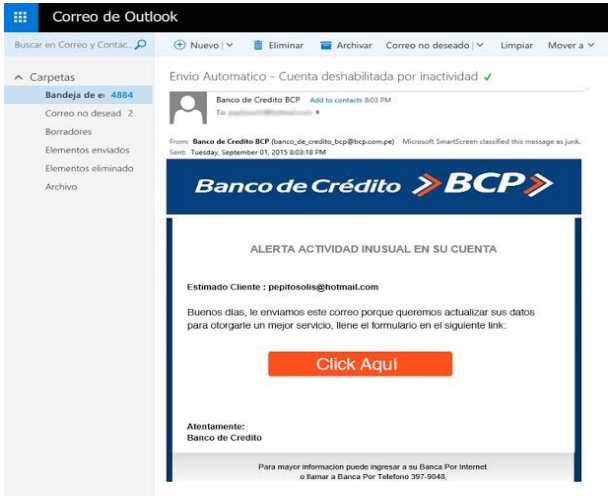
Detección de sitios web de phishing aplicando algoritmos de clasificación

Nombre: Fabian Farez

Contexto

- **Phishing:** estafas por Internet
- Uso de correo electrónico, mensajes
- Robar información personal
- Métodos Machine Learning

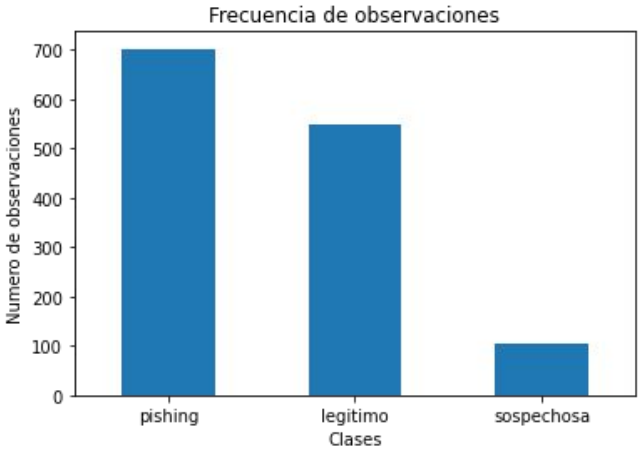


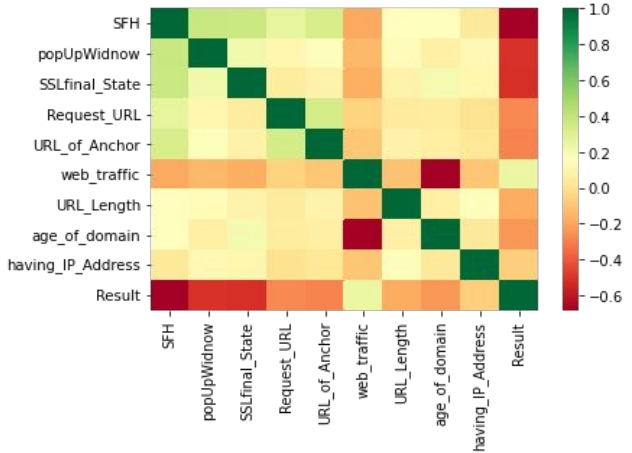
Objetivo:

- Evaluar clasificadores: Árbol de clasificación,SVM, métodos Ensemble(RandomForest, Gradient Boost, XG Boost)
- Clustering: K-means, Jerárquico Aglomerativo, DBSCAN

Dataset

- 1353 instancias
- 10 atributos
- Phishing, legítimo, sospechoso
- Datos desbalanceados(703,548,103)
- Penalización por pesos a clase mayoritaria
- Baja correlación entre atributos





Clasificación

Arboles clasificación

Desbalanceo	Accuracy	Recall	F1-score
-1	0.87	0.87	0.87
0	0.90	0.90	0.90
1	0.84	0.84	0.84
Balanceo	**Accuracy**	**Recall**	**F1-score**
-1	0.90	0.86	0.88
0	0.88	1	0.96
1	0.85	0.88	0.86

