Acta Informatica Pragensia, 2018, 7(2), 152–165
DOI: 10.18267/j.aip.120
Original Article

Review of Latent Dirichlet Allocation Methods Usable in Voice of Customer Analysis

Lucie Sperkova [1]

Abstract

The aim of the article is to detect and review existing topic modelling methods of Latent Dirichlet Allocation and their modifications usable in Voice of Customer analysis. Voice of Customer is expressed mainly through textual comments which often focus on the evaluation of products or services the customer consumes. The most studied data source are customer reviews which contain next to the textual comments also ratings in form of scales. The aim of the topic models is to mine the topics and their aspects the customers are evaluating in their reviews and assign to them a particular sentiment or emotion. The author completed a systematic literature review of peer-reviewed published journal articles indexed in leading databases of Scopus and Web of Science and concerning the current use of Latent Dirichlet Allocation model variants in Voice of Customer textual analysis for performing the tasks of aspect detection, emotion detection, personality detection and sentiment assignation. In total, 38 modifications of the LDA model were identified with the reference to their first application in the research of text analytics. The review is intended for researchers in customer analytics the field of sentiment or emotion detection, and moreover as results from the review, for studies in personality recognition based on the textual data. The review offers a basic overview and comparison of LDA modifications which can be considered as a knowledge baseline for selection in a specific application. The scope of the literature examination is limited to the period of years 2003–2018 with the application relevant to the analysis of Voice of Customer subjective textual data only which is closely connected to the area of marketing or customer relationship management.

Keywords: Aspect detection, LDA, Sentiment, Text analytics, Topic models, VoC.

1 Introduction

In recent years the text analytics is increasingly performed for analysis of customers' opinions and expressed sentiment and emotions. The textual comments from customers are coming from various channels such as emails, feedbacks from surveys, posts from social networks or online reviews as a Voice of Customer (VoC). VoC provides valuable input and feedback about the customers for marketers who want to understand their behaviour.

The customers can evaluate and have opinions about many different topics like products or services their purchased and consumed or events they attended, and their aspects which can represent different features of products or services within one textual contribution. Part of the contribution can be positive or negative even without the subjective opinion of the

[1] Department of Information Technologies, Faculty of Informatics and Statistics, University of Economics, Prague,
W. Churchill Sq. 1938/4, 130 67 Prague, Czech Republic
✉ lucie.sperkova@vse.cz

contributor. Thus, even the overall sentiment of the whole comment is positive; it can also contain negative aspects. Moreover, in reviews where next to the textual comment is also rating in form of the structured numerical Likert-type scale, the rating does not necessarily reflect the evaluation of the customer in his textual expression. At the same time, the consumer can give a full positive rating, but also write a comment evaluating the negative aspects of the product or vice versa.

The aim of aspect detection techniques is the extraction of single topics about them the expressed opinions and emotions are further evaluated. The assumption is that the expressed aspects are linked with the expressed sentiment (Liu, 2015). Many types of research in aspect detection is focused on the online reviews (Hu & Liu, 2004; Titov & McDonald, 2008; Wang et al., 2010; Jo & Oh, 2014; Wallace et al., 2014; Büschken & Allenby 2016; Dong et al., 2018) There are many techniques described in (Liu, 2015) for aspect detection, however, the predominant approach is topic modelling. Topic models can solve the issue with the subjectivity detection of different topics within one contribution as the topic may affect the polarity of sentiment within the same domain. For example, in the domain of restaurants, the adjective "cheap" is positive when discussing food, but negative when talking about the decorations and atmosphere (Brody & Elhadad, 2010). Many otherwise neutral terms acquire a sentiment polarity in the context of a specific aspect. Latent Dirichlet Allocation (LDA) is a probabilistic approach, one of the most performed in topic modelling. During the years, different modification of LDA methods emerged in the area of textual analytics, mainly sentiment analysis.

As there exist higher tens of different variations of LDA (see Jelodar et al., 2017) in different areas of study, the author focuses only on modifications performed on data from VoC domain (reviews, posts, emails, feedbacks) and usable for detection of Customer Experience elements. Customer Experience has a high managerial impact (see Lemon & Verhoef 2016) in customer relationship management. Among these elements, the author includes satisfaction (presented as a sentiment), emotions and personality traits. The elements reflect the perceptions of the target objects which are contained in the customer's opinions and represent entities and their aspects of the Customer Experience that the customer has an opinion about. In other words, Customer Experience is contained in VoC textual data.

The aim of analysing VoC data with LDA method and its modification is to facilitate the manual analysis of the content with the growing number of the textual data. Automated analysis can lead to higher reliability of detection of specific information from VoC. Voice of Customer in its textual form allows to understand the customer himself and focus on his individuality.

2 Latent Dirichlet Allocation

Latent Dirichlet Allocation (LDA) developed by (Blei et al., 2003) is an unsupervised learning model based on Bayesian networks searching for the semantic structure of the text set by mining the topics in the text. It assumes that each document consists of a mixture of latent topics where each topic has its own multinomial distribution over a fixed Bag-of-Words vocabulary.

LDA overcomes the disadvantage of another frequently used approach in topic modelling – Latent Semantic Analysis (LSA) (Deerwester, 1988), which is its dependency on annotated training data and proneness to overfitting. The probabilistic variation of LSA – Probabilistic LSA (PLSA) (Hofmann, 1999) learns a distribution over the topic for each document in the training set. The number of model parameters then grow linearly with the size of the training data. LSA does not solve the problem with different meanings of the same word since the meaning of the word can be conditioned by other words in the document. Failure arises from the fact that every word has only one assigned point in the semantic space of LSA. In PLSA the topic distribution is only learned from those documents that are in the training set, so it cannot generate topics from the previously unseen document.

