Evolving long-term dependency rules in Lifelong Learning Models

Muhammad Taimoor Khan
FAST-NUCES, Peshawar
Pakistan
taimoor.muhammad@gmail.com

Sonam Yar
University of Peshawar, Peshawar
Pakistan
sonam.yar@gmail.com

Shehzad Khalid
Bahria University, Islamabad
Pakistan
shehzad_khalid@hotmail.com

Furqan Aziz
IMSciences, Peshawar
Pakistan
furqan.aziz@imsciences.edu.pk

ABSTRACT: *Topic models are extensively used for text analysis to extract prominent concepts as topics in a large collection of documents about a subject domain. They are extended with different approaches to suit various application areas. Automatic knowledge-based topic models are recently introduced to specifically meet the processing needs of large-scale data having many subject domains. The model automatically learns rules across all domains and use them to improve the results of the current domain by purposefully grouping words into topics to better represent the underlying concept. The existing models apply thresholds on evaluation criteria to learn rules, however, being automatic it may learn wrong, irrelevant or inconsistent rules as well. In this research article the proposed model learns rules and monitor their contributions towards the quality of results. As the model learns new rules, the existing rules undergo refinement and detachment procedures to retain reliable rules only. Experimental results on user reviews from Amazon.com shows improvement in the quality of topics by using fewer rules which advocates the quality of rules and help avoid performance bottleneck at high experience.*

Key words: Learning, Knowledgement, Document Management, Text Analysis

Received: 27 April 2016, Revised 27 May 2016, Accepted 5 June 2016

1. Introduction

Web 2.0 has emerged the era of Big social data produced online on various platforms. It consists of a variety of issues discussed there in and is streaming in continuously at high rate. Since it is impossible for users to read through these documents and develop their understanding about it in a reasonable time, therefore, machine learning techniques are employed. Topic models are popularly used for text analysis to perform various Natural Language Processing (NLP) tasks. They explore unknown domains and identify the hidden thematic structures to represent the key concepts discussed as topics, where the topics consists of contextually co-related words [1]. Unlike dictionary and template based techniques, topic models consider the context which is important in specialized domains. For example the terms *raw* and *data* are semantically different but may be grouped together due to contextual relevance in specific domain. Topic models assign topics to all words in all documents at random. The inference technique converges the distribution by updating the word topic assignment to the most probable topic based on the current distribution. The process continues until the topics stabilize with words. The words in a topic represents a concept specific to the subject domain. For example, the topics “dna, movement, disease, organism”and “computer, network, messaging, processing, storage”represent biology and computing accordingly. Topics in a topic model may represent chapters when applied to a book, policies when applied to political speeches and research areas when applied to scientific journals. In

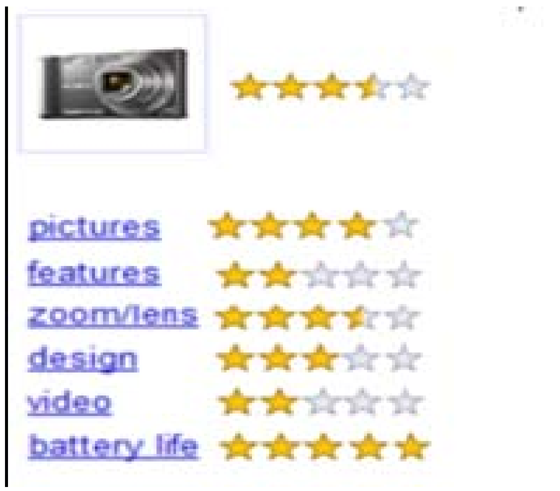

Figure 1. extracted aspects as sentiment targets

commercial domains, topics are considered to be product aspects while the terms in topic are synonyms and near-synonyms used to refer to it. They are used as sentiment targets in Aspect-based sentiment analysis as shown in Fig. 1.

Topic models are widely used with unsupervised approach to identify the major concepts in an unknown domain. The number of topics control the depth of analysis as fewer number of topics focus on global concepts only, while high number of topics identify local concepts as well. But unsupervised topic models with higher number of topics may produce incoherent topics as well. The words in incoherent topics do not make sense to be conceived as a concept, due to various reasons discussed in [2]. The incoherent topics also results in because of the subject domain having fewer sample documents or having noise. Topic coherence is a measure of accuracy for extracted topics which refers to the compactness of terms in a topic. In order to improve with the accuracy of topic models as topic coherence, they are extended with different approaches. Supervised topic models are trained on labeled data to improve accuracy for sensitive domains. Hybrid models are trained on a small labeled dataset to help the model extract optimum initial values for hyper-parameters. Transfer learning topic models are trained on one domain having more consistent sample documents and is applied on another relevant domain. Semi-supervised topic models require domain experts to provide seed terms with topics and the model populate the words in topics based on the distribution of seed words. But labeled datasets do not exist for fresh domains discussed online and domain experts are expensive to find for each subject domain. These models improved the accuracy of the model by producing more coherent topics, however, they cannot be used without manual guidance. Fig. 2 shows working of the semi-supervised topic models.

