

A HYBRID DEEP NEURAL NETWORK FOR ASPECT BASED SENTIMENT ANALYSIS ON RAJYA SABHA QUESTIONS

Shreyas R Hegde

Senior Software Engineer, Electronics for Imaging
Bengaluru, Karnataka, India 560066
Shreyas.hegde91@gmail.com

Yogesh R Gaikwad

Head of Engineering, Prudential
Singapore, 828781
Yogesh.r.gaikwad@gmail.com

Abstract

In the modern era, Technology and Politics are strongly connected and have become inseparable. This paper proposes a hybrid approach by using unsupervised and supervised techniques for analyzing the innumerable opinions expressed in the Rajya Sabha Question Hour to extract aspects, sentiments and perform Aspect Based Sentiment Analysis (ABSA). An unsupervised Attention-Based Aspect Extraction-Long Short-Term Memory (ABAE-LSTM) network is used to identify the ‘N’ cluster of aspects present in the corpus and categorize the aspect terms present in the questions into one of these aspect categories. Supervised Neural Networks, Target Dependent-LSTM (TD-LSTM), Target Control-LSTM (TC-LSTM) and Attention Term Aspect Extraction-LSTM (ATAE-LSTM) are employed to perform ABSA. Experiments conducted on the unsupervised ABAE-LSTM show high coherence scores between the aspect terms, and this also creates a gold standard training data set with aspect and sentiment labels in the domain. Results from supervised techniques then display promising accuracy for all three models.

Keywords: Aspect Based Sentiment Analysis; Attention-Based Aspect Extraction; Long Short-Term Memory; Target Dependent; Target Control; Attention Term Aspect Extraction.

1. Introduction

1.1. Overview

The Rajya Sabha, also known as the ‘Council of States’ is the upper house of Parliament of India. The ‘Question Hour’ has paramount importance in the parliamentary proceedings because the members can obtain information through questions on everyday life matters for which Ministries collectively answer to the legislature. This paper mainly studies two categories: Starred and Unstarred. Starred Questions are the ones to which answers are given orally on the floor of the House. Unstarred questions are ones to which written answers will be provided on the table of the House at the end of the hour.

NLP is a powerful tool that can be used in the political domain to unbridle the potential of enormous unstructured data. The field of NLP began in the 1940s, during the second world war when people started realizing the importance of a machine-based translation of language from one to another [NLP - overview]. The recent advances in deep learning-based NLP methods, mimic neuron functionalities of the human brain to provide better performance and accurate results. From a political campaign to social good, AI is proliferating the political process faster than we imagine [How AI will decide your future Prime Minister].

In recent years, there are lots of researches in the areas of sentiment analysis and aspect extraction in the Indian political domain ranging from analyzing parliamentary speeches to analyzing political tweets. However, much work has not been done on analyzing aspects and sentiments from the Rajya Sabha Question Hour. This paper uses the dataset from [Q & A Discussed in Parliament of India | Kaggle] which contains ~88000+ Question-Answers of the Rajya Sabha Question hour from 2009-2017 categorized by question data, answer date, ministry, etc. This paper employs a hybrid of unsupervised and supervised learning techniques, mainly because there is no gold standard data set with training labels available for the domain of this study and the observations in the dataset used for this work are unlabeled.

Text is arguably the most important artifact of political analysis, and even today political science students read a lot of historical political documents [Monroe and Schrod, 2008]. With increasing political texts being digitalized, it is important to view texts as data rather than texts as texts. This will not only enable better analysis and results; it will significantly help in avoiding preserving tons of unstructured data. Keyword extraction approaches [Rose et al., 2010; Timonen et al., 2012; Khurana et al., 2017] explains the important topics of the question, but they do not indicate the broader aspect the question is talking about. E.g. let us consider the question *'whether Government proposes to place doctors on board in Duronto Express?'*, keyword extraction techniques identify the keywords as *'doctors; express'* but furthermore a rule-based technique would be needed to categorize these keywords. Unsupervised ABAE proposed in the paper classifies the aspect terms as *'Duronto; express'* and categorizes these aspect terms into the aspect category *'Railways'*. Similarly, let us consider another question *'the number of roads constructed till 31 December 2009 under Pradhan Mantri Gram Sadak Yojana (PMGSY)?'*, keyword extraction technique identifies the keyword *'roads'* whereas the unsupervised ABAE proposed categorizes the aspect terms identified *'Yojana; PMGSY'* into aspect category *'Schemes'*. Sentiment labels can be manually annotated in the unsupervised phase, but in this paper, we use VADER (Valence Aware Dictionary and Sentiment Analyzer) [Hutto and Gilbert, 2014] to determine sentiment labels and avoid the manual task of annotating sentiments. Supervised neural networks TC-LSTM, TD-LSTM, ATAE-LSTM then uses the training set created from the unsupervised phase of the study to perform ABSA.

1.2 Significance of the work

In political science, NLP methods have been extensively used for many analyses types and tasks, including inferring the policy position of actors from textual evidence, detecting topics in political texts, and analyzing stylistic aspects of political texts [Glavas et al., 2019].

Some of the major significance of this work is:

- The unsupervised ABAE neural network model proposed in the paper creates a gold standard train dataset in the Indian parliamentary domain for aspect extraction tasks, which can be used by other researchers.
- Political Science students and scientists working on the Rajya Sabha analysis can benefit from this study as they will be able to quickly gain insights on the political aspects being discussed.
- The study can not only aid policymakers and political science students, but even the general public can visualize the various aspects being discussed and how the importance of aspects has changed over the years.
- Empirically understand the sentiment of the Question relating to an aspect and identify the focus shift over the years concerning these aspects.

The rest of the paper is organized as follows: Section 2 describes related works for aspect extraction and sentiment analysis tasks; section 3 describes the hybrid neural network model, unsupervised ABAE-LSTM network followed by supervised TD-LSTM, TC-LSTM and ATAE-LSTM network models; experiments and results of the proposed model are discussed in section 4; finally, in section 5 conclusion is provided with future scope.

2. Related Work

AI in politics promises to de-facto reinstates a more direct and representative democracy by strengthening the bond between government officials and the people [A.I. Will Mark A Turning Point in the History of Politics | by Sukhayl Niyazov | Towards Data Science]. AI in politics will enable decision making based on data and less political. The Fundamental characteristic of every democracy is to learn from mistakes, unlearn, relearn, and adapt to changing times, essentially mapping to the reinforcement learning model of Machine Learning.

Over recent years, NLP with Deep Learning has produced excellent results in a variety of domains. [Zhang et al., 2018] provides an overview of the deep learning techniques. Deep Learning can exploit a greater learning power of artificial neural networks using networks of multiple layers. Neural Networks mimic the structure of the biological brain and consists of an enormous amount of knowledge processing units arranged systematically in multiple layers, working simultaneously. Connection weights are adjusted between neurons to perform learning tasks. Deep learning uses a torrent of non-linear processing units from multiple layers to extract and transform features. Lower layers generally closer to the input stream and learn from simple features whereas higher layers assimilate more complex information derived from lower layers. Most of the deep learning models use word embeddings as inputs. Word embedding is a representation where words are transformed to vectors of real numbers and involve embedding from a higher dimensional sparse vector space to a lower dimension dense vector space.

Aspect Based Sentiment Analysis is a subcategory of sentiment analysis, which considers the aspect words (or target words) in a sentence to identify the aspect associated with that sentence. Aspect level sentiment analysis is considered to be a challenging task because modeling a semantic relationship of an aspect with its surrounding words is always difficult [Shen et al., 2020]. Therefore, it is important that the neural networks model built captures the relationship between aspect words and context words. Aspect-based sentiment classification on Twitter is studied by [Vo and Zhang, 2015] using additional rich automatic features obtained by using

unsupervised learning. The study showed that multiple pooling functions, sentiment lexicons, and multiple embeddings can extract rich information and help improve performance. [Tang et al., 2016] proposes that LSTM's can capture syntactic relations between context and target words in an easier way using Target-dependent LSTM and Target-connection LSTM by considering Target's to extend LSTM. A study by [Ruder et al., 2016] proposes a model to leverage Intra and Inter sentence relations to use a hierarchical and bidirectional LSTM for aspect-based sentiment classification. The structure of the bi-directional review layer of LSTM which outputs a probability distribution over sentiments is fed with concatenated inputs from the forward and backward embedding layer of LSTM. An effective way to enforce neural models attend to the related part of the sentence was proposed by [Wang et al., 2016], attention-based LSTM with aspect embeddings.

Aspect-based Sentiment Analysis has colossal scope in a variety of fields, especially in the political domain it can be stupendous. The study on transcribed speeches from the 2016 US elections [Gold et al., 2018] shows the manual annotation of aspects is reliable in an unmarked schema. The corpus used in this study consists of speeches taken from the presidential debates of 2016. Marked annotation schema with three layers, Entity, Aspect, and Sentiment layer is used on eight pre-defined categories. The annotation schema used in this study is limited to nouns and adjectives. There is a lot of political sentiment analysis carried out in different countries across a variety of language corpus [Elghazaly et al., 2016; Rauh, 2018; Sidarenka, 2019]. Unsupervised Lexicon and Rule-based, Naive Bayes, SVM, and dictionary-based approaches have been evaluated to classify the polarity of Political tweets during Delhi 2014 elections and general state elections in 2016 [Singhal et al., 2015; Sharma and Moh, 2016]. The study on Indian Parliamentary (Lok Sabha) debates by [Rohit and Singh, 2018] analyses the stance of speeches into four categories of Issue, Blame, Appreciate, Call for Action. The study uses a manually annotated dataset into one of the above stances for its analysis. The study uses a key phrase extracting technique TextRank for finding keywords in the corpus and VADER for sentiment analysis. A recent study [Ansari et al., 2020] on political sentiment orientation analysis on tweets about the general election 2019 uses tf-idf for feature extraction and further employs Logistic Regression, Random Forest Classifier, LSTM, to extract the most common words floating around.

