Semantic similarity in Q&A using Deep learning techniques

Master's Thesis Project report submitted

in partial fulfillment for the award of the degree of

Masters of Technology

in

Computer Science and Engineering

by

Sandesh C

(12CS30041)

Under the supervision of

Professor Pawan Goyal

Department of Computer Science and Engineering

IIT Kharagpur

Spring Semester, 2016-17

April 28, 2017

DECLARATION

I certify that

(a) The work contained in this report has been done by me under the guidance of my supervisor.

(b) The work has not been submitted to any other Institute for any degree or diploma.

(c) I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

(d) Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the thesis and giving their details in the references. Further, I have taken permission from the copyright owners of the sources, whenever necessary.

Date: April 28, 2017 (Sandesh C)
Place: Kharagpur (12CS30041)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

IIT KHARAGPUR

KHARAGPUR - 721302, INDIA

CERTIFICATE

This is to certify that the project report entitled "**Semantic similarity in Q&A using Deep learning techniques**" submitted by **Sandesh C** (Roll No. 12CS30041) to IIT Kharagpur towards partial fulfillment of requirements for the award of degree of Masters of Technology in Computer Science and Engineering is a record of bona fide work carried out by him under my supervision and guidance during Spring Semester, 2016-17.

Professor Pawan Goyal
Department of Computer Science and Engineering
IIT Kharagpur
Kharagpur - 721302, India

Date: April 28, 2017
Place: Kharagpur

Abstract

Community Question Answering (CQA) forums have since long been plagued with the problem of answer reranking, to automate the process of finding good comments to a question. Here in this work, we take up the problem of Question-Comment similarity with a simple approach where a question-comment pair is represented as concatenation of: distributed paragraph vector representations of question text, comment text and centroidal comment (of that question), along with various syntactic and metadata features. A multilayer perceptron is used to compute the similarity scores for such a question-commnet pair.

Despite it's simplicity the model attains competitive results compared to the best submissions at SemEval '16 Task 3 - Community Question Answering (Subtask A).

Acknowledgements

First and foremost I would like to thank my guide, Prof. Pawan Goyal, for providing an opportunity to work on a challenging and relevant problem, that has blossomed to be a great learning experience. I also thank Prof. Pawan Goyal, for the constant support and guidance over the course of the project.

I would like to thank our faculty advisor, Prof. Rajat Subhra Chakraborty, for his continued efforts to help and enrich the academic experience our batch on various occasions. Also, to acknowledge, the Department of Computer Science and Engineering, IIT Kharagpur, has been helpful through providing easy access to substantial resources for computation.

Finally, I would like to thank my parents, family and friends for their continued belief and support throughout the course of the programme.

Contents

List of Tables

Abbreviations

CQA	**C**ommunity **Q**uestion **A**nswering
QL	**Q**atar **L**iving
SGD	**S**tochastic **G**radient **D**escent
PV	**P**aragraph **V**ector
CBOW	**C**ontinuous **B**ag-**O**f-**W**ords
DM	**D**istributed **M**emory
DBOW	**D**istributed **B**ag-**O**f-**W**ords
MAP	**M**ean **A**veraged **P**recision
MRR	**M**ean **R**eciprocal **R**ate
AvgRec	**Av**era**g**e **Rec**all
P	**P**recision
R	**R**ecall
Acc	**Acc**uracy

Introduction

CQA forums such as Stack Overflow[1] and Qatar Living[2], are gaining popularity online. These forums are seldom moderated, quite open, and thus they typically have little restrictions, if any, on who can post and who can answer a question. On the positive side, this means that one can freely ask any question and can then expect some good, honest comments. On the negative side, it takes effort to go through all possible comments and to make sense of them. For example, it is not unusual for a question to have hundreds of comments, which makes it very time-consuming for the user to inspect and to winnow through them all. The present work is intended to automate the process of finding good comments to questions in a community-created discussion forum, by automatically ranking the existing comments.

SemEval Tasks[3] (Semantic Evaluation) are an ongoing series of evaluations of computational semantic analysis systems. The **SemEval Task 3** in particular deals with semantic comparison for words and texts in the domain of Community Question Answering (CQA). In essence, the main CQA task can be defined as follows: *"given (i) a new question and (ii) a large collection of question-comment threads created by a user community, rank the comments that are most useful for answering the new question"*.

[1]https://stackoverflow.com/
[2]http://www.qatarliving.com/
[3]http://alt.qcri.org/semeval2017/

Subtask A – Question-Comment Similarity. In this project we address the CQA task of finding question-comment similarity, by exploiting the semantic similarity in Q&A using Deep learning techniques. In particular we focus on a single subtask under **SemEval - Task 3**, namely the **Subtask A**, defined below.

Subtask A *Given a question from a question-comment thread, rank the comments as per their relevance (similarity) with respect to the question.*

The thesis is further organized as follows: Chapter 2 addresses the recent works pertaining to the task of finding question-comment similarity, and also sheds light on a few such works that use Deep Learning methodologies to solve this problem; Chapter 3 then provides a detailed explanation of the multilayer perceptron based prediction model approach adopted in our work, using distributed document representations and various syntactic, metadata features. Finally, Chapter 4 tabulates the results obtained with our approach on SemEval '16 Task 3 - Subtask A dataset, which we shall see are competitive with the best results published at the same event.

Literature Survey

The tasks falling under the Community Question & Answering section of SemEval goes in the direction of passage reranking, where automatic classifiers are normally applied to pairs of questions and comment passages to derive a relative order between passages. In other words, this is the task of Answer re-ranking.

In recent years, many advanced models have been developed for automating answer selection, producing a large body of work. For instance, Wang et al. (2007) proposed a probabilistic quasi synchronous grammar to learn syntactic transformations from the question to the candidate answers; Heilman and Smith (2010) used an algorithm based on Tree Edit Distance (TED) to learn tree transformations in pairs; Wang and Manning (2010) developed a probabilistic model to learn tree-edit operations on dependency parse trees; and Yao et al. (2013) applied linear chain CRFs with features derived from TED to automatically learn associations between questions and candidate answers. One interesting aspect of the above research is the need for syntactic structures; this is also corroborated in [Severyn and Moschitti (2012); Severyn and Moschitti (2013)]. Note that answer selection can use models for textual entailment, semantic similarity, and for natural language inference in general.

Quite a few works in this domain have also started to adopt Deep Learning Techniques to solve the problem of answer re-ranking. For eg. Lin and Wang (2015) treated the answer selection task as a sequence labeling problem and proposed recurrent convolutional neural networks to recognize good comments. In a follow-up work, Zhou et al. (2015) included long-short term memory (LSTM) units in their convolutional neural network to learn the classification sequence for the thread. In parallel, Barrón-Cedeno et al. (2015) exploited the dependencies between the thread comments to tackle the same task. This was done by designing features that look globally at the thread and by applying structured prediction models, such as Conditional Random Fields Lafferty et al. (2001).

Notably, at SemEval '16 Task 3 – Subtask A, whose dataset has been used for training and evaluation in this project, provided a great set of Tree Kernel based approaches, that proved to give best results, out performing various LSTM based approaches. Although the reason for this could be for the lack of substantial data for an LSTM to be trained over.

Filice et al. (2016), being the best submission at SemEval '16 Task 3 – Subtask A, gave an SVM learning algorithm that operates on a linear combination of kernel functions, each one applied over a specific representation of the targeted examples: (i) feature vectors containing linguistic similarities between the texts in a pair; (ii) shallow syntactic trees that encode the lexical and morpho-syntactic information shared between text pairs; (iii) feature vectors capturing task-specific information. On similar lines, Joty et al. (2016) trained a binary SVM classifier using partial tree kernels defined over shallow syntactic trees, along with few other metadata and ranking features.

Mihaylov and Nakov (2016) on the other hand constructs various syntactic and metadata features and trains an SVM to the solve the classification problem of comment relevance. This idea of using various helpful features from metadata and constructing scores for the similarity between the syntax of question and comment text, is quite simplistic and has been adopted in our approach as well. Very similar to this approach, Mihaylova et al. (2016) builds variety of features – like question and comment metadata, question and comment lexical features, distance measures between the question and the comment, text readability measures applied to the question and to the comment, lexical semantics vectors for the question and for the comment, features modeling the likelihood of a user being a troll and so on – to be trained on an SVM classifier. Notably, both these approaches build word embeddings as a means to understand the sematics of the question/comment text.

Note that Filice et al. (2016), Joty et al. (2016), Mihaylov and Nakov (2016), Mihaylova et al. (2016) were top submissions at SemEval '16 Task 3 – Subtask A. AlessandroMoschitti et al. (2016) provide the task description and evaluation results paper published for SemEval '16 Task 3, which is a detailed entry on the task, data and submissions at the event.

Building on the success of previous attempts, we shall tread in the direction of exploring Deep Learning Techniques to effectively solve the problem of finding Question - Comment similarity. More specifically we intend to exploit the lexical, semantic similarities between question-comment pairs and information in metadata, all the while designing a simple approach. Note that we use the terms relevant-comment and answer interchangeably thoughout the document.

Approach

For this task, we adopt a neural approach to open-domain non-factoid QA developed by Bogdanova and Foster (2016); which focused on “answer re-ranking”, i.e. given a list of candidate answers to a question, order the answers according to their relevance to the question. This approach was simplistic owing to the fact that it required no feature engineering. Question-answer pairs are represented by concatenated distributed representation vectors and a multilayer perceptron is used to compute the score for an answer (the probability of an answer being the best answer to the question). Despite its simplicity, their work achieved state-of-the-art performance on the Yahoo! Answers dataset of manner or How questions introduced by Jansen et al. (2014). This improved performance was attributed to the use of paragraph vector representations instead of averaging over word vectors, and to the use of suitable data for training these representations. This project aims at improving the simplistic model proposed by Bogdanova and Foster (2016) with a few enhancements to achieve state-of-art performance at the **SemEval Task 3 - Subtask A** of finding Question – Comment similarity.

It is for this reason we use Paragraph Vectors (Le and Mikolov (2014)) for quantifying the question-comment text documents. Paragraph Vector is an unsupervised framework that learns continuous distributed vector representations for pieces of texts. The texts can be of variable-length, ranging from sentences to documents.

The name Paragraph Vector is to emphasize the fact that the method can be applied to variable-length pieces of texts, anything from a phrase or sentence to a large document.

3.1 Learning Algorithm

We used a simple feedforward neural network, i.e. a multilayered perceptron, to predict the best answer, similar to the apporach in Bogdanova and Foster (2016). As shown in Figure 3.1, the first layer of network takes the vector representation for a question-comment pair *(q, c)* as input, which is a concatenation of the distributed representations q and c, the question and the comment text respectively. Each representation is a real-valued vector of a fixed dimensionality d, which is a parameter to be tuned. The input layer is concatenated with another d dimensional vector, namely the centroidal comment, which is centroid of the distributed representation of all comments to the question q (subsection 3.3.1). This is further concatenated with another set of features, that capture syntactic and metadata information, generated from the pair (q, c) as described in section 3.3. The latter two enhancements is the reason our approach shall improve upon the performance achieved by Bogdanova and Foster (2016).

The input layer is then followed by one or more hidden layers; the number of layers and units in each of these layers being the parameters to be experimentally tuned. We consider the activation function to be a parameter as well. Finally, a softmax layer is used to compute the output probability p, i.e. the probabilities p1 and p2 of the negative (i.e. not a relevant comment) and positive (i.e. relevant comment) classes respectively. For each question, all its user-generated comments are ranked according to their probability of being the best answer, as predicted by the network.

Given a question-comment pair (q, c), the possible values for the ground-truth label are 1 (relevant) and 0 (not relevant). The network is trained by minimizing the

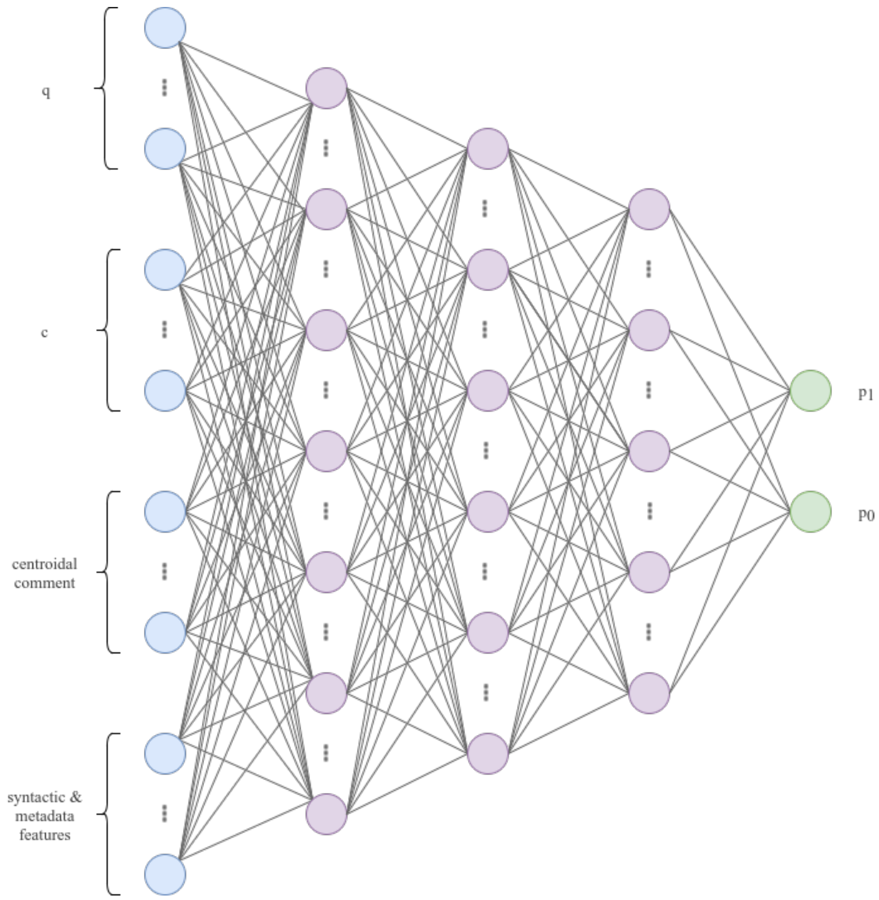

FIGURE 3.1: Architecture of proposed Feedforward Neural Network

L2-regularized cross-entropy loss function between the ground-truth labels and the network predictions on the training set. We use either stochastic gradient descent (SGD) or Adam solver and early stopping to minimize the loss over the training set.

3.2 Document Representations

This approach requires question-comment pairs to be represented as a fixed-size vector. We experimentally evaluate the Paragraph Vector model (PV) proposed by Le and Mikolov (2014). The PV is an extension of the widely used continuous bag-of-words (CBOW) and skip-gram word embedding models, known as word2vec.

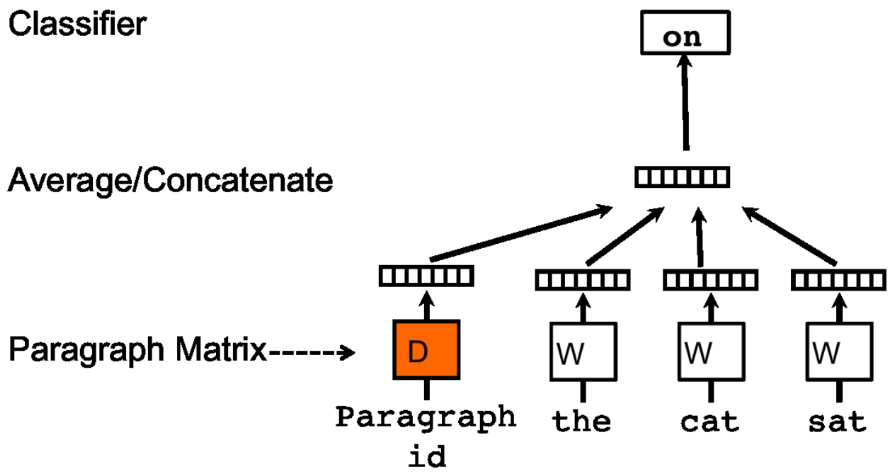

FIGURE 3.2: Distributed Memory (DM) framework for learning paragraph vector. In this model, the concatenation or average of word vectors with a context of few words is used to predict the next word. The paragraph vector represents the missing information from the current context and can act as a memory of the topic of the paragraph.

However, in contrast to CBOW and skip-gram models that only learn word embeddings, the PV is able to learn representations for pieces of text of arbitrary length, e.g. sentences, paragraphs or documents. The types of PV include (1) the distributed memory (DM) model, that predicts the next word in a text window using the concatenation of the word vectors of previous words and the paragraph vector; (2) the distributed bag-of-words (DBOW) model, that – similar to the skip-gram model – predicts words (in a small window) randomly sampled from the paragraph, given the paragraph vector. We experiment with both DM and DBOW models. Figure 3.2 and Figure 3.3 provide an illustration for these paragraph vector models. Also, note that we shall use the terms paragraph vector (PV) and document vector/representation interchangeably.