Knowledge-based topic models uses knowledge rules to improve the accuracy for topics extracted. They are initially proposed with semi-supervised approach. The rules are manually provided by domain experts as word pairs having positive or negative co-relations, called mustlinks and cannotlinks respectively. They have reduced the effort time required to train a semi-supervised model, but the model is still not usable with the guidance of domain expert. Knowledge-based topic models associate the accuracy and performance of the model the quality and quantity of the rules provided. The rules add a bias into the inference technique that is more valuable for domains with fewer samples or noise. Despite of the fact that the extensions of topic models

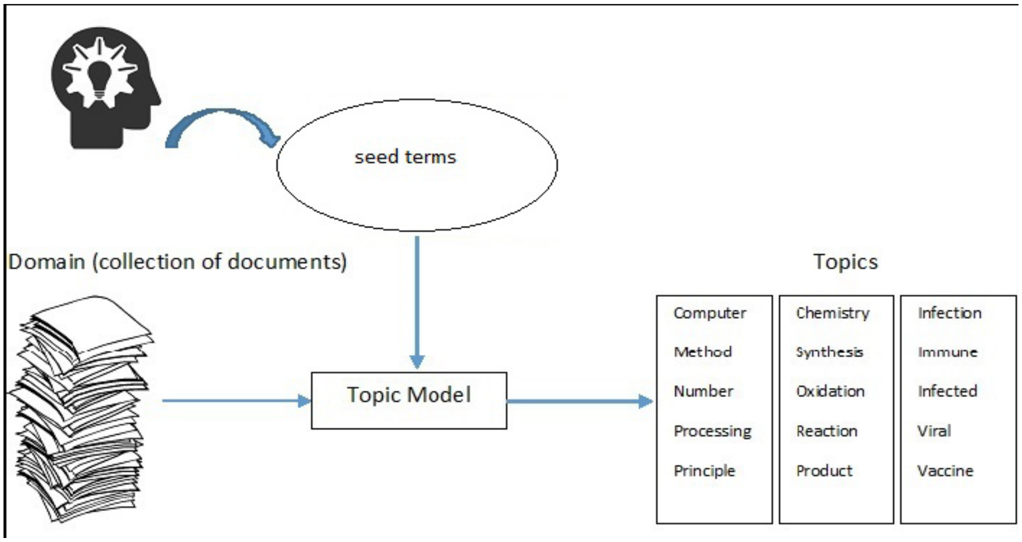

Figure 2. Semi-supervised Topic Model, using seed terms from domain expert to extract topics

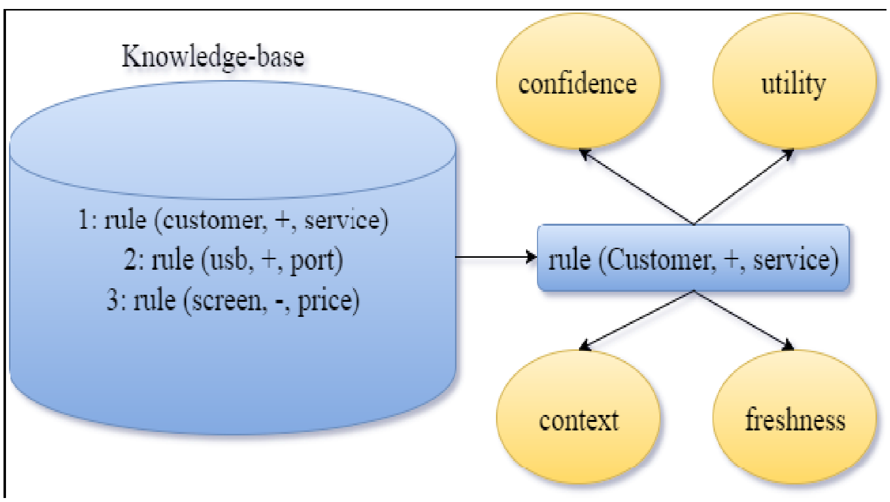

Figure 3. Knowledge abstraction and presentation in knowledge module

discussed with different approaches has achieved their goal of accuracy and performance for offline subject domains but they cannot be applied to big social data due to their dependency on manual guidance. They follow a single-shot learning approach and do not carry what is learnt from one domain, to another. Therefore, they cannot be applied to big social data that has many unknown domains. It is streaming in from the online sources and it is hard to find domain experts or labeled data for it. The issue is addressed with Automatic Knowledge-based topic models.

