Project Report

Understanding Semantic Similarity between Sentences

Project	Semantic Textual Similarity
Baseline Model	Siamese Recurrent Architectures for Learning Sentence Similarity + Augment word representations with character level features
Team	Arvin-Sachin Group - Sachin Raja (2020801001), Arvin Goyal (2019900039)
Mentor	Prashant Kodali
Version	Final Report

1. INTRODUCTION

Paraphrasing methods identify, generate, or extract phrases, sentences that convey almost the same information. Different worded sentences may bear the similar meaning and can be identified by paraphrase identification. Paraphrase detection has importance as it contributes to various NLP tasks like Text summarization, Document Clustering, Question Answering, natural language inference, information Retrieval, Plagiarism Detection, Text Simplification.

Paraphrase detection remains a hard problem because of limited availability of annotated data, variable sentences length and complex sentence structures. Earlier vector representation of sentences was done using bag-of-words or TF-IDF based methods which were high dimensional and sparse in nature which made it difficult to learn inherent patterns. Recently, neural embeddings have proved to be effective for paraphrase detection and other semantics based NLP tasks.

Formally, the paraphrase or sentence similarity detection problem as defined as follows: Given a pair of sentences, X (x_1, x_2.....x_{T_a}) and Y (y_1, y_2.....y_{T_b}) where both sentences have different lengths, denoted by T_a and T_b, the goal is to predict a similarity score between these two sentences. This problem is generally tackled as a supervised regression or classification problem.

In order to compute similarity between a pair of objects, generally Siamese style networks are employed. A Siamese neural network is an artificial neural network that uses the same weights while working in tandem on two different input vectors to compute comparable output vectors. In order to compute similarity between two sentences, a vector representation for two-sentences is required. For variable length sentences, recurrent neural network architectures such as RNNs and LSTMs have proved effective for tasks such as text classification and language translation. Another class of architectures used for computing sentence embeddings are transformers. Another aspect of capturing similarity is representation of individual tokens and the third major aspect being the choice of the similarity function. In this work, we evaluate different choices of representations and similarity measures to train and test sentence similarity models, which is based on siamese class of networks.

2. RELATED WORK

At a high level, there are three major dimensions of works in order to understand semantic similarity between sentences - (a) Word-based sentence similarity, (b) Structure-based sentence similarity and (c) Vector-based sentence similarity.

Word similarity method utilizes similarity between sentences' words in order to compute sentence similarity. Some of the methods in this category of solutions include creation of a similarity matrix between all words of sentence 1 and sentence 2 by measuring cosine distance between every pair of word vectors, and then sentence similarity is computed by aggregating information in this matrix using pooling or other similar techniques. Some of the proposed methods in this class include max similarity method, identifying similar and dissimilar parts of the sentence and word sense disambiguation.

Another class of methods utilize sentence structure in order to understand semantic similarity between sentences. These basically use information about grammatical formation of a sentence, part-of-speech analysis and/or word orders to understand sentence similarity. However, the basic assumption behind structure based methods is that people generally use similar structure to represent similar sentences. This however, may not be always true as in many applications, sentences are not well-structured and do not follow grammatical rules.

Most recent methods for computing sentence similarity are vector-based. This involves two steps - generation of a vector representation for a sentence and then using some distance or similarity measure to understand similarity between a pair of sentences.

Different methods in this class primarily use different ways of obtaining sentence vectors including - (a) distributional sentence vectors using matrix factorization, (b) averaging of word vectors such as Glove embeddings, (c) learnt vectors for example, MaLSTM method and (d) skip-thought vectors, which aim at predicting the surrounding sentences given a sentence. Once sentence vectors are obtained, similarity can be computed by using a distance measure on top such as euclidean, manhattan or cosine distances.

Method	WordNet	Word embed-ing	LSA	Word-order	Using struc-ture	Vector based	Trained-net-work	Data set	results
Atish 2018[30]	√			√		√		Pilot	0.837
Distri-butional sentence similarity[38]			√			√	√	MSRP	80.41
Grammar-based[35]	√				√			MSRP Pilot	71.02 --
skip-thought		√				√	√	MSRP SICK	75.8 0.8655
Wang[31]		√				√	√	MSRP	78.4
Max similarity[29]	√		√					MSRP	70.3
Similarity matrix[24]	√					√		MSRP	74.1
WordNet and Word embedding	√	√		√				MSRPPilot	71.6 0.852

3. DATASETS

3.1 MSRPC

Microsoft Research Paraphrase Corpus (MSRPC) is a text file containing 5800 pairs of sentences which have been extracted from news sources on the web, along with human annotations indicating whether each pair captures a paraphrase/semantic equivalence relationship in binary form. Last published: March 3, 2005.

The dataset consists of 5,801 sentence pairs. The average sentence length is 21, the shortest sentence has 7 words while 36 is the longest sentence length. 3,900 are labeled as being in the paraphrase relationship. 4,076 training pairs are present in training file while 1,725 pairs present in test file. All sentences were labeled by two annotators who agreed in 83% of the cases. A third annotator resolved conflicts.

3.2 Semeval STS

Nearly 14 thousand sentence pairs are linked above with gold human annotated STS labels. It contains test sets with gold standard labels for the 2016 Semantic Textual Similarity (STS) shared task.

Each evaluation set gold standard label file has the following format:

- One numeric STS label between 0 and 5 per line, which we plan to model using regression.
- Blank lines indicate that the pair was not included in the official scoring

3.3 SICK

The Sentences Involving Compositional Knowledge (SICK) data set consists of 10,000 English sentence pairs, built starting from two existing paraphrase sets: the 8K ImageFlickr data set (http://nlp.cs.illinois.edu/HockenmaierGroup/data.html) and the SEMEVAL-2012 Semantic Textual Similarity Video Descriptions data set (http://www.cs.york.ac.uk/semeval-2012/task6/index.php?id=data). Each sentence pair is annotated for relatedness in meaning and for the entailment relation between the two elements.

