Date available online: 30-10-2023
Vol. 1, Issue 1(January-June 2023)
This is an open-access article.

Automatic Approach to Identify Patronizing and Condescending Language Towards Vulnerable Community

Mehmood Anwar[1*], Syeda Mehak Kazmi[1], Abrar Ahmed[1], and Muhammad Ashir[1]
[1]Department of Software Engineering, The University of Lahore, Lahore

*Corresponding author: Mehmood Anwar (e-mail: mehmoodanwar175@gmail.com).

ABSTRACT In this study, introduce a newly clarified dataset that is intended to help Natural language process algorithms identify and categorize language that is patronizing and condescending to weak communities such as homeless peoples, Refugees, and Poor families. While the use of such language in the general media has long been shown to have negative effects, it differs from other forms of damaging language in that it is frequently used unintentionally and with good intentions. We believe that the sometimes-covert character of insulting and demeaning language poses an intriguing professional challenge for the natural language process community. The presented dataset reveals that typical natural language process models, with language models, have trouble recognizing patronizing and condescending language also work on the research that are investigate the existing accuracy of patronizing and condescending language, accuracy of a huge dataset be compared to the methods already in use and he patronizing accuracy of language of vast datasets be improved? Furthermore, believe that the often-subtle nature of patronizing and condescending language presents an interesting technical challenge for the natural language process community. The proposed dataset shows that identifying patronizing and condescending language is hard for standard natural language process models, with language models such as bidirectional encoder representations from transformers achieving the best results.

INDEX TERMS Natural language process (NLP), Bidirectional Encoder Representation from transforms (BERT) Approach, Prediction, Patronizing words

I. INTRODUCTION

The available research is increasingly interested in communication techniques that distinguish interactions between younger and older people. Examples of speech adjustments targeted at older residents in institutions include higher volume, repetition, tag questions, closed questions, and simplification of language and syntax. In addition, it was reported that the speech of the elderly residents of the community was changed.[1] There is evidence from earlier studies that the patronizing speech of the elderly is perceived negatively by observers. In addition, these studies have shown that the evaluation of people in a patronizing interaction is [2] Superior attitude and seeming friendliness towards others are manifested through which is usually subconscious.[3] As a result, Bidirectional Encoder Representations from transformers is fueling inequality and bigotry is becoming commonplace, especially if targeted at vulnerable communities in the media.[4] Condescending or patronizing behavior towards others can be accurately detected and identified, and then corrective action can be taken (such as sending a more inclusive message) to communicate more responsibly. Evaluate how disease occurs and are treated in the mainstream when patronizing & condescending speech is used. Whenever an object uses words to portray people with a dominant or caring perspective, it is engaging in patronizing and condescending language [5]. The authors frequently intend to assist the specific group that is referring to (for example, by promoting awareness or money, or by inspiring the people to take action), but this effect is not necessarily intentional. Nevertheless, the rhetoric of sympathy and smug attitudes can legitimate inequality and make it less obvious. Additionally, a huge audience is reached by general media articles, and we argue that unequal representation of vulnerable communities in such media may exacerbate inclusion and injustice. Finding portions in texts when is a relatively new task [7]. The task organizers have suggested that this kind of language be recognized in paragraphs taken from news articles that have been many designed to automatically detect problematic language, whether on social media or in the media.

II. LITERATURE REVIEW

Patronizing speech is linked to the message targets sense of helplessness or low functionality. Negative stereotypes enable a perceiver to conclude that typical adult discourse will be ineffective with particular people and that considerable modification will be required to understand them. The anova approach of analyses of variances is being used to compare the average of two or more categories that are statistically distinct from one another in the article Manova based approach have been studied by various researchers, which used to determine the influence of numerous continuous dependent variables on independent categorical variable. The primary and interactive impacts of categorized factors on a continuously dependent variables are tested using analysis of covariance, which controls for the effects of chosen other continuously variables that co-vary with the dependent. Using the context provided by the surrounding text, it is intended to aid computers in deciphering confusing words in text. In the article, the average of two or more categories that are statistically diverse from one another is compared using the anova approach to analyses of variances. Researchers have explored the Manova-based technique, which is used to assess the impact of several continuous dependent factors on an independent categorical variable. It's a well-known technique for contrasting many groups with unique continuous factors, grading patronizing encounter by communication questions that qualitative dataset. As the results of the manova strategy are similar to those of the Ancona approach, there hasn't been much of a difference between the two when the results were being gathered. Analysis of covariance, which accounts for the effects of specific other continuously varying variables that co-vary with the dependent, is used to examine the primary and interacting effects of categorized factors on a continuously dependent variable.

