ISSN 2320 - 2602
Volume 3, No.1, January 2014
International Journal of Advances in Computer Science and Technology
Available Online at http://warse.org/pdfs/2014/ijacst04312014.pdf

Entity Disambiguation With Comparable Entity Mining from Comparative Questions

S.Deepa [1], Ms.S.Vasumathi Kannagi [2]
[1]ME student , Department of Computer Science and engineering, Info Institute Of Engineering , Kovilpalayam Coimbatore-641107, Tamilnadu, India, sdeepa153@gmail.com.
[2]Assistant Professor, Department of Computer Science and engineering, Info Institute Of Engineering , Kovilpalayam Coimbatore-641107, Tamilnadu, India, vasumathi.arun@gmail.com

ABSTRACT

Comparing substitute option is one of the necessary processes for making decision to take out each day. Though, it is not forever simple task to compare and other alternatives for each and every comparison. To perform this earlier work used a novel method to automatically extract equivalent entities from relative questions that are posted by online for each and every user .But the existing work it becomes difficult task to solve the entity ambiguity problem. To conquer these difficulty proposition numerous disambiguation formula/features and utilize a Markov logic network(MLN) to representation of interweaved constraints. It is one of the major types of entity linking method with genetic material state relating. Proposed MLN which is the combination of first order logic (FOL) and Markov networks with combination of NIL-filtering and entity disambiguation stages. For entity disambiguation problem the representation capture the entity information from background knowledge with familiar entities as well as the constraints while connecting an entity.

Keywords: Markov Logic Network (MLN), Comparative Questions, Information Extraction, Bootstrapping, Sequential Pattern Mining and Comparable Entity Mining.

1. INTRODUCTION

In this article, PSO based ontology system is anticipated for knowledge illustration and examination in overkill of user profiles in local repositories as fine as global knowledge phase. PSO based ontology system analysis ontological user profiles from mutually a world knowledge base and user local instance repositories. The ontology model is evaluated by comparing it against existing models in web information gathering, it shows that PSO ontology better results.

Many of the existing question answering systems methods uses exterior information and tools for respond analytical. For reference it uses entity

taggers, WordNet, specific parsers and ontology list. Though, at the latest TREC-10 QA assessment, the attractive scheme second-hand immediately single resource. The obvious authority of such patterns stunned numerous. To handle this accordingly determined to examine their possible by acquiring patterns routinely and to determine their accurateness.

In the world wide wed the comparison of search results with similar concept or similar information via search the relevant pages regarding the targeted products, discover contending products, understand writing review, and recognize pros. In this paper focal point of discovering set of comparable pair of entities .Generally it becomes complex to choose if together entities are equivalent for a variety of reasons. To overcome this problem entity linking helps to study the possible information from background knowledge many disambiguation move toward have been planned to deal with the entity ambiguity difficulty. For instance, Dredze et al [1] proposed the disambiguation mission as a ranking difficulty and developed features to link Wikipedia entries.

Zhang et. al. [3] second-hand be automatically generate the quantity to instruct a dual classifier to reduce ambiguity. Dai et al. [2] composed exterior information for every entity and intended likelihoods stating the correspondence of the present textbook by means of the information to get better the disambiguation presentation. In adding together to the entity ambiguity difficulty, the EL task in Text Analysis Conference (TAC) 2009 establish the nonappearance concern McNamee et.al [4] for entities that include no equivalent entry in the KB a NIL be supposed toward exist returned.

To contract with the nonappearance problem, Bunescu and Pasca et.al [5] filtered away linked talk about whose scores be less than a predetermined threshold. Li et al. [6] trained a separate binary classifier to validate linked mentions. To conquer the difficulty of entity disambiguation wished-for MLN for comparable entity mining.

