,09/0,07	0,06/0,05	0,10 /0,06	0,10/0,06	0,08/0,05	0,08/0,05	0,02 / 0,02	0,04 / 0,04
10-eeg eye state	0,10/0,04	0,12/0,04	0,10/0,04	0,13 / 0,03	0,11/ 0,03	0,11/0,04	0,11/0,04	0,13 /0,05	0,13 /0,04	0,12/0,04	0,11/0,04	0,12/ 0,03	0,08 /0,05	0,10/0,05
11-seeds	0,79 / 0,05	0,78/0,06	0,77/0,07	0,76/0,07	0,77/0,06	0,78/0,07	0,79 /0,06	0,75/0,07	0,75/ 0,05	0,76/0,06	0,76/0,07	0,77/0,07	0,67 /0,06	0,68/0,07
12-user knowledge	0,66/ 0,04	0,68 /0,05	0,64/0,05	0,63/0,05	0,66/ 0,04	0,63/0,06	0,67 /0,06	0,58/0,08	0,64/ 0,04	0,65/ 0,04	0,46/0,10	0,56/0,09	0,45 /0,07	0,47/0,07
13-wine quality white...	0,10 / 0,02	0,10 / 0,02	0,11 / 0,02	0,10 /0,03	0,09/0,03	0,10 / 0,02	0,10 /0,03	0,09/0,03	0,10 / 0,02	0,10 /0,03	0,10 /0,03	0,07/0,03	0,06 /0,03	0,06 /0,03
14-phoneme	0,32 /0,06	0,32 /0,06	0,28/0,08	0,29/0,06	0,30/0,06	0,34 / 0,04	0,30/0,07	0,25 /0,10	0,26/0,09	0,27/0,09	0,31/0,08	0,30/0,08	0,25 /0,07	0,31/ 0,05
15-mfeat fourier	0,45/ 0,03	0,46 /0,04	0,40/ 0,03	0,40/ 0,03	0,40/0,04	0,45/ 0,03	0,46 /0,04	0,40/ 0,03	0,40/0,04	0,40/0,05	0,43/ 0,02	0,44/ 0,03	0,30 / 0,03	0,30 /0,05
16-ionosphere	0,62/ 0,06	0,63 / 0,06	0,59/0,08	0,62/0,07	0,62/0,08	0,61/0,08	0,63 /0,07	0,59/ 0,06	0,56/0,07	0,58/0,08	0,53/0,12	0,50/0,13	0,21 /0,07	0,23/ 0,05
17-heart disease va...	0,05/0,04	0,06 /0,04	0,05/0,04	0,06 /0,06	0,04/0,04	0,05/0,04	0,06 / 0,03	0,05/0,06	0,04/0,05	0,04/0,05	0,03/0,05	0,03/0,04	0,02 / 0,03	0,02 / 0,03
18-thyroid ann	0,47 /0,12	0,52/0,08	0,61/0,11	0,76 / 0,05	0,61/0,09	0,69/0,13	0,70/0,11	0,68/0,13	0,50/0,09	0,57/0,12	0,71/0,14	0,72/0,11	0,74 / 0,07	0,70/0,15
19-connect. vowel...	0,30/ 0,02	0,31 / 0,02	0,27/0,03	0,28/0,03	0,27/0,03	0,31 / 0,02	0,31 / 0,02	0,27/0,03	0,28/0,03	0,27/0,03	0,29/0,03	0,30/ 0,02	0,17 /0,04	0,18/0,03
20-statlog german cre...	0,12 /0,06	0,10/0,05	0,09/0,06	0,10/ 0,04	0,11/0,06	0,12 /0,07	0,10/0,06	0,08/0,05	0,08/0,06	0,08/0,06	0,10/0,06	0,12 /0,07	0,11/0,05	0,07 / 0,04
21-turkiye student	0,19/ 0,01	0,20 / 0,02	0,18 /0,02	0,19/0,02	0,18 /0,02	0,19/ 0,01	0,20 /0,02	0,19/0,02	0,19/0,02	0,19/0,03	0,19/0,03	0,20 /0,02	0,19/0,02	0,19/0,02
22-ringnorm	0,40 /0,03	0,41 /0,02	0,36/0,04	0,36/0,03	0,37/0,03	0,38/0,03	0,40 /0,03	0,35/0,04	0,37/0,03	0,35/0,03	0,36/0,04	0,38/0,04	0,00 / 0,00	0,00 / 0,00
23-ozone eighthr	0,12 /0,05	0,09/ 0,04	0,10/0,06	0,10/0,06	0,11/ 0,04	0,13 /0,06	0,10/0,06	0,06/0,05	0,07/0,05	0,06/0,05	0,12 /0,06	0,10/0,05	0,06/ 0,04	0,04 / 0,03
24-yeast 4class	0,30 /0,05	0,29/0,05	0,28/ 0,04	0,29/ 0,04	0,29/ 0,04	0,30 / 0,04	0,30 / 0,04	0,29/0,05	0,27/ 0,04	0,28/0,05	0,29/0,05	0,30 / 0,04	0,16 /0,05	0,16 /0,06
25-parkinsons	0,47/ 0,07	0,48 /0,08	0,45/0,09	0,47/0,11	0,46/ 0,07	0,47/0,10	0,50 /0,08	0,47/0,10	0,45/0,10	0,46/0,10	0,47/0,09	0,47/ 0,07	0,39/0,10	0,38 /0,09
26-flare	0,41 /0,04	0,40/ 0,03	0,38 /0,04	0,38 /0,04	0,40/0,04	0,40/0,05	0,40/ 0,03	0,38 /0,04	0,38 /0,04	0,39/0,04	0,39/0,04	0,38 /0,04	0,38 /0,05	0,41 /0,04
27-autoUniv au6 cd1 4...	0,01 / 0,02	0,01 / 0,02	0,01 / 0,02	0,01 / 0,01	0,02 / 0,02	0,01 / 0,02	0,01 / 0,02	0,01 /0,03	0,01 / 0,02	0,01 / 0,02	0,01 / 0,02	0,01 / 0,02	0,01 / 0,02	0,01 / 0,02
28-glass	0,44 /0,06	0,43/ 0,05	0,41/0,07	0,42/0,06	0,41/ 0,05	0,43/0,06	0,44 /0,06	0,40/0,06	0,39 /0,06	0,40/0,06	0,41/0,06	0,45 / 0,05	0,42/ 0,05	0,40/0,06
29-ecoli	0,65/0,05	0,63/ 0,04	0,64/0,06	0,65/0,05	0,62/0,05	0,65/0,06	0,65/0,05	0,64/0,06	0,61 /0,06	0,64/0,05	0,67 / 0,04	0,66 /0,07	0,64/0,05	0,64/0,05
30-thyroid newthyroid	0,71 /0,08	0,72 /0,07	0,68/0,08	0,69/0,10	0,67/0,10	0,70/0,10	0,70/0,08	0,68/0,09	0,68/0,07	0,70/0,08	0,67/0,11	0,71 /0,11	0,01/ 0,01	0,00 / 0,00
31-wilt	0,06/ 0,09	0,05 / 0,07	0,17/0,18	0,32/0,20	0,13/0,12	0,13/0,17	0,14/0,18	0,24/0,20	0,22/0,16	0,22/0,16	0,14/0,18	0,14/0,18	0,49 /0,10	0,46 /0,10
32-optdigits	0,51 /0,02	0,50/0,02	0,44/0,03	0,45/0,02	0,44/0,02	0,51 / 0,01	0,51 /0,02	0,45/0,03	0,43/0,02	0,44/0,02	0,49/ 0,01	0,49/0,02	0,36 /0,03	0,39/0,03
33-pima indians diabe...	0,27/0,07	0,27/0,07	0,28/ 0,05	0,29 / 0,05	0,28/0,06	0,29 /0,08	0,28/0,06	0,29 /0,06	0,28/0,06	0,28/ 0,05	0,28/0,06	0,31 /0,07	0,20/0,08	0,18 /0,07
34-twonorm	0,43/0,02	0,44 / 0,01	0,41/0,02	0,40/0,02	0,42/ 0,01	0,43/0,02	0,44 / 0,01	0,41/0,02	0,42/0,02	0,41/0,02	0,44 /0,02	0,44 / 0,01	0,40/0,03	0,39 /0,04
35-volcanoes e1	0,10/0,06	0,11 /0,08	0,06/ 0,04	0,12 /0,08	0,11 /0,07	0,07/0,05	0,08/0,06	0,04 / 0,03	0,09/0,06	0,06/0,06	0,05/ 0,04	0,07/0,05	0,08/0,07	0,08/0,06
36-nursery	0,60/ 0,03	0,59/ 0,02	0,61/ 0,03	0,58/ 0,03	0,58/ 0,03	0,62 / 0,03	0,61/ 0,03	0,59/0,04	0,59/0,04	0,60/0,04	0,62 / 0,03	0,62 /0,04	0,55 /0,05	0,55 /0,04
37-statlog heart	0,45/0,08	0,46/0,09	0,43/0,08	0,43/0,07	0,44/0,07	0,46/ 0,05	0,48 /0,08	0,46/0,07	0,44/0,08	0,45/0,08	0,46 / 0,06	0,48 / 0,06	0,43/ 0,06	0,41 /0,07
38-statlog australian...	0,60/0,07	0,62 / 0,05	0,59/0,07	0,58/0,07	0,59/0,07	0,62 /0,06	0,64 / 0,05	0,58/0,10	0,57 /0,08	0,58/0,07	0,61/0,07	0,62 /0,06	0,62 /0,06	0,58/ 0,05
39-molecular splice j...	0,45/0,04	0,47 /0,05	0,41/0,05	0,42/0,04	0,42/0,05	0,44/0,04	0,47 /0,05	0,39/0,06	0,39/0,07	0,41/0,06	0,42/0,05	0,43/ 0,03	0,16 / 0,03	0,17/ 0,03
40-first order theore...	0,12 /0,02	0,12 /0,02	0,10/0,02	0,09/0,02	0,09/0,02	0,12 /0,02	0,12 /0,02	0,09/0,02	0,10/0,02	0,09/0,02	0,10/0,02	0,11/0,02	0,05 /0,02	0,06/0,02
41-page blocks	0,42/ 0,06	0,45 / 0,06	0,45/0,07	0,56 / 0,06	0,47/0,08	0,51 / 0,05	0,50 / 0,06	0,49/0,12	0,42/0,09	0,47/0,11	0,48/0,09	0,51 /0,10	0,27 /0,14	0,29/0,14
42-artificial charact...	0,16/ 0,01	0,16/0,02	0,16/0,02	0,14/ 0,01	0,15/0,02	0,17 /0,02	0,16/0,02	0,14/0,02	0,15/0,02	0,14/0,02	0,16/0,02	0,17 /0,02	0,06 /0,02	0,06 / 0,01
43-vertebra column 2c	0,49 / 0,08	0,50 / 0,08	0,48/0,09	0,48/ 0,08	0,46/0,09	0,48/0,09	0,49 /0,11	0,45 / 0,08	0,47/ 0,08	0,46/0,10	0,48/0,09	0,46/0,10	0,48/ 0,07	0,46/0,10
44-pendigits	0,58/0,02	0,59 / 0,02	0,51/ 0,02	0,50/0,03	0,50/ 0,02	0,58/ 0,02	0,59 / 0,02	0,50/0,03	0,49/0,03	0,51/0,03	0,55/ 0,02	0,57/ 0,02	0,41/0,04	0,37 /0,04
45-mammographic mass	0,57 /0,06	0,56/0,05	0,55/0,07	0,56/ 0,04	0,55/0,05	0,57 /0,06	0,56/0,06	0,55/0,06	0,54/ 0,04	0,54/0,05	0,55/0,06	0,57 /0,07	0,51/0,08	0,48 /0,05

Tabela 24 – Estratégias comparadas nos conjuntos de dados 46-90 com aprendiz C4.5w. Medida: $ALC-\mu_{\kappa}/\sigma_{\kappa}$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
46-spambase	0,55/ 0,03	0,56/0,03	0,54/0,04	0,54/ 0,03	0,54/0,04	0,55/0,05	0,57/0,03	0,55/ 0,03	0,55/ 0,03	0,55/0,04	0,56/0,04	0,54/0,04	0,46/0,06	0,20/0,10
47-horse colic surgic...	0,56/ 0,06	0,55/0,08	0,54/0,10	0,54/0,08	0,54/0,08	0,58/0,07	0,57/0,07	0,56/0,08	0,54/0,08	0,54/0,08	0,58/0,07	0,57/0,08	0,58/0,06	0,58/0,06
48-iris	0,86/0,04	0,86/0,05	0,83/0,05	0,85/0,05	0,84/ 0,04	0,86/0,05	0,86/0,05	0,84/0,05	0,83/0,05	0,85/0,05	0,85/0,06	0,85/0,07	0,73/0,05	0,74/0,06
49-volcanoes b5	0,17/ 0,07	0,18/0,08	0,19/0,14	0,25/0,09	0,22/0,08	0,11/ 0,07	0,15/0,11	0,14/0,13	0,16/0,10	0,18/0,13	0,14/0,08	0,18/0,12	0,06/0,06	0,06/0,07
50-statlog image segm...	0,72/ 0,02	0,73/0,02	0,67/ 0,02	0,65/0,04	0,66/0,03	0,71/ 0,02	0,73/0,02	0,67/0,03	0,67/0,03	0,66/0,04	0,69/0,04	0,72/0,03	0,37/0,07	0,38/0,04
51-voting	0,79/ 0,05	0,80/ 0,05	0,79/0,06	0,80/0,06	0,79/0,08	0,81/0,06	0,80/ 0,05	0,78/0,06	0,76/0,06	0,77/ 0,05	0,82/0,06	0,82/0,06	0,72/0,07	0,74/0,08
52-wholesale channel	0,67/0,07	0,67/0,08	0,68/0,07	0,67/0,09	0,66/0,08	0,67/0,09	0,70/0,08	0,69/0,07	0,68/ 0,05	0,69/0,06	0,68/0,08	0,69/0,08	0,68/0,07	0,53/0,09
53-cardiotocography 3...	0,75/ 0,03	0,75/ 0,04	0,68/0,10	0,77/0,07	0,76/0,06	0,66/0,05	0,72/ 0,04	0,58/0,14	0,64/0,09	0,67/0,09	0,63/ 0,04	0,65/0,05	0,05/0,15	0,07/0,17
54-hill valley withou...	0,00/0,00	0,00/0,01	0,00/0,01	0,00/0,02	0,00/0,02	0,00/0,00	0,00/0,01	0,00/0,01	0,00/0,01	0,00/0,01	0,00/0,03	0,00/0,02	0,01/0,03	0,03/0,03
55-autoUniv au7 300 d...	0,06/ 0,02	0,06/ 0,02	0,06/ 0,02	0,06/0,03	0,06/0,03	0,07/0,02	0,06/0,02	0,05/0,03	0,05/0,02	0,05/0,02	0,06/0,02	0,06/ 0,02	0,07/0,03	0,07/0,03
56-texture	0,59/0,02	0,59/0,02	0,50/0,03	0,49/0,04	0,51/0,03	0,58/ 0,02	0,59/0,02	0,50/0,03	0,50/0,03	0,49/0,03	0,57/ 0,02	0,58/ 0,02	0,31/0,03	0,30/0,03
57-autoUniv au7 700	0,08/ 0,04	0,10/0,03	0,06/0,05	0,08/ 0,04	0,09/ 0,04	0,08/0,05	0,10/0,04	0,07/0,06	0,07/ 0,04	0,08/0,05	0,08/0,05	0,06/0,05	0,11/0,04	0,07/ 0,04
58-musk	0,03/ 0,03	0,02/0,06	0,11/0,07	0,12/0,09	0,12/0,08	0,04/0,04	0,02/0,06	0,10/0,07	0,13/0,07	0,14/0,09	0,05/0,07	0,02/0,05	0,04/ 0,02	0,06/ 0,03
59-kr vs kp	0,72/0,04	0,71/0,05	0,68/0,05	0,69/0,05	0,64/0,07	0,72/0,03	0,73/0,04	0,63/0,09	0,63/0,07	0,66/0,07	0,70/ 0,04	0,69/0,06	0,49/0,05	0,50/ 0,04
60-autoUniv au1 1000	0,07/0,05	0,09/ 0,04	0,06/0,05	0,07/ 0,04	0,07/0,05	0,07/0,06	0,09/0,06	0,05/0,06	0,06/0,05	0,06/0,05	0,07/0,06	0,07/0,05	0,14/0,06	0,13/0,07
61-monks1	0,42/0,09	0,39/0,09	0,52/0,10	0,73/0,08	0,52/0,11	0,44/0,12	0,40/ 0,08	0,52/0,12	0,52/0,13	0,53/0,13	0,44/0,13	0,38/ 0,07	0,40/0,09	0,32/0,09
62-car evaluation	0,27/0,10	0,24/0,08	0,29/ 0,08	0,35/0,09	0,30/0,09	0,28/0,11	0,27/0,09	0,34/ 0,08	0,35/0,09	0,35/0,08	0,30/0,08	0,28/0,10	0,25/0,10	0,26/0,09
63-tic tac toe	0,19/0,05	0,17/0,06	0,17/0,05	0,24/0,06	0,20/0,08	0,22/0,07	0,19/0,05	0,21/0,11	0,20/0,07	0,20/0,07	0,16/0,08	0,22/0,08	0,02/0,02	0,04/ 0,04
64-thyroid sick euthy...	0,51/0,14	0,55/0,12	0,59/0,17	0,64/0,11	0,56/0,13	0,56/0,16	0,62/0,11	0,46/0,14	0,41/0,12	0,44/0,13	0,57/0,15	0,63/0,12	0,06/0,04	0,07/ 0,07
