

# Programming Project

Σύστημα Συστάσεων

## Περιεχόμενα

Εισαγωγή.....	3
Αποτελέσματα .....	4

# Εισαγωγή

Αυτή η αναφορά παρουσιάζει την αξιολόγηση ενός συστήματος συστάσεων ταινιών με χρήση συνεργατικού φιλτραρίσματος. Δοκιμάστηκαν δύο σταθμίσεις ομοιότητας – Pearson και Cosine - σε διάφορες μεθόδους πρόβλεψης. Τα αποτελέσματα αξιολογούνται με βάση το **Μέσο Απόλυτο Σφάλμα (MAE)**, την Μέση **Ακρίβεια (Macro Average Precision)** και την Μέση **Ανάκληση (Macro Average Recall)**.

Διεξήχθησαν πειράματα με αναλογίες **T = 90%, 80%, 70% και 50%**.

Γειτονική χρήστες (N): Δοκιμάστηκαν 5, 10, 15, 20, 25.

## Μέθοδοι Πρόβλεψης

1. Σταθμισμένος μέσος όρος με βάση την ομοιότητα
2. Σταθμισμένος μέσος όρος που δίνει μεγαλύτερο βάρος στα δημοφιλή αντικείμενα
3. Σταθμισμένος μέσος όρος που δίνει μεγαλύτερο βάρος στα λιγότερο δημοφιλή αντικείμενα

## Μετρικές Αξιολόγησης

1. Μέσο Απόλυτο Σφάλμα (MAE) (Χαμηλότερο είναι καλύτερο)
2. Μέση Ακρίβεια (Precision) (Υψηλότερη είναι καλύτερη)
3. Μέση Ανάκληση (Recall) (Υψηλότερη είναι καλύτερη)

Για να εκτελέσετε τον κώδικα, «τρέξτε» την εντολή

- `python main.py`

Ο κώδικας είναι διαθέσιμος και στο αποθετήριο μου στο [GitHub](#)

# Αποτελέσματα

## Pearson

- Υψηλό MAE (~2.0 έως 3.1), δείχνοντας χαμηλή ακρίβεια προβλέψεων.
- Η ακρίβεια παραμένει χαμηλή (~0.5) για όλες τις τιμές του N.
- Η ανάκληση βελτιώνεται ελαφρώς καθώς αυξάνεται το N, με μέγιστη τιμή στο N = 25.
  - Μικρή επίδραση της στάθμισης δημοφιλών/λιγότερο δημοφιλών αντικειμένων.

## Cosine

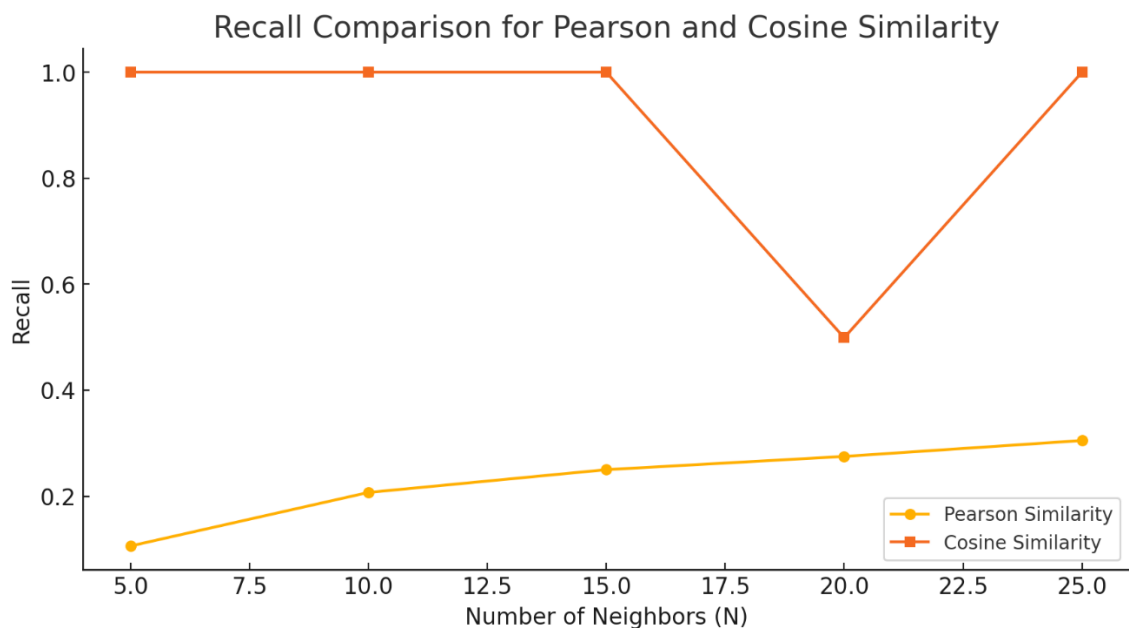
- Εξαιρετικά χαμηλό MAE (~0.03 έως 0.22), δείχνοντας υψηλή ακρίβεια προβλέψεων.
- Η ακρίβεια και η ανάκληση είναι σταθερά 1.0
- Η απόδοση παραμένει σταθερή για διαφορετικές τιμές του N