In this paper current an approach move toward for automatically learning such mining comparators beginning comparative questions and additionally, make available and grade comparable entities intended for a user's input entity suitably beginning the web. It is very useful method for help to users to choose alternative choices by suggestive of similar entities based on additional users' previous desires. To mine comparators pairs result first need to detect whether the question is present in comparator or not

Richardson and Domingos et .al [7] developed markov logic network based joint model which combine first order logic (FOL) and Markov networks. The model captures the contextual information of the recognized entities for entity disambiguation as well as the constraints when linking an entity mention to a KB entry. Our method uses the machine learning based weakly supervised method for bootstrapping to formulate a huge tagged corpus preliminary through simply a small number of examples of QA pairs. Comparable methods have been investigated expansively in the field of information extraction. These methods are significantly aided by the information that there is no necessitate in the direction of corpus, whereas the profusion of data on the web makes it easier to conclude dependable statistical estimates.

2. RELATED WORK

In conditions of discovering associated substance for an entity, our employment is comparable to the investigate on recommender systems, which suggest substance to a consumer. Recommender systems mostly rely on similarities among items and their arithmetical correlation in consumer log data [8]. While consideration of Amazon, the principle of commendation is to attract their customers to append additional substance items. Still these types of questions posted by web users are complex to be predicting basically based on item similarity among them. They are comparable but also dissimilar so request assessment with every other. It is obvious that comparator mining and item recommendation are related other than not the similar.

Our effort on comparator mining is associated to the investigate on entity and relative extraction in information extraction [9]. Jindal and Liu [10], [11] also proposed a comparator mining methods for mining relative sentences and relationships. Both class and sequential rules learned to annotate the result of news and review domain to mine relative sentences as well as relationship. The similar methods followed by author [10] also applied to comparative question identification. Though, their methods characteristically can accomplish elevated precision but endure from low recall [11].

Solving Entity Linking problem for Mineral Industry Research Laboratory proposed an MLN. With a joint conclusion procedure can carry out together tasks concurrently to let alone this kind of inaccuracy proliferation by Poon and Domingos et.al [12] . Joint inferences have developed into well-liked lately, since they make it probable for features and constraints to be communal amongst tasks. For instance, word sense disambiguation (WSD) solved by using representation of joint model by Che and Liu [13] and integrated parsing as well as entity recognition in a joint representation by Finkel and Manning et.al [14].

3. WEAKLY SUPERVISED AND MARKOV-LOGIC NETWORK COMPARABLE ENTITY MINING

Markov logic network (MLN) to representation of interweaved constraints. It is one of the major types of entity linking method with genetic material state relating. Proposed MLN which is the combination of first order logic (FOL) and Markov networks with combination of NIL-filtering and entity disambiguation stages. The representation captures the background information of the familiar entities for entity disambiguation as well as consideration of entity linking in the Knowledge Base (KB) .For instance, an individual declare preserve simply be linked to a KB entry when the state has not been familiar as an NIL. The KB bases the formula are demonstrated with four keywords: constants, variables, functions, and predicates. Whereas constants are referred to as objects in the database entries, that related variables are denoted as x and y for selected objects. Relationship among the data objects are represented as predicates. A world is an obligation of reality values to everyone probable view atoms is also referred to as predicates. Knowledge Base (KB) is an incomplete requirement of a world; every particle in it is accurate, false or unidentified.

A Markov Logic Network (MLN) characterizes the joint distribution of a set of variables $X = (X_1, X_2, \ldots\ldots X_n) \in x$ as a result of factors:

$$P(X = x) = \frac{1}{Z}\prod\nolimits_{Z} f_k(x_k)$$

Where every factor f_k is a non-negative purpose of a separation of the variables x_k , and Z is normalization constant.

As extended as intended for every one $P(X = x) > 0$, for everyone x the distribution can be consistently represent as a log-linear representation:

$P(X = x) = \frac{1}{Z}\exp(\sum_i w_i g_i(x))$, Where $g_i(x)$ is the features are subjective functions of the variables situation. An MLN L is a set of pairs (F_i, w_i) , where F_i is a principle in FOL and w_i is a real numeral represent a weight. Mutually with a predetermined position of constants, it describe a Markov network,$M_{L,C}$ where contains single node for every probable preparation of every predicate appear in L. The assessment of the node is 1 if the ground predicate is true, and 0 or else. The probability distribution in excess of probable worlds is known by $P(X = x) = \frac{1}{Z}\exp(\sum_i \sum_j w_i g_j(x))$ where Z is the separation function, F is the set of every one first order formula in the MLN, is the set of groundings of the i^{th} first-order formula, and $g_j(x) = 1$ if the j^{th} ground formula is true and $g_j(x) = 0$ or else.