This dataset not only contain scoring labels for semantic relatedness but also for entailments (neutral, entailment or contradiction) for every pair of sentence, which we plan to model using regression.

4. METHODOLOGY

4.1 BASELINE

We use “Siamese Recurrent Architectures for learning Sentence Similarity” by Jonas Mueller and Aditya Thyagarajan as a baseline model for our work. At a high level, their proposed MaLSTM (Manhattan LSTM) model uses two LSTMs to extract features for each of the sentence in a given pair, over which a similarity measure (exp of Manhattan distance between the h’s of two LSTMs) is applied. This model is trained in a supervised manner. We implemented both classification and regression models for the baseline to predict sentence similarity classification results on MSRPC as well as regression scores on SICK datasets.

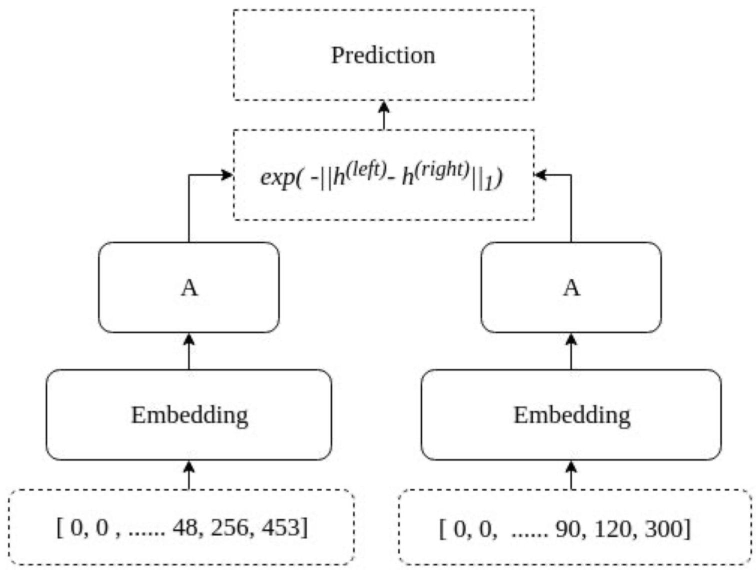
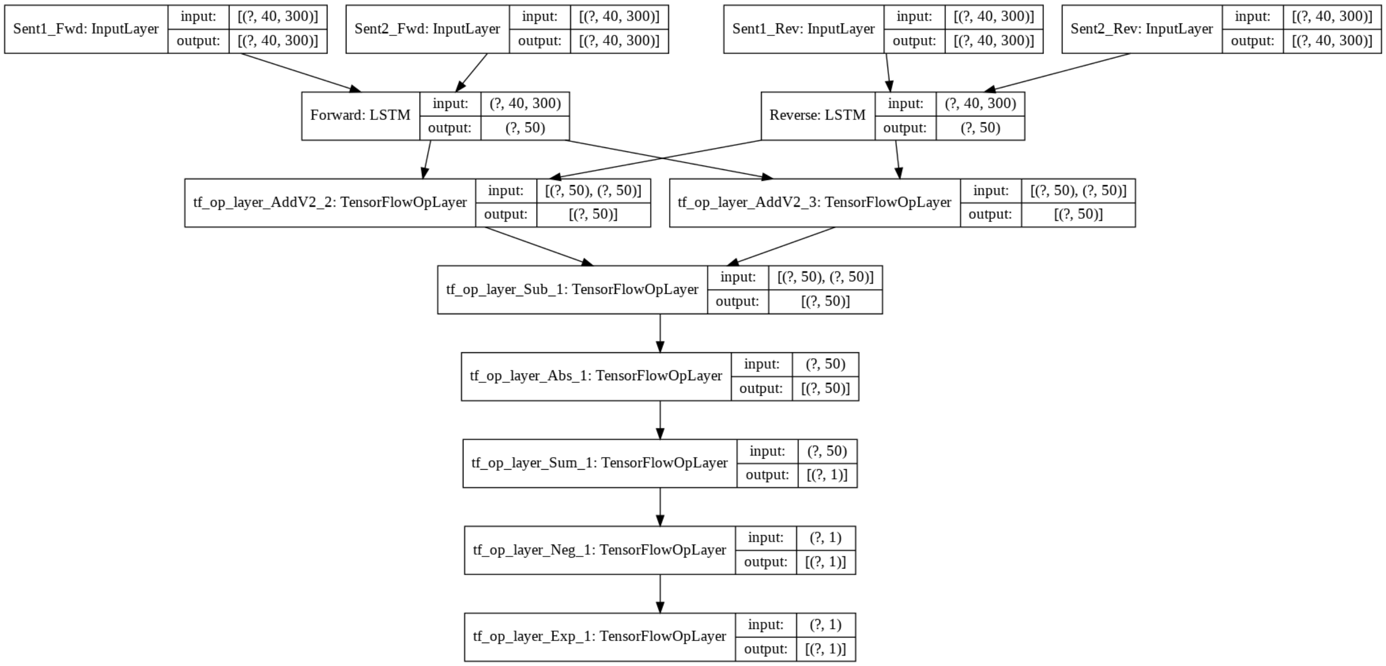


Figure 1: Basic structure of the Siamese neural network. Unit A is changed over the architectures.