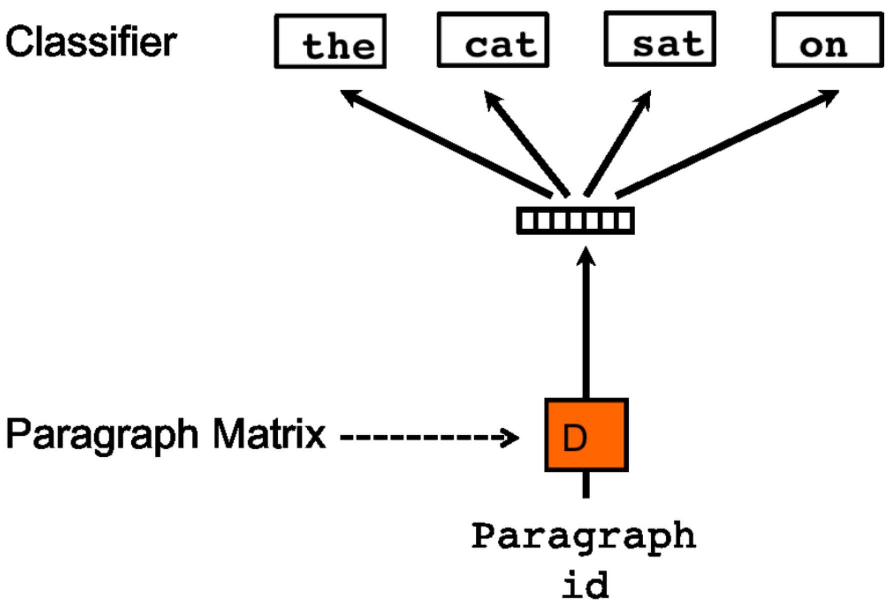

Figure 3.3: Distributed Bag of Words (DBOW) version of paragraph vectors. The paragraph vector is trained to predict the words in a small window

3.3 Feature Set

Apart from paragraph vectors of the Question (q) and Comment (c) that the feed-forward network takes as input, we describe below the surplus features incorporated in our model:

3.3.1 Centroidal Comment

In order to rank the comments, it is only intuition that we must use the information in other comment texts to capture the inter-dependency at thread level and accurately provide relative relevance scores; since the output probability is the inverse rank for comment texts. It is for this reason we introduced the centroidal comment, denoted by avg_com_q, which is computed as:

$$avg\_com_q = \frac{\sum\limits_{c \in q} c}{||\sum\limits_{c \in q} c||} \tag{1}$$

3.3.2 Syntactic and Metadata Features

We used several semantic vector similarity and metadata feature groups as mentioned in Mihaylov and Nakov (2016) to better capture the relevance of each comment to a question text. For the ease of the reader, we shall describe the same feature groups below.

Note that for the similarity measures mentioned below, we used cosine similarity:

$$sim(u,v) = 1 - \frac{u.v}{||u||.||v||} \tag{2}$$

Semantic Word Embeddings. We used semantic word embeddings obtained from Word2Vec models trained on the unannotated data set from QatarLiving. For each piece of text we constructed the centroid vector from the vectors of all words in that text (excluding stopwords).

$$centroid(w_{1...n}) = \frac{\sum_{i=1}^{n} w_i}{n} \tag{3}$$

We construct centroid vectors (3) from the question text (subject + body) and the comment text to design various features as described below.

Semantic Vector Similarities. We used various similarity features calculated using the centroid word vectors on the question text (subject + body) and on the comment text, as well as on parts thereof:

Question to Answer similarity. We assume that a relevant answer should have a centroid vector that is close to that for the question. We used the question text to comment text vector similarities.

Maximized similarity. We ranked each word in the comment text to the question text centroid vector according to their similarity and we took the average similarity of the top N words. We took the top 1, 2, 3, 4 and 5 words similarities as features.

The assumption here is that if the average similarity for the top N most similar words is high, then the comment might be relevant.

Aligned similarity. For each word in the question text, we chose the most similar word from the comment text and we took the average of all best word pair similarities as suggested in Tran et al. (2015).

Part of speech (POS) based word vector similarities. We performed part of speech tagging using the Stanford tagger Toutanova et al. (2003), and we took similarities between centroid vectors of words with a specific tag from the comment text and the centroid vector of the words with a specific tag from the question text. The assumption is that some parts of speech between the question and the comment might be closer than other parts of speech.

Word clusters (WC) similarity. We clustered the word vectors from the Word2Vec vocabulary in 1,000 clusters using K-Means clustering. We then calculated the cluster similarity between the question body word clusters and the answer text word clusters. For all experiments, we used clusters obtained from the Word2Vec model trained on QatarLiving forums with vector size of 100, window size 10.

LDA topic similarity. We performed topic clustering using Latent Dirichlet Allocation (LDA) as implemented in the gensim toolkit Rehurek and Sojka (2010) on Train1 + Train2 + Dev questions and comments. We built topic models with 100 topics. For each word in the question text and for the comment text, we built a bag-of-topics with corresponding distribution, and calculated similarity. The assumption here is that if the question and the comment share similar topics, they are more likely to be relevant to each other

Paragraph Vector similarities. The similarity between the distributed vector representations of question text (q), answer text (a) and the centroidal comment (avg_com_q), taken two at a time are also included.

Metadata. In addition to the semantic features described above, we also used some common sense metadata features:

Answer contains a question mark. If the comment has an question mark, it may be another question, which might indicate a bad answer.

Answer length. Assumption here is that longer answers could bring useful details.

Question length. If the question is longer, it may be more clear, which may help users give a more relevant answer.

Question to comment length. If the question is long and the answer is short, it may be less relevant.

The answer's author is the same as the corresponding question's author. If the answer is posted by the same user who posted the question and it is relevant, why has he/she asked the question in the first place?

Answer rank in the thread. Earlier answers could be posted by users who visit the forum more often, and they may have read more similar questions and answers. Moreover, discussion in the forum tends to diverge from the question over time.

Question category. We took the category of the question as a sparse binary feature vector (a feature with a value of 1 appears if question is in the category). The assumption here is that the question-comment relevance might depend on the category of the question.

Comments by the same User. The number of comments by the author of a given comment to the same question and the order of the comments (first, second, ...) is also included as a feature. If the author produced an incomplete answer in the first attempt, he/she might be obliged to produce another comment subsequently.

Time difference between Question and Comment posting. Immediate comments could reflect incomplete answers to longer questions, while comments posted after substantial time might reflect well-thought answers.

Experiments

4.1 Data

Though Bogdanova and Foster (2016) experiments with the *Yahoo! Answers dataset*[4], we have used the data provided as a part of the popular **SemEval '16 Task 3** for **Subtask A** (Nakov et al. (2016)). Table 1 contains the statistics about the forementioned dataset. This dataset contains about 42K (q, c) pairs to learn from; spreading over about 5.4K questions. We shall refer to this data as the CQA-QL corpus in future. Further we also use a large unannotated dataset, released by the same source, from Qatar Living with 189,941 questions and 1,894,456 comments; which is used for unsupervised learning/training of domain-specific word/document embeddings.

Category	Train (Part-I)	Train (Part-II)	Train+Dev+Test (from SemEval 2015)	Dev	Test	Total
Questions	**1,411**	**379**	**2,480+291+319**	**244**	**327**	**5,451**
Comments	**14,110**	**3,790**	**14,893+1,529+1,876**	**2,440**	**3,270**	**41,908**
-Good	*5,287*	*1,364*	*7,418+813+946*	*818*	*1,329*	*17,975*
-Bad	*6,362*	*1,777*	*5,971+544+774*	*1,209*	*1,485*	*18,122*
-Potentially	*2,461*	*649*	*1,504+172+156*	*413*	*456*	*5,811*

Table 1: Statistics on English CQA-QL corpus
from SemEval-2017 Task 3 (Subtask A)

[4] http://webscope.sandbox.yahoo.com/

4.2 Experimental Setup

We use the gensim[5] implementation of DM and DBOW paragraph vector models. The data for training the unsupervised *doc2vec* model (PV model) is the forementioned large unannotated dataset from Qatar Living forums. Each paragraph (q or c) was converted to lowercase, tokenized by space character and cleaned of stop words before training *doc2vec* models. The parameters of training these models being the window size (maximum distance between the predicted word and context words used for prediction within a document) and number of epochs of training, were cross-validated to give low errors on the training dataset. We further use normalized versions of the document vector representations thus generated, to be fed as inputs to the feedforward neural network described in section 3.1.

For the implementation of the feedforward neural network as described in section 3.1, we used the popular python library *scikit-learn*[6]'s *MLPClassifier*[7].

4.3 Results

4.3.1 Document Vector Representations

For training the document vectors, each question/comment text was treated as a document/paragraph and assigned a label, which was used as a key to retrieve the document vector. Furthermore, post training the doc2vec model is able to infer a document vector for any new question/comment text whose vocabulary is from the original corpus. The errors were computed as averaged squared error over all question/comment text, by computing squared error between the document vector learnt by the model corresponding to the text's label and the document vector

[5]https://radimrehurek.com/gensim/models/doc2vec.html
[6]http://scikit-learn.org/stable/index.html
[7]http://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html

inferred from the corresponding question/comment text. These squared errors are computed for normalized document vectors. For comparison purposes squared error between any two random (normalized) document vector is tabulated beside these errors (averaged over as many iterations as the number of question/comment text). Experiments show that 100-dimensional PV trained over the ~2.3M samples from the unannotated QL corpus, gives sufficiently low errors. Further more, PV-DBOW proves to outperform the PV-DM representations as seen in Table 2. It contains few of the best results and has the rows sorted by the value of column *'Ratio'*, as it is an indicator of how good the representation is. The complete list of experiments are tabulated under Table 11 in Appendix A.

Category	**Window Size**	**Epochs**	**Normalized Sq. Error (A)**	**Norm. Sq. Error (Random) (B)**	**Ratio (B/A)**
PV-DBOW	10	5	0.14	0.80	5.89
PV-DBOW	10	10	0.14	0.83	5.84
PV-DM	10	5	0.21	0.99	4.67
PV-DM	15	10	0.22	0.98	4.47

TABLE 2: Training document vector representations – Best results

To decide upon the dimension of the paragraph vectors we conduct similar experiments by testing on PV-DBOW document vectors of size 100 and 200. The best results tabulated in Table 3, show that 100 dimensional paragraph vectors prove to be a better choice, owing to having both: higher accuracy and low computational complexity. The complete list of experiments are tabulated under Table 12 in Appendix A.

Dimension	**Window Size**	**Epochs**	**Normalized Sq. Error (A)**	**Norm. Sq. Error (Random) (B)**	**Ratio (B/A)**
100	10	5	0.14	0.80	5.89
100	10	10	0.14	0.83	5.84
200	10	10	0.15	0.84	5.65
200	10	5	0.15	0.82	5.47

TABLE 3: Training PV-DBOW document vectors of sizes 100 & 200 – Best results

4.3.2 SemEval Task 3 – Subtask A

The training data comprises of 38,638 comments spanning over 5,124 questions. The neural net input is a tuple of the form $(q, c, avg\_ans_q, ft_{(q,c)})$, where,

avg_com_q is (normalized) average over the PV of all comments to question q

$ft_{(q,c)}$ is feature vector corresp. to the pair (q,c) as described in subsection 3.3.2

SemEval Task 3 has as an official evaluation measure used to rank the participating systems, the metric of Mean Average Precision (MAP), calculated for the ten comments a participating system has ranked highest. Further metrics such as Mean Reciprocal Rank (MRR) and Average Recall ($AvgRec$) for top-10 results; Precision (P), Recall (R), F_1 (with respect to the Good/Relevant class) and Accuracy (Acc) are also reported.

4.3.2.1 Preliminary experiments with (q, c) inputs

Initially experiments were conducted with only (q, c) concatenation as input to the neural nets. The nets were trained while varying solvers, activation functions, hidden layer configurations. The best performance for each parameter configuration is as tabulated in Table 4, while the complete results are tabulated in Appendix B.

Category	**Solver**	**Activation**	**MAP**	**AvgRec**	**MRR**	**P**	**R**	**F_1**	**Acc**
PV-DBOW	Adam	logistic	**70.49**	82.92	77.62	66.01	55.08	60.05	70.21
PV-DBOW	SGD	relu	70.19	82.51	77.16	63.27	59.37	61.26	69.48
PV-DBOW	SGD	logistic	70.18	82.42	77.32	64.12	58.77	61.33	69.88
PV-DBOW	SGD	tanh	70.09	82.45	76.62	63.39	59.14	61.19	69.51
PV-DBOW	Adam	relu	69.93	81.06	76.94	59.98	57.41	58.67	67.13
PV-DBOW	Adam	tanh	69.80	82.31	76.35	63.86	55.46	59.36	69.14
PV-DM	SGD	relu	65.78	78.55	74.58	57.93	53.57	55.67	65.32

TABLE 4: Preliminary experiments using only (q, c) inputs – Best results

PV-DBOW clearly outperforms PV-DM representations in these preliminary runs. Building on this, further experiments where conducted using only the PV-DBOW representations.

4.3.2.2 Improvement with inclusion of Centroidal comment

As described in subsection 3.3.1, additionally, to capture the relative goodness of a comment with respect to other comments of the same question, avg_com_q (normalized post averaging over the PV of all comments to question q) was fed as an input to the neural net. The best results for these experiments are tabulated in Table 5. Complete results are tabulated under Appendix C.

Category	Solver	Activation	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	relu	**73.06**	84.16	79.61	66.07	57.86	61.69	70.8
PV-DBOW	SGD	tanh	71.88	83.43	79.11	66.84	56.58	61.29	70.95
PV-DBOW	SGD	logistic	71.79	83.46	79.13	66.52	55.3	60.39	70.52
PV-DBOW	Adam	logistic	71.77	83.42	78.96	66.35	57.26	61.47	70.83
PV-DBOW	Adam	tanh	71.69	83.39	78.64	65.5	57.71	61.36	70.46
PV-DBOW	Adam	relu	71.61	82.63	79.45	62.98	60.05	61.48	69.42

TABLE 5: Experiments using (q, c, avg_com_q) inputs – Best results

Clearly there is a significant improvement in MAP scores after inclusion of the centroidal comment of a question as an input feature.

4.3.2.3 Further improvement with Syntactic and Metadata Features

In addition, we experimented by including a few more features as described in subsection 3.3.2; which includes various thread level metadata features, and features that capture syntactic similarities between the question and comment text. The best results are tabulated in Table 6. Complete results are tabulated under Appendix D. To note, we excluded the *ADAM* solver from parameters, as *SGD* performs significantly better on all runs (can be seen from Table 5 as well).

Category	Solver	Activation	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	tanh	**77.74**	88.2	85.58	70.81	63.51	66.96	74.53
PV-DBOW	SGD	relu	77.43	87.96	84.81	70.57	62.6	66.35	74.19
PV-DBOW	SGD	logistic	77.26	87.87	85.51	71.49	63.58	67.3	74.89

TABLE 6: Experiments using $(q, c, avg\_com_q, ft_{(q,c)})$ inputs – Best results

4.3.2.4 Analysis with $(q, c, ft_{(q,c)})$ inputs

To get more insight into the contribution of each set of features, we remove the avg_com_q features and test the model. The best results are tabulated in Table 7, while the complete results can be found in Appendix E. In this case, we do not use the similarities of (q, avg_com_q) and (c, avg_com_q) pairs, in our syntactic feature set contained in $ft_{(q,c)}$. That is, we are removing all contributions of centroidal comments from our feature set.

Category	Solver	Activation	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	logistic	**77.22**	87.44	84.67	69.62	62.75	66.01	73.73
PV-DBOW	SGD	tanh	77.14	87.44	84.65	69.31	63.21	66.12	73.67
PV-DBOW	SGD	relu	77.03	87.39	84.8	70.32	62.75	66.32	74.10

TABLE 7: Experiments using $(q, c, ft_{(q,c)})$ inputs – Best results

4.3.2.5 Analysis with only $(ft_{(q,c)})$ inputs

Finally, to see more clearly how much impact the $ft_{(q,c)}$ feature set has, we experimented with these features alone as our inputs. Though this feature set doesn't manage to clock the best results, it manages to out perform the feature set containing concatenation of document vector representations for question text, comment text and centroidal comment (subsubsection 4.3.2.2). The best resutls of these experiments are in Table 8, while the complete results are tabulated under Appendix F.

Category	Solver	Activation	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	tanh	**75.11**	86.63	81.78	69.93	59.14	64.08	73.06
PV-DBOW	SGD	relu	74.84	86.25	81.61	70.09	59.59	64.42	73.24
PV-DBOW	SGD	logistic	74.33	85.98	81.35	68.96	59.67	63.98	72.69

TABLE 8: Experiments using $(ft_{(q,c)})$ inputs – Best results

Discussion and Future Work

Summarizing our efforts, experiments using $(q, c, avg\_com_q, ft_{(q,c)})$ prove to produce the best results with a MAP score of 77.74, only being second to the best submission (primary) at SemEval '16 Task 3 – Subtask A, who managed to achieve a MAP of 79.19. Table 9 summarizes the best results produced for each type of feature set attained by us, for comparison with best results and baselines at SemEval '16 Task 3 – Subtask A, which is tabulated under Table 10. Appendix G contains the complete list of submissions published at the event, for our subtask. Note that our model performs slightly better than the approach suggested by Mihaylova et al. (2016) on which our symantic and metadata features are based on.

