Roadmap for Polarity Lexicon Learning and Resources: A Survey

Swati Sanagar and Deepa Gupta

Abstract Sentiment analysis opens door for understanding opinions conveyed in text data. Polarity lexicon acts as heart in sentiment analysis tasks. Polarity lexicon learning is explored using multiple techniques over years. This survey paper discuss polarity lexicon in two aspects. The first part is literature study which depicts from initial techniques of polarity lexicon creation to the very recent ones. The second part reveal facts about available open source polarity lexicon resources. Also, open research problems and future directions are unveiled. This informative survey is very useful for individuals entering in this arena.

Keywords: sentiment analysis; polarity lexicon; survey; transfer learning; sentiment lexicon.

1 Introduction

Today's internet era has provided easy access to huge online information. People around the world digitize their diverse opinions and experiences on internet. These text data experiences and opinions are analysed to extract valuable information using Sentiment analysis techniques. The word sentiment/opinion is individuals' view, understanding, or experience about some entity. Sentiment analysis is computational study of people's opinion and emotion about some event, topic, object, individual, or organization etc. This research primarily started since early 1990's when standard information retrieval was observed insufficient and research for more in-depth information [18] and individual's point of views

Swati Sanagar
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Amrita University, Bangalore, India. e-mail: swatisanagar@yahoo.co.in

Deepa Gupta
Department of Mathematics, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Amrita University, Bangalore, India. e-mail: g_deepa@blr.amrita.edu

J.M. Corchado Rodriguez et al. (eds.), *Intelligent Systems Technologies and Applications 2016*, Advances in Intelligent Systems and Computing 530,
DOI 10.1007/978-3-319-47952-1_52

[48] was explored. This initial research direction is explored for semantic orientation of adjectives [17] and also subjectivity [49] etc. In subsequent years, it received attention and is explored for product reputation mining [29], sentiment classification [35] etc. Sentiment analysis research has flourished over decade and it has unlocked numerous opportunities such as investor's opinion analysis for stock market [7] etc.

Consider following example from consumer review category.

- This APP is *amazing*!! *Very easy* to control and *user friendly*.
- *Waste of money*. This sweatshirt is *not well-made* at all.

The first example from Android Apps domain contain words/phrases '*amazing', 'very easy',* and *'user friendly'* which indicate positivity. The second example from Clothes domain contain words/phrases '*waste of money', 'not well-made*' which denote negativity. These are *opinion words* that convey individual's opinion and it is the only means to detect underlying sentiment in text data. These opinion words or phrases are compiled together to form lexicon. Polarity lexicon stores set of opinion words and their polarity orientation or scores. It is an important resource in sentiment analysis which avoids recreation efforts and help in faster processing. Most importantly, it is building block of most sentiment analysis applications**.** Sentiment analysis approaches are broadly categorised as machine learning/statistical, lexicon based, and hybrid. Machine learning/statistical approaches solve sentiment analysis as a regular text classification problem using machine learning algorithms**.** It makes use of syntactic and/or linguistic features such as part of speech patterns, pruning, n-grams etc. e.g. unigram, bigram features used with Naïve Bayes classifier [21] and enhanced with Support Vector Machine [39]. These features include opinion words as major contribution. Lexicon based approaches utilise and/or generate polarity lexicon which can be stored and reused for domain related sentiment analysis task. These approaches [26] are either corpus-based that use linguistic information or knowledge/dictionary based that use existing knowledge bases. Lexicon based approaches are focused on polarity lexicon learning. Hybrid approach is combination of machine learning/statistical and lexicon based approaches.

As per our knowledge this is first attempt of writing survey on polarity lexicon learning. This work is organised in five sections. Section 2 gives general polarity lexicon creation process. This survey paper discuss polarity lexicon learning in two ways, literature study in section 3 and facts about available open source polarity lexicon in section 4. The last section concludes with open research problems and future research direction.

2 General Process for Polarity Lexicon Learning

General steps of polarity lexicon learning are given in Fig. 1. The process starts from corpora selection and follows multiple steps.

- Corpora is unstructured text either labelled or unlabeled.
- Data pre-processing is used for data cleaning and may involve multiple steps such as tokenization, spell-check, removal of stop word, part-of-speech tagging, lemmatization, chunking, parsing, etc. Most of the times data pre-processing decision is taken by observing data.
- Opinion words/phrases are selected based on decision made for model. It may include n-grams, part-of-speech, syntactic patterns, collocations etc.
- Term weighting is performed for opinion words/phrases using different techniques that may include binary, frequency, or tf-idf based.
  Knowledge/Corpus based lexicon creation technique is/are applied to opinion word/phrase to assign weight based on lexical affinity technique or to assign polarity based on keyword spotting technique.

