



Grupo 32 - Lab 1

Presentación de los resultados

FutAlpes F.C.

Predicción Market Value jugadores



Objetivo:

- Predecir el valor de mercado de jugadores de fútbol usando un modelo de regresión.
- Analizar el desempeño del modelo y sus limitaciones.



Análisis cuantitativo

MAE:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i|$$

Es la diferencia entre el valor pronosticado y el valor real en cada punto pronosticado

Resultado:

Train: 9947305.145865792

Test: 9726188.189205592

- El MAE \approx 9.7 millones significa que, en promedio, las predicciones del valor de mercado de un jugador se equivocan por unos 9 a 10 millones de euros.
- No hay overfitting: el train y el test dieron resultados similares



RMSE

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2}$$

Mide la diferencia promedio entre los valores predichos y los valores reales

Resultado:

```
Train: 14945347.801968534  
Test: 14445171.50092685
```

- El RMSE penaliza más los errores grandes. Ya que es cerca al doble de mayor que el MAE indica que existen outliers, como es el caso de estrellas de fútbol con contratos de miles de millones de euros.



R²

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}$$

indica la proporción de la varianza de la variable dependiente que es explicada por las variables independientes de un modelo

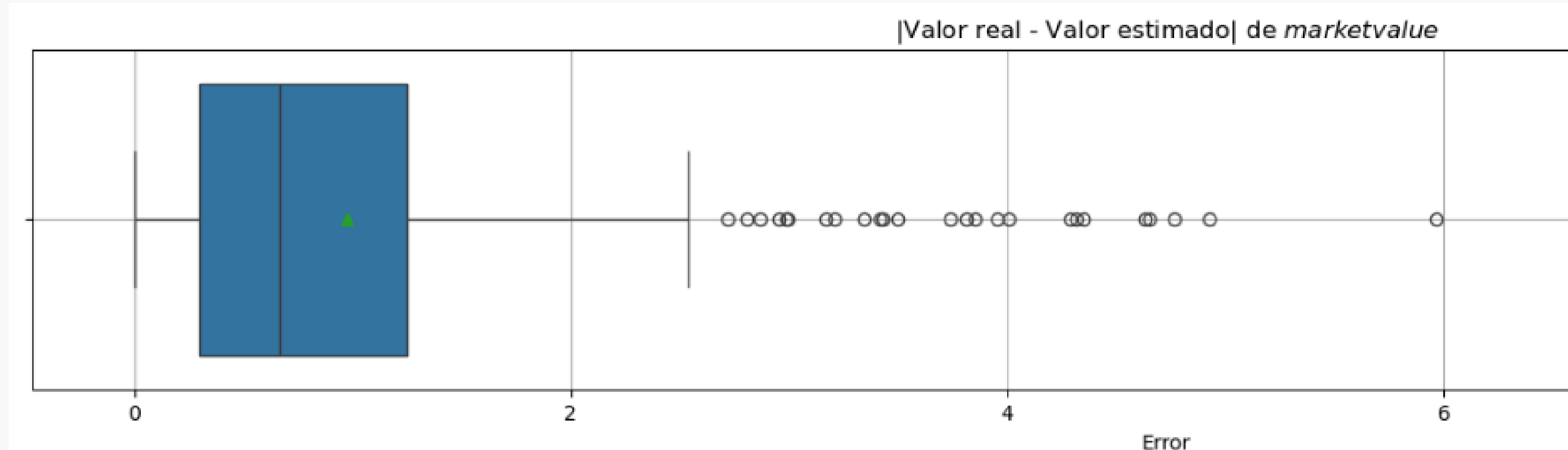
Resultado:

0.3972149850433435

- El modelo logra explicar aproximadamente el 40% de la variabilidad del valor de mercado.
- Las estadísticas de desempeño explican una parte relevante del valor, pero hay un 60% de la variación que se debe a factores fuera del modelo.