In their architecture, authors use a recurrent neural network which learns a mapping from the variable sequence length of word vectors into a fixed dimensional vector which is referred to as a sentence embedding. The same RNN weights are shared to compute the sentence embeddings for both the sentences in the pair. Finally exponential of Manhattan distance between the two vectors is computed and the mean squared error with scaled ground truth (between 0 to 1) is back-propagated to learn the LSTM weights.

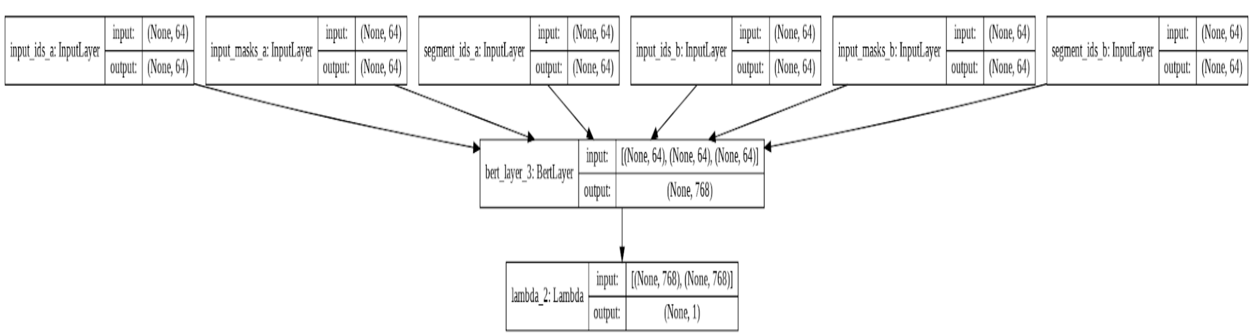
4.2 OUR EXTENSION

For our baseline, we implement the recursive neural network using a Bi-directional LSTM. The architectural diagram is shown below:



In our work, we also implemented euclidean and cosine distance measures to analyze performance. Further, we L2 normalized the sentence vectors before computing the distance. This provided a better spread of the values and demonstrated approximately a 3% increase in Spearman's correlation values on SICK dataset. While the baseline reports best results with Manhattan distance, we obtained the best results with the Euclidean distance measure.

As a next step, we wanted to try out attention mechanism for this problem. In order to achieve that, we decided to employ transformer architecture with self as well as cross attention using BERT model. In order to understand the impact of self attention and cross attention, we tried out two different variants - using parallel bert, where we fed the two sentences independently through the same weights of the BERT layer; (SBERT) and also using a single bert to implement cross-attention by using <CLASS> and <SEP> tokens to separate the two sentences.The architectural diagram is shown below:



From the output of a single BERT, we obtained a vector that captured information about both the sentences collectively. We passed this through a dense vector with a single unit and sigmoid activation to restrict the values between 0 and 1. For backpropagation, we used an SGD optimizer with root mean squared error as the empirical loss function.

In the following section, we describe our quantitative and qualitative results.

5. RESULTS AND ANALYSIS

5.1 MSRPC DATASET

Classification report				
	Precision	Recall	F1-score	Support
0	0.65	0.17	0.26	578
1	0.69	0.95	0.8	1147
Accuracy			0.69	1725
Macro avg	0.67	0.56	0.53	1725
Weighted avg	0.68	0.69	0.62	1725

Confusion Matrix		
	Actual 1	Actual 0
Predicted 1	96	482
Predicted 0	52	1095

Tables 1 and 2 above demonstrate the classification results of the Bi-directional LSTM model using Manhattan distance measure on the MSRPC dataset. One of the reasons that we observe a low true positive score can be attributed to the fact that there is a significant class imbalance between the true class and the false class in terms of number samples.

5.2 SICK DATASET

In the table below, we compare the quantitative results of different techniques for obtaining sentence vectors and three different distance measures to quantify similarity score. We also evaluated on different activation functions to obtain the output.

With respect to sentence vectorization, we used Bi-directional LSTMs and BERT models. Further, within BERT, we experimented with both – self attention only along with self attention + cross attention techniques. For a single BERT with cross attention, we experimented with sigmoid as well as RELU activation functions. Detailed comparison with respect to correlation scores and inference times is shown in the table below:

Model Detail	Pearson Score	Spearman Corr	MSE	1 epoch training time	1500 sample inference time
Manhattan Distance using Unidirectional LSTM (base model)	0.501	0.474	1.278	5.0 sec	0.525 sec
Manhattan Distance + Bidirectional LSTM (base model)	0.826	0.759	0.351	3.0 sec	0.458 sec
Manhattan Distance + Bidirectional LSTM + l2 normalization	0.852	0.795	0.292	3.0 sec	0.505 sec
Euclidean Distance + Bidirectional LSTM + l2 normalization	0.863	0.809	0.268	3.0 sec	0.458 sec
Cosine similarity + Bidirectional LSTM + l2 normalization	0.855	0.802	0.284	2.0 sec	0.436 sec
Two parallel SBERT Embedding(one for each sentence) + Manhattan Distance + norm	0.779	0.695	0.4	912 sec	79.25 sec
Single BERT (cross sentence attention) + Manhattan Distance on output sequence	0.874	0.829	0.253	741 sec	53.09 sec
Single BERT (cross sentence attention) + pooled output + norm + relu + exp	0.887	0.835	0.214	711 sec	57.95 sec
Single BERT (cross sentence attention) + pooled output + norm + sigmoid	0.898	0.848	0.203	790 sec	63.97 sec

From the table above, the importance of self and cross attention becomes evident. It can be seen that addition of self and cross attentions through a transformer-based architecture using BERT results in an improvement of 7.2% and 8.9% Pearson's and Spearman's correlation scores. It can be further observed that just by L2 normalization of the sentence embedding before computing the Manhattan distance further improves the correlation score by around 3 percent. Further, while the baseline paper reports the best results with the use of Manhattan distance measure, we obtained slightly better results with Euclidean distance measure. We attribute this to the L2 normalization of the sentence vectors. However, for the sake of consistency with respect to the original paper, we do a detailed comparative analysis against Manhattan distance measure. Lastly, it can be inferred from the table that due to a significantly large size of the BERT model, the training and inference time also get significantly increased.

