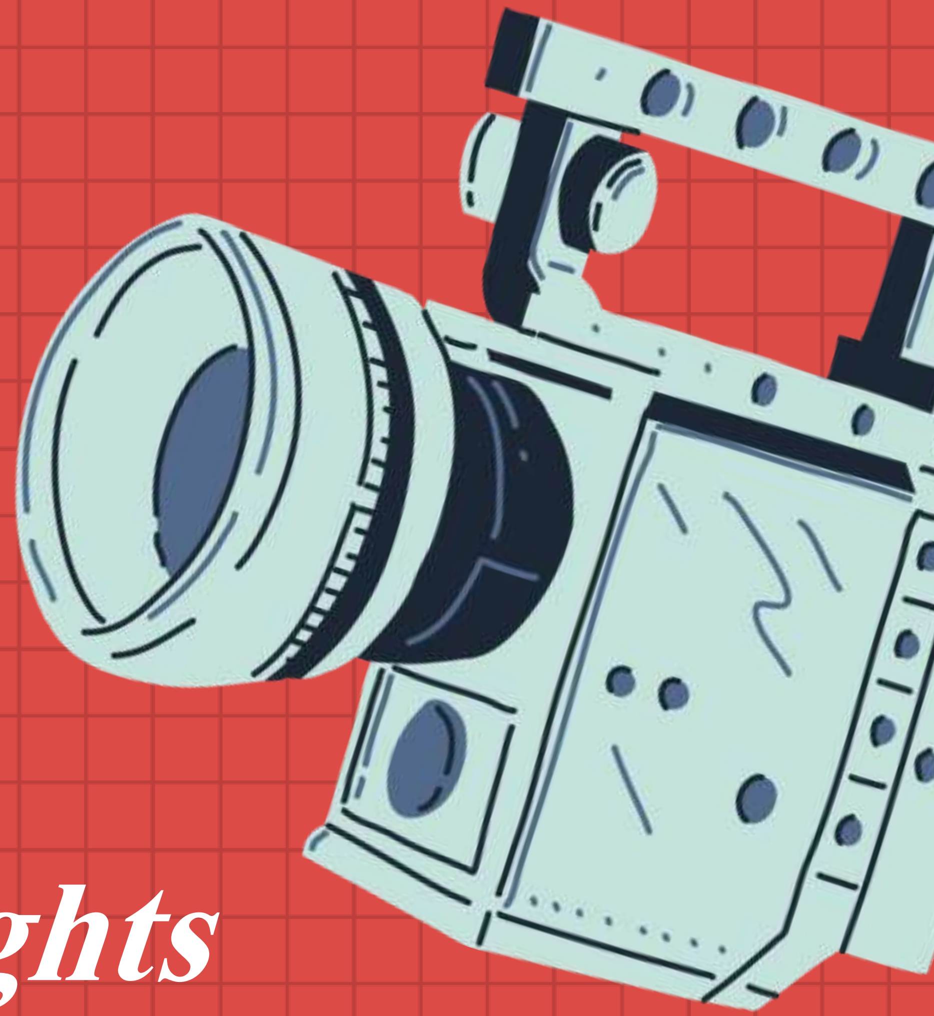
