Relations between UN Speeches, International Trade and World Happiness

Alina Baciu, Sion Coolwijk, Zsofia Katona, and Anya Nikiforova

University of Amsterdam
Fundamentals of Data Science: Assignment 1
Group A5

Abstract. Advancements in machine learning can aid deepen our understanding of complex political and social issues by allowing us to uncover global trends given the wealth of available information. This paper aims to aid in that task by focusing on identifying common emotions in UN General Debate Speeches and determining whether trading activity or emotions in speeches is a better predictor of a country's happiness levels. Exploratory sentiment analysis of UN General Debate Speeches was conducted and the emotion features were used to built predictive decision tree classifiers. We found that while both the volume of trade and emotions can be used as predictors of a country's general level of happiness, the trading activity is a more promising indicator across all evaluation metrics used, including demonstrating consistently higher scores in accuracy, precision, and recall.

Keywords: Sentiment Analysis · Decision Tree Classifier · Happiness Metrics

1 Introduction

1.1 Speeches, Trade and Happiness

There is a wealth of data available on political speeches from around the world which the United Nations (UN) diligently maintains and archives [5]. These speeches offer insights into the viewpoints of governments on a wide range of topics, yet they are usually ignored in the study of international politics [3]. Though delivered by real human beings, these speeches are at risk of becoming mere historical texts. Fortunately, advancements in machine learning offer unique opportunities to delve deep into these speeches, such as detecting common emotions in large amounts of text, which may not be apparent to human readers [1]. Whether the emotions identified though ML techniques can be used as predictor for happiness is the focus of our first inquiry. Additionally, the country's economic standing could also be a reliable predictor of happiness. This research aims to address the following research questions:

- What are the most common emotions in political speeches per regions in the world, and how does this evolve over time?
- Is the trading activity of a country or the emotions in its UNGD speech a better predictor of the happiness levels in the country?

1.2 Approach and Relevance

In this paper, we explore how the emotions in political speeches vary across different regions of the world over time. To achieve this, emotions expressed in speeches from the UN General Debate Corpus (UNGDC) dataset [5] are employed as predictive indicators for the country's happiness levels. Research is conducted into the relationship between emotional factors and the happiness of a country's population, as well as how happiness levels are influenced by trading activity. Ultimately, our objective is to identify the most reliable predictor of happiness between these two key features. We aim to provide insights for informed decision-making and policy recommendations that can aid speechwriters, inform political and economic decision-makers, with respect to the overall well-being of citizens and communities.

1.3 Roadmap

The report will systematically guide you through the research process, which includes data understanding, pre-processing, modeling, evaluation, and deployment. First, the methodology section will provide insight into our data exploration and preparation steps, followed by an explanation of our modeling and feature extraction techniques. Then the results section will provide insights into our findings and data visualizations. The discussion will include a critical analysis of our work, providing insight into its limitations, and proposing steps for future research. Finally, the conclusion will summarize our key findings and provide a brief overview of the paper.

2 Methodology

2.1 CRISP-DM

To begin the research process and address our research questions, it was decided to adopt this CRISP-DM framework [7]. This approach involves a structured sequence of steps, allowing us to progress methodically while remaining flexible to revisit previous stages as needed. This consists of the following steps: understanding needs, data understanding, data pre-processing, data exploration, modeling, evaluation, deployment.

2.2 Materials and Data Pre-processing Steps

Several datasets were utilized to conduct our research. The first dataset, known as the United Nations General Debate Corpus 1970-2022 [5], served as a one of the key components of our analysis. In addition, we incorporated data on international imports and exports from various countries obtained from the Correlates of War Project's Trade Data Set, which covers data up until 2014 [2]. Lastly, we integrated a dataset focused on the latest data on happiness of populations, known as the World Happiness Report 2023 [4]. Specifically, we referred to Chapter 2 titled *World happiness, trust, and social connections in times of crisis*[1]. Our final pre-processed dataset resulted from merging these three data sets based on the year and country, culminating in a comprehensive data set encompassing 131 countries and data spanning from 2005 to 2014.

