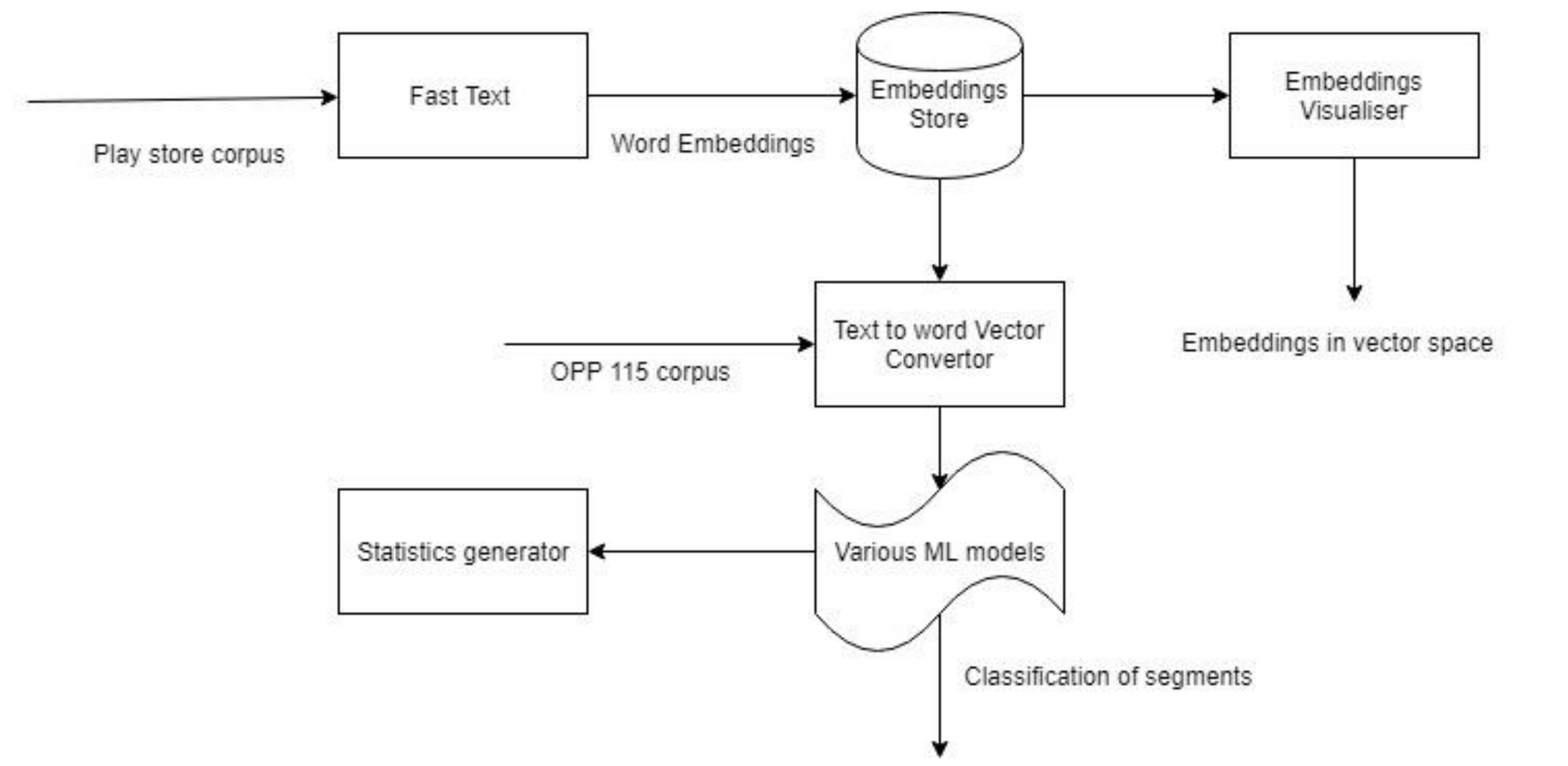
Quantifying in-domain Distributed Word Representations from Privacy Policies