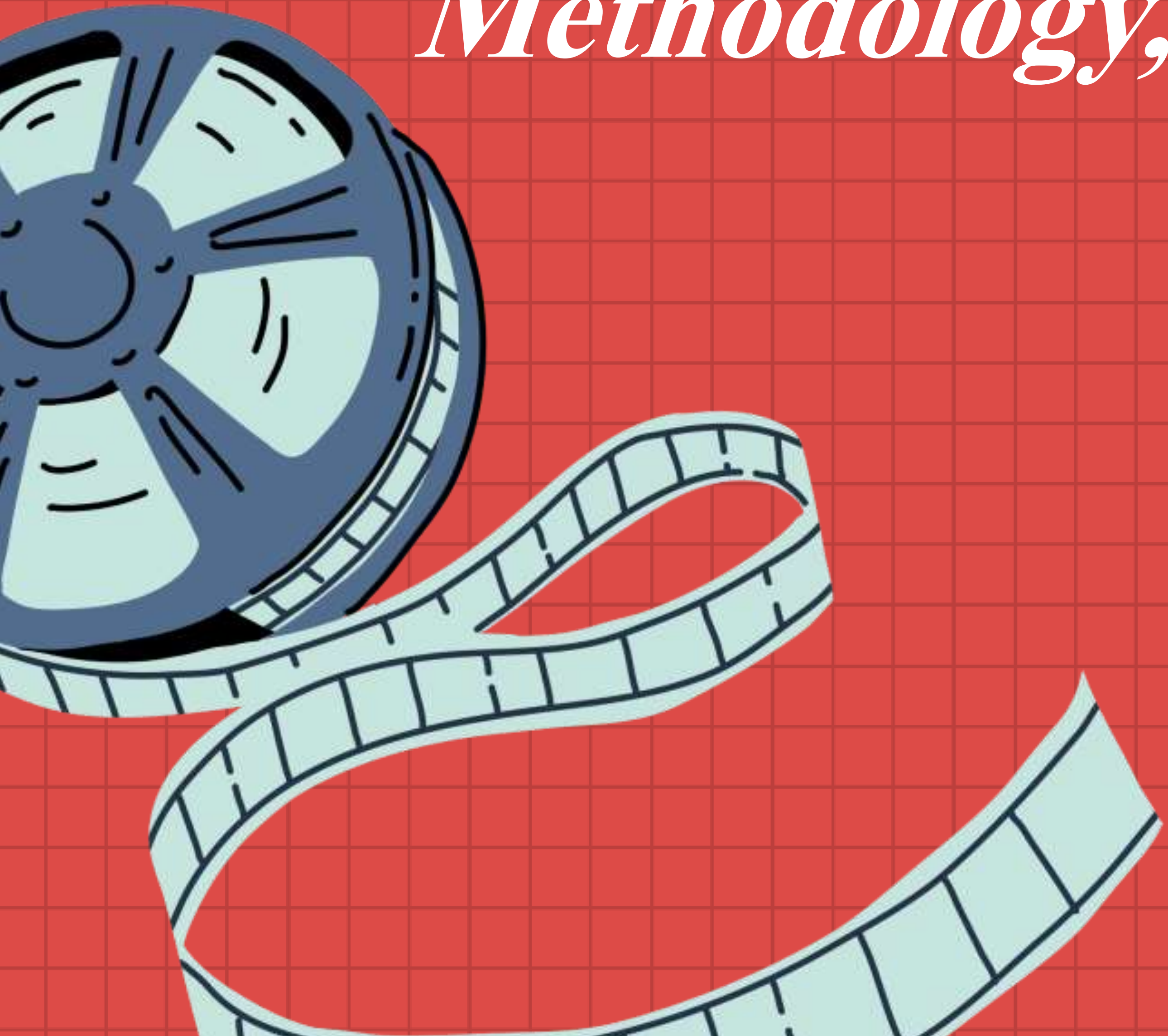
Northeastern University

