Multi-Modal Machine Learning for Political Video Advertisement Analysis: Integrating Audio, Textual, and Visual Features

Moulik Kumar
Indian Institute of Technology, Kharagpur

Satish Gopalani
PubMatic Inc

Pranav Gupta
PubMatic Inc

ABSTRACT

This paper presents a novel framework for the automated classification, tagging, and issue-level sentiment analysis of video advertisements using advanced machine-learning techniques. The proposed multi-pass approach leverages audio transcription, Optical Character Recognition (OCR), and video feature extraction to achieve high accuracy in distinguishing between political and non-political content. The research introduces robust meth- ods for candidate identification for political videos using phrase matching and fuzzy logic, as well as issue tagging and senti- ment analysis utilizing natural language processing algorithms. The system demonstrates significant improvements over existing methods, achieving 99.2% accuracy in political ad classification when combining audio and OCR data. Furthermore, the developed issue-level sentiment analysis provides granular insights into the emotional tone of political messaging. This research contributes to the growing field of content moderation in digital advertising, offering valuable insights for publishers, researchers, and policymakers in the realm of political communication.

General Terms

Political advertising, Video classification, Machine learning, Natural language processing, Content moderation

Keywords

Political advertising, Video classification, Machine learning, Natural language processing, Content moderation, Candidate verification, Sentiment analysis

1. INTRODUCTION

Video classification techniques have evolved rapidly in recent years, leveraging advances in machine learning and computer vision to automatically categorize visual content. These methods, which analyze features from video frames, audio, and metadata, find applications across various domains, including content moderation, recommendation systems, and advertising [1].

In the digital advertising landscape, video classification has become increasingly crucial. The Advertisement Technology (AdTech) industry relies on these techniques to ensure brand safety, improve ad targeting, and enhance user experience. As digital ad spending continues to grow, reaching $455.3 billion globally in 2021 and surpassing $1 trillion by 2025, the need for sophisticated content analysis tools has never been greater [2].

Political advertising, particularly in video format, represents a significant and growing segment of digital advertising. In the United States, political ad spending is projected to reach $12.3 billion in 2024, a 30% increase from $9.6 billion in 2020. Digital media is expected to account for approximately $3.5 billion of this spending, nearly tripling from the previous election cycle [3]. This surge in digital political advertising necessitates advanced classification, verification, and analysis systems.

This research focuses on developing a robust framework for classifying political video ads, verifying candidates, and analyzing issue-level sentiment using advanced text processing and machine learning techniques. The study addresses not only the binary classification of political versus non-political content but also the more nuanced tasks of candidate/party identification, political issue tagging, and sentiment detection at the issue level. This comprehensive approach aims to provide publishers with granular control over their ad inventory, enabling them to curate content that aligns with their brand values and audience expectations.

By implementing such a system, publishers can maintain platform integrity, ensure regulatory compliance, and create a more transparent and effective digital advertising ecosystem. This research contributes to the growing body of work on content moderation in the digital age, with a specific focus on the unique challenges posed by political video advertising.

2. CURRENT APPROACHES

The classification and tagging of political video advertisements have traditionally been performed manually, a process that is both time-consuming and prone to human error. The rapid growth in political advertising, particularly on digital platforms, has necessitated the development of automated methods to handle the sheer volume of content effectively.

Recent research has focused on using deep learning techniques for video classification. Dhakal (2019) [4] investigated various deep-learning methods for classifying political video advertisements by combining text features with video classification techniques. This study highlights the potential of deep learning to significantly improve the accuracy of political ad classification

Another study by Belhaouari et al. (2023) [5] provides a comprehensive review of deep learning models for video classification, emphasizing the importance of network architecture, evaluation criteria, and benchmark datasets. This review indicates that advanced deep learning methods, such as Convolutional Neural Networks (CNNs) and transformers, have made substantial progress in the automatic classification of videos.