Box plot del error del modelo:



```
count    4.600000e+02
mean      9.726188e+06
std       1.069171e+07
min       3.467107e+04
25%       2.939849e+06
50%       6.647965e+06
75%      1.251084e+07
95%      2.952409e+07
99%      4.699135e+07
max       9.458072e+07
Name: market_value, dtype: float64
```

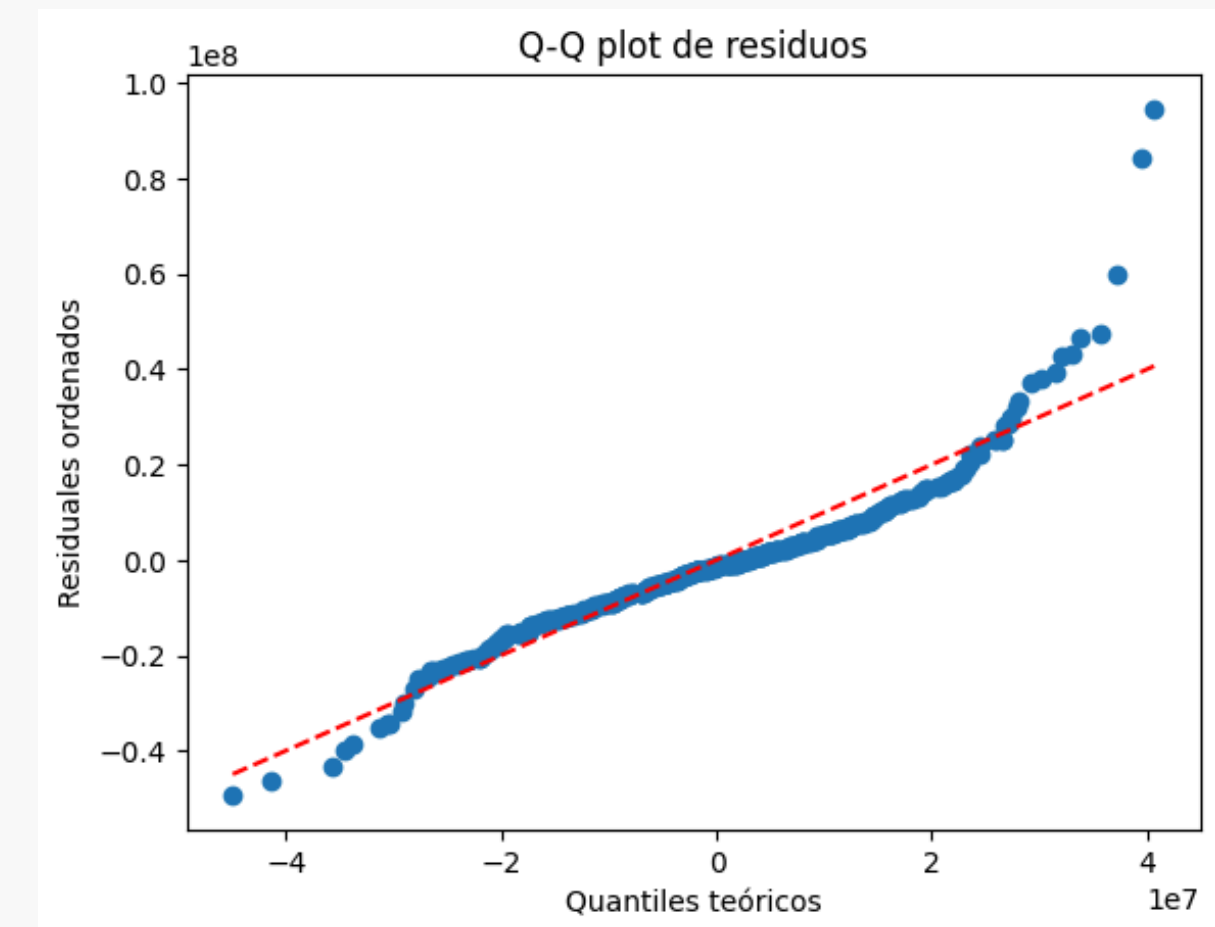
