



Department of Computer Science & Engineering
University of Moratuwa

FINAL YEAR PROJECT - GROUP 43

**PARTY BASED SENTIMENT ANALYSIS OF LEGAL OPINION
TEXTS**

Supervisors:

Dr. Shehan Perera

Dr. Nisansa de Silva

Group Members:

S.R. Isanka (160222M)

D.V.D.R. Karunarathna (160280L)

M.M.C.Ruchini (160537H)

February, 2021

Declaration

“We declare that this is our own work and this thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, we hereby grant to the University of Moratuwa the non-exclusive right to reproduce and distribute our thesis/dissertation, in whole or in part in print, electronic or other medium. We retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

Signature: Date:

Signature: Date:

The supervisor/s should certify the thesis/dissertation with the following declaration.

The above candidate has carried out research for the Masters/MPhil/PhD thesis/ Dissertation under my supervision.

Name of the supervisor:

Signature of the supervisor: Date:

Acknowledgments

First and foremost, we would like to thank Dr. Adeesha Wijayasiri, Coordinator of Final Year Projects, Department of Computer Science and Engineering, University of Moratuwa, for providing us with this valuable opportunity to put into practice what we learned and research new concepts and ideas.

Next, we would like to express our deep sense of gratitude and a profound feeling of admiration to Dr. Shehan Perera, Department of Computer Science and Engineering, University of Moratuwa, our project supervisor, who provided us with the project opportunity and guided throughout the whole duration of this research, supported to overcome every hindrance, sharing his expertise and experiences with the project.

Also, we would like to acknowledge with much appreciation Dr. Nisansa De Silva, Department of Computer Science and Engineering, University of Moratuwa, our project supervisor(external) whose work laid the foundation for our project. His support, advice, guidance and expertise massively helped us to successfully complete this project. We would also like to express our heartfelt gratitude to Dr. Nisansa De Silva for her cordial encouragement and making our first research experience a remarkable one.

We would also like to gratefully acknowledge and to offer our gratitude to Mr. Gathika Ratnayaka for providing constant support for the implementation of our project and sharing his expertise and experiences with the project.

Furthermore, we would gratefully acknowledge Ramesh Pathirana and his team, post-graduates in the Faculty of Law, University of Colombo for providing the advice, guidance related to the legal domain and the support in data annotation process.

We would like to express our gratitude towards Prof. Sanath Jayasena and Mrs. Vishaka Nanayakkara for providing valuable feedback during the project evaluations.

At last but not least we would like to express our very profound gratitude who advised, helped and motivated us to complete this project successfully and providing us with unfailing support and continuous encouragement throughout our years of study and through the process of researching and writing this thesis.

Contents

1	Introduction	12
1.1	Overview	12
1.2	Motivation	14
1.3	Problem Statement	15
1.4	Objectives of the Research	16
2	Literature Review	17
2.1	Legal Information Extraction	17
2.2	Benchmark Datasets for ABSA tasks in different domains	18
2.3	Earlier approaches to ABSA tasks	19
2.4	Deep Learning for Aspect Based Sentiment Analysis	19
2.4.1	Deep Neural Network Layers	20
2.4.2	Convolutional Neural Network Model (CNN)	20
2.4.3	Recurrent Neural Network (RNN) and Long-short Term Memory (LSTM) Models	21
2.4.4	Attention Mechanism for Classification	21
2.4.5	Memory Network for Aspect Sentiment Classification	23
2.4.6	Bidirectional Encoder Representations from Transformers [BERT]	23
2.4.7	Graph Convolutional Network [GCN]	24
2.5	Evaluation Techniques	24
2.6	Discussion	25
3	Research Methodology	27
3.1	Creation of Dataset (SigmaLaw-ABSA)	29

3.1.1	Data Collection	29
3.1.2	Data Annotation	30
3.2	Rule-Based Approach	32
3.2.1	Introduction	32
3.2.2	Co-reference Resolution	32
3.2.3	Generating sub sentences	34
3.2.4	Extracting phrases	35
3.2.5	Phrase-level Sentiment Annotator	38
3.2.6	Sentiment Classification	38
3.3	Ensemble Learning Approach	40
3.3.1	Research on ABSA Architectures	40
3.3.2	Ensemble Model	43
3.4	Deep Learning-based Approach	44
3.4.1	Introduction	44
3.4.2	Word Embedding Layer	44
3.4.3	RNN Layer	46
3.4.4	Position Aware Attention Mechanism	46
3.4.5	Graph Convolution Network	48
3.4.6	Sentiment Classification	50
3.4.7	Model Training	50
4	Experiments	51
4.1	Experimental Setup and Dataset	51
4.1.1	Dataset Statistics	51
4.2	Experimental Results	53
4.2.1	Rule-Based Approach	53
4.2.2	Research on ABSA deep-learning based Architectures	55
4.2.3	Deep Learning Approach	58
4.2.3.1	Word Embedding Models Comparison	58
4.2.3.2	RNN Models Comparison	58
4.2.3.3	Overall Performance	60
4.2.3.4	Ablation Study	60

5	Discussion	63
5.1	Multi Functionality of SigmaLaw-ABSA Dataset	63
5.1.1	T1 - Aspect Extraction	63
5.1.2	T2 - Aspect Term Polarity	64
5.1.3	T3 - Aspect Category Extraction	64
5.1.4	T4 - Aspect Category Polarity	64
5.2	Future Work	65
5.3	Individual Contribution	65
5.4	Conclusions	67

List of Figures

1.1	Research Outline	16
3.1	Constituency Parser Tree for Example1	31
3.2	Rule-Based Model Architecture	33
3.3	Co-reference Resolution	34
3.4	Constituency Parser Tree for a sentence	35
3.5	Main phrases of a sentence	36
3.6	The structure of verb phrase: type 01	37
3.7	The structure of verb phrase: type 02	37
3.8	Process of dividing sentence	40
3.9	Input and Output of the Deep Learning Model	41
3.10	Stacking Ensemble Model	44
3.11	Overall architecture	45
3.12	Relative Distance	47
3.13	Dependency Graph	47
3.14	Sentiment Graph	49
4.1	Word Count Frequency Plot	53
4.2	Precision Variation	55
4.3	Recall Variation	57
4.4	F1 score Variation	57
5.1	Inputs of the system	78
5.2	Output results	79

List of Tables

2.1	Review of Existing Models	26
3.1	Sentiment Classes Mapping	38
4.1	Dataset Statistics	52
4.2	Statics for the Overall Sentiment of Sentences	52
4.3	Results of Proposed Approach	54
4.4	Results using Stanford Sentiment Annotator	55
4.5	Performances of Different Models on SigmaLaw-ABSA	56
4.6	Word embedding models comparison	59
4.7	Different word embedding strategies comparison of BERT model	59
4.8	RNN models comparison	60
4.9	Performances of Different Models on SigmaLaw-ABSA	61
4.10	Results of ablation study	62
5.1	Individual Project Contribution	66

Abstract

Factual scenario analysis of previous court cases holds a significant importance to lawyers and legal officers whenever they are handling a new legal court case. Legal officials are required to analyse previous court cases and statutes to find arguments and evidence before they represent a client at a trial. As the number of legal cases increases, legal professionals typically endure heavy workloads on a daily basis, and they may be overwhelmed and may be unable to obtain quality analysis. In this analysis process, identifying advantageous and disadvantageous relevant to legal parties can be considered a critical and time consuming task. By automating this task, legal officers will be able to reduce their workload significantly. Our ultimate goal of this research is to introduce a system to predict sentiment value of sentences in legal documents in relation to its legal parties. To achieve this task, we use a fine-grained sentiment analysis technique called Aspect-Based Sentiment Analysis(ABSA). To the best of our knowledge, this would be the first study that brings the concepts of Aspect Based Sentiment Analysis to the legal domain. To the best of our knowledge, there is no publicly available dataset for the aspect (party)based sentiment analysis for legal opinion texts. Hence, we created a dataset(SigmaLaw-ABSA) which consists 2000 legal opinion text fetched from court cases in order to train the models and to conduct the experiments. Then we developed a rule-based model which is primarily built around a phrase-level sentiment annotator and using rationally built rules to perform Party-Based Sentiment Analysis. The complexity of the structure of the sentences in legal texts has prompted to develop first a rule-based approach over other approaches. Developing this model, we got a much understanding on the legal domain, complex structure of the legal texts and many other benefits for the future works. But in the evaluation process, we experienced some major limitations as it's significantly depends on the phrase level sentiment annotator and

there's a huge amount of scenarios which can not cover all using manual rules. Then we developed a deep-learning based approach to mitigate those limitation and to perform the task efficiently. After going through the different existing architectures for other domains which have their own capabilities of achieving accurate output overcoming different limitations, we came up with a new architecture which can be introduced as a combination of all the capabilities of the above models. We evaluated this model and existing ABSA models on the SigmaLaw-ABSA dataset and experiments showed that our model outperforms the state-of-the-art models for the SigmaLaw-ABSA dataset.

A List of Abbreviations

ABSA - Aspect-Based Sentiment Analysis
PBSA - Party-Based Sentiment Analysis
NLP - Natural Language Processing
RNN - Recurrent Neural Network
LSTM - Long Short Term Memory
CNN - Convolutional Neural Network
GCN - Graph Convolutional Network
SVM - Support Vector Machine
DNN - Deep Neural Network
GRU - Gated Recurrent Unit
TD-LSTM - Target-Dependent LSTM
TC-LSTM - Target-Connection LSTM
AE-LSTM - Aspect Embedding LSTM
AT-LSTM - Attention-based LSTM
ATAE-LSTM - Attention-based with Aspect Embedding LSTM
HEAT - HiErarchical ATtention
RAM - Recurrent Attention on Memory
IAN - Interactive Attention Networks
BERT - Bidirectional Encoder Representations from Transformers
Bert-spc - BERT for Sentence Pair Classification
Cabasc - Content Attention Model for Aspect-Based Sentiment Analysis
Aen-Bert - Aspect Level Sentiment Classification with Deep Memory Network
Bert-LCF - Local Context Focus Mechanism
ASGCN - Aspect-specific Graph Convolutional Networks

S-LSTM - Sentence-State LSTM

SDGCN - Sentiment Dependencies with Graph Convolutional Networks

SBAR - Subordinating Conjunction

NP - Noun Phrase

VP - Verb Phrase

PP - Prepositional Phrase

RNTN - Recursive Neural Tensor Network

Chapter 1

Introduction

1.1 Overview

A document which elaborates opinions and arguments related to the previous court cases is known as a legal opinion document. Presently, lawyers and legal officers are confronted with an enormous amount of legal documents. Lawyers and legal officials have to spend considerable effort and time to obtain the required information manually from those documents when dealing with new legal cases [1, 2]. New court cases are also on the rise. Legal officials are therefore under considerable pressure to formulate case opinions in a timely manner [3]. Hence, it provides much convenience to those individuals if there is a way to automate the process of extracting information from legal opinion texts [4]. The process of automating the extraction of legal information has spread to a wide range. In this study, we focus on finding advantages and disadvantages facts or arguments in court cases which one of the most critical and time-consuming tasks in court case analyse. As this task plays an important role in the legal information extraction process, automation of this task makes a huge contribution to the legal domain.

Opinion mining or sentiment analysis is identifying opinions and then classifying them into several polarity levels (Positive or Neutral or Negative) using computational linguistics and information retrieval [5]. When it comes to the legal domain, sentiment

analysis becomes a much challenging area because of the domain-specific meanings or behavior of words in the legal opinion texts. The complexity and the length of the sentences also increases the difficulty. Sentiment Analysis(SA) can be divided into 4 levels; document-level sentiment analysis, sentence-level sentiment analysis, phrase-level sentiment analysis, and aspect level sentiment analysis [6]. Document Level SA considers the whole document is about an entity and classifies whether the expressed sentiment is positive, negative or neutral; Sentence Level SA determines the sentiment of each sentence, Phrase Level SA [7] focus on finding out the sentiment of phrases; Entity or Aspect-Based SA performs finer-grained analysis in which all entities and their aspects should be extracted and the sentiment on them should also be determined [8].

Sentences in a legal case usually contain two or more members/entities which belong to the main legal parties (plaintiff, petitioner, defendant, and respondent). Extracting opinions with respect to each legal party cannot be performed only by using document-level or sentence-level or phrase-level sentiment analysis. This is because those approaches only take the sentiment of the whole document, sentence or phrase without paying attention to any entity or aspect [9]. Therefore, aspect-based sentiment analysis is the most appropriate solution to perform party-based sentiment analysis. A number of researchers have addressed the Aspect Based Sentiment Analysis in different domains such as restaurants, hotels, movies, product reviews, government services, mobile phones, and telecommunication [10]. However, to the best of our knowledge, there is no research conducted on ABSA in the legal domain. Therefore, this can be considered as a significant research gap in the area of aspect-based sentiment analysis in the legal domain. When we consider the legal domain, legal parties can be considered as aspects, thus Aspect Based Sentiment Analysis in the legal opinion texts can also be named as party-based sentiment analysis.

Aspect extraction and sentiment classification are the main two subtasks in ABSA [9]. When applying ABSA into the legal domain, identifying legal parties and assigning sentiment level for each legal party in a legal opinion text can be taken as the main subtasks. For a given legal opinion sentence in a court case, party identification task deals with extracting all persons or organizations belong to any legal party[11]. In certain cases, every person or organization mentioned in a legal text would not be included for the main legal parties. As an example judge is also a person entity, but he does

not belong to any legal party. Moreover, the same party can be referred to in many ways. Both tasks hold some significant changeable processes because of the complex structure in the legal domain, unlike other domains. In this study, we only focus on the sentiment classification for each legal party and we provide legal party entities along with the sentence as inputs to the system. since, the system focuses only on identifying the sentiment from each sentence in relation to the given aspects, aspects have been manually extracted by the human annotators.

1.2 Motivation

Aspect Based Sentiment Analysis can be taken as one of the most important research areas in the legal domain, as it can be used to perform sentiment analysis with respect to each legal party in a court case and figure out which legal party has obtained the most winning arguments. When a lawyer or legal officer receives a new legal case, he has to go through all previous related court cases to gather the required information [12, 13]. However, doing this process manually consumes a significant amount of time and effort.

The main motivation for the project is the absence of a system or a research methodology that can perform Party-Based Sentiment Analysis. There are a few systems implemented to assist lawyers and other legal professionals in finding previous court cases. FindLaw[14], WestLaw[15] and BAILII[16] are the some systems which are capable of retrieving a large amount of legal cases. While these systems developed with the intention helping the legal officials be more effective in their profession, they do not provide a well-structured representation of legal information. Our system will help reduce the time and effort a lawyer has to put in to find advantageous and disadvantageous facts relevant to legal parties. Moreover, aspect-based sentiment analysis is a research area which has rarely been touched in the legal domain. Hence, it is clear that developing a way to perform party-based sentiment analysis in the legal domain holds significant demand and importance.

1.3 Problem Statement

Lawyers and legal officers have to spend their valuable time searching for required information describing legal opinions from the previous court cases and identifying arguments, counter-arguments, facts, and evidence using these cases to proceed with new legal cases [12, 13]. When analyzing the previous court cases, identifying the advantageous and disadvantageous arguments or facts in a court case with respect to each legal party and detecting contradictory statements in a court case can be considered as significant tasks. In order to perform these tasks, analyzing the sentiment of sentences in a legal document holds significant importance when automating the legal information extraction process. When a single sentence in a legal document involves multiple legal parties, the sentiment of the sentence should be addressed relevant to each party. In the legal domain, sentiment analysis has been carried out only at a phrase/sentence level. Phrase level sentiment analysis is not sufficient to figure out the sentiment of a sentence for a particular party.

Example 1

- Sentence 1.1: *Sentence 1.1: In 2008, federal officials received a tip from a confidential informant that Lee had sold the informant ecstasy and marijuana.*

For example, consider Example 1 taken from Lee v. United States [17] which consists of two legal parties; petitioner Lee and government representing federal officials. The sentence in the example mentioned that officials received a tip about Lee’s illegal works. When considering the context of this sentence, we can clearly see that the context has a positive sentiment regarding government (here federal officials are taken as a part of the government) and negative sentiment regarding person Lee. The phrase-level sentiment of the sentence is only acquired using the existing information extraction systems for the legal domain [7, 18]. For example, consider Example 1; the parse tree and sentiment analysis of the sentence is shown in Figure 1. Here the parse tree is only showing the overall sentiment of the sentence but it doesn’t provide a sentiment for a particular party. There are many studies for the aspect-based sentiment analysis in other domains except the legal domain. The objective of this research is to introduce a novel approach to analyse the party based sentiment for legal opinion text which means it detects sentiment

level (positive, negative, neutral) of sentences in legal documents with respect to each legal party mentioned in the document.