Grigsby and Fowler (2020) explored the political ideology of ads by analyzing image and text content, focusing on the type of ad sponsor and the ideological leanings of the ad. Their work underscores the importance of accurately classifying political ads to understand digital campaigning's characteristics and its impact on political communication [6].

Despite these advancements, the task of political ad classification still faces challenges due to the nuanced and varied nature of political content. Current approaches often require large labelled datasets for training, which may not always be available. Moreover, the ability to handle multi-modal data—combining text, audio, and visual features—remains a critical area for improvement.

3. PROPOSED SOLUTION

3.1 Problem Overview

This research aims to develop a comprehensive system for analyzing political video advertisements. The main objectives of the system are to:

(1) Distinguish between political and non-political video ads
(2) Identify candidates mentioned in political ads
(3) Detect the political issues discussed in the ads
(4) Analyze the sentiment associated with each issue and the overall ad tone

These objectives present a complex challenge that requires analyzing multiple aspects of video content, including spoken words, on-screen text, and visual elements.

3.2 Framework Approach

To address the challenge of political ad classification, this paper proposes a multi-pass framework that leverages different types of data extracted from video advertisements. This approach is based on analysis of political ad characteristics and regulatory requirements. The dataset indicates that approximately 90% of political video ads in the U.S. contain spoken political commentary, making audio analysis an effective first pass. The second pass incorporates OCR to capture text-based disclaimers, which are mandated for political ads [7] and present in 95% of the samples. The approach consists of three main stages, each incorporating additional data to refine the classification:

(1) **First Pass: Audio-based Classification**
  —In this initial stage, the system transcribes the audio content of the video advertisement.
  —The analysis processes the transcribed text to make a preliminary classification.
  —This pass is computationally efficient and allows for quick filtering of clearly non-political content.

(2) **Second Pass: Audio and OCR-based Classification**
  —For ads that aren't accurately classified in the first pass, the system incorporates Optical Character Recognition (OCR) data.
  —The process extracts and analyzes text visible in the video frames, particularly focusing on the final frames where disclaimers often appear.
  —This additional textual information helps to catch political ads that might not have clear audio indicators.

(3) **Third Pass: Comprehensive Audio, OCR, and Video Feature Classification**
  —In the final pass, the system incorporates visual features extracted from the video frames.
  —This comprehensive analysis combines all available data: audio transcription, OCR text, and video features.
  —This pass is designed to make final decisions on ads that were ambiguous in the previous stages.

3.3 System Architecture

Figure 1 illustrates the architecture of the 3-pass classification system. The process flow can be summarized as follows:

(1) Input video advertisements are processed to extract audio transcriptions.
(2) Initial classification is performed using audio features.
(3) For ambiguous cases, OCR is applied to extract on-screen text, and classification is refined.
(4) In the final pass, video features are incorporated for comprehensive classification.
(5) The output is a binary classification of each ad as either political or non-political.

3.4 Subsequent Analysis

Following the classification of an advertisement as political, the system performs additional analyses, including:

—**Candidate Identification:** Detecting and identifying political candidates mentioned in the ad.
—**Issue Detection:** Identifying the political issues discussed in the ad.
—**Sentiment Analysis:** Analyzing the sentiment associated with each detected issue and the overall ad tone.

These subsequent steps, while crucial for comprehensive political ad analysis, are separate from the initial classification framework and will be discussed in detail in later sections of this paper.

4. DATA PREPARATION

This study utilizes a dataset of 2,071 video creatives obtained from a leading adtech company, comprising 1,082 non-political and 989 political advertisements. This dataset was carefully curated to represent the diverse range of content typically encountered by publishers in real-world scenarios.

The video creatives exhibit significant variety in terms of content, duration, production quality, target audience, and advertising objectives. This diversity ensures that the model is trained on a representative sample of advertisements, enhancing its ability to generalize to new, unseen content.

By maintaining a near-balanced distribution between political and non-political ads, the methodology minimizes potential biases in model training and evaluation. This approach enables the system to develop a nuanced understanding of the subtle distinctions between political and non-political content, particularly in cases where ads may share similar language or imagery.