65-bupa	0,13/0,06	0,12/0,09	0,13/0,06	0,15/0,06	0,11/0,06	0,13/0,07	0,12/0,06	0,13/0,07	0,12/ 0,05	0,13/0,06	0,14/0,07	0,13/0,07	0,12/0,07	0,14/0,05
66-abalone 3class	0,30/ 0,03	0,31/0,03	0,28/0,04	0,31/0,04	0,30/ 0,03	0,31/0,03	0,32/0,02	0,29/0,06	0,29/0,04	0,29/0,04	0,31/0,04	0,30/0,04	0,20/0,09	0,19/0,09
67-banana	0,07/0,10	0,05/0,07	0,17/0,10	0,18/0,13	0,19/0,13	0,10/0,12	0,08/0,09	0,16/0,11	0,14/0,10	0,14/0,10	0,10/0,11	0,11/0,11	0,00/0,00	0,00/0,00
68-lymphography	0,50/0,10	0,49/0,09	0,46/0,08	0,46/0,09	0,49/0,09	0,49/0,09	0,49/0,08	0,45/0,08	0,45/0,10	0,46/0,09	0,49/0,06	0,47/0,06	0,46/0,07	0,45/0,06
69-volcanoes a3	0,27/0,09	0,30/0,05	0,27/0,10	0,30/0,06	0,26/0,08	0,21/0,09	0,25/0,09	0,21/0,12	0,23/0,12	0,22/0,14	0,20/0,10	0,26/0,08	0,15/0,12	0,18/0,13
70-leaf	0,32/0,03	0,32/0,03	0,29/ 0,03	0,29/0,04	0,28/ 0,03	0,32/0,03	0,32/0,03	0,29/0,04	0,28/ 0,03	0,28/ 0,03	0,32/0,03	0,31/0,04	0,24/ 0,03	0,21/0,02
71-heart disease hung...	0,48/0,10	0,48/0,10	0,45/0,09	0,46/0,09	0,46/0,10	0,48/0,10	0,50/0,08	0,47/0,09	0,45/ 0,07	0,45/ 0,08	0,47/0,09	0,47/0,09	0,38/0,10	0,36/0,13
72-saheart	0,19/0,08	0,17/ 0,07	0,18/ 0,07	0,19/0,08	0,19/ 0,06	0,21/0,09	0,18/ 0,07	0,20/0,08	0,19/ 0,07	0,19/0,09	0,21/0,08	0,19/ 0,07	0,06/0,08	0,06/0,07
73-statlog vehicle si...	0,41/0,03	0,41/0,04	0,36/0,04	0,37/0,04	0,37/ 0,03	0,41/0,04	0,41/0,03	0,37/0,04	0,37/0,04	0,37/0,04	0,38/0,04	0,39/0,05	0,31/0,07	0,30/0,07
74-spect heart	0,23/0,10	0,22/0,09	0,24/0,09	0,24/0,10	0,22/0,09	0,24/0,10	0,22/0,09	0,25/0,10	0,24/ 0,08	0,25/0,08	0,24/0,09	0,23/0,14	0,15/0,10	0,17/0,11
75-indian liver patie...	0,10/0,05	0,13/0,06	0,12/ 0,05	0,13/ 0,05	0,12/0,07	0,14/0,05	0,14/0,07	0,11/ 0,05	0,14/0,05	0,14/0,05	0,12/0,07	0,13/0,07	0,13/0,06	0,10/0,07
76-banknote authentic...	0,77/0,03	0,76/0,02	0,73/0,04	0,74/0,05	0,72/0,05	0,76/0,04	0,76/0,03	0,70/0,06	0,70/0,04	0,69/0,06	0,71/0,06	0,72/0,06	0,50/0,12	0,48/0,13
77-magic	0,35/0,05	0,37/0,05	0,37/0,06	0,34/0,05	0,33/0,05	0,34/0,05	0,37/0,05	0,33/0,06	0,34/ 0,04	0,33/0,05	0,34/0,06	0,35/0,05	0,12/0,05	0,02/0,03
78-thyroid hypothyroi...	0,30/0,16	0,43/0,14	0,55/0,18	0,61/0,11	0,49/0,14	0,32/0,17	0,49/0,17	0,35/0,15	0,32/0,15	0,33/0,17	0,32/0,17	0,49/0,17	0,00/0,00	0,00/0,00
79-vertebra column 3c	0,63/0,04	0,63/0,05	0,62/0,06	0,63/0,05	0,61/0,06	0,60/0,09	0,63/0,06	0,58/0,08	0,61/0,06	0,61/0,06	0,59/0,07	0,61/0,07	0,62/0,07	0,59/0,05
80-hepatitis	0,31/0,12	0,32/0,11	0,30/0,12	0,28/0,13	0,31/0,15	0,28/0,14	0,29/0,12	0,24/0,12	0,26/0,12	0,26/0,12	0,27/0,17	0,28/0,14	0,15/0,07	0,16/ 0,09
81-balance scale	0,46/0,04	0,45/0,04	0,42/0,06	0,44/ 0,03	0,41/0,05	0,46/0,05	0,47/0,03	0,40/0,07	0,40/0,05	0,42/0,05	0,44/0,07	0,45/0,05	0,39/0,04	0,39/0,05
82-robot failure lp5	0,34/0,06	0,34/0,06	0,32/0,06	0,31/0,06	0,31/0,07	0,32/ 0,05	0,34/0,06	0,31/0,07	0,31/0,06	0,32/ 0,05	0,32/0,06	0,33/0,07	0,29/0,05	0,30/0,07
83-autoUniv au7 cpd1...	0,09/0,05	0,07/0,05	0,08/0,06	0,07/0,05	0,08/0,05	0,09/0,05	0,08/0,04	0,08/0,06	0,07/0,04	0,07/0,04	0,08/ 0,03	0,08/0,07	0,08/0,05	0,09/0,04
84-wdbc	0,75/0,05	0,75/ 0,04	0,73/0,06	0,73/ 0,04	0,74/ 0,04	0,77/0,04	0,76/0,04	0,75/ 0,04	0,73/ 0,04	0,74/ 0,04	0,76/0,04	0,76/0,04	0,54/0,05	0,34/0,10
85-robot nav sensor r...	0,89/ 0,02	0,91/0,02	0,89/0,03	0,89/0,03	0,86/0,04	0,90/ 0,02	0,92/0,02	0,86/0,04	0,86/0,03	0,88/0,04	0,89/0,04	0,91/0,03	0,10/0,14	0,10/0,16
86-habermans survival	0,13/0,09	0,12/0,09	0,12/0,08	0,10/0,09	0,13/0,07	0,07/ 0,06	0,12/0,08	0,08/ 0,07	0,12/0,08	0,12/0,09	0,09/ 0,07	0,13/0,11	0,09/ 0,07	0,06/0,07
87-sythetic control	0,58/0,04	0,59/0,03	0,55/ 0,02	0,54/0,04	0,54/0,03	0,59/0,03	0,59/0,03	0,55/0,03	0,56/0,03	0,55/0,04	0,57/ 0,02	0,57/0,03	0,38/0,06	0,38/0,07
88-monks3	0,87/0,04	0,88/0,03	0,87/0,04	0,90/0,02	0,84/0,04	0,88/0,04	0,90/0,02	0,85/0,05	0,83/0,04	0,84/0,04	0,87/0,08	0,88/0,05	0,73/0,09	0,55/0,05
89-movement libras	0,32/0,04	0,33/0,04	0,29/0,04	0,29/ 0,03	0,28/ 0,03	0,32/0,04	0,33/0,04	0,28/ 0,03	0,27/ 0,03	0,28/0,04	0,31/0,04	0,32/ 0,03	0,25/ 0,03	0,24/0,03
90-heart disease clev...	0,22/0,05	0,22/0,05	0,21/ 0,04	0,22/0,05	0,23/0,04	0,23/0,06	0,22/0,06	0,23/0,05	0,23/0,04	0,22/0,05	0,22/ 0,04	0,21/0,05	0,16/0,05	0,19/0,05

Tabela 25 – Estratégias comparadas nos conjuntos de dados 1-45 com aprendiz NB. Medida: $ALC-\mu_K/\sigma_K$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
1-connect. mines vs...	0,40/0,08	0,40/0,08	0,36/ 0,07	0,40/0,10	0,37/0,09	0,36/0,08	0,41 /0,10	0,33 / 0,07	0,36/0,09	0,34/0,09	0,37/0,09	0,37/0,09	0,39/ 0,07	0,41 /0,10
2-wine quality red	0,10/0,07	0,09/0,06	0,11 /0,05	0,11 /0,06	0,12 /0,06	0,09/0,05	0,11 /0,06	0,09/0,06	0,10/0,07	0,10/0,06	0,09/0,05	0,11 /0,05	0,00 / 0,04	0,00 / 0,03
3-qsar biodegradatio...	0,36/0,07	0,38 /0,07	0,36/0,08	0,37/0,07	0,36/ 0,06	0,35/0,07	0,38 /0,07	0,40 / 0,06	0,37/ 0,06	0,37/ 0,06	0,35/0,07	0,36/0,07	0,27/0,08	0,23 /0,10
4-waveform v2	0,59 /0,02	0,59 /0,02	0,55 /0,02	0,59 /0,02	0,57/0,03	0,57/ 0,01	0,57/0,02	0,56/0,02	0,57/0,03	0,58/0,03	0,57/ 0,01	0,57/0,02	0,58/0,02	0,58/0,02
5-steel plates fault...	0,17/0,06	0,15 /0,05	0,19/0,07	0,15 /0,09	0,17/0,08	0,19/0,05	0,19/0,04	0,23 /0,07	0,16/0,04	0,16/0,04	0,20/0,04	0,19/0,05	0,20/ 0,02	0,21 / 0,02
6-wine	0,90/ 0,03	0,91 /0,04	0,85 /0,06	0,91 /0,04	0,88/0,04	0,90/ 0,03	0,90/ 0,03	0,89/0,04	0,89/0,04	0,90/ 0,03	0,90/ 0,03	0,90/ 0,03	0,88/0,04	0,89/ 0,03
7-volcanoes d1	0,03 /0,04	0,04 /0,06	0,01/0,01	0,01/0,02	0,02/0,03	0,01/0,01	0,01/0,01	0,00 /0,01	0,03 /0,05	0,03 /0,05	0,01/0,01	0,01/0,01	0,00 /0,00	0,01/ 0,00
8-mushroom	0,82/ 0,01	0,82/ 0,01	0,72 /0,09	0,90/0,03	0,81/0,03	0,93 /0,02	0,90/0,03	0,92/0,02	0,80/0,03	0,92/0,03	0,93 /0,02	0,93 /0,02	0,92/0,02	0,93 / 0,01
9-climate simulation...	0,07 /0,07	0,05 /0,06	0,03/0,05	0,04/0,05	0,05 /0,07	0,03/0,04	0,02/0,03	0,02/0,03	0,04/0,04	0,04/0,04	0,04/0,04	0,04/0,06	0,01 / 0,01	0,01 / 0,02
10-eeg eye state	0,01 / 0,01	0,03/0,03	0,04/0,04	0,04/0,04	0,02/ 0,02	0,05/0,04	0,05/0,04	0,07/0,05	0,04/0,03	0,04/0,03	0,05/0,04	0,04/0,04	0,09 /0,05	0,08 /0,06
11-seeds	0,80/ 0,05	0,80/ 0,05	0,78/ 0,05	0,80/0,06	0,79/ 0,05	0,82 / 0,05	0,81/ 0,05	0,80/ 0,05	0,77/ 0,05	0,78/ 0,05	0,82 / 0,05	0,82 / 0,05	0,69 /0,08	0,73/0,09
12-user knowledge	0,63/ 0,04	0,64 / 0,04	0,62/ 0,04	0,63/0,05	0,62/0,05	0,62/0,05	0,64 / 0,03	0,57 /0,11	0,62/ 0,04	0,62/ 0,04	0,59/0,07	0,58/0,08	0,60/0,14	0,64 /0,09
13-wine quality white...	0,05/0,04	0,05/0,03	0,06/0,04	0,07 /0,04	0,06/0,04	0,05/0,04	0,05/0,04	0,06/0,04	0,07 /0,05	0,07 /0,05	0,06/0,04	0,04/0,04	0,01 / 0,02	0,01 / 0,01
14-phoneme	0,29 /0,10	0,26/0,09	0,26/0,09	0,26/0,10	0,25/0,10	0,28/ 0,06	0,23/0,10	0,23/0,10	0,25/0,11	0,24/0,11	0,25/0,11	0,21 /0,11	0,28/0,17	0,31 / 0,08
15-mfeat fourier	0,56 / 0,02	0,56 / 0,03	0,48/ 0,03	0,54/ 0,03	0,50/ 0,03	0,50/ 0,03	0,51/ 0,03	0,43/0,04	0,49/ 0,03	0,49/ 0,03	0,50/ 0,03	0,52/ 0,03	0,34 /0,05	0,34 /0,05
16-ionosphere	0,69 /0,07	0,68/0,07	0,60/0,09	0,66/0,07	0,66/0,08	0,69 /0,07	0,68/ 0,06	0,66/ 0,06	0,63/0,07	0,65/0,07	0,69 /0,07	0,69 /0,07	0,57 /0,18	0,60/0,15
17-heart disease va...	0,00/ 0,00	0,00/0,01	0,00/0,01	0,00/0,01	0,00/ 0,00	0,00/ 0,00	0,00/0,01	0,00/0,01	0,00/0,01	0,00/0,01	0,00/0,01	0,00/0,01	0,00/0,00	0,00/ 0,00
18-thyroid ann	0,50/0,13	0,51/0,13	0,56/0,17	0,70/ 0,08	0,61/0,10	0,70/0,14	0,58/0,18	0,65/0,16	0,46 /0,11	0,50/0,11	0,62/0,23	0,60/0,21	0,76 / 0,08	0,73 /0,16
19-connect. vowel...	0,20 / 0,03	0,19/ 0,03	0,20 /0,04	0,20 /0,04	0,18/0,04	0,19/ 0,03	0,19/ 0,03	0,18/0,05	0,19/0,05	0,19/0,05	0,18/0,04	0,18/ 0,03	0,12 /0,04	0,15/0,05
20-statlog german cre...	0,21 /0,07	0,19/ 0,06	0,17/ 0,06	0,17/0,07	0,20 / 0,06	0,18/0,07	0,18/0,08	0,15 / 0,06	0,18/0,07	0,17/0,07	0,18/0,07	0,19/ 0,06	0,19/ 0,06	0,18/0,06
21-turkiye student	0,18/0,02	0,18/0,02	0,17 /0,02	0,17 /0,02	0,17 /0,02	0,18/0,02	0,18/0,02	0,18/ 0,01	0,17 / 0,01	0,17 / 0,01	0,18/0,02	0,18/0,02	0,19 / 0,01	0,20 / 0,01
22-ringnorm	0,41/0,04	0,42/0,04	0,41/0,07	0,48 /0,05	0,43 /0,04	0,37/ 0,03	0,39/ 0,03	0,39/0,05	0,42/0,06	0,41/0,06	0,37/ 0,03	0,38/ 0,03	0,02 / 0,03	0,02 / 0,03
23-ozone eighthr	0,06/0,06	0,06/0,05	0,07/0,06	0,09 /0,07	0,08 /0,05	0,07/0,08	0,04/0,06	0,02 /0,04	0,04/0,05	0,04/0,05	0,08 /0,07	0,05/0,05	0,03/ 0,02	0,02 / 0,02
24-yeast 4class	0,23 /0,05	0,22/0,06	0,21/0,05	0,22/0,05	0,23 /0,05	0,22/0,05	0,22/0,04	0,21/0,05	0,21/0,06	0,20/0,06	0,21/0,05	0,22/0,05	0,03 /0,04	0,05/0,08
25-parkinsons	0,44/ 0,09	0,44/ 0,09	0,45/0,11	0,46/0,10	0,43/0,10	0,49/0,10	0,47/0,10	0,47/0,14	0,43/0,11	0,43/0,11	0,50 /0,11	0,47/ 0,09	0,36/0,15	0,32 /0,11
26-flare	0,44/0,04	0,44/0,04	0,44/0,04	0,44/ 0,03	0,42 /0,04	0,45/0,04	0,44/ 0,03	0,45/0,04	0,43/0,05	0,44/0,05	0,44/0,04	0,45/ 0,03	0,46 / 0,03	0,47 / 0,03
27-autoUniv au6 cd1 4...	0,00 / 0,02	0,00 /0,04	0,00 /0,03	0,00 /0,03	-0,01 /0,03	0,00 /0,03	0,00 /0,03	-0,01 /0,03	0,00 /0,03	0,00 /0,03	0,00 / 0,02	0,01 /0,03	0,00 /0,03	0,00 /0,03
28-glass	0,33/ 0,06	0,34/0,07	0,35/0,08	0,36/ 0,05	0,35/0,07	0,36/0,08	0,36/ 0,06	0,35/0,07	0,32 /0,07	0,33/0,08	0,36/0,07	0,34/0,07	0,38 / 0,06	0,38 / 0,06
29-ecoli	0,65 / 0,03	0,64/ 0,04	0,65 / 0,04	0,66 /0,05	0,63/ 0,04	0,64/0,05	0,64/ 0,04	0,64/0,05	0,61 /0,06	0,63/0,05	0,64/0,06	0,64/0,07	0,62/ 0,04	0,62/0,05
30-thyroid newthyroid	0,83 /0,06	0,83 /0,04	0,76/0,07	0,83 /0,07	0,79/0,06	0,79/0,05	0,79/0,06	0,72/0,11	0,78/0,06	0,78/0,06	0,77/0,06	0,77/0,06	0,01/ 0,01	0,00 / 0,00
