Conversational Humor Analysis: Developing Data, Annotation Schema and Models

Thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science
in
Computational Linguistics by Research

by

Vaishnavi Pamulapati
20161114
vaishnavi.p@research.iiit.ac.in

International Institute of Information Technology, Hyderabad
Deemed to be University
Hyderabad - 500 032, INDIA
Dec 2021
It is certified that the work contained in this thesis, titled “Conversational Humor Analysis: Developing Data, Annotation Schema and Models” by Vaishnavi Pamulapati, has been carried out under my supervision and is not submitted elsewhere for a degree.

Date

Adviser: Prof. Radhika Mamidi
To my family and friends.
Acknowledgments

“I want to thank me. I want to thank me for believing in me. I want to thank me for doing all this hard work. I want to thank me for having no days off. I want to thank me for never quitting.”- Snoop Dogg.

Jokes aside (though you’re about to read about Conversational Humor), I am proud of who I was, am, and will be. I wholeheartedly appreciate myself. That being said, the people I cherish played a quintessential role in shaping me.

First and foremost, I would like to express my gratitude towards Prof. Radhika Mamidi for being my counsel from my very first year. Be it course projects, when I was the Cultural Secretary, or a research student, you have always given me words of encouragement and guidance. This work would not have been possible without your unwavering support.

The courage to take on an unconventional choice stems from the guarantee that my family will support and encourage me no matter what. Collectively we have been through a lot and are now much closer to each other than I would have ever imagined. Cheers to all our efforts to make this happen.

Every time I have a nightmare or I need sound emotional advice, I turn to my sister from another mother, Samhitha. The strength that I have in me to deal with matters of the heart is all thanks to her. I will forever be indebted to you and the time you give me.

All those years in college, I would rush to your rooms, Geethu, Sanju, and Sravs. The walks we shared in the middle of the night, nimbu panis we gulped down in the hot summer, and the frequent stifled giggles in class make me miss you all till my heart aches with tears.

Niaaaa. If you did not live 2 minutes from my house in 2020, I would have gone clinically insane. All my worries and anxiety washed away when we would train like P.V. Sindhu. We are the badminton players, ra. Meghna thalli, I am so lucky to have found a friend in you that trusts and instills your belief in me when I constantly doubt myself.

The more I meet people, the more I realize how lucky I was to have you in my life for three years, Ravi. I’d like to thank you for always being by my side with a shoulder for me to weep my woes on. Seniors who I never thought I would be grateful for: Surya, Suhan and Riya. Without your help and your tales of experience, I would not have completed this thesis in the year 2021. This is a fact.

These five years of college surpassed every initial expectation I had. If there is anything I would go back to college for, it’d be the people. The same people that I am adamant on not leaving for life. I am an amalgamation of the people that I love and it is evident in the little gestures and language I use every
day. This way I constantly carry you with me and I would not have it any other way. The greatest lesson that 2020 has taught me is that the strength and significance of friendship withstand all inconsequential matters like money, appearance, or name.

I thank you all from the bottom of my heart.

Figure 1: Fin.
Abstract

Conversational humor (CH) is a sub-domain of humor where the participants (speakers or listeners) engage in different types of humor such as retorts or teasing for various purposes. There are shared complexities between this phenomenon and the larger domain of humor, but several features are unique to CH. To overcome these complexities, we focus on two aspects of comprehending CH. First, we formulate a schema for a methodological approach to annotating CH using a famous Telugu stage play, *Kanyasulkam*, as a medium of analysis. We then collect data on humorous and non-humorous conversations and perform experiments using the state-of-the-art NLP models to detect the occurrence of CH.

Contemporary work that focuses on shedding light on the purposes of CH by interlocutors considers only a few techniques or types. These analyses show the speakers’ intention, but it does not construct CH’s structure. In our work, we devise a schema that includes a hierarchical approach with different levels such as monologue/dialogue, benign/non-benign, and identifies the types and techniques involved in CH as well.

In the pursuit of teaching a machine a language, we must bear in mind that for a model to perform reasonably well in tasks of any domain, we must have ample data to feed the model. However, for low resource languages such as Telugu, this becomes a difficult task. For this reason, annotators are employed to help with providing metadata to a dataset of a specific domain so that the model can learn the patterns belonging to it and produce satisfactory results. To this end, a work of literature is fully annotated by A1 and A2 in our study. We then calculate the agreement between their annotations to give an objective measure of the validity of the hierarchical schema.

Many researchers in NLP have provided profound insights with the aim of detecting humor in various forms of text (tweets, movie scripts, jokes). Nevertheless, few studies have focused on text classification of conversational humor due to the complexity of the domain. We attempt to facilitate research in this direction by using several popular models to classify humorous and non-humorous conversations automatically. Recurrent Neural Networks (RNNs) such as Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (BiLSTM) learn sentence embeddings and are then used to classify text. In contrast, we use Text Graphical Convolutional Networks (GCN) that simultaneously learns the class a conversation belongs to and word embeddings based on word co-occurrence and document-word
relations. In order to make the most out of the well-acclaimed pre-trained models, we fine-tune Fast-
Text word embeddings and different BERT (Bidirectional Encoder Representations from Transformers)
models to generate sentence embeddings. We further use these models to classify text and compare their
performance based on standard evaluation metrics.
# Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Motivation</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Key Contributions</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Thesis overview</td>
<td>5</td>
</tr>
<tr>
<td>2 Background and Related Work</td>
<td>6</td>
</tr>
<tr>
<td>2.1 Introduction to Pragmatics</td>
<td>6</td>
</tr>
<tr>
<td>2.1.1 Violation</td>
<td>7</td>
</tr>
<tr>
<td>2.1.2 Flouting</td>
<td>8</td>
</tr>
<tr>
<td>2.2 Pragmatics of Humorous Conversation</td>
<td>8</td>
</tr>
<tr>
<td>2.2.1 Conversational Humor</td>
<td>8</td>
</tr>
<tr>
<td>2.3 Culture’s influence on perception of humor</td>
<td>9</td>
</tr>
<tr>
<td>2.4 Computational Humor Analysis</td>
<td>10</td>
</tr>
<tr>
<td>3 A Novel Annotation Schema for Conversational Humor using a stage play, Kanyasulkam</td>
<td>13</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>13</td>
</tr>
<tr>
<td>3.2 Related Work</td>
<td>14</td>
</tr>
<tr>
<td>3.3 Data and Annotation</td>
<td>14</td>
</tr>
<tr>
<td>3.4 Annotation Schema</td>
<td>16</td>
</tr>
<tr>
<td>3.4.1 Level-1: Non-Conversational (NC)/ Conversational Humor (C)</td>
<td>16</td>
</tr>
<tr>
<td>3.4.2 Level-2: Monologue (M)/ Dialogue (D)</td>
<td>18</td>
</tr>
<tr>
<td>3.4.3 Level-3: Benign (B)/ Non-Benign (NB)</td>
<td>18</td>
</tr>
<tr>
<td>3.4.4 Level-4: Types of conversational humor</td>
<td>19</td>
</tr>
<tr>
<td>3.4.4.1 Teasing (T)</td>
<td>19</td>
</tr>
<tr>
<td>3.4.4.2 Retort (R)</td>
<td>19</td>
</tr>
<tr>
<td>3.4.4.3 Banter (Ba)</td>
<td>20</td>
</tr>
<tr>
<td>3.4.4.4 Schadenfreude (S)</td>
<td>20</td>
</tr>
<tr>
<td>3.4.5 Techniques:</td>
<td>21</td>
</tr>
<tr>
<td>3.4.5.1 Dramatic Irony (DIm)</td>
<td>21</td>
</tr>
<tr>
<td>3.4.5.2 Satire (Str)</td>
<td>22</td>
</tr>
<tr>
<td>3.4.5.3 Sarcasm ( Src)</td>
<td>23</td>
</tr>
<tr>
<td>3.4.5.4 Fallacious Reasoning (PR)</td>
<td>23</td>
</tr>
<tr>
<td>3.4.5.5 Utilizing a Foreign Language (FL)</td>
<td>24</td>
</tr>
<tr>
<td>3.4.5.6 Allusion (A)</td>
<td>25</td>
</tr>
<tr>
<td>3.4.5.7 Profanity (P)</td>
<td>25</td>
</tr>
</tbody>
</table>
CONTENTS

3.4.5.8 Hyperbole/ Exaggeration (Ex) ........................................... 25
3.4.5.9 Other identified techniques (O) ....................................... 25
3.5 Disagreement Analysis ....................................................... 26
3.6 Conclusion and Future Work .............................................. 27

4 Data Collection Methods for Conversational Humor Recognition .................................................. 29
4.1 Introduction ........................................................................... 29
4.2 Humorous Data ..................................................................... 30
4.2.1 Source and Structure of data ............................................. 30
4.2.2 Filters to make humorous data homogeneous ................. 31
4.3 Non-Humorous Data ............................................................. 32
4.3.1 Source and Structure of data ............................................. 32
4.4 Attempts to homogenize humorous and non-humorous data .... 34
4.5 Conclusion ........................................................................... 36

5 Conversational Humor Recognition ............................................... 37
5.1 Introduction ........................................................................... 37
5.1.1 Outline of Methodology .................................................. 39
5.2 Text GCN ............................................................................ 39
5.2.1 Heterogenous graph ....................................................... 39
5.2.2 GCN for Text Classification ............................................ 40
5.3 FastText ............................................................................... 41
5.3.1 Word Embeddings .......................................................... 41
5.3.2 FastText Classifier .......................................................... 42
5.4 BERT ................................................................................... 43
5.4.1 Language Modeling .......................................................... 43
5.4.2 Linear Classifier ................................................................. 45
5.5 Results and Discussion ....................................................... 46
5.6 Conclusion ........................................................................... 48

6 Conclusions and Future Work ...................................................... 49
6.1 Conclusion ........................................................................... 49
6.2 Future Work ........................................................................ 50

Bibliography ............................................................................... 52
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Fin.</td>
<td>vi</td>
</tr>
<tr>
<td>3.1 Hierarchical schema for Conversational humor</td>
<td>16</td>
</tr>
<tr>
<td>3.2 Levels of Discourse</td>
<td>21</td>
</tr>
<tr>
<td>3.3 Distribution of Techniques by A1 and A2</td>
<td>27</td>
</tr>
<tr>
<td>4.1 A tweet’s reply</td>
<td>33</td>
</tr>
<tr>
<td>4.2 A tweet’s retweet</td>
<td>33</td>
</tr>
<tr>
<td>5.1 Methodology</td>
<td>38</td>
</tr>
<tr>
<td>5.2 Heterogeneous Graph; nodes denoted by ‘O’ are document nodes; black bold edges communicate edges between documents and words whereas light grey edges between words</td>
<td>40</td>
</tr>
<tr>
<td>5.3 Word n-gram features are embedded and averaged to form the hidden variable</td>
<td>43</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Humor Tagset</td>
<td>17</td>
</tr>
<tr>
<td>3.2 Cohen’s Kappa for Level-1</td>
<td>26</td>
</tr>
<tr>
<td>4.1 Statistics of Dataset</td>
<td>36</td>
</tr>
<tr>
<td>5.1 FastText nearest neighbors of word ‘aNxulO’</td>
<td>46</td>
</tr>
<tr>
<td>5.2 Performance of architectures implemented for Conversational Humor recognition</td>
<td>47</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Humor has been a phenomenon shared by all human beings since the beginning of our time. Since Plato’s times, many endeavors to comprehend humor use methodologies from diverse fields, including linguistics, psychology, sociology, neurology, and philosophy. Questions such as how humor is expressed, when it is triggered, or what causes some communities to find a joke humorous while others are offended, have made scholars explore humor.

Cognitive psychologists deem humor to arise from bisociation – perception of an idea in two frames of reference [52]. Social psychologists speculate on the interpersonal functions of humor [61]. Linguists had focused on the ambiguity of lexemes used in jokes or riddles [77]. Counseling psychologists have found that engaging in humor could facilitate insights [78]. Sociologists attempt to comprehend the root of jokes (power relations, authority figures) [43]. We deduce that the reasons, attributes, and functions of humor are as many as the depth of the field itself.

Initially, studies dedicated to humor were relatively few and considered laughter in a negative light. Plato believed that laughter was an irrational emotion that overrode self-control. A prominent philosophical thinker, Epictetus, advises in his writings, “Let not your laughter be loud, frequent and uncontrolled.” This notion of laughter and humor had pervaded early European culture as well. This emphasis that humor and laughter primarily originated from hostility led to the research and literature on ‘superiority theory’ that people found humor in condemning or looking down upon another. However, in the 18th-century, scholars began to examine the non-hostile occurrences of humor in human interaction and argued that humor arises when an expectation is violated (incongruity theory) [69].