Features	**MAP**	**AvgRec**	**MRR**	**P**	**R**	**F_1**	**Acc**
(q, c)	**70.49**	82.92	77.62	66.01	55.08	60.05	70.21
(q, c, avg_com_q)	**73.06**	84.16	79.61	66.07	57.86	61.69	70.8
$(ft_{(q,c)})$	**75.11**	86.63	81.78	69.93	59.14	64.08	73.06
$(q, c, ft_{(q,c)})$	**77.22**	87.44	84.67	69.62	62.75	66.01	73.73
$(q, c, avg\_com_q, ft_{(q,c)})$	**77.74**	88.2	85.58	70.81	63.51	66.96	74.53

TABLE 9: Best results corresponding to each of the feature sets

Starting with concatenation of document representations for question and comment texts, while achieving a MAP of 70.49, we analyse the contribution of each feature

Submission	MAP	AvgRec	MRR	P	R	F_1	Acc
Kelp-primary Filice et al. (2016)	79.19	88.82	86.42	76.96	55.30	64.36	75.11
ConvKN-contrastive1 Joty et al. (2016)	78.71	88.98	86.15	77.78	53.72	63.55	74.95
SUper_team-contrastive1 Mihaylova et al. (2016)	77.68	88.06	84.76	75.59	55.00	63.68	74.50
ConvKN-primary Joty et al. (2016)	77.66	88.05	84.93	75.56	58.84	66.16	75.54
SemanticZ-primary Mihaylov and Nakov (2016)	77.58	88.14	85.21	74.13	53.05	61.84	73.39
Baseline (chronological)	59.53	72.60	67.83	—	—	—	—
Baseline (random)	52.80	66.52	58.71	40.56	74.57	52.55	45.26
Baseline (all 'true')	—	—	—	40.64	100.00	57.80	40.64
Baseline (all 'false')	—	—	—	—	—	—	59.36

TABLE 10: Best submissions and Baselines for SemEval '16 Task 3 – Subtask A

set for their impact in providing more accurate question-comment similarity scores. While with inclusion of the centroidal comment we see an improvement of ~2.5% in the MAP score, there has been a significant rise in the MAP score of ~7% with the introduction of syntactic and metadata features to this concatenation of document representations. Interesting fact to note is that the syntactic and metadata features alone (which also contain the pair-wise similarity scores amongst question text, comment text, centroidal comment) give of about ~4.5% better MAP than that with the concatenation of question, comment text embeddings. Finally, the fact remains that the best results are obtained with the usage of all the fore-mentioned feature sets.

It is interesting to note that our best results have a higher recall value compared to the best submissions at the event (Table 10). This could possibly be due to the fact that our method, with the document vector representations is able to capture more semantic similarities to factor into the relevance scores. Infact our method produces the best F_1 score out of all the submissions at the event (tabulated in Appendix G).

The model, with each inclusion of a new set of features has responded on the postive side with an improvement in the MAP score. Further work on improving these results are hence highly likely to be fruitful. The inclusion of various other features related to syntactic similarity amongst the question, comment text and possibly the centroidal comment, could impact highly on the MAP; owing to the substantial rise in score post inclusion of $ft_{(q,c)}$ feature set. Using syntax trees of question/comment/centroidal-comment text, to find more similarity metrics, on the lines of the tree kernel approach as adopted by Filice et al. (2016) (Best submission at SemEval '16 Task 3 – Subtask A), Joty et al. (2016) is such an example.

Appendix A

Training PV-DM and PV-DBOW

Category	Window Size	Epochs	Normalized Sq. Error (A)	Norm. Sq. Error (Random) (B)	Ratio (B/A)
PV-DBOW	10	5	0.14	0.80	5.89
PV-DBOW	10	10	0.14	0.83	5.84
PV-DBOW	15	5	0.15	0.80	5.50
PV-DBOW	10	3	0.14	0.78	5.48
PV-DBOW	15	10	0.15	0.82	5.47
PV-DBOW	15	3	0.15	0.78	5.12
PV-DBOW	20	5	0.16	0.80	5.09
PV-DBOW	20	10	0.16	0.82	5.04
PV-DBOW	20	3	0.16	0.77	4.71
PV-DM	10	5	0.21	0.99	4.67
PV-DM	15	10	0.22	0.98	4.47
PV-DM	15	5	0.23	0.98	4.29
PV-DM	20	10	0.23	0.98	4.28
PV-DM	25	10	0.24	0.98	4.17
PV-DM	20	5	0.24	0.98	4.09
PV-DM	15	20	0.24	0.98	4.03
PV-DM	25	5	0.25	0.98	3.95
PV-DM	20	20	0.25	0.98	3.95
PV-DM	25	20	0.25	0.98	3.89
PV-DBOW	10	1	0.19	0.73	3.75
PV-DM	15	30	0.26	0.96	3.69
PV-DM	20	30	0.27	0.97	3.64
PV-DM	25	30	0.27	0.97	3.60
PV-DBOW	15	1	0.21	0.72	3.48
PV-DBOW	20	1	0.22	0.72	3.23
PV-DM	10	1	0.33	0.98	2.94
PV-DM	15	1	0.36	0.97	2.72
PV-DM	20	1	0.37	0.97	2.59
PV-DM	25	1	0.38	0.97	2.52

TABLE 11: Training document vector representations – All results

Dimension	**Window Size**	**Epochs**	**Normalized Sq. Error (A)**	**Norm. Sq. Error (Random) (B)**	**Ratio (B/A)**
100	10	5	0.14	0.80	5.89
100	10	10	0.14	0.83	5.84
200	10	10	0.15	0.84	5.65
100	15	5	0.15	0.80	5.50
100	10	3	0.14	0.78	5.48
200	10	5	0.15	0.82	5.47
100	15	10	0.15	0.82	5.47
200	15	10	0.16	0.84	5.18
100	15	3	0.15	0.78	5.12
100	20	5	0.16	0.80	5.09
100	20	10	0.16	0.82	5.04
200	15	5	0.16	0.82	5.01
200	10	3	0.16	0.80	4.96
100	20	3	0.16	0.77	4.71
200	20	10	0.18	0.84	4.71
200	20	5	0.18	0.82	4.58
200	15	3	0.18	0.80	4.54
200	20	3	0.19	0.79	4.17
100	10	1	0.19	0.73	3.75
100	15	1	0.21	0.72	3.48
200	10	1	0.23	0.76	3.33
100	20	1	0.22	0.72	3.23
200	15	1	0.24	0.75	3.06
200	20	1	0.26	0.74	2.86

TABLE 12: Training PV-DBOW document vectors of sizes 100 & 200 – All results

Appendix B

Preliminary experiments

Category	Solver	Activation	Hidden Layer	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	Adam	logistic	[500, 250, -]	70.49	82.92	77.62	66.01	55.08	60.05	70.21
PV-DBOW	SGD	relu	[100, -, -]	70.19	82.51	77.16	63.27	59.37	61.26	69.48
PV-DBOW	SGD	logistic	[50, -, -]	70.18	82.42	77.32	64.12	58.77	61.33	69.88
PV-DBOW	SGD	logistic	[100, -, -]	70.13	82.42	77.23	63.53	58.84	61.09	69.54
PV-DBOW	SGD	tanh	[500, 100, 100]	70.09	82.45	76.62	63.39	59.14	61.19	69.51
PV-DBOW	Adam	logistic	[50, 50, 50]	70.08	82.52	76.8	60.43	65.84	63.02	68.59
PV-DBOW	SGD	relu	[50, -, -]	70.07	82.40	77.01	63.74	58.99	61.27	69.69
PV-DBOW	SGD	tanh	[250, 100, -]	70.07	82.37	76.86	63.10	58.54	60.73	69.24
PV-DBOW	Adam	logistic	[100, 50, 50]	70.07	82.51	76.89	63.25	59.97	61.57	69.57
PV-DBOW	SGD	logistic	[250, -, -]	70.06	82.36	77.36	63.96	58.62	61.17	69.76
PV-DBOW	Adam	logistic	[100, 100, 100]	70.06	82.50	76.9	64.04	57.49	60.59	69.60
PV-DBOW	SGD	logistic	[500, -, -]	70.05	82.40	77.2	63.69	58.47	60.97	69.57
PV-DBOW	SGD	tanh	[250, 50, 50]	70.04	82.41	76.89	63.39	59.14	61.19	69.51
PV-DBOW	Adam	logistic	[250, 250, 250]	70.03	82.48	76.97	62.17	62.45	62.31	69.30
PV-DBOW	Adam	logistic	[250, 250, 50]	70.00	82.48	76.94	62.05	62.00	62.02	69.14
PV-DBOW	Adam	logistic	[500, 100, -]	69.99	82.44	76.92	62.81	60.87	61.83	69.45
PV-DBOW	Adam	logistic	[250, 100, -]	69.98	82.40	76.7	61.64	62.75	62.19	68.99
PV-DBOW	Adam	logistic	[250, 50, -]	69.98	82.39	76.66	61.94	63.81	62.86	69.36
PV-DBOW	Adam	logistic	[500, 50, -]	69.97	82.43	76.75	63.46	60.50	61.94	69.79
PV-DBOW	Adam	logistic	[250, 250, 100]	69.97	82.43	76.84	63.43	58.99	61.13	69.51
PV-DBOW	SGD	tanh	[250, 100, 50]	69.96	82.35	76.66	63.34	58.77	60.97	69.42
PV-DBOW	SGD	tanh	[100, 100, 50]	69.96	82.37	76.99	63.70	58.77	61.14	69.63
PV-DBOW	SGD	tanh	[100, -, -]	69.96	82.30	76.67	63.40	58.39	60.79	69.39
PV-DBOW	Adam	logistic	[500, 250, 50]	69.95	82.38	76.69	64.47	57.49	60.78	69.85
PV-DBOW	SGD	tanh	[500, 500, 250]	69.94	82.41	76.76	63.39	59.14	61.19	69.51
PV-DBOW	SGD	tanh	[500, 500, 100]	69.94	82.35	76.89	63.47	58.84	61.07	69.51
PV-DBOW	SGD	tanh	[500, 500, -]	69.94	82.38	76.43	63.47	58.84	61.07	69.51
PV-DBOW	Adam	logistic	[100, 100, 50]	69.94	82.40	76.51	64.06	57.94	60.85	69.69
PV-DBOW	Adam	logistic	[100, 50, -]	69.94	82.43	76.74	62.79	60.95	61.86	69.45
PV-DBOW	Adam	relu	[500, 50, -]	69.93	81.06	76.94	59.98	57.41	58.67	67.13
PV-DBOW	SGD	tanh	[500, 500, 50]	69.92	82.36	76.5	63.44	59.14	61.21	69.54

Table 13: Preliminary experiments using only (q, c) inputs – All results (MAP>0.6)

Category	Solver	Activation	Hidden Layer	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	tanh	[250, 50, -]	69.92	82.25	76.76	63.30	58.92	61.03	69.42
PV-DBOW	SGD	tanh	[100, 100, -]	69.92	82.41	76.74	63.49	58.62	60.95	69.48
PV-DBOW	Adam	logistic	[500, 500, 500]	69.92	82.46	76.77	60.98	65.84	63.31	68.99
PV-DBOW	Adam	logistic	[250, 100, 50]	69.92	82.38	76.74	62.24	62.75	62.50	69.39
PV-DBOW	Adam	logistic	[50, 50, -]	69.92	82.41	76.87	63.33	60.42	61.84	69.69
PV-DBOW	SGD	tanh	[50, 50, -]	69.90	82.36	76.77	63.73	58.84	61.19	69.66
PV-DBOW	SGD	tanh	[50, -, -]	69.90	82.35	76.95	64.07	58.77	61.30	69.85
PV-DBOW	SGD	relu	[500, -, -]	69.89	82.40	76.58	63.77	59.07	61.33	69.72
PV-DBOW	SGD	tanh	[500, 100, -]	69.89	82.31	76.79	63.38	58.62	60.91	69.42
PV-DBOW	SGD	tanh	[250, -, -]	69.87	82.38	76.94	63.21	58.69	60.87	69.33
PV-DBOW	SGD	tanh	[100, 100, 100]	69.87	82.35	76.59	63.65	58.77	61.11	69.60
PV-DBOW	SGD	tanh	[500, 100, 50]	69.86	82.33	76.5	63.65	59.29	61.39	69.69
PV-DBOW	Adam	logistic	[100, 100, -]	69.86	82.35	76.28	61.92	62.53	62.22	69.14
PV-DBOW	SGD	relu	[250, -, -]	69.85	82.40	76.55	63.69	58.99	61.25	69.66
PV-DBOW	Adam	logistic	[100, -, -]	69.84	82.32	76.54	64.30	56.66	60.24	69.60
PV-DBOW	SGD	tanh	[500, 250, 100]	69.83	82.33	76.2	63.56	59.07	61.23	69.60
PV-DBOW	SGD	tanh	[500, 250, 50]	69.83	82.29	76.52	63.50	59.29	61.32	69.60
PV-DBOW	SGD	tanh	[500, 250, -]	69.83	82.35	76.68	63.39	58.77	60.99	69.45
PV-DBOW	Adam	logistic	[500, -, -]	69.83	82.30	76.51	59.83	68.92	64.06	68.56
PV-DBOW	Adam	logistic	[50, -, -]	69.83	82.33	76.55	63.85	58.47	61.04	69.66
PV-DBOW	SGD	tanh	[500, 250, 250]	69.82	82.34	76.61	63.41	58.69	60.96	69.45
PV-DBOW	SGD	tanh	[500, 50, 50]	69.82	82.32	76.63	63.09	58.92	60.93	69.30
PV-DBOW	SGD	tanh	[500, -, -]	69.82	82.35	76.68	63.39	58.24	60.71	69.36
PV-DBOW	Adam	logistic	[500, 500, 50]	69.82	82.11	76.28	63.11	58.69	60.82	69.27
PV-DBOW	Adam	logistic	[250, 250, -]	69.81	82.29	76.51	61.49	64.41	62.92	69.14
PV-DBOW	Adam	tanh	[100, -, -]	69.80	82.31	76.35	63.86	55.46	59.36	69.14
PV-DBOW	Adam	tanh	[50, -, -]	69.80	82.32	76.64	62.64	61.70	62.17	69.48
PV-DBOW	SGD	tanh	[500, 50, -]	69.79	82.29	76.37	63.38	58.99	61.11	69.48
PV-DBOW	SGD	tanh	[100, 50, 50]	69.79	82.39	76.52	63.39	58.77	60.99	69.45
PV-DBOW	SGD	tanh	[250, 250, 50]	69.75	82.22	76.47	63.16	58.69	60.84	69.30
PV-DBOW	SGD	tanh	[50, 50, 50]	69.75	82.24	76.66	63.69	58.99	61.25	69.66
PV-DBOW	SGD	tanh	[250, 250, 100]	69.74	82.32	76.6	63.23	58.99	61.04	69.39
PV-DBOW	SGD	tanh	[250, 250, -]	69.74	82.24	76.32	63.30	58.54	60.83	69.36
PV-DBOW	SGD	tanh	[250, 100, 100]	69.74	82.28	76.39	63.22	58.84	60.95	69.36
PV-DBOW	Adam	tanh	[500, -, -]	69.74	82.28	76.48	62.74	62.45	62.59	69.66
PV-DBOW	SGD	tanh	[100, 50, -]	69.73	82.31	76.44	63.56	58.92	61.15	69.57
PV-DBOW	SGD	tanh	[500, 500, 500]	69.71	82.21	76.48	63.31	59.07	61.11	69.45
PV-DBOW	Adam	logistic	[250, -, -]	69.66	82.20	76.19	64.70	55.30	59.63	69.57
PV-DBOW	Adam	tanh	[250, -, -]	69.64	82.18	76.23	63.31	58.69	60.91	69.39
PV-DBOW	SGD	tanh	[250, 250, 250]	69.62	82.14	76.16	63.35	58.92	61.05	69.45
PV-DBOW	Adam	logistic	[500, 250, 250]	69.59	82.05	76.23	62.96	60.87	61.90	69.54
PV-DBOW	Adam	relu	[500, 500, 500]	69.53	81.34	75.64	58.63	61.85	60.20	66.76
PV-DBOW	Adam	logistic	[500, 250, 100]	69.52	81.95	75.94	64.39	55.38	59.55	69.42
PV-DBOW	Adam	logistic	[500, 500, 100]	69.48	81.96	76.16	62.54	61.93	62.23	69.45
PV-DBOW	Adam	logistic	[500, 500, -]	69.48	82.06	75.89	64.93	54.33	59.16	69.51
PV-DBOW	Adam	logistic	[500, 100, 100]	69.44	82.00	75.84	62.71	60.50	61.59	69.33
PV-DBOW	Adam	logistic	[500, 100, 50]	69.38	81.92	76	62.71	60.35	61.50	69.30
PV-DBOW	Adam	logistic	[250, 100, 100]	69.31	81.87	75.65	65.36	52.52	58.24	69.39
PV-DBOW	Adam	logistic	[250, 50, 50]	69.30	81.96	75.76	63.25	59.97	61.57	69.57
PV-DBOW	Adam	logistic	[500, 500, 250]	69.27	81.89	75.7	59.82	68.55	63.88	68.50
PV-DBOW	Adam	logistic	[500, 50, 50]	69.19	81.87	75.65	62.85	60.35	61.57	69.39
PV-DBOW	SGD	relu	[500, 500, 100]	68.98	80.95	76.17	59.01	58.39	58.70	66.61
PV-DBOW	Adam	relu	[500, 500, 100]	68.90	80.85	76.14	58.95	57.26	58.09	66.42
PV-DBOW	Adam	relu	[500, 250, 50]	68.89	80.52	77.5	57.56	56.13	56.84	65.35
PV-DBOW	Adam	relu	[500, 100, -]	68.80	80.94	75.7	59.12	56.36	57.70	66.42
PV-DBOW	Adam	relu	[500, 250, 250]	68.72	81.00	75.47	58.79	61.93	60.32	66.88
PV-DBOW	Adam	relu	[500, -, -]	68.71	80.81	76.88	57.29	58.24	57.76	65.38
PV-DBOW	SGD	relu	[500, 100, 100]	68.61	80.71	77.62	57.70	56.66	57.18	65.50

Preliminary experiments using only (q, c) inputs – All results (*MAP*>0.6)

