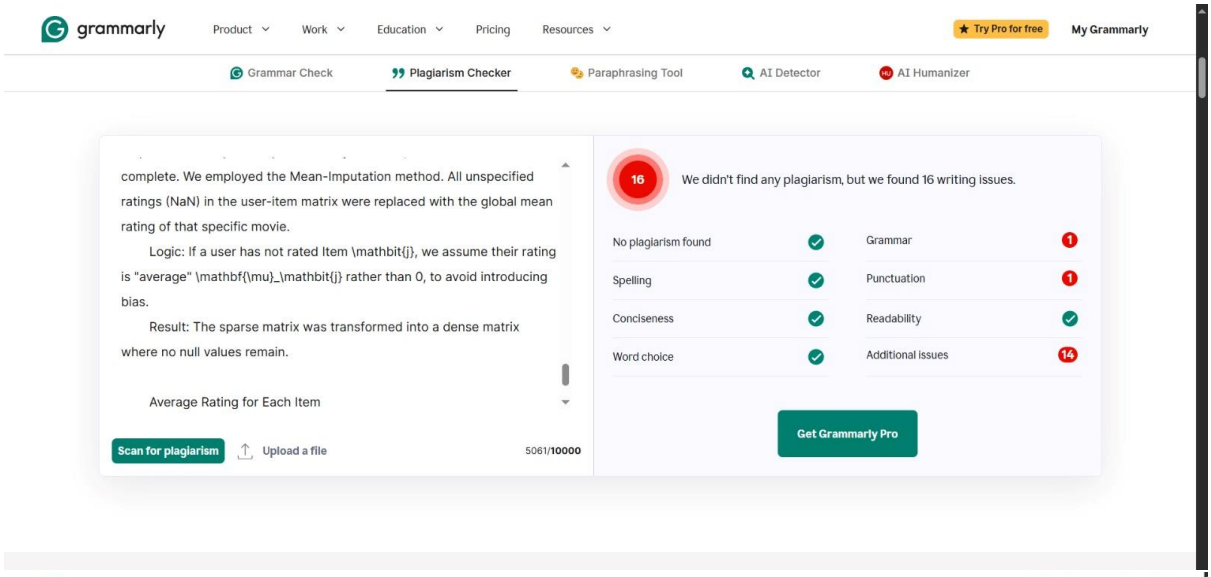
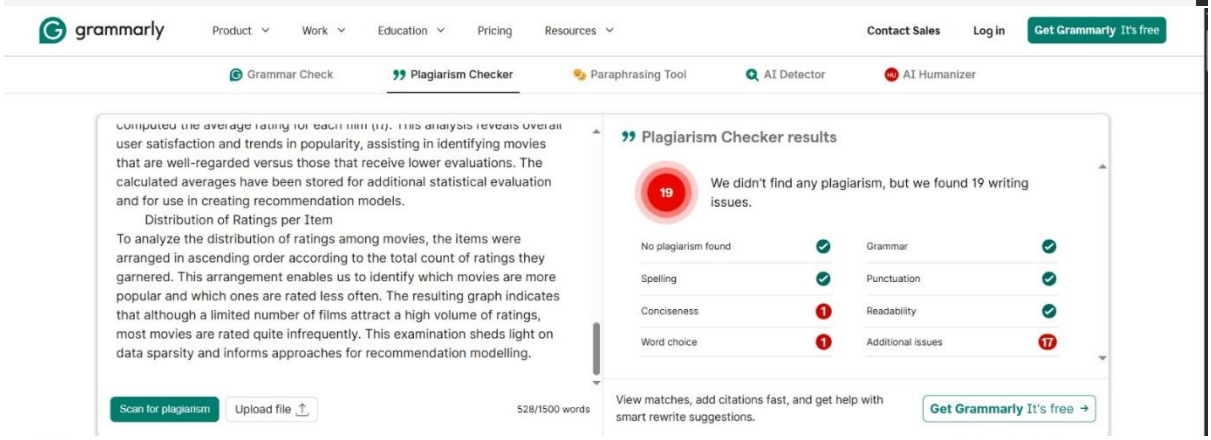
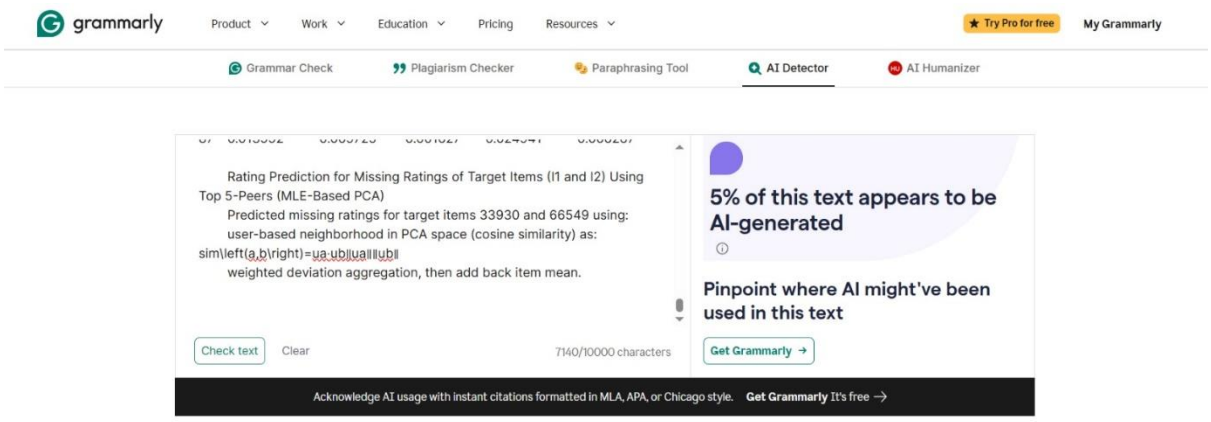