Humor takes many forms in real life, whether it arises due to someone slipping on a banana peel (slapstick humor), in the form of knock-knock jokes (canned jokes), memes, the anecdotal narration of a funny incident, or quips made about a member present during a conversation (conversational humor). Many factors contribute to something to be deemed humorous. According to McGraw [65], psychological distance influences how humorous an individual finds a presumably humorous event/joke. For example, sexist jokes are humorous to those who are not affected by them [116]. Culture plays a sig-
nificant role in finding a joke humorous, too; it can signify whether an individual 'gets the joke' or 'is in the know' [104]. If a reference/allusion is made to a famous work of literature in a joke, it is natural that those exposed to it will understand and appreciate the humor.

A significant subset of humor research is dedicated to verbally expressed humor. It is the humor conveyed in language, as opposed to physical or visual humor, but not necessarily playing on the form of language [87]. Contrary to verbal humor, internet memes (visual) have been a recently developed form of humor but are vastly enjoyed by people worldwide, paving the way for a dedicated branch of humor research: Meme studies. Humor in AI is rapidly advancing, from generating jokes of a fixed structure [1] to generating memes based on popular templates[1] where the latter uses a combination of Image Processing and Natural Language Processing.

1.1 Motivation

The comprehension of humor may seem to be an insurmountable task, with all its complexities spanning numerous fields. In some ways, it may seem futile to dissect humor and understand humor using scientific methodologies [114], but we cannot deny its countless applications. In the domain of conversational humor, significant studies are dedicated to the pragmatics involved in communication, including interlocutors’ intentions to create in-group solidarity or out-group hostility, their shared knowledge base [54], and how individuals use conversational humor to establish their identity when power relations in a workplace are involved [96].

There are plenty of studies that provide a schema for humor, but devising taxonomies is a relatively new direction in conversational humor. Dynel [28] lays out different types and techniques of conversational humor after thoroughly examining several studies in verbal humor. However, the research study does not use a conversational humor dataset to verify the typology presented.

For a machine to learn conversational humor, a dataset containing several documents of conversational humor can be fed to it. This is feasible in languages such as English, where there is no paucity of data for almost any domain. However, for low resource languages such as Telugu, an annotated dataset with metadata is required for the machine to comprehend conversational humor. It can be argued that English data could be translated to Telugu and used for this task. This ignores two vital aspects. Machine translation tools are majorly trained on news articles. They can perform well on simple one-liners, which are universally humorous (due to global culture), but conversations are far more complex. The second aspect remains that Telugu culture-specific norms would not be represented. Hence, this calls for formulating an annotation schema for conversational humor for a machine to learn its patterns.

Virtual assistants such as Siri, Alexa, and Cortana are presently displaying humorous retorts but there is still a long way for such systems to understand and generate conversational humor. As opposed to annotating low resource data, which is a time-consuming task, if the system learned how to detect conversational humor automatically from data given, this would reduce manual intervention efforts. For this purpose, in this thesis, state-of-the-art NLP models such as Text GCN, FastText, and BERT are used to detect conversational humor. We analyze the edge that pre-trained language models such as the FastText library and BERT offer in terms of language understanding. Using these models, we experiment on the collected data of humorous and non-humorous conversations to further the endeavors in this domain.

1.2 Key Contributions

The main contributions presented in this thesis can be categorized into four categories which are briefed below.

- A hierarchical schema for a fine-grained annotation of Conversational Humor is presented. A prominent 19th century stage play from Telugu, *Kanyasulkam* is annotated. In order to validate the schema, two annotators A1 and A2 fully annotate the play. This is to substantiate the work across cultures at multiple levels. A total of 6,645 segments were to be annotated, resulting in 1,881 conversationally humorous segments.

- Using the Benign Violation Theory, the benignity or non-benignity of the segment is established. In addition to different types of humor, the techniques utilized by the speakers to generate Conversational Humor are identified. The four types of Conversational Humor identified were Teasing, Retort, Banter and Schadenfreude. Among the many techniques identified, Sarcasm, Satire, Dramatic Irony, Fallacious Reasoning, and Use of Foreign Language were few. Furthermore, the inter-annotator agreement is calculated to assess the accuracy and validity of the dataset. An in-depth analysis of the disagreement is performed to understand the subjectivity of humor better.

- A dataset of 8k+ conversations is contributed to the domain of humor research. Telugu jokes from various sources have been collected to provide a dataset of 6.1k+ jokes. As the central topic of this thesis is conversational humor, only jokes that have a conversational structure have been used, adding up to a total of 2k+ conversational jokes. For non-humorous data, tweets and their replies have been scraped using Twitter API. After preprocessing and further filtration, the dataset amounts to 6.2k+ non-humorous conversations.

- Several experiments have been designed to run on state-of-the-art NLP models to classify the conversations into humorous or non-humorous categories. Two types of NLP models have been...
used, pre-trained models such as FastText and BERT, and non-pre-trained models such as Text GCN. A comprehensive analysis is performed to gain an insight into the data as well as the models.
1.3 Thesis overview

- Chapter 2 provides the essential background and related research studies that cover pragmatics of language, Grice’s maxims, pragmatics of humorous conversation, role of culture in the perception of Humor, and work done computationally in the domain of Humor and Conversational Humor.

- In Chapter 3, we explore the hierarchical framework for Conversational Humor mentioned. Each level of the annotation schema is described with suitable illustrative examples. We analyzed several issues regarding the complexities of the domain and the role of culture in the perception of Humor.

- Chapter 4 describes our efforts in collecting, filtering, and pre-processing humorous data (jokes) and non-humorous data (conversations scraped from Twitter). We also discuss the measures taken so that the models do not learn the differences in syntax of both humorous and non-humorous data and further use this to classify the text.

- In Chapter 5, our methodology in training different types of NLP models for text classification is discussed. We compare various architectures to provide insights regarding their efficiency and mechanisms. We also discuss the advantages of using pre-trained models in comparison to models that require training from scratch.

- In Chapter 6, we conclude this thesis and briefly lay out a summary of the insights gleaned. We also provide possible directions for future work on Conversational Humor.
Chapter 2

Background and Related Work

2.1 Introduction to Pragmatics

Pragmatics is a sub-field of linguistics where the primary focus is the effect of context on language usage. It also focuses on other aspects of communication such as utterance meaning, intention, and inference [53]. Before we delve any further, the notion of ‘conversation’ must be defined. According to Richards [86], a conversation is a face-to-face oral interaction between two or more participants. However, a conversation involves more than information exchange between these participants. Several aspects are brought into a conversation such as shared assumptions and expectations about what the conversation is about, how it develops, and the contribution they are expected to make.

How participants organize a conversation indicates several aspects regarding their role in the setting or their intentions [35]. An individual learns the capability to navigate conversations, to learn constructive communication strategies from a young age by observing interactions or by erring and learning when their goal is not achieved efficiently.

On that account, a participant’s communicative competence can be defined as the adequate understanding of syntax, morphology, phonology of the language and the possession of social knowledge about the appropriate norms respective to the culture [45] to achieve a conversational goal. In order to be communicatively competent, a participant need not utter the literal meaning of what they intend to say. Due to the universal principles common to all participants in this world, parts of the utterance are implied and hence not said out loud [41]. Take this conversation, for example,

Frank walks into a bar, insults his friend Kermit and storms out without waiting for a response. Kermit looks at the bartender.

Kermit: What’s wrong with him?
Bartender: Are you free for a couple of hours?
The bartender implies that to list out what’s wrong with Frank, it will take a couple of hours to do so and hence asks Kermit if he’s free. Grice [41] argues that the reason we use conversational implicatures can be attributed to the ‘Cooperative Principle.’ It is presumed that the other participants are cooperatively navigating through the conversation to achieve their goals effectively. Four categories or maxims contribute to the cooperative principle: Maxim of Quantity, Maxim of Quality, Maxim of Relation, and Maxim of Manner [41].

- **Quantity**: *Make your contribution as informative as (neither more nor less informative than) is required.*
- **Quality**: *Try to make your contribution one that is true.*
- **Relation**: *Be relevant.*
- **Manner**: *Be perspicuous: avoid obscurity of expression, avoid ambiguity, be brief (avoid unnecessary prolixity), be orderly.*

An example of a conversation following the Maxim of Quality is given below.

John: *Do you know where the Eiffel Tower is?*
George: *It’s located in Paris.*

George does not contribute what he believes to be false and unsubstantiated, therefore follows the Maxim of Quality. The Maxim of Quantity is also followed as the contribution is as informative as required. In everyday conversation we do not obey the maxims of cooperative principle for different reasons. Grice [41] identified several types of not following the maxims. Two categories, violation and flouting, are the most relevant to this thesis.

### 2.1.1 Violation

When a speaker disobeys a maxim intentionally but does so in a quiet or unostentatious manner, it is called a violation of a Grice maxim. A speaker may violate a maxim for various purposes [115]. A speaker could deliberately mislead the addressee or lie, and the addressee is not aware that the speaker is not observing the maxim [41]. An example from *Kanyasulkam*,

Puuta: yeVkkada vAdu?
maXu: mIrannamaniRi yikkaderaMte ceVvinibeVttarugaxA!

**Translation:**

Puuta: *Where is he?*
Madhu: *You don’t listen to me, I have already told you that the person you’re looking for is not here!*
Context: Giri and Rama are hiding under Madhu’s bed. Puta is looking for Giri and the latter is hiding because he is scared of Puta. The audience, Giri and Rama know that Madhu is lying about Giri’s whereabouts. She violates the Maxim of Quality by not stating the truth.

2.1.2 Flouting

According to Cutting [23], flouting occurs when a speaker blatantly [41] violates a maxim with the intention to mislead and expects the hearer to recognize this violation, supplementing to the conversational implicatures. The flouting of a maxim can also produce a comic effect of irony or sarcasm. Take this conversation as example [58].

The child walks into the kitchen and takes a handful of popcorn.

Father: *I thought you were practicing the violin.*
Son: *I need to get the stand.*
Father: *Is it under the popcorn?*

The father’s first utterance is indirect and generates a conversational implicature that the son should be practicing. The son explains why he has not been. After which the father flouts the Maxim of Quality as he knows that the stand could not be under the popcorn, which is not new knowledge to the son too. The blatant violation of the maxim is to imply that the son should be practicing.

2.2 Pragmatics of Humorous Conversation

Grice’s maxims have been substantially pathbreaking in the field of pragmatics, calling for debates and discussion on their robustness. As seen in the above example, conversations exhibiting irony or sarcasm are considered to violate or flout a Grice’s maxim(s), therefore proving to be against the Cooperative Principle (CP) of conversation. But this analysis has been challenged by some scholars [6] [83] and modified by annexing to the maxims to encapsulate more experiences [7].

2.2.1 Conversational Humor

Dynel [28] differentiates between verbally expressed humor and conversational humor. Conversational Humor (CH) is a subset of Verbal Humor. Verbal humor exists in the verbality of what is being spoken, as opposed to Situational Humor. An example of the latter is physical slapstick (someone slipping on a banana peel), or practical jokes (chair removed from someone about to sit). Canned jokes (such as light-bulb jokes, knock-knock jokes) are a part of verbal humor, where they are not context dependent, i.e., they can be removed from a conversation and still perceived humorous. For example,

*I ate a clock yesterday; it was very time-consuming.*
Whereas CH heavily depends on the context, the participants’ behavior, personalities and so on. The above section depicts the dense complexities of the pragmatics of humorous conversation. An example from *Kanyasulkam*, the stage play:

giri: *ni mArtarikinannucUswe kittaxu. aMxucAwa ninnu PeVyil ceSAdugAni lekuMte nuvevevizti PeVyil kAvadavezvizti! awanikI nAkU yeVMxuku viroXaAM-vovCciMxo weVlisiMxA? awanu ceVppexaMwA wappulawadaka. axi nenu nyUsu pAparlo yekeSAnu. appatnuMcI nenaMte vAdikkadupudu.

Translation:

Giri: *Your schoolteacher cannot bear to see me well. That’s the reason he failed you in your test; because there is no possibility of you failing. Do you know why your teacher and I have our differences? All he teaches his students are mistakes. I reported this and got this published in a newspaper. Since then, he hates me.*

**Context:** If this utterance was removed from the ensuing conversation in the play, we could take it at its face value that Giri and the addressee’s schoolteacher had a dispute. The addressee is Giri’s pupil, Venka. However, if we take the context into account, we would deduce that Giri has made up a story, therefore lying to Venka, attempting to escape the accountability of not teaching Venka well. This irony (violating the Maxim of Quality) is recognized by the audience and causes Dramatic Irony [25].