Automatic knowledge-based topic models are equipped with a learning module that enables the model to learn rules for itself. It does not require guidance from domain experts and learns rules after each task. The learning of the model depends on the number of tasks it has completed and therefore, the accuracy of the model improves with experience by following a Lifelong Machine Learning (LML) approach. LML models aim at designing and developing computational systems and algorithms that can learn like humans. They are also known as Human-like learning or Never-ending learning models. These models are specifically designed for the needs of large-scale data having many unknown domains. The model exploits the huge volume of data and the variety in its domains, for learning rules. The rules learnt from many subject domains are utilized to improve the accuracy of an individual domain. The rules learnt across many domains help the model extract global rules that are generally true and are supported by many domains. This is used as the evaluation criteria for learning rules to which thresholds are applied to learn mustlinks and cannotlinks separately. Fig. 3 represents rules as mustlinks and cannotlinks in the knowledge-base along with the characteristics associated. It also benefits from the fact that that many domains have overlapping vocabulary, that is having common aspects in multiple subject domains. For example the two subject domains about laptop and phone, have overlap in their vocabulary for words about screen and battery etc. It is different from transfer learning that is trained for a single target domain that is known prior to analysis. The accuracy and performance of LML models depend on the quality and quantity of rules learnt.

Automatic knowledge-based topic models learns as word pairs i.e. mustlinks and cannotlinks. In order to learn rules, the model generates candidate rules and pass them through an evaluation criteria. It allows each candidate rule to have a confidence value. A pair of thresholds are applied on the confidence value to learn mustlinks and cannotlinks separately. The rules learnt are utilized when a contextually relevant scenario arise. While transferring the impact of rules into the task in hand, mustlink words reinforce each other within a topic while the cannotlink words subside each other for the same topic. Since the rules learnt may not be utilized by the next task and is rather dependent on the similarity of context, the rules learnt are retained and maintained for long term. The evaluation criteria is based on Frequent Itemset Mining (FIM) that creates a performance bottleneck at high experience. The existing models do not analyze the usefulness of the rules learnt in order to monitor their performance. The automatic learning module is expected to learn wrong, irrelevant, inconsistent and contradictory rules. Treating all the rules equally, cannot resolve the dispute in conflicting rules. The contribution of this research paper is an efficient evaluation criterion for learning rules as mustlinks and cannotlinks. The rules are added to the knowledge-base in abstract form through long term dependency mechanism to bring persistence in learning patterns. The utility and freshness of rules are continuously monitored and are updated after each task through refinement mechanism. The conflict between contradictory rules is resolved through confidence of a rule. If a rule is learnt but has its confidence dropped in the following domains is filtered out. Similarly high confidence rule that has limited utility is filtered as well for irrelevant. It keeps a check of the rules after being learnt and the knowledge-base stays consistent to perform well at high experience. (experiment / dataset / results)

2. Literature Review

The data produced online has developed the need for more knowledge-based systems that could learn for a specific task [3]. In support of content based analysis, the structural information of the platform i.e. user profiles and their relationships can also be used for better analysis [4, 5]. The structural or meta information varies with the platform from which the data is acquired. Transcribed text contains words that are spoken but doesn't mean anything semantically e.g. ummm, Aahhh etc. It requires procedures specific to the needs of transcribed text to enhance its usefulness [3, 6, 7]. Both the aspects and sentiments can be viewed conceptually and contextually for specialized domains [8, 9]. Micro-blogging is another popular platform of content analysis that has a character limit and therefore, usually miss the context. Such content need to be associated to relevant entity to better understand their context e.g. a recent event, person, place etc. Meta information like tweets, re-tweets, hashtags and emoticons can be used to help in analysis [10]. The online social media platforms are cheap and efficient source of identifying key issues or concerns in a region and the orientation of public sentiments towards it. With real time analysis, current issues can be monitored continuously. In case of offline domains, the number of sampled documents can be doubled by using inverse of each document, however, the using the inverse of only selected documents improved results [11]. The concepts extracted from a large collection of document represents characteristics of the problem in hand, for which the data is extracted which are further analyzed through aspect-based sentiment analysis models [12, 13, 14, 15, 16].

Topic models are extensively used for unknown domain exploration and text analysis [1]. It generates topics representing the key concepts in the data. Topic models with unsupervised approach have the freedom to identify the unknown concepts. In commercial domains, the concepts represent product aspects. Supervised models only learns about the aspects for which the model is trained and ignore the less frequent or uncommon aspects for which the model is not trained. Similarly semi-supervised models also direct the model towards aspects about which manual guidance is provided [17]. However, unsupervised topic models can extract the uncommon, less known, infrequent aspects as well at compromised accuracy. The aspect-sentiment joint model has topics representing aspects as well as sentiments by usin a Probabilistic Latent Semantic Analysis (pLSA) [18]. MGLDA (Multi-grain LDA) topic models are used to extend the topics to focus on local concepts or aspects as well along with global or popular concepts through different window sizes as conceptual document [19]. Review format specific models only process content that has review document supported with pros and cons where aspects from pros and cons section are used to help extract less common aspects from the review document [20]. MaxEnt-LDA distinguished the separation between topics having aspects and sentiment labels by utilizing a small labeled dataset as a hybrid approach [21, 22].