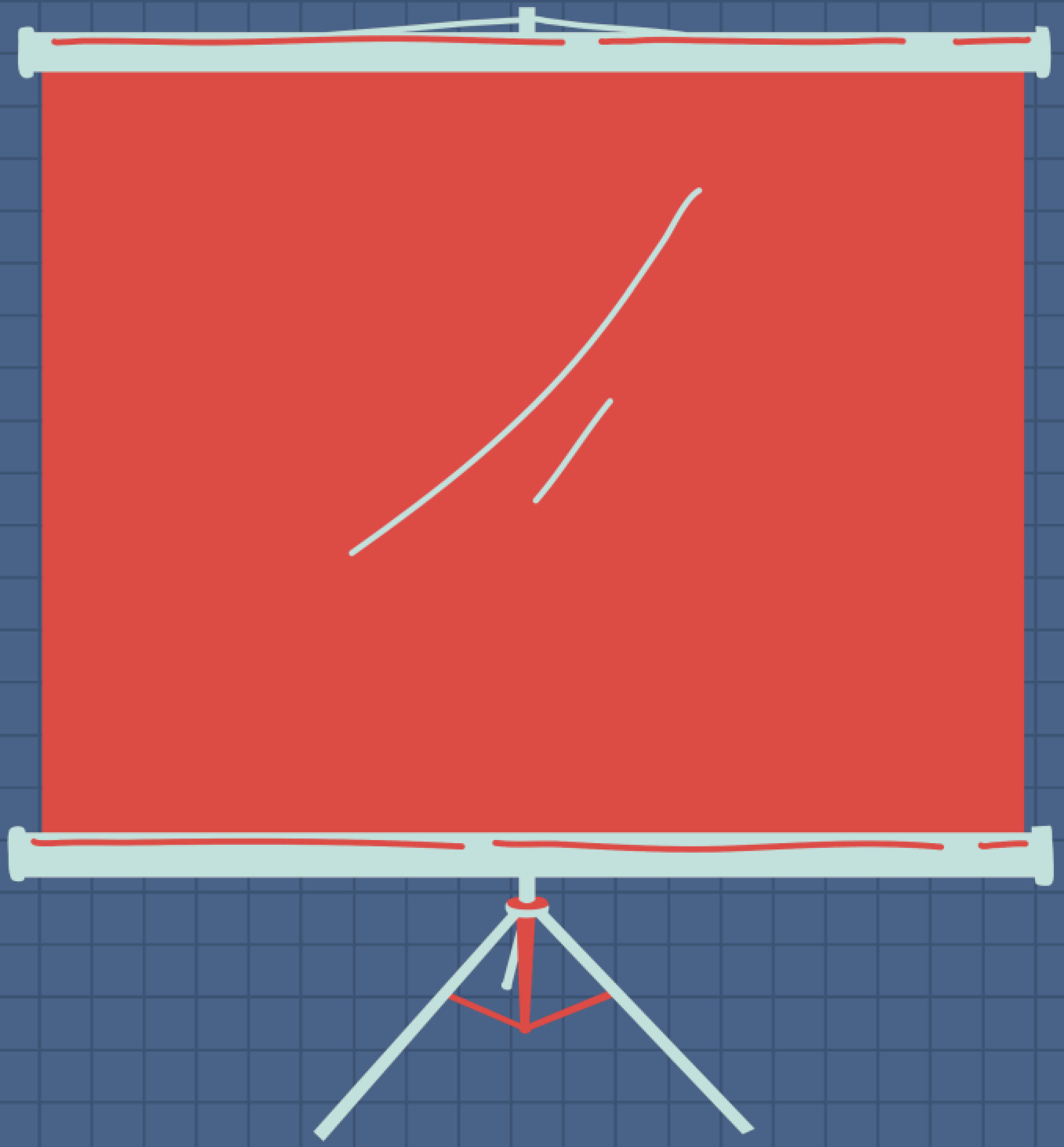
Ad Videos Sentiment Analysis:
Logistic Regression, TF-IDF, and Predictive Analytics.