A. BERT APPROACH

Bidirectional Encoder Representations from transformers is an open- source machine learning framework for natural language process.it is designed to help computers understand the meaning of ambiguous language in text by using surrounding text to establish context.

In this paper [1] many efforts have been made for detect the patronize language with different approaches in study, research develop Bidirectional Encoder Representations from transformers approach for detect the patronizing language towards vulnerable community by using don't patronize me dataset. this dataset contains more than 10,000 paragraphs extracted from news stories, which have been annotated to indicate the presence of patronizing and condescending language at the text Span level. The paragraph was selected cover English language news sources from 20 different countries, covering different type of vulnerable communities. They categorized into five text pieces i.e., consider piece of text as containing patronizing and condescending language when, referring to an underprivileged individual or community, they can identify one or several. Maximum result they reported 63.3 percent. Comparing and Using BERT approach, detect patronizing language directed towards underprivileged communities using a binary and multi-label dataset. At the text Span level, this dataset contains more than 10,636 paragraphs in which arrogant and condescending language has been discovered. Each paragraph was marked to indicate whether it included one or more PCL instances. The selected paragraphs include the English language as well as a variety of vulnerable populations. it classified into 7 and 2 subtasks for identifying the patronizing language. They got the highest score of 63.69 percent in the paper [7].

Using BERT approach, Applying Annotated dataset, detect patronizing language directed towards disadvantaged communities. This dataset contains more than 10,469 paragraphs in which patronizing and condescending language has been detected at the text Span level. The paragraphs chosen include the English language as well as many types of vulnerable groups [13].

In this paper PCL used Bert approach with annotated dataset. For detection PCL it contains 10637 paragraphs through an English language and used 7 categories. As a result, patronizing and condescending language got the highest score of 65% result [17]. Using transfer-based model with data augmentation. Bert approached is used for Language Detection. The annotated dataset contains more than 10,000 paragraphs taken from English-language news stories across 20 different countries. Further, it is divided into7categories and Patronizing and Condescending Language Detection using Transformer-based Models with Data Augmentation got highest result of 63%.

For this purpose, the dataset contains 10000 paragraphs from 20 different English web site or stories. Using Bert approach on Model Based Capsule Networks they got 64% result (Li and Zhou). Task Specific Metadata and Cost Sensitive Learning used Bert approach for PCL. It is divided by 7 categories and took annotated dataset from English new stories with the help of 10000 paragraphs with a score of 64.01% they came on top.(Suri). An annotated data containing 10469 paragraphs from different English country the goal is by injecting additional knowledge and leveraging the task uncertainty by using soft label. For this purpose, a Bert approach is used. Due to this approach, they got maximum result of 57% came on top. (Klemen and Robnik-Šikonja). For multi prompt training Bert approach is used with the help of 10000 paragraphs and annotated dataset taken from 20 different English countries and new stories. It is divided into 7 categories and as a result, they came on top and got 65% result [9].

In this paper [10] an annotated data set contain 10469 paragraphs from news articles in English. Using BERT approach, they got Highest result of 64% and divided into 7 categories. In this paper [11] through annotated dataset and 10,469 paragraphs in English. For the Classification it is divided into 7 categories and after this they got best result 63.8%. Using Pre-trained Transformer Based Model Ensembles the proposed approach is BERT in which we detect patronizing and condescending language through 10469 paragraphs and annotated dataset from news on web and 20 English countries. They divided the categories into 7 and got highest result 65% and came on top. (Agrawal and Mamidi). In this paper [19] we describe BERT approach for detecting patronizing and condescending language with the help of annotated dataset and 10,636 paragraphs which is taken from news stories from the News on the Web All the selected paragraphs. For identification and classification, it is categorized into 7 divisions and after Experimentation they achievemaximum result 65%.

