

Golf Analytics

Using Analysis on Downswing Features to Construct
a Virtual Coaching App

Erica King, Amy Wang

Mentored by Dr. Dahbura and Tad Berkery

Sports Analytics Research Group ([sports-analytics.cs.jhu](https://sports-analytics.cs.jhu.edu))



JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Introduction

A golf swing consists mostly of a backswing and a downswing. Although both are important in determining the final shot, we are taking a closer look at the downswing.

By analyzing features in the downswing (club speed, ball speed, etc.), we aim to discover which will most enhance carry (the distance the ball travels).

Golfers of all skill levels can then benefit from in-depth coaching that is both tailored to their golfing patterns and backed by research.

Objectives

- What aspects of the down swing should an amateur golfer focus on improving to best increase their carry?
- How can we utilize our research to provide customized coaching for users?

Materials and Methods

We gathered data from the Trackman simulators at Golf Galaxy.

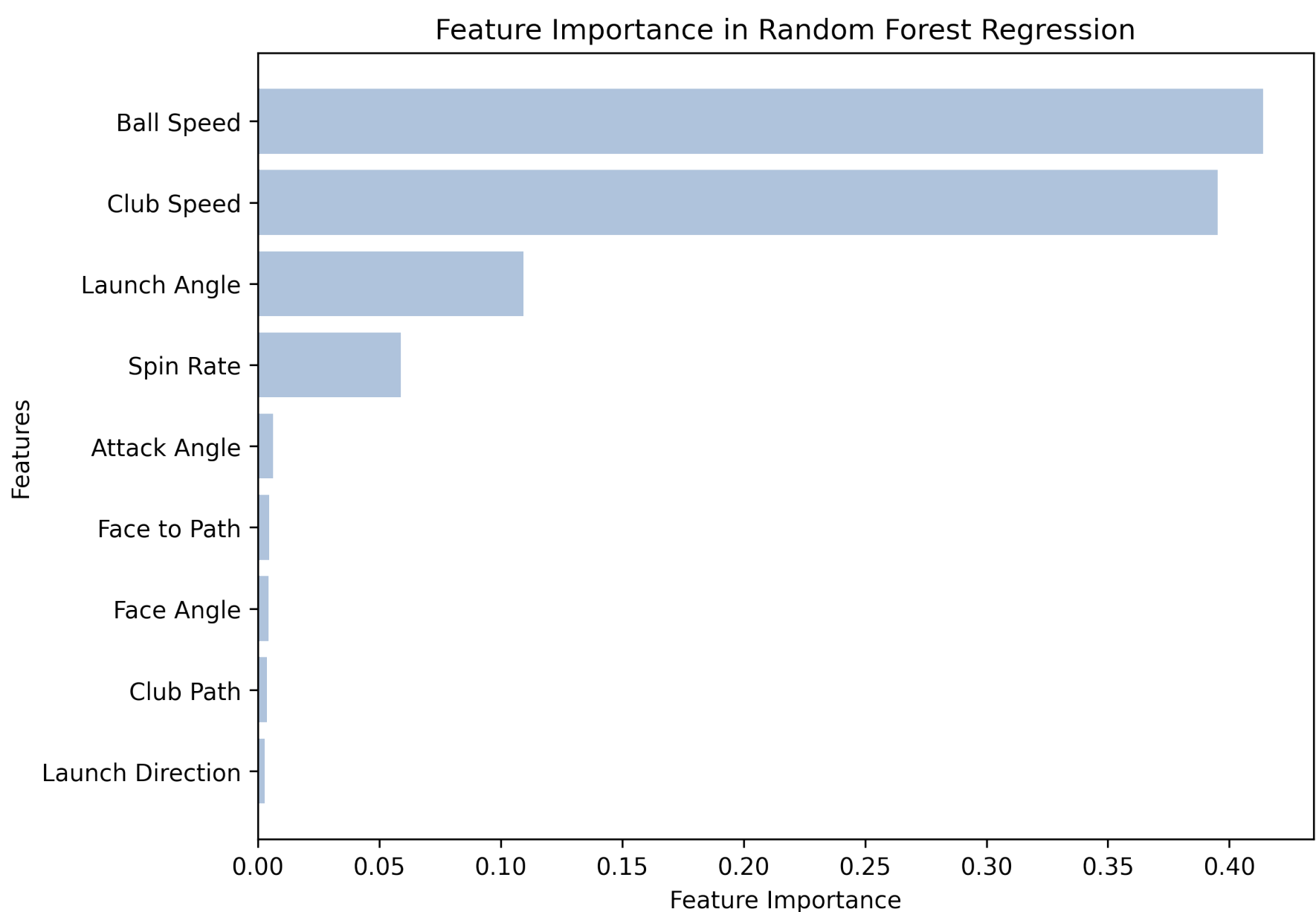
Then, we took a statistical approach by using machine learning algorithms such as random forest regression, support vector regression, and Pearson correlation coefficient calculations.

