Predicting Baseball Hall Of Fame Induction

Michael Hirsch

ILLC, University of Amsterdam

Abstract

Every year, the Baseball Writers Association of America votes on a new Hall of Fame class. Each ballot consists of 10 votes, and players need to appear on 75% of ballots in order to be inducted into the hall. Players have 15 years to get inducted, and are no longer eligible if that time period has passed. Here we attempt to classify hall of fame batters based on their career statistics using decision trees and an ensemble method, random forests.

1. Introduction

Major League Baseball (MLB) has been keeping fantastic records of batting, picthing, and fielding statistics since its innagural season in 1869. Recently, with the advent of sabermetrics by the Society for American Baseball Research (SABR), many new metrics building on traditional statistics were created. Such metrics caught the eye of stasticians, as these new metrics allowed for the creation of even more powerful predictive models.

In this paper, we are investigating what makes a hall of fame batter. There are several ways in which a player can be inducted, and we only concern ourselves with players inducted from Baseball Writers Association of America (BBWAA) ballots. Attention is paid to 18 different batting statistics over the course of the player's carreer:

1. G: games played
2. G_batting: games in which player batted
3. AB: at-bats
4. R: runs
5. H: hits
6. X2B: doubles
7. X3B: triples

8. HR: homeruns
9. RBI: runs batted in
10. SB: stolen bases
11. CS: caught stealing
12. BB: walks
13. SO: strike outs
14. IBB: intentional walks
15. HBP: hit by pitches
16. SH: sacrifice hits
17. SF: sacrifice flys
18. GIDP: grounded into double plays

lobortis. Nunc feugiat faucibus varius. Ut sed purus nunc. Ut eget eros quis lectus mollis pharetra ut in tellus. Pellentesque ultricies velit sed orci pharetra et fermentum lacus imperdiet. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Suspendisse commodo ultrices mauris, condimentum hendrerit lorem condimentum et. Pellentesque urna augue, semper et rutrum ac, consequat id quam. Proin lacinia aliquet justo, ut suscipit massa commodo sit amet. Proin vehicula nibh nec mauris tempor interdum. Donec orci ante, tempor a viverra vel, volutpat sed orci.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Pellentesque quis interdum velit. Nulla tincidunt sem quis nisi molestie nec hendrerit nulla interdum. Nunc at lectus at neque dapibus dapibus sit amet in massa. Nam ut nisl in diam consectetur dignissim. Sed lacinia diam id nunc suscipit vitae semper lorem semper. In vehicula velit at tortor fringilla elementum aliquam erat blandit. Donec pretium libero et neque vehicula blandit. Curabitur consequat interdum sem at ultrices. Sed at tincidunt metus. Etiam vulputate, lacus eget fermentum posuere, ante mi dignissim augue, et ultrices felis tortor sed nisl.

Maecenas [1] fermentum [2] urna ac sapien tincidunt

- Bullet point one
- Bullet point two

1. Numbered list item one
2. Numbered list item two

1.1. Subsection One

Quisque elit ipsum, porttitor et imperdiet in, facilisis ac diam. Nunc facilisis interdum felis eget tincidunt. In condimentum fermentum leo, non consequat leo imperdiet pharetra. Fusce ac massa ipsum, vel convallis diam. Quisque eget turpis felis. Curabitur posuere, risus eu placerat porttitor, magna metus mollis ipsum, eu volutpat nisl erat ac justo. Nullam semper, mi at iaculis viverra, nunc velit iaculis nunc, eu tempor ligula eros in nulla. Aenean dapibus eleifend convallis. Cras ut libero tellus. Integer mollis eros eget risus malesuada fringilla mattis leo facilisis. Etiam interdum turpis eget odio ultricies sed convallis magna accumsan. Morbi in leo a mauris sollicitudin molestie at non nisl.