### 2.3 Culture’s influence on perception of humor

An individual’s culture is an essential part of their everyday life, influencing their language, accent of a language, sense of humor, the way their interpersonal relationships are maintained, choice of humor when strangers are present [38], and stylistic choices when creating humor. Humor transcends all cultures [36], but the perception of humor can vary [61]. The latter is proved by numerous intracultural studies on humor perception.

Hiranandani and Yue [120] found that Indian students rated the importance of humor higher than their contemporaries, Hong Kong Chinese students. They also concluded that both groups laid more emphasis on self-enhancing and affiliative humor as compared to self-deprecating humor. These studies point to the importance of comprehending the intersection of culture and humor as they contribute to the understanding of human behavior.

In the domain of Conversational Humor, the interpersonal relationships between participants become crucial to analyze. Intracultural humor studies provide necessary insights to the people belonging to the relevant culture. Levisen [59] recognizes a key concept in Danish conversational humor called “sort humor”. He shows that sort humor is a collaborative and highly culturally specific concept which
reflects Danish communicative values in CH using suitable examples and analysis. An interesting humor research study investigates two groups of people: Danes (Danish people) working in France and the French working in Denmark [60]. Their work substantiates Levisen’s work [59] that Danes are more inclined and more open with their use of irony in the workplace as compared to the French who engage in wordplay during conversation.

Apart from intracultural and intercultural studies, there have been efforts to understand the topics used in CH universally. Quite a few studies have focused on analyzing the connection between power asymmetry and humor contributing to workplace culture. Workplace humor can strengthen the bond between colleagues [95], but it can also be used to establish or maintain power [94]. Many themes in humor are universal—for instance, sex, trickster tales, political satire, witty or unconventional responses from children, and joking about stupidity. As having a child, and a child grasping the realities and norms of the world over time is universal, this is perceived as humorous around the world, irrespective of culture as seen from the example provided (edited for brevity) below.

Son: Mommy, where do babies come from?
Mother: Well, they grow in your belly.
Son: But how do they get there?
Mother: Well, it’s a kind of miracle. God starts a seed there and it grows.
Son: Mommy, how does the baby get out? Does it pop out of your belly?
Mother: Well, it comes out from your private area.
Son: Mommy, THIS I gotta see.
Mother: Uh, no you don’t.
Son: Yes, Mommy, THIS I gotta see.
Mother: Well, maybe when you’re older you can be a doctor and deliver babies.
Son: But that’s impossible, I don’t even know how to drive!

The child does not possess the knowledge of childbirth. Therefore, he assumes that ‘delivering a baby’ is similar to ‘delivering food’, where someone drives a vehicle to do so. This gap in knowledge is humorous to those who understand the process of childbirth.

2.4 Computational Humor Analysis

One of the cornerstone studies in humor recognition using empirical methods is Mihalcea and Strapparava’s work [67]. For humorous data, 16k one-liners from various jokes websites were collected. Whereas for non-humorous data, different sources and domains such as proverbs, news headlines from Reuters and BBC National Corpus sentences were collected. Three sets of experiments were performed

[10]

on this data. In the first experiment, stylistic figures such as alliteration, antonymy and adult slang were considered to classify humorous and non-humorous data. However, in the second experiment, content-based features were taken into account and Naïve Bayes and SVM classifiers were utilized, producing high accuracy. Motivated by the latter, stylistic and content-based features were combined to classify the text, which improved the accuracy further.

Khandelwal, et. al [49] collected 10k+ tweets from different domains such as sports, entertainment and politics. The focus of this study was to detect humor in code-mixed tweets. Annotators categorized a tweet as humorous or non-humorous. Features such as N-grams, Bag of Words, Common tags and hashtags and classifiers such as Kernal SVM, Random Forest and Extra tree were employed. Sane, et. al [91] learned the word embeddings (using Word2Vec and FastText) of code-mixed data using around 200k tweets and then used these embeddings for classification of humorous and non-humorous data annotated by Khandelwal, et. al [49]. Several models were used for binary classification such as CNN, BiLSTM, and Attention-based BiLSTM.

Reyes and Rosso [84] computationally analyzed customer reviews from Amazon for verbal irony. They noted that several instances of ironic reviews went viral on social media such as Youtube, Wikipedia, and BBC, thus inadvertently marketing the product and reaching far more people, which expressed the commercial value of this research. Although they considered each review a standalone utterance for irony detection and hence was not context-dependent, features used for detection such as funny profiling, positive/negative profiling, and pleasantness profiling established this as a unique contribution to the field.

Joshi, et. al [47] manually annotated conversations from the famous TV show Friends and used two sequence labeling algorithms to detect sarcasm. Instead of considering one utterance where sarcasm lies, this study looked at the previous utterances as well. An example from their paper,

Chandler is at the table. Ross walks in, who looks very tanned.

Chandler: Hold on! There is something different.
Ross: I went to that tanning place your wife suggested.
Chandler: Was that place... The Sun?

The sarcasm in the last utterance is not perceived without the background contained in the previous utterance. Unigrams of spoken words is used as a lexical feature, positive and negative score of the utterance and its antecedent are used as conversational context features and speaker of the utterance and the speaker of the previous utterance are used as speaker context features.

Chen and Lee [18] collected Ted Talk transcripts for the prediction of audience’s laughter. In the transcripts, the last utterance before the transcript marking “(Laughter)” was labelled as humorous. Semantic structural features such as incongruity, ambiguity, interpersonal effect and phonetic pattern
Semantic distance features which were learned from a KNN were also utilized. An end-to-end CNN was used to predict laughter or non-laughter of the utterance. Using varied-sized filters and dropout regularization, CNN’s performance was improved, achieving 56% of laughter detection.

Speech from a popular TV show called Daria was used for the task of sarcasm detection [80]. Prosodic features such as pitch, energy, speaking rate and deltas were utilized. Using logistic regression as a classifier, 80% accuracy was achieved and found that pitch was the crucial feature of sarcasm. However, this research study has several limitations such as, a relatively small dataset (150 total sentences, half comprising sarcastic, while the other half non-sarcastic) and utterances by one speaker only.

To detect humor or techniques such as sarcasm, tweets’ hashtags such as “#sarcasm”, and “#sarcastic” were used [39]. Posts on Reddit and their comments have been used to detect irony in a conversation [109]. Four subreddits (user communities to discuss a certain topic) were considered for the dataset. (a) progressive[^2] and conservative[^3] political views as topics of subreddit (b) atheism[^4] and Christianity[^5] subreddits. The replies/comments were examined along with the main thread, to detect verbal irony. Features such as sentiment (positive or negativity of the reply and post), the topic of the subreddit (one of the four), noun phrases extracted from comment texts, and finally noun phrases from the comments and the thread. Using SVM as a classifier, although a hike was seen on the overall recall, a significant improvement in precision was not observed.

[^2]: https://www.reddit.com/r/progressive/
[^3]: https://www.reddit.com/r/Conservative/
[^4]: https://www.reddit.com/r/atheism/
[^5]: https://www.reddit.com/r/Christianity/
Chapter 3

A Novel Annotation Schema for Conversational Humor using a stage play, Kanyasulkam

3.1 Introduction

Humor and its dependence on society and culture have been the focus of research since times immemorial [81]. From finding theories to define humor [82] [66] [9] to an analysis of the perception of humor in jokes [81], humor studies have been proved to be an essential aspect of linguistic as well as sociological, psychological and philosophical research. Many papers discuss types of humor [28] [3] [13], but this work stands apart. It focuses on creating an annotation schema for conversational humor with a stage play as the medium of analysis while claiming that this schema can be used across languages. Conversational humor is the spontaneous or pre-constructed interactional humor. The interlocutors intend to amuse the listener directly or shift to a humorous frame where there is humor beyond what the literal verbalizations convey [28]. Stage play is chosen as the medium of analysis since ‘conversational humor’ is an umbrella term that covers various semantic and pragmatic types of humor that occur in interpersonal conversation, both real-life and fictional [28].

There are key differences between plays and other forms of discourse, like transcribed recordings of actual conversations or novels, justifying our use of stage play in this thesis. The differences include but are not limited to pauses, pause fillers, and discourse markers as essential features of characterization in a play, unlike their use in actual conversations. In a play, there is more character-character interaction than in novels, which have more narration from one point of view [106]. But the annotation schema presented here does not restrict its application on plays alone but can also cover novels, TV shows, movies, etc; essentially any genre that involves a premeditated conversation.

This work also focuses on how humor’s form and function are influenced multiculturally by annotating one of the most famous plays of the Telugu culture, Kanyasulkam. Studies show that culture plays a vital role in conversational humor in some distinct ways like the need for shared knowledge and standard references, and others more indirect, like how the importance given to language awareness by
any culture dictates the preference for wit and linguistic play [70]. *Kanyasulkam* is a play set in the 19th century Vijayanagaram which uses humor to talk about the social evils prevalent in the society. However, while the author talks about child marriages, widow re-marriage, and the Nautch question, we also see him discuss customs and traditions, superstitions, use of English and the fascination towards it, etc. Thus, making it culturally relevant and further justifying the use of *Kanyasulkam* in validating the role of culture on humor.

Persona identification is an important application of the schema proposed. For instance, if character A has a tendency to sarcastically tease character B on most occasions, we gain an insight into A’s sense of humor (SOH) as well the social function performed by A. While there have been several studies that suggest that an SOH indicates positive personality traits such as self-actualising, self-acceptance, and others [62] [4], the social function performed by A also provides an understanding of A’s overall role in the story, therefore the character’s persona.

### 3.2 Related Work

Interest in the study of humor has faced steady growth since 1970 [63]. This interest in humor studies has led to a great deal of research on humor types and functions. In his paper on the issues in conversational joking, [74] talks about the structure of humorous discourse, the forms of conversational humor and its interpersonal functions, i.e., aggression vs. rapport. Two of Marta Dynel’s studies, one based on a popular English sitcom, Friends [29] and another on the sitcom, House [58], are deemed relevant to this study. While the former analyses cultural references, the latter attempts to extract universal communicative phenomena that cause humor.

Dirk Delabastita [24] presents in her work, an overview of the humorous scenes with bilingual and translation-based situations from Shakespeare’s plays. Levisen [58] uses Natural Semantic Metalanguage to compare the Danish concept of ‘sort humor’ (a highly culturally specific way of Danish communication) and the English, ‘black humor.’ To recognize humor and irony in tweets, Antonio Reyes et al. [83] analyze humor and irony to recognize these concepts in tweets. Agnese Augello et al. [10] have worked on building a chatbot that recognizes and generates humorous expressions. There have been continuous efforts in the field of computer science for the comprehension [14], detection [105], and production [44] of conversational humor.

### 3.3 Data and Annotation

In the play, to mark the entrance/exit of characters or the end of a scene, phrases such as ‘characters exit’ are omitted from the dataset as it does not contribute to the conversation between characters. For instance,
The full text of the Telugu play, Kanyasulkam, is annotated by two people, A1 and A2. For the preprocessing of the data, the whole text was split first by each character’s dialogue, and each utterance by the character was further split into single sentences/segments. In the presence of poems, lists, etc. the utterance remains as is, and this final output is used for annotation giving a total of 6645 segments to be annotated. After developing the gold standard corpus, 2710 utterances were classified as humorous, 1782 were given the tag dialogue, 1881 conversational, and 892 benign. The annotation was done with appropriate checkpoints after every 2000 segments to identify any new techniques or revise the schema.

Kanyasulkam was first published in 1892, approximately 130 years ago. There are many differences between the classical text and the text that the annotators are exposed to in the present day. Firstly from a linguistic perspective, archaic words were abundant throughout the text. The translation for a few words is unavailable on Google Translate thus far. ‘Aboru’ (translation: ‘word’ in the usage of the phrase ‘I give you my word’). Secondly, the cultural shift was considerable. Allusions/references made in the text are mostly to epics such as Ramayana and Mahabharata. But in the present day allusions are made to movies, songs, celebrities, and so on. For instance,

**girisam: nalacariwralo xamayaMwi reVMdopeVMdli sAtiMcinapaxyaM caxuvu**

**Translation:**

Girisam: Read the poem about Damayanti’s second marriage written in *Nalacharithra*.

**Context:** Widow remarriage was prohibited in the then day. Girisam is attracted to Bucchamma who is a widow. To convince her to remarry, he asks his pupil to read a poem that describes Dhamayanthi’s (Mahabharata character) second marriage.

