CAPP 30255: Advanced Machine Learning: Measuring Political Compatibility with node2vec and moral values scale

Andrei Bartra and Oscar Noriega

May 27, 2021

Abstract

Elections season is characterized by high information variance where much of the opinion formation process comes from feedback loops within social networks. In this scenario, it is not easy to have unbiased tools to introspect about ideas' compatibility and those of candidates. In this project, we propose a tool based on the Relative Universalism Index proposed by Enke [2020] to compare moral values foundations of users and candidates. To increase the precision of the estimates, we propose an extension of the Moral Foundations Dictionary (MFD 2.0) by estimating a word embedding using US political speeches from November 2020 to April 2021.

Github Repository: AML Project

Keywords— moral values, word embedding, node2vec, deep learning, political compatibility

1 Introduction

Elections season is characterized by variance in political information and non-uniform exposure to all the candidates' proposals and discourses[1], which may lead citizens to choose a candidate with a biased subset of the political information.

This project aims to give citizens an additional piece of political information to understand candidates' overall political stance and find compatibility with them based on their media-covered speeches. Of course, this subset of data is a narrow approach to the overall political compatibility with a candidate. The goal is to give information independent from the resonance box of the local social network to trigger self-examination.

The *compatibility* measure is the Relative Universalism Index (RUI), borrowed from Enke [2020], where he contrast universalist values (fairness and care) against communal values (loyalty and authority). The novel methodology of Enke [2020] deviates from the classic right/left, liberal/conservative spectrum. According to Enke [2020], the moral values scale is more stable, simple, and less contradictory

There are two main ways to measure the RUI: (i) measuring the relative frequency of words related to the moral foundations in speeches by matching against the Moral Foundations Dictionary (MFD 2.0) , or (ii) measuring moral values using the Moral Foundations Questionnaire (MFQ). The problem is that we can not gather the RUI from all the political personalities using the MFQ [2]. Comparing the RUI obtained by word frequency of political speeches with user RUI obtained from the questionnaire directly may raise doubts on the comparison. Therefore, the solution is to measure RUI's with word frequencies, for both users and speakers by allowing open questions in the MFQ.

One challenge with this approach is that the MFD, in its extended version, has only 2,103 words. Having a small list is not a problem for political personalities because it is possible to gather several speeches and observe enough realizations of those words. However, gathering the same amount of information from a user may be unattainable. One way to approach this is by expanding the MFD using Natural Language Processing techniques (NLP).

A word embedding is a vector representation of words in a latent space, where relative distances between words ingrain information about the semantic and syntactic characteristics. This methodology is an efficient way to expand the dictionary because words close to each other are either words with similar meanings or frequently related words.

For this projects we fit a custom node2vec (Grover and Leskovec [2016]) embedding. Then, we fit a classifier using the embedding of the words in the MFD. This classifier is then used to predict the moral foundation label for the rest of the words. Conveniently, the output vector for each word is a vector with the probability distribution of each moral foundation, which we can use as weights in the RUI estimation.

The methodology follows Garten et al. [2016] methodology for moral sentiment analysis in social media posts. It uses the MFD to measure the distance of words in Twitter posts to the dictionary's words within an embedded word representation. The main difference is using a different embedding methodology (node2vec instead of word2vec), another data set (speech transcripts instead of tweets), and focus on the RUI estimate instead of the whole moral foundation spectrum.

This report is organized in the following way: In section 2: Moral Foundations, we explain the Moral Foundations theory and the application that inspired this project. In Section 3: Data cleaning, we give a description of the data and its cleaning process. Then, in section 4: Node2vec, we give a detailed explanation of the node2vec methodology. We then pass to section 5: Moral Foundation Classifier, describing the moral foundation classifier model. Section 6: Relative Universalism Index Estimate describes the steps to estimate the RUI for each political personality speech. In section 7:

[1] Citizens may be more exposed to biased information aligned with their social networks. Moreover, not all topics are covered equally.

[2] For reasons we believe are apparent

User Interface, we show a proof of concept of how a user could get estimates of its RUI and political compatibility. Finally, in section 8: Effort, we discuss conclusions and lessons from the project.

2 Moral Foundations Theory

In his paper "Moral values and voting", Enke [2020] explores heterogeneity in moral values and their role in US presidential elections voting behaviors, paying close attention to 2016. The primary mechanism behind Enke's analysis relies on the assumption that "... voters aim to minimize the distance between their own moral type and the weighted average type of the candidate and their party."

As the main theoretical framework for being able to compute a measure of moral values, Enke based his work on the Moral Foundations Theory (MFD) developed in Haidt [2007] and Haidt [2012]. This theory was developed to understand the high observed variance in moral values across cultures and how they somehow seem to be anchored in some "core" elements that imply certain similarities between them.

Based on this notion, they defined five foundations of moral concerns:

- **Care/Harm**: Protection of others. Includes notions of kindness and empathy
- **Fairness/Cheating**: Ideas related to justice, rights and equality.
- **Loyalty/Betrayal**: Relevance of standing by your group (ranging from family, community, country, etc.)
- **Authority/Subversion**: Respect of hierarchical society structures and the importance of traditions
- **Sanctity/Degradation**: Mainly concepts related to religious notions of "purity"

The first two foundations are understood as being more "universalist", in the way that they are usually required to be applied to everyone, regardless of any differentiation or group belonging. Simultaneously, the next two are considered more "communal", where they are applied more to a particular group, nation, institution, or so. However, the fifth foundation cannot be easily categorized into this universalist-communalist dichotomy, so it is not considered for the subsequent analysis.

This communal/universalist classification was used by Enke [2020] to analyze the speeches, statements, debates, and press releases of the US presidential candidates since 2008, by constructing an index of relative universalism using relative frequency of terminology associated with both universalism and communality.