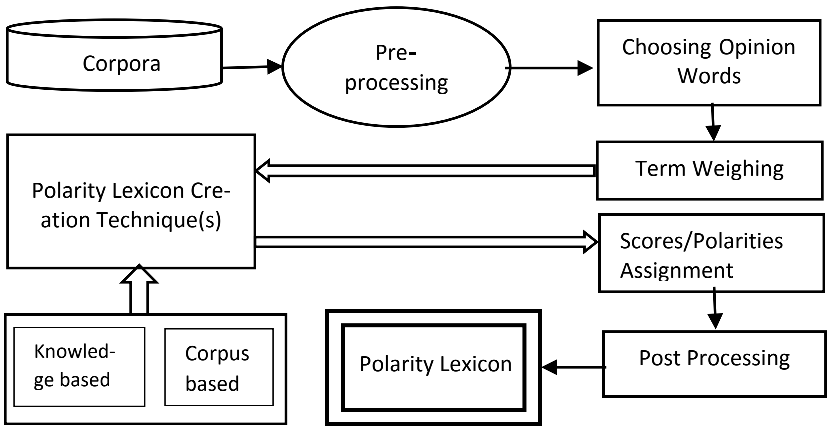

Fig. 1 General procedure for polarity lexicon learning.

– Raw polarity lexicon may be post-processed for pruning and relevancy to get final polarity lexicon.

3 General Classification of Polarity Lexicon Learning Techniques

General classification of polarity lexicon learning approaches from our viewpoint are given in Fig. 2. Polarity lexicon learning approaches include learning from existing knowledge bases, learning using linguistic information from corpus, or learning manually. Manual approach is labour intensive, time consuming, and non-scalable which is inadequate for information era. Other two approaches are classified depending on usage of labelled data and method used for polarity lexicon learning. Literature study based on this is discussed in subsequent subsections.

3.1 Learning from Knowledge base/Dictionary

Learning from knowledge bases/dictionaries approaches exploit linguistic resources such as language dictionaries, thesaurus, WordNet, and other lexicon resources. Polarity lexicon learns from single knowledge base or

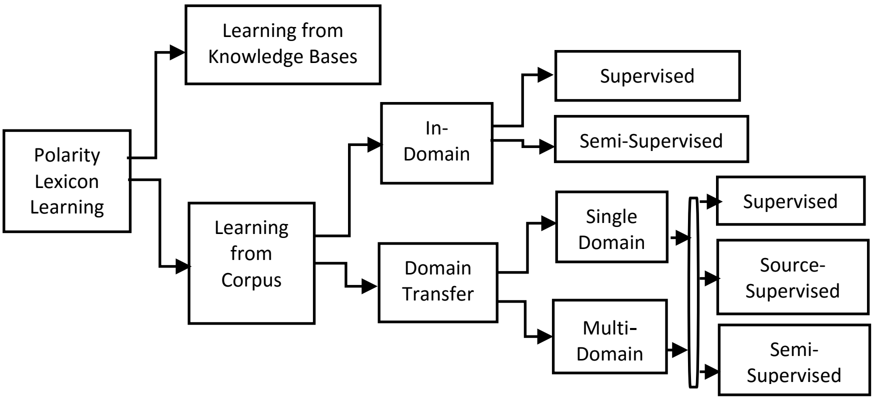

Fig. 2 Polarity lexicon learning approaches

incorporate knowledge from multiple knowledge bases. Learning from single knowledge base observed using three major approaches that includes relation based, graph based, and hybrid.

