IPHV7I20047X
International Journal Of Advanced Research and Innovation -Vol.7, Issue .II
ISSN Online: 2319 – 9253
Print: 2319 – 9245

ENTITY DISAMBIGUATION WITH PREDICTIVE MINING OF COMPARABLE ENTITIES

#1 BADUGU KIRAN KUMAR - M.Tech Student,
#2 Dr.M.S.S.Sai- Professor &HOD,
Department of CSE,
KKR&KSR INSTITUTE OF TECHNOLOGY&SCIENES, GUNTUR,A.P.

Abstract: - Comparing entities is an important part of decision making. Several approaches have been reported for mining comparable entities from Web sources to im-prove user experience in comparing entities online. However, these efforts extract only entities explicitly compared in the corpora, and may exclude entities that occur less-frequently but potentially comparable. To build a more complete comparison machine that can infer such missing relations, here we develop a solution to predict transitivity of known comparable relations. 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, Sequential Pattern Mining and Comparable Entity Mining..*

I.INTRODUCTION

To assist decision making, it is useful to compare enti-ties that share a common utility but have distinguish-ing peripheral features. For example, when deciding on a new mobile device to purchase, a customer benefits from knowing products with similar specifications, e.g., iPhone, Nexus One and Blackberry.

One possible approach is comparable entity mining, which extracts comparable pairs that are explicitly compared on the Web corpus. However, these tech-niques are limited by their ability to mine only enti-ties explicitly compared in Web sources, excluding en-tities that are potentially comparable but are not cur-rently explicitly compared in the corpora. However, for a fully-functional comparison suggestion system, such comparisons should not be disregarded. In fact, such missing links for comparable entities are inevitable even with large datasets.

An orthogonal approach is predictive mining, which can complement existing mining approach. It expands the known comparable relations using transitivity to infer the unknown relations. We stress that the two approaches are clearly different: for the task of classify-ing missing links into comparable and non-comparable ones, the former leads to zero precision and recall,

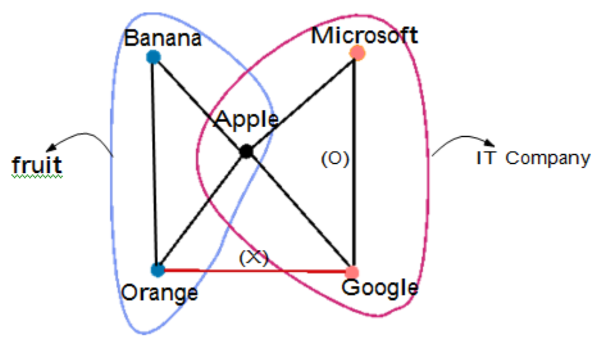

Figure 1: Two different sub-graphs of “fruit” and “IT Company” are connected by a bridge node “Apple”; (o) – comparable edge, (x) – non-comparable edge whereas the predictive mining can classify them with reasonable accuracy.

We first consider a comparable entity graph (CE-graph) containing these comparable entities and binary relations. It is an undirected graph G = (V,E) where V is a set of named-entities, E is a set of edges where (vi,vj) ∈ E indicates that vi and vj are comparable. An initial CE-graph can be con-structed with entity pairs that are explicitly compared and mined by using techniques and resources proposed in comparable entity mining (Jindal and Liu 2006; Li et al. 2010; Jain and Pantel 2011). For an uncon-nected pair of nodes in a CE-graph, we should next determine the comparability of the pair, i.e., we should predict a link between the nodes if the pair is compa-rable.

To infer such transitivity, two challenges must be overcome. First, ambiguous entities may serve as bridge nodes in a graph, which connect two semantically dif-ferent subgraphs in a CE-graph. For example, Apple is the bridge node that connects two subgraphs, (Fruit: Apple, Banana, Orange) and (IT company: Apple, Mi-crosoft, Google) (Fig. 1). Bridge nodes may cause an in-correct prediction deduced from a graph topology, such as (Orange, Google).