Describe four predicates to confine the accepted questions environment information, together with question location, Question Interaction (QI), Tissue Type and Question ontology. The formula describing the relation of and hasquestionInfo and islinkedto is defined as follows: $\text{hasquestionInfo}(i, id, +sd) \Rightarrow \text{islinkedTo}(i, id)$ .At this time, can perceive that in attendance is an added parameter $(+sd)$ indicate in hasquestionInfo.sd consequent to id locates. The "+ " details in the beyond method indicates that necessity study a split weight for every grounded variable (sd). For example, : $\text{hasquestionInfo}(i, id, 0)$and

$\mathrm{hasquestionInfo}(i, id, 1)$ are specified two dissimilar weights in our MLN model following preparation
.

Correlation information from knowledge base (KB) approach interacts with entity one to entity two to solve a disambiguating an entity problem. The QI information stored in the backend database with correlation measure. Based on this result and candidate KB entry distribution result , the id to associated with the majority unambiguous entries is the mainly probable id to be linked to i .Additional describe the subsequent formula to confine the dependence that an entity be supposed to be linked to id_2 if one more entity have be linked to id_1 structure a correlation with id_2 .Filtering the subsequent mention type persons belong to classes with the intention of are not in the database curation objective; called NILs. In linking question with gene are stored to KB Database and NIL filter apply the QI interaction to solve the entity disambiguation problem. The subsequent formula to make sure to, every time the entity is linked to a KB entry id , it be supposed to be an entity appropriate for linking,

$\mathrm{islinkedTo}(i, id) \Rightarrow \mathrm{issuitableForlinking}(i)$

$\exists w.\, \mathrm{hasWord}(w) \Lambda \mathrm{QIKeyword}(w)$
$\Lambda \mathrm{islinkedTo}(i, id_1)$
$\Lambda \mathrm{hascandidate}(j, id_2)$
$\Lambda \mathrm{isQIPair}(id_1, id_2) \Longrightarrow \mathrm{islinkedTo}(j, id_2)$ formula(1)

The steps involved in this Markov Logic Network are defined below:

Input : A Markov network represents the joint distribution of a set of variables $X = (X_1, X_2, \ldots X_n) \in x$, L is set of pairs (F_i, w_i)

Output: Find disambiguation result (F_i, w_i)

Step 1: Define or found the set of disambiguation pairs from using Markov Logic Network (MLN).

Step 2: Find the set of disambiguation result (F_i, w_i) where F_i a formula in FOL is and w_i is a real number represented a weight.

$\exists w.\, \mathrm{hasWord}(w) \Lambda \mathrm{QIKeyword}(w)$
$\Lambda \mathrm{islinkedTo}(i, id_1)$
$\Lambda \mathrm{hascandidate}(j, id_2)$
$\Lambda \mathrm{isQICPartner}(id_1, id_2) \Longrightarrow \mathrm{islinkedTo}(j, id_2)$
formula(1)

Step 3:If it is $\mathrm{if}\ (F_i, w_i) > C$ then defines a Markov network ,$M_{L,C}$ where contains one node for each possible grounding of each predicate appearing in L
.

Step 4: The value of the node is 1 if the ground predicate is true, else 0 otherwise

Step 5: Find the probability distribution over possible worlds is given by ,

$P(X = x) = \frac{1}{z} \exp(\sum_i \sum_j w_i g_j(x))$

Step 6: In the step $g_j(x) = 1$ if the jth ground is true and $g_j(x) = 0$ otherwise.

