

Análise Comparativa Entre os Classificadores MLP e Random Forest sob a Perspectiva de um Teste Bayesiano

Henrique Luiz Voni Giuliani, Ricardo Suyama, Denis Gustavo Fantinato
Universidade Federal do ABC (UFABC)
Santo André - SP, Brasil 09210-580
Email: {henrique.voni, ricardo.suyama, denis.fantinato}@ufabc.edu.br

I. BACKGROUND, MOTIVAÇÃO E OBJETIVOS

Os testes de abordagem bayesiana representam uma proposta alternativa que visa suprir as deficiências dos testes de hipótese nula e da interpretação dos p -values [1], [2]. O teste bayesiano leva em consideração o desempenho dos classificadores sobre determinados conjuntos de dados (informação *a priori*) para – dado um novo conjunto de dados – calcular (i) a probabilidade do classificador A ser melhor que o classificador B [$p(A)$], (ii) a probabilidade do classificador B ser melhor que o classificador A [$p(B)$] e (iii) a probabilidade dos classificadores possuírem uma equivalência prática de desempenho [$p(ropo)$]. Sob essa perspectiva, o presente trabalho analisa o desempenho de dois classificadores: *Multilayer Perceptron* (MLP) e Random Forests.

II. CONTRIBUIÇÕES E MÉTODOS

Neste projeto, os classificadores MLP e Random Forest são comparados à luz do teste bayesiano utilizando 5 diferentes bases de dados: (i) Heart-disease; (ii) Breast Cancer; (iii) Iris; (iv) Glass e (v) Mushroom. Para o teste bayesiano, assume-se que os dados de cada dataset possuem distribuição gaussiana.

III. RESULTADOS, DISCUSSÃO E CONCLUSÕES

Para cada dataset, utilizou-se a técnica de *Cross-Validation* em k -folds, sendo considerado aqui $k = 50$ folds. Os parâmetros estruturais dos classificadores MLP e Random Forest foram assumidos (i) 10, 20 e 40 neurônios em uma camada para a MLP e (ii) 10, 20 e 40 árvores de decisão como classificadores-base para a Random Forest. A configuração selecionada em cada dataset foi aquela com maior valor médio da métrica *f1-score*.

A Fig. 1 ilustra, por dispersão, a distribuição das probabilidades. O resultado obtido ressalta que as Random

Forests tem um considerável ganho de desempenho em relação à MLP, graças a seus elementos aleatórios. A MLP, por sua vez, não tem desempenho distante (note na Fig. 1 que a densidade se aproxima do ponto central), mas exige que os dados a ela apresentados possuam informações mais relevantes.

Consideramos, por fim, a análise bayesiana promissora na comparação dos classificadores. Experimentos com maior número de datasets serão considerados para uma análise futura mais profunda.

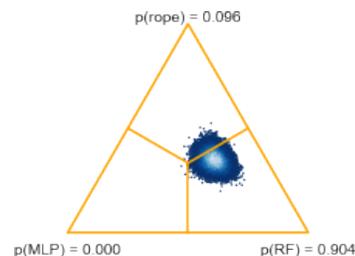


Figura 1. Distribuição de probabilidades de desempenho

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

REFERÊNCIAS

- [1] A. Benavoli, G. Corani, J. Demsar, and M. Zaffalon, “Time for a change: A tutorial for comparing multiple classifiers through bayesian analysis,” *Journal of Machine Learning Research*, vol. 18, 06 2016.
- [2] G. Corani, A. Benavoli, J. Demšar, F. Mangili, and M. Zaffalon, “Statistical comparison of classifiers through bayesian hierarchical modelling,” *Machine Learning*, vol. 106, no. 11, pp. 1817–1837, Nov 2017. [Online]. Available: <https://doi.org/10.1007/s10994-017-5641-9>