

Escolha de classificadores para métodos de Ensemble em classificação de áudios

Isabelle Rodrigues Vaz de Melo

Programa de engenharia elétrica, Universidade Federal do Rio de
Janeiro, Brasil

isabelle.melo@coppe.ufrj.br

Sumário

1- Introdução ao problema estudado

2 - Métodos de Ensemble

3 - Classificação de áudios

4 - Métodos de escolha de classificadores

5 - Resultados encontrados

6 - Referências

Introdução ao problema estudado

Motivação:

- . Em problemas de classificação, podem ser utilizados diversos classificadores e comparar seus desempenhos para escolher um modelo final
- . Uma das estratégias para obter bons desempenhos é utilizar métodos de Ensemble que fusionam os preditores

Objetivos:

- . Estudar possíveis métodos de escolha de classificadores para fazer parte do Ensemble
- . Aplicação das metodologias a 2 trabalhos prévios de classificação de áudio

Métodos de Ensemble

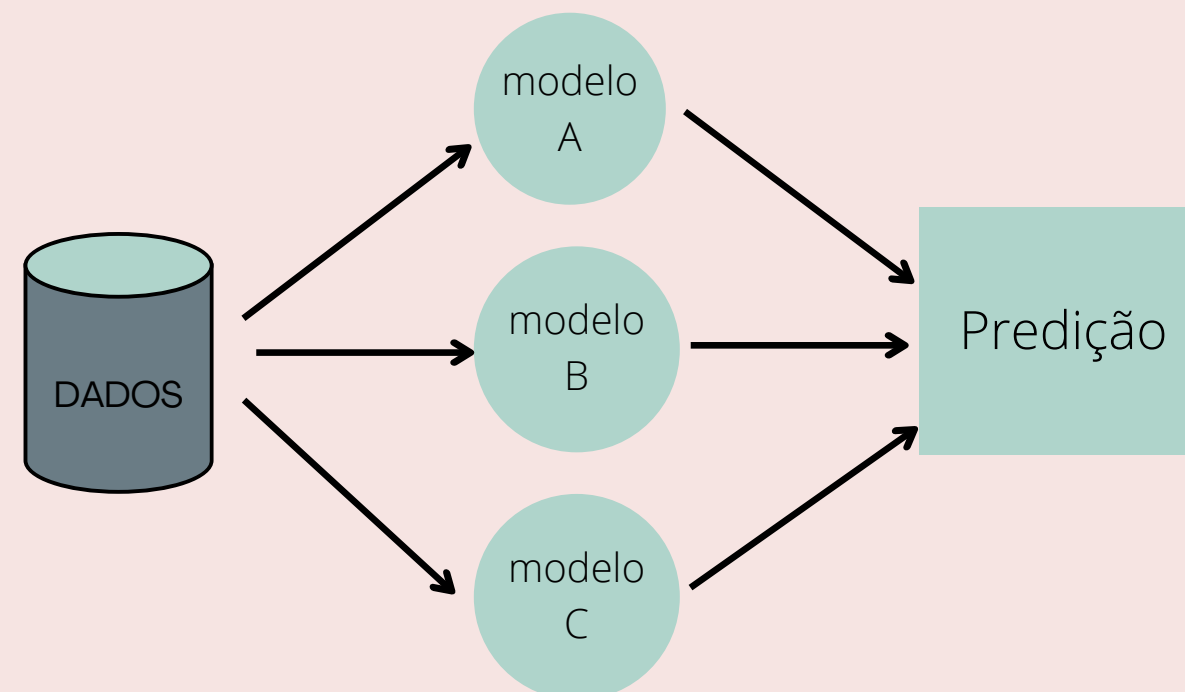
Introdução

.Técnicas que combinam diversos classificadores/modelos a fim de produzir um classificador final

. Fusão de classificadores/modelos

. 3 tipos principais: Bagging, Boosting e Stacking

. Custosos computacionalmente!



