

Batting Order Setup in One Day International Cricket

Masoumeh Izadi, Simranjeet Narula

Television Content Analytics, 1 Changi North Street 1, Singapore 498789
masoumeh.izadi@tvconal.com, simran.narula@tvconal.com

Abstract

In the professional sport of cricket, batting order assignment is of significant interest and importance to coaches, players, and fans as an influencing parameter on the game outcome. The impact of bat-ting order on scoring runs is widely known and managers are often judged based on their perceived weakness or strength in setting the batting order. In practice, a combination of experts' intuitions plus a few descriptive and sometimes conflicting performance statistics are used to assign an order to the batters in a team line-up before the games and in player replacement due to injuries during the games. In this paper, we propose the use of learning methods in automatic line-up order assignment based on several measures of performance and historical data. We discuss the importance of this problem in designing a winning strategy for cricket teams and the challenges this application introduces to the community and the currently existing approaches in AI.

Introduction

The main duty of a batsman is to score runs at a good rate while conserving his position not to get dismissed. Crafting an optimal batting line-up from a team squad that takes advantage of every player's potential in scoring runs is a hard problem. In practice, a combination of expert intuition and descriptive performance statistics such as batting average, high score, not outs, strike rate, and run rate are used to form the team line-up for a match. Among these statistics, the most widely used one is the player's batting average, which is computed as the total number of runs divided by times dismissed. In test cricket, where there are practically unlimited overs, a batting average approximates expected runs. Therefore, in test cricket, it is somewhat reasonable to use the batting average to assign bat-ting order. However, in the One-Day International (ODI) format, as the game progresses and the number of balls left decreases, it is important to use batsmen who has higher run rate earlier in the line-up. In ODI cricket, the run rate and the av-

erage are sometimes conflicting and can lead to different batting order (Beaudoin and Swartz 2003). Therefore, captains and team managers sometimes struggle to find the best place to assign players in the line-up, especially if the players are new (from other formats). It is also hard to evaluate the captain's policy on the line up versus alternative strategies without a precise model.

Advances in technologies for data collection provide new data sources that might be correlated with players' performance and the game outcome. However, the analytics infrastructure which is instrumental in using these data to support important decisions in cricket is still fragmented. In this paper, we present a new AI application for enhancing experts' intuitions on line-up configuration considering different and possibly conflicting metrics with the aid of learning algorithms and historical data. In this application, we consider player's portfolio as suggested in the literature (Barr and Kantor), plus other parameters that might be important in batting performance.

Background

Cricket has a long history and is the world's second most popular sport with an estimate of 2.5 billion fans. Cricket is played between two teams of 11 players on a 22-yard-long pitch. A game can last anywhere from one afternoon to 4 or 5 days. In one-day matches, each team bats only in one innings with 50 overs. Each over has a set of six ball deliveries. Sometimes more than 300 balls may be delivered in one ODI innings due to penalties. A coin flip prior to the game decides which team has the option to bat first. Bowlers from the fielding side bowl balls with alternate overs bowled by different bowlers from opposite ends. Batsmen play in pairs, one on strike facing the bowler and one at the bowler's end. After hitting the ball, a striker may score from 0 to 6. A batsman can be dismissed in various ways. A team's innings ends when 10 batsmen are out or they bat-tered for 50 overs. The batting team attempt to score runs, while the bowling team tries to end the inning for the bat-ting team through the process of taking wickets (dis-

missing the striking batsman). The team with the most runs or higher number of wickets left wins the game. There are many rules related to batting, bowling, and fielding restrictions that are particularly designed for ODIs.

Modelling batters expected score is fundamental to understanding the effect of cricket batting order. Researchers have explored using dynamic programming methods in estimating a player's expected runs (Norman & Clarke 2010, Preston & Thomas 2000). These studies hoped that the association between batter's order and expected score could be used to optimize the total score and to develop ranking systems. The quantification of the impact of a specific batting order is hard in the absence of a complete model of cricket games. To find the line-up that results in the best expected total score, a simulation model of a real game is needed. However, it is still intractable to compute the expected score for all 11! order combinations. Ovens and Bukiet considered an approach using a Markov chain for a subset of 163,324 orders, where only three or four batsmen in the lineup change position. Swartz et al. also studied simulated annealing to reduce number of choices for batting strategies, however, strong assumptions limits the applicability of this approach. Batsman-specific models have been suggested to gain insight on the best order for a batsman (Koulis et.al, 2014). However, this still doesn't solve the complete line-up setup for a squad.

Data and Challenges

The data contains the scorecards and player data of all ODI matches from 1971 to present. We only considered the matches played in the last 15 years to extract players who are more recent and have played continuously for at least two years. The collective data consists of records from 1242 matches that includes both male and female cricket matches and from 1654 players (58.52% male). To see the variation of batting order in each individual player in our data, we assigned the average of these orders (rounded to the closest integer) to the player and we call that number the player's "usual order". We observed that some players typically stay within their usual class, while some moved substantially between classes. We considered 30 correlated features from the pool of 86 features on batting and bowling. To build a winning lineup configuration from a fixed set of team's squads one faces the following challenges.

Appropriate Batting Metrics

Currently used metrics such as batting average, economy rate, and strike rate do not adequately capture the richness of the underlying dynamics in ODI. Advanced metrics that indicate the batting pressure, consistency over time and over the batter's career, and player's contribution into team's success should be properly defined and incorpo-

rated in models. In preliminary investigations, we considered the average run contribution (proportion of the team's runs scored by the player) to evaluate player's performance when a player played within his usual class or out-side that class. We observed some players do normally better outside their usual class. Such observations might imply suboptimal assignments and must be investigated.

Player Prognosis

The expected number of balls faced and runs to be scored for the batters need to be estimated. Models built on historical observations needed to compute the expected deliveries before dismissal for each striking batsman in a lineup considering the opponent team and its current bowler, the non-striking partner, the cricket ground, and the remaining resources such as wickets, balls, and other batting metrics. Such prognoses are necessary to build a simulation model for the entire game and use that to assess different lineups.

Batting Order Optimization

Insights about the expected team outcomes with respect to the total runs scored and the wickets left over a fixed horizon of 300 ball deliveries can help to develop optimization strategies. We believe generative models can shed light on the expected outcomes for a team and players with respect to the choice of batting order. The challenge remains in building transition models that capture the team's status changes after each player selection to bat. It should be noted that sometimes there will be players who are new to ODI format with not enough history. Such players should be mapped to the most similar players and the generative models should be generalized to such players.

References

- Barr GDI. And Kantor BS. A criterion for comparing and selecting batsmen in limited overs cricket. In *Journal of the Operations Research Society* 55:1266-1274
- Beaudoin D. and Swartz T.B. 2003. The best batsmen and bowlers in one-day cricket. In *South African Statistical Journal* 2003, 37:203-22.
- Koulis T., Muthukumarana S., Briercliffe C. 2014. A Bayesian stochastic model for batting performance evaluation in one-day cricket. *Journal of Quantitative Analysis in Sports* 10(1): 1-13.
- Norman J.M. and Clarke S.R. 2010. Optimal batting orders in cricket. *Journal of the Operational Research Society* 61:980-986
- Ovens M. and Bukiet B. 2006. A mathematical model approach to one-day cricket batting orders. *Journal of Sports Science and Medicine* 5: 495-502
- Preston I. and Thomas J. 2000. Batting strategy in limited overs cricket. *The Statistician* 49: 95-106
- Swartz T.B., Gill P.S., Beaudoin D. and de Silva B.M. 2006. Optimal batting orders in one-day cricket. *Computers & Operations Research* 33:1939-1950