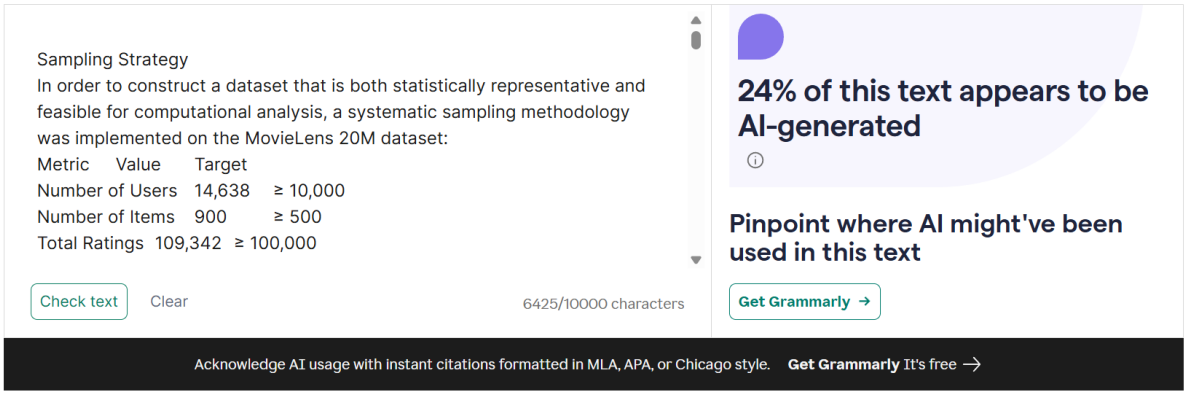
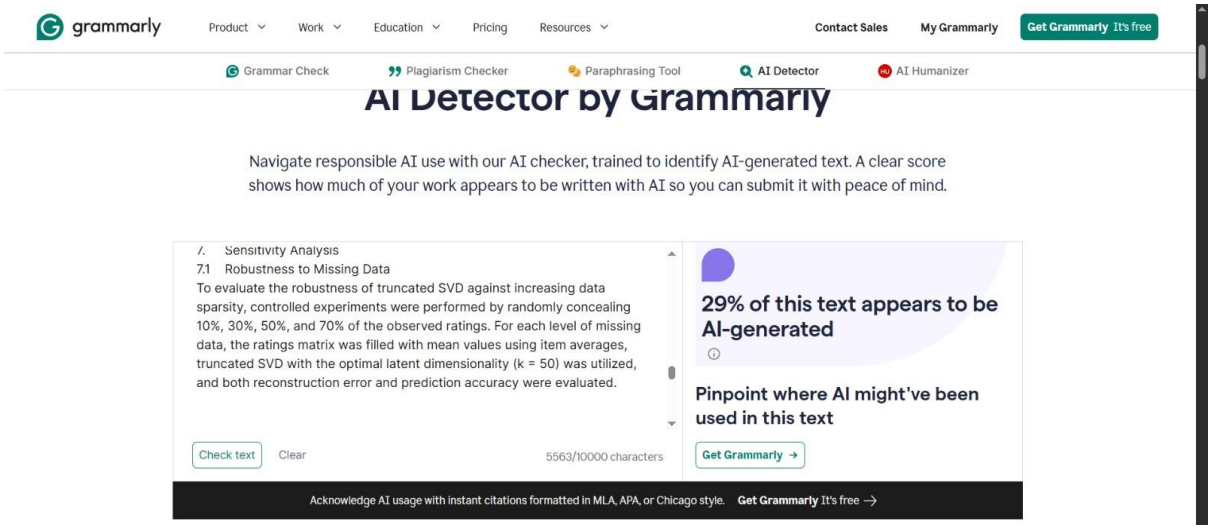
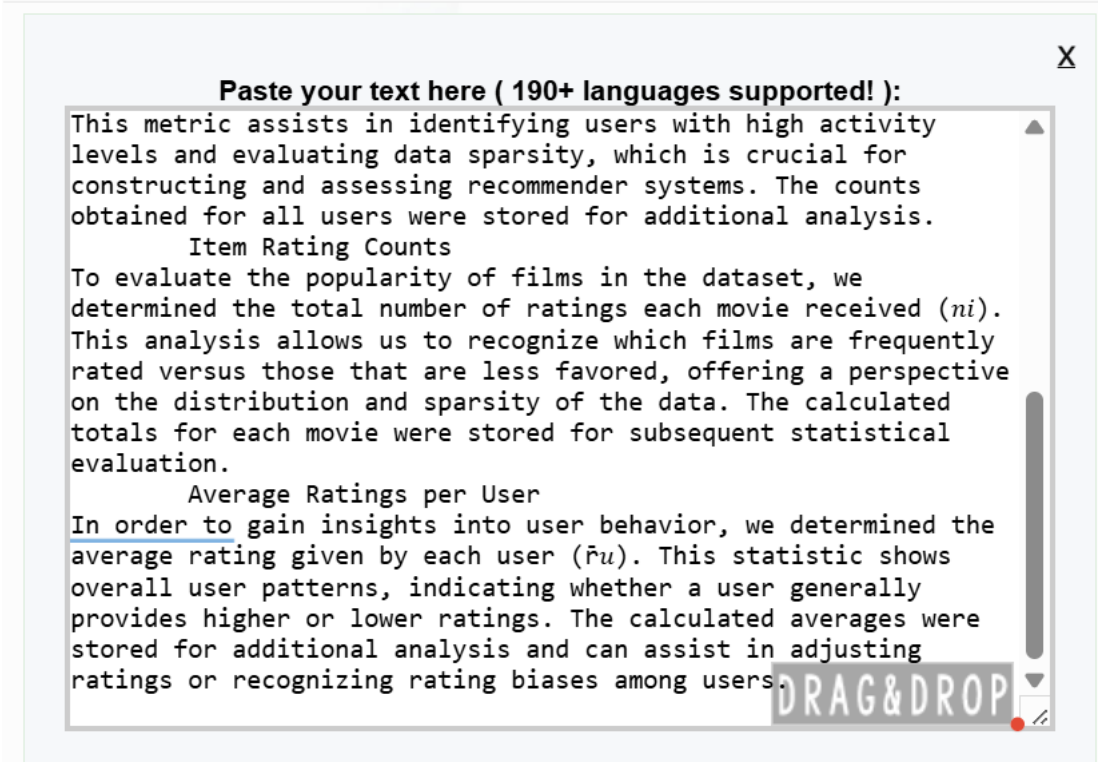

