Pronósticos deportivos

Para calcular los pronósticos deportivos es necesario desarrollar distintos algoritmos que, al final, logren acertar de mejor manera los deportes.

Este proyecto presenta 2 secciones y 4 etapas en cada una. Están divididas como sigue:

1. **Modelo tradicional**
   - I **Algoritmo G&P** A partir de este algoritmo se obtiene el vencedor de un enfrentamiento. Se denota como V a la probabilidad de victoria del vencedor.
   - II **Empates** Si bien ya se obtuvo el vencedor del enfrentamiento en el paso anterior, en el fútbol también existe la probabilidad de empate. Por esto, se debe cacular la probabilidad de empate de un enfrentamiento. Se denota como E a la probabilidad de empate del encuentro.
   - III **Pronósticos** El mejor pronóstico posible es una combinación entre las probabilidades halladas de los 2 puntos anteriores(C_V^E).
   - IV **Test** Para esta etapa se requiere calcular los pronósticos y luego el porcentaje de acierto.
2. **Extensión del modelo**
   - I **Algoritmo G&P** A partir de este algoritmo se obtiene el vencedor de un enfrentamiento. Se denota como V a la probabilidad de victoria del vencedor.
   - II **Empates** Si bien ya se obtuvo el vencedor del enfrentamiento en el paso anterior, en el fútbol también existe la probabilidad de empate. Por esto, se debe cacular la probabilidad de empate de un enfrentamiento. Se denota como E a la probabilidad de empate del encuentro.
   - III **Pronósticos** El mejor pronóstico posible es una combinación entre las probabilidades halladas de los 2 puntos anteriores(C_V^E).
   - IV **Test y calibración** Para esta etapa se requiere calcular el porcentaje de acierto, realizar distintas simulaciones y luego calibrar I, II y III. Y testear nuevamente.

1 Modelo tradicional

I. Algoritmo G&P (Ganar & Perder)

Este es un algoritmo que fue diseñado para el pronósticos de resultados deportivos. Se utilizan probabilidades anidadas y una base de datos que contiene probabilidades (ganar de local, ganar de visitante, perder de local, perder de visitante, empatar de local, empatar de visitante) de los últimos 10 partidos para cada equipo.

Primero, se debe obtener el enfrentamiento. Es decir, el equipo local y el equipo visitante. Segundo, se espera predecir cual de los dos equipos será el vencedor. Para esto, se deben calcular dos probabilidades: Ganar el enfrentamiento (P[G]) o Perder el enfrentamiento[P[P]) (son dos probabilidades independientes, por lo que se deben hallar ambas).

Para calcular la probabilidad de ganar el enfrentamiento se utilizan las probabilidades condicionadas: Si quiero saber quién ganará entre el equipo local y el equipo visitante, resto la probabilidad de ganar del local y la probabilidad de ganar del visitante. De igual forma, para calcular la probabilidad de perder el enfrentamiento se resta la probabilidad de perder del local y la probabilidad de perder del visitante.
Matemáticamente:
Probabilidad de ganar: $P[G] = G$
Probabilidad de perder: $P[P] = P$

$$G = P[G|equipo\_local] - P[G|equipo\_visitante] \tag{1}$$

$$P = P[P|equipo\_local] - P[P|equipo\_visitante] \tag{2}$$

De (1) y (2) se pueden derivar distintas formas de calcular G y P. Por ejemplo, ponderando, promediando, etc. (**Verlo en extensiones del modelo**)

Para saber a cual equipo corresponde G y P, se hace el siguiente análisis:

A partir de (1) se puede deducir que:

$$G \begin{cases} \text{Si } G > 0, & \text{el equipo local gana } (G+) \\ \text{Si } G < 0, & \text{el equipo visitante gana } (G-) \\ \text{Si } G = 0, & \text{ningún equipo gana } (G_0) \end{cases}$$