Treatments	**Response 1**	**Response 2**
Treatment 1	0.0003262	0.562
Treatment 2	0.0015681	0.910
Treatment 3	0.0009271	0.296

Table 1: Table caption

1.2. Subsection Two

Donec eget ligula venenatis est posuere eleifend in sit amet diam. Vestibulum sollicitudin mauris ac augue blandit ultricies. Nulla facilisi. Etiam ut turpis nunc. Praesent leo orci, tincidunt vitae feugiat eu, feugiat a massa. Duis mauris ipsum, tempor vel condimentum nec, suscipit non mi. Fusce quis urna dictum felis posuere sagittis ac sit amet erat. In in ultrices lectus. Nulla vitae ipsum lectus, a gravida erat. Etiam quam nisl, blandit ut porta in, accumsan a nibh. Phasellus sodales euismod dolor sit amet elementum. Phasellus varius placerat erat, nec gravida libero pellentesque id. Fusce nisi ante, euismod nec cursus at, suscipit a enim. Nulla facilisi.

Integer risus dui, condimentum et gravida vitae, adipiscing et enim. Aliquam erat volutpat. Pellentesque diam sapien, egestas eget gravida ut, tempor eu nulla. Vestibulum mollis pretium lacus eget venenatis. Fusce gravida nisl quis est molestie eu luctus ipsum pretium. Maecenas non eros lorem, vel adipiscing odio. Etiam dolor risus, mattis in pellentesque id, pellentesque eu nibh. Mauris nec ante at orci ultricies placerat ac non massa. Aenean imperdiet, ante eu sollicitudin vestibulum, dolor felis dapibus arcu, sit amet fermentum urna nibh sit amet mauris. Suspendisse adipiscing mollis dolor quis lobortis.

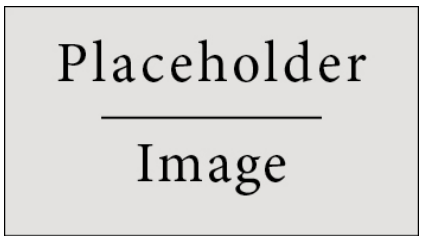

Figure 1: Figure caption

$$e = mc^2 \tag{1}$$

2. The Second Section

Reference to Section 1. Etiam congue sollicitudin diam non porttitor. Etiam turpis nulla, auctor a pretium non, luctus quis ipsum. Fusce pretium gravida libero non accumsan. Donec eget augue ut nulla placerat hendrerit ac ut mi. Phasellus euismod ornare mollis. Proin tempus fringilla ultricies. Donec pretium feugiat libero quis convallis. Nam interdum ante sed magna congue eu semper tellus sagittis. Curabitur eu augue elit.

Aenean eleifend purus et massa consequat facilisis. Etiam volutpat placerat dignissim. Ut nec nibh nulla. Aliquam erat volutpat. Nam at massa velit, eu malesuada augue. Maecenas sit amet nunc mauris. Maecenas eu ligula quis turpis molestie elementum nec at est. Sed adipiscing neque ac sapien viverra sit amet vestibulum arcu rhoncus.

Vivamus pharetra nibh in orci euismod congue. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Quisque lacus diam, congue vel laoreet id, iaculis eu sapien. In id risus ac leo pellentesque pellentesque et in dui. Etiam tincidunt quam ut ante vestibulum ultricies. Nam at rutrum lectus. Aenean non justo tortor, nec mattis justo. Aliquam erat volutpat. Nullam ac viverra augue. In tempus venenatis nibh quis semper. Maecenas ac nisl eu ligula dictum lobortis. Sed lacus ante, tempor eu dictum eu, accumsan in velit. Integer accumsan convallis porttitor. Maecenas pretium tincidunt metus sit amet gravida. Maecenas pretium blandit felis, ac interdum ante semper sed.

In auctor ultrices elit, vel feugiat ligula aliquam sed. Curabitur aliquam elit sed dui rhoncus consectetur. Cras elit ipsum, lobortis a tempor at, viverra vitae mi. Cras sed urna sed eros bibendum faucibus. Morbi vel leo orci, vel

faucibus orci. Vivamus urna nisl, sodales vitae posuere in, tempus vel tellus. Donec magna est, luctus non commodo sit amet, placerat et enim.

[1] J. M. Smith, A. B. Jones, Book Title, Publisher, 7th edition, 2012.

[2] A. B. Jones, J. M. Smith, Article Title, Journal Title 13 (2013) 123–456.