Data-driven rating system makes it easier to select sports teams
15 January 2021

Picking the right sports team for a particular event or to play in certain conditions is many a selector's nightmare.

But research by Te Herenga Waka—Victoria University of Wellington Ph.D. graduate Dr. Ankit Patel could make their tricky task easier.

Ankit, who received his Ph.D. in statistics last month, has crunched many numbers and invented a framework for calculating sport-based player ratings.

He says the rating systems need to be "meaningful"—i.e. their evaluations of team and player performances are reliable, robust, intuitive, and transparent.

"The idea is it can effectively predict match outcomes and develop team-selection strategies," says Dr. Patel. "This work is currently being applied across multiple sporting codes, but specifically cricket and rugby."

A lot of the detail of the system and who is using it has to remain confidential, given the potential for abuse by gambling syndicates and the commercial sensitivity of the research.

The intellectual property for the system is owned by DOT Loves Data, whose founder and chief operating officer, Dr. Paul Bracewell, was one of Dr. Patel's supervisors and is an adjunct researcher in the University's School of Mathematics and Statistics.

Dr. Patel says one of the challenges in developing the system was the number of parameters involved that affect a player's performance, and how these have varying impacts at different stages of any game.

"So, for example, in cricket, if someone hits a six at the beginning in the first five overs of the first innings, their ratings might go up by two. But if they hit a six on the second-to-last ball of the second innings, and they need seven runs to win, it could go up by 20.

"The system's ability to map back to real-world outcomes and account for match conditions and context is extremely important. Therefore, the key is the ability for the system to produce ratings that are statistically reliable and robust, but also interpretable and intuitive."
As well as his Ph.D., Dr. Patel has a conjoint Bachelor of Commerce/Bachelor of Science in finance, marketing and statistics, a Master of Science in statistics, and a Master of Applied Statistics from the University.

He says there has been significant growth in demand for data-driven rating systems to assess performance in the past decade.

"This has been experienced across many industries, but is most evident within the sporting industry."

To be of use, his model "had to be robust, and yield good performance where data is drawn from a wide range of probability distributions that are largely unaffected by outlying figures, small departures from model assumptions, and small sample sizes.

"It had to be reliable, with ratings producing accurate and highly informative predictions that are well calibrated and transparent in terms of being interpretable and easy to communicate.

"Finally, it had to be intuitive, with ratings that had to relate to real-world observable outcomes and the context to which the system is being applied."

Dr. Patel's ensemble forecasting strategy was tested and validated by constructing both team and individual player-based rating systems within the cricketing world.

It builds on work he did several years ago when he presented papers at the University of the Sunshine Coast, Queensland, for the 14th Australasian Conference on Mathematics and Computers in Sport (Mathsport), part of the Australia–New Zealand Industrial and Applied Mathematics organisation.

His work then was recognised with the Neville de Mestre Best Student Paper and Presentation Award.

In his second paper, he constructed a method for improving estimates of the expected total in the first innings of a T20 game of cricket. This resolved issues relating to the context of the game that had previously affected prediction accuracy.

Dr. Patel has joined the firm Precision Data to help build its data science and advanced analytics capability.