5.3 QUALITATIVE ANALYSIS ON SICK DATASET

In this work, we do a four-fold qualitative analysis:

- Comparison of distance measures (L1, L2, Cosine)
- Comparison of various implementations of BERT
- Comparison of BERT against Bi-directional LSTM
- Analysis of failure cases of the best-performing BERT model.

5.3.1 COMPARISON OF DISTANCE MEASURES

Table below demonstrates some examples from SICK dataset where high variance between the predicted values were observed. It can be inferred that for most cases, where differences in predictions were observed, L2 distance predicted the closest value when compared against the ground truth and that too specifically for non-similar sentences.

Sentence A	**Sentence B**	**GT**	**L1 Sim**	**L2 Sim**	**Cos Sim**	Deviation	
			Pred	Pred	**Pred**	Inter-pred	GT
A child is playing with a water spout outdoors and the rest of his family is watching	A man is standing beside a birdcage which is large and colorful	**1.40**	3.42	1.53	1.49	0.90	1.17
Three small dogs are sniffing at something	Butter is being chopped into a container by a man	**1.00**	1.44	1.79	3.11	0.72	1.33
A man is shooting a shotgun	Someone is playing the guitar	**1.00**	1.57	1.40	2.98	0.71	1.21
The sun is shining on the face of the girl who is in a yellow dress	A young black child is standing on the edge of a body of water near some buckets	**2.00**	2.34	3.48	1.81	0.70	0.88
A person is rinsing a steak with water	A man is buttering a slice of bread	**1.70**	2.95	1.26	2.31	0.69	0.84
A woman is taking two eggs out of a pot of water	A person is buttering a tray	**1.80**	3.19	1.57	2.30	0.66	0.86
A bike rider in a black and red uniform is standing on a dirt bike	Five adults are sitting on stone steps	**1.10**	3.09	1.48	2.24	0.66	1.34
A homeless man is holding up a sign and is begging for money	The cat is hungrily drinking milk	**1.00**	2.12	0.65	1.74	0.62	0.80
A man with tattoos is lounging on a couch and is looking for a pencil	A cow is eating hay	**1.00**	2.46	1.09	2.26	0.61	1.12
One boy in orange shorts is standing on a rock cliff over the water and the other boy in black shorts is jumping of it into the water	A baby is sneezing and frightening another baby	**1.00**	2.66	1.22	1.81	0.59	1.07
A group of children in a church basement is playing guitars and tambourines	A group of children in a church basement is playing maracas and tambourines	**4.20**	3.41	4.73	4.56	0.59	0.59
A boy is happily playing the piano	A white bird is landing swiftly in the water	**1.00**	2.82	1.47	1.74	0.58	1.17
A lion is slowly moving around	A person is slicing a carrot into pieces	**1.00**	1.26	2.28	2.63	0.58	1.21
A man in full gear is wearing a helmet with sponsor logos and is riding a red sport motorcycle	A woman is tapping her fingers	**1.00**	2.38	0.98	1.48	0.58	0.84
Two cars for racing are on a road in front of a grassy parking area	A white car and a yellow car are racing down the track	**3.10**	2.90	1.79	3.09	0.57	0.76
A group of children in a church basement is playing guitars and tambourines	A group of children is playing tambourines	**4.00**	3.02	4.25	4.20	0.57	0.59
The military officer is shouting at the recruits	A rollerblader is performing a trick on a ramp	**1.00**	2.92	1.59	1.94	0.56	1.28
A fearful little boy is on a climbing wall	The man is sitting next to a birdcage	**1.10**	2.91	1.65	1.91	0.55	1.19
There is no man cutting a box	A homeless man is holding up a sign and is begging for money	**1.10**	2.82	1.50	2.04	0.54	1.16
A barefoot man in pajamas is looking toward the sky and is standing on the tennis court	A man is wearing a blue shirt and walking barefoot on a tennis court	**3.40**	2.69	3.56	3.99	0.54	0.54