For the technical aspects of the CRISP-DM process, we relied on the Python programming language. Additionally, GitHub was used to create and manage a repository, and Git enabled collaborative work and tracking version history in our repository.

Feature Engineering As part of our pre-processing, the export and import variables were normalised in respect to GPD per capita, which is a feature from the World Happiness data set. The export and import features of the model were initially defined as the total imports/exports of the respective country in millions of current US dollars. The GPD per capita feature was originally provided in a logarithmic form and was transformed back into monetary value. The rescaling of these variables allowed us to ensure ease of comparison and interpretation.

In order to measure the emotional affect of the UNGD speeches, the NRCLex model was applied [2]. This lexicon lists around 27,000 English words and connects them to emotions and sentiments. For our data, we removed the two sentiments and chose to only keep the emotions for both data exploration and the models (Fig.1).

It is also important to note that we chose to set the target variable to be binary with each country per year being classified as 'happy' or 'not happy' using the median as the boundary.

2.3 EDA

The summary statistics of the pre-processed data set were first examined, as well as the first five rows (see Appendix A, Figure 1 3) of the table. This includes key features related to trading activity (import and export), emotional sentiment scores from the speeches, and the year and country. These statistics have been visualized in a table (see Appendix A, Figure 2) displaying descriptive information and statistics on the mean, standard deviation, quantiles, and other key measurements.

Next, Figure 1 represents a correlation matrix, where blue shades indicate positive correlations, and red shades indicate negative correlations. This visualization allows us to identify which pairs of variables are strongly correlated and may impact our research questions.

To further explore the relationships between variables, a scatter plot matrix was created (Figure 2). This matrix provides a visual overview of how all the continuous features relate to each other.

[1] Link to Chapter 2: https://happiness-report.s3.amazonaws.com/2023/WHR+23_Statistical_Appendix.pdf

[2] https://pypi.org/project/NRCLex/

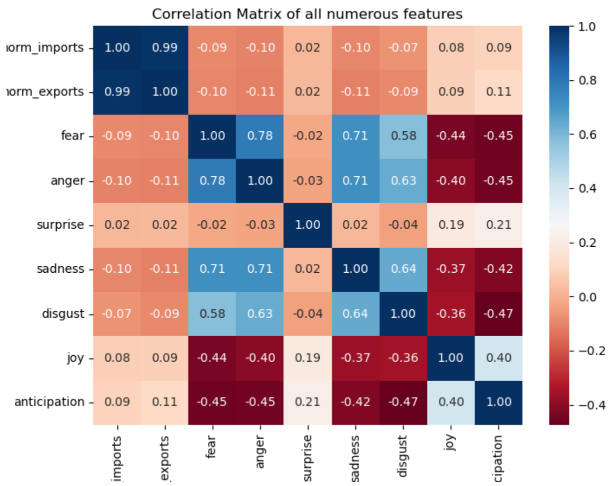

Fig. 1. This heatmap displays correlations between dataset features. Blue shades indicate positive correlations, while red shades indicate negative correlations

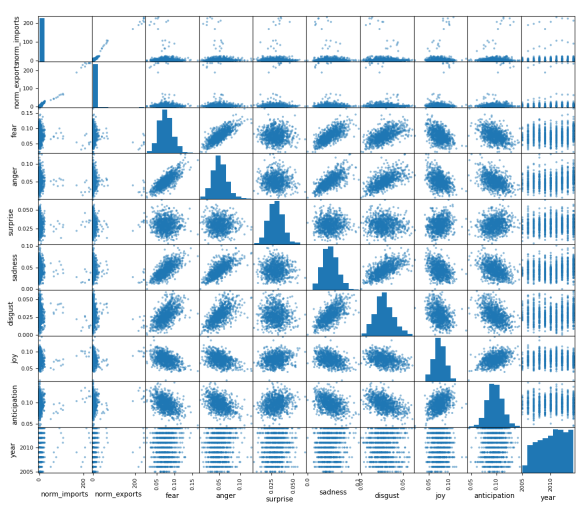

Fig. 2. This matrix of scatter plots provides a visual overview of the relationships between variables in the dataset.