1.4 Objectives of the Research

The main objective of this research is to develop a methodology to perform Party-Based Sentiment Analysis in legal opinion texts efficiently. This research will introduce the paradigm of Aspect Based Sentiment Analysis to the Legal Domain. The sub-objectives of the research can be listed as follows:

- **Create a dataset for evaluations**

To the best of our knowledge, there is no publicly available dataset for the aspect (party) based sentiment analysis for legal opinion texts. Hence, we have to prepare a suitable dataset in order to train the models and to conduct the experiments.

- **Design and develop an efficient approach**

- Detect sentiment value for each member of parties (petitioner and defendant) included in a sentence
- Provide an overall sentiment for main two parties

Figure 1.1 illustrates the objective methodology mapping or the outline of the research.

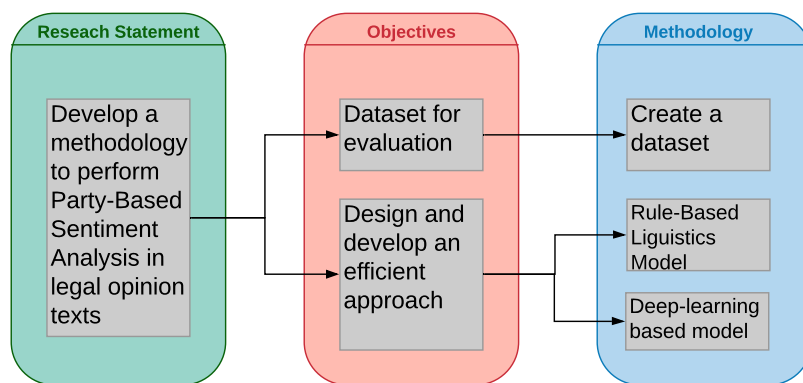


Figure 1.1: Research Outline

Chapter 2

Literature Review

2.1 Legal Information Extraction

Legal Domain can be considered as a domain which can be made more productive and effective with the introduction of Artificial Intelligence. This fact is evidenced by the number of research that have been conducted over many years on automatic legal information extraction. Past literature related to the legal domain covers the areas of information organization [2, 19, 20], information extraction [21], and information retrieval [22]. Few studies have been published on tasks of legal jargon embeddings in vector spaces [21, 23]. However, sentiment analysis in the legal domain can be considered as an area which was not touched until Gamage et al. developed an automatic sentiment annotator for the legal domain [7] in 2018. Their approach is based on transfer learning and it is a domain adaptation task which uses the Stanford Sentiment Annotator [18] as the base model. Stanford Sentiment Annotator is based on a Recursive Neural Tensor Network (RNTN) model [18]. The Stanford Model [18] has been widely used due to its ability to consider the compositionality between textual units when determining the sentiment of the text. Socher et al claim that the Stanford Model outperformed other models when it comes to detecting negation relationships. However, bias towards the movie domain has been identified as the major drawback of Stanford Sentiment Annotator when it is being used in the legal domain [7]. Moreover, Ratnayaka et al have proposed methodologies to identify relationships among sentences in the legal

documents [12, 24]. They have demonstrated that sentiment analysis can be used to identify sentences that provide different opinions on the same topic (contradictory opinions) within a legal opinion text.

2.2 Benchmark Datasets for ABSA tasks in different domains

In recent years, a large amount of research effort has been devoted to Aspect-Based Sentiment Analysis over many different domains. There are many publicly available datasets such as movie reviews, products and restaurants to evaluate ABSA tasks. Ganu et al. [25] created a dataset of restaurant reviews for the task of improving rating predictions. The dataset was annotated on six aspect categories with overall sentiment polarity. The two major steps of Aspect-Based Sentiment Analysis are aspect term extraction and sentiment classification. Hence, this work can not be identified as a fully completed ABSA dataset, as it does not use related sentiment polarity for an aspect. It only includes the aspect category. The semEval 2014 [26], an international workshop in the domain of Natural Language Processing (NLP), introduced datasets that are annotated with four fields as aspect term, aspect term polarity, aspect category and aspect category polarity of each sentence. They introduced datasets on restaurant and laptop domains. As a continuation and improvement from SemEval 2014, SemEval 2015 [27] introduced datasets giving a new definition to the aspect category as a combination of the type of entity and type of attribute for restaurants, hotels and laptop domains. Moreover, SemEval 2016 [10] introduced Multilingual datasets, a total of 39 datasets from 7 domains and 8 languages for the ABSA task. It included datasets for the domains of restaurants, laptops, hotels, mobile phones, museums, digital cameras, and telecommunication in English, Spanish, Arabic, Chinese, Turkish, French, Dutch, and Russian languages. Smadi et al. [28] created a human-annotated Arabic language book review dataset (HAAD). In the annotation process, they used 14 categories along with 4 polarity types as positive, negative, neutral and conflict. For the ABSA of IT product reviews, Tamchyna et al. [29] introduced a dataset in the Czech language. As there were no existing studies on ABSA of Bangla text Rahman et al. [30] create two datasets on cricket and restaurant domains. To the best of our knowledge, there is no publicly available

dataset for the legal domain in the field of ABSA.

2.3 Earlier approaches to ABSA tasks

From the early 2000s, Sentiment analysis has attracted a great deal of attention in natural language processing. That is because of the explosive growth of social media, which produces large amounts of opinionated data. Consequently, now it is an active research area in data mining, Web mining, and information retrieval. Since the term “opinion” is critical to several activities, the interest in sentiment analysis extends to many domains. Previous works on aspect-based sentiment analysis focus mainly on training sentiment classifiers based on bag-of-words feature extraction and features on sentiment lexicons [31]. Conventional representation methods consist of statistic-based methods [32] and rule-based [33] methods. In past literature there can be seen statistical-based approaches which use MaxEnt-LDA [34] and SVM [32]. These techniques mainly depend on labor-intensive feature engineering and excess linguistic resources.

2.4 Deep Learning for Aspect Based Sentiment Analysis

Deep learning also can be considered as a machine learning approach that is based on artificial neural networks with representation learning where neural networks consist of several layers and input data is analyzed and classified in those layers. In these networks, it follows a process named backpropagation which means the output of one layer is fed into the next layer of the neural network [35]. As an example of deep learning, deep neural networks (DNN) can be introduced. A deep neural network is a subfield of artificial neural networks that consists of multiple layers of neurons or connected processors where each neuron is activated by environmental sensors or else by the outputs from the previous neurons [36]. In general, deep learning approaches divide the used datasets into three sub parts named train, test, and validation datasets as in machine learning approaches [37]. Lecun et al. [35] mentioned that the training process is nothing more than a transition of input data into vector scores without considering the input type.

2.4.1 Deep Neural Network Layers

If we consider deep neural networks in the area of natural language processing, deep learning methods can be segmented into three modules such as dense word embeddings, and output [37]. If we consider word embeddings first, they represent words as d -dimensional vector spaces and encode them as dense numerical vectors [38]. Hidden layers come up with several architectures and each and every hidden layer consists of stacked multiple neurons such that non-linear outputs are generated [39]. Finally, the distributed probability of classes or labels are represented by the third module, output. Lecun et al. [35] proposed an approach that distributed represented DNNs have capability of generalizing new combinations of learned features rather than rely only on what has been learned from the training process. Rojas-Barahona [38] argue that compared to typical machine-learning approaches, deep-neural-networks try to learn representations or features automatically. Further, according to Araque et al. [40], deep neural networks do not need a lot of feature engineering as traditional machine learning models. DNN models can be applied to perform aspect based sentiment analysis (ABSA) tasks. Hence, in the below sections, we reviewed some related models such as convolution neural networks (CNN) models, resurrect neural networks (RNN), and so on.

2.4.2 Convolutional Neural Network Model (CNN)

Reviewing past literature, CNN shows success in many natural language processing approaches including sentiment analysis. The ability of extracting essential n -gram features from an input sentence and formation of “informative latent semantic representation” of a given input sentence can be identified as the main strength of architecture of CNN [41]. Compared to non-linear CNN models, linear models are able to better fit data as Conditional Random Field and do not need labor-intensive feature engineering [42]. The study of Toh et al. [43] used deep CNN to determine the aspect category and according to the authors, the performance of the system highly depends on CNN features. Xue et al. used two CNNs to aspect term extraction and sentiment polarity detection for multilingual aspect-based sentiment analysis [44]. Gu et al. [45] introduces a cascaded CNN model for the same task with two CNN levels. The two CNNs for aspect extraction and sentiment detection are organized in a cascaded manner. The

proposed methodology minimizes the feature engineering and specially cascaded CNN model reduced the elapsed time with comparison to SVM models.

2.4.3 Recurrent Neural Network (RNN) and Long-short Term Memory (LSTM) Models

Recently RNN based approaches have shown better results on aspect-based sentiment classification tasks due to its ability to capture the sequential nature of languages. Compared to CNN models, the flexible computation phases of RNN models allow the ability to collect language con-text dependencies and also model different text lengths [46]. While CNN has specific parameters at its every layer, RNN has the same set of parameters and hence decreases the number of parameters required to learn. Further RNN shows promising results compared to CNN in computing sequential data due to the ability to have memory on previous computations. However, a major drawback of simple RNN is having a vanishing gradient problem [47]. Hence, to overcome this limitation Long-short term memory (LSTM) networks and gated recurrent units (GRU) have been introduced. Basic LSTM can not be directly used to ABSA as it can not specially focus on the target sentiment when modeling the semantic representation of a sentence. Hence Tang et al. [48] proposed improved basic LSTM and proposed two LSTM networks by considering the target word. TD-LSTM uses two LSTM networks in order to extract important information from the left and right sides of the target. After that, it concatenates the hidden states of two LSTM networks and finally classifies the polarity labels by feeding them to a softmax layer. However, this model cannot incorporate the interactions between context and the aspect target. TC-LSTM incorporates semantic relatedness of the context words and the target by extending TD-LSTM with a target connection component. Although it improves the LSTM architecture it is often impossible to distinguish between various sentiment polarities at a fine-grained level.

2.4.4 Attention Mechanism for Classification

The main aim of the ABSA is to classify the sentiment levels of context with respect to each aspect terms included text. For that, it is important to have a proper methodology to represent the relationships among the aspect terms and the context. The Main draw-

back of the traditional RNN network is encoding irrelevant features for information-rich inputs [49]. Every word in a sentence is not equally important for any task. So as a solution researchers employed an attention mechanism to learn the model of the parts which should have given special focus. This attention mechanism computes attention weights of each lower level for upper-level representation, gets the sum of calculated weighted vectors. Wang et al. [50] proposed AT-LSTM and ATAE-LSTM incorporating attention mechanisms to model relationships between aspect and context. AT-LSTM incorporates the relationship between the word and the target only at the attention layer while the ATAE-LSTM network incorporates the relationship at both the input and attention layer. Each input word vector is appended with aspect embedding and attention mechanism is used along with LSTM. In order to better understand target information, Cheng et al. [51] introduced the HiErarchical ATtention (HEAT) network with sentiment attention and aspect attention. The sentiment features and aspect details are captured with respect to the extracted aspect details and target aspect. The single attention mechanism is not sufficient to get all the important information of long sentences and it is difficult to model the relationship between context and targets when they are scattered over long distances. As a solution to this, researchers have adopted multiple attention mechanisms for the better coordination of targets and contexts. Chen et al. [49] designed the RAM model by adopting multiple attentions to extract important information from memory. This model first utilizes a bidirectional LSTM to produce memory from the input and weighted the memory slices according to their relative distances to the aspect target. Then multiple attentions on position-weighted memory are combined non-linearly with a RNN to predict the final sentiment polarity. IAN is proposed by Ma et al. [52] utilizes a bidirectional attention mechanism and learns the attention for the contexts and targets separately via interactive learning. By improving the effectiveness of attention mechanism He et al. [53] proposed a methodology for target representation that effectively expresses the semantic meaning of the aspect target by incorporating an attention mechanism that integrates syntactic knowledge obtained from a dependency parser. Li et al. [54] proposed a Hierarchical Attention based Position-aware Network by introducing position embeddings to learn the position-aware representations of sentences and generate the target-specific representations of contextual words.

2.4.5 Memory Network for Aspect Sentiment Classification

A growing body of literature shows that researchers have devoted their attention to memory networks and could obtain better results in question and answer. Deep memory networks were developed by Tang et al. [46] for aspect-based sentiment analysis by capturing the information related to context words. To select the relevant information towards the target, an attention mechanism was employed. This utilizes a deep memory network and multiple attention is paid to word embeddings. For sentiment prediction, the output of the last attention is fed to the softmax layer. Here the results are different, not combined. Their proposed approach has higher performance due to the usage of multiple computation layers in memory. However building an adequate attention value and aspect- word representation was still a challenging task. To incorporate the related information on neighboring aspects for sentiment classification of the target aspect Majumder et al. [55] proposed IRAM model.

2.4.6 Bidirectional Encoder Representations from Transformers [BERT]

In recent years there has been growing interest in BERT based models for NLP tasks. Sentiment analysis also achieved great improvements with BERT models. Bert is a language representation model which can be fine-tuned easily, by adding just one output layer. BERT can be used to create the best models for a wide range of tasks including sentiment analysis. It can bring a considerable improvement in classification performance under extremely small data in the task of text multi-classification. Bert could be used for our sentiment analysis task by using the fine tuned model of that [56]. By analysing the past literature, we could identify several models which have used BERT pre-trained models to perform state of the art results. Devlin et al. [56] have proposed an approach that uses a pre-trained BERT model where the BERT model is finetuned by adding just one output layer. Here, fine tuning can be done without modifying the architecture specifically. AEN (Attentional Encoder Network) [46] is another BERT based model that tries to eliminate repetition and takes attention based encoders into consideration to represent between the target and context. Further in this model, it introduces a label smoothing regularization mechanism to reduce the issue of the label unreliability. Moreover, Zeng et al. [57] proposed a model named BERT-LCF which takes both

local context and global context of a sentence into consideration. The key fact of this architecture is that it considers the local context has more important information than the global context. Hence, it can be concluded that using BERT pre-trained models have improved the performance of models related to the area of aspect-based sentiment analysis.

2.4.7 Graph Convolutional Network [GCN]

GCN is successful in handling graph data that contains rich relationship information. Although attention-based models have shown promising results over many ABSA tasks, they are not adequate to catch syntactic dependency between aspect and the context words within the sentence[58]. The important feature of GCN is it has the ability to draw syntactically related terms to the target aspect and manipulate multi-word associations and syntactical knowledge in long-range with GCN layers [59]. Zhang et al proposed ASGCN adopting GCN for ABSA. The authors conclude GCN improves overall efficiency by exploiting both syntactic knowledge and long-range word dependency. Zhao et al. [60] introduced the SDGCN model with the aim of modeling sentiment dependencies within a sentence among different target aspects. Their approach first implements a bidirectional attention mechanism with position encoding in order to model aspect specific representation for the relation of context word with its aspects. Then, to capture the dependencies over various aspects within the sentence GCN was employed. They claim that dependencies among various target aspects within a sentence highly influence the ABSA tasks and the GCN module can improve the effectiveness of tasks.

2.5 Evaluation Techniques

There are multiple evaluation methods that are generally used for the evaluation of the accuracy of the predicted values by the classifier. In order to perform this task, the original dataset divides into a test and training set. Then the classifier is trained using the training data set and evaluated using a testing data set which was held out. When observing the past literature related to the area of Aspect Based Sentiment Analysis (ABSA), efficiency of the classification model has been measured by using precision, recall, f-score, and accuracy. Corresponding evaluation measurement equations are shown

below.

TP - True Positive, TN - True Negative, FP - False Positive, FN - False Negative

$$Precision = \frac{TP}{TP + FP} \quad (2.1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2.2)$$

$$F1score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2.3)$$

2.6 Discussion

A summarization of some related work in the area of aspect-based sentiment analysis over different domains is described in the below Table 2.1. The classification accuracy of a model might vary based on the techniques used in the model. Different models have used different techniques to enhance the performance of the prediction of results. Hence, in this section we review some models with their performances and limitations. Description of each model is provided in further sections.

Studies related to legal information extraction emphasize the importance of sentiment analysis in the legal domain and describe how sentiment analysis can facilitate other legal information tasks such as contradiction detection [11] and discourse analysis [5]. Also, it can be seen that Aspect Based Sentiment Analysis has been discussed in other domains over many years. When these factors are taken into account, it is evident that carrying out sentiment analysis at aspect level has many advantages over sentence level/document level sentiment analysis.

Table 2.1: Review of Existing Models

Related Work	Dataset	Limitations
TD-LSTM [48]	Tweets	Can arise a problem after the model takes a sentiment feature that is located much away from the target, where it is needed to be propagated by word to the target. Hence, there is a probability of losing that feature.
ATAE-LSTM [50]	SemEval 2014	Different aspects are input separately i.e. if there are more than one aspect in a sentence, that should be processed separately for each aspect.
IAN[52]	SemEval 2014	It is difficult to do memory optimization because encoded hidden layers should be kept simultaneously in memory in order to perform attention mechanisms.
MemNet [46]	SemEval 2014	Word vectors sequence is used as memory that might not be able to synthesize phrases like original sentence features.
RAM [49]	SemEval 2014	Always the memory is treated with a fixed number of attention, that will make it unreasonable in several cases.
Bert-spc [56]	SemEval 2014	Different aspects are input separately i.e. if there are more than one aspect in a sentence, that should be processed separately for each aspect.