One among early papers using this approach [19] take initial small set of labelled adjectives as opinion words. These words are used to expand polarity lexicon using WordNet. It adds synonyms to same polarity and antonyms to opposite polarity set. Adjectives are strong representatives of sentiment, but verbs and nouns also express sentiments. This fact is used to extend approach by considering verbs, nouns in addition to adjectives [22] as opinion words. WordNet-Affect [44] is created with similar criteria by including hyponymy, meronymy, holonymy relations using WordNet and other resources. First the core synsets are identified and then extended using relations. Basic assumption that synonym of positive word is also positive may not hold in every case. Some cross-verification mechanism is needed. Approach of finding synonym of polarity opinion word is extended [23] using probabilistic approach to confirm polarities of synonyms. Opinion word is assigned polarity based on its maximum closeness with polarity class across synsets. Initial set of polarity words is extended using synonym, antonym, hypernymy, hyponymy search [11] and used to generate feature vectors for polarity opinion words using glosses. Word feature vectors cosine normalised tf-idf score is assigned as polarity strength to word. Orientation of opinion words is considered subjective. Approach is further progressed [12] by considering subjectivity to assign positive, negative, and objective scores to each opinion word. Also, new words are added using WordNet graph navigation. Refined SentiWordnet is extended by adding objectivity along with subjectivity [13]. Positive, negative, and objective scores are assigned to each synsets available in WordNet by combining results from eight ternary classifiers. Labelled adjectives are used [1] to add more opinion words using different WordNet relations, POS-pruning, polarity overlapping etc. in multiple steps. Score is assigned to word depending on frequency of word in polarity class in multiple runs.

Graph based approaches often learn polarity lexicon using label propagation approach. Word is considered as node and relation as edge. Initially nodes are unlabelled, later on they are labelled with polarity. Synonym relation between adjectives is used as edge [20]. WordNet distance based measure is used for assigning semantic orientation to adjectives. It is used to calculate distance between adjectives using geodesic measure and in turn to calculate score for it. Also, connected syntactic components such as noun and verb are explored. Labelled polarity words are used to extract synonym graph from WordNet [38]. It is explored using three graph approaches Mincut, Randomized Mincut, and label propagation. Markov random walk model is used to create polarity lexicon [16] and two words are connected using WordNet synonym and hypernym relation. Polarity is assigned to words based on mean hitting time.

Hybrid approach is used Roget-like thesaurus to extract synonyms and antonyms from initial labelled opinion words. It used additional affix pattern rules for antonym generation [30] to increase polarity lexicon coverage. Two approaches are combined for large corpora [36]. Initial labelled word list is expanded using synonym, antonym relation from WordNet and conjunctions. Adjacency matrix is created for adjectives and synonym relation as edge weight. A constrained symmetric nonnegative matrix factorization (CSNMF) method is applied to matrix and an iterative process is used to cut graph of adjectives into positive and negative sets where each adjectives is assigned positive and negative score. Approach uses WordNet and corpora jointly for improved performance.

Usage of multiple knowledge bases is less explored in literature. Multiple knowledge bases such as General Inquirer, SentiWordnet, Subjectivity Clues and Moby thesaurus are combined using max rule, sum rule etc. [34] and evaluated on multiple domains.

Initial development of polarity lexicon learning from knowledge bases is observed mainly for relation based approach in 2000s. Subsequently graph based and hybrid approaches are evolved. Initial approaches assigned polarity to opinion words and later approaches assigned scores.

3.2 Learning from Corpus

Learning from corpus approach utilize linguistic knowledge exploiting corpora and explore using statistical and semantic methods. It is divided into two main approaches in-domains learning and domain transfer learning. This approach learns opinion words and their orientation by unveiling facts from corpus.

3.2.1 In-Domain Learning

In-domain polarity lexicon learning uses domain corpora for training and for testing. It is supervised when labelled training data is used and semi-supervised when minimal labelled information is used.

Supervised polarity lexicon learning models provide highly accurate results. These models are mostly based on syntactic relation, semantic orientation etc. Supervised model [17] used knowledge of conjunction and morphological relation. This knowledge is combined in log linear model to check if two conjoined adjectives have same orientation using labelled adjectives. Many conjunction rules are used in addition to other rules [9] to extract orientation of word in context using initial labelled instances. Labelled opinion words are used [43] to build a phrase based polarity lexicon. Point wise Mutual Information (PMI) approach is used to calculate semantic orientation score of word using opinion word orientation. This is based on co-occurrence. Rule based approach based on dependency relation [37] used minimal labelled seed input. This double propagation approach extract opinion words and assign polarity using conjunction, negation etc. rules utilising other opinion words. Domain specific semi-supervised polarity lexicon creation framework [52] is applied to ten domains. Dependency relation and few labelled instances are used to design rules for polarity assignment.

Polarity lexicon creation approach is scaled to multiple domains [42] to improve scope. Scores are assigned using semantic orientation calculator using linguistic features such as negation, intensification etc. Multi-domain knowledge is used to construct polarity lexicon [46] to create domain dependent and domain independent polarity lexicon. Scores are assigned based on frequency proportionality in polarity context.

