# EVALUATION OF THE EMOTIONALITY OF PUBLIC DISCOURSE ON THE WAR BETWEEN RUSSIA AND UKRAINE USING AN EMOTIONAL CLASSIFIER

Finn Fassbender, Frieder Wizgall, Georg Tirp, Nina Lutz & Simon Krülle

In this project we classified tweets and news articles on the war between Ukraine and Russia according to emotions. Looking at the timelines of different emotions, we traced back certain ones that stood out to major war events. We also compared differences in emotions between different media forms. While this is only correlational data, it might still provide some interesting insight into how the war is perceived in the public eye, specifically on social media and in major news outlets.

## 1. Datasets

In order to train our classifier we used a training set of English Twitter messages with six basic emotions: anger, fear, joy, love, sadness, and surprise (Saravia et al., 2018).

As our main dataset we used a set of 454 Million tweets on the war between Ukraine and Russia (Chen & Ferrara, 2022), that span from the start of the war till the first of October and are updated regularly. The tweets are hereby represented as tweet ID's, which can be transcribed to JSON documents containing the raw text and metadata. We used a GUI program called Hydrator that works on the Twitter API to transcribe the ID's.

In order for us to handle the classification on our computers we extracted 5000 tweets per day from the dataset to get a more appropriate sized set of about 1 Million tweets.

There was a loss of about 19% due to deleted tweets. We also used the language labels from the tweet metadata to only select english tweets, which make up 70% of the dataset. Furthermore we removed any replies or retweets, because we want to rate the emotion of original opinions and not distort our results with reactions to other tweets. Lastly we replaced all usernames, hashtags and links with empty space and sorted out all tweets with less than three words. In conclusion our prepared dataset makes up about 55% of the remaining tweets or about 2227 tweets per day.

Future improvements to our approach could work with a much bigger sample of tweets, since we only used a small subset. This would require a lot more computational power but could make the emotion classification way more accurate. We could also train our classifier on multiple languages and explore the difference in emotion between tweets of different languages.

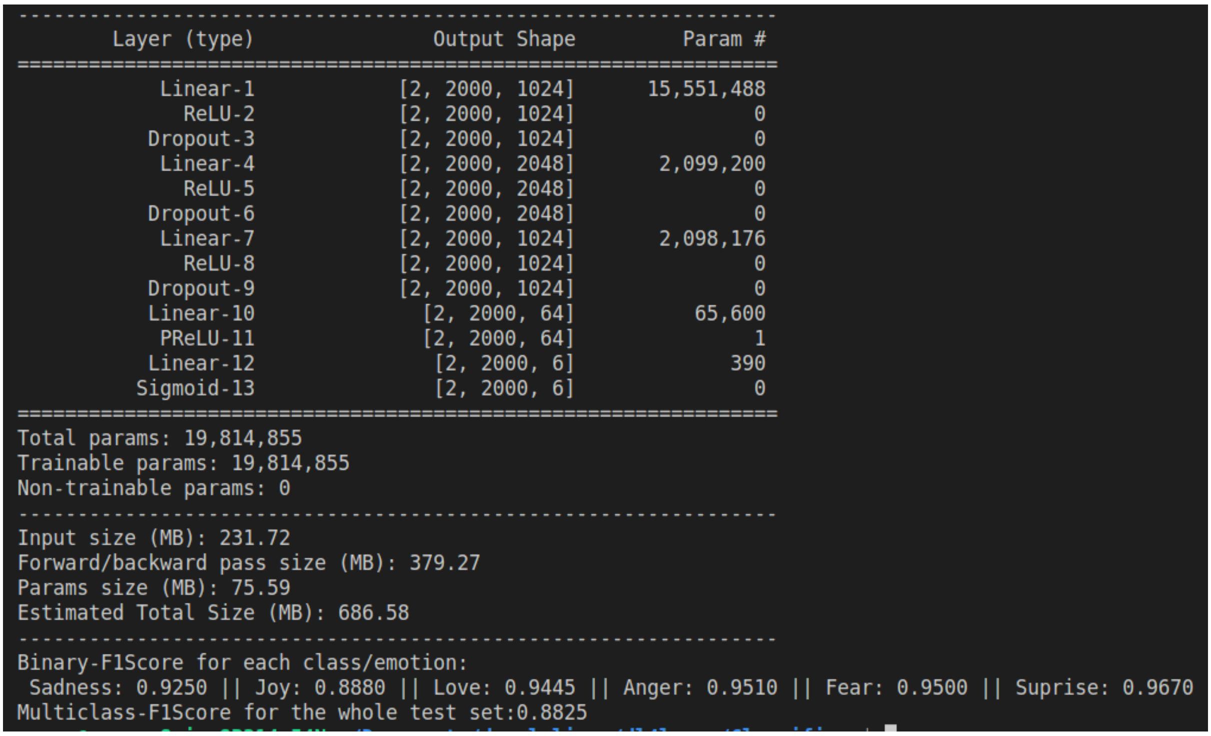
To compare tweets with articles we used a dataset containing 407 news articles from NYT and Guardians related to ongoing conflict between Russia and Ukraine (Khawaja 2022).