Tabla.3. Matriz de confusión clase Desbalanceada

Desbalanceo	1	-1
1	TP=95	FN=18
-1	FP=18	TN=140

Tabla.4.Matriz de confusión clase balanceadas

Balanceo	1	-1
1	TP=99	FN=14
-1	FP=17	TN=141

SVM

```
              precision    recall  f1-score
          -1       0.87      0.88      0.88
           0       1.00      0.19      0.32
           1       0.76      0.86      0.80
```

```
[121, 16]
[18, 116]
```

```
[4, 17]
[0, 250]
```

```
[97, 16]
[31, 127]
```

c>5

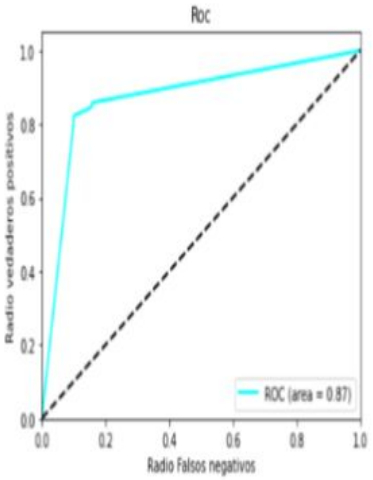
```
              precision    recall  f1-score
          -1       0.91      0.86      0.88
           0       0.95      0.95      0.95
           1       0.84      0.89      0.87
```