The relevant terms to be used as representative for either universal or communal values correspond to those found in the Moral Foundations Dictionary (MFD 2.0), a list of keywords created by psychologists that can directly be associated with each of the MFT foundations either as a virtue or a vice.

$$\text{RUI} = \frac{f_{\text{care}} + f_{\text{fairness}} - f_{\text{In-Group}} - f_{\text{Authority}}}{|D|} \tag{1}$$

where

$$f_{\text{moral value}} = \frac{\frac{\sum_{z=1}^{N_f^v} w_z}{N_f^v} + \frac{\sum_{z=1}^{N_f^m} w_z}{N_f^m}}{2} \tag{2}$$

RUI stands for the Relative Universalism Index, where f_{value} is the weighted average frequency of a value, being N_f^v the total sum of *vice* unique words for foundation f in entire document set (N_f^m for virtue words) and w_z the frequency of word z in the text.

Therefore, this relative universalism index[3] provides a proxy of each candidate's moral alignment, identifying that Donald Trump's rhetoric during the 2016 presidential campaign was considerably leaned towards communal values (or a negative relative universalism index). These values were in line with the nationalist movement that took place in America and across Europe, mentioned by Graham et al. [2018]. The most salient manifestation ended up being Trump's victory in 2016 over Hillary Clinton, who had a more relative universalist speech.

This approach, as Enke mentions, allows us to have a one-dimensional measurement of morality, but rather than taking it in absolute terms (i.e., stating that candidate A has "more moral" than candidate B), we have a measurement of the type of morals deemed important by the individual.

3 Data & Cleaning

3.1 Data

For our project, we required two sources of data: The aforementioned Moral Foundations Dictionary and political speech transcripts to analyze. In this section, we briefly describe the data.

3.1.1 Moral Foundations Dictionary 2.0 (MFD)

The MFD is a set of keywords labeled with the 5 Moral Foundations Values (care, fairness, in-group, authority, purity). According to their positive or negative association, the keywords are flagged with virtue or vice flag, respectively, to the 5 MFT foundations (care, fairness, in-group, authority, purity).

The original MFD constructed by Jesse Graham and Jonatan Haidt in 2009 contains 318 words for the vice/virtue categorization for the five moral foundations. We considered this to be quite restricting for our project. We found out that in 2019, Jeremy Frimer, a psychology professor from the University of Winnipeg, developed with collaboration from Graham and Haidt a more extensive version of the MFD, dubbed MFD 2.0.

MFD 2.0 contains 2,103 words for the aforementioned moral foundations' classification. The number of words is nearly seven times the number of words from the original MFD. A comparison between these two FD versions is summarized in Table 1:

Table 1: Moral Foundations Dictionaries comparison

Moral Foundation:	Care		Fairness		Loyalty		Authority		Sanctity		
	Virtue	Vice	Virtue	Vice	Virtue	Vice	Virtue	Vice	Virtue	Vice	**Average**
Original MFD	16	35	26	18	29	23	45	37	35	54	**32**
MFD 2.0	182	288	115	236	142	49	301	130	272	388	**210**

In order to reproduce Enke [2020] proposed universalism scale, we consider only the foundation pairs (care, fairness) and (in-group, authority) in order to define universal and communal moral values, respectively.

3.1.2 Political speeches

In order to estimate political characters RUI, our analysis makes use of data on political speeches from press conferences, rallies, debates, and official announcements transcripts from Rev, a service specialized in audio-to-text conversion. Their transcript data is publicly available on their website and continuously updating (as of May 26, 2021, it contained 3,570 political transcripts).

[3] Under the Enke [2020] context, for example, this index of relative universalism, rather than stating if the candidate holds more "universal" moral values compared to communal ones, measures how important universalism values are in his/her speech in terms of relative frequency. This implies that a candidate who potentially speaks against universal values for a long time in a speech would have a high index value.

For our analysis, we developed a web-scraper using Selenium to download the Rev website transcripts. We considered a time frame of approximately five months from late November 2020 to mid-April 2021, for which we gathered 1,319 transcripts using our web-scraper.

3.2 Data cleaning

The structure of each transcript is divided by speaker interventions during an event, as illustrated in Figure 1. Instead of treating a speech as a single sequence of words, we kept the structure to help us identify the most informative parts of said speech and drop relatively irrelevant interventions.

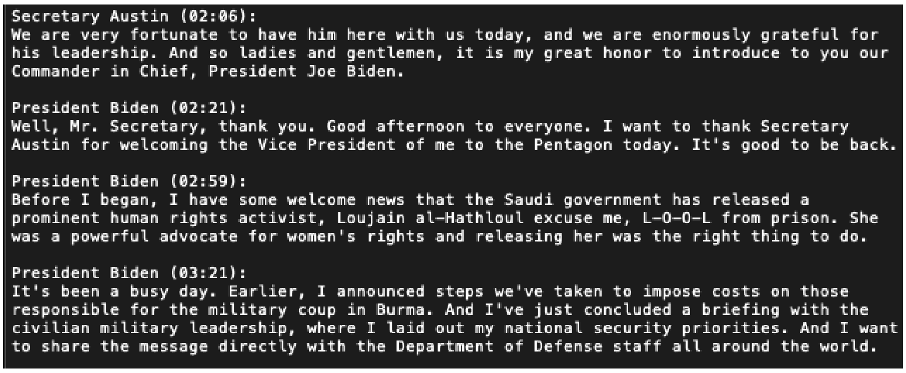

```
Secretary Austin (02:06):
We are very fortunate to have him here with us today, and we are enormously grateful for
his leadership. And so ladies and gentlemen, it is my great honor to introduce to you our
Commander in Chief, President Joe Biden.

President Biden (02:21):
Well, Mr. Secretary, thank you. Good afternoon to everyone. I want to thank Secretary
Austin for welcoming the Vice President of me to the Pentagon today. It's good to be back.

President Biden (02:59):
Before I began, I have some welcome news that the Saudi government has released a
prominent human rights activist, Loujain al-Hathloul excuse me, L-O-O-L from prison. She
was a powerful advocate for women's rights and releasing her was the right thing to do.

President Biden (03:21):
It's been a busy day. Earlier, I announced steps we've taken to impose costs on those
responsible for the military coup in Burma. And I've just concluded a briefing with the
civilian military leadership, where I laid out my national security priorities. And I want
to share the message directly with the Department of Defense staff all around the world.
```

Figure 1: Example of a Rev transcript structure

For example, in Figure 2, we can see that the interventions of Joe Biden and Anderson Cooper do not reflect any valuable information for trying to identify the RUI for any of them. Analyzing the interventions, we found out that starting at a 20-word count, we started to see some terms contained in the MFD and more relevant ideas or statements. So, we filtered out all the interventions with a word count of less than 20.

