Sports Analytics, Golf, and Gameforge: Innovative Analytics for Recommender Systems

The Analytics Revolution on the PGA Tour

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Systems Engineering

By
Samuel Roberts

November 1, 2021

Technical Team Members:
Rose Dennis
Zachary Kay
Rachel Kreitzer
Jerry Lu
Thomas Twomey
Steven Wasserman

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Hannah Rogers, Department of Engineering and Society

William Scherer, Department of Engineering Systems and Environment

Introduction

When most people think of athletes, golfers are not necessarily the first idea that will spring to mind. One reason for that also contributes to the sport's popularity: it can be played by anyone, and one's skill is not reliant on their physical athleticism. Many of the world's greatest golfers from the 1970s and 1980s more closely resembled a suburban dad than they did their counterparts in the NFL and NBA. However, the vast majority of PGA Tour golfers today are unmistakably athletic. This is one example of how the sport has evolved and grown as time has gone on. Another example of this is the idea of performance analytics. Metrics can be used by golfers both to help evaluate weak points in a player's game and to help predict how young players will grow and improve.

The metric revolution in sports over the past twenty years has certainly hit the PGA Tour, as advanced statistics are around every corner. With the advancement of metrics, professionals have been able to pinpoint what the strengths and weaknesses of their games are, and adjust accordingly (Arastey, 2020). Furthermore, a few golfers have completely reinvented their games in pursuit of the best chance of winning. The proposed STS project investigates the PGA Tour to see how the best players in the world have adopted analytics to evaluate and adapt to the times in recent years.

Golf is one of the last major sports where college recruiting analytics have not become commonplace; most coaches simply recruit through "the eye test." This process is incredibly subjective and prone to variance; there are many potentially fantastic golfers who are slipping through the cracks. A group of UV.a. students began to address this problem last year (Bassilios et al, 2021). The technical project detailed within will attempt to build & improve upon these

predictive analytics to identify the most skilled junior golfers who may be otherwise falling through the cracks in the recruiting process.

Technical Topic

GameForge is a small company that specializes in golf analytics. They are attempting to expand into the golf recruiting space, which is by and large unoccupied at this point in time. This technical project will expand upon the work done by a 2020-2021 capstone project working on the same problem (Bassilios et al, 2021). They created various models, which assisted in predicting whether a high school player would go to Division 1 (the highest level of college golf) and how they would affect the capability of a college team if added to the roster. This year, this group plans to implement three different improvements to the overall GameForge recruiting system.

The first improvement involves a college recommender system. This system predicts both the likelihood of a player choosing a particular school and the likelihood of that school to accept them. These things involve both qualitative and quantitative metrics. The predictors of a recruit choosing a particular school are largely qualitative; they include academic strength of a school, location, and other such values. The prediction of whether a particular school will accept them is much more metric-based; it will likely use the win-shares model from last year's team. Our final deliverable for this improvement will be two percentages: the likelihood that a recruit would be accepted to a college team, and the likelihood that a recruit would come to a school if accepted.

The second part of this project focuses on a junior golfer ranking system. Plenty of individual junior tours, such as the AJGA, have their own ranking system for players who play

their events (“Rolex AJGA Rankings”). However, there is not a simple ranking exclusively for junior golfers that uses each tour’s data. This group will use two main factors in making the ranking system. First of all, there will be a combination of many major golf rankings, such as the AJGA, GolfStat, and WAGR junior ranking systems. These metrics for present performance will be combined with statistics that last year’s group found that are predictive of future performance to create one overall ranking system that shows the junior players most primed to make an impact at the intercollegiate level.

The final part of this project looks at classifying players based on their strengths and weaknesses. The idea behind this phase is to show what types of holes players excel at in order to build a strong all-around team. To do this, the team will use metrics created last year that show players’ mean scores on holes of varying lengths and pars; the potential classifications are par 3s, short and long par 4s, and short and long par 5s. From there, players will be compared to the overall average scores and variance in scores to see where they truly shine. A player will be “tagged” if they are above average or exceptional in each area. This system will be used to evaluate teams as a whole and show coaches what kind of players they are missing. For example, if a particular team has no players who are above average in par 3 scoring, that could be an area for coaches to explore in future recruits.

STS Research Topic

The best male golfers analyzed in the technical portion will eventually rise to the professional level, where they will find it much harder to improve. At earlier levels, success is dependent on pure talent. However, at the PGA Tour level, all golfers are extremely talented; the

best of the best must maximize their skill set to best perform. The research question addressed with this STS research will be "How has the rise of analytics and metrics in golf changed how players train and prepare at the top level?" Within this question, there are two topics in particular that are worth exploring. One aspect is the new metric "strokes gained," and how it affects the way professional golfers evaluate their game as a whole. The other main topic is the omnipresent race to gain more distance on golf shots, and how players feel that improves their game. These two ideas for professional player improvement go hand-in-hand, as they were both brought to the forefront in the early 2000s, and players have been chasing them ever since.

The strokes gained metric is mainly only present on the professional level; this is because it is mainly reliant on comparing a player's performance to the average Tour player using precise GPS shot data called ShotLink, which was introduced in 2003 (Arastey, 2021). This system takes how you'd expect an average Tour player to play on a specific hole; if your player performs better than average on a specific shot, they get a certain number of strokes that they "gained" on the field for that shot ("Strokes gained: how it works"). For example, if a player hits an exceptionally long drive, that could be worth a few decimal points in strokes that they've "gained" in the long run. Players can use this to separate out strokes gained for separate aspects of the game; they could just look at your first shot on each hole to see the strokes gained off the tee, and they could isolate each shot on green to see the total strokes gained putting. Many pros use these stats to see in which areas of their game they excel, and where they need to improve.

