



# Vaccine-lies & COVID-19 NLP & ML model

João Carlos Machado Rocha Pires (UP201806079)

João Francisco de Pinho Brandão (UP201705573)

Pedro Miguel Afonso Teixeira (UP201505916)

Machine Learning Complements - 2021-22 - M.EIC



# Agenda

- Context and definition
- Goals
- Summary of the original paper
- Models used in the original paper
- Original paper - CNN Configurations
- Proposal
- Experimental setup
- Results of the proposed variant
- Time series data
- Conclusions and future work



# Context and definition

- Dataset of COVID-19 & Vaccine-lies tweets pre-labeled with the following categories:
  - No Stance
  - Not Relevant
  - Reject
  - Accept
- Tweets identified by id (Twitter API to get tweets' content)
- Paper with an approach applied to another dataset, using NLP and ML
  - “Using Deep Learning to Correlate Reddit Posts with Economic Time Series during the COVID-19 Pandemic” (P2)



# Goals

- Create an NLP/ML model to classify COVID-19 tweets
- Use original paper approach as baseline
- Extensions and improvements to that approach -> better metrics and performance

## Steps:

- Get the tweets content (Twitter API)
- Analyse and prepare the data
- Use different models, with different configurations
- Evaluate the performance of the models



# Summary of the original paper

- Reddit posts about employment in the USA during the COVID-19 pandemic
- Reddit posts manually labeled
- Unemployment time series data
- 6 models to determine the class of each Reddit post
  - Unemployment
  - Cut/Furlough
  - Employment
  - Other
- Comparison of predictions with time series data



## Models used in the original paper

| Model               | Details                                   |
|---------------------|---|
| Random Forest       | GloVe 300, Stopword Removal, Regex        |
| Logistic Regression | GloVe 300, Stopword Removal, Regex        |
| Linear SVM          | GloVe 300, Stopword Removal, Regex        |
| CNN 1               | GloVe 300                                 |
| CNN 2               | GloVe 300, Stopword Removal, Regex        |
| CNN 3               | GloVe 300, Stopword Removal, Regex, VADER |



# Original Paper - CNN Configurations

## Layers:

1. Convolution 1D
2. Max Pool
3. Dense F.C.
4. Dropout
5. Dense F.C.
6. Dropout
7. Softmax Output

## Training:

- Stochastic gradient descent
- Learning rate = 0.0001
- 150 training epochs



# Proposal

## Original paper:

- no personal information collected
- classification on a two-class score and four-class score
- 6 models
- VADER sentiment analysis as CNN layer
- Text input = post title + first 25 words of the post text

## Our approach:

- verified information (tweet's author account verified by Twitter)
- classification on a four-class score
- 8 models (6 original models + Naive Bayes + KNN)
- VADER sentiment analysis at the pre-processing stage
- Text input = whole tweet





# Experimental setup

- Jupyter Notebook (code + annotations)
- Individual configuration of Twitter API access
- Data analysis and preparation
- Model setup and run
- Python libraries used:
  - Pandas
  - Tensorflow/Keras
  - NLTK (Stop-Word removal)
  - Sklearn
  - vaderSentiment (VADER sentiment analysis)



# Results of the proposed variant

| Model               | F1 Score (Macro) | F1 Score (Weighted) |
|---------------------|------------------|---------------------|
| Random Forest       | 19.083           | 20.349              |
| Logistic Regression | 18.060           | 19.744              |
| Linear SVM          | 18.061           | 19.703              |
| Naive Bayes         | 12.667           | 12.502              |
| <b>KNN</b>          | <b>23.257</b>    | <b>24.681</b>       |
| CNN 1               | 17.020           | 20.229              |
| CNN 2               | 17.237           | 20.331              |
| CNN 3               | 16.321           | 18.887              |



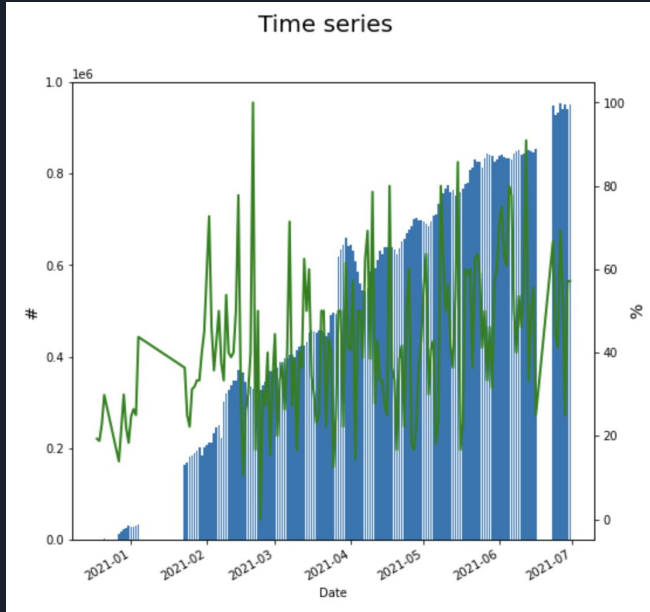
# Time series data

- No time series data about disinformation found online
- **Solution:** World Vaccination progress time series data
  - Data taken from [kaggle](#)

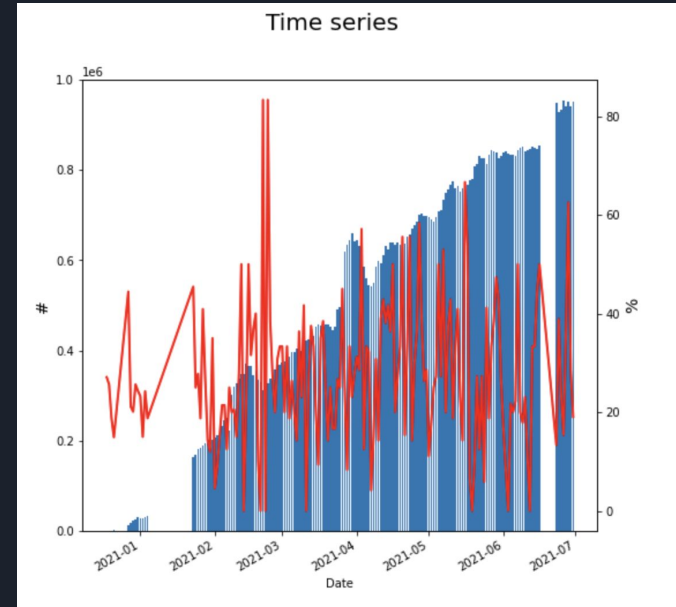
## Steps:

- Find timeframe for tweets data
- Filter time series data using that time interval
- Compare monthly vaccinations with Accept and Reject rates (same timeframe)
- Check if there's a correlation between the information

# Time series data (Results)



Acceptance rate vs World Vaccination progress



Rejection rate vs World Vaccination progress



# Time series data (Result Analysis)

- Peak for both rates in an early stage of the vaccination process
- Lot of variation in the rates in an early stage
- Rejection rate quite close to 0 much more times than the acceptance rate as the vaccination process moves forward
- Too many outer variables to have trustable comparison



# Conclusions and future work

- + Application of some variants to the original approach
- + Fulfillment of original models configurations (GloVe 300, Stop-Word and Regex Removal, VADER)
- + Comparison of our best scoring model predictions with a time series data
  - Low F1 score
  - Time series data not the ideal for our dataset subject

## Future work:

- Improve NLP-related tasks
- Try different CNN configurations/parameters
- Find semantically richer time series data