Métodos de Ensemble

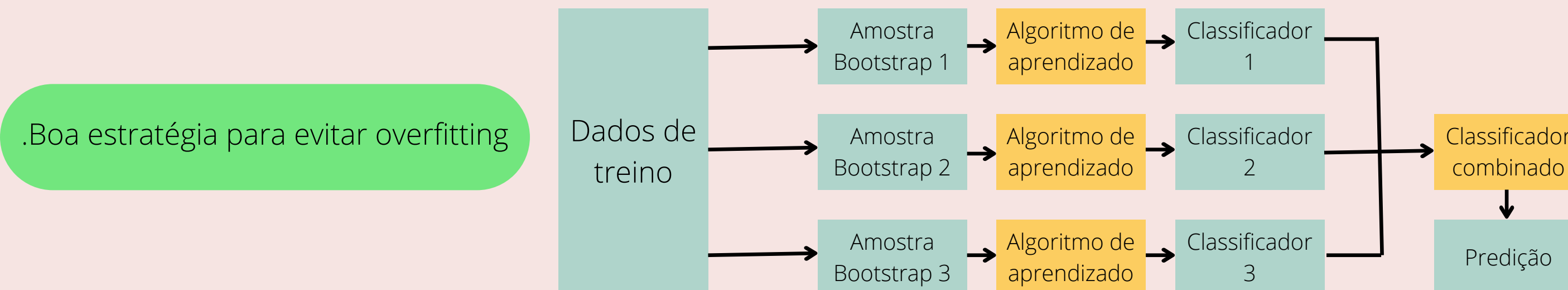
Bagging

.Bagging = Bootstrap aggregation

. Normalmente envolve um único algoritmo de Aprendizado de Máquina

. Escolha aleatória de amostras por Bootstrap

. Resultado final por votação

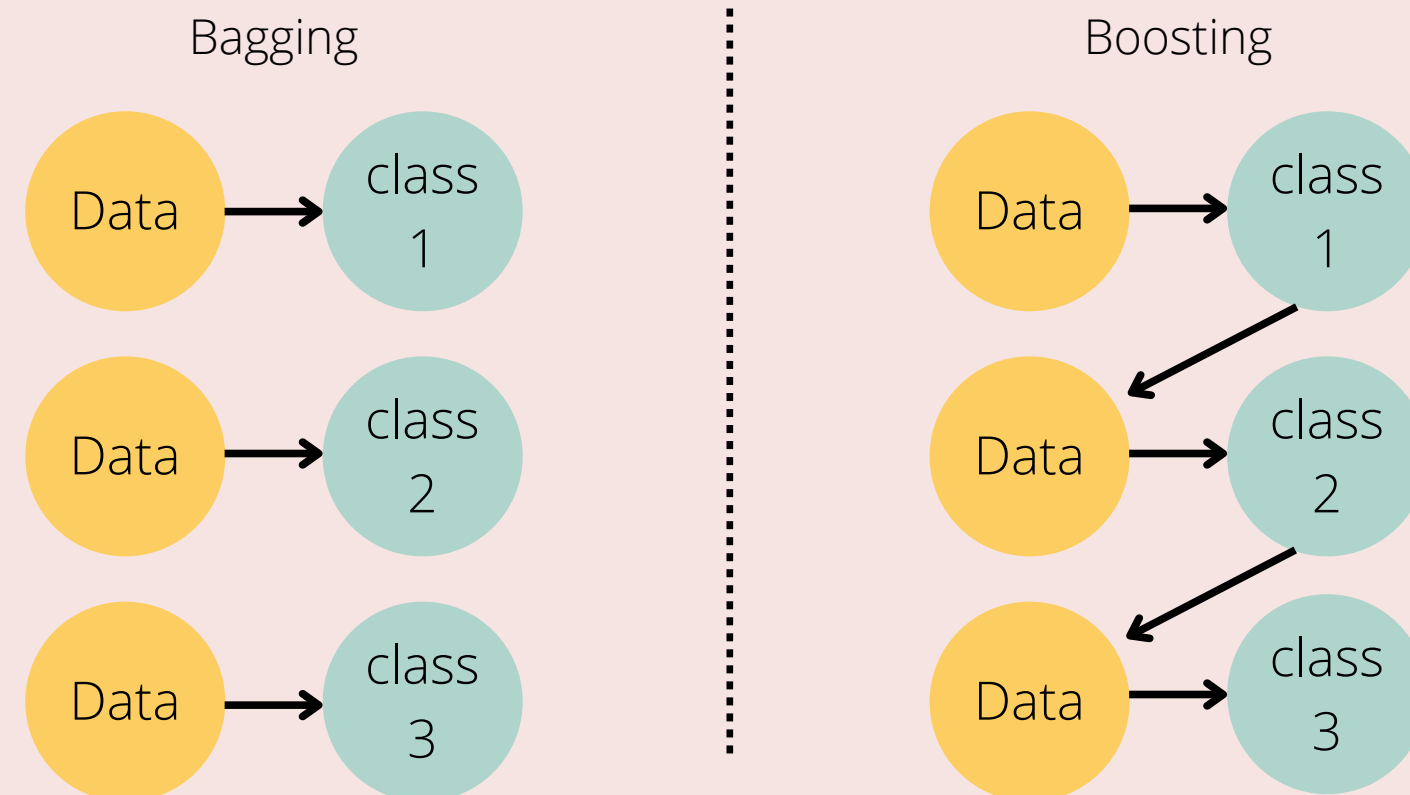


Métodos de Ensemble

Boosting

- . Boosting: Semelhante ao Bagging
- . Treino sequencial de "weak-learners"
- . XGboost, GradientBoost, AdaBoost, etc

. No treino em sequência, identifica os pontos classificados errados e atribuí um peso de votação maior às amostras classificadas corretamente



