Running 1,947,792 Regressions using Transition Matrix and Audio Classification for Presidential Speech Analysis

By Magilan Vasanthakumar

1 Introduction

Speech can deeply resonate with people, particularly in the context of presidential speeches, as it can significantly influence their public image. This influence lies within the emotional spectrum, making it a relevant topic for a behavioral study. By examining the impact of speech on people's thoughts, we can gain insights into how emotions conveyed through presidential speeches affect public perception. In this research, we expand upon the traditional survey methods by integrating machine learning. We extract emotions from audio files and fit them in a large-scale analysis involving 1,947,792 regressions using transition matrices.

2 Outcome Data

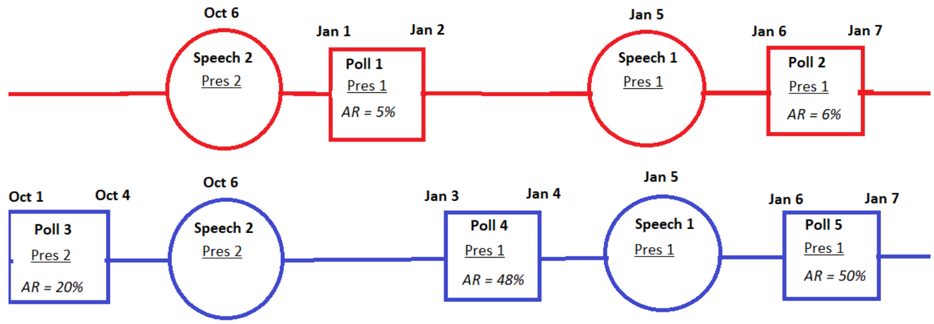
The objective of this research is to examine the subconscious impact of emotions conveyed through speech on listeners' opinion. For instance, individuals may not favor a president who optimistically presents bad news as it shows a lack of awareness. The analysis will involve conducting a regression analysis, where the dependent variable is the average change in approval rates, and the independent variable is the transitional elements of emotions. Approval rates are collected from polls that rate the president's likability at that time. Roper Center is an organization that compiled approval rate polls collected from hundreds of organizations. Attached below is the clean version of their data.

	Organization	Polling Start	Polling End	Approve	Disapprove	No Opinion	Sample Size	Sample Type	President
0	American Research Group	2023-05-17	2023-05-20	41	52	7	1100	NATIONALADULT	Joe Biden
1	Associated Press/NORC Center	2023-05-11	2023-05-15	40	59	1	1680	NATIONALADULT	Joe Biden
2	Reuters/Ipsos	2023-05-05	2023-05-07	40	54	6	1022	NATIONALADULT	Joe Biden
3	ABC News / Washington Post	2023-04-28	2023-05-03	36	56	8	1006	NATIONALADULT	Joe Biden
4	Fox	2023-04-21	2023-04-24	44	55	1	1004	REG_VOTERS	Joe Biden
...	...	...	...	...	...	...	...	...	...
6524	Gallup Organization	1938-07-04	1938-07-11	52	40	7	Not reported	Not reported	Franklin D. Roosevelt
6525	Gallup Organization	1938-05-29	1938-06-04	54	46	Not reported	Not reported	Not reported	Franklin D. Roosevelt
6526	Gallup Organization	1938-05-22	1938-05-27	54	46	Not reported	Not reported	Not reported	Franklin D. Roosevelt
6527	Gallup Organization	1937-10-30	1937-11-04	63	37	Not reported	Not reported	Not reported	Franklin D. Roosevelt
6528	Gallup Organization	1937-08-04	1937-08-09	60	40	Not reported	Not reported	Not reported	Franklin D. Roosevelt

The data spans from 1937 to 2023 and contains relevant information on each poll, including details about its duration, source, and sample size. The approval rates in this dataset are represented in percentage form. To preprocess approval rates for regression analysis, presidential speech videos are required. The attached data below was obtained through web scraping from the Miller Center, an organization that archives presidential speech videos.

