

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Problem Statement	2
1.3	Objectives	3
1.4	Roadmap	4
2	Background and Related Work	5
2.1	Classification	5
2.2	Regression	7
2.3	Representations	8
2.4	Neural Networks	9
2.5	Applications in human language	11
2.6	Related work	13
3	Proposed Methodology	15
3.1	Method	15
3.2	Static word embeddings	16
3.3	Contextualized word embeddings	17
4	Experiments	18
4.1	Data	18
4.2	Representation models	19
4.3	Sentiment Experiments	20
4.4	Sentiment Discussion	23
4.5	Emotion Experiments	24
4.6	Emotion Discussion	26
4.7	Qualitative Analysis	26
5	Conclusions and Further Research	29
	Bibliography	31

List of Tables

2.1	Example of contingency table in a binary classification problem.	6
4.2	Emotion datasets properties. Words in NRC lexicon may appear in multiple emotions.	19
4.1	Sentiment datasets properties.	19
4.3	General purpose static embeddings ROC AUC, F1 and Kappa scores from multiple train and test domains	21
4.4	Edinburgh static embeddings ROC AUC, F1 and Kappa scores from multiple train and test domains	21
4.5	BERT contextualized embeddings ROC AUC, F1 and Kappa scores from multiple train and test domains, varying between CLS and AVG representations.	22
4.6	Winner configuration from all models ROC AUC, F1, and Kappa scores from multiple train and test domains. AVG represents BERT model with Average to Average embeddings meanwhile CLS/AVG represents BERT with CLS to Average embeddings.	22
4.7	Anger transfer learning between word and tweet domains using BERT and Edinburgh (Edin) representations. Metrics shown are Pearson Correlation (COR), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE)	24
4.8	Fear transfer learning between word and tweet domains using BERT and Edinburgh (Edin) representations. Metrics shown are Pearson Correlation (COR), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE)	25
4.9	Sadness transfer learning between word and tweet domains using BERT and Edinburgh (Edin) representations. Metrics shown are Pearson Correlation (COR), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE)	25
4.10	Joy transfer learning between word and tweet domains using BERT and Edinburgh (Edin) representations. Metrics shown are Pearson Correlation (COR), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE)	25
4.11	Examples of classification errors of both models in tweet prediction, trained on words.	27
4.12	Examples of classification errors of BERT AVG representation at word prediction, trained on tweets.	27
4.13	Examples of classification errors of Word2Vec Edinburgh representation at word prediction, trained on tweets.	28
4.14	Tweets with the presence of words related to “unhelpful”	28

List of Figures

2.1	ROC curve showing True Positive vs. False Positive rate at different classification thresholds. The area under this curve is known as the ROC AUC score.	7
2.2	Example of two sentences yielding the same bag of words representation. . .	9
2.3	Fully connected network with an input of size 3, a hidden layer of size n , and an output layer of size 2. f and g are the activation functions of neurons. W^1 and W^2 represent the weight matrices. \vec{b}^1 and \vec{b}^2 are the biases of each layer.	10
3.1	Method for transferring sentiment knowledge between words, tweets and movie reviews	16