The basic generative process of LDA closely resembles PLSA (Lu, 2010) but learns topic distribution as a random vector of parameters and models it by a Dirichlet prior. Each parameter can have a different value. In LDA, however, most researchers use the same value for all parameters (symmetric Dirichlet distribution). To obtain the distributions, commonly used methods of derivation are a variational inference (Blei et al., 2003) and Gibbs sampling (Griffiths & Steyvers, 2004). LDA model has a high extensibility and steady mathematical basis.

3 Methodology and Results

The aim of this review is to detect the existing topic modelling methods of Latent Dirichlet Allocation and their modifications usable in Voice of Customer analysis for marketing, specifically in the detection of Customer Experience elements. The research draws on academic peer-reviewed journal articles and conference proceedings searched in online databases of Scopus and Web of Science. Other document types as book chapters, theses, reviews, letters, reports were omitted from the results. The systematic search was performed using Boolean operators with combinations of keywords including "Latent Dirichlet Allocation", "LDA", "joint model", joint topic model", "topic modelling", "sentiment", "aspect" and "emotion". Articles with no applicability to current research were removed from the list. Figure 1 shows the research methodology following the PRISMA principles (Moher et al., 2009).

There was identified 3281 papers indexed in years since 2004 with the most articles written in 2017 (482) and 2016 (473). The LDA algorithm for topic discovery is relatively new, designed in 2003 by (Blei et al., 2003). Literature review covers the whole history of the research in topic modelling with LDA method since its design in 2003 to 2018 (March). Within the screening task (see task 2 in Figure 1), the keyword "joint model" had the strongest power on downsizing the sample as LDA modification are almost always joint models. As there exist various application of LDA including different areas of research (for example for image classification), the author had to exclude articles which have not to correspond to the area of analysing the Voice of Customer textual data. The databases results were refined by the subject area of the research to computer science, mathematics, engineering, social science, decision sciences, neurocomputing and business, management and accounting as other subject areas have no connection to VoC. This refinement along with screening the abstracts limited the results to 81 documents. The rest of the articles were downloaded for full-texts and manually sorted and assessed for eligibility (task 3 in Figure 1) according to their content. Based on the references in detected journal papers were further also assessed articles not discovered by keyword search but contained in selected databases. In total, 49 articles relevant to this study were identified and assessed as eligible for this review.

The exhaustive list of existing LDA variations was prepared in (Jelodar et al., 2017). The authors included all scholar papers to the review (author of this paper focus only on peer-reviewed journals from recognized databases Scopus and Web-of-Science) from all areas of research they found. However, the systematic approach of the literature review is missing and not all modification of LDA detected by the author of this study are incorporated. The performance of the models is not discussed along with the description of the models. Authors see the connection of topic modelling with different disciplines such as politics, linguistic or medicine, but not marketing or customer relationship management. This gap is bridged in this work.