Vinayshekhar BK,Abhilasha Ravichander, Peter Story and Norman Sadeh

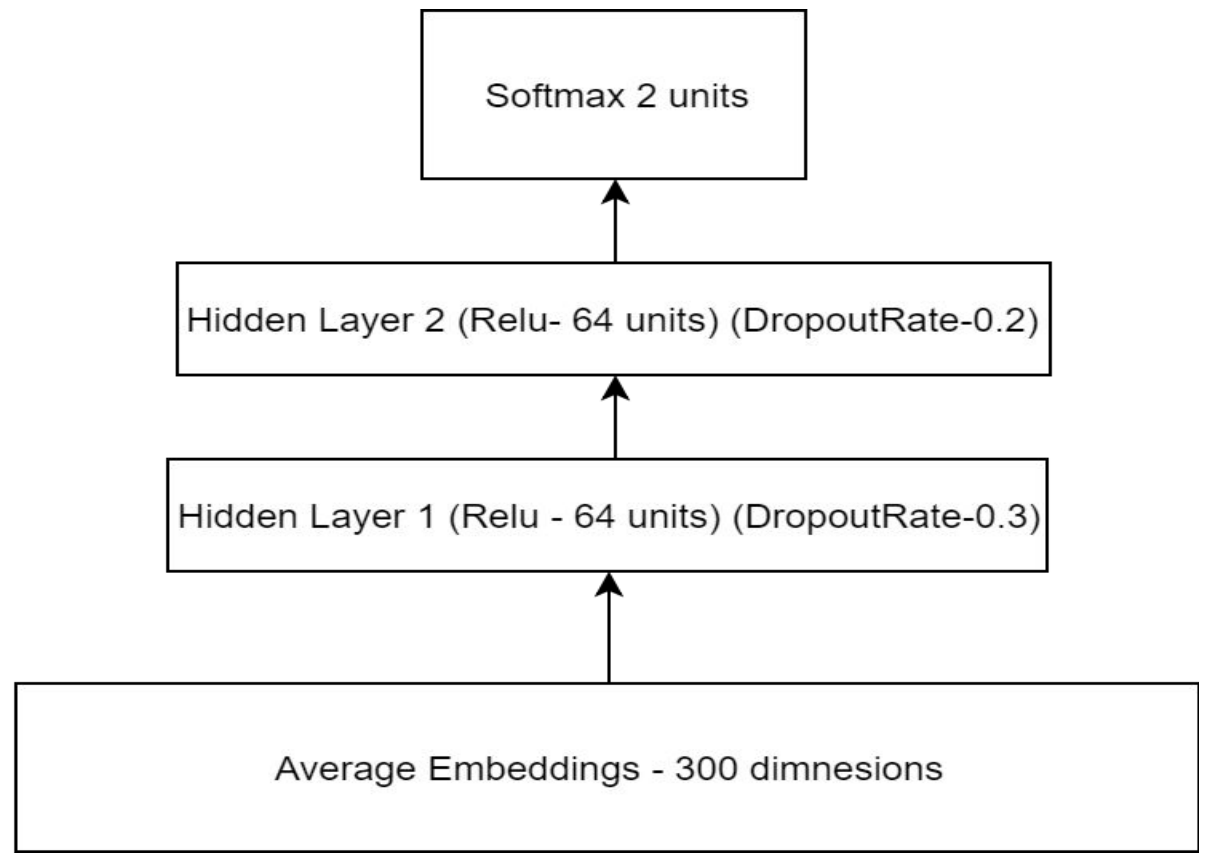
Motivation

Privacy policies are documents which describe what companies do with users' data.These are very long and often written in a way that is difficult for an average user to understand. In the recent past there has been interest in the Natural Language Processing (NLP) community to analyze these policies using machine learning algorithms. We want to improve research in this domain by building word embeddings. We use 150000 privacy policies to build word vectors in an unsupervised manner.