In order to reduce the effort and cost of manual guidance by domain expert, knowledge-based topic models are introduced with a semi-supervised approach. The domain expert loads the model with domain specific rules which the model uses to guide the reference towards coherent topics. DF-LDA use knowledge rules as mustlinks and cannotlinks to improved the results while reduced the efforts of domain expert [23]. These models are rigid about the knowledge rules provided, that are believed to be correct and the model has no mechanism of verifying them. Despite of reduced manual guidance, the model still requires domain expert for each task and cannot be extended to big data that have many unknown domains. The issue is addressed through Automatic knowledge-based topic models that are specifically designed for big social data. They have an automatic learning mechanism used as a replacement for domain expert. The learning module extract rules for itself and then use those rules to improve its results [24]. Being unsupervised, the model can be directly applied to big data while having a learning module it produces coherent topics to utilize the automatically generated rules. These models are specifically designed for big data and exploit the huge volume of data and the variety of domains to its advantage for learning good quality rules by clustering prominent words. The learning mechanism is improved by using Multi-support Frequent Itemset Mining (MS-FIM) [25] for generating better rules [26]. Multi-generalized Polya Urn (M-GPU) model [27] is used to transfer the impact of rules into the current task. Knowledge-based recommendation systems are developed to learn according to specific tasks performed on online social data [28, 29, 30].

There are certain limitations in the existing Automatic knowledge-based topic models. Since the learning mechanism is automatic, the model is expected to learn wrong, irrelevant, inconsistent and contradictory rules as well. Similarly the rules having confidence above the threshold having varying confidence and needs to be prioritized accordingly. Once a rule is learnt, there is no mechanism to monitor its contribution and apply a policy accordingly. Therefore, a model is proposed that prioritize knowledge rules, monitor their contributions and follow a refinement or detachment mechanism to keep track of the rules, resolve disputes among conflicting rules and keep the knowledge-base consistent by weeding out wrong or irrelevant rules.

3. Proposed Model

LML models learn like humans, having long term dependency learning. The learning from one task may not necessarily be used in the next task. Since it is not known when a rule learnt is to be used, LML models need to maintain long-term dependency learning. Therefore, the rules that fulfil the evaluation criteria are stored in the knowledge-base. The rules stored are abstracted

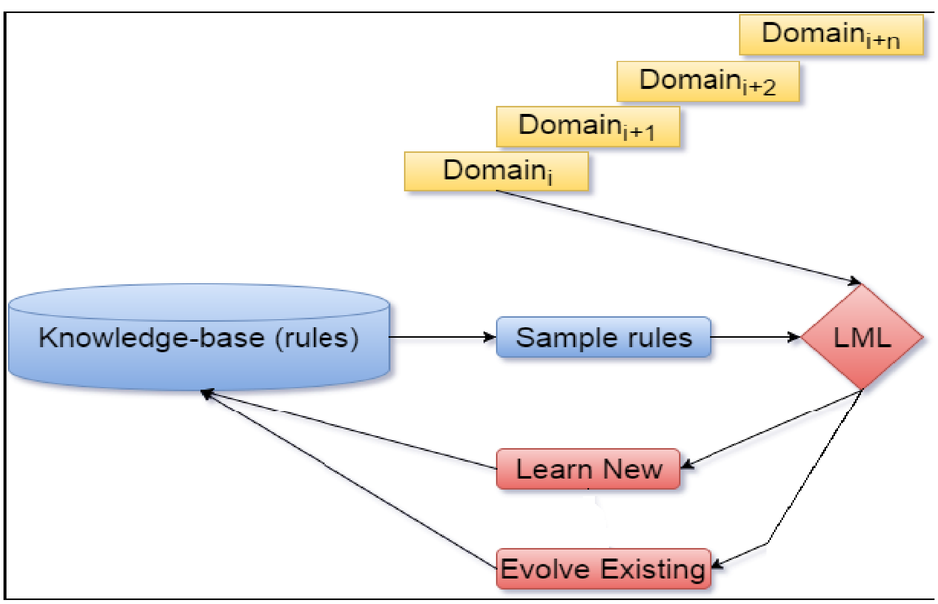

Figure 4. Knowledge-evolving LML

and additional features are added Fig. 4. Knowledge-evolving LML for resourcefulness. Working of the proposed model is shown in Fig. 4. This is because the learning from previous tasks is utilized when a contextually similar task arrives, no matter where they are placed in spatial and temporal terms. Thus the learning and transfer techniques do not follow a sequential approach. For example, if the models learn some knowledge from the Laptop domain may not be utilized for the next many domains about *fashion, clothing* and *cooking*. However, it can be used for a domain *Tablet* for having similar context to it. The rules i.e. mustlinks and cannotlinks are captured along with their context which represents the top few words from vocabulary of the domain from which the rule was learnt. The context helps to resolve word sense disambiguation, as a word have different semantic meanings in different domain scenarios. The use of context information ensures that the rule is used in the same context from it was extracted. Thus the overlap between the context of a rule and the vocabulary of the current domain must be greater than , as shown in Eq.