grammarly
Product
Work
Education
Pricing
Resources
Try Pro for free
My Grammarly
Grammar Check
Plagiarism Checker
Paraphrasing Tool
AI Detector
AI Humanizer
complete. We employed the Mean-Imputation method. All unspecified ratings (NaN) in the user-item matrix were replaced with the global mean rating of that specific movie.
Logic: If a user has not rated Item \mathbit{j}, we assume their rating is "average" \mathbf{\mu}_\mathbit{j} rather than 0, to avoid introducing bias.
Result: The sparse matrix was transformed into a dense matrix where no null values remain.
Average Rating for Each Item
Scan for plagiarism
Upload a file
5061/10000
16
We didn't find any plagiarism, but we found 16 writing issues.
No plagiarism found
Grammar
1
Spelling
Punctuation
1
Conciseness
Readability
Word choice
Additional issues
14
Get Grammarly Pro


grammarly
Product
Work
Education
Pricing
Resources
Contact Sales
Log in
Get Grammarly It's free
Grammar Check
Plagiarism Checker
Paraphrasing Tool
AI Detector
AI Humanizer
user satisfaction and trends in popularity, assisting in identifying movies that are well-regarded versus those that receive lower evaluations. The calculated averages have been stored for additional statistical evaluation and for use in creating recommendation models.
Distribution of Ratings per Item
To analyze the distribution of ratings among movies, the items were arranged in ascending order according to the total count of ratings they garnered. This arrangement enables us to identify which movies are more popular and which ones are rated less often. The resulting graph indicates that although a limited number of films attract a high volume of ratings, most movies are rated quite infrequently. This examination sheds light on data sparsity and informs approaches for recommendation modelling.
Scan for plagiarism
Upload file
528/1500 words
Plagiarism Checker results
19
We didn't find any plagiarism, but we found 19 writing issues.
No plagiarism found
Grammar
Spelling
Punctuation
Conciseness
1
Readability
Word choice
1
Additional issues
17
View matches, add citations fast, and get help with smart rewrite suggestions.
Get Grammarly It's free