Aspect Extraction and Aspect based Sentiment Analysis have been one of the most sought-after areas in NLP and AI. ABSA models are easily scalable, provide real-time analysis, provides a deeper understanding of the subject and consistent results. There is a tremendous amount of ongoing research on various deep learning models for ABSA. However, the Indian political domain is an area where Aspect Extraction and Sentiment Analysis have not been explored to their full extent. From the studies, it is evident that there is a strong need for the creation of a gold standard training dataset on the Indian parliamentary domain which can be used by further researchers. Most of the existing studies on the Indian parliamentary speeches or debates rely on a human-annotated dataset for aspect and sentiment analysis. Taking this into account, this study strongly focuses on unsupervised deep learning-based techniques to extract aspects and sentiments for the Rajya Sabha corpus. This dataset can also be a baseline dataset for studies across Indian parliamentary data analysis. The comprehensive study described in the paper follows a hybrid model, where the aspect terms and sentiment labels extracted from the unsupervised model then input into supervised algorithms for aspect-based sentiment analysis.

3. Methodology

The structured data used for the work [Q & A Discussed in Parliament of India | Kaggle], contains ~88000+ Question-Answers of the Rajya Sabha Question hour from 2009-2017 categorized by question data, answer date, and ministry, etc. The study of sentiment analysis and aspect extraction is one of the under-researched topics in the Indian Parliamentary dataset. As mentioned in the previous sections too, the creation of a gold standard training dataset with aspect and sentiment labels in this research area is the need of the hour. The study is based on qualitative data (expressed in words) on the Rajya Sabha Question Hour.

3.1. Dataset Preprocessing and Word embeddings

The 'question_description' column in the dataset is structured in the format: "(a) whether India and Russia have agreed and signed trade targeted to reach at \$20 billion by 2015; (b) whether both sides call for diversification of trade into high technology areas; (c) whether India is one of the very few countries with whom Russian trade is growing rather than declining; and (d) if so, the details of decisions taken, strategy adopted to improve trade between two countries?". From the example question, it can be noted that the intent and aspect can be understood by the '(a)' subpart of the question. To improve the attention focus on a lesser number of words from the question set, '(a)' part of the question is extracted as a new column, which can be treated as the 'primary question' and the subsequent parts '(b)', '(c)', '(d)' as follow up questions. In the corpus certain words in the question always tend to occur in pairs. It is important to consider these words together during analysis. The n-gram collocations are modeled using distributed representations of words, phrases, and Pointwise Mutual Information in Colocation Extraction [Bouma, 2009; Demeester et al., 2016].

3.2. Unsupervised ABAE Neural Network Model

A corpus consisting of (word id, frequency) is created using the list of lemmatized data words. Further, the tokenized list of words from this corpus is then mapped into a dictionary to build a bag-of-words representation. Given the input words of a sentence, the following steps are performed:

- Construct Sentence Embedding using attention mechanism by filtering away non-aspect words.
- From a linear combination of Aspect embeddings, reconstruct sentence embeddings.

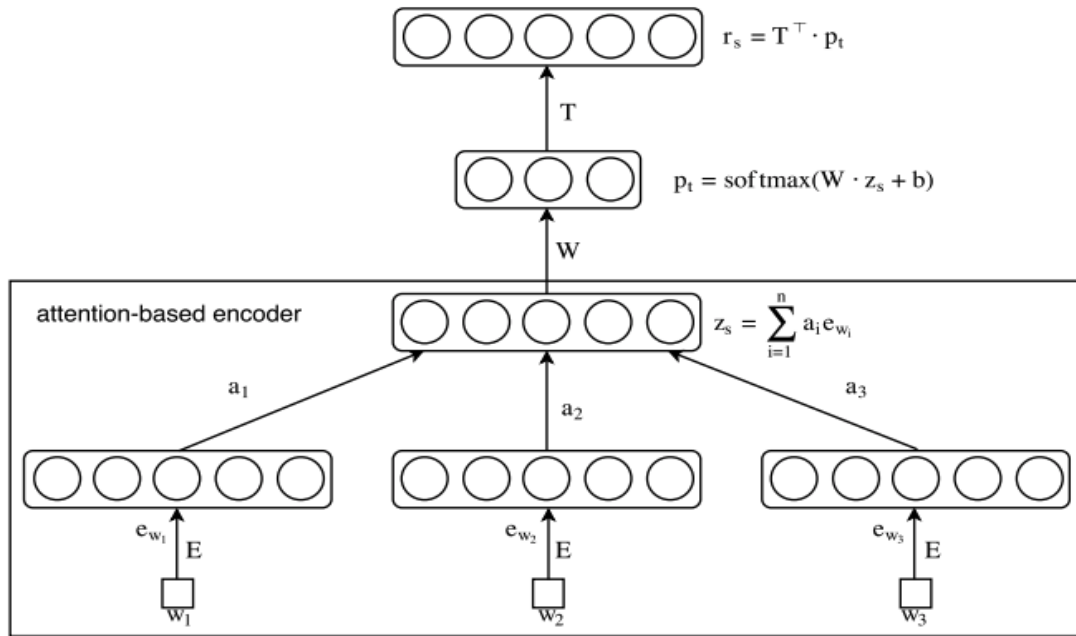


Fig. 1. An illustration of Attention Based Aspect Extraction [He et al., 2017]

3.2.1 Constructing Sentence Embedding

Sentence embedding is defined as a weighted sum of word embeddings. For each word in the input sentence, a positive weight is generated which indicated whether the word being processed is the right word to pay attention or not. The positive weight is computed by the context of the sentence and the embedding of the word.

$$e_w = \text{word_emb}(\text{sentence}_{\text{input}}) \quad (1)$$

$$y_s = \text{Average}()(e_w) \quad (2)$$

$$\text{att}_{\text{weights}} = \text{Attention}(\text{name} = ' \text{att}_{\text{weights}} ')([e_w, \text{att}_{\text{weights}}]) \quad (3)$$

$$z_s = \text{WeightedSum}()([e_w, \text{att}_{\text{weights}}]) \quad (4)$$

In Eq. (1), ' e_w ' represents the embedding vector of a word in the sentence, ' y_s ' in Eq. (2) is the average of word in a sentence and ' z_s ' in Eq.(4) is the sentence embedding defined as weighted sum of word embeddings.

3.2.2 Reconstructing Sentence with Aspect Embedding

The reconstruction process of Sentence contains two steps of transition, similar to an auto-encoder. A dense layer is applied to the sentence embedding with the aspect size. Further, a softmax non-linearity layer is applied to yield normalized non-negative weights, where the weighted matrix and bias are learned as a part of the training.

$$p_t = \text{Dense}(\text{aspect}_{\text{size}})(z_s) \quad (5)$$

$$p_t = \text{Activation}(' \text{softmax} ', \text{name} = ' p_t ')(p_t) \quad (6)$$

$$r_s = \text{WeightedAspectEmb}(\text{aspect}_{\text{size}}, \text{embedding}_{\text{dimension}}, \text{name} = ' \text{aspectEmbedding} ') \quad (7)$$

The encoded and decoded layer ' p_t ' and ' r_s ' are defined in Eq. (5), Eq. (6) and Eq. (7) respectively.

3.2.3 Loss Function

The objective of training the unsupervised attention-based model is to reduce errors in reconstructing sentences. A max-margin loss similar to [He et al., 2017] is employed in this paper. Randomly ' n ' sentences are sampled as negative and representation of it is computed by average of its word embeddings. This is to ensure the

reconstructed sentences are like the target sentence. An orthogonal regularization term is added to avoid redundancy of the aspect embedding matrix during training.

$$BatchesPerEpoch = len(TrainingSample) + aspect_{size} \quad (8)$$

$$Total_{Loss} = Total_{Loss} + BatchesPerEpoch \quad (9)$$

$$MaxMargin_{Loss} = MaxMargin_{Loss} + (BatchMaxMargin_{Loss} + BatchesPerEpoch) \quad (10)$$

Total loss during training, 'Eq. (9)' is calculated as batch loss divided over batches per epoch 'Eq. (8)'. Since the objective is to make reconstructed sentence embedding similar to target sentence embedding, a Max Margin loss (hinge loss) 'Eq. 10' is adopted.