$$\text{overlap} = \begin{cases} \dfrac{(\#(contextr_i \cap vocab_j)}{\# \; context} > \in & use \\ else & ignore \end{cases} \tag{1}$$

LML models consists of four components that are knowledge representation, extraction, transfer and maintenance or retention. The knowledge extraction mechanism consists of techniques that generate knowledge and evaluate it on different criteria. The candidate knowledge rules are passed to various filters. The proposed model uses $nPMI\ (w_1;w_2)$ as evaluation criterion to identify confidence of a candidate rule with words w_1 and w_2 as given in Eq.

$$k\ Rule\ (w_1, w_2) = n\ PMI\ (w_1, w_2) \tag{2}$$

Where

$$n\ PMI\ (w_1, w_2) = - \frac{PMI\ (w_1, w_2)}{\log P(w_1, w_2)} \tag{3}$$

The *n PMI* values lies in [-1, 1] range on which a pair of thresholds are applied at t1 and t2 to learn mustlinks and cannotlinks as rules. *n PMI* is a light-weight rule extraction technique that evaluate both candidate mustlinks and cannotlinks together with a confidence score to avoid performance bottleneck. The candidate rules having confidence as *n PMI* score above t_1 are learnt as mustlinks or below t_2 are learnt as cannotlinks while others are ignored as shown in Eq.

$$n\ PMI\ (w_1, w_2) = \begin{cases} > t1 & mustlinks \\ < t2 & cannotlinks \\ otherwise & ignore \end{cases} \tag{4}$$

The extracted knowledge rules are reduced to an abstract form and are represented accordingly. The representation of knowledge is focused on getting maximum utilization at higher performance, as shown in Fig. 3. The knowledge transfer mechanism is concerned about transferring the bias from knowledge into the current task. Generalized Polya Urn model is used for adding the rule bias into the topic model [27]. It helps to push mustlink words up for the same topic while push cannotlink words up for different topics. The maintenance and retention module focuses on providing long-term persistence until a contextually similar task arrives. Improving the mentioned components to enhance the accuracy and performance of the knowledge-based model is called knowledge engineering. For a set of documents in domain $D^t = \{D_1, D_2, ..., D_n\}$ and *Knowledge - base* as repository of long-term dependency rules. Knowledge-base consists mustlinks and cannotlinks extracted through Eq. 2, 3 and 4. Working of the proposed model is shown in Algorithm 1. *V* is vocabulary of current task and *N* is the iterations of Gibbs Sampler used for inference [26]. In step 2, the rules relevant to task are sampled through the overlap in context of rule and vocabulary of the current task. Step 3 has the sampled rules used in Automatic knowledge-based topic model to extract topics with improved accuracy with the help of the rules. Before proceeding to the next task, the knowledge-base undergo a series of steps to keep rules consistent. The relevant rules that helped in current task are upgraded in step 4 with their confidence updated and utility increased. In step 5, the rules that didn't made contributions are updated by having their freshness reduced, where the rules with low freshness and limited utility is eventually filtered out as irrelevant. In step 6 the knowledge-based is updated to be consistent and ready for processing the next task, for which the process repeats.

Algorithm 2 explains the working of refinement and detachment mechanism keeps the knowledge-base consistent. It has three procedures i.e. *refineRules*, *updateRules* and *rule* Detachment. The process is repeated after each task.The *refineRules* procedure

Algorithm 1 proposedModel(D^t, V, N, Knowledge-base)

1: **procedure** PROPOSEDMODEL

2: relRules contextOverlap(V, Knowledge-base)

3: topics AKBTM(D^t, V, N, relRules)

4: Rulesrefined refineRules(D^t, V, topics, relRules)

5: Rulesupdated updateKB(D^t, knowledge-base)

6: Knowledge-base Rulesrefined U Rulesupdated

7: **end procedure**

in steps 1 - 7, sample relevant rules that have contributed in performing the current task for refinement. In steps 3 and 4, the confidence and utility of these rules are increased as a token of their quality and relevance. In procedure *updateRules*, steps 8 - 12 all the rules in the knowledge-base are updated by having their age increased by one. In other words, all the rules get older by a task, in Step 10. In *ruleDetachment* procedure steps 13 - 24, all the rules undergo a detachment mechanism, that filters the rules which have failed to maintain their confidence, step 15, 16. They are removed for being wrongly learnt or having inconsistent confidence in different tasks. Similarly, the rules having high confidence but fail to prove their utility over a longer period of time are filtered out as irrelevant, in steps 19, 20. The rules that have retained their place in the knowledgebase by proving both relevant and consistent are prioritized by these two attributes in steps 17 and 21. Thus, the rules are consistently competing for their place in the knowledge-base while during their life as a rule, they are ranked to decide their contribution in a task and to resolve conflict if contradictory rules are learnt.

