Identifying User Intent in Social Media Comments using BERT

Nidhi Malik[a], Yash Srivastava[b], Priyansh Aggarwal[c], Dev Sharma[d], Ujjwal Mehra[e]
[a,b,c,d,e]The NorthCap University, Gurugram
nidhimalik@ncuindia.edu

yash20csu130@ncuindia.edu

priyansh20csu209@ncuindia.edu

dev20csu032@ncuindia.edu

ujjawal20csu157@ncuindia.edu

Abstract

We have developed a newer system known as “Intent Recognition in Social Media Comments” to help us see more into the intentions of people interacting on social media. The Natural Language Processing (NLP) and machine learning is applied by our software for user comments grouping them into genres such as praise, criticism, suggestion or warning with high precision. A BERT model combined with an in-house built neural network architecture enhances interactivity by decoding human speech and their emotions concerning what they say. This helps businesses to understand where emerging market trends are going. One key attribute of our product is its fast rate in processing and analyzing user feedback that allows real-time users experiences on social media platforms. Additionally, it detects and removes offensive content from social media comments for online safety purposes. By doing these things in a proactive way, the platforms can minimize risks considerably while encouraging better digital environments that will build trust and loyalty among online communities. “Intent recognition in social media comments” initiative can be seen as one step closer to understanding user intent within this dynamic digital interactions landscape.

Keywords: BERT; Intent Recognition; Social Media; Natural Language Processing; Transformer Models

1. Introduction

1.1. Background

In this digital era of today, social media platforms have emerged as the main avenues where people can voice their opinions and interact with one another. Successful engagement and encouraging a safe and constructive online environment requires looking beyond user comments to appreciate the underlying emotions and intentions. This goes beyond content management and brand reputation, entering into an area of online security and fostering respectful conversations. Appreciating the subtleties of user expression is therefore essential due to the all-pervasive role that social media plays in our lives. It implies not only comprehending the textual meanings but also detecting hidden messages conveyed by emojis, pictures or videos shared. The proficiency in understanding these slightly hidden signs is instrumental in building more secure and rewarding online societies. In addition, it helps companies understand their customer’s requirements better as well as market trends.

1.2. Objective

The proposed work here has an innovative manner of identifying those who are the driving force behind the temperament of social media posts and comments. It examines numerous facets of linguistic and machine learning, which are used to correctly arrange the subject's intentions, such as compliments, crude remarks, sales pitches, or insults. The model employs the BERT language model tuned to chatbots with an artificial neural network of its own.

This project transcends the simple understanding of what users want. Wider benefits are derived from this.Enhanced

Social Media Interactions.Social media channels will be more effective in real-time communication. Improved User Experience:It makes websites and apps easier and more enjoyable to use. Refined Marketing Strategies: This improves on targeting in businesses' marketing campaigns and addressing consumer needs. Effective Brand Reputation Management:A company's reputation must be carefully managed so that it can continue being trusted by its customers.Improved Online Safety: It helps quickly identify and remove potentially harmful content, creating a safer online space.Foster Trust and Loyalty:* By proactively addressing user needs and concerns, it builds trust and fosters strong online communities.

The technology is dwarfs cover many elements apart from social media. To accomplish this, the United Nations has taken measures to enforce health, education, and governance. For example, in cases with health care, it takes measuring public opinion about a particular treatment, and providers to have better user experience while at the same time patients doctors. The lecturers may also be able to learn from students as they are conversing and plan to standardize their online learning experiences. Analogously, it may be used by policy makers to predict the way the mass public will feel on particular issues so that the policymakers would base their decisions upon real facts and will promote transparency in governance.

2. Literature Survey

In the first section of this paper, I wanted to show you what the old techniques have been used for interpreting the meaning of the expressions and the results in social media by examining users ' inputs and scoring sentiments. On this article, this time our attention is centered on both sides of results of the issues regarding those obstacles.

