

# Analisi dell'orientamento politico, del sentiment e delle emotions delle principali testate giornalistiche italiane

Progetto di esame a cura di Alessia Lusci e Nicholas Garau  
Web Analytics e Analisi Testuale - A.A. 2020/2021  
Università degli Studi di Cagliari



# Obbiettivi



Predizione su  
orientamento politico

Sentiment analysis

Emotions analysis



# Scraping

- *Estrazione tweet partiti politici*
  - destra: FDI, Lega, Casapound
  - neutro: Forza Italia, +Europa, Mov5 Stelle
  - sinistra: Rifondazione, PD, LEU
- *Estrazione tweet testate giornalistiche*
  - Repubblica, il Fatto Quotidiano, il Giornale, l'Unione Sarda, il Messaggero, Libero, il Post, il Corriere della Sera, la Stampa

## Tools

- **snscraper**
- **requests**





# Scraping

- *Creazione corpus per orientamento per frequency distribution*
- *Filtraggio notizie per categoria (politica ed economica ✓)*
- *Creazione dataset con assegnazione orientamento (per target in modellizzazione)*

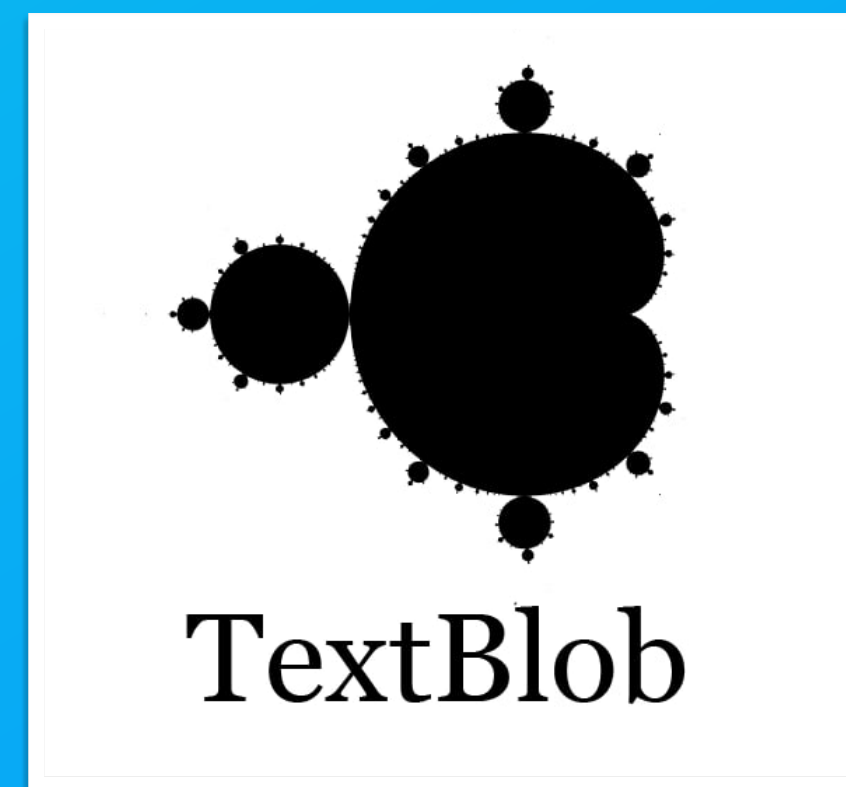
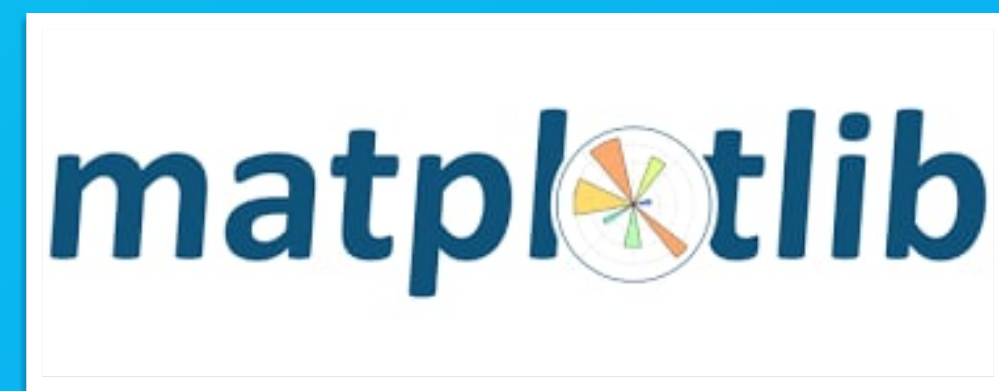
## Tools