## 2. Creating the Emotional Classifier

While working on the emotional classifier, we tried different things before we reached a classifier that was fit for the task.

### a) First Try: Linear Model

Our first idea was to create a linear neural network with a pytorch library. We wanted to make a dictionary with training data where words are encoded as one-hot vectors and every sample gets encoded so that the data would be a 2D matrix. Because of the many features and the large data set, the data is saved in a sparse format to not exceed the working memory. Then, that sparse-data is feeded into the linear NN in batches. A picture of the model summary can be seen below. We used CrossEntropy as the loss function and Adam as the optimizer. For the regularization, Dropoutlayers (details in the model summary) and a weight-decay of 0.0001 were used in the optimizer parameters. F1 scores are also stated in the summary.



```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Linear-1          [2, 2000, 1024]      15,551,488
              ReLU-2          [2, 2000, 1024]               0
           Dropout-3          [2, 2000, 1024]               0
            Linear-4          [2, 2000, 2048]       2,099,200
              ReLU-5          [2, 2000, 2048]               0
           Dropout-6          [2, 2000, 2048]               0
            Linear-7          [2, 2000, 1024]       2,098,176
              ReLU-8          [2, 2000, 1024]               0
           Dropout-9          [2, 2000, 1024]               0
           Linear-10            [2, 2000, 64]          65,600
            PReLU-11            [2, 2000, 64]               1
           Linear-12             [2, 2000, 6]             390
          Sigmoid-13             [2, 2000, 6]               0
================================================================
Total params: 19,814,855
Trainable params: 19,814,855
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 231.72
Forward/backward pass size (MB): 379.27
Params size (MB): 75.59
Estimated Total Size (MB): 686.58
----------------------------------------------------------------
Binary-F1Score for each class/emotion:
 Sadness: 0.9250 || Joy: 0.8880 || Love: 0.9445 || Anger: 0.9510 || Fear: 0.9500 || Suprise: 0.9670
Multiclass-F1Score for the whole test set:0.8825
```

FIGURE 1: MODEL SUMMARY (LINEAR MODEL)

Because of the many parameters, a lengthy training time and high-performing hardware were needed. A further flaw of that first model was that context was not taken into account as it just counts words and looks for keywords. Sentences like “I am not angry” would be labeled as ‘angry’. While the F1 score is acceptable for rather simple training data, the performance is pretty bad for ambiguous short texts.

### b) Second Try: LSTM Neural Network

To overcome the context problem, we wanted to amend our model. After some research on emotion classification, we wanted to use LSTM layers implemented with Pytorch. We again went for a dictionary made by training data where words are encoded as one-hot vectors.