A man is skateboarding on a half pipe	The kids are playing outdoors near a man with a smile	**1.60**	3.28	2.44	2.00	0.53	1.11
A shirtless man is escorting a horse that is pulling a carriage along a dirty road	A shirtless man is leading a horse that is pulling a carriage	**4.30**	3.00	3.95	4.25	0.53	0.78
A dog is running on the sand and chasing a ball	A brown and white dog is biting a dirty tennis ball in a dirt field	**3.40**	3.39	3.93	2.64	0.53	0.54
A man is wearing a blue shirt and walking barefoot on a tennis court	A barefoot man in pajamas is looking toward the sky and is walking on the tennis court	**3.60**	2.96	3.99	4.15	0.53	0.54
A little girl is smiling and running outside	The surfer is riding a small wave	**1.00**	2.81	2.44	1.56	0.52	1.37
The man is singing heartily and playing the guitar	A bicyclist is holding a bike over his head in a group of people	**1.00**	1.78	1.28	2.55	0.52	1.02
A woman is playing with a brown dog on a garden path	A woman is playing with a brown cat on a garden path	**3.50**	2.99	4.26	3.61	0.52	0.53
Two men are fighting	A couple of boys are playing a video game	**1.30**	2.75	1.53	1.83	0.52	0.90
A big brown and white spotted dog is lying on a jacket on the street	The dog is having a rest in the living room	**2.70**	3.24	2.64	1.98	0.52	0.52
A man is wearing a blue shirt and walking barefoot on a tennis court	A barefoot man in pajamas is looking toward the stars and is walking on the tennis court	**3.20**	3.16	4.10	4.36	0.51	0.84
A shirtless man is escorting a horse that is pulling a carriage along a paved road	A shirtless man is leading a carriage that is being pulled by a horse	**4.60**	3.10	4.19	4.15	0.51	0.94
A badger, which is shrewd, is digging the earth	A woman is slicing tofu	**1.00**	1.70	2.68	2.84	0.50	1.50
A cat is playing keyboards	There are no men fighting	**1.10**	1.91	0.75	1.66	0.50	0.61
A bike is being ridden over a monkey	A person is riding a dirt bike down a dirt hill	**2.80**	2.42	3.61	2.76	0.50	0.52
Someone is playing the guitar	A shirtless woman is jumping over a log	**1.00**	1.38	1.59	2.53	0.50	0.97
The man is doing a wheelie on a mountain bike	A man is stopping in the middle of a road	**1.80**	3.19	1.98	2.48	0.50	0.90
There is no man playing a game on the grass	A man is playing the guitar	**1.70**	2.02	3.19	2.33	0.49	0.95
A black and brown cat is eyeing a fly	The man is eating cereal	**1.00**	1.34	1.53	2.46	0.49	0.92
The girl is cheering the man in the blue and white uniform	A soccer player is being tackled by his opponent	**1.19**	2.76	1.57	2.21	0.49	1.11
A cat is playing with a watermelon	A woman is looking at the view of a city	**1.00**	1.41	1.87	2.59	0.48	1.07
A shirtless man is escorting a horse that is pulling a carriage along a paved road	A shirtless man is leading a horse that is pulling a carriage	**4.40**	3.13	4.09	4.19	0.48	0.77
People are riding two camels on the sand	Two people are seated on a camel and another camel is in the foreground	**3.90**	3.57	2.95	2.41	0.48	1.04
The boy is playing in the mud	A little boy is topless and is serious	**1.60**	2.81	3.65	2.53	0.48	1.47
A large brown dog is jumping over a red hurdle	A young man is running away from the fishing line	**1.00**	2.92	1.80	2.15	0.47	1.37
A horse is standing	Someone is holding a hedgehog	**1.00**	2.97	1.82	2.34	0.47	1.45
A man is putting garlic on some bread slices	A person is bowling the ingredients to the man at the mixer	**2.20**	2.44	1.94	3.07	0.46	0.54

A man is mowing a lawn	A black and white dog is playing with a tattered volleyball in a brown field	**1.00**	2.53	1.41	1.94	0.46	1.06
A white car is being driven by the man	A person playing football is running past an official carrying a football	**1.00**	2.23	1.14	1.69	0.45	0.82
A young man is sitting on a bench	An elderly woman is sitting on a bench and is wearing a gray jacket and black pants	**2.80**	3.18	2.09	2.60	0.45	0.48
A woman is eating a cupcake	A man is slicing a potato	**1.30**	3.04	2.08	2.10	0.45	1.19

5.3.2 COMPARISONS OF BERT IMPLEMENTATIONS

Table below shows some qualitative examples of using self attention vs using self attention + cross attention. It can be inferred from the table that incorporating cross attention with a sigmoid activation function gives the best results in terms of mean squared error as well as in terms of correlation scores.

Sentence A	**Sentence B**	**GT**	**SBERT Pred**	**BERT Seq L1 dist Pred**	**BERT Pool Sig Pred**	**Deviation**	
						Inter-pred	**GT**
Two crocodiles are floating in a green colored swimming pool near some playing kids	Two poodles are in the snow and one is jumping high	**1.10**	3.56	1.80	1.36	**0.95**	**1.48**
The brown dog is playing outdoors	Two people are arguing near a crowd	**1.00**	3.58	1.91	1.36	**0.94**	**1.59**
Two dogs are playing on the beach	Two people are driving a jeep and a lady is sitting on the top of it	**1.00**	3.52	1.88	1.37	**0.92**	**1.56**
Two men are walking through the grass	Two men are standing in deep water	**2.40**	3.82	2.12	1.71	**0.91**	**0.93**
The people are walking on the road beside a beautiful waterfall	There is no brown dog and black dog playing in the sand	**1.00**	3.52	1.81	1.42	**0.91**	**1.55**
A large brown dog is jumping over a red hurdle	A young man is running away from the fishing line	**1.00**	3.39	1.81	1.37	**0.87**	**1.47**
A person is kicking a soccer ball between their feet	One brown and black dog is running through the leaves	**1.00**	3.37	1.81	1.36	**0.86**	**1.46**
There is no girl playing a flute	Two baby pandas are playing	**1.00**	3.39	1.90	1.37	**0.86**	**1.49**
A woman is cutting a fish	The man is slicing potatoes	**2.00**	3.38	1.88	1.37	**0.85**	**0.88**
A boy is standing in a room by a lamp light	A small group of people are standing and two are sitting on the couch	**2.20**	3.59	1.84	1.73	**0.85**	**0.87**
A person is rinsing a steak with water	A man is buttering a slice of bread	**1.70**	3.53	1.87	1.65	**0.84**	**1.06**
A woman is surfing	A fearful little boy is on a climbing wall	**1.00**	3.36	1.96	1.36	**0.84**	**1.48**
A monkey is brushing the dog	One man is jumping off a rock wall and another man is dropping a rope	**1.00**	3.33	1.86	1.36	**0.83**	**1.45**
Two kids are doing martial arts on a blue mat	A tan dog is running through the brush	**1.00**	3.30	1.82	1.36	**0.83**	**1.43**

