

BIENVENIDOS

Lab de Proyecto

Master Class IA

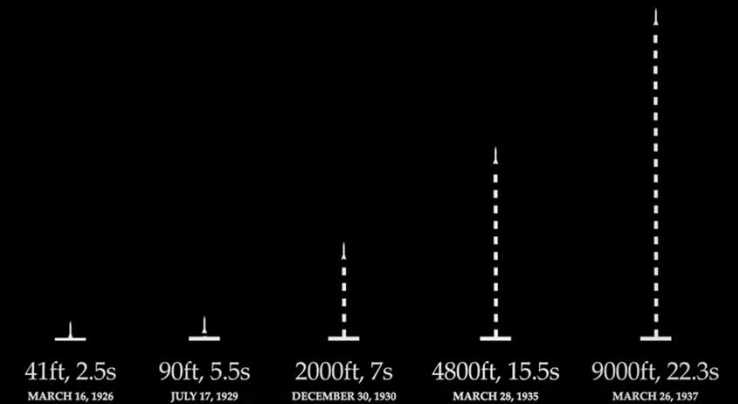




Carl Friedrich Gauß (1777 - 1855)

“Ich habe fleissig
sein müssen; wer
es gleichfalls ist,
wird eben so weit
kommen.”

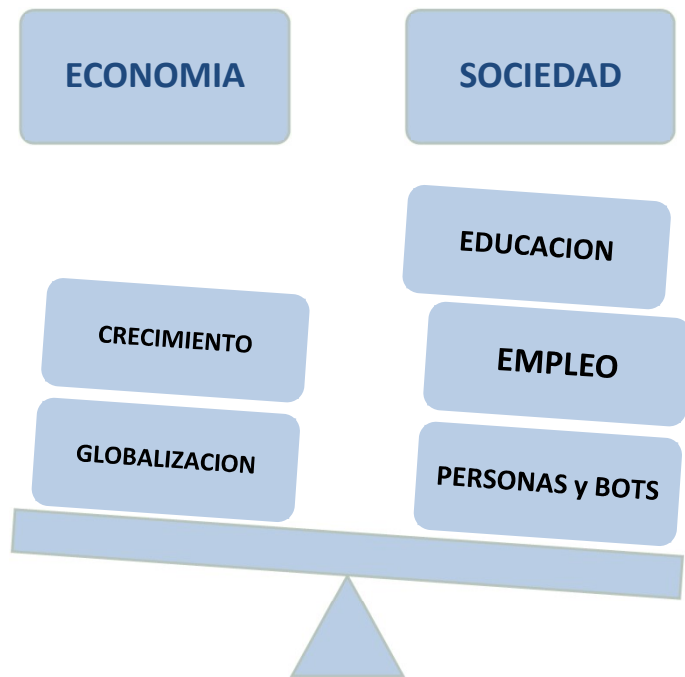
La historia de la Ciencia



Over the next two decades, he continued experimenting with liquid-fueled rockets, making rapid progress and reaching increasingly greater heights.

Robert H. Goddard
(1882 - 1945)

El mundo esta cambiando



INTELIGENCIA ARTIFICIAL MACHINE LEARNING

No solo es posible, sino que
es indispensable

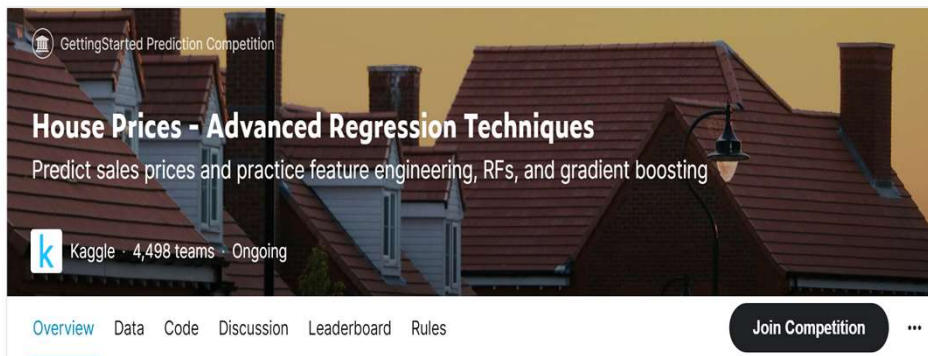
Artificial Intelligence needs all of us | Rachel Thomas