Other than references, the text is embellished with many obsolete proverbs such as

**rama: veVrYrYikuxiriMxi, rokali walaki cuttamannAtta!**

**Translation:**

Rama: You have gone crazy, it’s like ‘he asked you to wrap a pestle around your head!’

**Context:** Madhuravani asks Rama to cancel the wedding between a nine-year old and an old man. Rama then calls her crazy as he has already prepared all the arrangements for the marriage.

To overcome these challenges the annotators relied on their parents’/grandparents’ knowledge to help with the translation and meaning of the play.
3.4 Annotation Schema

**Types:** Teasing(T), Retort(R), Banter(Ba), Schadenfreude(S)

**Techniques:** Dramatic Irony(DIrn), Sarcasm(Src), Satire(Str), Fallacious Reasoning(FR), Exaggeration(Ex), Use of foreign language(FL), Allusion(A), Profanity(P), Other stylistic figures(O) (refer to Table 3.1). *Note:* A segment can be annotated with more than one technique. The complete tagset can be seen in Table 3.1

3.4.1 Level-1: Non-Conversational (NC)/ Conversational Humor (C)

This study uses Dynel’s [28] conceptualization of conversational humor. As the dataset used here is a play, it is of primal importance to note that the speaker’s intent may not be to cause humor. However, a third-party present, or the metarecipients of the conversation, the audience/readers, may find it humorous. It is a common phenomenon to make the audience laugh at the expense of a fictional character. This study makes a distinction between conversational and non-conversational humor. In the latter, the humor does not exist in the realm of verbality but rather in the domain of the situation (slapstick humor, a character’s trait such as miserliness, stupidity, etc.) For example, when a cowardly character is badmouthing his rival and the latter appears just then.

<table>
<thead>
<tr>
<th>Tag</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Conversational</td>
<td>Conversation has one speaker and two listeners.</td>
</tr>
<tr>
<td>NC</td>
<td>Non-Conversational</td>
<td>The Three Stooges getting poked in the eye or thrown pies at their faces.</td>
</tr>
<tr>
<td>M</td>
<td>Monologue</td>
<td>Only one speaker present and no listeners.</td>
</tr>
<tr>
<td>D</td>
<td>Dialogue</td>
<td>Conversation has one speaker and three listeners.</td>
</tr>
<tr>
<td>B</td>
<td>Benign</td>
<td>[A short person can’t reach a shelf by a wide margin] A friend says, “If only you were an inch taller.”</td>
</tr>
<tr>
<td>NB</td>
<td>Non-Benign</td>
<td>“The woman who is yelling in the street is a rascal that bites men”</td>
</tr>
<tr>
<td>T</td>
<td>Teasing</td>
<td>[A woman spills her drink] Her boyfriend says, “Let me grab a sippy cup for you”</td>
</tr>
<tr>
<td>R</td>
<td>Retort</td>
<td>“I’m sorry but I don’t speak bullshit.”</td>
</tr>
<tr>
<td>Ba</td>
<td>Banter</td>
<td>A series of teases and retorts between speakers.</td>
</tr>
<tr>
<td>S</td>
<td>Schadenfreude</td>
<td>“Somebody stole my lunch out of the fridge at work today. The worst part about it... I’m working from home.”</td>
</tr>
<tr>
<td>DIrn</td>
<td>Dramatic Irony</td>
<td>[A character is known to be promiscuous] He says, “None can be loyal to a woman as I am”</td>
</tr>
<tr>
<td>Sar</td>
<td>Sarcasm</td>
<td>[Torrential rain on an expected sunny day] “Oh what warm weather!”</td>
</tr>
<tr>
<td>Str</td>
<td>Satire</td>
<td>“People say jokes are dead. But one can be found alive and kicking in the White House.”</td>
</tr>
<tr>
<td>FR</td>
<td>Fallacious Reasoning</td>
<td>“I never generalize because everyone who does is a hypocrite.”</td>
</tr>
<tr>
<td>Ex</td>
<td>Exaggeration</td>
<td>“How are you still hungry? You have a bottomless pit for a stomach.”</td>
</tr>
<tr>
<td>FL</td>
<td>Use of Foreign language</td>
<td>[A mother asks her son to come home] He replies, “Je ne comprends pas!”</td>
</tr>
<tr>
<td>A</td>
<td>Allusion</td>
<td>“Don’t act like Romeo in front of her!”</td>
</tr>
<tr>
<td>P</td>
<td>Profanity</td>
<td>“The idiotic excuse of a brother I have has no sense of decency!”</td>
</tr>
<tr>
<td>O</td>
<td>Other identified techniques</td>
<td>“She was as tall as a six-foot-two-inch tree.”</td>
</tr>
</tbody>
</table>

Table 3.1: Humor Tagset
3.4.2 Level-2: Monologue (M)/ Dialogue (D)

The notion of “dialogue” is taken in the Socratic dialogue sense: a conversation between two or more people. In contrast, there are several definitions and types of monologues present: Dramatic monologue [34], soliloquy [97], and inner monologue [71]. This study defines a monologue as “utterances by a single person/character in real-life/fictional with the assumption by the speaker that there are no listeners present to hear their thoughts”.

The distinction between humor found in a dialogue and a monologue is made because it is recognized that if a person speaks to themselves with no listeners present, it gives rise to certain types and techniques of humor compared to those occurring in a dialogue. Take retorts, for instance. A retort is a sharp or witty remark in response to another’s utterance [89]). Hence, for this type of conversational humor, the first turn in an adjacency pair must transpire [89].

3.4.3 Level-3: Benign (B)/ Non-Benign (NB)

The Benign Violation Theory examines the intersection at which the listener perceives a violation in a joke [64]. A joke is not a passive entity but is expressed by a person and perceived by another. Hence, the relationship between the joke-teller, the joke, and the joke-listener must be explored [48].

In conversational humor, there is no notion of a ”joke” (”canned joke” in [28]). Humor is caused by the interlocutor’s spontaneous speech that may or may not be humorous to the listeners present. For instance, in derision, reprimands, or put-downs, the speaker or listener may not find it humorous; instead, a third-party present (or an audience) may find the utterance humorous. In this study, it is important to note that the relationship between the author of the play and the audience is not examined, but the latter is given the role of a passive listener.

karata: bAvA yIsammaMXaM ceswe nI koVMpakI aggeVtteswAnu.
agni: vIlYlammA SiKAwaraga, prawIgAdixakoVdukU wiMdipowullAga
nAyIMuijeri nannanevAIYe

Translation:

Karata: Brother-in-law, if you agree to this proposal, I will set your house on fire.
Agni: (An expletive directed at Karata’s mother), every son of a donkey, comes to
my house to eat like a glutton and ends up criticizing me.

Context: Here, the interlocutors present in the scene do not find Agni’s utterance humorous as he only intends to ridicule Karata and Venkamma. However, the metarecipients, the audience, are bound to find it amusing [28].

This study augments the BVT by modifying the factors by which a joke can be labeled as benign: (a) two contradictory norms of the relevant culture (b) a weak commitment to the violated norm, or (c)
the social distance between the interlocutors and the content of what is uttered (d) the intention of the humor causer understood by the listener whether benign or not [111] [112]. By these four conditions, the above example is labeled non-benign as it goes against the salient norm of respecting a guest, and there exists no norm that states to insult a guest in the Telugu culture blatantly. Furthermore, Agni’s intention is to solely deride his guest’s behavior.

3.4.4 Level-4: Types of conversational humor

3.4.4.1 Teasing (T)

In this study, teasing is considered to transpire when the speaker intends to be playful, to only nip at the present listener non-aggressively. The main objective is to develop/strengthen the bond between the speaker and the listener(s). Other than this benevolent intention, the speaker also uses an element of “pretense” to tease [21] [28].

\[
\text{maXu: anyAyaM mAtalu AdakaMdi, Ayana yaMwa caxuvukunnAdu, Ayanaki yaMwapraKyAwi vuMxi! nedorepo goVppa vuxyogaM kAnEyyuMxi.}
\]

Translation:

Madhu: Don’t be unfair. He is a very learned man. He has a lot of fame as well. Very soon, he’ll land himself a good job.

Context: Madhu teases her client Ramappanthulu by praising another client of hers, Girisam. Rama detests and is jealous of Girisam. Knowing this, Madhu seeks to elicit a reaction by exaggerating (a common technique used in teasing) Girisam’s strength playfully.

3.4.4.2 Retort (R)

A retort takes place at the second turn in an adjacency pair where the purpose is to out-challenge or outwit the other interlocutor(s) [28] of the conversation by making a quick comeback (utilizing the other’s behavior, personality, past, etc.)

\[
\text{lubXA: mAmagAru hAsyAnikaMtunnArugAni, ninnoVxulwArA?}
\]

rAma: alAgaddi peVtuMdi!

\[
\text{maXu: [...]}
\]

\[
\text{gaddi gAdixalu wiMtAyi;manuRyulu winaru.}
\]

Translation:

Lubdha: Uncle is just pulling your leg. Do you think he will forsake you?

Rama: Teach her a lesson like that (Idiom with the literal meaning of ‘feed her grass’)

Madhu: [...] Grass is eaten by donkeys, not people.
**Context:** Rama is reading a letter written by Girisam where the latter refers to him as a donkey. In response to Rama’s suggestion of teaching Madhu a lesson, she mocks him indirectly by referring to the letter when she says, “Grass is eaten by donkeys, not by people”.

### 3.4.4.3 Banter (Ba)

If there is a continuous exchange of retorts and teasing in a multi-turn conversation, it is called banter. It is important to note that Banter cannot be a hierarchical category encompassing Retort and Teasing as they can also occur independently.

maXu: wAkattuvswuva wappiMcuku pArIpoweno? kukkanA, nakkA, kAxugaxA goVlusulvesi kattadAniki?
maXu: valalo muwyapu cippalupadiwe lABaMgAni, nawwagullalupadiwe mowacetu.
karata: yaMwasepU dabbu, dabbenA?
snehaM, valapU, anevi vuMtAyA?
maXu: snehaM mIlAAtivArIcota

**Translation:**

Madhu: What if you run away after pawnng it?
You aren’t a dog or a fox to tie you with chains.
Karata: Can any living being be freed from your trap?
Its hold is stronger than that of any chain.
Madhu: Only if pearls are trapped, it is of any use. Getting a hold on rocks/shells will only increase my burden.
Karata: Why are you always concerned about money? What about friendship, justice, etc.?
Madhu: Friendship with people like you (with sarcasm)

**Context:** Madhu is hesitant to depart from her necklace, which Karata is asking for. Karata teases her by flattery and hopes it will help in achieving his goal. However, Madhu retorts by indirectly comparing him to a weed/stone. Subsequently, in response to his reprimand that she always talks about money, she retorts using sarcasm once more.

### 3.4.4.4 Schadenfreude (S)

Schadenfreude is a German word that refers to the pleasure derived from another’s misfortune. It is the “malicious joy” evoked by the downfall of others, mostly high achievers.
Figure 3.2: Levels of Discourse

UTa: Ave VXavavuMte nAkeM kAvAli, vuMdakuMte nAkeM kAvAli.
Adu nIkiccina yiravayi rUpAyaiU yicceVy.
MaXu: yavadi kiccAvo vANNe adagavamMnA.
UTa: ve VXavakanabadiwe sigapAyixIsi clpurugattawo moVwwuxunu, yeVkkadax-AcAvevizti?

Translation:

Puta: I couldn’t care less about that idiot’s whereabouts. Just give me the 20 rupees that he gave you.
Madhu: Ask the person you gave it to.
Puta: I will cut his hair and thrash him with a broom if I find him, where did you hide him?

Context: Puta comes to Madhu’s house, searching for Girisam, who has run away with her money. On getting no help from Madhu in finding his whereabouts, Puta is immensely angered, and humor is found in Girisam’s plight.

3.4.5 Techniques:

3.4.5.1 Dramatic Irony (DIrn)

In a stage play setting, there exist two or more levels of discourse (refer to Fig. 3.2), the author-audience/reader, and the character-character level [99]. When the character is portraying pretense with one character, another character may or may not be in on it, but the readers necessarily are. Hence, other than the knowledge that exists between the characters, the audience is also privy to knowledge only they possess [55].

girI: AlrEt - gAni - nAkikkada cAlA vyavahAramulalo naRtaMvaswuMx - mu-nasabugAri pillalki Salavullo pATAIcEVPiweV PipI rupij yiswAvazmnAru; ayinA nI viRayavEz yaMwa lAs vaccinA nenu ker ceVyyanu.

Translation:
Giri: It’s alright. But I’ll incur a lot of losses here. The village head has promised to give 50 rupees for tutoring his kids over the vacation. However, I do not care about any loss when it comes to you.

