Comprehensive Project : SEM-8 Progress Report 1

A Project Report Submitted in Partial Fulfillment of the Requirements for Award
of
the Degree of Bachelor of Technology in Information and Communication
Technology

Submitted by
Ishan Mistry - 18BIT033
Sagar Sinha - 18BCP096
Sanskar Bhuwania - 18BCP099
Anurag Pal - 18BCP007

Under the Supervision and Guidance of
Dr. Paawan Sharma
Department of Information and Communication Technologies
Pandit Deendayal Energy University

Under the Supervision and Guidance of
Dr. Santosh Kumar Bharti
Department of Computer Science and Engineering
Pandit Deendayal Energy University

Submitted to
Department of Information and Communication Technology
School of Technology
Pandit Deendayal Energy University (PDEU)
Gandhinagar, INDIA, 382007

Declaration

I hereby declare that the project work entitled **"Project Title"** is an authentic record of my own work carried out in xyz industry/xyz university/xyz research institute as requirement of B. Tech dissertation for the award of **Bachelor of Technology in Information and Communication Technology**. I have duly acknowledged all the sources from which the ideas and extracts have been taken. The project is free from any plagiarism and has not been submitted elsewhere for any degree, diploma and certificate.

Signature:...............................

Your Name

(Roll No.)

Department of Information and Communication Technology
PDEU

Dr. Ganga Prasad Pandey
H. O. D.   e-mail:

Certificate

This is to certify that the project entitled **"Title of the Project"** submitted by **Your Name**, Roll No. xyz123 to the Department of Information and Communication Technology under School of Technology, PDPU in partial fulfillment of the requirements for award of the degree of **Bachelor of Technology in Information and Communication Technology** embodies work carried out under the guidance and supervision of Dr./ MR. XXX YYYY, designation of Industry mentor/Professor, xyz industry/Department of XYz.

....
Dr. Ganga Prasad Pandey
(HOD, I.C.T. Department, PDEU.)

DR./Mr. xxx yyy
Designation of Industry
mentor/ if Inhouse write about
Inhouse mentor

Phone: xxx - xxxx
Fax: xxxx - xxx
e-mail: ABC@xyz.com

Certificate

This is to certify that the project entitled **"Identification of the critical point in a conversation and use it to identify and prevent cancel culture using NLP techniques"** submitted to the Department of Information and Communincation Technology under School of Technology, Pandit Deendayal Petroleum University in partial fulfillment of the requirements for award of the degree of **Bachelor of Technology in Information and Communication Technology** is a record of work carried out by **Ishan Mistry**, Roll No. 18BIT033 under my supervision.

Signature ....
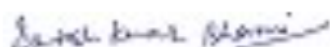Dr Paawan Sharma
Phone No:
Email ID:

DR./Mr. xxx yyy
Designation of Industry
mentor/ if Inhouse write about
Inhouse mentor

Phone: xxx - xxxx
Fax: xxxx - xxx
e-mail: ABC@xyz.com

Certificate

This is to certify that the project entitled **"Identification of the critical point in a conversation and use it to identify and prevent cancel culture using NLP techniques"** submitted to the Department of Information and Communincation Technology under School of Technology, Pandit Deendayal Petroleum University in partial fulfillment of the requirements for award of the degree of **Bachelor of Technology in Information and Communication Technology** is a record of work carried out by **Ishan Mistry**, Roll No. 18BIT033 under my supervision.

Signature ....
Dr Santosh Kumar Bharti
Phone No:
Email ID:

Certificate of Approval

The forgoing project entitled **"Project Title"** submitted by **Your Name**, Roll No XYX123 to the Information and Communication Technology under School of Technology, PDPU is hereby approved as project work carried out and presented in a manner satisfactory to warrant its acceptance as a prerequisite of **Bachelor of Technology in Information and Communication Technology** degree for which it has been submitted. It has been understood that by this approval of the undersigned do not necessarily endorse or approve every statement made, opinion expressed or conclusion drawn therein but approve only for the purpose for which it is being submitted.