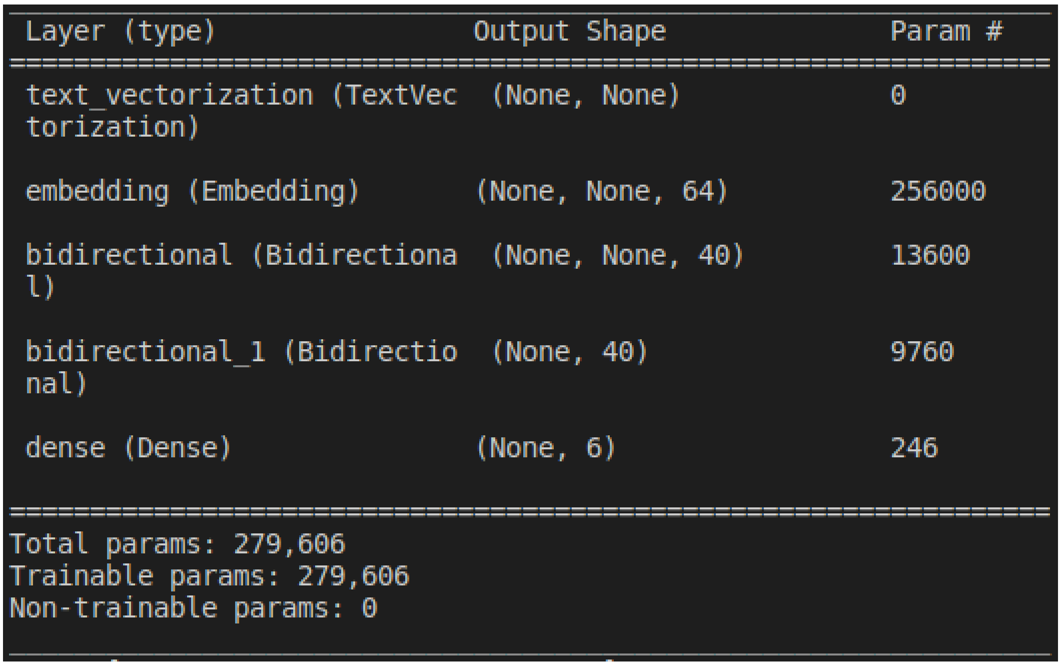
This time however, every sentence gets sequenced, so that every word gets encoded to a vector, producing a 3D data matrix. Due to the same reasons as before, data had to be saved in a sparse format (dim(1) dense and dim(2) sparse). The sparse-data was fed into the LSTM model in batches. The loss function CrossEntropy and the Adam optimizer were selected, same as before.

Due to performance issues the trial had to be canceled. We therefore have no performance data available.

Our missing expertise on RNN/LSTM implementations and on network architectures, that would be compatible with the hardware accessible to us, raised some troubles. Without the sparse format, the training data (120 GB) is way too large, but the pytorch sparse format is in beta and cannot be fed directly into the LSTM layers. Consequently, every batch of data has to be converted back to dense matrices in the Dataloader, which causes significant performance issues. That combined with our hardware limits increased the training time to a point where we decided to cancel this trial.

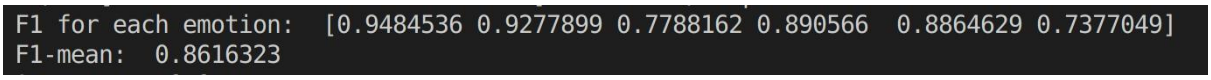
### c) Third try: LSTM model with tensorflow

To overcome the issues described above, we used the same pytorch version, but this time with a tensorflow library. We proceeded as follows: The most common 4000 words of the tweets are put into a dictionary and are assigned numbers. Then, the tweets are fed in batches to an embedding layer of the RNN which encodes each word of the tweets into one hot encoding according to the dictionary. Next, two LSTM layers with a dropout of 0.4 and an L2 regularization of 0.01 are trained, concluding in an output of the 6 different emotions. As before, CrossEntropy was chosen for the loss function and Adam for the optimizer. Regularization was achieved with Dropoutlayers (see model summary for details) and L2-norm punishment with a lambda value of 0.01.



```
 Layer (type)                Output Shape              Param #
=================================================================
 text_vectorization (TextVec  (None, None)              0
 torization)

 embedding (Embedding)       (None, None, 64)          256000

 bidirectional (Bidirectiona  (None, None, 40)          13600
 l)

 bidirectional_1 (Bidirectio  (None, 40)                9760
 nal)

 dense (Dense)               (None, 6)                 246

=================================================================
Total params: 279,606
Trainable params: 279,606
Non-trainable params: 0
```

