Prospectus

Recommender System for Golf Athletes
(Technical Topic)

Actor Network Theory and Recommender System Biases
(STS Topic)

By

Josh Barnard

11/20/20

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.

Signed: ____

Approved: ____ Date_______________
Ben Laugelli, Department of Engineering and Society

Approved: ____ Date_______________
William Scherer, Department of Engineering Systems and Environment

Sociotechnical Problem

In the US, there are millions of high school athletes wanting to play their sport at the collegiate level. Athletes in high revenue-generating sports, such as football or basketball, have intense tracking of players and their metrics, with plenty of models and game film for coaches to watch. However, these are not the only collegiate sports, and the others are are not given the same treatment from the recruiting community, due to their relative smaller size and difficulty to monetize the recruiting market. This does not mean that these sports have less athletes pursing collegiate or professional careers though. In the case of collegiate golf, there are tens of thousands of high school golfers vying to be in the 6% of golfers to play in college, and fewer receive scholarships (NCAA, 2020). Currently, the only recruiting process is through private recruiting consultants or to contacting the college programs themselves (NHSGA, 2020). We are proposing a recruitment tool that will pair prospective high school golfers with programs that would fit them well and would improve the program simultaneously. However, predictive recruiting tools come with a multitude of socio-technical issues, as I will demonstrate with a program that Amazon developed to improve the quality of their hires. We have to recognize the biases that were present in the data being used, prior to building the predictive model, else we risk continuing and further building these biases. In attempting to fully solve the golf recruiting issue, we must first fully understand both the social and technological aspects of the problem and proposed solution. In order to accomplish this, I will draw upon the Science, Technology, and Society framework of Actor Network Theory (ANT). Using ANT, we can better understand both the underpinned issues of using systems we do not fully comprehend to give recommendations we do not understand. In the past scholars have attempted to explain the failures of artificial intelligence based recommender system, focusing largely on biased data—but have missed some key aspects of the failure. When these systems have failed there has been a distinct lack of transparency and understanding of the systems being created, and an absence of human input in the recruiting loop. By showing that the problem here is socio-technical in nature, and gaining this understanding we can deliver a fuller solution to the problem. The technical side of the project will tackle the fact that this novel approach to golf recruiting will reduce the information frictions present in the current state of affairs. Alone this solution is inadequate, as the biases created by only using past data will likely trend towards pushing recommendations of affluent student athletes, a disparity we do not wish to continue—therefore it is necessary to explore and account for the societal aspects of the AI recommender system.

Technical Problem

College golf is a growing sport, with over 100 new teams in women's golf and 46 new men's teams being added in the NCAA over the last decade (Williams, 2019). Playing collegiately is important for aspiring professional golfers, especially with especially with movements such as PGA Tour University—a program allowing select college players direct access to the professional scene (Romine, 2020). However, the current state of college golf recruiting is an extremely fragmented system that has to abide by an ever-changing set of rules set forth by sports governing bodies, with policies being put into place that push the timetable of when players can commit to a school. This policy puts more financial and time strains on the coaches of many programs, where recruiting now has to occur simultaneously with early fall practices, orienting new players on their current team, and watching rankings of players they may want to recruit down the road. Some schools will have no problems with this as they have assistant coaches who will be able to bear some of this burden, giving a distinct advantage to schools willing to hire full time assistants for the team (Ryan, 2019). This process is also difficult from the student side, attempting to balance hopes of professional play with the realities of not making it to the tour and knowing the importance of academics (Cummings, 2018). Additionally, some need the assistance of scholarships, which have a set limit on the number of scholarships that can be given, at the D1 level only 4.5 for men and 6 for women (Drotar). The biggest challenge that students face is being known and followed by programs—a task placed on the athlete themselves, excluding top players (Richardson).

Currently, the large majority of connections that get made between colleges and athletes come from athletes. The closest thing that exists as an aid to both sides of recruiting is using either online services or consultants that will help players reach out to schools and have knowledge of the process and how to effectively go about it. There are very expensive personal consultants who generally have relationships with a program that they will be able to better pair players to as they get to know both sides relatively intimately (NHSGA, 2020). College coaches generally are limited to looking at ranking data such as AJGA or WAGR ranks. This ranking data is only high level—there is no way to discern what skills a player may need to develop or their ability in different course lengths. The current system therefore excludes a multitude of players who have the talent to play in college but are overlooked because they have limited exposure. The smaller programs are more likely to be absent from these recruiting consultants, and face additional difficulty finding these athletes with their smaller budget and resources.

There exists a massive information friction between players and coaches, and no wide spread technology to match otherwise unnoticed mid-level players to programs has been developed. We are proposing an athlete-program matching system that will give both sides recommendations, based on a number of developed models and both sides' preferences. This will create greater fit of players to the universities they will attempt to contact, as well as optimizing the portfolio of players in a college program. We will use predictive modeling and possibly machine learning to predict player success in college using publicly available data from AJGA and WAGR, with private metrics from our client, Gameforge. These data flows are shown below in Figure 1. To validate our design, we will attempt to show that there are players who could have had great success at the DI level, but were relegated to DII, DIII, or junior colleges.