Category	**Solver**	**Activation**	**Hidden Layer**	**MAP**	**AvgRec**	**MRR**	**P**	**R**	**F_1**	**Acc**
PV-DBOW	Adam	relu	[250, -, -]	68.61	80.53	75.89	59.04	55.53	57.23	66.27
PV-DBOW	SGD	relu	[500, 250, -]	68.55	80.81	76.23	58.50	55.15	56.78	65.87
PV-DBOW	SGD	relu	[500, 100, -]	68.53	80.42	77.34	58.76	55.53	57.10	66.09
PV-DBOW	SGD	relu	[500, 500, -]	68.45	80.52	75.42	58.70	56.36	57.50	66.15
PV-DBOW	SGD	relu	[500, 250, 100]	68.28	80.12	75.33	57.53	57.19	57.36	65.44
PV-DBOW	SGD	relu	[500, 50, -]	68.23	80.55	75.46	57.71	57.19	57.45	65.57
PV-DBOW	Adam	relu	[250, 250, 100]	68.23	81.51	76.37	60.77	50.94	55.42	66.70
PV-DBOW	Adam	relu	[250, 250, 250]	68.16	80.73	76.21	61.63	47.86	53.88	66.70
PV-DBOW	SGD	relu	[500, 250, 250]	68.05	80.13	75.65	57.47	57.04	57.25	65.38
PV-DBOW	Adam	relu	[500, 500, 250]	68.00	80.35	76.37	60.21	56.13	58.10	67.09
PV-DBOW	Adam	relu	[500, 250, 100]	67.96	79.63	75.73	58.29	61.10	59.66	66.42
PV-DBOW	Adam	relu	[500, 100, 50]	67.95	80.25	75.87	58.99	58.99	58.99	66.67
PV-DBOW	SGD	relu	[500, 250, 50]	67.90	80.00	74.37	58.47	56.36	57.39	65.99
PV-DBOW	Adam	relu	[50, -, -]	67.89	80.20	74.61	57.46	53.88	55.61	65.05
PV-DBOW	Adam	relu	[500, 250, -]	67.84	80.15	74.97	60.37	53.88	56.94	66.88
PV-DBOW	Adam	relu	[250, 250, 50]	67.84	80.29	74.02	58.92	51.92	55.20	65.75
PV-DBOW	SGD	relu	[500, 100, 50]	67.80	80.27	75.18	58.28	57.71	58.00	66.02
PV-DBOW	Adam	relu	[250, 50, -]	67.80	80.43	74.84	58.99	53.35	56.03	65.96
PV-DBOW	Adam	relu	[500, 500, 50]	67.79	79.93	75.16	60.69	51.92	55.96	66.79
PV-DBOW	Adam	relu	[250, 50, 50]	67.77	80.28	74.79	61.44	50.11	55.20	66.94
PV-DBOW	Adam	relu	[100, -, -]	67.64	79.63	75.58	57.76	56.28	57.01	65.50
PV-DBOW	Adam	relu	[250, 100, 50]	67.59	80.06	75.18	58.27	52.22	55.08	65.38
PV-DBOW	SGD	relu	[500, 500, 500]	67.57	79.98	74.65	58.20	56.58	57.38	65.84
PV-DBOW	SGD	relu	[500, 500, 50]	67.50	80.02	75.24	57.78	56.43	57.10	65.54
PV-DBOW	Adam	relu	[250, 100, 100]	67.49	80.09	75.43	57.14	62.60	59.75	65.72
PV-DBOW	Adam	relu	[500, 50, 50]	67.45	79.80	73.88	56.75	63.28	59.84	65.47
PV-DBOW	Adam	relu	[100, 100, -]	67.36	79.85	73.98	57.34	58.16	57.75	65.41
PV-DBOW	SGD	relu	[500, 50, 50]	67.35	79.94	75.36	58.47	58.47	58.47	66.24
PV-DBOW	SGD	relu	[50, 50, -]	67.31	79.78	74.39	57.64	56.21	56.91	65.41
PV-DBOW	SGD	relu	[250, 250, 50]	67.26	79.35	73.95	56.71	55.98	56.34	64.74
PV-DBOW	Adam	tanh	[500, 500, 500]	67.24	79.53	75.14	59.31	58.24	58.77	66.79
PV-DBOW	SGD	relu	[500, 500, 250]	67.23	79.55	73.58	57.84	56.36	57.09	65.57
PV-DBOW	SGD	relu	[250, 250, 100]	67.21	79.67	75.01	57.56	56.43	56.99	65.38
PV-DBOW	SGD	relu	[250, 100, -]	67.07	79.89	74.35	56.55	53.27	54.86	64.37
PV-DBOW	Adam	relu	[500, 500, -]	67.07	79.23	73.13	57.47	57.86	57.67	65.47
PV-DBOW	SGD	relu	[50, 50, 50]	67.00	80.21	74.36	57.96	59.97	58.95	66.06
PV-DBOW	SGD	relu	[250, 250, -]	66.93	79.42	74.6	57.80	55.23	56.48	65.41
PV-DBOW	SGD	relu	[250, 100, 50]	66.85	79.09	73.52	55.92	54.70	55.31	64.07
PV-DBOW	Adam	relu	[250, 250, -]	66.84	80.03	73.9	57.09	57.56	57.32	65.17
PV-DBOW	SGD	relu	[250, 50, 50]	66.80	79.15	73.88	57.35	56.66	57.00	65.26
PV-DBOW	Adam	relu	[250, 100, -]	66.73	79.36	73.37	57.25	57.64	57.44	65.29
PV-DBOW	Adam	tanh	[500, 500, 250]	66.58	78.78	73.56	58.74	57.64	58.18	66.33
PV-DBOW	Adam	tanh	[500, 250, 100]	66.52	78.88	72.88	58.47	55.08	56.72	65.84
PV-DBOW	Adam	relu	[500, 100, 100]	66.51	79.38	72.41	58.22	61.02	59.59	66.36
PV-DBOW	Adam	relu	[50, 50, -]	66.43	78.86	73.6	55.62	55.83	55.73	63.94
PV-DBOW	SGD	relu	[250, 100, 100]	66.42	79.08	73.86	57.31	54.55	55.90	65.02
PV-DBOW	SGD	relu	[100, 50, -]	66.40	79.27	74.39	55.75	56.88	56.31	64.13
PV-DBOW	SGD	relu	[250, 50, -]	66.29	79.63	73.97	57.04	54.55	55.77	64.83
PV-DBOW	SGD	relu	[250, 250, 250]	66.27	79.16	72.39	57.32	53.65	55.42	64.92
PV-DBOW	Adam	tanh	[500, 500, 100]	66.20	78.90	73.43	56.96	57.86	57.41	65.11
PV-DBOW	Adam	tanh	[250, 250, 250]	66.15	78.23	73.69	58.16	56.28	57.21	65.78
PV-DBOW	Adam	tanh	[500, 500, -]	66.02	78.72	74.35	59.88	58.16	59.01	67.16
PV-DBOW	Adam	tanh	[250, 250, 100]	66.01	78.76	72.89	58.31	55.98	57.12	65.84
PV-DBOW	Adam	relu	[100, 50, -]	65.96	78.62	74.06	55.73	54.55	55.13	63.91
PV-DBOW	Adam	tanh	[500, 250, 50]	65.82	78.58	72.06	60.67	56.06	58.27	67.37
PV-DBOW	SGD	relu	[100, 100, -]	65.81	78.63	72.44	54.78	55.15	54.97	63.27
PV-DBOW	SGD	relu	[100, 50, 50]	65.81	78.70	72.68	55.45	55.91	55.68	63.82
PV-DM	SGD	relu	[100, -, -]	65.78	78.55	74.58	57.93	53.57	55.67	65.32

Preliminary experiments using only (q, c) inputs – All results (MAP>0.6)

Category	Solver	Activation	Hidden Layer	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	Adam	tanh	[250, 50, -]	65.74	78.64	72.75	56.78	59.22	57.97	65.11
PV-DBOW	Adam	tanh	[500, 500, 50]	65.72	79.09	72.68	58.79	58.62	58.70	66.48
PV-DM	SGD	relu	[50, -, -]	65.66	78.12	74.01	54.68	51.39	52.99	62.94
PV-DBOW	Adam	tanh	[500, 50, -]	65.56	78.40	73.92	56.98	57.41	57.20	65.08
PV-DBOW	Adam	tanh	[500, 250, 250]	65.50	78.26	73.79	57.45	59.22	58.32	65.60
PV-DBOW	Adam	relu	[100, 100, 100]	65.44	78.36	72.81	55.82	55.91	55.86	64.10
PV-DBOW	SGD	relu	[100, 100, 100]	65.42	78.31	73.9	53.91	55.53	54.71	62.63
PV-DBOW	Adam	tanh	[100, 100, 50]	65.32	78.20	73.38	55.83	55.83	55.83	64.10
PV-DBOW	Adam	tanh	[250, 250, 50]	65.30	78.53	74.63	57.34	55.53	56.42	65.14
PV-DBOW	SGD	relu	[100, 100, 50]	65.27	77.98	72.78	53.98	55.08	54.53	62.66
PV-DM	SGD	relu	[250, -, -]	65.10	77.88	72.85	55.12	50.19	52.54	63.15
PV-DBOW	Adam	tanh	[100, 50, 50]	65.00	77.90	71.93	56.50	57.26	56.88	64.71
PV-DBOW	Adam	tanh	[500, 100, 100]	64.97	77.39	71.12	56.95	57.64	57.29	65.08
PV-DM	SGD	relu	[500, -, -]	64.89	78.16	73.25	56.72	57.19	56.95	64.86
PV-DBOW	Adam	tanh	[250, 100, -]	64.83	77.97	72.33	56.28	58.01	57.13	64.62
PV-DBOW	Adam	tanh	[250, 100, 50]	64.55	77.72	70.64	57.34	54.93	56.11	65.08
PV-DBOW	Adam	tanh	[250, 100, 100]	64.33	77.45	71.85	57.66	59.44	58.54	65.78
PV-DBOW	Adam	tanh	[500, 250, -]	64.32	77.18	72.2	54.30	56.51	55.38	63.00
PV-DBOW	Adam	relu	[100, 100, 50]	64.31	78.01	72.71	56.13	56.13	56.13	64.34
PV-DBOW	Adam	tanh	[100, 100, -]	64.30	77.32	71.16	55.37	57.41	56.37	63.88
PV-DBOW	Adam	tanh	[500, 100, 50]	64.16	77.17	71.75	54.49	53.88	54.18	62.97
PV-DBOW	Adam	tanh	[100, 100, 100]	64.14	77.05	71.56	55.32	52.82	54.04	63.49
PV-DBOW	Adam	tanh	[100, 50, -]	64.13	77.26	72.96	53.76	55.91	54.81	62.54
PV-DBOW	Adam	tanh	[500, 100, -]	63.97	76.83	70.39	56.60	54.85	55.71	64.56
PV-DBOW	Adam	tanh	[50, 50, -]	63.76	75.78	70.44	54.67	56.36	55.50	63.27
PV-DBOW	Adam	relu	[100, 50, 50]	63.64	77.27	70.43	56.84	60.95	58.82	65.32
PV-DBOW	Adam	tanh	[500, 50, 50]	63.58	76.75	70.43	55.45	55.15	55.30	63.76
PV-DBOW	Adam	tanh	[250, 250, -]	63.47	76.82	70.07	56.76	57.19	56.97	64.89
PV-DM	SGD	relu	[50, 50, -]	63.45	75.82	72.47	51.59	51.32	51.45	60.64
PV-DBOW	Adam	relu	[50, 50, 50]	63.33	76.39	70.17	54.41	56.13	55.26	63.06
PV-DBOW	Adam	tanh	[50, 50, 50]	63.10	76.69	70.62	54.28	56.36	55.30	62.97
PV-DBOW	Adam	tanh	[250, 50, 50]	62.92	76.19	70.82	55.08	55.91	55.49	63.55
PV-DM	SGD	relu	[500, 250, 50]	62.86	75.86	72.08	52.27	55.46	53.82	61.31
PV-DM	SGD	relu	[500, 500, 250]	62.37	75.42	71.01	51.35	54.40	52.83	60.52
PV-DM	SGD	relu	[500, 100, 100]	62.35	75.31	71.02	51.45	54.70	53.03	60.61
PV-DM	SGD	relu	[500, 50, -]	62.15	75.60	70.78	51.38	56.13	53.65	60.58
PV-DM	SGD	relu	[500, 500, 500]	62.04	75.40	71.33	49.86	52.52	51.15	59.24
PV-DM	SGD	relu	[500, 100, 50]	61.94	75.24	70.9	52.23	53.80	53.00	61.22
PV-DM	SGD	relu	[500, 250, 100]	61.93	75.17	70.47	50.32	53.20	51.72	59.63
PV-DM	SGD	relu	[500, 500, 100]	61.92	75.05	70.94	51.58	52.90	52.23	60.67
PV-DM	SGD	relu	[500, 100, -]	61.84	74.96	70.66	52.07	54.85	53.43	61.13
PV-DM	SGD	relu	[500, 500, -]	61.75	75.33	70.17	51.30	54.85	53.02	60.49
PV-DM	SGD	relu	[500, 500, 50]	61.74	75.25	71.03	52.46	56.13	54.23	61.50
PV-DM	SGD	relu	[250, 250, -]	61.68	74.24	70.98	48.77	55.30	51.83	58.23
PV-DM	SGD	relu	[500, 250, 250]	61.66	75.23	71.18	52.28	53.50	52.88	61.25
PV-DM	SGD	relu	[500, 250, -]	61.64	75.13	70.24	52.17	55.98	54.01	61.25
PV-DM	SGD	relu	[500, 50, 50]	61.61	74.98	69.61	51.62	53.88	52.72	60.73
PV-DM	SGD	relu	[250, 50, -]	61.54	74.13	72.24	51.00	55.46	53.14	60.24
PV-DM	SGD	relu	[250, 250, 50]	61.46	74.42	69.09	47.80	53.05	50.29	57.37
PV-DM	SGD	relu	[100, 50, 50]	61.31	74.49	69.56	50.68	55.98	53.20	59.97
PV-DM	SGD	relu	[250, 50, 50]	61.28	73.86	69.05	49.51	53.57	51.46	58.93
PV-DM	SGD	relu	[100, 50, -]	61.08	74.75	69.63	49.80	55.30	52.41	59.17
PV-DM	SGD	relu	[250, 100, -]	60.93	74.37	68.79	51.53	55.83	53.59	60.70
PV-DM	SGD	relu	[100, 100, -]	60.89	74.71	68.59	50.58	58.69	54.34	59.91
PV-DM	SGD	relu	[250, 100, 100]	60.77	74.05	69.81	51.00	53.65	52.29	60.21
PV-DM	SGD	relu	[250, 250, 100]	60.47	74.17	68.62	50.48	55.53	52.88	59.79
PV-DM	SGD	relu	[50, 50, 50]	60.09	74.15	68.04	50.65	58.62	54.34	59.97

Preliminary experiments using only (q, c) inputs – All results (MAP>0.6)

Appendix C

Experiments post inclusion of Average Answer

Category	Solver	Activation	Hidden Layer	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	relu	[250, -, -]	73.06	84.16	79.61	66.07	57.86	61.69	70.80
PV-DBOW	SGD	relu	[100, -, -]	72.96	84.18	80.81	66.78	56.88	61.44	70.98
PV-DBOW	SGD	relu	[500, -, -]	72.53	84.29	78.88	65.36	60.35	62.75	70.89
PV-DBOW	SGD	tanh	[500, 100, -]	71.88	83.43	79.11	66.84	56.58	61.29	70.95
PV-DBOW	SGD	tanh	[250, 250, 250]	71.82	83.40	78.87	66.55	56.58	61.16	70.80
PV-DBOW	SGD	logistic	[100, -, -]	71.79	83.46	79.13	66.52	55.30	60.39	70.52
PV-DBOW	Adam	logistic	[50, -, -]	71.77	83.42	78.96	66.35	57.26	61.47	70.83
PV-DBOW	SGD	tanh	[500, 500, 500]	71.72	83.37	78.78	66.75	56.96	61.47	70.98
PV-DBOW	SGD	tanh	[250, 100, 50]	71.71	83.29	78.48	66.55	56.73	61.25	70.83
PV-DBOW	SGD	tanh	[50, -, -]	71.71	83.39	79.09	66.28	56.06	60.74	70.55
PV-DBOW	SGD	tanh	[250, 250, 100]	71.69	83.33	78.79	66.61	56.13	60.92	70.73
PV-DBOW	SGD	tanh	[100, 100, -]	71.69	83.38	79.13	66.52	55.76	60.66	70.61
PV-DBOW	SGD	tanh	[50, 50, 50]	71.69	83.25	78.47	66.73	56.58	61.24	70.89
PV-DBOW	Adam	tanh	[250, -, -]	71.69	83.39	78.64	65.50	57.71	61.36	70.46
PV-DBOW	SGD	logistic	[250, -, -]	71.68	83.38	78.83	66.64	55.76	60.71	70.67
PV-DBOW	SGD	tanh	[100, 50, 50]	71.67	83.26	78.84	66.87	56.36	61.17	70.92
PV-DBOW	Adam	logistic	[250, -, -]	71.67	83.40	78.68	66.76	53.80	59.58	70.34
PV-DBOW	SGD	tanh	[500, 500, 50]	71.66	83.33	78.75	66.76	56.81	61.38	70.95
PV-DBOW	SGD	tanh	[250, -, -]	71.66	83.36	79.01	66.49	55.98	60.78	70.64
PV-DBOW	SGD	tanh	[250, 250, 50]	71.65	83.33	78.64	66.52	56.21	60.93	70.70
PV-DBOW	SGD	tanh	[250, 250, -]	71.65	83.24	78.77	66.61	56.28	61.01	70.76
PV-DBOW	SGD	tanh	[100, 100, 100]	71.64	83.28	78.71	66.67	56.58	61.21	70.86
PV-DBOW	Adam	logistic	[250, 250, -]	71.64	83.34	78.65	65.62	59.59	62.46	70.89
PV-DBOW	Adam	logistic	[100, 100, -]	71.63	83.31	78.76	65.98	57.79	61.61	70.73
PV-DBOW	SGD	logistic	[500, -, -]	71.62	83.36	78.93	66.34	55.30	60.32	70.43
PV-DBOW	Adam	logistic	[500, -, -]	71.62	83.34	78.79	66.03	57.04	61.20	70.61
PV-DBOW	Adam	logistic	[100, 100, 50]	71.62	83.31	78.73	65.19	58.62	61.73	70.46
PV-DBOW	SGD	tanh	[500, 250, 50]	71.61	83.32	78.96	66.70	56.21	61.00	70.80
PV-DBOW	SGD	tanh	[250, 50, -]	71.61	83.35	78.69	66.02	56.13	60.68	70.43
PV-DBOW	Adam	relu	[500, 250, 250]	71.61	82.63	79.45	62.98	60.05	61.48	69.42