2.4 Research Method

The machine learning models we opted for in the prediction tasks are decision trees, which are known for their relatively high interpretability [6]. As mentioned before, we constructed two models, both predicting the binary happiness class of a country: one with predictors being the emotion scores in the country's UNGD speech in the corresponding year and one with predictors the normalized trading volumes in the same year. The models were trained using the GINI Index, which measures the probability of a randomly chosen instance to be misclassified and is minimized during the training process. Importantly, our data set was randomly split into a training and test set (80% and 20%, respectively), so that we can ensure that the models' performance is measured on unseen data in the evaluation phase.

In the training phase, we conducted 10-fold cross-validation to select the best hyperparameter settings for our models. The following hyperparameters were selected in this manner: the tree's maximum depth and the minimum number of samples in each leaf (both out of values 1, 3, 5, 7 and 10) as well as the minimum number of samples at each split in the tree (out of values 5, 10, 15, 20, 50). Importantly, we searched for the best *combination* of these settings, so that we opted for a grid search setup in our cross-validation process, where the performance of the models on the training set was evaluated with each combination based on accuracy.

Once the best hyperparameter settings were found for both models, we trained the models on the complete training dataset with these settings. The trained models were evaluated on the testing set based on five numerical metrics (accuracy, precision, recall, F1 score, and AUC) as well as the shape of their ROC curve.

3 Results

3.1 Emotions Over Time

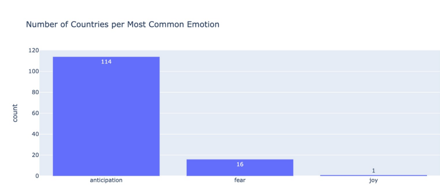

Fig. 3. Most common emotions in political speeches

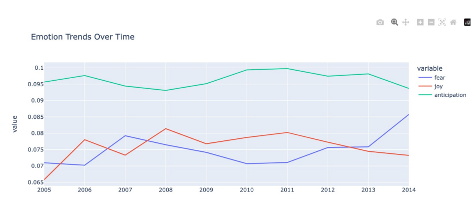

Fig. 4. Most common emotions over time

The most common emotions in political speeches per country were anticipation, fear and joy. Figure 4 depicts that the anticipation emotion indicated small variation throughout the years, while joy and fear showed notable changes, with joy dominating for a period, being later surpassed by fear.

3.2 Machine Learning Model Performance

The best hyperparameter settings found for both models during cross-validation are summarized in Table 1 below.

Table 1. The table displays the best hyperparameter settings found during cross-validation.

Hyperparameter	UNGD Emotions model	Trading Activity model
Minimum depth	3	10
Minimum samples at a split	5	15
Minimum samples in a leaf	7	3

The performance of of the two models obtained on the testing set, can be seen in Table 2 below, while Figure 5 displays the ROC curves produced on the same testing set. As can be seen, the decision tree trained on trading activities outperformed the UNGD Emotions model by a wide margin, which is clear based on the numerical metrics (17-38 percentage points difference in the various measures, with the Trading Activity model performing better in each one) as well as the shape of the ROC curves. In this latter aspect, the UNGD Emotions model showed only slightly better performance (AUC=0.58) than a classifier which selects a random class with equal probability (AUC=0.5).

Table 2. The table demonstrates the results obtained on the testing set for both classification models.

