CCSU
DEPARTMENT OF MATHEMATICAL SCIENCES

COLLOQUIUM

Friday, April 29
11:45 AM – 12:15 PM
Davidson Hall, Room 207

IDENTIFYING THE ATTRIBUTES THAT CONTRIBUTE TO SUCCESS IN THE PGA TOUR'S FEDEX CUP USING SHOTLINK DATA

PHILIP D. HICKEY

(Data Mining MS Thesis Presentation)

CENTRAL CONNECTICUT STATE UNIVERSITY

Abstract: This thesis analyzes the shots taken by 1,875 professional golfers on the PGA tour from 2007 to 2014 to determine the attributes of the golf game that contributed most to success in the FedEx Cup playoffs. The data was obtained from the CDW ShotLink® database and was analyzed using the R statistical programming language and SQL. Golf is currently experiencing a statistics revolution thanks in large part to the PGA Tour's decision in the mid 90's to completely overhaul the way it tracked and disseminated scoring data. Beginning in the 2004 season ShotLink® was introduced which utilized a laser measuring system to track the starting and ending position for every shot by every player at every tour event (excluding the 4 major championships). In partnership with CDW every golf course is mapped prior to the event in order to calculate exact distance between any two coordinates on a course to within 1 inch. In addition to the PGA and CDW staff it takes roughly 350 volunteers per event to score a tournament. The Random Forest algorithm was used to identify and rank attributes of the golf game for both the regular and playoff seasons in terms of their importance with respect to a player's finish position in competition. Linear regression was also used to identify correlations between specific variables and success in the regular season and/or playoff season with the goal of identifying variables that contribute more to the success of FedEx Cup winners in the playoffs compared with the regular season.

For further information:
gotchevi@ccsu.edu 860-832-2839
http://www.math.ccsu.edu/gotchev/colloquium/