TABLE 14: Experiments using (q, c, avg_com_q) inputs – All results (MAP>0.6)

Category	Solver	Activation	Hidden Layer	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	tanh	[250, 100, -]	71.60	83.28	78.31	66.55	56.28	60.99	70.73
PV-DBOW	SGD	tanh	[500, 500, -]	71.58	83.28	78.42	66.52	56.21	60.93	70.70
PV-DBOW	SGD	tanh	[500, 100, 50]	71.58	83.21	78.53	66.70	56.81	61.36	70.92
PV-DBOW	SGD	tanh	[100, 50, -]	71.57	83.28	78.88	66.49	56.13	60.87	70.67
PV-DBOW	Adam	logistic	[50, 50, 50]	71.57	83.25	78.64	64.72	60.87	62.74	70.61
PV-DBOW	SGD	tanh	[500, 500, 250]	71.56	83.27	78.65	66.64	56.51	61.16	70.83
PV-DBOW	SGD	tanh	[500, 250, 250]	71.56	83.29	78.52	66.81	56.66	61.32	70.95
PV-DBOW	SGD	tanh	[500, -, -]	71.56	83.28	78.76	66.76	55.91	60.85	70.76
PV-DBOW	SGD	tanh	[500, 100, 100]	71.54	83.29	78.52	66.97	56.28	61.16	70.95
PV-DBOW	SGD	tanh	[250, 50, 50]	71.54	83.29	78.78	66.11	56.51	60.93	70.55
PV-DBOW	SGD	tanh	[100, -, -]	71.54	83.24	78.62	66.25	55.23	60.24	70.37
PV-DBOW	SGD	tanh	[50, 50, -]	71.54	83.25	78.58	66.58	56.66	61.22	70.83
PV-DBOW	Adam	tanh	[500, -, -]	71.52	83.35	78.54	68.50	46.80	55.61	69.63
PV-DBOW	Adam	logistic	[100, 100, 100]	71.52	83.27	78.52	66.49	55.38	60.43	70.52
PV-DBOW	SGD	tanh	[500, 50, -]	71.50	83.29	78.6	66.67	56.73	61.30	70.89
PV-DBOW	SGD	tanh	[500, 250, -]	71.48	83.25	78.69	66.64	56.36	61.07	70.80
PV-DBOW	SGD	tanh	[500, 250, 100]	71.47	83.24	78.41	66.46	56.21	60.91	70.67
PV-DBOW	SGD	tanh	[250, 100, 100]	71.47	83.18	78.47	66.40	56.36	60.97	70.67
PV-DBOW	SGD	logistic	[50, -, -]	71.47	83.28	78.62	66.73	55.98	60.88	70.76
PV-DBOW	Adam	logistic	[100, -, -]	71.47	83.27	78.63	66.86	53.12	59.20	70.24
PV-DBOW	SGD	relu	[50, -, -]	71.45	83.28	78.38	66.11	56.36	60.84	70.52
PV-DBOW	Adam	logistic	[50, 50, -]	71.44	83.25	78.38	66.76	54.40	59.95	70.46
PV-DBOW	SGD	tanh	[500, 50, 50]	71.43	83.22	78.5	66.90	56.43	61.22	70.95
PV-DBOW	SGD	tanh	[500, 500, 100]	71.42	83.15	78.38	66.81	56.51	61.23	70.92
PV-DBOW	Adam	relu	[500, 100, -]	71.40	82.81	79.44	63.41	58.16	60.68	69.36
PV-DBOW	Adam	logistic	[500, 500, 250]	71.40	83.20	78.36	67.03	56.13	61.10	70.95
PV-DBOW	SGD	tanh	[100, 100, 50]	71.39	83.16	78.34	66.73	56.73	61.33	70.92
PV-DBOW	Adam	logistic	[500, 250, -]	71.38	83.14	78.68	67.85	51.92	58.82	70.46
PV-DBOW	Adam	logistic	[250, 250, 100]	71.38	83.19	78.73	65.24	60.72	62.90	70.89
PV-DBOW	Adam	logistic	[500, 100, 100]	71.33	83.15	78.77	66.23	57.11	61.33	70.73
PV-DBOW	Adam	tanh	[50, -, -]	71.31	83.17	78.16	65.99	56.06	60.62	70.40
PV-DBOW	SGD	relu	[500, 50, -]	71.28	82.42	79.71	61.80	59.89	60.83	68.65
PV-DBOW	Adam	logistic	[500, 500, 50]	71.28	83.10	78.52	65.50	60.72	63.02	71.04
PV-DBOW	Adam	logistic	[250, 250, 50]	71.25	83.11	78.58	65.91	58.92	62.22	70.92
PV-DBOW	Adam	logistic	[500, 50, 50]	71.23	83.28	78.68	63.57	63.43	63.50	70.37
PV-DBOW	Adam	logistic	[250, 50, 50]	71.22	83.17	78.76	67.22	51.39	58.25	70.06
PV-DBOW	Adam	logistic	[500, 500, -]	71.17	83.15	78.78	65.48	60.80	63.05	71.04
PV-DBOW	Adam	logistic	[500, 250, 50]	71.14	83.01	78.31	66.60	52.22	58.54	69.94
PV-DBOW	Adam	logistic	[250, 50, -]	71.13	83.12	78.68	65.15	60.35	62.66	70.76
PV-DBOW	Adam	logistic	[500, 500, 500]	71.10	83.13	78.21	66.91	55.68	60.78	70.80
PV-DBOW	Adam	logistic	[500, 100, -]	71.10	83.14	78.68	66.35	57.11	61.38	70.80
PV-DBOW	Adam	logistic	[250, 100, -]	71.10	83.04	78.53	66.48	54.03	59.61	70.24
PV-DBOW	Adam	logistic	[500, 100, 50]	71.09	83.11	78.61	68.99	48.38	56.88	70.18
PV-DBOW	Adam	logistic	[250, 100, 50]	71.08	83.09	78.67	66.17	57.11	61.31	70.70
PV-DBOW	Adam	logistic	[500, 250, 250]	71.06	83.04	78.17	65.62	58.01	61.58	70.58
PV-DBOW	Adam	logistic	[250, 250, 250]	71.03	82.98	78.17	66.76	52.75	58.93	70.12
PV-DBOW	Adam	logistic	[500, 250, 100]	71.00	83.07	78.6	66.58	55.91	60.78	70.67
PV-DBOW	Adam	logistic	[500, 50, -]	70.93	82.96	78.19	67.28	51.99	58.66	70.21
PV-DBOW	Adam	logistic	[250, 100, 100]	70.93	82.97	78.31	65.50	60.42	62.86	70.98
PV-DBOW	Adam	logistic	[500, 500, 100]	70.88	83.05	78.39	66.54	54.78	60.09	70.43
PV-DBOW	Adam	logistic	[100, 50, 50]	70.83	83.00	78.22	64.37	62.53	63.44	70.70
PV-DBOW	Adam	relu	[250, 100, 100]	70.15	81.72	77.38	58.63	64.94	61.62	67.13
PV-DBOW	SGD	relu	[500, 500, -]	70.06	81.79	78.32	62.04	60.27	61.15	68.87
PV-DBOW	SGD	relu	[500, 250, 250]	70.00	81.99	77.32	61.09	60.12	60.60	68.23
PV-DBOW	SGD	relu	[500, 100, -]	69.90	81.98	78.02	61.45	59.37	60.39	68.35
PV-DBOW	SGD	relu	[250, 250, 50]	69.88	81.25	77.23	60.56	56.96	58.70	67.43
PV-DBOW	SGD	relu	[500, 250, 100]	69.77	82.00	76.84	60.97	58.54	59.73	67.92
PV-DBOW	SGD	relu	[250, 250, 250]	69.76	81.39	77.73	60.29	58.62	59.44	67.49
PV-DBOW	SGD	relu	[250, 250, -]	69.74	81.84	76.81	60.16	57.71	58.91	67.28
PV-DBOW	Adam	relu	[500, -, -]	69.73	81.77	76.76	62.68	58.39	60.46	68.96
PV-DBOW	Adam	relu	[500, 50, 50]	69.46	82.04	77.34	61.98	60.35	61.15	68.84
PV-DBOW	SGD	relu	[500, 250, -]	69.44	81.52	76.18	61.30	58.16	59.69	68.07
PV-DBOW	SGD	relu	[500, 500, 250]	69.41	81.62	77.64	60.05	57.11	58.54	67.13
PV-DBOW	SGD	relu	[250, 100, -]	69.39	81.52	76.73	59.00	59.22	59.11	66.70
PV-DBOW	Adam	relu	[250, 100, -]	69.38	81.63	76.97	60.92	56.66	58.71	67.61
PV-DBOW	SGD	relu	[250, 50, -]	69.33	81.17	78.16	59.57	58.77	59.17	67.03
PV-DBOW	Adam	relu	[250, -, -]	69.32	81.48	76.87	60.46	57.41	58.90	67.43
PV-DBOW	Adam	relu	[500, 500, -]	69.27	81.24	76.81	57.47	65.99	61.44	66.33
PV-DBOW	Adam	relu	[500, 500, 500]	69.24	81.71	77.13	62.33	58.39	60.30	68.75
PV-DBOW	Adam	relu	[100, 100, -]	69.17	81.21	76.6	60.00	56.88	58.40	67.06
PV-DBOW	Adam	relu	[500, 500, 250]	69.08	81.85	76.13	60.68	65.61	63.05	68.75
PV-DBOW	Adam	relu	[100, 100, 100]	69.08	81.05	76.42	59.51	54.63	56.96	66.45

Experiments using (q, c, avg_com_q) inputs – All results (MAP>0.6)

Category	Solver	Activation	Hidden Layer	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	Adam	relu	[250, 250, 50]	69.02	80.47	76.05	60.03	55.38	57.61	66.88
PV-DBOW	SGD	relu	[500, 50, 50]	68.90	81.14	76.18	60.38	59.74	60.06	67.71
PV-DBOW	Adam	relu	[500, 100, 100]	68.89	80.91	75.52	62.33	54.03	57.88	68.04
PV-DBOW	Adam	relu	[500, 50, -]	68.85	81.47	76.52	62.55	59.44	60.96	69.05
PV-DBOW	SGD	relu	[250, 250, 100]	68.83	81.16	74.95	61.59	57.56	59.51	68.17
PV-DBOW	SGD	relu	[250, 50, 50]	68.80	80.70	76.93	59.86	56.88	58.33	66.97
PV-DBOW	Adam	relu	[500, 250, 100]	68.79	81.09	76.69	60.80	52.29	56.23	66.91
PV-DBOW	Adam	logistic	[100, 50, -]	68.77	81.34	75.84	62.20	52.75	57.08	67.77
PV-DBOW	SGD	relu	[100, 100, -]	68.70	81.13	77.42	58.21	57.64	57.92	65.96
PV-DBOW	Adam	relu	[500, 500, 50]	68.69	81.08	76.01	59.50	61.25	60.36	67.31
PV-DBOW	Adam	relu	[250, 250, 250]	68.64	80.77	75.48	61.41	56.88	59.06	67.95
PV-DBOW	Adam	relu	[250, 50, 50]	68.61	81.03	76.95	61.16	57.71	59.39	67.92
PV-DBOW	Adam	relu	[250, 250, 100]	68.59	80.82	76.68	57.77	60.72	59.21	65.99
PV-DBOW	Adam	relu	[250, 50, -]	68.59	81.24	76.08	63.08	54.63	58.55	68.56
PV-DBOW	Adam	tanh	[500, 250, 100]	68.58	80.40	77.83	60.71	58.01	59.33	67.68
PV-DBOW	SGD	relu	[250, 100, 100]	68.57	80.66	75.94	59.95	58.24	59.08	67.22
PV-DBOW	Adam	relu	[500, 500, 100]	68.55	81.43	74.75	65.69	53.72	59.11	69.79
PV-DBOW	SGD	relu	[500, 100, 50]	68.41	80.64	77.55	58.72	57.49	58.10	66.30
PV-DBOW	Adam	relu	[250, 250, -]	68.34	80.20	75.26	60.39	51.62	55.66	66.57
PV-DBOW	Adam	relu	[50, -, -]	68.30	80.92	76.38	58.55	54.85	56.64	65.87
PV-DBOW	SGD	relu	[500, 100, 100]	68.27	80.53	75.99	59.91	58.24	59.06	67.19
PV-DBOW	SGD	relu	[500, 500, 100]	68.26	80.76	75.83	61.03	57.26	59.08	67.77
PV-DBOW	Adam	relu	[250, 100, 50]	68.25	81.19	75.69	59.09	59.67	59.38	66.82
PV-DBOW	SGD	relu	[500, 250, 50]	68.24	81.22	76.21	61.37	61.32	61.35	68.59
PV-DBOW	Adam	relu	[100, -, -]	68.24	80.63	75.7	59.95	58.24	59.08	67.22
PV-DBOW	SGD	relu	[500, 500, 500]	68.18	80.54	75.2	59.22	61.63	60.40	67.16
PV-DBOW	Adam	relu	[500, 250, 50]	68.16	81.52	75.63	63.03	51.32	56.57	67.98
PV-DBOW	Adam	relu	[500, 250, -]	68.02	81.25	74.16	62.04	53.12	57.24	67.74
PV-DBOW	Adam	relu	[500, 100, 50]	67.96	81.36	74.41	64.52	52.82	58.09	69.02
PV-DBOW	SGD	relu	[500, 500, 50]	67.94	80.61	75.36	60.08	58.09	59.07	67.28
PV-DBOW	Adam	tanh	[50, 50, -]	67.92	80.10	75.67	57.01	56.88	56.95	65.05
PV-DBOW	SGD	relu	[50, 50, 50]	67.89	80.29	76.84	58.81	58.77	58.79	66.51
PV-DBOW	Adam	tanh	[100, -, -]	67.80	80.26	74.38	59.24	57.41	58.31	66.64
PV-DBOW	SGD	relu	[250, 100, 50]	67.67	80.05	74.68	58.82	55.68	57.21	66.15
PV-DBOW	Adam	relu	[100, 100, 50]	67.67	79.96	74.64	59.98	49.29	54.11	66.02
PV-DBOW	Adam	tanh	[500, 500, 50]	67.48	79.90	75.93	59.47	59.29	59.38	67.03
PV-DBOW	Adam	tanh	[500, 500, 100]	67.30	79.66	73.96	59.73	57.49	58.59	66.97
PV-DBOW	Adam	tanh	[500, 250, 250]	67.29	79.40	73.76	60.64	59.59	60.11	67.86
PV-DBOW	SGD	relu	[50, 50, -]	67.10	79.81	75.27	58.40	56.73	57.56	65.99
PV-DBOW	SGD	relu	[100, 50, -]	67.08	79.32	75.14	57.07	56.51	56.79	65.05
PV-DBOW	Adam	tanh	[250, 100, 50]	67.04	79.12	74.4	59.90	53.72	56.64	66.57
PV-DBOW	Adam	relu	[100, 50, -]	66.97	79.27	75.04	56.00	54.10	55.03	64.07
PV-DBOW	SGD	relu	[100, 50, 50]	66.96	79.70	74.92	58.33	55.30	56.78	65.78
PV-DBOW	Adam	tanh	[500, 250, 50]	66.91	79.66	74.67	60.11	58.62	59.35	67.37
PV-DBOW	Adam	tanh	[500, 100, 100]	66.89	79.68	75.45	58.27	56.21	57.22	65.84
PV-DBOW	Adam	tanh	[250, 250, -]	66.73	79.60	74.55	60.18	56.73	58.40	67.16
PV-DBOW	SGD	relu	[100, 100, 100]	66.61	79.33	74.06	58.67	57.56	58.11	66.27
PV-DBOW	Adam	tanh	[500, 250, -]	66.57	78.98	76.67	58.43	54.25	56.26	65.72
PV-DBOW	Adam	tanh	[100, 100, 100]	66.56	79.21	74.25	58.73	57.19	57.95	66.27
PV-DBOW	Adam	relu	[50, 50, 50]	66.55	79.13	73.14	58.26	58.09	58.18	66.06
PV-DBOW	Adam	relu	[50, 50, -]	66.50	79.27	73.82	56.87	52.29	54.49	64.50
PV-DBOW	Adam	tanh	[250, 100, -]	66.41	79.07	73.35	59.32	54.10	56.59	66.27
PV-DBOW	SGD	relu	[100, 100, 50]	66.35	79.98	73.89	59.94	58.09	59.00	67.19
PV-DBOW	Adam	tanh	[250, 250, 250]	66.33	79.23	74.53	59.17	57.79	58.47	66.64
PV-DBOW	Adam	relu	[100, 50, 50]	66.15	78.99	70.95	57.11	60.72	58.86	65.50
PV-DBOW	Adam	tanh	[500, 500, 250]	66.14	79.36	74.14	60.85	57.19	58.96	67.65
PV-DBOW	Adam	tanh	[500, 100, -]	66.11	78.30	73.24	56.85	55.91	56.37	64.83
PV-DBOW	Adam	tanh	[100, 100, -]	66.09	79.17	74.31	59.69	52.14	55.66	66.24
PV-DBOW	Adam	tanh	[250, 250, 50]	66.07	79.14	73.87	58.60	53.35	55.85	65.72
PV-DBOW	Adam	tanh	[500, 50, -]	65.88	78.72	73.37	58.39	55.23	56.77	65.81
PV-DBOW	Adam	tanh	[250, 250, 100]	65.88	79.05	73.31	58.86	54.48	56.58	66.02
PV-DBOW	Adam	tanh	[500, 100, 50]	65.86	79.34	74.55	59.29	57.86	58.57	66.73
PV-DBOW	Adam	tanh	[500, 500, 500]	65.76	78.48	73.9	60.56	57.19	58.82	67.46
PV-DBOW	Adam	tanh	[500, 500, -]	65.67	78.41	72.29	57.50	56.81	57.15	65.38
PV-DBOW	Adam	tanh	[250, 100, 100]	65.67	78.16	72.83	55.60	56.73	56.16	64.01
PV-DBOW	Adam	tanh	[250, 50, 50]	65.51	78.45	73.46	59.17	56.06	57.57	66.42
PV-DBOW	Adam	tanh	[100, 100, 50]	65.37	77.66	72.16	57.58	55.46	56.50	65.29
PV-DBOW	Adam	tanh	[100, 50, -]	65.37	78.44	72.16	58.42	55.61	56.98	65.87
PV-DBOW	Adam	tanh	[250, 50, -]	64.97	78.18	73.01	57.89	54.10	55.93	65.35
PV-DBOW	Adam	tanh	[500, 50, 50]	64.94	77.65	72.93	57.72	53.42	55.49	65.17
PV-DBOW	Adam	tanh	[100, 50, 50]	64.45	77.27	71.17	55.49	56.66	56.07	63.91
PV-DBOW	Adam	tanh	[50, 50, 50]	64.03	77.22	71.21	57.60	55.61	56.58	65.32