grammarly
Product
Work
Education
Pricing
Resources
Try Pro for free
My Grammarly
Grammar Check
Plagiarism Checker
Paraphrasing Tool
AI Detector
AI Humanizer
Rating Prediction for Missing Ratings of Target Items (I1 and I2) Using Top 5-Peers (MLE-Based PCA)
Predicted missing ratings for target items 33930 and 66549 using:
user-based neighborhood in PCA space (cosine similarity) as:
sim\left(a,b\right)=ua.ub||ua||ub||
weighted deviation aggregation, then add back item mean.
Check text
Clear
7140/10000 characters
5% of this text appears to be AI-generated
Pinpoint where AI might've been used in this text
Get Grammarly
Acknowledge AI usage with instant citations formatted in MLA, APA, or Chicago style. Get Grammarly It's free


Sampling Strategy
In order to construct a dataset that is both statistically representative and feasible for computational analysis, a systematic sampling methodology was implemented on the MovieLens 20M dataset:
Metric Value Target
Number of Users 14,638 ≥ 10,000
Number of Items 900 ≥ 500
Total Ratings 109,342 ≥ 100,000
Check text
Clear
6425/10000 characters
24% of this text appears to be AI-generated
Pinpoint where AI might've been used in this text
Get Grammarly →
Acknowledge AI usage with instant citations formatted in MLA, APA, or Chicago style. Get Grammarly It's free →


grammarly
Product
Work
Education
Pricing
Resources
Contact Sales
My Grammarly
Get Grammarly It's free
Grammar Check
Plagiarism Checker
Paraphrasing Tool
AI Detector
AI Humanizer
AI Detector by Grammarly
Navigate responsible AI use with our AI checker, trained to identify AI-generated text. A clear score shows how much of your work appears to be written with AI so you can submit it with peace of mind.
7. Sensitivity Analysis
7.1 Robustness to Missing Data
To evaluate the robustness of truncated SVD against increasing data sparsity, controlled experiments were performed by randomly concealing 10%, 30%, 50%, and 70% of the observed ratings. For each level of missing data, the ratings matrix was filled with mean values using item averages, truncated SVD with the optimal latent dimensionality (k = 50) was utilized, and both reconstruction error and prediction accuracy were evaluated.
Check text
Clear
5563/10000 characters
29% of this text appears to be AI-generated
Pinpoint where AI might've been used in this text
Get Grammarly →
Acknowledge AI usage with instant citations formatted in MLA, APA, or Chicago style. Get Grammarly It's free →


X
Paste your text here (190+ languages supported!):
This metric assists in identifying users with high activity levels and evaluating data sparsity, which is crucial for constructing and assessing recommender systems. The counts obtained for all users were stored for additional analysis.
Item Rating Counts
To evaluate the popularity of films in the dataset, we determined the total number of ratings each movie received (ni). This analysis allows us to recognize which films are frequently rated versus those that are less favored, offering a perspective on the distribution and sparsity of the data. The calculated totals for each movie were stored for subsequent statistical evaluation.
Average Ratings per User
In order to gain insights into user behavior, we determined the average rating given by each user (r̄u). This statistic shows overall user patterns, indicating whether a user generally provides higher or lower ratings. The calculated averages were stored for additional analysis and can assist in adjusting ratings or recognizing rating biases among users.
DRAG&DROP

100% Unique

Total 1876 chars , 277 words, 14 unique sentence(s).

Results	Query	Domains (original links)
Unique	The dataset contains ratings on a 0.5 to 5 scale	-
Unique	To standardize the data for statistical analysis and subsequent recommendation algorithms, the ratings were rescaled	-
Unique	R_new= ((R_old- 〖Min〗_old)/(〖Max〗_old- 〖Min〗_old))×(〖Max〗_new- 〖Min〗_new) 〖Min〗_new R_new= ((R_old-0.5)/4.5)×(4) 1 This transformation ensures that all ratings are	-
Unique	The rescaled ratings now have a minimum value of 1 and a maximum value of	-
Unique	Statistical Analysis Before conducting dimensionality reduction, a comprehensive statistical analysis was carried out to gain	-
Unique	User Rating Counts To gain insight into user behavior within the dataset, the total amount	-
Unique	This metric assists in identifying users with high activity levels and evaluating data sparsity, which	-
Unique	The counts obtained for all users were stored for additional analysis	-
Unique	Item Rating Counts To evaluate the popularity of films in the dataset, we determined the	-
Unique	This analysis allows us to recognize which films are frequently rated versus those that are	-
Unique	The calculated totals for each movie were stored for subsequent statistical evaluation	-