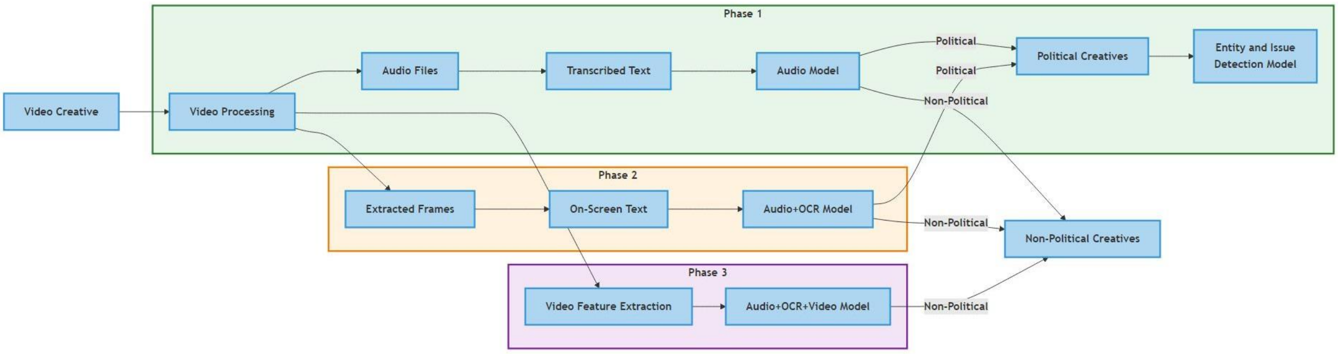

Fig. 1: 3-Pass Classification System Architecture

All creatives were anonymized and processed in compliance with relevant data protection regulations and ethical guidelines for research.

5. METHODOLOGY

5.1 Audio Transcription and Analysis

The audio files were transcribed using OpenAI's Whisper model [8], an open-source, state-of-the-art automatic speech recognition (ASR) system. Whisper is renowned for its robust performance across diverse accents, dialects, and noisy environments, making it ideal for the analysis of political advertisements where clarity and accuracy of transcription are paramount.
Key steps in the audio analysis include:

(1) Transcription: Converting spoken words to text using ASR.
(2) Text Preprocessing: Cleaning and normalizing the transcribed text, including removing punctuation, converting to lowercase, and handling common abbreviations [9].
(3) Feature Extraction: Applying TF-IDF (Term Frequency-Inverse Document Frequency) vectorization to convert the text into a numerical format suitable for machine learning algorithms [10].

$$\text{TF-IDF}(t, d, D) = \text{TF}(t, d) \times \text{IDF}(t, D)$$

$$\text{TF}(t, d) = \frac{\text{Number of times term } t \text{ appears in document } d}{\text{Total number of terms in document } d}$$

$$\text{IDF}(t, D) = \log\left(\frac{N}{n_t}\right)$$

N = Total number of documents in the corpus

n_t = Number of documents containing the term t

5.2 Optical Character Recognition (OCR)

For extracting text from video frames, the study employed an accessible OCR engine [11] known for its effectiveness in detecting and recognizing text in various orientations and styles. This tool is particularly useful for political ads, which often include essential on-screen text such as disclaimers and candidate names in the final frames [7].
The OCR process involves:

(1) Frame Extraction: Selecting key frames from the video, with a focus on the final frames where political disclaimers often appear [12].
(2) Text Detection: Identifying regions in the frames that contain text.
(3) Text Recognition: Converting the detected text regions into machine-readable text.
(4) Text Processing: Applying similar preprocessing steps as used for the audio transcription.