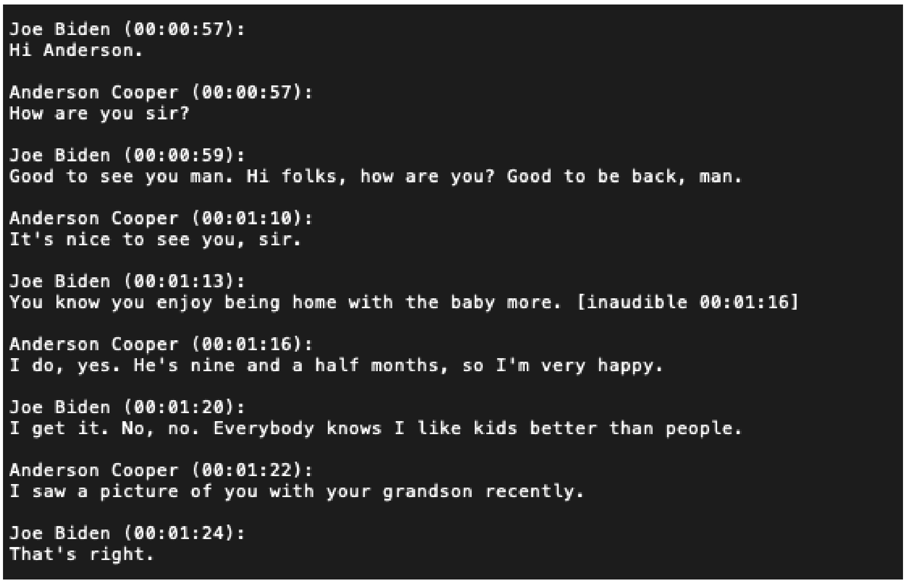

```
Joe Biden (00:00:57):
Hi Anderson.

Anderson Cooper (00:00:57):
How are you sir?

Joe Biden (00:00:59):
Good to see you man. Hi folks, how are you? Good to be back, man.

Anderson Cooper (00:01:10):
It's nice to see you, sir.

Joe Biden (00:01:13):
You know you enjoy being home with the baby more. [inaudible 00:01:16]

Anderson Cooper (00:01:16):
I do, yes. He's nine and a half months, so I'm very happy.

Joe Biden (00:01:20):
I get it. No, no. Everybody knows I like kids better than people.

Anderson Cooper (00:01:22):
I saw a picture of you with your grandson recently.

Joe Biden (00:01:24):
That's right.
```

Figure 2: Example of irrelevant interventions on a transcript

Finally, we kept only those speakers who had a count of at least 268 interventions in our data set. This number is somewhat arbitrary, and we considered it to include the interventions we had of Barack Obama. While the threshold is arbitrary, the reason behind this filter is to guarantee that we have enough data on each speaker in order to estimate their RUI distribution as accurately as possible.

This data cleaning process left us with a data set of 26,183 interventions by 25 speakers with an average of 101.23 words per intervention.[4]. Figure 3

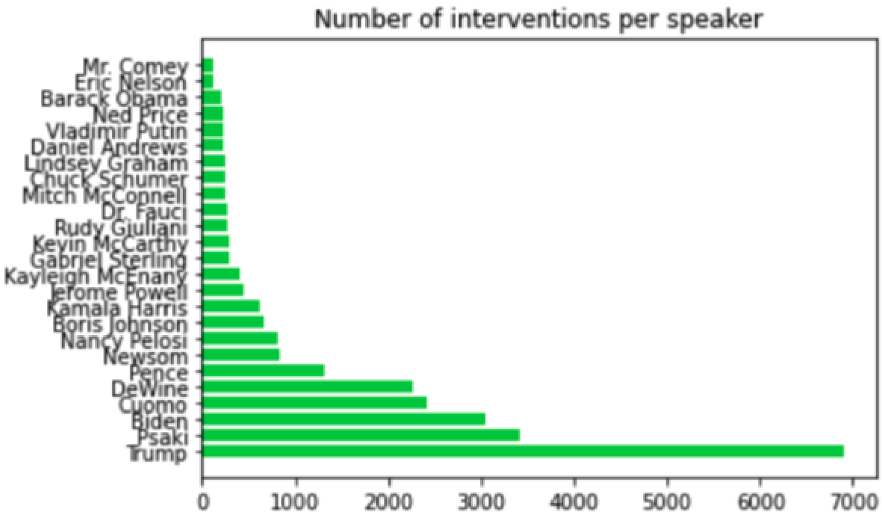


Figure 3: Number of interventions by speaker in our sample

3.3 Network representation

Once we got our data cleaned up, we require setting our data in a network representation to use it as an input for the node2vec algorithm,

We structured the data in a Weighted Directed Acyclic Graph of all the word pairs in our corpus, using the Networkx package. We achieved this by organizing each intervention as a series of bi-gram tuples to form the network nodes and edges and recording the frequencies. In this case, each word acts as a node, the vertices are the sentences' adjacencies, and weights are the frequency each node appears.

Once we transformed our data, we passed that to Networkx to create the network structure, which turned out with 24,888 nodes and 605,714 edges at first. It is important to notice that node2vec requires a strongly connected graph[5], which was not the case for our network. We kept the biggest strongly connected component of our network. Luckily, we did not lose many words and ended up with 24,629 nodes and 605,388 edges for our strongly connected network.