	date	speaker	header	source	summary	transcript	link	path
0	February 21, 2023	Joe Biden	February 21, 2023: Remarks on the One-Year Ann...	The White House	Speaking at the Royal Castle in Warsaw, Poland...	THE PRESIDENT: Hello, Poland! One of our great...	https://dvydlafohk4qa.cloudfront.net/audio/bid...	[redacted]
1	February 07, 2023	Joe Biden	February 7, 2023: State of the Union Address	The White House	In his State of the Union Address, President J...	Mr. Speaker. Madam Vice President. Our First L...	https://dvydlafohk4qa.cloudfront.net/audio/bid...	[redacted]
2	September 21, 2022	Joe Biden	September 21, 2022: Speech before the 77th Ses...	The White House	President Joe Biden addresses the 77th session...	Thank you.Mr. President, Mr. Secretary-General...	https://dvydlafohk4qa.cloudfront.net/audio/bid...	[redacted]
3	September 01, 2022	Joe Biden	September 1, 2022: Remarks on the Continued Ba...	The White House	President Joe Biden speaks in Philadelphia, Pe...	THE PRESIDENT: My fellow Americans, please, if...	https://dvydlafohk4qa.cloudfront.net/audio/bid...	[redacted]
4	May 24, 2022	Joe Biden	May 24, 2022: Remarks on School Shooting in Uv...	The White House	President Biden makes an impassioned plea to s...	Good evening, fellow Americans.I had hoped, wh...	https://dvydlafohk4qa.cloudfront.net/audio/bid...	[redacted]
...	...	...	...	...	...	...	...	...
387	July 19, 1940	Franklin D. Roosevelt	July 19, 1940: Democratic National Convention	National Archives	President Roosevelt accepts the Democratic Par...	Members of the Convention-my friends:\nIt is v...	https://d4q9blt8qjhv3.cloudfront.net/audio/spe...	[redacted]
388	June 10, 1940	Franklin D. Roosevelt	June 10, 1940: Stab in the Back Speech	National Archives	The President condemns the actions of Mussolin...	President Newcomb, my friends of the Universit...	https://d4q9blt8qjhv3.cloudfront.net/audio/spe...	[redacted]
389	May 26, 1940	Franklin D. Roosevelt	May 26, 1940: Fireside Chat 15: On National De...	National Archives	Responding to Hitler's invasion of Denmark, No...	My friends:\nAt this moment of sadness through...	https://d4q9blt8qjhv3.cloudfront.net/audio/spe...	[redacted]
390	September 03, 1939	Franklin D. Roosevelt	September 3, 1939: Fireside Chat 14: On the Eu...	National Archives	Hours after Great Britain and France declare w...	My fellow Americans and my friends:\nTonight m...	https://d4q9blt8qjhv3.cloudfront.net/audio/spe...	[redacted]
391	July 08, 1938	Franklin D. Roosevelt	July 8, 1938: Dedication of a Memorial to the ...	National Archives	President Franklin Roosevelt celebrates the 15...	Governor Davey, Senator Bulkley, Chairman Whit...	https://d4q9blt8qjhv3.cloudfront.net/audio/spe...	[redacted]

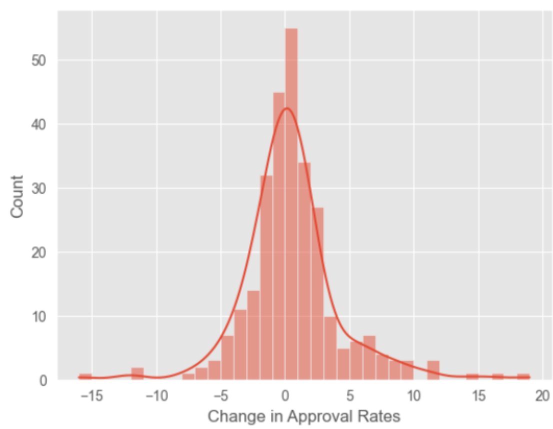
The provided data includes information such as the date of the speech, the transcript of the video, and web links to the audio. Additionally, there is a column called ”Path,” containing the path names to the audio files in the local directory. To combine this data with the first dataset, a differencing operation is performed on the approval rates. Specifically, the difference between the current approval rate and the previous approval rate is calculated, but only under specific conditions. These conditions involve filtering the first dataset to ensure that two polls are a month apart from each other, collected from the same organizations and featuring the same president. Furthermore, two polls must be between the date of a speech. Since the polls are collected from multiple organizations, it is possible that a poll collected from one organization may overlap with another poll collected from another organization by the duration of in-between dates. To address this, the differences are averaged by presidential speech. An example of this process is provided below.



Red represents organization 1, and blue represents organization 2. Poll 1 and Poll 2 fulfill the specific criteria: they are collected from the same organization, pertain to the same president, occur between Speech 1, and are a month apart. As a result, the difference in approval rates is calculated for these polls. Similarly, Poll 4 and Poll 5 meet the same requirements, thus the difference in approval rates is computed as well. However, Poll 3 and Poll 4 do not satisfy any of these conditions, so Speech 2 is removed from the data. Since the change of approval rates for Polls 2-1 and Polls 5-4 are mapped to the same speech, the averages of these differences are taken for better accuracy and to avoid potential biases. Attached below is the first difference data, which captures the changes in approval rates based on the specified criteria for organization, president, and speech timing.

	speaker	path	avg_approval_rate_change
0	Joe Biden	[redacted]	-2.666667
1	Joe Biden	[redacted]	1.200000
2	Joe Biden	[redacted]	-0.400000
3	Joe Biden	[redacted]	0.500000
4	Joe Biden	[redacted]	-1.000000
...	...	...	...
273	Franklin D. Roosevelt	[redacted]	0.000000
274	Franklin D. Roosevelt	[redacted]	0.000000
275	Franklin D. Roosevelt	[redacted]	-3.000000
276	Franklin D. Roosevelt	[redacted]	-4.000000
277	Franklin D. Roosevelt	[redacted]	6.500000