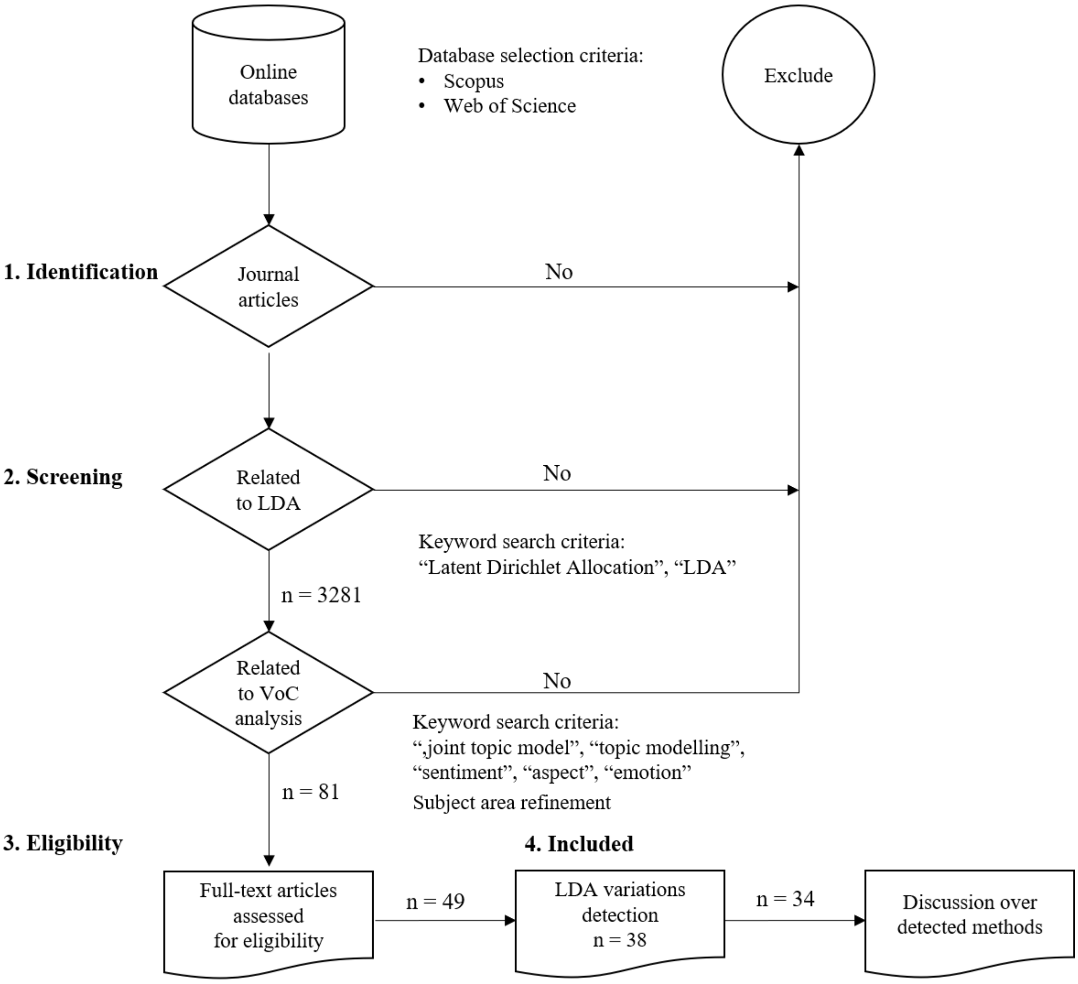

Fig. 1. *Selection criteria and evaluation framework. Source Author.*

The paper refers in Table 1 only to those journal papers where the given variation of LDA was first applied. Other articles are excluded from the review. It is common that researchers further modify already modified LDA with refinements or omission of some parameters. Some of these modifications are absent in the list as they do not change the nature of the original model or improve the model significantly. Table 1 shows the result of 38 modifications of the LDA algorithm from 34 papers with the short description of the principle and performance of the model. There exists more modification of LDA in research, but not related to the sentiment analysis, emotion detection or aspect detection. The further section discusses detected methods.

4 Discussion over Detected Modifications of LDA

Many researchers in text analytics prefer to extend and apply LDA in sentiment analysis and several variations of the model emerged over the years to get more accurate results or refine LDA to the needs of the specific task. This chapter discusses the different modification of LDA which are summarised in Table 1 along with the short description of the principle they work on and given performance stated in the papers. The performance is measured in papers differently with various statistical measures. Not all researchers provide the specific numbers of the results and just describe the performance in the text. Also, it is not possible to compare the performance of all methods among themselves as every method is dedicated to a slightly different task.

The original LDA model is built based on the Bag-of-Words (BoW) principle, where the words in the document are treated as unordered unigrams ignoring the complex and structural relationships between the words. The Collocation model (LDA-Col) in (Griffiths et al., 2007) is a variation of LDA with the effort to overcome the Bag-of-Words assumption of LDA with incorporating a hidden variable bearing the information if the word is a part of a collocation with another world. This word tends to follow the previous word with high frequency. Collocations carry more semantic information then if the words are treated as unigrams. This model performs as well as the standard Hidden Markov Model (HMM) proposed for identifying syntactic classes in (Griffiths et al., 2005).

According to (Titov & McDonald, 2008a), standard LDA is not suitable for aspect detection in the review. LDA tends to capture more global topics in the data than rateable aspects relating to the review. Their multi-grain topic model (MG-LDA) detects two layers of themes – global and local, where local topics corresponding to rateable aspects. MG-LDA distinguishes dozens of local topics, but mapping many-to-one between them and rateable aspects is not explicit. Because the MG-LDA cannot learn the whole topic distribution for a review, in (Titov & McDonald, 2008b) they extended the model to Multi-Aspect Sentiment model (MAS) and derived such a mapping with aspect-specific evaluation provided along with the review text. Only those aspects rated by users in some review are detected in the text and sentiment for every aspect is aggregated. This supervised setting is not practical in real applications.

Lin and He (2009) proposed unsupervised four-level Joint Sentiment/Topic model (JST) model based on LDA more suitable for the sentiment classification task. The only minimal prior information being incorporated to mine the pairs of sentiment and topic. Sentiment labels are associated with documents, under which topics are associated with sentiment labels and words are associated with both sentiment labels and topics. In further research, the same authors (Lin et al., 2012) implemented a reparametrized version of JST, Reverse-JST (RJST) where sentiment label generation is dependent on topics. Dong et al. (2018) propose the Unsupervised Topic-Sentiment Joint probabilistic model (UTSJ) to detect the misleading reviews. UTSJ model first considers the topic and then the sentiment of each review consistently to the written expression habit of the contributor.

Brody and Elhadad (2010) detect aspects throughout the local version of the LDA operating at the sentence rather than document level. For the opinion words, they consider only the adjectives, which may potentially miss opinion words with other Part-of-Speech tags and use only a small number of aspects that directly correspond to the topics. This approach solves the problem of frequent-term methods. They use morphological indicators of negation to automatically create a basic set of highly relevant positive and negative adjectives, which are guaranteed to be relevant to the aspect. These automatically derived basic sets achieve comparable results as when using manually created sets. The author refers in Table 1 to this approach as LocLDA same as (Zhao et al. 2010).

The first modification of LDA for joint sentiment and topic modelling performed probably (Blei & McAuliffe 2007) with supervised LDA (sLDA) model. Model infers topics in a given classification or regression problem. Zhao et al. (2010) proposed a MaxEnt-LDA. A hybrid model of maximum entropy (MaxEnt) and LDA discovers separately aspects and aspect-specific opinion words jointly but does not separate positive and negative polarity. MaxEnt-LDA represents a rather fine-grained model which can find specific aspect word, a general aspect word, an opinion word specific to the aspect, a generic opinion word, or a commonly used background word.

Naïve Bayes and Latent Dirichlet Allocation (NB-LDA) (Zhang et al., 2013) does not assume documents as Bag-of-Words but rather Bag-of-Sentences. Each sentence has assigned a latent sentiment label drawn from the distribution of sentiment over the document. The words or features in the sentence are generated by the latent sentiment label in a Naïve Bayes manner.

Other variations of LDA models also try to solve the problem of finding multiword phrases as topics. Constrained-LDA model (Zhai et al., 2011), and LDA(p_GPU) model based on the generalized Pólya urn model (GPU) (Fei, Chen & Liu 2014) find phrases first and then run two special topic models, respectively. Constrained-LDA models similarity with must-links and cannot-links. LDA(p_GPU) treats phrases as individual terms and allows their component words to have some connections or co-occurrences with them.

