INTENT CLASSIFICATION AND OOS PREDICTION

Natural Language Processing

Amanda Oliveira - Benedek Lellei - Filippo Comastri

THE TASK

CLASSIFY QUERIES INTO DIFFERENT INTENT CLASSES

Queries given to AI assistant:

Ex:

Query: "define ambivalence for me please",
Intent class: "definition"

Some Query are out of scope:

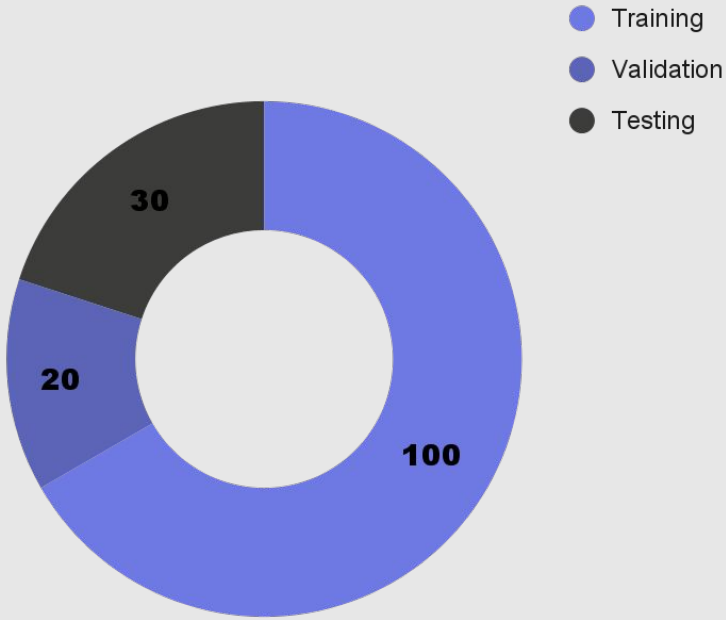
oos: "am kind of busy now"

,

[NLP]

THE DATASET

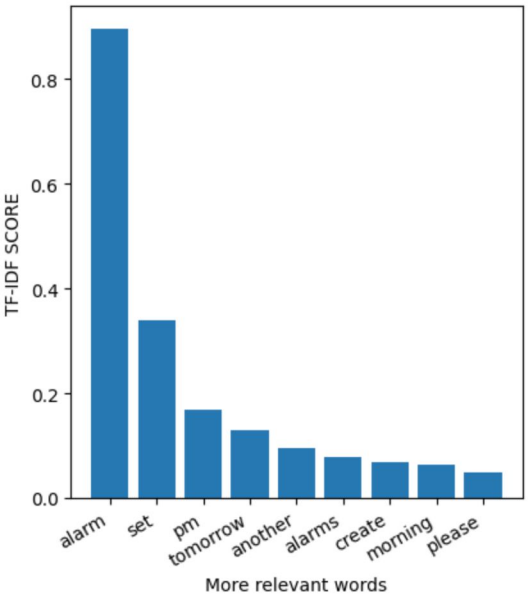
- Language: English.
- Contains **23,700** queries.
- Where **22,500 are in-scope** covering **150 intentens** (Grouped by 10 Domains).
- Only **1,200** are classified as **Out-of-Scope.**
- Two paths:
  - Distinguishing only between intent classes
  - Including OOS as an other intent

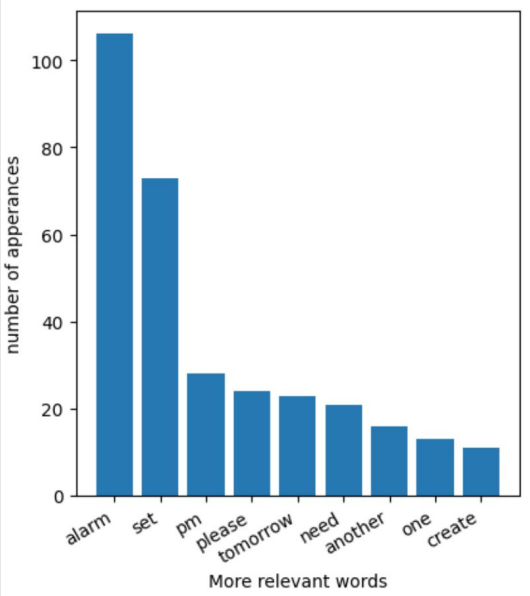


DATA ANALYSIS

USING TF-IDF & BAG OF WORDS

Intent: “Alarm”





FEATURE EXTRACTION METHODS

SPARSE REPRESENTATION

- - Bag of Words
- - TF- IDF

DENSE REPRESENTATION – WORD2VEC

- - Pre-trained on Google News
- - Trained with our dataset

CLASSIFIER MODELS & RESULTS

NATUR
PROCE
(NLP)

Baseline
Naïve Bayes and BOW

Accuracy: 0.865
Precision: 0.875
Recall: 0.865
F1: 0.863

<<<<

CLASSIFIER MODELS & RESULTS

	ACCURACY	PRECISION	RECALL	F1
L. REGRESSION	0.899	0.904	0.899	0.898
SVM	0.912	0.919	0.912	0.912
DECISION TREE	0.783	0.792	0.783	0.783
RANDOM FOREST	0.399	0.418	0.399	0.378
NÄIVE BAYES	0.732	0.741	0.732	0.724

ERROR ANALYSIS

The Out-Of-Scope class overlap with the other classes

Calculating the misclassification with the confusion matrix we obtained the following results:

Percentage of queries misclassified as oos:

13.3% - 'yes'
6.6% - 'new_card'
20.0% - 'restaurant_reviews'

.

Percentage oos queries misclassified as in scope:

2.0% - 'travel_suggestion'
2.1% - 'what_can_i_ask_you'
1.9% - 'directions'

CONCLUSIONS

Highlights:

- Sparce representation provided a better performance than word embeddings.
- Pre-trained word embeddings worse than self-trained.
- When including the out of scope class for training the predictions were slightly worse.

>>>>

THANKS!

ANY QUESTIONS?

First assignment - Natural Language Processing - Feup - 22/23

[NLP]