FIGURE 2: MODEL SUMMARY (LSTM MODEL WITH TENSORFLOW)



```
F1 for each emotion:  [0.9484536 0.9277899 0.7788162 0.890566  0.8864629 0.7377049]
F1-mean:  0.8616323
```

FIGURE 3: F1 VALUES (LSTM MODEL WITH TENSORFLOW)

In this trial, we struggled with lacking theoretical knowledge concerning NN architecture and hyperparameter tuning. Another challenge was to get over an F1 score of ~0.86. Also, the runtime for evaluating the tweets is still lengthy with approximately seven hours for one million tweets.

### 3. Results & Interpretation

First, we applied our emotional classifier on a dataset of tweets (Chen & Ferrara, 2022) tracking the twitter discourse on the war. We observed different timelines for sadness, joy, love, anger, fear, and surprise. This can be seen in figure 1.

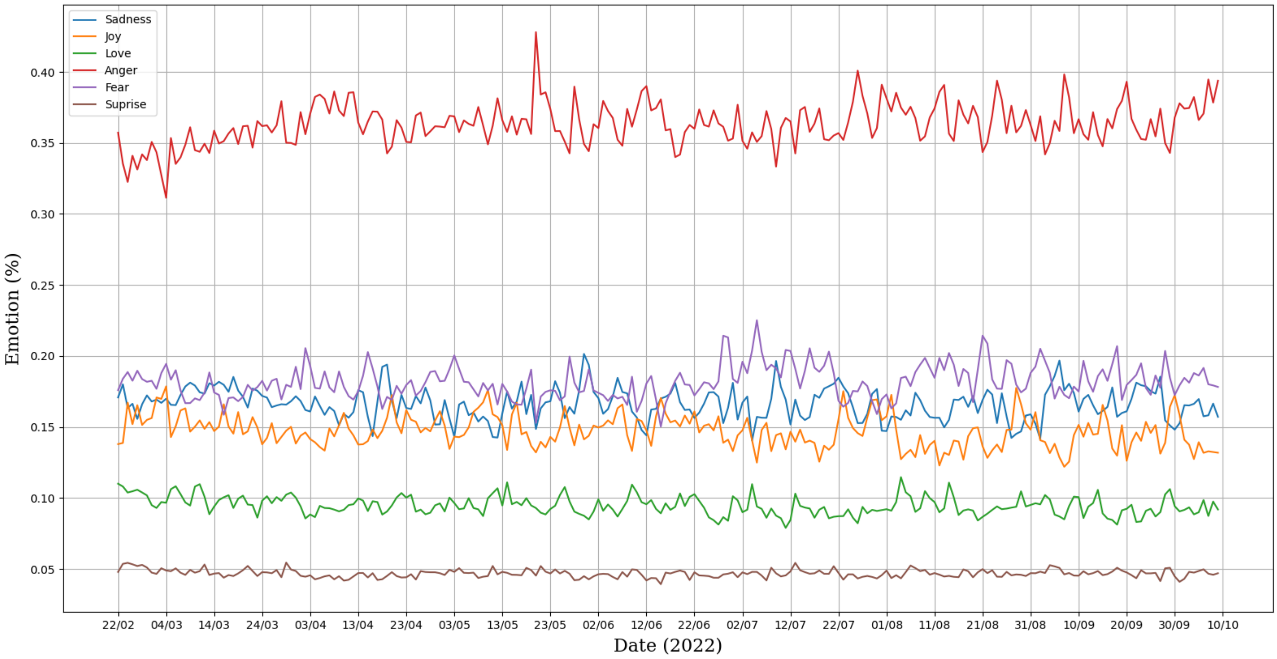


FIGURE 4: TIMELINES FOR SADNESS, JOY, LOVE, ANGER, FEAR, SURPRISE IN TWEETS ON THE WAR

Without trying to ascribe any causal relationship between war events and emotional peaks and valleys, it might still be interesting to mention some correlations we observed.

