High Throughput Screening Optimization

High Throughput Screening is the process of examining a large number of possible anti-bacterial chemicals at the same time. The process can screen up to forty-five chemical agents simultaneously. Unfortunately, the process for each plate of chemicals is very time consuming. If we could find better possible anti-bacterial agents, then the discovery of effective bactericides would be significantly accelerated. The high throughput screening process rests on our ability to detect and quantify complex patterns that emerge in very large databases. This project explores a family of effective chemical compounds related by certain properties to find the set of underlying patterns that might make good predictors of bactericide effectiveness in other chemical compounds. Using pattern analysis and complex search methods, this project has general application to problems involving elements with a differing number of features as well as elements that exist in three dimensional or even a multi-dimensional space.

We decided to apply some fuzzy data mining techniques to a large Oracle database of previous high throughput screening results. The results of this data mining were clusters of molecule descriptions centered on their effectiveness values. The distance from the centroid of the cluster was a measure of the molecule as a possible candidate.

The total project cost $385,000. In the first month of production, the savings were $1.3M in clinical research time with an addition reduction of over seven months of exploratory time. These savings came from two areas: the large reduction in clinical research time, and added efficiencies during the analysis phase. Clinical research time can cost upwards of $1100/hr, so minimizing this quickly increases savings. By creating a screening system that was much more effective and efficient than its previous version, the clinical researchers were able to quickly isolate over twenty compounds that proved effective antibodies.