In-domain polarity learning is most obvious, highly accurate, and popular approach. Linguistic knowledge is explored using various perspective such as occurrence of conjunction, negation, intensification etc. and techniques such as rule based, co-occurrence, dependency relation etc.

3.2.2 Domain Transfer Learning

Domain transfer learning is based on storing knowledge while solving one problem and transferring knowledge to other domains. Transferred knowledge is learned either from single or multiple domains. Supervised transfer learning for multi-domain and single domain uses labelled data from both training i.e. source and target i.e. test domains. Source-supervised learning utilises labelled data only from source domain. Semi-supervised approach uses only minimal supervised information.

Supervised single domain knowledge transfer approach for POS feature ensemble and sample selection of target and source instances [51] are checked for closeness using principal component analysis. Based on this word scores of source domain are tuned to adapt to target domain. A source supervised structured correspondence learning model is used to identify frequently occurring features from source and target domains called pivot features [4]. These features are used to build correspondence with non-pivot features from source and target domains. Success of this method depends on choice of pivot features in both domains, based on which the algorithm learns a projection matrix that maps features from target domain into the feature space of source domain. Geodesic flow kernel adaptation algorithm used [15] source supervised approach. Domain independent opinion words available in labelled data of source domain and having similar word distribution as target domain are selected as Landmarks. Semi-supervised knowledge transfer approach for single domain is not observed.

Very few multi-domain knowledge transfer approaches are observed in literature. A supervised corpus based approach is explored by learning domain specific and domain independent lexicons [45]. An improved polarity lexicon is constructed by acquiring knowledge from multiple source domains which is transferred to multiple similar target domains. Polarity lexicon learned using multiple source supervised domains using various approaches [2] such as majority voting, weighted voting etc. is transferred to target domain. A semi-supervised multi-domain transfer learning approach [40] uses minimal seed words as labelled input and learns seed words and polarity lexicon from multiple source domains using iterative learning process. This process embeds Latent Semantic Analysis approach that assign score to opinion words. This process learn seed words applicable to group of domains belonging to a particular category. Learned seed words are used to create polarity lexicon for target domains. Also, source domain polarity lexicons are adapted to target domain.

A major advantage of domain transfer polarity lexicons is that it capture domain specific effects. Polarity lexicon customized to new domain can help avoid individual polarity lexicon construction from scratch.

Exploration of learning from knowledge bases is seen since early 2000. Learning from knowledge bases make easy and quick access to large number of sentiment words. Since most of these knowledge bases are generalized they end up creating generic polarity lexicon. **T**hey lack in domain orientations and do not provide solution to domain specific task. Corpus based approaches have gained popularity and importance, as they overcome this drawback and capture domain specific effects. Supervised corpus based approaches are highly accurate, but bottleneck for these techniques is availability of labelled data. This bring limitations when it comes to scaling up to multiple domains. Unsupervised approaches provide solution in this scenario. Although, scaling up also depends on other factors such as algorithm limitations. Domain transfer learning has capability to overcome these shortcomings. It show path towards development of polarity lexicon that preserves domain specific characteristics and still provide solution to multiple domains. This evolving futuristic approach is bringing revolution in polarity lexicon learning methods.

Table 1 Open-source polarity lexicon resource

Details → Lexicon↓	Size	Domain	Polarity	Method	Applicable	Corpora used
ANEW (1999)[5]	1,040	General	9 point score	Manual	emotion and attention study	Brown Corpus
General Inquirer(GI) (2000)[14]	11,788	General	Categorical 182 tag categories	Manual	Content Analysis, social cognition, (social science)	Harvard , Lasswell dictionary
WordNet-Affect (2004)[41]	2,874 SS[a] 4,787 WD[c]	General	+ / - /neu/ ambiguous	Semi-automatic	Multi-Category SA[b]	Based on WordNet
Hu & Liu (2004)[19]	6,800	Social Media	+/-	Semi-automatic	Social Media data analysis	Social Media
MPQA (2005)[50]	8,221	General	+ / - /neu Subjectivity	Semi-automatic	Subjectivity clues analysis	News documents
Micro-WNOp (2007)[28]	1,105 SS	General	Pos/neg (0 to1)	Manual	General Purpose SA	WordNet
SentiWordnet 1.0(2006) 3.0(2010)[3]	1,17,659	General	(0 to 1)	Semi-automatic	General Purpose SA	Based on WordNet
LabMT (2011)[10]	10,222	Mixed	Rank	Semi-automatic	Ranking, time series analysis	Twitter postings
AFINN (2011)[33]	2,477	Micro-blogging	-5 to +5	Manual	Micro-blogging SA	Twitter postings
NRC Word-Emotion(2011)[32]	14,182	General	8 emotion, (+ / -)	Manual	Multi-category SA	Macquarie Thesaurus
Warninger (2013)[47]	13,915	General	9 point scale	Manual	emotion and attention SA	Anew, Categorical norms
Sentiment140 (2013)[31]	62,468 U[d], 677,698 B[e], 480,010 P[f]	Micro-blogging	Context Frequency (+/-), score	Auto -matic	Twitter based SA	Twitter
NRC MD[g] Twitter (2014)[24]	1,515	Micro-blogging	(-1 to1), (0 to 1)	Semi-automatic	Micro-blogging SA	Sentiment140 & Hashtag sentiment
NRC Laptop (2014)[25]	26,577 U, 1,55,167 B	Consumer Product (Laptop)	Context Frequency (+/-), score	Auto -matic	Consumer product (Laptop) review SA	Amazon Consumer reviews