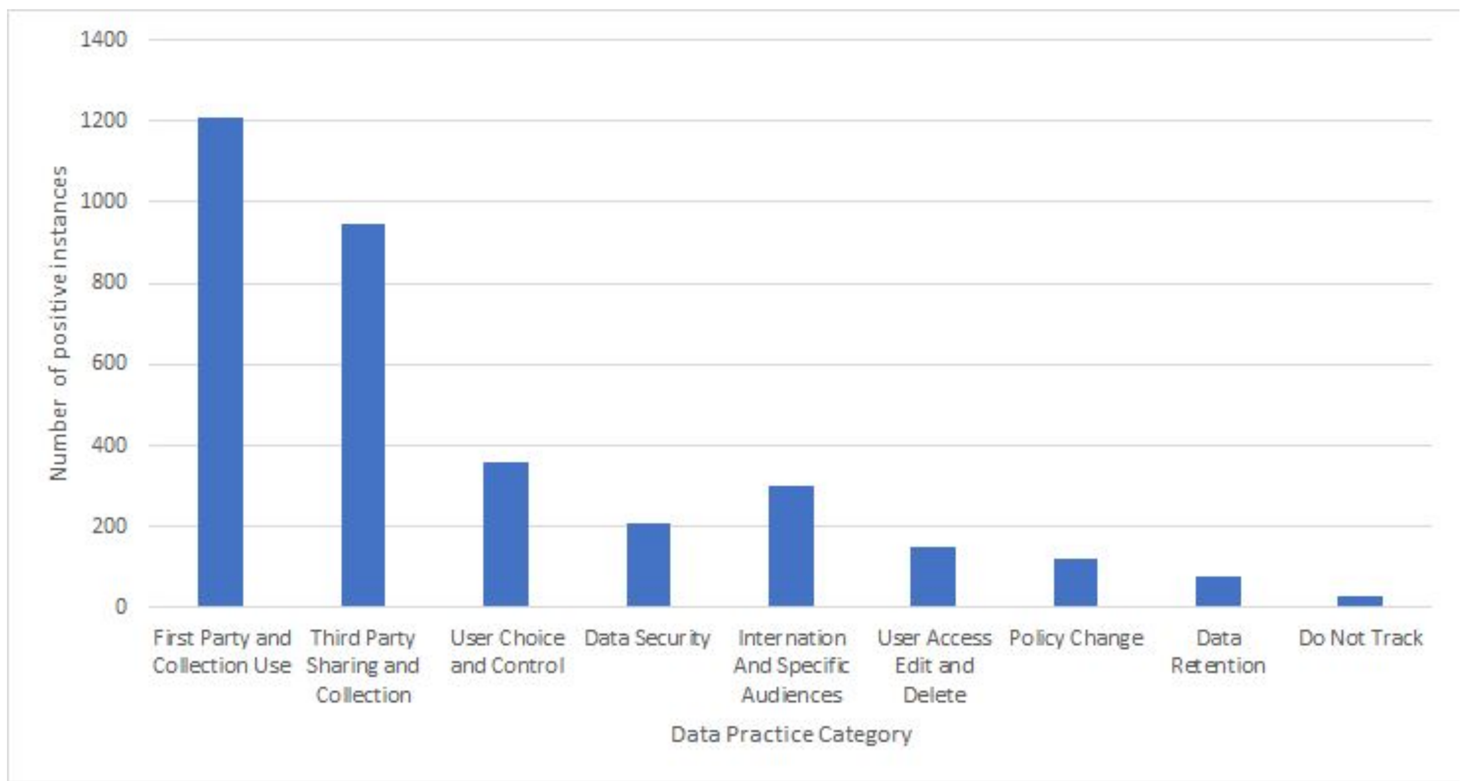
System Architecture



Deep CBOW Model



Corpus Stats and Results



Category	Avg F1	Std Dev	Test-F1
First Party Collection/Use	0.814	0.01	0.801
Third Party Sharing/Collection	0.791	0.02	0.79
User Choice Control	0.692	0.07	0.712
Data Security	0.838	0.01	0.837
Intl and Specific Audiences	0.898	0.03	0.871
Access, Edit and Delete	0.757	0.04	0.823
Policy Change	0.917	0.08	0.875
Data Retention	0.55	0.05	0.58
Do Not Track	0.949	0.07	0.941

Table 2: Feed-forward Network's result on the dev and test set across different data-practices in the OPP-115 corpus.

