

SISTEMAS DE RECOMENDACIÓN

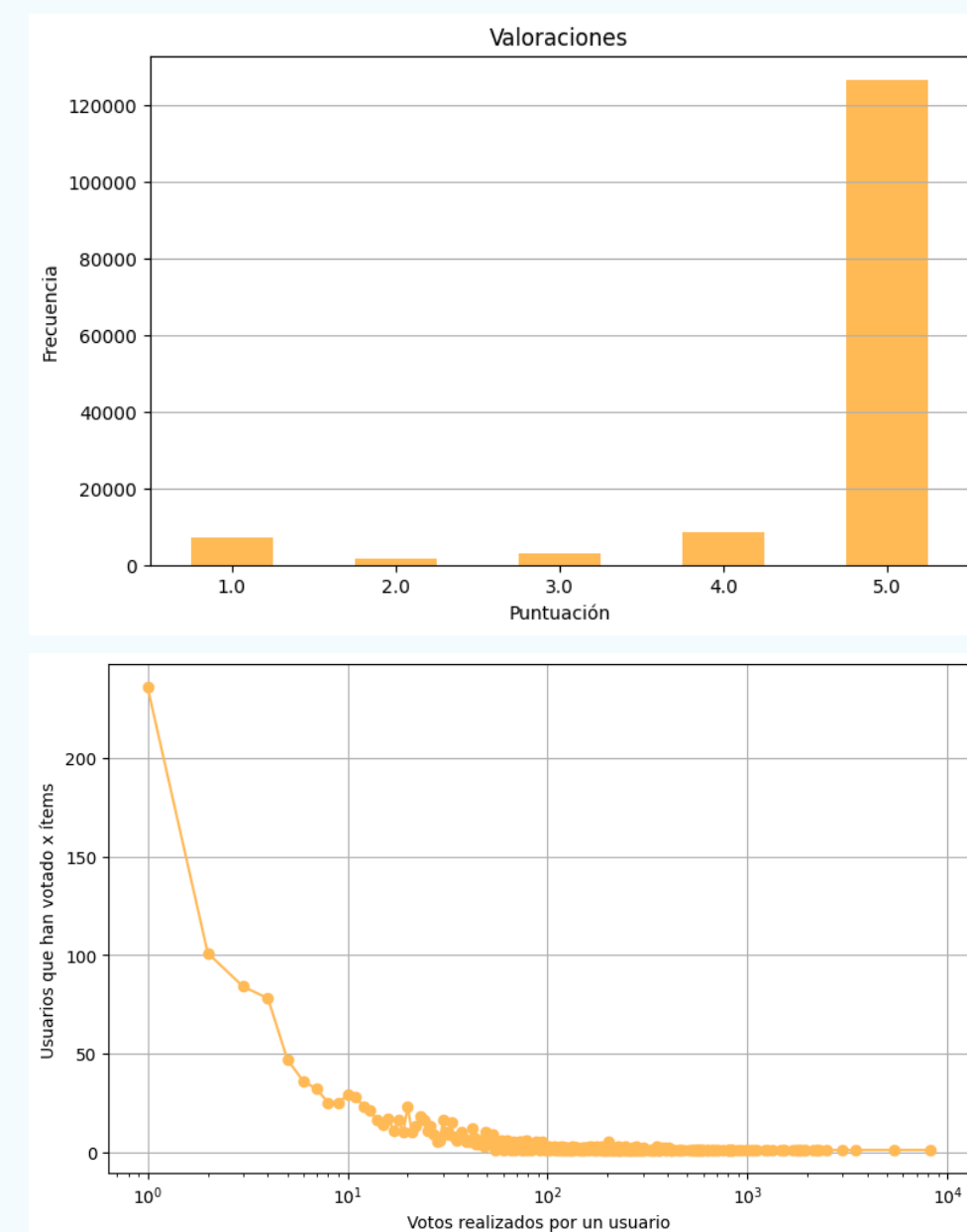
TARJETAS DE REGALO DE amazon

Análisis del dataset

Se trata de un dataset con votaciones de usuarios de Amazon a tarjetas de regalo.

