

Bert for Sentiment Analysis Classification

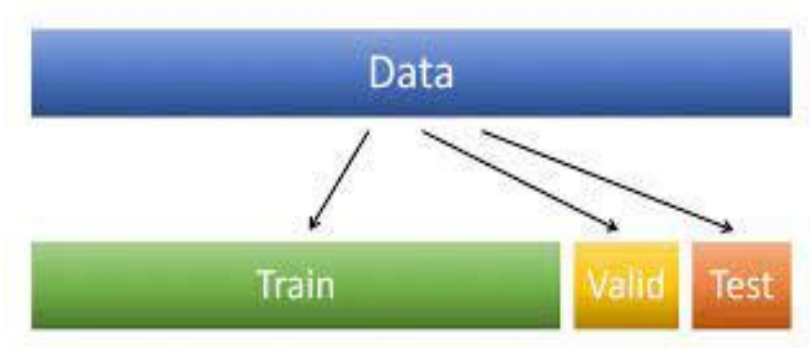
Francesco Pugliese, PhD

neural1977@gmail.com

Methodological Approaches for the Sentiment Analysis

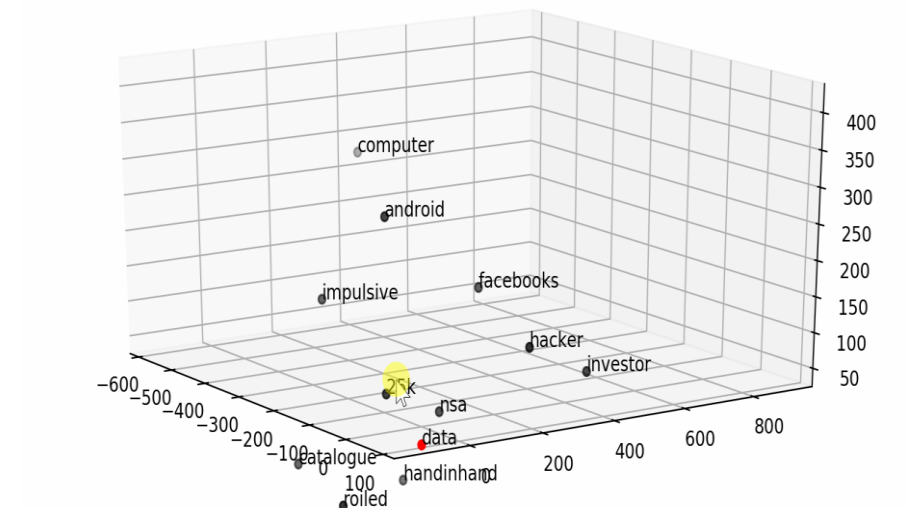
- 1) Techniques based on the **Lexicon Dictionary (Sentix, ecc)**: we assign some **scores** to some words. **Features** are **extracted words** from sentences within the text and they are assigned to one score which will define the sentiment of the whole sentence. Features are not chosen from a model but by a domain expert.
- 2) Techniques based on the **Machine Learning**: the features needed to assign a sentiment class are extracted automatically by a **Machine Learning (ML) / Deep Learning (DL)** model which is trained on a **Training Set**.

word	pos	neg	polarity	intensity
riforma	0.125	0	1	0.125
fallimento	0.021	0.375	-0.929	0.376
dat	0.063	0.104	-0.312	0.121
disoccupazione	0	0.625	-1	0.625
giovanil	0.125	0	1	0.125
aumento	0.208	0.083	0.516	0.224



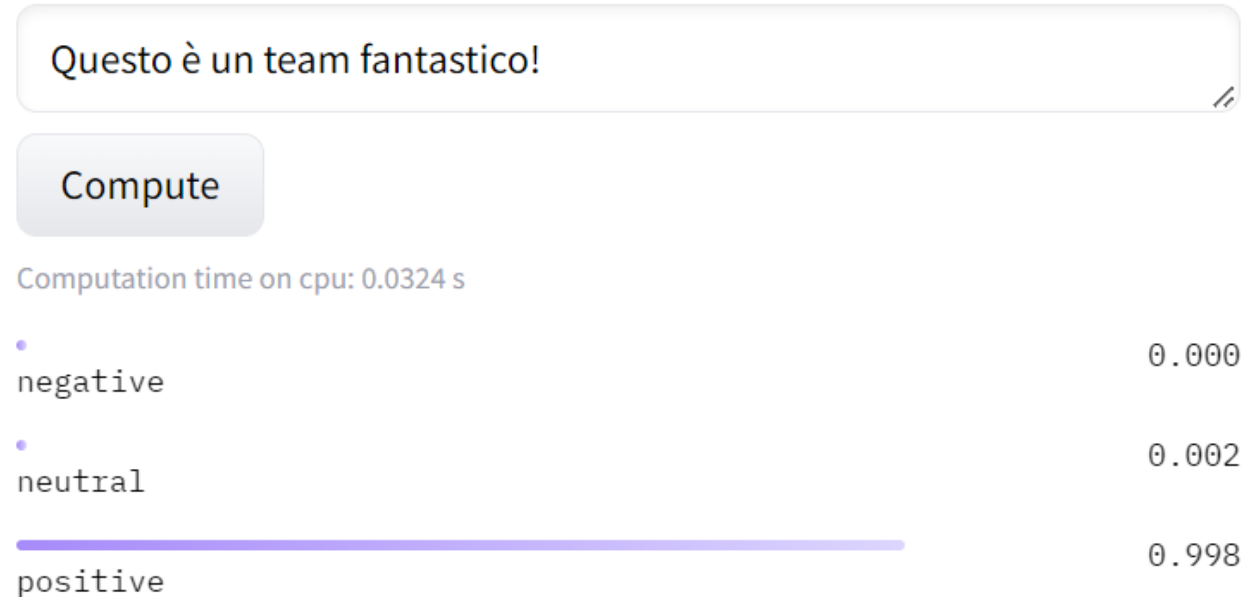
BERT - Bidirectional Encoder Representations from Transformers