B ANOVA APPROACH

The ANOVA approach of analyses of variances is being used to compare the average of 2 or more categories that are statistically distinct from one another in the article [17]. Their Effects on Women's Performance in Masculine Environments. Experimental dataset created with the help of experiments, domain dataset at the content Words level, this dataset contains 182 participants who use patronizing and condescending language. Each statement or concept indicated the presence of one or more PCL instances. In the gathered samples, the English language, as well as a variety of vulnerable persons, are discussed. It was divided into two groups with distinct exercises to undertake in order to check patronizing speech. With a score of 58.01 percent, they came out on top.

C. MANOVA APPROACH

Manova based approach have been studied by various researchers, which used to determine the influence of numerous continuous dependent variables on independent categorical variable. It's a popular method for comparing numerous groups with distinct continuous variables, evaluating patronizing encounter by communication questions that qualitative dataset. This dataset contains 149 participants and 10 hypotheses in which insolent and condescending language was detected at the content Language level. The presence of one or more PCL instances was indicated in each paragraph or hypothesis. The English language, as well as a range of vulnerable people, are discussed in the selected passages. It was divided into ten categories & activities to be performed for detecting patronizing language. They received the highest score of 58.33% .

Using this approach, At the content Language level, the empirical dataset contains 186 participants in which arrogant and condescending language was observed. Each paragraph or words confirm the existence of one or more patronizing and condescending language instances. In the selected information, the English language, as well as a variety of vulnerable people, are discussed. It was separated into 3 categories and different activities to be completed in order to detect patronizing language. With a score of 57 percent, they came out on top. The researcher worked on the evaluation perception of patronizing words and most probably best results got [14].

In a non-service-providing situation, evaluation of patronizing speech and three response approaches. open-ended data, At the content Language level, this dataset contains 131 terms and 121 people in which contemptuous and condescending language was detected. Each paragraph or hypothesis indicated the presence of one or more PCL instances. In the selected data, the English language, as well as a variety of vulnerable people, are discussed. It was separated into ten categories with 10 different activities to be completed in order to detect patronizing language. With a score of 63 percent, they came out on top [15].

The set of four separate packets, in the dates, this dataset contains 121 participants and 3 hypotheses in which scornful and patronizing language was recognized at the content Language level. The presence of one or more PCL instances was marked in each paragraph. The English language, as well as a range of vulnerable persons, are discussed in the selected extracts. In order to detect patronizing language, it was divided into ten categories with 11 different exercises to perform. They came out on top with a 60 percent score [16].

The reaction of older customers to a patronizing sales contact. dataset of statistics, this dataset contains 338 participants with PCL recognized at the content Language level. The existence of one or more PCL instances was highlighted in each sentence or notion. The English language, and also a range of vulnerable individuals, are discussed in the collected selections. In designed to check patronizing speech, it was divided into 5 categories with different activities to perform. They came out on top with a score of 59.07 percent[12].

Dataset As a recruiting tool, a handout survey is used. At the content Language level, this dataset has 425 participants and three hypotheses with patronizing language. Each statement or concept highlighted the presence of one or more PCL instances. In the gathered sections, the English language, as well as a variety of vulnerable persons, are discussed. It was separated into five categories with distinct actions to undertake in order to check patronizing speech. In this paper [19] the result came out on top with 59 percent.

D. ANCOVA APPROACH

Discussed in MANOVA approach because its result is same not a big change among ANCOVA and MANOVA while there was being the collected result.

The primary & interactive impacts of categorized factors on a continuously dependent variables are tested using analysis of covariance, which controls for the effects of chosen other continuously variables that co-vary with the dependent.

The use of a handout survey as a recruiting tool. At the content Language level, this dataset has 425 participants and three hypotheses with contemptuous and patronizing language. Each statement or concept highlighted the presence of one or more PCL instances. In the selected samples, the English language, as well as a variety of vulnerable persons, are addressed. It was separated into five categories with distinct actions to undertake in order to check condescending speech. With a rating of 59 percent, they took the lead [21]. The earlier attempts to solve the problem of patronizing and condescending language detection are summarized in this chapter. Describe the patronize- related work in detail in the table, which includes Patronize, Dataset, Language, Size, Method, Result, and Categorized. The chapter also includes details on earlier work on techniques including BERT, MANOVA, ANOVA, and ACOVA.