Figure 4 provides an illustration of our data network structure.

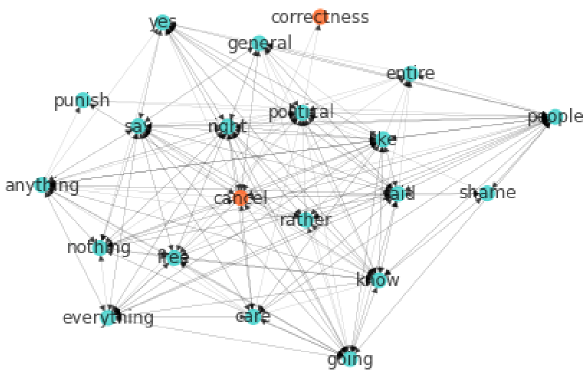


Figure 4: Network representation of the shortest paths between the words "correctness" and "cancel"

[4]For this total, we also excluded generic speakers (identified as "Speaker 3", "Crowd", and such) and dropping stop words

[5]A graph where it is possible to get from one node to any node

4 Node2vec

A standard approach for word embedding is the word2vec algorithm (Mikolov et al. [2013]). word2vec inputs sequences of indexed words through a two-layered neural network and outputs a vector representation of the prediction of the *next word*.

The main advantage of this approach is that it is simple and does an excellent job ingraining the syntactic structures of speeches. However, in our context, these advantages are not optimal. Our primary interest is to recognize words frequently related to each other and with similar meanings but avoid them being constrained by the specific constructs of the political personalities' speeches. In other words, we do not want to confuse the political compatibility of a user with idiolect compatibility.

Our approach to this challenge is to use a node2vec instead of word2vec. Node2vec starts with a weighted directed graph representation of words (see section 3). The advantages are that instead of a linear sequence of each intervention, the embedding process considers other potential sequences. Although not perfect, at least it contributes to breaking the idiolect dependency of the embedding.

Node2vec works by taking biased random walks through the graph representation of the corpus. The idea is to control the extent to which the random walk behaves like a Depth for Search Algorithm (DFS): Traversing the maximum distance possible from the original node). Or, like a Breath for Search (BFS) Algorithm that prioritizes covering the nodes in the neighborhood.

Considering a walk from node w_0 to w_1 and candidates $[x_1, x_2, x_3, w_0]$ for the next step. Te next node is chosen at random with a probability function $f(\alpha_{p,q}) = \frac{b\alpha_{pq}}{\mathbb{N}}$. Where b is the weight of the edge, $\mathbb{N}$ is a normalization constant (sum of all numerators) and $\alpha_{p,q}$ is the un-normalized weight that controls the BFS or DFS behavior with the following rule:

$$\alpha_{p,q} = \begin{cases} \frac{1}{p} & \text{if } \ d(w_0, x) = 0 \\ 1 & \text{if } \ d(w_0, x) = 1 \\ \frac{1}{q} & \text{if } \ d(w_0, x) = 2 \end{cases}$$

If the distance of the original node w_0 to candidate node is $d(w_0, x) = 0$ (returning), the probability will be proportional to $\frac{1}{p}$. If $d(w_0, x) = 0$ (adjacent to w_0), the probability will be proportional to 1. Finally, if $d(w_0, x) = 2$ (adjacent to the current node w_1), the probability will be proportional to $\frac{1}{q}$.

p is called the *return parameter*, as it controls the likelihood to return immediately. q is the *in-out parameter*, as it controls the BFS-DPS behavior of the walk. This process is illustrated in Figure 5.

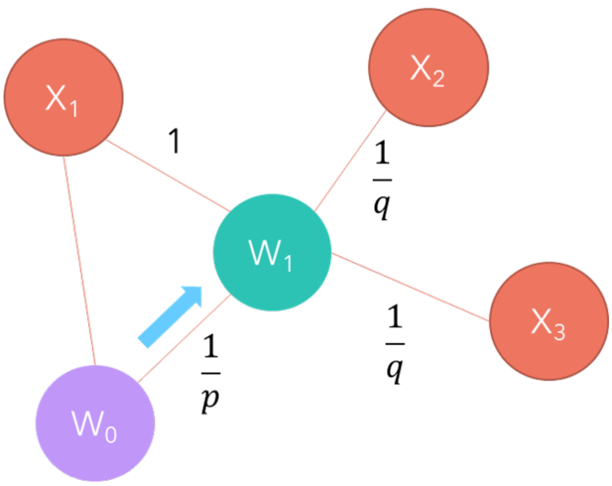
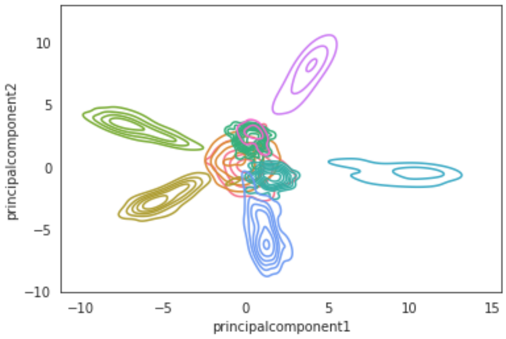


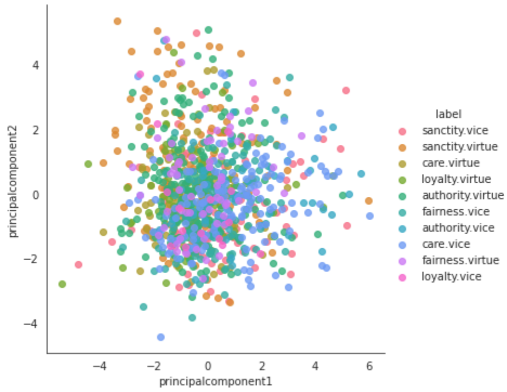
Figure 5: Illustration of a biased random walk procedure in *node2vec*. Own elaboration

For this project, the sequences were generated with 30 walks per node of length 80 using the largest strongly connected component of the network structure of the corpus. The embedding was trained with windows of size 15 with batches of size 10.