31-wilt	0,00 / 0,00	0,00 / 0,00	0,00 /0,01	0,01/0,03	0,00 /0,01	0,00 / 0,00	0,00 / 0,00	0,00 / 0,00	0,01/0,01	0,01/0,01	0,00 / 0,00	0,00 / 0,00	0,19 /0,14	0,28 /0,16
32-optdigits	0,65 / 0,03	0,65 / 0,03	0,61/0,04	0,65 / 0,03	0,62/0,04	0,55/0,04	0,58/ 0,03	0,50/0,04	0,59/0,05	0,59/0,05	0,56/ 0,03	0,57/ 0,02	0,45 / 0,03	0,50/0,04
33-pima indians diabe...	0,24/ 0,08	0,26/0,09	0,28 /0,09	0,28 /0,11	0,24/0,11	0,26/0,09	0,25/ 0,08	0,26/0,09	0,26/0,09	0,26/0,10	0,24/0,10	0,27/0,09	0,07 /0,09	0,07 /0,10
34-twonorm	0,66/0,03	0,63/0,03	0,67 /0,06	0,73 /0,03	0,67 /0,04	0,56/ 0,02	0,55/ 0,02	0,54/0,04	0,65/0,04	0,66/0,04	0,56/ 0,02	0,55/0,03	0,44/0,03	0,43 /0,04
35-volcanoes e1	0,02/0,02	0,04 /0,04	0,01 /0,02	0,01 /0,02	0,03 /0,05	0,01 /0,02	0,02/0,02	0,01 / 0,01	0,03 /0,04	0,03 /0,04	0,01 / 0,01	0,01 / 0,01	0,01 / 0,01	0,01 / 0,01
36-nursery	0,67/ 0,02	0,66/ 0,02	0,55 /0,08	0,62/0,03	0,60/ 0,02	0,69 /0,03	0,69 / 0,02	0,68/ 0,02	0,61/0,03	0,68/ 0,02	0,69 / 0,02	0,70 / 0,02	0,63/0,04	0,63/0,03
37-statlog heart	0,47/0,10	0,46/0,10	0,48/0,10	0,52 /0,13	0,44/0,10	0,44/0,08	0,46/0,09	0,44/0,09	0,46/0,12	0,46/0,12	0,41 /0,08	0,42/ 0,07	0,41 /0,10	0,42/ 0,07
38-statlog australian...	0,63/0,05	0,63/0,05	0,56 /0,10	0,59/0,05	0,60/0,07	0,64/0,06	0,67 / 0,04	0,59/0,10	0,60/0,07	0,61/0,06	0,63/0,13	0,66/ 0,04	0,64/0,12	0,67 /0,05
39-molecular splice j...	0,57/0,04	0,60 / 0,03	0,47/0,08	0,60 /0,05	0,55/0,05	0,60 / 0,03	0,60 /0,04	0,59/0,04	0,51/0,10	0,58/0,04	0,60 / 0,03	0,60 /0,04	0,46/0,04	0,38 / 0,02
40-first order theore...	0,02 /0,02	0,01 / 0,01	0,02 / 0,01	0,02 /0,02	0,02 /0,02	0,01 / 0,01	0,01 / 0,01	0,02 /0,02	0,02 / 0,01	0,02 / 0,01	0,02 /0,02	0,02 /0,02	0,02 /0,02	0,04 /0,02
41-page blocks	0,40/ 0,07	0,40/0,10	0,44/0,14	0,60 / 0,06	0,43/0,08	0,34/0,14	0,32/0,12	0,41/0,13	0,33/0,13	0,34/0,14	0,36/0,13	0,33/0,10	0,12 /0,16	0,12 /0,16
42-artificial charact...	0,01 / 0,01	0,01 / 0,01	0,02 /0,02	0,02 /0,02	0,02 /0,02	0,01 / 0,01	0,01 / 0,01	0,02 /0,02	0,02 /0,02	0,02 /0,02	0,01 / 0,01	0,01 / 0,01	0,01 / 0,01	0,01 / 0,01
43-vertebra column 2c	0,45/ 0,07	0,47 /0,09	0,42/0,08	0,43/0,11	0,40 /0,08	0,46/0,08	0,46/0,08	0,43/0,08	0,40 / 0,07	0,40 / 0,07	0,44/0,08	0,47/0,08	0,48 /0,10	0,47 /0,09
44-pendigits	0,61 /0,04	0,63 / 0,03	0,55/0,04	0,60/0,05	0,60/ 0,03	0,55/0,04	0,59/ 0,03	0,47 /0,06	0,57/0,04	0,58/0,04	0,54/0,04	0,57/0,05	0,54/ 0,03	0,53/ 0,03
45-mammographic mass	0,53 /0,08	0,50/0,07	0,53 /0,07	0,54 /0,08	0,51/0,08	0,52/0,07	0,49/0,07	0,50/0,08	0,51/0,08	0,50/0,09	0,50/ 0,06	0,51/ 0,06	0,37 /0,09	0,39/0,08

Tabela 26 – Estratégias comparadas nos conjuntos de dados 46-90 com aprendiz NB. Medida: $ALC-\mu_{\kappa}/\sigma_{\kappa}$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
46-spambase	0,65/0,04	0,66 /0,04	0,64/0,04	0,68 /0,03	0,63/0,06	0,56/0,04	0,56/ 0,03	0,58/ 0,03	0,65/0,04	0,66 /0,04	0,56/0,04	0,55/0,04	0,36/0,09	0,13 /0,10
47-horse colic surgic...	0,57/0,06	0,58/0,07	0,55 /0,08	0,58/0,07	0,56/0,07	0,60/0,07	0,61 /0,06	0,61 /0,07	0,57/0,08	0,60/0,06	0,60/ 0,05	0,62 /0,05	0,61 /0,06	0,61 /0,06
48-iris	0,90 /0,04	0,89 /0,04	0,89 /0,05	0,89 /0,04	0,88/0,05	0,89 /0,05	0,89 /0,04	0,88/ 0,04	0,88/ 0,04	0,89 /0,04	0,89 /0,04	0,89 /0,04	0,75 /0,06	0,79/0,07
49-volcanoes b5	0,08/0,09	0,08/0,09	0,09/0,12	0,18 /0,14	0,12 /0,11	0,04/0,07	0,07/0,10	0,06/0,10	0,09/0,12	0,09/0,12	0,05/0,08	0,06/0,09	0,01 /0,01	0,03/ 0,06
50-statlog image segm...	0,75/ 0,02	0,76/ 0,02	0,68/0,04	0,73/0,03	0,71/0,03	0,75/ 0,02	0,77 /0,03	0,70/0,05	0,69/0,03	0,73/0,03	0,75/0,03	0,77 /0,02	0,51/0,05	0,48 /0,08
51-voting	0,71/0,08	0,71/0,08	0,72/0,09	0,67 /0,08	0,71/0,08	0,82/ 0,06	0,78/0,07	0,83 /0,06	0,70/0,08	0,78/ 0,06	0,83 /0,06	0,83 /0,06	0,84 /0,06	0,83 /0,06
52-wholesale channel	0,73/0,08	0,73/0,08	0,72/ 0,07	0,75 /0,06	0,73/ 0,07	0,74 /0,08	0,72/0,08	0,73/0,08	0,72/ 0,07	0,73/ 0,07	0,72/0,09	0,72/0,08	0,73/0,09	0,64 /0,16
53-cardiotocography 3...	0,54/ 0,04	0,54/ 0,04	0,48/0,06	0,57 /0,09	0,56 /0,07	0,55/ 0,04	0,56 /0,04	0,54/0,12	0,50/0,08	0,50/0,08	0,54/ 0,04	0,55/0,05	0,11 /0,22	0,11 /0,23
54-hill valley withou...	0,00 /0,00	0,00 /0,00	0,00 /0,00	0,00 /0,01	0,00 /0,00	0,00 /0,01	0,00 /0,01	0,00 /0,03	0,00 /0,00	0,00 /0,00	-0,01 /0,03	-0,01 /0,03	0,00 /0,03	0,00 /0,04
55-autoUniv au7 300 d...	0,08/0,03	0,07 /0,03	0,08/ 0,02	0,07 /0,04	0,07 /0,02	0,09/0,03	0,07 /0,03	0,08/0,03	0,07 /0,03	0,07 /0,03	0,09 /0,04	0,09 /0,04	0,09 /0,03	0,10 /0,03
56-texture	0,66 /0,01	0,66 /0,01	0,59/0,03	0,61/0,03	0,60/0,02	0,63/0,03	0,65/0,02	0,57/0,04	0,58/0,03	0,58/0,04	0,63/0,03	0,66 /0,02	0,41 /0,04	0,43/0,04
57-autoUniv au7 700	0,04 /0,04	0,04 /0,05	0,03/ 0,03	0,02/0,05	0,04 / 0,03	0,03/0,04	0,03/0,04	0,01 /0,05	0,02/0,04	0,02/0,05	0,02/0,04	0,02/0,04	0,04 /0,04	0,02/0,04
58-musk	0,16/0,08	0,10 /0,08	0,26/0,09	0,39/0,07	0,31/0,07	0,43 / 0,06	0,20/0,07	0,39/ 0,06	0,26/ 0,05	0,29/ 0,06	0,42/0,07	0,48 /0,08	0,27/ 0,06	0,29/0,09
59-kr vs kp	0,54/ 0,03	0,53/0,04	0,44/0,06	0,40 /0,07	0,46/0,06	0,62 /0,04	0,63 / 0,03	0,62 /0,05	0,47/0,05	0,60/0,05	0,61/0,04	0,62 /0,04	0,55/0,05	0,52/0,04
60-autoUniv au1 1000	0,00 /0,00	0,00 /0,00	0,00 /0,01	0,00 /0,00	0,01/0,01	0,00 /0,00	0,00 /0,00	0,00 /0,00	0,00 /0,01	0,00 /0,01	0,00 /0,00	0,00 /0,00	0,00 /0,00	0,04 /0,03
61-monks1	0,40/0,07	0,41/0,07	0,40/0,07	0,41/0,08	0,40/0,08	0,40/0,07	0,40/0,07	0,42/0,07	0,43 / 0,06	0,43 / 0,06	0,39/0,10	0,42/0,07	0,17/ 0,06	0,06 /0,08
62-car evaluation	0,41/0,06	0,40/0,05	0,35 /0,10	0,43/0,05	0,40/0,05	0,43/0,07	0,45/0,06	0,50 / 0,04	0,39/0,07	0,46 /0,06	0,44/ 0,04	0,43/0,05	0,46 /0,04	0,44/ 0,04
63-tic tac toe	0,23 / 0,04	0,22/0,06	0,18/0,06	0,15 /0,09	0,19/0,05	0,21/0,08	0,23 / 0,04	0,20/0,05	0,21/0,06	0,23 /0,05	0,23 /0,07	0,24 /0,06	0,17/0,06	0,18/0,06
64-thyroid sick euthy...	0,32/0,16	0,30/0,14	0,40/0,20	0,41/0,16	0,32/0,19	0,49/0,20	0,47/0,13	0,56/0,11	0,16 / 0,10	0,24/0,11	0,53/0,15	0,57/0,15	0,62 / 0,10	0,58 /0,15
65-bupa	0,02/0,03	0,01 / 0,02	0,01 / 0,02	0,03 /0,04	0,01 / 0,02	0,02/0,03	0,03 /0,08	0,04 /0,07	0,02/0,03	0,02/0,03	0,01 / 0,02	0,02/0,04	0,03 /0,05	0,02/0,04
66-abalone 3class	0,31/0,03	0,31/0,03	0,30/0,04	0,31/0,04	0,30/0,03	0,31/ 0,02	0,32 / 0,02	0,27/0,08	0,30/0,05	0,30/0,04	0,31/0,03	0,32 /0,03	0,25/0,05	0,24 /0,05
67-banana	0,00 / 0,01	0,00 / 0,01	0,04/0,05	0,04/0,04	0,04/0,05	0,00 / 0,01	0,01/0,03	0,05 /0,07	0,03/0,04	0,03/0,04	0,01/0,03	0,01/0,02	0,05 /0,06	0,05 /0,06
68-lymphography	0,64 /0,10	0,64 /0,09	0,63/0,07	0,61/0,07	0,60 /0,08	0,63/0,07	0,64 /0,08	0,63/ 0,06	0,61/0,08	0,63/ 0,06	0,62/0,07	0,63/ 0,06	0,62/0,07	0,62/0,07
69-volcanoes a3	0,14/0,08	0,19 /0,11	0,12/0,12	0,15 /0,12	0,15 /0,11	0,07/0,07	0,09/0,08	0,12/0,12	0,12/0,12	0,13/0,13	0,05/0,06	0,06/ 0,05	0,03 /0,06	0,03 / 0,05
70-leaf	0,40 /0,04	0,40 /0,04	0,37/0,04	0,39/ 0,03	0,37/0,04	0,39/0,04	0,40 / 0,03	0,38/0,04	0,37/0,04	0,38/0,04	0,40 / 0,03	0,40 /0,04	0,28/0,04	0,26 /0,05
71-heart disease hung...	0,51 /0,12	0,50 /0,13	0,47/0,11	0,49/0,14	0,48/0,12	0,49/0,12	0,50 /0,10	0,45/0,11	0,47/0,10	0,47/0,10	0,46/0,10	0,47/0,10	0,27 / 0,09	0,30/ 0,09
72-saheart	0,21/0,10	0,19/0,11	0,18/ 0,07	0,23 /0,10	0,18/ 0,07	0,21/ 0,07	0,19/0,08	0,18/0,08	0,19/0,08	0,19/0,08	0,23 /0,08	0,21/ 0,07	0,02 /0,11	0,03/ 0,07
73-statlog vehicle si...	0,19 /0,05	0,19 / 0,04	0,21/0,06	0,22/0,06	0,22/ 0,03	0,22/0,07	0,20/ 0,04	0,23/0,06	0,21/ 0,04	0,21/ 0,04	0,21/ 0,04	0,22/0,05	0,27 /0,05	0,25 /0,06
74-spect heart	0,31/0,13	0,29/0,13	0,29/0,14	0,27 /0,15	0,29/0,12	0,33/ 0,09	0,29/0,11	0,35 /0,11	0,32/0,12	0,33/0,10	0,34 / 0,09	0,34 /0,10	0,30/ 0,08	0,30/ 0,09
75-indian liver patie...	0,11/0,09	0,13 /0,09	0,10/0,07	0,12 /0,07	0,11/0,09	0,12 /0,07	0,10/0,08	0,08/0,07	0,12 /0,07	0,12 /0,07	0,11/0,08	0,12 /0,08	0,03 / 0,04	0,03 / 0,05
76-banknote authentic...	0,72 / 0,04	0,70 / 0,04	0,62 /0,08	0,69/0,05	0,67/0,06	0,70 / 0,04	0,69/ 0,04	0,67/0,07	0,66/0,08	0,66/0,08	0,67/0,09	0,68/0,07	0,65/0,06	0,66/0,06
77-magic	0,24/0,07	0,24/0,08	0,24/0,09	0,29 /0,07	0,24/0,07	0,28/0,10	0,28/0,07	0,29 /0,08	0,28/0,07	0,28/ 0,06	0,28/0,08	0,28/0,09	0,08/ 0,06	0,07 /0,11
78-thyroid hypothyroi...	0,38/0,18	0,50/ 0,16	0,66/ 0,13	0,76 /0,18	0,48/0,20	0,61/0,26	0,60/0,18	0,68 /0,22	0,31 /0,19	0,34/0,19	0,61/0,28	0,54/0,34	0,60/0,31	0,54/0,34
79-vertebra column 3c	0,55/ 0,03	0,55/0,05	0,56/0,06	0,58/0,05	0,55/0,05	0,60 /0,05	0,60 /0,05	0,59/0,04	0,54 /0,05	0,57/0,04	0,61 /0,04	0,60 /0,04	0,60 / 0,03	0,58/0,04
80-hepatitis	0,47/0,17	0,49 /0,17	0,47/ 0,15	0,49 /0,18	0,47/0,18	0,48/ 0,16	0,47/0,19	0,47/0,18	0,45 / 0,16	0,47/ 0,16	0,45 /0,18	0,45 /0,19	0,46/0,19	0,45 /0,18
81-balance scale	0,06 / 0,05	0,07/ 0,05	0,19 /0,07	0,23 /0,10	0,14/0,09	0,08/0,07	0,09/0,07	0,19 /0,09	0,17/0,08	0,17/0,08	0,08/0,07	0,08/0,06	0,09/0,08	0,09/0,07
82-robot failure lp5	0,45/0,06	0,45/0,06	0,42 /0,06	0,45/0,08	0,43/0,06	0,46 /0,06	0,46 /0,05	0,43/ 0,05	0,45/0,07	0,45/0,07	0,47 /0,05	0,46 /0,05	0,45/0,05	0,46 /0,06
83-autoUniv au7 cpd1...	0,07/0,05	0,07/0,05	0,06 / 0,04	0,06 / 0,04	0,08/ 0,04	0,07/ 0,04	0,08/0,05	0,09/0,05	0,07/0,05	0,07/0,05	0,07 / 0,04	0,08/0,05	0,08/ 0,04	0,10 /0,05
84-wdbc	0,84/0,04	0,84/0,03	0,81/0,05	0,86 /0,03	0,84/0,03	0,85 / 0,02	0,85 /0,03	0,84/0,03	0,84/0,04	0,85 /0,03	0,85 / 0,02	0,85 /0,03	0,82/0,05	0,77 /0,13
85-robot nav sensor r...	0,89/0,02	0,90/ 0,01	0,85/0,03	0,92 /0,02	0,88/0,02	0,91/0,02	0,92 / 0,01	0,85/0,05	0,85/0,04	0,86/0,04	0,91/0,02	0,92 / 0,01	0,43 /0,36	0,47/0,38