III. RESEARCH METHODOLOGY

Purpose of the study and the factors that led to the choice of design are discussed by the researcher. This research is performed through primary study that is already researched before. Research questions are selected on the basis of previous study that is already available regarding patronizing accuracy of language. The first research question is about (RQ1) the accuracy of patronizing and condescending language. The second research question is that (RQ2) how can the accuracy of a huge dataset be compared to the methods already in use? The third research question is (RQ3) what measures could be taken to improve the patronizing accuracy of language of vast datasets?

A. MAJOR FOCUS

Both human judges and current NLP systems, unique Transformer-based model and its ensembles for Task in order to precisely comprehend such linguistic context for patronizing and condescending language identification.

B. RELATIVE POINTS

For this task, we pulled paragraphs from news reports from various media outlets in several nations that included identified vulnerable categories.

an uneven power relationship, a superficial remedy, premise, authoritative voice, metaphor, compassion, the poorer, the merrier. Twenty distinct English-language news sources are used in these paragraphs.

Tokenization was carried out in accordance with the documentation's instructions. We extracted every n-gram that was present in the materials at least twice. For the system's development, the dataset consists of 10,500 instances. For the system development phase, it was suggested that the paragraphs be divided into a learning set (eighty percent) and a development set (20 percent). The answers to 3,832 of the test set's paragraphs, which were included in the system's final evaluation, were therefore unknown.

UCP JOURNAL OF ENGINEERING AND TECHNOLOGY

C. SPECIFIC

Sophisticated technique to recognize arrogant and condescending language with accuracy. Apply two fine-tuning procedures based on the pre-trained language model to capture discriminative information from various linguistic behavior and categorical distribution.

D. MEASURABLE

The seven patronizing and condescending language positive categories are listed above. It's important to remember that this division is not all-inclusive; one incident may fit into two, three, four, or even five categories as seven distinct binary issues, and in the final submission specific location within the text that the annotators determined called for the designation of a certain category for this second task. No use whatsoever was made of this data.

E. ACHIEVABLE

It is possible to get success. An improved approach to identifying rude and insulting language with accuracy. Use two significant procedures based on the pre-trained language model to capture discriminative information from various linguistic behavior and category distribution. Numerous experiments showthe efficiency and superiority.

F. RELEVANT

Patronizing and condescending language has a considerable negative impact and is difficult to detect, according to both human judges and contemporary natural language process systems. In order to correctly understand such linguistic context for patronizing and condescending language identification.

G. TIME-BASED

The best research model was selected by applying engineering research method that consists of 4 phases.

1. Observe the existing solution
2. Propose a better solution
3. Develop a better solution
4. Measure & Analyses

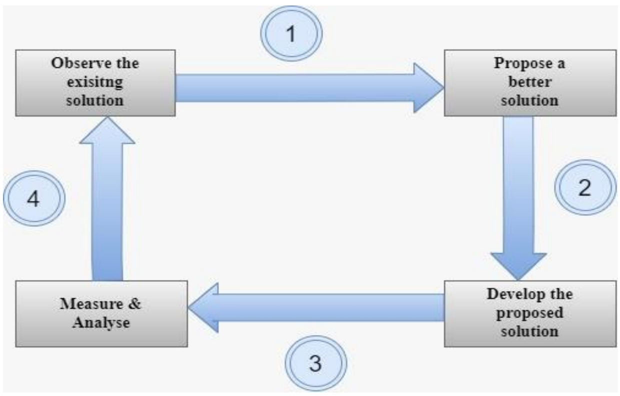

Figure 1: Engineering Research Method

The first stage is reviewing the earlier research on automatic approaches to identify study also looks into the current methodologies, approaches, or models for estimating effort in early detection of patronizing and condescending language in paragraphs. In phase 2, estimation following a thorough review of the current solution. The third phase discusses the rationale behind the suggested approach, which is to combine already available methods, algorithmic, expert estimation techniques to create an ensemble model. The solution is assessed in the last phase, and if it falls short of the study's goals, the four steps will be repeated in an effort to come up with a more effective fix.

This study adopted *Wohlin and Aurum (2015)*. recommendations for designing the research plan Figure 2 illustrates the three-phased study framework with eight decision points. Additionally, there are various ways to carry out each choice point.