Methodology, Results & Predictive Insights



Mohit Ravindra Kamble
Northeastern University
July 15, 2024



Objective

The goal of this task was to utilize video advertisements and their corresponding text descriptions and speech captions to answer 21 binary (yes/no) questions.

The results were documented, agreement percentages calculated, and F1 scores, precision, and recall metrics computed against the provided ground-truth data.

A classifier was developed to maximize these metrics.

This objective guided the methodology and the evaluation process.

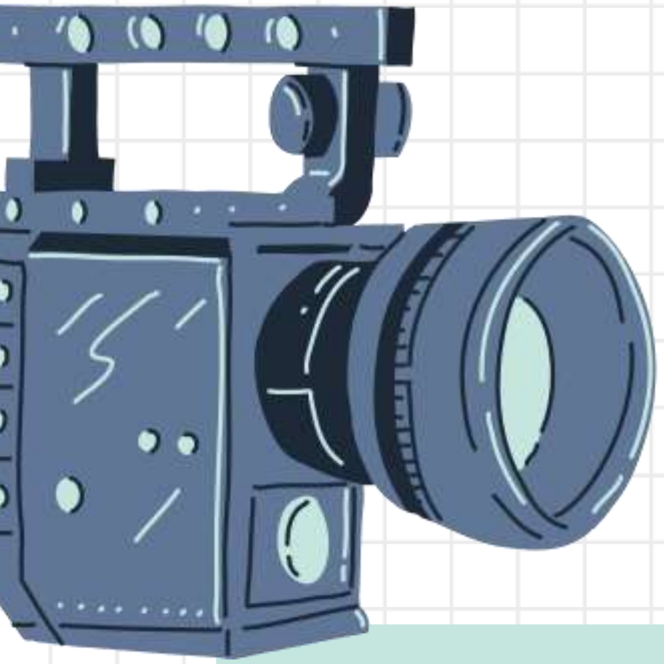
Dataset Overview

- **The dataset for this task consisted of 150 video advertisements from various companies.**

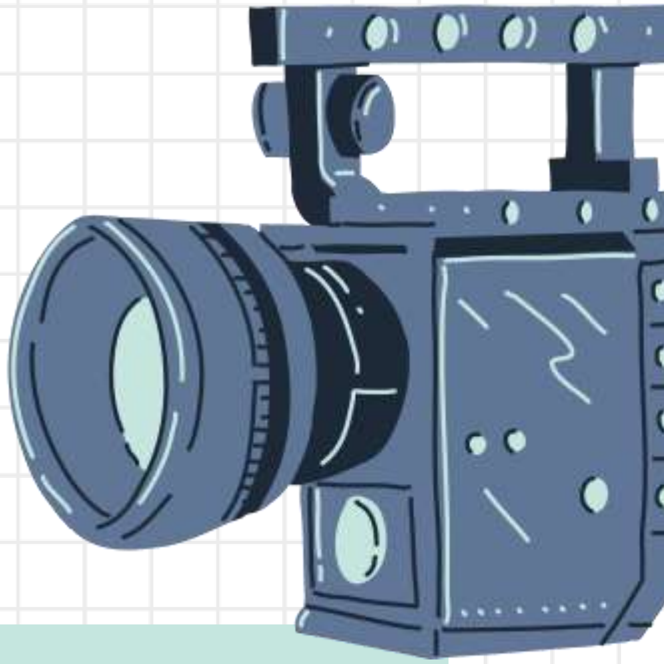
- **The textual data included:**
  - **Ad Campaign Descriptions:** Provided by the companies.
  - **Transcriptions:** Generated from the speech in the videos.
  - **On-Screen Text:** Extracted from the video content.

- **The ground truth dataset consisted of answers from human coders to 21 yes/no questions.**

- **The majority vote among the human coder's answers was considered the ground truth for evaluation purposes.**



Methodology Overview



The methodology for this task involved four key steps

Data Preprocessing	Text Processing	Model Training	Evaluation Metrics
Cleaning and merging datasets.	Using a TF-IDF Vectorizer to represent text data.	Training Logistic Regression models for each question.	Calculating agreement percentage, precision, recall, and F1 scores.

Methodology Justification:

- Logistic regression was chosen due to its simplicity and effectiveness in binary classification tasks.
- It works well when the relationship between features and the target variable is approximately linear, making it suitable for the provided dataset.

Data Preprocessing

Data preprocessing was a crucial step to ensure the quality of the input data.

The steps included:

- **Merging Datasets:** Combined the sample & ground-truth datasets on the common identifier ***creative_data_id***.

- **Dropping Unwanted Columns:** Removed irrelevant columns to streamline the data.

- **Cleaning Text Data:** Performed text cleaning by converting text to lowercase. This step ensured the textual data was uniform and free of noise.