- **urllib**
- **requests**
- **NLTK**



# Analisi e modellizzazione: **i tools**🔧

- text2emotion
- NLTK



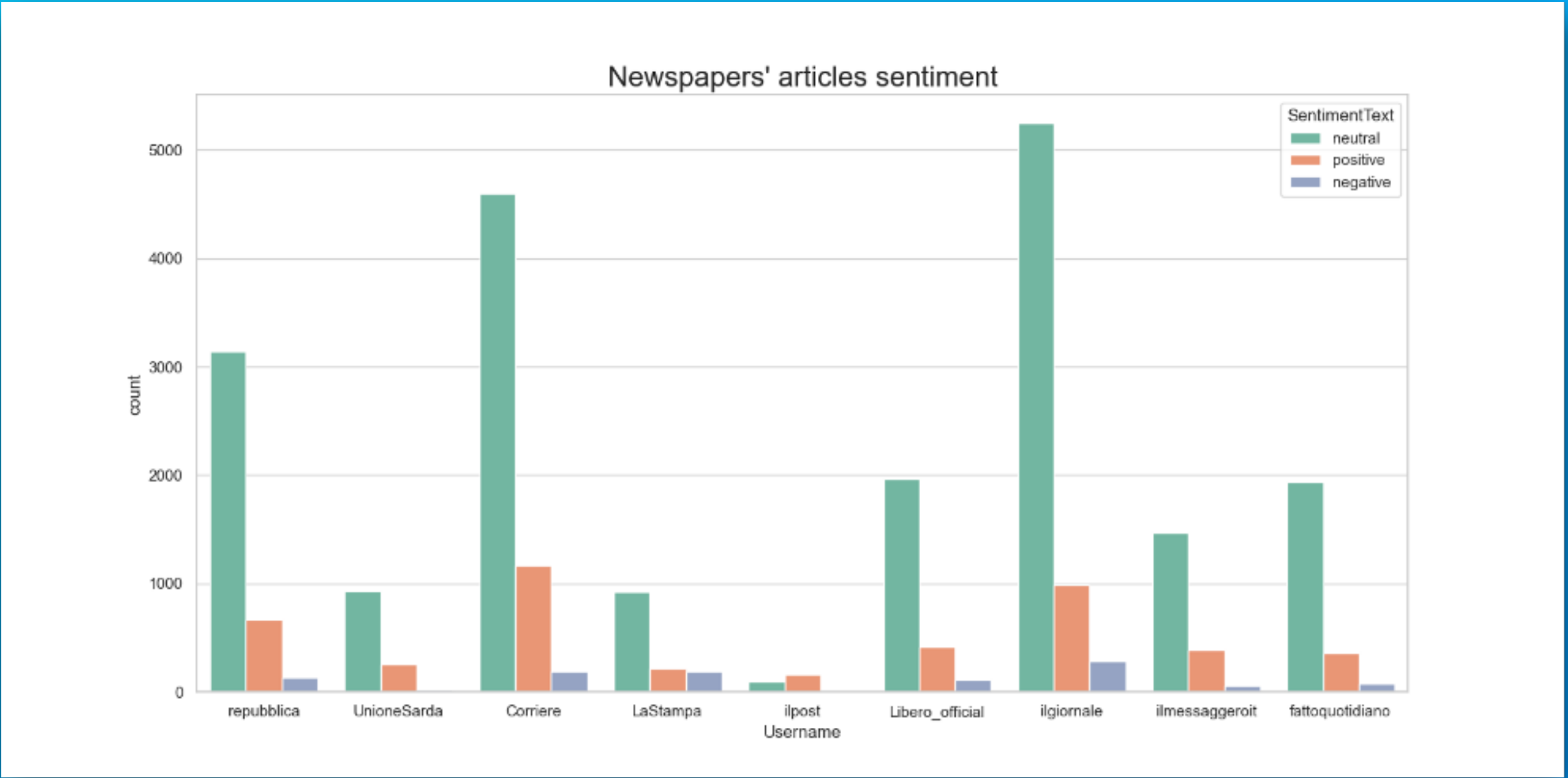
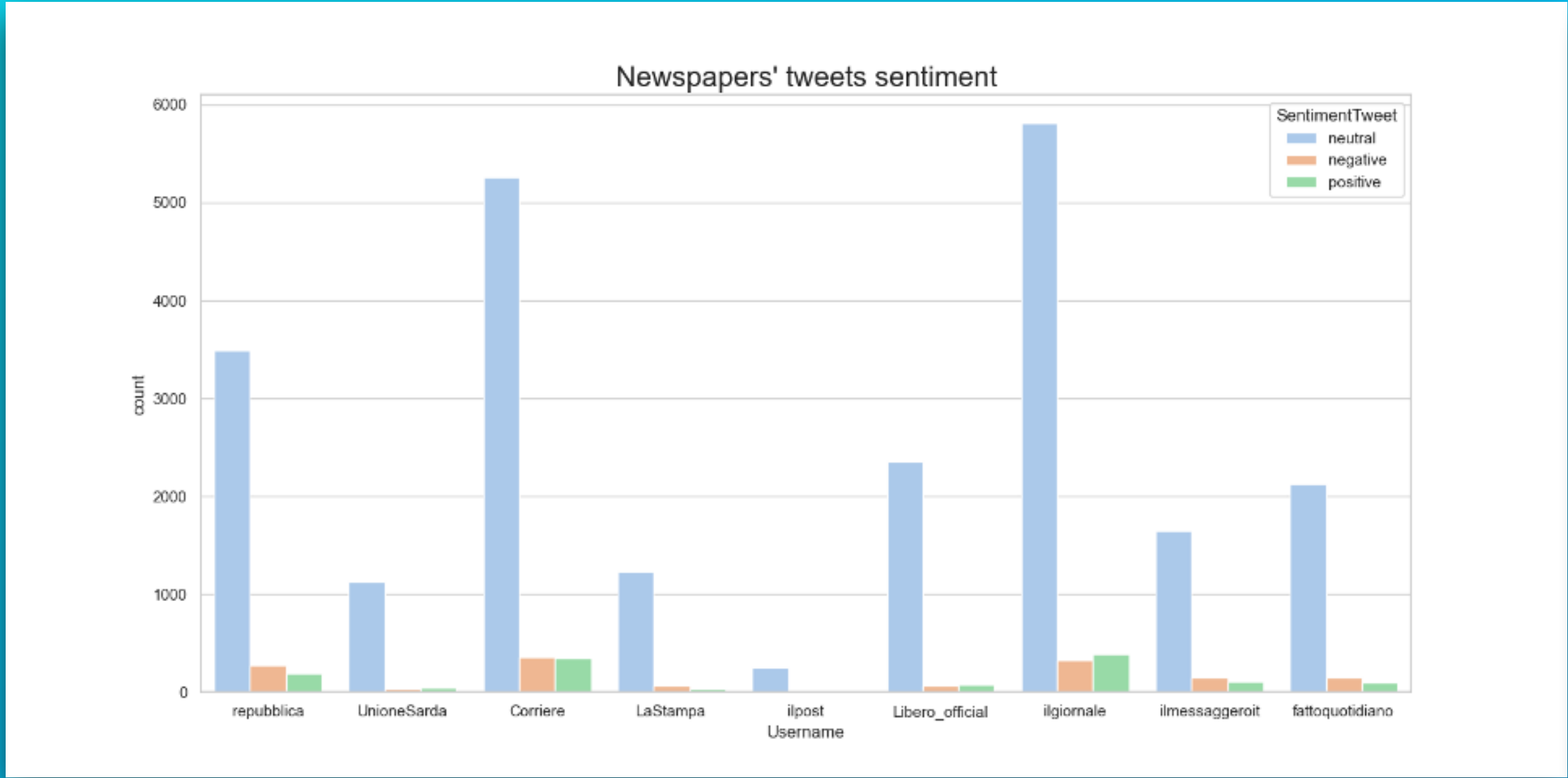
# Sentiment & emotions analysis