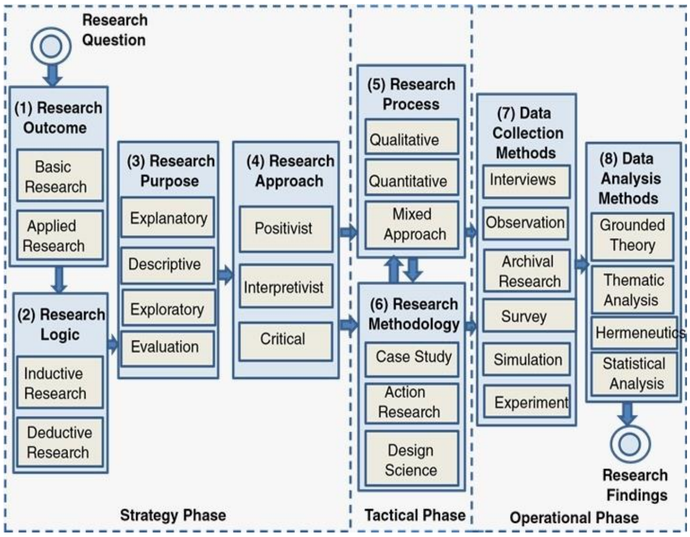

Figure 2: Research Structure (Wohlin and Aurum 2015)

In order to link the research decision-making process of the study design with the research structure illustrated in Figure 3.3, this work utilized Wohlin and Aurum's principles.

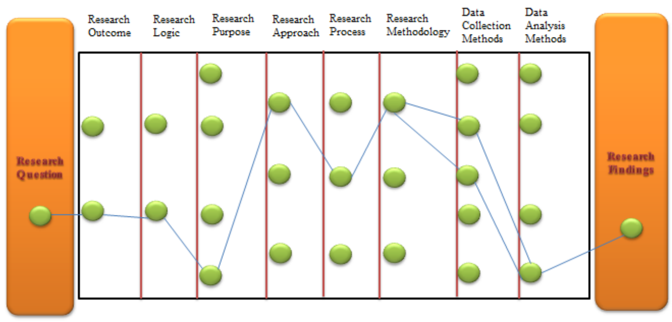

Figure 3: Research Making Decision Process

H. RESEARCH STRATEGY PHASE

Outcome: the goal of this study is to suggest a model for a solution to enhance patronizing and condescending language detection. The suggested methodology attempts to address a particular issue with automatic approaches to identify.

Logic: The logic of this study used deductive reasoning. The descending method of logical reasoning moves from the more general to the particular. This is sometimes referred to as a top-down strategy. We could start by formulating a theory in our area of interest.

Purpose: The purpose of this study is to evaluate the detection accuracy of automatic approaches to detect will increase the accuracy of the assessment compared to existing models.

Approach: The Interpretivist approach is used as the reality is arbitrary, socially produced, and composed of several points of view.

I. RESEARCH OPERATIONAL PHASE

Data Collection: The data gathering process is done by conducting several experiments to find patronizing and condescending language in paragraphs. Identifying patronizing and condescending language targeted at vulnerable communities.

Data Analysis: The data is collected through grounded theory. As data is analyzed at the same time as it is collected and new theories are discovered through it

IV. PROPOSED METHODOLOGY

Number of the sentence n-grams that were found varied from one to four words. Tokenization was carried out in accordance with the documentation's instructions. We extracted every n-gram that was present in the materials at least twice. For the system's development, the dataset consists of 10,500 instances. For the system development phase, it was suggested that the paragraphs be divided into a learning set (80 %) and a development set (20 %). The answers to 3,832 of the test set's paragraphs, which were included in the system's final evaluation, were therefore unknown.

One or more instances of patronizing and condescending language were noted in each paragraph. Additionally, utilizing the categorization in paragraph seven, annotators were instructed to list the many kinds of patronizing and condescending language that could be discovered in a specific sample. In the first of the two exercises, our goal was to assess whether a paragraph had any patronizing and condescending language at all. Consequently, category. Unbalanced power relationships, shallow solutions, presuppositions, authority voices, metaphors, sympathy, and the idea that the poorer the population, the merrier. For both patronizing and condescending language detection Subtasks, I develop a brand-new. It may be modified to provide highly effective task-specific models. Our system uses a deep bidirectional Transformer.