- Usuarios: 1.548
- Ítems: 128.877
- Ratings: 147.136
- Rating medio: 4.67

Tenemos una matriz de ratings altamente dispersa.



KNN

Puesto que la matriz es altamente dispersa es difícil encontrar vecinos de calidad luego con k mayores que 20 se obtienen iguales o peores resultados.

Además, no se observan diferencias entre los diferentes métodos de predicción (average, weighted average y deviation from mean predictions) y tampoco utilizando diferentes métricas (jsmd y correlación) en el cálculo del MAE y RMSE.

En este método se obtiene un MAE de 0.15, RMSE de 0.21.

En cuanto a la precisión, recall y f1 se obtienen los mejores resultados usando el método de predicción average prediction con la métrica jsmd obteniendo un mismo valor para los tres de 0.0581.



Matrix factorization

Sin bias

Parámetros óptimos:

- Factores latentes: **14**
- Learning rate: **0.0001**
- Regularization: **0.2**
- Iteraciones: **50**

Métricas:

- MAE: **0.99**
- RMSE: **1.65**
- Precision: **0.1438**
- Recall: **0.1438**
- F1: **0.1438**

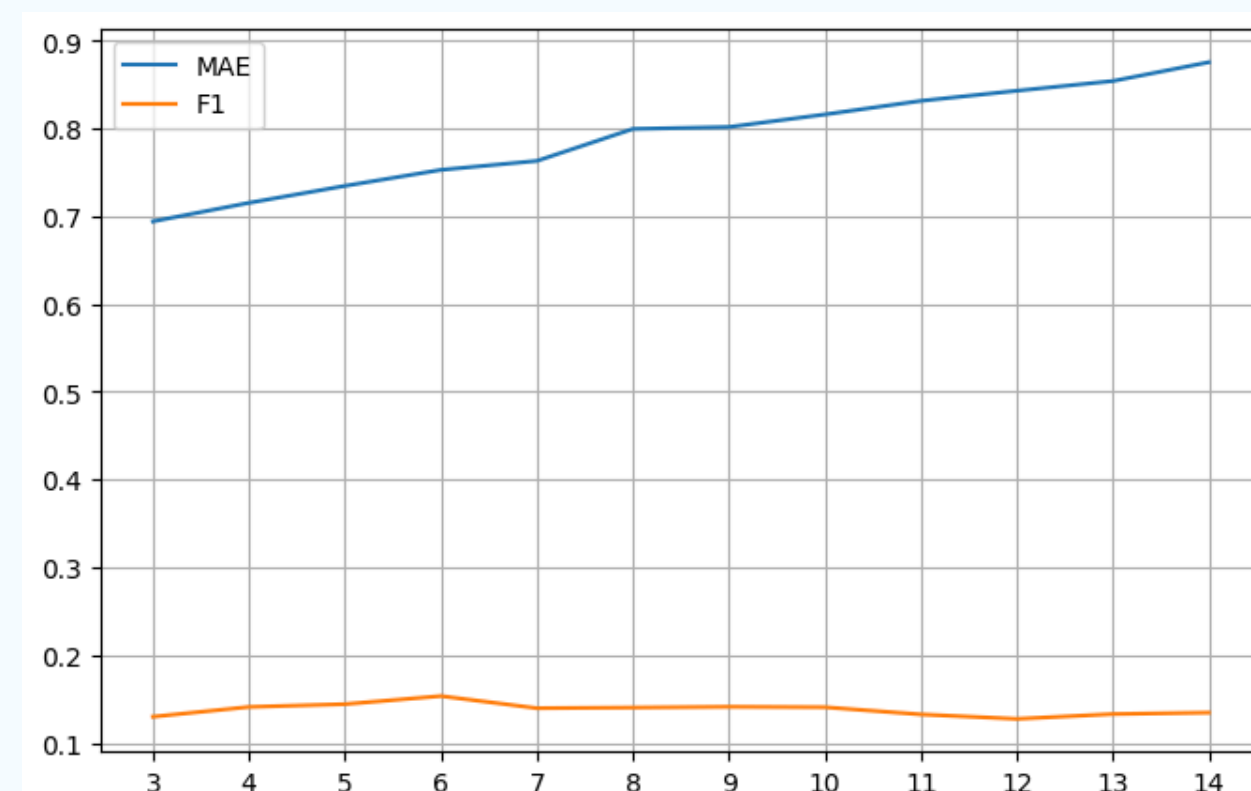
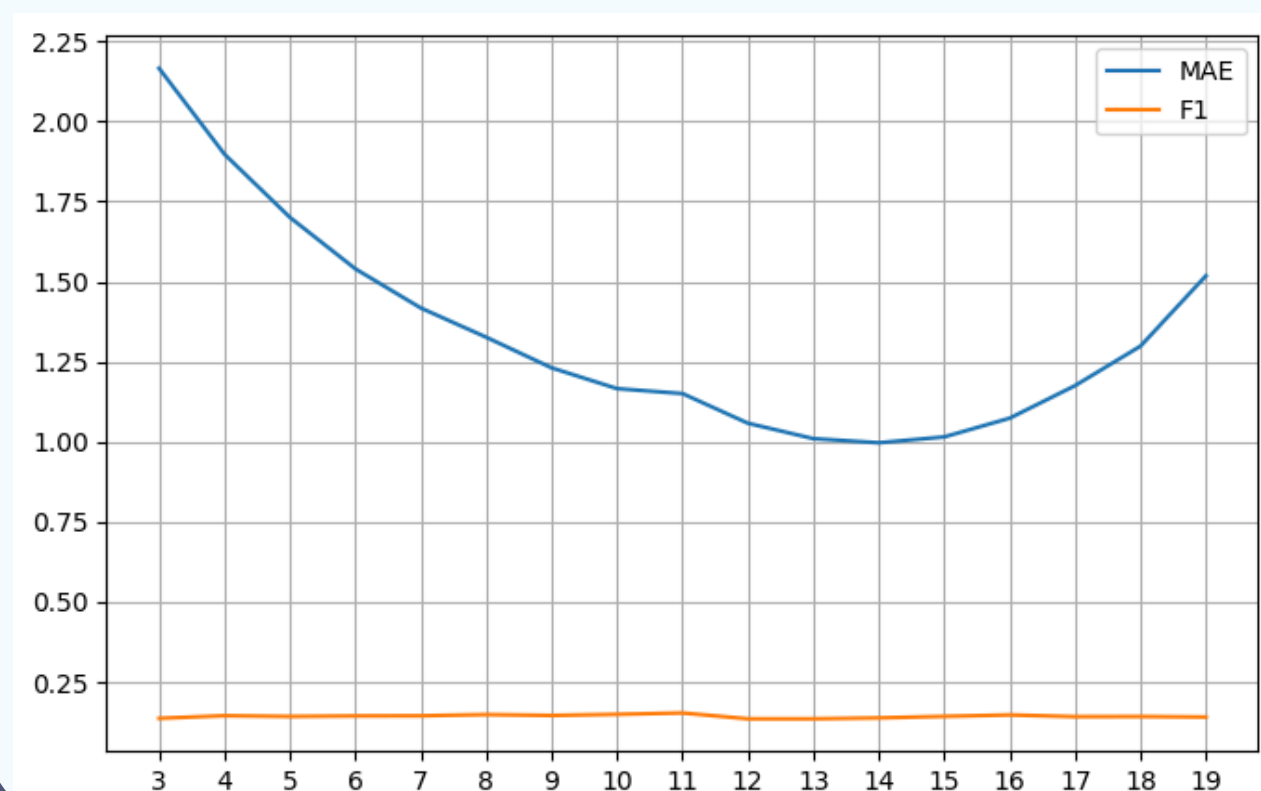
Con bias

Parámetros óptimos:

- Factores latentes:
- Learning rate: 0.0001
- Regularization: 0.2
- Iteraciones: 50

Métricas:

- MAE: **0.63**
- RMSE: **1.19**
- Precision: **0.1406**
- Recall: **0.1406**
- F1: **0.1406**