Chapter 3

Research Methodology

The ultimate goal of legal texts analysis is to automatically determine which legal party has an advantage over a particular argument. Party-based sentiment analysis will play a key role when it comes to developing such methodologies. Building a complete system to achieve this ultimate goal requires a lot of research. In this study, our main concern is to research on identifying a fine-grained sentiment analysis technique to achieve this task.

As the starting point, we have to create a dataset for our research as there are no publicly available dataset for Aspect (Party) Based Sentiment Analysis for legal opinion texts. It is not much effective to use a dataset created for a particular domain to evaluate ABSA tasks in a different domain. This implies that domain-specific datasets need to be created to evaluate ABSA tasks. Therefore, creating a publicly available dataset for the research of ABSA for the legal domain can be considered as a task with significant importance. In this study, we introduce a manually annotated legal opinion text dataset (SigmaLaw-ABSA) intended towards facilitating researchers for ABSA tasks in the legal domain. SigmaLaw-ABSA consists of legal opinion texts in the English language which have been annotated by human judge and the dataset is prepared to support various research tasks, including aspect extraction, polarity detection, party identification, and party-polarity detection.

As the second step of the research, we developed a rule-based model to perform party-

based sentiment analysis in legal opinion texts in order to determine sentiment levels with respect to a particular legal party. The complexity of the structure of the sentences in legal texts has prompted to develop first a rule-based approach over other approaches. The rule-based model is primarily built around a phrase-level sentiment annotator [7] which is specifically developed for the legal domain. We use rationally built rules with the annotator to make the system aspect-specific and also to integrate domain knowledge. Along with the annotator and rules, the Stanford NLP library [61] is used to analyse the grammatical relationships of words in a sentence. Developing this model, we got a much understanding on the legal domain, complex structure of the legal texts and many other benefits for the future works. However, it significantly depends on the phrase-level sentiment annotator and also manually created rules may not cover all the sentence patterns. That has been the main drawback of the system.

To address the drawback in the Rule-based approach, we have we have decided to incorporate deep learning-based methodologies to our study. As the first step of designing a deep learning architecture, we did research on different ABSA architectures in different domains. While gaining the knowledge we analyzed the advantages and drawbacks of those architectures. As the sentences in legal documents are often long and have a complex semantic structure, the existing model architectures do not perform well for the legal domain. Hence, we identified several techniques in the literature that can be optimized and applied to our domain aggregate our research problem into the deep learning approach specifically for the legal domain. The proposed model first uses a pre-trained BERT model further fine-tuned for a legal corpus for a strong word embedding. Then model employs the position aware attention mechanism to capture the critical parts of the sentence relevant to aspects with incorporating position information using the dependency tree. Because multiple legal party members are involved in a single sentence, GCN is employed over the attention mechanism to model the inter-dependencies between members. Section 3.4 describes the implementation details of deep learning-based model.

The overall structure of the methodology is organized as follows. Section 3.1 describes creation of dataset (SigmaLaw-ABSA). Section 3.2 describes the implementation process of rule-based approach and section 3.3 discusses the research on existing deep-learning ABSA architectures. Lastly, section 3.4 describes the proposed deep-learning

ABSA architecture specifically for legal domain.

3.1 Creation of Dataset (SigmaLaw-ABSA)

There are many studies for the Aspect-Based Sentiment Analysis in other domains except the legal domain. Common publicly accessible datasets such as restaurants, hotels, movies, and product reviews typically satisfy the requirements of the researchers to perform their studies in the field of ABSA . To the best of our knowledge, there is no publicly available dataset for the legal domain in the field of ABSA so far made. So this implies the need to create a dataset for the legal domain ABSA tasks.

There is no publicly available dataset for the legal domain in the field of ABSA. Consequently, no existing system or research methodologies which intend to effectively analyse the sentiment of legal opinion texts with respect to each legal party in a court case. Therefore, this study aims to address this research gap by introducing the dataset *SigmaLaw-ABSA* by extracting aspect categories and corresponding polarities of legal texts from the United States supreme court cases.

3.1.1 Data Collection

This section describes the process of collecting data from legal cases. The court cases were fetched from the *SigmaLaw - Large Legal Text Corpus and Word Embeddings dataset*¹ which contains a large legal data text corpus. The legal data corpus consists of 39,155 legal cases including 22,776 taken from the United States Supreme Court. For the data collection process, about 2000 sentences were gathered to annotate and court cases were selected without targeting any specific category. Here we consider 2 types of sentences when preparing the dataset;

- Original Sentence extracted from the court case
- Meaningful sub-sentences extracted from the original sentence

Reviews, dialogues, or informal text normally contain shorter sentences where whole sentences can be used in the dataset without any concern. Nevertheless, more com-

plex sources, such as legal documents contain longer sentences and more subordinate clauses [62]. The subordinate clauses of a sentence can be defined as sub-sentences of that sentence [24]. Sub-sentences can be generated by splitting the sentence using subordinate clauses. In the process of generating sub-sentences, the constituency parser[61] in the Stanford CoreNLP² library is used to achieve this task due to the proven performance of the parser [24].

The Constituency Parser returns the parse tree of the sentence after annotation process. The parse tree of the sentence will be built according to the grammatical structure of the sentence. Figure 3.1 illustrates the parse tree for example 1 given above. In the parse tree, subordinating conjunctions label with the SBAR tag. The sub-sentences splitting occurs by recognizing associated terms with the SBAR tag. According to the defined annotation process, *"In 2008, federal officials received a tip from a confidential informant"* and *"Lee had sold the informant ecstasy and marijuana"* are the sub-sentences of the Example 1 which splitting by identifying the associated term *that* with the SBAR tag.

There are some cases which give meaningless sub-sentences after splitting by subordinating conjunctions. In order to extract only meaningful sub-sentences, each sub-sentence is examined by the authors. In this process, all meaningless sub-sentences are eliminated, considering a sub-sentence without a subject as meaningless.

3.1.2 Data Annotation

Reviewing the past literature we found that, there is no publicly accessible tool that assists efficient legal text annotation. The authors, a group of undergraduates and post-graduates in the faculty of law, University of Colombo Sri Lanka were involved in the data annotation process. The annotated dataset contains entities of different parties, their polarities, aspect category (Petitioners and defendants), and the category polarities. The annotators' task was to figure out the aspect categories of each party and give a polarity label for each of them. In this process, we have used polarity types as *Negative*, *Positive*, and *Neutral*.

Three human annotators analysed the collected data. As a result, a rating process is

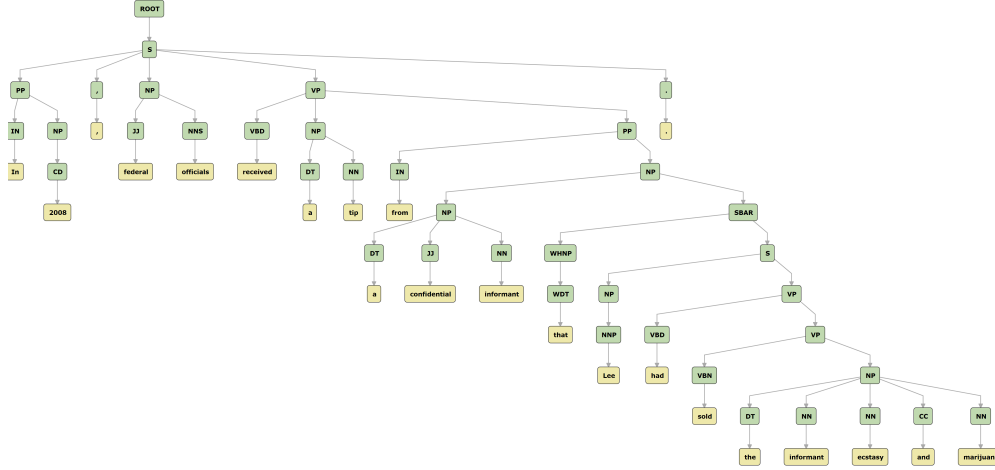


Figure 3.1: Constituency Parser Tree for Example1

needed to determine the final sentiment value. Fleiss' kappa [63] was selected to measure for assessing the reliability of agreement between raters (inter-rater reliability). We have obtained the kappa value of 0.59 which belongs to the moderate agreement level and located at the narrow margin of the substantial agreement (0.61). In this dataset, each sentence is annotated under the following fields.

- **Party:** Petitioner and defendant are the main parties of a court case and each party may consist of several members. Therefore, in this section, we use list of list data structure and a member who represents the petitioner party or the defendant party goes under the petitioner list (first inner list) or defendant list (second inner list) respectively. Further, when preparing the dataset we considered all required human entity pronouns.
- **Sentiment:** Party based sentiment polarity values are annotated under this field. This field also uses a list of list data structure with the size equal to the list of lists used in the party field. Polarity values for all the parties under petitioner and defendant are listed with respect to listed parties. In the annotation process used polarity types: *Negative*, *Positive*, *Neutral* as -1, +1, 0, respectively.
- **Overall sentiment:** This section is used to indicate the overall sentiment of the sentences without considering the legal parties related to it. All the sentences are given one of the sentiment levels: *Positive*, *Negative*, *Neutral* as +1, -1, 0 respectively.

3.2 Rule-Based Approach

3.2.1 Introduction

The ultimate goal of legal texts analysis is to automatically determine which legal party has an advantage over a particular argument. Party-based sentiment analysis will play a key role when it comes to developing such methodologies. In this study, we propose a rule-based sentiment analysis approach to determine the sentiment value of each party in legal opinion texts[64]. The complexity of the structure of the sentences in legal texts has prompted to select the rule-based approach over other approaches. Languages being used are sometimes mixed up with several origins (i.e English, Latin, etc.) and in certain cases, the meaning of the words and context differs with domain interpretations [22]. Considering these facts, rationally built rules hold significant importance as it can be considered as an easier way to integrate domain specific knowledge with NLP tasks.

This approach is primarily built around a phrase-level sentiment annotator [7] which is specifically developed for the legal domain. We use rationally built rules with the annotator to make the system aspect-specific and also to integrate domain knowledge. Along with the annotator and rules, the Stanford NLP library [61] is used to analyse the grammatical relationships of words in a sentence. The whole sentence is divided into phrases that are smaller instances of the same sentence and process each phrase and allocate the sentiment for each party mentioned in those phrases. After that, we combine the output of phrases into the main output for the input sentence. The overall system developed in this study assigns sentiment scores for input sentences with respect to each legal party and provides the sentiment levels as Negative or Non-negative. Figure 3.2 illustrates the overall methodology of our approach and the following subsections outline the steps taken in this study.

3.2.2 Co-reference Resolution

Co-reference resolution is an important step that deals with language understanding tasks in information extraction. Co-reference resolution [65, 66] is the function of identifying all expressions in a text that corresponds to the same real-world object. In a court case, there can be some sentences which have used pronouns to mention members

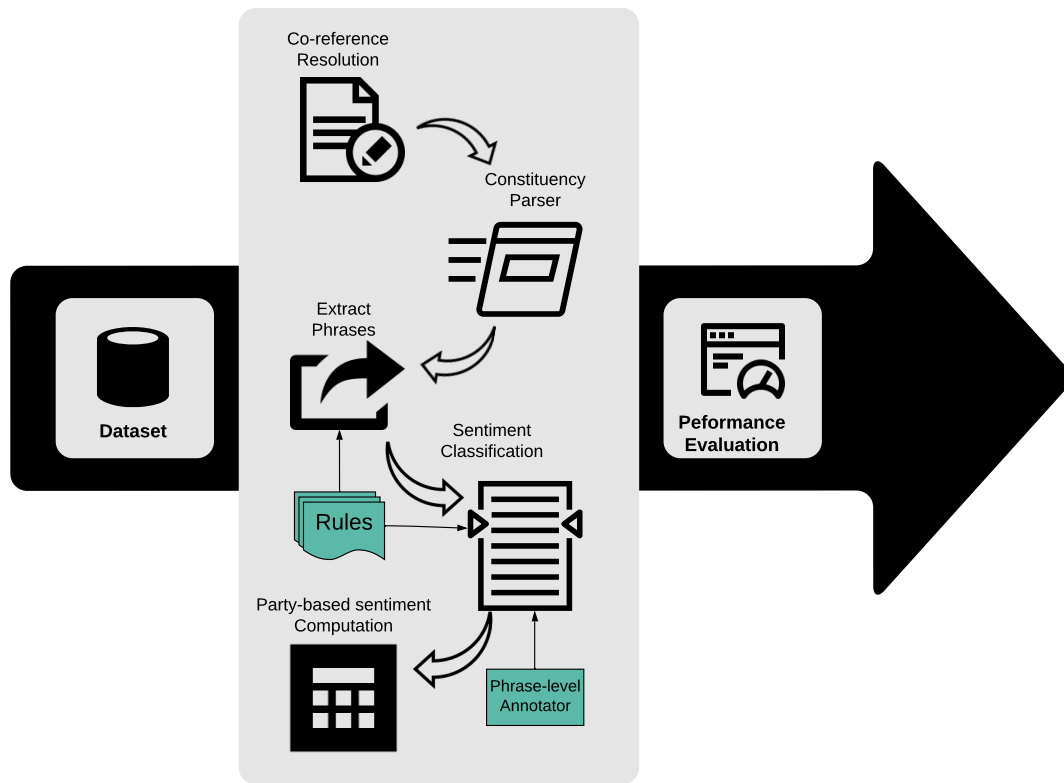


Figure 3.2: Rule-Based Model Architecture

of a particular party. Before calculating the sentiment polarity values corresponding to different parties, it is important to find out how different aspects (entities of parties) are demonstrated in the sentence.

Example 2

- Sentence 2.1: *During the plea process, Lee repeatedly asked his attorney whether he would face deportation; his attorney assured him that he would not be deported as a result of pleading guilty.*

Consider Example 3 taken from Lee v. United States [17] which consists of two legal parties; Lee and attorney. In this sentence used different pronouns his, him and he to identify the same person Lee. In certain cases, the sentences contain only pronouns without the name of the person or organization. In such instances, it is not possible for the system to identify for which party does that the person or organization provided by the pronoun belongs. Therefore, in such cases, we provide a cluster of sentences in

a court case as the input after checking whether all pronouns have their origin entity. The cluster is made up of adjacent sentences of the considered sentence (from the dataset). As a future work, we plan to modify the whole dataset using co-referencing, so this problem will be automatically solved. The goal of using co-reference resolution for our approach is to identify sets of co-referring expressions in a sentence of court cases referring to the same aspect (party) and cluster them in order to get overall polarity value towards a particular party. In our approach we used Stanford Co-reference Resolution [61] model due to its proven performance over many NLP tasks. As shown in Figure 3.3, the Stanford co-reference model has identified all entities mentioned in the sentence and the pronouns such as his, him and he as Lee.

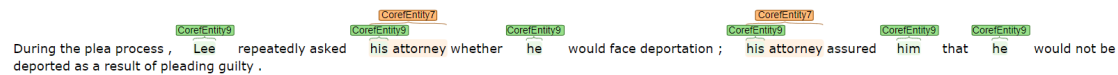


Figure 3.3: Co-reference Resolution

3.2.3 Generating sub sentences

We used the constituency parser [61] of Stanford CoreNLP³ for the process of generating sub sentences. Legal documents usually contain longer sentences with many subordinate sentences. When a sentence has several subordinate sentences becomes even more complex for a machine to understand [24]. Hence the sentence is split using subordinating conjunctions and subordinate clauses are quoted as sub sentences. Once the sentence is annotated using Stanford CoreNLP Constituency Parser, parse tree of the sentence will be built according to the grammatical structure of the sentence. In here *SBAR* tags are used to label the subordinate conjunctions. Therefore the splitting of sentences in to sub sentences occur by capturing the related words with the *SBAR* tag.

Example 4 sentence 4.1 shows the complete sentence and sentence 4.2 and 4.3 are the sub sentences split by using the constituency parser. Figure 3.4 illustrates the constituency parse tree of the sentence 4.1 which is generated by Stanford CoreNLP parser.