Phrase-LDA (PLDA) (Zhan & Li, 2011) is a basic LDA model which supposes that each phrase of the review is related to the single topic. Tang et al. (2016) combined PLDA with Labeled LDA (Ramage et al. 2009) to Labeled Phrase LDA (LPLDA) where each document is assumed as bag-of-phrases and the target topics are restricted of each labelled document to the label set of the document. Moghaddam and Ester (2012) further extended the PLDA to learn both aspects and ratings from phases simultaneously in Separate-PLDA (S-PLDA) and Dependency-PLDA (D-PLDA) where the dependency between the aspects and ratings is added.

Sentence-LDA (Jo & Oh, 2011) and Sentence-Constrained-LDA (SC-LDA) model (Büschken & Allenby 2015) are same models constraining the LDA so that words within a sentence are generated from one topic. In the same study, Jo and Oh (2011) extended Sentence-LDA to discover sentiments related to each aspect in a sentence with unsupervised Aspect and Sentiment Unification Model (ASUM) (Jo & Oh, 2011). The model discovers the pairs of sentiment and aspect (senti-aspect) by generating the distribution over sentiments of the review and aspect detection from the distribution of given sentiment over aspects. ASUM is similar to JST, but unlike JST it constraints the individual words in a single sentence to the same language model.

Li et al. (2010) developed Sentiment-LDA (SLDA) and Dependency-Sentiment-LDA (DSLDA) models to find aspects with positive and negative sentiments. The dependency model assumes that the sentiments of words form a Markov chain where the sentiment of a word is dependent on the previous word. It does not find aspects independently and does not separate aspect words and sentiment words. Sauper, Haghighi and Barzilay (2011) combined topic modelling with a hidden Markov model (HMM-LDA) same as (Griffiths et al., 2005). HMM models the word sequences with types (aspect word, sentiment word, or background word).

Factorial-LDA (f-LDA) in (Wallace et al. 2014) assumes words in document reflect a mixture of vectors of latent topics rather than single topics as in classic LDA. Each word is associated with one variable representing its *aspect* and one its *sentiment*. The used dataset was manually labelled by experts.

Wang et al. (2018) suppose that individual words in a document have either strong or weak ability to convey objective facts or subjective opinions depending on the assigned topic. The supervised model called identified objective-subjective latent Dirichlet allocation (iosLDA) incorporates the simple Pólya urn (SPU) model with a probabilistic generative process to obtain the Bag-of-Discriminative-Words (BoDW) representation for the documents. Each document has two BoDW representations regarding objective and subjective senses respectively. These representations are then employed in the joint objective and subjective classification. iosLDA has better computational performance than sLDA, especially with the growing number of topics.

Xu et al. (2018) propose a Time-User Sentiment/Topic Latent Dirichlet Allocation (TUS-LDA) which simultaneously models sentiments and topics for VoC posts. TUS-LDA aggregates posts in the same time slices or from the same users as pseudo-documents to mitigate the context problem. They use document-level sentiment distribution to detect the sentiment and words of a post and within the specific sentiment, the topic is then drawn from time-level or user-level sentiment/topic distribution.

4.1 Modifications with Assigned Prior Knowledge

The main issue of aspect-based modelling is that it needs a large volume of data and a significant amount of tuning to achieve reasonable results (Liu, 2016). The results of unsupervised LDA may be non-existing or non-interpretable topics. Andrzejewski et al. (2009) thus incorporate prior domain knowledge in their Dirichlet Forrest-LDA (DF-LDA) model. The knowledge is in the form of must-links and cannot-links. Must-link states that two words should belong to the same topic whereas a cannot-link indicates that two words should not be on the same topic. The shortcoming is that the model cannot solve multiple meaning of the words because of the characteristic of the links to enforce transitivity. Chen et al. (2013a) overcome this drawback by MC-LDA (LDA with M-set and C-set). MC-LDA adds a new latent variable in LDA to distinguish multiple senses. To exploit the knowledge from multiple domains in (Chen et al. 2013b) authors created Multi-Domain Knowledge LDA (MDK-LDA) which they further improved in (Chen et al., 2013c) to identify domain independent wrong knowledge especially the incorrect semantic relationships. The General Knowledge based LDA (GK-LDA) generates lexical relation sets which are automatically learned from the documents. Further, Chen et al. (2014) designed Automated Knowledge LDA (AKL), which can exploit the automatically learned prior knowledge and deal with the issue of incorrect knowledge to produce superior aspects.

4.2 Modifications of LDA Used in Emotion Mining

Joint topic modelling is also employed in emotion detection. Emotion topic model (ETM) (Bao et al., 2009, 2012) is a joint emotion-topic model adding an intermediate layer into LDA, where a topic represents a component of emotion. The Emotion-LDA (Rao et al., 2014a) captures social emotions at the topic level and generates topics without the supervision of emotion labels. Multi-label supervised topic model (MSTM) (Rao et al., 2014a) is an extension of the supervised LDA (Blei & McAuliffe, 2007). The model first generates a set of topics from words and then samples emotions from each topic. Sentiment latent topic model (SLTM), also from (Rao et al., 2014b) generates topics directly from social emotions. The affective topic model (ATM) (Rao et al., 2014c) generates emotion lexicon (emotion-annotated topics) and predicts emotions to unlabelled documents. The model is intended to predict emotions from reader's perspective, not the writer's point of view, which is necessary for Customer Experience, but can be used as an emotion lexicon.

Liu, Wang & Jiang (2016) use LDA in their research on personality recognition. They understand the issue as a multi-label classification problem as every person can have more personality traits with different strength. They use a probabilistic topic model based on LDA to predict the personality traits within the framework of the Five-Factor Model (McCrae & John, 1992). Each topic is characterised by five Gaussian distributions over five personality traits as users with different personalities may publish various topics. Model jointly integrates the personality traits through mixture Gaussian distributions and predicts the personality strength simultaneously.