This embedding is not good enough to identify "regions" of the moral foundations. However, after training a neural network, the regions are clear:



(a) After Neural Network Training



(b) Before Neural Network Training

Figure 6: Principal components of the Embedding. Own elaboration

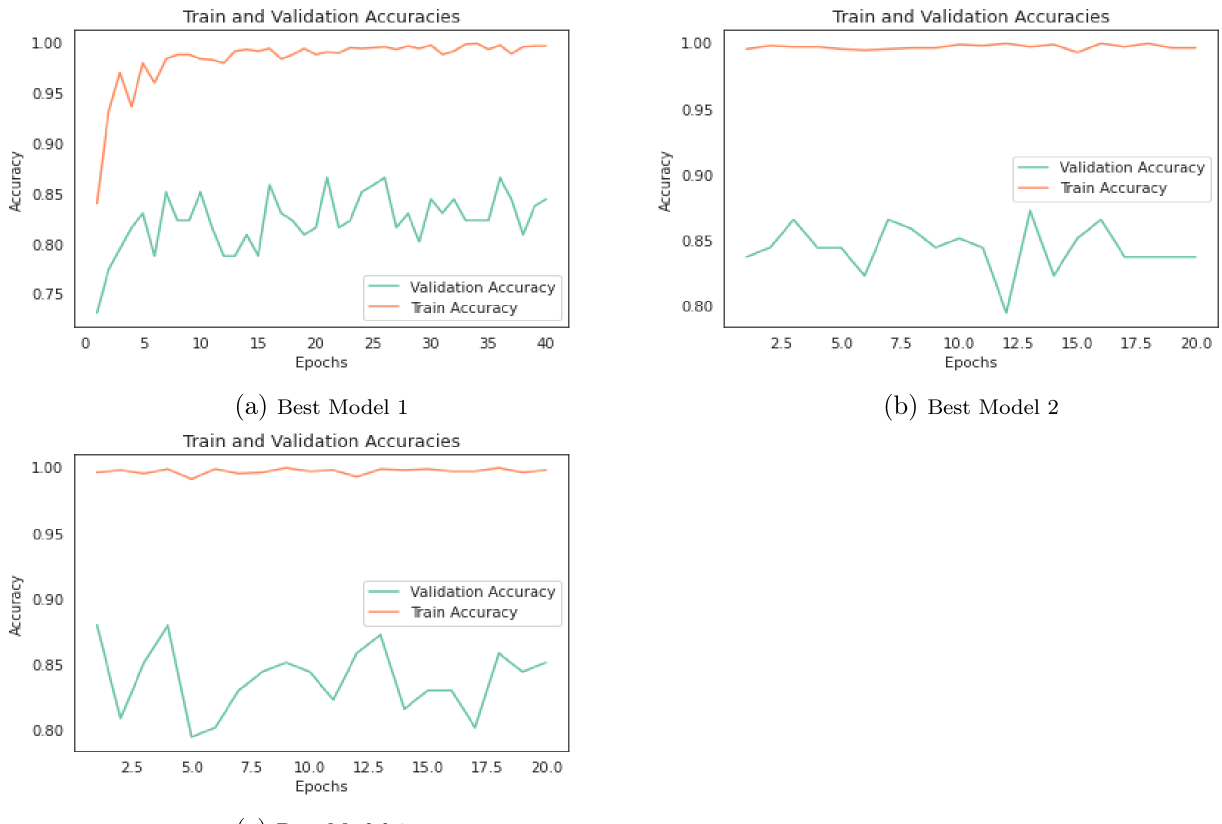
5 Moral Foundation Classifier

We can proceed to fit a moral foundation tagger with a successfully trained embedding for the words in the MFD that we can use to estimate the moral foundation probability distribution of words outside the dictionary.

The process goes a follows:

1. Following Enke [2020], we filter the words labeled as sanctity/degradation.
2. In Table 1 it can be seen that the categories are highly unbalanced. We balanced the label distribution following a random (over/under)sampling procedure to fix the examples of each label to 185.
3. The samples were partitioned following a standard 80%, 10%, 10% partition for train, validation, and test samples, respectively. In order to keep the balance, we did this stratifying by label category.
4. We matched the words in the MFD words with our embedding and fed the vocabulary keys into a data loader setting the batch size to 4
5. We fit a multi-layer neural network with the following specifications:
    - The embedding layer is set to be trainable.
    - 5 hidden linear layers with sizes: [300, 400, 500, 100, 20]
    - 5 corresponding non-linear activation functions: [Tanh, ReLU, ReLU, CELU, CELU]
    - 5 corresponding dropout layers with the following probabilities: [0.9, 0.8, 0.5, 0.3, 0.05]
    - Negative Log-Likelihood Loss function
    - Trained for three stages, keeping the model with the best validation loss for the next stage.
    - The first stage was trained with a learning rate of 1e-3 for 40 epochs. The next stage with 1e-4 of learning rate for 20 epochs, and the last stage with a learning rate of 1e-5 for 20 epochs.

Figure 7 shows the train and validation accuracies for each of our three stages. At first, we trained the model for 40 epochs and kept the best performing model (Best Model 1), then we re-train that model for 20 epochs and look for any improvement, keeping the best performing (Best Model 2), and finally, we do the same process a third time (Best Model 3).