Second, the sparseness of an initial CE-graph offers little structural information for link prediction. For example, in a CE-graph obtained from Microsoft Live Search versus query logs collected over one month, thenumber of entity pairsexplicitly compared is only 0.03% of all possible pairs of entities in the versus query logs (i.e., 5,129 pairs among about 14 million possible pairs of 5,368entities). Later we empirically show that apply-ing generic link prediction algorithms to such a sparse graph achieves very low recall for prediction.

To predict the missing links considering these chal-lenges, the three criteria listed below are required for a possible solution to properly expand known relations using transitivity.

• Graph structure: To infer transitivity of links in the given graph, graph structure should be considered to reflect how likely the two nodes are to be connected via neighbors.

• Attributes: To determine whether two nodes are comparable, attributes (e.g., semantics) of nodes should be considered.

• Disambiguation: Graphs inevitably include am-biguous nodes, which should be disambiguated to prevent generation of heterogenous clusters.

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 h uge 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.

II. RELATED WORK

We survey two research areas related to our work: (1) comparable entity mining that complements our predic-tion work and (2) link prediction methods that compete with ours work.

IPHV7I20047X
International Journal Of Advanced Research and Innovation -Vol.7, Issue .II
ISSN Online: 2319 – 9253
Print: 2319 – 9245

2.1 Comparable Entity Mining

Several approaches exist to extract comparable entities from various web corpus. Jindal and Liu proposed su-pervised mining of comparable entities from compar-ative sentences (Jindal and Liu 2006); their method uses a class sequential rule (CSR) to classify sentences into comparative or non-comparative. This method re-quires a comparative keyword set for training sequential rules; but keyword sets should be manually defined. To overcome this drawback, Li et al. proposed a weakly-supervised bootstrapping method to identify compara-tive questions and extract comparable entities (Li et al. 2010). Recently, Jain and Pennacchiotti used pattern learning methods to extract comparable entities from both query logs and web documents. Their experiments showed that query logs are superior to web documents as resources from which to extract comparable entities (Jain and Pennacchiotti 2010).

The above entity mining techniques focus on mining comparable pairs readily observed in the web corpus, but our work focuses on predicting pairs that cannot be observed from it. Our prediction work thus comple-ments existing comparable entity mining; when used together, both approaches achieve the goal of obtaining a comparable entity set.

2.2 Link Prediction

In this section, we describe two main link prediction approaches–(1) using graph structure and (2) using clustering.

Using graph structure This approach uses graph structure to solve link prediction problems. The type of graph structure used includes node neighbors and the ensemble of all possible paths. We do not cover these methods in detail, since they have been well-studied and evaluated in a previous survey (Liben-Nowell and Kleinberg 2003).

Using clustering For this approach, we specifically discuss three methods that come closest to meeting the three criteria listed above (Table 1).

MC-Cluster is a generic graph clustering that con-siders the presence of bridge nodes (Scripps and Tan 2006). In the given graph, MC-Cluster generates clus-ters in which every pair of nodes in an Must-Link (ML) edge belong to the same cluster, and any nodes in an Cannot-Link (CL) edge cannot in the same cluster. A node is identified as a bridge node when it is connected to two nodes by ML edges and the two nodes are con-nected by a CL edge. The graph is disambiguated by cloning the bridge node to several nodes such that each belongs to one cluster.

TP-Cluster was devised for a word sense induction problem that clusters semantically-similarwordsamong a list of wordsthat co-occurredwith a givenword. (Bor-dag 2006). To disambiguate words, TP-Cluster creates all possible triplets from a list of words that the tar-get word is co-occurred with. Each triplet contains the intersection of the co-occurrence list from each word; the intersection is used as a feature of the triplet. Two triplets are merged if they share similar features, un-til the algorithm converges. TP-Cluster is a context-based algorithm, so when clustering it considers node attributes, not graph structure.

SA-Cluster is a graph clustering algorithm that con-siders both structural and attributal properties (Zhou, Cheng, and Yu 2009). SA-Cluster converts attributes to structural properties by inserting attribute nodes that are connected to all nodes that have the corre-sponding attribute. They exploit a unified random walk distance on the augmented graph. We discov-ered that SA-Cluster shows a poor performance for our problem due to bridge nodes, and attribute nodes that are commonly-shared by many nodes. As a result, the augmented graph becomes very densely-connected, and may become one big heterogenouscluster instead of sev-eral homogenous clusters.