De igual forma, a partir de (2) se puede deducir que:

$$P \begin{cases} \text{Si } P > 0, & \text{el equipo local pierde } (P+) \\ \text{Si } P < 0, & \text{el equipo visitante pierde } (P-) \\ \text{Si } P = 0, & \text{ningún equipo pierde } (P_0) \end{cases}$$

Se supone que G y P son independientes, por ende, pueden contradecirse. Por ejemplo, si el equipo local tiene la probabilidad de ganar el enfrentamiento ($G+$), pero también tiene probabilidad de perder el enfrentamiento($P+$). Con este ejemplo, se observa que es necesario encontrar una probabilidad conjunta a partir de Ganar(G) y Perder(P) con el objetivo de definir cuál equipo es el vencedor del enfrentamiento.
A partir de lo anterior, se deben combinar todos los casos entre G y P.

casos = { (G+, P+), (G+, P-), (G+, P_0), (G-, P+), (G-. P-), (G-, P_0), (G_0, P+), (G_0, P-), (G_0, P_0) }
Hay un total de 9 casos. Sin embargo, en unos casos hay contradicciones (como el ejemplo anterior). Por ende, para encontrar el vencedor de cada caso, se debe tener en cuenta estas contradicciones:

1. (G+, P+)
    - (G+ > P+) $\longrightarrow$ Vence local
    - (G+ < P+) $\longrightarrow$ Vence visitante
    - (G+ = P+) $\longrightarrow$ No hay vencedor
2. (G+, P-) $\longrightarrow$ Vence local
3. (G+, P_0) $\longrightarrow$ Vence local
4. (G-, P+) $\longrightarrow$ Vence visitante
5. (G-, P-)
    - (G- > P-) $\longrightarrow$ Vence visitante
    - (G- < P-) $\longrightarrow$ Vence local
    - (G- = P-) $\longrightarrow$ No hay vencedor
6. (G-, P_0) $\longrightarrow$ Vence visitante
7. (G_0, P+) $\longrightarrow$ Vence visitante
8. (G_0, P-) $\longrightarrow$ Vence local
9. (G_0, P_0) $\longrightarrow$ No hay vencedor

De los 13 casos presentados, en 10 ocasiones vence el local y el visitante (5 veces cada uno) y en 3 ocasiones no hay vencedor. Para cada caso se deben utilizar metodologías y paramétros arbitrarios que deben ser calibrados, luego de que se calibren, se obtiene la probabilidad del vencedor.

Si bien hay 13 opciones donde hay distintos vencedores, se pueden calcular los 3 casos (>, < y =) para los 9 casos iniciales. Posteriormente, se aplica un algoritmo para cada uno. 1 = 0.4 2 = 0.3

1. **(G+, P+)**

    1.1. (G+ > P+) $\longrightarrow$ Vence local

    $$\max \quad V = \alpha G_+ - \beta P_+ \qquad \text{s.t} \qquad \begin{array}{l} \alpha + \beta = 1 \\ G_+ - P_+ \geq V > 0 \end{array}$$

    Combinando las restricciones:

    $$\max \quad V = \alpha G_+ - (1-\alpha) P_+ \qquad \text{s.t} \qquad G_+ - P_+ \geq \alpha G_+ - (1-\alpha)P_+ > 0$$

    $$\max \quad V = \alpha G_+ - (1-\alpha) P_+ \qquad \text{s.t} \qquad \frac{G_+}{P_+ + G_+} \geq \alpha > \frac{P_+}{P_+ + G_+} \equiv \text{a} \geq \alpha > \text{b}$$

    1.2. (G+ < P+) $\longrightarrow$ Vence visitante

    $$\max \quad V = \beta P_+ - \alpha G_+ \qquad \text{s.t} \qquad \begin{array}{l} \alpha + \beta = 1 \\ P_+ - G_+ \geq V > 0 \end{array}$$

Combinando las restricciones:

$$\max \quad V = \beta P_+ - (1-\beta)G_+ \qquad \text{s.t} \qquad P_+ - G_+ \geq \beta P_+ - (1-\beta)G_+ > 0$$

$$\max \quad V = \beta P_+ - (1-\beta)G_+ \qquad \text{s.t} \qquad \frac{P_+}{P_+ + G_+} \geq \beta > \frac{G_+}{P_+ + G_+} \equiv \text{a} \geq \beta > \text{b}$$