....
Signature of Panel Members

Acknowledgement

Write your acknowledgement here. For example: The success and final outcome of this project required a lot of guidance and assistance from many people. It is my privilege to pledge the following few lines of dedication to those who helped me directly or indirectly in completing my project.

Words are ..

I owe my ..

Furthermore, I am thankful to and fortunate enough to the faculties of ICT. Department for all the knowledge they have imparted, which I could put to use in this project...

Most importantly, last but not the least, I appreciate the co-operation and benevolence of my friends for helping me with the data recording and family for cheering me up all through out this journey.

Contents

Chapter 1

Introduction

1.1 Problem statement

"Identification of the critical point in a conversation and use it to identify and prevent cancel culture using Natural Language Processing techniques"

1.2 Motivation

It is not uncommon to observe how differences in the eccentricities of people to converse and present ideas conduces to cause confusion or disagreement even when the topic is something that both the individuals are quite familiar with. This phenomenon is almost never a sudden jump from a categorization like "**Strongly agree**" to "**Strongly Disagree**" but rather there are many such sporadic instances/points sprinkled throughout the conversation that lead the other individual into believing that they are against them.The aim of the undergoing study is to identify these clues and help mitigate such situations proactively.

If left unattended, such conversation could become lopsided easily and in the process, as people vehemently express their opinions – people with differing views might feel or even be subject to some verbal attacks or scrutiny.Such conversations are practically dead ends where they **could have been fruitful if moderated properly**.

Cancel culture or call-out culture is a modern form of ostracism in which someone is thrust out of social or professional circles – whether it be online, on social media, or in person. This is getting quite exorbitant these days because of the anonymity bestowed upon us by the current online platforms that affords them this extraneous power over others. We do realize that such events can have its own consequences and do surely take a toll on the victim mentally and emotionally.

One of the key features of this project is to **identify and quantify the latent patterns** in the conversational data. This study could be a general eye-opener which will help us quantify the internal biases (confirmation bias, selection bias, etc.) that we subconsciously fall prey to. These insights will help us better understand how we communicate and whether we are victims to such biases.

This project takes off as a general idea/study, but we would hone in on a particular problem for validating our ideas. We would try to give special emphasis on the

generalizability of the project to the extent such that all we would need are little tweaks could be transferred to other domains as well.One such example could be, a model that is trained to identify the vagueness or relevance of the sentences can later on be used a mould for a model that needs to predict whether there is scope for contention or conflicts later on due to these misunderstandings?

The study revolves around modelling a critical point that would be pivotal to understand these complex social interactions online.Complex in nature as they can be viewed very differently depending on the context of the problem at hand:

- Points of collaboration in a conversation
- Resolving conflicts
- Facilitating discussions that are vague
- Steering the conversations - amplify the seemingly weaker ideas
- Preventing cancel culture
- Preventing Groupthink
- Identifying suicidal tendencies

Hence, having a common ground for analysing such situations could be a great starting point for researchers in various fields such as Natural Language Understanding Computational Linguistics,Social Sciences,research psychologists A system that can be used to transfer its learned insights from one domain to another would be interesting to observe - maybe some new insights that we did not know earlier could be brought to light.

1.3   Objectives

Modeling the entire conversations in a computationally efficient manner, preserving the various relationships(both intrinsic and extrinsic) between the sentences of the conversation. the so that we can later on perform analysis on the various kinds of conversations and henceforth use these observed patterns to train a generative model.

Once we have a set framework of how to store these sentences we can have our own custom metrics and study all the various categories of conversations - Conflicts/collaboration/mental health related, etc. and try to encode them into a single metric that we would aim to optimize.This final metric will take into consideration all the features(Relevance to the topic, history, sentiment etc.) that will be extracted from the EDA of the conversations.

This critical point in a conversation can help us pinpoint the causes(sentences) of contention and groupthink, and thereafter predict a sentence that could moderate the discussion which otherwise would have turned into a heated argument or herd thinking."