A dog is playing on the green grass	The man is not playing a guitar	**1.00**	3.31	1.85	1.37	**0.82**	**1.44**
A child is smiling at the camera and swimming underwater	An animal is barking at a ball	**1.10**	3.29	1.82	1.36	**0.82**	**1.34**
The woman is picking up the kangaroo	A man is standing on the top of a roof and playing a violin	**1.00**	3.28	1.84	1.36	**0.82**	**1.42**
Mimes are performing on a stage	No little dog is running on the sand	**1.00**	3.28	1.88	1.36	**0.81**	**1.43**
A little girl and a woman wearing a yellow shirt are getting splashed by a city fountain	A woman is playing the trumpet	**1.70**	3.26	1.82	1.37	**0.81**	**0.92**
A dog in a colored coat is running across the yard	The flute is being played by one man	**1.00**	3.25	1.81	1.37	**0.81**	**1.40**
Two men are fighting	A couple of boys are playing a video game	**1.30**	3.32	1.96	1.40	**0.81**	**1.23**
A young man is pushing a motocross bike down a dirt hill	A dog is swimming after a tennis ball	**1.00**	3.24	1.79	1.37	**0.80**	**1.39**
Little kids are playing in a water fountain in front of few people	A monkey is brushing the dog	**1.30**	3.24	1.82	1.37	**0.80**	**1.16**
A man is singing to a woman	A funny crowd of people is doing a skit of two boxing men	**1.30**	3.25	1.88	1.36	**0.79**	**1.17**
A girl is standing in a group and is wearing a black shirt and pink beads	A small white dog is jumping up in the snow	**1.10**	3.21	1.77	1.37	**0.79**	**1.29**
The man is playing a guitar	There is no hockey player in a yellow jersey guarding the goal	**1.10**	3.26	1.94	1.37	**0.79**	**1.35**
Three children are running up hill	A pet dog is standing on the bank and is looking at another brown dog in the pond	**1.00**	3.24	1.87	1.38	**0.79**	**1.40**
A man is in a parking lot and is playing tennis against a large wall	The snowboarder is leaping fearlessly over white snow	**1.00**	3.21	1.80	1.36	**0.79**	**1.37**
A great singer is dancing on the ceiling	A child is making a snow ball	**1.00**	3.21	1.84	1.36	**0.78**	**1.38**
A man is sitting on the grass and drinking from a water bottle	The child is delightedly playing with toys outdoors	**1.10**	3.19	1.81	1.36	**0.78**	**1.28**
A girl is wearing a t-shirt and has her mouth open.	A brown and white dog is holding a baseball in its mouth	**1.60**	3.18	1.79	1.37	**0.77**	**0.93**
A man is mowing a lawn	A black and white dog is playing with a tattered volleyball in a brown field	**1.00**	3.17	1.82	1.36	**0.77**	**1.36**
A couple is running towards the ocean	A snowboarding man is jumping through the air	**1.30**	3.16	1.86	1.36	**0.76**	**1.12**
A woman is taking two eggs out of a pot of water	A person is buttering a tray	**1.80**	3.60	1.80	2.35	**0.75**	**1.09**
Two children are playing soccer in the park	There is no white dog wearing a Christmas reindeer headband and playing with a brown dog in the grass	**1.10**	3.30	1.86	1.59	**0.75**	**1.37**
A fearful little boy is on a climbing wall	The man is sitting next to a birdcage	**1.10**	3.19	1.83	1.44	**0.75**	**1.29**
A dog, which is small, is playing on the green grass	A woman wearing a blue shirt and high heels is standing on the sidewalk next to the man	**1.00**	3.15	1.92	1.37	**0.75**	**1.37**
A family is buying something at the vending machine	A man in a suit is standing at a microphone and singing	**1.00**	3.12	1.82	1.36	**0.74**	**1.33**
A dog is near a ball colored in red, which is in the air	A man is playing a guitar	**1.00**	3.13	1.89	1.37	**0.74**	**1.35**

The man is playing the drums	The black bird is sitting in a leafless tree	**1.00**	3.12	1.88	1.36	**0.74**	**1.34**
A woman is looking into the distance and people are walking between buildings behind	A man is playing a flute	**1.10**	3.13	1.87	1.38	**0.74**	**1.26**
A boy is happily playing the piano	A white bird is landing swiftly in the water	**1.00**	3.09	1.85	1.35	**0.73**	**1.32**
A person is dancing on a roof	A white dog is standing on fallen leaves	**1.20**	3.11	1.84	1.41	**0.72**	**1.17**
A woman, who is shoeless, is sitting on a blanket under a lavender umbrella	A small dog is lying on the bed	**1.00**	3.06	1.85	1.37	**0.71**	**1.31**
The people are standing at a parade	A white bird is landing swiftly in the water	**1.10**	3.07	1.88	1.36	**0.71**	**1.23**
A fish is being sliced by a man	A cat is jumping into a box	**1.00**	3.04	1.82	1.36	**0.71**	**1.29**
A baby is sneezing and scaring another baby	A young child with black hair is deleting a picture from the camera	**1.00**	3.05	1.87	1.36	**0.70**	**1.30**
Different things from a war are being shown to some people by a veteran	Three friends are making faces for the camera	**1.40**	3.08	2.07	1.38	**0.70**	**1.05**
A dog with golden fur is in the water	A woman is talking on a telephonic device	**1.00**	3.02	1.85	1.37	**0.70**	**1.28**
A woman is bowling two eggs to a break dancer	A baby is playing with a doll	**1.00**	3.01	1.81	1.37	**0.69**	**1.27**

5.3.3 COMPARISON OF BI_DIRECTIONAL LSTM AND BERT WITH CROSS-ATTENTION

It can be inferred from the table below that BERT with cross attention and sigmoid activation performs better with respect to reducing false positives and for relatively longer sequences in the dataset.