Experiments using (q, c, avg_com_q) inputs – All results (MAP>0.6)

Appendix D

Further experiments with Syntactic & Metadata features

Category	Solver	Activation	Hidden Layer	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	tanh	[250, 250, 50]	77.74	88.20	85.58	70.81	63.51	66.96	74.53
PV-DBOW	SGD	relu	[250, 100, 50]	77.43	87.96	84.81	70.57	62.60	66.35	74.19
PV-DBOW	SGD	logistic	[500, 100, 100]	77.26	87.87	85.51	71.49	63.58	67.30	74.89
PV-DBOW	SGD	logistic	[100, 100, 50]	77.26	87.98	85.1	71.22	63.13	66.93	74.65
PV-DBOW	SGD	logistic	[500, 500, 100]	77.23	87.87	85.51	70.90	63.05	66.75	74.46
PV-DBOW	SGD	tanh	[250, 100, -]	77.22	88.01	84.81	70.46	63.36	66.72	74.31
PV-DBOW	SGD	relu	[250, 250, 250]	77.21	87.88	84.88	70.71	62.68	66.45	74.28
PV-DBOW	SGD	relu	[250, 50, -]	77.21	87.85	84.7	71.17	62.98	66.83	74.59
PV-DBOW	SGD	tanh	[250, 250, -]	77.21	87.87	84.7	70.66	63.05	66.64	74.34
PV-DBOW	SGD	tanh	[50, 50, -]	77.18	87.93	85.37	70.96	63.43	66.98	74.59
PV-DBOW	SGD	relu	[250, 50, 50]	77.16	87.97	84.79	70.65	62.68	66.43	74.25
PV-DBOW	SGD	tanh	[500, 50, -]	77.16	87.92	84.98	71.49	63.96	67.51	74.98
PV-DBOW	SGD	tanh	[500, 250, 50]	77.15	87.95	84.93	71.34	63.51	67.20	74.80
PV-DBOW	SGD	tanh	[500, -, -]	77.15	87.86	84.83	70.13	63.05	66.40	74.07
PV-DBOW	SGD	tanh	[500, 100, 50]	77.12	87.96	84.94	71.25	63.96	67.41	74.86
PV-DBOW	SGD	tanh	[250, 50, -]	77.11	87.76	85.36	70.83	63.58	67.01	74.56
PV-DBOW	SGD	tanh	[250, 100, 50]	77.10	87.85	84.85	71.24	64.11	67.49	74.89
PV-DBOW	SGD	relu	[250, 250, 100]	77.09	87.79	84.72	70.32	61.85	65.81	73.88
PV-DBOW	SGD	tanh	[500, 500, 250]	77.09	87.92	85.06	70.24	63.05	66.46	74.13
PV-DBOW	SGD	tanh	[250, 250, 100]	77.09	87.89	84.69	70.81	62.60	66.45	74.31

TABLE 15: Experiments using $(q, c, avg\_com_q, ft_{(q,c)})$ inputs – All results

Category	**Solver**	**Activation**	**Hidden Layer**	**MAP**	**AvgRec**	**MRR**	**P**	**R**	**F_1**	**Acc**
PV-DBOW	SGD	tanh	[100, -, -]	77.07	87.85	84.9	71.92	63.21	67.28	75.02
PV-DBOW	SGD	tanh	[250, 250, 250]	77.05	87.91	84.69	70.23	63.36	66.61	74.19
PV-DBOW	SGD	tanh	[500, 250, 250]	77.04	87.99	85.3	70.89	63.96	67.25	74.68
PV-DBOW	SGD	tanh	[500, 500, 500]	77.03	87.85	84.91	70.69	63.88	67.11	74.56
PV-DBOW	SGD	tanh	[250, 100, 100]	77.03	87.86	84.79	70.46	62.83	66.43	74.19
PV-DBOW	SGD	logistic	[500, 500, 500]	77.02	87.71	85.03	70.88	62.83	66.61	74.40
PV-DBOW	SGD	tanh	[50, 50, 50]	77.00	87.79	84.92	70.84	63.43	66.93	74.53
PV-DBOW	SGD	logistic	[50, 50, 50]	77.00	87.75	85.18	71.09	62.90	66.75	74.53
PV-DBOW	SGD	logistic	[500, 250, 50]	76.98	87.71	84.87	71.28	62.75	66.75	74.59
PV-DBOW	SGD	logistic	[250, 250, 50]	76.98	87.73	84.98	70.79	62.90	66.61	74.37
PV-DBOW	SGD	logistic	[100, 50, -]	76.97	87.79	84.75	71.30	63.36	67.09	74.74
PV-DBOW	SGD	logistic	[500, 250, 250]	76.96	87.69	84.95	70.47	62.30	66.13	74.07
PV-DBOW	SGD	relu	[250, 100, -]	76.95	87.60	84.56	70.65	61.78	65.92	74.04
PV-DBOW	SGD	tanh	[500, 100, -]	76.95	87.81	84.34	70.65	63.21	66.72	74.37
PV-DBOW	SGD	logistic	[500, 500, -]	76.95	87.65	85.1	70.43	62.38	66.16	74.07
PV-DBOW	SGD	tanh	[500, 500, -]	76.93	87.94	84.63	70.51	63.51	66.83	74.37
PV-DBOW	SGD	tanh	[250, 50, 50]	76.92	87.76	84.9	70.50	63.66	66.90	74.40
PV-DBOW	SGD	tanh	[100, 50, -]	76.92	87.86	84.98	70.58	62.83	66.48	74.25
PV-DBOW	SGD	relu	[250, 250, -]	76.90	87.73	84.98	71.14	62.68	66.64	74.50
PV-DBOW	SGD	relu	[500, 500, 500]	76.89	87.66	83.14	70.03	63.13	66.40	74.04
PV-DBOW	SGD	tanh	[500, 100, 100]	76.89	87.95	84.98	70.96	63.43	66.98	74.59
PV-DBOW	SGD	logistic	[100, 50, 50]	76.89	87.67	84.93	71.26	62.68	66.69	74.56
PV-DBOW	SGD	tanh	[50, -, -]	76.88	87.64	84.46	71.15	63.66	67.20	74.74
PV-DBOW	SGD	logistic	[250, 50, 50]	76.88	87.63	84.85	70.99	62.60	66.53	74.40
PV-DBOW	SGD	tanh	[250, -, -]	76.87	87.82	84.94	70.32	62.75	66.32	74.10
PV-DBOW	SGD	tanh	[100, 100, 100]	76.85	87.65	84.99	71.26	62.68	66.69	74.56
PV-DBOW	SGD	logistic	[250, 100, -]	76.85	87.62	85.08	70.68	62.23	66.19	74.16
PV-DBOW	SGD	logistic	[500, 100, -]	76.84	87.67	84.98	70.55	62.38	66.21	74.13
PV-DBOW	SGD	tanh	[100, 100, -]	76.80	87.75	84.99	70.63	62.98	66.59	74.31
PV-DBOW	SGD	logistic	[500, 500, 250]	76.80	87.62	84.95	70.57	62.60	66.35	74.19
PV-DBOW	SGD	logistic	[500, 250, 100]	76.80	87.60	84.87	70.40	62.45	66.19	74.07
PV-DBOW	SGD	logistic	[50, 50, -]	76.77	87.44	84.94	70.41	62.30	66.11	74.04
PV-DBOW	SGD	relu	[100, 50, 50]	76.76	87.52	84.37	71.06	63.73	67.20	74.71
PV-DBOW	SGD	logistic	[500, 50, -]	76.76	87.51	84.82	70.45	62.08	66.00	74.01
PV-DBOW	SGD	logistic	[250, 100, 50]	76.76	87.62	84.67	70.93	62.98	66.72	74.46
PV-DBOW	SGD	tanh	[500, 50, 50]	76.75	87.70	85.12	70.54	63.05	66.59	74.28
PV-DBOW	SGD	logistic	[100, 100, -]	76.75	87.59	84.68	71.36	62.60	66.69	74.59
PV-DBOW	SGD	relu	[500, -, -]	76.74	87.71	84.19	71.59	62.38	66.67	74.65
PV-DBOW	SGD	tanh	[100, 100, 50]	76.74	87.72	84.6	70.99	62.23	66.32	74.31
PV-DBOW	SGD	relu	[250, 250, 50]	76.73	87.70	84.26	70.44	62.23	66.08	74.04
PV-DBOW	SGD	logistic	[500, 50, 50]	76.73	87.44	85.07	70.04	61.40	65.44	73.64

Experiments using $(q, c, avg\_com_q, ft_{(q,c)})$ inputs – All results

Category	**Solver**	**Activation**	**Hidden Layer**	**MAP**	**AvgRec**	**MRR**	**P**	**R**	$\mathbf{F_1}$	**Acc**
PV-DBOW	SGD	logistic	[100, 100, 100]	76.73	87.62	84.49	70.98	62.75	66.61	74.43
PV-DBOW	SGD	tanh	[500, 250, 100]	76.72	87.75	84.76	70.65	63.21	66.72	74.37
PV-DBOW	SGD	logistic	[500, 500, 50]	76.70	87.60	84.64	70.69	62.60	66.40	74.25
PV-DBOW	SGD	logistic	[250, 100, 100]	76.70	87.55	84.86	70.75	62.60	66.43	74.28
PV-DBOW	SGD	logistic	[500, 250, -]	76.69	87.49	84.93	70.13	62.53	66.11	73.94
PV-DBOW	SGD	logistic	[250, 50, -]	76.69	87.50	84.75	70.54	62.53	66.29	74.16
PV-DBOW	SGD	relu	[500, 250, 250]	76.68	87.64	84.17	71.40	62.75	66.80	74.65
PV-DBOW	SGD	tanh	[500, 250, -]	76.68	87.79	84.68	70.50	63.66	66.90	74.40
PV-DBOW	SGD	logistic	[500, 100, 50]	76.67	87.51	84.89	70.47	62.30	66.13	74.07
PV-DBOW	SGD	tanh	[500, 500, 50]	76.66	87.83	84.14	71.20	63.43	67.09	74.71
PV-DBOW	SGD	relu	[50, 50, 50]	76.64	87.43	84.63	70.42	62.15	66.03	74.01
PV-DBOW	SGD	tanh	[100, 50, 50]	76.64	87.77	84.46	71.28	63.13	66.96	74.68
PV-DBOW	SGD	logistic	[500, -, -]	76.58	87.34	84.45	70.40	61.93	65.89	73.94
PV-DBOW	SGD	logistic	[250, 250, 250]	76.57	87.50	84.57	70.56	62.23	66.13	74.10
PV-DBOW	SGD	logistic	[100, -, -]	76.57	87.53	84.58	70.95	62.30	66.35	74.31
PV-DBOW	SGD	relu	[100, 100, -]	76.56	87.63	84.67	70.48	61.63	65.76	73.91
PV-DBOW	SGD	relu	[100, 50, -]	76.56	87.44	84.81	70.71	62.30	66.24	74.19
PV-DBOW	SGD	logistic	[250, 250, -]	76.55	87.48	84.8	70.63	62.60	66.37	74.22
PV-DBOW	SGD	relu	[500, 100, 100]	76.54	87.68	83.85	70.98	63.13	66.83	74.53
PV-DBOW	SGD	logistic	[50, -, -]	76.52	87.40	84.72	70.70	61.93	66.02	74.10
PV-DBOW	SGD	relu	[500, 500, -]	76.49	87.62	84.1	71.38	63.05	66.96	74.71
PV-DBOW	SGD	relu	[500, 50, -]	76.49	87.65	84.12	71.69	63.05	67.09	74.86
PV-DBOW	SGD	relu	[500, 100, -]	76.46	87.59	84.26	71.45	63.66	67.33	74.89
PV-DBOW	SGD	tanh	[500, 500, 100]	76.46	87.71	83.81	70.53	63.21	66.67	74.31
PV-DBOW	SGD	logistic	[250, 250, 100]	76.45	87.39	84.67	69.90	62.38	65.92	73.79
PV-DBOW	SGD	relu	[50, -, -]	76.43	87.30	84.08	71.74	62.45	66.77	74.74
PV-DBOW	SGD	relu	[500, 250, -]	76.42	87.49	84.02	70.18	62.68	66.22	74.01
PV-DBOW	SGD	relu	[250, 100, 100]	76.41	87.61	84.26	71.55	62.83	66.91	74.74
PV-DBOW	SGD	relu	[500, 500, 250]	76.39	87.61	84.29	70.92	62.38	66.37	74.31
PV-DBOW	SGD	relu	[500, 500, 100]	76.39	87.54	83.89	71.42	62.60	66.72	74.62
PV-DBOW	SGD	relu	[50, 50, -]	76.38	87.44	84.46	71.39	62.90	66.88	74.68
PV-DBOW	SGD	logistic	[250, -, -]	76.37	87.29	84.33	70.21	61.70	65.68	73.79
PV-DBOW	SGD	relu	[100, 100, 50]	76.36	87.53	83.91	70.87	62.23	66.27	74.25
PV-DBOW	SGD	relu	[500, 250, 50]	76.33	87.57	84.27	71.09	62.53	66.53	74.43
PV-DBOW	SGD	relu	[500, 500, 50]	76.31	87.56	83.85	71.59	63.13	67.09	74.83
PV-DBOW	SGD	relu	[100, 100, 100]	76.29	87.38	83.68	70.38	61.85	65.84	73.91
PV-DBOW	SGD	relu	[500, 250, 100]	76.28	87.49	83.82	71.09	63.28	66.96	74.62
PV-DBOW	SGD	relu	[500, 50, 50]	76.25	87.37	83.85	71.13	61.93	66.21	74.31
PV-DBOW	SGD	relu	[500, 100, 50]	76.24	87.21	83.6	70.67	62.00	66.05	74.10
PV-DBOW	SGD	relu	[250, -, -]	76.16	87.23	83.69	70.79	61.10	65.59	73.94
PV-DBOW	SGD	relu	[100, -, -]	75.90	87.03	83.7	71.16	61.63	66.05	74.25