- *Calcolo sentiment analysis su articoli* ✓
- *Calcolo sentiment analysis sui tweets dei partiti* ✓
- *Detecting emotions sugli articoli dei giornali* ✓
- *Detecting emotions sui tweets dei partiti politici* ✓
- *Prevalenza di sentiment ed emotions neutrali*
- *Es:*
  - *emotions neutrale per l'area politica: **sinistra 88%** - **area politica neutro 85%** - **destra 91%***



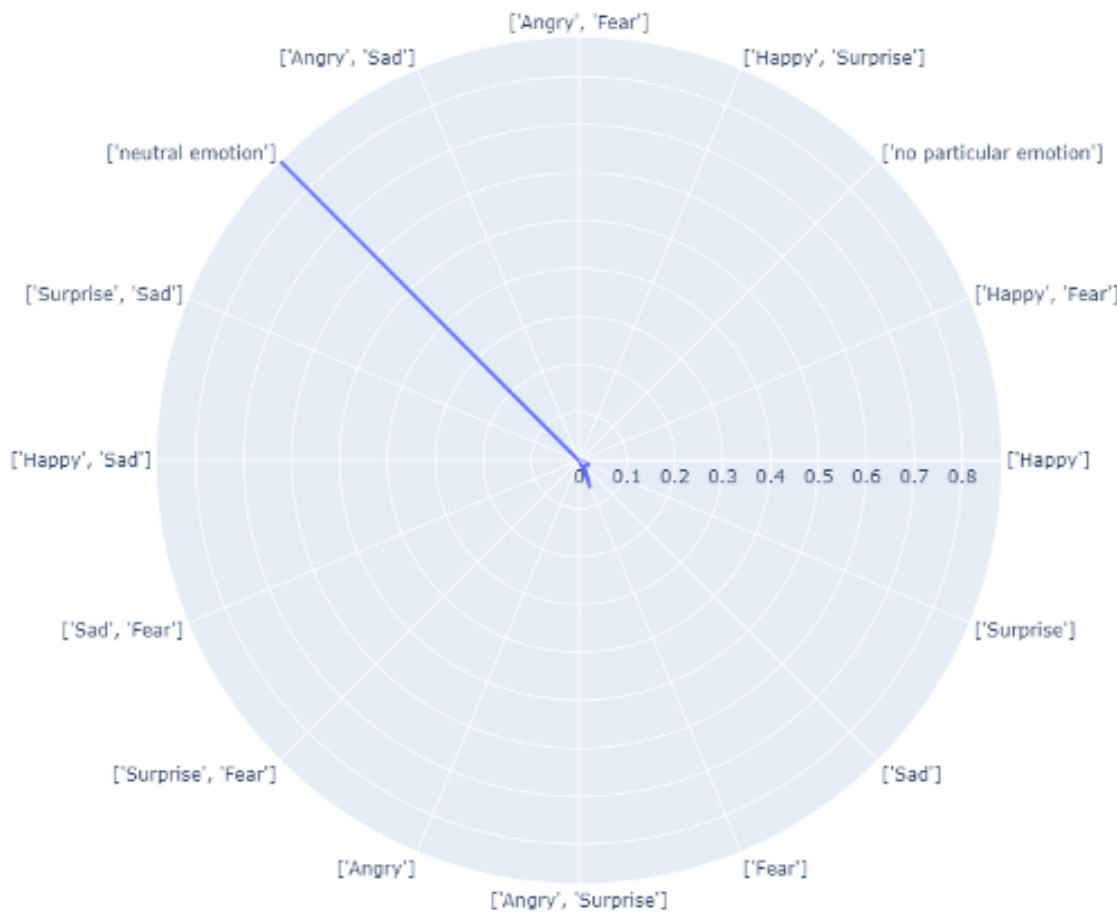


