

Golf Forecasting: Week 6 Status Update

Brett Mele

Project Timeline

Task	Week 1 (6/25)	Week 2 (7/2)	Week 3 (7/9)	Week 4 (7/16)	Week 5 (7/23)	Week 6 (7/30)	Week 7 (8/6)	Week 8 (8/13)	Week 9 (8/20)	Week 10 (8/26)
Research and project scope	✓	✓	✓							
Project overview, plan, goals & KPIs			✓							
Data collection			✓	✓						
Exploratory data analysis				✓	✓					
Data preparation				🟡	✓					
Model experimentation					🟡	🟢	🟢			
Finalize models, aggregate outputs							🟢			
Prototype app/dashboard							🟢	🟢	🟢	
Deployment								🟢	🟢	
Final Paper & Presentation									🟢	🟢

- ✓ = complete
- 🟢 = on track
- 🟡 = delayed

Accomplishments

- Engineering
 - Completed data collection
 - Set up data infrastructure
 - Cleaned and processed data into “gold” format for modeling
- Modeling
 - Finished initial exploratory data analysis
 - Developed model for adjusted strokes gained (strokes gained relative to the average golfer)
 - Developed model for overall latent skill (in terms of sg), along with aging curves
 - Tested models for projecting skill into future seasons

Next Steps

1. Commit to approach to model golfer skill in future seasons (likely approach A)
 - a. use aging curve in latent skill model and adjust estimate based on most recent performance
 - b. use model that includes the latent skill estimate at the end of the season, combined with moving averages of performance, to predict performance in next season
2. Create aging curves for components of overall skill (off tee, approach, around green putting)
3. Create model to adjust overall skill estimates in future seasons based on component estimates in future seasons
 - a. Not all golfers will have accurate component estimates due to the limited availability of shot tracking data, so need to develop approach to combine the two that accounts for the confidence in each type of estimate
4. Model tournament outcomes (and earnings) using predicted future skill
 - a. Figure out how to model tournament participation
 - b. Simulation to determine results of tournaments
5. Build prototype shiny application

Summary of Risks and Issues

Risk/Issue	Description	Severity
Deployment of engineering pipeline	Time constraints may limit ability to automate ELT pipeline (not necessary to complete project objectives)	
Modeling of future tournament outcomes	Simulation approach will require a way to figure out what players will participate in each event, since the field for each event is unknown.	