86-habermans survival	0,02 /0,05	0,05 /0,07	0,03/ 0,04	0,03/0,05	0,03/0,05	0,03/ 0,04	0,04/0,06	0,05 /0,07	0,05 /0,06	0,05 /0,06	0,03/0,06	0,06 /0,09	0,03/0,05	0,02 / 0,04
87-synthetic control	0,82/ 0,02	0,82/ 0,02	0,75 /0,08	0,82/0,03	0,81/0,03	0,84 /0,03	0,85 / 0,02	0,80/0,03	0,80/0,03	0,84 / 0,02	0,81/0,06	0,83/0,05	0,75 /0,06	0,78/0,07
88-monks3	0,72/0,09	0,75/ 0,06	0,73/0,08	0,73/0,13	0,76 /0,08	0,75/0,09	0,75/0,07	0,78 /0,07	0,74/0,08	0,74/0,08	0,75/0,09	0,76 / 0,06	0,70/0,09	0,62 /0,10
89-movement libras	0,25/0,04	0,25/ 0,03	0,27 /0,04	0,28 /0,04	0,26/0,04	0,25/ 0,03	0,25/ 0,02	0,27 /0,04	0,27 / 0,03	0,27 / 0,03	0,24 / 0,03	0,24 /0,04	0,26/0,04	0,26/0,04
90-heart disease clev...	0,29 /0,05	0,27/0,05	0,27/ 0,04	0,28 /0,06	0,28 /0,05	0,27/0,05	0,27/ 0,04	0,26/ 0,04	0,27/0,05	0,26/0,05	0,26/ 0,04	0,26/ 0,04	0,25 /0,05	0,25 /0,05

Tabela 27 – Estratégias comparadas nos conjuntos de dados 1-45 com aprendiz SVM. Medida: $ALC-\mu_{\kappa}/\sigma_{\kappa}$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
1-connect. mines vs...	0,38/0,07	0,37 /0,07	0,38/0,10	0,46 /0,08	0,42 /0,06	0,38/0,10	0,38/ 0,06	0,40/0,09	0,40/0,08	0,40/0,08	0,42 /0,09	0,39/0,15	0,40/ 0,06	0,40/0,07
2-wine quality red	0,13/0,05	0,12/ 0,04	0,14 /0,05	0,13/0,06	0,16 / 0,04	0,13/0,07	0,12/0,05	0,12/0,06	0,14 /0,05	0,14 /0,05	0,14 /0,05	0,13/0,06	0,03/0,06	0,00 /0,05
3-qsar biodegradatio...	0,38/0,08	0,41/0,08	0,40/0,08	0,46 / 0,05	0,38/0,09	0,39/ 0,05	0,41/0,06	0,41/ 0,05	0,42 /0,07	0,42 /0,07	0,37/0,07	0,39/0,07	0,35/ 0,05	0,33 /0,06
4-waveform v2	0,51/ 0,03	0,51/ 0,03	0,54/ 0,03	0,52/0,04	0,51/0,04	0,52/0,04	0,51/ 0,02	0,50/0,04	0,51/ 0,03	0,51/ 0,03	0,49 /0,04	0,50/0,04	0,55 /0,09	0,56 /0,04
5-steel plates fault...	0,72/ 0,03	0,73/ 0,03	0,72/0,05	0,69/0,09	0,71/0,05	0,77 / 0,03	0,76 / 0,03	0,65/0,12	0,70/ 0,03	0,70/ 0,03	0,66/0,13	0,65/0,12	0,63 /0,12	0,69/0,16
6-wine	0,87/ 0,03	0,87/ 0,03	0,85/0,05	0,80/0,06	0,87/ 0,03	0,86/ 0,03	0,87/ 0,03	0,76 /0,08	0,88 / 0,03	0,88 / 0,03	0,78/0,06	0,82/0,05	0,77/0,06	0,79/0,06
7-volcanoes d1	0,01/0,03	0,02 /0,02	0,00 /0,01	0,00 /0,01	0,02 /0,02	0,00 /0,01	0,00 / 0,00	0,00 /0,01	0,01/0,01	0,01/0,01	0,00 / 0,00	0,00 / 0,00	0,00 /0,01	0,00 /0,01
8-mushroom	0,75/ 0,02	0,75/ 0,02	0,81 /0,03	0,81 /0,07	0,80/0,03	0,76/0,03	0,78/ 0,02	0,58 /0,25	0,79/0,03	0,79/0,03	0,64/0,26	0,73/0,17	0,65/0,21	0,75/0,07
9-climate simulation...	0,03 /0,04	0,01/0,03	0,00 /0,01	0,00 /0,01	0,03 /0,04	0,00 /0,01	0,00 / 0,00	0,00 / 0,00	0,01/0,03	0,01/0,03	0,00 / 0,00	0,00 / 0,00	0,00 /0,01	0,00 /0,01
10-eeg eye state	-0,04 / 0,03	-0,01/0,04	0,02/0,04	0,05 /0,04	-0,02/ 0,03	0,03/0,05	0,03/0,04	0,05 /0,04	0,01/0,04	0,01/0,04	0,05 /0,05	0,04/0,04	0,04/0,04	0,04/0,04
11-seeds	0,81 / 0,05	0,81 / 0,05	0,78/0,06	0,75 /0,06	0,80/ 0,05	0,80/ 0,05	0,82 / 0,05	0,78/ 0,05	0,78/0,06	0,78/ 0,05	0,79/ 0,05	0,79/ 0,05	0,75 /0,06	0,76/ 0,05
12-user knowledge	0,54/0,03	0,54/0,03	0,57 /0,04	0,55/0,06	0,54/0,04	0,57 / 0,02	0,56/0,03	0,53/0,07	0,53/0,04	0,53/0,04	0,54/0,03	0,56/ 0,02	0,46/0,09	0,43 /0,06
13-wine quality white...	0,04/ 0,03	0,05/0,04	0,06 /0,04	0,06 /0,04	0,06 /0,04	0,06 /0,04	0,04/0,04	0,03/ 0,03	0,04/0,04	0,04/0,04	0,04/ 0,03	0,05/0,04	0,02 / 0,02	0,02 / 0,03
14-phoneme	0,27/0,10	0,22/0,12	0,24/0,10	0,24/0,10	0,25/0,11	0,29/0,09	0,27/ 0,08	0,15 /0,11	0,21/0,10	0,21/0,10	0,21/0,12	0,23/0,09	0,30 /0,09	0,30 / 0,07
15-mfeat fourier	0,30/ 0,02	0,32/ 0,02	0,34 /0,03	0,35 /0,04	0,32/0,05	0,32/0,03	0,32/ 0,02	0,30/0,09	0,33/0,04	0,33/0,04	0,29/0,06	0,31/0,04	0,19 /0,03	0,21/0,03
16-ionosphere	0,71 / 0,04	0,70/ 0,04	0,70/0,07	0,70/0,06	0,72 /0,05	0,71 / 0,03	0,70/ 0,04	0,61 /0,13	0,71 / 0,04	0,71 /0,05	0,61 /0,17	0,66/0,06	0,62/0,10	0,62/0,06
17-heart disease va...	0,00/ 0,03	0,01/ 0,03	0,01/0,04	0,02 / 0,02	0,02 / 0,03	0,00/0,04	0,01/ 0,03	0,02 /0,04	0,01/ 0,03	0,01/ 0,03	-0,01 /0,04	0,01/ 0,03	0,00/ 0,03	0,01/ 0,03
18-thyroid ann	0,08 / 0,05	0,11/0,06	0,31/ 0,05	0,32/ 0,05	0,18/0,08	0,27/ 0,05	0,25/0,06	0,31/0,06	0,08 /0,07	0,08 /0,07	0,34 / 0,04	0,34 / 0,05	0,22/0,08	0,20/0,06
19-connect. vowel...	0,14/ 0,02	0,13/ 0,02	0,20 /0,03	0,19/ 0,02	0,17/ 0,02	0,13/ 0,02	0,13/ 0,02	0,20 / 0,02	0,18/ 0,02	0,18/ 0,02	0,15/0,03	0,14/ 0,02	0,10 / 0,01	0,10 / 0,02
20-statlog german cre...	0,06/0,06	0,06/0,06	0,07 /0,06	0,07 /0,06	0,07 /0,06	0,05 /0,05	0,05 /0,04	0,06/0,06	0,05 /0,06	0,05 /0,06	0,05 /0,04	0,05 /0,06	0,05 /0,04	0,05 /0,04
21-turkiye student	0,06 / 0,02	0,06 / 0,02	0,07/ 0,02	0,08 / 0,02	0,06 / 0,02	0,06 / 0,02	0,06 / 0,02	0,07/0,03	0,07 / 0,02	0,07 / 0,02	0,06 / 0,02	0,07/0,03	0,08 / 0,02	0,09 /0,03
22-ringnorm	0,85 / 0,02	0,84/ 0,02	0,84/0,03	0,84/0,06	0,85 /0,03	0,57/0,37	0,79/0,03	0,23/0,34	0,85 / 0,02	0,85 / 0,02	0,20 /0,33	0,20 /0,33	0,62/0,14	0,65/0,05
23-ozone eighthr	0,05/0,05	0,05/0,06	0,06/0,06	0,09 /0,07	0,05/0,07	0,04/0,04	0,05/0,07	0,05/0,05	0,01 / 0,02	0,01 / 0,02	0,07 /0,06	0,05/0,06	0,06/0,05	0,04/0,04
24-yeast 4class	0,26 / 0,04	0,24/0,05	0,26 /0,06	0,27 / 0,03	0,26 /0,05	0,25/0,05	0,25/0,05	0,25/0,05	0,23/0,05	0,23/0,05	0,26 /0,06	0,25/0,05	0,17 /0,07	0,17 /0,06
25-parkinsons	0,38/0,11	0,40/ 0,10	0,44/0,11	0,49 /0,11	0,40/0,12	0,43/0,12	0,41/ 0,09	0,42/0,12	0,39/0,11	0,39/0,11	0,47 /0,12	0,47 /0,11	0,35/0,13	0,30 /0,13
26-flare	0,34/ 0,02	0,34/0,03	0,35/0,03	0,37 /0,03	0,36 /0,04	0,35/0,03	0,34/ 0,02	0,32/0,06	0,34/0,04	0,34/0,04	0,35/0,04	0,33/0,04	0,26 /0,09	0,30/0,07
27-autoUniv au6 cd1 4...	-0,01 / 0,02	-0,01 / 0,02	-0,01 /0,03	-0,01 / 0,02	-0,01 / 0,02	-0,01 / 0,02	-0,01 / 0,02	0,00 / 0,02	-0,01 / 0,02	-0,01 / 0,02	-0,01 / 0,02	-0,02 / 0,02	-0,02 / 0,02	-0,02 / 0,02
28-glass	0,32/0,07	0,30/0,07	0,33/0,08	0,34 /0,06	0,33/0,06	0,32/0,06	0,32/0,06	0,35 /0,06	0,31/0,07	0,31/0,07	0,32/ 0,05	0,33/ 0,05	0,28/0,07	0,24 /0,06
29-ecoli	0,65/0,04	0,63/0,05	0,65/0,04	0,65/ 0,03	0,66/0,05	0,66/0,04	0,66/ 0,03	0,65/0,04	0,62 /0,07	0,62 /0,07	0,67 /0,04	0,67 /0,05	0,65/0,04	0,65/0,04
30-thyroid newthyroid	0,78/0,07	0,78/0,06	0,80/0,06	0,78/0,08	0,77/0,07	0,82 / 0,05	0,82 / 0,05	0,77/0,07	0,76 /0,06	0,76 /0,06	0,82 / 0,05	0,82 / 0,05	0,79/ 0,04	0,78/ 0,05
31-wilt	0,04 / 0,05	0,06/ 0,06	0,15/0,11	0,38 /0,10	0,10/0,09	0,23/0,08	0,25/0,07	0,27 /0,08	0,10/0,08	0,10/0,08	0,23/ 0,06	0,26/ 0,06	0,27 /0,08	0,26/ 0,06
32-optdigits	0,55/ 0,02	0,55/ 0,02	0,55/0,03	0,53/0,05	0,55/0,03	0,56 / 0,02	0,56 / 0,01	0,51/0,10	0,52/0,03	0,52/0,03	0,56 / 0,02	0,57 / 0,02	0,42 / 0,02	0,44/0,03
33-pima indians diabe...	0,22/ 0,08	0,24/0,09	0,24/0,09	0,27 / 0,08	0,23/0,10	0,20/ 0,08	0,20/ 0,08	0,21/0,10	0,25 / 0,08	0,25 / 0,08	0,17/0,09	0,22 / 0,08	0,09/ 0,08	0,07 / 0,07
34-twonorm	0,81/ 0,02	0,81/ 0,02	0,80/0,03	0,85 /0,03	0,83 /0,04	0,76/0,05	0,77/0,03	0,48 /0,17	0,83 /0,03	0,83 /0,03	0,58/0,13	0,61/0,11	0,67/0,05	0,66/0,08
35-volcanoes e1	0,03 /0,03	0,04 /0,05	0,01/ 0,01	0,02/0,02	0,03 /0,04	0,01/ 0,01	0,01/0,02	0,00 / 0,01	0,03 /0,03	0,03 /0,03	0,01/ 0,01	0,01/ 0,01	0,01/ 0,01	0,01/0,02
36-nursery	0,44 /0,05	0,44 /0,04	0,55 / 0,03	0,55 / 0,03	0,51/ 0,02	0,46/0,04	0,46/0,04	0,50/0,07	0,53/ 0,03	0,53/ 0,03	0,52/0,07	0,49/0,09	0,51/0,07	0,50/0,11
37-statlog heart	0,53 /0,08	0,50/0,09	0,48/0,09	0,53 /0,08	0,50/0,08	0,53 / 0,07	0,51 / 0,07	0,45 /0,09	0,51/0,08	0,51/0,08	0,50/ 0,07	0,49/ 0,07	0,46/ 0,06	0,47/0,09
38-statlog australian...	0,54/ 0,05	0,55/ 0,05	0,54/0,06	0,57 /0,06	0,54/0,08	0,54/ 0,05	0,56 / 0,04	0,49 /0,11	0,51/0,07	0,51/0,07	0,52/0,06	0,52/0,08	0,52/0,08	0,51/0,06
39-molecular splice j...	0,43/0,04	0,45/ 0,03	0,45/0,04	0,49 / 0,03	0,42/0,04	0,45/0,04	0,46 / 0,03	0,44/0,07	0,41/0,08	0,41/0,08	0,41/0,14	0,46 /0,09	0,16/0,04	0,15 / 0,03
40-first order theore...	0,02 /0,02	0,01 / 0,01	0,01 /0,02	0,02 /0,02	0,02 /0,02	0,01 /0,02	0,01 / 0,01	0,01 /0,02	0,02 /0,03	0,02 /0,03	0,01 / 0,01	0,01 / 0,01	0,02 /0,02	0,04 /0,03
41-page blocks	0,31/ 0,07	0,29/ 0,07	0,42/0,08	0,49 /0,08	0,30/0,09	0,39/0,12	0,41/ 0,07	0,51 /0,10	0,23 /0,10	0,23 /0,10	0,44/ 0,07	0,45/0,08	0,40/0,08	0,31/0,09
42-artificial charact...	0,06/0,02	0,06/0,02	0,09/0,02	0,10 /0,02	0,09/0,02	0,07/0,02	0,06/0,02	0,10 /0,02	0,09/0,02	0,09/0,02	0,07/0,02	0,08/0,02	0,00 / 0,01	0,01/ 0,01
43-vertebra column 2c	0,47/0,08	0,46/0,09	0,48 /0,08	0,50 /0,08	0,44/0,09	0,48 /0,08	0,46/ 0,07	0,41 /0,10	0,44/0,09	0,44/0,09	0,45/0,09	0,46/ 0,07	0,46/0,09	0,43/0,10
44-pendigits	0,63/ 0,02	0,64 / 0,02	0,61/0,03	0,58/0,04	0,60/0,03	0,64 / 0,02	0,64 / 0,02	0,62/0,04	0,59/0,03	0,59/0,03	0,62/ 0,02	0,64 /0,03	0,44/ 0,02	0,41 / 0,02
45-mammographic mass	0,47/0,07	0,47/0,06	0,42/0,07	0,46/0,06	0,46/0,06	0,48/0,07	0,48/0,07	0,40/0,10	0,49 / 0,05	0,49 / 0,05	0,43/0,06	0,41/0,07	0,35 /0,10	0,37/0,10

Tabela 28 – Estratégias comparadas nos conjuntos de dados 46-90 com aprendiz SVM. Medida: $ALC-\mu_{\kappa}/\sigma_{\kappa}$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
46-spambase	0,50/ 0,04	0,49/ 0,05	0,51/ 0,05	0,54 / 0,05	0,50/0,06	0,52 /0,06	0,49/ 0,05	0,48/0,10	0,49/ 0,05	0,49/ 0,05	0,48/0,09	0,47/0,08	0,35/0,09	0,23 /0,14
47-horse colic surgic...	0,33 /0,09	0,38/0,09	0,39/0,07	0,43 /0,08	0,37/0,09	0,37/0,07	0,38/0,08	0,40 /0,08	0,33 /0,09	0,33 /0,09	0,39/0,09	0,35/0,10	0,39/ 0,06	0,40 / 0,06
48-iris	0,83/ 0,03	0,83 / 0,03	0,81/0,06	0,77 /0,06	0,84 /0,04	0,82/0,04	0,83 / 0,03	0,78/0,07	0,84 /0,04	0,84 /0,04	0,81/0,04	0,82/ 0,03	0,80/0,05	0,80/0,04