5.3 Video Feature Extraction

In the third pass, the system extracts visual features directly from the video content, focusing on the last 7.5 seconds of each video. This choice is motivated by the common practice in political advertisements to display crucial information towards the end of the ad [7]. The key features extracted include:

(1) Color Analysis:
—Percentage of red pixels
—Percentage of blue pixels
—Combined percentage of red and blue pixels
(2) Face Detection:
—Total number of faces detected
—Average number of faces per frame
—Average number of faces per second
—Maximum number of faces in any single frame
—Face density (faces per second)
(3) Video Duration:
—Total length of the video in seconds

Fig. 2: Example Political Advertisement [13]

Red-Blue %:	21.51%
Red %:	7.82%
Blue %:	13.69%

The study utilizes OpenCV [14] for video processing and the Haar Cascade classifier [15] for face detection, allowing for efficient processing of large numbers of video files while capturing key visual characteristics that may indicate political content.

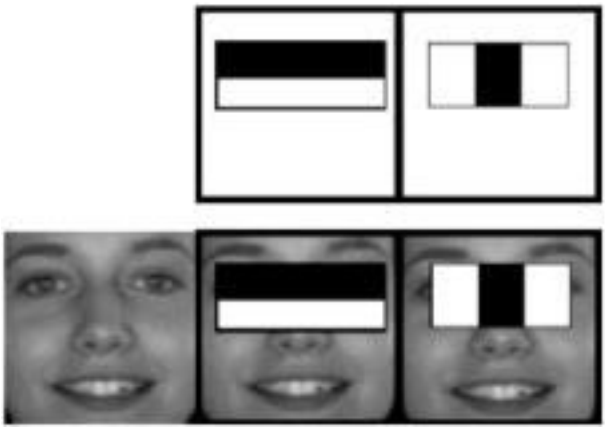

Fig. 3: Haar Cascade Algorithm [16]

6. MODEL TRAINING AND EVALUATION

6.1 Feature Engineering

After extracting features from audio transcriptions, OCR text, and video content, the study combines these multi-modal features into a unified representation for each advertisement. The feature set includes:

(1) TF-IDF vectors from audio transcriptions
(2) TF-IDF vectors from OCR-extracted text
(3) Video features (color analysis, face detection metrics, and duration)

6.2 Classification Models

The study experiments with several machine learning algorithms for the binary classification task of distinguishing between political and non-political ads. The models evaluated include:

—Support Vector Machines (SVM): Effective for high-dimensional spaces but may struggle with large datasets.
—Logistic Regression: Offers high explainability through accessible regression coefficients for TF-IDF features, and is efficient, but may underperform with complex non-linear relationships.
—Random Forest: Handles non-linear relationships well and is less prone to overfitting, but can be computationally intensive.
—Gradient Boosting Machines: Often achieves high accuracy but may be prone to overfitting without careful tuning.
—Neural Networks: Capable of capturing complex patterns but requires substantial data and computational resources.

Each model is trained and evaluated using stratified k-fold cross-validation [17] to ensure robust performance estimates. The results show that Logistic Regression offered the best balance between performance and explainability. The ability to interpret the regression coefficients for TF-IDF features in Logistic Regression provides valuable insights into the most influential terms for political ad classification.

6.3 Evaluation Metrics

To assess the performance of the study's models, the following evaluation metrics were used:

(1) Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

(2) Precision:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

(3) Recall (also known as Sensitivity or True Positive Rate):

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

(4) F1-Score:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

(5) Area Under the Receiver Operating Characteristic Curve (AUC-ROC):

$$\text{AUC-ROC} = \int_0^1 TPR(T) \times FPR'(T)\, dT \tag{5}$$

Where TPR is the True Positive Rate and FPR is the False Positive Rate at threshold T.

Where:

—TP = True Positives
—TN = True Negatives
—FP = False Positives
—FN = False Negatives

6.4 Classification Model Results

The multi-pass approach demonstrated significant improvements over single-modality classification methods. Table 1 presents the performance metrics for the best-performing model at each stage of the classification process.

The results show that incorporating OCR and video features progressively improves the classification performance, with the final model achieving 99.5% accuracy.