Surrounding March 4th, a decrease of anger and an increase in joy can be observed. At this time, Russia failed to conquer Kiev, which led to a lengthened line of tanks outside of Kiev. Ukraine started their first counteroffensive concerning Kiev. This was the first time numerous weapons were announced and delivered regularly, which foreshadowed a lengthy war. That could very well have been the cause for an increase in fear, as can be seen in the graph.

After that, anger levels were rising. It was during those days the brutality of the offensive became clear, and the Siege of Mariupol took place, where about 22,000 civilians were killed.

Anger levels dropped again surrounding March 27th. At that time it became clear that Russia wouldn't capture Kiev, but despite a slow fallback, Kiev was still contested. Ukraine started several smaller counter offensives, and on March 29th negotiations between Ukraine and Russia took place in Istanbul.

In the days surrounding April 3rd, anger peaked. A related event might be the discovery of bodies from the Bucha Massacre.

The following drop in anger on April 13th correlates with the sinking of the Moskwa.

Another anger peak on May 17th coincides with the capitulation of troops in the Azovstal iron and steel works, who were taken into war captivity by the Russians.

Furthermore, we observed varying mean percentages for the emotions mentioned above, as can be seen in figure 2. Anger is distinctly the prevalent emotion, appearing in more than 35 percent of all tweets. The second most frequent emotion is fear, closely followed by sadness. Both can be found in just below 20 percent of all tweets. All of these are not surprising, as the conflict sparked a lot of anger not only from both parties, but also the rest of the world. It is also self-explanatory that any war raises fear and sadness, mostly in the directly affected populations, but not limited to those.

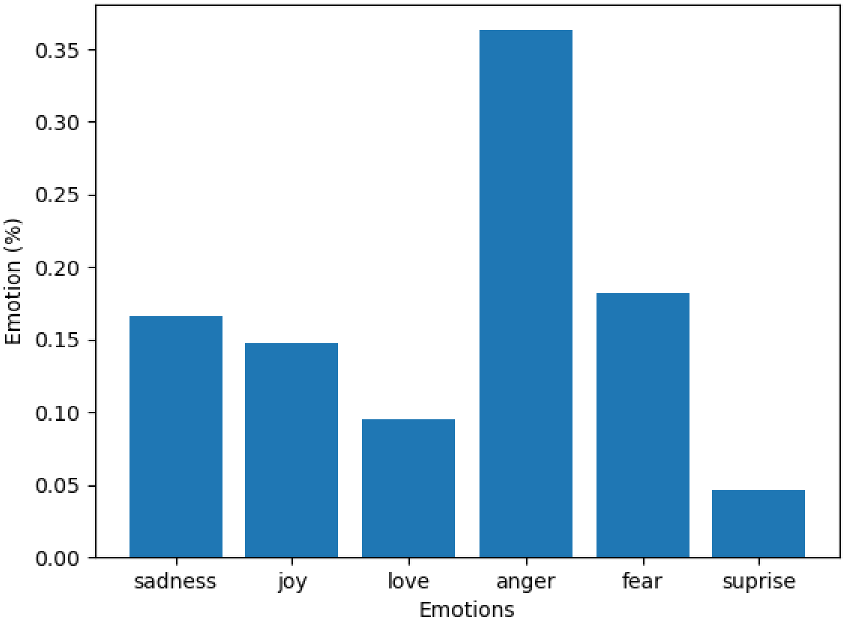


FIGURE 5: MEAN EMOTIONS IN TWEETS ON THE WAR

Secondly, we used the same classifier to evaluate a collection of news articles from NYT and Guardians related to the war (Khawaja, 2022).