Performance metric	UNGD Emotions model	Trading Activity model
Accuracy	0.58	0.78
Precision	0.58	0.75
Recall	0.41	0.79
F1-score	0.48	0.77
AUC	0.58	0.79

4 Discussion

Key Findings and their Interpretation The most common emotion across the years turned out to be anticipation. Given that the speakers are given the choice of topic as well as a way to make their voices heard, it makes sense why

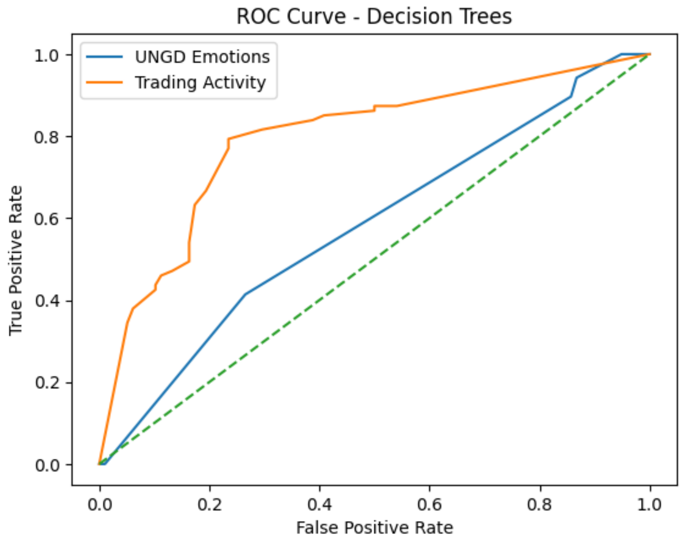

Fig. 5. The figure shows the ROC curves obtained on the testing set for both models. The dashed green line demonstrates the performance of a hypothetical "gambling" classifier, which chooses one class or the other with 50% probability.

anticipation could be the leading emotion. In regards to fear and joy, their slight fluctuations can be attributed to the ever-changing political landscape and the well-being of their country at the time the speech is given.

The Trading Activity model showed consistently higher performance in all chosen metrics as compared to the UNGD Emotions model (Table 2). The UNGD Emotions model produced a smaller, easily comprehensible tree with, while the Trading Activity model resulted in a much larger tree. The fact that the economic activity model was a better predictive model can be explained by the fact that economic factors are a much more representative measure of the well-being of a nation and their standard of life as compared to a emotions contained in once speech.

4.1 Practical Applications and Limitations

The way that the economic and political life of a country affects its citizens' well-being is a complex and multidimensional topic. Before training the models, we had to simplify our knowledge to fit a specific relationship. It must be noted that in reality, the attitude of a government does not necessarily reflect the mood of its people and the volume of trade does not necessarily reflect the financial well-being of a country.

We hope that this research can be used in combination with other findings and have a variety of implications and can allow professionals in different industries from public policy to medical industries make more informed decisions on how to improve the well-being of their clients and citizens. Even though this model can predict the general happiness level of a country, it cannot be applied to an individual level without further consideration.

4.2 Future Work

Multiple things can be looked at to build on top of the questions we aimed to answer and the techniques that we used. In particular, different feature engineering techniques and more complex classifiers can be applied to answer the more intricate insight of the general research question.

The research question can be built on top of the World Happiness Report by adding other metrics and measure of happiness to provide a more complete picture in both an explanatory and predictive research. In terms of the techniques used, while the NRCLex is a well-established resource in sentiment analysis, there are many more tools and resources that have become available. Comparing how the features would look like using other models and how that would compare to the model used in this paper is one way that this paper can be built on. Furthermore, other predictive algorithms that are built on decision trees, such as an ensemble methods like a random forest algorithm or a gradient boosting algorithm could be trained to see if they enhance predictive performance.

5 Conclusion

Using data available from the UNGD speeches, the World Happiness Report 2023 and the War Project's Trade Data Set, our paper aimed to contribute to the research of the relationship between happiness with the financial and political state of a country. Our paper had two focus points. Looking at the most common emotions in political speeches per regions in the world, we found anticipation, fear and joy to be the most common emotions. Despite some small variation, the emotions manifested in UN speeches are consistently stable across the years. We noted an increase in fear in from 2013 and 2014 and a slight decrease in joy during that same time period (Fig. 4). The second part of this paper looked at whether the trading activity of a country or the emotions in UNGD speeches was a better predictor of the happiness levels in the respective the country. We found the volume of economic activity to be a better indicator of a country's general happiness level.