<https://www.fast.ai/>

Lab de Proyecto

Reto 1

Reto 2



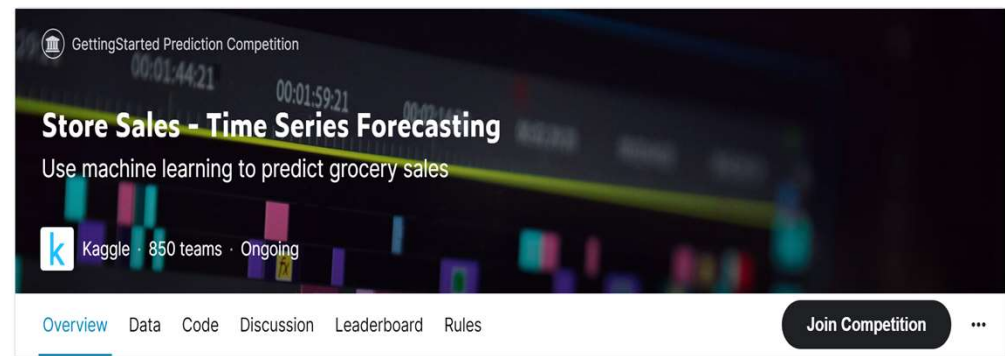
GettingStarted Prediction Competition

House Prices - Advanced Regression Techniques

Predict sales prices and practice feature engineering, RFs, and gradient boosting

Kaggle · 4,498 teams · Ongoing

[Overview](#) [Data](#) [Code](#) [Discussion](#) [Leaderboard](#) [Rules](#) [Join Competition](#) ...



GettingStarted Prediction Competition

Store Sales - Time Series Forecasting

Use machine learning to predict grocery sales

Kaggle · 850 teams · Ongoing

[Overview](#) [Data](#) [Code](#) [Discussion](#) [Leaderboard](#) [Rules](#) [Join Competition](#) ...



GettingStarted Prediction Competition

Digit Recognizer

Learn computer vision fundamentals with the famous MNIST data

Kaggle · 1,775 teams · Ongoing

[Overview](#) [Data](#) [Code](#) [Discussion](#) [Leaderboard](#) [Rules](#) [Join Competition](#) ...



GettingStarted Prediction Competition

Natural Language Processing with Disaster Tweets

Predict which Tweets are about real disasters and which ones are not

Kaggle · 850 teams · Ongoing

[Overview](#) [Data](#) [Code](#) [Discussion](#) [Leaderboard](#) [Rules](#) [Join Competition](#) ...

Equipos

GRUPO 1

DIEGO MARZO
JULIAN CORIANO

GRUPO 2

RAUL PULIDO
RACHELE RENDINA
RODRIGO ANCHELEGRES

GRUPO 3

JAFET BENITEZ
JOSE MARIA CALVO
JUAN BADAL

GRUPO 4

NATALIA SISAMON
JAVIER RABAL

GRUPO 5

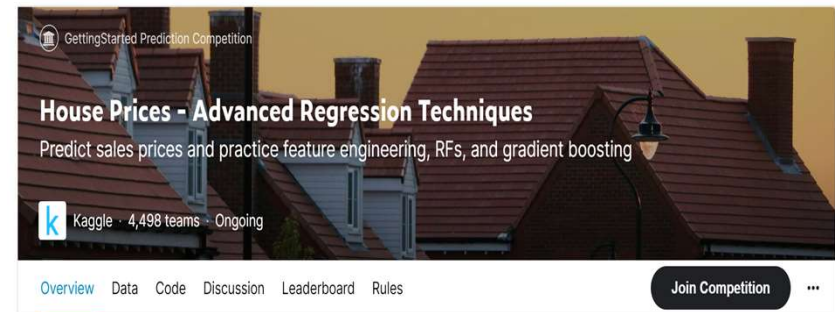
FRANCISCO RUIZ



[cpl-ceste/midsia2022_lab_proy](https://github.com/cpl-ceste/midsia2022_lab_proy)