(a) Best Model 1

(b) Best Model 2

(c) Best Model 3

Figure 7: Train and Validation accuracies for our Best Models

The final result was an embedding matrix of size 24,629×8, which allowed us to classify all the relevant words in our corpus to the 8 Moral Foundations categories. The test accuracy is 85.1%. The confusion matrix of the test sample is shown in figure 8

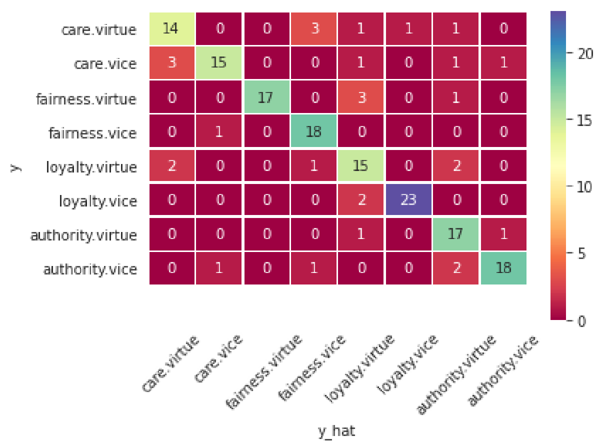


Figure 8: Confusion matrix of the test sample of the Moral Foundations Classifier. Own elaboration

6 Relative Universalism Index Estimate

In Enke [2020], it is shown that the RUI index of a person may vary according to the political context. This variability must not be confused with variance on the estimate, as this variance shows consistency with expected political behavior. Therefore, a distribution of RUI per speaker across all their interventions would be more informative than a point RUI index estimate over the whole body of interventions of a speaker.

The first step is to set the probability of the words on the MFD to 1 for moral foundations labeled words. This step is to keep consistency with the original methodology; this also helps make the distribution less centered, and, at least for those words, we clear the bias and variance from our model.

On the other hand, some words are either too common or have very ambiguous moral foundations distribution. Having a high number of such words tends to narrow the distribution around zero making the estimates hard to compare. To solve this, we propose using the Herfindahl Index (HHI) to discard the words that tend to have a uniform distribution.

The HHI is used in various contexts to measure how concentrated is a distribution of proportions on a subset of categories. To visualize the methodology in figure 9 we illustrate how the HHI works for the case of two categories. The vectors of proportions form a line because they have to sum to 1. Establishing a cut-off on the sum of squares is, in practice defining a circle (or hypersphere) around the origin, where all the vectors inside the circle are the discarded vectors.

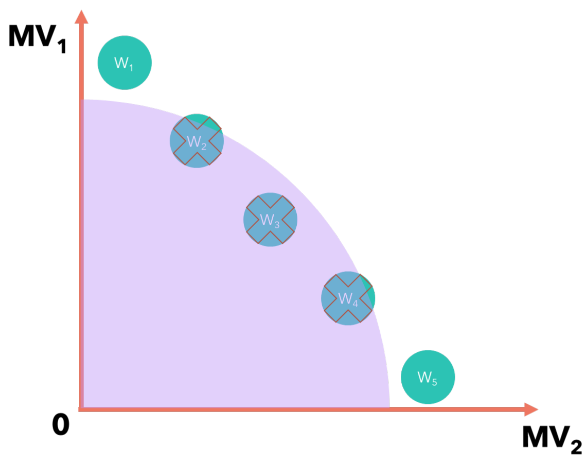


Figure 9: Illustration of HHI with two categories. Own elaboration

For this exercise, we discarded the bottom 80% of the words in the HHI index. We were expanding the MFD from 2,103 words to 17,240 from the 24,629 original words in the vocabulary.

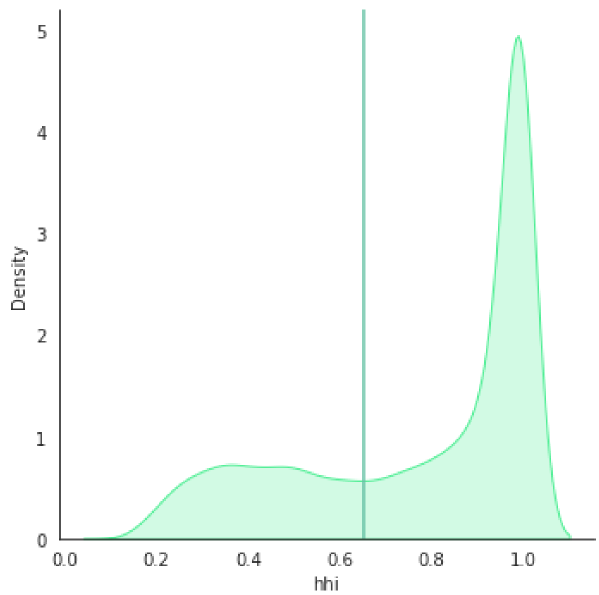


Figure 10: HHI cut-off. Own elaboration

Considering this final set of words and their distributions, we estimate the RUI index for each intervention following equations 1 and 2. Instead of adding the frequencies, we add the probabilities to lower the weight of our estimates relative to the words in the original MFD.

The initial distribution of RUI shows long tails. The presence of outliers can distort the estimates. Therefore we decided to cut outliers by considering RUI above -4 and below 4.

With our RUI estimates we can now calculate the individual RUI densities for each of the speakers of our data. Since the final objective is to compare a user's estimated RUI with other political figures to determine who is more similar, we require the individual distributions.

In order to have a correct visualization of these distributions, in Figure 15 we show the densities for six different characters. From there, we can see that Trump's RUI distribution leans towards communal values (in line with the results of Enke [2020]), while Barack Obama is a little leaned towards universalist values.

The previous graph would allow the user to compare their RUI value with the distributions of the speakers. Given an RUI value, one can see which character is more likely to hold the same moral values stance.

