Federal Shutdown Predictions
Modeled from House Representative's Social Media Presence

Brandon Graviola, Anh Nguyen, & Sarah Wooster

Background

Appropriation Bills are the foundation of federal funding, as they are allocations to specific purposes and agencies. Prior to the end of the United States Government's fiscal year, bills are voted on to provide federal funding for the upcoming year. The fiscal runs from October 1st to September 30th.[1] For a bill to pass, it must be approved by a majority in the House of Representatives, and to prevent a funding gap, the bills must be passed before the end of the fiscal year. If all twelve Appropriation Bills do not pass before the start of the new fiscal year, the government must halt its operations until funding is allocated. This is referred to as a government shutdown.

Training	**Testing**
Agriculture	Labor-HHS-Education
Defense	Transportation-HUD
Energy-Water	Financial Services
Homeland Security	Commerce-Justice-Science
Interior-Environment	
Legislative Branch	
Military Construction-Veterans Affair	
State-Foreign Operations	

Fig. 1. Test/Train Split

On September 28th, 2023, four bills were voted on, with one failing.[2] These votes occurred two days before the fiscal year deadline. With a shutdown looming and still seven bills to pass, the House passed a continuing resolution to prevent funding stoppage. The failure of the Agriculture Appropriation Bill and the lack of urgency for other bills to be voted on caused many to fear a potential shutdown once the resolution ran out. Our team identified the lack of information regarding a shutdown and aimed to alleviate concerns by creating models to predict the upcoming votes' outcomes. We sought to predict House Representatives' votes on upcoming Appropriation Bills based on their social media presence on X, formerly known as Twitter, before prior Appropriation votes. We claim that the attitudes and sentiments of Representatives prior to votes may indicate how they will vote, predicting future bill outcomes.

Data Collection

At the time of the project creation, eight of the twelve Appropriation Bills were voted on, with one failing. Agriculture, Defense, Energy-Water, Homeland Security, Interior-Environment,

[1] *Funding gaps and shutdowns in the federal government*.
[2] Appropriations status table - CRS reports

Legislative Branch, Military Construction-Veterans Affairs, and State-Foreign Operations had been put to a vote.[3] These votes were used as the training dataset, with the remaining four bills used for testing and predicting a potential shutdown (see **Figure 1**). The set of swing voters used was identified by GovTrack as "statistically notable votes are the votes that are most surprising, or least predictable, given how other members of each voter's party voted and other factors." [4] Thirty-one House Representatives were identified as statistically notable across the eight bills, of which twenty-six were Republican and five were Democrat (see **Fig. 2**).

Representative	Party Affiliation
Juan Ciscomani	R
David Valadao	R
Young Kim	R
Mariannette Miller-Meeks	R
Ashley Hinson	R
Zach Nunn	R
Randy Feenstra	R
Tracey Mann	R
John Moolenaar	R
Brad Finstad	R
Michelle Fischbach	R
Don Bacon	R
Thomas Kean	R
Nicholas LaLota	R
Andrew Garbarino	R
Michael Lawler	R
Frank Lucas	R
Lori Chavez-DeRemer	R
Brian Fitzpatrick	R
Nancy Mace	R
Monica De La Cruz	R
Dan Newhouse	R
Ken Buck	R
Jared Golden	D
Marie Gluesenkamp Perez	D
Tim Burchett	R
Marcus Molinaro	R
Vicente Gonzalez	D
Mary Peltola	D
Donald Davis	D
Marjorie Greene	R

Fig. 2. Swing Votes

Our data was collected using a tweet scraper, referred to as Tweet Flash, hosted on Apify.[5] The data provided via CSVs contained various variables, including retweets, mentions, and the tweet itself. To assess the attitudes and sentiments before each vote, a timeline of one week prior was identified; for example, the Military Construction-Veterans Affair's vote occurred on July 28th, so its designated timeline was July 21st to July 28th. The specified time frame per tweet and the username were inputted. The bills that were voted on within the same time frame were duplicated, despite containing the same tweets, as their votes across the bills were inconsistent.

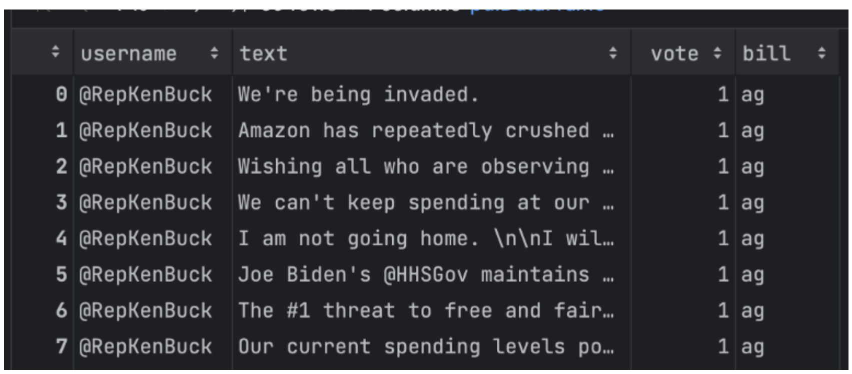
---

[3] Appropriations status table - CRS reports
[4] *H.R. 4368: Agriculture, Rural Development, Food and Drug Administration, and ... -- House vote #507*
[5] *Tweet Flash*