Algorithm 2 RefineRules (D^t, N, topics, V, Rules)

1: **procedure** REFINERULES

2: *Rules* $\leftarrow$ relevantRules(V, Knowledge-base);

3: **for** (rule r : Rules/) **do**

4: $v \leftarrow$ nPMI(r) //update confidence

5: $\rho \leftarrow +1$ //increase utility

6: **end for**

7: **end procedure**

8: **procedure** UPDATERULES

9: **for** (rule r : All Rules) **do**

10: $\lambda \leftarrow +1$ //Rules gets older by a task

11: **end for**

12: **end procedure**

13: **procedure** RULEDETACHMENT

14: for (rule r : All Rules) do

15: **if** $v_r < t_1 \,\|\, v_r > t_2$ **then**

16: drop *r*

17: **else***rank* $\leftarrow$ *prioritize*(v, ρ)

18: **end if**

19: **if** $\frac{\rho}{\lambda} < \omega$ then

20: drop *r*

21: else*rank* $\leftarrow$ prioritize(v, ρ)

22: **end if**

23: **end for**

24: **end procedure**

4. Experiment And Results

The proposed model is experimented for improvement in topic coherence through the use of the long-term dependency rules learning. The experiments are performed on Amazon. com real users dataset as large-scale data. The dataset consists of 100 domains with 50 each about electronic and non-electronic subject domains. Each domain has thousands of documents which consists of actual user reviews about the product of domain. The initial parameters, number of topics and number of top words

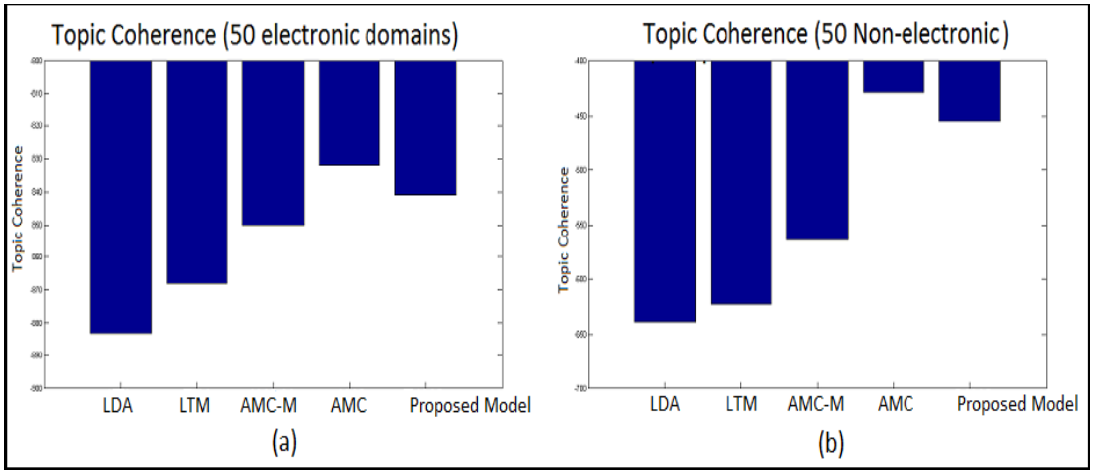

Figure 5. Comparison of topic coherence of the proposed model with Lifelong learning models

in a topic for calculating topic coherence, Gibbs sampler iterations are kept same with the models used in literature. The value of t_1 for learning mustlinks is set to 0.8 while the value of t_2 for learning cannotlinks is set to -0.9. The value of $\in$ is set to 0.3 as minimum overlap between context of a rule and vocabulary of the current task. The value of ω for weeding out irrelevant rules is kept at 0.02. The experiments are performed on HP Pavilion Dm4 machine having Intel 2.4GHz Core i5, with 8GB RAM and 500GB Hard drive.