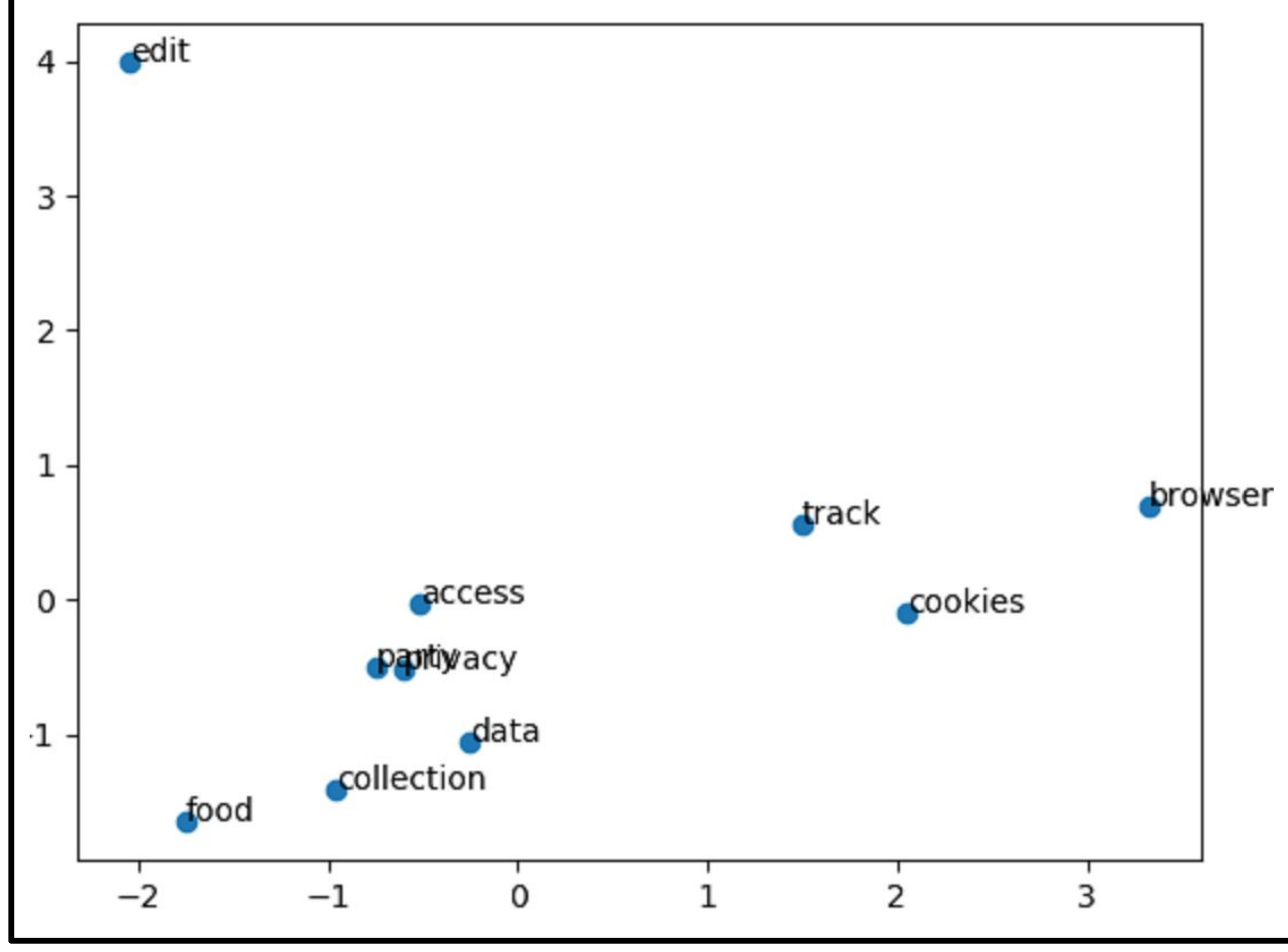
By using in-domain word embeddings we improve the performance of our classifiers.

