



**CALIFORNIA STATE SCIENCE FAIR
2004 PROJECT SUMMARY**

Name(s) Tony Wu	Project Number S1223
Project Title A Category Oriented Web Search Engine Based on Round Robin Learning and Ranking Algorithm	
<p style="text-align: center;">Abstract</p> <p>Objectives/Goals I hoped to construct a category oriented web search engine through Round Robin learning and ranking algorithm. Such a search engine was designed to classify and rank the Web pages efficiently and to produce more effective search results than existing search engines.</p> <p>Methods/Materials More than 40,000 Web pages were loaded into a group of databases and stored as the URLs and the hits of terms in these Web pages. 500 URLs of these Web pages were indexed by the aggregate measure of subject and related keywords as training and testing data, which were used for calculation of the optimal decision boundary and Euclidean distance. A program was developed based on the algorithms, which allows users to automatically classify and rank the stored Web pages according their search queries and their selection of category.</p> <p>Results The search engine has shown the effectiveness in categorizing search results and ranking the relevance of returned Web pages to a search query through sufficient experimental data. Experimental data also demonstrated that the data structure based on the new classification and ranking algorithm has resulted in satisfying system performance with little cost in terms of data storage space and search speed.</p> <p>Conclusions/Discussion Through Round Robin Learning, a category oriented search engine can be constructed technically and economically. The study has shown the effectiveness of the search engine at an inexpressive cost compared with existing commercial search using link-based algorithms. Such a category oriented search engine has potential in uses of hunting terrorists and academic research. I plan to expand this project by adding more categories and testing the search engine under other conditions such as varying the user groups and the volume of Web pages.</p>	
Summary Statement The project presents new classification and ranking algorithms for building a category oriented search engine based on Round Robin Learning approach.	
Help Received First and foremost, I would like to thank Professor Jeffrey D. Ullman, Professor Gio Wiederhold, and Dr. Jan Jannink in Computer Science Department, Stanford University for their advice. Additional thanks to Mr. Tinh Tran and Mr. Robert Ferazzi for proof-reading my paper.	