**Context:** Girisam lies that he has been offered a valuable job opportunity but that he will reject it as he genuinely cares for his pupil Venka. The audience knows that this is untrue as, before this conversation, Girisam was plotting to take advantage of Venka’s economic resources.

As mentioned above, a requisite component of teasing is that of pretense. Nevertheless, teasing is not mandatorily marked with dramatic irony as there is a difference in intent. In dramatic irony, the intent is to dupe the listener by pretending to have values, attributes, etc. that the speaker does not possess. Whereas, in teasing, the motivation is to reduce the social distance or to poke fun at the listener benignly.

### 3.4.5.2 Satire (Str)

Satire has been defined as the ridicule of a subject (a person, situation, or an institution) to point out its faults [12]. It does not need to be present at the scene of action.

**SiRyudu:** ArneVllalomAtu poVswakaMpattukuMte koVvwaSlokAlu pAwaSlokAlu oVkkalA gkanapadAvyi. mAguruvugAriki xoVMdakAya kUra yiRTaM lexu, guruvugAri peVIYIAM peVratlo xoVMdapAxuMxani rojU AkUreV voVMduwuMxii. bawikunnavaIYla yiRTaveVz yiA yeduswUMte caccinavAdi yiRTAyiriRTAlwo yeMpani? _ylcaxuvikkadiwo cAliMci girISaMgAri xaggira nAlugiMgilIRu mukkalu nercukuMrAnu._

**Translation:**

Student: If I open my books once in 6 months, then the poems I have already learned, and the new ones all look the same. My teacher does not like ivy gourd curry. But his wife makes the same curry everyday owing to the ivy gourd plant in their house. If the likes of a person who is very much alive are not cared about, how do the likes of someone dead matter? I should stop these lessons here and learn a few English words from Mr. Girisam.

**Context:** Karata’s student is asked to learn a poem by heart where the poet talks about his likes and dislikes of flowers and nature. The student is fed up by this mode of learning and feels it is pointless to learn about a dead person’s likes and dislikes when his own guru’s likes are not cared about by the latter’s wife. This example is a satire on the education being provided to the student by Karata. The element of satire being used here is Chatuvu (ridicule).
3.4.5.3 Sarcasm (Src)

The difference between irony and sarcasm is fuzzy and is often misunderstood, given that they are inevitably bound to each other. However, the relationship between them remains unclear to native speakers but is highlighted when a comparison is drawn between cultures or linguistic communities [76]. Attardo, in his study, cites that sarcasm is an overly aggressive type of irony with more explicit markers or cues and a clear target [7].

Translation:

Clerk: What is his surname and address?
Bhima: What is it, man?
Agni: His name is Girisam. I do not know anything else.
Kale: Great. Sounds good. Let’s get it published in the newspaper that someone kidnapped Mr. Agni’s daughter. And hence, anyone who knows his name and village should immediately inform us.

Context: Agni goes to register a complaint against Girisam who runs away with the former’s daughter. When Agni states that he knows only the first name and nothing else, the officer sarcastically praises him and suggests that it would be great to publish this news in the Gazette and ask the public’s help to get to know Girisam’s details.

3.4.5.4 Fallacious Reasoning (FR)

A fallacy is defined as an argument that has faulty reasoning [37] either by intentional pretense by the speaker or by genuine ignorance.
Translation:

Girisam: Very good. If marriage is a good thing, and since the more you do, the more you achieve, a young girl should be married to an old man and once he dies, another man and if he dies, then another one and so on while collecting a thousand from the first guy, then the next, then another, like butter on bread and bread on butter, collect all the Kanyasulkam (bride price) and finally if she gets married to a wise guy like me, isn’t that enjoyable?

Context: Girisam pretends to agree that selling young girls for marriage is good for society when widow re-marriage is allowed. He argues that for every man that dies naturally with old age over time if the child is married and re-married to other older men, the father of the child gets money until the girl can marry a sensible person like Girisam. It is evident that this is an example of "non sequitur" fallacy, where the premises are true, but the conclusion is false.

3.4.5.5 Utilizing a Foreign Language (FL)

Several studies have attempted to understand the motivations for using a foreign language to produce humor [100] [50]. Grosjean [42] states that situations, messages, attitudes, and emotions influence foreign language use. In Kanyasulkam, English is used sporadically only by one character, Girisam, to achieve his objective: to portray and distinguish himself among the characters as well-educated.

agni: oVkkaxammidl yivvanu.
ixAMwA topI vyavahAraMlA kanapaduwuMxi.
karata: [...] 
girISaM: xislj bArbaras, cUCArAmdI, jeVMtilmen anagA peVxamaniRini yalA aMtunnAro!

Translation:

Agni: I will not spare even one penny
Karata: [...] 
Girisam: This is barbarous. Did you see how he is talking to a gentleman, meaning, learned person!

Context: Upon being accused of cheating by Agni, Girisam is angered. Here, knowing fully well that the listeners do not understand English, Girisam still chooses to talk in English and then condescendingly explains what he means. He does this to establish superiority over others as people who knew English in those times were held in high regard.
3.4.5.6 *Allusion (A)*

Direct or indirect reference to an object or circumstance from a different context is defined as an allusion [72].

"karata: Ayanexo kurYrYavAdiwo yiMgilIRu mAtaMte puccakAyalaxoVMgaMte bujAIwaduvuzkunnattu nmlxα peVttukumAveM?"

**Translation and context:**
Karata: If he is talking to his student in English, why are you getting involved like the watermelon thief rubbing his shoulders (idiom) meaning, why are you letting yourself be caught red-handed by getting angry and proving that you do not understand a word of it?

3.4.5.7 *Profanity (P)*

Profanity is defined as language that is considered as strongly impolite, rude, or socially offensive [31].

"girisam: rAskevl vulakalexu palakalexu sarekaxA moVхаM pakkaki wippi kadup-pagiletattu navvuwunnAdu."

**Translation and context:**
Girisam: Not only did that rascal fail to support me during my lecture, he turned to his side and laughed almost until his stomach burst.

3.4.5.8 *Hyperbole/ Exaggeration (Ex)*

It is the representation of an entity as more dramatic, better, or worse than it really is. Hyperbole is a figure of speech using exaggeration.

"girisam: nene xAni hajbеVMdnEvuMte, nilabaddapAtuna nI waMdrini rивAlvarwo RUt ceSivuMxunu."

**Translation and context:**
Girisam: If I were her husband, I would have shot your father with a revolver from where I stood.

Although exaggeration necessarily has a pretense factor, any segment where exaggeration is identified, ”dramatic irony” is not marked.

3.4.5.9 *Other identified techniques (O)*

Such as simile, metaphor, etc. are also marked during annotation.

"maXu: catlaki cAva nalupu, maniRiki cAva weVlupU. (simile)"

**Translation and context:**
Madhu: A person’s death is marked by white, like how a tree’s death is marked by black.
3.5 Disagreement Analysis

The validity of the tag set and their definitions are measured using Cohen’s Kappa ($\kappa$) [22]. Although the annotators were asked to mark all levels of the hierarchical scheme, the inter-annotator agreement (IAA) for level 2 (Monologue/Dialogue) was not measured as the definition for these categories provided no ambiguity.

The annotation for level 1, Conversational vs. Non-Conversational Humor, gave a Kappa value of 0.48 (moderate agreement). The disagreement emerged could be attested to the variation in the perception of humor (Table 3.2). For instance, A2 could have found the character’s trait (Non-conversational) humorous, whereas A1 identified a verbal technique in the speaker’s utterance, causing disagreement. Annotation of level 3, Benign vs. Non-Benign Humor gave a Kappa value of 0.42 (moderate agreement). The disagreement exhibited can be due to the difference in perception of the benignity of the utterance. A1 could be aware of a salient norm that can be violated (Section 3.4.3), whereas A2 is not producing disagreement.

<table>
<thead>
<tr>
<th>A1</th>
<th>A2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Verbal</td>
</tr>
<tr>
<td>Verbal</td>
<td>1559</td>
</tr>
<tr>
<td>Situational</td>
<td>191</td>
</tr>
<tr>
<td>Null</td>
<td>290</td>
</tr>
<tr>
<td>Total</td>
<td>2040</td>
</tr>
</tbody>
</table>

Table 3.2: Cohen’s Kappa for Level-1

Annotation of level 3 of the schema, types of Conversational Humor, resulted in a Kappa value of 0.49 (moderate agreement). Most of the disagreement at level 3, (Null, <some_type>) or (<some_type>, Null), is due to the failure or success of labelling the type by one annotator or a difference in perception of humor itself. A significant overlap of types can only be observed at (Retort, Teasing), which occurs at 11 segments. The dissimilarity of perception of the speaker’s intent causes this overlap. Annotator A1 perceived that the speaker intends to outwit or challenge the listener, whereas A2 perceived that the speaker only intends to pull the listener’s leg.

For the final level of the schema, techniques of Conversational Humor (Figure 3.3), it is worthy to note that each segment can be marked with one or more techniques, but it is not mandatory. If level 1, 2, and 3 (Conversational/Non-conversational, Monologue/Dialogue, Benign/Non-Benign) are marked, then level 4 (Type, Technique) is to be tagged compulsorily. For this type of data, it was considered best to use Krippendorff’s Alpha ($\alpha$) [56] for measuring the agreement between both annotators, resulting in an alpha value of 0.691. According to Krippendorff [56], tentative conclusions are acceptable where $\alpha \geq 0.667$.  

26
The low agreement value can be attested to the following observations. The role of the culture of
the then period, and knowledge of the language itself to recognize allusion and wordplay respectively
contribute majorly. Failure to understand the invalidity of the argument presented by the speaker leads
to a Null tag in place of Fallacious Reasoning by the annotator. The common feature of dramatic irony
with both exaggeration (Section 4.5.6) and Fallacious Reasoning is the presenting of a false statement
to the listener, causing a grey area for annotation.

Finally, the knowledge base possessed, culture exposed to, the emotion experienced at the time of an-
notation influences the individual’s subjectivity of humor [8] [46] [61], contributing to the disagreement
in the annotation of humor categories.

3.6 Conclusion and Future Work

This chapter of the thesis describes the work done on developing a fine-grained hierarchical annota-
tion schema for Conversational Humor. The annotation was performed on a relevant dataset, a prominent
Telugu play called Kanyasulkam. The inter-annotator disagreement highlighted the complexity of the
task as well as the domain itself. As mentioned in the introduction, the schema can be utilised for per-
sona identification, a use case especially beneficial for the literary field. For example, when analysing a
Shakespearean character, analysing their sense of humor may help the researcher recognize their perti-
nent traits.

This study also finds that inclusion of cultural nuances in the play has a significant effect on the
perception of humor. Further, this annotation schema can be applied to other culturally significant
works by utilizing the analysis provided in this work. However, when applying this schema to other
works, it is to be noted that the types and techniques listed here are non-exhaustive and more can be
added based on the language and cultural significance of the data being annotated. If a computer were to
generate humor, the traditional meaning would mean that it could generate "jokes" (knock-knock jokes, etc.) but would fail to generate conversational humor. It is believed that this work will aid in automating this process.
Chapter 4

Data Collection Methods for Conversational Humor Recognition

4.1 Introduction

For the task of humor recognition, several well-used datasets are available. As stated in section 2.4, Mihalcea and Strapparava collected 16,000 distinct one-liners for humor recognition. These were mined from jokes websites. Correspondingly, for non-humorous data, they mined proverbs and news headlines. Weller and Seppi collected jokes from Reddit and used the number of upvotes to demarcate whether a joke was funny or not-funny.

The Short Jokes dataset, found on Kaggle (a website for data resources), contains 230k+ jokes and have no restriction on the joke types. The jokes’ length varies from 10 to 200 characters. Chen and Soo, on an attempt to close the gap between humorous and non-humorous data, used news data from WMT162 that had the same distribution of words as the Short Jokes dataset. In contrast to English jokes, PTT jokes in Chinese are frequently used in humor recognition tasks. PTT is a bulletin board system where users converse about varied topics. A board called joke contains humorous data of different domains.

To the best of our knowledge, there exists no single source or dataset for Telugu jokes. One could argue that the Short Jokes dataset in English could be translated to Telugu and used. Culturally relevant features such as references made to popular English movies/songs/TV shows or wordplay may or may not be understood worldwide. These jokes containing popular references could be filtered out, leaving only universally jokes as a dataset for humor recognition. Free language translators such as Google or Bing Translate could be used to convert these remaining jokes to Telugu. Examples from the Short Jokes dataset:

https://www.kaggle.com/abhinavmoudgil95/short-jokes  
https://term.ptt.cc/
(a) I want to die peacefully in my sleep, like my grandfather... Not screaming and yelling like the passengers in his car.