Details → Lexicon↓	Size	Domain	Polarity	Method	Applicable	Corpora used
SenticNet (2014)[6]	30,000 concepts	General	Affective scores -1 to 1	Auto-matic	Concept level SA	WordNet-Affect, Open Mind
SentiSense (2014)[8]	2,190 SS 5,496 WD	General	14 emotion categories	Semi-automatic	Intensity, emotion SA	WordNet 3.0
Yelp Restaurant(2014)[25]	39,274 + 276,651	Restaurant related	Frequency (+/-), score	Auto - matic	Restaurant review SA	Yelp Restaurant
Loughran McDonald Master(2015)[27]	85,131	Finance	7 category & (+,-)	Semi-automatic	Finance domain	2of12inf list

[a]Synsets, [b]Sentiment Analysis, [c]Words, [d]Unigram, [e]Bigram, [f]Pair of uni/bigram, [g]MaxDiff

4. Existing Open Source Polarity Lexicon Resources

Many polarity lexicon are created for different tasks representing different domains. Open source resources make research valuable, freely available, and receive wide acceptance from society. Countable, but widely useful open source polarity lexicons have been created over two decades. A brief informative description of major available open-source polarity lexicons for English language is given in Table 1. It describes key facts about polarity lexicon such as size, creation year, polarity measure used, application etc. In Table2 examples containing almost same opinion words are given for polarity lexicon listed in Table1. This describes variation in polarity/score assignment across the listed polarity lexicon. Every polarity lexicon have their own scales to assign polarity/score/category to words/phrases. Most of the examples are self-explanatory. Consider example from *LabMT* that stores rank from high as 1 and grows low. It store rank of word in different context such as word's rank in happiness index, word's rank in twitter frequency index etc. The word *happy* is ranked 4 indicating it as highly *happy word* compared to word *sad* which is ranked much low according to *happiness index. Warninger* lexicon is extended using earlier *Anew* polarity lexicon by size and adding emotional norm dif-

Table 2 *Examples from open-source polarity lexicon*

Polarity Lexicon	Example and Description
ANEW	Word — Valence(M, SD)[h] — Arousal(M, SD) — Dominance(M, SD) happy — 8.25, 1.39 — 7.00, 2.73 — 6.73, 2.28 sad — 1.61, 0.95 — 4.13, 2.38 — 3.45, 2.18
General Inquirer (GI)	Happy: H4Lvd[i], positive, pleasure, emotion, related adjective: Joyous, pleased Sad: H4Lvd, negative passive, pain, emotion, related adjective: unhappy
WordNet-Affect	Affective category labels are emotion, behaviour, attitude, cognitive state Word — Tag — Senses joy — positive — joy, elated, gladden, gleefully sad — negative — sadness, unhappy, sadden, deplorably
Hu & Liu	happy: positive; sad: negative
MPQA	happy: strong subjective, positive; sad: strong subjective, negative
Micro-WNOp	good: adjective, positive (1), negative(0) by 3 annotators; ugly: adjective, positive(0,0), negative(0.75, 1)
SentiWordnet	Word — Positive — Objective — Negative — POS happy — 0.875 — 0.125 — 0 — Adjective sad — 0.125 — 0.125 — 0.75 — Adjective
LabMT	Word — Happiness — Twitter — GBks[j] — NYT[k] — Music Lyrics happy — 4 — 65 — 1372 — 1313 — 375 sad — 10091 — 306 — 3579 — 3441 — 526
AFINN	happy: positive, score(3); sad: negative, score(-2)
NRC Wd Emotion	happy: anticipation, joy, positive, trust; sad: not available
Warninger	Word — Valence — Word — Arousal — Word — Dominance happiness — 8.48 — calm — 1.67 — uncontrollable — 2.18
Sentiment140	Word — PMI[l] score — Count in‘ +’context — Count in‘- ’ context happy — 1.196 — 19174 — 6087 sad — -2.735 — 1442 — 23342
NRC Laptop	Word/Phrase — PMI score — Count in‘ +’context — Count in‘- ’ context happy — 1.121 — 3,646 — 268 sad — -1.342 — 74 — 64 best_laptop — 3.462 — 287 — 2