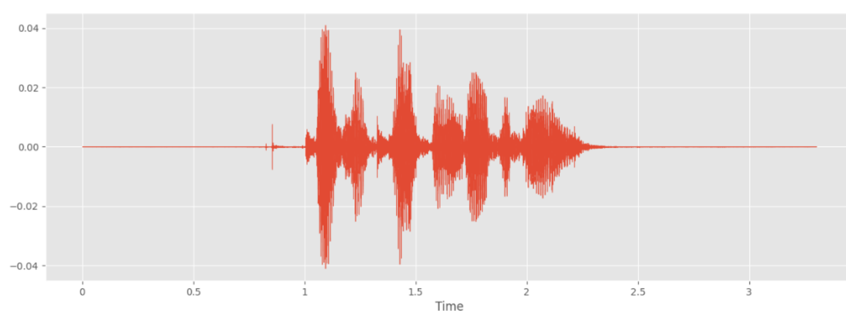
The polls were filtered to maintain a one-month time difference, precisely 30 days, to ensure an adequate number of samples while minimizing bias. The determination of the appropriate number of samples was made by examining the distribution. The filtering process involved removing samples that had the potential to disrupt the distribution. The final choice of a one-month interval resulted in a distribution that maintained sufficient smoothness and resembled a bell-curve shape. Attached below is the distribution.



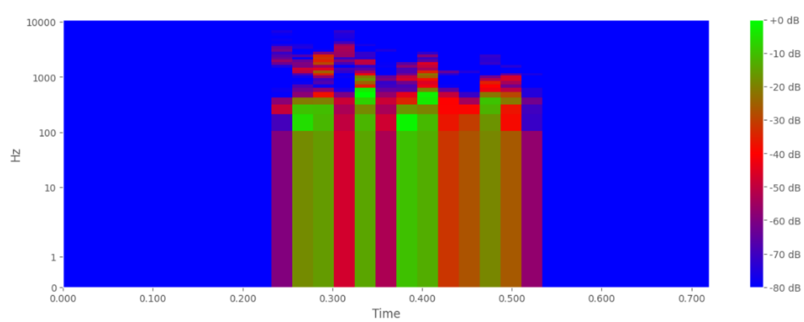
The distribution of the average change in approval rate is normal, with a mean of 0.45 and a variance of 14. This indicates that while there are small overall changes in the approval rate, the variations in these changes are substantial percentage-wise. Conducting a regression analysis enables us to explore and understand this high magnitude of variance.

3 Labeled Data

The independent variable in this research is the transitional elements of emotions, which are extracted from the audio files using a Neural Network. Neural Network is a subset of machine learning, where the algorithm is trained to learn patterns in the data. For supervised learning, labeled data is required, where each sample is associated with a specific class. The labeled data used for training the Neural Network is collected from three sources: CREMA-D, Savee, and Ravee. CREMA-D is a dataset of 7,442 original clips from 91 actors. The SAVEE database was recorded from four native English male speakers, who are also postgraduate students and researchers at the University of Surrey aged from 27 to 31 years.The RAVDESS is a dataset collected by researchers at the University of Ryerson which was recorded from 24 professional. The compiled data is filtered to meet certain criteria: all speakers are male (considering all the presidents are male), there is one channel of sound (normal intensity), and there are six labeled emotions (happy, sad, angry, fearful, disgust, and neutral). Refer to the first audio file from the labeled data as 'example', which is mapped to the emotion "neutral." When played, this example will feature a man saying "kids are talking by the door" in a neutral tone. To load this file and process it for analysis, it is essential to understand that all forms of digital information, including videos, images, and text, can be represented and translated into numerical arrays. At a fundamental level, numbers are constructed using binary sequences consisting of zeroes and ones. The digitized representation of the 'example' is as follows:



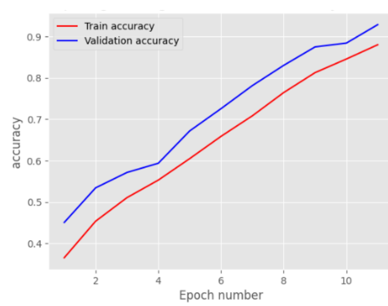
The waveform of audio is constructed by an array of numbers. Now consider this: if audio can be digitized to an array of numbers and an image can be digitized to an array of numbers, then it must be possible to transform an array of numbers that can form audio into an array of numbers that can form an image. Simply put, this process maps sound to colors, which is the core concept of spectrograms. The spectrogram is not necessarily needed to preprocess audio, but it tends to provide strong results for machine learning. To start, audio files are loaded, and their samples and sample rates are obtained. Samples represent the values that make up the waveform, while sample rates indicate the frequencies at which the samples are taken. By resampling the audio to a specific sample rate, the audio is effectively 'filtered' to ensure consistency in the audio channel. An optimal sample rate is 44100 Hertz, widely considered to be optimally supported and compatible with most audio playbacks. After digitizing the audio file, a Short Time Fourier Transformation (STFT) is applied to break down the audio into its fundamental frequency components. Frequencies are represented as unique sine and cosine functions. The audio is divided into smaller segments composed of frequencies, and these segments are then overlapped within a smaller domain, resulting in a 2D matrix known as the spectrogram. In this case, a MEL Spectrogram is used, which is similar to a traditional spectrogram but employs MEL filterbanks. These filters capture the non-linear relationship between frequency and human perception of pitch. As a final step, a layer of color is added to the spectrogram, with colors representing the intensity of decibels. The attached image shows the MEL Spectrogram of the example audio file.