Step 7:Return the best probability result for each pairs (F_i, w_i)

Step 8:Then now apply bootstrapping procedure

In our disambiguation move toward, rely on background knowledge k , such as an entity's populated location id . Describes a variety of aspect of the entity's ambiguous background knowledge entry, id. Every time the entity is discussed, a number of this aspect determination be state as well. Using k can write formula similar to the subsequent for disambiguation.

hasCandidate(i, id)
hasquestionInfo(i, id, sd)
hasWord(w): the abstract contain a word w.
QIKeyword(w), isQIPartner(id1, id2)
hasQIPartnerRank (i, id, r), hasGOTermRank(i, id, r),
hasTissueTermRank(i, id , r)
hasPrecedingWord(i, w, l), hasFollowingWord(i, w, l)
hasUnigramBetween(i, j, w)

Variable Type

i: an integer, which refers to the ith question mention in the
id: an EntrezQuestion ID, which refers to a linked KB entry.
sd: an integer, which refers to the sentence distance.
w: a word.
r: an integer, which refers to the rank of the matching.
l: an integer, which refers to a context window length

A collection of sequence patterns is specified as S an indicative extraction pattern (IEP) ,condition it be able to be used to identify comparative questions and extract comparators in them through elevated consistency. Primary will properly describe the consistency attain of a sample. Formerly a question that matches to the user pattern in IEP then it is classified as comparative question and token sequence after that the extraction of patterns becomes result. If the question matches several IEP patterns for user given question the longest or highest IEP is selected or manually select patterns with keyword .Demonstrate how to obtain IEPs automatically by means of a bootstrapping process with smallest regulation by taking benefit of a large unlabeled question collection. The proposed weakly supervised indicative extraction pattern (IEP) is based on two explanation: If a sequential pattern be able to be second-hand to extract numerous dependable comparator pairs, it is extremely probable to be there an IEP. If a comparator pair can be extracted by an IEP, the pair is consistent. The method aspires to study sequential patterns which are able to be used to recognize comparative question and extract comparators concurrently. The sequence patterns is specified as S as a sequence S where s_i can be a word or a representation of symbol denote moreover a comparator ($c) , or the beginning (#start) or the end of a question(#end).
A collection of sequence patterns is specified as S an indicative extraction pattern (IEP) ,condition it be able to be used to identify comparative questions and extract comparators in them through elevated consistency.

Input: CP, G
Initialize solution $Q \leftarrow \{\}, P \leftarrow \{\}$ $P_{new} \leftarrow \{\}$ $CP_{new} \leftarrow \{\}$
Repeat
$P \leftarrow P + P_{new}$
$Q_{new} \leftarrow$ compartiveQuestionidentify(CP_{new})
$Q \leftarrow Q + Q_{new}$
For $q_i \in G$ do
If ismatchexistingpatterns(p, q_i) then
$Q \leftarrow Q - q_i$
End if
End for
$p_{new} \leftarrow$ mineGoodpatterns(Q)
$cp_{new} \leftarrow \{ \quad \}$
For $q_i \in G$ do
$cp \leftarrow$ extractcomparablerpatterns(p, q_i)
If $cp \neq NULL$ and $cp \notin CP$ then
$CP_{new} \leftarrow CP_{new} + \{CP\}$
End if
End for
Until $P_{new} = \{\}$
Return P

3.1. Patterns Generation and evaluation

To produce sequential patterns, become accustomed the exterior text pattern mining technique introduced. For some specified comparative question and its pairs, questions of each comparator are replaced with representation $Cs. Together symbols, #start and #end, are emotionally involved to the start and the end of every sentence in the question. To decrease variety of series information and extract possible patterns, expression chunking is practical. After that, the next three kinds of sequential patterns are generated beginning series of questions:

Lexical patterns point toward sequential patterns containing only the representation of symbols and of only words. They generate sequential patterns using suffix tree algorithm among consideration of two constraints that is β not more than one $C, and its occurrence in compilation be supposed to exist additional than an empirically resolute number β.
Generalized patterns are able to be as well precise simplify lexical patterns by replacing one or additional words/phrases by means of their POS tags. 2n - 1 generalized patterns can be fashioned beginning a lexical pattern containing N words exclusive of $Cs.