1.4   Literature review

We tried to follow the major contributions in this area so far and tried to come up an idea for the baseline model.

1.4.1 Discovering Useful Sentence Representations from Large Pretrained Language Model

This paper was particularly interesting as it was a successful attempt at linking unstructured data like Knowledge Graphs (comprising of the conversations insights in our case) to some structured data (like sentence spaces for each conversation).

Wishlist from author's pov:

1. Preserve semantic relations in sentence space

2. Graph encoding to encode semantics of the graph

3. Comparison of several sentence embeddings

4. Novel alignment task for efficacy of sentence embeddings and how they can be used for **relation extraction**.

Sentence Embedding Methods described in the paper:

1. Bag of Words - an efficient sentence classification model by representing a sentence as the average of its component word representations.

2. GloVe - Mean (using vector mean as the representation of a sentence)

3. GloVe - DCT (emphasis here is on the sequence/arrangement of the words in a sentence)

4. Skip Thought - First one to try RNN for sentence encoding - encoder/decoder setup to predict k sentences to left and to the right

5. Skip Thought - Disagrees to the objective function used in Skip Thought and says that "semantic info is not so related to the form of words" - Used Window but now to identify the correct center sentence

6. LASER - Outperforms BERT-like architectures, for our case,in the task of sentence similarities.

7. Parameter free - GEM - They make use of the geometry of the words in sentences. Firstly, they create an orthogonal basis for words to capture meaning and syntax, secondly they perform QR factorization and then re-weighting by 3 metrics : words' novelty, significance and uniqueness

Knowledge Graph Embedding methods:

Knowledge Graph could be used for **distant supervision** (for weak annotation) but assumption is that if two entities are present in the KG triple then there has to be a relation between them. This is not always the case as the sentences can be quite obscure and complex.

Motivation to study these is to for **statistical representation learning of higher dimensional graphs into low dimensions**.

3 Main categories of KGs:

1. Translation-based models(The hero of this paper) The intuition behind translation-based models is we wish to have low-dimensional, dense representations of h, r, t such that $h + r \approx t$. Emphasis here is on the space entities - words and the scoring function used to identify true relations from noise.

2. Sematic matching models,

3. Graph structure model

Models

1. TransE - Simplest model - entities and relations in the same low dim space and uses a distance based scoring function. It was unable to encode entities with one to many relations.

2. TransH - resolves the problem in TransE . Here, each relation has its own hyperplane(assuming that hyperplane has some fixed embedding style).

3. TransR - goes one step further and each relation has its own distinct embedding space. Many parameters but good performance.

Objectives :

Dataset $\rightarrow$ function $\rightarrow$ Embedding (imagine f as compression)
    The aim is to exploit the invariances(only when there are differences, we can cluster) between the embeddings. The degree of invariance is a measure of amount of data required.

Method for checking Cluster-ability

They have worked with spatial histograms and then doing PCA with projecting it onto 2 new dimensions. Then these transformed points are split into n equal bins, counted them and joint probability mass function is calculated. This process is applied for 500 uniformly generated points. KL Divergence for each spatial histogram is used to compare the difference.

    Great results on GEM+GLOVE , firstly because orthogonal basis used - the vectors will tend to be very different from each other. Secondly, it is observed that all sentence and vector representation.

Semantic mapping capacity:

Basically testing if the knowledge graph and the sentence embedding follow similar structure in low dimensional spaces.

    Trying to learn W from the batches of sentence s_i and triples (s_i, t_i) , such that $\sum_{i=1}^{n} ||Wf_1(s_i) - f_2(t_i)||^2$ is minimized.

1.4.2   A Review of Relational Machine Learning for Knowledge Graphs

ML: Traditional machine learning algorithms take as input a feature vector, which represents an object in terms of numeric or categorical attributes. The main learning

task is to learn a mapping from this feature vector to an output prediction of some form. This could be class labels, a regression score, or an unsupervised cluster id or latent vector (embedding).