With the topic of emotion and personality detection is also related work of (Huang et al., 2017). Their multimodal joint sentiment topic model (MJST) inserts an additional sentiment layer into LDA and takes multimodal data such as emoticon or personality and text into consideration while inferring message sentiment. Joint Author Sentiment Topic Model (JAST) in (Mukherjee et al., 2014) uses LDA to learn the distribution of author-specific topic preferences and emotional attachment to topics. The model uses HMM to capture the short-range syntactic and long-range semantic dependencies in reviews to detect consistency in author writing style. JAST jointly discovers the topics in a review, topic ratings, author preferences for the topics and the overall review rating from the author's point of view.

Tab. 1. *Variations of LDA model in current research of topic modelling. Source Author.*

Model	**Author(s)**	**Principle**	**Performance in tasks**
HMM-LDA	Griffiths et al., 2005	The model considers both word sequence and word-bag.	Mean accuracy of 48%(0.06), 48%(0.05), 46%(0.08) where number in parentheses show standard error in document classification
LDA-Collocation	Griffiths et al., 2007	The model considers whether or not the word is a part of a collocation with the previous word.	More accurate results than LSA
Multi Grain-LDA	Titov & McDonald, 2008a	The model captures two layers of themes – global and local, where local topics correspond to rateable aspects.	Best accuracy of 74.8% in multi-aspect discovery on hotel review unigrams
Multi-Aspect Sentiment model	Titov & McDonald, 2008b	Extension of MG-LDA for sentiment predictors.	The average precision 75.8% and 85.5% for aspects service and location on hotel review data which is comparable to supervised methods
Supervised-LDA	Blei & McAuliffe, 2007	The model infers topics appropriate for use in a given classification or regression problem.	sLDA overperformed LDA in aspect discovery
Labeled LDA	Ramage et al., 2009	Topics of each document are restricted to its labels.	LLDA outperformed SVM in multi-label text classification
Joint Sentiment/Topic model (JST)	Lin & He, 2009	Topics are generated conditioned on a sentiment polarity.	The best overall accuracy 86.2% in document classification based on the filtered subjectivity lexicon and the subjective dataset. JST outperformed Reverse-JST
Reverse-JST	Lin et al., 2012	Reparametrized JST. Sentiments are generated conditioned on a topic.	The best overall accuracy 75% in document classification based on the filtered subjectivity lexicon and the subjective dataset.

Model	Author(s)	Principle	Performance in tasks
Emotion Topic model	Bao et al., 2009	In the model, the topic acts as an important component of emotion.	The model outperformed the baseline model with the improvement of 34% in accuracy in social emotions prediction
Dirichlet Forrest-LDA	Andrzejewski et al., 2009	The models incorporate prior domain knowledge in the form of must-links and cannot-links.	DF-LDA overperformed LDA in identifying good topics
LocLDA	Brody & Elhadad, 2010	The model detects aspects throughout the local version of the LDA	The model can compete to manual seed set when is not available
Maximum Entropy-LDA	Zhao et al., 2010	The model discovers both aspects and aspect-specific opinion words.	Results similar to LocLDA on restaurant dataset, but it can model aspects and opinion worlds separately
Sentiment-LDA	Li et al., 2010	The model classifies the overall sentiment polarity for the document, but also calculate the polarity for each topic.	Accuracy around 60–65% outperformed lexicon-based methods in detecting topics and sentiment simultaneously
Dependency-Sentiment-LDA	Li et al., 2010	The model considers the dependency between aspects and their ratings.	Improves S-LDA by 3–5% in detecting topics and sentiment simultaneously
Sentence-LDA	Jo & Oh, 2011	The model imposes a constraint that all words in a sentence are generated from one topic.	S-LDA performed better than LDA in aspects discovery on restaurant and electronic review data
Aspect and Sentiment Unification Model	Jo & Oh, 2011	Extension of Sentence-LDA that incorporates both aspect and sentiment into pairs.	ASUM outperformed JST with accuracies 78%, 79%, 84% and 86% in senti-discovery on restaurant and electronic review data
Phrase-LDA (PLDA)	Zhan & Li, 2011	The basic LDA model which learns general aspects from opinion phrases.	The precision of 0.83 outperformed LDA on labelled product review data
Constrained-LDA	Zhai et al., 2011	Modelling similarity with must-links and cannot-links.	Outperformed LDA and multilevel mLSA by a large margin (up to 10%) in detecting topics in product review dataset
Separate-PLDA	Moghaddam & Ester, 2012	Extension of Phrase-LDA learns both aspects and ratings from phrases.	The precision of 0.83 outperformed LDA, PLDA on labelled product review data
Dependency-PLDA	Moghaddam & Ester, 2012	Extension of Phrase-LDA adding the dependency between ratings and aspects.	The precision of 0.87 outperformed LDA, PLDA, S-PLDA on labelled product review data
MC-LDA	Chen et al., 2013a	Latent variable distinguishes the multiple sense with M-set and C-set.	MC-LDA outperformed DF-LDA, LDA in identifying good topics on product review data
MDK-LDA	Chen et al., 2013b	The model identifies a right lexical relation sets for each word.	MDK-LDA overperformed LDA DF-LDA, LDA, LDA-GPU in identifying good topics on product review data
GK-LDA	Chen et al., 2013c	The model identifies wrong knowledge during the modelling.	GK-LDA overperformed MDK-LDA, DF-LDA, LDA, LDA-GPU in identifying good topics on product review data