Sentence A	**Sentence B**	**GT**	**Manh Pred**	**BERT Pool Sig Pred**	**Deviation**	
					Inter-pred	**GT**
Different things from a war are being shown to some people by a veteran	Three friends are making faces for the camera	**1.40**	3.51	1.38	1.06	1.49
A person is scrubbing a zucchini	The woman is cutting cooked octopus	**1.19**	3.30	1.39	0.95	1.50
A child is playing with a water spout outdoors and the rest of his family is watching	A man is standing beside a birdcage which is large and colorful	**1.40**	3.42	1.56	0.93	1.43
A boy is standing next to the opening of a fountain	The man is standing on a rocky mountain and gray clouds are in the background	**1.40**	3.25	1.45	0.90	1.31
Two crocodiles are floating in a green colored swimming pool near some playing kids	Two poodles are in the snow and one is jumping high	**1.10**	3.11	1.36	0.88	1.44
A man is skateboarding on a half pipe	The kids are playing outdoors near a man with a smile	**1.60**	3.28	1.59	0.85	1.19

A bike rider in a black and red uniform is standing on a dirt bike	Five adults are sitting on stone steps	**1.10**	3.09	1.40	0.85	1.42
A woman is eating a cupcake	A man is slicing a potato	**1.30**	3.04	1.37	0.83	1.23
A group of children in a church basement is playing maracas and tambourines	Tambourines are being played by a group of children	**4.60**	2.84	4.46	0.81	1.25
A horse is standing	Someone is holding a hedgehog	**1.00**	2.97	1.37	0.80	1.42
A man in blue has a yellow ball in the mitt	A man is jumping rope outside	**1.20**	3.04	1.47	0.79	1.32
A surfer is riding the wave	There is no man tying a shoe	**1.00**	2.94	1.37	0.78	1.40
Two people are riding a bike	A practicing jumper is tossing a person's snowboard into the air	**2.00**	2.93	1.37	0.78	0.80
The military officer is shouting at the recruits	A rollerblader is performing a trick on a ramp	**1.00**	2.92	1.36	0.78	1.38
A large brown dog is jumping over a red hurdle	A young man is running away from the fishing line	**1.00**	2.92	1.37	0.78	1.39
A boy is happily playing the piano	A white bird is landing swiftly in the water	**1.00**	2.82	1.35	0.74	1.31
A fearful little boy is on a climbing wall	The man is sitting next to a birdcage	**1.10**	2.91	1.44	0.73	1.30
A shirtless man is escorting a horse that is pulling a carriage along a dirty road	A shirtless man is leading a horse that is pulling a carriage	**4.30**	3.00	4.45	0.73	0.93
A little girl is smiling and running outside	The surfer is riding a small wave	**1.00**	2.81	1.36	0.72	1.30
There is no man cutting a box	A homeless man is holding up a sign and is begging for money	**1.10**	2.82	1.39	0.72	1.24
Two men are walking through the grass	Two men are standing in deep water	**2.40**	3.14	1.71	0.71	0.71
A shirtless man is escorting a horse that is pulling a carriage along a paved road	A shirtless man is leading a horse that is pulling a carriage	**4.40**	3.13	4.50	0.68	0.90
The man is using a sledgehammer to break a concrete block that is on another person	A man and a woman are walking together through the woods	**1.00**	2.74	1.38	0.68	1.26
A shirtless man is escorting a horse that is pulling a carriage along a paved road	A shirtless man is leading a carriage that is being pulled by a horse	**4.60**	3.10	4.46	0.68	1.07
Two men are fighting	A couple of boys are playing a video game	**1.30**	2.75	1.40	0.67	1.03
A group of children in a church basement is playing guitars and tambourines	A group of children is playing tambourines	**4.00**	3.02	4.35	0.67	0.74
A man is in a parking lot and is playing tennis against a large wall	The snowboarder is leaping fearlessly over white snow	**1.00**	2.68	1.36	0.66	1.21
A boy is playing a game with wooden blocks	Some cheerleaders are dancing	**1.00**	2.67	1.36	0.66	1.21
A brown and white dog is running through the river	An animal is emerging from a lake	**2.50**	3.59	2.29	0.65	0.79
One boy in orange shorts is standing on a rock cliff over the water and the other boy in black shorts is jumping of it into the water	A baby is sneezing and frightening another baby	**1.00**	2.66	1.36	0.65	1.20
A person is rinsing a steak with water	A man is buttering a slice of bread	**1.70**	2.95	1.65	0.65	0.88
People are sitting against a wall	A woman is wearing a blue shirt with a white vest and a white cap and is talking and marching	**1.40**	2.74	1.45	0.64	0.95