Just like in the tweets, anger dominates in the news articles. It does so even stronger, occurring in more than half of all articles. This raises a question on the neutrality of newspapers, which might be a lot more biased than they are perceived. Obviously, the Western population is also biased, so that might be reflected in those articles. In comparison to anger, the other emotions fall far behind, the closest one being fear with just above 20 percent. A high amount of fear in the tweets already suggested high levels of fear in the population, so it makes sense that this is reflected in the news. Interestingly, the percentage of joy in the articles trumped sadness, even if not by far. Both emotions were found in about ten percent of all evaluated articles. Surprise could be determined in about five per cent of them, and the least dominant emotion was love with about one per cent. The fact that love was more visible in the tweets than in the news articles seems logical, as tweets are often more directed towards specific people whereas articles tend to address a larger population. It seems that the more personal the messages, the more there is a want and need to express love. Levels of joy were also higher in tweets than in articles, which can be explained with a similar reasoning, and/or with a bias of negative news.

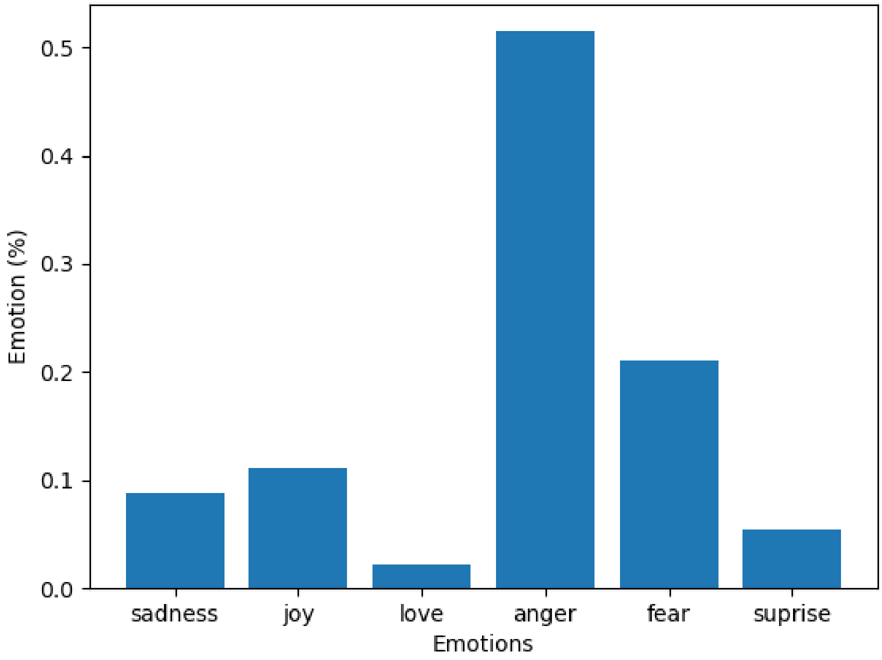


FIGURE 6: MEAN EMOTIONS IN NEWS ARTICLES ON THE WAR

We also created a word cloud to visualize the most frequently used words in Khawaja (2022). Synonyms were merged to the most prominent representative of the word group and interjections, prepositions and pronouns were removed.

This might not be the most scientific method but it can help visualize the datasets. We could further create another word cloud with the tweet data and compare the most prominent words used in both datasets. This might give us another visual clue why, for example, the emotion love is more often represented in the tweet dataset.