- Sentence 3.1: *In 2008, a confidential informant told federal officials that Lee had sold the informant approximately 200 ecstasy pills and two ounces of hydroponic marijuana over the course of eight years.*
- Sentence 3.2: *In 2008, a confidential informant told federal officials*
- Sentence 3.3: *Lee had sold the informant approximately 200 ecstasy pills and two ounces of hydroponic marijuana over the course of eight years.*

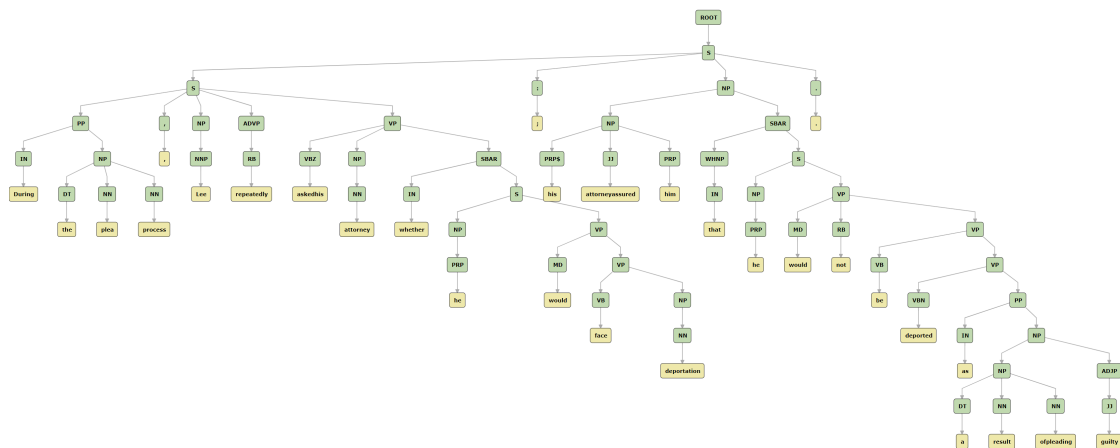


Figure 3.4: Constituency Parser Tree for a sentence

After dividing the sentence into sub sentences, each sub sentence is processed separately. The grammatical structure of sub-sentences is of utmost importance in the process of sentiment detection. There are mainly two types of parsing in NLP; dependency and constituency. Dependency parsing focuses on relations between words while constituency parsing focuses on identifying phrases and their recursive structure [67]. As the phrase level sentiment annotator is used for the sentiment allocation, constituency parser is the most appropriate parser for this process. It outputs the constituency-based parse tree for a given sentence and the syntactic structure of the parse tree is based on the phrase structure grammar or context free grammar [68]. Applying different linguistics rules to parse tree, the phrases of each sub sentence are extracted.

35

and a lexicon of words and symbols [69]. Even though there are many rules defined in the phrase structure grammar, we selected the most important rules which can be used in our process. Mainly the sentence (S) is divided into two phrases (Noun Phrase(NP), Verb Phrase(VP)) [69]. The figure 3.5 shows a rule which states that a sentence can include a noun phrase followed by a verb phrase:

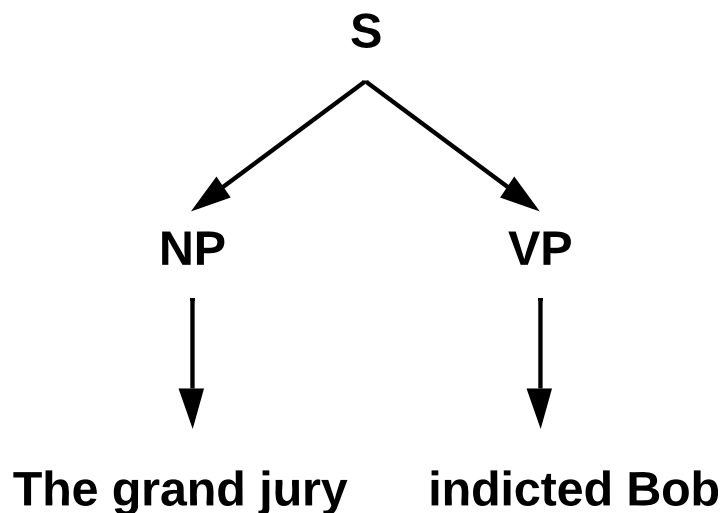


Figure 3.5: Main phrases of a sentence

A verb phrase contains a verb accompanied by a variety of other things; for example, as shown in the figures, one type of verb phrase contains a verb followed by a noun phrase (figure 3.6) Or the verb phrase may be followed by a noun phrase with a prepositional phrase (figure 3.7):

The basic pattern of a simple sentence is subject-verb-object-adverbial [69]. The subject is typically a noun phrase that refers to a person, place, or thing. The verb identifies an action or a state of being. An object is given an action and typically follows a verb [70]. Having regard to the structure, we have come up with a principal rule which considers that the sentiment of the verb phrase totally affects the object. Depending on this rule, only the noun phrase and verb phrase of a sub sentence are extracted.

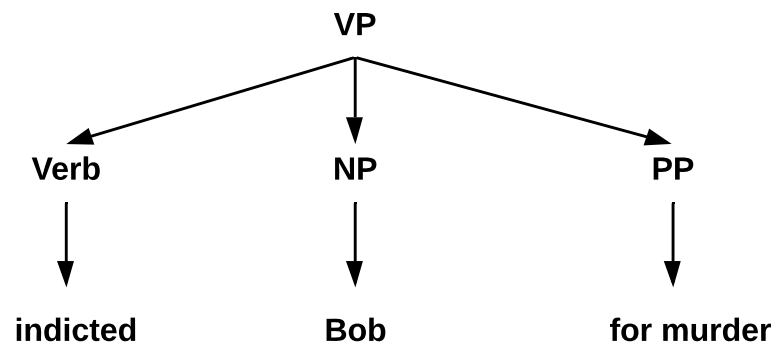


Figure 3.6: The structure of verb phrase: type 01

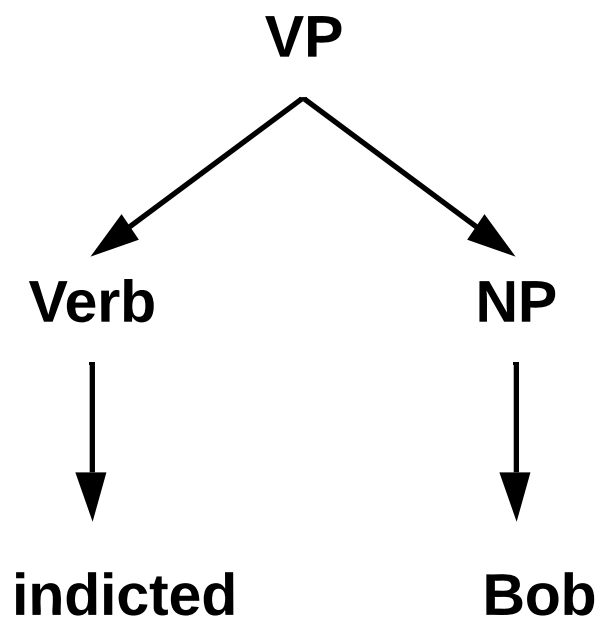


Figure 3.7: The structure of verb phrase: type 02

Table 3.1: Sentiment Classes Mapping

Outputs of Stanford Sentiment Annotator	Outputs of Phrase Level Sentiment Annotator for legal domain [7]
Very negative, negative	negative
Neutral, Positive, Very Positive	Non negative

3.2.5 Phrase-level Sentiment Annotator

For our study we used the phrase level sentiment annotator for the legal texts developed by [7] in 2018. Their model has been developed specifically for the legal domain based on transfer learning, and it can be considered as a domain adaptation task which uses the Stanford Sentiment Annotator [18] as the base model. In their approach Sentiment of a phrase is classified into one of negative or non-negative classes without considering any specific party involved in the sentence. In their study, through the proposed methodology they could achieve 6% improvement of the sentiment classification in the legal domain when compared with the accuracy of the Stanford model. Stanford Sentiment Annotator is based on a Recursive Neural Tensor Network (RNTN) model [18]. The Stanford Model has been widely used in different studies related to sentiment analysis. Socher et al. [18] claim that the Stanford Model outperformed other models when it comes to detecting negation relationships. In the Stanford model, they used 5 different classes as very negative, negative, neutral, positive and very positive for detecting polarity levels. However, in the legal domain, the fundamental requirement is to decide whether a given sentence of a court case is opposing plaintiff's/petitioner's claim or not [7]. So this model [7] mapped the five sentiment polarity classes provided in the Stanford Model [18] to the two classes as negative and non-negative. The mapping used for [7] is shown in Table 3.1. In our approach we also selected the two output classes used by [7] as it is more relevant to the legal domain.

3.2.6 Sentiment Classification

The sentiment allocation is processed using phrase-level sentiment annotator [7] and rationally built rules. The input sentence is divided into phrases using the above NLP

tools and rules as shown in figure 3.8. This whole process is a sort of divide and conquer model. The complete sentence is divided into phrases(noun and verb phrases) that are smaller instances of the same sentence and process each phrase and allocate the sentiment for each party mentioned in those phrases. In sentiment allocation process, first we take the contextual sentiment score and sentiment level of each verb phrases using the phrase level annotator and assign those values to the party mentioned in the verb phrase and assign opposite values to the noun phrase. Here, we assume that the sentiment of the verb phrase totally affects the object(explained in extracting phrases section) and that one phrase can have one party member. After that, we can combine the output of phrases into the main output for the input sentence.

After detecting the contextual sentiment scores of each and every small phrase of sub sentences, the next step is to calculate the sentiment value of the entire sentence in relation to each party mentioned in the sentence. When we get the sentence as a whole, one phrase can have one party, multiple phrases can have the same party or multiple phrases can have different parties. Therefore when detecting the final sentiment we grouped each aspect (party) and their sentiment scores. When assigning the final sentiment for a sentence we followed the sentiment score averaging method.

In this approach we determined the sentiment polarity scores of each small phrase of sentence relevant to each party. In here we used sentiment scores of sub phrases as sentiment weights of those phrases for our calculations. For the detection of final sentiment for the whole sentence, we got the average of each negative and non-negative phrases and assigned the highest averaged sentiment as the final sentiment for the relevant entities of parties. In certain cases, more than one member from a legal party (here we consider petitioner and defendant as main two parties) may be included in a single sentence. After getting the sentiment levels for the members of each legal party mentioned in the sentence we use the same sentiment score averaging method to obtain the sentiment level for the main two legal parties(petitioner and defendant).

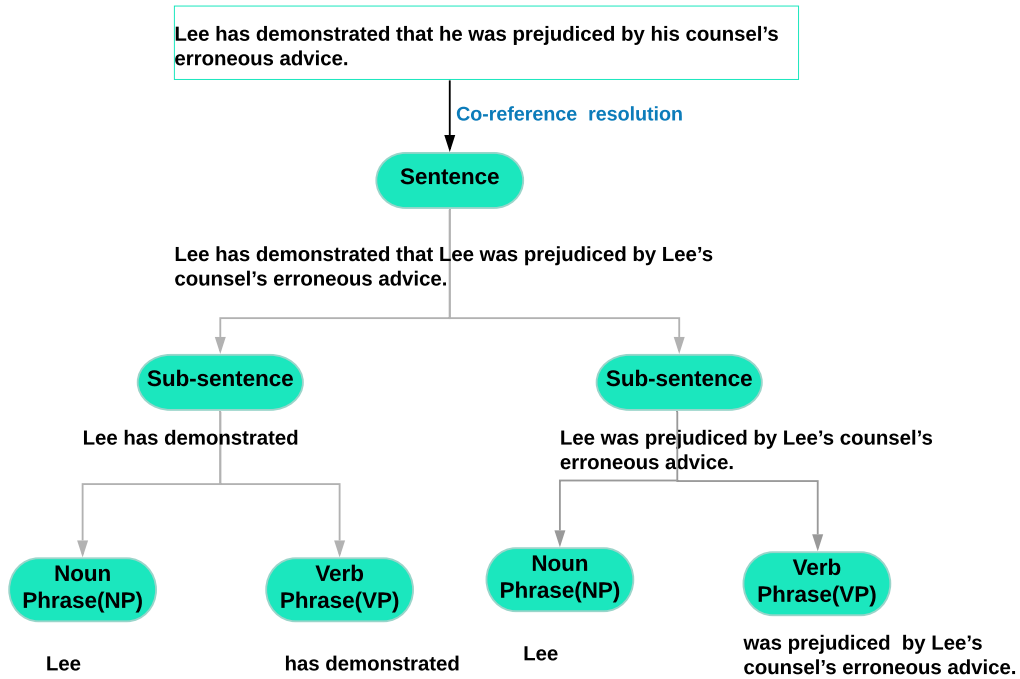


Figure 3.8: Process of dividing sentence

3.3 Ensemble Learning Approach

3.3.1 Research on ABSA Architectures

Our research process has followed an empirical research method. The main procedure that we have followed is referring to the existing literature of aspect-based sentiment analysis and understanding them in order to use the knowledge gained by them to use in our research domain. Instead of going straight to one architecture, we did research on different architectures using our dataset. While gaining the knowledge we analyzed the advantages and drawbacks of those architectures. By considering these facts we have decided whether we go through with those methods or not. When we aggregate our research problem into the deep learning approach, the sentence and the mentioned parties are inputs, and sentiments of each party are the output as shown in figure 3.9.

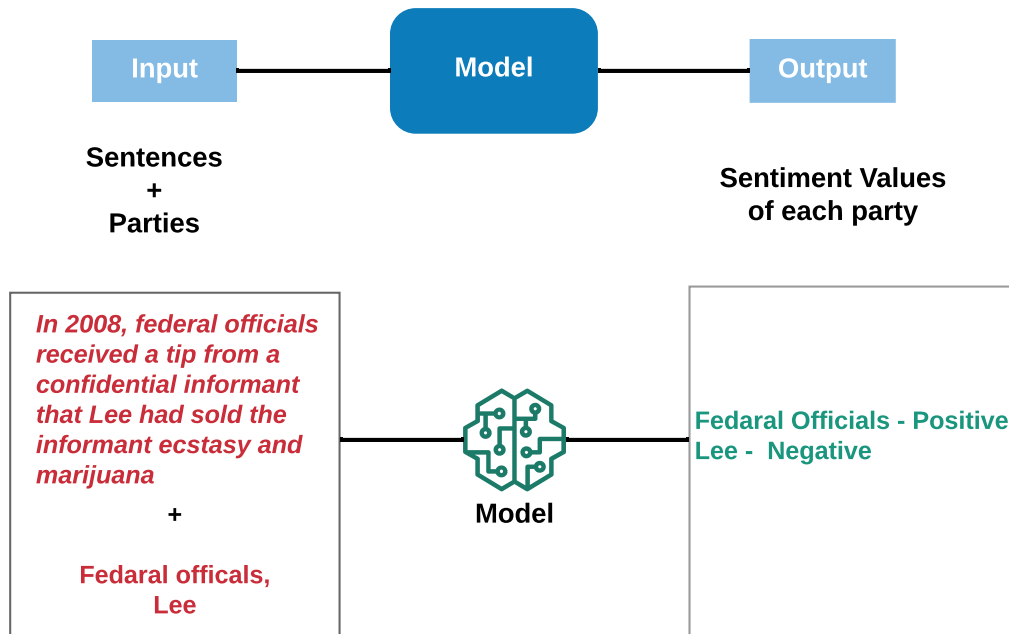


Figure 3.9: Input and Output of the Deep Learning Model

We have gone through different architectures of ABSA in other domains and identified several techniques in the literature that can be optimized and applied to our domain to perform party-based sentiment analysis. We implemented them to the legal domain and evaluated using our created dataset of SigmaLaw-ABSA.

- Target-Dependent Long Short-Term Memory (TD-LSTM)[48] - This model extends LSTM (long short-term memory) by considering the target word. It uses two LSTM networks to extract important information from the right and left side of the given target. Then concatenate the hidden states of two LSTM networks and finally classify the polarity labels by feeding them to a softmax layer.
- Target-Connection Long Short-Term Memory (TC-LSTM)[48] - This model incorporates semantic relatedness of the target with its context words by extending TD-LSTM with a target connection component.
- LSTM with Aspect Embedding (AE-LSTM)[50] - This model uses an LSTM network and learns an embedding vector for each aspect. Then each word input

vector is appended with input aspect embedding.

- Attention-based LSTM (AT-LSTM)[50] - This model uses an LSTM network with an attention mechanism to incorporate the interaction among words and the target.
- Attention-based LSTM with Aspect Embedding (ATAE-LSTM)[50] - This network incorporates the relationship among words and the target at both the input and attention layer. Each input word vector is appended with aspect embedding and attention mechanism is used along with LSTM.
- Interactive Attention Networks (IAN)[52] - This model was designed to create representations separately for contexts and targets by interactively learning attention in the target and context.
- MemNet[46] - This utilizes a deep memory network and multiple attention is paid to word embedding. For sentiment prediction, the output of the last attention is fed to the softmax layer. Here the results of different attentions are not combined.
- Content Attention Model for Aspect-Based Sentiment Analysis (Cabasc)[71] - This model uses two attention enhancing mechanisms. The sentence-level content attention mechanism is responsible for capturing important details of the aspect from a global perspective and context attention mechanism is responsible for taking into consideration at the same time the order of terms and their associations, by combining them into a set of customized memories.
- Recurrent Attention on Memory (RAM)[49] - This model first utilizes a bidirectional LSTM to produce memory from the input and weighted the memory slices depending on the relative distance to the aspect. Then multiple attentions on position-weighted memory are non-linearly combined with a recurrent neural network to predict the final sentiment polarity.
- Bert-spc (BERT for Sentence Pair Classification) [72] - This model uses a pre-trained BERT model where the BERT model is finetuned by adding just one output layer. Here, fine tuning can be done without modifying the architecture specifically.
- Aen-Bert (Aspect Level Sentiment Classification with Deep Memory Network)

[73]- This model tries to eliminate repetition and takes attention based encoders into consideration to represent between the target and context. Further in this model, it introduces a label smoothing regularization mechanism to reduce the issue of the label unreliability.