Specialized patterns a pattern be able to universal even though a question is relative, For this cause, carry out pattern specialization by addition POS tags to all comparator slots .
According to our primary supposition, a reliability score $R^k(p_i)$ for a contestant pattern p_i at iteration k might be definite as follows

$$R^k(p_i) = \frac{\sum_{\forall cp_j \in cp^{k-1}} N_Q(p_1 \rightarrow cp_j)}{N_Q(p_1 \rightarrow cp_j)}$$

Where candidate pattern p_i can extract identified consistent comparator pairs cp_j, cp^{k-1}indicates the reliable comparator pair depository accumulated awaiting the $(k-1)^{th}$ iteration. $N_Q(x)$ means the numeral of questions rewarding a condition x. The condition $p_i \rightarrow cp_j$ specifies that cp_j can be extracted from a question by applying pattern pi whereas the condition $p_i \rightarrow *$ specifies some question containing pattern p_i .

3.2. Comparator Extraction

Comparator extraction used a random based strategy to perform comparator, it randomly choose a pattern amongst patterns which be able to be useful to the question. Another type of strategy is Maximum length strategy. These strategies select a maximum pattern for given a question which is able to be applied to the question comparator extraction. From the discussion above comparator extraction in this work uses a maximum length method is able to exist exactly enclosed which means that the model is additional appropriate intended for the query.

3.3. Comparable ranking methods

The major importance of comparable based ranking methods is to compare the extra attractive entity for an entity if it is compared with the entity further regularly. Based on this insight, describe a straightforward ranking function $R_{freq}(c, e)$ which ranks the comparator results corresponding to the amount of time when the comparator$c$ is compare toward the user's key e in relative questions collection Q:

$$R_{freq}(c, e) = N(Q_{c,e})$$

where $Q_{c,e}$ is a set of questions from the comparator$c$ is compare toward the user's key e can be extracted as a comparator couple .Describe one more ranking function R_{rel} by combination of dependability scores predictable in comparator mining stage

$$R_{rel}(c, e) = \sum_{q \in Q_{c,e}} R(p_{q,c,e})$$

where $p_{q,c,e}$ way the model that is preferred to mine comparator pair of comparator c is compare toward the user's key e from question q in comparator mining phase. This ranking function determination is present denoted as Reliability-based system.

3.4. Graph-Based Ranking

Although regularity is well-organized for comparator ranking, the frequency-based technique can experience whilst an effort occur infrequently in question collection; for instance, understand the case that all probable comparators to the effort are compared simply on one occasion in questions. In this case, the Frequency-based method might be unsuccessful to create a significant ranking end result. Then, Representability is supposed to moreover be considered. For instance, when individual requirements to buy a smart phone and allowing for "iphone-89","iphone 87" is the primary lone he/she needs to evaluate. It uses a graph-based Page Ranking method to compare questions. If a comparator is compared to numerous additional significant comparators which are able to be moreover compared to the input entity, it would be considered as a precious comparator in ranking. Based on this scheme, examine Page Rank algorithm to rank comparators for a known input entity which merge regularity and representability.

4. EXPERIMENTAL RESULTS

All experimentation was conducted on concerning questions that are mined beginning Yahoo! Answers' question name field. The motivation to facilitate used simply a name field is that they obviously convey a major purpose of an asker by means of a structure of straightforward questions in all-purpose .Physically constructed keyword set which contains upto 53 words such as "otherwise" and "rather," which are superior indicators of comparative questions. Categorizes of each and every questions set into SET-A and SET-B one or more keywords from each set ,it randomly selected other than earlier selected questions beginning every Yahoo! Answers category with atleast one keyword present as mentioned above. It contains 765 comparative questions and 1,456 noncomparative questions. For comparative question identification experiments were conducted for each set category separately. Whereas comparator extraction is applied only for SET-B. All the left behind unlabeled questions that is SET-R used for weakly supervised method.
Table 1 shows experimental result in the category of Identification, extraction and all results. Identification says that the comparative questions are identified correctly, Extraction only says that the in which the comparator extracts the question correctly extracted are used as input, and All indicate the back-to-back performances whilst question detection outcome were second-hand in comparator extraction. Reminder that the outcome of WSN-MLN technique on our collections are extremely comparable to what is reported in their manuscript and the figure 1,2,3 values are tabulated in 1.