49-volcanoes b5	0,04 /0,05	0,04 / 0,04	0,10/0,08	0,09/0,08	0,07/0,09	0,05/0,05	0,04 /0,05	0,11 /0,07	0,05/0,07	0,05/0,07	0,10/0,07	0,11 /0,08	0,05/0,05	0,06/ 0,04
50-statlog image segm...	0,60/ 0,02	0,59 / 0,02	0,60/0,03	0,58/0,03	0,61 /0,03	0,61 /0,03	0,59/ 0,02	0,59/0,06	0,58/0,04	0,58/0,04	0,60/0,03	0,60/0,03	0,23 /0,03	0,23 /0,04
51-voting	0,73/ 0,06	0,73/ 0,06	0,71/0,08	0,76 /0,08	0,71/0,09	0,73/ 0,06	0,75 / 0,05	0,47 /0,14	0,68/0,08	0,68/0,08	0,56/0,17	0,55/0,17	0,55/0,16	0,53/0,15
52-wholesale channel	0,63 /0,08	0,62/0,09	0,62/ 0,07	0,63 /0,08	0,63 / 0,07	0,64 / 0,07	0,63 / 0,07	0,59/ 0,07	0,60/0,09	0,60/0,09	0,62/0,08	0,62/ 0,07	0,55 /0,09	0,57/ 0,07
53-cardiotocography 3...	0,75/ 0,03	0,74/0,04	0,73/0,10	0,73/0,07	0,76/0,06	0,78 / 0,03	0,77 / 0,03	0,65 /0,10	0,65 /0,08	0,65 /0,08	0,65 /0,08	0,67/0,07	0,66/0,09	0,65 /0,10
54-hill valley withou...	0,00/ 0,01	0,00/ 0,01	0,00/0,02	0,00/ 0,01	0,00/ 0,01	0,00/0,02	0,00/ 0,01	0,00/ 0,01	0,00/ 0,01	0,00/ 0,01	0,00/ 0,01	0,00/ 0,01	0,00/ 0,01	0,00/ 0,01
55-autoUniv au7 300 d...	0,02 /0,03	0,01 /0,03	0,02 / 0,02	0,02 /0,03	0,02 / 0,02	0,01 /0,03	0,01 /0,03	0,01 / 0,02	0,02 /0,03	0,02 /0,03	0,01 /0,03	0,01 / 0,02	0,02 / 0,02	0,02 / 0,02
56-texture	0,48/ 0,02	0,50/ 0,02	0,51/0,03	0,50/0,03	0,50/0,03	0,50/ 0,02	0,50/ 0,02	0,49/0,04	0,48/0,03	0,48/0,03	0,50/ 0,02	0,52 / 0,02	0,21/0,03	0,19 / 0,02
57-autoUniv au7 700	0,00/ 0,03	-0,01 /0,04	0,00/0,04	0,03 / 0,03	0,01/ 0,03	0,02 /0,04	-0,01 /0,05	0,00/0,04	0,00/ 0,03	0,00/ 0,03	0,00/0,04	0,01/0,04	0,02 /0,04	0,01/ 0,03
58-musk	0,65/ 0,06	0,66/0,07	0,71 /0,15	0,61/0,20	0,68/0,10	0,72 / 0,06	0,65/0,08	0,42/0,15	0,56/0,11	0,56/0,11	0,26/0,16	0,31/0,26	0,19 /0,27	0,33/0,29
59-kr vs kp	0,40/ 0,04	0,40/0,06	0,41/ 0,04	0,45 / 0,04	0,41/0,06	0,37 /0,09	0,41/0,06	0,41/0,09	0,42/0,05	0,42/0,05	0,44 /0,07	0,44 /0,07	0,44 /0,05	0,42/ 0,04
60-autoUniv au1 1000	0,01/ 0,02	0,01/0,03	0,01/0,03	0,03/0,04	0,01/0,04	0,01 / 0,02	0,00 /0,03	0,01/ 0,01	0,02/0,03	0,02/0,03	0,02/0,03	0,01/0,04	0,04 /0,04	0,06 /0,04
61-monks1	0,28/ 0,06	0,29/0,09	0,32/0,08	0,37 /0,09	0,33/0,07	0,36 /0,09	0,32/ 0,06	0,34/0,10	0,33/0,07	0,33/0,07	0,32/0,10	0,29/0,11	0,13 /0,09	0,15/0,09
62-car evaluation	0,30/0,07	0,27/0,07	0,35 /0,06	0,37 /0,07	0,30/0,06	0,33/ 0,05	0,31/ 0,04	0,22/0,13	0,33/0,07	0,33/0,07	0,21 /0,10	0,23/0,11	0,29/0,11	0,29/0,10
63-tic tac toe	0,09/ 0,06	0,04 / 0,04	0,20 /0,07	0,19 /0,07	0,16/0,08	0,17/0,07	0,10/0,07	0,17/0,08	0,18/0,07	0,18/0,07	0,17/ 0,06	0,19 /0,07	0,13/ 0,06	0,10/0,07
64-thyroid sick euthy...	0,10/0,08	0,09/0,08	0,22/0,12	0,32 /0,09	0,14/0,13	0,19/0,14	0,14/0,14	0,29/0,14	0,08 / 0,07	0,08 / 0,07	0,29/0,16	0,31/0,12	0,34 /0,14	0,32 /0,13
65-bupa	0,06/0,07	0,00 /0,10	0,07/0,08	0,06/0,08	0,04/0,07	0,10 /0,08	0,07/0,06	0,04/0,08	0,05/ 0,05	0,05/ 0,05	0,07/0,08	0,08 /0,07	0,07/0,09	0,08 /0,07
66-abalone 3class	0,24/ 0,04	0,24/ 0,04	0,24/ 0,04	0,24/0,05	0,25 / 0,04	0,25 / 0,04	0,24/ 0,03	0,24/0,05	0,25 / 0,04	0,25 / 0,04	0,25 / 0,04	0,25 / 0,04	0,11/0,08	0,07 /0,09
67-banana	0,39/0,06	0,40/0,07	0,42/0,08	0,42/0,10	0,43 /0,06	0,41/0,08	0,41/0,07	0,37/0,09	0,43 /0,06	0,43 /0,06	0,39/0,10	0,37/0,11	0,16/ 0,05	0,15 / 0,04
68-lymphography	0,51/0,07	0,50/0,07	0,53 /0,06	0,53 /0,07	0,49/ 0,05	0,51/ 0,05	0,51/0,06	0,48/0,08	0,47 /0,08	0,47 /0,08	0,48/0,10	0,49/0,10	0,52/0,06	0,53 /0,07
69-volcanoes a3	0,10/ 0,05	0,13/0,08	0,14/0,08	0,16 /0,07	0,11/0,09	0,07 / 0,05	0,09/0,07	0,18 /0,06	0,07 /0,07	0,07 /0,07	0,14/ 0,05	0,11/0,06	0,11/0,07	0,11/ 0,05
70-leaf	0,09/0,02	0,10/0,02	0,12/0,02	0,13 /0,02	0,10/0,02	0,09/0,02	0,10/ 0,01	0,13 /0,02	0,12/0,02	0,12/0,03	0,10/0,02	0,10/0,02	0,03/ 0,01	0,02 / 0,01
71-heart disease hung...	0,48 /0,10	0,47 /0,10	0,44/0,09	0,47 /0,10	0,44/0,13	0,45/0,09	0,44/0,10	0,40/0,11	0,46/ 0,08	0,46/ 0,08	0,41/0,09	0,43/ 0,08	0,37 /0,12	0,37 /0,12
72-saheart	0,14/0,09	0,12/0,08	0,15 /0,09	0,15 /0,08	0,14/0,06	0,14/0,10	0,11/0,09	0,13/0,08	0,13/0,09	0,13/0,09	0,11/0,08	0,08/0,07	0,01 / 0,04	0,01 / 0,03
73-statlog vehicle si...	0,33/ 0,03	0,33/0,04	0,34/ 0,03	0,35 /0,04	0,33/0,04	0,35 /0,04	0,35 / 0,03	0,32 /0,04	0,33/0,04	0,33/0,04	0,35 / 0,03	0,35 / 0,03	0,34/0,04	0,32 / 0,03
74-spect heart	0,20 / 0,07	0,16/0,09	0,17/0,08	0,19 /0,09	0,17/0,09	0,17 / 0,07	0,17/0,10	0,15/0,10	0,19 / 0,07	0,19 / 0,07	0,16/0,10	0,15/ 0,07	0,11 / 0,06	0,13/0,07
75-indian liver patie...	0,01/0,04	0,02/0,03	0,02/0,04	0,02/0,03	0,01/ 0,02	0,00 /0,03	0,02/0,03	0,02/ 0,02	0,04 /0,06	0,04 /0,06	0,01/0,03	0,02/0,03	0,03/0,04	0,01/ 0,02
76-banknote authentic...	0,83/ 0,03	0,83/ 0,02	0,84 / 0,03	0,88 / 0,03	0,84 /0,04	0,81/ 0,03	0,83/ 0,03	0,83/0,04	0,83/0,04	0,83/0,04	0,80 /0,06	0,81/0,04	0,82/0,04	0,82/0,04
77-magic	0,36 / 0,05	0,33/0,06	0,36 /0,09	0,37 /0,07	0,36 /0,07	0,33/ 0,05	0,34/ 0,05	0,34/0,09	0,36 /0,07	0,36 /0,07	0,32/0,09	0,31/0,09	0,20 /0,09	0,22/0,07
78-thyroid hypothyroi...	0,19 /0,13	0,30/0,12	0,42/0,14	0,53 /0,11	0,25/0,15	0,37/0,11	0,38/0,11	0,46/0,15	0,21/0,14	0,21/0,14	0,53 / 0,06	0,47 / 0,07	0,46/0,08	0,48/0,09
79-vertebra column 3c	0,55/0,06	0,57/0,05	0,57/0,05	0,57/0,06	0,56/0,05	0,59 /0,05	0,59 / 0,04	0,52 /0,07	0,53/0,06	0,53/0,06	0,56/0,06	0,56/0,05	0,56/ 0,04	0,55/ 0,04
80-hepatitis	0,38/0,15	0,39 /0,16	0,34/ 0,12	0,39 /0,13	0,38/0,15	0,38/0,13	0,37/0,15	0,32/0,13	0,31 /0,14	0,31 /0,13	0,34/0,13	0,34/ 0,12	0,32/ 0,11	0,31 / 0,12
81-balance scale	0,62/ 0,03	0,62/0,04	0,63/0,04	0,64 /0,04	0,59 /0,05	0,63/ 0,03	0,64 / 0,03	0,62/0,04	0,61/0,05	0,61/0,05	0,63/ 0,03	0,63/ 0,03	0,60/0,05	0,61/0,04
82-robot failure lp5	0,22/0,05	0,23/0,06	0,23/0,05	0,19/ 0,04	0,18 /0,05	0,22/ 0,04	0,23/0,05	0,19/0,05	0,19/0,06	0,19/0,06	0,20/ 0,04	0,22/ 0,04	0,25 /0,05	0,27 /0,05
83-autoUniv au7 cpd1...	0,01/0,03	0,00 /0,03	0,02 /0,03	0,02 /0,03	0,02 /0,03	0,02 /0,03	0,00 /0,03	0,02 / 0,02	0,02 /0,03	0,02 /0,03	0,01/0,04	0,01/0,03	0,01/ 0,02	0,02 /0,03
84-wdbc	0,78/ 0,04	0,79/ 0,04	0,78/0,05	0,80/0,08	0,80/0,05	0,81 / 0,03	0,82 / 0,04	0,69 /0,09	0,78/ 0,04	0,78/ 0,04	0,72/0,05	0,76/0,05	0,75/ 0,04	0,73/0,06
85-robot nav sensor r...	0,68 / 0,03	0,67/ 0,03	0,67/0,04	0,67/0,04	0,66 / 0,03	0,67 / 0,03	0,68 / 0,03	0,61/0,04	0,62/0,05	0,62/0,05	0,67/ 0,03	0,68 / 0,03	0,41 /0,05	0,42/0,06
86-habermans survival	0,07 /0,06	0,06 /0,06	0,03/0,05	0,05/0,04	0,06 /0,06	0,04/0,05	0,06 /0,07	0,02/ 0,03	0,06 /0,07	0,06 /0,07	0,02/ 0,03	0,02/0,04	0,01 / 0,03	0,02/0,04
87-systhetic control	0,76/ 0,02	0,75/ 0,02	0,79 /0,03	0,67/0,09	0,77 / 0,02	0,77 / 0,02	0,76/ 0,02	0,62 /0,14	0,77 / 0,02	0,77 / 0,02	0,74/0,10	0,77 /0,04	0,69/0,04	0,66/0,05
88-monks3	0,60/0,06	0,60/0,06	0,59/0,07	0,67 /0,06	0,61/ 0,05	0,63/ 0,05	0,62/0,06	0,64 /0,06	0,60/0,07	0,60/0,07	0,64 /0,07	0,64 /0,06	0,62/0,06	0,58 /0,10
89-movement libras	0,24/ 0,02	0,25 / 0,02	0,29/0,04	0,30 /0,03	0,28/0,04	0,24/ 0,02	0,25 / 0,02	0,30 /0,03	0,27/0,03	0,27/0,04	0,26/0,03	0,22/ 0,02	0,22 / 0,02	0,21 / 0,02
90-heart disease clev...	0,21/0,06	0,22 /0,06	0,21/0,05	0,22 /0,05	0,21/0,05	0,22 /0,05	0,20/0,06	0,19/ 0,04	0,20/0,06	0,21/0,06	0,21/ 0,03	0,20/ 0,04	0,17 / 0,04	0,17 /0,05

Tabela 29 – Estratégias comparadas nos conjuntos de dados 1-45 com aprendiz RFw. Medida: $ALC-\mu_{\kappa}/\sigma_{\kappa}$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
1-connect. mines vs...	0,43/ 0,06	0,42/ 0,06	0,49/0,10	0,48/0,09	0,48/0,08	0,43/0,07	0,43/0,07	0,51 /0,09	0,46/0,07	0,50 /0,08	0,41 / 0,05	0,41 /0,07	0,42/0,07	0,42/ 0,06
2-wine quality red	0,17/0,04	0,17/0,04	0,17/ 0,03	0,18 / 0,03	0,18 / 0,03	0,18 / 0,03	0,17/ 0,03	0,18 / 0,03	0,17/ 0,03	0,17/ 0,03	0,18 / 0,03	0,18 /0,04	0,04 /0,04	0,04 /0,04
3-qsar biodegradatio...	0,46/ 0,04	0,47/0,05	0,48/0,06	0,49/0,05	0,46/0,08	0,47/ 0,04	0,47/ 0,04	0,51 /0,08	0,48/0,05	0,51 /0,05	0,46/0,05	0,47/ 0,04	0,34/0,06	0,32 /0,07
4-waveform v2	0,54/0,02	0,53 /0,02	0,62 /0,03	0,62 /0,02	0,60/0,03	0,53 /0,02	0,53 /0,02	0,63 /0,03	0,60/0,03	0,61/0,03	0,53 /0,02	0,54/ 0,01	0,54/ 0,01	0,54/ 0,01
5-steel plates fault...	0,30/ 0,03	0,29/ 0,03	0,38/0,07	0,40/0,06	0,34/0,07	0,30/ 0,03	0,31/ 0,03	0,46 /0,08	0,29/0,04	0,42 / 0,03	0,29/0,04	0,30/ 0,03	0,26 /0,04	0,30/0,05
6-wine	0,87 / 0,03	0,87 / 0,03	0,90/ 0,02	0,91 / 0,03	0,89/0,04	0,88/ 0,03	0,89/ 0,03	0,93 / 0,03	0,89/0,04	0,91 / 0,03	0,88/ 0,03	0,89/ 0,03	0,87 / 0,03	0,87 / 0,03
7-volcanoes d1	0,11 /0,04	0,12 /0,04	0,06/0,04	0,08/0,04	0,09/0,05	0,10/0,05	0,10/0,04	0,04 / 0,02	0,09/0,05	0,08/0,04	0,10/0,03	0,11 /0,05	0,07/0,04	0,08/ 0,02
8-mushroom	0,88/ 0,01	0,88/ 0,01	0,90/0,02	0,90/ 0,01	0,87 /0,03	0,91 / 0,01	0,90/ 0,01	0,94 / 0,01	0,87 /0,03	0,88/0,02	0,90/ 0,01	0,91 / 0,01	0,89/0,02	0,90/0,02
9-climate simulation...	0,09 /0,07	0,07/0,05	0,03/0,04	0,03 / 0,03	0,04/0,05	0,06/0,04	0,06/0,05	0,02 / 0,03	0,02 / 0,02	0,03 / 0,03	0,07/0,05	0,07/0,05	0,07/0,04	0,08 /0,05
10-eeg eye state	0,16 / 0,03	0,17 / 0,03	0,16 / 0,03	0,15/0,04	0,15/ 0,03	0,17 / 0,03	0,18 / 0,03	0,16/0,04	0,16/ 0,03	0,16/ 0,03	0,17 / 0,03	0,17 / 0,03	0,12 / 0,03	0,12 / 0,03
11-seeds	0,82/0,05	0,81/0,05	0,83/0,05	0,85 / 0,04	0,82/0,05	0,84/ 0,04	0,84/ 0,04	0,87 /0,05	0,80 /0,05	0,83/0,06	0,85 / 0,04	0,84/0,05	0,85 /0,05	0,84/ 0,04
12-user knowledge	0,66/ 0,03	0,67/0,04	0,69/0,04	0,69/0,04	0,67/0,04	0,69/ 0,03	0,69/ 0,03	0,72 /0,04	0,66/0,04	0,72 /0,04	0,68/ 0,03	0,69/ 0,03	0,62 /0,06	0,65/0,05
13-wine quality white...	0,11/0,03	0,12/0,03	0,13 /0,03	0,13 /0,03	0,13 /0,03	0,12/0,03	0,12/ 0,02	0,13 /0,04	0,13 /0,03	0,13 /0,03	0,11/0,03	0,12/0,03	0,05 / 0,02	0,06/0,03