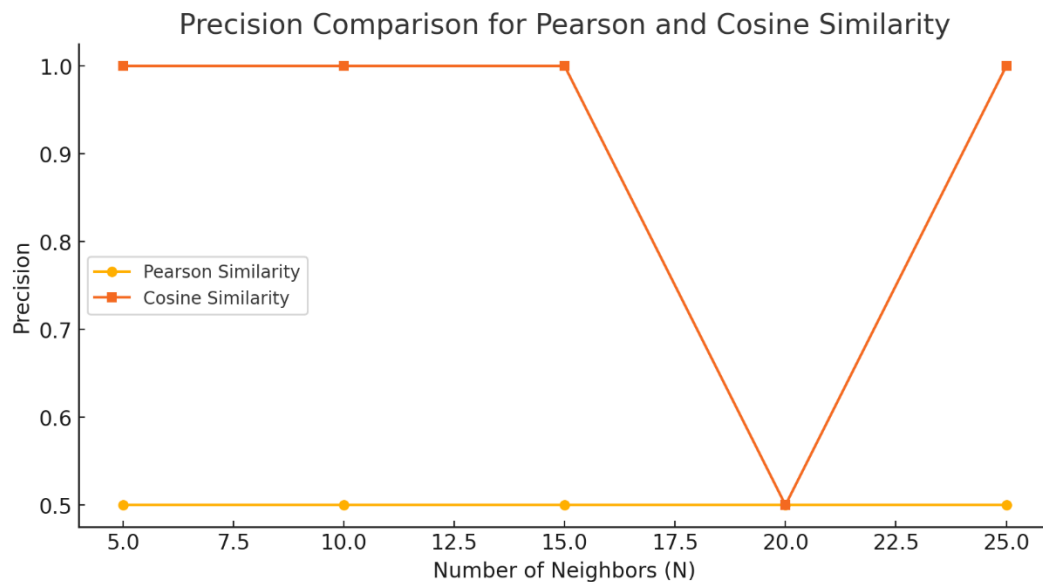
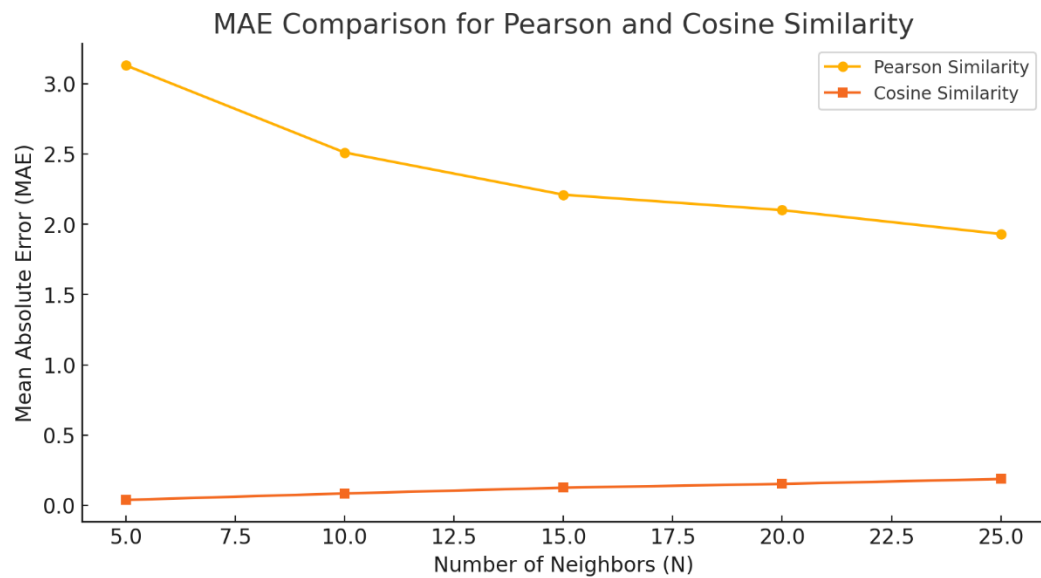
## Σημαντικές Παρατηρήσεις