Table 1: Performance Comparison between Weakly supervised model (WSM) and Weakly supervised model with Markov logic network (WSM-MLN)

Results	Identification only		Extraction only		All	
	Weakly supervised model(WSM) and Weakly supervised model with Markov logic network(WSM-MLN)					
Recall	0.817	0.915	0.760	0.854	0.760	0.870
Precision	0.833	0.925	0.716	0.925	0.776	0.916
F-score	0.825	0.935	0.833	0.889	0.768	0.936

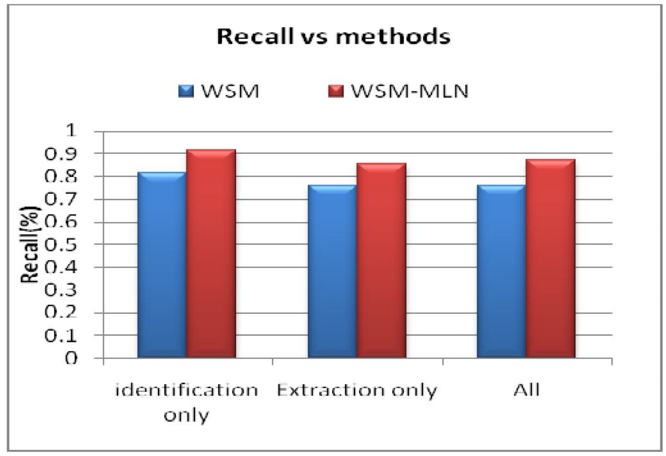

Figure 1: Recall vs. types

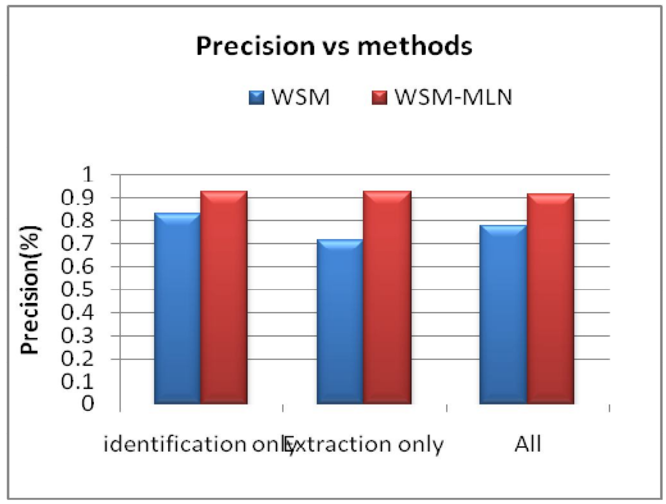

Figure 2: Precision vs. types

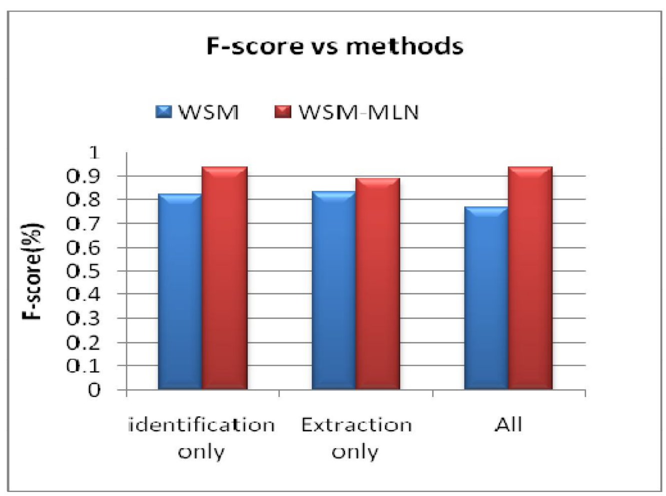

Figure 3:F-Score vs. types

In addition we analyze the result of pattern generalization and specialization. Table 5 demonstrates the results. Regardless of the simplicity of our methods, they considerably contribute to show improvements. These outcomes show the significance of learning patterns flexibly to confine a variety of comparative question expressions. TABLE 2: Effect of Pattern Specialization and Generalization in the End-to-End Experiments, these values are showed in figure 4, 5, 6.

