

## Lista de Exercício 1

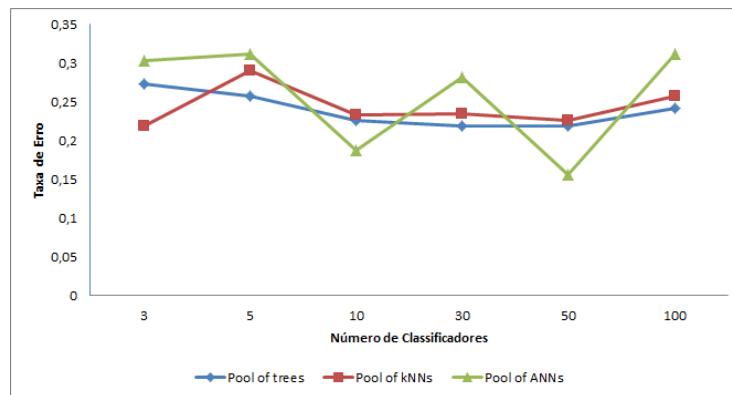
### Questão 1

Implemente o Bagging e teste o seu desempenho com 3 classificadores diferentes. Comente os resultados.

### Resposta

O Bagging foi implementado em Matlab e foram escolhidos os seguintes classificadores para avaliação de desempenho:

- Árvore de decisão: Foi utilizado o algoritmo CART disponível no Matlab, com os valores padrão de configurações;
- kNN: Foi escolhido o valor de  $k = 5$  por alcançar taxa de precisão comparável com os valores maiores de  $k$  e por possuir menor tempo de processamento. A distância utilizada foi a euclidiana (valor padrão do Matlab);
- RNA: Foi escolhida a rede MLP com uma camada oculta com 10 neurônios (arquitetura padrão do Matlab) utilizando a função de treinamento Levenberg-Marquardt (função de treinamento padrão do Matlab).



**Figura 1:** Gráfico de erro para o Bagging com diferentes tipos de classificadores e quantidades diversas de classificadores. Esses resultados foram obtidos através da base Ecoli da UCI.

O desempenho do Bagging foi avaliado para um *pool* com 10, 30, 50 e 100 classificadores. O método de avaliação utilizado foi o de validação cruzada *KFold*, com  $K = 10$ , onde a menor taxa de erro é escolhida para a análise. A base utilizada para a avaliação foi a base *Ecoli* da UCI. A figura 1 ilustra o desempenho do Bagging obtido.

Podemos observar a partir dos resultados que houve melhora na taxa de erro para os classificadores instáveis, que são as CARTs e as RNAs, com o uso do Bagging e variando a quantidade de classificadores utilizados. Houve uma melhora gradual da taxa de erro para as CARTs ao aumentar o número de classificadores até 50. No caso das RNAs, foram geradas as melhores taxas de erro do experimento para as quantidades 10 e 50 de classificadores, e houve pequenas degradações na taxa de erro em relação à quantidade inicial de 3 classificadores para os valores 5 e 100. Analisando o classificador estável kNN, houve degradação da taxa de erro quando se aumenta o número de classificadores do Bagging, comparando com o valor inicial de 3 classificadores.

Os resultados apresentados não apresentam divergência com os conceitos ensinados na aula, os quais afirmam que o Bagging é eficiente para melhorar os resultados de classificadores instáveis e ineficiente para melhorar os resultados de classificadores estáveis.

## Questão 2

Na técnica Rotation Forest, faz alguma diferença aplicar um extrator de características diferente do PCA? Justifique.

### Resposta

Sim, e a diferença pode ser boa ou ruim para os resultados, dependendo do método escolhido para rotacionar os eixos do conjunto de características. Ainda não é certo afirmar que o PCA é o melhor método para efetuar a rotação dos dados. Em [1], é apresentado um estudo no qual as características extraídas através do PCA levam à melhores resultados que outros extratores como o *non-parametric discriminant analysis (NDA)* e as projeções aleatórias. Neste estudo, esse questionamento é levantado, considerando que para o problema de classificação, métodos lineares de extração de características podem ser mais apropriados. Na conclusão do estudo, é dito que até agora o PCA parece ser o melhor método para a rotação dos eixos do conjunto de características.

## Questão 3

Sabe-se que o Random Subspace não é indicado em bases de poucas características. Faz sentido gerar e adicionar novas características para aplicar o Random Subspace? Quais estratégias poderiam ser utilizadas para fazer isto (adicionar novas características dadas as características existentes)?

## Resposta

Não faz sentido. O algoritmo proposto para seleção de múltiplas variáveis utilizando Random Subspace concentra-se em dados de alta dimensionalidade e no problema *small sample size (SSS)*, onde o número de características é maior que a quantidade de exemplos para aprendizado [2]. A ideia da proposta é tentar reduzir o espaço de características, tentando selecionar as mais relevantes, aumentando o número de exemplos por conjunto de características. Ao adicionar novas características a partir das existentes e em seguida utilizar o Random Subspace, os subconjuntos de características vão acabar convergindo para as características originais ou correlacionadas a elas. A adição de novas características seria considerado um trabalho desnecessário, algo como criar um problema que não existe com o objetivo de aplicar a solução escolhida.

## Referências

- [1] Kuncheva L. I. and Rodríguez J. J., An Experimental Study on Rotation Forest Ensembles, Springer, 2007.
- [2] Lai, Reinders and Wessels, Random Subspace Method for multivariate feature selection, Pattern Recognition Letters, Elsevier, 2006.