- Η ομοιότητα Cosine υπερτερεί της Pearson, καθώς ο συντελεστής Pearson παρουσιάζει υψηλό MAE και χαμηλή ακρίβεια ενώ ο Cosine εμφανίζει πολύ χαμηλό MAE και τέλεια ακρίβεια/ανάκληση.
- Στην ομοιότητα Pearson, η αύξηση του N βελτιώνει την ανάκληση αλλά όχι σημαντικά το MAE. Στην ομοιότητα Cosine, η απόδοση παραμένει σταθερή, δείχνοντας ότι το N έχει μικρή επίδραση

## Γραφήματα

Παρακάτω υπάρχουν τρία γραφήματα που συγκρίνουν τα MAE, Precision και Recall για Pearson και Cosine similarity αντίστοιχα.





Στιγμιότυπα από την εκτέλεση του κώδικα:

```
T=50%, N=5, Similarity=cosine, Method=predict_weighted_unpopular
MAE: 0.09615263213546653, Precision: 0.5, Recall: 0.4968696642003415
T=50%, N=10, Similarity=pearson, Method=predict_weighted_average
MAE: 2.67368658249088, Precision: 0.5, Recall: 0.17385642194367573
T=50%, N=10, Similarity=pearson, Method=predict_weighted_popular
MAE: 2.677356408311431, Precision: 0.5, Recall: 0.17385642194367573
T=50%, N=10, Similarity=pearson, Method=predict_weighted_unpopular
MAE: 2.674576487296139, Precision: 0.5, Recall: 0.17385642194367573
T=50%, N=10, Similarity=cosine, Method=predict_weighted_average
```

```
T=90%, N=25, Similarity=pearson, Method=predict_weighted_unpopular
MAE: 2.1390008429916496, Precision: 0.5, Recall: 0.27634854771784234
T=90%, N=25, Similarity=cosine, Method=predict_weighted_average
MAE: 0.14832220110413552, Precision: 1.0, Recall: 1.0
T=90%, N=25, Similarity=cosine, Method=predict_weighted_popular
MAE: 0.14783178410570944, Precision: 1.0, Recall: 1.0
T=90%, N=25, Similarity=cosine, Method=predict_weighted_unpopular
MAE: 0.17947242271620797, Precision: 1.0, Recall: 1.0
```

Αξίζει να σημειωθεί πως περαιτέρω βελτιώσεις είναι απαραίτητες για τη δημιουργία ενός πιο ισορροπημένου και αξιόπιστου συστήματος συστάσεων.