Métodos de Ensemble

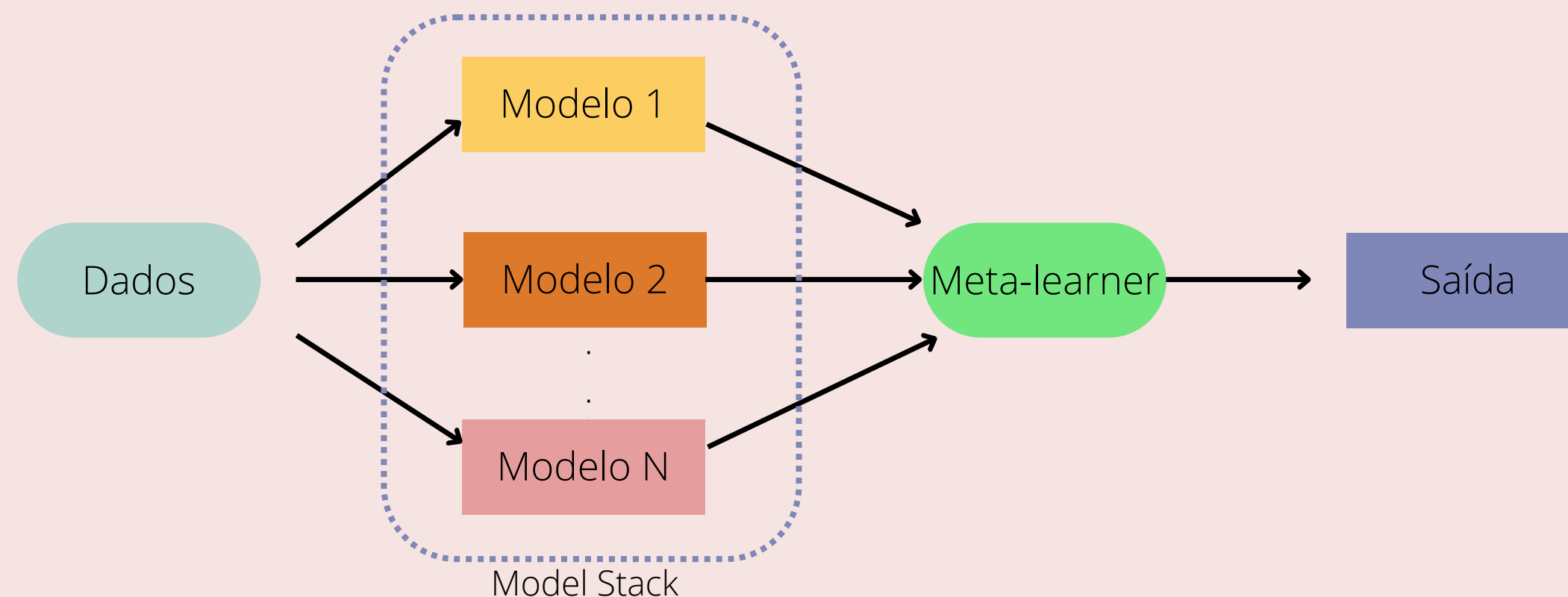
Stacking

- . Stacking : combinação de diferentes classificadores
 - . Modelo utilizado no projeto
- . Utilização de modelos de base e um modelo mais robusto utilizado no final, chamado de meta-learner
 - . Metamodelo é treinado nas saídas do modelo inicial

Métodos de Ensemble

Stacking

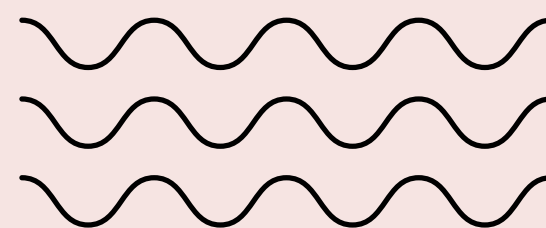
- . Consiste em tomar a amostra inteira, tomar rótulos via modelos de Machine Learning, sem combinação
- . Posteriormente, utilizando esses rótulos como entradas de um “novo conjunto de dados”, prever a variável resposta dos dados originais



Classificação de áudios

Introdução

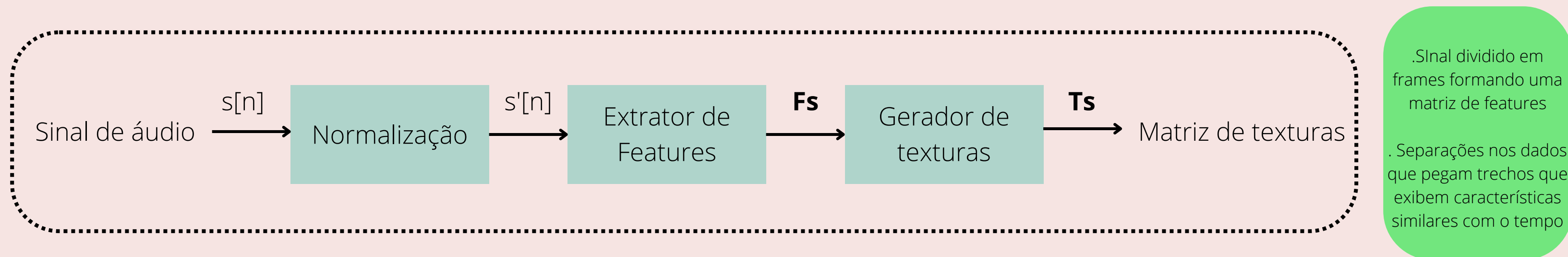
- . Base de dados GTZAN :
- . Áudios coletados entre 2000 e 2001 a partir de rádio, CDs e gravações
 - . Possui 1000 registros (30s/áudio)
- . 10 gêneros musicais (Blues, Clássica, Country, Disco, Hip-hop, Jazz, Metal, Pop, Reggae, Rock)



Classificação de áudios

Introdução

- . Processamento de sinal de áudios
- . O sinal de áudio bruto é processado para obter um conjunto de features que resume informações temporais



Classificação de áudios

Introdução

. SOMP faz uma combinação linear da saída do conjunto de dados para prever a saída
. Apenas inclui as features que combinadas melhor definem a saída

- . Redução de features escolhida no problema: Structured Orthogonal Matching Pursuit (SOMP)
- . SOMP: método de regressão linear esparsa modificada (combina saídas de modelos)
- . Validação cruzada: Nested 10-fold (100 amostras de conjuntos de treinamento, validação e teste)
- . 10 modelos testados: Gaussian Naive Bayes, Multiple Logistic regression, Multiple Layer Perceptron, KNN, SVM-radial, SVM-linear, QDA, LDA, LGBost, Random Forests

Classificação de áudios

Treino e teste de classificadores

. Melhores hiperparâmetros encontrados:

| GNB | LDA | QDA | GB | SVM-l | SVM-r | RF | MLP | KNN | MLR |
|------------------------------|--|--------------------|---|----------------------------------|--|------------------|--|----------------|-----------------------------------|
| var_smoothing: 0.00000001 | solver: 'lsqr' shrinkage: 0.1 | reg_param : 0.1 | n_estimators: 200 learning_rate: 0.1 | max_iter: 1000 C: 0.01 | C: 1 kernel: 'rbf' gamma: 'auto' | n_estimators: 50 | max_iter: 1000 hidden_layer_sizes : [25] activation: 'tanh' | n_neighbors: 5 | max_iter: 10000 C: 0.01 |

. Mediana de acurácias encontradas nas classificações:

| GNB | LDA | QDA | GB | SVM-l | SVM-r | RF | MLP | KNN | MLR |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 64.0% | 80.0% | 80.0% | 83.0% | 81.0% | 86.0% | 77.0% | 82.5% | 79.0% | 81.0% |