References

1. Alm, C.O., Roth, D., Sproat, R.: Emotions from text: machine learning for text-based emotion prediction. In: Proceedings of human language technology conference and conference on empirical methods in natural language processing. pp. 579–586 (2005)
2. Barbieri, K., Keshk, O.M.G.O.: Correlates of war project trade data set codebook (2016), https://correlatesofwar.org.
3. Baturo, A., Dasandi, N., Mikhaylov, S.J.: Understanding state preferences with text as data: Introducing the un general debate corpus. Research & Politics **4**(2), 2053168017712821 (2017). https://doi.org/10.1177/2053168017712821
4. Helliwell, J.F., Layard, R., Sachs, J.D., Aknin, L.B., De Neve, J.E., Wang, S., (Eds.): World happiness report 2023 (11th ed.). Sustainable Development Solutions Network https://worldhappiness.report/ed/2023/
5. Jankin, S., Baturo, A., Dasandi, N.: United Nations General Debate Corpus 1946-2022 (2017). https://doi.org/10.7910/DVN/0TJX8Y, https://doi.org/10.7910/DVN/0TJX8Y
6. Kingsford, C., Salzberg, S.L.: What are decision trees? Nature biotechnology **26**(9), 1011–1013 (2008)
7. Schröer, C., Kruse, F., Gómez, J.M.: A systematic literature review on applying crisp-dm process model. Procedia Computer Science **181**, 526–534 (2021)

6 Appendix A

	happiness_classification	norm_imports	norm_exports	fear	anger	surprise	sadness	disgust	joy	anticipation	year	country
0	not-happy	13.643921	11.979343	0.044010	0.029340	0.034230	0.031785	0.026895	0.056235	0.100244	2005	Mexico
1	not-happy	13.469564	15.552570	0.060606	0.028520	0.033868	0.035651	0.024955	0.057041	0.112299	2005	Japan
2	happy	6.874739	7.213294	0.072381	0.049524	0.040000	0.038095	0.034286	0.097143	0.091429	2005	Belgium
3	not-happy	6.624081	4.183923	0.059840	0.043883	0.019947	0.041223	0.031915	0.070479	0.093085	2005	Pakistan
4	happy	12.099429	11.116681	0.070225	0.043539	0.036517	0.036517	0.028090	0.081461	0.102528	2005	France

Fig. 6. This figure presents a subset of the final dataset, showcasing the first five

	norm_imports	norm_exports	fear	anger	surprise	sadness	disgust	joy	anticipation	year
count	921.000000	921.000000	921.000000	921.000000	921.000000	921.000000	921.000000	921.000000	921.000000	921.000000
mean	6.456758	5.973886	0.075533	0.051372	0.029842	0.044563	0.027387	0.076754	0.096618	2010.248643
std	21.974137	23.279404	0.019541	0.017523	0.008499	0.015038	0.010827	0.014812	0.014985	2.612691
min	0.027052	0.009347	0.021505	0.005025	0.004808	0.000000	0.000000	0.015038	0.045113	2005.000000
25%	0.912420	0.489041	0.062077	0.039352	0.024217	0.033898	0.019980	0.067293	0.086667	2008.000000
50%	1.978003	1.436391	0.074246	0.049784	0.029661	0.043554	0.026506	0.076726	0.096386	2011.000000
75%	4.269860	3.763795	0.087719	0.061644	0.035262	0.054902	0.034137	0.085995	0.106713	2012.000000
max	230.704658	241.990881	0.165254	0.126783	0.063679	0.100962	0.061275	0.143275	0.146186	2014.000000

Fig. 7. This figure displays a summary of descriptive statistics for the dataset.