



AIE425 Intelligent Recommender Systems, Fall Semester

25/26

FINAL COURSE PROJECT PLAGRISM REPORT

GROUP 4

221101140 NOURELDEEN AHMED MAHER MESBAH

221101573 YOUSEF MOHAMED ABDELWHAB

221101030 YOUSEF ZAKARIA SOBHY

221100244 BASSANT KAMAL MESILAM

SUBMISSION DATE: MONDAY, JANUARY 5,2026

As we don't have access to Paid Plagiarism Checker Tools , we will be taking 5 random samples from the text inside the report and uploading each of them to 2 different Free tools for checking (Grammarly & ZeroGPT)

Plagiarism Checker

Ensure every word is your own with Grammarly's AI-powered plagiarism checker, which uses advanced AI to detect plagiarism in your text and check for other writing issues.

The screenshot shows the Grammarly interface for a sample text. On the left, there's a sidebar with 'Methods Compared' (Mean-Fill Point 9 vs MLE Point 4) and 'Prediction Results' (MLE is better). At the bottom are 'Scan for plagiarism' and 'Upload a file' buttons, and a progress bar at 6154/10000.

Results Summary: We didn't find any plagiarism, but we found 19 writing issues.

Category	Count
No plagiarism found	✓
Grammar	10
Spelling	1
Punctuation	✓
Conciseness	✓
Readability	✓
Word choice	4
Additional issues	4

Get Grammarly Pro button is visible at the bottom right.

Plagiarism Checker

Ensure every word is your own with Grammarly's AI-powered plagiarism checker, which uses advanced AI to detect plagiarism in your text and check for other writing issues.

The screenshot shows the Grammarly interface for a sample text. On the left, there's a sidebar with 'Aspect' and 'Winner' columns, and a note about Maximum Likelihood Estimation. At the bottom are 'Scan for plagiarism' and 'Upload a file' buttons, and a progress bar at 5134/10000.

Results Summary: We didn't find any plagiarism, but we found 15 writing issues.

Category	Count
No plagiarism found	✓
Grammar	7
Spelling	✓
Punctuation	1
Conciseness	✓
Readability	✓
Word choice	3
Additional issues	4

Get Grammarly Pro button is visible at the bottom right.

Plagiarism Checker

Ensure every word is your own with Grammarly's AI-powered plagiarism checker, which uses advanced AI to detect plagiarism in your text and check for other writing issues.

2.3 Feature Extraction (Content-Based)
To understand streamer content, we extracted features from:

1. Text: Game titles, genres, and themes (from IGDB).
- o Technique: TF-IDF Vectorization (max_features=1000, min_df=2).
2. Numerical: Popularity metrics (Average Viewers, Followers).
- o Technique: Log-transformation (\log_{10}) followed by Min-Max Scaling to handle power-law usage distribution.
3. Combination: Features were weighted (90% Text / 10% Numerical) to prioritize content relevance over raw popularity.

Scan for plagiarism Upload a file 8409/10000

Get Grammarly Pro

28 We didn't find any plagiarism, but we found 28 writing issues.

No plagiarism found	✓	Grammar	9
Spelling	3	Punctuation	8
Conciseness	✓	Readability	✓
Word choice	1	Additional issues	7

Plagiarism Checker

Ensure every word is your own with Grammarly's AI-powered plagiarism checker, which uses advanced AI to detect plagiarism in your text and check for other writing issues.

optimized SVD), delivering statistically significant improvements over non-hybrid baselines

7. Appendices
Appendix A: Key Code Snippets
Snippets from

- Collaborative.py
- Hybrid.py
- Content_based.py

Scan for plagiarism Upload a file 6494/10000

Get Grammarly Pro

! We have found plagiarism in your text and have also detected 17 writing issues.

Plagiarism found	!	Grammar	3
Spelling	2	Punctuation	✓
Conciseness	1	Readability	✓
Word choice	✓	Additional issues	11

Plagiarism Checker

Ensure every word is your own with Grammarly's AI-powered plagiarism checker, which uses advanced AI to detect plagiarism in your text and check for other writing issues.

The screenshot shows a document being analyzed for plagiarism. The document content includes sections on start users, cascade hybrid architectures, future directions, and future work. Below the document, there are buttons for 'Scan for plagiarism' and 'Upload a file', and a character count of 3148/10000. To the right, a summary section indicates '7' writing issues found. A table details the findings across categories: Plagiarism found (green checkmark), Grammar (red circle with '2'), Spelling (green checkmark), Punctuation (green checkmark), Conciseness (green checkmark), Readability (green checkmark), Word choice (green checkmark), and Additional issues (red circle with '5'). A 'Get Grammarly Pro' button is also present.

Trusted GPT-5, ChatGPT and AI Detector tool by ZeroGPT

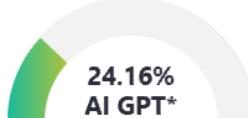
ZeroGPT the most Advanced and Reliable Chat GPT, GPT5 & AI Content Detector

The screenshot displays the ZeroGPT interface. On the left, there are several icons representing different features. The main area shows a list of items for comparison, including 'Top-5: Less computation, good for sparse users' and 'Top-10: Better accuracy, recommended for active users'. Below this is a section titled '1.6-' comparing 'point 9' and 'point 4'. It lists 'Methods Compared': 'Mean-Fill Point 9: k-NN with cosine similarity in 5D latent space' and 'MLE Point 4: Reconstruction formula ($\mu_i + \mathbb{I}_{t,u,p} \times W_{i,p}$) in 5D latent space'. Under 'Prediction Results (Top-5 PCs)', it shows a table:

User	Item	Mean-Fill (Point 9)	MLE (Point 4)	Better
U1	I1	3.30 (err=0.39)	3.69 (err=0.00)	MLE
U1	I2	3.10 (err=0.64)	3.74 (err=0.00)	MLE

At the bottom, there are buttons for 'Detect Text' and 'Upload File', and a character limit message: '6,154/15,000 Characters' and 'Check 350,000 characters, [Upgrade Here](#)'.