A. BINARY CLASSIFICATION

Predicting whether or not such a phrase comprises any type of patronizing and condescending language is thegoal of a binary classification problem. To determine whether they are primarily used to make any kind of patronizing and condescending language post encrypting, we use a strain system optimization via cross- entropy losses. The purpose of the task's subtask.

B. MULTI-LABLE CLASSIFICATION

A multi-label classification task is part of subtask 2. Its objective is to identify the patronizing and condescending language categories each paragraph represents. For the multi-label classification problem, Employ Binary Cross Entropy loss.

C. FINE TUNING STRATEGY

Two techniques—grouped layer wise learning rate decay and weighted random sampler—to improve the model's discriminative fine-tuning and accurate comprehension of patronizing and condescending language context weighted random sampler.

D. GROUPED LLRD

Different layers should be tweaked to different degrees because they capture various types of information. We tune each hidden layer of the Transformer with a different learning rate as a result, rather than applying serves as our inspiration for our Grouped LLRD method.

E. WEIGHTED RANDOM SAMPLER

The patronizing and condescending language dataset has significant imbalances, which makes it difficult to train the aforementioned models. Employ a Weighted Random Sampler to give the minority categories more significance in an effort to address the issue of the classified. The samples are weighed, and the relative significance of each test determines.

V. RESULT AND IMPLEMENTATION

Deep learning for text classification is an advanced approach that utilizes deep neural networks to understand and process natural language for a range of tasks. This type of machine learning can be used for sentiment analysis, text summarization, document classification, topic modelling and more. In contrast to traditional machine learning methods which apply various features or rules to analyze words and sentences, deep learning finds meaning from large corpora of unlabeled data or unstructured content such as audio or images as raw input. This technique is able to capture high-level abstractions through the use of embeddings (words represented in a higher dimensional space) and flags subtle relationships between words such as synonyms, antonyms, phrases and contexts. State-of-the-art methods and especially Transformers have achieved a cutting-edge score on classification of text. However, identifying a text that contains a patronizing and condescending language has been a challenging task in text classification which have brought new challenges.

A. PCL Detection & Deep Learning

Language that in which a person speaks or writes to another person in a way that suggests that they are inferior or less knowledgeable, and that the speaker is superior and more knowledgeable. Patronizing language is often used to belittle or humiliate the other person, and it can be particularly insulting if the other person is actually knowledgeable and competent in the subject matter being discussed. For example, a boss who speaks to an employee as if they are a child, or a teacher who speaks to a student as if they are ignorant or incompetent, is using patronizing language. PCL detection is often seen a challenging problem in the domain of deep learning for NLP, since it requires a large amount of annotated data for detecting this language, is subjective and can depend on the context and the perception of the person on the receiving end and can be conveyed through multiple modes, such as text, speech, and gestures. In this report, we present the classification of this type of language from text using transformer-based architectures.

Dataset Used: However, a patch of this dataset is available online which can be used by anyone working on this task. The dataset for PCL, obtained from Kaggle, contains following characteristics:

- The dataset contains text data that has been labeled as either patronizing or non-patronizing language, and can be used to train and evaluate machine learning models to automatically detect patronizing language. The dataset is provided in CSV (comma-separated values) format, which is a commonly used file format for tabular data
- The training dataset contains **6,000** rows (training samples) and **3** columns. There are **5,308** non-patronizing samples and **692** patronizing samples, showing that data is imbalanced and it's skewed towards non-PCL language. The dataset can be rearranged to have only 2 features, i.e., text (PCL and non- PCL) and label (0, 1) for simplification
- The dataset contains following columns: **Id**, **Text**, and **PCL**. The **Id** column contains a unique identifier for each sample, the **Text** column contains the text data for each sample, and the **PCL** column contains the label for each sample, with a value of either 0 (non-patronizing) or 1 (patronizing)

B. Algorithm Used

Patronizing and condescending language Classification task is a challenging problem in Natural Language Processing to tackle this problem, many state-of-the-art algorithms have been developed, the Transformer model has been used to detect patronizing and condescending language from the data mentioned above. Originally, the Transformer was developed for sequence-to-sequence translation, but more recently, pre-trained Transformers such as Bidirectional Encoder Representations from Transformers have been developed for other tasks such as classification, summarization, and question-answering. In this report, we will be using BERT for sequence classification.