House Prices

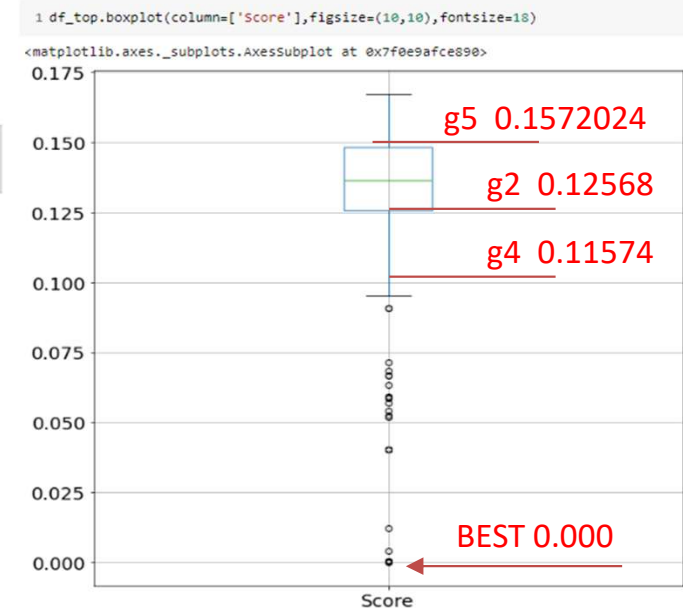
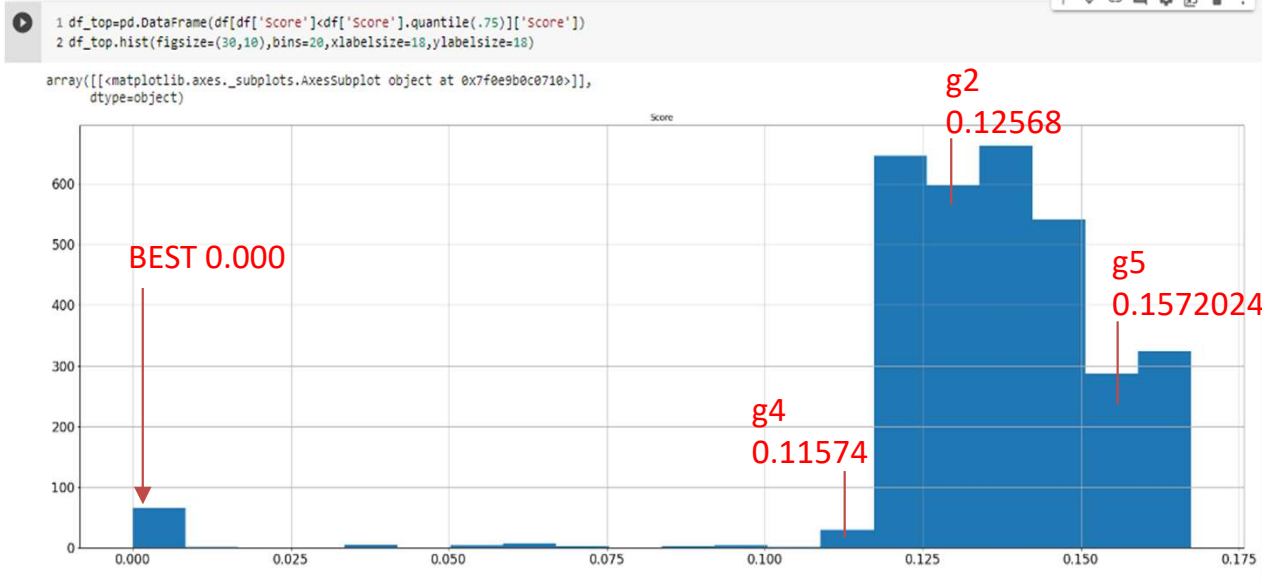
4238 participantes -> ~1060 en cada cuartil



```
1 for q in [0.25, 0.50, .75, 1]:  
2     print(f'quantile {q} ->'+str(df['Score'].quantile(q)))
```

```
quantile 0.25 ->0.1294  
quantile 0.5 ->0.1443  
quantile 0.75 ->0.16743  
quantile 1 ->24.69626
```