Métodos de escolha de classificadores

Introdução

- . Qual o melhor método para selecionar N classificadores dentre os 10 ($N < 10$) para um Ensemble de stacking?
 - . Usar todos os classificadores pode ser muito custoso computacionalmente!
 - . 3 novos métodos investigados: Fk-score, F(1-k)-score e SOMP
 - . Stackings com 3,5 e 8 classificadores (de 10 totais)
 - . 48 stackings avaliados por mediana de acurácia(meta-learners)

Métodos de escolha de classificadores

Método Fk-score

. Método Fk-score: média harmônica acurácia e kappa

$$F_{\kappa} = \frac{2 a \kappa}{(a + \kappa)}$$

$$k = \frac{p_o - p_e}{1 - p_e}$$

.P0 = concordância empírica
.Pe= concordância aleatória

. kappa avalia concordância entre avaliadores e elimina o efeito de coincidência por sorte em classificação

. Os classificadores que obtiverem os maiores Fk são selecionados como candidatos ao Ensemble

Métodos de escolha de classificadores

Método F(1-k)-score

. Método F(1-k)-score: média harmônica acurácia e 1-kappa

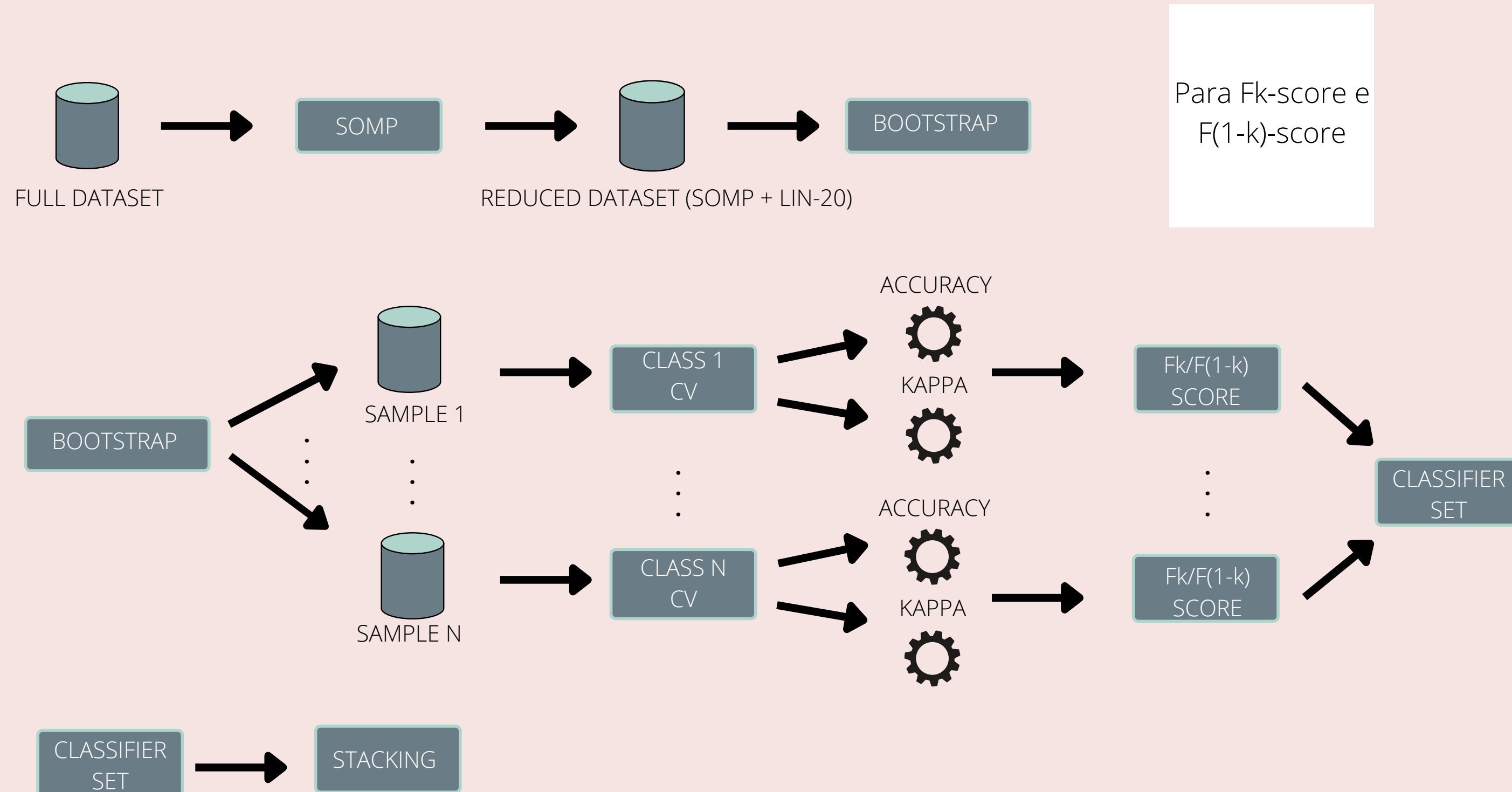
$$F_{1-\kappa} = \frac{2 a (1-\kappa)}{(a+(1-\kappa))}$$

$$k = \frac{p_o - p_e}{1 - p_e}$$

. Os classificadores que obtiverem os maiores F(1-k) são selecionados como candidatos ao Ensemble

