CCSU

DEPARTMENT OF MATHEMATICAL SCIENCES

COLLOQUIUM

Thursday, March 21

3:00 – 4:00 PM

Maria Sanford Hall, Room 101

ANALYSIS OF THE IMPACT OF WEATHER ON RUNS SCORED IN BASEBALL GAMES AT FENWAY PARK

STEVEN CULTRERA

(Data Mining MS Thesis Presentation)

CENTRAL CONNECTICUT STATE UNIVERSITY

Abstract: The goal of this thesis is to determine the impact of weather on the number of runs scored in a baseball game. To accomplish this, a dataset was extracted using 40 years of baseball data from Major League Baseball games played at Fenway Park in Boston, MA. This data was combined with weather data measured five miles away at Logan International Airport. Several data mining techniques were used to analyze the data including unsupervised k-means clustering and principal components analysis (PCA), as well as supervised decision trees and exhaustive linear model search.

Several interesting results were obtained through exploratory data analysis (EDA). The orientation of the park, it's location on the New England coast, and the prevailing wind conditions mean there are many games played during favorable hitting conditions. We also see there is no difference in average runs scored between warmer months and cooler months.

But using linear models, both the wind blowing toward the outfield and the temperature were found to account for just over 2% of the variance of total runs scored by both teams in the games played during the 40-year span. And it is shown that this variance is not explained by players' increased intent to hit home runs while the wind is blowing in a favorable direction.

Additionally, there are about 14.5% more runs scored during warm games when the wind is blowing out as compared to cooler games when the wind is blowing in toward home plate.

For further information:

gotchevi@ccsu.edu 860-832-2839

http://www.math.ccsu.edu/gotchev/colloquium/