To: Izbassarova Bibigul

From: Akmarzhan Abylay

Subject: Using statistical tools to predict potential archers

Date: March 25, 2020

Executive Summary[1]

This decision memo proposes predicting high potential archers (HPA) given students' physical and mental fitness using simple statistical tools to improve the quality of the admission mechanism in an Akmola archery school in Nur-Sultan. Due to high demand in Nur-Sultan's archery club (i.e., there is only one) and the limited spot availability, coaches have to turn down some talented students. Instead, they accept students who later drop out because of the first-come-first-serve acceptance mechanism. My proposal will allow prioritizing higher potential students (i.e., both physical and mental fitness) and reduce the bias due to the time-dependent acceptance, which could improve the Akmola archery club's performance on regional and national competitions in future.

Problem: Lack of proper admission process in Akmola archery club

Archery is a highly competitive sport, and different clubs want to choose and focus on higher potential students to make it easier to train them and get results. The recruitment happens once a year, and there are usually 10-15 open spots, while there are around 50-60 interested applicants, on average. The Akmola archery club opened in 2018, and there was no time to develop an appropriate admission process, which is why they accept on the first-come-first-serve basis. After 10-15 students are recruited, the other potentially better 35-50 students are automatically turned down because of the lack of free spots. However, after a few weeks, some students stop attending archery training and even drop out of the club, which results in a waste of coaches' time, since they could have focused on other students to make more progress. Those students usually don't have much progress since they lack initial body strength and motivation, while other students who see improvement and have excellent physical abilities grow more interested in archery and continue training. On average, half of the original students drop out, and by that time (i.e., around November), no new students want to enroll as school gets more intense.

To effectively choose potential archery students and ensure attendance and interest, a better strategy in this situation would be to leverage higher potential students and prioritize them over other applicants.

Proposal

I propose to use a method called "logistic regression" to predict the chances of being an HPA through means of physical and mental fitness, so more talented and motivated students will be prioritized.

There are a lot of complex factors that define an HPA, which is why the coaches need first to develop a comprehensive list of factors that they think are essential for

[1] **#audience** - used knowledge of an audience's perspective to tailor a message in a subtle manner (i.e., tried to use simplified language and didn't use technical terms, tried to tailor the work to fit the receiver's expectations).

excelling in archery. Although athletes mostly build up their strength and technique through constant training, the initial physical fitness is always crucial since it gives a considerable advantage when first starting. Not only physical but also mental fitness matters. Since archery is a highly competitive sport, there is a lot of psychological pressure from opponents, and it takes time and effort to get better. Additional factors that affect becoming HPA include motivation to participate in archery, self-awareness, and other mental fitness factors.

To come up with a comprehensive list, coaches can consult other coaches from other regions and also analyze some of the research papers on the topic, which include essential factors, such as:

- The motivation behind starting archery;
- Stress Management;
- Upper Muscle Strength and Core Muscle Strength assessed through push-up and sit-up tests;
- Hand Grip strength using an adjustable dynamometer or through trying on pulling different weight bows (Muazu et al., 2019) [1].

The training data (i.e., the information that is used to build the model) could be obtained from current and previous high/low potential archers in all regions of Kazakhstan. Since all archers have "sportsman diaries," which include all kinds of information, including physical fitness factors at that point and progress since the beginning of a career, as well as their thoughts (i.e., mental state), it wouldn't be a problem to gather training data. Each row in the data frame could correspond to each archer, and the high (i.e., "1") or low (i.e., "0") potential archers could be classified through the achievement of a master of sports (or a candidate) title.

The "logistic regression" would be implemented through the use of any applicable software (e.g., R, Python) and further assistance of the author. The training data would be used in identifying which of the provided factors are the most important in predicting top students. After the in-person meeting, the coaches will input the results of the physical examination into the model. It will output student's chances of being an HPA in a range of 0-100% given their initial physical fitness. Since no newbie is advanced enough to become a master of sports right away, this forecasting is needed to identify the most well-fit students, and this probability is just an estimate, which could be wrong. However, in this specific context, as a result of the lack of appropriate admission process, this method is still better than shooting in the dark with choosing the first students who apply. Based on the model output, all of the prospective students could be ranked in a descending order to identify the highest fitness students.

The prospective students always meet in-person with the coach and try out archery before deciding to join. During that time, it would be best to both collect physical fitness information as well as to conduct a quick mental fitness/motivation survey (e.g., could be both in-person, or individually filled out). Based on these two measures, it would be easier to choose more appropriate high potential students by inputting the data into the logistic regression model.

Below is a brief introduction comics (that I drew) to help better understand which steps to follow in logistic regression as the receiver is a fan of manga[2]. You can see

[2] **#interpetivelens** - the coach doesn't have any exposure to data science and doesn't understand

better quality file here[3].

Figure 1: Logistic Regression Comics.

Conclusion

The Akmola archery club administration and coaches could improve their admission process by predicting the chances of becoming high potential archers (i.e., 0-100% values) combined with motivation tests using available information on the athletes' physical and mental fitness. The proposed strategy is better than the previously implemented first-come-first-serve admission process, which could allow prioritizing higher potential students and reduce bias due to the time-dependent acceptance, as well as prospectively improve the club's performance[4].

Word count: 998

technical concepts very well, which might affect her interpretation of the analysis; she also loves anime/manga (e.g., her favourite BNHA heroes which I drew here) and is a visual learner, which is why portraying the required steps for a complex process (i.e., logistic regression) through comics could help her understand and interpret it better.

[3] **#communicationdesign** - effectively applied principles of communication design (i.e., followed all guidelines to create a catchy illustration which doesn't only help the coach better understand the concept, but also memorize it better through known figures - anime characters).

[4] **#composition** - throughout the assignment, I tried to communicate as precisely and parsimoniously as possible (e.g., the executive summary is very brief and informative - only 4 sentences).

Annex

This proposal is addressed to my coach/administrator in an Akmola archery club, which I have been a part of since its foundation. I know this person well and rather than asking specific questions, I have been involved in this problem myself, observing numerous students being turned down, and seeing those who were accepted dropping out. I have met a lot with the coach and we discussed this problem thoroughly. Then I found a relevance to the LBA and she gave me permission to use this information in the assignment.

I realized I don't have any recent photos with her, which is why I am going to include some of the other photos from several months ago. We are highlighted in red in the second photo.

Figure 2: #1 - Coach and I.

Figure 3: #2 - Coach and I.

References

[1] Muazu, M. R., Majeed. A. A., Taha, Z., Chang, S. W., Nasir, A. F. Abdullah M.R. (2019). A machine learning approach of predicting high potential archers by means of physical fitness indicators. PLoS One, 14(1). Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6317817/.