Métodos de escolha de classificadores

Metodologia para Fk e F(1-k)-score

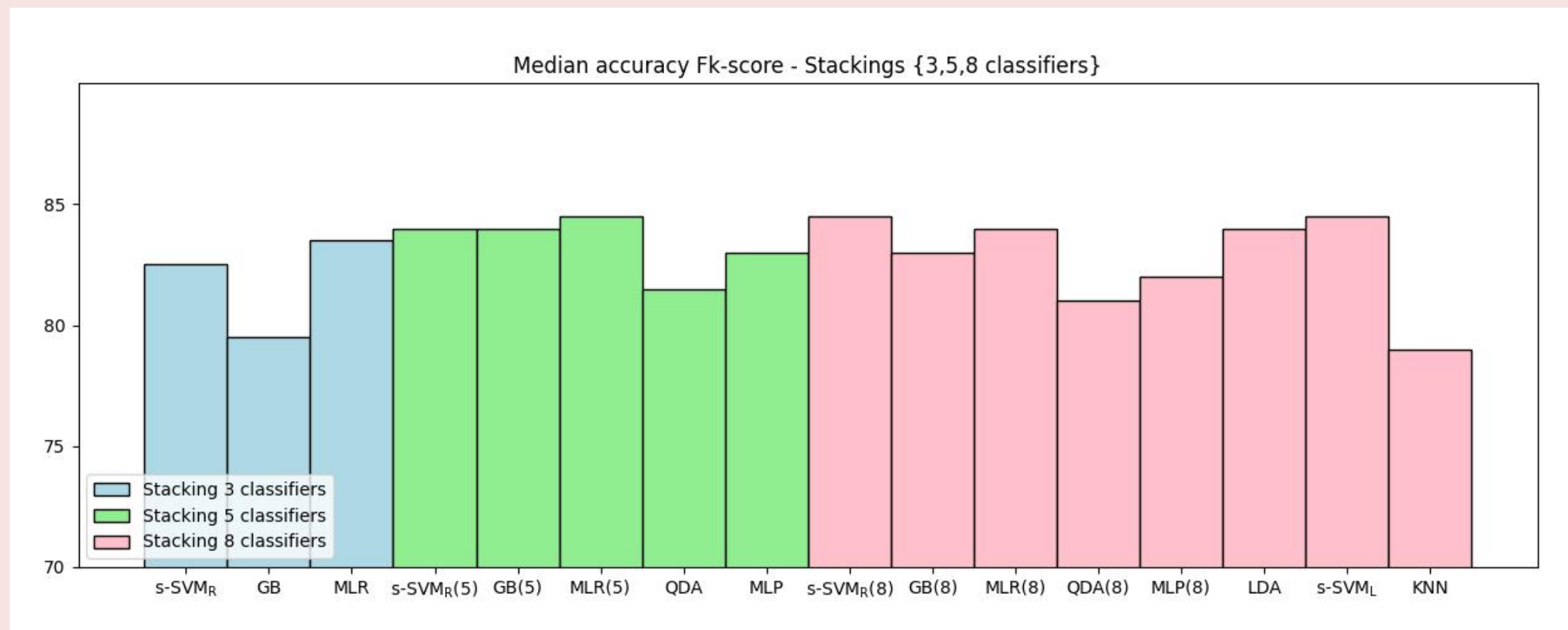


[illegible]

Resultados encontrados

Método Fk-score

. Fk-score:

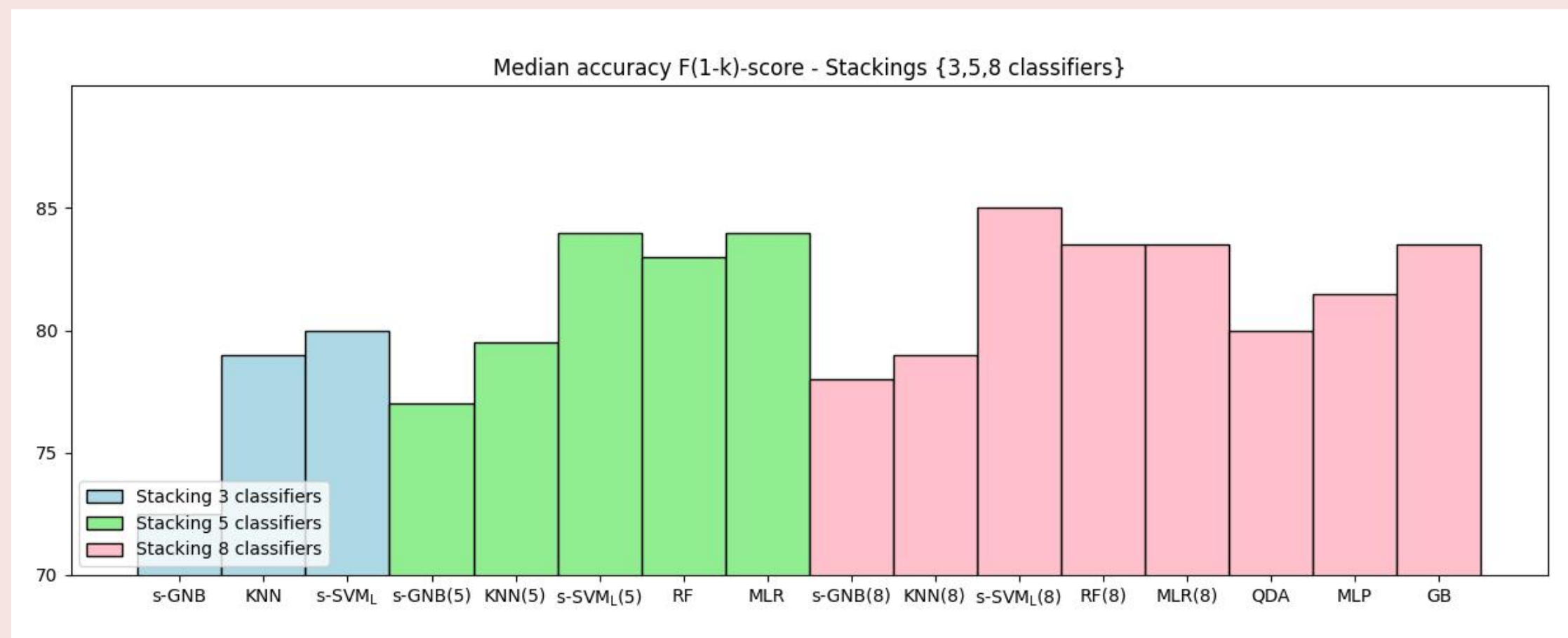


| Média 3 clf | Média 5 clf | Média 8 clf | Média global |
|-------------|-------------|-------------|--------------|
| 81.8% | 83.4% | 82.8% | 82.7% |