Macro F1:

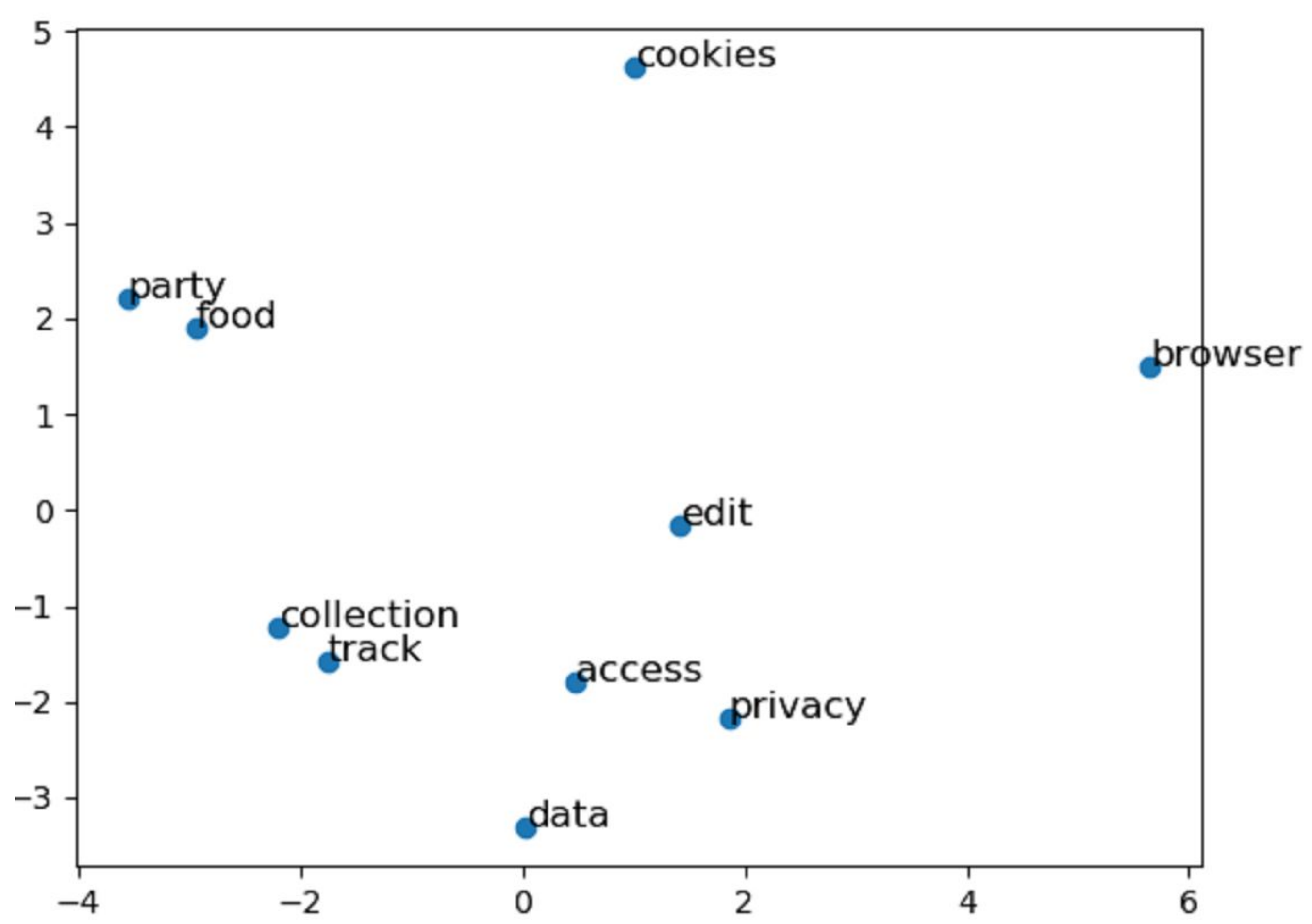
Glove Embeddings(Previous best): 0.76

Fast Text Embeddings: 0.8

Fast Text Word Embeddings



Glove Word Embeddings



Amount of data needed

More training data is needed for the Data Retention class which has a lower number of examples.