It should be noted that this is a 32x32x3 image. This implies that there are 1024 pixels with 3 combinations of colors (red, blue, green). The size of the image is set by selecting the length of the Fourier transformation and the number of Mel spectrogram bins. For a 32-pixel width, the audio is divided into 32 segments and overlapped. The x-axis represents time within 44100/32 samples of the resampled audio. For a 32-pixel height, 32 unique filtered frequency ranges are applied over the overlapping waves. The y-axis represents the logarithmic scale of these frequency levels, aiding visualization. Hertz is the number of cycles per second. From the image, we can observe that green represents the loudest and blue the quietest. The decibels are in log scaling, where $\log(1) = 0$ dB, representing 100% or maximum loudness. Green appears brighter at higher hertz, indicating higher frequency levels and higher pitches of sounds. However, lower levels of frequency do not directly translate to quieter noise; they imply lower pitch but higher bass. The same preprocessing method is used for all labeled files and speech audio files. The pixels are normalized to prevent any value from overwhelming the others, and these preprocessed images are then used as inputs for our machine-learning algorithm.

4 Neural Network

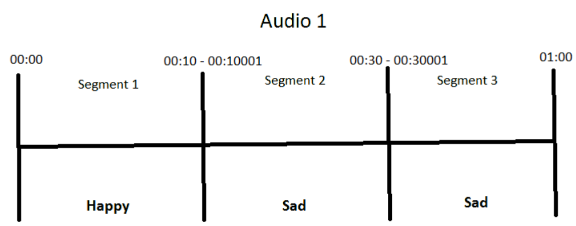
I constructed a Convolution Neural Network(CNN) algorithm, which is similar to a normal Neural Network, but creates filters to identify patterns in a frame of pixels. For instance, suppose there is a sequence in a frame: ”red pixel, red pixel, blue pixel.” If the sliding window is set to two pixels, the CNN applies weighted filters to capture the patterns of ”red pixel, red pixel” and then ”red pixel, blue pixel.” By sliding the filter across each set of pixels, CNN can extract features and adjust the weights accordingly to capture sequential patterns. Then, a normalization layer is applied to ensure that no values overwhelm others. Next, pooling is performed, creating partitions and averaging the values within each partition. This process helps lower the dimensions while retaining essential information. Another convolution layer is then applied before flattening the matrix into a one-dimensional array for the Dense Layers. These Dense layers are comprised of neurons. Neurons are defined as functions multiplied by weights. The function transforms a linear relationship into a non-linear relationship. Consider the function as a constraint to prevent overfitting. The weights are used to adjust for each iteration. After going through the layers, the model tries to predict the classes for each sample. By comparing the predicted labels with the actual labels, the algorithm can identify the errors resulting from misclassifications. Adam is a gradient-based optimizer that takes advantage of these errors by guiding the model toward minimizing them in subsequent iterations. It does so by adjusting the weights toward the opposite direction of the gradient vector. Attached below is the result of the training.



Epoch refers to one complete cycle of training data during the training process. It is important to note that the labeled data was divided into three sets: training data, test data, and validation data. The model is trained using the training data, and during each epoch, it assesses consistency by predicting both the training and validation data. The goal is not to fit the data but rather generalize it. In the provided plot, the validation accuracy runs parallel to the training accuracy. This parallel behavior indicates that the model is learning effectively and not overfitting. If the validation accuracy were to spike up at some epoch, it might indicate that the model is learning too quickly, which can be a sign of overfitting. The test accuracy achieved is 93%, which is considered good, given that the training and validation accuracy from the final iteration were also close to 93%. If the test accuracy were close to 100%, it could raise concerns about overfitting. Based on these results, it is now possible to confidently use this trained model to classify presidential speech files effectively.

5 Transition Matrix

For each speech, the signal is split into segments based on a threshold. Split at any point of time where loudness is 30db or lower for 0.001 seconds. A human can hear sounds from 0db to 130db. A whisper is 30db and a normal voice is 60db. The lower the duration, the more segments obtained from an audio file. An example is shown below.