Experiments using $(q, c, avg\_com_q, ft_{(q,c)})$ inputs – All results

Appendix E

Analysis with $(q, c, ft_{(q,c)})$ inputs

Category	Solver	Activation	Hidden Layer	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	logistic	[500, 500, 100]	77.22	87.44	84.67	69.62	62.75	66.01	73.73
PV-DBOW	SGD	logistic	[250, 100, 50]	77.22	87.45	84.94	69.59	62.68	65.95	73.70
PV-DBOW	SGD	tanh	[500, 500, -]	77.14	87.44	84.65	69.31	63.21	66.12	73.67
PV-DBOW	SGD	logistic	[500, 250, -]	77.09	87.37	84.74	69.75	62.98	66.19	73.85
PV-DBOW	SGD	relu	[500, 250, -]	77.03	87.39	84.8	70.32	62.75	66.32	74.10
PV-DBOW	SGD	tanh	[50, 50, -]	77.02	87.40	85.01	69.51	61.93	65.50	73.49
PV-DBOW	SGD	tanh	[50, -, -]	77.02	87.41	84.7	70.14	62.23	65.95	73.88
PV-DBOW	SGD	tanh	[250, 50, 50]	77.00	87.37	84.78	69.62	62.75	66.01	73.73
PV-DBOW	SGD	logistic	[500, 100, 100]	77.00	87.24	84.68	69.80	62.60	66.01	73.79
PV-DBOW	SGD	logistic	[500, 50, -]	76.97	87.27	84.51	69.46	62.30	65.69	73.55
PV-DBOW	SGD	logistic	[100, 100, -]	76.97	87.32	84.67	69.83	63.05	66.27	73.91
PV-DBOW	SGD	tanh	[500, 500, 250]	76.93	87.38	84.92	69.84	62.90	66.19	73.88
PV-DBOW	SGD	tanh	[250, 100, 50]	76.92	87.25	84.45	69.33	62.60	65.80	73.55
PV-DBOW	SGD	relu	[500, 500, 500]	76.91	87.47	84.17	70.32	63.28	66.61	74.22
PV-DBOW	SGD	logistic	[500, 100, 50]	76.91	87.20	84.76	69.70	62.30	65.79	73.67
PV-DBOW	SGD	logistic	[250, 250, 100]	76.90	87.26	84.81	70.06	63.21	66.46	74.07
PV-DBOW	SGD	relu	[500, 100, 100]	76.88	87.23	84.73	69.89	62.00	65.71	73.70
PV-DBOW	SGD	relu	[250, 250, 50]	76.88	87.41	84.84	69.97	63.28	66.46	74.04
PV-DBOW	SGD	relu	[500, 500, 100]	76.87	87.43	84.42	69.03	62.38	65.53	73.33
PV-DBOW	SGD	logistic	[100, 100, 50]	76.87	87.32	84.52	69.73	62.75	66.06	73.79

TABLE 16: Experiments using $(q, c, ft_{(q,c)})$ inputs – All results

Category	Solver	Activation	Hidden Layer	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	logistic	[500, 500, 250]	76.85	87.24	84.58	69.66	62.53	65.90	73.70
PV-DBOW	SGD	logistic	[250, 250, -]	76.84	87.21	84.66	69.87	62.83	66.16	73.88
PV-DBOW	SGD	logistic	[100, 50, 50]	76.84	87.23	84.5	69.99	62.83	66.22	73.94
PV-DBOW	SGD	tanh	[500, 100, 50]	76.83	87.24	84.4	69.10	62.08	65.40	73.30
PV-DBOW	SGD	logistic	[500, 500, 50]	76.83	87.22	84.74	69.66	62.53	65.90	73.70
PV-DBOW	SGD	tanh	[250, 250, 250]	76.82	87.32	84.58	69.48	62.53	65.82	73.61
PV-DBOW	SGD	tanh	[250, 250, 100]	76.82	87.23	84.32	69.44	62.90	66.01	73.67
PV-DBOW	SGD	tanh	[100, 100, -]	76.82	87.34	84.52	69.62	62.75	66.01	73.73
PV-DBOW	SGD	tanh	[100, 100, 100]	76.81	87.21	84.52	69.22	63.28	66.12	73.64
PV-DBOW	SGD	tanh	[500, 500, 100]	76.80	87.31	84.77	69.30	62.15	65.53	73.43
PV-DBOW	SGD	tanh	[250, 50, -]	76.80	87.19	84.45	70.25	63.43	66.67	74.22
PV-DBOW	SGD	relu	[250, 250, 100]	76.79	87.31	84.6	69.56	62.08	65.61	73.55
PV-DBOW	SGD	logistic	[500, 50, 50]	76.79	87.12	84.35	69.92	62.45	65.98	73.82
PV-DBOW	SGD	logistic	[250, 250, 50]	76.79	87.26	84.37	69.71	62.68	66.01	73.76
PV-DBOW	SGD	relu	[500, 250, 250]	76.78	87.21	84.02	70.14	62.23	65.95	73.88
PV-DBOW	SGD	relu	[250, 250, -]	76.78	87.27	84.53	69.20	62.38	65.61	73.43
PV-DBOW	SGD	tanh	[500, 50, -]	76.78	87.32	84.81	69.75	62.98	66.19	73.85
PV-DBOW	SGD	logistic	[250, 100, 100]	76.77	87.17	84.63	69.60	62.53	65.87	73.67
PV-DBOW	SGD	relu	[50, -, -]	76.76	87.22	84.54	70.07	61.47	65.49	73.67
PV-DBOW	SGD	tanh	[250, -, -]	76.76	87.25	84.51	69.71	62.68	66.01	73.76
PV-DBOW	SGD	logistic	[500, 500, -]	76.76	87.24	84.44	69.68	62.60	65.95	73.73
PV-DBOW	SGD	logistic	[500, 100, -]	76.75	87.16	84.56	69.82	62.68	66.06	73.82
PV-DBOW	SGD	tanh	[250, 100, -]	76.74	87.22	84.63	69.78	62.90	66.17	73.85
PV-DBOW	SGD	logistic	[250, 100, -]	76.72	87.19	84.46	69.98	63.51	66.59	74.10
PV-DBOW	SGD	logistic	[500, 250, 100]	76.70	87.16	84.45	69.58	62.30	65.74	73.61
PV-DBOW	SGD	logistic	[250, 250, 250]	76.70	87.20	84.55	69.63	62.98	66.14	73.79
PV-DBOW	SGD	relu	[100, 50, 50]	76.69	87.18	84.15	69.66	61.32	65.23	73.43
PV-DBOW	SGD	tanh	[500, 250, 100]	76.69	87.09	84.11	69.55	62.90	66.06	73.73
PV-DBOW	SGD	logistic	[250, 50, 50]	76.69	87.26	84.52	69.75	62.98	66.19	73.85
PV-DBOW	SGD	logistic	[50, 50, 50]	76.69	87.25	84.32	70.08	62.90	66.30	74.01
PV-DBOW	SGD	tanh	[100, 50, -]	76.67	87.16	84.3	69.63	63.13	66.22	73.82
PV-DBOW	SGD	tanh	[50, 50, 50]	76.67	87.28	84.92	70.28	62.45	66.14	74.01
PV-DBOW	SGD	logistic	[100, 100, 100]	76.67	87.17	84.33	69.65	63.21	66.27	73.85
PV-DBOW	SGD	tanh	[500, 500, 500]	76.66	87.16	84.23	69.46	62.98	66.06	73.70
PV-DBOW	SGD	tanh	[250, 250, 50]	76.66	87.26	84.63	69.15	62.23	65.50	73.36
PV-DBOW	SGD	relu	[250, 250, 250]	76.65	87.19	83.83	69.54	63.05	66.14	73.76
PV-DBOW	SGD	tanh	[100, 50, 50]	76.65	87.10	84.15	69.85	63.28	66.40	73.98
PV-DBOW	SGD	relu	[500, 100, -]	76.64	87.16	84.26	69.70	62.15	65.71	73.64
PV-DBOW	SGD	tanh	[500, -, -]	76.64	87.22	84.59	69.86	63.13	66.32	73.94
PV-DBOW	SGD	tanh	[500, 250, -]	76.62	87.32	84.29	69.81	63.51	66.51	74.01
PV-DBOW	SGD	tanh	[100, 100, 50]	76.60	87.18	84.68	69.31	62.53	65.74	73.52

Experiments using $(q, c, ft_{(q,c)})$ inputs – All results

Category	**Solver**	**Activation**	**Hidden Layer**	**MAP**	**AvgRec**	**MRR**	**P**	**R**	**F_1**	**Acc**
PV-DBOW	SGD	logistic	[100, 50, -]	76.60	87.10	84.04	69.52	62.98	66.09	73.73
PV-DBOW	SGD	relu	[500, 250, 100]	76.59	87.35	83.84	69.35	62.83	65.93	73.61
PV-DBOW	SGD	relu	[250, -, -]	76.58	87.09	84.34	70.14	62.75	66.24	74.01
PV-DBOW	SGD	relu	[100, 100, 100]	76.58	87.05	84.1	69.40	62.30	65.66	73.52
PV-DBOW	SGD	logistic	[500, 250, 250]	76.58	87.10	84.29	69.51	62.60	65.87	73.64
PV-DBOW	SGD	relu	[500, 50, -]	76.57	87.11	83.79	70.79	62.90	66.61	74.37
PV-DBOW	SGD	relu	[100, -, -]	76.57	86.99	83.73	69.99	62.30	65.92	73.82
PV-DBOW	SGD	tanh	[500, 500, 50]	76.57	87.21	84.26	68.99	61.93	65.27	73.21
PV-DBOW	SGD	relu	[500, 100, 50]	76.55	87.13	84.15	69.37	61.85	65.39	73.39
PV-DBOW	SGD	relu	[500, 50, 50]	76.55	87.20	84.08	69.74	62.60	65.98	73.76
PV-DBOW	SGD	logistic	[500, 500, 500]	76.55	87.10	84.57	69.46	62.45	65.77	73.58
PV-DBOW	SGD	logistic	[500, 250, 50]	76.55	87.13	84.36	69.61	62.38	65.79	73.64
PV-DBOW	SGD	relu	[250, 100, -]	76.54	87.12	84.42	69.99	62.30	65.92	73.82
PV-DBOW	SGD	relu	[250, 50, -]	76.54	87.21	84.42	69.43	62.38	65.72	73.55
PV-DBOW	SGD	logistic	[250, 50, -]	76.53	87.12	84.25	69.69	62.45	65.87	73.70
PV-DBOW	SGD	logistic	[50, 50, -]	76.51	87.15	84.2	69.88	62.68	66.08	73.85
PV-DBOW	SGD	relu	[500, 250, 50]	76.50	87.13	83.92	69.71	62.68	66.01	73.76
PV-DBOW	SGD	tanh	[500, 250, 50]	76.50	87.12	83.87	69.31	62.53	65.74	73.52
PV-DBOW	SGD	tanh	[500, 50, 50]	76.50	87.15	83.85	68.61	62.00	65.14	73.03
PV-DBOW	SGD	tanh	[100, -, -]	76.50	87.05	83.71	69.49	62.38	65.74	73.58
PV-DBOW	SGD	tanh	[500, 100, 100]	76.49	87.13	84.31	69.45	62.60	65.85	73.61
PV-DBOW	SGD	tanh	[250, 250, -]	76.48	87.22	84.4	69.26	62.90	65.93	73.58
PV-DBOW	SGD	tanh	[250, 100, 100]	76.48	87.07	84.08	68.63	62.38	65.35	73.12
PV-DBOW	SGD	tanh	[500, 100, -]	76.47	87.29	84.32	69.11	63.13	65.99	73.55
PV-DBOW	SGD	relu	[100, 100, 50]	76.46	87.03	83.7	69.39	62.08	65.53	73.46
PV-DBOW	SGD	tanh	[500, 250, 250]	76.45	87.21	84.27	69.09	62.90	65.85	73.49
PV-DBOW	SGD	relu	[500, 500, -]	76.44	87.16	83.77	69.70	62.30	65.79	73.67
PV-DBOW	SGD	relu	[250, 100, 100]	76.44	86.97	84.35	70.03	62.23	65.90	73.82
PV-DBOW	SGD	relu	[250, 100, 50]	76.44	86.99	83.81	69.59	62.15	65.66	73.58
PV-DBOW	SGD	relu	[50, 50, 50]	76.42	87.13	84.06	70.06	62.15	65.87	73.82
PV-DBOW	SGD	relu	[50, 50, -]	76.42	86.99	84.12	69.09	60.87	64.72	73.03
PV-DBOW	SGD	relu	[100, 100, -]	76.38	87.05	83.68	69.17	62.30	65.56	73.39
PV-DBOW	SGD	relu	[100, 50, -]	76.37	87.02	83.76	69.51	62.45	65.79	73.61
PV-DBOW	SGD	relu	[500, 500, 250]	76.34	87.15	83.73	69.38	62.75	65.90	73.61
PV-DBOW	SGD	relu	[500, 500, 50]	76.34	87.09	83.46	69.50	62.23	65.66	73.55
PV-DBOW	SGD	logistic	[50, -, -]	76.30	87.00	83.47	69.96	62.38	65.95	73.82
PV-DBOW	SGD	relu	[250, 50, 50]	76.16	87.03	83.51	69.62	62.60	65.93	73.70
PV-DBOW	SGD	relu	[500, -, -]	76.14	87.04	83.65	70.14	62.23	65.95	73.88
PV-DBOW	SGD	logistic	[250, -, -]	76.00	86.77	83.55	69.70	62.30	65.79	73.67
PV-DBOW	SGD	logistic	[500, -, -]	75.95	86.74	83.68	69.44	62.23	65.63	73.52
PV-DBOW	SGD	logistic	[100, -, -]	75.92	86.74	83.46	69.73	62.23	65.77	73.67

Experiments using $(q, c, ft_{(q,c)})$ inputs – All results

Appendix F

Analysis with only $(ft_{(q,c)})$ inputs

Category	Solver	Activation	Hidden Layer	MAP	AvgRec	MRR	P	R	F_1	Acc
PV-DBOW	SGD	tanh	[250, 100, 50]	0.7511	0.8663	81.78	0.6993	0.5914	0.6408	0.7306
PV-DBOW	SGD	relu	[500, 100, 100]	0.7484	0.8625	81.61	0.7009	0.5959	0.6442	0.7324
PV-DBOW	SGD	tanh	[500, 500, 50]	0.7484	0.8633	81.16	0.6969	0.5952	0.642	0.7303
PV-DBOW	SGD	tanh	[500, 500, 500]	0.7481	0.8634	81.86	0.6935	0.6027	0.6449	0.7303
PV-DBOW	SGD	tanh	[500, 500, 100]	0.7477	0.8636	81.69	0.6968	0.5967	0.6429	0.7306
PV-DBOW	SGD	relu	[250, 100, 50]	0.7476	0.8627	81.67	0.7006	0.6005	0.6467	0.7333
PV-DBOW	SGD	tanh	[500, 100, 100]	0.7475	0.8629	81.46	0.6928	0.5922	0.6385	0.7275
PV-DBOW	SGD	tanh	[100, 50, -]	0.7475	0.8628	81.61	0.6946	0.5869	0.6362	0.7272
PV-DBOW	SGD	relu	[500, 500, 50]	0.7473	0.863	81.6	0.6994	0.5899	0.64	0.7303
PV-DBOW	SGD	tanh	[500, 500, 250]	0.7469	0.8629	81.4	0.6927	0.6005	0.6433	0.7294
PV-DBOW	SGD	relu	[500, 250, -]	0.7467	0.8626	81.66	0.6988	0.5899	0.6397	0.73
PV-DBOW	SGD	tanh	[250, 250, 50]	0.7467	0.8627	81.69	0.6937	0.5914	0.6385	0.7278
PV-DBOW	SGD	relu	[250, 250, 250]	0.7466	0.8624	81.37	0.6952	0.5922	0.6396	0.7287
PV-DBOW	SGD	relu	[250, 250, 100]	0.7466	0.8613	81.3	0.6999	0.5967	0.6442	0.7321
PV-DBOW	SGD	tanh	[500, 100, -]	0.7464	0.8623	81.85	0.6921	0.5937	0.6391	0.7275
PV-DBOW	SGD	tanh	[250, 50, 50]	0.7464	0.8624	81.42	0.6981	0.602	0.6465	0.7324
PV-DBOW	SGD	relu	[500, 50, 50]	0.7463	0.8625	81.58	0.6962	0.6035	0.6465	0.7318
PV-DBOW	SGD	tanh	[500, 250, 100]	0.7462	0.8616	81.55	0.6931	0.5914	0.6382	0.7275
PV-DBOW	SGD	tanh	[500, 50, -]	0.7461	0.8619	81.43	0.6953	0.5974	0.6427	0.73
PV-DBOW	SGD	tanh	[500, 50, 50]	0.7458	0.8621	81.24	0.6932	0.5967	0.6413	0.7287