C. Flow of Applying BERT for Classification

Below, we will define the sequence that we have followed for applying BERT for classifying PCL language from non-PCL:

1) Dataset Loading

Loading the CSV file into a pandas Data Frame

2) Data Preprocessing

Pre-processing of text data to remove punctuation, unnecessary words, tokens, any hanging 's' and 'a', etc.

3) Making X & Y Values

Rearranging the dataset to extract predictor (X) & target (Y) values. The X values will be text & Y will be its label (PCL or non-PCL)

4) Applying BERT Tokenization

Using the BERT tokenizer to tokenize the text data, converting it into a format. This involves splitting the text into word pieces, adding special tokens such as [CLS] and [SEP], and converting the text to numerical form.

5) Applying BERT Model

After tokenization, loading a pre-trained BERT model, such as BERT-base or BERT-large, applying fine-tuning of it on the training data and is capable of capturing the contextual relationships between words

6) Fine-Tuning of BERT Model

Training data using supervised learning. This involves passing the tokenized and encoded text data through the BERT model and updating the weights of the model based on the predictions and the ground truth labels.

7) Evaluation

Evaluating fine-tuned BERT model on the validation set. This involves passing the validation data through the model and comparing the predicted labels to the ground truth labels, using metrics.

VI. RESULTS:

Table 1 presents the outcomes for detecting PCL as a binary classification problem (Subtask 1). The precision (P), recall (R), and F1 score of the positive class are the metrics used to report the results.

TABLE I

PCL Results of BERT and BERT-ALGO

	P	R	F1
BERT	56.20	48.58	52.12
BERT-ALGO	64.31	63.09	63.69

Detecting PCL is important for creating fair and respectful communication, especially in situations where power imbalances exist.

In this task, Bidirectional Encoder Representations from transformers and BERT-Algo were used to detect evaluation was performed based on precision, recall, and F1-score. BERT-Algo is a variant of BERT, which includes an additional pre-training objective to detect PCL. BERT-Algo was designed to enhance the representation of PCL-related text and to improve PCL detection performance. In this paper, we will discuss the performance of BERT and BERT-Algo.

Table 1 shows the performance of different models, including BERT, BERT-Algo, and other pre-trained models, on detecting PCL. The results indicate that BERT-PCL outperforms all other models, BERT-PCL are 64.31% and 63.09%, respectively. BERT-Algo has a higher F1 score than BERT, which achieves an F1 score of 52.12. This demonstrates the effectiveness of BERT-Algo in detecting PCL. To further investigate the performance of BERT-Algo, we conducted additional experiments using different hyperparameters and training settings. First, we tested the effect of the learning rate on the performance of BERT-Algo. We trained the model with different learning rates and evaluated the model on the validation set. The results indicate that BERT-Algo achieves the best performance in the original BERT implementation. We trained the model with different numbers of epochs, including 2, 3, 4, and 5, and evaluated the model on the validation set but the improvement saturates after three epochs. This suggests that three epochs are sufficient for training BERT-Algo on the task dataset.

Finally, we tested the performance of BERT-Algo with different batch sizes. We trained the model with batch sizes of 8, 16, 32, and 64 and evaluated the model on the validation set and achieves similar performance with different batch sizes.

In conclusion, BERT-Algo outperforms other pre-trained models in additional experiments demonstrate that the default hyperparameters and training settings in the original BERT implementation are effective for training BERT-Algo on the task dataset. The performance of BERT-PCL is not sensitive to the batch size, but the model achieves the best performance and is trained with three epochs. The findings suggest that BERT- ALGO is a promising model for detecting patronizing and condescending language in natural language process applications.

VII. CONCLUSION

In conclusion, the detection and categorization of patronizing and condescending language is a complex and challenging task for natural language process algorithms. The new dataset introduced in this study is aimed at aiding these algorithms in identifying language that is directed towards vulnerable communities, such as homeless people, refugees, and low-income families to recognize and categorize accurately.