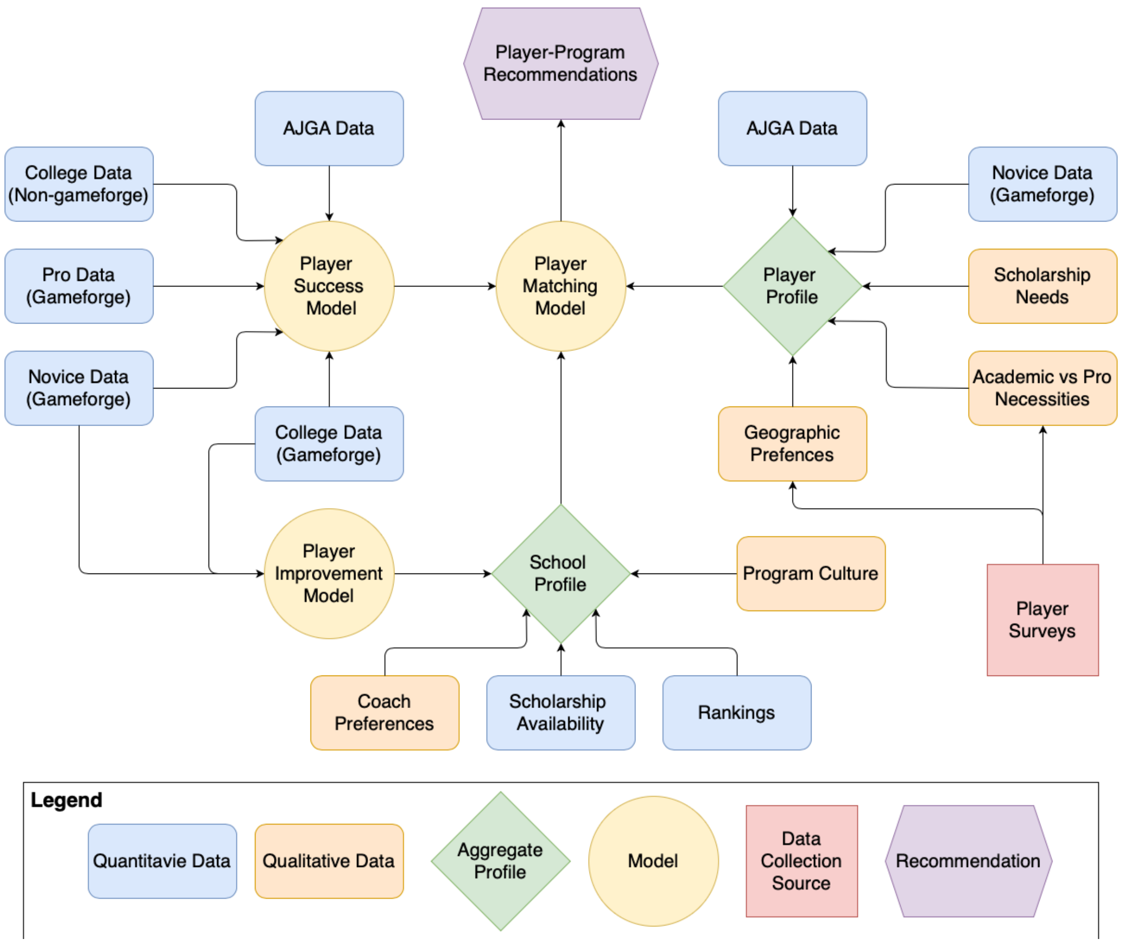

Figure 1: Data Flows for the Player-Program Recommender

STS Problem

With the growing complexity of the job market, and the introduction of the internet, the number of potential candidates for jobs have increased dramatically, especially for large companies like Amazon. To attempt to narrow the pipeline and only have qualified individuals apply, Amazon attempted to implement a machine learning based algorithm that would only present recruiters with applications it believed would be high performers in those jobs. This tool demonstrated significant biases—specifically against women—and often recommended unqualified candidates and was eventually scrapped by executives (Dastin 2018). With around 40% of US companies completely outsourcing their recruiting practices (Cappelli, 2019), and many still having difficulty finding “good” applicants despite low unemployment (Knowledge@Wharton, 2019), many have turned to AI based recruitment systems as a silver bullet for recruiting. However, as seen with Amazon, AI systems are not always extremely effective and can introduce more problems than they attempt to solve. (Bogen, 2019). This was the case when Amazon attempted to implement their own AI recruiting tool—when a realization that the system demonstrated extreme bias towards female applicants (Dastin, 2018). The failure of this tool has largely been attributed to the data that was fed into the machine learning algorithms— the resumes and skills of high performers at Amazon (Black & van Esch, 2020). While scholars do acknowledge the potential benefits of combining the use of traditional and AI recruitment tools—they fail to consider the complexity and opacity of the AI-tool leading to an obfuscation of the recruiter’s understanding of the system as an important additional factor in the downfall of the implementation of the recruitment tool. If we continue to only believe that the data fed into the algorithm was responsible for the failure, we will fail to account for unnecessary complications in the system, and will gain understanding of the importance of the recruitment tool’s interaction with the rest of the system.