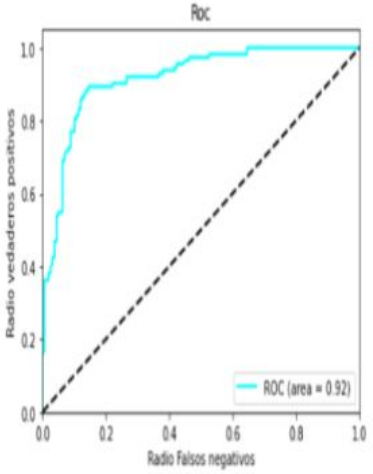
```
[118, 19]
[12, 122]
```

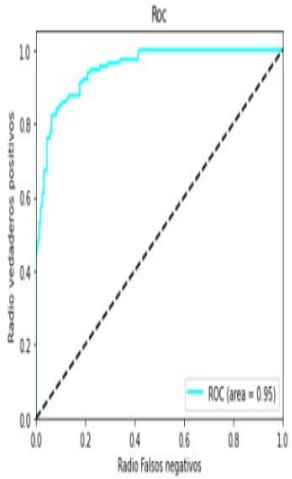
```
[20, 1]
[1, 249]
```

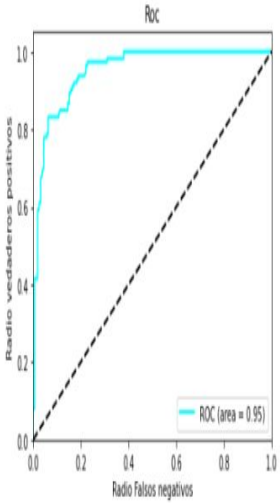
```
[101, 12]
[19, 139]
```

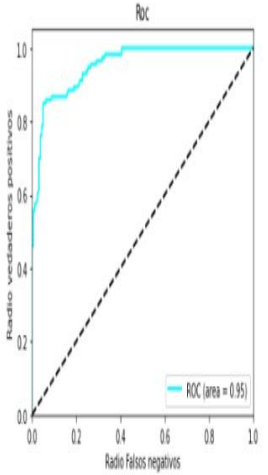
Curva ROC y AUC

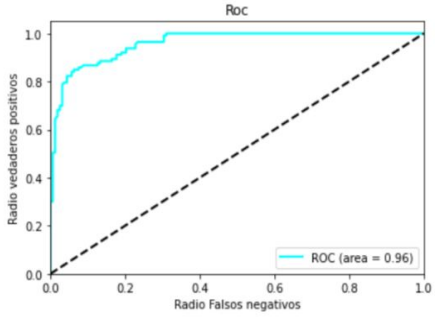










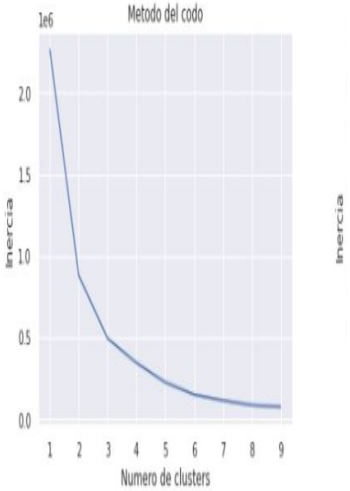


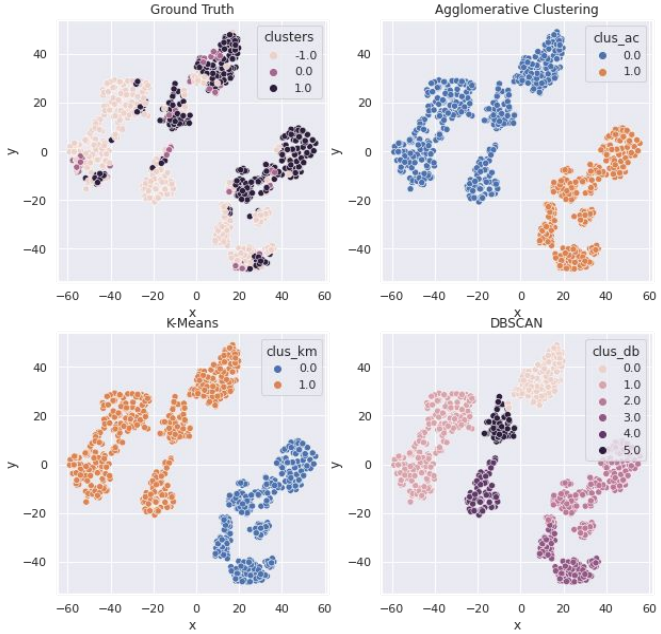
Arbol clasificación: AUC=87
SVM: AUC=87
RandomForest: AUC=95
Gradient Boost: AUC=95
XG Boost: AUC=95