14-phoneme	0,37/ 0,04	0,35/0,06	0,38 / 0,05	0,41 /0,06	0,38 /0,06	0,34/0,07	0,35/0,07	0,36/0,12	0,33 /0,08	0,35/0,07	0,36/0,06	0,37/ 0,05	0,36/ 0,05	0,37/0,07
15-mfeat fourier	0,46/0,02	0,46/0,02	0,56 /0,02	0,56 /0,03	0,54/0,03	0,46/0,02	0,45/0,02	0,57 /0,03	0,51/0,03	0,55/0,02	0,45/ 0,01	0,45/0,02	0,25 /0,02	0,25 / 0,01
16-ionosphere	0,64/0,06	0,63/0,06	0,72/0,08	0,72/0,06	0,70/0,06	0,66/0,06	0,66/0,06	0,76 /0,07	0,68/0,06	0,73 / 0,05	0,67/0,06	0,67/0,06	0,58/ 0,05	0,56 / 0,05
17-heart disease va...	0,06 /0,04	0,05/ 0,03	0,05/0,04	0,05/0,05	0,05/ 0,03	0,06 /0,05	0,05/0,04	0,05/0,04	0,05/0,04	0,05/0,04	0,05/ 0,03	0,05/0,04	0,03 /0,04	0,03 /0,04
18-thyroid ann	0,35/0,09	0,39/0,08	0,72/0,07	0,72/0,06	0,54/0,11	0,74/0,04	0,71/ 0,03	0,85 / 0,03	0,34 /0,08	0,56/0,06	0,75 /0,04	0,74/0,04	0,73/0,04	0,72/0,04
19-connect. vowel...	0,36/ 0,02	0,36/ 0,02	0,40/0,03	0,41 /0,03	0,39/0,03	0,36/ 0,02	0,36/ 0,02	0,41 / 0,02	0,38/0,03	0,41 / 0,03	0,36/ 0,02	0,36/ 0,02	0,20 / 0,02	0,20 / 0,02
20-statlog german cre...	0,15 /0,05	0,13/ 0,04	0,14/ 0,04	0,13/0,05	0,14/0,05	0,15 /0,05	0,14/0,04	0,14/ 0,04	0,12 /0,05	0,12 /0,05	0,14/0,04	0,14/0,06	0,15 /0,04	0,13/0,04
21-turkiye student	0,13/ 0,01	0,13/ 0,01	0,14/0,02	0,15 / 0,01	0,13/0,02	0,13/ 0,01	0,13/ 0,01	0,14/0,02	0,14/0,02	0,14 /0,02	0,13/ 0,01	0,13/ 0,01	0,10 / 0,01	0,11/ 0,01
22-ringnorm	0,54/ 0,02	0,56/ 0,02	0,61/0,04	0,63 /0,03	0,57/0,04	0,55/0,03	0,55/0,03	0,65 / 0,02	0,57/0,04	0,58/0,04	0,55/ 0,02	0,55/0,03	0,13 / 0,02	0,13 / 0,02
23-ozone eighthr	0,07/ 0,04	0,06/ 0,04	0,07/0,05	0,08/0,05	0,06/0,07	0,08/0,05	0,07/0,05	0,11 /0,05	0,03 /0,04	0,05/0,04	0,11 /0,07	0,10/0,05	0,09/0,05	0,07/ 0,03
24-yeast 4class	0,31/0,04	0,31/0,04	0,33 /0,05	0,34 /0,04	0,33 /0,04	0,31/ 0,03	0,32/0,04	0,33 /0,05	0,31/0,04	0,33 / 0,03	0,31/0,04	0,31/0,04	0,13/0,04	0,10 /0,05
25-parkinsons	0,52/ 0,09	0,53 / 0,09	0,60/0,10	0,60/0,11	0,55/0,10	0,56 / 0,09	0,55/0,10	0,62 /0,12	0,54/0,12	0,61 /0,11	0,57/ 0,09	0,57/ 0,09	0,53/ 0,08	0,50 / 0,09
26-flare	0,42/ 0,03	0,43/ 0,03	0,45/0,05	0,46 /0,04	0,45/0,05	0,43/ 0,03	0,43/ 0,02	0,47 /0,04	0,43/0,04	0,46 / 0,03	0,43/ 0,03	0,44/ 0,03	0,40 / 0,03	0,41/0,04
27-autoUniv au6 cd1 4...	0,01 / 0,01	0,01 / 0,01	0,01 / 0,01	0,02 / 0,01	0,02 /0,02	0,01 /0,02	0,01 /0,02	0,02 / 0,01	0,02 / 0,01	0,02 /0,02	0,01 / 0,01	0,01 /0,02	0,02 / 0,01	0,02 / 0,01
28-glass	0,51/ 0,05	0,51/ 0,05	0,54 /0,06	0,53/0,06	0,52/ 0,05	0,51/ 0,05	0,51/ 0,05	0,54 / 0,05	0,50/0,06	0,55 /0,06	0,52/ 0,05	0,49/ 0,05	0,45/ 0,05	0,44 /0,05
29-ecoli	0,68/ 0,04	0,68/ 0,04	0,68/0,05	0,70/0,05	0,67/0,05	0,70/0,05	0,70/0,05	0,72 / 0,05	0,66 /0,06	0,71 /0,05	0,70/0,05	0,71 /0,05	0,69/0,05	0,68/0,05
30-thyroid newthyroid	0,79 /0,08	0,80/0,07	0,84/0,07	0,85 /0,07	0,81/0,08	0,82/ 0,06	0,82/0,07	0,86 /0,07	0,80/0,08	0,83/0,07	0,82/0,07	0,83/0,07	0,81/ 0,06	0,80/ 0,06
31-wilt	0,17 /0,11	0,20/ 0,09	0,34/0,15	0,46/0,15	0,21/0,12	0,50/0,10	0,50/ 0,09	0,54 /0,12	0,24/0,11	0,29/0,11	0,49/0,11	0,51 / 0,09	0,42/ 0,09	0,43/ 0,09
32-optdigits	0,63 / 0,01	0,63 / 0,01	0,56/0,02	0,58/ 0,01	0,56/0,02	0,63 / 0,01	0,63 / 0,01	0,56/0,02	0,55/0,02	0,55/0,02	0,63 / 0,01	0,63 / 0,01	0,39 /0,02	0,42/0,02
33-pima indians diabe...	0,33/0,07	0,34/0,07	0,35/0,07	0,35/0,06	0,35/0,07	0,35/0,06	0,35/0,07	0,35/0,07	0,36 /0,06	0,37 /0,06	0,34/ 0,05	0,33/ 0,05	0,27 /0,08	0,29/0,11
34-twonorm	0,68/0,02	0,69/ 0,01	0,79/0,02	0,81 / 0,01	0,76/0,03	0,69/0,03	0,69/0,02	0,82 / 0,01	0,75/0,03	0,76/0,03	0,69/0,03	0,69/0,03	0,61/0,02	0,60 /0,03
35-volcanoes e1	0,15/ 0,05	0,16 /0,06	0,14/0,06	0,13/ 0,04	0,15/ 0,05	0,12 /0,06	0,14/0,06	0,12 / 0,05	0,14/0,06	0,14/ 0,05	0,14/ 0,05	0,14/ 0,05	0,15/ 0,05	0,16 /0,07
36-nursery	0,62 /0,02	0,62 / 0,01	0,60/0,02	0,61/0,02	0,58/0,02	0,62 /0,02	0,62 /0,02	0,59/0,03	0,59/0,03	0,60/0,03	0,62 / 0,01	0,62 / 0,01	0,56 /0,03	0,57/0,04
37-statlog heart	0,50 /0,08	0,51/0,08	0,54/0,09	0,55 /0,08	0,53/ 0,07	0,52/ 0,07	0,52/ 0,07	0,55 / 0,07	0,55 /0,08	0,56 /0,08	0,51/0,08	0,52/ 0,07	0,50 / 0,07	0,51/0,11
38-statlog australian...	0,56 /0,06	0,59/ 0,04	0,63 /0,06	0,63 /0,05	0,61/0,07	0,57/0,05	0,59/0,05	0,64 / 0,04	0,60/0,07	0,63 /0,06	0,58/0,06	0,59/0,05	0,57/0,05	0,56 /0,05
39-molecular splice j...	0,32/0,03	0,36/0,03	0,51 /0,06	0,45/0,05	0,40/0,11	0,30/0,04	0,36/0,03	0,53 /0,04	0,39/0,17	0,44/0,13	0,30/0,04	0,33/0,04	0,11 / 0,01	0,13/ 0,01
40-first order theore...	0,14 / 0,02	0,14 / 0,02	0,12/ 0,02	0,14 /0,03	0,12/0,03	0,14 / 0,02	0,14 / 0,02	0,12/0,03	0,12/ 0,02	0,13/ 0,02	0,14 / 0,02	0,14 /0,03	0,06 / 0,02	0,07/ 0,02
41-page blocks	0,47/0,05	0,48/0,07	0,67 /0,05	0,63/0,07	0,52/0,09	0,60/0,06	0,61/ 0,04	0,69 /0,08	0,44 /0,10	0,55/0,05	0,60/0,05	0,61/ 0,04	0,52/0,09	0,51/0,11
42-artificial charact...	0,22/ 0,01	0,21/ 0,01	0,23 /0,02	0,23 /0,02	0,21/0,02	0,21/0,02	0,22/ 0,01	0,23 /0,02	0,21/0,03	0,22/0,03	0,21/0,02	0,21/0,02	0,05 / 0,01	0,07/ 0,01
43-vertebra column 2c	0,49/0,07	0,50/0,08	0,55 /0,07	0,53/ 0,06	0,50/0,09	0,50/0,07	0,51/ 0,06	0,54 / 0,06	0,50/0,08	0,53/0,07	0,52/ 0,06	0,53/ 0,05	0,48 / 0,06	0,48 / 0,06
44-pendigits	0,70 / 0,01	0,70 / 0,01	0,64/0,02	0,64/0,02	0,62/0,02	0,70 / 0,01	0,71 / 0,01	0,65/ 0,01	0,61/0,02	0,63/0,02	0,70 / 0,01	0,70 / 0,01	0,48/0,02	0,42 /0,03
45-mammographic mass	0,55/0,07	0,55/0,07	0,56/0,06	0,57/0,08	0,56/0,06	0,58 /0,06	0,57/ 0,05	0,59 /0,07	0,55/ 0,05	0,57/ 0,05	0,56/0,08	0,57/ 0,05	0,57/0,07	0,54 /0,08

Tabela 30 – Estratégias comparadas nos conjuntos de dados 46-90 com aprendiz RFW. Medida: $ALC-\mu_{\kappa}/\sigma_{\kappa}$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
46-spambase	0,64/0,03	0,64/0,03	0,73/0,02	0,71/0,03	0,69/0,04	0,64/0,04	0,66/0,03	0,75/0,02	0,69/0,04	0,70/0,04	0,66/0,03	0,67/0,02	0,55/0,06	0,43/0,10
47-horse colic surgic...	0,50/0,06	0,52/0,07	0,59/0,07	0,59/0,06	0,57/0,08	0,52/0,07	0,51/0,06	0,60/0,07	0,55/0,08	0,59/0,07	0,50/0,06	0,52/0,07	0,53/0,09	0,54/0,06
48-iris	0,89/0,04	0,88/0,04	0,87/0,04	0,89/0,04	0,87/0,04	0,89/0,04	0,89/0,04	0,89/0,04	0,87/0,04	0,89/0,04	0,89/0,04	0,89/0,04	0,87/0,04	0,87/0,04
49-volcanoes b5	0,19/0,05	0,18/0,05	0,22/0,06	0,23/0,05	0,18/0,05	0,22/0,05	0,24/0,04	0,24/0,05	0,17/0,06	0,18/0,06	0,22/0,04	0,24/0,04	0,15/0,06	0,18/0,05
50-statlog image segm...	0,77/0,01	0,78/0,02	0,80/0,02	0,81/0,02	0,79/0,03	0,79/0,01	0,79/0,01	0,82/0,03	0,77/0,02	0,82/0,02	0,79/0,01	0,79/0,01	0,44/0,04	0,43/0,03
51-voting	0,79/0,05	0,78/0,05	0,82/0,06	0,82/0,07	0,80/0,07	0,81/0,05	0,80/0,05	0,84/0,06	0,79/0,07	0,83/0,06	0,81/0,06	0,80/0,06	0,82/0,05	0,82/0,06
52-wholesale channel	0,72/0,08	0,72/0,07	0,74/0,08	0,75/0,08	0,73/0,07	0,74/0,07	0,74/0,07	0,76/0,08	0,73/0,07	0,75/0,07	0,74/0,06	0,74/0,07	0,72/0,06	0,72/0,06
53-cardiotocography 3...	0,57/0,03	0,57/0,03	0,71/0,07	0,72/0,03	0,63/0,06	0,63/0,03	0,64/0,03	0,80/0,03	0,52/0,07	0,67/0,05	0,63/0,03	0,64/0,03	0,57/0,05	0,58/0,04
54-hill valley withou...	0,02/0,05	0,03/0,05	0,02/0,06	-0,01/0,05	0,02/0,05	0,00/0,05	0,03/0,05	0,00/0,03	0,02/0,05	0,01/0,06	0,00/0,04	0,01/0,05	0,01/0,03	0,00/0,03
55-autoUniv au7 300 d...	0,07/0,02	0,06/0,01	0,08/0,03	0,07/0,02	0,08/0,03	0,07/0,02	0,06/0,01	0,07/0,03	0,08/0,03	0,08/0,03	0,06/0,02	0,07/0,02	0,05/0,01	0,04/0,02
56-texture	0,67/0,01	0,67/0,01	0,61/0,02	0,62/0,02	0,61/0,02	0,67/0,01	0,68/0,01	0,62/0,02	0,58/0,02	0,61/0,02	0,67/0,01	0,67/0,01	0,39/0,03	0,37/0,03
57-autoUniv au7 700	0,06/0,02	0,09/0,03	0,10/0,04	0,08/0,04	0,10/0,03	0,07/0,03	0,08/0,03	0,10/0,04	0,09/0,04	0,10/0,04	0,09/0,03	0,07/0,03	0,09/0,03	0,08/0,02
58-musk	0,19/0,04	0,13/0,03	0,34/0,10	0,34/0,08	0,29/0,07	0,31/0,05	0,16/0,05	0,31/0,14	0,21/0,06	0,22/0,05	0,32/0,06	0,30/0,05	0,14/0,02	0,08/0,04
59-kr vs kp	0,56/0,03	0,55/0,03	0,62/0,05	0,63/0,04	0,57/0,06	0,63/0,03	0,56/0,02	0,67/0,03	0,58/0,04	0,63/0,03	0,56/0,03	0,57/0,03	0,46/0,06	0,46/0,07
60-autoUniv au1 1000	0,05/0,03	0,07/0,03	0,06/0,04	0,06/0,03	0,07/0,05	0,06/0,04	0,06/0,05	0,11/0,06	0,05/0,04	0,07/0,05	0,07/0,04	0,06/0,03	0,11/0,04	0,10/0,04
61-monks1	0,53/0,08	0,52/0,06	0,64/0,06	0,68/0,07	0,59/0,07	0,61/0,06	0,59/0,05	0,67/0,08	0,59/0,08	0,73/0,07	0,61/0,07	0,60/0,06	0,43/0,12	0,48/0,10
62-car evaluation	0,34/0,05	0,34/0,05	0,44/0,05	0,47/0,04	0,39/0,06	0,36/0,04	0,33/0,04	0,50/0,04	0,39/0,08	0,47/0,06	0,36/0,05	0,35/0,05	0,34/0,05	0,35/0,06
63-tic tac toe	0,24/0,03	0,22/0,04	0,28/0,07	0,28/0,05	0,27/0,05	0,24/0,03	0,24/0,05	0,32/0,09	0,27/0,06	0,31/0,06	0,24/0,05	0,25/0,04	0,12/0,08	0,10/0,04
64-thyroid sick euthy...	0,44/0,11	0,44/0,09	0,67/0,09	0,63/0,09	0,43/0,16	0,63/0,07	0,61/0,08	0,73/0,05	0,29/0,12	0,46/0,07	0,64/0,07	0,64/0,10	0,61/0,12	0,58/0,16
65-bupa	0,19/0,06	0,19/0,08	0,21/0,07	0,22/0,08	0,21/0,08	0,18/0,06	0,19/0,06	0,21/0,09	0,24/0,08	0,23/0,07	0,19/0,07	0,21/0,07	0,17/0,05	0,18/0,04
66-abalone 3class	0,32/0,03	0,33/0,03	0,32/0,05	0,32/0,05	0,33/0,03	0,33/0,03	0,33/0,03	0,30/0,10	0,32/0,04	0,33/0,04	0,32/0,03	0,32/0,04	0,22/0,08	0,23/0,09
67-banana	0,50/0,03	0,49/0,03	0,45/0,05	0,37/0,09	0,46/0,04	0,42/0,08	0,44/0,07	0,28/0,12	0,46/0,04	0,49/0,04	0,38/0,09	0,40/0,10	0,18/0,09	0,19/0,08
68-lymphography	0,55/0,07	0,53/0,08	0,51/0,07	0,53/0,07	0,51/0,06	0,54/0,07	0,53/0,07	0,53/0,05	0,50/0,08	0,52/0,06	0,53/0,05	0,54/0,06	0,51/0,07	0,52/0,06
69-volcanoes a3	0,30/0,06	0,31/0,04	0,32/0,06	0,32/0,05	0,29/0,07	0,30/0,06	0,32/0,05	0,31/0,07	0,27/0,09	0,31/0,06	0,31/0,06	0,32/0,06	0,27/0,06	0,28/0,06