Model	**Author(s)**	**Principle**	**Performance in tasks**
Naïve Bayes LDA	Zhang et al., 2013	The model assumes that each sentence instead of the word has a latent sentiment label.	The accuracy 71.85% outperformed JST on movie review dataset in sentiment classification on product review dataset
Joint Author Sentiment Topic Model	Mukherjee et al., 2014	The model jointly discovers the topics, author preferences for the topics, topic ratings and the overall review rating from the point of view of an author.	The accuracy of 87.69% on movie review dataset outperformed JST
LDA(p_GPU)	Fei et al., 2014	The model treats phrases as individual terms and allows their component words to have some connections or co-occurrences with them.	Outperformed LDA significantly t based on a paired t-test ($p < 0.01$) on product review dataset
Automated Knowledge-LDA	Chen et al., 2014	The model exploits the automatically learned prior knowledge and also deal with the issue of incorrect knowledge to produce superior aspects	AKL outperformed LDA, GK-LDA, MC-LDA significantly t based on a paired t-test ($p < 0.0001$) in topic detection on product review data
Factorial-LDA	Wallace et al., 2014	The model associates each review with a joint distribution over aspects and sentiment.	f-LDA predicts user ratings in doctor review dataset with lower error than LDA
Emotion-LDA	Rao et al., 2014a	The model captures emotions at the topic level and generates topics without the supervision of emotion labels.	ELDA gained similar performance results as ETM in social emotion classification of online news with an accuracy of 55.6%
Multi-label Supervised Topic Model	Rao et al., 2014b	The model associates each topic with social emotions jointly.	MSTM outperformed ETM by 0.44% of averaged accuracy and with more stable performance
Affective Topic Model	Rao et al., 2014c	Te model associates each topic with word tokens and social emotions jointly and predicts probabilities of emotions contained in unlabelled documents.	ATM outperformed ETM, and MSMT with the highest accuracy of 74.78% in detecting social emotions from news articles
Sentence-Constrained LDA	Büschken & Allenby, 2015	Words within a sentence pertaining to the same topic.	More accurate prediction of consumer rating in reviews than LDA.
Probabilistic Topic-LDA	Liu et al., 2016	The model deals with multiple continuous labels through a mixture of Gaussian distributions and jointly predicts their strength.	PT-LDA overperformed sLDA in topic detection and support vector regression methods in personality traits detection on a Facebook dataset with RMSE (0.481, 0.636, 0.754, 0.557, 0.763) for different personalities
Labeled Phrase LDA	Tang et al., 2016	Each document is assumed as bag-of-phrases and the target topics are restricted of each labelled document to the label set of the document	LPLDA is more efficient in scalability and finding the most appropriate phrases than PLDA

Model	Author(s)	Principle	Performance in tasks
Multimodal Joint Sentiment-Topic model	Huang et al., 2017	The model takes emoticons as items and introduces personality factor to adaptively adjust sentiment polarities of messages.	MJST outperformed SVM, JST, S-LDA, D-PLDA with the best accuracy 70.75% in sentiment classification on microblogging messages
Unsupervised Topic-Sentiment Joint probabilistic model	Dong et al., 2018	The model classifies reviews according to the extracted topic features as well as corresponding sentiment features.	Precision 0.87 and F-measure 0.85 on hotel review balanced dataset outperformed and precision 0.84 and F-measure 0.85 on hotel review unbalanced dataset overperformed JST in detecting deceptive reviews
objective–subjective-LDA	Wang et al., 2018	The model jointly discovers latent topics and their objectivity or subjectivity.	iosLDA overperformed sLDA with accuracy 81.5% on multi-domain sentiment dataset, 79% on the Twitter dataset in sentiment detection
Time-User Sentiment/Topic-LDA	Xu et al., 2018	The model captures the sentiment-aware topics from posts, but also monitor the variations of sentiment-aware topics over time.	In detecting bursty sentiment-aware topics gained precision < 0.86 TUS-LDA outperformed JST, ASUM in sentiment classification

5 Conclusions

This paper serves as a review of modification of LDA methods for analysing the customer textual data VoC which author systematically identified in Scopus and Web of Science databases. The topic models especially based on LDA have been intensively studied during the last two decades and widely applied in many text analytics tasks for aspect detection as is seen from the review. LDA, unlike the other aspect detection methods, provides a more objective approach to the analysis of the contexts of reviews because of its mathematical characteristics. It can be used in any language.

As LDA can find topics without a prior knowledge, the results of detection can contain topics unobservable by supervised methods or LSA. Due to this reason is LDA also language-independent. However, the results may contain non-existing or non-interpretable words which LSA prevent with lexicons. Thus, LDA or its unsupervised variations, in the discriminative tasks in discovering the topics and latent sentiments where the number of classes is known a priori can have worse predictive results than supervised methods. Therefore, researchers incorporate in LDA modifications some prior knowledge with Bayesian techniques.

Most supervised extensions of LDA (e.g. sLDA, JST, MG-LDA) utilize the Bag-of-Topics (latent topics representing the document), Bag-of-Sentence or similar representation of one document for the prediction of its corresponding label, in which the proportion of topics (instead of the words represented as Bag-of-Words) in the document is considered to be the predictive feature. Modifications of LDA usually contain joint modelling of both aspects and sentiment words (Blei & McAuliffe, 2007; Lin & He, 2009; Brody & Elhadad, 2010, Dong et al., 2018), and joint modelling of aspects and sentiment ratings on the aspects (Titov & McDonald, 2008; Lu et al., 2009; Wang et al., 2010; Lakkaraju et al., 2011; Moghaddam & Ester, 2011). However, most of these modifications model both aspects and sentiments, although they may not separate the two types of words (Jo & Oh, 2011; Mei et al., 2007; Lin & He, 2009). The most recent models detect also subjectivity of the topics or monitor the sentiment over the time.