6.5 Feature Importance Analysis

To understand the contribution of different features to the classification decision, the study analyzed the feature importance scores. Tables 2 and 3 show the top 10 most significant TF-IDF features from audio transcriptions and OCR-extracted text, respectively. These tables highlight the most influential terms in determining whether an advertisement is political.

7. CANDIDATE IDENTIFICATION AND VERIFICATION

Following the classification of an ad as political, the proposed system performs candidate identification and verification. This process involves several steps:

7.1 Key Phrase Extraction

The study defines a set of key political phrases, each associated with a weight reflecting its importance in identifying candidates. These phrases are matched within the ad text using regular expressions [18], with specific instructions on where to look for candidate names relative to the phrase (before, after, or both directions) and within a specified character limit.

7.2 Named Entity Recognition

The study utilized Named Entity Recognition (NER) [19] to identify potential candidate names within the text. This step helps in capturing names that might not be directly associated with the predefined key phrases.

7.3 Fuzzy Matching

To account for variations in name spelling and common nicknames, the study employs fuzzy matching techniques [20]. This includes:

(1) Levenshtein Distance [21]:

$$lev_{a,b}(i,j) = \begin{cases} \max(i,j) & \text{if } \min(i,j) = 0, \\ \min \begin{cases} lev_{a,b}(i-1,j)+1 \\ lev_{a,b}(i,j-1)+1 \\ lev_{a,b}(i-1,j-1)+[a_i \neq b_j] \end{cases} & \text{otherwise.} \end{cases} \quad (6)$$

Where $[a_i \neq b_j]$ is 1 if $a_i \neq b_j$, and 0 otherwise.

(2) Jaro-Winkler Similarity [22]:

$$JW(s_1, s_2) = J(s_1, s_2) + lp(1 - J(s_1, s_2)) \quad (7)$$

Where J is Jaro Similarity, l is common prefix length (max 4), p is scaling factor (usually 0.1).

(3) Metaphone algorithm [23]: Applies phonetic rules to convert names to codes. Key rules:
—Drop duplicate adjacent letters, except C.
—Drop first letter if word begins with KN, GN, PN, AE, WR.
—Convert X to S at start, else to KS.
—Transform specific letter combinations.

A final ”match score” was utilized by proportionately weighing the matching scores of the first, last, and compound names:

$$\begin{aligned} \text{Match Score} = {} & 0.3 \cdot \text{first\_phonetic} + 0.3 \cdot \text{last\_phonetic} \\ & +0.1 \cdot \text{first\_fuzzy} + 0.1 \cdot \text{last\_fuzzy} \\ & +0.1 \cdot \text{full\_name} + 0.1 \cdot \text{jaro} \end{aligned} \quad (8)$$

7.4 Verification Against Official Database

Detected candidate names are verified against the Federal Election Commission (FEC) candidate database [24]. This step ensures that the identified names correspond to actual political candidates, reducing false positives.

8. ISSUE DETECTION AND SENTIMENT ANALYSIS

8.1 Issue Detection

The proposed system employs a multi-faceted approach to detect political issues discussed in the advertisements:

—**Predefined Issue Dictionary**
The study maintain a comprehensive dictionary of politi- cal issues (e.g., ”healthcare”, ”immigration”, ”environmen- tal politics”), each associated with relevant keywords and phrases.

—**Machine Learning Classification**
A multi-output classifier (MultiOutputClassifier with LinearSVC) [25] is trained on a labeled dataset to predict the presence of predefined issues in the ad text.

—**Topic Modeling**
The study uses Latent Dirichlet Allocation (LDA) [26] to detect potential new or emerging issues not covered by the predefined set. This allows the system to adapt to evolving political discourse.

8.2 Sentiment Analysis

For each detected issue, as well as for the overall ad, the study performs sentiment analysis using a multi-faceted approach:

—**Issue-Specific Sentiment**
Each predefined issue has associated lists of positive and negative phrases. The study counts occurrences of these phrases to determine the sentiment expressed towards each specific issue.