In the provided image, it is observed that between 10 seconds and 10.001 seconds, the decibel levels are lower than 30 dB, resulting in splitting segments. Each segment is then classified using the neural network model. Using the sequence of emotions obtained from the classification, a transition matrix is constructed. The transition matrix is represented as a 2D matrix, where each element represents the probability of transitioning from one state to another state.Below is an example of a transition matrix for one of the audio files:

	disgust	happy	neutral	sad	angry	fearful
disgust	0.236304	0.231023	0.147855	0.234323	0.108251	0.042244
happy	0.166023	0.241313	0.187741	0.239382	0.124517	0.041023
neutral	0.046051	0.084274	0.674649	0.133318	0.051117	0.010592
sad	0.164476	0.220432	0.212802	0.262823	0.116575	0.022891
angry	0.151543	0.228675	0.219601	0.227768	0.136116	0.036298
fearful	0.185304	0.268371	0.175719	0.185304	0.108626	0.076677

Consider the element in the first row and first column of the transition matrix. The probability of the speech conveying ”disgust” to ”disgust” in this audio file is approximately 24%. After flattening the transition matrix into a one-dimensional array for each audio, a new data is constructed consisting of transition elements. It is then merged with the outcome data. The final data is provided below:

	avg_approval_rate_change	disgust_to_disgust	disgust_to_happy	disgust_to_neutral	disgust_to_sad	disgust_to_angry	disgust_to_fearful	happy_to_disgust	happy_to_happy	happy_to_neutral	...
0	-2.666667	0.580607	0.179907	0.066589	0.110981	0.049065	0.012850	0.265378	0.309315	0.163445	...
1	1.200000	0.236304	0.231023	0.147855	0.234323	0.108251	0.042244	0.166023	0.241313	0.187741	...
2	-0.400000	0.290224	0.240326	0.084521	0.183299	0.158859	0.042770	0.259887	0.227119	0.073446	...
3	0.500000	0.116822	0.182243	0.364486	0.126168	0.130841	0.079439	0.064234	0.255474	0.357664	...
4	-1.000000	0.293651	0.218254	0.103175	0.214286	0.091270	0.079365	0.324176	0.181319	0.076923	...
...	...	...	...	...	...	...	...	...	...	...	...
273	0.000000	0.190229	0.168399	0.166320	0.073805	0.270270	0.130977	0.230855	0.174251	0.185350	...
274	0.000000	0.119850	0.183521	0.273408	0.168539	0.209738	0.044944	0.081690	0.147887	0.408451	...
275	-3.000000	0.167453	0.523585	0.122642	0.141509	0.023585	0.021226	0.060646	0.689770	0.110297	...
276	-4.000000	0.122056	0.218415	0.319058	0.169165	0.128480	0.042827	0.098018	0.190529	0.352423	...
277	6.500000	0.166667	0.111111	0.322222	0.066667	0.200000	0.133333	0.081197	0.192308	0.354701	...

’...’ implies that there are more variables in this data. The final data includes the outcome variable and 36 transition elements. Listed under the ’Definition’ section are the interpretations of each transition element. Do not fixate on these interpretations. There is always room for more interpretations.

6 Channel Regression

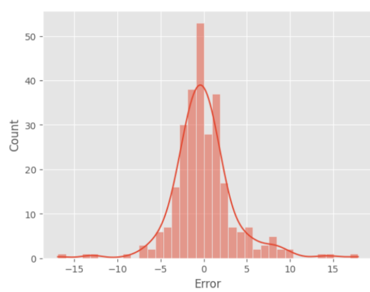
The following analysis involves a regression referred to as the ’happy channel regression’. This regression only includes transitional elements that channel happiness as the independent variables. Attached below are the Ordinary Least Square(OLS) regression, the distribution of the error, and the distribution of the variance.

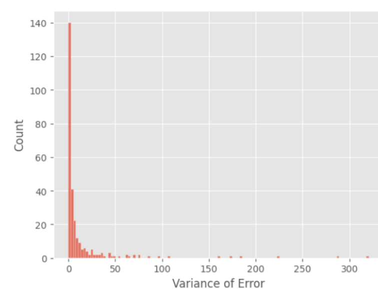
OLS Regression Results

Dep. Variable:	avg_approval_rate_change	R-squared:	0.035
Model:	OLS	Adj. R-squared:	0.014
Method:	Least Squares	F-statistic:	1.644
Date:	Tue, 18 Jul 2023	Prob (F-statistic):	0.135
Time:	16:20:42	Log-Likelihood:	-757.61
No. Observations:	278	AIC:	1529.
Df Residuals:	271	BIC:	1555.
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.3878	0.595	2.334	0.020	0.217	2.558
neutral_to_happy	4.7224	4.198	1.125	0.262	-3.542	12.987
happy_to_happy	-5.4536	3.991	-1.367	0.173	-13.310	2.403
sad_to_happy	4.7280	5.708	0.828	0.408	-6.510	15.966
angry_to_happy	-11.2176	5.346	-2.098	0.037	-21.743	-0.692
fearful_to_happy	3.5533	3.313	1.073	0.284	-2.969	10.075
disgust_to_happy	0.7041	4.842	0.145	0.884	-8.828	10.236

