

## Identifying common and specific genes in field of cancer

D. C. Macedo<sup>1</sup>, E.C.M.Ishikawa, C.B. Santos, S.N.Matos, H.B.Borges and A.C.Francisco

<sup>1</sup>Administration Department, State University of Centro-Oeste, Brazil

Identifying common and specific genes in field of cancer is a hard task. Thus, was developed a method to identify a common and specific genes using dimensionality reduction through DRM-F method. This research made the comparison of two methods of reduction: the Method of Attributes Selection and the Method based on Framework, called DRM-F. This comparison aimed to evaluate the proposed method with the existing method in data mining, Attribute Selection. The DRM-F method based on Framework the adapted Ben-Abdhalal et al method was used. (2004). These two methods were applied in the field of gene expression and three bases were also used, as follows DLCBL, DLCBL tumor about leukemia ALL / AML containing lymphoma data. These sets have already been used by Borges and Nievola in their studies. The three sets of data were extracted from the analysis from the biomedical data repository Kent Ridge. Analyzing the results obtained, using as criteria for evaluating the Cross Validation, where it was found that the use of the methods resulted in an improvement in the values of accuracy rate when compared with the bases possessing all the attributes in the domain of gene expression. In this domain the best reduction method was using the Wrapper approach in the three bases. Nevertheless, it is noteworthy that the proposed method outperformed the 80% rate of accuracy results, which may not be considered a reduction method with poor performance. The results were approximate to the original base set containing all the attributes. Taking into account that the search criterion of the proposed method is based on the identification of common and specific attributes among the analyzed bases of the chosen domain. The proposed method aims to search the equivalence and generalization of attributes in the study domain. By means of effective results for the proposed (DRM-F) method in generation of predictive models it was decided to make a biological analysis for the common attributes (common Entrez IDs) identified among the three bases to seek insights or even new discoveries and biological meanings of these attributes. Thus, the common Entrez IDs among the three bases DLBCL, DLBCL - Tumor and ALL / AML were subjected to an analysis.

### Biography:

Dayana Carla de Macedo received the Technologist Degree in Food Technology from Federal Technological University of Parana, Ponta Grossa in 2009, the Administrator Degree in Administration from State University of Ponta Grossa in 2010, Paraná and the Master Degree in Production Engineering from Federal Technological University of Parana, Ponta Grossa in 2012. Doctoral Degree in Production Engineering. Her current research interest are data mining and framework in cancer field.