Apart from mining the aspects and their sentiment, there exist also models which can extract emotions and personality traits from the comments (Bao et al., 2009; Rao et al., 2014a, 2014b, 2014c; Huang et al., 2017). Customer personality and emotions determine how the customer react. The ability to measure customer reactions to the company and its offering is a key element of understanding the customer and managing his experience. For managing relationships with the customers offers Voice of Customer a rich data source next to other customer structure data.

The LDA methods present the opportunity to extract information from customer textual data without prior knowledge and determine the importance of the aspects' impact on overall customer satisfaction. Customer satisfaction is critical in the current customer-oriented approach in business. LDA allows for exploration of dynamics over time due to data at a highly granular temporal level. The amount of Voice of Customer data coming through company channels is growing significantly and their automatic analysis facilitates their evaluation and fast reaction. This study serves as a knowledge base of LDA modifications that may be used to guide future research in the application the appropriate method.

References

Andrzejewski, D., Zhu, X., & Craven, M. (2009). Incorporating domain knowledge into topic modelling via Dirichlet forest priors. In *Proceedings of the 26th annual international conference on machine learning* (pp. 25-32). New York: ACM. doi: 10.1145/1553374.1553378

Bao, S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D., & Yu, Y. (2009). Joint emotion-topic modeling for social affective text mining. In *Ninth IEEE International Conference on Data Mining* (pp. 699-704). New York: IEEE. doi: 10.1109/ICDM.2009.94

Bao, S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D., & Yu, Y. (2012). Mining social emotions from affective text. *IEEE transactions on knowledge and data engineering*, 24(9), 1658-1670. doi: 10.1109/TKDE.2011.188

Blei, D. M., **& McAuliffe, J. D.** (2007). Supervised topic models. In *Advances in neural information processing systems 20* (pp. 121-128). Neural Information Processing Systems Foundation.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993-1022.

Brody, S., & Elhadad, N. (2010). An unsupervised aspect-sentiment model for online reviews. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 804-812). Stroudsburg: Association for Computational Linguistics.

Büschken, J., & Allenby, G. M. (2016). Sentence-based text analysis for customer reviews. *Marketing Science*, 35(6), 953-975. doi: 10.1287/mksc.2016.0993

Chen, Z., Mukherjee, A., & Liu, B. (2014). Aspect extraction with automated prior knowledge learning. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (Vol. 1, pp. 347-358). Stroudsburg: Association for Computational Linguistics.

Chen, Z., Mukherjee, A., Liu, B., Hsu, M., Castellanos, M., & Ghosh, R. (2013). Leveraging Multi-Domain Prior Knowledge in Topic Models. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence* (pp. 2071-2077). Burlington: Morgan Kaufmann Publisher.

Chen, Z., Mukherjee, A., Liu, B., Hsu, M., Castellanos, M., & Ghosh, R. (2013). Discovering coherent topics using general knowledge. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management* (pp. 209-218). New York: ACM. doi: 10.1145/2505515.2505519

Chen, Z., Mukherjee, A., Liu, B., Hsu, M., Castellanos, M., & Ghosh, R. (2013). Exploiting domain knowledge in aspect extraction. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (pp. 1655-1667). Stroudsburg: Association for Computational Linguistics.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407. doi: 10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9

Dong, L. Y., Ji, S. J., Zhang, C. J., Zhang, Q., Chiu, D. W., Qiu, L. Q., & Li, D. (2018). An unsupervised topic-sentiment joint probabilistic model for detecting deceptive reviews. *Expert Systems with Applications*, 114, 210-223. doi: 10.1016/j.eswa.2018.07.005

Fei, G., Chen, Z., & Liu, B. (2014). Review topic discovery with phrases using the pólya urn model. In *Proceedings of the 25th International Conference on Computational Linguistics: Technical Papers* (pp. 667-676). Stroudsburg: Association for Computational Linguistics.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, 101(suppl 1), 5228-5235. doi: 10.1073/pnas.0307752101

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological review*, 114(2), 211. doi: 10.1037/0033-295X.114.2.211

Griffiths, T. L., Steyvers, M., Blei, D. M., & Tenenbaum, J. B. (2005). Integrating topics and syntax. In *Proceedings of the 18th Annual Conference on Neural Information Processing Systems* (pp. 537-544). Neural information processing systems foundation.

Hofmann, T. (1999). Probabilistic latent semantic analysis. In *Proceedings of the 15th Conference on Uncertainty in artificial intelligence* (pp. 289-296). Burlington: Morgan Kaufmann Publishers.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177). New York: ACM. doi: 10.1145/1014052.1014073

Huang, F., Zhang, S., Zhang, J., & Yu, G. (2017). Multimodal learning for topic sentiment analysis in microblogging. *Neurocomputing*, 253, 144-153. doi: 10.1016/j.neucom.2016.10.086

Jelodar, H., Wang, Y., Yuan, C., & Feng, X. (2017). Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. *ArXiv preprint.* ArXiv:1711.04305. Retrieved November 30, 2018, from: https://arxiv.org/abs/1711.04305

Jo, Y., & Oh, A. H. (2011). Aspect and sentiment unification model for online review analysis. In *Proceedings of the 4th ACM international conference on Web search and data mining* (pp. 815-824). New York: ACM. doi: 10.1145/1935826.1935932

Lakkaraju, H., Bhattacharyya, C., Bhattacharya, I., & Merugu, S. (2011). Exploiting coherence for the simultaneous discovery of latent facets and associated sentiments. In *Proceedings of the 2011 SIAM international conference on data mining* (pp. 498-509). Society for Industrial and Applied Mathematics.

Li, F., Huang, M., & Zhu, X. (2010). Sentiment Analysis with Global Topics and Local Dependency. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence* (Vol. 10, pp. 1371-1376). AAAI Press.