Resultados encontrados

Método F(1-k)-score

. F(1-k)-score:



| Média 3 clf | Média 5 clf | Média 8 clf | Média global |
|-------------|-------------|-------------|--------------|
| 77.2% | 81.5% | 81.8% | 80.1% |

Resultados encontrados

Método SOMP

. SOMP:

. Foram variados 4 pesos n : {0.01, 0.1, 1, 10}



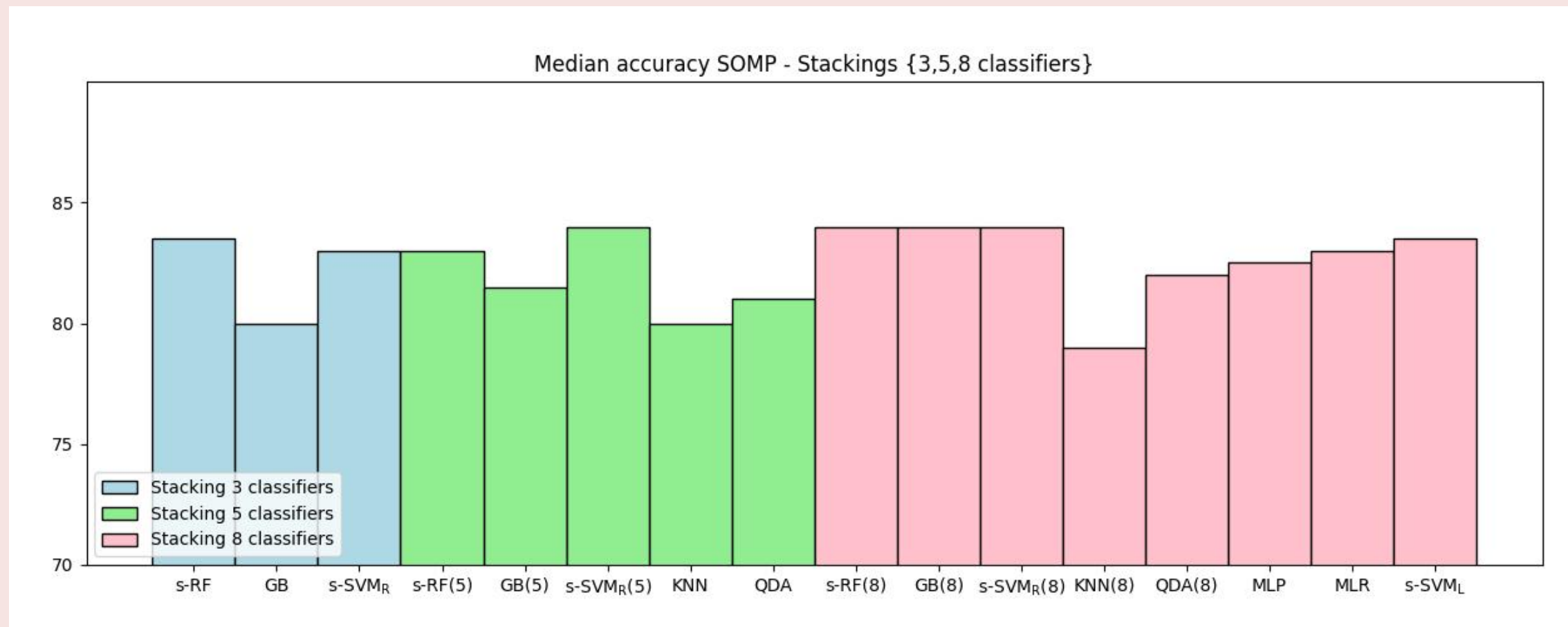
Melhor parâmetro encontrado $n=10$

Classificadores vencedores (do melhor para o pior): {Random Forests, LGBBoost, SVM, KNN, QDA, MLP, MLR, SVM-linear}

Resultados encontrados

Método SOMP

. SOMP:

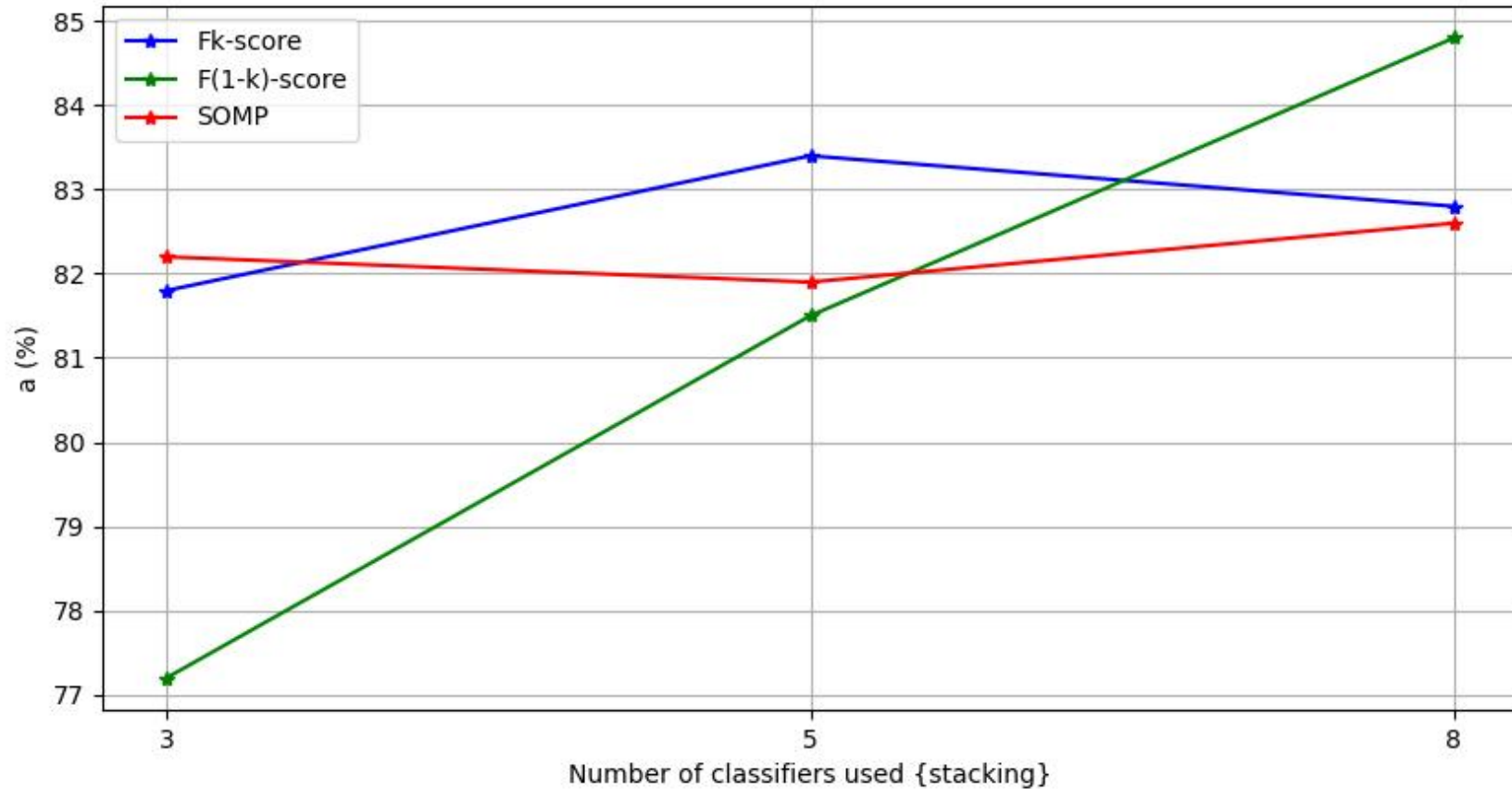


| Média 3 clf | Média 5 clf | Média 8 clf | Média global |
|-------------|-------------|-------------|--------------|
| 82.2% | 81.9% | 82.6% | 82.2% |