TABLE 2: Effect of Pattern Specialization and Generalization in the End-to-End Experiments

Methods	**Recall**		**Precision**		**F-Score**	
	Weakly supervised model(WSM) Weakly supervised model with Markov logic network(WSM-MLN)					
Original patterns	0.689	0.815	0.449	0.760	0.544	0.750
Specialized	0.731	0.850	0.602	0.810	0.665	0.851
Generalized	0.760	0.860	0.776	0.854	0.768	0.825

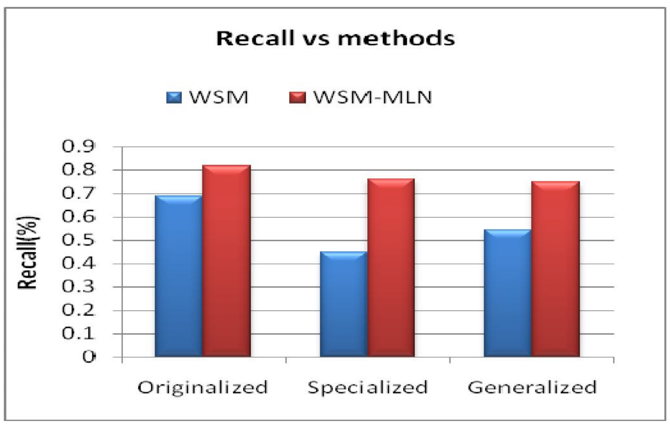

Figure 4: Effect of Pattern vs. recall

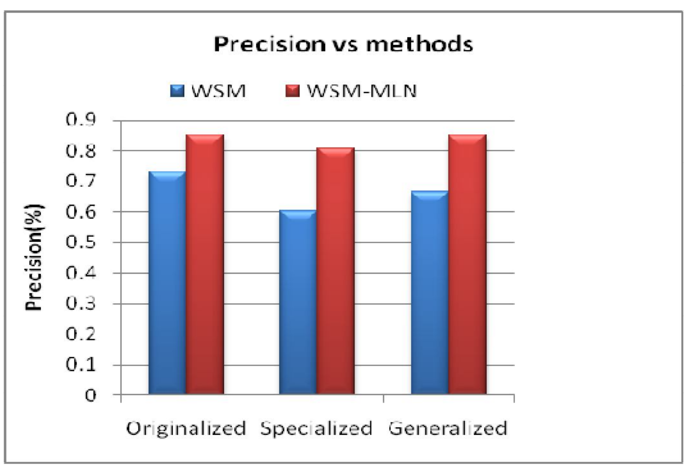

Figure 5: Effect of Pattern vs. precision

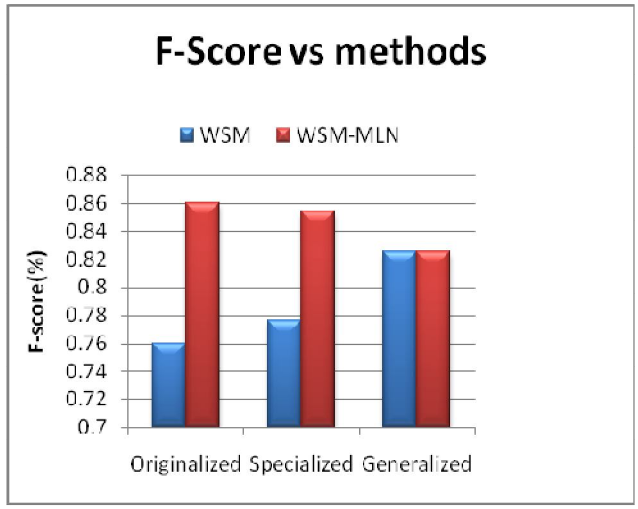

Figure 6: Effect of Pattern vs. F-score

5. CONCLUSION AND FUTURE WORK

In this paper current an original entity disambiguation by means of weakly supervised process to recognize comparative questions and extract comparator pairs concurrently. It depends on insight of key patterns that are generated by high-quality comparative question detection pattern be supposed to extort good comparators, and a good quality comparator pair be supposed to suggest itself in good comparative questions to bootstrap the extraction process. By leveraging huge quantity of unlabeled data and the bootstrapping procedure with insignificant management .The investigational outcome demonstrate that our method is effectual in together comparative question detection and comparator extraction. It considerably improve recall in together tasks whilst maintain elevated precision. Our examples demonstrate that these comparator pairs replicate interested in comparing which is actually wanted by user. Our comparator mining outcome can be second-hand for a commerce exploration or product recommendation organization. For instance, automatic proposition of comparable entities can help out users in their assessment activities earlier than building their acquire decision. In addition, our outcome can make available helpful information to companies which would like to recognize their competitors. In future work also map to extend technique to summarize answers pooled by a specified comparator pair.

REFERENCES

1. Dredze, Mark, Paul McNamee, Delip Rao, Adam Gerber and Tim Finin. 2010, **” Entity Disambiguation for Knowledge Base Population”,** *In: Proceedings of the 23rd*

International Conference on Computational Linguistics (Coling 2010), Beijing.