SRL: In statistical relational learning (SRL), the representation of an object can contain its relationships to other objects. Thus the data is in the form of a graph, consisting of nodes (entities) and labelled edges (relationships between entities). The main goals of SRL include prediction of missing edges, prediction of properties of nodes, and clustering nodes based on their connectivity patterns.

AIM: Review a variety of techniques from the SRL community and explain how they can be applied to large-scale knowledge graphs (KGs), i.e., graph structured knowledge bases (KBs) that store factual information in the form of relationships between entities.

KG:

KG Representation:

- Follow Resource Description Framework (RDF) standard and represent facts in the form of binary relationships, in particular (subject, predicate, object) (SPO) triples, where subject and object are entities and predicate is the relation between them. The existence of a particular. SPO triple indicates an existing fact, i.e., that the respective entities are in a relationship of the given type.

KG (KB) construction:

1. classify KB construction methods into four main groups:

   (a) in curated approaches, triples are created manually by a closed group of experts;

   (b) in collaborative approaches, triples are created manually by an open group of volunteers;

   (c) in automated semi structured approaches, triples are extracted automatically from semi structured text (e.g., infoboxes in Wikipedia) via hand-crafted rules, learned rules, or regular expressions;

   (d) in automated unstructured approaches, triples are extracted automatically from unstructured text via machine learning and natural language processing techniques

   The best is automatic KG esp. "Read the Web" approach. Eg- NELL and Knowledge Vault. "Noise" in such automatically extracted facts by using the knowledge from existing, high quality repositories.

2. KBs, can also be classified based on whether they employ a fixed or open lexicon of entities and relation

   (a) 1. Schema-based approach

   (b) 2. Schema free approach (the problem is we cannot suss that if there are two triples represented in different ways: like short name is used, do they belong to the same entity or not)

Let, $E = \{e1, e2, ...\}$ be the set of all entities and R $= \{r1, r2, ...\}$be the set of all relation types in a knowledge graph. We model each possible triple $X_{ijk} = (e_i, r_j, e_k)$ over this set of entities and relations as a binary random variable $Y_{ijk} = 0, 1$ that indicates its existence.

1.5   Plan of execution

No.	Task	Objectives
		Tentative Workflow
1	Scope and Definition	- Feasibility - Areas to cover
2	Literature Review	- Conventional Techniques - NLP based models - Linguistical Modelling - Analysis for such similar tasks - Benchmarks - Evaluation
3	Datasets	New/Existing Datasets: - Where used? - Biased towards certain types? - Pre-processing done?
4	Previous Methods	- Implement/Run/Experiment - Test for our case - Limitations- Reasons -if not doing well?
5	Proposed Model/Representation	- Description - How do we try to capture everything in the talk? - Modelling the conversations/Compressing it - Rationale of building it up this way - How do we tackle these limitations?
6	Model Implementation and Training	- Data Gathering - Data Cleaning - Distributed Representations - Explore - Working on various spaces - individually - How do we unify these models? Or do we work on one model and try to train it to learn various ranges for choosing?
7	Debugging	Debug the model by analyzing where exactly the model goes wrong? Any patterns? - New features
8	Training again	Making these updates and training again
9	Validation	Validation and measuring performance
10	Benchmark	Ideate/Formulate a human Benchmark : - Form or something interesting - How do you filter out the noise? - Choose ques that can produce the best bang for the buck i.e shorter convos yet meaningful - Probabilistic analysis of the results we have - How much of it is by chance? Can you trust this human-based data?
11	Conclusions	Consolidate everything and start with completing the draft
12	Paper Submission	The work on this would run in parallel as it might be tough to recollect it later on. So, this phase is to be used to just put together the pieces collected so far.