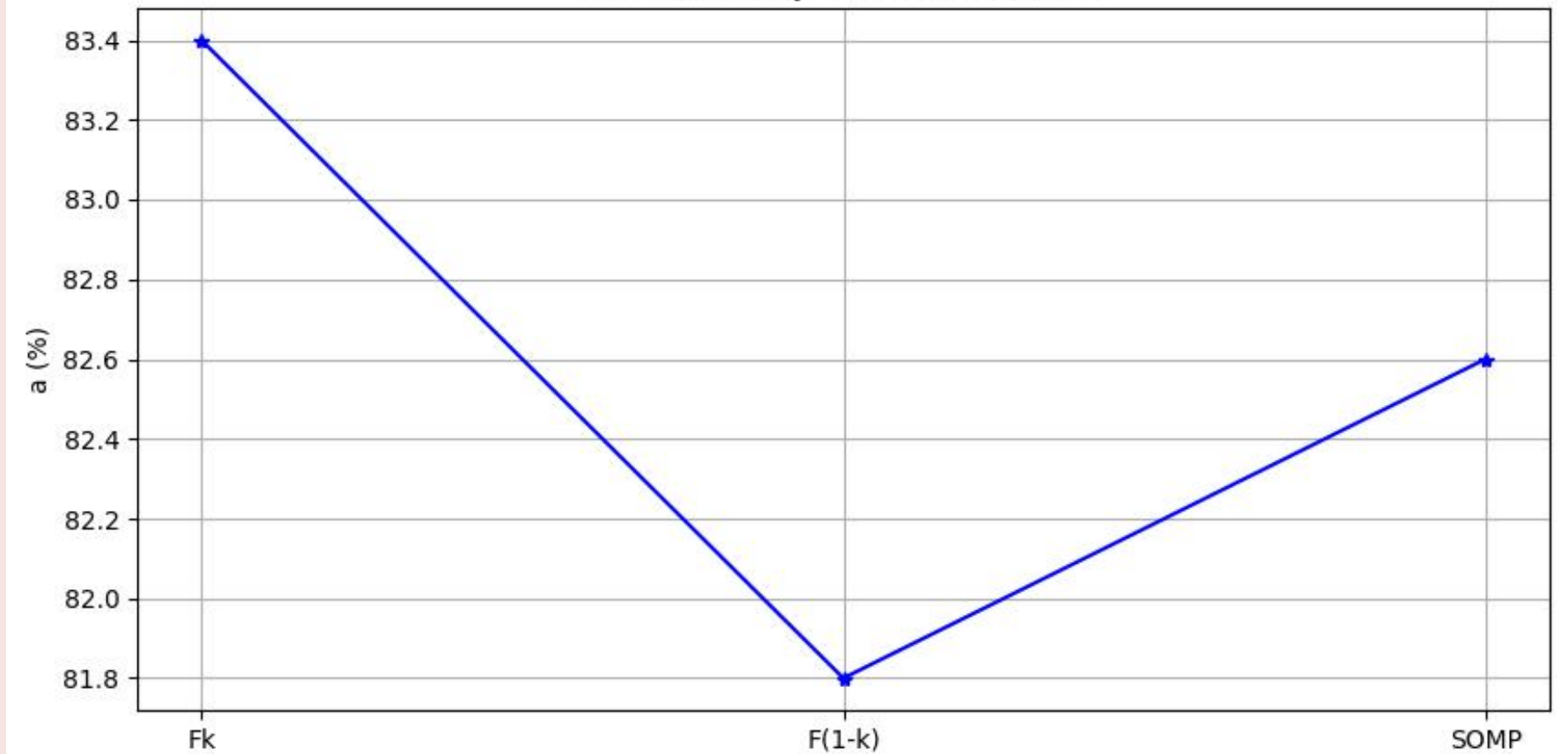
Resultados encontrados

Métodos vencedores e comparação com classifcadores isolados

Accuracy by numbers of classifiers used {Fk, F(1-k), SOMP}



Mean Accuracy {Fk, F(1-k), SOMP}

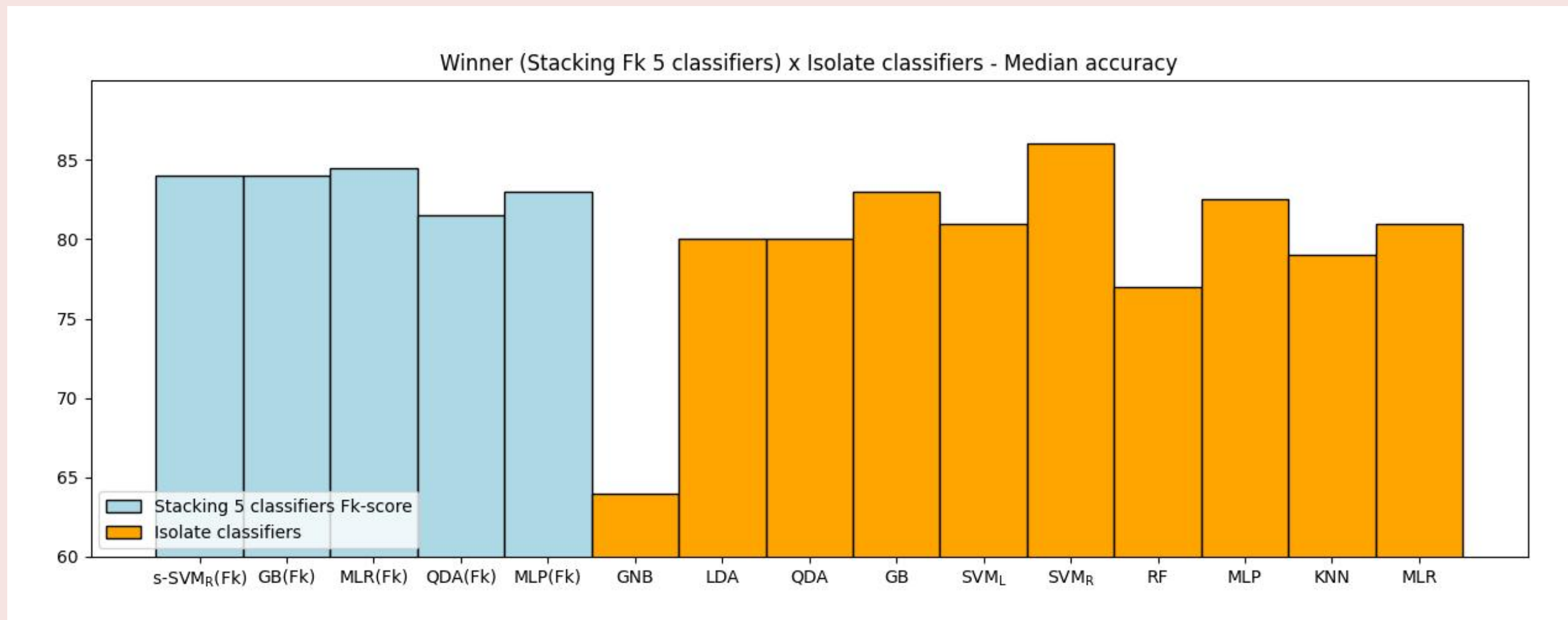


Médias vencedores

| Fk | F(1-k) | SOMP |
|-------|--------|-------|
| 83.4% | 81.8% | 82.6% |

Resultados encontrados

Métodos vencedores e comparação com classifcadores isolados



Resultados encontrados

Conclusões do trabalho

- . Os métodos vencedores para a escolha de classificadores foram o Fk-score e o SOMP
- . O F(1-k)-score obteve resultados ligeiramente piores que Fk e SOMP, entretanto os resultados não são ruins comparados aos classificadores isolados
- . Nem sempre o método de Stacking performa melhor que um único classificador isolado
 - . A codificação do SOMP ainda deve ser modificada
- . É possível se obter resultados interessantes e menos computacionalmente custosos escolhendo determinados classificadores ao Stacking

Referências

- [1] V. H. Da Silva Muniz, J. B. de Oliveira and S. Filho, "Robust Handcrafted Features for Music Genre Classification"
- [2] V. H. Da Silva Muniz, J. B. de Oliveira and S. Filho, "Feature Vector Design for Music Genre Classification," *2021 IEEE Latin American Conference on Computational Intelligence (LA-CCI)*, 2021, pp. 1-6, doi: 10.1109/LA-CCI48322.2021.9769848.
- [3] Xiaoshuang Shi a , Fuyong Xing b , Zhenhua Guo d, Hai Sua , Fujun Liuc , Lin Yanga,c,* , "Structured orthogonal matching pursuit for feature selection", 2019, ISSN 0925-2312