Lin, C., & He, Y. (2009). Joint sentiment/topic model for sentiment analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management* (pp. 375-384). New York: ACM. doi: 10.1145/1645953.1646003

Lin, C., He, Y., Everson, R., & Ruger, S. (2012). Weakly supervised joint sentiment-topic detection from text. *IEEE Transactions on Knowledge and Data engineering*, 24(6), 1134-1145. doi: 10.1109/TKDE.2011.48

Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge: CUP.

Liu, Y., Wang, J., & Jiang, Y. (2016). PT-LDA: A latent variable model to predict personality traits of social network users. *Neurocomputing*, 210, pp. 155-163. doi: 10.1016/j.neucom.2015.10.144

Lu, Y., Mei, Q., & Zhai, C. (2011). Investigating task performance of probabilistic topic models: an empirical study of PLSA and LDA. *Information Retrieval*, 14(2), 178-203. doi: 10.1007/s10791-010-9141-9

Lu, Y., Zhai, C., & Sundaresan, N. (2009). Rated aspect summarization of short comments. In *Proceedings of the 18th international conference on World wide web* (pp. 131-140). New York: ACM. doi: 10.1145/1526709.1526728

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of personality*, 60(2), 175-215. doi: 10.1111/j.1467-6494.1992.tb00970.x

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* (pp. 171-180). New York: ACM.

Moghaddam, S., & Ester, M. (2012). On the design of LDA models for aspect-based opinion mining. In *Proceedings of the 21st ACM international conference on Information and knowledge management* (pp. 803-812). New York: ACM. doi: 10.1145/2396761.2396863

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7), e1000097. doi: 10.1371/journal.pmed.1000097

Mukherjee, S., Basu, G., & Joshi, S. (2014). Joint author sentiment topic model. In *Proceedings of the 2014 SIAM International Conference on Data Mining* (pp. 370-378). Society for Industrial and Applied Mathematics.

Ramage, D., Hall, D., Nallapati, R., & Manning, C. D. (2009). Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (pp. 248-256). Stroudsburg: Association for Computational Linguistics.

Rao, Y., Li, Q., Mao, X., & Wenyin, L. (2014). Sentiment topic models for social emotion mining. *Information Sciences*, 266, 90-100. doi: 10.1016/j.ins.2013.12.059

Rao, Y., Lei, J., Wenyin, L., Li, Q., & Chen, M. (2014). Building emotional dictionary for sentiment analysis of online news. *World Wide Web*, 17(4), 723-742. doi: 10.1007/s11280-013-0221-9

Rao, Y., Li, Q., Wenyin, L., Wu, Q., & Quan, X. (2014). Affective topic model for social emotion detection. *Neural Networks*, 58, 29-37. doi: 10.1016/j.neunet.2014.05.007

Sauper, C., Haghighi, A., & Barzilay, R. (2011). Content models with attitude. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 350-358). Stroudsburg: Association for Computational Linguistics.

Tang, Y. K., Mao, X. L., & Huang, H. (2016). Labeled phrase latent Dirichlet allocation. In *Proceedings of the International Conference on Web Information Systems Engineering* (pp. 525-536). Cham: Springer. doi: 10.1007/978-3-319-48740-3_39

Titov, I., & McDonald, R. (2008). Modeling online reviews with multi-grain topic models. In *Proceedings of the 17th international conference on World Wide Web* (pp. 111-120). New York: ACM. doi: 10.1145/1367497.1367513

Titov, I., & McDonald, R. (2008). A joint model of text and aspect ratings for sentiment summarization. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 308-316). Stroudsburg: Association for Computational Linguistics.

Wallace, B. C., Paul, M. J., Sarkar, U., Trikalinos, T. A., & Dredze, M. (2014). A large-scale quantitative analysis of latent factors and sentiment in online doctor reviews. *Journal of the American Medical Informatics Association*, 21(6), 1098-1103. doi: 10.1136/amiajnl-2014-002711

Wang, H., Lu, Y., & Zhai, C. (2010). Latent aspect rating analysis on review text data: a rating regression approach. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 783-792). New York: ACM. doi: 10.1145/1835804.1835903

Wang, H., Wu, F., Lu, W., Yang, Y., Li, X., Li, X., & Zhuang, Y. (2018). Identifying objective and subjective words via topic modeling. *IEEE transactions on neural networks and learning systems*, 29(3), 718-730. doi: 10.1109/TNNLS.2016.2626379

Xu, K., Qi, G., Huang, J., Wu, T., & Fu, X. (2018). Detecting bursts in sentiment-aware topics from social media. *Knowledge-Based Systems*, 141, 44-54. doi: 10.1016/j.knosys.2017.11.007

Zhai, Z., Liu, B., Xu, H., & Jia, P. (2011). Constrained LDA for grouping product features in opinion mining. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 448-459). Berlin: Springer. doi: 10.1007/978-3-642-20841-6_37

Zhan, T. J., & Li, C. H. (2011). Semantic dependent word pairs generative model for fine-grained product feature mining. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 460-475). Berlin: Springer.

Zhang, Y., Ji, D. H., Su, Y., & Wu, H. (2013). Joint naive bayes and lda for unsupervised sentiment analysis. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 402-413). Berlin: Springer.

Zhao, W. X., Jiang, J., Yan, H., & Li, X. (2010). Jointly modeling aspects and opinions with a MaxEnt-LDA hybrid. In *Proceedings of the Conference on Empirical Methods in Natural Language* Processing (pp. 56-65). Stroudsburg: Association for Computational Linguistics.

The article has been reviewed. | **Received:** 27 August 2018 | **Accepted:** 3 December 2018

Academic Editor: Stanislava Mildeova