- **BERT** is the **state-of-art** in **Natural Language Processing** of ML/DL models. It was created by the **Google AI Language** group in **2018** (Devlin, et al., 2018). It is based on neural networks with attention mechanisms which are called **Transformers**. They use multi-head attention in a bidirectional mode:
- Strong points of BERT:
 1. **Contextualized Word Embeddings**, that is it recognizes words and its context
 2. **Many pre-trained models**
 3. **More accurate syntactical / semantical analysis of the sentence structure**



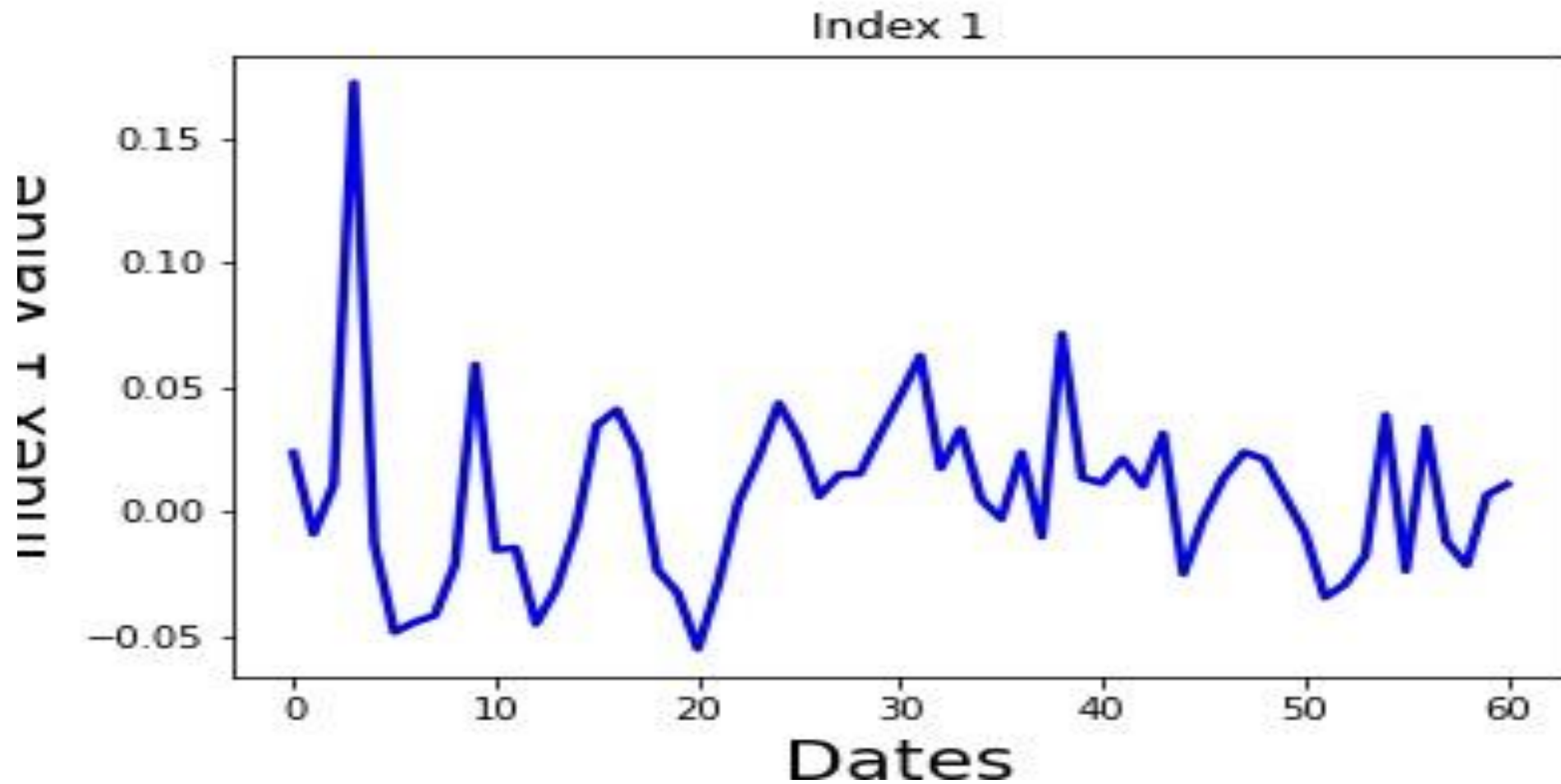
Language Modeling by Bert and Hugging Face models

- **Bert** model, which we use, has been trained combining two datasets of tweets extracted from **Sentipolc EVALITA 2016** (Barbieri, et al., 2016). In total, the dataset is made of **45 thousands preprocessed tweets**. This model has achieved **82%** (state-of-art) of Test Set accuracy which is the **20%** of the entire Dataset.
- It is possible to re-train with the technique “**Transfer Learning**” on new tweets labeled from us, achieving optimal results.



Results: Index 1

$$Index\ 1 = \frac{N_p - N_n}{Tot}$$



References

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

<https://huggingface.co/neuraly/bert-base-italian-cased-sentiment>

Barbieri, F., Basile, V., Croce, D., Nissim, M., Novielli, N., & Patti, V. (2016, December). Overview of the evalita 2016 sentiment polarity classification task. In *Proceedings of third Italian conference on computational linguistics (CLiC-it 2016) & fifth evaluation campaign of natural language processing and speech tools for Italian. Final Workshop (EVALITA 2016)*.