Polarity Lexicon	Example and Description
NRC MD Twitter	happy : 0.953(0 to 1), 0.734(-1 to 1); sad : 0.219(0 to 1), -0.562(-1 to 1)
SenticNet	related concepts for celebrate_special_occasion are celebrate (holiday, occasion, birthday, wedding, express appreciation) Concept pleasantness attention sensitivity aptitude polarity happy 0.894 0 0 0 0.298 sad - 0.919 0 0 0 -0.306 celebrate-special_occasion 0.93 0.724 0.0 0.0 0.551
SentiSense	depression: Sadness category; exultation: Joy; adorable: Love
Yelp Restaurant	Word PMI score Count in‘ +’context Count in‘- ’ context happy 0.825 15185 2118 sad -0.977 51 43
LoughranMc-Doald Master	Happy : Positive added in 2009; closed : Negative added in 2009

[h](Mean, Standard deviation), [i]Harvard-4 & Lasswell dictionary, [j]Google Books, [k]New York Times, [l]Point-wise Mutual Information

ferentiation by gender, age etc. This comprises most of the available open source lexicon which we observed from our exhaustive search.

From available polarity lexicon, we observed various different facts and how open source polarity lexicon changed over time. Overall, the first decade from 1999 to 2010 has displayed initial and important polarity lexicon creation efforts, mostly carried out manually and semi-automatically. Trend of touching different areas of study is observed. First decade denotes focus on many different areas, but mainly based on general context. Current decade since 2011 represent information era where development of polarity lexicon is around social media contents. Moreover, usability of these polarity lexicon is not limited to study, but are vastly used for research, commercial, and social purposes. Many useful polarity lexicons are small in size and are having potential to expand. Many available polarity lexicon are created manually which brings limitation on extending it to other domains. Automated polarity lexicon creation need more exploration. Available polarity lexicons have applications in different areas such as psychology or health analysis, subjectivity analysis, emotion analysis, ranking based, affect based, mood based learning, content analysis and for general context analysis etc.

5 Open Research Problems and Future Directions

In todays' information era huge unlabelled information is available compred to minimal labelled information. Sentiment analysis research has grown lateral to many other research areas. But the vertical research growth is restricted due to lack in fundamental research which includes polarity lexicon learning. Research in polarity lexicon learning for sentiment analysis has not matured up to the mark. Considering these facts following research direction are also open research problems in learning polarity lexicon for sentiment analysis.

- Semi-supervised and unsupervised seem to be key approaches. Using them to build polarity lexicon scalable to multiple domains with minimal efforts.
- Learning some mechanism to build domain specific and generic lexicon applicable to many similar, dissimilar, and unseen domains.
- Reducing rebuilding efforts of polarity lexicon learning using some intermediate resources such as seed words etc. with key focus on utilizing learned knowledge using some technique such as transfer learning.
- Extension of available polarity lexicon in terms of size, scope, and other aspects to improve usability.
- Collaboration of available polarity lexicon in terms of usability, score/polarity etc.

References

1. Andreevskaia A, Bergler S (2006). Mining WordNet for fuzzy sentiment: Sentiment tag extraction from WordNet glosses. In: EACL'06: Proceedings of the European Chapter of the Association for Computational Linguistics, vol. 6, pp. 209-16.
2. Anthony A, Gamon M (2005). Customizing sentiment classifiers to new domains: A case study. In: Proceedings of international conference on recent advances in natural language processing, vol. 1(3.1), Bulgaria, pp. 2-1.