- Bert-LCF (Local Context Focus Mechanism) [57] - In this model it takes both local context and global context of a sentence into consideration. The key fact of this architecture is that it considers the local context has more important information than the global context.
- SDGCN (Sentiment Dependencies with Graph Convolutional Networks) [60]- This model aims to model sentiment dependencies within a sentence among different target aspects. First implements a bidirectional attention mechanism with position encoding in order to model aspect specific representation for the relation of context word with its aspects. Then, to capture the dependencies over various aspects within the sentence GCN was employed.

3.3.2 Ensemble Model

We developed a stacking ensemble model as shown in Figure 3.10, using the highest four accuracy models (ATAE-LSTM, RAM, Bert-LCF, SDCCN) that we gained for the SigmaLaw-ABSA dataset. Here we used the stacking ensemble method. The predictions gained from deep learning-based sub-models are combined with a learning algorithm. In this stacking ensemble model first, get the outputs of trained four sub-models as the input to the stacking ensemble model and for better output prediction learn the best way to combine input predictions.

The training dataset was prepared by giving a testing dataset to each submodel and collecting the set of predictions. We obtained three predictions for each sentence in the test set for one model which consists of probabilities relevant to the three classes (positive, negative and neutral). Then logistic regression meta learner was trained using the created dataset and the entire model was evaluated using the test set.

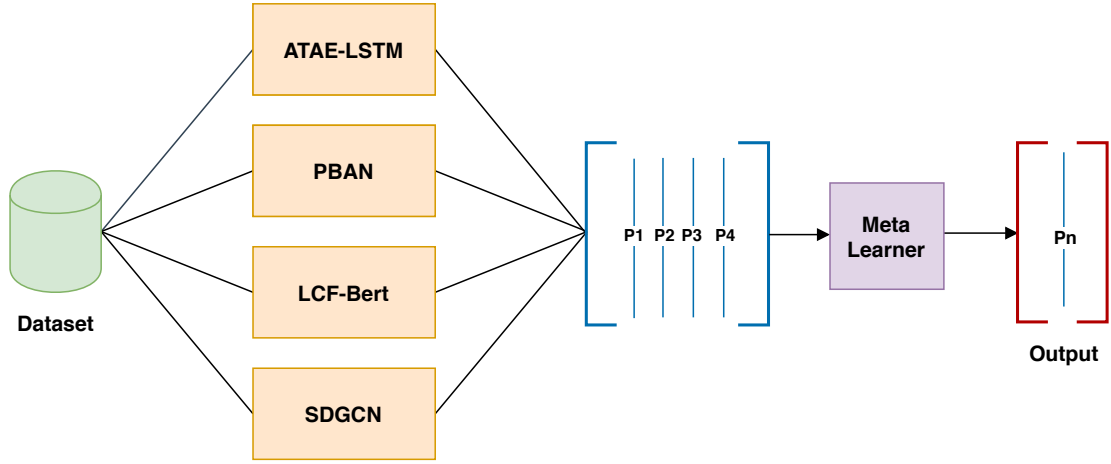


Figure 3.10: Stacking Ensemble Model

3.4 Deep Learning-based Approach

3.4.1 Introduction

The ultimate goal of our proposed approach is to detect the sentiment polarity of legal texts with respect to each legal party mentioned in the sentence. Legal texts usually consist of multiple legal parties having different inter-dependencies among them. Hence the sentiment classifier should be developed in order to classify sentiment polarity values of multiple legal parties. After going through the different architectures which have their own capabilities of achieving accurate output overcoming different limitations, we came up with a new architecture which can be introduced as a combination of all the capabilities of the above models. In our approach, positive, negative and neutral are considered as sentiment polarities. The overall architecture of our proposed model is illustrated in Fig. 3.11. To perform the aspect sentiment classification, our model architecture is designed with the following layers; word embedding layer, RNN layer, position aware attention mechanism, GCN layer and sentiment classification layer.

3.4.2 Word Embedding Layer

Word embedding layer maps each word to a high dimensional vector space. It is widely known that a strong word embedding is much important for composing a strong and

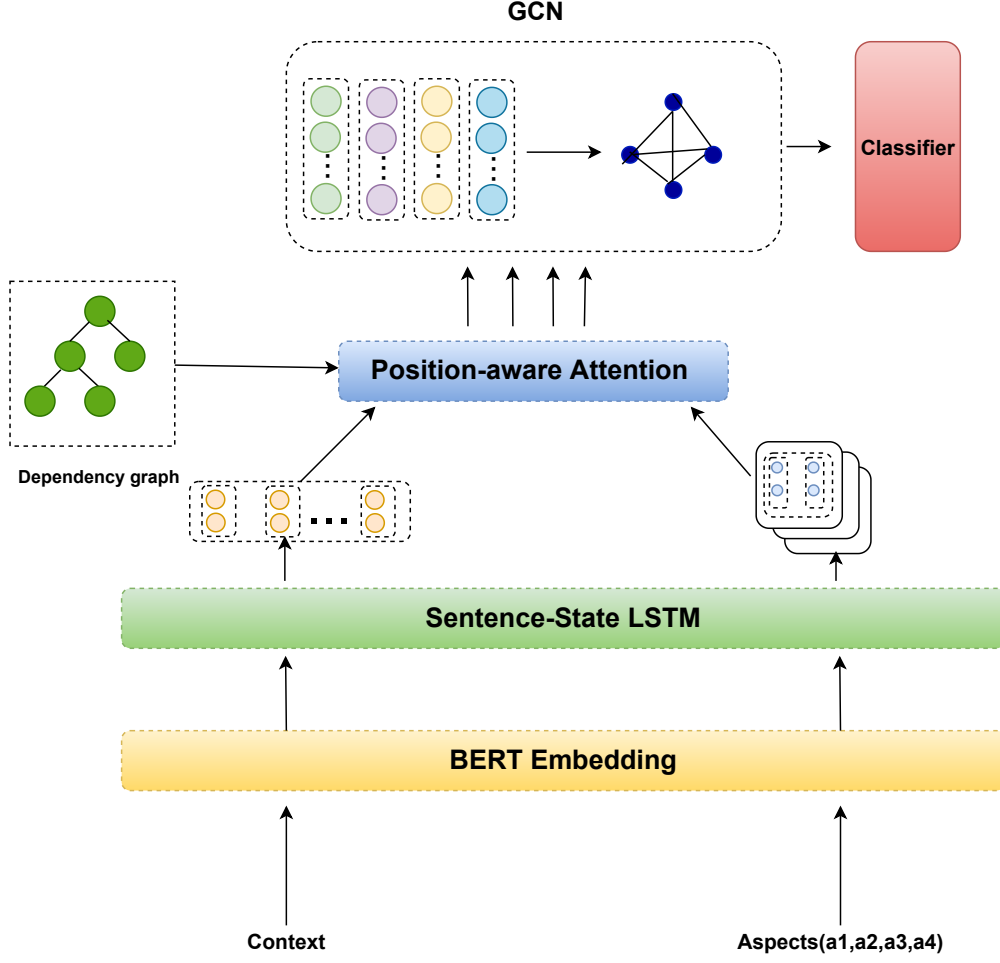


Figure 3.11: Overall architecture

efficient text representation at a higher stage[74]. We used a pre-trained BERT model⁴ further fine tuned for a legal corpus to obtain the word embedding [75]. Since other publicly available pre-trained BERT models are trained using corpora such as Wikipedia and books, we used the Bert model specifically fine-tuned for legal domain using a corpus of criminal court cases to obtain high accuracy[75].

Given a N-word input sentence $S = \{w_{s1}, w_{s2}, ..w_{sN}\}$ includes K aspect terms $W_a = \{W_{a1}, W_{a2}, ..W_{aK}\}$ and each aspect contains M_i words; $M_i \in [1; N)$. We use the above BERT model to get word embedding of the input sentence and all the aspect terms in

the sentence. First, we construct the input as “[CLS] + input + [SEP]” and feed it into BERT tokenizer to tokenize the input into tokens that correspond to BERT’s vocabulary. After mapping the token strings to their vocabulary indices, indexed tokens are fed into the BERT model. Each word of the context and aspects will be represented by a 768 dimensional embedding vector.

3.4.3 RNN Layer

In order to capture the contextual details for every word, on the top of the embedding layer we use Sentence-State LSTM(S-LSTM) [76]. Most of existing model architectures use LSTM, bi-LSTM, and Bi-GRU as the encoder. LSTM processes sequential data while maintaining long-term dependencies. However, while encoding long sentences it directs to less performance. In our case, the sentences of the legal documents are comparatively longer than in other domains. Therefore aiming to address these limitations of existing deep-learning approaches, we leverage a sentence state LSTM (S-LSTM) [77] to capture contextual information due to its proven performance. Instead of sequentially processing words, the S-LSTM simultaneously models the hidden states of all words in each recurrent time stage.

After feeding the word embeddings of a sentence to the S-LSTM model, it returns the contextual state H_i^t of the sentence which consists of a sub hidden state h_i^t for each word w_i and a sentence-level sub hidden state s^t as shown in the equation 3.1.

$$H^t = < h_0^t, h_1^t, h_1^t, h_2^t, \dots, h_{n-1}^t, s^t > \quad (3.1)$$

In our architecture, we use S-LSTM in order to get contextual hidden output of the sentence and contextual hidden outputs of aspects.

3.4.4 Position Aware Attention Mechanism

In a sentence, the sentiment polarity is heavily associated with the sentence’s aspect-words and opinion terms[78]. Hence, the method that we adopt to rely on these aspects-terms is quite important in the process of sentiment analysis. The main weakness of

RNN models is inability to understand the most critical parts of the sentence for sentiment analysis[79]. As a solution to this, we employ an attention mechanism which can grab the most important parts in a sentence. However, every word in a sentence is not equally important for determining sentiment polarity. Words which are closer to the target or having modifier relation to the target word should be given higher weights [53]. To ease this problem, we used an attention mechanism incorporating position information of each word in the sentence based on the current aspect term. We use position information here assuming that the aspect sentiment polarity is mainly influenced by the context words that are situated very close to the target aspect.

Lee was found guilty because the attorney had provided constitutionally ineffective assistance
 [0,6] [1,5] [2,4] [3,3] [4,2] [5,1] [6,0] [7,1] [8,2] [9,3] [10,4] [11,5]

Figure 3.12: Relative Distance

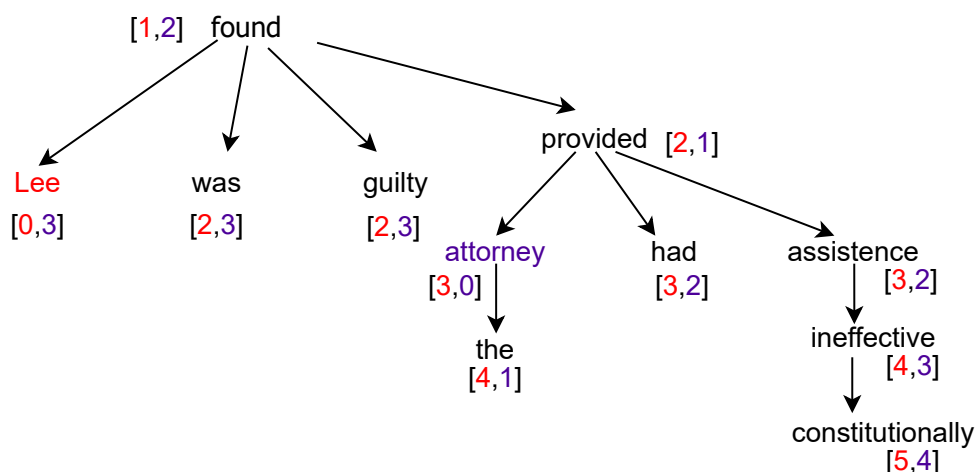


Figure 3.13: Dependency Graph

Here we used the bidirectional attention mechanism introduced by Zhao et al. [60] with two attention modules as context to aspect attention module and aspect to context attention module. We followed the same methodology for the calculation of attention weights. However, for position-aware representation, we used the distances along the dependency tree instead of relative distances used in their approach. In our approach as the distance, the length of the path from the specific word to the aspect in the dependency tree is experimented to encode the syntactic structure of the legal text. Fig. 3.12

illustrates the example sentence with relative distances to aspects and Fig. 3.13 shows distances along the dependency tree. When considering the two types of distances, we can see opinion words such as guilty and ineffective words are closer to the relevant aspects in the Fig. 3.13. The sentences in the court cases are comparatively much longer than other domains. Hence, opinion words are sometimes not much closer to the target. Therefore it is not suitable to get the relative distance between each word and the current aspect for position representation.

The final output of the attention mechanism will be Aspect-specific representation between the target aspect(party) and context words as $X = [x_1, x_2, \dots, x_K]$. Here K denotes the number of aspects.

3.4.5 Graph Convolution Network

In order to capture the inter-dependencies between multiple aspects/parties in a sentence, we have used GCN in our study as Zhao et al.[60] did. GCN can be identified as a basic and efficient convolution neural network running on graphs which has the ability to collect interdependent knowledge from rich relational data[60]. As the first stage of implementing the GCN layer, it is needed to construct a graph named sentiment graph where we considered a node and an edge as a party (aspect) mentioned in the sentence and the inter-dependency relation between 2 nodes respectively. If there is an edge between two nodes of the graph, it implies that there is a relation between corresponding aspects in the sentence. As shown in the Fig. 3.14, when creating the sentiment graph, we defined a fully-connected graph by taking an aspect is connected to every aspect of the sentence.

GCN generates a new vector representation for each node by discovering all relevant information about the neighboring nodes of the selected node. Moreover, when generating the new vector representation, it is needed to put attention on the information of the node itself. For that, we assume that each node has a self-loop. The new representation for a node can be defined as shown in the equation 3.2.

$$x_v^1 = ReLU(\sum_{u \in N(v)} W_{cross} x_u + b_{cross}) + ReLU(W_{self} x_v + b_{self}) \quad (3.2)$$

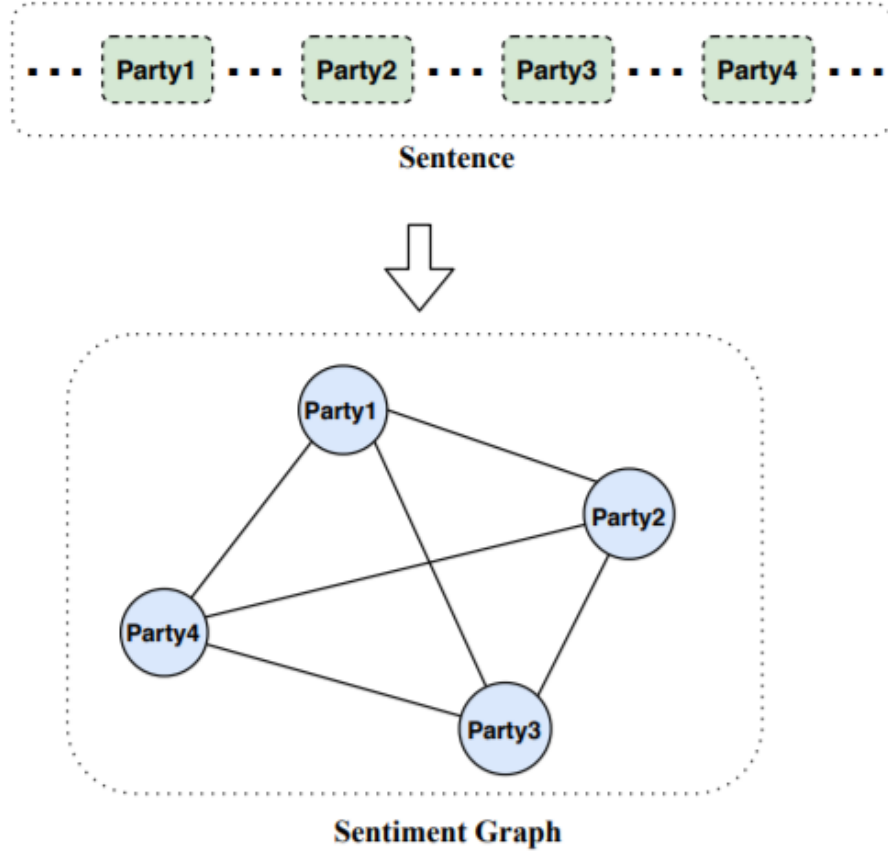


Figure 3.14: Sentiment Graph

where given v node $N(v)$ defines the all neighbors of v , $W_{cross}, W_{self} \in R^{d_m \times d_n}$, $b_{cross}, b_{self} \in R^{d_m \times 1}$, x_u is the u th aspect-specific representation taken from the output of attention layer and $d_m = d_n =$ dimension of hidden state.