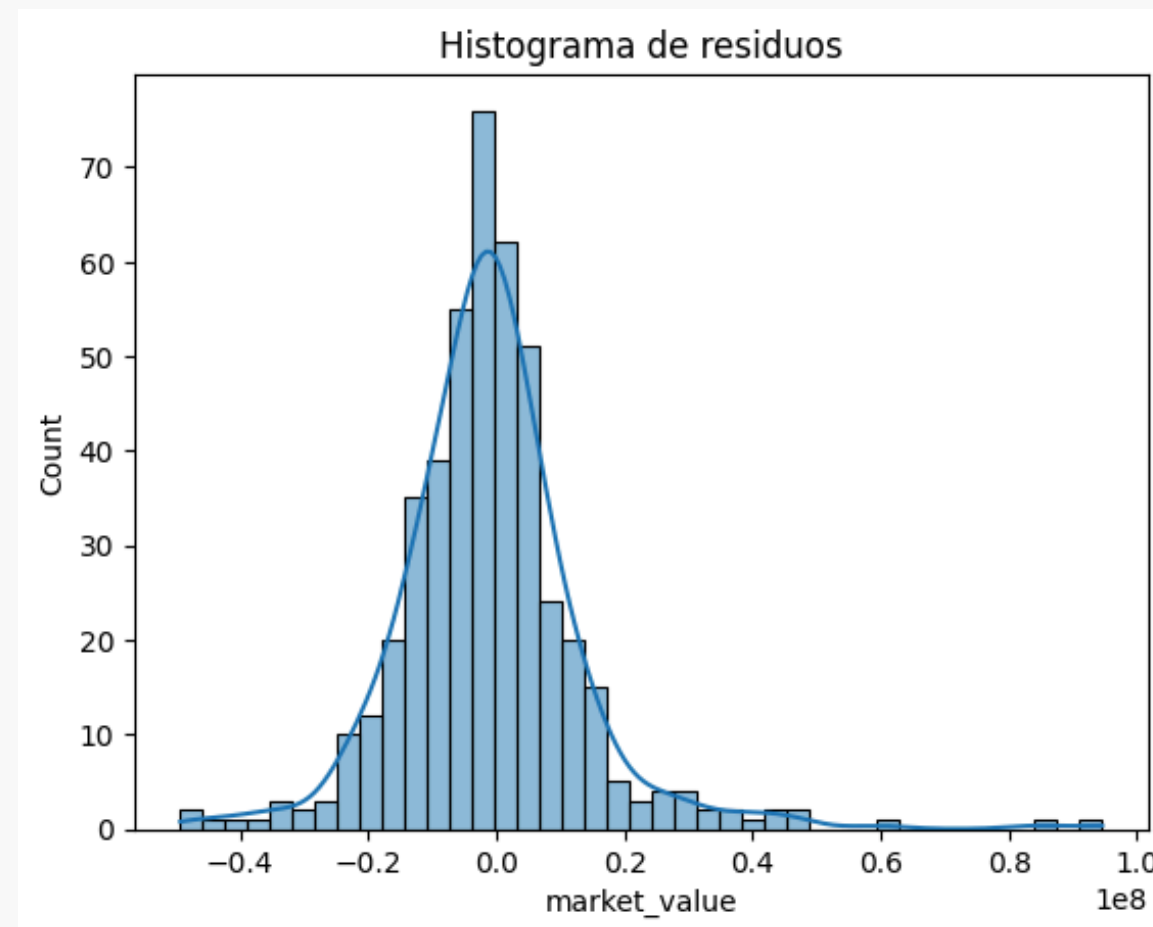
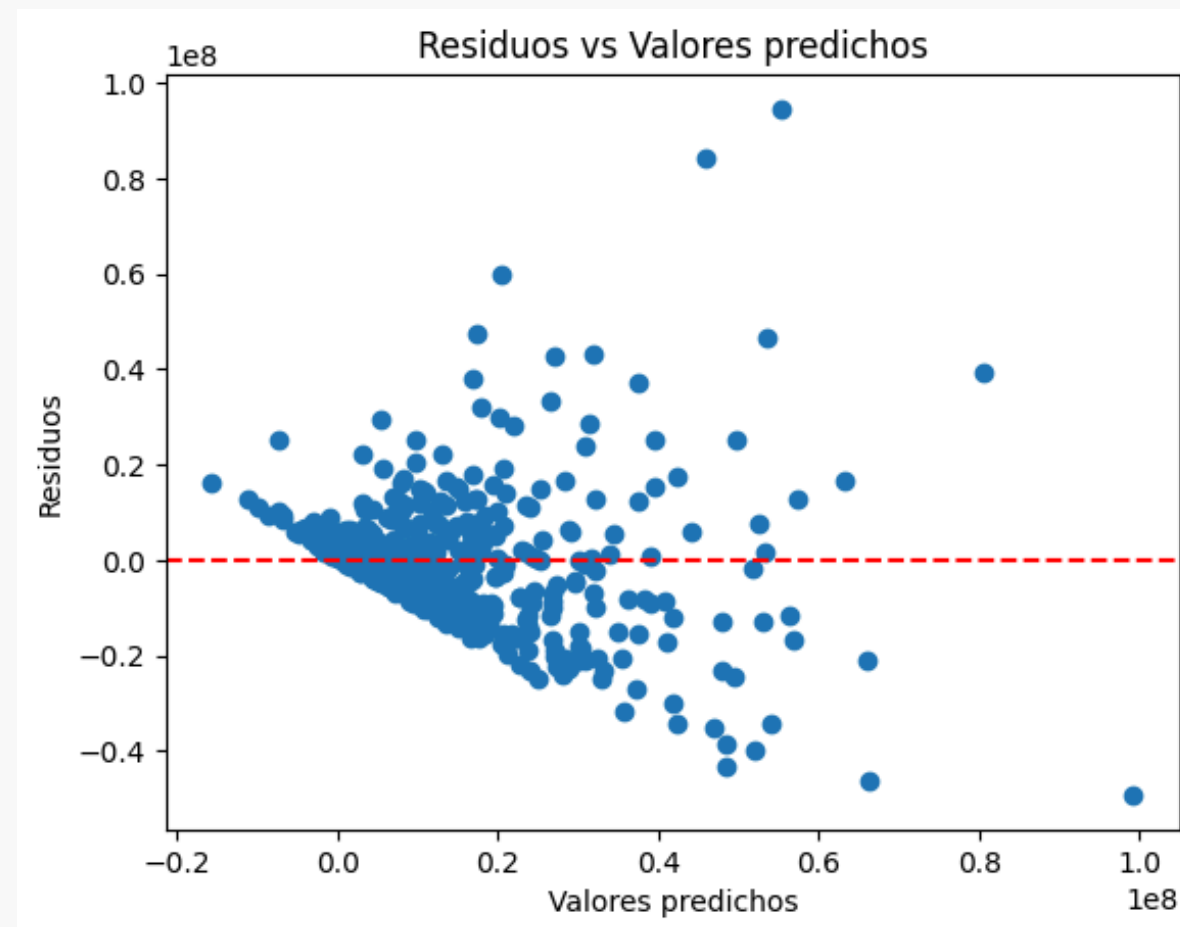
- En promedio, el modelo se equivoca en unos 9.7 millones, aunque hay mucha variabilidad (desviación estándar de 10.6 millones). La mitad de los errores están por debajo de 6.6 millones, y el 75% no supera los 12.5 millones