A man is climbing a rope	There is no man holding a mask in his raised hand	**1.20**	2.65	1.38	0.63	1.03
Two children are stretching over some metal bars	A bike rider in a black and red uniform is standing on a dirt bike	**1.30**	2.67	1.42	0.63	0.97
A boy is wearing an orange shirt and a striped tie	A girl in an orange shirt and clown makeup is standing in a park and others are looking on	**1.60**	3.12	1.88	0.62	1.09
A woman is cutting a fish	The man is slicing potatoes	**2.00**	2.60	1.37	0.62	0.62
A laughing child is holding a water gun and getting sprayed with water	A person in a blue jacket is jumping onto a tall cement wall	**1.10**	2.61	1.40	0.60	1.09
A man is not playing a guitar	A famous singer is dancing on the ceiling	**1.80**	2.56	1.36	0.60	0.62
Mimes are performing on a stage	No little dog is running on the sand	**1.00**	2.55	1.36	0.60	1.13
A great singer is dancing on the ceiling	A child is making a snow ball	**1.00**	2.55	1.36	0.59	1.12
A little boy is sticking his tongue out for the camera and another boy is looking on	Two kids are looking up at the camera and one is sticking out his tongue	**4.40**	3.04	4.22	0.59	0.97
A man is mowing a lawn	A black and white dog is playing with a tattered volleyball in a brown field	**1.00**	2.53	1.36	0.58	1.11
Two men are doing a skit, which is very funny, with a boxer in front of a crowd of people	A rabbit is playing with a toy rabbit	**1.00**	2.58	1.42	0.58	1.15
A woman is playing with a brown dog on a garden path	A woman is playing with a brown cat on a garden path	**3.50**	2.99	4.15	0.58	0.58
A person is reading the email	A person with a green shirt is jumping high over the grass	**1.10**	2.53	1.37	0.58	1.03
The girl is carrying a sign and a group of people is following her	A woman is cleaning a man's face	**1.00**	2.51	1.36	0.58	1.10
The people are standing at a parade	A white bird is landing swiftly in the water	**1.10**	2.52	1.36	0.58	1.02
A little boy is sticking his tongue out for the camera and another boy is looking on	Two young boys are looking up at the camera and one is sticking out his tongue	**3.90**	3.10	4.24	0.57	0.62
A baby is sneezing and scaring another baby	A young child with black hair is deleting a picture from the camera	**1.00**	2.51	1.36	0.57	1.10
A woman is surfing	A fearful little boy is on a climbing wall	**1.00**	2.50	1.36	0.57	1.09

5.3.4 FAILURE CASES

Sentence A	**Sentence B**	**GT**	**BERT Pool Sigmoid**	
			Pred	**Error**
The snowboarder is jumping off a snow covered hill	The person is not standing on white ice	**1.50**	3.24	1.74

An elder man is sitting on a bench and is wearing a gray jacket and black pants	An elderly man is sitting on a bench	**4.50**	2.98	1.52
There is no woman using a sewing machine	A girl is brushing her hair	**1.10**	2.43	1.33
A dog is standing on concrete and is holding a blue ball	The large dog is walking outside and is carrying a colorful toy in its mouth	**4.10**	2.81	1.29
A man is petting a dog near a stone path	A woman is playing with a brown dog on a garden path	**1.90**	3.15	1.25
A man is performing a jump on a bicycle	A bicyclist is performing a trick over a heavily graphitized wall	**3.60**	2.36	1.24
A biker is riding away from a fence	A man is dancing on the road	**1.20**	2.42	1.22
The man is doing exercises	A person is putting meat into a skillet	**1.10**	2.32	1.22
A cyclist is performing a jump on a bicycle	A bicyclist is performing a trick over wall full of graffiti	**3.80**	2.60	1.20
The boy is playing in the mud	A little boy is topless and is serious	**1.60**	2.75	1.15
A brown dog is gnawing a metallic post that is stuck in the ground	A brown puppy is biting a pole	**3.80**	2.65	1.15
A young girl is playing on the edge of a fountain and an older woman is watching her	A young girl is playing on the edge of a fountain and an older woman is not watching her	**2.80**	3.92	1.12
Two dogs are playing outside	A dog is running through a field and is chasing a ball	**2.70**	3.76	1.06
A bicyclist is performing a trick over a clean wall	A cyclist is performing a jump on a bicycle	**4.40**	3.37	1.03
A brown puppy is biting a stick	A brown puppy is gnawing a metallic post that is stuck in the ground	**2.90**	3.92	1.02
A brown puppy is gnawing a metallic post that is stuck in the ground	A brown puppy is biting a pole	**4.50**	3.49	1.01
A black and white dog is carrying a huge stick on the green grass	A black and white dog with a large branch is running in the field	**4.50**	3.52	0.98
Two people are seated on a camel and another camel is in the foreground	People are riding two camels at the beach	**4.10**	3.15	0.95
A dog is fetching a stick out of very clear water	A small black and white dog is biting a stick and is swimming	**3.80**	2.85	0.95
The orange rider is not driving a motorcycle on one wheel	The orange rider is driving a motorcycle on one wheel	**4.50**	3.56	0.94

6. CONCLUSION

Measuring sentence similarity between a pair of sentences has been one of the most challenging tasks in NLP. Our work primarily highlights the importance of vector normalization, choice of distance measure and importance of self and cross-attention models to obtain vector representation of sentences eventually used for paraphrase detection.

6. References

Siamese Recurrent Architectures for Learning Sentence Similarity
Jonas Mueller (Computer Science & Artificial Intelligence Laboratory Massachusetts Institute of Technology)
Aditya Thyagarajan (Department of Computer Science and Engineering M. S. Ramaiah Institute of Technology)

Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks
Nils Reimers and Iryna Gurevych
Ubiquitous Knowledge Processing Lab (UKP-TUDA), Department of Computer Science, Technische Universit¨at Darmstadt

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding
Jacob Devlin, Ming-Wei Chang, Kenton Lee and Kristina Toutanova (Google AI Language)

GloVe: Global Vectors for Word Representation
Jeffrey Pennington, Richard Socher, Christopher D. Manning
Computer Science Department, Stanford University, Stanford, CA 94305

Agirre, E., and Cer, D. 2013. SEM 2013 shared task: Semantic Textual Similarity. SemEval 2013.

Long Short-Term Memory 1997
Hochreiter, S., and Schmidhuber (Neural Computation 9(8):1735–1780.)

7. LINK TO CODE AND SAVED MODELS

https://drive.google.com/drive/folders/1D7I8iCLtLbLMxCTqN4O1naH9dGFse9Zs?usp=sharing