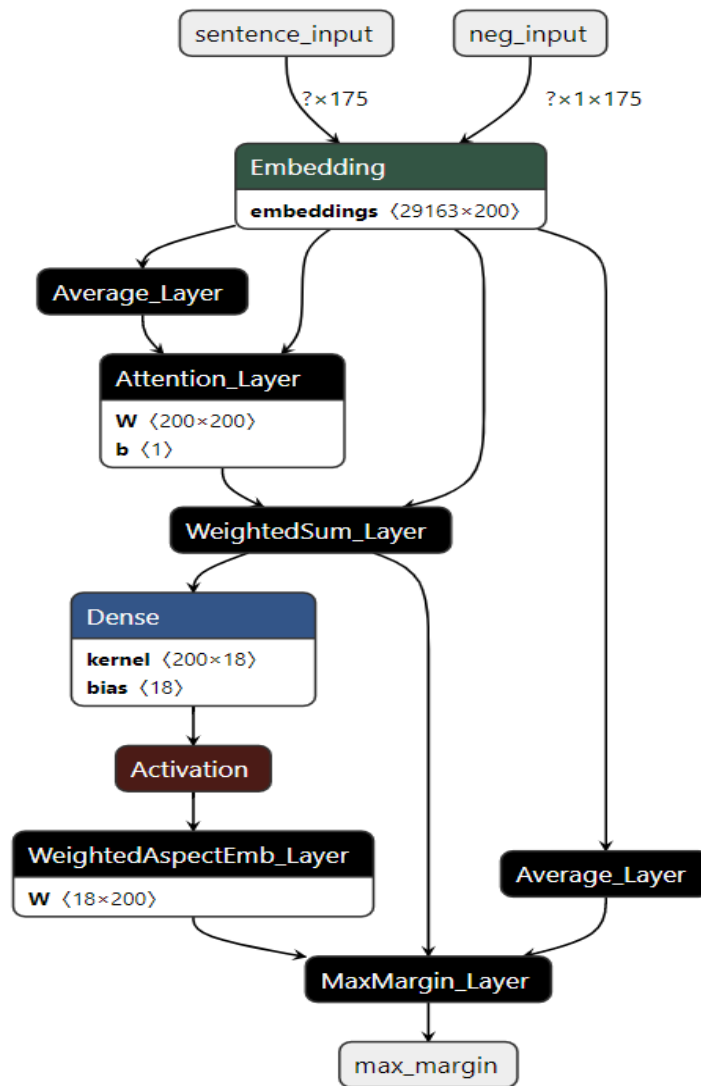


Fig. 2. Summary of inputs, weights, bias, and layers of the Unsupervised ABAE Model.

3.2.4 Creating Sentiment Labels

To analyze the polarity of the question and answer this study uses a parsimonious rule-based model VADER to assess sentiment [Hutto and Gilbert, 2014]. VADER has proven to achieve high accuracy in sentiment-based tasks. VADER provides a normalized, weighted score also known as a compound score where valence scores of each word are summed up and normalized between -1 (extremely negative) and +1 (extremely positive). Since the language in political communication generally carries a key essence of diplomacy, in this study a low valence compound score of -0.2 and +0.2 is categorized as neutral, greater +0.2 as positive, and lesser than -0.2 is considered as negative.

E.g., the question “whether his Ministry is unwilling to let minority schools off the hook when it comes to setting aside 25 percent of seats for disadvantaged children?” which has a compound score of -0.4019 is considered as negative. Similarly, a question “whether it is a fact that Indian pharmaceutical industry is estimated

to have an annual average growth of, 12.3 percent in the coming years?” having a compound score 0.3818 is considered positive.

3.3. Supervised ABSA

The aspect category labels created using the unsupervised ABAE and sentiment labels which are created using VADER [Hutto and Gilbert, 2014] are used as input seeds for performing supervised ABSA.

3.3.1 Data Preprocessing and Word Embeddings for supervised ABSA

A word corpus and character corpus are created from the words and characters from all the questions from the dataset. Similarly, aspect corpus and aspect character corpus are created from the aspect words. Word and corpus vocabularies are then built for each of these corpora. In the preprocessing stage, the embedding matrix for words is created using both word2vec [Mikolov et al., 2013] and GloVe vectors. GloVe i.e. Global Vectors for Word Representation is a global log bilinear regression model for unsupervised learning of word representations which is proven to outperform other models on word analogy [GloVe: Global Vectors for Word Representation]. Based on the aspect word position, separate vocabularies are created for words that appear on the left and right sides of the aspect word.

3.3.2 TD-LSTM and TC-LSTM

This work uses target-dependent LSTM for performing the target-dependent sentiment analysis. Target dependent LSTM's are known to capture the connection between context and target words when generating representations for the sentences. Two LSTM networks are used, a left LSTM which models the words appearing to the left of the target word, and a right LSTM which models the words appearing on the right of the target word. Left LSTM is input with the Target label along with context words preceding it. The input to the right LSM is succeeding context words along with the Target string. The left LSTM is run from left to right, and the Right LSM is run from Right to left. This is because the TD-LSTM uses the target string as the last unit could better utilize the semantics of the target word. TC-LSTM (Target-Connection LSTM) incorporates a target connection component to extend TD-LSTM concatenating word and aspect embeddings. The sentence vectors from the LSTM layers are then inputted to a dense layer with a rectified linear activation function to induce nonlinearity into the model. A final fully connected dense layer with a softmax activation function is added to represent the probability distribution of sentiment labels.