Translation: nenu nixralo praSAMwaMgA canipovAlanukuMtunnAnu, nA wAwa lAgA ... wana kARulo prayANikulAL aruswU kAxu.

(b) Why do you never see elephants hiding in trees? Because they are good at it.

Translation: enugulu ceVtlalo xAkkovadaM mIru eVMxuku cUdaru? eVMxukaMte vAru xAnilo maMcivAru.

Example (a) is an example of a correctly translated humorous one-liner. The information, tone and the humor remain intact. On the other hand, in the first part of example (b), the question is translated with all the information intact. However, the punch line, or second sentence, is translated incorrectly. The translated phrase is ‘Because, at it, they are good people.’ The words ‘they are good’ are taken as a phrase for translation. Even if the translation contained all the information, it may not be natural. This could cause the joke to seem absurd or non-humorous as the delivery of the joke is improper. However, in case of a successful translation, the nuances in culturally specific jokes are lost as we have filtered these jokes out, thus failing to capture culturally relevant phenomena in humor. In summation, these factors emphasize the importance of gathering humorous data in Telugu.

### 4.2 Humorous Data

#### 4.2.1 Source and Structure of data

Although *Kanyasulkam* is an incredible collection of conversational humor [75], the difference in culture and language makes it less relevant to present times. Therefore, attempts were made to collect recent data. There are several sources for conversational humor (refer section 2.4) but the paucity of data in Telugu makes this difficult. To present verifiable results, sufficient data is required. As stated in the Introduction above, translation of English conversational humor is inefficacious.

To compile humorous data for the task of humor recognition, jokes from various sources were scraped including jokes section of news websites like Times of India (Samayam) and other miscellaneous sources such as Blogspot websites. As the area of focus is Conversational humorous data, the dataset was filtered to assemble jokes that followed a conversational format. Though extensive attempts to obtain conversational data from movies or TV shows were made, due to reasons such as unavailability of transcripts, manual transcription needed, and unavailability of multilingual OCRs, this direction...
proved to be unfeasible. Despite jokes being used, in the final dataset several conversational features are intact. All instances of the final humorous data used are of the same format as:

Speaker 1: utterance 1
Speaker 2: utterance 2
...  
Speaker n: utterance n.

Therefore, features such as turn taking and sequence organization are present. Turn-taking organization is where participants alternate their utterances, minimizing the noise arising from clashing of utterances to have a smooth or effective communication [90]. Sequence organization is the organization of these turns. If the conversational goal is to seek information, the first turn is the question, and the second turn is the answer [92]. For example,

Barwa: namnu kukkakannA hInaMgA cUswunnAv... emanukuMtunnAv.. nA guriMcI...
BArya: muMxu mIru moVragadaM Apeswe bAvuMtuMxi

Translation:

Husband: You regard me less than a dog ... what do think I am capable of?
Wife: Firstly, it’ll be nice if you could stop barking like one.

4.2.2 Filters to make humorous data homogeneous

Initially, a total of 6k+ jokes were collected. Utilizing string similarity, which is based on longest common subsequence metric, duplicate jokes were removed [67]. To minimize non-canonical forms of Telugu words written in Roman script, only those jokes written in Telugu script were considered. Thus, leaving 5,115 jokes. Several filters were applied to make a homogeneous conversational humorous dataset as the model could distinguish humorous and non-humorous data based on structural features, rather than semantic. Using the below mentioned techniques, a total of 2,047 conversational jokes were compiled.

• Jokes that follow a story structure, with no conversations but those that only describe a situation, or a sequence of events were removed.

Translation:
A very intelligent person decided to publish in a newspaper that he wished to get married. As everyone might be annoyed at the number of similar advertisements, he decided to write differently and settled on “Wife wanted”. The next day he received thousands of letters. All of them had the same content “Don’t be late at all, come and take my wife”.

- Jokes that are conversational but are not of $<$speaker$>$: $<$utterance$>$ are converted. For instance, the below joke is converted to

  ”are.. accaM pulilAne vuMxi mI kukka...!” annAdu sureR
  ”axi pulenAmDi bAbU.. I maXya premA xomA aMtU wirigi wirigi axi kukkalA wayArEMxi..!” samAXAnaM iccAdu maheR.

  sureR: are.. accaM pulilAne vuMxi mI kukka...!
  maheR: axi pulenAmDi bAbU.. I maXya premA xomA aMtU wirigi wirigi axi kukkalA wayArEMxi..!

Translation:

Suresh: Hey! Your dog is exactly like a tiger!
Mahesh: That is (emphasis) a tiger. It has been going around talking and thinking about love and has turned into a dog!

4.3 Non-Humorous Data

4.3.1 Source and Structure of data

Compiling the non-humorous dataset was not an easy task. The MULTIWOZ 2 [16] dataset was used initially as it contains human-to-human interactions and encompasses several domains such as hotels, restaurants, taxi services, etc. The original corpus is in English and could not be used directly. Hence, Google Translate was used. After examination of the translated corpus, the translated conversation had all the information intact but proved to be ineffective as Google Translate did not produce a natural translation. For example (abridged for brevity purposes),

**English version of Google Translate’s output:**

A: *Can I possess the postcode too?*
B: *Parkside’s police station’s postcode is CB11JG. Can I give you any help?*
A: *Do you have the address to Parkside police station? If not, do I please have the address?*
B: *Yes, Parkside address.*
A: *Great. Thank you for contacting Cambridge Towninfo Center.*
Figure 4.1: A tweet’s reply

B: *You are great. Farewell (noun form).*

A: *We are happy to help (the implied ‘you’ in English makes it unnatural in Telugu). Good day! (Literal translation makes it unnatural)*

This example shows us that the machine translation systems currently are not trained on conversations and therefore cannot recognize and translate the mannerisms used in the relevant cultures. Universally accepted norms such as ‘Goodbye!’ are transliterated and can be understood as the word is borrowed from English. But, for example, ‘Good day!’, or ‘Happy to help!’ when directly translated without finding the nearest natural counterpart in the culture, cannot be comprehended.

As pre-trained models such as FastText word embeddings and language model BERT are trained on human-written Wikipedia articles, it was able to discern human-written jokes from machine-translated conversations easily. This is explained further in Chapter 5 Section 5.5. To overcome this challenge, potential sources for non-humorous conversations were further explored, leading to finally deciding on scraping Tweets and their replies as conversations. A reply to a tweet engages with the original tweet’s user and their audience (see Fig. 4.1). Any subsequent replies are either directed at the original tweet’s user or at the other replies’ users. Hence, instead of retweets\(^6\) (see Fig. 4.2) replies were used as it is a conversation between several users.

---

\(^6\) A tweet’s retweet is a reshare of the tweet with or without a caption. So essentially, a retweet is only quoting the original tweet to the retweeter’s audience.
Every tweet is assigned several attributes such as tweet ID, user ID, timestamp, etc. Using a tweet’s ID, conversation ID and timestamp, tweets and their replies were compiled to form a conversation. A combination of Twint\(^7\), an advanced Twitter scraping tool, and Twitter’s API v2 endpoint\(^8\) were used to assemble the data.

To avoid the non-canonical forms of words obtained from scraping Telugu tweets written in roman script, ‘lang: te’ (aiming to fetch tweets written in Telugu script) was used as a filter when scraping tweets. In addition, in order to build a multidomain non-humorous dataset\(^67\), hashtags such as ‘#cinema’, ‘#politics’, ‘#cooking’, etc. were used as filters when collecting tweets.

Usernames were not used as speakers in the conversation for two reasons. Firstly, for the sake of anonymity, and secondly, usernames and names contain numbers and special characters. Consequently, to generate a natural conversation, the usernames were replaced by common Telugu names. Using the tweet’s author ID, speaker identity was preserved. For instance,

User123: Hello, how is today’s weather?
User456: It seems to be very sunny.
User123: Oh, wonderful!

In both instances, ‘User123’ is replaced by the same common Telugu name. This resulted in a corpus of 10,156 conversations. Additional filters were applied to improve the dataset. After manually checking the corpus, conversations that were humorous or contained profanity were removed to avoid ambiguity whether the conversation was humorous\(^32\)[121], finally resulting in 6,202 non-humorous conversations.

4.4 Attempts to homogenize humorous and non-humorous data

To make the non-humorous data and humorous data as similar as possible in structure, several changes were made to the collected jokes and non-humorous conversations (statistics in Table 4.1).

- Preprocessing steps such as removal of URLs, hashtags, and emojis were performed. In the case of a tweet, if emojis comprised the entirety of an utterance, the predecessor utterance was demarcated the end of the conversation, as removing that utterance is essentially changing the conversation. Emojis in jokes were used redundantly. For instance, the emoji of a man would be placed next to the Telugu word translating to ‘man’ and hence, were removed.

\(^7\)https://github.com/twintproject/twint
\(^8\)https://developer.twitter.com/en/docs/twitter-api/early-access
• The average number of utterances of all the conversational jokes was calculated to be approximately 3. Hence, the non-humorous conversations were trimmed to result in an average of 3 utterances to prevent the model learning patterns based on structural features.

• As previously mentioned, only those conversational jokes and tweets in Telugu script were included in the data. For the tweets, the filter ‘lang: te’ was used for this purpose. This filter is applied on an average to the characters of a tweet. Therefore, if most characters are in the Telugu script, then it is considered ‘lang: te’. Due to this, a tweet may not entirely be in Telugu script but may contain words written in roman script. These words can belong to either English or Telugu languages. For example, the words underlined are written in Roman script, some are English words while others are Telugu.

  kArina: pravarwana nAdu ala nedu ila
  wAhira: mohan bAbu vEeVsAr wammudiki cuttaM. kAbatti A saport evp-pudU ceswAdu.
  kArina: Han avunu...viRu wife is jagan anna’s cousin sister
  amoli: Jagan anna enti sir.. Jagan gadu anandi

Translation:

  Kaarina: [his] behavior today is different from yesterday’s.
  Thaahira: Mohan Babu is YSR younger brother’s relative. Hence, he will always support like that.
  Kaarina: Yes, Vishu’s wife is Jagan’s cousin sister.
  Amoli: Why are you calling him ‘Jagan anna’ (anna is used for elder brothers in Telugu culture, or is used as a sign of respect), call him ‘Jagan gadu’ (‘gadu’ is reserved for those not deemed respectable).

These words written in Roman script were detected using the Python package Langdetect\(^9\). Subsequently, a Python API, Google Transliteration\(^10\) was used.

• In Telugu culture, it is common that the name of the target/butt of the jokes is named Subbarao, or Apparao and so on. These are similar in theme to Sardarji jokes\(^11\). These names, found abundantly in the humorous dataset, replaced speaker’s names in the non-humorous dataset. This was to prevent the model learning that conversations involving speakers having these names implied humor, which is not the case.

\(^9\)https://github.com/shuyo/language-detection
\(^10\)https://pypi.org/project/google-transliteration-api/
\(^11\)https://en.wikipedia.org/wiki/Sardarji_joke
### 4.5 Conclusion

A dataset of conversations taken from different domains (cinema, politics, cooking) which are humorous or non-humorous is presented in this thesis. The challenges faced and the processing steps to compile a dataset for the task of humor recognition is described in detail. The dataset of jokes (conversational and non-conversational) and non-humorous conversations is uploaded online[12](https://bit.ly/3uKg1N7).

<table>
<thead>
<tr>
<th></th>
<th>Humorous Data</th>
<th>Non-Humorous Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collected Data</td>
<td>6,107</td>
<td>10,156</td>
</tr>
<tr>
<td>Post-filtration</td>
<td>2,047</td>
<td>6,202</td>
</tr>
</tbody>
</table>

Table 4.1: Statistics of Dataset
Chapter 5

Conversational Humor Recognition

5.1 Introduction

Most computational studies on humor use English data as it is more readily available and there exists no scarcity owing to its large number of speakers. But, in a country such as India where people are exposed to diverse cultures, it is important that humor is studied at an inter- or intra-cultural level. Several attempts have been made to tackle the task of humor recognition from this point of view for Hindi-English code-mixed data \[49\][91]. To the best of our knowledge, other than the work presented in this thesis, there have been no previous attempts at detecting conversational humor or humor in Telugu texts.

This subdomain of humor that involves the psychology or emotions of the interlocuters, and current events of the world, provides unique complexities during its analysis (computational or theoretical). But the phenomenon is so ubiquitous in everyday life that it demands further exploration and research that overcome these complexities. Moreover, with the rapid development of virtual assistants like Siri, Alexa or Cortana, the need for understanding and analyzing conversational humor for a better interactive experience is imminent.