We can expand the neighborhood for each node by stacking multiple gcnn layers. Each GCN layer gets the input as the output from the previous layer and returns the new node representation. In our case, we use two GCN layers after the experiments(see Eq. 3.3).

$$x_v^2 = ReLU(\sum_{u \in N(v)} W_{cross}^1 x_u^1 + b_{cross}^1) + ReLU(W_{self}^1 x_v^1 + b_{self}^1) \quad (3.3)$$

3.4.6 Sentiment Classification

Once the output of the GCN layer(x) is obtained, it is fed to a softmax layer to obtain a probability distribution over polarity decision space of C classes(where W and B are the learned weight and bias):

$$z = \text{Softmax}(Wx + b) \quad (3.4)$$

3.4.7 Model Training

The model is trained by the gradient descent algorithm with cross entropy loss and L2 regularization.

$$Loss = - \sum_{c=1}^C y \log \hat{y} + \lambda ||\theta||^2 \quad (3.5)$$

C denotes the number of classes(3 in our case) y is the true label and predicted label denotes in \hat{y} . θ denotes all the parameters that need to regularized, and λ is the coefficient of L2-regularization.

Chapter 4

Experiments

4.1 Experimental Setup and Dataset

All the experiments and evaluations were carried out on the created dataset SigmaLaw-ABSA [80] consists of 2000 human-annotated legal sentences taken from previous court cases which were fetched from the *SigmaLaw - Large Legal Text Corpus and Word Embedding dataset*. The SigmaLaw-ABSA dataset is annotated by legal experts and it contains entities of different parties, their polarities, aspect category (Petitioner and defendant), and the category polarities. The dataset has been designed to perform various research tasks in the legal domain including aspect extraction, polarity detection, aspect category identification, aspect category polarity detection.

4.1.1 Dataset Statistics

This section discloses preliminary results of some statistical analysis performed on this collection of data. After the annotation process, our dataset contains 2000 total sentences which include 1007 full sentences and 993 sub-sentences. Overall, it contains 642 positive 978 negative and 380 neutral sentences. For the Polarity classification of aspect types; petitioner and defendant, it counts 335 positive, 618 negative, 198 neutral for the petitioner and 285 positive, 244 negative, 106 neutral for the defendant. Furthermore, the overall summary of the contents of the dataset is listed in Table 4.1.

Table 4.1: Dataset Statistics

Polarity	Petitioner	Defendant	Total
Positive	473	393	866
Negative	734	364	1098
Neutral	254	143	397

Furthermore, the dataset contains the overall sentiment of sentences as one of its features. The polarity wise details for full sentences and sub-sentences are reported in Table 4.2.

Table 4.2: Statics for the Overall Sentiment of Sentences

Polarity	Full Sentences	sub-sentences	Total
Positive	357	285	642
Negative	517	461	978
Neutral	133	247	380
Total	1007	993	2000

Figure 4.1 shows word-count frequencies of all sentences along with *positive*, *negative* and *neutral* sentences separately. The sentences of Legal documents are often long and have a complex semantic structure. This plot also highlights that sentences in legal documents are long. Difficulties of handling long sentences and understanding domain-specific phrases emphasize the complexity of the data annotation in the legal domain and the importance of this dataset.

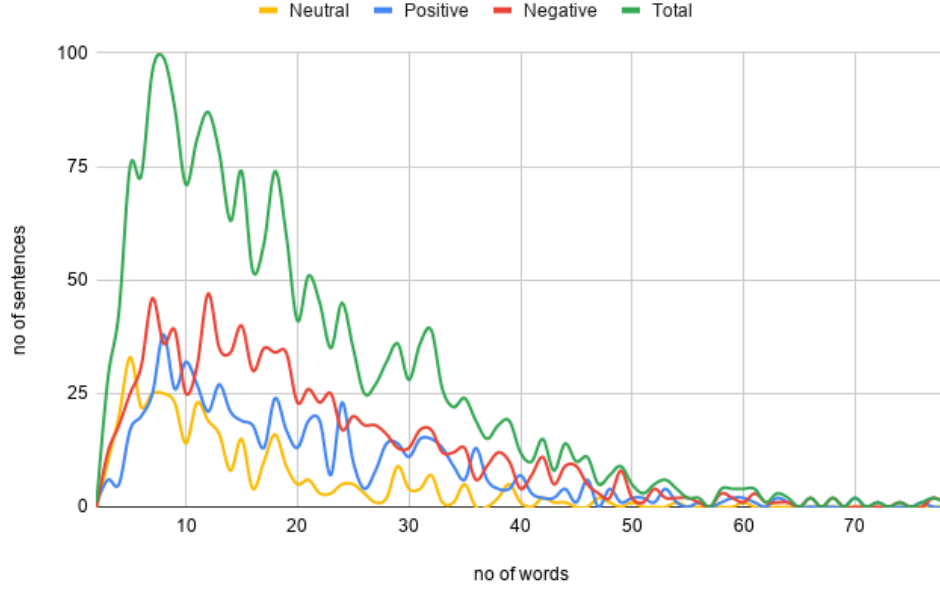


Figure 4.1: Word Count Frequency Plot

4.2 Experimental Results

This section discloses experimental results of the relevant methodologies described in chapter 3.

4.2.1 Rule-Based Approach

In the experiments, we compare the sentiment class of each party picked by human judges and the proposed rule-based method which makes use of the sentiment annotator by Gamage et al. [7]. The results have been evaluated considering the standard measures of precision and recall. Precision, recall, and f1 score are calculated using the following formulas. For the considered class C (C can be negative /non-negative), let S be the set of data points in the data-set which belong to the class C and Y be the set of data points that are classified as belonging to class C (classified by the proposed approaches).

$$Precision = \left(\frac{S \cap Y}{Y} \right) \quad (4.1)$$

Table 4.3: Results of Proposed Approach

	Recall	Precision	F1 Score
Petitioner-Negative	0.7257	0.6805	0.7023
Petitioner-Non negative	0.6389	0.6525	0.6455
Defendant-Negative	0.5857	0.6721	0.6267
Defendant-Non negative	0.6295	0.8491	0.7229

$$Recall = \left(\frac{S \wedge Y}{S} \right) \quad (4.2)$$

$$F1 \text{ score} = \left(\frac{2 \times Precision \times Recall}{Precision + Recall} \right) \quad (4.3)$$

The results acquired for precision, recall, and f-score from our proposed approach are shown in table 4.3 and this can be considered as the benchmark for the SigmaLaw-ABSA dataset.

Furthermore, we get the results from adopting the same rule-based classification methodology using Stanford Stanford Sentiment Annotator [18] instead of the phrase-level annotator by Gamage et al. mentioned in part D in the methodology section. The Stanford sentiment annotator [18] performed the sentiment classification with five classes as shown in table 3.1. But in our approach and also in the study by Gamage et al. [7], only negative and non-negative classes are considered. Therefore, we aggregate the output sentiment classes of the Stanford model [18] into two classes following the same approach proposed in the study by Gamage et al. [7] (See Table 4.4). According to the results rule base approach using phrase-level sentiment annotator by Gamage et al. [7] outperforms the same rule base approach using the Stanford sentiment. Figure 4.2, 4.3, 4.4 shows the precision, recall and f1 score distribution for two approaches respectively.

Table 4.4: Results using Stanford Sentiment Annotator

	Recall	Precision	F1 Score
Petitioner-Negative	0.6018	0.6044	0.6031
Petitioner-Non negative	0.5706	0.6035	0.5864
Defendant-Negative	0.6571	0.5246	0.5834
Defendant-Non negative	0.5245	0.7576	0.6198

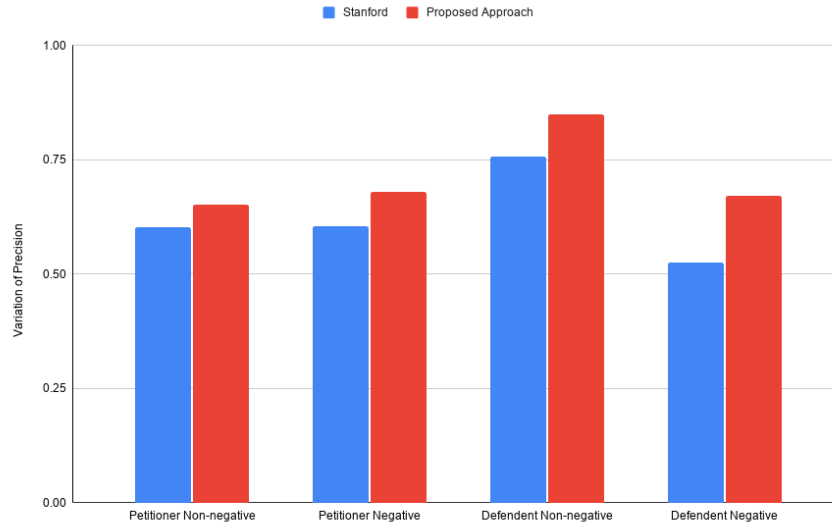


Figure 4.2: Precision Variation

4.2.2 Research on ABSA deep-learning based Architectures

This section discloses the experimental results of the research that we were carried out to find the best components in existing ABSA architectures. As the first step, preprocessing was performed on the dataset. After this, we classified the sentiment polarity using some popular models available for ABSA using deep learning techniques as mentioned in section 3.3 in chapter Methodology.

After evaluating all the models on the SigmaLaw ABSA dataset we have identified four

Table 4.5: Performances of Different Models on SigmaLaw-ABSA

Model	Accuracy	F1 score
TD-LSTM [48]	0.6512	0.5647
TC-LSTM [48]	0.6182	0.5438
AE-LSTM [50]	0.6228	0.5588
AT-LSTM [50]	0.6272	0.5592
ATAE-LSTM [50]	0.6542	0.5802
IAN [52]	0.6332	0.5650
PBAN [81]	0.6332	0.5650
MemNet [46]	0.5389	0.4361
Cabasc [71]	0.6123	0.5643
RAM [49]	0.6639	0.6022
SDGCN [60]	0.6781	0.6121
ASDGCN [59]	0.6699	0.6001

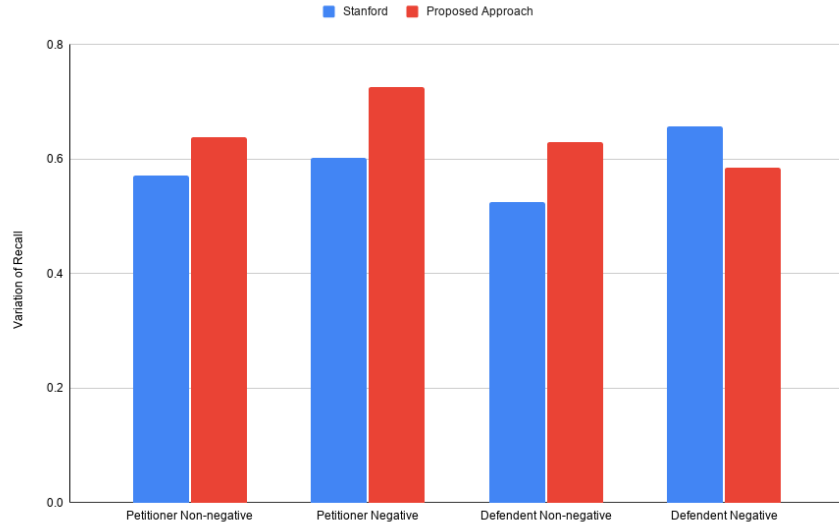


Figure 4.3: Recall Variation

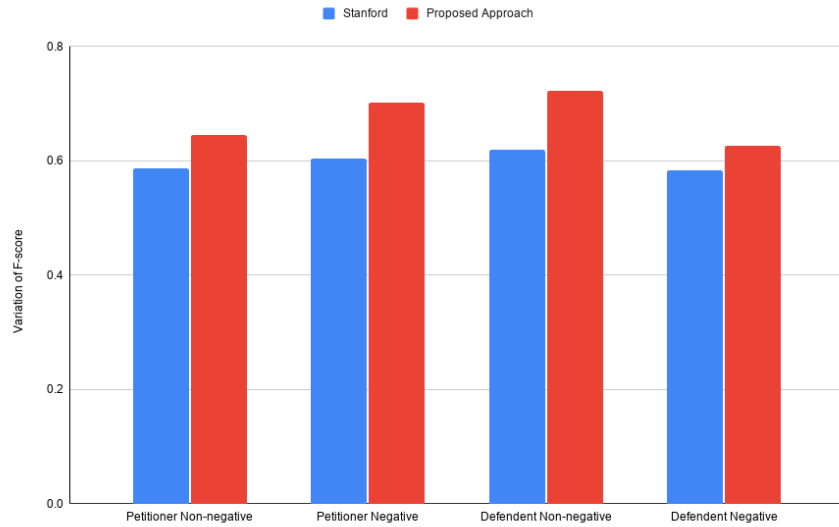


Figure 4.4: F1 score Variation

models that have the highest accuracy as best architectures that fit our problem. They are ATAE-LSTM, RAM, Bert-LCF, and SDGCN. But each technique alone is difficult to be applied in the party based sentiment analysis system. Due to the complexity of the language in the legal domain, single technique cannot tackle well. Therefore, we focus

on ensemble techniques to improve the accuracy of the model. The implementation details of the ensemble model is provided in section 3.31. The meta learner outperformed each of the sub-models and we obtained the following results for accuracy and f1-score.

- Stacked test Accuracy: 0.6921
- Stacked test f1 score: 0.6049

4.2.3 Deep Learning Approach

The experiments that we are carried out for the proposed deep-learning based approach in section 3.4 is described below. We also conducted component-wise experiments to identify which methods are most effective for legal text.

4.2.3.1 Word Embedding Models Comparison

In our experiments We tried two word embedding methods: 300-dimensional GloVe embeddings [82] and BERT representations [72]. In BERT representation, two different BERT models were tried for the embedding layer: base uncased English model and pre-trained BERT model specially fine-tuned for the legal corpus with dimension 768. Table 4.6 shows the comparison of the results of above models. The legal-BERT model outperformed the other models. The BERT models use 12 layers of transformer encoders, and each output per token from each layer of these and initial input embedding can be used as a word embedding. We tried various vector combinations of hidden layers to get state of art results. Table 4.7 illustrates the result of various word-embedding strategies using the BERT model for legal domain and summation of last four hidden layers outperformed the other strategies.

4.2.3.2 RNN Models Comparison

LSTM, Bi-LSTM, Bi-GRU and Sentence-State LSTM(S-LSTM) models are tried as the encoder for our approach. S-LSTM outperformed because it has strong representation power compared to RNNs [76]. The sentences of Legal documents are often long and have a complex semantic structure. As S-LSTM offers richer contextual information exchange with more parallelism compared to BiLSTMs, it obtained higher result than

Table 4.6: Word embedding models comparison

Model	Accuracy	F1 score
Glove	0.6615	0.5798
BERT-base	0.6997	0.6193
BERT-legal domain	0.7068	0.6281

Table 4.7: Different word embedding strategies comparison of BERT model

Strategy	Accuracy	F1 score
Initial embedding	0.6670	0.5705
Last hidden layer	0.6921	0.6193
Sum all layers	0.6954	0.6098
Sum last 4 layers	0.7086	0.6281
Concat last 4 layers	0.6987	0.6105

Table 4.8: RNN models comparison

model	Accuracy	F1 score
LSTM	0.6721	0.5964
Bi-LSTM	0.6987	0.6204
Bi-GRU	0.6854	0.6045
S-LSTM	0.7086	0.6281

other models. Table 4.8 shows the comparison of the results of the models listed above.

4.2.3.3 Overall Performance

The experimental results generated on different existing models using the SigmaLaw-ABSA dataset [80] are shown in Table 5.1. By analysing the obtained results, we can conclude that in the legal domain our proposed model outperforms every other existing models. We claim that it is mainly due to the complexity and the length of the sentences in the legal domain as it makes it difficult to those models to understand the sentence well.

4.2.3.4 Ablation Study

In order to study the efficiency of the various modules in our proposed approach, we conducted an ablation study on the SigmaLaw-ABSA dataset. Results shown in the Table 4.10 and it shows that removing both attention mechanisms and GCN drops the F1 score by 0.0617. By introducing the attention mechanism(with dependency tree distance) to baseline f1 score increase by 0.0439. This verifies the significance of the position-aware attention mechanism. The results gained from using the dependency tree distance to calculate position weights shows higher performance than the calculating position weights through relative distances. This shows the impact of the syntactic information introduced by the dependency trees.