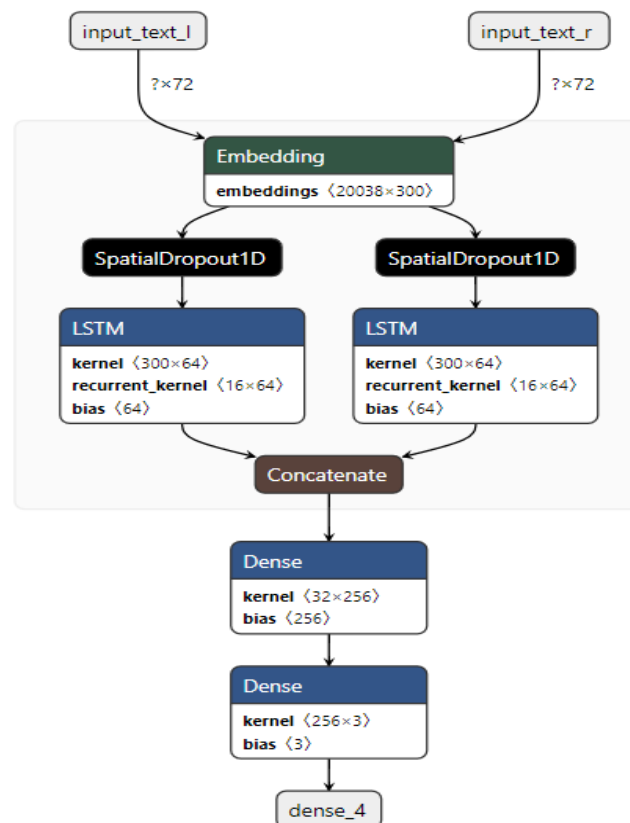


Fig. 3. Model Summary with weight, bias, and layers of TD-LSTM

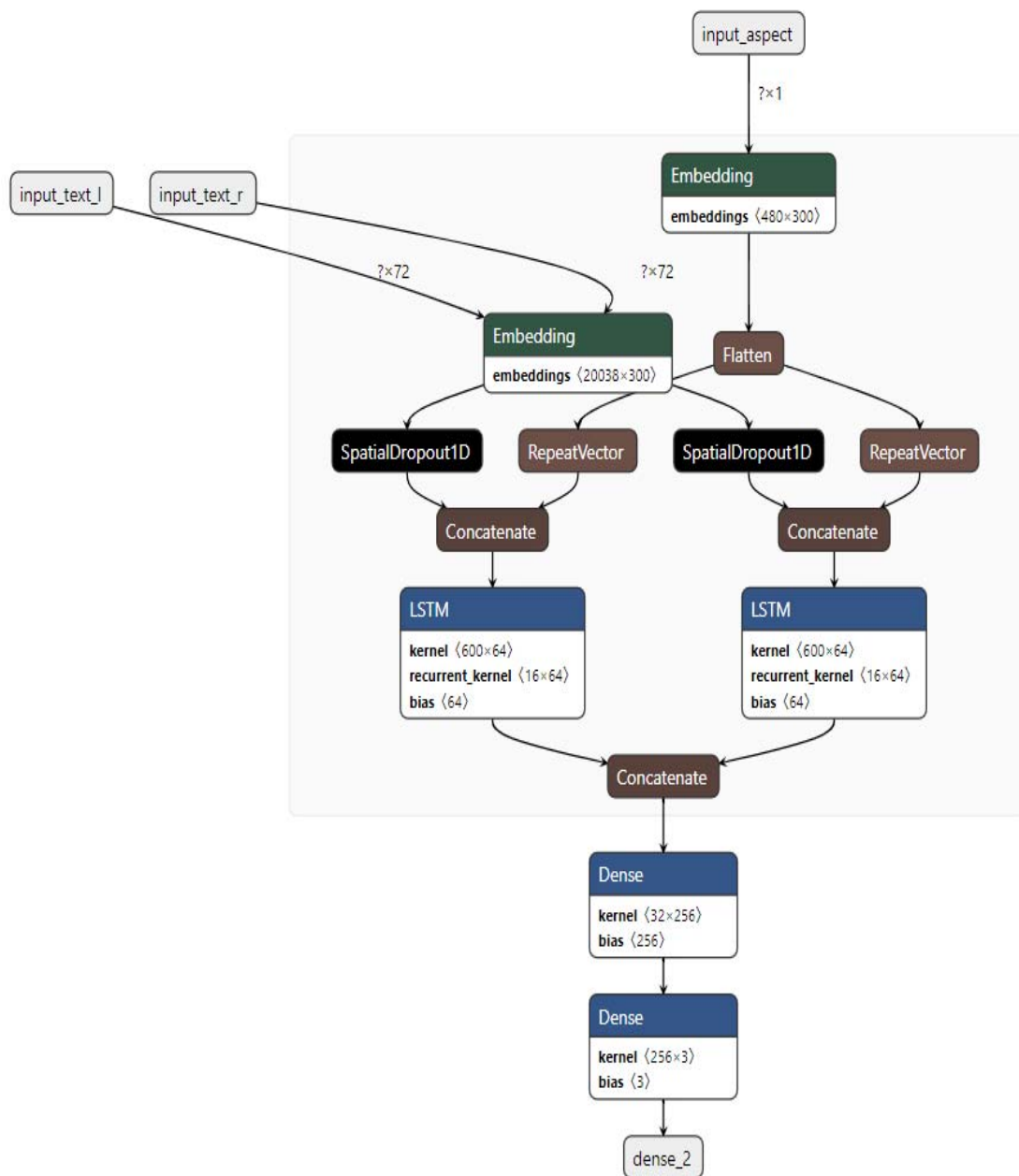


Fig. 4. Model Summary with weight, bias, and layers of TC-LSTM

3.3.4 ATAE-LSTM

One of the shortcomings of encoder-decoder architecture where one set of LSTM's learns to encode a set of input sequences and other sets of LSTM's read internal representation to decode it to output sequence suffers from the constraint that all input sequences should be encoded to the fixed-length vector [Attention in Long Short-Term Memory Recurrent Neural Networks]. Attention within sequences frees the encoder-decoder from having fixed-length vector representation, which makes the case for evaluation of ATAE-LSTM in this study. In ATAE-LSTM, aspect embeddings play a role in computing attention weights. To better utilize aspect information, aspect embedding is appended to each input word. The sentence vectors from the LSTM layers are then inputted to a dense layer with a rectified linear activation function to induce nonlinearity into the model. A final fully connected dense layer with a softmax activation function is added to represent the probability distribution of sentiment labels.

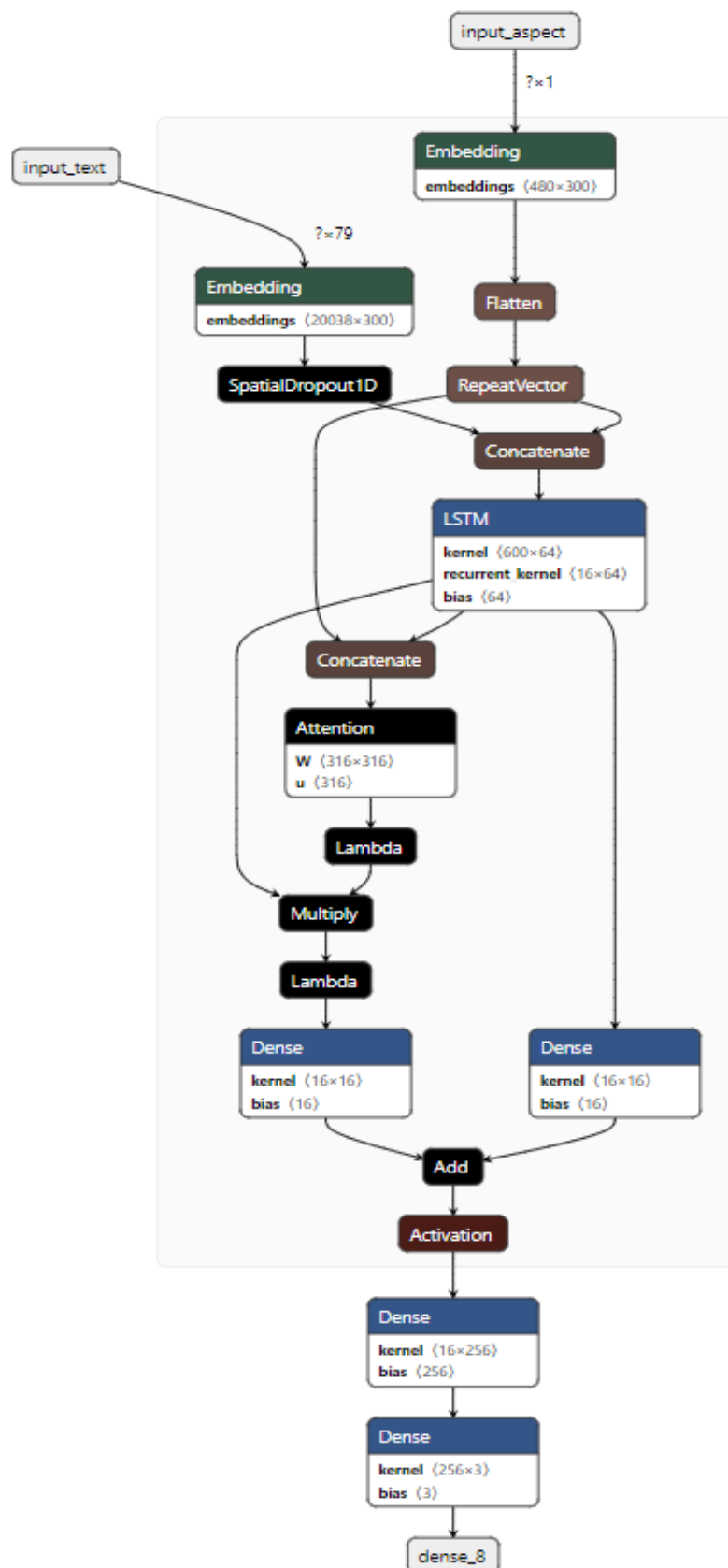


Fig. 5. Model Summary with weight, bias, and layers of ATAE-LSTM

4. Experiments and Results

4.1 Extracting Aspect Terms using unsupervised ABAE

Batch size is an important hyperparameter on which the dynamics of training and learning algorithms are influenced. It defines the number of examples from the training dataset that will be propagated through the network. It is used in estimating the accuracy of the estimate of error gradient when training neural networks and is extensively used to control the stability of the network [How to Control the Stability of Training Neural Networks With the Batch Size]. Batch sizes experiments with batch sizes of 64, 128, 256, and 512 did not have much difference concerning total loss and max-margin loss.

In this paper, a mini-batch method is followed to determine the loss. The mini-batch size is calculated by performing an integer division on the length of the training sample with the batch size. In this study, optimization algorithms AdaGrad, AdaDelta, RmsProp, SGD, Adam, and AdaMax are experimented with [Soydaner, 2020].

| | <i>AdaGrad</i> | <i>AdaDelta</i> | <i>RmsProp</i> | <i>SGD</i> | <i>Adam</i> | <i>AdaMax</i> |
|---------------|----------------|-----------------|----------------|------------|-------------|---------------|
| Learning Rate | 0.01 | 1 | 0.001 | 0.01 | 0.001 | 0.002 |
| rho | - | 0.95 | 0.9 | - | - | - |
| momentum | - | - | - | 0 | - | - |
| epsilon | 1.00E-06 | 1.00E-06 | 1.00E-06 | - | 1.00E-08 | 1.00E-08 |
| beta_1 | - | - | - | - | 0.9 | 0.9 |
| beta_2 | - | - | - | - | 0.999 | 0.999 |
| decay | - | - | - | 0 | - | - |
| nesterov | - | - | - | FALSE | - | - |
| clipnorm | 10 | 10 | | 10 | 10 | 10 |
| clipvalue | 0 | 0 | 0 | 0 | 0 | 0 |

Table 1. Hyperparameters defined for the experiment on various optimization algorithms

Max Margin Loss

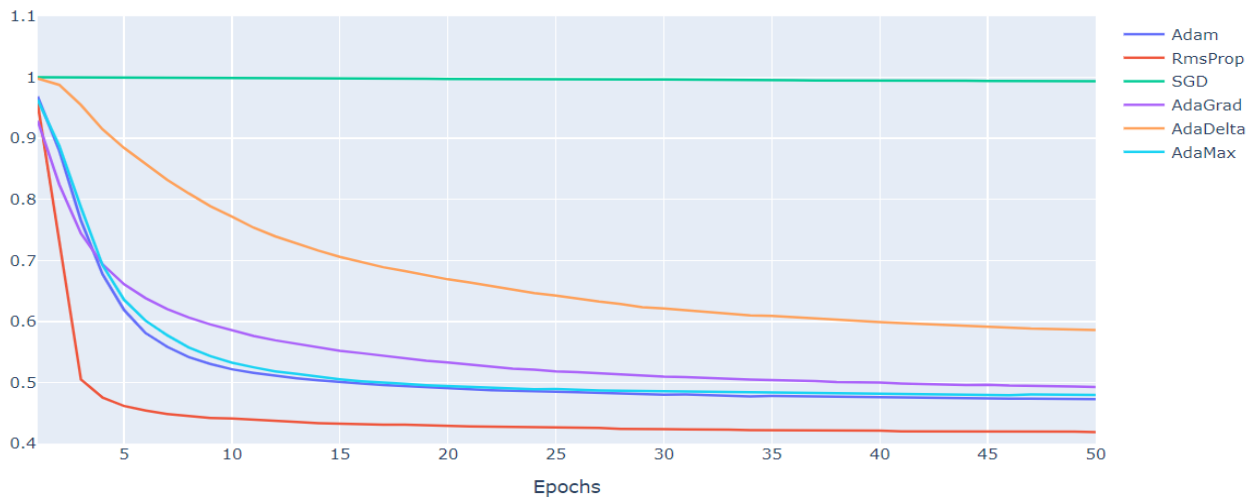


Fig. 6. Max Margin loss during training the model with various optimizers

It can be observed that the max-margin or hinge loss for SGD is constant throughout all the epochs. Hinge loss is not differentiable and is not a good loss function for methods that are not differentiable like SGD. It penalizes the data points lying on the wrong side of the hyperplane in a linear way [What Are Loss Functions in ML And Why Are They Important]. Learning rate, a configurable variable is one of the important hyperparameters which adjusts the weight of the neural network which apropos to the loss of gradients. How rapidly, or how slowly will the model move towards an optimal weight is determined by the learning rate [Understanding Learning Rate. Originally published at OpenGenus IQ. | by Aditya Rakhecha | Towards Data Science]. Learning rates can make huge differences in the accuracy of models. In this study, we experiment with three learning rates for Adam (0.001, 0.01, and 0.1).