To fully comprehend the downfall of this program, I will draw on Actor Network Theory (ANT), and argue that it was the not only biased data but the use of this data in conjunction with an unnecessary punctualization of the tool creating an overt opacity in the design and lack of human element in their recruiting that caused this program from Amazon to fail. Thomas

Cressman's overview of Actor Network Theory describes the way that networks are formed by a variety of actors by a network builder, the creator of the technology, and that some of these actors can be punctualizations, entire networks that because of their self-obviousness can be converted into a single actor in another network (Cressman, 2009). Additionally, in a network an actor can be rogue, destabilizing and changing the network. To support my argument, I will analyze the Amazon recruitment software failure, drawing on the works of scholars who have analyzed the issue previously, while using Actor Network Theory to add to analysis, arguing that punctualization in a network can result in that actor becoming a rouge actor.

Conclusion

In this paper both the technical and social efforts needed to implement a novel recruitment system will be explored. The technical solution will present a recommender system for high school golf athletes and college programs. This will be a highly technical system drawing on predictive modeling, AI, and machine learning. This heavily data driven approach needs to have a social element, which will consist of attempting to recognize and account for biases in our data early on. In the social solution we will explore the case of the AI recruitment tool built by Amazon, that failed to perform. I will first draw on other scholars work to explain the background and some of the identified factors in the failure of the system. Then I will draw on Actor Network Theory to more fully explain the extent of the failure—arguing the opacity of the was in fact a rouge actor. The combination of the two parts of this paper will inform the reader why both efforts are necessary and will aid in the solution of this socio-technical problem.

References

Black, J., & Van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? *Business Horizons,63*(2), 215-226. doi:https://doi.org/10.1016/j.bushor.2019.12.001

Cross, S., Braswell, P. (2019, December 10). A data-driven approach to hiring more diverse talent. *Harvard Business Review.* Retrieved from https://hbr.org/2019/12/why-isnt-your-organization-isnt-hiring-diverse-talent

Bogen, M. (2019, May 6). All the ways hiring algorithms can introduce bias. *Harvard Business Review.* Retrieved from https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias

Cappelli, P. (2019). Your approach to hiring is all wrong. *Harvard Business Review*: *May-June 2019 Issue.* Retrieved from https://hbr.org/2019/05/recruiting#your-approach-to-hiring-is-all-wrong

Cressman, D. (2009). A brief overview of actor-network theory: Punctualization, heterogeneous engineering & translation. *CPROST Catalogue Number: 09-01,* 1-12.

Cummings, N. (2018, November 01). Player blog: Must-ask questions of the recruiting process. *Amatuergolf.com* [Web log post]. Retrieved October 26, 2020, from https://www.amateurgolf.com/golf-tournament-news/22143/Player-blog--Must-ask-questions-of-the-recruiting-process

Drotar, B. (n.d.). The myth of the college golf scholarship. *Therecruitingcode.com* [Web log post]. Retrieved October 26, 2020, from https://therecruitingcode.com/the-myth-of-the-college-golf-scholarship/

Dastin, J. (2018, October 10). Amazon scraps secret ai recruiting tool that showed bias against women. *Reuters.* Retrieved from https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

Knowledge@Wharton (Producer). (2019, October 4). How companies can fix the hiring process. *Knowledge*@Wharton [Audio podcast]. Retrieved October 26, 2020, from https://knowledge.wharton.upenn.edu/article/fix-the-hiring-process/

National Collegiate Athletic Association (NCAA). (2020, April 16). Estimated probability of competing in college athletics. *NCAA.* Retrieved October 26, 2020, from http://www.ncaa.org/about/resources/research/estimated-probability-competing-college-athletics

National High School Golf Association (NHSGA). (2020, August 06). Looking to get recruited to play college golf?. *NHSGA.* Retrieved October 26, 2020 from https://www.highschoolgolf.org/high-school-golf-recruiting/

Richardson, T. (n.d.). How to get a golf scholarship - An expert's advice. *Nextgengolf.com* [Web log post]. Retrieved October 26, 2020, from https://blog.nextgengolf.org/college-golf/how-to-get-a-golf-scholarship-an-experts-advice

Romine, B. (2020, June 01). What is PGA Tour university? A detailed look at the Tour's new major. *Golf Channel.* Retrieved October 26, 2020, from https://www.golfchannel.com/news/what-pga-tour-university-detailed-look-tours-new-major

Ryan, B. (2019, May 01). NCAA Invokes new recruiting rules changes effective June 15. *Junior Golf Mag.* Retrieved October 26, 2020, from https://www.juniorgolfmag.net/blog/ncaa-invokes-new-recruiting-rules-changes-effective-june-15/

Williams, J. (2019, December 27). NCAA research shows major growth in number of women's college golf teams since 2009. *Golfweek.* Retrieved from https://golfweek.usatoday.com/2019/12/27/ncaa-research-womens-college-golf-teams-growth/