

CASE STUDY

Data Mining & Analysis



Overview

We have a strong experience in the practice of Data Mining -- discovering useful patterns in data and successfully harnessing the information gained with Research Companies. We also have experience in active researchers in Data Mining, and providing training to research staff.

Leading Clients

Literature Mining using Probabilistic Latent Semantic Analysis (PLSA)

Customer: Johnson & Johnson, New Jersey

We implemented a distributed computing solution using Matlab's distributed computing capabilities on a matrix of terms counts in documents. Matlab along with the high performance computing environment formed the back end, with all the literature data in a SQL server database, all being driven by an ASP.net website. We not only got better results than the customer's existing method, we also reduced the time greatly, from several hours or days per search to less than 15 minutes.

Array CGH Data Analysis Customer: National Human Genome Research Institute	Gene Expression Analysis Customer: Brigham's Women Hospital
We implemented Matlab programs for loading custom microarray data, preprocessing it, and created chromosomal and segmentation plots. The need for five different applications was eliminated with the single integrated solution - Matlab.	We analyzed gene expression data from three different projects involving cell lines given two distinct treatments. The entire analysis was done using a combination of Dchip and Matlab.

Currently, we are working with customers on projects involving mass spectrometry, high throughput screening, 3-d image analysis, wavelet analysis, distributed computing, and pk/pd modeling. The core areas of Tanisha Systems are training and consulting. We give our customers the option to "do-it-yourself" after training in our topics, or get the tools built by our contract consultants and get trained later in the use of those tools, along with the underlying technologies