| | | |
|-----------|---|----------|
| epsilon | - | 1.00E-08 |
| beta_1 | - | 0.9 |
| beta_2 | - | 0.999 |
| clipnorm | - | 10 |
| clipvalue | - | 0 |

Table 2. Learning rate is changed and experimented with the above-defined hyperparameter setting

Max Margin Loss vs Learning Rate

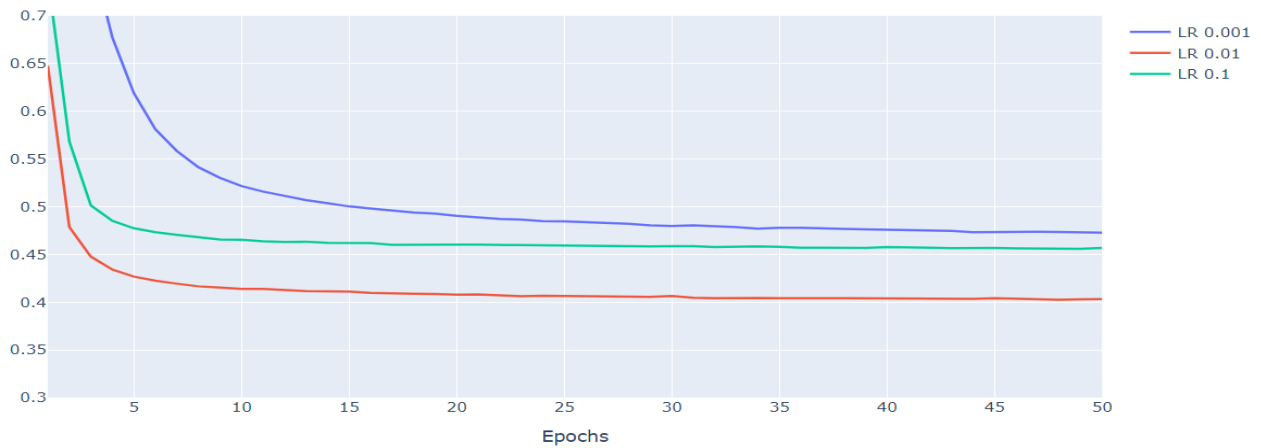


Fig. 7. Max Margin loss convergence for different learning rates of Adam

Based on the experiments the optimal hyper-parameter setting selected for the model are aspect size = 18 (number of categories); Batch size = 512; Epochs = 50; Adams optimizer = 0.01; Embedding dimension = 200; Orthogonal Regularization = 0.1; Negative Size = 1.

A cluster map with appropriate aspect names is created for the aspect categories classified by the unsupervised attention model. i.e. Aspect categories clustered by {aspect 0: banks, psbs, ..., aspect 1: naxalities, encounters ..., aspect 2: power, mw ... aspect 16: doctors, specialists..., aspect 17: water, groundwater ...} are redefined in a cluster map to provide human-understandable aspect names.

```
cluster_map = {
    0: 'Banking and Finance', 1: 'Violence', 2: 'Power', 3: 'Information governance',
    4: 'Telecommunication', 5: 'Education', 6: 'Railways', 7: 'Foreign Affairs', 8: 'Ore',
    9: 'Misc', 10: 'Commodities', 11: 'Aviation', 12: 'Disabled, Children and Women welfare', 13: 'SME and Tourism',
    14: 'Schemes', 15: 'Road Infrastructure', 16: 'Organisation and Companies', 17: 'Water'
}
```

Fig. 8. Cluster map for the aspect categories.

4.1.1 Visualizing the attention layer

Predicated labels and a list of attention words (sorted by the highest attention weights) are computed for each 'primary_question'. Based on the aspect category, the attention words are validated against the aspect term list for that category to determine the best aspect term. E.g., let us consider the question "the loan schemes available for the poor but meritorious students for completion of their education and the provisions therein?" which is classified into aspect category Education. Aspect category Education has aspect terms ['learning', 'graduation', 'taught', 'students', 'iits', 'teacher']. Below are some of the example question texts with attention weight to the words.

Question: "the quantity of rice and vegetables being exported and their volume and value, product-wise?"

| quantity | rice | vegetables | exported | volume | value | product |
|----------|-------|------------|----------|--------|-------|---------|
| 0.048 | 0.356 | 0.356 | 0.048 | 0.048 | 0.048 | 0.048 |

Aspect Category: Commodities; Aspect Term: rice

Question: “the details of central universities which have been given minority status, State-wise?”

| details | central | universities | given | minority | status | state | wise |
|---------|---------|--------------|-------|----------|--------|-------|-------|
| 0.048 | 0.048 | 0.356 | 0.048 | 0.356 | 0.048 | 0.048 | 0.048 |

Aspect Category: Education; Aspect Term: universities

Question = “the impact of drought on the Rabi sowing in various States?”

| impact | drought | rabi | sowing | various | states |
|--------|---------|-------|--------|---------|--------|
| 0.04 | 0.294 | 0.294 | 0.294 | 0.04 | 0.04 |

Aspect Category: Commodities; Aspect Term: rabi

Question: “whether it is a fact that Government is considering to install oxygen bars at major airports in the country?”

| considering | install | oxygen | bars | major | airports | country |
|-------------|---------|--------|-------|-------|----------|---------|
| 0.075 | 0.075 | 0.075 | 0.075 | 0.075 | 0.552 | 0.075 |

Aspect Category: Aviation; Aspect Term: airports

4.1.2 Evaluating Aspect Term/Category quality

The below table presents aspect terms inferred for all 18 aspect categories using the unsupervised attention-based aspect extraction model.

| Aspect Category | Aspect Terms |
|------------------------|--|
| Banking and Finance | banks, psbs, borrowers, write, defaulters, npa, nationalised, nationalized, npas, performing_assets_npas, defaulted, rbi, borrowing, lenders, performing_assets, loans, bank, waived, loan, sbi, lending, nbfc, corporates, debts, borrowings, rrbs, atms, deposits, credit, waiver |
| Violence | naxalites, encounters, civilians, insurgency, naxalite, terrorists, encounter, militants, maoists, firing, killings, bsf, extremists, fighting, crpf, naxals, jawans, militant, naxalism, paramilitary, maoist, troops, innocent, itbp, rifles, killing, attacks, policemen, forces, strikes |
| Power | power, mw, thermal, energy, megawatt, megawatts, ntpc, solar, hydro, reactors, plants, generation, electricity, reactor, generating, nuclear, renewable, ultra mega, nhpc, umpps, wind, kudankulam, rooftop, hydel, grid, npcil, atomic, fuel, jaitapur, ramagundam |
| Information governance | constitutional, lokpal, petition, judgement, hearing, petitions, appeal, rajya_sabha, honble, cic, cpc, observations, anomalies, ceo, judgment, legislations, commissions, verdict, reply, constitution, composition, legislative, appeals, court, ordinance, lok_sabha, filing, passed, hon_ble, writing |
| Telecommunication | phone, calls, sim, mobile, unsolicited, customers, roaming, signals, phones, sms, networks, mobiles, vodafone, telephones, telephony, portability_mnp, landline, internet, messages, voice, providers, gsm, telecom, cable, towers, operators, subscribers, cellular, misleading_advertisements, computers |
| Education | learning, graduation, taught, students, iits, teacher, universities, iims, affiliated, courses, admissions, textbooks, campuses, cbse, studying, sanskrit, academic, admission, schools, colleges, coaching, nits, degrees, navodaya, pursue, institutes, secondary, exams, teachers, undergraduate |
| Railways | railway, railways, trains, rail, tracks, platforms, emu, suburban, railwaysa, broad_gauge, junction, train, passenger, shatabdi, divisions, narrow_gauge, rangia, station, gauge_conversion, ticketing, stalls, duronto, compartments, rajdhani, irctc, superfast, coach, coaches, ticket, secunderabad |
| Foreign Affairs | washington, brazil, malaysia, israel, germany, maldives, canada, afghanistan, brics, pact, saarc, russia, asean, uk, korea, kingdom, iran, mauritius, america, negotiations, fta, vietnam, paris, dialogue, bilateral, africa, counterpart, indonesia, ties, summit |