(a) RUI Distribution with Outliers

(b) RUI Distribution without Outliers

Figure 11: RUI distribution, with and without outliers. Own elaboration

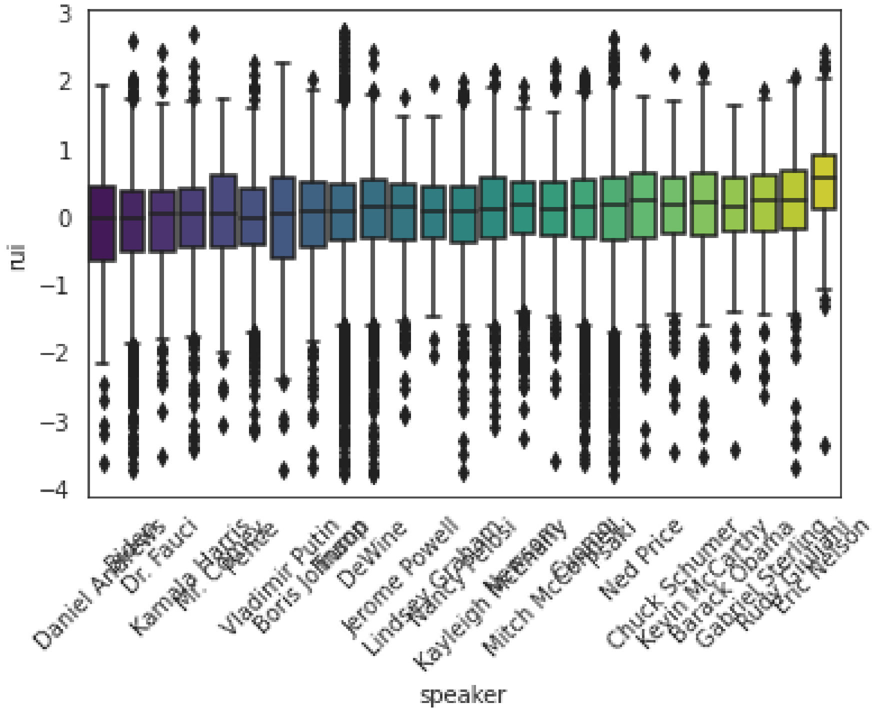


Figure 12: RUI distributions for all the speakers

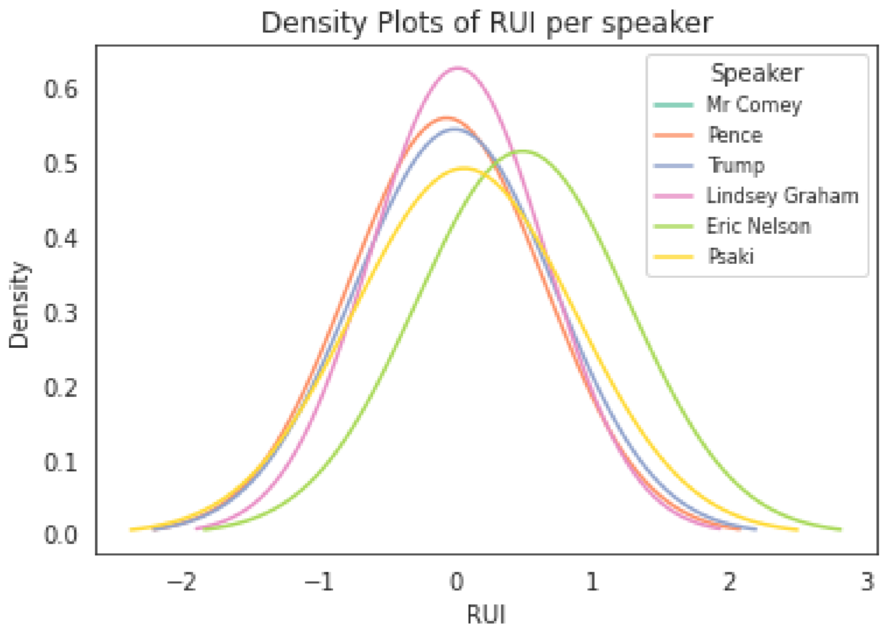


Figure 13: RUI estimated density for certain speakers

7 User Interface

Finally, we added proof of concept of how the user interface could look like. We adapted the Moral Foundations Questionnaire to allow open questions.

The original questionnaire had questions in the form:

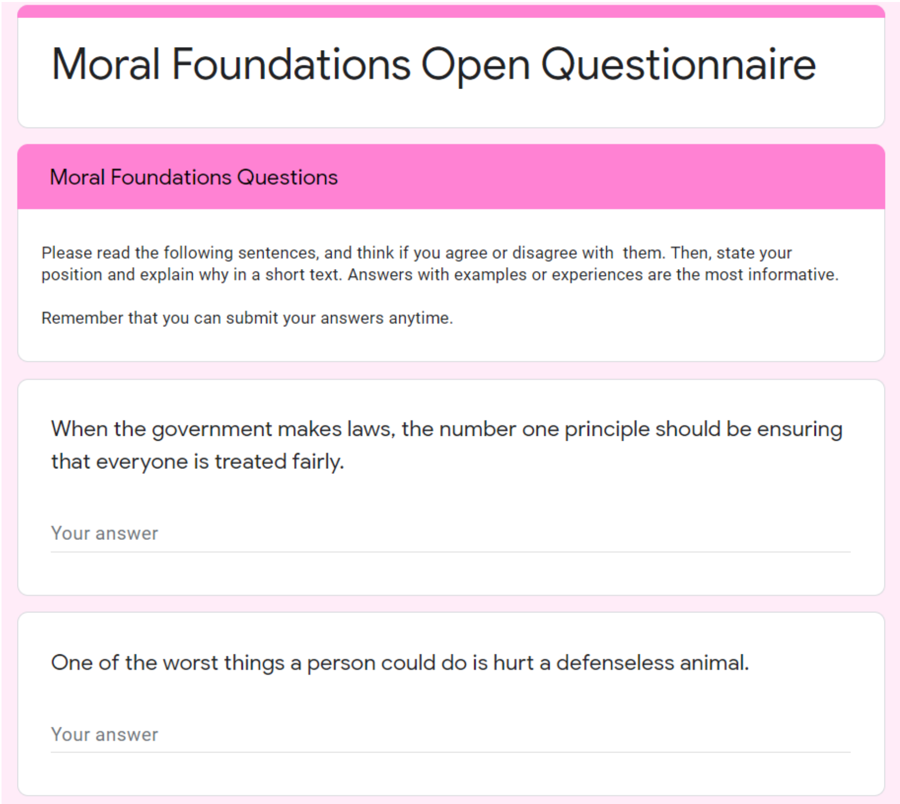
Please read the following sentences. On a scale from 1 to 5, please select the number that most represent you, being 1, disagree entirely, and 5 agree entirely.

