Text Emotion Analysis

Authors: **Wenjun SUN, Adrien CLAUDEL, Guillaume RICHEZ**

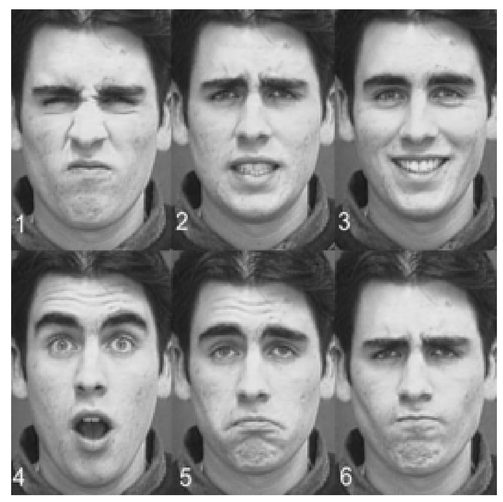
Supervisor: **Marianne CLAUSEL**

Examiner: **Miguel COUCEIRO**

UNIVERSITÉ DE LORRAINE

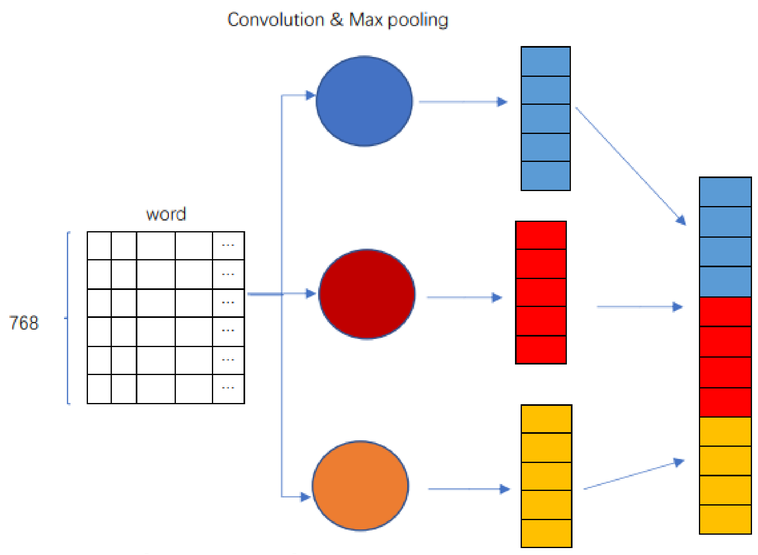
Abstract

Text emotion analysis refers to mining and analyzing the opinions and emotions oftexts through computer technology. It has become one of the most active re-searchfields in NLP, and has expanded from computer science to management and sociol-ogy, such as marketing, communication, health science, and even his-tory. Existingresearch has produced a large number of techniques that can be used for multipletasks in sentiment analysis, including supervised and unsuper-vised methods. Thesupervised method uses supervised machine learning meth-ods and feature combi-nations. Unsupervised methods include different meth-ods using emotion dictionar-ies, grammatical analysis, and syntactic patterns. This paper analyzes the emotionrecognition effects of various classifiers based on the ISEAR dataset. At the sametime, we also performed a cluster analysis based on the French corpus.

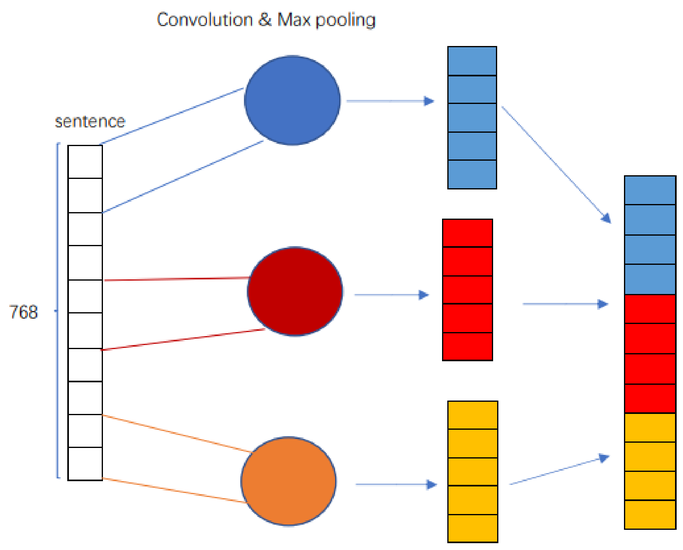


Corpora & Experiments

- Classification Task:
  - Corpora: **ISEAR**
  - Word-Embedding: **DistilBERT**
  - Classical supervised Methods:
    - **SVM**
    - **Random Forest**
    - **Logistic Regression**
  - Neural Networks:
    - **TextCNN**:



    - **TextCNN(Sentence)**:



    - **LSTM**

- Clustering Task
  - Corpora: **French Corpus** & **French Twitter**
  - Models:
    - **fr_core_news_sm**
    - **fr_dep_news_trf**

Results & Conclusions

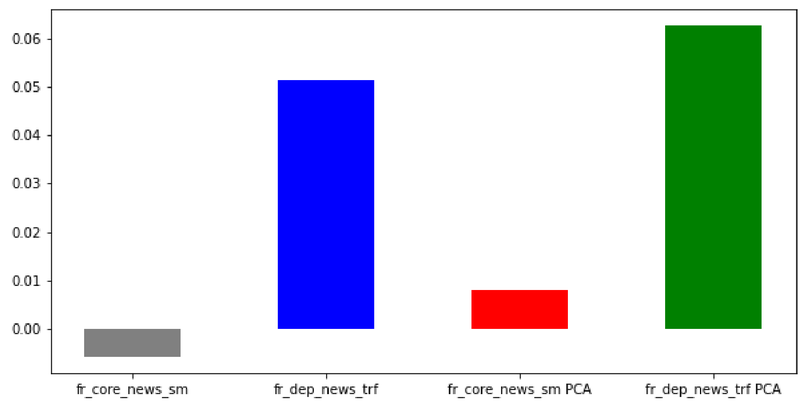
- Classification Task:
  - ClassicalsupervisedMethods:

method	accuracy
SVM	0.53
Random Forest	0.45
Logistic Regression	0.52

  - NeuralNetworks:

method	accuracy
Linear network	0.3783
TextCNN(word-vector)	0.6216
TextCNN(sentence-vector)	0.6164
LSTM	0.6320

- Clustering Task:



- Conclusion
  - Based on the ISEAR data set, we compared the performance of statistical classification methods and neural networks in emotion recognition.
  - We used the DistilBert langage model to obtain word and sentence embeddings.
  - Obviously, the performance of neural network is better on this task.
  - Some classes may be confused as (anger, disgust) or (shame, guilt)
  - For the French corpus, we achieved clustering task