| | |
|--------------------------------------|---|
| Ore | lignite, bauxite, ore, mines, leases, coal, manganese, mcl, mining, gmhc, mahanadi_coalfields, iron_ore, vedanta, mine, miners, coalfields, minerals, extraction, excavation, mineral, posco, drilling, lease, bccl, captive, stone, sand, blocks, hills, forests |
| Misc | years, last, three, year, yearwise, five, two, wise, programme, current, months, past, mohlai, rests, fdr, mtr, saurashtra, chichira, postgraduates, ucil, inpossession, shatab, ucbs, naicker, sovereignty, corresponding, kalam, tangdar, magwani, interested |
| Commodities | onion, onions, wheat, rice, pulses, potato, rabi, produces, bumper, soyabean, quintal, paddy, cereals, mustard, edible, tomato, vegetables, edible_oils, dal, sugarcane, maize, horticultural, oilseeds, fruits, grains, kharif, arhar, acreage, sowing, vegetable |
| Aviation | flights, air, flight, ai, airlines, aircrafts, runway, carrier, carriers, landing, fleet, airports, kingfisher, airline, aircraft, airport, cargo, aai, jet_airways, pilots, jet, flying, minute, privatise, fares, dreamliner, trivandrum, planes, baggage, boarding |
| Disabled, Children and Women welfare | infants, children, kids, aged, anaemia, destitute, malnutrition, mothers, malnourished, marriage, child, beggars, abused, begging, underweight, married, leprosy, juveniles, prevalence, elderly, female, abuse, trafficked, orphanages, juvenile, marriages, mortality, adult, hiv, rape |
| SME and Tourism | handicraft, apparel, handicrafts, leather, textiles, competitiveness, cottage, msme, tourism, textile, promoting, intensive, industry, thrust, attracting, oriented, handloom, agro, cluster, innovation, handlooms, cooperatives, inclusive, fpis, cruise, powerloom, processing, dairy, clusters, exclusive |
| Schemes | yojana, bima, swarozgar, rsby, ujjwala, pmjdy, sgsh, deen_dayal_upadhyaya, jyoti, nirman, rashtriya_swasthya_bima, iay, pmksh, grameen, jan_dhan, awaas, swarozgar, rajiv, grameen_vidyutikaran, nirmal, awas, swarnjayanti, gramsh, saansad_adarsh, sagy, aay, deendayal, sabla, abhiyan, rggvy |
| Road Infrastructure | highway, highways, lane, nh, roads, road, laning, stretches, expressways, nhs, stretch, ring, expressway, lanes, km, bypass, kms, nhai, nhdp, flyovers, length, bridges, widening, toll, golden_quadilateral, kilometers, toll plazas, rampur, bridges_robs, bro |
| Organization and Companies | doctors, specialists, engineers, contractual, unani, specialist, esi, dispensaries, safdarjung_hospital, nursing, empanelled, sciences_aiims, cghs, ayurvedic, deputation, employee, hospital, retired, retirement, employeees, dispensary, resigned, employees, inducted, ayurveda, esic, hands, quit, homoeopathy, cabin_crew |
| Water | water, groundwater, lakes, contamination, river, ponds, ground, polluted, arsenic, harvesting, effluents, rivers, drinking, rain, tributaries, yamuna, potable, ganga, depletion, fluoride, reservoirs, flood, surface, flows, depleting, soil, toxic, flowing, narmada, glaciers |

Table 3. Top 30 aspect terms per category classified by unsupervised aspect-based attention model

Coherence Score is used to evaluate the quality of aspect terms. The cosine similarity score between two words can be calculated from the keyed vector model of word2vec. A list of lists is created to store aspect terms, where each sublist represents an aspect category. Each aspect term in the sublist is compared with every other word in the sub list to calculate the average coherence score for the sub list. The average of the coherence across all sublists is returned as the final coherence score for the aspects categorized by the unsupervised attention model.

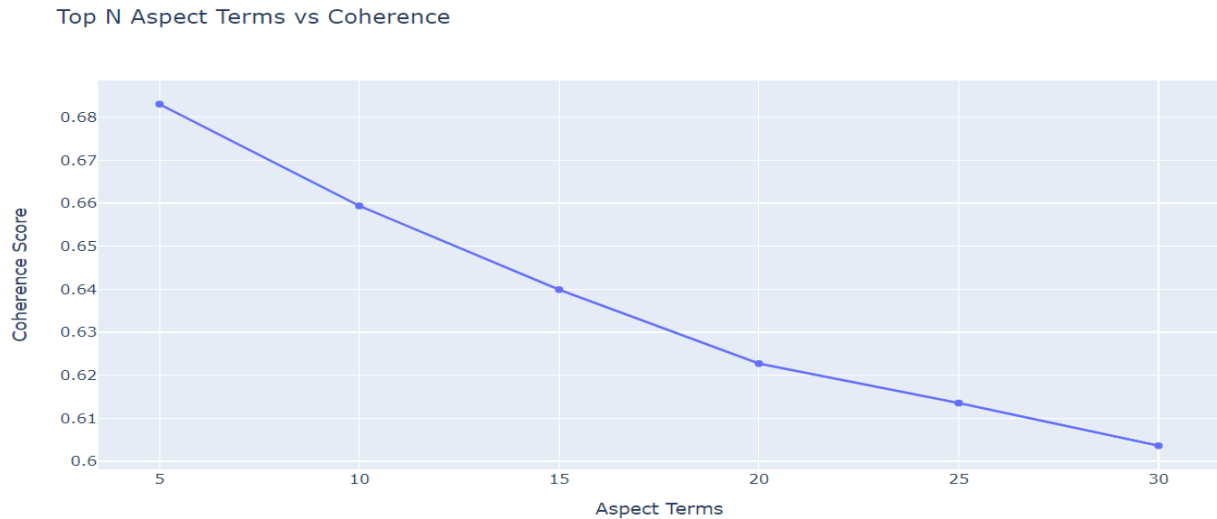


Fig. 9. Top 'N' Aspect Terms vs Coherence Score from the Unsupervised ABAE model

We compared the unsupervised ABAE coherence score with the coherence score generated from LDA MALLET topic modeling.

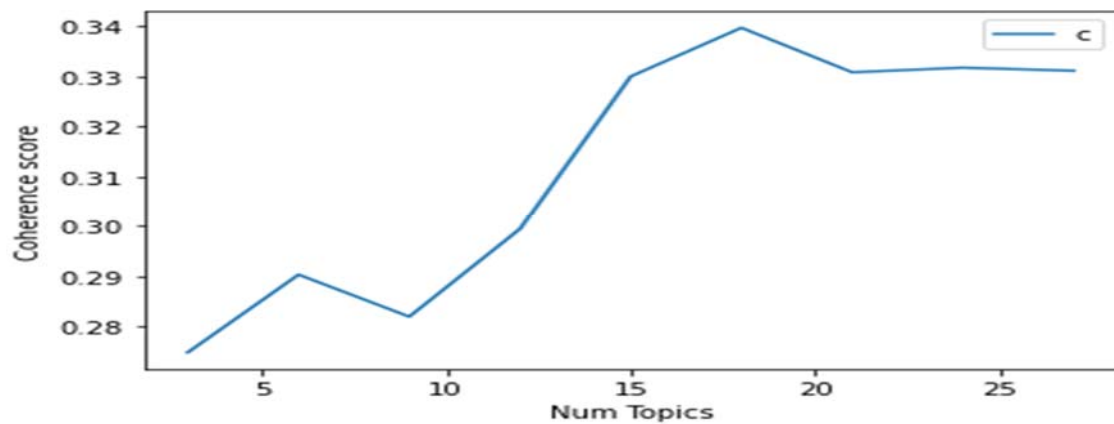


Fig. 10. Coherence Score of LDA MALLET topic model

The aspect terms obtained from our unsupervised ABAE model are relevant to the dataset and the domain of the study. The coherence score obtained for the top 30 aspect term is ~ 0.6 , with an improvement of ~ 0.27 as compared to the LDA MALLET model. E.g. terms classified as aspect category 'SCHEME' by LDA would be {national, rural, scheme, employment, government, act, gandhi, fact, labour, states, guarantee, areas, years, mahatma, funds, country}. It can be observed that terms classified by unsupervised ABAE {yोजना, bima, swarozgar, rsby, ujjwala, pmjdy, sgys, deen_dayal_upadhyaya, jyoti, nirman, rashtriya_swasthya_bima, iay, pmksy, grameen, jan_dhan, awaas, swarojgar, rajiv, grameen_vidyutikaran, nirmal, awas, swarnjayanti, gramin, saansad_adarsh, sagy, aay, deendayal, sabla, abhiyan, rggvy} is more relevant to the domain.

4.2 Supervised ABSA

The aspect category labels and sentiment labels (generated using VADER, as described in section 3.2.4) is used as an input seed for performing supervised ABSA.

4.2.1 Preparing Input and Building Base network

For TD-LSTM, input_l, and input_r which are embedding for words located at the left of the aspect word and right of the aspect words are created. Spatial Dropouts are added to the embedding layer. Adding spatial dropout is considered to be an effective mechanism to prevent overfitting of the model [Gal and Ghahramani, 2016]. Two hidden LSTM layers are created, one each for left input text embeddings and right text embeddings.

Similarly, for TC-LSTM, embedding for words on the left of the aspect word and words appearing on the right of the aspect word are created. A spatial dropout is added to both the embedding layers to prevent overfitting. Additionally, an aspect embedding layer is used to consider the aspect words. By making use of RepeatVector the squashed single feature vectors in the aspect embedding layer are reconstructed and concatenated to the left and

right input embedding layers. Two hidden LSTM layers are created, one each for left input text embeddings and right text embeddings.

For ATAE-LSTM, two embedding layers of input text and aspect text are initially created. Spatial dropout is added to the input text embedding layer and RepeatVector for the aspect embedding layer as both are concatenated into an LSTM layer. Hidden vectors and final hidden state values from the LSTM layer are captured. To implement the attention mechanism, hidden vectors from the LSTM and the aspect embeddings are concatenated into a dense attention layer. Another dense layer consisting of the final hidden state values of LSTM is created.