Also, we can see that the model shows higher results with the introduction of the GCN layer. Therefore we can conclude that the GCN layer contributes significantly to in-

Table 4.9: Performances of Different Models on SigmaLaw-ABSA

Model	Accuracy	F1 score
TD-LSTM [48]	0.6512	0.5647
TC-LSTM [48]	0.6182	0.5438
AE-LSTM [50]	0.6228	0.5588
AT-LSTM [50]	0.6272	0.5592
ATAE-LSTM [50]	0.6542	0.5802
IAN [52]	0.6332	0.5650
PBAN [81]	0.6332	0.5650
MemNet [46]	0.5389	0.4361
Cabasc [71]	0.6123	0.5643
RAM [49]	0.6639	0.6022
SDGCN [60]	0.6781	0.6121
ASDGCN [59]	0.6699	0.6001
SigmaLaw-PBSA	0.7086	0.6281

Table 4.10: Results of ablation study

Setting	Accuracy	F1 score
Base	0.6287	0.5664
Base + Attention(relative distance)	0.6689	0.5989
Base + Attention(dependency tree distance)	0.6793	0.6103
Base + Attention(dependency tree distance) + 1_layer GCN	0.6938	0.6215
Base + Attention(dependency tree distance) + 2_layer GCN	0.7086	0.6281

crease the results since it helps to capture the inter-dependencies among multiple aspects and relationships between words at long ranges.

Chapter 5

Discussion

5.1 Multi Functionality of SigmaLaw-ABSA Dataset

Aspect-Based Sentiment Analysis (ABSA) is a combination of several core sub-tasks. Aspect term extraction, aspect term polarity, category identification, and aspect category polarity can be taken as major sub-tasks. In addition to these sub-tasks, the complexity of the ABSA process depends on the domain and language. As a result of that, much research has been conducted related to these core sub-tasks, domains, and languages. Considering all facts, the datasets of the ABSA field should be multi-tasks. The SigmaLaw-ABSA dataset is prepared to cover all of the above core sub-tasks relevant to the legal domain in the language of English. All sub-tasks are performed on the sentences in court cases. In this section, we discuss the multi-functionality of SigmaLaw-ABSA dataset.

5.1.1 T1 - Aspect Extraction

For the given legal opinion sentence in a court case, this aspect extraction task concerns the extraction of all persons or organizations that belong to any legal party. Identifying only members of legal parties is a bit of a challenging process. Hence, the aspect extraction model needs better training. As one of its features, this dataset provides all members of parties in sentences.

5.1.2 T2 - Aspect Term Polarity

This task is based on allocating sentiment value (*positive*, *negative*, and *neutral*) to the aspects extracted from the T1 task. The neutral case arises where the content of the sentence is not affected by any party positively or negatively.

5.1.3 T3 - Aspect Category Extraction

The aspect extraction (T1) task focuses on extracting all persons or organizations belong to any party without any categorization. For lawyers and legal officials, it's beneficial to identify the party of persons or organizations. Therefore, this task deals with categorizing the aspects extracted from a legal opinion text under the petitioner and defendant parties.

5.1.4 T4 - Aspect Category Polarity

This task explores the possibilities of allocating sentiment values for the aspect categories of sentences. T1, T2, and T3 tasks can be applied together for research inspecting how aspects and their polarities influence the polarity of aspect categories. Identifying the sentiment value of the legal texts with respect to the petitioner and defendant parties is an important task when analysing the court cases. The research of identifying sentiment of the parties can be carried out by manipulating the output from the previous tasks.

Example 4

- Sentence 4.1: *According to Lee, the lawyer assured him that if deportation was not in the plea agreement, "the government cannot deport you."*

Consider Example 2 taken from Lee v. United States[17] which consists of two legal parties; petitioner Lee and the defendant government. The output from T1, T2, and T3 tasks are $\{Lee, lawyer, government\}$, $\{positive, neutral, negative\}$, and $\{petitioner-[Lee, lawyer], defendant-[government]\}$. The output data are sufficient to classify the sentiment values of both petitioner and defendant parties. The output data from tasks T2 and T3 can be combined as follows: $\{petitioner-[positive, neutral], defendant-[negative]\}$. We can predict the polarity of parties by adding aspect polarity values of

both parties separately. Then we can get a positive value $((+1)+0)$ for the petitioner party and a negative value (-1) for the defendant party. Researchers can do further experiments on predicting aspect category polarity more accurately.

After analysing the subtasks of ABSA relevant to the legal domain, we hope that our dataset will be useful to achieve these subtasks. Both T2 and T3 tasks are covered by the field of party which uses the list of lists as its data structure. Here, all aspects belonging to both parties are listed separately. The first and second lists define aspects of petitioner and defendant parties respectively. Utilizing one feature, we have achieved two tasks. Sentiment values of aspects are also defined just like aspects as the list of lists. Task T3 can be performed directly using this dataset. As explained in the task T4, data included in party and sentiment features can be used to carry out those experiments. Besides that, we have indicated the overall sentiment of the sentence as one of the features. The purpose of indicating the overall sentiment is that any researcher can use this dataset to do experiments on identifying the relationship between sentiment analysis and the Aspect-Based Sentiment Analysis fields. Concluding all facts we explored, SigmaLaw-ABSA dataset has proved its multi-functionality.

5.2 Future Work

So far several approaches are being researched to identify the best model architecture for party-based sentiment analysis of legal opinion texts. Improving the overall system implementation and performance can also be considered as a future work of this research study. Furthermore, as the major future work in the legal domain, it can be mentioned that develop a system that can predict the winning party of a given case.

5.3 Individual Contribution

Table 5.1: Individual Project Contribution

Index Number	Contribution	Task Completed
160222M	100%	<ul style="list-style-type: none"> • Dataset - Data collection and Annotation • Rule Based Approach - Extracting phrases and Sentiment Classification tasks • Ensemble Learning Approach - Implementation of IAN, MemNet, PBAN, SDGCN models • Deep Learning Based Approach - Implementation of Word Embedding Layer and RNN Layer • Research paper publications
160280L	100%	<ul style="list-style-type: none"> • Dataset - Data collection and Annotation • Rule Based Approach - Generating sub sentences • Ensemble Learning Approach - Implementation of TD LSTM, TC LSTM, AEN, Bert-SPC, and Bert-LCF • Deep Learning Based Approach - Graph Convolution Network (GCN) and Front-end development • Research paper publications
160537H	100%	<ul style="list-style-type: none"> • Dataset - Data collection and Annotation • Rule Based Approach- Implementing Co-reference Resolution and setting up Phrase level sentiment annotator • Ensemble Learning Approach - Implementation of ATAE-LSTM, RAM, ASDGCN and stacking ensemble models. • Deep Learning Based Approach - Implementation of Position aware attention mechanism 66 • Research paper publications

5.4 Conclusions

In conclusion, the primary research contribution of this study would be proposing methodologies to perform Party-Based Sentiment Analysis for Legal Opinion Text. To the best of our knowledge, this would be the first study that brings the concepts of Aspect Based Sentiment Analysis to the legal domain. In this study, sentiment polarity (positive, negative, neutral) of sentences in legal documents with respect to each legal party mentioned in the document will be extracted. This project aims to mitigate the issues in existing work on Phrase-Level Sentiment Analysis in the legal domain, by proposing aspect level sentiment analysis to facilitate legal information tasks such as party identification, discourse analysis and detection of contradictory statements. Therefore the proposed study has the potential to contribute to the research community in a significant manner while solving a real-world issue.

To evaluate the proposing systems we have created a SigmaLaw-ABSA dataset[80] which consists of 2000 sentences taken from previous court cases. The court cases were collected from the SigmaLaw - Large Legal Text Corpus and Word Embeddings dataset. The SigmaLaw-ABSA dataset has been designed to perform various research tasks in the legal do-main including aspect extraction, polarity detection, aspect category identification, aspect category polarity detection and this dataset will provide significant importance for the research of the Party-Based Sentiment Analysis for legal opinion texts.

Then we introduced a Rule-Based Approach for Party-based Sentiment Analysis of Legal Opinion Texts[64]. By introducing this approach , we identified how rule-based approaches can be effectively developed and applied to sentiment analysis at aspect level. In this study, sentiment polarity (positive, negative, neutral) of sentences in legal documents with respect to each legal party mentioned in the document will be extracted. Furthermore, proposing a benchmark for party based sentiment analysis in legal opinion texts also can be considered as a main contribution of this approach. This rule-based system was primarily built around the phrase-level sentiment annotator developed by Gamage et al. [7] specifically for the legal domain and we used Stanford NLP library [61] and rationally built rules for aspect-level sentiment classification. The proposed approach consists with several major steps as co-reference resolution, extract-

ing phrases using constituency parser, sentiment classification of phrases and overall sentiment polarity computation.

Moreover, we introduced a deep learning based approach to perform party based sentiment analysis in legal opinion texts. The model first uses a pre-trained BERT model further fine tuned for a legal corpus for a strong word embedding. Then model employs the position aware attention mechanism to capture the critical parts of the sentence relevant to aspects with incorporating position information using the dependency tree. Because multiple legal party members are involved in a single sentence, GCN is employed over the attention mechanism to model the inter-dependencies between members. Experiments were carried out using the SigmaLaw-ABSA dataset and our experimental results demonstrates that our proposed approach outperformed all other existing ABSA models.

Bibliography

- [1] K. Sugathadasa, B. Ayesha, N. de Silva, A. S. Perera, V. Jayawardana, D. Lakmal, and M. Perera, “Legal Document Retrieval using Document Vector Embeddings and Deep Learning,” in *Science and information conference*. Springer, 2018, pp. 160–175.
- [2] V. Jayawardana, D. Lakmal, N. de Silva, A. S. Perera, K. Sugathadasa, and B. Ayesha, “Deriving a Representative Vector for Ontology Classes with Instance Word Vector Embeddings,” in *2017 Seventh International Conference on Innovative Computing Technology (INTECH)*. IEEE, 2017, pp. 79–84.
- [3] P. Jackson, K. Al-Kofahi, A. Tyrrell, and A. Vachher, “Information extraction from case law and retrieval of prior cases,” *Artificial Intelligence*, vol. 150, no. 1-2, pp. 239–290, 2003.
- [4] M.-F. Moens, C. Uyttendaele, and J. Dumortier, “Intelligent information extraction from legal texts,” *Information & Communications Technology Law*, vol. 9, no. 1, pp. 17–26, 2000.
- [5] S. Moralwar and S. Deshmukh, “Different approaches of sentiment analysis,” *International Journal of Computer Sciences and Engineering*, vol. 3, no. 3, pp. 160–165, 2015.
- [6] B. Liu and L. Zhang, “A survey of opinion mining and sentiment analysis,” in *Mining text data*. Springer, 2012, pp. 415–463.
- [7] V. Gamage, M. Warushavithana, N. de Silva, A. S. Perera, G. Ratnayaka, and

- T. Rupasinghe, “Fast approach to build an automatic sentiment annotator for legal domain using transfer learning,” *arXiv preprint arXiv:1810.01912*, 2018.
- [8] K. Schouten and F. Frasincar, “Survey on aspect-level sentiment analysis,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 3, pp. 813–830, 2015.
- [9] N. K. Laskari and S. K. Sanampudi, “Aspect based sentiment analysis survey,” *IOSR Journal of Computer Engineering (IOSR-JCE)*, vol. 18, no. 2, pp. 24–28, 2016.
- [10] M. Pontiki, D. Galanis, H. Papageorgiou *et al.*, “Semeval-2016 task 5: Aspect based sentiment analysis,” 01 2016, pp. 19–30.
- [11] S. Brüninghaus and K. D. Ashley, “Improving the representation of legal case texts with information extraction methods,” in *Proceedings of the 8th international conference on Artificial intelligence and law*, 2001, pp. 42–51.
- [12] G. Ratnayaka, T. Rupasinghe, N. de Silva, M. Warushavithana, V. Gamage, and A. S. Perera, “Identifying relationships among sentences in court case transcripts using discourse relations,” in *2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer)*. IEEE, 2018, pp. 13–20.
- [13] G. Ratnayaka, T. Rupasinghe, N. de Silva, M. Warushavithana, V. Gamage, M. Perera, and A. S. Perera, “Classifying Sentences in Court Case Transcripts using Discourse and Argumentative Properties,” *ICTer*, vol. 12, no. 1, 2019.
- [14] “caselaw: cases and codes - findlaw caselaw,” 2020. [Online]. Available: <https://caselaw.findlaw.com/>
- [15] “Westlaw uk - online legal research from sweet maxwell - westlaw uk,” 2020. [Online]. Available: <https://legalsolutions.thomsonreuters.co.uk/en.html>
- [16] “Bailii | ials,” 2020. [Online]. Available: <http://ials.sas.ac.uk/digital/bailii>
- [17] “Lee v. United States,” in *US*, vol. 432, no. No. 76-5187. Supreme Court, 1977, p. 23.

- [18] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Y. Ng, and C. Potts, “Recursive deep models for semantic compositionality over a sentiment treebank,” in *EMNLP*, 2013.
- [19] V. Jayawardana, D. Lakmal, N. de Silva, A. S. Perera, K. Sugathadasa, B. Ayesha, and M. Perera, “Word Vector Embeddings and Domain Specific Semantic based Semi-Supervised Ontology Instance Population,” *International Journal on Advances in ICT for Emerging Regions*, vol. 10, no. 1, p. 1, 2017.
- [20] —, “Semi-Supervised Instance Population of an Ontology using Word Vector Embeddings,” in *2017 Seventeenth International Conference on Advances in ICT for Emerging Regions (ICTer)*. IEEE, 2017, pp. 1–7.
- [21] K. Sugathadasa, B. Ayesha, N. de Silva, A. S. Perera, V. Jayawardana, D. Lakmal, and M. Perera, “Synergistic Union of Word2Vec and Lexicon for Domain Specific Semantic Similarity,” in *2017 IEEE International Conference on Industrial and Information Systems (ICIIS)*. IEEE, 2017, pp. 1–6.
- [22] —, “Legal Document Retrieval using Document Vector Embeddings and Deep Learning,” in *Science and information conference*. Springer, 2018, pp. 160–175.
- [23] J. J. Nay, “Gov2vec: Learning distributed representations of institutions and their legal text,” *arXiv preprint arXiv:1609.06616*, 2016.
- [24] G. Ratnayaka, T. Rupasinghe, N. de Silva, V. S. Gamage, M. Warushavithana, and A. S. Perera, “Shift-of-perspective identification within legal cases,” *arXiv preprint arXiv:1906.02430*, 2019.
- [25] G. Ganu, N. Elhadad, and A. Marian, “Beyond the stars: improving rating predictions using review text content.” in *WebDB*, vol. 9. Citeseer, 2009, pp. 1–6.
- [26] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, “SemEval-2014 task 4: Aspect based sentiment analysis,” in *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*. Dublin, Ireland: Association for Computational Linguistics, Aug. 2014, pp. 27–35. [Online]. Available: <https://www.aclweb.org/anthology/S14-2004>

- [27] M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar, and I. Androutsopoulos, "Semeval-2015 task 12: Aspect based sentiment analysis," in *Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015)*, 2015, pp. 486–495.
- [28] M. AL-Smadi, O. Qawasmeh, B. Talafha, and M. Quwaider, "Human annotated arabic dataset of book reviews for aspect based sentiment analysis," 08 2015.
- [29] A. Tamchyna, O. Fiala, and K. Veselovská, "Czech aspect-based sentiment analysis: A new dataset and preliminary results," in *ITAT*, 2015.
- [30] M. Rahman, E. Kumar Dey *et al.*, "Datasets for aspect-based sentiment analysis in bangla and its baseline evaluation," *Data*, vol. 3, no. 2, p. 15, 2018.
- [31] D. Rao and D. Ravichandran, "Semi-supervised polarity lexicon induction," in *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, 2009, pp. 675–682.
- [32] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, "Target-dependent twitter sentiment classification," in *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies*, 2011, pp. 151–160.
- [33] X. Ding, B. Liu, and P. S. Yu, "A holistic lexicon-based approach to opinion mining," in *Proceedings of the 2008 international conference on web search and data mining*, 2008, pp. 231–240.
- [34] X. Zhao, J. Jiang, H. Yan, and X. Li, "Jointly modeling aspects and opinions with a maxent-lda hybrid." *ACL*, 2010.
- [35] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [36] S. Poria, E. Cambria, A. Gelbukh, F. Bisio, and A. Hussain, "Sentiment data flow analysis by means of dynamic linguistic patterns," *IEEE Computational Intelligence Magazine*, vol. 10, no. 4, pp. 26–36, 2015.