Omnibus:	51.815	Durbin-Watson:	1.991
Prob(Omnibus):	0.000	Jarque-Bera (JB):	326.834
Skew:	0.531	Prob(JB):	1.07e-71
Kurtosis:	8.205	Cond. No.	33.9





The OLS regression reveals a low adjusted R-squared value. The investigation involves plotting the error, which follows a normal distribution with a mean approximately close to 0 and a variance of around 13. While the mean is desirable, we further examine the variation by plotting the variance. Thankfully, since the error term is normally distributed with a mean approximately close to zero, we can determine that the variance is equal to the square of the error. However, we observe that the variance of the error appears to follow a Pareto distribution, indicating the presence of heteroskedasticity. To address this issue, we will implement a Weighted Least Square (WLS) regression with weights of $1/\sigma^2$. Attached below is the WLS regression:

WLS Regression Results

Dep. Variable:	avg_approval_rate_change	R-squared:	0.846
Model:	WLS	Adj. R-squared:	0.843
Method:	Least Squares	F-statistic:	248.3
Date:	Thu, 20 Jul 2023	Prob (F-statistic):	4.99e-107
Time:	12:13:20	Log-Likelihood:	-473.71
No. Observations:	278	AIC:	961.4
Df Residuals:	271	BIC:	986.8
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.3277	0.044	30.123	0.000	1.241	1.414
neutral_to_happy	4.7599	0.394	12.088	0.000	3.985	5.535
happy_to_happy	-5.9521	0.326	-18.232	0.000	-6.595	-5.309
sad_to_happy	4.6389	0.457	10.159	0.000	3.740	5.538
angry_to_happy	-10.3271	0.593	-17.419	0.000	-11.494	-9.160
fearful_to_happy	3.3284	0.222	15.022	0.000	2.892	3.765
disgust_to_happy	0.7647	0.531	1.441	0.151	-0.280	1.809

Omnibus:	1484.504	Durbin-Watson:	1.909
Prob(Omnibus):	0.000	Jarque-Bera (JB):	43.256
Skew:	0.252	Prob(JB):	4.05e-10
Kurtosis:	1.134	Cond. No.	104.

Multiple channels of Weighted Least Squares (WLS) regression were also conducted, resulting in consistent findings concerning the adjusted R-squared and other descriptive statistics. This consistency supports the notion that WLS is an appropriate method for the analysis, allowing for the execution of 1,947,792 regressions.

7 Running 1,947,792 Regressions

In happy channel regression, the coefficient for 'happy_to_happy' is negative, significant, and with high magnitude. This intuitively makes sense since people do not prefer a president who acts naive and optimistic as that shows a lack of awareness. As a Canadian, I direct your attention to Prime Minister Trudeau, who has been involved in many scandals yet acts as if Canada is a utopia. Now consider the possibility that the coefficient is positive for a different combination of independent variables. To validate the coefficient, it is necessary to control all variations in emotions. Unfortunately, all 36 transition elements cannot be added into a regression due to overfitting. Therefore, every combination of regression will be run. With 36 independent variables and the selection of only 6 of them, there is a total of $\frac{36!}{6!(36-6)!} = 1947792$ regressions. Attached below is data that contains coefficient values for each WLS regression.

	disgust_to_disgust	disgust_to_happy	disgust_to_neutral	disgust_to_sad	disgust_to_angry	disgust_to_fearful	happy_to_disgust	happy_to_happy	happy_to_neutral	happy_to_sad	...
0	2.683400	-2.036382	1.304571	4.093549	1.946790	-3.158929	NaN	NaN	NaN	NaN	...
1	0.753274	-2.200337	1.731012	3.951207	1.384749	NaN	3.484827	NaN	NaN	NaN	...
2	3.375693	3.030881	1.429987	3.879779	1.498929	NaN	NaN	-4.330282	NaN	NaN	...
3	3.100782	-0.795904	-2.201698	4.568588	1.587253	NaN	NaN	NaN	4.050934	NaN	...
4	3.274422	-2.216850	1.180035	10.689298	1.391485	NaN	NaN	NaN	NaN	-6.933394	...
...	...	...	...	...	...	...	...	...	...	...	...
1947787	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
1947788	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
1947789	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
1947790	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
1947791	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...

Attached below is a data that contains t-values for each regression.

	disgust_to_disgust	disgust_to_happy	disgust_to_neutral	disgust_to_sad	disgust_to_angry	disgust_to_fearful	happy_to_disgust	happy_to_happy	happy_to_neutral	happy_to_sad	...
0	20.836077	-9.744369	6.965881	20.891832	9.285924	-4.588683	NaN	NaN	NaN	NaN	...
1	1.435845	-3.968597	3.297576	7.853056	2.249645	NaN	8.274001	NaN	NaN	NaN	...
2	7.478317	4.892129	3.422661	8.397173	2.366246	NaN	NaN	-13.193949	NaN	NaN	...
3	15.506422	-3.003802	-4.261256	25.902961	6.226874	NaN	NaN	NaN	9.388931	NaN	...
4	29.041583	-15.477857	8.911672	60.440097	8.668038	NaN	NaN	NaN	NaN	-46.091086	...
...	...	...	...	...	...	...	...	...	...	...	...
1947787	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
1947788	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
1947789	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
1947790	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...
1947791	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...