3. Baccianella S, Esuli A, Sebastiani F (2010). SentiWordNet 3.0: An enhanced lexical resource for Sentiment Analysis and Opinion Mining. In: Proceedings of the seventh conference on international Language Resources and Evaluation, vol. 10, pp. 2200-4.
4. Blitzer J, Dredze M, Pereira F (2007) Biographies, Bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification In: Proceedings of 45th annual meeting of the Association of Computational Linguistics, Prague, Czech Republic, June 2007, pp. 440–47.
5. Bradley MM, Lang PJ (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical Report C-1, Center for Research in Psychophysiology, University of Florida.
6. Cambria E, Olsher D, Rajagopal D (2014). SenticNet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis. In: Twenty-eighth AAAI conference on artificial intelligence, Quebec City 2014, pp. 1515-21.
7. Das S, Chen M (2001). Yahoo! for Amazon: Extracting market sentiment from stock message boards. In: Proceedings of the Asia pacific finance association annual conference, vol. 35, p. 43.
8. deAlbornoz JC, Plaza L, Gervas P (2012). SentiSense: An easily scalable concept-based affective lexicon for sentiment analysis. In: LREC, pp. 3562-3567.
9. Ding, X, Bing L, Yu PS (2008) A holistic lexicon-based approach to opinion mining. In: Proceedings of the conference on Web search and Web Data Mining, pp. 231-40.
10. Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM (2011). Temporal Patterns of Happiness and Information in global social network: Hedonometrics and Twitter. *PLoS one,* 6(12):e26752. doi:10.1371/journal.pone.0026752
11. Esuli A, Sebastiani F (2005). Determining the semantic orientation of terms through gloss classification. In: Proceedings of the 14th ACM international conference on Information and knowledge management, pp. 617-24.
12. Esuli A, Sebastiani F (2006a). Determining term subjectivity and term orientation for opinion mining. In: Proceedings of conference of the European chapter of the Association for Computational Linguistics.
13. Esuli A, Sebastiani F (2006b). SentiWordNet: A publicly available lexical resource for opinion mining. In: Proceedings of Language Resources and Evaluation, vol. 6, pp. 417-22.
14. General Inquirer http://www.wjh.harvard.edu/~inquirer/ Accessed on 2016 Jan 20.
15. Gong B, Grauman K, Sha F (2013). Connecting the dots with landmarks: discriminatively learning domain-invariant features for unsupervised domain adaptation. In: ICML’13: Proceedings of the 30th international conference on Machine Learning (ICML), Atlanta, pp. 222-30.
16. Hassan A, Radev D (2010). Identifying text polarity using random walks. In: Proceedings of annual meeting of the Association for Computational Linguistics.
17. Hatzivassiloglou V, McKeown K (1997). Predicting the semantic orientation of adjectives. In: Proceedings of 8th conference of Association for Computational Linguistics, pp. 174-81.
18. Hearst MA (1992). Direction-based text interpretation as an information access refinement. Text-based intelligent systems: Current research and practice in information extraction and retrieval, 1:257-74.
19. 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, ACM, pp. 168-77.
20. Kamps J, Marx MJ, Mokken RJ, Rijke MD (2004). Using Wordnet to measure semantic orientations of adjectives. In: Proceedings.of LREC’2004, pp. 1115-18.