# Sentiment & emotions analysis

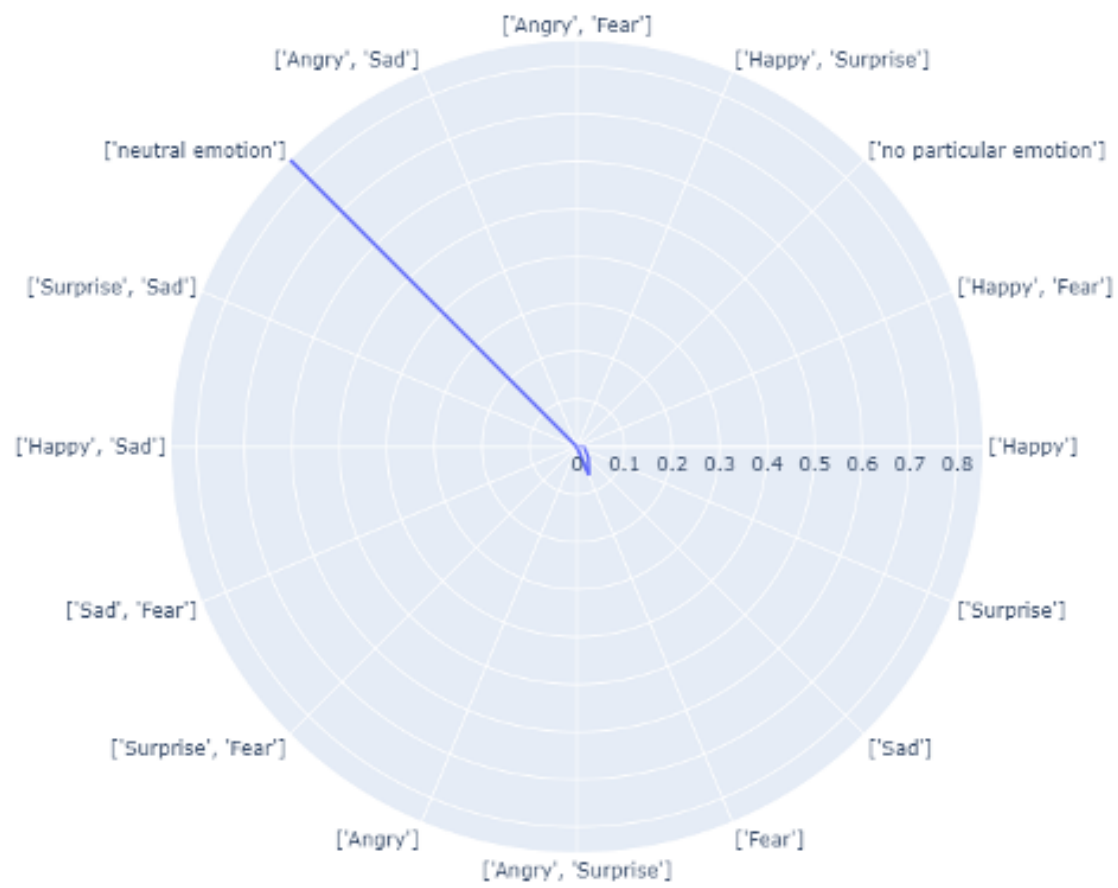


# Sentiment & emotions analysis

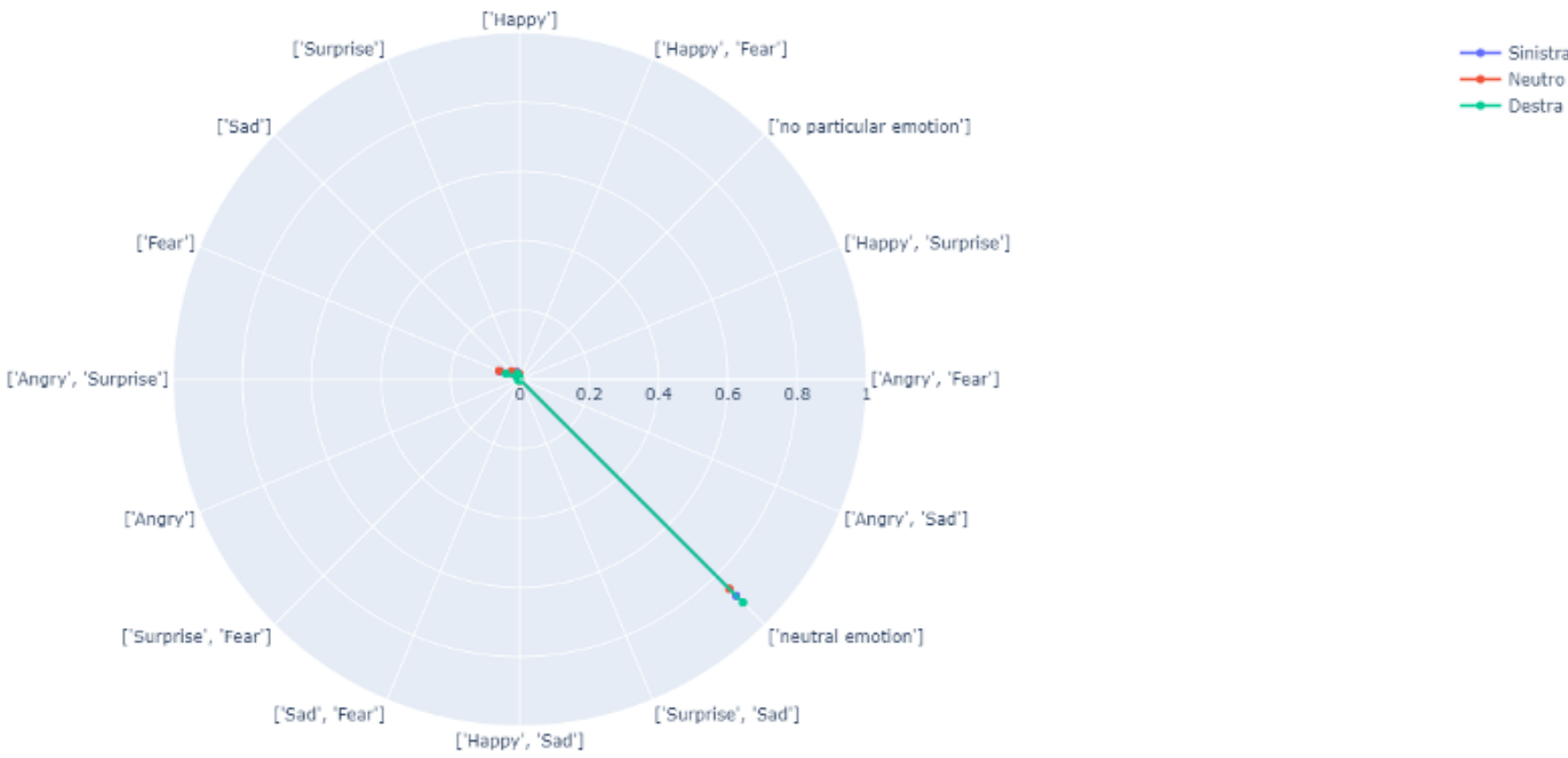
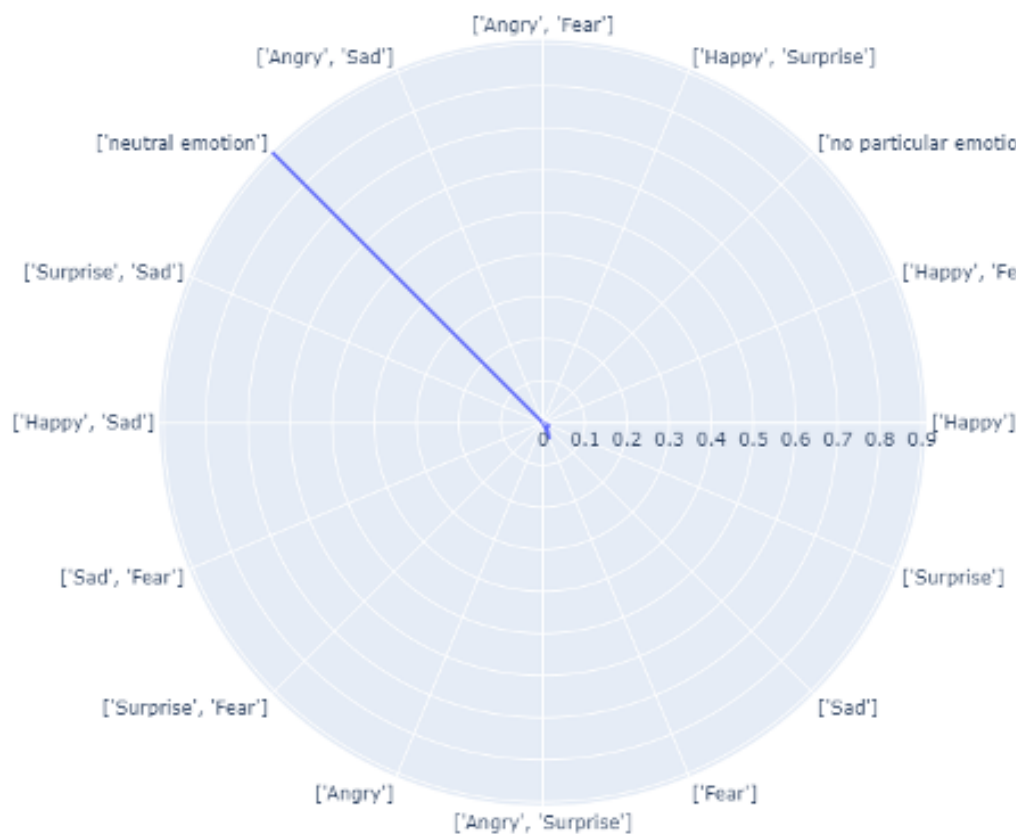
Spider plot for sinistra