XG Boost: AUC=96

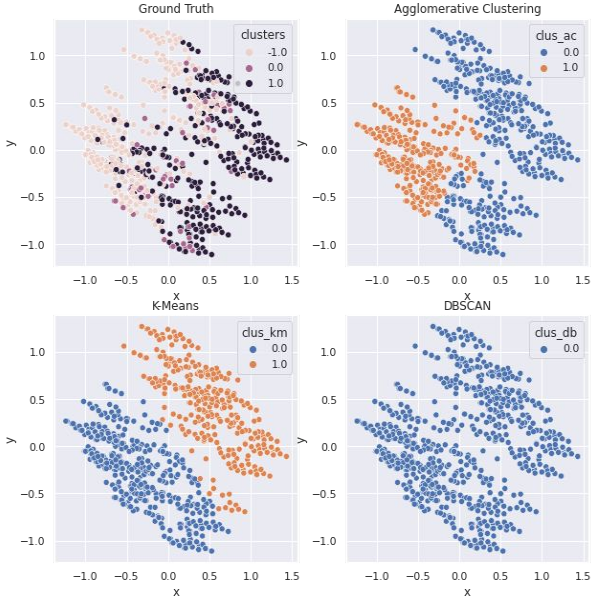
Clustering

TSNE



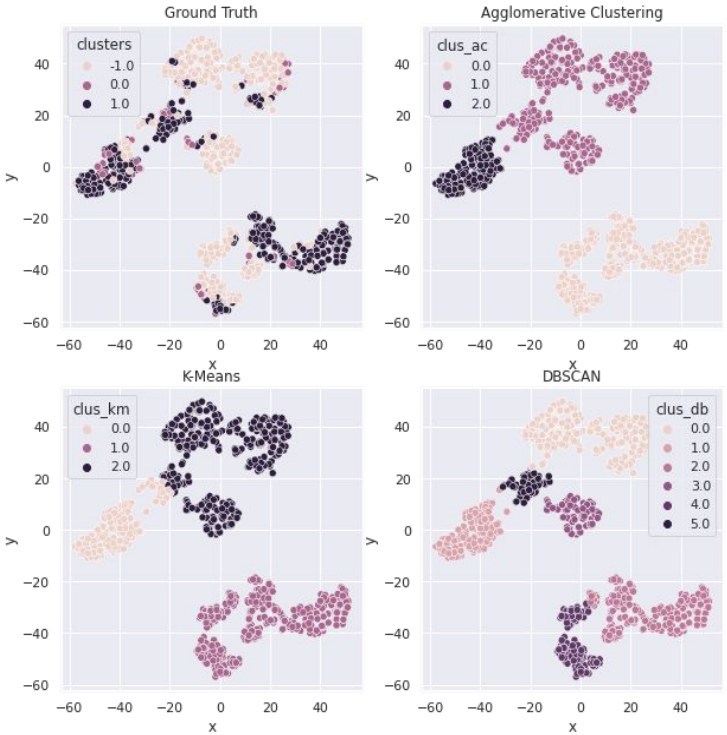


PCA

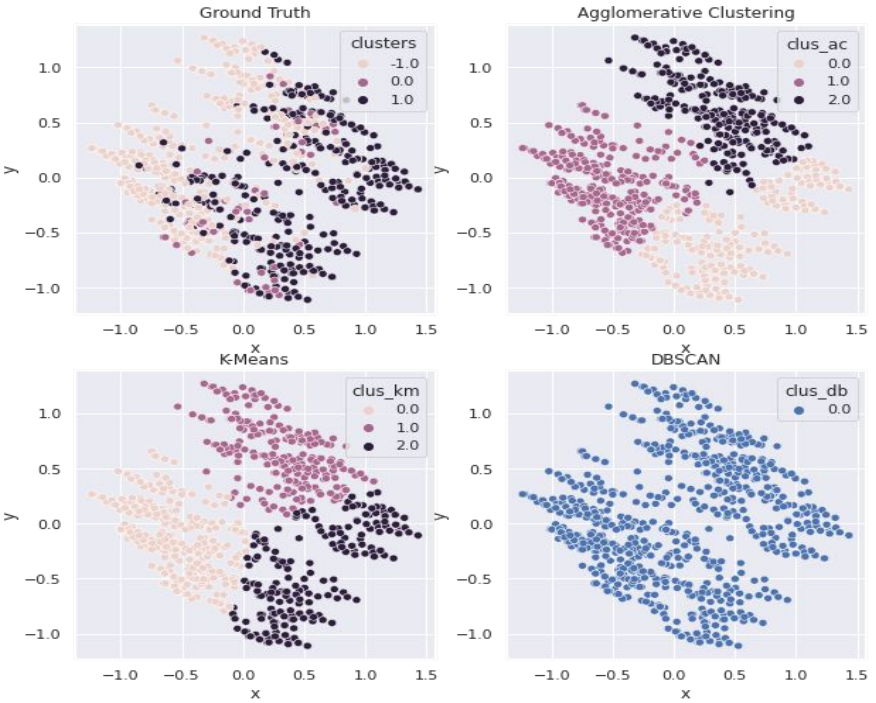


Clustering

TSNE



PCA



Conclusiones

- El phishing es una de las grandes amenazas en la actualidad y que afecta enormemente a los usuarios, a consecuencia de ataques, robos o estafas mediante páginas web. Sin embargo el uso de machine learning y técnicas de clasificación y agrupación nos ayudan a detectar estos sitios.
- Los resultados entre estos clasificadores que se obtuvieron tienen una ligera diferencia debido a la forma de implementación que tienen y las características de los dataset con los que aprenden y luego comparan para determinar si una página es pishing o no, los métodos ensemble superan a los métodos tradicionales.
- Al aplicar diferentes algoritmos de clusterización podemos observar que utilizando reducción de dimensionalidad TSNE se logra una mejor visualización y separación de grupos en el dataset.