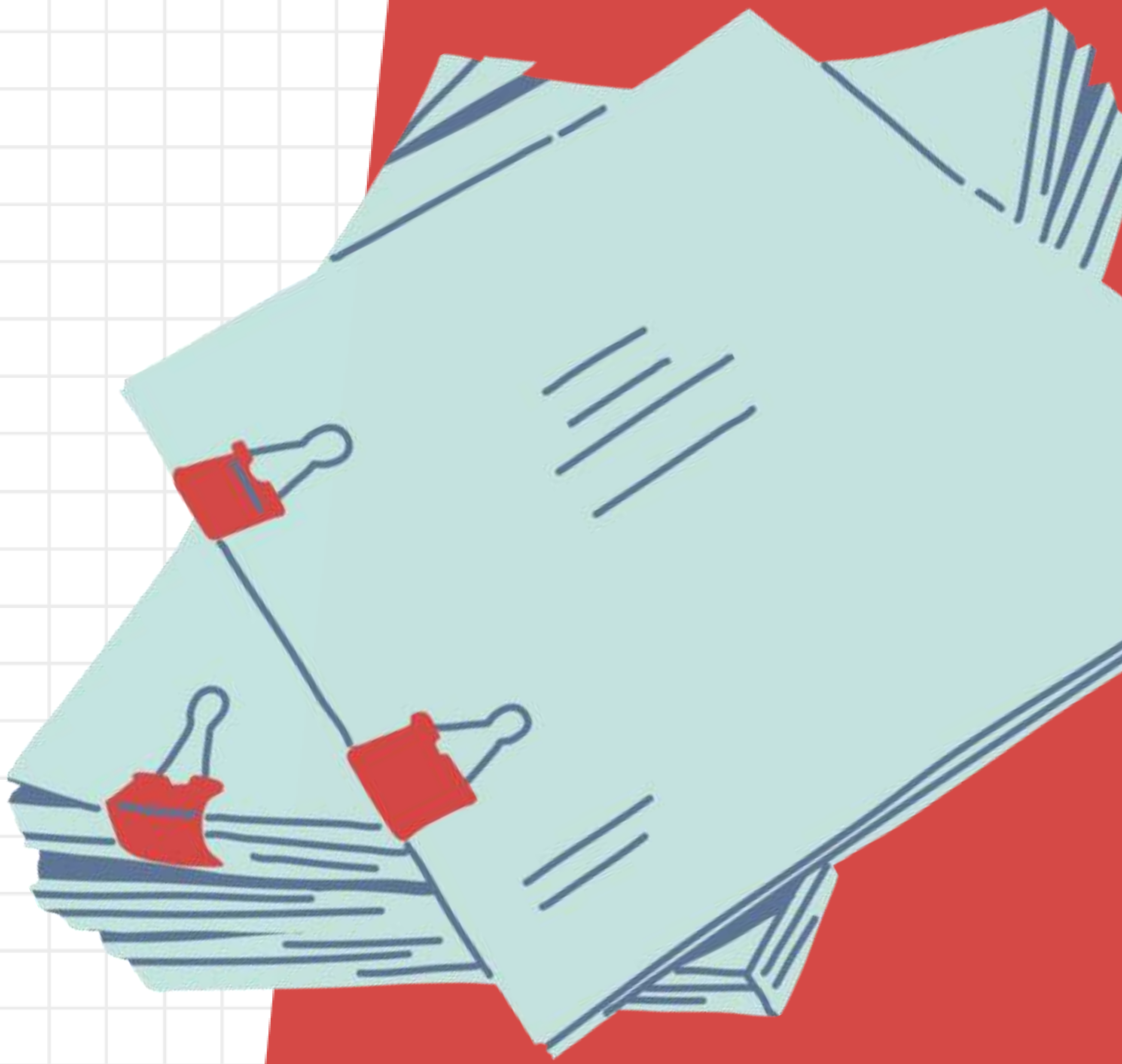
Text Processing

Text data was processed using a ‘TF-IDF Vectorizer’ which transforms text into numerical features.

This involved:

- **TF-IDF Vectorizer:** The Term Frequency-Inverse Document Frequency (TF-IDF) Vectorizer was used to extract features from the cleaned speech data.

- **Feature Representation:** Selected the 5000 most frequent terms as features. This representation helps in capturing the importance of terms relative to the entire corpus.



Model Training

Model Training Using Logistic Regression

For each of the 21 questions, a separate Logistic Regression model was trained.