Methods

Our team divvied up data collection for the thirty-one swing voters, reflected in the organization of our CSV files and Jupyter notebooks. Each team member collected data for 10-11 representatives. To prepare the data for modeling, the *username* and *text* column—the tweet itself—were retained from the original scraped dataset. A function was used to iterate through each bill and the corresponding vote, assigning a value to the *vote* column. A value of "0" indicated a "Nay" or "Abstain" from the representative, while "1" indicated a "Yea". The *vote* column would be used as the response variable.

Once the data was prepared, the final product for each representative contained four columns (see **Fig. 3**): *username*, *text*, *vote*, and *bill*. Each representative contained the tweets one week prior to the vote with the corresponding vote for all bills. In order to fit a model, a vectorizer was used to convert the *text* column into a matrix of features. Vectorizers work by creating a vocabulary of unique words in the given document and recording the frequency of the word per row.



	username	text	vote	bill
0	@RepKenBuck	We're being invaded.	1	ag
1	@RepKenBuck	Amazon has repeatedly crushed …	1	ag
2	@RepKenBuck	Wishing all who are observing …	1	ag
3	@RepKenBuck	We can't keep spending at our …	1	ag
4	@RepKenBuck	I am not going home. \n\nI wil…	1	ag
5	@RepKenBuck	Joe Biden's @HHSGov maintains …	1	ag
6	@RepKenBuck	The #1 threat to free and fair…	1	ag
7	@RepKenBuck	Our current spending levels po…	1	ag

Fig. 3. Example of Ken Buck's dataset.

To create models for all thirty-one representatives, we created a function to iterate through the representative names in each batch of voters and fit a model for each person's tweets. The model was fitted using the *vote* column as the response variable and the features collected via a vectorizer as the predictors.

A variety of models were fitted to the thirty-one swing voters, including Random Forest, Logistic Regression, Gradient Boosting, the combination of AdaBoost with Random Forest using a soft voting classifier, the combination of Gradient Boosting with Random Forest using a soft voting classifier, and Support Vector Classification (SVC). Cross-validation using five folds was used to evaluate each model's robustness on the entire dataset. Confusion matrices and metrics from a classification report, including individual and weighted precision, recall, and F1 scores for each class were used to analyze class imbalance issues.

Our team aimed to fit the same class of model to all thirty-one swing voters, which proved difficult with varying training set sizes for each voter. Representatives who voted one way for most of the appropriations suffered from a major class imbalance in their resulting models. Voters who posted fewer tweets in the timeframe of interest faced lower model accuracies. The most challenging part of our project was dealing with these difficulties across a large number of models.

Random over-sampling and synthetic minority over-sampling techniques (SMOTE) were solutions we explored to address the class imbalance. We chose to implement the former, as it was easier to implement and SMOTE did not provide significant improvements when tested on a control model.

Our team also chose to implement a TF-IDF vectorizer instead of a Count Vectorizer as TF-IDF vectorizers are known to perform relatively well on short documents (such as tweets) and can reduce noise.

Other pre-processing methods for natural language processing such as stemming lemmatization were tested on our control model (a Logistic Regression model) but did not significantly improve our results. Thus, we opted out of using these techniques.

Through lengthy experimentation with many models across many voters, our top four contenders were Random Forest, AdaBoost with Random Forest, Logistic Regression, and SVC. Hyperparameter tuning was applied to all four models, which all but Logistic Regression benefited from. Across the board, these models performed relatively similarly, which made fine-tuning for model performance difficult.

Our team chose to use the cross-validation scores to decide which model we should use to make our final predictions. We wanted a model with a relatively good cross-validation average and variance, preferring models with lower variance to reduce the chance of overfitting.

Taking the average cross-validation and variance for the 10-11 swing voters for each batch of models, we found that AdaBoost with Random Forest and SVC were the top contenders for all three batches. SVC, however, had higher accuracy scores and lower variance across the board. A shortcoming to be addressed in the future is the class imbalance that persists in all of these models, which again, could be a reflection of our data and model fitting.

Another shortcoming in our methodology is our model fitting, specifically making decisions to fine-tune based on the results of the majority. Alongside our choice to use the same model type for all representatives, being unable to fine-tune each model one by one makes it difficult to account for complexities specific to each person. Future studies would benefit from personalizing each model to each representative, at the cost of computational expense.