Los 75% top ~3000 mejores



Digit Recognizer

1376 participantes -> 350 en cada cuartil



Digit Recognizer
Learn computer vision fundamentals with the famous MNIST data

Kaggle 1,775 teams Ongoing

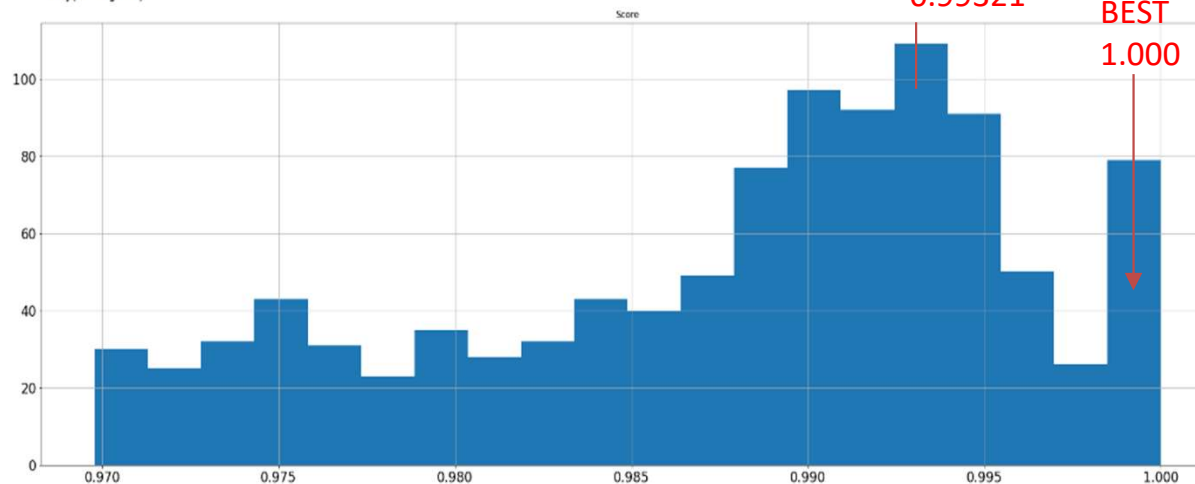
Overview Data Code Discussion Leaderboard Rules

Join Competition

```
1 for q in [0.25, 0.50, .75, 1]:  
2     print(f'quantile {q} ->'+str(df['Score'].quantile(q)))  
  
quantile 0.25 ->0.9697724999999999  
quantile 0.5 ->0.985885  
quantile 0.75 ->0.9926  
quantile 1 ->1.0
```

Los 75% top ~1000 mejores

```
[40] 1 df_top=pd.DataFrame(df[df['Score']>df['Score'].quantile(.25)][df['Score']])  
2 df_top.hist(figsize=(30,10),bins=20,xlabelsize=18,ylabelsize=18)  
  
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x7f0e9b557cd0>]],  
      dtype=object)
```



```
[41] 1 df_top.boxplot(column=['Score'],figsize=(10,10),fontsize=18)
```



Store Sales

630 participantes -> 150 en cada cuartil

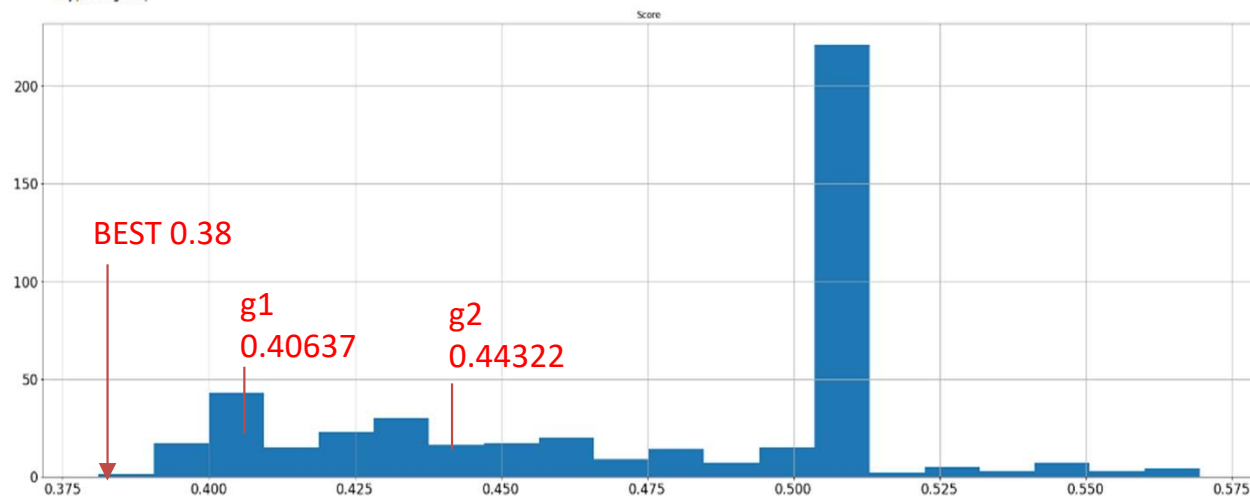
```
1 for q in [0.25, 0.50, .75, 1]:
2     print(f'quantile {q} ->'+str(df['Score'].quantile(q)))

quantile 0.25 ->0.45383249999999997
quantile 0.5 ->0.5109
quantile 0.75 ->0.5706875
quantile 1 ->4.53053
```