In the domain of Conversational Humor, Ahuja, et. al \[2\] initialized the word embeddings with random or GloVe embeddings and used different deep learning architectures such as LSTM, biLSTM, CNN and FastText for classification of utterances where conversational humor lies. Ted Talks and the popular TV show Friends’ transcripts are used as their dataset. Chaudhary, et. Al \[17\] used features at several levels such as morpho-syntactic, lexico-semantic and pragmatic level and architectures including Logistic Regression, Gaussian Naïve Bayes, and SVM. They utilized an existing dataset on Kaggle that comprises 200k+ Reddit jokes in English.

Weller and Seppi \[113\] employ a Transformer architecture for humor detection and explain the advantages over traditional classifiers such as CNN and compare their results with human annotated data.
Existing datasets such as Short Jokes dataset on Kaggle and Pun of the Day were used. Apart from these they had mined Reddit jokes and used the number of upvotes to demarcate the humorous jokes from the non-humorous jokes.

Annamoradnejad and Zoghi [5] use BERT to generate embeddings for their combined dataset of several existing humorous datasets, including the dataset provided by Khandelwal, et. Al [49]. Different classifiers such as SVM and XLNet are compared with their proposed model, the latter giving an F1 score of 98.2 percent.

Yao, et. Al [119] proposed a Graphical Convolutional Networks architecture (GCN) for text classification. The framework builds a text graph based on word co-occurrence using the dataset’s text. Similar to the above researchers, five existing datasets from various domains (medical, news, and movie reviews) were classified into their respective classes, beating state-of-the-art architectures’ performance by a significant margin (except on one dataset). These aforementioned studies provided motivation to incorporate BERT, FastText and TextGCN architectures in the experiments for detecting conversational humor.

---

1This work has been submitted to a conference workshop and is currently under review.
5.1.1 Outline of Methodology

There are two approaches that are chosen as part of our proposed methodology (Fig. 5.1). The first is the use of pre-trained models, which is then further used for our downstream task of conversational humor detection. The second is by learning the word embeddings of the conversations using a heterogeneous text graph over which a convolutional neural network is optimized to classify the text. For the first approach we use FastText and BERT’s pre-trained models (non-English or dedicated to Indian languages such as Multilingual-BERT or MuRIL) as they have been trained on a large quantity of data. Subsequently, for the second approach, the Text Graphical Convolutional Network (GCN) framework proposed by Yao, et. Al [119] is implemented and fine-tuned.

5.2 Text GCN

5.2.1 Heterogenous graph

There have been several attempts at learning word representations by mapping words and the documents they belong to to a graph and learning the word-word and word-document relations using different features. Instead of learning words’ embeddings by only using unsupervised methods which use unlabeled and unstructured data, Tang, et. Al [103] also make use of pre-labeled text. They use a semi-supervised model to learn word representations by building a heterogeneous text network. In this network there are edges capturing the information between word-word, word-document and word-label. Le and Mikolov [57] embed arbitrary pieces of text such as documents and sentences using Paragraph Vectors. By using sentence or document embeddings, they learn the representations of words belonging in them.

In the paper by Yao, et. Al [119], using unsupervised learning, they build a heterogenous graph of words and documents. First, this graph is built using word co-occurrence and document-word relations, after which a Text Graphical Convolutional Network is learnt on the corpus. A vital difference must be made between the graphical frameworks earlier stated and [119]. The former learns word embeddings and then builds text representations, while the latter learns word and document embeddings simultaneously for the purpose of text classification.

The heterogeneous graph comprises of nodes that are interconnected. Let G (V, E) be a graph where V are the nodes and E are the edges between them. The words and documents of the text make up the nodes or V. Therefore, the total number of documents and the total number of distinct or unique words in the corpus sum up to be the total number of nodes in the heterogeneous text graph (Fig. 5.2).

Using term frequency-inverse document frequency (TF-IDF), edges are constructed between words and documents. It was found that using term frequency alone did not help with the overall task’s per-
performance. Similarly, based on word co-occurrence, edges are constructed between words to make word-word relations. Instead of focusing on word co-occurrence in a single document (in our case, a conversation), a fixed sliding window on all the documents of the corpus is used to calculate global co-occurrence information.

Using PMI (point-wise mutual information), word associations are calculated. These word associations help us capture the relation between word nodes. PMI is a popular measure used to calculate the association between two entities, in this case words \[20]. Below is the mathematic formula for the same.

\[
PMI(x, y) = \log \frac{P(x, y)}{P(x)P(y)} \tag{5.1}
\]

\[
P(x, y) = \frac{W(x, y)}{W}; P(x) = \frac{W(x)}{W} \tag{5.2}
\]

\(P(x, y)\) is the probability of the word pair \((x, y)\), and similarly, \(p(x)\) is the probability of the word \(x\). \(W(x, y)\) is the number of sliding windows that contain the word pair \((x, y)\), \(W(x)\) is the number of sliding windows that contain the word \(x\), and finally, \(W\) is the total number of sliding windows of the corpus. A positive PMI implies a high association between words. Whereas a negative PMI implies little or no relation between words \(x\) and \(y\). Yao, et. Al \[119]\ considered only positive values, and only these edges were considered for the heterogenous text graph.

5.2.2 GCN for Text Classification

The Graphical Convolutional Network (GCN) proposed by Kipf and Welling \[51]\ provided a method by which text classification is performed on graphically structured data. A two layered GCN is applied
on the built heterogenous graph. A single layer GCN captures information about a node’s immediate neighbors, further stacking of such layers integrates a larger neighborhood in its understanding.

Convolutional Neural Network (CNN) works with Euclidean structured pixels of an image, where there is a regular pattern to the graph. The same filter is used for different neighborhoods for convolution \[117\]. However, when dealing with text, the number of neighbors a node is connected to varies. The PMI and TF-IDF weights are updated in the adjacency matrix which is further modified by degree matrix $D$.

Let the feature matrix be $X$, which is initialized as an identity matrix. Hence, all nodes (unique words and documents) are initialized with one-hot vectors. As in any feed forward neural network, the feature representation at current layer $i + 1$ is dependent on layer $i$. This way, on further stacking of layers, a forward pass or message passing between nodes of the graph occurs. Yao, et. al \[119\] experimented with 2 layers of GCN and observed that further stacking of layers did not improve performance. The layers’ weight matrices $W_0$ and $W_1$ are learnt by using gradient descent. The output of the second layer is fed to a softmax layer for determining the document class (in our case, a humorous or a non-humorous conversation). The resultant equation is provided below:

$$Z = softmax(A_1ReLU(A_1XW_0)W_1)$$ (5.4)

Briefly, a GCN is a multilayered neural network which produces word embedding vectors of the nodes based on properties of their neighbors. A single layer of GCN considers only the immediate neighbor’s information, whereas a two-layer GCN passes messages between nodes that are two hops away. After experimenting, the first layer’s embedding size was limited to 200, whereas the window size was set to 20, along with learning rate as 0.02 and dropout rate as 0.5. The training and test data were split into 80:20 ratio and the number of training epochs was stopped early if after 10 successive epochs the validation loss did not decrease.

### 5.3 FastText

#### 5.3.1 Word Embeddings

Before we delve into the workings of FastText \[15\], it is important to understand its parallel contemporary efforts, Word2Vec. Word2Vec learns words’ embeddings from a large corpus of unstructured data. It uses two methods, CBOV (Continuous Bag of Words) or Skip-gram model, both of which use a window of context to learn the representation of a word. These embeddings are then projected into $n$-dimensional vector space. As the context of a word is considered, a word and its synonyms will have a similar embedding due to their appearance in shared contexts.
An important feature of Word2Vec is that it takes the entire word or token into consideration when learning its representation. Consequently, a noun’s singular and plural form are both represented in space as distinct words (with a short distance between them). Therefore, for the machine to learn that the noun ‘leaf’ and ‘leaves’ are closely related there must be sufficient data where ‘leaf’ and ‘leaves’ are used in similar contexts.

On the other hand, FastText uses sub-word information to capture a word’s representation. A word is split into character n-grams, and these collectively make up the embedding for a word. The number of n-grams can be chosen as per our requirement for the task. For example, if we consider trigrams and the word ‘machine’, we would be left with ‘<ma’, ‘mac’, ‘ach’, ‘chi’, ‘hin’, ‘ine’, ‘ne>’. The extra characters ‘<’, ‘>’, added to the first and the last trigram are useful to demarcate a word’s boundaries and hence help the machine in understanding prefixes and suffixes. These trigrams are trained using the Skip-gram model. A word’s embeddings are then a summation of the character n-grams that make up that word. Hence, the machine does not require extra data to learn the embedding for the word ‘leaves’.

Moreover, FastText is highly advantageous for an agglutinative language such as Telugu [102]. In an agglutinative language, a lexeme is attached with suffixes which carry information such as gender, number (singular/plural) or tense. For instance, the word below is made up of smaller morphemes.

**intikochindhannamaata**

**Morphemes that make up the word:**

*illu + ki + ochuta + i + anamata + emphasis*

**Translation:**

*house + to + come + past tense marker + apparently + emphasis*

It has been observed that FastText performs significantly better at capturing word embeddings for such languages than Word2Vec [15]. Facebook has released its pre-trained FastText word embeddings for 157 languages including Indian languages like Telugu. These word embeddings are learnt on Common Crawl and Wikipedia data, both of which are readily available.

5.3.2 FastText Classifier

There are a couple of unique features of the FastText Classifier that allow it to perform on par in terms of accuracy and quicker as compared to other contemporary text classifiers. The classifier considers a text given as a bag of n-grams. These n-grams are hashed so that their lookup is easier and faster. After fetching the n-grams’ embeddings, they are averaged to form the hidden variable of the linear
Figure 5.3: Word n-gram features are embedded and averaged to form the hidden variable classifier (see Fig. 5.3). This is to potentially share information between features and classes so that the information gleaned about one class can be used for the detection of other classes.

The bag-of-words technique does not take the explicit order of the entire text into account as it would be computationally expensive to do so. Therefore, bag of n-grams is used as it takes the local context of a word into account which preserves the word order partially. Instead of learning the word vectors using the text’s labels, pre-trained vectors are used for better performance. FastText fetches the pre-trained word vectors of the bag of n-grams and averages it to result in the text’s representation.

The output of the linear classifier is fed to a hierarchical softmax \[40\] based on the Huffman coding tree \[68\] to reduce the time complexity of the overall algorithm. The tree is organized based on categories. Due to this structuring, the computational complexity for training reduces from linear to logarithmic time with respect to the number of classes. The probability of a node is the product of the probability of its parent nodes, and hence, is lesser. This allows us to discard many subtrees based on their low probability while performing a depth first search to hit the node with the maximum probability while testing. The model was trained for 15 epochs, with the dimension of word embeddings set to 300, the number of word n-grams set to 2 (bigrams), learning rate set to 1.0, and loss set to hierarchical softmax.

5.4 BERT

5.4.1 Language Modeling

First and foremost, an important distinction must be made between FastText and BERT. FastText is a Facebook library that offers a word embedding algorithm. However, BERT is a language model. Given a corpus of text, a word embedding algorithm learns a distributed representation of a word based on a window of context. If two or more words have similar word embeddings, it is deduced that they are either related semantically or syntactically (synonyms or singular/plural noun forms).
On the contrary, a language model learns the likelihood of a word’s appearance in a sequence based on its context (unidirectionally or bidirectionally). After training, when a sequence $w_1, \ldots, w_i, w_{i+1}, \ldots, w_n$ is given to the model, it returns a word at $w_i$ that has the maximum probability of occurring at that position.

The paper *Attention is All You Need* proved to be a cornerstone contribution to the field of NLP as well. Unlike BERT, the Transformer proposed by this paper is not a language model but rather translates a sequence to another. Briefly, a Transformer contains a stack of encoders and decoders, that have many components including multi-head attention layers, feed forward networks, and layer-normalization networks.

RNNs or LSTMs are used for Sequence Modeling as well but exhibit a disadvantage of not retaining relevant information when given a long sequence. This disadvantage is avoided by a Transformer’s architecture due to a mechanism called attention. This mechanism aims at capturing the dependency between a word and the rest of the sequence. For example, let us consider the below two sequences.

*The animal did not cross the street because it was tired.*

*The animal did not cross the street because it was inundated.*

The word ‘it’ in the first sentence refers to the animal, whereas in the second sentence ‘it’ refers to the street. An encoder’s attention layer bakes the dependency between the word ‘it’ and the other words in the sequence. This attention layer is multi-headed, and hence calculates this dependency several times using weight matrices. The output of a single encoder’s multi-headed attention layer is sent to a feed-forward network, whose output is then sent to the successive encoder, which repeats this cycle.