The index in the data denotes a regression (1 indicates regression 1). When a coefficient is marked as NaN, it implies that the corresponding variable was not included in that specific regression. To conduct inference, the average statistics are calculated, skipping the NaN values. Below are the average statistics:

	disgust_to_disgust	disgust_to_happy	disgust_to_neutral	disgust_to_sad	disgust_to_angry	disgust_to_fearful	happy_to_disgust	happy_to_happy	happy_to_neutral	happy_to_sad	...
coeff	0.044659	-0.398350	-0.172118	0.774230	-0.049292	-1.606438	-0.020197	-0.635862	0.606988	-0.087721	...
t	0.602987	-1.881921	-0.767213	3.005307	0.087916	-1.755848	0.485901	-3.035089	1.777324	0.146443	...

To interpret a coefficient, it is stated that on average, controlling for all variations of conveyed emotions in a speech, a one-unit increase in a transition element implies that the average change in approval rates will increase by β percent. The Definition section serves as a reference when identifying patterns from this data. Furthermore, specific rules are followed for interpreting these coefficients. First, only variables that are significant at a 95% confidence level are considered. Second, only variables with absolute coefficients greater than one are taken into account. Third, priority is given to the effects of transitional elements on the speaker's behavior based on overwhelming magnitudes. These rules are set due to the presence of 36 independent variables, with the aim to focus on the most prominent likelihood of these effects. Below is a subset of variables that fulfill the specified requirements and are thus prioritized for further analysis and interpretation:

	coeff	t
disgust_to_happy	-2.390100	-11.291524
disgust_to_neutral	-1.032708	-4.603276
disgust_to_sad	4.645377	18.031845
disgust_to_fearful	-9.638625	-10.535086
happy_to_happy	-3.815171	-18.210532
happy_to_neutral	3.641925	10.663942
happy_to_angry	6.551711	15.583186
happy_to_fearful	13.523076	12.345104
neutral_to_disgust	3.685323	18.786923
neutral_to_sad	-1.022315	-4.498721
neutral_to_angry	-5.130778	-17.553554
neutral_to_fearful	-18.686601	-20.174953
sad_to_disgust	2.589066	12.503465
sad_to_happy	-1.564187	-9.047883
sad_to_sad	-1.373836	-5.286535
sad_to_angry	6.349106	16.112308
sad_to_fearful	3.150173	2.668857
angry_to_disgust	4.329726	19.475880
angry_to_happy	-6.627076	-22.513568
angry_to_sad	5.468927	18.037085
angry_to_angry	-5.315446	-15.301736
fearful_to_disgust	2.851462	13.906386
fearful_to_neutral	-2.929438	-14.524537
fearful_to_sad	1.473174	7.750466
fearful_to_angry	-3.116959	-12.442668
fearful_to_fearful	3.705838	7.509334

In this subset, there are a total of 13 positive coefficients and 13 negative coefficients. The positive coefficients are mapped to the so-called 'good speaker', while the negative coefficients are mapped to the so-called 'bad speaker'. . To be clear with the rules, consider the coefficients for Disgust_to_Fearful and Sad_to_Sad. Disgust_to_Fearful suggests that the bad speaker incites fear in people of their actions for their personal interest. Sad_to_Sad suggests that the bad speaker probably feels empathetic over victims of a tragic event appearing relatable. However, there is a contrast in information between these two coefficients, leading to examining magnitudes. The coefficient for Disgust_to_Fearful is nine times bigger than Sad_to_Sad, indicating that Disgust_to_Fearful has a more prominent effect on the bad speaker's behavior. In light of this observation, alternative interpretations for Sad_to_Sad should be explored such that it aligns with the overall traits of this speaker. One possibility is that the bad speaker's display of empathy towards victims may come across as mere pity used to create a false sense of trust. Notably, there appears to be a manipulative pattern in the prominent traits of the bad speaker. For instance, the variable with the highest magnitude for the bad speaker is Neutral_to_Fear, which has a coefficient two times bigger than Disgust_to_Fear. This suggests that the bad speaker is skilled at creating a sense of worry around any discussion topic. Additionally, the coefficients for Anger_to_Anger and Neutral_to_Anger, both with similar levels of magnitude, reveal that the bad speaker evokes frustration among the people upon an event to further advance their personal agenda. Furthermore, Disgust_to_Fearful suggests that the bad speaker employs rhetorical remarks to provoke people into becoming sentimental over said event. Moreover, the high magnitudes in Anger_to_Happy and Fearful_to_Anger indicate that the bad speaker justifies their actions by priding themselves as a contributor to 'Making America Great Again'. These transition elements collectively contribute to painting a comprehensive picture of the bad speaker's manipulative tendencies. To conclude, the following key characteristics of each speaker are highlighted:

- A bad speaker can appear open-minded to options, but not empathetic enough to understand the needs of the people. Despite warnings about potential threats, they manipulate people's fears to serve their schemes. The bad speaker uses emotions for personal gains, often resorting to excuses when confronted with corrupt practices. When fighting for their defence, they resort to aggression and display emotional instability.