Spider plot for neutro

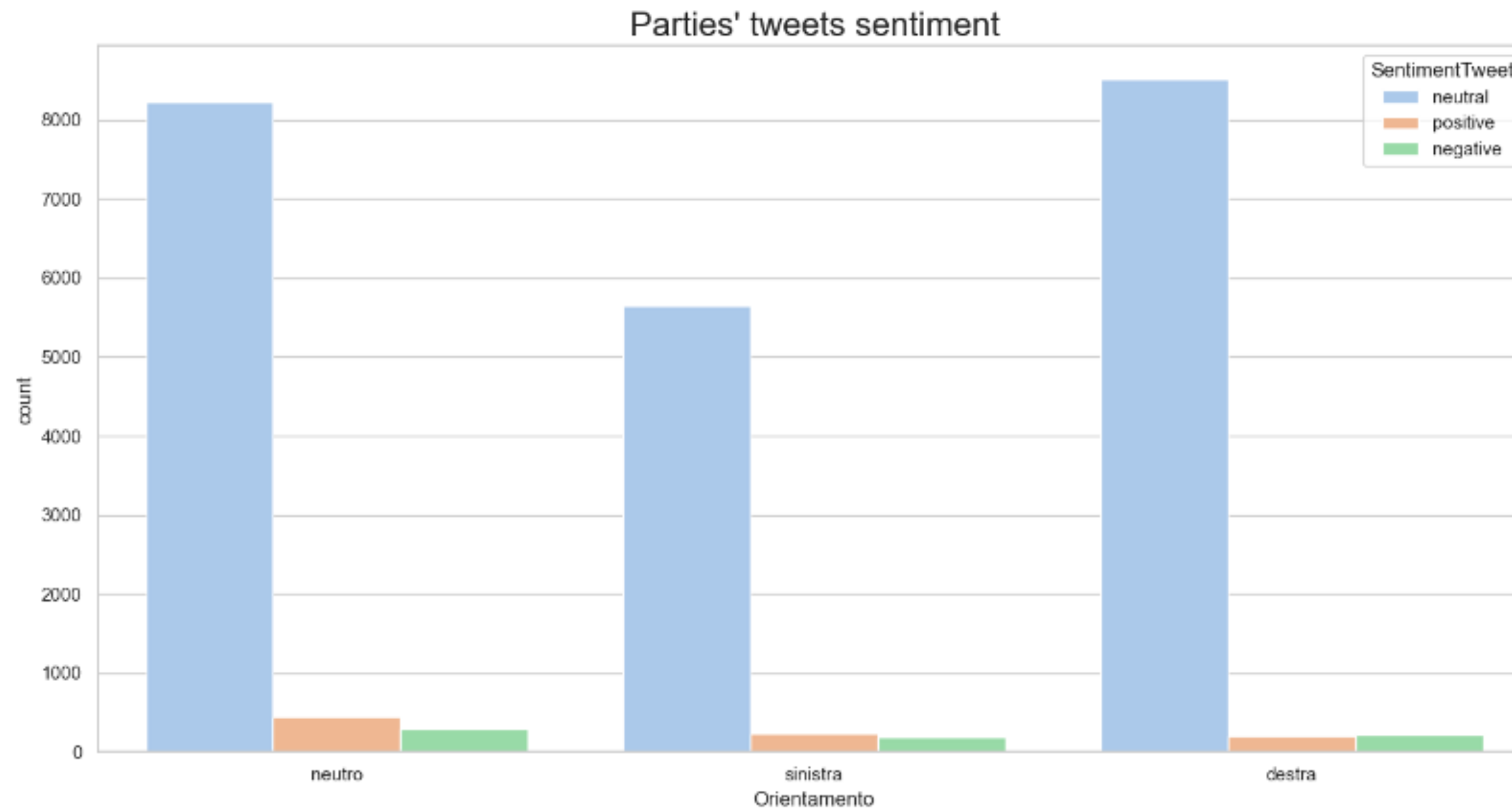


Spider plot for destra





# Sentiment & emotions analysis



# Random forest

- 6000 tweets per area politica (18000 osservazioni)
- 80% training set / 20% test set
- Tfidf vectorizer pre fitting
- File .pickle per salvataggio modello
- Il report ➡➡