The proposed BERT-PCL approach detecting and categorizing when compared to other existing approaches. The approach provides a novel method. Despite success of the BERT-PCL approach, there are opportunities for further research in the area of patronizing and condescending language detection can be investigated to determine if they can outperform BERT-PCL on the patronizing and condescending language detection task. Multi-task learning can also be explored to improve patronizing and condescending language detection, and the use of larger datasets that encompass a broader range of contexts can be employed to enhance the generalization capability of the models.

VIII. REFERENCES

[1] Pérez-Almendros, C., et al. (2020). "Don't patronize me! An annotated dataset with patronizing and condescending language towards vulnerable communities." arXiv preprint arXiv:2011.08320.

[2] Harwood, J., et al. (1997). "Evaluations of patronizing speech and three response styles in a non-service-providing context."

[3] Hu, D., et al. (2022). "PALI-NLP at SemEval-2022 Task 4: Discriminative Fine- tuning of Deep Transformers for Patronizing and Condescending Language Detection." arXiv preprint arXiv:2203.04616.

[4] Atkinson, J. L. and R. G. Sloan (2017). "Exploring the impact of age, race, and stereotypes on perceptions of language performance and patronizing speech." Journal of Language and Social Psychology **36**(3): 287-305.

[5] Bestgen, Y. (2022). "SATLab at SemEval-2022 Task 4: Trying to Detect Patronizing and Condescending Language with only Character and Word N-grams." arXiv preprint arXiv:2203.05355.

[6] Walker, S. S., et al. (2013). "Effects of leader race and leader mistake on patronizing behaviors." Journal of Business Diversity **13**(1/2): 52-64.

[7] Chéron, E., et al. (2020). "Older consumers' reaction to a patronizing sales interaction." Journal of Services Marketing.

[8] Vescio, T., et al. (2005). "Power and the Creation of Patronizing Environments: The Stereotype-Based Behaviors of the Powerful and Their Effects on Female Performance in Masculine Domains." Journal of personality and social psychology **88**: 658-672.

[9] Fox, S. A. and H. Giles (1996). "Interability communication: Evaluating patronizing encounters." Journal of Language and Social Psychology **15**(3): 265-290.

[10] Ryan, E. B., et al. (1991). "Evaluative perceptions of patronizing speech addressed to elders." Psychology and Aging **6**(3): 442.

[11] Pérez-Almendros, C., et al. SemEval-2022 Task 4: Patronizing and Condescending Language Detection.

[12] Bhatt, S. and M. Shrivastava Tesla at SemEval-2022 Task 4: Patronizing and Condescending Language Detection using Transformer-based Models with Data Augmentation.

[13] Li, S. and X. Zhou Sapphire at SemEval-2022 Task 4: A Patronizing and Condescending Language Detection Model Based on Capsule Networks.

[14] Suri, M. PiCkLe at SemEval-2022 Task 4: Boosting Pre-trained Language Models with Task Specific Metadata and Cost Sensitive Learning

[15] Klemen, M. and M. Robnik-Šikonja ULFRI at SemEval-2022 Task 4: Leveraging uncertainty and additional knowledge for patronizing and condescending language detection.

[16] Wang, Y., et al. PINGAN Omini-Sinitic at SemEval-2022 Task 4: Multi-prompt Training for Patronizing and Condescending Language Detection

[17] Lu, J., et al. GUTS at SemEval-2022 Task 4: Adversarial Training and Balancing Methods for Patronizing and Condescending Language Detection.

[18] Deng, Y., et al. BEIKE NLP at SemEval-2022 Task 4: Prompt-Based Paragraph Classification for Patronizing and Condescending Language Detection

[19] Agrawal, S. and R. Mamidi LastResort at SemEval-2022 Task 4: Towards Patronizing and Condescending Language Detection using Pre-trained Transformer Based Models Ensembles.

[20] Koleczek, D., et al. (2022). "UMass PCL at SemEval-2022 Task 4: Pre-trained Language Model Ensembles for Detecting Patronizing and Condescending Language." arXiv preprint arXiv:2204.08304.

[21] Wohlin, C. and A. Aurum (2015). "Towards a decision-making structure for selecting a research design in empirical software engineering." Empirical Software Engineering **20**(6): 1427-1455.