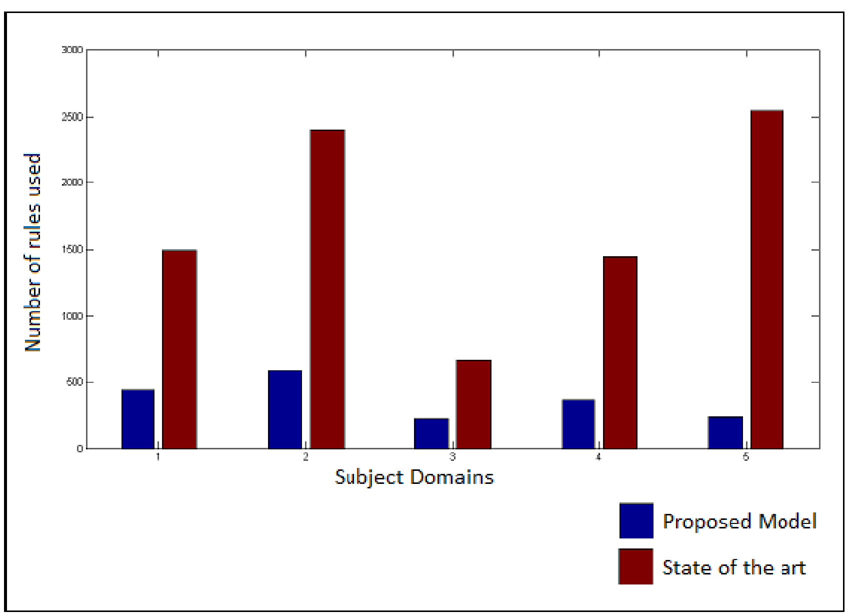

Figure 6. Comparison of rules used by the proposed model and state of the art model in selected domains

Experimental results performed on model shows that the proposed model has higher topics coherence. Comparison of topic coherence of the proposed model with state of the art models is given in Fig. 5. The improvement in topic coherence averagingto 34 points for all domains. It is achieved by using 21% fewer number of rules as compared to state of the art model as shown in Fig. 6. Improving results with 1000 rules lesser than the other models, is a token of the quality of rules learnt by the proposed model. This attributes to the refinement and detachment mechanism of the proposed model that assumes that a wrong or irrelevant rules may make it to the knowledge-base, therefore, instead of tightening the evaluation mechanism a monitoring and refinement mechanism is applied to keep a check on the quality of the rules. It keeps the knowledge-base consistent which results in improvement in performance of the model as well. Table I has the mustlinks and cannotlinks at high rank in the knowledge-base.

Type	Knowledge Rules
mustlinks	(tech, support), (cable, usb), (batter, charge),(video, card), (high, long),(connection, vga), (operating, system), (cd, dvd), (mac, support), (pro, con), (windows, xp), (money, worth), (inch, picture),(monitor, macbook), (customer, support), (big, deal)
cannotlinks	(sound, money), (feature, battery), (price, device),(screen, sound), (signal, easy), (home, gps), (review, screen), (worse, digital), (price, easy), (gas, screen), (battery, price), (feature, battery), (hotel, traffic), (driver, cooler), (monitor, processor), (monitor, speed)

Table 1. High Priority Must Links And Cannot Links

5. Conclusions

Lifelong learning models are used with large-scale data to process them with intuition, however, without involving manual support. The model is highly scalable and can be applied to any type of data. The model helps improve results for each task while learn something for it or update the state of its previous knowledge about it as a self rectifying mechanism. The existing Lifelong learning topic models have a rule extraction mechanism but do not have a monitoring and refinement mechanism to watch the contribution of a rule and treat it accordingly. This has led these models process incorporate increasing number of rules as the model grows in experience. The proposed model has comparatively produced better accuracy by using fewer number of rules. It has given a new perspective to lifelong learning models that by supporting the learning mechanism with additional techniques the quality of rules can be improved which results in improvement in results of the model. It has improved topic coherence by 34 points while utilizing 21% fewer rules as compared to the other model.

In future more variety of learning techniques can be incorporated into the lifelong learning model that would not only help in grouping words in the topics but will also identify the nature of a topic. Such a model can be used to associate topics with each other and explore their relationships.

References

[1] Blei, D. M., Ng, A. Y., Jordan, M. I. (2003). Latent dirichlet allocation, *The Journal of Machine Learning Research,* 3, p. 993–1022.

[2] Mimno, D., Wallach, H. M. Talley, E. Leenders, M. McCallum, A. (2011). Optimizing semantic coherence in topic models, *In*: Proceedings of the Conference on Empirical Methods in Natural Language Processing, p. 262–272, Association for Computational Linguistics.

[3] Katz, G.Ofek, N., Shapira, B. (2015). Consent, *Knowledge- Based Systems*, 84 (C) 162–178.

[4]Tang, J., Chang, Y., Liu, H. (2014). Mining social media with social theories: a survey, *ACM SIGKDD Explorations Newsletter*, 15 (2) 20–29.

[5] Guellil, I., Boukhalfa, K (2015). Social big data mining: A survey focused on opinion mining and sentiments analysis, *In*: Programming and Systems (ISPS), 2015 12th International Symposium on, p. 1–10, IEEE, 2015.