Another recent aspect of the technological innovation of golf is the quest to gain distance on shots. It obviously gives a competitive advantage to golfers to hit the ball farther; if a player's ball is closer to the hole after their tee shot, it will be easier to get in the hole in fewer strokes.

Players have become more and more focused on increasing their swing speed, as a faster swing means a farther shot (Bowden, 2013). This has led to more and more players going to the gym and lifting weights to become stronger. This is a new development; back in the 1980s and 1990s, it was common for pro golfers to have the build of an average person, and some golfers who were quite overweight. That all changed with one of the most influential athletes of all time, Tiger Woods.

Woods is well known for his impact on the racial diversity of the PGA Tour and his overall dominance. One aspect of his game that is less remarked on is his commitment to personal fitness. He was one of the first golfers to commit to lifting weights and becoming stronger (McCormick, 2021). Today, many golfers are closer in resemblance to other professional athletes due to their workout regimens. Some golfers have taken this mission to increase distance to the extreme. One example of this would be Bryson DeChambeau. In late 2019 and early 2020, he gained 40 pounds of mass in an effort to gain distance on his tee shots (Harig, 2020). This change completely changed the way he played the game of golf; he now has the longest average driving distance on the PGA Tour by a decent margin ("Golf stats and records"). It's worth exploring if changes such as this one are useful in the long term for PGA pros.

Thomas S. Kuhn writes about paradigm shifts in his paper, "The Structure of Scientific Revolutions." In it, he describes how his graduate physics problem was radically changed by new scientific discoveries, and how these paradigm shifts can change the way you see the world at large. A similar thing has happened over the past few decades in golf. The research will use

this framework to show how the way in which golfers analyze the game and their own play has evolved as time has gone on and new paradigms come to light.

Conclusion

This thesis will combine two approaches to quantifying player success in the game of golf. The analytics revolution has been famously advanced in sports with more widespread popularity such as baseball, basketball, and football. Books such as “Moneyball” and sites such as Pro Football Focus have pushed the idea of sports metrics and analytics into the public eye. Golf, being a less-popular sport with plenty of traditions and taboos, has lagged behind in this sense. These analyses will help to show that golf can be quantified in a similar way and metrics can be created to measure player success.

References

Arastey, G. M. (2020, January 23). The increasing presence of data analytics in golf. Sport Performance Analysis. Retrieved November 18, 2021, from https://www.sportperformanceanalysis.com/article/increasing-presence-of-data-analytics-in-golf.

Bowden, J. (2016, March 6). *Carry distance vs. Swing Speed Chart*. GolfWRX. Retrieved November 18, 2021, from https://www.golfwrx.com/64715/carry-distance-vs-swing-speed-chart/.

Bassilios, M. (2021, July 16). *Developing a Recommendation System for Collegiate Golf Recruiting*. IEEE Xplore. Retrieved November 1, 2021, from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9483777&casa_token=5M0gJBxlKQYAAAAA%3AI7t307cgTc1zyK1V3fTTLoxoeOr4PBV5XbZtmPT8ekEpbmOE5qV1N3Ye2-oKKJVqTjKm_-OgP3o.

Broadie, M. N. (2014). *Every shot counts: Using the revolutionary strokes gained approach to improve your golf performance and strategy*. *Google Books*. Gotham. Retrieved October 31, 2021.

Christie, C. A., & Hetu, F. E. (1998, June 1). *Effects of conditioning on physical fitness and club head speed in mature golfers - fred E. hetu, Chris A. Christie, Avery D. Faigenbaum, 1998*. SAGE Journals. Retrieved November 1, 2021, from https://journals.sagepub.com/doi/abs/10.2466/pms.1998.86.3.811?casa_token=oD0iguYa

diQAAAAA%3AjKnC5G_KDn746BuUZSO0xlDO3qbELli86cMu5rMzS6uJMBc3XLs HPah5Y2zsyYNSkMocWEz-Rp3xPw&journalCode=pmsb.

Golf stat and records: PGA tour. PGATour. (n.d.). Retrieved November 1, 2021, from https://www.pgatour.com/stats.html.

Harig, B. (2020, June 19). *Bryson DeChambeau's eat whatever he wants, whenever he wants plan seems to be working*. ESPN. Retrieved November 1, 2021, from https://www.espn.com/golf/story/_/id/29335275/bryson-dechambeau-eat-wants-whenever-wants-plan-seems-working.

McCormick, J. (2021, March 9). *McCormick: Tiger Woods was major influence on fitness in golf*. Florida Today. Retrieved December 18, 2021, from https://www.floridatoday.com/story/life/wellness/2021/03/09/mccormick-tiger-woods-major-influence-fitness-golf/4627236001/.

PGA Tour. (2016, June 3). *Strokes gained: How it works*. PGA Tour. Retrieved November 1, 2021, from https://www.pgatour.com/news/2016/05/31/strokes-gained-defined.html.

Rolex AJGA Rankings. AJGA. (n.d.). Retrieved November 1, 2021, from https://www.ajga.org/rankings.

Ryan, S. (2021, January 11). *The golf stats that matter most on the PGA tour*. GolfDigest.com. Retrieved November 18, 2021, from https://www.golfdigest.com/story/golf-stats-that-matter-most-which-skills-correlate-to-success-on-the-pga-tour.

Urbaczewski, A., & Elmore, R. (n.d.). *Loss aversion in professional golf*. SAGE Journals.

Retrieved November 1, 2021, from

https://journals.sagepub.com/doi/full/10.1177/1527002520967403?casa_token=jW48G5u

oLrEAAAAA%3ARbFxuDpCRCwIJRp3E-s7hMF7ijBBy9_7mybrcG2hvAfLil4bBGNa

PnSiSoEg-W3poBCfxoRoJXRkkQ