- A good speaker is a realist aware of the injustice in America. Despite feeling powerless and even distressed at times, they shoulder the responsibility as president suggesting a sense of duty and desire to make a difference. The speaker is not afraid to express their disapproval of misconduct and is deeply affected by the suffering of the people.

8 Conclusion

This research proposes innovative ideas to expand sampling methodologies for behavior studies. As mentioned earlier, the audio files were collected through web scraping, which offers flexibility and freedom in working with data. Leveraging machine learning techniques, valuable information can be extracted from unstructured data, such as natural languages, images, and sounds. Incorporating audio classification into regression analysis opens up new possibilities and sub-fields within the economic domain. The concept of Transition Matrix for regression analysis is also a novel approach. By running numerous combinations, all variations of conveyed emotions in a speech can be effectively captured. Although the interpretations of these variables are open-ended and explained through common logic, this type of research has the potential to advance our understanding of behavioral theories. Overall, I hope that this research inspires further exploration and development in this area.

9 Definitions

- **Neutral to Neutral**: the likelihood that the speaker acts calm and composed when presenting news
- **Neutral to Anger**: the likelihood that the speaker displays a lack of control to remain calm towards an issue.
- **Neutral to Happy**: the likelihood that the speaker delivers news with joy and excitement.
- **Neutral to Sad**: the likelihood that the speaker expresses their sentiment of news with grief and sorrow.
- **Neutral to Fearful**: the likelihood that the speaker highlights the importance of any potential threats from an event.
- **Neutral to Disgust**: the likelihood that the speaker expresses strong disapproval of current events
- **Anger to Neutral**: the likelihood that the speaker can control themselves when being sentimental about a certain topic
- **Anger to Anger**: the likelihood that the speaker is aggressive and confrontational regarding any news
- **Anger to Happy**: the likelihood that the speaker channels hate into prideful messages with conviction.
- **Anger to Sad**: the likelihood that the speaker shows empathy towards victims of controversial events.
- **Anger to Fearful**: the likelihood that the speaker is pessimistic towards the future.
- **Anger to Disgust**: the likelihood that the speaker shows repulsion over violations or social misconduct.
- **Happy to Neutral**: the likelihood that the speaker is unbiased in favour of positive news
- **Happy to Anger**: the likelihood that the speaker creates a sense of intensity towards a positive delivery
- **Happy to Happy**: the likelihood that the speaker is blinded with optimism to show awareness or empathy.
- **Happy to Sad**: the likelihood that the speaker is sentimental over the best of news
- **Happy to Fearful**: the likelihood that the speaker is doubtful for news to be too good to be true

- **Happy to Disgust**: the likelihood that the speaker feels a sense of betrayal of positive expectation.
- **Sad to Neutral**: the likelihood that the speaker does not get overwhelmed by sentimentality of an event
- **Sad to Anger**: the likelihood that the speaker feels frustrated and powerless over a particular issue.
- **Sad to Happy**: the likelihood that the speaker channels tragedy into hopeful messages.
- **Sad to Sad**: the likelihood that the speaker appears compassionate or understanding of the people's struggles
- **Sad to Fearful**: the likelihood that the speaker feels vulnerable over the anticipation of negative outcomes.
- **Sad to Disgust**: the likelihood that the speaker is deeply disappointed over the actions of themselves or others.
- **Fearful to Neutral**: the likelihood that the speaker remains open to possibilities regardless of the circumstances
- **Fearful to Anger**: the likelihood that the speaker is confrontational over an alarming threat.
- **Fearful to Happy**: the likelihood that the speaker displays optimism over the worst case possible.
- **Fearful to Sad**: the likelihood that the speaker is apologetic and shoulders responsibilities.
- **Fearful to Fearful**: the likelihood that the speaker is aware of the potential threat and commits to addressing them
- **Fearful to Disgust**: the likelihood that the speaker displays distress over precarious situations.
- **Disgust to Neutral**: the likelihood that the speaker can be reasonable over actions that they disapprove
- **Disgust to Anger**: the likelihood that the speaker acts outraged over the violation of the moral value
- **Disgust to Happy**: the likelihood that the speaker employs rhetorical techniques to provoke thought or create a humorous effect.
- **Disgust to Sad**: the likelihood that the speaker recognizes injustice with an event
- **Disgust to Fearful**: the likelihood that the speaker attempts to exploit the people's insecurities with threats
- **Disgust to Disgust**: the likelihood that the speaker shares their disapproval over opposing issues.

10 Reference

- Presidential Approval. Roper Center. (n.d.). https://ropercenter.cornell.edu/presidential-approval
- Presidential speeches. Miller Center. (n.d.). https://millercenter.org/the-presidency/presidential-speeches