[6] Takeuchi, H., Yamaguchi, T (2014). Text mining of businessoriented conversations at a call center, *In:* Data Mining for Service, p. 111–129, Springer.

[7] Cailliau, F., Cavet, A (2013). Mining automatic speech transcripts for the retrieval of problematic calls, *In:* Computational Linguistics and Intelligent Text Processing, p. 83–95, Springer.

[8] Gangemi, A. , Presutti, V., Reforgiato Recupero, D. (2014). Frame-based detection of opinion holders and topics: a model and a tool, *Computational Intelligence Magazine, IEEE*, 9 (1) 20–30.

[9] Weichselbraun, A. Gindl, S. Scharl, A. (2014). Enriching semantic knowledge bases for opinion mining in big data applications, Knowledge-based systems, 69, p. 78– 85.

[10] Ofek, N., Rokach, L., Mitra, P (2014). Methodology for connecting nouns to their modifying adjectives, *In:* Computational Linguistics and Intelligent Text Processing, p. 271–284, Springer, 2014.

[11] Xia, R., Zong, C., Hu, X., Cambria, E. (2013). Feature ensemble plus sample selection: domain adaptation for sentiment classification, Intelligent Systems, IEEE, 28 (3) 10–18, 2013.

[12] Katz, G., Elovici, Y., Shapira, B. (2014). Coban: A context based model for data leakage prevention, *Information Sciences*, V 262, p. 137–158.

[13] Hai, Z., Chang, K., Kim, J.-J., Yang, C. C. (2014). Identifying features in opinion mining via intrinsic and extrinsic domain relevance, *Knowledge and Data Engineering, IEEE Transactions on*, 26 (3) 623–634.

[14] Poria, S., Gelbukh, A., Hussain, A., Howard, N. Das, D., Bandyopadhyay, S (2013). Enhanced senticnet with affective labels for concept-based opinion mining, *IEEE Intelligent Systems*, 2, p. 31–38,

[15] Machova, K., Marhefka, L (2014). Opinion classification in conversational content using n-grams, *In:* Recent Developments in Computational Collective Intelligence, p. 177–186, Springer, 2014.

[16] Medhat, W. Hassan,A. Korashy, H.(2014). Sentiment analysis algorithms and applications: A survey, *A in Shams Engineering Journal*, 5 (4) 1093–1113.

[17] Titov, I., McDonald, R. T. (2008). A joint model of text and aspect ratings for sentiment summarization, *In:* ACL, vol. 8, p. 308–316, Citeseer.

[18] Mei,Q., Ling, X. ,Wondra, M Su, H. Zhai, C. (2007). Topic sentiment mixture: modeling facets and opinions in weblogs, *In:* Proceedings of the 16th international conference on World Wide Web, p. 171–180, ACM.

[19] Titov, I., McDonald, R. (2008). Modeling online reviews with multi-grain topic models, *In:* Proceedings of the 17th International Conference on World Wide Web, p. 111–120, ACM.

[20] Branavan, S., Chen, H., Eisenstein, J., Barzilay, R (2009). Learning document-level semantic properties from freetext annotations, *Journal of Artificial Intelligence Research*, p. 569–603.

[21] Zhao, W. X. Jiang, J. Yan, H. Li, X. (2010). Jointly modeling aspects and opinions with a maxent-lda hybrid, *In:* Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, p. 56–65, Association for Computational Linguistics.

[22] Griffiths, T. L. Steyvers, M. Blei, D. M., Tenenbaum,J. B.(2004). Integrating topics and syntax, *In:* Advances in Neural information processing systems, p. 537–544.

[23] Andrzejewski, D., Zhu, X (2009). Latent dirichlet allocation with topic-in-set knowledge, *In:* Proceedings of the NAACL HLT 2009 Workshop on Semi-Supervised Learning for Natural Language Processing, p. 43–48, Association for Computational

Linguistics, 2009.

[24] Chen, Z. Mukherjee, A. Liu, B. (2014). Aspect extraction with automated prior knowledge learning, *In:* Proceedings of ACL, p. 347–358, 2014.

[25] Liu, B., Hsu, W. , Ma, Y. (1999). Mining association rules with multiple minimum supports, *In:* Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data mining, p. 337–341, ACM, 1999.

[26] Chen Z., Liu, B (2014). Mining topics in documents: standing on the shoulders of big data, *In:* Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data mining, p. 1116–1125, ACM, 2014.

[27] Mahmoud, H. (2008). P´olya urn models. CRC Press, 2008.

[28] Burke, R. (2002). Hybrid recommender systems: Survey and experiments, User modeling and user-adapted interaction, 12 (4) 331–370.

[29] Burke, R. (2007). Hybrid web recommender systems, *In:* The adaptive web, p. 377–408, Springer.

[30] Tang, J., Hu, X., Liu, H. (2013). Social recommendation: a review, *Social Network Analysis and Mining,* 3 (4) 1113–1133.