

Proyecto Shark

Introducción

Los datos se han convertido en una importante materia prima para las mejoras en diferentes industrias a nivel mundial tanto que a día de hoy los deportes lo han incluido para mejorar su desempeño. Se puede ver en diferentes equipos europeos como Manchester City o Leicester City que cuentan con un equipo dedicado al análisis de números y estadísticas generadas por sus jugadores en cada partido.

Otro caso similar es como un juego de simulación de fútbol (FIFA o FC) ha sido capaz de predecir el ganador del mundial en los últimos cuatro campeonatos, esto a partir de los datos recolectados en su videojuego que, aunque su finalidad es entretener y vender grandes cantidades de copias también es una enriquecida base de datos junto con inteligencia artificial convirtiéndolo en una herramienta de simulación muy precisa.

Es por eso que decidí realizar un proyecto combinando dos cosas que me gustan, el análisis y el fútbol, y decidí empezar un análisis exploratorio basado en el equipo del cual soy hinchas y suelo mirar sus partidos: el Junior de Barranquilla, un equipo perteneciente a Barranquilla, ciudad ubicada en la costa colombiana. Este proyecto se basó en el torneo de apertura del FPC (Fútbol Profesional Colombiano) del 2024, torneo posterior a la décima estrella lograda en el torneo anterior.

Consideraciones

1. El análisis al momento de realizar este informe abarca hasta la fecha 11 debido a que solo se han jugado esa cantidad de fechas.
2. Este primer informe del análisis contiene información únicamente de estadísticas a nivel general y no enfocadas a cada partido.

Descripción y obtención de datos

Los datos utilizados para el análisis provienen de la página FBREF un portal de estadísticas de fútbol de diferentes países y ligas, realizando una extracción de datos con web scraping, una metodología utilizada para la extracción de datos de sitios web de manera automatizada.

Se realizó la extracción de una tabla principal que contiene las fechas jugadas y los respectivos resultados de cada enfrentamiento junto con otros datos relevantes, el dataframe utilizado de esta tabla durante el análisis lleva por nombre Partidos.

La otra extracción es la de las estadísticas por jugador según cada uno de los partidos jugados hasta la fecha 11, esta tabla contiene datos de los jugadores de Junior (titulares y cambios), edad, minutos jugados y demás estadísticas, los dataframes usados para el análisis llevan el nombre de su respectiva Fecha.

Objetivos de análisis

Generales:

1. Encontrar insights relevantes acerca del rendimiento del equipo, que pueda servir para generar propuestas de mejora.

Específicos:

1. Realizar un análisis exploratorio sobre el rendimiento general como equipo local y visitante.
2. Realizar análisis para la información detallada de cada partido.

Análisis de datos

El análisis de datos se realizó con Python, por medio de librerías como numpy, pandas y matplotlib. Para iniciar se generaron las estadísticas generales relevantes como cantidad de partidos como local y visitante y cantidad de goles en contra y a favor y se sacó el rendimiento del equipo hasta la fecha 11 dando que el equipo tuvo un rendimiento del 54.5%, indicando que de los puntos posibles a conseguir (33 puntos) se habían conseguido poco más de la mitad (18 puntos).

Partidos	
Jugados	11
Jugados como local	5
Jugados como visitante	6
Ganados	5
Empatados	3
Perdidos	3

Goles	
A favor	15
En contra	11
Diferencia	+ 4

Rendimiento General
54.5 %

Luego se decidió revisar por separado el rendimiento del equipo en partidos como local y como visitante:

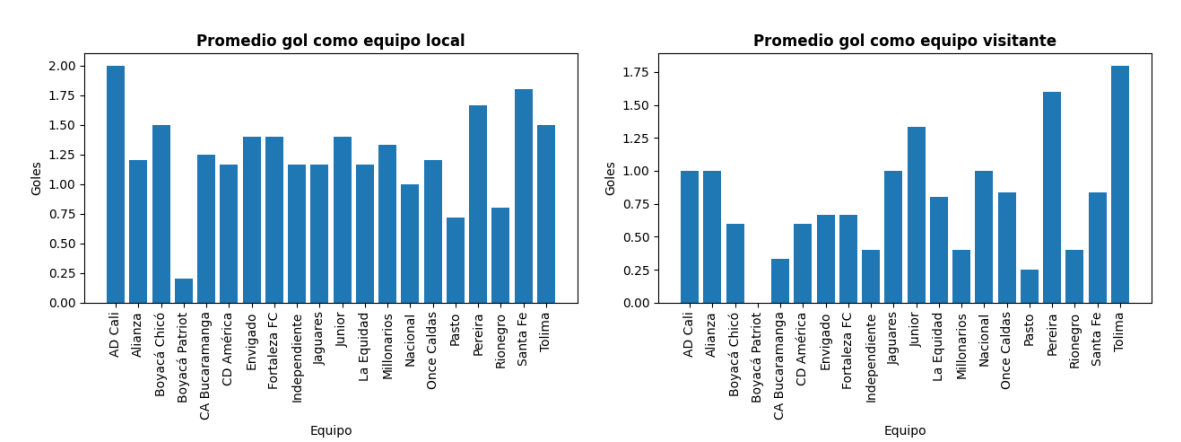
Estadísticas como equipo local	
Partidos Jugados	5
Partidos Ganados	3
Promedio gol a favor	1.4
Promedio gol en contra	0.2
Rendimiento	66.6 %