Bernoulli Matrix Factorization

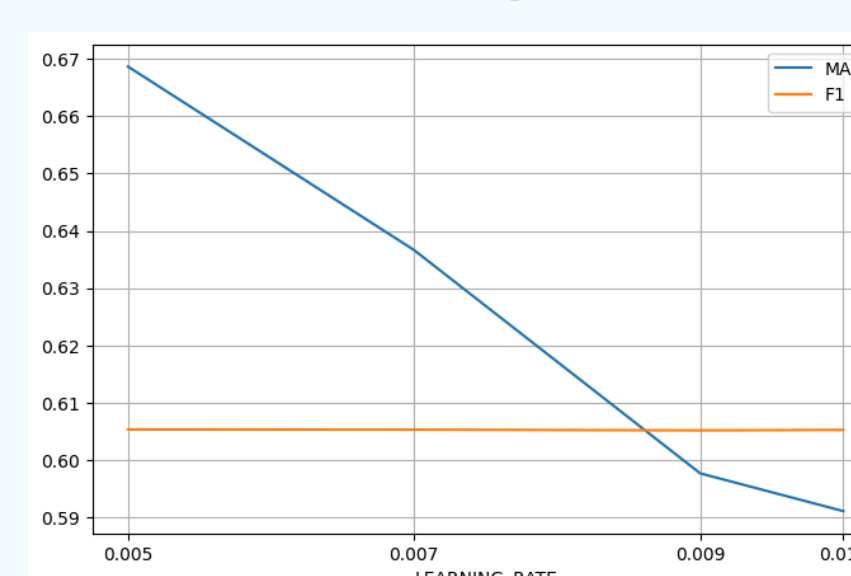
Parámetros óptimos:

- Factores latentes: **7**
- Learning rate: **0.1**
- Regularization: **0.9**
- Iteraciones: **30**

Métricas:

- MAE: **0.487**
- RMSE: **0.824**
- Precision: **0.915**
- Recall: **0.562**
- F1: **0.606**
- nDCG: **0.510**

Optimización del learning rate



Neuronal Collaborative Filtering

Generalized Matrix Factorization

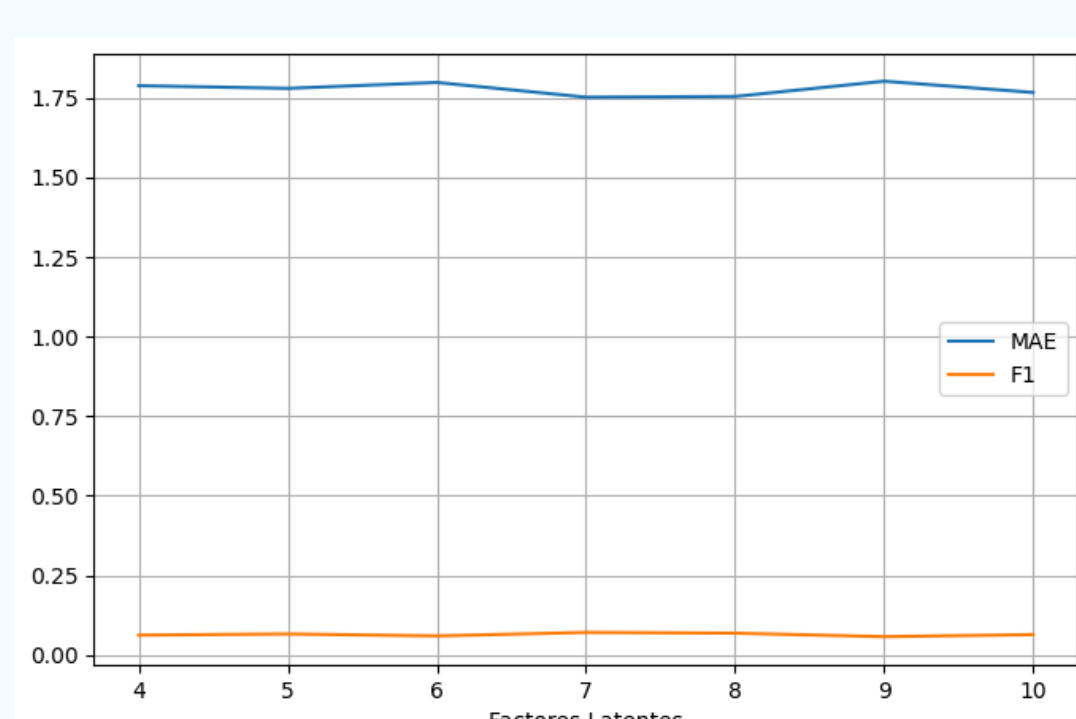
Parámetros óptimos:

- Factores latentes: **7**
- Epochs: **11**

Métricas:

- MAE: **1.778**
- RMSE: **1.921**
- Precision: **0.220**
- Recall: **0.217**
- F1: **0.062**

Optimización de los factores latentes



Multi Layer Perceptron

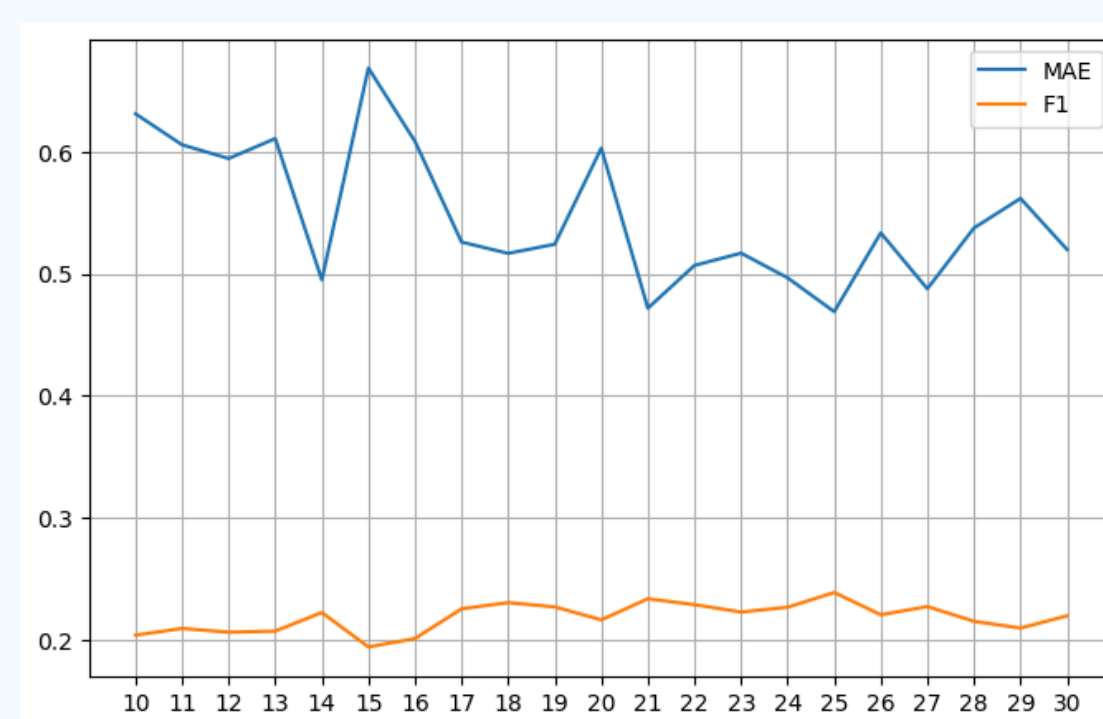
Parámetros óptimos:

- Neuronas: **50 y 25**
- Factores latentes: **20**
- Epochs: **18**

Métricas:

- MAE: **0.388**
- RMSE: **0.982**
- Precision: **0.371**
- Recall: **0.233**
- F1: **0.239**

Optimización de las neuronas



Conclusiones

1. Entre las principales conclusiones podemos destacar que la matriz de ratings era demasiado dispersa lo que causaba muchos problemas en el método de KNN.
2. Comparando los dos modelos de redes neuronales, MLP obtiene unos resultados muy superiores a los de GMF debido a que MLP puede comprender interacciones no lineales en contraste con la linealidad de GMF.