EXISTING METHOD:

Comparator mining is related to the research on entity and relation extraction in information extraction Specifically, the most relevant work is mining comparative sentences and relations. Their methods applied class sequential rules (CSR) and label sequential rules (LSR) learned from annotated corpora to identify comparative sentences and extract comparative relations respectively in the news and review domains. The same techniques can be applied to comparative question identification and comparator mining from questions.

PROPOSED METHOD:

we present a novel weakly supervised method to identify comparative questions and extract comparator pairs simultaneously. We rely on the key insight that a good comparative question identification pattern should extract good comparators, and a good comparator pair should occur in good comparative questions to bootstrap the extraction and identification process. By leveraging large amount of unlabeled data and the bootstrapping process with slight supervision to determine four parameters.

III. IMPLEMENTATION

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 ,X ,⋯. .X) ∈ x as a result of factors:

P(X = x) = 1 f (x)

Where every factor f is a non-negative purpose of a separation of the variables x , andZ 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) = exp(Σ w g (x)) ,

Where g (x) is the features are subjective functions of the variables situation. An MLN L is a set of pairs (F ,w) , where F is a principle in FOL andw is a real numeral represent a weight. Mutually with a predetermined position of constants, it describe a Markov network,M , 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) = exp(Σ Σ w g (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 first-order formula, and g (x) = 1 if the j ground formula is true and g (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:

hasquestionInfo(i,id,+sd) ⟹ islinkedTo(i,id) .At this time, can perceive that in attendance is an added parameter (+sd) indicate in hasquestionInfo.sd consequent toid locates. The "+ " details in the beyond method indicates that necessity study a split weight for every grounded variable (sd). For example, : hasquestionInfo(i,id,0)and 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 , theid 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 if one more entity have be linked toid structure a correlation with id. 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 entryid , it be supposed to be an entity appropriate for linking,

islinkedTo(i,id) ⇒ issuitableForlinking(i)

∃w.hasWord(w)ΛQIKeyword( )

ΛislinkedTo i,id) Λhascandidate(j, id)

ΛisQIPair(id ,id) ⟹ islinkedTo(j, id) 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 , X ,⋯.. X) ∈ x , L is set of pairs (F ,w) Output: Find disambiguation result (F ,w)

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 ,w) where F a formula in FOL is and w is a real number represented a weight.

∃w.hasWord(w)ΛQIKeyword(w)

ΛislinkedTo(i,id)

Λhascandidate(j, id)

ΛisQICPartner(id ,id) ⟹ islinkedTo(j, id)

formula(1)

Step 3:If it is if (F ,w) > then defines a Markov network ,M , 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 ,

$$P(X = x) = \exp(\Sigma \ \ \Sigma \ w\, g (x))$$

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

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

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 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
$P \leftarrow P + P_{new}$
$Q_{new} \leftarrow \text{compartiveQuestionidentify}(CP_{new})$
$Q \leftarrow Q + Q_{new}$
For $q_i \in G$ do
If $\text{ismatchexistingpatterns}(p, q_i)$ then
$Q \leftarrow Q - q_i$
End if
End for
$p_{new} \leftarrow \text{mineGoodpatterns}(Q)$
$cp_{new} \leftarrow \{\ \ \}$
For $q_i \in G$ do
$cp \leftarrow \text{extractcomparablerpatterns}(p, q_i)$
If $cp \neq \text{NULL}$ 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.

IPHV7I20047X
International Journal Of Advanced Research and Innovation -Vol.7, Issue .II
ISSN Online: 2319 – 9253
Print: 2319 – 9245

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 (p) for a contestant pattern p at iteration k might be definite as follows

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

Where candidate pattern p can extract identified consistent comparator pairs cp ,cp indicates the reliable comparator pair depository accumulated awaiting the (k −1) iteration. N (x) means the numeral of questions rewarding a condition x. The condition p → cp specifies that cp can be extracted from a question by applying pattern pi whereas the condition p →∗ specifies some question containing pattern p .