The final encoder’s output has the measure of how dependent a word is to another and therefore, has a better understanding of which word ‘it’ refers to. For the first sentence, the dependency calculated between ‘it’ and ‘animal’ will be greater as compared to ‘it’ and ‘street’ and vice-versa for the second sentence. This final output is sent to the first decoder that generates the target sequence one word/token at a time.

But as NLP has numerous tasks such as classification, question-answering, and multiple-choice selection, this Transformer that was originally intended for machine translation was refactored for downstream tasks by Open AI. This refactored Transformer is used for language modeling by using just the decoder. The output of this decoder is then sent to a feed forward neural network and a softmax layer which can be fine-tuned to perform the downstream task.

BERT is a combination of features that are described above along with characteristics that are unique to itself. BERT can be summed up as a pre-trained model on massive unstructured data using bidirec-
tional encoder representations from Transformers. Thus, it shares features with Vanilla Transformer and Open AI Transformer.

BERT considers the complete context surrounding a word and is trained on unlabeled data. BERT contains a stack of encoders such as the Vanilla Transformer that has multi-headed attention layers. It is layered by a feedforward neural network and softmax layer for the purpose of downstream tasks. However, Vanilla Transformers are modeled to look at a word’s context unidirectionally. Therefore, to mitigate this disadvantage, BERT uses masking of words/tokens at random so the model can learn the probability of the word appearing in its entire surrounding context over iterations. BERT is also trained on a two-sentence task. Given a sentence A and B, the model learns the likelihood of B following A. This way the model learns the relationship between sentences as well.

This pre-trained model can be plugged in for many downstream tasks by taking the pooled output of BERT and passing it to a neural network suitable for the task at hand. An important feature of BERT to be noted is that it does not consider whole words like Word2Vec but rather ‘word pieces’, and hence is useful for agglutinative languages such as Telugu and can handle unknown or erroneous spelled words.

Numerous pre-trained models are available on HuggingFace. Instead of restricting to using BERT pre-trained models, BERT’s cousins AlBERT and DistilBERT are also experimented with. AlBERT and DistilBERT attempt to reduce the computational complexity. Specified below are the pre-trained models used for the purposes of this thesis:

- BERT: Multilingual base model (cased, trained on top 104 languages with the largest Wikipedia), MuRIL (Multilingual Representations for Indian Languages, trained on 17 Indian languages),
- AlBERT: Indic-BERT (ALBERT model pre-trained only on 12 major Indian languages),
- DistilBERT: A distilled version of Multilingual base model (cased)

### 5.4.2 Linear Classifier

The pooled output of the BERT (or AlBERT or DistilBERT) encoders is directly sent to a classifier and thereafter to a softmax layer. A Pytorch classifier that applies a linear transformation to the given data using the following equation:

$$y = xA^T + b \quad (5.5)$$

The classifier outputs logits which are normalized by the softmax layer which essentially produces the predicted value of whether the conversation is humorous or non-humorous. As the classifier learns the appropriate weights, the loss function guides the model whether the predicted values are far from the true values. Cross Entropy loss function is chosen for this purpose. The lesser the Cross Entropy loss, the nearer the predicted values and true values the model is learning during training.
5.5 Results and Discussion

After collecting and pre-processing the tweets and their replies to form non-humorous conversations as stated in Chapter 4, experiments were run with Text GCN, FastText and various BERT models (refer to Table 5.2). The respective model is trained on 80% of the data and is tested on 20% of unseen data. As there are 2,047 instances of humorous conversations whereas 6,202 total instances of non-humorous conversations, this makes it an unbalanced dataset. The weights of the classes are taken into consideration. Accuracy and F1 score are used as evaluation metrics.

Highest accuracy results are obtained by Multilingual BERT by Google. Multilingual BERT is trained on 104 languages using Wikipedia dumps of the respective language. This model is pre-trained with both objectives: Masked Language Modeling (MLM) and Next Sentence Prediction. Additionally, it is observed that FastText and BERT models perform comparatively better than Text GCN. The low F1 score implies a low precision and low recall. This means that the model does not predict that the text is humorous well and sufficiently. In Section 5.2.1 It is specified that the edges between word-word nodes are given weights based on PMI (point-wise mutual information). The word pair with the highest PMI (12.34) is

\(<\text{aMxulO}, \text{kaxA}>\)

Translation:

\(<\text{in that, right (in the sentence 'you finished your homework, right?')}\>

It is evident that both words have no syntactic or semantic relation. Thus, the heterogenous graph does not capture word relations well. In comparison, FastText’s nearest neighbors defined in the pre-trained model for the word ‘aMxulO’ (translates to ‘in that’) are shown in Table 5.1.

<table>
<thead>
<tr>
<th>Nearest Neighbors</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>iMxulO</td>
<td>in this</td>
</tr>
<tr>
<td>vAtiLO</td>
<td>in those</td>
</tr>
<tr>
<td>xIniLO</td>
<td>in this</td>
</tr>
<tr>
<td>xAniLO</td>
<td>in that</td>
</tr>
<tr>
<td>axi</td>
<td>that</td>
</tr>
<tr>
<td>vItiLO</td>
<td>in these</td>
</tr>
</tbody>
</table>

Table 5.1: FastText nearest neighbors of word ‘aMxulO’

By inspecting the nearest neighbors of the queried word, FastText captures syntactic (plural of ‘in that’ is ‘in those’) and semantic relations (antonym of ‘in that’ is ‘in this’). This highlights the importance of word embeddings for the model’s overall performance for the task to be carried out. Word representations is a key aspect that contributes to the effectiveness and performance of text classification [98][110].
As mentioned in Chapter 4, section 4.3.1, conversational text written in Roman script was transliterated to Telugu script using Google Transliterate API. The API produces an array of most probable transliterations of the input given, after which the first element is considered by default. However, at times, the API does not produce accurate results with a slight margin of error. Thus, potentially producing word(s) that do not exist in the language’s vocabulary. Take the below utterance from a non-humorous conversation for instance. It is a Telugu utterance written in Roman script.

Sir mundu tanatho pracharaniki ranichina aa janasena candidate meda katina charyalu tiskovali

Google Transliterate API’s result:

sar muMxu wanawo pracArAniki rANiMcina A janasena kAMdidet mIxa kaTina caryalu wIsukovAli

Translation:

Sir, even if we first let her come to the campaign, we must take strict actions against that Janasena (a political party) candidate.

Here, the only error is found in the word ‘rANiMcina’ should be ‘rAniMcina’, an error difference of one character. But when Text GCN or word embedding algorithms such as Word2Vec encounters this incorrectly transliterated word, and if it does not occur frequently in similar contexts in the corpus, it will treat it as an out-of-vocabulary word. Hence, not capturing this unknown word’s relation with the correct word ‘rAniMcina’.

FastText and BERT fill this inadequacy by taking sub-word or word-piece information. As the difference between the incorrectly transliterated word and the actual word from the language is one-character, both FastText and BERT will produce word embeddings which will be closer as compared to what TextGCN would produce. For a language such as Telugu, where suffixes are attached to carry semantic and syntactic information, breaking a word into sub-words or word-pieces becomes crucial.

Another advantage that transfer learning applications such as FastText and BERT models provide is that they are pre-trained on vast amounts of Wikipedia data and can therefore generalise a word’s meaning in a context better.

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Accuracy</th>
<th>F1 Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text GCN</td>
<td>0.592</td>
<td>0.374</td>
</tr>
<tr>
<td>FastText</td>
<td>0.973</td>
<td>0.946</td>
</tr>
<tr>
<td>Multilingual BERT</td>
<td>0.993</td>
<td>0.985</td>
</tr>
<tr>
<td>MuRIL</td>
<td>0.988</td>
<td>0.977</td>
</tr>
<tr>
<td>Indic-BERT</td>
<td>0.992</td>
<td>0.982</td>
</tr>
<tr>
<td>Multilingual DistilBERT</td>
<td>0.990</td>
<td>0.980</td>
</tr>
</tbody>
</table>

Table 5.2: Performance of architectures implemented for Conversational Humor recognition
5.6 Conclusion

In this chapter, the problem of conversational humor detection in Telugu is addressed. Different word embedding algorithms or language models, coupled with different classifiers are used to resolve the work at hand. The performances of the various models are evaluated and analyzed to glean insights regarding the mechanisms employed. For low-resource Indian languages such as Telugu, the hurdles that lack of data pose are avoided by effective use of pre-trained models on a substantial amount of Telugu data.

The BERT model trained on 104 languages, Multilingual BERT base (cased) by Google, delivered the best performance with an accuracy of 99.3% and an f1 score of 98.5%. Comparatively, FastText comes close with merely a 2% difference in accuracy, 97.3% and an f1 score of 94.6%. State-of-the-art results are thus produced by utilizing transfer learning techniques and methodologies.
Chapter 6

Conclusions and Future Work

6.1 Conclusion

To the best of our knowledge, the work described in this thesis is the first attempt in computational humor research in Telugu and moreover in the sub-domain of Conversational Humor. Conversational Humor occurs when interlocutors of a conversation use several types and techniques for various social and psychological purposes. Conversations are analyzed theoretically as well as computationally for humor using appropriate and adequate data.

An annotation schema for Conversational Humor is presented by utilizing a relevant humorous stage play, *Kanyasulkam*, written 130 years ago, as a medium of analysis. The play is fully annotated by two annotators who are fluent in Telugu. It is a hierarchical framework that distinguishes between humor that is conversational or non-conversational, and if it is conversational, a difference is made between humor that arises in a monologue or a dialogue.

Apart from this, the benignity or non-benignity of a speaker’s humor is also considered. To this end, the Benign Violation Theory by Peter McGraw is taken into account where the benignity of an utterance is dependent on factors including norms of the relevant culture and psychological distance between the topic of humor and the interlocutors. Numerous types and techniques are identified as well and are described with illustrative examples. An in-depth disagreement analysis is also reported to comprehend the complexities of the domain better.

For the computational task of detection of conversational humor, the challenges faced, the motivations behind scraping conversational jokes and tweets have been elucidated in this thesis. Google Translate service’s room for improvement in translation of conversations has been explained using examples. 2,047 conversational jokes form the humorous dataset whereas 6,202 tweets and their replies form the non-humorous dataset. Pre-processing steps and filters are applied to both datasets. Attempts are also made to homogenize both datasets in structure to essentially make them differ exclusively semantically.
Finally, using the assimilated humorous and non-humorous datasets several experiments are run on diverse models for conversational humor detection. Text GCN, FastText and BERT models are employed and compared for this purpose. For the Text GCN, first a heterogenous graph is built where word-word relations and document-word relations are learnt based on certain features after which GCN is applied. As transfer learning is popularly used in the field of NLP currently, FastText pre-trained word embeddings and BERT pre-trained models are experimented with. Compared to Text GCN that built word representations using only the data provided, the accuracy is comparatively much lower (59.2%) whereas FastText and Multilingual BERT provide a state-of-the-art accuracy of 97.3% and 99.3% respectively.

In summary, as mentioned above, theoretically and computationally conversational humor is delved into to further the efforts in humor research. A novel annotation schema is formulated that considers the structure and semantics of conversational humor using a critically well acclaimed piece of literature. In order to use relevant data of the present times, conversational jokes and tweets written recently are scraped. This way conversations that are pre-meditated (stage play and jokes) and spontaneous (tweets and replies) are assimilated to produce a holistic dataset. Taking advantage of transfer learning high accuracies are obtained for the classification of conversational humor.

6.2 Future Work

Following the schema proposed the types, techniques, benignity or non-benignity of humor used by a speaker can be identified in a conversation. Using these, information regarding their personality can be gleaned and hence, aid in persona identification. Telugu movie scripts could be analyzed to comprehend the trends in the types of humor used in the Telugu culture, the influencing factors, and the importance of shared knowledge of culture in the perception of humor. Instead of using premeditated conversations, real-time conversations transcribed from humorous Telugu interviews would capture the essence of conversational humor better.

Machine Translation systems like Google Translate could be additionally trained on conversations and the relevant culture’s mannerisms for better translation. Detection of humor in conversations could be taken one step further to detect a particular technique(s) or type(s) of humor. This pre-trained model can be used to develop a tagged dataset which could lead to the generation of conversational humor. Although generation of conversational humor is extremely complex, information from the ongoing conversation could be used to generate humor. For example, callback humor, where a previously mentioned topic is uttered again to produce humor. The latter can be integrated in Dialogue Systems to make the conversations generated by computers more human-like.
Related Publications


Bibliography


