

Artificial Neural Networks and Ranking Approach for Probe Selection and Classification of Microarray Data

CMPE 59H - BIOINFORMATICS
BUSE BUZ
ONUR POYRAZ

Introduction

- ▶ We already know that there is a strong relationship between gene structure and some cancer types.
- ▶ In this study, gene expression values come from cancer patients with either acute lymphocytic leukemia (ALL) or acute myeloid leukemia (AML).
- ▶ A dataset which used to classify the type of cancer based on measurements of patients' gene expressions is provided.

Methods:

What they did:

- ▶ Feature selection for finding more related genes
 - ▶ GINI : Measures the inequality among features of a frequency distribution –best for in our case
 - ▶ Rank features
 - ▶ Fisher scoring
 - ▶ Kruskall
 - ▶ Stepwise regression :
- ▶ Artificial Neural Network for classification

Methods:

What we did:

- ▶ We are going to use these models and compare all of them,
 - ▶ PCA to reduce the dimension of feature set
 - ▶ K means clustering
 - ▶ Bayesian optimization
 - ▶ Gaussian process
 - ▶ Multilayer Perceptron
 - ▶ Neural Network
 - ▶ SVMs

Evaluation

- ▶ Accuracy_kmeans: 44%
- ▶ Accuracy_bayes: 44%
- ▶ Accuracy_gaussian: 58%
- ▶ Accuracy_SVM: 58%
- ▶ Accuracy_neural: 64%
- ▶ Accuracy_multilayer: 70%

Conclusion

Predicted/True	ALL	AML
ALL	20	0
AML	12	2

$$\text{Precision} = 20/20 = 1$$

$$\text{Recall} = 20/32 = 0.625$$

$$\text{F-measure} = 0.76$$

References

- ▶ Artificial Neural Networks and Ranking Approach for Probe Selection and Classification of Microarray Data
AW Faria et al.
2013 BRICS Congress on Computational Intelligence & 11th Brazilian Congress on Computational Intelligence
- ▶ Monitoring Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression
T. R. Golub et al.
Science DOI: 10.1126/science.286.5439.531