Based on the various experiments conducted for supervised ABSA, the optimal hyper-parameters selected are: Word embedding dimension = 300; Aspect embedding dimension = 300; Embedding type = glove and word2vec; LSTM units = 16; Dense units = 256; Spatial dropout = 0.2; Batch Size = 256; Optimizer = Adam; Learning rate = 0.001.

4.2.2 Model Accuracies and Model Loss

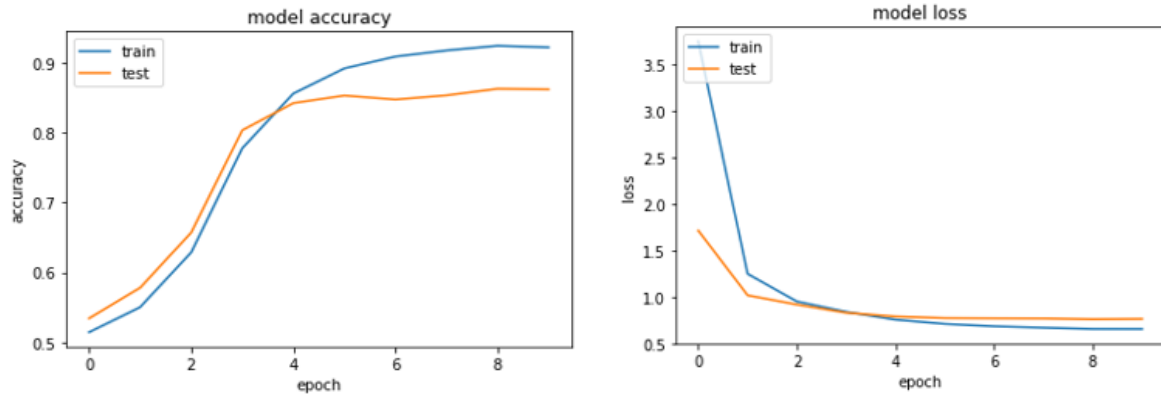


Fig. 11. TD-LSTM: Model accuracy and Model loss plots on train and validation dataset using word2vec embeddings.

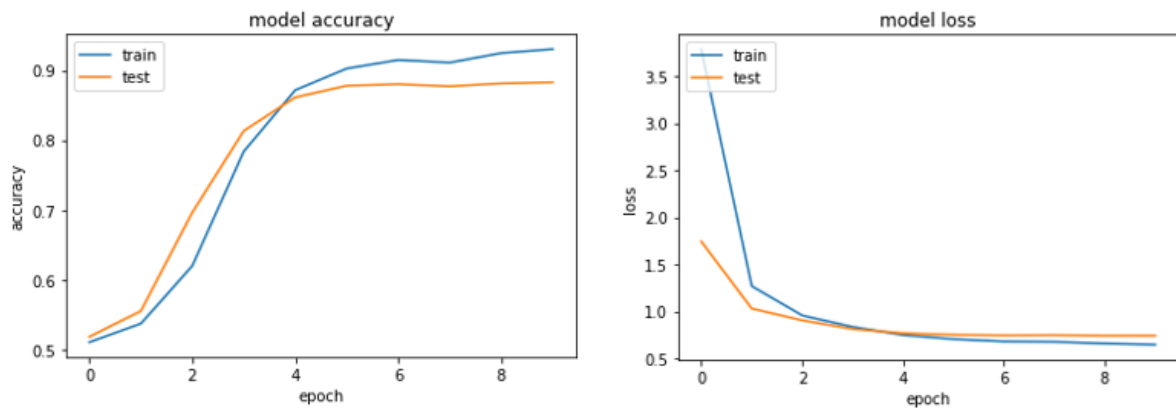


Fig. 12. TD-LSTM: Model accuracy and Model loss plots on train and validation dataset using GloVe embeddings.

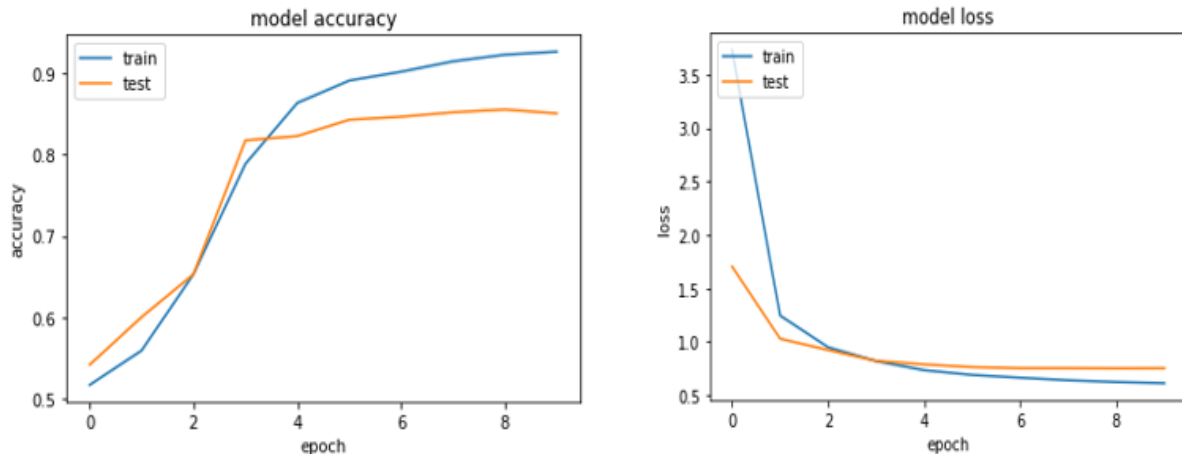


Fig. 13. TC-LSTM: Model accuracy and Model loss plots on train and validation dataset using word2vec embeddings.

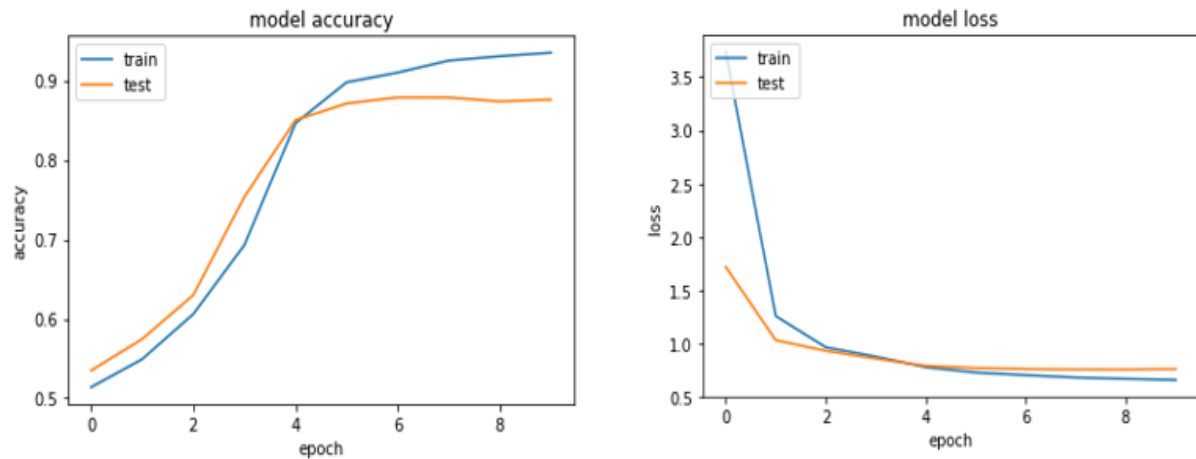


Fig. 14. TC-LSTM: Model accuracy and Model loss plots on train and validation dataset using GloVe embeddings.

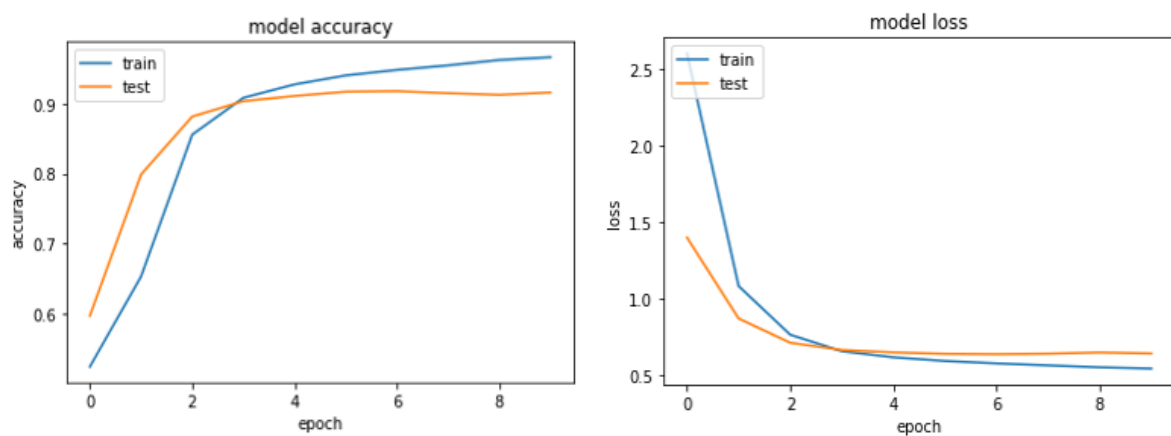


Fig. 15. ATAE-LSTM: Model accuracy and Model loss plots on train and validation dataset using word2vec embeddings.