70-leaf	0,44/0,04	0,44/0,03	0,50/0,04	0,52/0,04	0,49/0,05	0,44/0,04	0,44/0,03	0,51/0,04	0,49/0,05	0,50/0,05	0,44/0,03	0,44/0,03	0,34/0,03	0,33/0,03
71-heart disease hung...	0,47/0,10	0,50/0,10	0,51/0,10	0,53/0,11	0,51/0,10	0,50/0,10	0,48/0,08	0,53/0,09	0,51/0,08	0,52/0,09	0,47/0,07	0,49/0,08	0,47/0,08	0,46/0,10
72-saheart	0,22/0,07	0,19/0,06	0,21/0,06	0,21/0,07	0,20/0,06	0,21/0,09	0,21/0,05	0,23/0,09	0,21/0,08	0,22/0,08	0,21/0,05	0,21/0,05	0,08/0,07	0,09/0,06
73-statlog vehicle si...	0,45/0,03	0,46/0,03	0,48/0,04	0,47/0,04	0,45/0,04	0,46/0,03	0,47/0,03	0,49/0,03	0,45/0,04	0,48/0,03	0,45/0,03	0,46/0,03	0,36/0,03	0,35/0,04
74-spect heart	0,24/0,07	0,22/0,09	0,28/0,08	0,28/0,08	0,26/0,11	0,23/0,07	0,23/0,08	0,29/0,10	0,28/0,08	0,30/0,08	0,23/0,07	0,22/0,07	0,22/0,07	0,22/0,07
75-indian liver patie...	0,12/0,05	0,14/0,06	0,15/0,07	0,15/0,06	0,17/0,08	0,14/0,07	0,12/0,06	0,14/0,08	0,18/0,05	0,17/0,08	0,16/0,07	0,14/0,06	0,13/0,07	0,14/0,07
76-banknote authentic...	0,82/0,03	0,83/0,03	0,82/0,07	0,88/0,04	0,80/0,06	0,86/0,02	0,87/0,02	0,89/0,04	0,79/0,04	0,86/0,03	0,86/0,03	0,87/0,02	0,73/0,08	0,72/0,07
77-magic	0,36/0,04	0,38/0,04	0,38/0,05	0,38/0,05	0,35/0,05	0,38/0,04	0,38/0,06	0,39/0,06	0,35/0,05	0,35/0,05	0,39/0,05	0,38/0,06	0,18/0,06	0,17/0,08
78-thyroid hypothyroi...	0,61/0,09	0,63/0,10	0,76/0,09	0,78/0,05	0,64/0,11	0,77/0,05	0,77/0,04	0,84/0,04	0,56/0,15	0,67/0,07	0,78/0,04	0,78/0,04	0,78/0,04	0,79/0,03
79-vertebra column 3c	0,64/0,04	0,64/0,05	0,66/0,06	0,67/0,05	0,65/0,05	0,66/0,04	0,65/0,04	0,68/0,05	0,65/0,06	0,68/0,06	0,66/0,04	0,66/0,05	0,63/0,05	0,61/0,04
80-hepatitis	0,33/0,13	0,36/0,13	0,38/0,15	0,42/0,16	0,39/0,18	0,34/0,13	0,32/0,14	0,43/0,17	0,33/0,15	0,38/0,16	0,35/0,14	0,34/0,14	0,33/0,12	0,34/0,12
81-balance scale	0,55/0,04	0,55/0,04	0,55/0,04	0,56/0,05	0,54/0,05	0,57/0,03	0,55/0,03	0,57/0,05	0,54/0,05	0,57/0,04	0,56/0,03	0,56/0,03	0,49/0,05	0,48/0,05
82-robot failure lp5	0,44/0,05	0,43/0,05	0,52/0,07	0,52/0,06	0,49/0,06	0,44/0,05	0,44/0,05	0,54/0,06	0,50/0,06	0,54/0,06	0,44/0,05	0,44/0,05	0,40/0,04	0,38/0,05
83-autoUniv au7 cpd1...	0,09/0,04	0,08/0,03	0,07/0,04	0,07/0,03	0,07/0,03	0,07/0,03	0,08/0,04	0,07/0,03	0,07/0,03	0,07/0,03	0,08/0,02	0,07/0,03	0,07/0,03	0,08/0,03
84-wdbc	0,82/0,03	0,82/0,03	0,85/0,04	0,87/0,03	0,83/0,04	0,85/0,03	0,85/0,03	0,89/0,03	0,83/0,05	0,87/0,03	0,85/0,03	0,85/0,03	0,84/0,03	0,83/0,03
85-robot nav sensor r...	0,91/0,01	0,92/0,01	0,89/0,04	0,94/0,02	0,91/0,02	0,95/0,02	0,95/0,01	0,95/0,02	0,90/0,03	0,92/0,02	0,95/0,01	0,95/0,01	0,92/0,03	0,91/0,03
86-habermans survival	0,15/0,08	0,12/0,07	0,14/0,09	0,15/0,09	0,14/0,08	0,12/0,08	0,16/0,07	0,12/0,10	0,13/0,09	0,12/0,09	0,13/0,08	0,13/0,09	0,14/0,06	0,12/0,07
87-synthetic control	0,76/0,02	0,75/0,02	0,85/0,02	0,83/0,02	0,80/0,04	0,76/0,01	0,76/0,01	0,87/0,01	0,80/0,03	0,87/0,02	0,76/0,01	0,76/0,01	0,55/0,03	0,55/0,02
88-monks3	0,84/0,04	0,86/0,03	0,87/0,04	0,90/0,03	0,85/0,03	0,89/0,02	0,88/0,02	0,91/0,02	0,85/0,05	0,89/0,03	0,87/0,06	0,88/0,02	0,84/0,07	0,82/0,05
89-movement libras	0,43/0,03	0,43/0,03	0,48/0,05	0,48/0,04	0,44/0,05	0,43/0,03	0,43/0,03	0,49/0,05	0,44/0,05	0,49/0,05	0,43/0,04	0,44/0,03	0,31/0,03	0,31/0,03
90-heart disease cleve...	0,26/0,05	0,26/0,05	0,26/0,04	0,27/0,06	0,28/0,06	0,25/0,04	0,26/0,05	0,28/0,05	0,27/0,06	0,28/0,05	0,26/0,04	0,25/0,04	0,22/0,05	0,23/0,04

Tabela 31 – Estratégias comparadas nos conjuntos de dados 1-45 com aprendiz RoF. Medida: $ALC-\mu_{\kappa}/\sigma_{\kappa}$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
1-connect. mines vs...	0,47 /0,06	0,47 /0,06	0,43 /0,06	0,44/ 0,05	0,44/0,06	0,47 /0,07	0,47 /0,07	0,44/ 0,05	0,43 /0,06	0,43 /0,06	0,47 /0,07	0,47 /0,08	0,44/0,06	0,45/0,06
2-wine quality red	0,18/0,04	0,19 / 0,03	0,19 /0,04	0,19 / 0,03	0,19 / 0,03	0,19 / 0,03	0,19 / 0,03	0,19/0,04	0,18/ 0,03	0,18/ 0,03	0,18/0,04	0,20 /0,04	0,05 /0,04	0,05 /0,04
3-qsar biodegradatio...	0,49/ 0,04	0,51 /0,05	0,50/ 0,04	0,49/0,06	0,46/0,07	0,50/ 0,04	0,51 / 0,04	0,51 /0,05	0,48/0,05	0,48/0,05	0,52 /0,05	0,50/0,05	0,37 /0,06	0,37 /0,07
4-waveform v2	0,54/ 0,02	0,54/ 0,02	0,52/0,03	0,52/ 0,02	0,51 /0,03	0,53/ 0,02	0,53/ 0,02	0,52/ 0,02	0,52/0,03	0,52/0,03	0,53/ 0,02	0,53/ 0,02	0,55 / 0,02	0,55 / 0,01
5-steel plates fault...	0,44 / 0,03	0,42/ 0,03	0,38/0,04	0,41/0,04	0,37/0,05	0,42/ 0,03	0,43 / 0,03	0,43 /0,04	0,35/0,04	0,35/0,04	0,42/ 0,03	0,43 / 0,03	0,22 /0,04	0,27/0,06
6-wine	0,91/ 0,02	0,91/ 0,02	0,91/ 0,02	0,91/ 0,02	0,89 /0,03	0,92 / 0,02	0,92 / 0,02	0,91/ 0,02	0,90/0,03	0,90/0,03	0,92 / 0,02	0,92 / 0,02	0,91/ 0,02	0,91/ 0,02
7-volcanoes dl	0,09 /0,04	0,09 /0,04	0,05 / 0,02	0,05 / 0,03	0,08/0,04	0,09 /0,04	0,08/ 0,03	0,05 / 0,03	0,07/ 0,03	0,07/ 0,03	0,10 /0,05	0,09 /0,04	0,06/ 0,03	0,07/0,04
8-mushroom	0,89/ 0,01	0,89/ 0,01	0,87/0,03	0,89/0,02	0,85/0,02	0,93 / 0,01	0,92/ 0,01	0,92/ 0,01	0,84 /0,02	0,84 /0,02	0,93 / 0,01	0,93 / 0,01	0,90/ 0,01	0,90/ 0,01
9-climate simulation...	0,08/ 0,01	0,08/0,05	0,08/0,07	0,07/0,06	0,09 /0,08	0,08/0,05	0,08/ 0,04	0,09 / 0,04	0,06 /0,04	0,06 / 0,04	0,10 /0,05	0,09 /0,05	0,06 / 0,04	0,08/0,05
10-eeg eye state	0,20/0,04	0,21/ 0,02	0,20/ 0,03	0,20/ 0,03	0,19/ 0,03	0,22 / 0,03	0,20/ 0,03	0,20/0,04	0,19/0,04	0,19/0,04	0,22 / 0,03	0,20/ 0,03	0,11 / 0,03	0,11 /0,04
11-seeds	0,87/0,04	0,87/0,04	0,87/0,04	0,88/0,04	0,86/0,04	0,89 /0,04	0,89 /0,04	0,89 /0,04	0,84 /0,04	0,84 /0,04	0,89 /0,04	0,89 /0,04	0,88/ 0,03	0,88/ 0,03
12-user knowledge	0,69/0,03	0,69/0,03	0,68/0,03	0,69/0,03	0,66/0,03	0,71 / 0,02	0,71 / 0,02	0,70/0,03	0,66/0,03	0,66/0,03	0,71 /0,03	0,71 /0,03	0,62/0,04	0,61 /0,04
13-wine quality white...	0,12/0,03	0,12/0,03	0,11/ 0,02	0,12/0,03	0,11/ 0,02	0,13 /0,03	0,13 /0,03	0,11/0,03	0,11/ 0,02	0,11/ 0,02	0,14 / 0,02	0,14 / 0,02	0,07/ 0,02	0,06 / 0,02
14-phoneme	0,39/ 0,04	0,37/0,06	0,35/0,07	0,36/0,06	0,36/0,06	0,40 /0,05	0,39/0,06	0,37/0,09	0,32 /0,07	0,32 /0,07	0,40 / 0,04	0,39/ 0,04	0,39/0,06	0,37/0,06
15-mfeat fourier	0,46 / 0,02	0,46 / 0,02	0,40/0,03	0,42/ 0,02	0,39/ 0,02	0,46 / 0,02	0,46 / 0,02	0,40/ 0,02	0,39/ 0,02	0,39/ 0,02	0,45/ 0,02	0,46 / 0,02	0,24 / 0,02	0,24 / 0,02
16-ionosphere	0,68/0,05	0,67/ 0,04	0,70/0,05	0,71 / 0,04	0,67/ 0,04	0,70/ 0,04	0,71 / 0,04	0,72 / 0,04	0,65/0,05	0,65/0,05	0,70/ 0,04	0,71 / 0,04	0,57/0,05	0,56 /0,07
17-heart disease va...	0,07 /0,05	0,07 /0,04	0,07 /0,04	0,06/0,04	0,06/0,04	0,06/ 0,03	0,06/0,05	0,05/0,05	0,06/0,04	0,06/0,04	0,06/ 0,03	0,06/ 0,03	0,03 /0,04	0,04/0,04
18-thyroid ann	0,51/0,08	0,54/0,08	0,65/0,08	0,70/0,07	0,59/0,07	0,76 / 0,04	0,74/0,06	0,72/0,05	0,48 /0,08	0,48 /0,08	0,76 /0,05	0,75/0,06	0,65/0,05	0,66/ 0,04
19-connect. vowel...	0,38 /0,02	0,38 /0,02	0,36/0,02	0,37/0,02	0,35/0,02	0,38 /0,02	0,38 /0,02	0,36/0,02	0,34/0,02	0,34/0,02	0,38 /0,02	0,38 /0,02	0,23 /0,02	0,23 /0,02
20-statlog german cre...	0,16 /0,05	0,14/0,05	0,11 /0,05	0,13/ 0,04	0,14/ 0,04	0,13/ 0,04	0,15/0,05	0,13/0,05	0,13/0,05	0,13/0,05	0,13/0,05	0,17 /0,05	0,16 /0,04	0,14/0,05
21-turkiye student	0,13 / 0,01	0,13 / 0,01	0,11/ 0,01	0,11/0,02	0,11/ 0,01	0,12/ 0,01	0,13 / 0,01	0,11/ 0,01	0,11/ 0,01	0,11/0,02	0,12/ 0,01	0,13 / 0,01	0,07 / 0,01	0,09/0,02
22-ringnorm	0,66/ 0,02	0,66/0,03	0,62/ 0,02	0,63/ 0,02	0,61/ 0,02	0,67 / 0,02	0,67 / 0,02	0,63/0,03	0,60/0,03	0,60/0,03	0,67 / 0,02	0,67 / 0,02	0,14 /0,03	0,14 / 0,02
23-ozone eighthr	0,11/0,06	0,10/0,06	0,10/0,06	0,08/ 0,05	0,10/0,07	0,16 /0,08	0,12/0,06	0,12/ 0,04	0,06 /0,06	0,06 /0,06	0,15/0,08	0,15/0,07	0,10/ 0,05	0,10/0,06
24-yeast 4class	0,33 / 0,03	0,33 /0,04	0,32/ 0,03	0,33 / 0,03	0,32/0,04	0,33 /0,04	0,33 /0,04	0,32/ 0,03	0,30/ 0,03	0,30/ 0,03	0,33 /0,04	0,34 / 0,03	0,17/0,05	0,16 /0,05
25-parkinsons	0,58/0,10	0,60/ 0,08	0,59/0,10	0,60/0,10	0,56/0,10	0,62/ 0,09	0,59/0,10	0,63 /0,10	0,55/0,12	0,55/0,12	0,63 / 0,09	0,61 / 0,09	0,55/0,10	0,53 /0,10
26-flare	0,41/0,04	0,42/ 0,03	0,41/ 0,03	0,43 /0,05	0,40/0,05	0,42/ 0,03	0,42/ 0,03	0,41/ 0,03	0,39 /0,04	0,39 /0,04	0,43 /0,04	0,43 / 0,03	0,41/0,05	0,41/0,04
27-autoUniv au6 cd1 4...	0,01/ 0,01	0,01/ 0,01	0,01/ 0,01	0,01/0,02	0,01/0,02	0,01/ 0,01	0,01/ 0,01	0,01/0,02	0,01/0,02	0,01/0,02	0,01/ 0,01	0,01/ 0,01	0,01/ 0,01	0,01/ 0,01
28-glass	0,52/0,05	0,52/0,05	0,49/0,05	0,51/0,06	0,49/0,05	0,53 / 0,04	0,52/0,06	0,51/0,05	0,48 /0,06	0,48 /0,06	0,52/0,06	0,54 / 0,04	0,48 /0,06	0,48 /0,05
29-ecoli	0,71/ 0,04	0,71/ 0,04	0,72/0,05	0,74 /0,05	0,70/0,05	0,73/ 0,04	0,73/0,05	0,74 /0,05	0,69 /0,05	0,69 /0,05	0,73/0,05	0,74 /0,05	0,72/0,05	0,71/0,05
30-thyroid newthyroid	0,87/0,07	0,87/ 0,06	0,87/ 0,06	0,88/ 0,06	0,85 /0,07	0,88/ 0,06	0,88/ 0,05	0,89 / 0,06	0,85 / 0,06	0,85 / 0,06	0,89 / 0,06	0,88/ 0,06	0,88/ 0,06	0,86/ 0,06
31-wilt	0,21 /0,11	0,23/0,10	0,52/0,09	0,57/0,06	0,28/0,12	0,63 / 0,04	0,61/0,05	0,63 /0,05	0,33/0,13	0,33/0,13	0,62/0,06	0,61/ 0,04	0,43/0,05	0,42/0,05
32-optdigits	0,71 / 0,01	0,70 / 0,01	0,63/0,02	0,65/0,02	0,63/0,02	0,70 / 0,01	0,70 / 0,01	0,64/0,02	0,61/0,02	0,61/0,02	0,70 / 0,01	0,70 / 0,01	0,45 /0,02	0,49/0,03
33-pima indians diabe...	0,34/0,06	0,35/0,07	0,35/0,06	0,35 / 0,05	0,34/0,06	0,36 /0,06	0,35 / 0,05	0,36 / 0,05	0,34/ 0,05	0,34/ 0,05	0,35/ 0,04	0,35 / 0,05	0,25 /0,07	0,28/0,06
34-twonorm	0,80 / 0,01	0,80 /0,02	0,79/0,02	0,79/0,02	0,77/0,03	0,80 /0,02	0,80 /0,02	0,78/0,02	0,77/0,03	0,77/0,03	0,81 / 0,01	0,80 /0,02	0,70/0,02	0,69 /0,02