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 (c,e) which ranks the comparator results corresponding to the amount of time when the comparatorc is compare toward the user's key e in relative questions collection Q:

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

where Q , is a set of questions from the comparatorc is compare toward the user's key e can be extracted as a comparator couple .Describe one more ranking function R 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 , , 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.

IV.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

IPHV7I20047X
International Journal Of Advanced Research and Innovation -Vol.7, Issue .II
ISSN Online: 2319 – 9253
Print: 2319 – 9245

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) **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

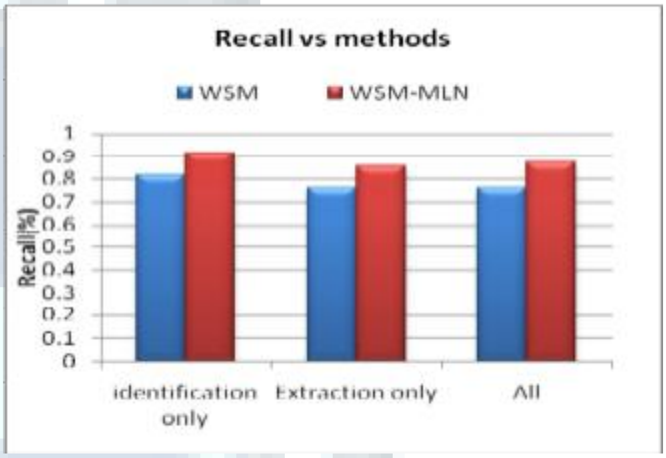

Figure 1: Recall vs. types

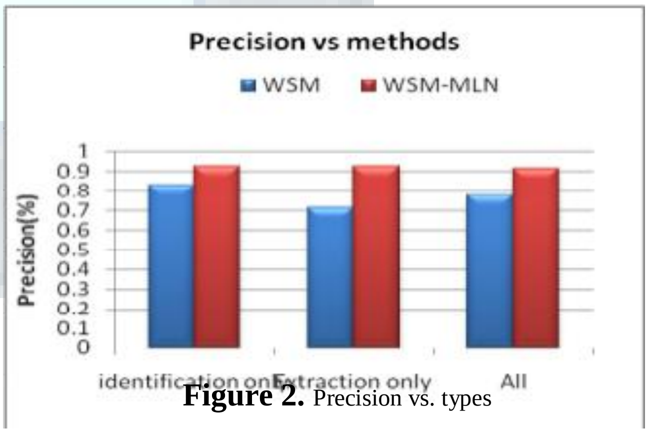

Figure 2. Precision vs. types

Figure 3: F-Score vs. types

V.CONCLUSION

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.

REFERENCES

[1]. Dai, Hong-Jie, Po-Ting Lai and Richard Tzong-in Information Retrieval (SIGIR '06), pp. 244-251, 2006.

[2]. Han Tsai," Multistage Gene Normalization and11. N. Jindal and B. Liu, "Mining Comparative SVM-Based Ranking for Protein Interactor Sentences and Relations," Proc. 21st Nat'l Conf. Extraction in Full-TextArticles",IEEE Artificial Intelligence (AAAI '06), 2006. 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"

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

[7]. 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.

[8]. Finkel, Jenny Rose and Christopher D. Manning, "Joint parsing and named entity recognition",

[9]. In: Proceedings of NAACL 2009., In: European Chapter of the Association for Computational Linguistics,2006.

[10]. 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.

AUTHOR'S PROFILE:

[1]. BADUGU KIRAN KUMAR, Pursuing M.Tech in Department of Computer Science & Engineering at KKR&KSR INSTITUTE OF TECHNOLOGY & SCIENES,Vinjanampadu Guntur.

[2]. Dr.M.S.S.Sai, working as Professor & Head of The Department of Computer Science & Engineering, KKR&KSR INSTITUTE OF TECHNOLOGY & SCIENES, Vinjanampadu Guntur.