Aims

Sports data mining has experienced rapid growth in recent years. Beginning with fantasy league players and sporting enthusiasts seeking an edge in predictions, tools and techniques began to be developed to better measure both player and team performance. These new methods of performance measurement are starting to get the attention of major sports franchises including baseball’s Boston Red Sox and Oakland Athletics as well as soccer’s AC Milan.

Before the advent of data mining, sports organizations relied almost exclusively on human expertise. It was believed that domain experts (coaches, managers and scouts) could effectively convert their collected data into usable knowledge. As the different types of data collected grew in scope, these organizations sought to find more practical methods to make sense of what they had. This led first to the addition of in-house statisticians to create better measures of performance and better decision-making criteria. The second step was to find more practical methods to extract valuable knowledge using data mining techniques. Sports organizations are sitting on a wealth of data and need ways to harness it.

This monograph will highlight current measurement inadequacies and showcase techniques to make better usage of collected data. Properly leveraging sports data mining techniques can result in better team performance by matching players to certain situations, identifying individual player contribution, evaluating the tendencies of opposition, and exploiting any weaknesses.

Audience

The primary audience for this monograph includes the following:

- **IT Academic Audience**: College professors, research scientists, graduate students, and select undergraduate juniors and seniors in computer science,
information management, information science and other related public policy disciplines who are interested in data mining and its applications in various emerging technology fields.

- **Sports Management Academic Audience**: College professors, research scientists, graduate students and select undergraduate juniors and seniors in sports education and management related fields who are interested in an overview of sports measurement techniques and application to the sports environment.

- **Sporting Industry Audience**: Executives, managers, analysts and researchers in the business of sports, research institutions that are actively conducting sports data mining research and industry analysts who are interested in identifying critical inventions and innovations that can lead to major commercial successes in the industry.

- **Sports Enthusiasts Audience**: Individuals and sports-related organizations that want to learn how to gain an edge using modern data mining tools and techniques to uncover hidden knowledge.

**Future**

Over the next several years, sports data mining practices will be faced with several challenges and obstacles. The most obvious of which is to overcome the years of resistance by the members of sporting organizations that would rather stick with a traditional way of doing things. Aside from the challenges that are faced, sports data mining currently sits at a pivotal junction in history with many opportunities just waiting to be grabbed. Some avenues of opportunity will be pursued quickly, while others may take years or decades to become fruitful. In any event, sports data mining today is still in its infancy. While some first steps were made with pioneers such as Dean Oliver and Bill James, the next few years will become a transition period as the technology begins to mature within the sporting community and become more commonplace. New metrics, algorithms and ways of thinking will begin circulating themselves as the field enters puberty and begins to mature. The coming decades will be fascinating to watch.

Robert P. Schumaker
Osama K. Solieman
Hsinchun Chen

Cleveland, OH
Tucson, AZ
Tucson, AZ
Sports Data Mining
Schumaker, R.P.; Solieman, O.K.; Chen, H.
2010, XIV, 138 p., Hardcover