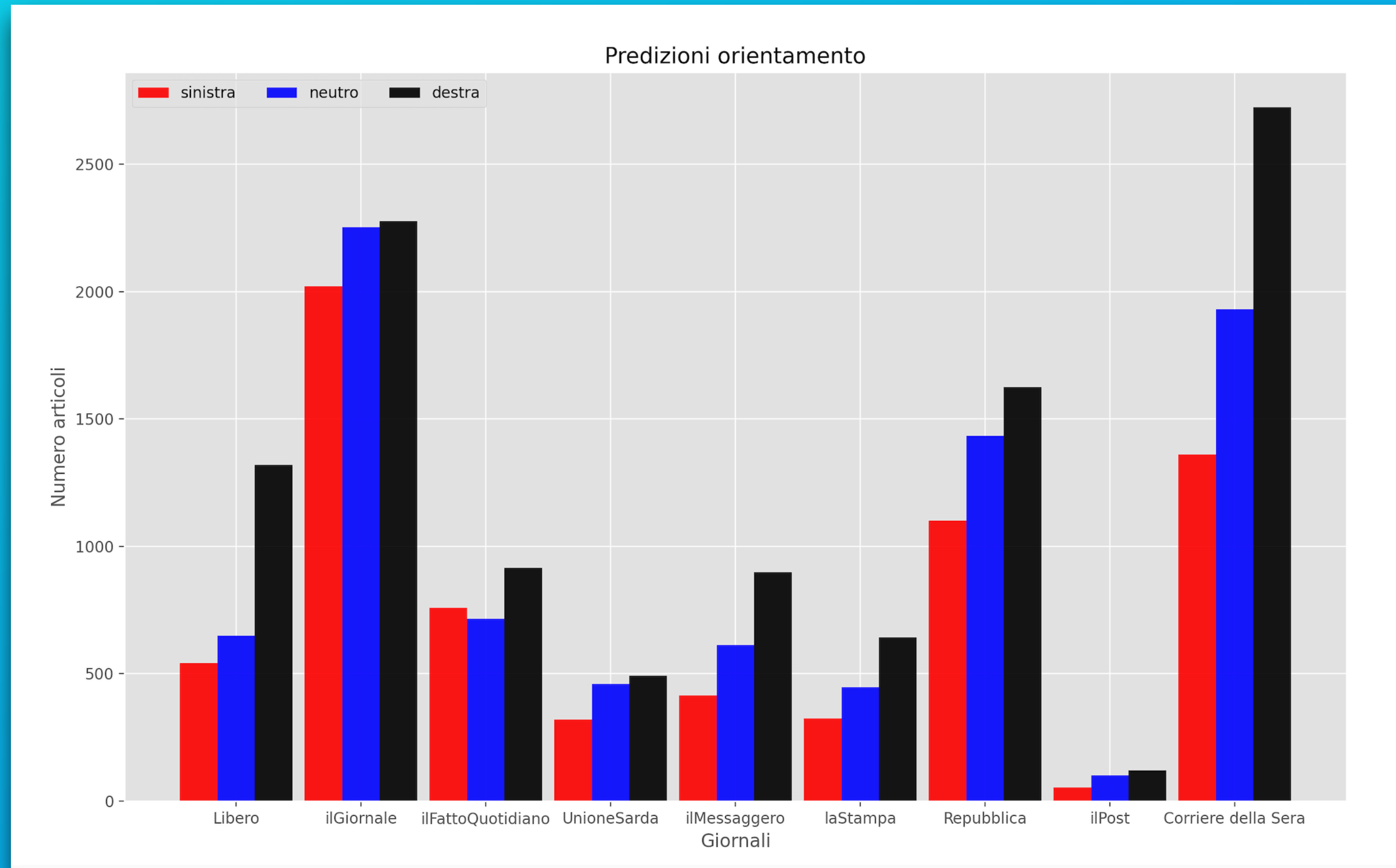
```
report :
          precision    recall  f1-score   support

   destra         0.83     0.85     0.84     1191
   neutro         0.78     0.78     0.78     1213
   sinistra        0.78     0.76     0.77     1196

 accuracy                   0.80     3600
 macro avg         0.80     0.80     0.80     3600
weighted avg         0.80     0.80     0.80     3600
```