To build and deploy the dashboard, we used Streamlit with a SQLite database and DeepSeek for the NLP Coach.

Results or Findings

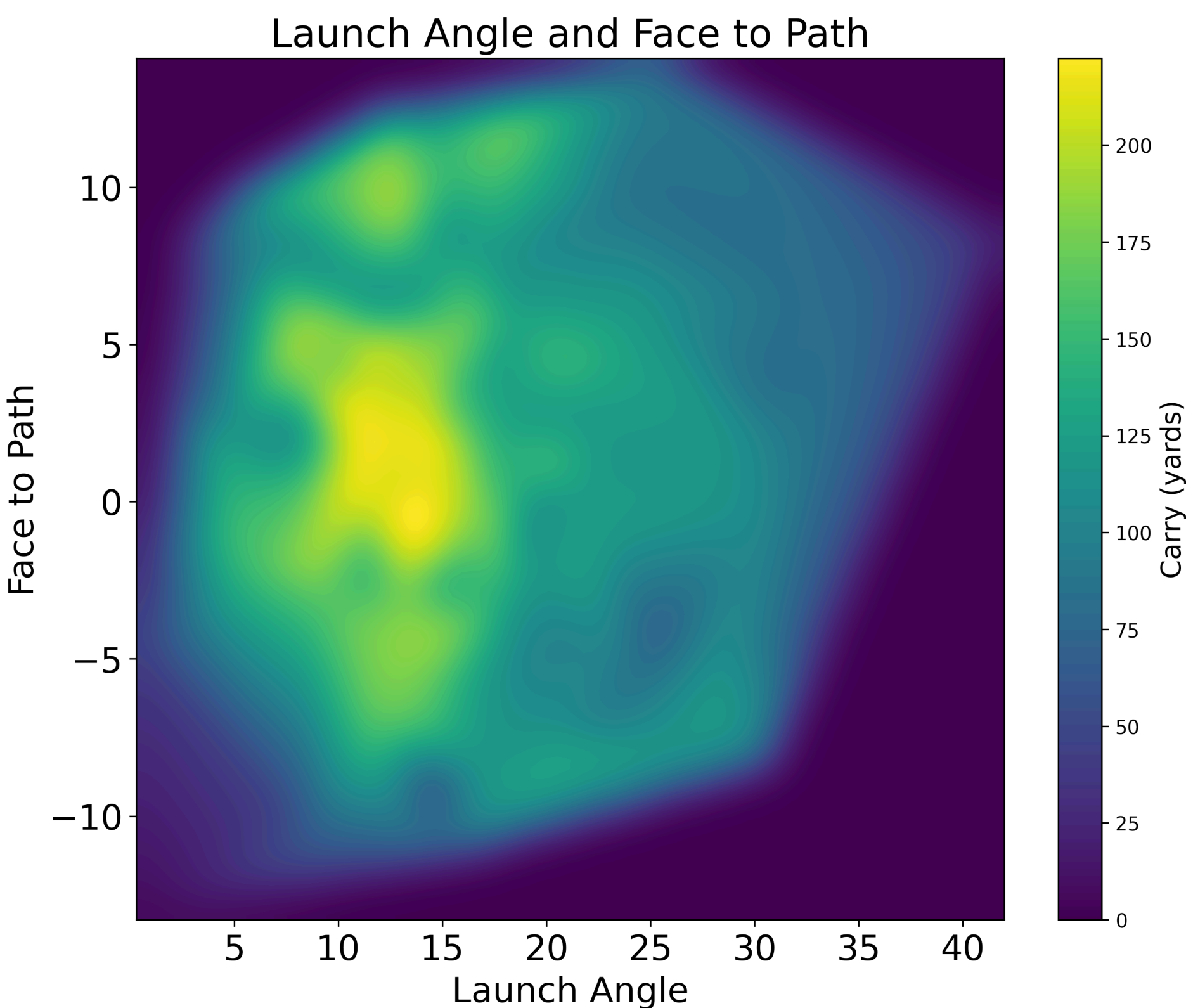
1 Finding Feature Importance

- Using random forest regression to find the order of importance across the 9 features
- Provides a general order of importance of the total influence (when considered alone and in interaction with other features).