We changed them to:

Please read the following sentences, and think if you agree or disagree with them. Then, state your position and explain why in a short text. Answers with examples or experiences are the most informative.

Having 30 open-ended questions makes the questionnaire more tedious. Therefore, we are allowing the users to exit the user to submit whenever they want. In the end, we only need a sample text. Filling all the questions is not necessary. Fixed order of the questions may induce bias; therefore, the questions are sorted randomly.

Figure 14: Modified MFQ



Then, the user runs the program showing a result like this one:

```
Compatibility with Ned Price: 0.458
Compatibility with Chuck Schumer: 0.497
Compatibility with Kevin McCarthy: 0.451
Compatibility with Barack Obama: 0.560
Compatibility with Gabriel Sterling: 0.532
Compatibility with Rudy Giuliani: 0.477
Compatibility with Eric Nelson: 0.417
Most compatible with Lindsey Graham, with 0.624 of compatibility
<matplotlib.lines.Line2D at 0x7fe1faf53410>
```

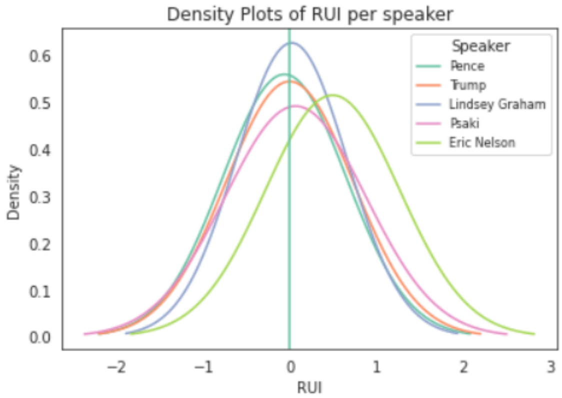


Figure 15: Example of how the result would be displayed

8 Effort

From the total effort delved into the development of this project (approx. 54 hours from each one of us; covering from formulating the project proposal to the writing of this final report), we estimate the following distribution for each component:

1. **Project formulation**, finding both an exciting and relevant question (it helped that both our countries, Mexico and Peru, face electoral processes this year) and discussing feasibility and scope of the project took around roughly **4%** of the total effort.

2. **Data collection**: From finding how, and from where, to gather the data for our original idea (tweets from Mexican political figures) and subsequent development of our Selenium web-scraper to download all our data from the Rev website, we estimate an approx. **12%** of total effort was put into this task, particularly on making sure that our crawler worked correctly and seamlessly.

3. **Data cleaning**, deciding the best structure of our data; experimenting with different intervention word counts and analyzing the structures of them; dealing with slight differences in some speaker names and compound words in the MFD took us around **10%** of the aggregate effort

4. **Mid-quarter presentation**, elaborating the slides, discussing a minimum viable product to show, and making sure everything worked well for our software demo **4%** of the effort

5. **NetworkX representation and node2vec model**, generating the necessary data network structure to work, figuring out the network visualizations, trying different node2vec specifications, trying parallelization in Andrei's computer, and general debugging took around **25%** of the effort.

6. **Neural Network definition and results**, deciding the correct structure of the NN as well as an adequate number of layers, as well as the nonlinearities to use; further data cleaning to partition our data and balance it for the four relevant foundations and correctly generating DataLoader objects; parameter tuning; PCA analysis and trying to find out why initially our PCA was terrible (that is where we found that including the sanctity foundation as a category, even though we were not using it was messing our results); generating distributions conditioned on HHI values took about **35%** of the effort

7. **Designing the Moral Foundations Questionnaire**, testing its functionality and analyzing the predicted results from our model took **4%**

8. **Final presentation and report**, updating slides, design structure, and rehearsals, this report writing around **6%** of the effort

   To successfully develop this project, we had to learn the networkX and node2vec functionality and necessary code of both packages. From the course homework, we already had learned how to implement a neural network using PyTorch and were familiar with it; from other courses, we already knew the usage of pandas, scikit-learn (for PCA analysis), seaborn, and matplotlib visualizations. Particularly, Andrei was who beforehand knew how to use Dask to parallelize our data loading/cleaning process, as well as Selenium; both of them were explained to Oscar, who had no previous experience with those packages, in order to know how to apply them in the context of this project.

   An essential factor of this project was that both of us were familiar with Enke's work from a previous Behavioral Economics course, so it was relatively easy to develop it as a basis for our work. Therefore, besides further reading about MFT and recent developments (such as the MFD 2.0), the vast majority of our time was devoted to coding tasks.

8.1 Final Task Distribution

- **Andrei Bartra**: Data extraction (Selenium web-scraper); Filtering; Network representation; Generating RUI distributions per speaker; User interface (open version of MFQ)
- **Oscar Noriega**: Data processing and speaker name adjustments; Moral Foundations Dictionary inclusion; Moral values prediction for words outside of the MFD

References

Benjamin Enke. Moral values and voting. *Journal of Political Economy*, 128(10):3679–3729, 2020.

Justin Garten, Reihane Boghrati, Joe Hoover, Kate M Johnson, and Morteza Dehghani. Morality between the lines: Detecting moral sentiment in text. In *Proceedings of IJCAI 2016 workshop on Computational Modeling of Attitudes*, 2016.

Jesse Graham, Jonathan Haidt, Matt Motyl, Peter Meindl, Carol Iskiwitch, and Marlon Mooijman. Moral foundations theory. *Atlas of moral psychology*, pages 211–22, 2018.

Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864, 2016.

Jonathan Haidt. The new synthesis in moral psychology. *science*, 316(5827):998–1002, 2007.

Jonathan Haidt. *The righteous mind: Why good people are divided by politics and religion.* Vintage, 2012.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.