- [37] H. H. Do, P. Prasad, A. Maag, and A. Alsadoon, “Deep learning for aspect-based sentiment analysis: a comparative review,” *Expert Systems with Applications*, vol. 118, pp. 272–299, 2019.
- [38] L. Rojas-Barahona, “Deep learning for sentiment analysis language and linguistics. compass 10: 701–719,” 2016.
- [39] Y. Goldberg, “A primer on neural network models for natural language processing,” *Journal of Artificial Intelligence Research*, vol. 57, pp. 345–420, 2016.
- [40] O. Araque, I. Corcuera-Platas, J. F. Sánchez-Rada, and C. A. Iglesias, “Enhancing deep learning sentiment analysis with ensemble techniques in social applications,” *Expert Systems with Applications*, vol. 77, pp. 236–246, 2017.
- [41] T. Young, D. Hazarika, S. Poria, and E. Cambria, “Recent trends in deep learning based natural language processing,” *ieee Computational intelligence magazine*, vol. 13, no. 3, pp. 55–75, 2018.
- [42] S. Poria, E. Cambria, and A. Gelbukh, “Aspect extraction for opinion mining with a deep convolutional neural network,” *Knowledge-Based Systems*, vol. 108, pp. 42–49, 2016.
- [43] Z. Toh and J. Su, “Nlangp at semeval-2016 task 5: Improving aspect based sentiment analysis using neural network features,” in *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)*, 2016, pp. 282–288.
- [44] L. Xu, J. Lin, L. Wang, C. Yin, and J. Wang, “Deep convolutional neural network based approach for aspect-based sentiment analysis,” *Advanced Science and Technology Letters*, vol. 143, pp. 199–204, 2017.
- [45] X. Gu, Y. Gu, and H. Wu, “Cascaded convolutional neural networks for aspect-based opinion summary,” *Neural Processing Letters*, vol. 46, no. 2, pp. 581–594, 2017.
- [46] D. Tang, B. Qin, and T. Liu, “Aspect level sentiment classification with deep memory network,” *arXiv preprint arXiv:1605.08900*, 2016.

- [47] B. Plank, A. Søgaard, and Y. Goldberg, “Multilingual part-of-speech tagging with bidirectional long short-term memory models and auxiliary loss,” *arXiv preprint arXiv:1604.05529*, 2016.
- [48] D. Tang, B. Qin, X. Feng, and T. Liu, “Effective lstms for target-dependent sentiment classification,” *arXiv preprint arXiv:1512.01100*, 2015.
- [49] P. Chen, Z. Sun, L. Bing, and W. Yang, “Recurrent attention network on memory for aspect sentiment analysis,” in *Proceedings of the 2017 conference on empirical methods in natural language processing*, 2017, pp. 452–461.
- [50] Y. Wang, M. Huang, X. Zhu, and L. Zhao, “Attention-based lstm for aspect-level sentiment classification,” in *Proceedings of the 2016 conference on empirical methods in natural language processing*, 2016, pp. 606–615.
- [51] J. Cheng, S. Zhao, J. Zhang, I. King, X. Zhang, and H. Wang, “Aspect-level sentiment classification with heat (hierarchical attention) network,” in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2017, pp. 97–106.
- [52] D. Ma, S. Li, X. Zhang, and H. Wang, “Interactive attention networks for aspect-level sentiment classification,” *arXiv preprint arXiv:1709.00893*, 2017.
- [53] R. He, W. S. Lee, H. T. Ng, and D. Dahlmeier, “Effective attention modeling for aspect-level sentiment classification,” in *Proceedings of the 27th international conference on computational linguistics*, 2018, pp. 1121–1131.
- [54] L. Li, Y. Liu, and A. Zhou, “Hierarchical attention based position-aware network for aspect-level sentiment analysis,” in *Proceedings of the 22nd conference on computational natural language learning*, 2018, pp. 181–189.
- [55] N. Majumder, S. Poria, A. Gelbukh, M. S. Akhtar, E. Cambria, and A. Ekbal, “Iarm: Inter-aspect relation modeling with memory networks in aspect-based sentiment analysis,” in *Proceedings of the 2018 conference on empirical methods in natural language processing*, 2018, pp. 3402–3411.

- [56] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” *arXiv preprint arXiv:1810.04805*, 2018.
- [57] B. Zeng, H. Yang, R. Xu, W. Zhou, and X. Han, “Lcf: A local context focus mechanism for aspect-based sentiment classification,” *Applied Sciences*, vol. 9, no. 16, p. 3389, 2019.
- [58] L. Yao, C. Mao, and Y. Luo, “Graph convolutional networks for text classification,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 7370–7377.
- [59] C. Zhang, Q. Li, and D. Song, “Aspect-based sentiment classification with aspect-specific graph convolutional networks,” *arXiv preprint arXiv:1909.03477*, 2019.
- [60] P. Zhao, L. Hou, and O. Wu, “Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification,” *Knowledge-Based Systems*, vol. 193, p. 105443, 2020.
- [61] C. D. Manning, M. Surdeanu, J. Bauer, J. R. Finkel, S. Bethard, and D. McClosky, “The stanford corenlp natural language processing toolkit,” in *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*, 2014, pp. 55–60.
- [62] R. M. Palau and M.-F. Moens, “Argumentation mining: The detection, classification and structure of arguments in text,” in *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, ser. ICAIL '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 98–107. [Online]. Available: <https://doi.org/10.1145/1568234.1568246>
- [63] J. L. Fleiss, “Measuring nominal scale agreement among many raters.” *Psychological bulletin*, vol. 76, no. 5, p. 378, 1971.
- [64] I. Rajapaksha, C. R. Mudalige, D. Karunaratna, N. de Silva, G. Rathnayaka, and A. S. Perera, “Rule-based approach for party-based sentiment analysis in legal opinion texts,” in *2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer)*, 2020, pp. 284–285.

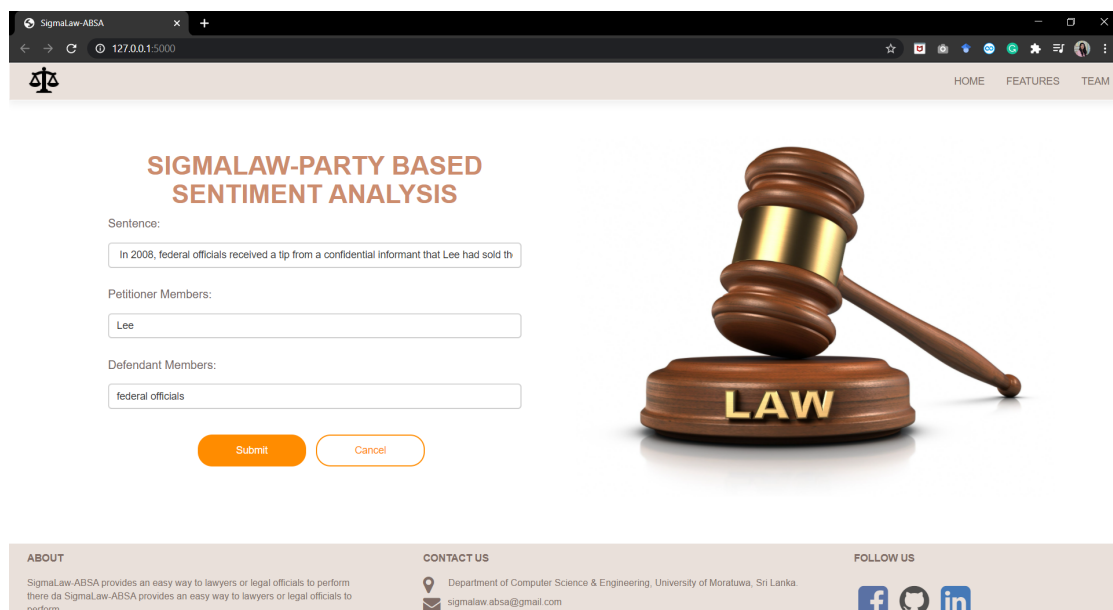
- [65] K. Clark and C. D. Manning, “Deep reinforcement learning for mention-ranking coreference models,” *arXiv preprint arXiv:1609.08667*, 2016.
- [66] —, “Improving coreference resolution by learning entity-level distributed representations,” *arXiv preprint arXiv:1606.01323*, 2016.
- [67] J. Zhou, Z. Zhang, and H. Zhao, “Limit-bert: Linguistic informed multi-task bert,” *arXiv preprint arXiv:1910.14296*, 2019.
- [68] M. Zhang, “A survey of syntactic-semantic parsing based on constituent and dependency structures,” *Science China Technological Sciences*, pp. 1–23, 2020.
- [69] D. Jurafsky and J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, 1st ed. USA: Prentice Hall PTR, 2000.
- [70] T. T. Thet, J.-C. Na, and C. S. Khoo, “Aspect-based sentiment analysis of movie reviews on discussion boards,” *Journal of information science*, vol. 36, no. 6, pp. 823–848, 2010.
- [71] Q. Liu, H. Zhang, Y. Zeng, Z. Huang, and Z. Wu, “Content attention model for aspect based sentiment analysis,” in *Proceedings of the 2018 World Wide Web Conference*, 2018, pp. 1023–1032.
- [72] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding.” ACL, Jun. 2019, pp. 4171–4186.
- [73] Y. Song, J. Wang, T. Jiang, Z. Liu, and Y. Rao, “Attentional encoder network for targeted sentiment classification,” *arXiv preprint arXiv:1902.09314*, 2019.
- [74] A. Mandelbaum and A. Shalev, “Word embeddings and their use in sentence classification tasks,” *arXiv preprint arXiv:1610.08229*, 2016.
- [75] G. Ratnayaka, N. de Silva, A. S. Perera, and R. Pathirana, “Effective Approach to Develop a Sentiment Annotator For Legal Domain in a Low Resource Setting,” *arXiv preprint arXiv:2011.00318*, 2020.

- [76] Y. Zhang, Q. Liu, and L. Song, “Sentence-state lstm for text representation,” in *ACL*, 2018.
- [77] P. Demotte, L. Senevirathne, B. Karunanayake, U. Munasinghe, and S. Ranathunga, “Sentiment analysis of sinhala news comments using sentence-state lstm networks,” in *2020 Moratuwa Engineering Research Conference (MER-Con)*, 2020, pp. 283–288.
- [78] K. Yadav, “A comprehensive survey on aspect based sentiment analysis,” *arXiv preprint arXiv:2006.04611*, 2020.
- [79] J. Zhou, J. X. Huang, Q. Chen, Q. V. Hu, T. Wang, and L. He, “Deep learning for aspect-level sentiment classification: Survey, vision, and challenges,” *IEEE Access*, vol. 7, pp. 78 454–78 483, 2019.
- [80] C. R. Mudalige, D. Karunarathna, I. Rajapaksha, N. de Silva, G. Ratnayaka, A. S. Perera, and R. Pathirana, “SigmaLaw-ABSA: Dataset for Aspect-Based Sentiment Analysis in Legal Opinion Texts,” *arXiv preprint arXiv:2011.06326*, 2020.
- [81] S. Gu, L. Zhang, Y. Hou, and Y. Song, “A position-aware bidirectional attention network for aspect-level sentiment analysis,” in *Proceedings of the 27th International Conference on Computational Linguistics*. ACL, Aug. 2018.
- [82] J. Pennington, R. Socher, and C. Manning, “GloVe: Global vectors for word representation.” Association for Computational Linguistics, Oct. 2014, pp. 1532–1543.

Appendix A

SigmaLaw-ABSA system provides functionalities to obtain the sentiment value for each legal party mentioned in the input sentence. User can input a sentence taken from a court case and petitioner and defendant party members mentioned in the sentence and then take the corresponding sentiment values as the result.

1. As shown in the Figure 5.1, input the sentence, petitioner party members and defendant party members and then click on SUBMIT button to get the result or click on CANCEL button to clear the form.

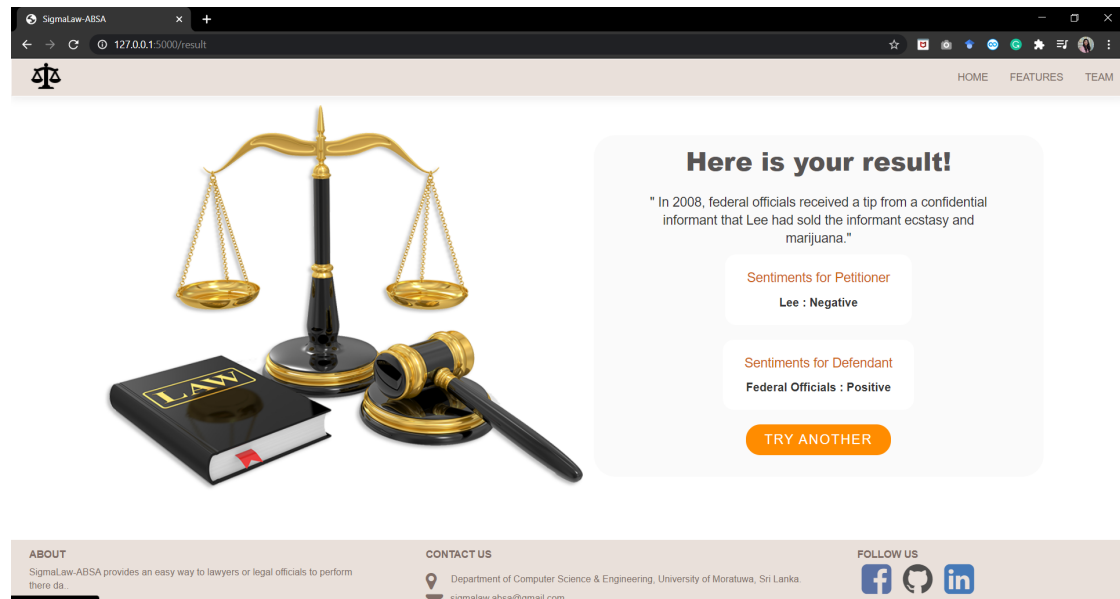


The screenshot displays the SigmaLaw-ABSA web application interface. The browser address bar shows the URL 127.0.0.1:5000. The page title is "SIGMALAW-PARTY BASED SENTIMENT ANALYSIS". The form includes a "Sentence:" field with the text "In 2008, federal officials received a tip from a confidential informant that Lee had sold th", a "Petitioner Members:" field with the text "Lee", and a "Defendant Members:" field with the text "federal officials". There are "Submit" and "Cancel" buttons. To the right of the form is a 3D rendering of a wooden gavel resting on a wooden block with the word "LAW" in gold letters. The footer contains an "ABOUT" section, a "CONTACT US" section with the address "Department of Computer Science & Engineering, University of Moratuwa, Sri Lanka" and email "sigmalaw absa@gmail.com", and a "FOLLOW US" section with social media icons for Facebook, Twitter, and LinkedIn.

Figure 5.1: Inputs of the system

2. The predicted results is shown in the result page (Figure 5.2) and you can try more

sentences by clicking on TRY ANOTHER button.



The screenshot displays the SigmaLaw-ABSA web application interface. The browser's address bar shows the URL `127.0.0.1:5000/result`. The page features a navigation bar with links for HOME, FEATURES, and TEAM. The main content area is divided into two sections. On the left, there is a 3D illustration of a golden scale of justice, a black book with 'LAW' written on it, and a gavel. On the right, a white box titled 'Here is your result!' contains the following text: 'In 2008, federal officials received a tip from a confidential informant that Lee had sold the informant ecstasy and marijuana.' Below this text, there are two sentiment analysis results: 'Sentiments for Petitioner' with 'Lee : Negative' and 'Sentiments for Defendant' with 'Federal Officials : Positive'. A 'TRY ANOTHER' button is located at the bottom of this box. The footer section includes an 'ABOUT' section, a 'CONTACT US' section with the address 'Department of Computer Science & Engineering, University of Moratuwa, Sri Lanka' and email 'sigmalaw.absa@gmail.com', and a 'FOLLOW US' section with social media icons for Facebook, GitHub, and LinkedIn.

Here is your result!

" In 2008, federal officials received a tip from a confidential informant that Lee had sold the informant ecstasy and marijuana."

Sentiments for Petitioner
Lee : Negative

Sentiments for Defendant
Federal Officials : Positive

TRY ANOTHER

ABOUT
SigmaLaw-ABSA provides an easy way to lawyers or legal officials to perform there da...

CONTACT US
Department of Computer Science & Engineering, University of Moratuwa, Sri Lanka.
sigmalaw.absa@gmail.com

FOLLOW US
f GitHub in

Figure 5.2: Output results

Appendix B - List of published papers

1. Conference Paper - "SigmaLaw-ABSA: Dataset for Aspect-Based Sentiment Analysis in Legal Opinion Texts"

Conference - 15th IEEE International Conference on Industrial and Information Systems (ICIIS-2020)

Best Paper Award (Computer, Embedded, and Intelligent systems and data Engineering Category)

Indexed in Scopus, IEEE conference proceedings

2. Conference Paper - "Rule-Based Approach for Party-Based Sentiment Analysis in Legal Opinion Texts"

Conference - 20th International Conference on Advances in ICT for Emerging Regions (ICTer)

Indexed in Scopus, IEEE conference proceedings

3. Journal Paper(Under review) - "Party-Based Sentiment Analysis of Legal Opinion Texts using a Linguistics Rule-Based Model"

Journal - International Journal on Advances in ICT for Emerging Regions

4. Conference Paper(Under review) - "A Deep Learning Model for Party-Based Sentiment Analysis in the Legal Domain"

Conference - 25th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-2021)

CORE Ranking - A