Local Binary Patterns-Based Statistical Data Analysis for Accurate Soccer Match Prediction

Authors : Mohammad Ghahramani, Fahimeh Saei Manesh

Abstract : Winning a soccer game is based on thorough and deep analysis of the ongoing match. On the other hand, giant gambling companies are in vital need of such analysis to reduce their loss against their customers. In this research work, we perform deep, real-time analysis on every soccer match around the world that distinguishes our work from others by focusing on particular seasons, teams and partial analytics. Our contributions are presented in the platform called "Analyst Masters." First, we introduce various sources of information available for soccer analysis for teams around the world that helped us record live statistical data and information from more than 50,000 soccer matches a year. Our second and main contribution is to introduce our proposed in-play performance evaluation. The third contribution is developing new features from stable soccer matches. The statistics of soccer matches and their odds before and in-play are considered in the image format versus time including the halftime. Local Binary patterns, (LBP) is then employed to extract features from the image. Our analyses reveal incredibly interesting features and rules if a soccer match has reached enough stability. For example, our "8-minute rule" implies if 'Team A' scores a goal and can maintain the result for at least 8 minutes then the match would end in their favor in a stable match. We could also make accurate predictions before the match of scoring less/more than 2.5 goals. We benefit from the Gradient Boosting Trees, GBT, to extract highly related features. Once the features are selected from this pool of data, the Decision trees decide if the match is stable. A stable match is then passed to a post-processing stage to check its properties such as betters' and punters' behavior and its statistical data to issue the prediction. The proposed method was trained using 140,000 soccer matches and tested on more than 100,000 samples achieving 98% accuracy to select stable matches. Our database from 240,000 matches shows that one can get over 20% betting profit per month using Analyst Masters. Such consistent profit outperforms human experts and shows the inefficiency of the betting market. Top soccer tipsters achieve 50% accuracy and 8% monthly profit in average only on regional matches. Both our collected database of more than 240,000 soccer matches from 2012 and our algorithm would greatly benefit coaches and punters to get accurate analysis.

Keywords : soccer, analytics, machine learning, database
Conference Title : ICMLC 2017 : International Conference on Machine Learning and Computing
Conference Location : Vancouver, Canada
Conference Dates : August 07-08, 2017