The steps included:

- **Classifier:** Logistic Regression was chosen due to its simplicity and effectiveness in binary classification tasks.
- **Training Process:** The models were trained on the processed text data for each question.
- **Evaluation Metrics:** Metrics such as accuracy, precision, recall, and F1 score were calculated to evaluate the performance of the models.

Evaluation Metrics

The performance of the models was evaluated using the following metrics:

Precision

The ratio of true positive predictions to the total predicted positives.

Recall

The ratio of true positive predictions to the total actual positives.

F1 Score

The harmonic mean of precision and recall, providing a single measure of a model's accuracy.

Agreement Percentage: The percentage of predictions that matched the ground truth. This is calculated as the accuracy of the model expressed as a percentage:

Agreement Percentage = Accuracy × 100

Results Overview and Insights

	agreement_percentage	precision	f1_score	recall
Is there a call to go online (e.g., shop online, visit the Web)?	0.74	0.55	0.63	0.74
Is there online contact information provided (e.g., URL, website)?	0.97	0.97	0.97	0.97
Is there a visual or verbal call to purchase (e.g., buy now, order now)?	0.69	0.79	0.61	0.69
Does the ad portray a sense of urgency to act (e.g., buy before sales ends, order before ends)?	0.75	0.57	0.65	0.75
Is there an incentive to buy (e.g., a discount, a coupon, a sale or "limited time offer")?	0.91	0.92	0.91	0.91
Is there offline contact information provided (e.g., phone, mail, store location)?	0.83	0.69	0.76	0.83
Is there mention of something free?	0.93	0.86	0.89	0.93
Does the ad mention at least one specific product or service (e.g., model, type, item)?	0.84	0.71	0.77	0.84
Is there any verbal or visual mention of the price?	0.71	0.80	0.64	0.71
Does the ad show the brand (logo, brand name) or trademark (something that most people know is the brand) multiple times?\n\nFor example, Nike ads often have the "swoosh" logo prominently displayed on shoes and apparel worn by celebrity athletes. The "Just Do It" slogan is another Nike trademark frequently included.	0.85	0.72	0.78	0.85
Does the ad show the brand or trademark exactly once at the end of the ad?	0.85	0.73	0.79	0.85
Is the ad intended to affect the viewer emotionally, either with positive emotion (fun, joy), negative emotion (sad, anxious) or another type of emotion? (Note: You may not personally agree, but assess if that was the intention.)	0.79	0.63	0.70	0.79
Does the ad give you a positive feeling about the brand?	0.83	0.68	0.75	0.83
Does the ad have a story arc, with a beginning and an end?	0.77	0.59	0.67	0.77
Does the ad have a reversal of fortune, where something changes for the better, or changes for the worse?	0.87	0.76	0.81	0.87
Does the ad have relatable characters?	0.99	0.99	0.99	0.99
Is the ad creative/clever?	0.82	0.86	0.80	0.82
Is the ad intended to be funny? (Note: You may not personally agree, but assess if that was the intention.)	0.81	0.65	0.72	0.81
Does this ad provide sensory stimulation (e.g., cool visuals, arousing music, mouth-watering)?	0.97	0.97	0.97	0.97
Is the ad visually pleasing?	0.69	0.79	0.57	0.69
Does the ad have cute elements like animals, babies, animated, characters, etc?	0.77	0.60	0.67	0.77

Results Overview and Insights

- **Agreement Percentages:**
  - High agreement percentages for clear questions like "*Is there online contact information provided?*" (**97%**) & "*Is there an incentive to buy?*" (**91%**).
  - Lower agreement percentages for subjective questions such as "*Is the ad visually pleasing?*" (69%) & "*Is there a visual or verbal call to purchase*" (**69%**).

- **Precision:**
  - High precision for objective questions like "*Does the ad have relatable characters?*" (**99%**) and "*Is there online contact information provided?*" (**97%**).
  - Lower precision for subjective questions like "*Is there a call to go online?*" (**55%**) and "*Does the ad portray a sense of urgency to act?*" (**57%**).

- **Recall:**
  - High recall for questions with clear textual indicators like "*Does the ad have relatable characters?*" (**99%**) and "*Is there online contact information provided?*" (**97%**).
  - Lower recall for questions like "*Is there a visual or verbal call to purchase?*" (**69%**) and "*Is there any verbal or visual mention of the price?*" (**71%**).