—**General Sentiment**
The study employs a pre-trained sentiment analysis model (DistilBERT fine-tuned on SST-2) [27] in combination with TextBlob [28] to assess the overall sentiment of the text.

—**Context-Aware Adjustment**
The sentiment scores are adjusted based on the presence of negation words (e.g., ”not”, ”never”) and other contextual cues.

9. CONCLUSION AND FUTURE WORK

This paper presents a comprehensive framework for the automated classification, candidate verification, and issue-level sentiment analysis of political video advertisements. The proposed multi-pass approach, combining audio transcription, OCR, and video feature extraction, achieves high accuracy (99.5%) in distinguishing between political and non-political content. The subsequent analyses of candidate identification and issue-level sentiment

Table 1. : Performance Metrics for Multi-Pass Classification

Pass	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Audio Only	0.980	0.975	0.985	0.980	0.995
Audio + OCR	0.992	0.990	0.994	0.992	0.998
Audio + OCR + Video	0.995	0.994	0.996	0.995	0.999

Table 2. : Top 10 Audio TF-IDF Features

Feature	IDF	Coefficient
vote	2.809	3.8929
run	3.016	2.7579
congress	3.115	2.5478
approve	2.943	2.4348
message	2.929	2.4309
state	3.308	2.2994
fight	3.107	2.2792
city	3.533	2.1901
conservative	3.654	2.0535
border	3.760	1.9333

Table 3. : Top 10 OCR TF-IDF Features

Feature	IDF	Coefficient
paid	2.211	5.4749
congress	2.744	5.3974
vote	3.328	2.9487
senate	4.142	2.2950
approved	3.323	2.1754
district	3.792	2.0947
elect	4.189	2.0017
state	3.397	1.9988
democrat	3.721	1.7732
vegas	3.825	1.7139

Table 4. : Top 8 Key Phrases for Candidate Identification

Key Phrase Pattern	Weight	Direction
I approve this message	1.0	Before
I'm [. . .] and I approve this message	1.0	Before
paid for by (. . .)	0.85	After
authorized by (. . .)	0.8	After
sponsored by (. . .)	0.8	After
vote for (. . .)	0.75	After
elect (. . .)	0.75	After
re-elect (. . .)	0.75	After

provide valuable insights into the content and tone of political messaging.
Key contributions of this work include:

(1) A novel multi-modal approach to political ad classification
(2) Robust methods for candidate verification using fuzzy matching and official databases
(3) Granular issue detection and sentiment analysis capabilities
(4) Insights into the most indicative features for political content identification

While the system demonstrates significant improvements over existing methods, there are several avenues for future work:

(1) Incorporating more advanced deep learning techniques for video analysis, such as 3D convolutional neural networks for spatiotemporal feature extraction [29].
(2) Expanding the system to handle multi-lingual content, enabling analysis of political ads in diverse linguistic contexts.
(3) Developing methods to track the evolution of political issues and sentiment over time, providing insights into changing campaign strategies and public opinion.
(4) Investigating the potential of transfer learning to adapt the model to different political systems and cultural contexts

As the landscape of political advertising continues to evolve, particularly in digital spaces, the need for sophisticated, automated analysis tools grows increasingly crucial. This work contributes to this field by providing a comprehensive, accurate, and interpretable system for political ad analysis, with potential applications in content moderation, campaign strategy, and political communication research.

10. REFERENCES

[1] W. Wu, H. Wang, Y. Ye, and Z. Zhang. A comprehensive survey of video-based action recognition using deep learning: Dataset, method, and challenge. *IEEE Transactions on Neural Networks and Learning Systems*, 33(9):4104–4124, 2022.

[2] eMarketer. Worldwide digital ad spending 2021. *eMarketer Insider Intelligence*, 2021.

[3] Advertising Analytics and Cross Screen Media. 2024 political advertising outlook, 2023.

[4] S. Dhakal. Deep learning approach for political video advertisement classification. In *Proceedings of the 2019 3rd International Conference on Deep Learning Technologies (ICDLT 2019)*, pages 31–35. Association for Computing Machinery, 2019.

