Deep Learning vs. Traditional Machine Learning for Sentiment Analysis: A Performance Comparison

Shlomit Finegold

Motivation

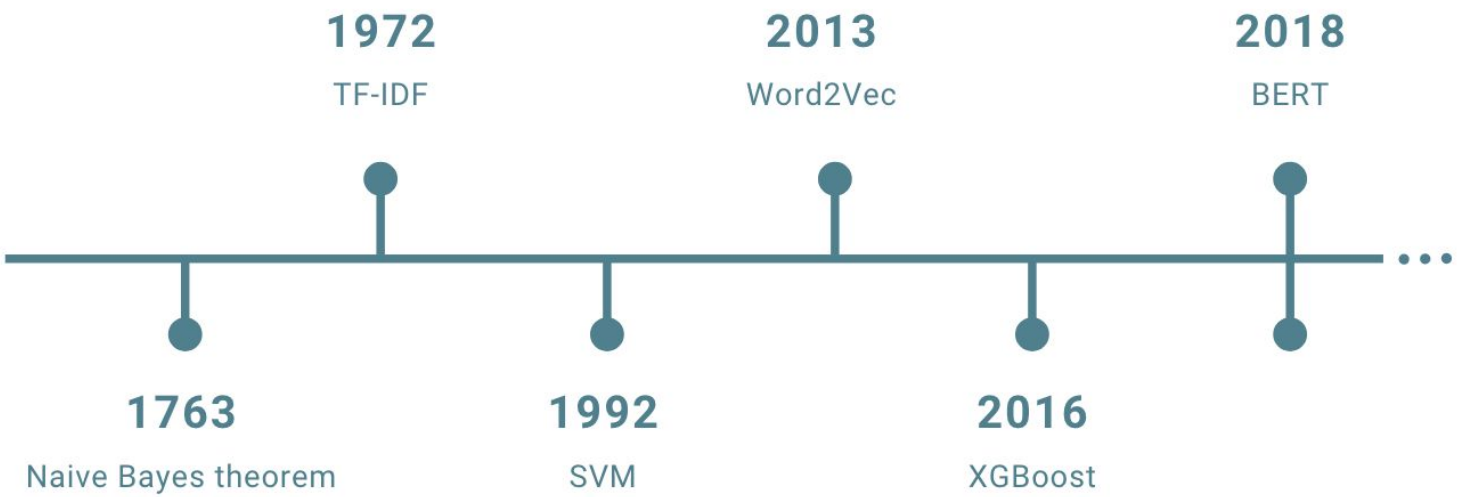
"The most successful entrepreneurs are those who can seamlessly blend the old with the new, finding value in both."

Reid Hoffman (LinkedIn co-founder)

Research aim

Can harnessing NLP and ML developments into low dimensionality matrix input of NLP based extracted features, be compatible with the most advanced DL techniques?

Literature review



Method

Data

- IMDb reviews database
- 50K records
- Perfectly balanced

	Traditional ML	SOTA DL
pre-processing	• Text cleaning (URLS, HTML tag) • Extracting numeric NLP based features • Scaling and standardization • Feature selection	• Text cleaning (URLS, HTML tags, punctuations) • Lower case • Lemmatization • Tokenization
Baseline models	• SVM • NB	• Pre-trained DistilBERT

Compared models	• XGBoost
	  ○ Grid search parameters optimization
	  ○ Thresholding score optimization
		• Fine tuned pre-trained DistilBERT
		  ○ Hyperparams tuning on validation set
		  ○ Retraining on training+validation set