1.3. (G+ = P+) $\longrightarrow$ No hay vencedor

2. **(G+, P-)** $\longrightarrow$ Vence local
   - 2.1. (G+ > P-)
   - 2.2. (G+ < P-)
   - 2.3. (G+ = P-)
3. **(G+, $\mathbf{P_0}$)** $\longrightarrow$ Vence local
4. **(G-, P+)** $\longrightarrow$ Vence visitante
   - 4.1. (G- > P+)
   - 4.2. (G- < P+)
   - 4.3. (G- = P+)
5. **(G-, P-)**
   - 5.1. (G- > P-) $\longrightarrow$ Vence visitante
   - 5.2. (G- < P-) $\longrightarrow$ Vence local
   - 5.3. (G- = P-) $\longrightarrow$ No hay vencedor
6. **(G-, $\mathbf{P_0}$)** $\longrightarrow$ Vence visitante
7. **($\mathbf{G_0}$, P+)** $\longrightarrow$ Vence visitante
8. **($\mathbf{G_0}$, P-)** $\longrightarrow$ Vence local
9. ($\mathbf{G_0}$, $\mathbf{P_0}$) $\longrightarrow$ No hay vencedor

II. Empates

Luego de obtener el vencedor (con su respectiva probabilidad V). Se debe buscar la probabilidad de empate (E).

$$E = P[E|equipo\_local] + P[E|equipo\_visitante] \tag{3}$$

De (3) Se pueden derivar distintas formas de calcular la probabilidad (promedio, ponderación, etc.). (**Verlo en extensiones del modelo**)

III. Pronósticos

Luego de tener las respectivas probabilidades de I y II (V y E), se debe hacer una combinación de ambas probabilidades (C_E^V). Esta combinación es el pronóstico.

$$\begin{gathered} \text{pronóstico} = C_E^V = \Phi \\ \Phi = V - E \end{gathered} \tag{4}$$

De (4) se infiere que:

$$\Phi \begin{cases} \text{Si } \Phi > 0, & \text{el pronóstico es para el vencedor} \\ \text{Si } \Phi < 0, & \text{el pronóstico es un empate} \\ \text{Si } \Phi = 0, & \text{no hay pronóstico} \end{cases}$$

De (4) se pueden derivar distintas formas de calcular Φ, como ponderando o con promedio. Además, se pueden cambiar los casos y hacerlo más estricto (con un parámetro n, donde en el ejemplo, n = 0). (**Verlo en extensiones del modelo**)

2 Extensiones del modelo

I. Algoritmo G&P (Ganar & Perder)

Ahora definimos la probabilidad de ganar y de perder como combinaciones: $G_{G|equipo\_local}^{G|equipo\_visitante}$ y $P_{P|equipo\_local}^{P|equipo\_visitante}$

Hay distintas formas de representar lo anterior, a continuación hay algunos ejemplos:

Modelo tradicional:

$$\begin{gathered} G = P[G|equipo\_local] - P[G|equipo\_visitante] \\ P = P[P|equipo\_local] - P[P|equipo\_visitante] \end{gathered}$$

Ponderación:

$$\begin{gathered} G = \alpha_1 P[G|equipo\_local] - \beta_1 P[G|equipo\_visitante] \\ P = \alpha_2 P[P|equipo\_local] - \beta_2 P[P|equipo\_visitante] \end{gathered}$$

Promedio:

$$G = \frac{P[G|equipo\_local] - P[G|equipo\_visitante]}{2}$$

$$P = \frac{P[P|equipo\_local] - P[P|equipo\_visitante]}{2}$$

Para saber a cual equipo corresponde G y P, se hace el siguiente análisis:

A partir de (1) se puede deducir que:

$$G \begin{cases} \text{Si } G > 0, & \text{el equipo local gana } (G+) \\ \text{Si } G < 0, & \text{el equipo visitante gana } (G-) \\ \text{Si } G = 0, & \text{ningún equipo gana } (G_0) \end{cases}$$

De igual forma, a partir de (2) se puede deducir que:

$$P \begin{cases} \text{Si } P > 0, & \text{el equipo local pierde } (P+) \\ \text{Si } P < 0, & \text{el equipo visitante pierde } (P-) \\ \text{Si } P = 0, & \text{ningún equipo pierde } (P_0) \end{cases}$$

Se supone que G y P son independientes, por ende, pueden contradecirse. Por ejemplo, si el equipo local tiene la probabilidad de ganar el enfrentamiento ($G+$), pero también tiene probabilidad de perder el enfrentamiento($P+$). Con este ejemplo, se observa que es necesario encontrar una probabilidad conjunta a partir de Ganar(G) y Perder(P) con el objetivo de definir cuál equipo es el vencedor del enfrentamiento.
A partir de lo anterior, se deben combinar todos los casos entre G y P.

casos = { (G+, P+), (G+, P-), (G+, P_0), (G-, P+), (G-. P-), (G-, P_0), (G_0, P+), (G_0, P-), (G_0, P_0) }
Hay un total de 9 casos. Sin embargo, en unos casos hay contradicciones (como el ejemplo anterior). Por ende, para encontrar el vencedor de cada caso, se debe tener en cuenta estas contradicciones:

1. (G+, P+)
   - (G+ $>$ P+) $\longrightarrow$ Vence local
   - (G+ $<$ P+) $\longrightarrow$ Vence visitante
   - (G+ $=$ P+) $\longrightarrow$ No hay vencedor
2. (G+, P-) $\longrightarrow$ Vence local
3. (G+, P_0) $\longrightarrow$ Vence local
4. (G-, P+) $\longrightarrow$ Vence visitante
5. (G-, P-)
   - (G- $>$ P-) $\longrightarrow$ Vence visitante
   - (G- $<$ P-) $\longrightarrow$ Vence local

- (G- = P-) $\longrightarrow$ No hay vencedor

6. (G-, P_0) $\longrightarrow$ Vence visitante
7. (G_0, P+) $\longrightarrow$ Vence visitante
8. (G_0, P-) $\longrightarrow$ Vence local
9. (G_0, P_0) $\longrightarrow$ No hay vencedor

De los 13 casos presentados, en 10 ocasiones vence el local y el visitante (5 veces cada uno) y en 3 ocasiones no hay vencedor. Para cada caso se deben utilizar metodologías y paramétros arbitrarios que deben ser calibrados, luego de que se calibren, se obtiene la probabilidad del vencedor.

II. Empates

Luego de obtener el vencedor (con su respectiva probabilidad V). Se debe buscar la probabilidad de empate (E).

$$E = P[E|equipo\_local] + P[E|equipo\_visitante] \tag{1}$$

De (3) Se pueden derivar distintas formas de calcular la probabilidad (promedio, ponderación, etc.). (**Verlo en extensiones del modelo**)

III. Pronósticos

Luego de tener las respectivas probabilidades de I y II (V y E), se debe hacer una combinación de ambas probabilidades (C_E^V). Esta combinación es el pronóstico.

$$\begin{gathered} \text{pronóstico} = C_E^V = \Phi \\ \Phi = V - E \end{gathered} \tag{2}$$

De (4) se infiere que:

$$\Phi \begin{cases} \text{Si } \Phi > 0, & \text{el pronóstico es para el vencedor} \\ \text{Si } \Phi < 0, & \text{el pronóstico es un empate} \\ \text{Si } \Phi = 0, & \text{no hay pronóstico} \end{cases}$$

De (4) se pueden derivar distintas formas de calcular Φ, como ponderando o con promedio. Además, se pueden cambiar los casos y hacerlo más estricto (con un parámetro n, donde en el ejemplo, n = 0). (**Verlo en extensiones del modelo**)