Análisis cualitativo

Modelo 1:

1. Supuestos regresión:



Las gráficas muestran que los residuos no son completamente aleatorios, presentando un patrón triangular y un sesgo ligero hacia la izquierda. Además, aunque se aproximan a una distribución normal en el centro, en los extremos se desvían significativamente.



Análisis cualitativo

Modelo 1:

2. Coeficientes del modelo:

```
Intercepto: 15309086.672879776

Coeficientes:
      Variable      Coeficiente
0          Goles  2.582669e+06
1 Acciones_que_crear_tiros  5.897993e+06
2 Errores_defnsivos_ocasion_tiro  1.213380e+06
3      Faltas_cometidas -2.455507e+06
4 Veces_que_fue_barrido_regate  2.310042e+06
5      Duelos_aereos_ganados -1.042333e+06
6      Valla_no_vencida  9.856381e+05
7      Tiros_Totales  3.868633e+04
8              npxG  1.299697e+06
9      Pases_intentados  4.516851e+06
10     Pases_progresivos -5.581843e+06
11     Regates_exitosos  2.850209e+06
12     Pases_largos_completados -1.099454e+06
13              xA  4.724469e+05
14              xAG  2.514656e+06
15     Pases_en_ultimo_tercio  2.961898e+06
16     Pases_al_hueco  2.551110e+05
17     Pases_centros -7.837361e+06
18     Pases_fuera_de_juego -8.779924e+05
19     Toques_en_zona_ofensiva  1.118016e+07
20     Toques_en_area_rival  5.513317e+06
21     Acarreos_ultimo_tercio  2.638988e+05
22     Malos_controles -4.581916e+06
23     Perdida_balon -1.252826e+06
24     Pases_progresivos_recibidos -7.493551e+06
```

El modelo captura tendencias generales en el valor de mercado, pero sus coeficientes no siempre reflejan la intuición y muestran limitaciones .

