Project Report : Predicting Win Probability in DoTA2 Using Logistic Regression

Paraskumar, Karthigeyan, Shanmuga Priya, Raju, Niranjan

06/08/2019

### Data Cleaning

The majority of the data cleaning ws carried out in Microsoft Excel. We had to decouple the hero selection and hence created a another set of 111 columns to display the enemies chosen by the other team. Also the region column had to be converted from integer to a string to increase readability.

### Data Understanding

dota2<-read.csv("dota2\_final\_v0.csv", colClasses = "factor")  
dota2.hero<-dota2  
dota2.hero$region<-NULL

Here the first column tells whether radian team won (=1) or lost the match (=0). The region tell us the geographical server on which the game was played. The columns names with “choose\_” before the hero name tells if the radiant team chose (=1) that particular hero or not (=0). The column names with “find\_” before the hero name tells if the dire team chost (=1) that particular hero or not (=0). The entire dataset contains

### Model

We made a logistic regression where we find how choosing a particular characters affects your probability of winning

#Logistic Regression Model

# Exponential of Coefficients