



*Grupo 27*

**FUTALPES F.C.**

# OBJETIVO

Entrenar un modelo de aprendizaje que sea capaz de predecir el precio de nuevos jugadores a partir de sus estadísticas. Así, el club será capaz de reforzarse en cada zona del campo.

# CONSIDERACION SOBRE EL MODELO

Como el grueso de futbolistas que conformaban la base de datos eran jugadores de campo, el modelo no hace un buen trabajo prediciendo el precio de arqueros. Por un lado, no existen columnas como “atajadas” que podrían indicar a detalle qué tan bueno es un portero. Además, a variables como “Valla\_no\_vencida”, el modelo le asigna coeficientes que no son consistentes con el comportamiento que debería tener esta en un arquero (debería ser un aspecto muy valioso). A continuación, se muestra el coeficiente asociado a dicha variable.

Valla\_no\_vencida -1.11669e+04

# RESULTADOS

## MÉTRICAS DE CALIDAD

- CV (log):  $\text{RMSE}_{\text{CV\_log}} \approx 0.98-0.99$ .
- Test (log):  $\text{RMSE}_{\text{test\_log}} \approx 1.08$ ,  $R^2_{\text{test\_log}} \approx 0.40$ .

El RMSE obtenido nos indica que el modelo suele equivocarse en la estimación del valor de un jugador por un factor de alrededor de 2.7x hacia arriba o hacia abajo. Es decir, que entre mayor sea el valor de dicho jugador, más amplio es el intervalo en el que se equivoca el modelo. Por otro lado, el  $R^2$  nos dice que el modelo es capaz de captar, aproximadamente, el 40% de la varianza del valor de mercado de un jugador.

# RESULTADOS

## COEFICIENTES DE LA REGRESIÓN

### POSITIVOS

- Toques en zona ofensiva
- Pases completados
- Toques en el área rival
- Años de contrato restantes
- npxG

### NEGATIVOS

- Edad
- Errores defensivos que ocasionaron tiro
- Malos controles

# CONCLUSIONES Y RECOMENDACIONES

1

Recomendamos añadir a la base de datos características que le aporten valor específicamente a los arqueros. Por ejemplo: “Número de atajadas”, “Porcentaje de tiros atajados”, “Errores que ocasionaron un gol”, etc.

2

El modelo funciona mejor con precios moderados; pues entre mayor sea el precio del futbolista, mayor será el rango de error del modelo.

3

Recomendamos el uso del modelo como una herramienta de estimación, por lo que es importante tomar en cuenta el error que puede cometer este para cada predicción.

4

Los residuos (error entre valor real y predicho) no se comportan de manera normal, por lo que es necesario tener esto en cuenta para interpretación estadística clásica.

5

Se transformó la variable objetivo a  $y=\log(1+market\_value)$  con el propósito de estabilizar la varianza y suavizar el efecto de outliers.

6

Del análisis de supuestos, rescatamos que el modelo es generalizable y apto para producir estimaciones consistentes.