21. Kang H, Yoo SJ, Han D (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews, Expert Syst. Appl., 39(5):6000-10.
22. Kim SM, Hovy E (2004). Determining the sentiment of opinions. In: Proceedings of the 20th international conference on Computational Linguistics. Association for Computational Linguistics, pp. 1367.
23. Kim SM, Hovy E (2006). Identifying and analyzing judgment opinions. In: Proceedings of Human Language Technology Conference of the North American Chapter of the ACL, pp. 200-7.
24. Kiritchenko, S., Zhu, X., Mohammad, S. (2014). Sentiment Analysis of Short Informal Texts. J. Artificial Intelligence Res. 50:723-62.
25. Kiritchenko S, Zhu X, Cherry C, Mohammad S (2014). Detecting Aspects and Sentiment in Customer Reviews. In: Proceedings of the 8th international workshop on Semantic Evaluation Exercises, SemEval-2014, Dublin, Ireland, pp. 437-42.
26. Liu B (2012). Sentiment analysis and opinion mining, Morgan & Claypool Publishers.
27. Loughran T, McDonald B (2011). When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks. J. of Finance, 66:1, 35-65.
28. Micro-WnOp, http://www-3.unipv.it/wnop/#Cerini07 Accessed on 2016 May 15.
29. Morinaga S, Yamanishi K, Tateishi K, Fukushima T (2002). Mining product reputations on the web. In: Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining, ACM, pp. 341-49.
30. Mohammad S, Dunne C, Dorr B (2009). Generating high-coverage semantic orientation lexicons from overtly marked words and a thesaurus. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, vol. 2, pp. 599-608.
31. Mohammad SM, Kiritchenko S, Zhu X (2013). NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. In: Proceedings of seventh international workshop on Semantic evaluation exercises (SemEval-2013), Atlanta, Georgia, USA.
32. Mohammad SM, Turney PD (2010). Emotions evoked by common words and phrases: using Mechanical Turk to create an emotion lexicon. In: Proceedings of the NAACL-HLT'10 workshop on computational approaches to analysis and generation of emotion in text, California, ACL, pp. 26-34.
33. Nielsen FÅ (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:1103.2903. http://arxiv.org/abs/1103.2903 Accessed on 2016 Jan 20.
34. Ohana B, Tierney B, Delany S (2011). Domain independent sentiment classification with many lexicons. In: WAINA'11: Advanced information networking and applications. IEEE workshops of international conference, pp. 632-37.
35. Pang B, Lee L, Vaithyanathan S (2002). Thumbs up? sentiment classification using machine learning techniques. In: Proceedings of conference on empirical methods in natural language processing, vol. 10, pp. 79-86.
36. Peng W, Park DH (2011). Generate adjective sentiment dictionary for social media sentiment analysis using constrained nonnegative matrix factorization. In: Proceedings of fifth international AAAI conference on weblogs and social media, pp. 273-80.
37. Qiu G, Liu B, Bu J, Chen C (2009). Expanding domain sentiment lexicon through double propagation. In: Proceedings of international joint conference on Artificial Intelligence, vol. 9, pp. 1199-04.
38. Rao D, Ravichandran D (2009). Semi-supervised polarity lexicon induction. In: Proceedings of the 12th conference of the European chapter of the ACL, pp. 675-82.
39. Rui H, Liu Y, Whinston A (2013). Whose and what chatter matters? The effect of tweets on movie sales. Decision Support Syst. 55(4):863-70.

40. Sanagar S, Gupta D (2015). Adaptation of multi-domain corpus learned seeds and polarity lexicon for sentiment analysis. In: Proceedings of international conference on computing and network communications, pp.50-58.
41. Strapparava C, Valitutti A (2004). WordNet-Affect: an affective extension of WordNet. In: LREC'04: Proceedings of 4th international conference on Language Resources and Evaluation, Lisbon, pp. 1083-86.
42. Taboada M, Brooke J, Tofiloski M, Voll K, Stede M (2011). Lexicon-based methods for sentiment analysis. Computational linguistics 37(2):267-307.
43. Turney PD (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: Proceedings of 40th meeting of the Association for Computational Linguistics, pp. 417–24.
44. Valitutti A, Strapparava C, Stock O (2004). Developing affective lexical resources. Psychology J. 2(1): 61-83.
45. Venugopalan M, Gupta D (2015). An enhanced polarity lexicon by learning-based method using related domain knowledge. Int. J. Information Processing & Management 6(2):61.
46. Vishnu KS, Apoorva T, Gupta D (2014). Learning domain-specific and domain independent opinion oriented lexicons using multiple domain knowledge In: Proceedings of 7th IEEE international conference on Contemporary Computing, pp. 318-23.
47. Warriner AB, Kuperman V, Brysbaert M (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. Behavior research methods, 5(4), pp. 1191-207.
48. Wiebe JM (1990). Identifying subjective characters in narrative. In: Proceedings of international conference on computational linguistics, vol. 2, pp. 401-6.
49. Wiebe J (2000). Learning subjective adjectives from corpora. In: Proceedings of national conference on artificial intelligence, pp. 735-40.
50. Wilson T, Wiebe J, Hoffmann P (2005). Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In: Proceedings of the conference on human language technology and empirical methods in Natural Language Processing, ACL, pp. 347-54.
51. Xia R, Zong C, Hu X, Cambria E (2013). Feature ensemble plus sample selection: domain adaptation for sentiment classification. Intel. Syst. 28(3):10-8.
52. Zhang Z, Singh PM (2014). Renew: A semi-supervised framework for generating domain specific lexicons and sentiment analysis In: Proceedings of the Association for Computational Linguistics, pp. 542–51.