Discriminar por rol permite que el modelo capture mejor las métricas relevantes de cada posición, evitando que estadísticas poco representativas distorsionen los resultados globales.



Análisis cualitativo

Modelo 2:

Portero

```
Coeficientes del modelo
Intercepto: 8287912.087912088
Goles: -461867.3585761244
Acciones_que_crean_tiros: -406934.3632525931
Errores_defnsivos_ocasion_tiro: -1522716.3288837569
Faltas_cometidas: -775049.7290385416
Veces_que_fue_barrido_regate: 24186.615428162597
Duelos_aereos_ganados: 266471.7285479243
Valla_no_vencida: -943345.4020202494
Tiros Totales: 2147491.800961479
npxG: -117090.3544661426
Pases_intentados: 1767497.0913835233
Pases_progresivos: -1704014.7292910325
Regates_exitosos: -221106.19353108032
Pases_largos_completados: -1768629.3913555827
xA: 35085.63836844962
xAG: 1963898.7734966609
Pases_en_ultimo_tercio: -145896.33306927484
Pases_al_hueco: 2121736.3138623573
Pases_centros: -8.149072527885437e-10
Pases_fuera_de_juego: -303390.7446469609
Toques_en_zona_ofensiva: -3402076.9606733313
Toques_en_area_rival: 2147491.8009614795
Acarreos_ultimo_tercio: 1849948.6414490098
Malos_controles: -247644.14028045483
Perdida_balon: 746411.1998713951
Pases_progresivos_recibidos: 0.0
Edad: -1979827.5228241016
```

Delantero

```
Coeficientes del modelo
Intercepto: 17944331.983805664
Goles: 6578868.750902541
Acciones_que_crean_tiros: 11143051.668455103
Errores_defnsivos_ocasion_tiro: 1783539.3726171793
Faltas_cometidas: -2528922.3573032552
Veces_que_fue_barrido_regate: 4631676.659153129
Duelos_aereos_ganados: -2959655.264097408
Valla_no_vencida: -1370596.7861181593
Tiros Totales: 303737.669588754
npxG: 2808031.2140261736
Pases_intentados: -2281828.4614215116
Pases_progresivos: -7866876.724900361
Regates_exitosos: 865342.2329886811
Pases_largos_completados: -832095.501078793
xA: -492507.7692132371
xAG: 565780.9314972355
Pases_en_ultimo_tercio: 4827629.200466455
Pases_al_hueco: -895521.7212061856
Pases_centros: -6065669.948304803
Pases_fuera_de_juego: -100634.33705562976
Toques_en_zona_ofensiva: 14926256.830553388
Toques_en_area_rival: 3843706.908166753
Acarreos_ultimo_tercio: 3549143.179975573
Malos_controles: -8591156.411331093
Perdida_balon: -1155628.275965618
Pases_progresivos_recibidos: -3419895.01238405
Edad: -4395566.220539284
AM: -2434410.26675897
```

La comparación entre los modelos de delanteros y porteros muestra un contraste claro, mientras el modelo para delanteros mejora notablemente la predicción al capturar métricas directamente relacionadas con su desempeño ofensivo, el de porteros empeora debido a la escasez de datos y a la falta de variables explicativas realmente representativas de su rol.



Conclusiones y recomendaciones

- El modelo tiene un desempeño aceptable en jugadores promedio, pero falla en los jugadores con precios muy altos.
- El modelo puede contribuir parcialmente al objetivo de FutAlpes F.C. ya que ayuda a estimar de forma aproximada el valor de los jugadores a partir de su rendimiento, con limitaciones de precisión importantes.
- No recomendamos usar este modelo en producción aún, sin embargo, es una buena base inicial y permite explorar la predicción de valores medios.