- **F1 Score:**
  - High F1 scores for balanced performance in questions like "*Is there mention of something free?*" (**89%**) and "*Does the ad have a reversal of fortune, where something changes for the better, or changes for the worse?*" (**81%**).
  - Lower F1 scores for subjective questions like "*Is there a call to go online?*" (**63%**) and "*Does the ad have cute elements like animals, babies, animated, characters, etc?*" (**67%**).

Insights on Model Performance

Effectiveness on Objective Content:

- The model shows strong performance on objective questions, evident from high agreement percentages, precision, recall, and F1 scores.
- This suggests it is effective at handling straightforward, unambiguous content.

Challenges with Subjective Content:

- Lower performance metrics on subjective questions indicate the model struggles with nuanced features that require deeper understanding.
- This is likely due to the complexity of subjective features and diverse human interpretations.

Objective Questions:

- Is there online contact information provided (e.g., URL, website)?
- Is there an incentive to buy (e.g., a discount, a coupon, a sale or "limited time offer")?
- Does the ad show the brand (logo, brand name) or trademark multiple times?
- Does the ad have relatable characters?
- Does this ad provide sensory stimulation (e.g., cool visuals, arousing music, mouth-watering)?

Subjective Questions:

- Is the ad intended to affect the viewer emotionally?
- Is the ad visually pleasing?
- Is the ad creative/clever?
- Is the ad intended to be funny?
- Does the ad give you a positive feeling about the brand?

#Note:

For few questions: The agreement percentages, precision, f1 score and recall are the same.
This means that the model has classified an equal amount of answers as false positives, as it classified false negatives.

Bonus Questions Analysis

Video Classification Issues:

- Some videos might not work well with the classifier due to:
  - **Ambiguous Content:** Videos with unclear or ambiguous messaging can lead to inconsistent answers.
  - **Complex Narratives:** Videos with complex or multi-layered narratives might confuse the classifier.

Human Coders' Responses Analysis:

- **Variation in Responses:** Human coder’s responses varied, reflecting subjective interpretations.
- **Consistency Issues:** Differences in interpretation and attention to detail among coders impacted ground-truth consistency.

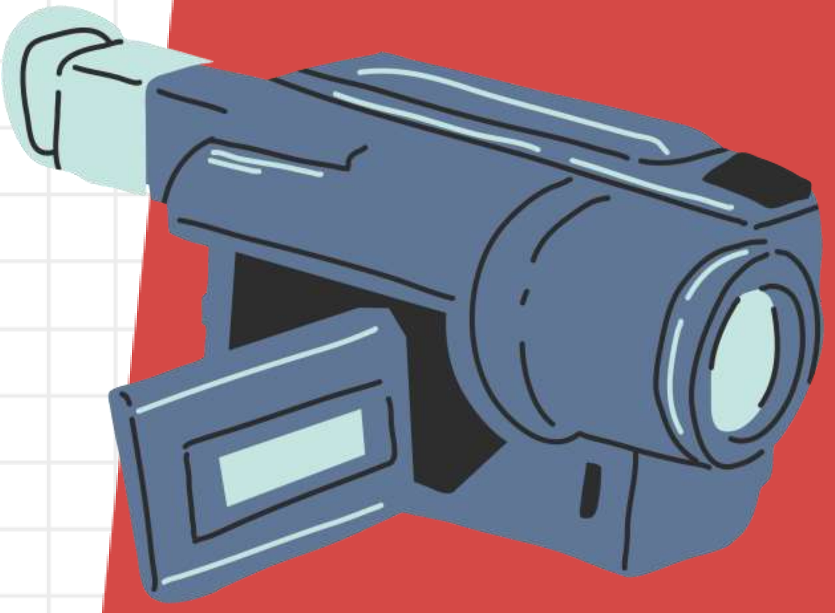
Observed Patterns and Anomalies:

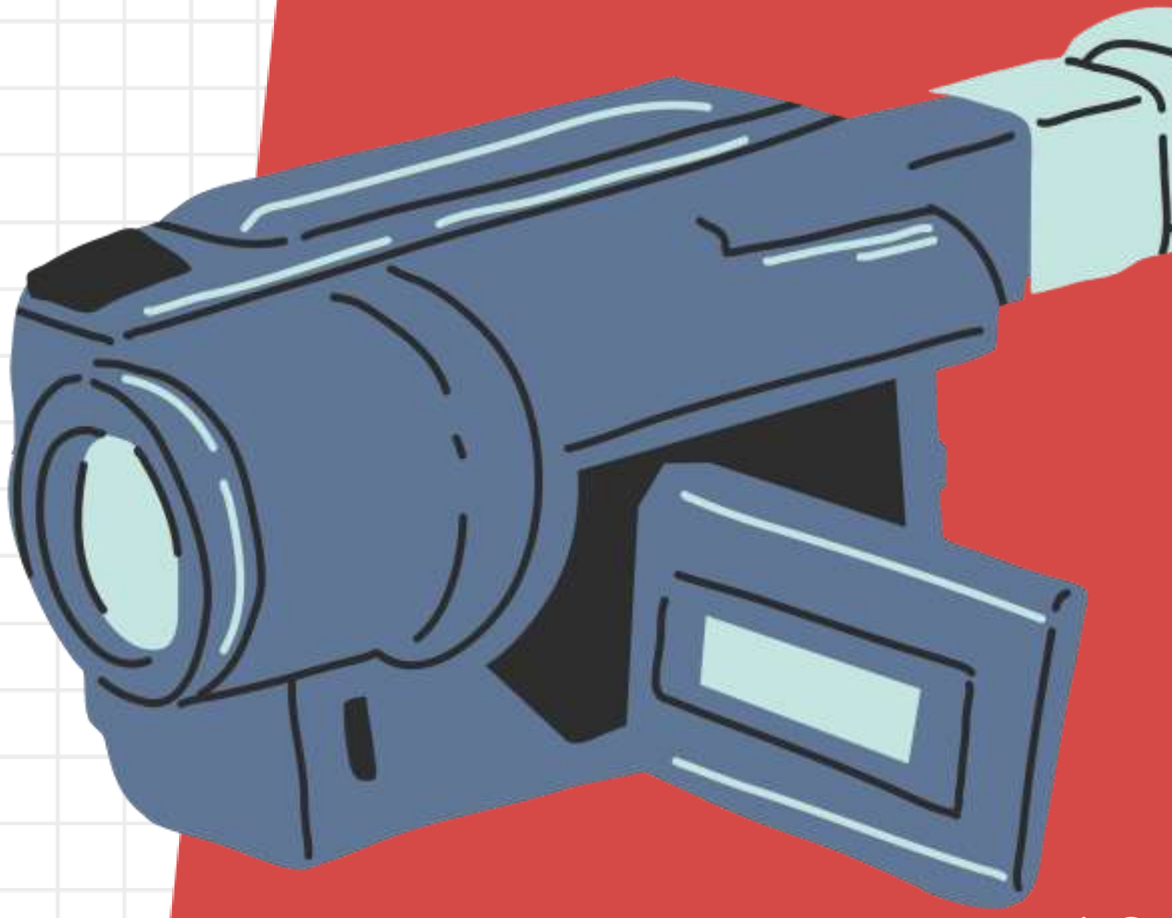
- **Recurrent Themes:** Certain themes or ad styles consistently yielded higher agreement percentages.
- **Anomalies:** Instances where ads had unexpected results, possibly due to abstract content or presentation styles.

Conclusion

Logistic Regression demonstrates viability for classifying video ads based on textual data, achieving reasonable agreement percentages and metrics. The model excels in objective classifications such as identifying online contact information and purchase incentives, highlighting its effectiveness in handling clear-cut, factual content.

It shows room for improvement in subjective assessments, where nuances like emotional impact and creative appeal are more challenging to interpret accurately. Addressing these subjective aspects could enhance the model's overall effectiveness in evaluating the holistic appeal and effectiveness of video advertisements.





Thank You