Estadísticas como equipo visitante	
Partidos Jugados	6
Partidos Ganados	2
Promedio gol a favor	1.33
Promedio gol en contra	1.66
Rendimiento	38.8 %

Aquí es donde resalta dos panoramas interesantes ya que el revisar las estadísticas de local estas reflejan un comportamiento positivo y dando un rendimiento de 66.6 % obteniendo 10 puntos de 15 posibles, lo que se puede considerar un rendimiento cercano a lo ideal ya que el promedio de goles demuestra que el equipo es capaz de marcar mas de 1 gol por partido y mantener el arco en 0.

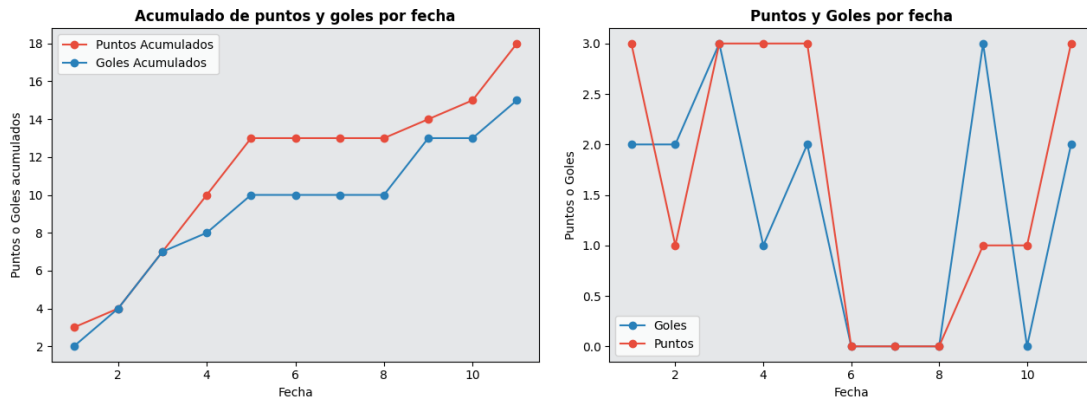
Sin embargo, para el panorama como visitante refleja la parte negativa pues solo se ganaron 2 partidos de 6 disputados y dejando un rendimiento del 38.8 %.

Se realizaron dos graficas con el fin de examinar dos puntos importantes; la cantidad de goles anotados comparada con los demás equipos y los puntos y goles conseguidos al pasar las fechas.



Para la grafica de los goles anotados como local y visitante se visualiza que Junior como local mantiene un promedio de goles similar a otros equipos siendo un punto a mejorar y como visitante si mantiene una cantidad de goles importante al compararlo con otros equipos.

Partidos como los de la Fecha 8 (vs Santa Fe) y 9 (vs Pereira) donde estos equipos marcan mayor cantidad de goles son partidos para analizar.



Para los goles anotados como local y visitante se resalta las fechas de la 5 a la 8 donde no se logra anotar ningún gol, pero no resalta algo más aparte de la correlación normal entre goles y puntos conseguidos.

Conclusiones Parciales

Teniendo en cuenta hasta la fecha 11:

1. Entre las estadísticas generales resalta, parcialmente, que el equipo mantiene un buen rendimiento jugando como local ganado 3 de 5 partidos, con promedio de gol a favor de 1.4 (siendo mínimo de un gol por partido) y promedio de gol en contra de 0.2.
2. Para los partidos jugados como visitante es donde se encuentra el mal rendimiento, ganando solo 2 de 6 partidos y con promedio de gol a favor y en contra similar.
3. Resultaría clave realizar un análisis para las fechas 5,6,7 y 8, donde se presenta una seguida de partidos sin puntos conseguidos.
4. La cantidad de goles anotados como equipo local es baja comparada con los demás equipos.

Debido a los puntos anteriores se debería profundizar en el rendimiento de los partidos a nivel visitante, donde se examine el rendimiento a nivel de jugadores y que patrones o similitudes se hayan.