TABLE 17: Experiments using $(ft_{(q,c)})$ inputs – All results

Category	**Solver**	**Activation**	**Hidden Layer**	**MAP**	**AvgRec**	**MRR**	**P**	**R**	$\mathbf{F_1}$	**Acc**
PV-DBOW	SGD	relu	[500, 250, 100]	0.7453	0.861	81.81	0.698	0.5929	0.6412	0.7303
PV-DBOW	SGD	relu	[100, 50, 50]	0.7453	0.8616	81.73	0.6983	0.6042	0.6478	0.733
PV-DBOW	SGD	tanh	[500, 250, -]	0.7452	0.8617	81.37	0.6896	0.5967	0.6398	0.7269
PV-DBOW	SGD	tanh	[100, 50, 50]	0.7452	0.8628	81.17	0.689	0.5869	0.6339	0.7245
PV-DBOW	SGD	tanh	[500, 250, 250]	0.7451	0.8618	81.61	0.6923	0.5959	0.6405	0.7281
PV-DBOW	SGD	tanh	[500, 100, 50]	0.7451	0.8613	81.58	0.6922	0.5974	0.6414	0.7284
PV-DBOW	SGD	relu	[250, 50, 50]	0.7449	0.861	81.38	0.6921	0.5937	0.6391	0.7275
PV-DBOW	SGD	tanh	[500, 500, -]	0.7449	0.8619	81.49	0.6912	0.5997	0.6422	0.7284
PV-DBOW	SGD	tanh	[100, 100, 100]	0.7448	0.8629	81.43	0.6968	0.5967	0.6429	0.7306
PV-DBOW	SGD	relu	[500, 500, -]	0.7446	0.8612	81.31	0.6968	0.5914	0.6398	0.7294
PV-DBOW	SGD	relu	[100, 100, 100]	0.7445	0.8609	81.31	0.7002	0.6027	0.6478	0.7336
PV-DBOW	SGD	relu	[500, 250, 250]	0.7443	0.86	81.15	0.6932	0.5899	0.6374	0.7272
PV-DBOW	SGD	relu	[500, 500, 250]	0.7442	0.862	81.29	0.6972	0.5892	0.6387	0.7291
PV-DBOW	SGD	relu	[250, 250, 50]	0.7442	0.8617	81.36	0.6955	0.5997	0.644	0.7306
PV-DBOW	SGD	tanh	[100, -, -]	0.7442	0.861	81.46	0.6981	0.5899	0.6395	0.7297
PV-DBOW	SGD	relu	[100, 100, 50]	0.7441	0.8612	81.32	0.6964	0.5989	0.644	0.7309
PV-DBOW	SGD	relu	[100, 100, -]	0.744	0.8603	81.3	0.6927	0.6005	0.6433	0.7294
PV-DBOW	SGD	tanh	[250, 250, 250]	0.7439	0.8612	81.39	0.7008	0.608	0.6511	0.7352
PV-DBOW	SGD	tanh	[250, 100, 100]	0.7439	0.8603	81.2	0.6961	0.605	0.6473	0.7321
PV-DBOW	SGD	relu	[500, 500, 500]	0.7438	0.8604	82.1	0.6904	0.5974	0.6406	0.7275
PV-DBOW	SGD	relu	[500, 500, 100]	0.7438	0.8608	81.32	0.6958	0.5869	0.6367	0.7278
PV-DBOW	SGD	tanh	[50, 50, -]	0.7437	0.8612	81.68	0.6914	0.5952	0.6397	0.7275
PV-DBOW	SGD	relu	[250, 100, 100]	0.7436	0.8601	81.53	0.699	0.5907	0.6403	0.7303
PV-DBOW	SGD	tanh	[250, -, -]	0.7436	0.8602	80.78	0.693	0.5997	0.643	0.7294
PV-DBOW	SGD	tanh	[500, -, -]	0.7435	0.8604	80.99	0.6945	0.5952	0.641	0.7291
PV-DBOW	SGD	relu	[500, 100, 50]	0.7434	0.86	81.56	0.693	0.6012	0.6438	0.7297
PV-DBOW	SGD	tanh	[250, 250, -]	0.7434	0.8615	80.91	0.6966	0.6132	0.6523	0.7343
PV-DBOW	SGD	relu	[500, 100, -]	0.7433	0.8606	81.19	0.697	0.5937	0.6412	0.73
PV-DBOW	SGD	relu	[50, 50, -]	0.7433	0.8591	81.28	0.6981	0.5952	0.6426	0.7309
PV-DBOW	SGD	logistic	[250, 100, 100]	0.7433	0.8598	81.35	0.6896	0.5967	0.6398	0.7269
PV-DBOW	SGD	relu	[500, 50, -]	0.743	0.8621	81.36	0.6948	0.5892	0.6376	0.7278
PV-DBOW	SGD	relu	[250, 50, -]	0.7427	0.8603	81.24	0.6957	0.6005	0.6446	0.7309
PV-DBOW	SGD	tanh	[250, 50, -]	0.7427	0.8611	81.11	0.6964	0.5989	0.644	0.7309
PV-DBOW	SGD	logistic	[100, 100, 50]	0.7427	0.8598	80.94	0.6938	0.6035	0.6455	0.7306
PV-DBOW	SGD	tanh	[500, 250, 50]	0.7425	0.86	81.13	0.6928	0.5974	0.6416	0.7287
PV-DBOW	SGD	logistic	[250, 250, 50]	0.7423	0.8594	81.35	0.6886	0.5974	0.6398	0.7266
PV-DBOW	SGD	relu	[250, 100, -]	0.7421	0.8601	81.38	0.69	0.5944	0.6386	0.7266
PV-DBOW	SGD	tanh	[100, 100, 50]	0.742	0.86	81.19	0.6961	0.5982	0.6435	0.7306
PV-DBOW	SGD	logistic	[50, 50, 50]	0.742	0.8592	81.13	0.692	0.602	0.6439	0.7294
PV-DBOW	SGD	relu	[250, 250, -]	0.7419	0.8592	81.21	0.6886	0.5989	0.6406	0.7269
PV-DBOW	SGD	relu	[500, 250, 50]	0.7418	0.8604	80.92	0.6965	0.5922	0.6401	0.7294

Experiments using $(ft_{(q,c)})$ inputs – All results

Category	**Solver**	**Activation**	**Hidden Layer**	**MAP**	**AvgRec**	**MRR**	**P**	**R**	**F_1**	**Acc**
PV-DBOW	SGD	logistic	[500, 250, 50]	0.7418	0.8569	82.12	0.6807	0.5952	0.6351	0.722
PV-DBOW	SGD	logistic	[250, 250, 250]	0.7417	0.8591	81.02	0.6893	0.5959	0.6392	0.7266
PV-DBOW	SGD	logistic	[250, 50, 50]	0.7417	0.8589	80.94	0.6885	0.6005	0.6415	0.7272
PV-DBOW	SGD	tanh	[250, 100, -]	0.7415	0.8589	80.61	0.6972	0.6012	0.6457	0.7318
PV-DBOW	SGD	logistic	[100, 100, -]	0.7415	0.8585	81.04	0.6947	0.6027	0.6454	0.7309
PV-DBOW	SGD	tanh	[100, 100, -]	0.7414	0.8588	81.01	0.6899	0.5892	0.6356	0.7254
PV-DBOW	SGD	logistic	[500, 500, 50]	0.7412	0.8562	81.93	0.6809	0.5989	0.6373	0.7229
PV-DBOW	SGD	logistic	[500, 500, -]	0.7412	0.8584	81.54	0.6873	0.5937	0.6371	0.7251
PV-DBOW	SGD	relu	[100, 50, -]	0.7411	0.8601	81.79	0.6891	0.5937	0.6378	0.726
PV-DBOW	SGD	tanh	[250, 250, 100]	0.7411	0.8583	80.67	0.6896	0.6035	0.6437	0.7284
PV-DBOW	SGD	logistic	[50, -, -]	0.7411	0.8583	81.34	0.6915	0.5937	0.6389	0.7272
PV-DBOW	SGD	logistic	[500, 50, -]	0.741	0.8574	81.88	0.6828	0.5959	0.6364	0.7232
PV-DBOW	SGD	logistic	[500, 250, 250]	0.7408	0.8573	81.85	0.6839	0.5959	0.6369	0.7239
PV-DBOW	SGD	relu	[50, 50, 50]	0.7406	0.8586	80.62	0.6911	0.5892	0.6361	0.726
PV-DBOW	SGD	logistic	[500, 500, 250]	0.7406	0.858	81.57	0.6839	0.5959	0.6369	0.7239
PV-DBOW	SGD	tanh	[50, -, -]	0.7403	0.8573	81.15	0.6922	0.5839	0.6335	0.7254
PV-DBOW	SGD	logistic	[500, 100, 50]	0.7402	0.8573	81.53	0.6799	0.5944	0.6343	0.7214
PV-DBOW	SGD	logistic	[500, 100, -]	0.7402	0.8578	81.08	0.6916	0.5922	0.638	0.7269
PV-DBOW	SGD	logistic	[100, 100, 100]	0.7402	0.8579	80.61	0.6934	0.6042	0.6458	0.7306
PV-DBOW	SGD	logistic	[100, 50, 50]	0.74	0.8577	80.66	0.6922	0.6057	0.6461	0.7303
PV-DBOW	SGD	tanh	[50, 50, 50]	0.7399	0.8562	80.51	0.6778	0.5952	0.6338	0.7205
PV-DBOW	SGD	logistic	[500, 50, 50]	0.7399	0.8581	81.11	0.6884	0.5952	0.6384	0.726
PV-DBOW	SGD	logistic	[500, 100, 100]	0.7398	0.8554	81.26	0.6795	0.5997	0.6371	0.7223
PV-DBOW	SGD	logistic	[500, 500, 500]	0.7397	0.8579	81.3	0.6855	0.5937	0.6363	0.7242
PV-DBOW	SGD	logistic	[500, -, -]	0.7393	0.8569	81.45	0.6843	0.5952	0.6366	0.7239
PV-DBOW	SGD	logistic	[250, 50, -]	0.7388	0.8572	81.29	0.6873	0.602	0.6418	0.7269
PV-DBOW	SGD	logistic	[250, 250, 100]	0.7386	0.8552	81.32	0.6834	0.6042	0.6414	0.7254
PV-DBOW	SGD	logistic	[500, 250, 100]	0.7384	0.8571	80.92	0.6845	0.5959	0.6372	0.7242
PV-DBOW	SGD	logistic	[250, -, -]	0.7384	0.8566	81.26	0.6842	0.5982	0.6383	0.7245
PV-DBOW	SGD	logistic	[500, 250, -]	0.7378	0.8569	80.81	0.6928	0.5974	0.6416	0.7287
PV-DBOW	SGD	logistic	[250, 100, -]	0.7378	0.8544	80.93	0.6814	0.605	0.6409	0.7245
PV-DBOW	SGD	logistic	[100, 50, -]	0.7378	0.8562	80.76	0.6928	0.6042	0.6455	0.7303
PV-DBOW	SGD	logistic	[50, 50, -]	0.7378	0.8574	80.84	0.6897	0.602	0.6428	0.7281
PV-DBOW	SGD	logistic	[500, 500, 100]	0.7377	0.8567	80.93	0.6824	0.5967	0.6367	0.7232
PV-DBOW	SGD	relu	[500, -, -]	0.7371	0.8568	80.9	0.6863	0.5794	0.6283	0.7214
PV-DBOW	SGD	logistic	[250, 100, 50]	0.7369	0.8561	80.96	0.6843	0.602	0.6405	0.7254
PV-DBOW	SGD	logistic	[100, -, -]	0.7369	0.8552	80.92	0.6912	0.5944	0.6392	0.7272
PV-DBOW	SGD	relu	[250, -, -]	0.7359	0.8557	80.8	0.688	0.5892	0.6348	0.7245
PV-DBOW	SGD	relu	[100, -, -]	0.7359	0.8558	80.4	0.6957	0.5779	0.6313	0.7257
PV-DBOW	SGD	logistic	[250, 250, -]	0.7358	0.8555	80.62	0.6864	0.6027	0.6418	0.7266
PV-DBOW	SGD	relu	[50, -, -]	0.7334	0.8532	80.66	0.6863	0.5696	0.6225	0.7193

Experiments using $(ft_{(q,c)})$ inputs – All results

Appendix G

SemEval '16 Task 3 – Subtask A

Submission	MAP	AvgRec	MRR	P	R	F_1	Acc
Kelp-primary	79.19	88.82	86.42	76.96	55.30	64.36	75.11
ConvKN-contrastive1	78.71	88.98	86.15	77.78	53.72	63.55	74.95
SUper_team-contrastive1	77.68	88.06	84.76	75.59	55.00	63.68	74.50
ConvKN-primary	77.66	88.05	84.93	75.56	58.84	66.16	75.54
SemanticZ-primary	77.58	88.14	85.21	74.13	53.05	61.84	73.39
ConvKN-contrastive2	77.29	87.77	85.03	74.74	59.67	66.36	75.41
ECNU-primary	77.28	87.52	84.09	70.46	63.36	66.72	74.31
SemanticZ-contrastive1	77.16	87.73	84.08	75.29	53.20	62.35	73.88
SUper_team-primary	77.16	87.98	84.69	74.43	56.73	64.39	74.50
MTE-NN-contrastive2	76.98	86.98	85.50	58.71	70.28	63.97	67.83
SUper_team-contrastive2	76.97	87.89	84.58	74.31	56.36	64.10	74.34
MTE-NN-contrastive1	76.86	87.03	84.36	55.84	77.35	64.86	65.93
SLS-contrastive2	76.71	87.17	84.38	59.45	67.95	63.41	68.13
SLS-contrastive1	76.46	87.47	83.27	60.09	69.68	64.53	68.87
MTE-NN-primary	76.44	86.74	84.97	56.28	76.22	64.75	66.27
SLS-primary	76.33	87.30	82.99	60.36	67.72	63.83	68.81
ECNU-contrastive2	75.71	86.14	82.53	63.60	66.67	65.10	70.95
SemanticZ-contrastive2	75.41	86.51	82.52	73.19	50.11	59.49	72.26
ICRC-HIT-contrastive1	73.34	84.81	79.65	63.43	69.30	66.24	71.28
ITNLP-AiKF-primary	71.52	82.67	80.26	73.18	19.71	31.06	64.43
ECNU-contrastive1	71.34	83.39	78.62	66.95	41.31	51.09	67.86
ICRC-HIT-primary	70.90	83.36	77.38	62.48	62.53	62.50	69.51
PMI-cool-primary	68.79	79.94	80.00	47.81	70.58	57.00	56.73
UH-PRHLT-contrastive1	67.57	79.50	77.08	54.10	50.11	52.03	62.45
Baseline (chronological)	59.53	72.60	67.83	—	—	—	—
Baseline (random)	52.80	66.52	58.71	40.56	74.57	52.55	45.26
Baseline (all 'true')	—	—	—	40.64	100.00	57.80	40.64
Baseline (all 'false')	—	—	—	—	—	—	59.36

Table 18: All submissions and Baselines at SemEval '16 Task 3 – Subtask A

Bibliography

AlessandroMoschitti, P. L., AbedAlhakimFreihat, W. H., Glass, J., and Randeree, B. (2016). Semeval-2016 task 3: Community question answering. *Proceedings of SemEval*, pages 525–545.

Barrón-Cedeno, A., Filice, S., Da San Martino, G., Joty, S. R., Màrquez, L., Nakov, P., and Moschitti, A. (2015). Thread-level information for comment classification in community question answering. In *ACL (2)*, pages 687–693. Citeseer.

Bogdanova, D. and Foster, J. (2016). This is how we do it: Answer reranking for open-domain how questions with paragraph vectors and minimal feature engineering.

Filice, S., Croce, D., Moschitti, A., and Basili, R. (2016). Kelp at semeval-2016 task 3: Learning semantic relations between questions and answers. *Proceedings of SemEval*, 16:1116–1123.

Heilman, M. and Smith, N. A. (2010). Tree edit models for recognizing textual entailments, paraphrases, and answers to questions. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 1011–1019. Association for Computational Linguistics.

Jansen, P., Surdeanu, M., and Clark, P. (2014). Discourse complements lexical semantics for non-factoid answer reranking. In *ACL (1)*, pages 977–986.

Joty, S., Moschitti, A., Al Obaidli, F. A., Romeo, S., Tymoshenko, K., and Uva, A. (2016). Convkn at semeval-2016 task 3: Answer and question selection for question answering on arabic and english fora. *Proceedings of SemEval*, pages 896–903.

Lafferty, J., McCallum, A., Pereira, F., et al. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the eighteenth international conference on machine learning, ICML*, volume 1, pages 282–289.

Le, Q. V. and Mikolov, T. (2014). Distributed representations of sentences and documents. In *ICML*, volume 14, pages 1188–1196.

Lin, X. Z. B. H. J. and Wang, Y. X. X. (2015). Icrc-hit: A deep learning based comment sequence labeling system for answer selection challenge. *SemEval-2015*, 210.

Mihaylov, T. and Nakov, P. (2016). Semanticz at semeval-2016 task 3: Ranking relevant answers in community question answering using semantic similarity based on fine-tuned word embeddings. *Proceedings of SemEval*, pages 879–886.

Mihaylova, T., Gencheva, P., Boyanov, M., Yovcheva, I., Mihaylov, T., Hardalov, M., Kiprov, Y., Balchev, D., Koychev, I., Nakov, P., et al. (2016). Super team at semeval-2016 task 3: Building a feature-rich system for community question answering. *Proceedings of SemEval*, pages 836–843.

Nakov, P., Màrquez, L., Magdy, W., Moschitti, A., Glass, J., and Randeree, B. (2016). SemEval-2016 task 3: Community question answering. In *Proceedings of the 10th International Workshop on Semantic Evaluation*, SemEval '16, San Diego, California. Association for Computational Linguistics.

Rehurek, R. and Sojka, P. (2010). Software framework for topic modelling with large corpora. In *In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. Citeseer.

Severyn, A. and Moschitti, A. (2012). Structural relationships for large-scale learning of answer re-ranking. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pages 741–750. ACM.

Severyn, A. and Moschitti, A. (2013). Automatic feature engineering for answer selection and extraction. In *EMNLP*, volume 13, pages 458–467.

Toutanova, K., Klein, D., Manning, C. D., and Singer, Y. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*, pages 173–180. Association for Computational Linguistics.

Tran, Q. H., Tran, V., Vu, T., Nguyen, M., and Pham, S. B. (2015). Jaist: Combining multiple features for answer selection in community question answering. In *Proceedings of the 9th International Workshop on Semantic Evaluation, SemEval*, volume 15, pages 215–219.

Wang, M. and Manning, C. D. (2010). Probabilistic tree-edit models with structured latent variables for textual entailment and question answering. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 1164–1172. Association for Computational Linguistics.

Wang, M., Smith, N. A., and Mitamura, T. (2007). What is the jeopardy model? a quasi-synchronous grammar for qa. In *EMNLP-CoNLL*, volume 7, pages 22–32.

Yao, X., Van Durme, B., Callison-Burch, C., and Clark, P. (2013). Answer extraction as sequence tagging with tree edit distance. In *HLT-NAACL*, pages 858–867. Citeseer.

Zhou, X., Hu, B., Chen, Q., Tang, B., and Wang, X. (2015). Answer sequence learning with neural networks for answer selection in community question answering. *arXiv preprint arXiv:1506.06490*.