[5] S. B. Belhaouari, A. Alshabani, A. T. Azar, and E. Almazrouei. A comprehensive review of deep learning models for video classification. *Applied Sciences*, 13(2):890, 2023.

[6] J. W. Grigsby and E. F. Fowler. Political advertising in the digital age: The political ideology of ads on facebook. *Political Communication*, 37(6):785–809, 2020.

[7] Stand by your ad act of 2002, 47 u.s.c. § 315 note (2002), 2002.

[8] OpenAI. Whisper: Openai's automatic speech recognition model [computer software], 2022. https://github.com/openai/whisper.

[9] H. Schu¨tze, C. D. Manning, and P. Raghavan. *Introduction to information retrieval*, volume 39. Cambridge University Press, 2008.

[10] A. Wendland, M. Zenere, and J. Niemann. Introduction to text classification: impact of stemming and comparing tf-idf and count vectorization as feature extraction technique. In *Systems, Software and Services Process Improvement: 28th European Conference, EuroSPI 2021, Krems, Austria, September 1–3, 2021, Proceedings 28*, pages 289–300. Springer International Publishing, 2021.

[11] JaidedAI. Easyocr: Ready-to-use optical character recognition with 80+ supported languages [computer software], 2020. https://github.com/JaidedAI/EasyOCR.

[12] M. K. Asha Paul, J. Kavitha, and P. A. Jansi Rani. Key-frame extraction techniques: A review. *Recent Patents on Computer Science*, 11(1):3–16, 2018.

[13] Joe Biden. Joe biden for president - joe biden: Keep up the fight. YouTube video, 2020. Accessed: 2024-09-10.

[14] G. Bradski. The opencv library [computer software], 2000. https://opencv.org.

[15] L. Cuimei, Q. Zhiliang, J. Nan, and W. Jianhua. Human face detection algorithm via haar cascade classifier combined with three additional classifiers. In *2017 13th IEEE interna- tional conference on electronic measurement & instruments (ICEMI)*, pages 483–487. IEEE, 2017.

[16] OpenCV Team. Face detection using haar cascades. https://docs.opencv.org/4.x/d2/d99/tutorial_js_face_detection.html, 2023. Accessed: [Insert access date here].

[17] T. Fushiki. Estimation of prediction error by using k-fold cross-validation. *Statistics and Computing*, 21:137–146, 2011.

[18] Y. Li, R. Krishnamurthy, S. Raghavan, S. Vaithyanathan, and H. V. Jagadish. Regular expression learning for information extraction. In *Proceedings of the 2008 conference on empirical methods in natural language processing*, pages 21–30, 2008.

[19] B. Mohit. Named entity recognition. In *Natural language processing of semitic languages*, pages 221–245. Springer Berlin Heidelberg, 2014.

[20] M. Cayrol, H. Farreny, and H. Prade. Fuzzy pattern matching. *Kybernetes*, 11(2):103–116, 1982.

[21] L. Yujian and L. Bo. A normalized levenshtein distance metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):1091–1095, 2007.

[22] Y. Wang, J. Qin, and W. Wang. Efficient approximate entity matching using jaro-winkler distance. In *International conference on web information systems engineering*, pages 231–239. Springer International Publishing, 2017.

[23] C. Snae. A comparison and analysis of name matching algorithms. *International Journal of Computer and Information Engineering*, 1(1):107–112, 2007.

[24] Federal Election Commission. Candidate and commit- tee viewer [data set]. https://www.fec.gov/data/candidates/.

[25] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, et al. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

H. Jelodar, Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, and L. Zhao. Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78:15169–15211, 2019.

[26] V. Sanh, L. Debut, J. Chaumond, and T. Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter, 2019. https://arxiv.org/abs/1910.01108.

[27] S. Loria. textblob documentation, 2018. https://textblob.readthedocs.io/en/dev/.

[28] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri. Learning spatiotemporal features with 3d convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, pages 4489–4497, 2015.