35-volcanoes e1	0,12/ 0,04	0,12/0,07	0,11/ 0,03	0,13/0,06	0,13/0,05	0,14 /0,06	0,14 /0,06	0,12/0,05	0,10 /0,05	0,11/0,05	0,14 /0,05	0,14 /0,06	0,14 / 0,04	0,16 /0,05
36-nursery	0,62 / 0,02	0,61/ 0,02	0,59/ 0,02	0,58/ 0,02	0,57/ 0,02	0,62 / 0,02	0,62 / 0,02	0,60/ 0,02	0,57/0,03	0,57/0,03	0,62 / 0,02	0,63 / 0,02	0,54 /0,04	0,55/0,04
37-statlog heart	0,53/0,08	0,54/ 0,07	0,53/ 0,07	0,54/0,08	0,52 / 0,07	0,55 /0,08	0,55 / 0,07	0,55 / 0,07	0,53/0,08	0,53/0,08	0,55 /0,08	0,54/ 0,07	0,55 / 0,07	0,54/0,07
38-statlog australian...	0,57/0,05	0,60 / 0,04	0,58/0,05	0,59/0,06	0,56/0,06	0,58/ 0,04	0,61 /0,05	0,57/0,05	0,54 /0,06	0,54 /0,06	0,58/0,06	0,59/0,06	0,56/ 0,04	0,54 /0,06
39-molecular splice j...	0,40/0,03	0,44 /0,02	0,38/0,04	0,40/0,04	0,37/0,04	0,40/0,03	0,43 /0,02	0,37/0,04	0,35/0,09	0,35/0,09	0,40/0,03	0,42/0,02	0,11 / 0,01	0,14/ 0,01
40-first order theore...	0,15 / 0,02	0,15 / 0,02	0,11/0,03	0,13/0,03	0,12/0,03	0,15 / 0,02	0,15 / 0,02	0,13/0,03	0,11/ 0,02	0,11/ 0,02	0,14/0,03	0,15 / 0,02	0,07 / 0,02	0,08/ 0,02
41-page blocks	0,53/0,05	0,55/0,06	0,58/0,05	0,67 /0,04	0,54/0,06	0,64/ 0,03	0,65/ 0,03	0,69 / 0,03	0,47 /0,08	0,47 /0,08	0,65/0,06	0,65/0,05	0,54/0,08	0,51/0,11
42-artificial charact...	0,25 /0,02	0,25 /0,02	0,22/0,02	0,22/0,03	0,21/ 0,01	0,25 /0,02	0,25 /0,02	0,22/0,02	0,20/0,02	0,21/0,02	0,25 /0,02	0,25 / 0,01	0,08 /0,02	0,09/0,02
43-vertebra column 2c	0,55/0,07	0,56/0,07	0,55/ 0,05	0,57/0,07	0,52 /0,08	0,58/0,06	0,58/0,06	0,55/0,07	0,52 /0,06	0,52 /0,06	0,59 / 0,05	0,59 /0,06	0,56/0,06	0,55/0,07
44-pendigits	0,78/ 0,01	0,78/ 0,01	0,72/0,02	0,74/0,02	0,70/0,02	0,79 / 0,01	0,79 / 0,01	0,74/ 0,01	0,69/0,02	0,69/0,02	0,78/ 0,01	0,78/ 0,01	0,59/0,03	0,54 /0,04
45-mammographic mass	0,54/0,06	0,55/0,07	0,55/0,06	0,56 /0,06	0,52/0,06	0,55/0,05	0,56 /0,05	0,56 /0,05	0,53/ 0,04	0,53/ 0,04	0,56 /0,05	0,56 /0,05	0,52/0,07	0,51 /0,07

Tabela 32 – Estratégias comparadas nos conjuntos de dados 46-90 com aprendiz RoF. Medida: $ALC-\mu_{\kappa}/\sigma_{\kappa}$. *Detalhes na Tabela 21.*

Conjunto de dados	ATUeuc	ATUman	EERacc	EERent	HS	HTUeuc	HTUman	Mar	Rnd	SGmulti	TUeuc	TUman	DWeuc	DWman
46-spambase	0,66/ 0,03	0,66/ 0,03	0,66/ 0,03	0,65/0,04	0,64/ 0,03	0,69/0,03	0,69/0,03	0,69/0,03	0,65/ 0,03	0,65/ 0,03	0,68/0,04	0,69/0,03	0,59/0,04	0,45/0,12
47-horse colic surgic...	0,51/ 0,05	0,52/0,07	0,51/0,09	0,53/0,07	0,50/0,07	0,54/0,05	0,53/0,07	0,53/0,08	0,49/0,07	0,49/0,07	0,54/0,06	0,54/0,06	0,51/0,07	0,53/0,05
48-iris	0,91/0,04	0,91/0,04	0,91/0,04	0,92/0,04	0,91/0,04	0,92/0,04	0,92/0,04	0,92/0,04	0,91/0,04	0,91/0,04	0,92/0,04	0,92/0,04	0,92/0,04	0,92/0,04
49-volcanoes b5	0,17/0,05	0,16/0,05	0,19/0,06	0,21/0,06	0,17/0,06	0,23/0,03	0,22/0,05	0,20/0,06	0,15/0,06	0,15/0,06	0,22/0,04	0,22/0,05	0,12/0,04	0,14/0,05
50-statlog image segm...	0,82/ 0,01	0,83/ 0,01	0,82/0,02	0,84/0,02	0,81/0,02	0,85/0,01	0,85/0,01	0,84/ 0,01	0,79/0,02	0,79/0,02	0,84/0,02	0,85/0,01	0,68/0,03	0,66/0,03
51-voting	0,75/0,07	0,76/0,06	0,76/0,06	0,77/0,06	0,73/0,07	0,80/0,05	0,79/ 0,05	0,80/0,05	0,73/0,08	0,73/0,08	0,80/0,05	0,79/0,06	0,80/0,06	0,80/0,06
52-wholesale channel	0,73/0,07	0,73/0,08	0,73/0,09	0,75/0,07	0,72/0,07	0,75/0,06	0,75/0,06	0,75/0,06	0,72/0,07	0,72/0,07	0,75/0,07	0,75/0,06	0,74/ 0,06	0,72/0,06
53-cardiotocography 3...	0,63/ 0,03	0,64/ 0,03	0,67/0,04	0,70/0,03	0,62/0,04	0,68/ 0,03	0,69/ 0,03	0,70/0,04	0,55/0,06	0,55/0,06	0,69/ 0,03	0,70/0,03	0,57/0,04	0,58/0,04
54-hill valley withou...	0,72/ 0,03	0,72/0,04	0,73/0,04	0,72/0,06	0,67/0,05	0,77/0,03	0,73/0,04	0,66/0,08	0,67/0,06	0,67/0,06	0,72/0,04	0,74/0,03	0,55/0,09	0,58/0,09
55-autoUniv au7 300 d...	0,07/0,02	0,07/0,02	0,06/ 0,02	0,06/ 0,02	0,07/0,02	0,07/0,02	0,07/0,02	0,06/ 0,02	0,06/ 0,02	0,06/ 0,02	0,07/0,02	0,07/0,02	0,04/0,01	0,04/0,02
56-texture	0,79/ 0,01	0,78/ 0,01	0,76/0,02	0,77/0,02	0,74/0,02	0,80/0,01	0,80/0,01	0,77/ 0,01	0,71/0,02	0,71/0,02	0,80/0,01	0,80/0,01	0,60/0,03	0,56/0,03
57-autoUniv au7 700	0,06/0,03	0,08/ 0,03	0,07/ 0,03	0,08/ 0,03	0,07/ 0,03	0,08/0,04	0,08/ 0,03	0,07/ 0,03	0,07/ 0,03	0,07/ 0,03	0,08/ 0,02	0,08/0,04	0,10/0,03	0,09/0,03
58-musk	0,24/ 0,03	0,16/ 0,03	0,34/0,07	0,36/0,07	0,32/0,06	0,36/0,06	0,24/0,09	0,34/0,16	0,26/0,05	0,26/0,05	0,33/0,11	0,36/0,06	0,11/ 0,03	0,04/0,04
59-kr vs kp	0,60/0,03	0,60/ 0,02	0,56/0,03	0,58/0,05	0,52/0,05	0,62/0,03	0,62/0,02	0,60/0,04	0,53/0,03	0,53/0,03	0,62/0,03	0,61/0,03	0,53/0,07	0,52/0,03
60-autoUniv au1 1000	0,06/0,04	0,09/0,04	0,07/0,04	0,06/0,03	0,07/0,05	0,07/0,04	0,08/0,04	0,07/0,05	0,06/0,04	0,06/0,04	0,08/0,04	0,07/ 0,03	0,12/0,04	0,11/0,04
61-monks1	0,54/0,07	0,55/0,05	0,55/0,06	0,58/0,05	0,53/0,05	0,61/0,04	0,61/0,04	0,60/0,05	0,53/0,06	0,52/0,06	0,60/0,05	0,60/0,05	0,52/0,08	0,51/0,06
62-car evaluation	0,39/0,05	0,38/0,05	0,42/0,05	0,44/0,05	0,39/ 0,04	0,39/0,05	0,39/0,05	0,46/0,04	0,39/0,06	0,39/0,06	0,39/0,05	0,39/0,05	0,37/0,05	0,37/0,05
63-tic tac toe	0,29/ 0,04	0,28/0,05	0,28/0,05	0,30/0,06	0,27/0,05	0,30/0,04	0,30/0,05	0,32/0,05	0,28/0,05	0,28/0,05	0,31/0,05	0,31/0,05	0,12/0,05	0,12/0,05
64-thyroid sick euthy...	0,54/0,09	0,55/0,09	0,67/ 0,05	0,69/0,09	0,55/0,11	0,73/0,07	0,71/ 0,07	0,71/0,09	0,44/0,12	0,44/0,12	0,72/0,07	0,71/0,08	0,65/0,15	0,65/0,09
65-bupa	0,20/0,08	0,21/0,07	0,22/0,09	0,20/0,07	0,20/0,07	0,22/0,08	0,20/ 0,06	0,22/0,09	0,21/ 0,06	0,21/ 0,06	0,21/0,08	0,21/0,07	0,19/0,06	0,19/0,06
66-abalone 3class	0,35/ 0,03	0,36/0,04	0,34/0,04	0,36/0,04	0,35/ 0,03	0,37/0,03	0,36/0,03	0,35/ 0,03	0,34/ 0,03	0,34/ 0,03	0,36/0,03	0,35/0,04	0,24/0,09	0,24/0,09
67-banana	0,60/0,04	0,58/ 0,04	0,52/0,07	0,51/0,09	0,51/0,05	0,59/0,05	0,58/0,06	0,48/0,10	0,53/ 0,04	0,53/ 0,04	0,55/0,10	0,54/0,10	0,22/0,07	0,22/0,07
68-lymphography	0,58/0,06	0,57/0,07	0,54/0,06	0,58/0,05	0,56/ 0,06	0,58/0,06	0,58/0,06	0,56/ 0,06	0,54/0,07	0,54/0,07	0,59/0,07	0,58/0,06	0,57/ 0,06	0,57/ 0,06
69-volcanoes a3	0,27/0,05	0,28/ 0,04	0,28/0,06	0,30/0,05	0,25/0,05	0,29/ 0,04	0,30/ 0,04	0,31/0,05	0,22/0,06	0,22/0,06	0,29/0,05	0,32/0,05	0,27/0,06	0,28/ 0,04
70-leaf	0,52/0,04	0,52/0,04	0,48/0,04	0,50/0,04	0,48/0,04	0,52/0,04	0,52/0,04	0,49/0,04	0,48/0,04	0,48/0,04	0,52/0,03	0,53/0,04	0,46/0,03	0,46/0,03
71-heart disease hung...	0,50/0,09	0,51/0,10	0,49/0,10	0,51/0,08	0,48/0,09	0,49/0,10	0,50/0,08	0,50/0,07	0,48/0,06	0,48/0,06	0,50/0,09	0,51/0,10	0,47/0,08	0,48/0,09
72-saheart	0,24/ 0,06	0,21/0,07	0,24/0,07	0,23/0,07	0,21/0,07	0,24/0,07	0,26/0,06	0,25/0,07	0,23/0,08	0,23/0,08	0,24/ 0,06	0,24/0,07	0,10/0,08	0,12/ 0,05
73-statlog vehicle si...	0,50/0,03	0,51/0,03	0,48/ 0,02	0,49/0,03	0,47/0,03	0,51/ 0,02	0,52/0,03	0,48/0,03	0,46/0,03	0,46/0,03	0,51/0,03	0,52/0,02	0,40/0,03	0,40/0,04
74-spect heart	0,25/ 0,07	0,23/0,09	0,23/0,06	0,26/0,10	0,23/0,10	0,24/0,10	0,24/0,09	0,26/0,09	0,24/ 0,07	0,24/ 0,07	0,24/0,09	0,25/0,09	0,23/0,07	0,23/0,09
75-indian liver patie...	0,12/0,05	0,15/ 0,05	0,15/0,06	0,15/0,06	0,15/0,08	0,18/0,07	0,16/0,07	0,16/0,07	0,17/0,05	0,17/0,06	0,17/0,06	0,17/0,07	0,15/0,07	0,14/0,06
76-banknote authentic...	0,91/ 0,02	0,90/ 0,02	0,89/0,03	0,91/ 0,02	0,87/0,04	0,94/0,02	0,93/0,01	0,92/ 0,02	0,86/0,03	0,86/0,03	0,93/0,02	0,93/0,02	0,87/0,03	0,87/0,03
77-magic	0,40/ 0,03	0,41/0,04	0,41/ 0,03	0,41/0,04	0,38/0,05	0,42/0,03	0,41/0,04	0,43/0,05	0,38/0,05	0,38/0,05	0,40/0,05	0,41/0,05	0,22/0,08	0,19/0,08
78-thyroid hypothyroi...	0,66/0,11	0,68/0,11	0,75/0,05	0,76/0,05	0,68/0,08	0,81/0,04	0,80/ 0,03	0,81/0,04	0,64/0,10	0,64/0,10	0,80/0,05	0,82/0,04	0,80/ 0,04	0,81/0,04
79-vertebra column 3c	0,67/0,05	0,67/0,05	0,68/0,05	0,69/0,05	0,66/0,04	0,70/0,05	0,71/0,04	0,71/0,04	0,66/0,06	0,66/0,06	0,70/0,05	0,71/0,05	0,69/0,05	0,68/0,05
80-hepatitis	0,38/0,14	0,40/0,15	0,38/0,14	0,38/ 0,13	0,37/0,16	0,37/0,15	0,41/0,15	0,39/0,15	0,32/0,13	0,32/0,13	0,40/0,15	0,38/0,15	0,36/0,14	0,35/ 0,12
81-balance scale	0,63/ 0,03	0,61/0,04	0,61/ 0,03	0,63/0,04	0,58/0,05	0,65/0,03	0,63/ 0,03	0,63/ 0,03	0,59/0,04	0,59/0,04	0,64/0,03	0,63/ 0,03	0,60/ 0,03	0,60/0,04
82-robot failure lp5	0,48/0,06	0,48/ 0,05	0,46/0,06	0,48/0,06	0,45/ 0,05	0,49/0,05	0,48/0,05	0,48/0,07	0,45/0,06	0,45/0,06	0,49/0,06	0,49/0,05	0,42/ 0,05	0,40/0,05
83-autoUniv au7 cpd1...	0,10/0,04	0,09/0,04	0,07/0,03	0,09/0,04	0,08/ 0,03	0,08/0,04	0,09/0,04	0,08/0,04	0,08/0,04	0,08/0,04	0,08/0,04	0,10/0,04	0,07/0,03	0,08/0,03
84-wdbc	0,85/0,03	0,86/0,03	0,85/0,03	0,86/ 0,02	0,85/ 0,02	0,88/0,03	0,88/0,03	0,88/0,02	0,84/0,03	0,84/0,03	0,88/0,02	0,88/0,03	0,87/0,03	0,87/0,03
85-robot nav sensor r...	0,79/0,01	0,76/ 0,02	0,77/ 0,02	0,78/0,03	0,74/0,03	0,78/ 0,02	0,79/0,02	0,77/0,03	0,73/0,03	0,73/0,03	0,78/ 0,02	0,78/ 0,02	0,53/0,08	0,55/0,07
86-habermans survival	0,13/ 0,07	0,12/ 0,07	0,12/0,08	0,13/0,09	0,12/0,08	0,11/0,10	0,15/0,09	0,11/0,08	0,13/ 0,07	0,12/ 0,07	0,12/ 0,07	0,14/0,08	0,13/ 0,07	0,12/0,08
87-synthetic control	0,83/ 0,01	0,83/0,02	0,81/0,02	0,82/0,01	0,79/0,02	0,84/0,01	0,84/0,01	0,82/0,01	0,79/0,02	0,79/0,02	0,84/0,01	0,84/0,01	0,64/0,02	0,64/0,02
88-monks3	0,79/0,04	0,82/0,03	0,82/0,03	0,82/0,04	0,75/0,04	0,86/0,02	0,85/ 0,02	0,86/0,02	0,75/0,05	0,75/0,05	0,85/0,03	0,86/0,03	0,84/0,04	0,83/0,03
89-movement libras	0,55/0,04	0,55/0,04	0,50/ 0,04	0,51/0,05	0,48/0,05	0,55/0,04	0,55/0,04	0,51/0,05	0,48/0,05	0,48/0,05	0,55/0,04	0,55/0,04	0,45/ 0,04	0,44/0,04
90-heart disease clev...	0,26/0,05	0,27/0,05	0,25/ 0,04	0,26/ 0,04	0,26/ 0,04	0,25/0,05	0,27/0,04	0,27/0,05	0,26/0,05	0,26/0,05	0,25/0,05	0,26/ 0,04	0,23/0,04	0,24/ 0,04