2 Finding Optimal Ranges under Pairwise Interaction:

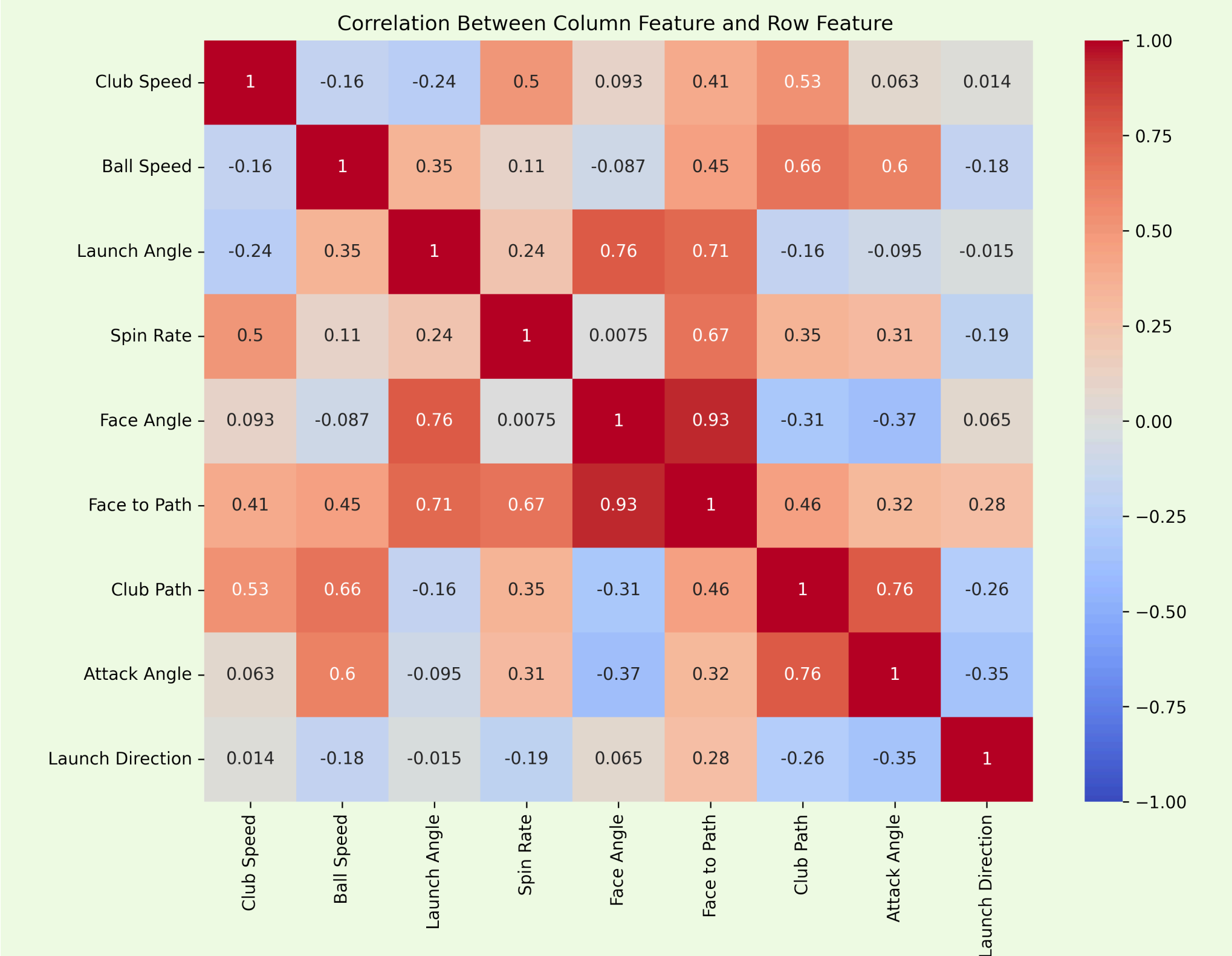
We found the relationships between each feature and carry using support vector regressions, from which we found the optimal range of values for each feature. Then, we used heatmaps to visualize these optimal ranges between pairs of features.



