

Classifying Disease Expression Profiles using Networks

Ofer Lavi, Gideon Dror and Ron Shamir

1 - School of Computer Science, Tel-Aviv University, Tel-Aviv, Israel,

2 - Department of Computer Science, Tel Aviv College

Analysis of gene expression using microarrays went into clinical use in the recent years after FDA approval for prediction of effectiveness of chemotherapy against various cancers. Prediction methods that are based on a computationally selected small set of genes show significantly higher success rates compared to previous, more traditional methods.

However, these methods have drawn critique from both biomedical and computational scientists. For the life-scientists, the fact that treatment decisions are made in a “black box” fashion, treating genes as numbers, and lacking any biological and cellular rationale, is dissatisfying. Additionally, from a computational perspective, these methods suffer from shortage of training data: thousands of expression values were measured for each patient, but only in a few hundred patients.

We are using additional information to overcome the shortcomings of the current methods for selecting diagnostic markers. By collecting published and experimental information, many biological systems can be summarized as signaling and interactions networks. Exploiting the finding that close genes in the network tend to have a similar contribution to learned classification models, we combine expression profiles of cases and controls with large scale protein-interaction networks in order to improve classification results. Our method is based on using SVM with a kernel tailored to integrate expression as well as network topology data. We demonstrate the newly derived method on a number of case-control studies from the literature.