Your Text is Most Likely Human written, may include parts generated by AI/GPT



Analysis and Comments

1. Top-10 PCs Generally Perform Better

Top-10 PCs achieve lower average error (0.39) compared to Top-5 PCs (0.51). This is expected because:

Trusted GPT-5, ChatGPT and AI Detector tool by ZeroGPT

ZeroGPT the most Advanced and Reliable Chat GPT, GPT5 & AI Content Detector

PCA MLE with Top-10 Principal Components is the recommended approach for rating prediction.
The computational overhead of MLE is justified by its significantly superior prediction accuracy and statistical validity.

Final Verdict
Aspect Winner
Prediction Accuracy MLE
Variance Explained MLE
Statistical Validity MLE
Computational Simplicity Mean-Fill
Overall MLE
Maximum Likelihood Estimation fundamentally transforms PCA-based recommendation by respecting the observed data structure, leading to dramatically better rating predictions

Detect Text 5,134/15,000 Characters
Check 350,000 characters, [Upgrade Here](#)

Your Text is Most Likely Human written, may include parts generated by AI/GPT



Prediction Results (Top-10 PCs)

User Item Mean-Fill (Point 11) MLE (Point 6) Better

U1 I1 3.55 (err=0.14) 3.69 (err=0.00) MLE

Trusted GPT-5, ChatGPT and AI Detector tool by ZeroGPT

ZeroGPT the most Advanced and Reliable Chat GPT, GPT5 & AI Content Detector

- Users: ~90,000 unique users.
- Items (Streamers): ~1,400 active streamers.
- Rating Scale: Implicit feedback converted to 1-5 scale (based on watch time/frequency).

2.3 Feature Extraction (Content-Based)
To understand streamer content, we extracted features from:

1. Text: Game titles, genres, and themes (from IGDB).
 - o Technique: TF-IDF Vectorization (max_features=1000, min_df=2).
2. Numerical: Popularity metrics (Average Viewers, Followers).
 - o Technique: Log-transformation (\log_{10}) followed by Min-Max Scaling to handle power-law usage distribution.
3. Combination: Features were weighted (90% Text / 10% Numerical) to prioritize content relevance over raw popularity.

Detect Text **Upload File** 8,409/15,000 Characters
Check 350,000 characters, [Upgrade Here](#)

Your Text is Most Likely Human written

17.63%
AI GPT*

1.9- Discussion and Conclusion for PART 3

a) Summary of Findings

Trusted GPT-5, ChatGPT and AI Detector tool by ZeroGPT

ZeroGPT the most Advanced and Reliable Chat GPT, GPT5 & AI Content Detector

The screenshot shows the ZeroGPT AI Detector tool interface. On the left, there is a vertical sidebar with icons for different features: a brain, a person, a document, a gear, a chart, a file, and a person with a gear. The main area displays a text snippet from a document:

games with missing metadata receive poorer recommendations.

6.3 Conclusion
The developed Hybrid Recommender System successfully meets the project objectives. It provides a robust solution that handles the spectrum of users from cold-start (via content-based) to power users (via optimized SVD), delivering statistically significant improvements over non-hybrid baselines

7. Appendices
Appendix A: Key Code Snippets
Snippets from

- Collaborative.py
- Hybrid.py
- Content_based.py

Below this, there are two buttons: "Detect Text" (in a blue button) and "Upload File" (with an upload icon). To the right, it shows "6,494/15,000 Characters" and "Check 350,000 characters, [Upgrade Here](#)".

The screenshot shows the AI detection results. At the top, it says "Your Text is Human written". Below that is a circular progress bar with a green segment. The text "11.33% AI GPT*" is displayed next to the bar. A cursor arrow is visible on the right side of the screen.

3- Implementation

3.1 System Architecture: Cascade Hybrid

Trusted GPT-5, ChatGPT and AI Detector tool by ZeroGPT

ZeroGPT the most Advanced and Reliable Chat GPT, GPT5 & AI Content Detector

 cold-start scenarios and filters irrelevant content, while collaborative filtering captures nuanced user preference patterns for ranking.

 Recommendations

Based on our findings, we recommend Truncated SVD with k=100 latent factors for production systems requiring scalability and real-time performance. For applications where prediction accuracy is paramount and computational resources are available, PCA MLE with Top-10 principal components provides optimal results. All deployed systems should incorporate content-based fallbacks to ensure coverage for cold-start users, and cascade hybrid architectures should be considered for balancing accuracy with computational efficiency.

 Future Directions

Future work should explore deep learning approaches for latent factor extraction, dynamic user profile modeling to capture preference drift over time, and integration of additional contextual signals such as temporal patterns and social connections. The successful implementation of the Twitch.tv hybrid system provides a foundation for extending these techniques to other streaming and content recommendation domains.

























































































<img alt="Icon of a document with a gear" data-bbox="131 3878 158