2. Dai, Hong-Jie, Po-Ting Lai and Richard Tzong-Han Tsai**," Multistage Gene Normalization and SVM-Based Ranking for Protein Interactor Extraction in Full-Text Articles",** *IEEE TRANSACTIONS ON COMPUTATIONAL BIOLOGY AND BIOINFORMATICS*, 7(3): 412-420, 2010.

3. Zhang, Wei, Jian Su, Chew Lim Tan and Wen Ting Wang**," Entity Linking Leveraging Automatically Generated Annotation",** *In: Proceedings of the 23rd International Conference on Computational Linguistics* ,Beijing,2010.

4. McNamee, Paul and Hoa Trang Dang**," Overview of the TAC 2009 Knowledge Base Population Track",** *In: Proceedings of the Second Text Analysis Conference (TAC 2009),* Gaithersburg, Maryland,2009.

5. Bunescu, R and M Pasca**," Using encyclopedic knowledge for named entity disambiguation"** *,In: European Chapter of the Association for Computational Linguistics*,2006.

6. Li, Fangtao, Zhicheng Zheng, Fan Bu, Yang Tang, Xiaoyan Zhu and Minlie Huang**," THU QUANTA at TAC 2009 KBP and RTE Track",** *In: Proceedings of Test Analysis Conference 2009 (TAC 09),* Gaithersburg, Maryland USA,2009.

7. Richardson, Matthew and Pedro Domingos **," Markov logic networks. Machine Learning,** *62(Special Issue: Multi-Relational Data Mining and Statistical Relational Learning):* 107-136,2006.

8. S. Li, C.-Y. Lin, Y.-I. Song, and Z. Li, **"Comparable Entity Mining from Comparative Questions,"** *Proc. 48th Ann. Meeting of the Assoc. for Computational Linguistics (ACL '10)*, 2010.

9. E. Riloff and R. Jones**, "Learning Dictionaries for Information Extraction by Multi-Level Bootstrapping,"** *Proc. 16th Nat'l Conf. Artificial Intelligence and the 11th Innovative Applications of Artificial Intelligence Conf. (AAAI '99/IAAI '99),* pp. 474-479, 1999.

10. N. Jindal and B. Liu, **"Identifying Comparative Sentences in Text Documents,"** *Proc. 29th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '06),* pp. 244-251, 2006.

11. N. Jindal and B. Liu, **"Mining Comparative Sentences and Relations,"** *Proc. 21st Nat'l Conf. Artificial Intelligence (AAAI '06)*, 2006.

12. Poon, H and P Domingos**, "Joint inference in information extraction",** *In, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press;* 2007.

13. Che, Wanxiang and Ting Liu**," Jointly Modeling WSD and SRL with Markov Logic"**, *In: Proceedings of the 23rd International Conference on Computational Linguistics,* Beijing, China,2010.

14. Finkel, Jenny Rose and Christopher D. Manning, **"Joint parsing and named entity recognition",** *In: Proceedings of NAACL 2009.*