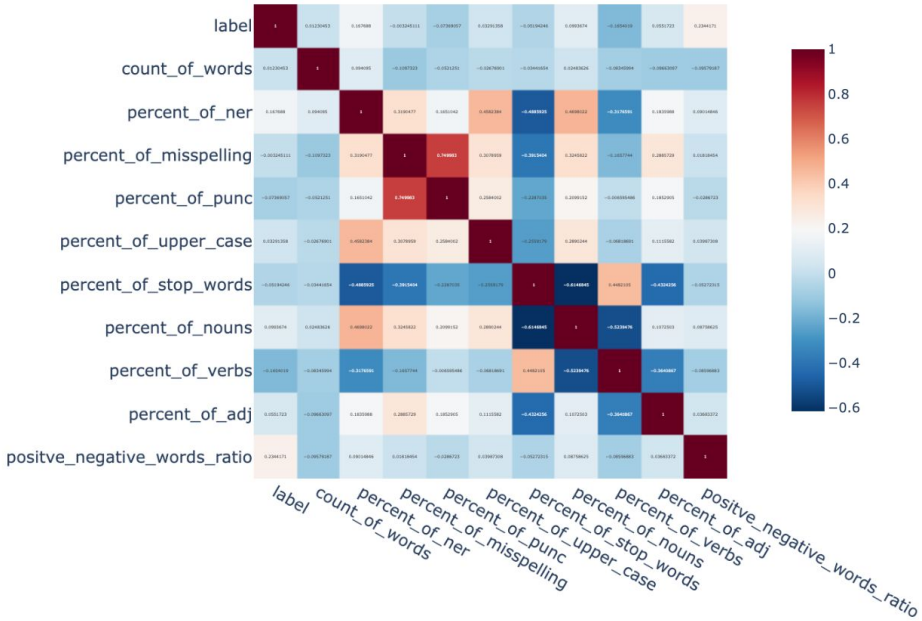
Method- Traditional ML

Features extraction:

- Count of words
- Percent of name entities (count of name entities / count of words)
- Percent of punctuation (count of punctuation / count of words)
- Percent of uppercase words (count of uppercase words / count of words)
- Percent of stop words (count of stop words / count of words)
- Percent of spelling mistakes (count of spelling mistakes / count of words)
- Percent of nouns (count of nouns / count of words)
- Percent of verbs (count of verbs / count of words)
- Percent of adjectives (count of adjectives / count of words)
- Ratio between common positive and negative words

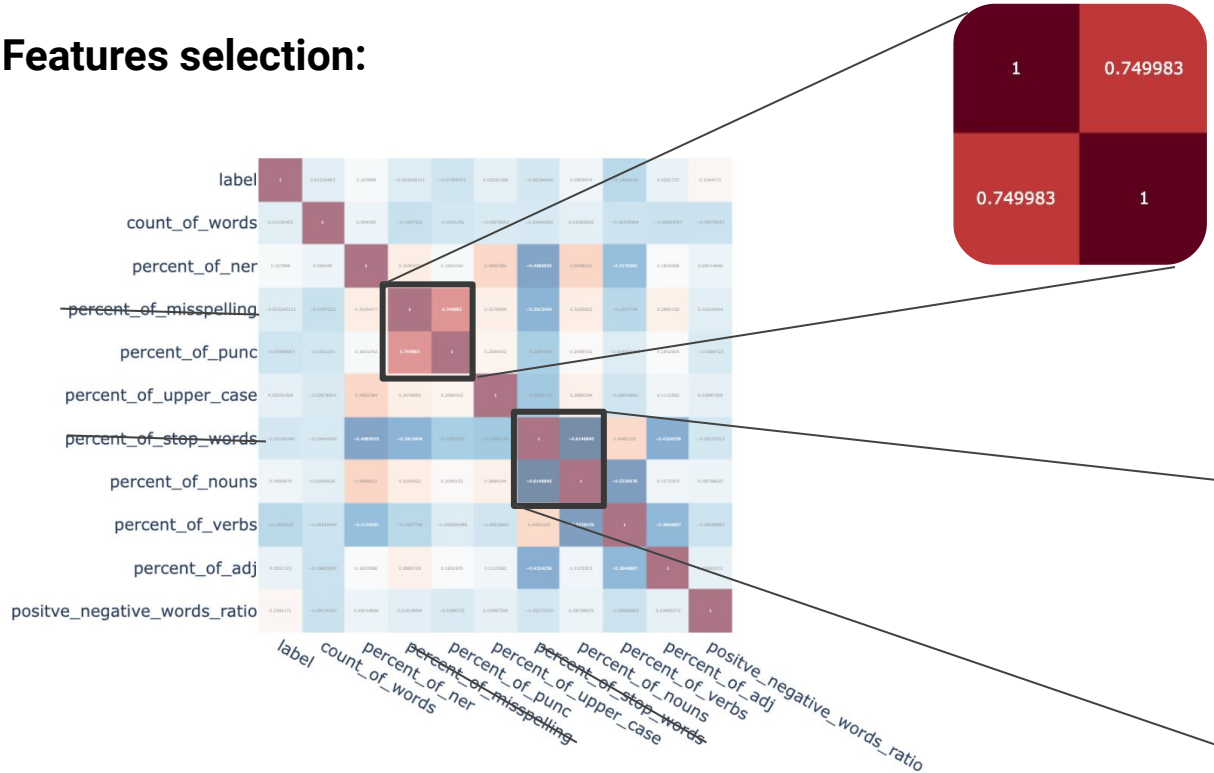
Method- Traditional ML

Features selection:



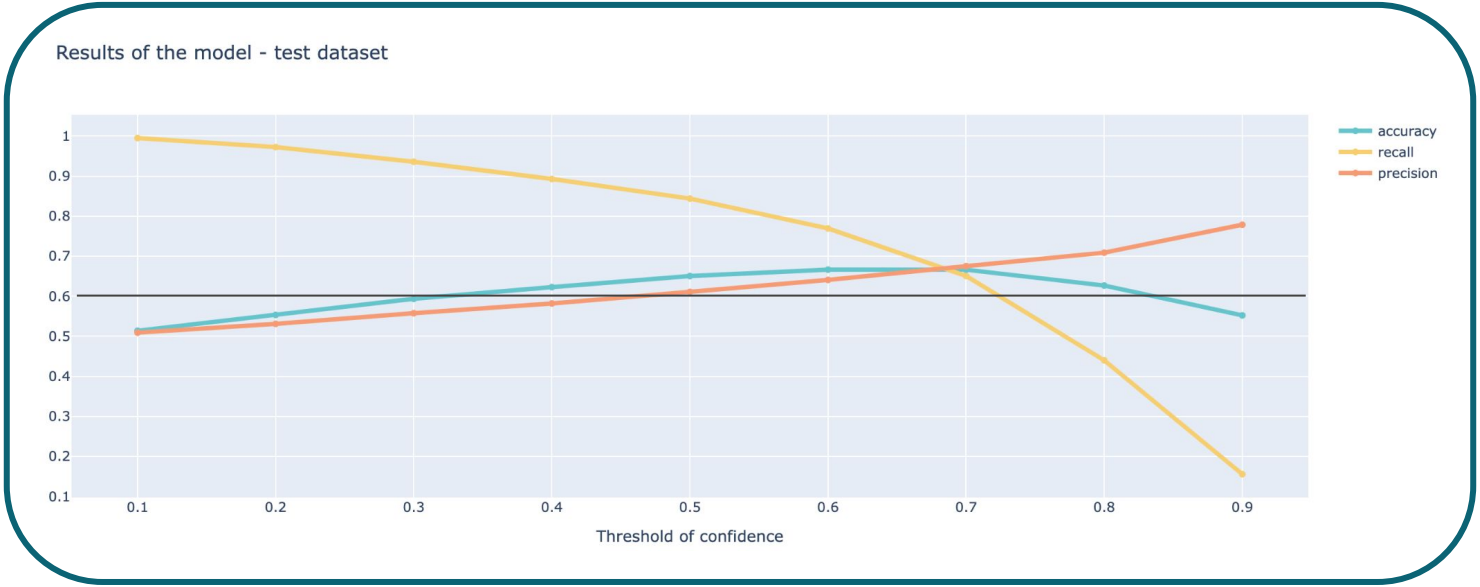
Method- Traditional ML

Features selection:



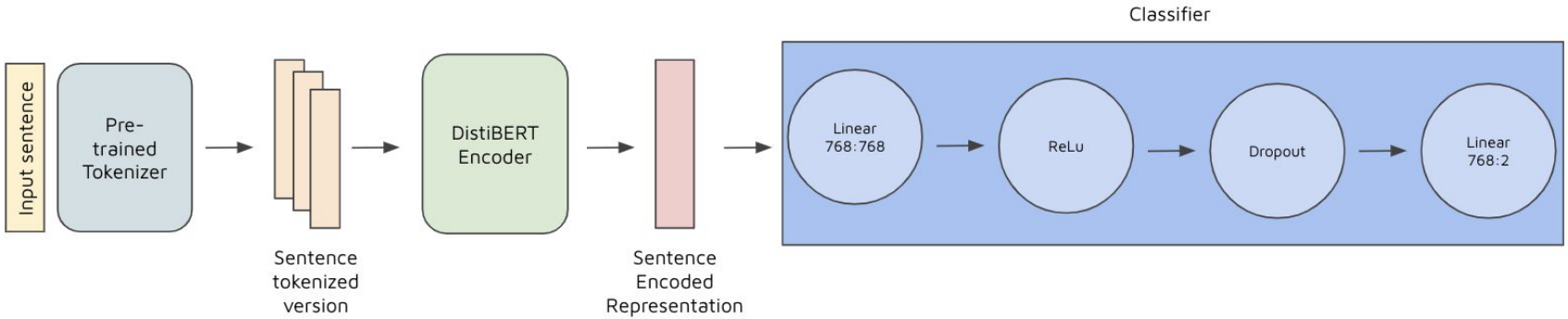
Method- Traditional ML

XGBoost prediction with thresholding optimization:



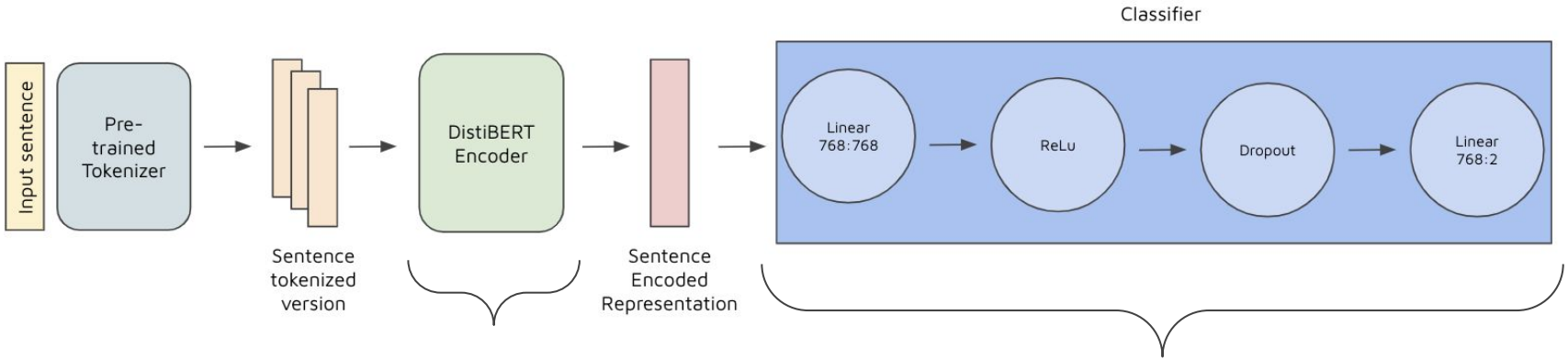
Method- SOTA DL

DistilBERT for classification architecture:



Method- SOTA DL

DistilBERT for classification architecture:



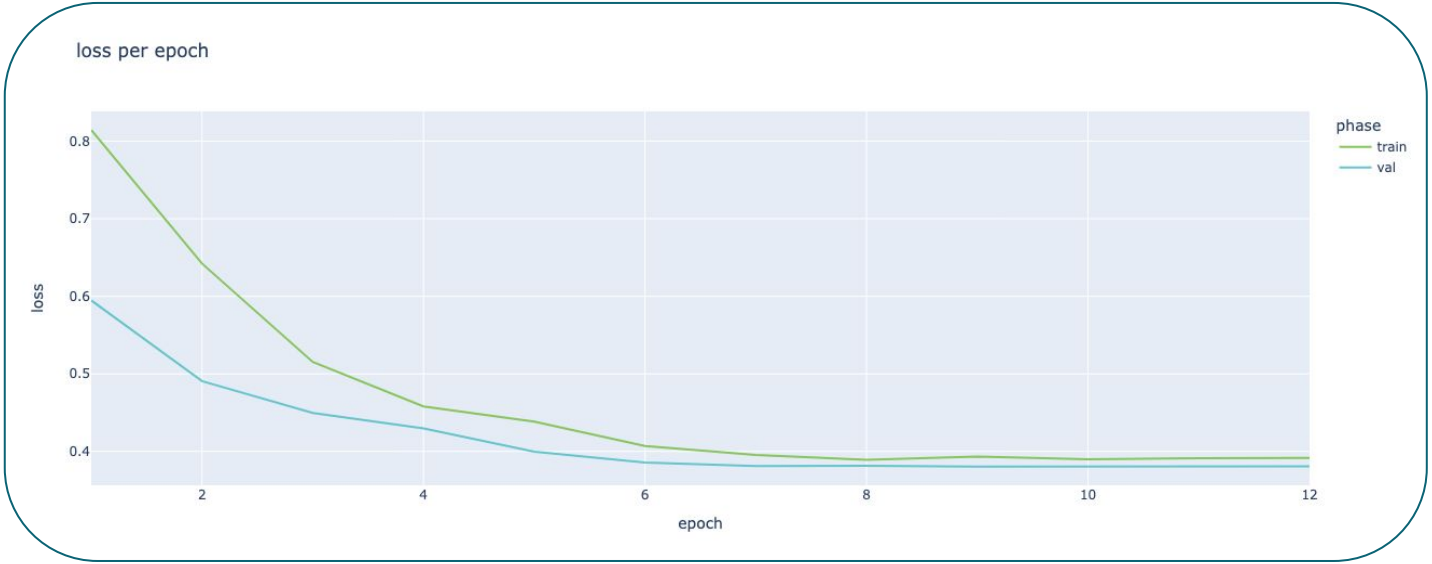
Freezed layers

Retrained for fine tuning

Method- SOTA DL

Hyperparameters optimization:

Learning rate set to 5e^-4 for 5 epochs, e^-4 for one epoch, Adam optimizer



Results

Method	Model	Accuracy (improve)	Recall (improve)	Precision (improve)
Traditional ML	SVM	0.649	0.697	0.639
	Naive Bayes	0.504	0.99	0.504
	XGBoost (ths=0.6 optimizing accuracy)	0.666	0.769	0.64
	XGBoost (ths=0.4 optimizing recall with min accuracy > 0.6)	0.622	**0.892 (14%)**	0.581
SOTA DL	Pre-trained DistilBERT	0.786	0.633	**0.909(27%)**
	Fine Tuned DistilBERT	**0.818(15%)**	0.748	0.868

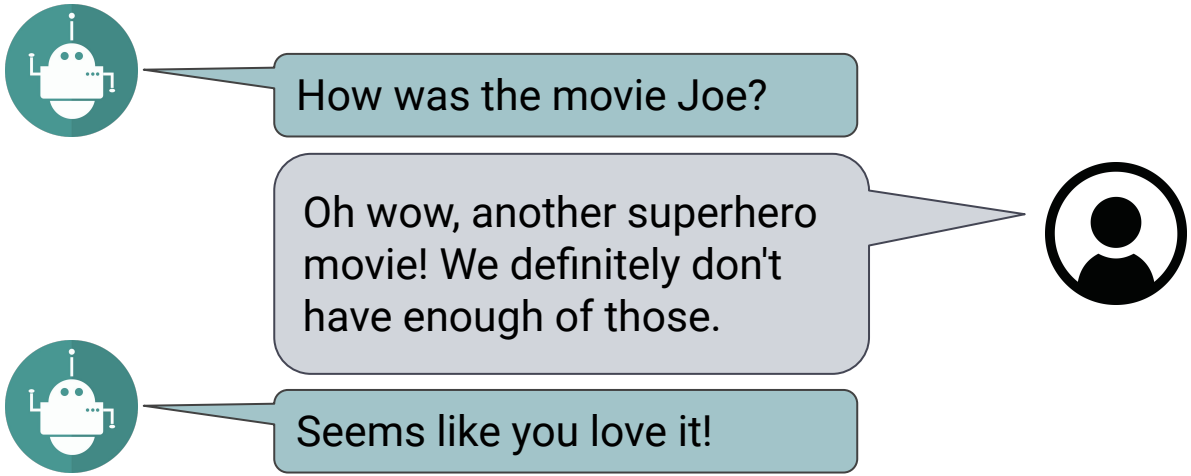
Conclusions

1. The traditional ML approach managed to achieve learning, though far from being compatible with the DL approach
2. Precision-Recall trade off:

   Recall → traditional ML, Precision → DL
3. Training convergence issues along the way

Discussion

In a world full of chatbots, can Bots developer actually rely on sentiment analysis?



Questions

Thanks for listening!