1.6 Dataset Collection

```python
import praw
reddit = praw.Reddit(client_id = 'qhpenVXe2q7xCilAKYZasg',
                     client_secret = '-iM2pYW8gY2BeIeReVQ6rU04Ag',
                     username = 'redacted',
                     password = 'redatced',
                     user_agent = 'testing')

subreddit = reddit.subreddit('unpopularopinion')

conversations = {}

new_opinions = subreddit.new(limit = 10)

for thread in new_opinions:
    print("\n Thread_id :",submission.id)
    if not thread.stickied and thread.num_comments:
        print("TITLE : {}, UP : {}, DOWN = {},
               #COMMENTS = {}".format( thread.title,
                                       thread.ups,
                                       thread.downs,
                                       thread.num_comments))

        thread.comments.replace_more(limit=0)
        for comment in thread.comments.list():
            print(20*'-')
            print("Parent ID : {} \t
                   Comment ID : {} \n".format(comment.parent(),
                                              comment.id))
            print("COMMENT: ", comment.body)

            if comment.id not in conversations:
                conversations[comment.id] = [comment.body,{}]
                if comment.parent() != thread.id:
                    parent = str(comment.parent())
                    conversations[parent][1][comment.id] =
                        [comment.ups,comment.body]


for post_id in conversations:
    message = conversations[post_id][0]
    replies = conversations[post_id][1]
    if len(replies) > 1:
        print('Original Message: {}'.format(message))
        print(35*'_')
        print('Replies:')
        for reply in replies:
            print(replies[reply])
```

The dictionary below can be called the pseudo-data just to have an intuition of how the we are storing the conversations :

```python
conversations = {post_id: [parent_content,
                           {reply_id:[votes, reply_content],
                            reply_id:[votes, reply_content],
                            reply_id:[votes, reply_content]}],
```

```
                    post_id: [parent_content,
                              {reply_id:[votes, reply_content],
                               reply_id:[votes, reply_content],
                               reply_id:[votes, reply_content]},
                    }
```

A demo piece of the data is as follows :

```
demo = [(key,value) for key,value in conversations.items()]
demo[:10]
```

Out[91]: [('hwn3ubb',
  ["Don't really think this is unpopular, or at least unpopular on reddit",
   {'hwn3yhx': [1, 'Welp I guess I saw this comment coming a mile away.']}]),
 ('hwn3ysy',
  ['Maybe the point you should try to make is to not hate other people's hobbies. People like to troll people who are into stuf
f they're not into. Its like when people who like rock music talk shit about rap. People who enjoy watching endless videos of t
heir favorite influencer talk shit about sports fanatics',
   {}]),
 ('hwn3yhx', ['Welp I guess I saw this comment coming a mile away.', {}]),
 ('hwn40vw',
  ['My ex best friend had unbearable issues that he projected onto others relentlessly. I swear he couldn\'t understand the mea
ning of the word Boundaries. I remember trying to push him to try therapy or something of the sort. And his response was, "Hone
stly I think I just need a good woman to stick it out and fix me."\n\nHe didn\'t understand why I got pissed off, explaining ho
w disgusting that mindset was. Way to trap someone into dealing with your bs that You yourself refuse to. People like that need
to grow tf up.\n\nWe stopped associating not long after. Though I feel bad for his current gf.',
   {}]),
 ('hwn1rwx',
  ['i fucking hate mustard, just my opinion tho',
   {'hwn1vdf': [1,
     'I figured thats why it would be a good unpopular opinion. Not a lot of people seem to like mustard.']}]),
 ('hwn1xsp',
  ['Hmm… I'll have to give this a try, but it'll have to be damn good to dethrone Heinz catsup.',
   {'hwn23y9': [1,
     "If you are a ketchup person, it probably won't be for you. Ketchup is so much sweeter than yellow mustard. Yellow musta
rd has more tang and bite to it, which is why I like it a lot more. It feels like its augmenting my fry, not blending into i
t."]}]),
 ('hwn25ou',
  ['I love mustard with fries. Sometimes I mix mustard and ketchup, but the mix is more mustard.',
   {'hwn2gyw': [1,
     "The rare mixer connoisseur! An intentionally mixer is a rare person indeed. \n\nIts got an interesting flavor when mixe
d together. Almost like a chili sauce? I can't quite describe it."],
    'hwn3rk2': [1, 'I do that for corn dogs so good']}]),
```