3 Highly Correlated Feature Groups (HCFGs)

Since improvement in highly correlated features happen concurrently, calculating all pairwise Pearson correlation coefficients allows us to use one feature as a proxy for its HCFG.

The HCFGs are 1) club speed & ball speed, 2) face to path & face angle, 3) attack angle & club path; only groups 1 and 2 are important as per the feature importance ranking.



4 Dashboard with NLP Coach: users can upload their own data and receive personalized feedback using our research.

Virtual Golf Coach

Loaded user data from database!

Your Shot History

	Shot Type	Carry (yards)	Club Speed (MPH)	Ball Speed (MPH)	Launch Angle (Deg)	Spin Rate (RPM)
0	Drive	1	30	30	-10	500

Ask your Virtual Coach:

How can I improve my accuracy?

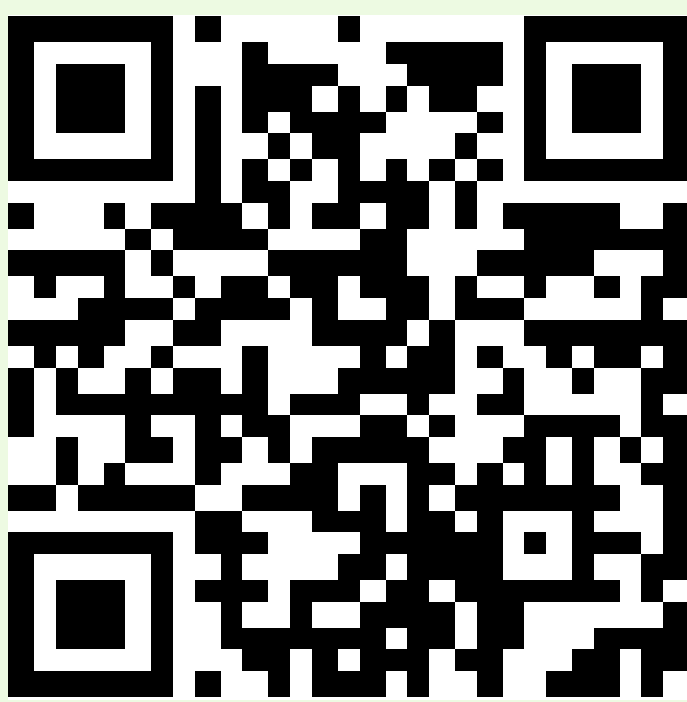
Get Advice

Conclusion

A player can best improve by focusing on **launch angle** and **face to path**.

Ball and club speeds are the most correlated to carry, but they are bounded by physical limitations.

Eliminating abnormal shots by aiming for 11° – 13° of launch angle and improve direction by achieving -1° – 2° of face to path is much more helpful.



Try out our
virtual
coaching
dashboard!