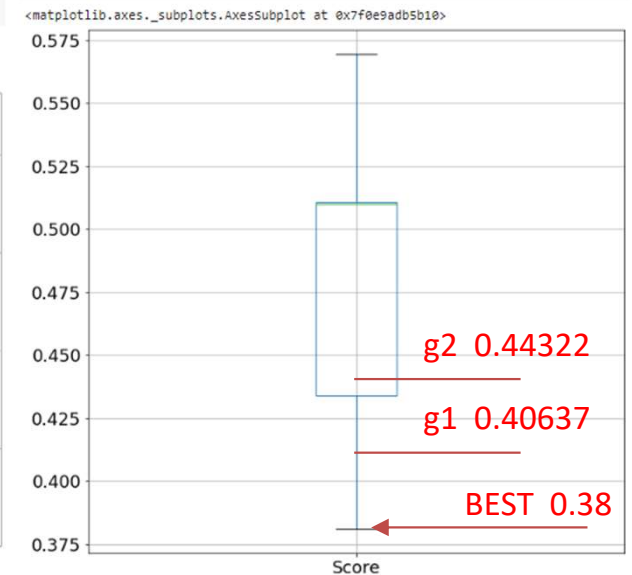
Los 75% top ~470 mejores

```
[57] 1 df_top=pd.DataFrame(df[df['Score']<df['Score'].quantile(.75)]['Score'])
      2 df_top.hist(figsize=(30,10),bins=20,xlabelsize=18,ylabelsize=18)

array([[<matplotlib.axes._subplots.AxesSubplot object at 0x7f0e9ae24b50>]],
      dtype=object)
```



```
1 df_top.boxplot(column=['Score'],figsize=(10,10),fontsize=18)
```



NLP Disaster Tweets

954 participantes -> 240 en cada cuartil



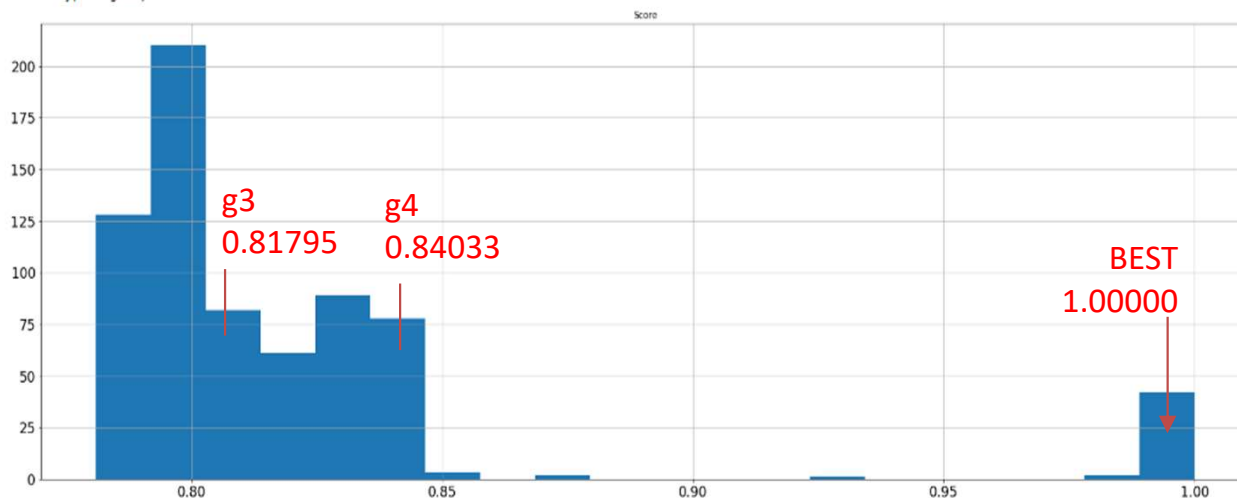
```
1 for q in [0.25, 0.50, .75, 1]:
2     print(f'quantile {q} ->'+str(df['Score'].quantile(q)))

quantile 0.25 ->0.78057
quantile 0.5 ->0.7965
quantile 0.75 ->0.82071
quantile 1 ->1.0
```

Los 75% top ~750 mejores

```
[63] 1 df_top=pd.DataFrame(df[df['Score']>df['Score'].quantile(.25)][['Score']]
      2 df_top.hist(figsize=(30,10),bins=20,xlabelsize=18,ylabelsize=18)

array([[<matplotlib.axes._subplots.AxesSubplot object at 0x7f0e9aaa4f10>]],
      dtype=object)
```



```
1 df_top.boxplot(column=['Score'],figsize=(10,10),fontsize=18)
```