Representative	Mean Test Prediction	Vote Prediction
Don Bacon	0.96	Yea
Ken Buck	0.18	Nay
Tim Burchett	0.5	Yea
Lori Chavez-DeRemer	1	Yea
Juan Ciscomani	1	Yea
Donald Davis	0.14	Nay
Monica De La Cruz	1	Yea
Randy Feenstra	1	Yea
Brad Finstad	1	Yea
Michelle Fischbach	1	Yea
Brian Fitzpatrick	1	Yea
Andrew Garbarino	0.88	Yea
Jared Golden	0.5	Yea
Vicente Gonzalez	0	Nay
Marjorie Greene	0.88	Yea
Ashley Hinson	1	Yea
Thomas Kean	1	Yea
Young Kim	0.96	Yea
Nicholas LaLota	1	Yea
Michael Lawler	0.97	Yea
Frank Lucas	1	Yea
Nancy Mace	1	Yea
Mariannette Miller-Meeks	0.95	Yea
Marcus Molinaro	0.96	Yea
John Moolenaar	1	Yea
Dan Newhouse	1	Yea
Zach Nunn	1	Yea
Mary Peltola	0.18	Nay
Marie Gluesenkamp Perez	0.87	Yea
David Valadao	1	Yea

Fig. 4. Model Predictions

In the final iteration of our model fitting, we compared the results of SVC with a linear vs. non-linear kernel, with the latter performing slightly better overall. Thus, we made our predictions by fitting SVC with a non-linear kernel to our entire dataset, using the last week of November as our test set. We decided to use this week as our test set because no appropriation votes have been scheduled as of this point in time. After fitting SVC models to all thirty-one voters, we took the mean prediction of each row as their final prediction, with a mean prediction greater or equal to 0.5 as "Yea."

Results

Fitting SVC models to the data associated with the previous eight appropriations to predict the outcome of a vote on our test data, the majority of our swing voters would vote "Yea" on a theoretical vote occurring on December 4th. Our models predicted that only four voters — Buck, Davis, Gonzalez, and Peltola — would vote "Nay" on the theoretical bill (see **Fig. 4**).

Comparing our predictions using SVC with a linear vs. non-linear kernel, a linear kernel model predicted an additional representative voting "Nay" — Burchett. With a non-linear kernel, Burchett is predicted to vote "Yay" with a mean prediction of 0.5.

Given that Random Forest and SVC with a non-linear kernel performed the best on our data set, this may suggest that the true classification boundary of our data may be complex or non-linear.

Future studies would benefit from exploring additional models that are suited for non-linear classification boundaries, experimenting more with hyperparameter tuning, and using data from personal Twitter accounts to increase the training set size.

Conclusion and Future Work

In this paper, we investigated whether swing votes in the U.S. House of Representatives can be predicted utilizing sentiment analysis of Tweets, with the intention of using such predictions to gauge the likelihood of a federal government shutdown. The results derived from our experiments provide valuable insight into the application of statistical models, particularly SVCs, in predicting voting outcomes. Our team experimented with both linear and non-linear classification boundaries in our SVC models and found that the latter performed best. Additionally, our SVC models' ability to accurately predict swing votes underscores the potential application of such models in other political contexts, such as elections.

Given the current House of Representatives size — 435 — and traditionally non-swing voters will vote on party lines, thirteen swing voters need to vote "Yea" for a bill to pass. Our model consistently predicted that the theoretical bill would pass, fluctuating between only four and five "Nays". This indicates that a government shutdown is unlikely; however, as noted previously, a major issue with passages is the scheduling of votes. Our model is dependent on the House's ability to convene for a vote, which has not occurred. Future research would benefit from additional data on the scheduling of bills. For now, our results can be interpreted as reassurance that another failure to pass, as what occurred with the Agriculture bill, would not occur.

Works Cited

Funding gaps and shutdowns in the federal government. US House of Representatives: History, Art & Archives. (n.d.). https://history.house.gov/Institution/Shutdown/Government-Shutdowns/

Appropriations status table - CRS reports. (n.d.). https://crsreports.congress.gov/AppropriationsStatusTable/

H.R. 4368: Agriculture, Rural Development, Food and Drug Administration, and ... -- House vote #507 -- Sep 28, 2023. GovTrack.us. (n.d.). https://www.govtrack.us/congress/votes/118-2023/h507

shanes. (n.d.). *Tweet Flash*. Apify. https://console.apify.com/actors/wHMoznVs94gOcxcZl/information/0.0.160/readme

Contributions

Brandon Graviola: Project research
Anh Nguyen: Model creation and validation
Sarah Wooster: Data cleaning and project ideation