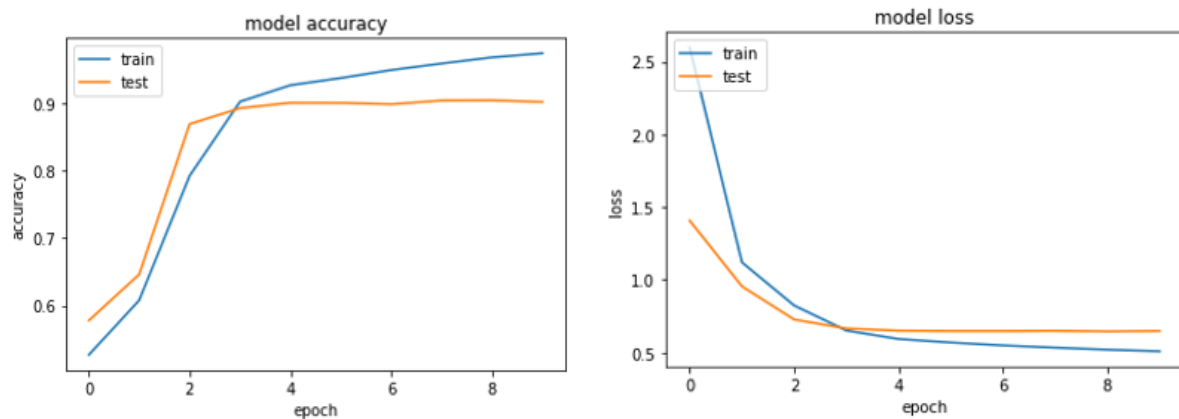


Fig. 16. ATAE-LSTM: Model accuracy and Model loss plots on train and validation dataset using Glove embeddings.

4.2.3 Evaluating the accuracy and F1 scores

In this study, F1 scores are used to evaluate how precise the model is. F1 scores are also a good indicator of how precise the model is. In the below figure we look at the accuracy and F1 scores for the evaluated supervised algorithms. The scores are based on the epoch where training and validation loss converge.

| Model | Embedding Type | Accuracy | Macro F1 |
|-----------|----------------|----------|----------|
| TC-LSTM | word2vec | 0.81 | 0.8 |
| | GloVE | 0.85 | 0.84 |
| TD-LSTM | word2vec | 0.84 | 0.82 |
| | GloVE | 0.86 | 0.85 |
| ATAE-LSTM | word2vec | 0.89 | 0.88 |
| | GloVE | 0.9 | 0.89 |

Table. 4. Accuracy and F1 scores of various supervised models implemented.

5. Conclusion and Future scope

5.1 Conclusion

In this paper, the study on aspect term detection, sentiment classification, and aspect-based sentiment analysis is carried out. Aspect terms are identified from the unsupervised ABAE model are classified into various aspect categories. The coherence score for the top 30 aspect terms in each aspect size ~ 0.6 , which is much better and accurate in comparison with the LDA MALLET topic modeling score. This paper also creates a gold standard training dataset with aspect and sentiment labels in the unsupervised ABAE model phase, which will be very much helpful for further researchers working in the Indian Political domain. Among the supervised aspect-based sentiment models, ATAE-LSTM is shown to achieve the highest accuracy on the domain dataset with an accuracy of ~ 0.9 and a Macro F1 score of ~ 0.89 .

5.2 Future Scope

Various optimizers with different hyperparameters can be experimented with in the supervised models used for the ABSA task. In this paper, only question texts are evaluated for aspects and sentiments while subjectively considering the aspects and sentiments of the answer text. Only a single aspect per question is considered in the current study. Future works can be carried on by considering various aspects present in the Q-A pair to identify the sentiments for these aspects.

References

- [1] A.I. Will Mark A Turning Point in the History of Politics | by Sukhayl Niyazov | Towards Data Science.
- [2] Ansari MZ, Aziz MB, Siddiqui MO, Mehra H, Singh KP. 2020. Analysis of Political Sentiment Orientations on Twitter. *Procedia Comput. Sci.* 167: 1821–1828.
- [3] Attention in Long Short-Term Memory Recurrent Neural Networks.
- [4] Bouma G. 2009. Normalized (Pointwise) Mutual Information in Collocation Extraction. *Proc. Ger. Soc. Comput. Linguist. (GSCS 2009)*: 31–40.
- [5] Demeester T, Rocktäschel T, Riedel S. 2016. Lifted rule injection for relation embeddings. *EMNLP 2016 - Conf. Empir. Methods Nat. Lang. Process. Proc.*: 1389–1399.
- [6] Elghazaly T, Mahmud A, Hefny H. 2016. Political Sentiment Analysis Using Twitter Data., p 1–5.
- [7] Gal Y, Ghahramani Z. 2016. A theoretically grounded application of dropout in recurrent neural networks. *Adv. Neural Inf. Process. Syst.*: 1027–1035.
- [8] Glavas G, Nanni F, Ponzetto SP. 2019. Computational analysis of political texts: Bridging research efforts across communities. *ACL 2019 - 57th Annu. Meet. Assoc. Comput. Linguist. Tutor. Abstr.*: 18–23.
- [9] GloVe: Global Vectors for Word Representation.
- [10] Gold D, Bexte M, Zesch T. 2018. Corpus of aspect-based sentiment in political debates. *KONVENS 2018 - Conf. Nat. Lang. Process. / Die Konf. zur Verarbeitung Nat. Spr.*: 89–99.
- [11] He R, Lee WS, Ng HT, Dahlmeier D. 2017. An unsupervised neural attention model for aspect extraction. *ACL 2017 - 55th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap. 1)*: 388–397.
- [12] How AI will decide your future Prime Minister.
- [13] How to Control the Stability of Training Neural Networks With the Batch Size.
- [14] Hutto CJ, Gilbert EE. 2014. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*. Proc. 8th Int. Conf. Weblogs Soc. Media, ICWSM 2014.
- [15] Khurana D, Koli A, Khatter K, Singh S. 2017. Natural Language Processing: State of The Art, Current Trends and Challenges.
- [16] Mikolov T, Chen K, Corrado G, Dean J. 2013. Efficient estimation of word representations in vector space. *1st Int. Conf. Learn. Represent. ICLR 2013 - Work. Track Proc.*: 1–12.
- [17] Monroe BL, Schrodt PA. 2008. Introduction to the Special Issue: The Statistical Analysis of Political Text. *Polit. Anal.* 16: 351–355.
- [18] NLP - overview.
- [19] Q & A Discussed in Parliament of India | Kaggle.
- [20] Rauh C. 2018. Validating a sentiment dictionary for German political language—a workbench note. *J. Inf. Technol. Polit.* 15: 319–343.
- [21] Rohit SVK, Singh N. 2018. Analysis of Speeches in Indian Parliamentary Debates.
- [22] Rose S, Engel D, Cramer N, Cowley W. 2010. Automatic Keyword Extraction from Individual Documents. In: *Text Mining: Applications and Theory*, p 1–20.
- [23] Ruder S, Ghaffari P, Breslin JG. 2016. A hierarchical model of reviews for aspect-based sentiment analysis. *EMNLP 2016 - Conf. Empir.*

- Methods Nat. Lang. Process. Proc.: 999–1005.
- [24] Sharma P, Moh TS. 2016. Prediction of Indian election using sentiment analysis on Hindi Twitter. Proc. - 2016 IEEE Int. Conf. Big Data, Big Data 2016: 1966–1971.
- [25] Shen C, Sun C, Wang J, Kang Y, Li S, Liu X, Si L, Zhang M, Zhou G. 2020. Sentiment classification towards question-answering with hierarchical matching network. Proc. 2018 Conf. Empir. Methods Nat. Lang. Process. EMNLP 2018: 3654–3663.
- [26] Sidarenka U. 2019. Sentiment analysis of German Twitter.
- [27] Singhal K, Agrawal B, Mittal N. 2015. Modeling Indian General Elections: Sentiment Analysis of Political Twitter Data. In: Mandal JK, Satapathy SC, Kumar Sanyal M, Sarkar PP, Mukhopadhyay A, editors. Information Systems Design and Intelligent Applications. New Delhi: Springer India, p 469–477.
- [28] Soydaner D. 2020. A Comparison of Optimization Algorithms for Deep Learning. Int. J. Pattern Recognit. Artif. Intell.
- [29] Tang D, Qin B, Feng X, Liu T. 2016. Effective LSTMs for target-dependent sentiment classification. COLING 2016 - 26th Int. Conf. Comput. Linguist. Proc. COLING 2016 Tech. Pap.: 3298–3307.
- [30] Timonen M, Toivanen T, Teng Y, Chen C, He L. 2012. Informativeness-based Keyword Extraction from Short Documents. In: KDIR 2012 - Proceedings of the International Conference on Knowledge Discovery and Information Retrieval., p 411–421.
- [31] Understanding Learning Rate. Originally published at OpenGenus IQ. | by Aditya Rakhecha | Towards Data Science.
- [32] Vo DT, Zhang Y. 2015. Target-dependent twitter sentiment classification with rich automatic features. IJCAI Int. Jt. Conf. Artif. Intell. 2015-Janua: 1347–1353.
- [33] Wang Y, Huang M, Zhao L, Zhu X. 2016. Attention-based LSTM for aspect-level sentiment classification. EMNLP 2016 - Conf. Empir. Methods Nat. Lang. Process. Proc.: 606–615.
- [34] What Are Loss Functions in ML And Why Are They Important.
- [35] Zhang L, Wang S, Liu B. 2018. Deep Learning for Sentiment Analysis : A Survey. Wiley Interdiscip. Rev. Data Min. Knowl. Discov.