Data mining technology created at NDSU licensed to Treeminer, Inc.

Published: 09 February 2011 09:16 AM

A novel vertical data mining method developed at NDSU offers significant accuracy and scalability advantages over current methods. University Distinguished Professor William Perrizo and his team have developed a technology tool to help government and businesses quickly process massive data sets. Treeminer, Inc., announced today that it has concluded a license agreement with the NDSU Research Foundation. The agreement gives Treeminer exclusive rights to further develop, market and sell the patented, award-winning data mining solutions. Treeminer will begin select demonstrations of the technology in the first quarter of 2011.

The amount of data available to businesses and governments is growing far faster than their ability to analyze the information. Significant advantage can be gained by being able to quickly make sense of millions or even billions of pieces of data, and applying the resulting knowledge. By organizing data vertically and then compressing it into a patented data structure called a pTree (predicate tree), dramatic reductions in analysis times can be gained over existing methods. Applications for the data mining technology based on pTree algorithms range from defense and intelligence to satellite image analysis, agriculture, computer network security, medical diagnostics, bioinformatics, resource allocation and many more.

"Today, data mining vendors are asking their customers to make trade-offs between the speed of analysis and its accuracy," observed Mark Silverman, CEO of Treeminer, Inc. "The truly novel approach taken by the team at NDSU enables incredibly dramatic decreases in analysis time while actually improving the accuracy of the analysis. We think that data mining technology will be a critical, fundamental building block technology across the information technology spectrum, and have formed the world's first Vertical Data Mining Company to bring this technology to a large and growing market."

Perrizo, distinguished professor of computer science, developed the patented algorithms and software on which the technology is based. "In the information science sphere, new approaches can sometimes effect increases in both the speed and accuracy of knowledge discovery. The pTree technology is an example of that," said Perrizo.

Representatives from the NDSU Research Foundation introduced Perrizo's work in data mining to Silverman, the entrepreneur who founded Treeminer. "Online analysis of large databases demands analysis that happens in seconds instead of hours," according to Silverman.

The technology developed by Perrizo and his team represents approximately a 15-year effort in data mining research. "Efforts such as Perrizo's illustrate the expertise available at NDSU that contributes to the body of knowledge in many areas, and contributes to the strength and vitality of state and national economic interests," said NDSU President Dean L. Bresciani. "We congratulate Bill and his team on reaching this milestone."

With the technology licensed to Treeminer, Inc., data mining operations on the pTree structure generated from large databases means data analysis becomes blazingly fast. The algorithms and software patented by Perrizo efficiently, accurately and elegantly mine the data for useful information, like an archeologist sifting through the dust to uncover hidden treasures.

The data is turned on its head, sliced up to change it into a vertical structure resulting in long, skinny pieces of data that are then compressed into a pTree to better manage it. "That wouldn't be all that useful if you had to uncompress every time you wanted to process the information. But we don't. We can process the compressed pTrees," said Perrizo. This approach enabled Perrizo to win the prestigious 2006 Knowledge Discovery and Data Mining Cup in solving what had been once characterized as the "Holy Grail" of Computer Aided Detection in medicine. Perrizo's team on software development for pTree technology includes Greg Wettstein, chief computational scientist and principal engineer for Research Computing Services at NDSU. "Dr. Wettstein is one of the best systems programmers in the world today and we are fortunate to have him on this team," said Perrizo.

"The coordinated efforts among NDSU researchers, the university's Technology Transfer Office and the NDSU Research Foundation help lay the groundwork for commercialization of discoveries developed at NDSU," noted Philip Boudjouk, vice president for Research, Creative Activities and Technology Transfer.

The NDSU Research Foundation's technology and licensing income from NDSU research discoveries has grown from $1.20 million in fiscal year 2006 to $1.88 million in fiscal year 2010. NDSU inventors, colleges and departments actively involved in developing innovations share in net revenue distributed by the Research Foundation.

"We are thrilled to see Dr. Perrizo's important work reach the market," said Dale Zetocha, executive director of the NDSU Research Foundation. "It represents a great opportunity to commercialize this research."