Joseph Schueder, Anthony Egbuniwe, Che’ Cobb

MSDS 6372

Project 2

Introduction

Kobe Bryant marked his retirement from basketball by scoring 60 points in his final game as a member of the Los Angeles Laker team on Wednesday, April 12, 2016. Starting to play professional basketball at the age of 17, Kobe earned the sport’s highest accolades throughout his long career. Using 20 years of data on Kobe's shots made and shots missed, can you predict which shots will be successful?

Section I. Data Description Exploratory Data Analysis

Our dataset includes 216 observation of monthly average ridership various metrics and three economic indicators from January 2001 to December 2018. Average metrics are the sum of a month’s total ridership, per type, divided by the total amount of day types in the month. The sum metrics are the products of the month’s average ridership, per type, multiplied by the number of day types in the month.

Data are collected and recorded by DAS, the faregate system BART uses. Data includes all revenue exits, discounted and full fare tickets.

Metrics include: 29 variables and 25697 observations in the Training dataset

A screenshot of a cell phone

Description generated with very high confidenceA screenshot of a computer

Description generated with high confidence

Section II. Exploratory Data Analysis

• **Address the need for any potential transformations**

**• Address and identify outliers**

• **Address and identify any multicollinearity**A screenshot of a cell phone

Description generated with high confidence

In review of the Tolerance results, we can see several variables – namely game\_event\_id, game\_id, period, playodds and shot\_id – having values well below our 0.1 cutoff value. This finding is echoed in review of the Variance Inflation results, where these same variables reveal values far larger than our 10 cutoff for this column. For the sake of completeness, we will also review the collinearity diagnostics.

A picture containing wall

Description generated with very high confidence

In review of the eigenvalue and condition index association, we can see a large deviation in the final three factors, with the eigenvalue resulting very close to zero and the condition index resulting quite large in comparison. So, we have found a prime case for multicollinearity

Section III. Questions to Answer

Do the odds of Kobe making a shot decrease with respect to distance he is from the hoop?

**Figure 3:** Predicted Shot Probability from distance to the Hoop

A picture containing text

Description generated with high confidence

Does the probability of Kobe making a shot descrease linearly with respect to the distance he is from the hoop?

Is the relationship between the distance Kobe is from the basket and the odds of him making the shot diﬀerent if they are in the playoﬀs.