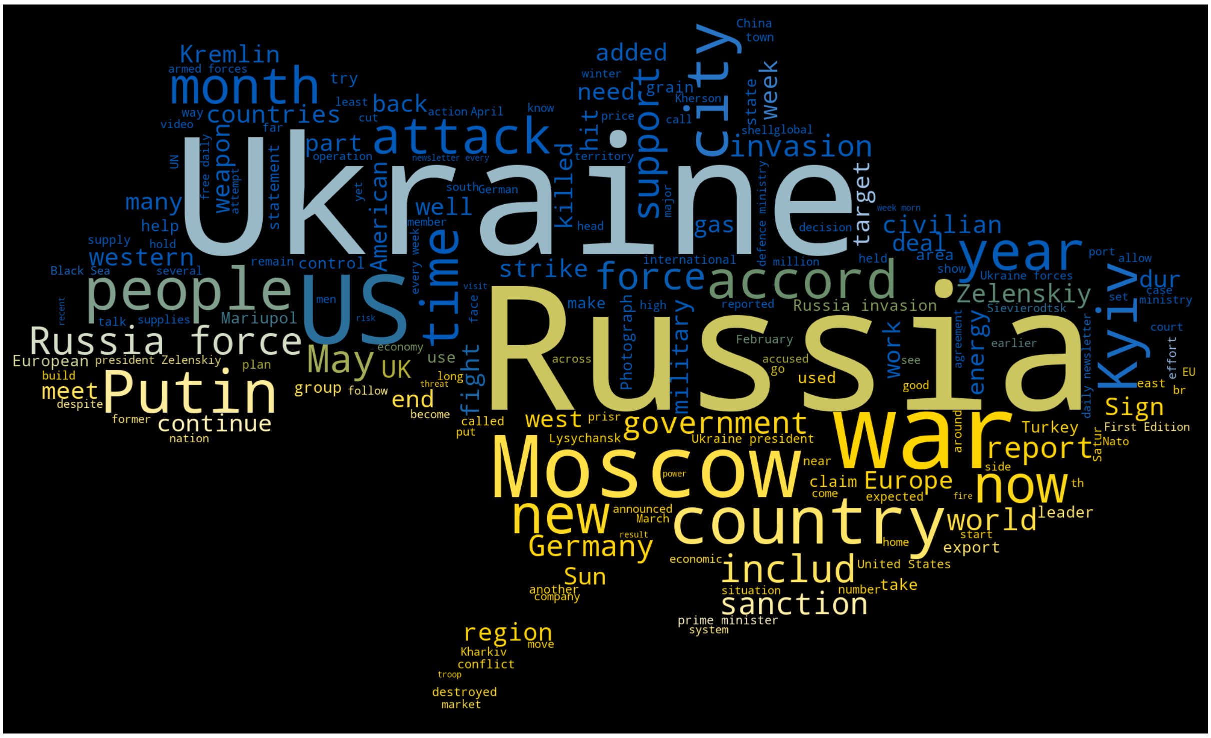


FIGURE 7: WORD CLOUD CREATED FROM NEWS ARTICLES ON THE WAR

All the programm code and the figures can be found at our public github repository (emotion.ruuk).

## SOURCES

Documenting the Now. 2020. Hydrator [Computer Software]. Retrieved from https://github.com/docnow/hydrator

Elvis Saravia, Hsien-Chi Toby Liu, Yen-Hao Huang, Junlin Wu, and Yi-Shin Chen. 2018. CARER: Contextualized Affect Representations for Emotion Recognition. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pages 3687–3697, Brussels, Belgium. Association for Computational Linguistics.

Emily Chen and Emilio Ferrara. 2022. Tweets in Time of Conflict: A Public Dataset Tracking the Twitter Discourse on the War Between Ukraine and Russia. arXiv:cs.SI/2203.07488

Emotion Classification of the Public Discourse on the War between Russia and Ukraine. 2022. https://github.com/dl4l-cog/emotion.ruuk

Hussain Shahbaz Khawaja. 2022. Russia Ukraine Conflict Articles: A Dataset. https://huggingface.co/datasets/hugginglearners/russia-ukraine-conflict-articles

Wikipedia contributors. (2022, October 13). Timeline of the 2022 Russian invasion of Ukraine. Wikipedia. https://en.wikipedia.org/wiki/Timeline_of_the_2022_Russian_invasion_of_Ukraine