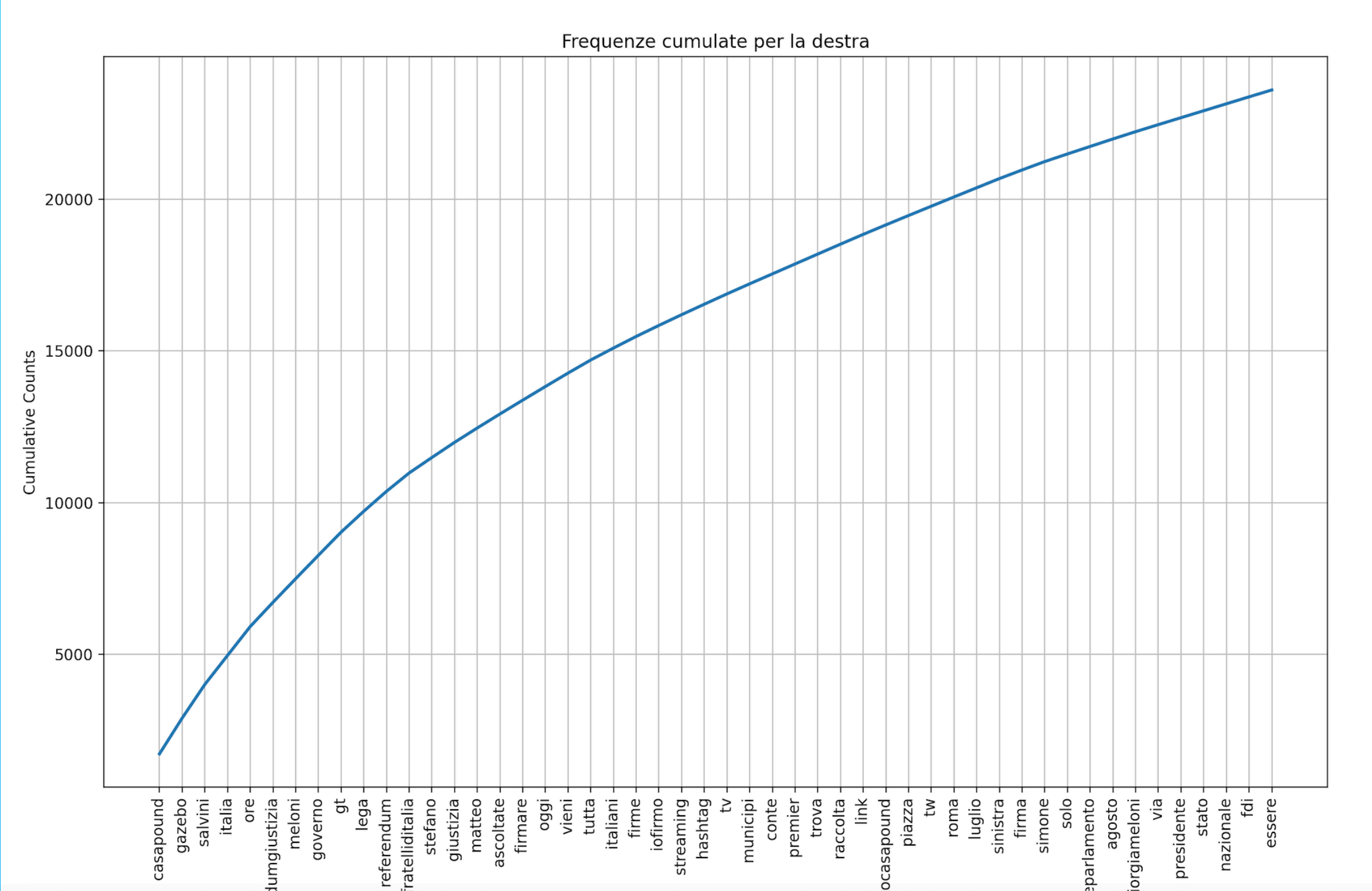
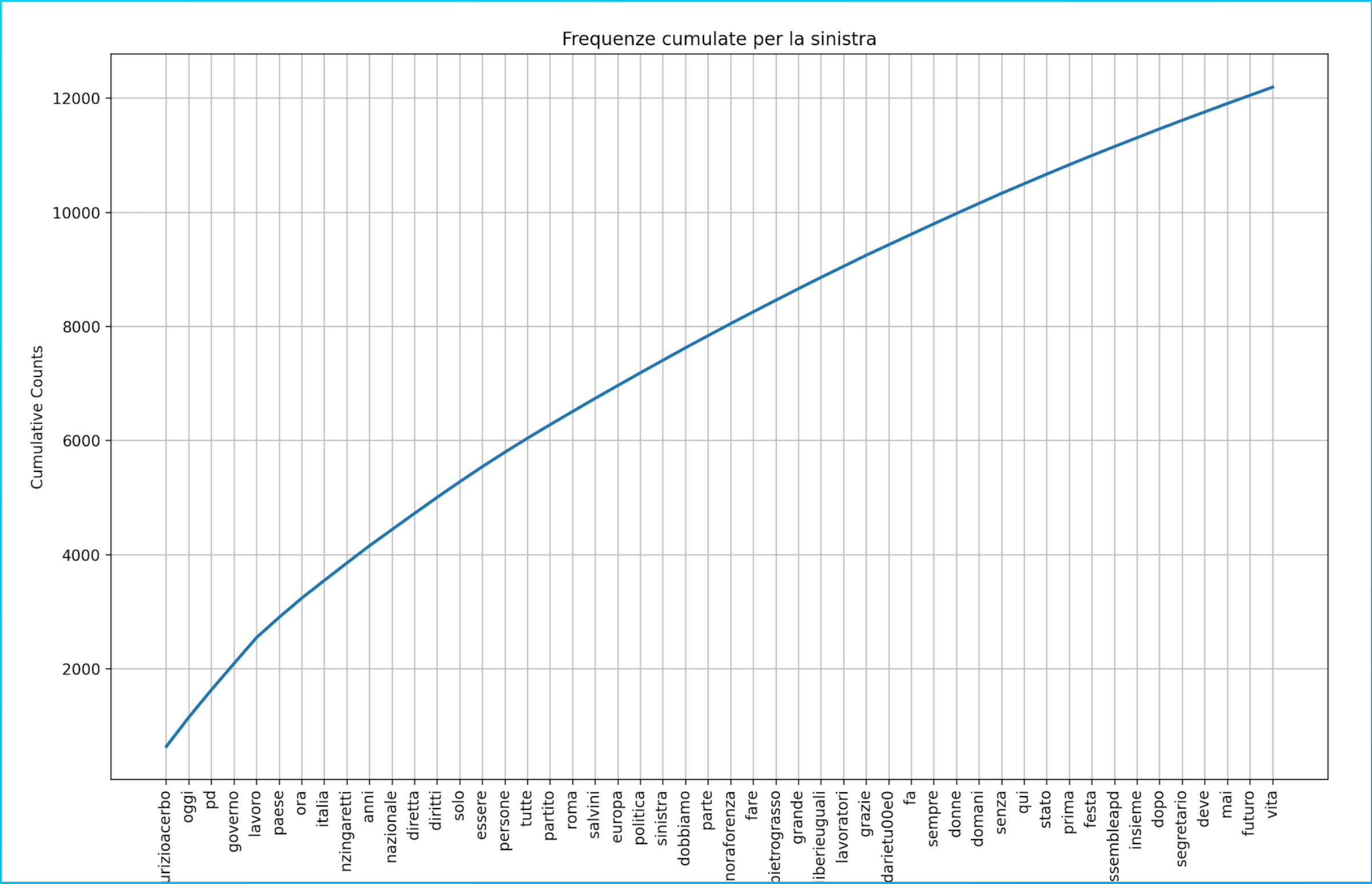


# Random forest





# Frequency distribution



# Grazie.

Alessia Lusci - [a.lusci2@studenti.unica.it](mailto:a.lusci2@studenti.unica.it)

Nicholas Garau - [n.garau6@studenti.unica.it](mailto:n.garau6@studenti.unica.it)