1. "Aspect-Based Sentiment Analysis Using Contextualized Word Embeddings: The increasing improvement of BERT's model which was presented by Akbar Karimi et al. [1]

The purpose of this research is to heighten the performance of BERT model in analyzing sentiments mentioned in reviews that contain item's attributes as well. BERT's covered layers prevail in building up the gradual accumulation and the hierarchical accumulation which aims to guide understanding. This pass roots emotionally tagging without the menace of the recurrent learning. On product review data sets, experiments was made and then, they proceeded even further and, of course, BERT proved to be even better for the task of sentiment analysis.

2. "SVM (Support Vector Machine) Classification by Yuzana Win as a separation system for detecting cyberbullying in Myanmar Language in Social Media"- Rina Usamawaty [2]

'Win has designed a method of doing language translation which understands what people talk about in networks of social media,' In doing so, the concept vectors of Word2vec method becomes the main parameters adjusting of SVMs. The system can tell from the smile on the face at the end of the musical experience that there is happiness.

be it a person pointing out a problem, voice dissatisfaction or even volunteer their ideology. It succeeds to compensate for some of the weaknesses yet it usually lets the grammar mistakes occur when dealing with more complicated sentence structures and unclear expressions.

3. An approach, BLIR or "BERT-based Language Model for Intent Recognition," by [Vasima] Khan, Tariq Azfar Meenai [3].

The research addresses the task of pinning down the users of any conversational system in order to find out how it can be understood better. On the basis of the characteristics of the rule-based and statistical implementations, choose individual applications. 2. Machine Learning (ML) Methods: - The implementation of techniques such as a Naïve classifier and a support vector machine (SVMs). - Along with the fine words tone range result, generate the reason of customer's feedback.

4. "The emergence of a hybrid model simultaneously addressing intent recognition and slot filling in spoken language understanding" Xiaodong Zhang and Houfeng Wang [4].

so saying this the research offers a simple but workable approach to the understanding that it is just the same as we can elsewhere say and it can be that it is not the case and that. The model, that is based on such as common bidirectional LSTMs and CRFs machineries, is used to learn the meaning of conversations. This is Smart Intent Management which covers: Open Intent Discovery facilitating a more thorough understanding of general user intent. Hence it goes beyond the limits of training examples and focuses on discerning general patterns of user intent Promts: - our model, while applicable in different spheres, gives the grounds for the search of a formula of efficient and high-quality goods production of our and all approaches.

5. This system uses algorithms to detect moving objects and their behavior, hat together create a Neuro-Network Based Surveillance System.Improvisation was a salient contribution of B.S. Harish and company [5]

The work concentrates on implementational framework of evolutional search algorithm which is a reflection of user's contextual search from natural language queries since it will be able to retrieve user intent from search phrases. By merging the approaches of: Machine Learning, Deep Learning and LSTM are examples of the proposed approach that becomes more knowledgeable from the user's input through the extraction of elements spectral. This approach is proved to be the best one in terms of improving the precision of the feature space, therefore the response of the search query becomes more relevant and provide fit answers for requests.

6. "AGIF: The expression "Adaptive Graph-Interactive Framework for Joint Multiple Intent Detection and Slot Filling" penned by Qin and colleagues [6]

Researchers have designed an algorithm which is able to capture the said speech and in that event it can decide what it intends to convey. The algorithm, through this complex neural network structure is a recurrent model, can consequently incorporate semantic and purpose breaks (tone of language) in the overall processing. In their work, the researchers point out the importance of intent-based as opposed to destination-based slots and employ in particular specific network elements intended to fit such issues as (Gated Recurrent Units and Max pooling layers) to handle them and pick up the patterns in time, so as to do timely classifications.

7. "BERT: In Google AI's research paper entitled "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding" [7], Jacob Devlin and other authors describe the BERT model.

This article presents a new neural network model that tackles two tasks: We have done division of texts into two types which are: contracts, commands or questions while genetic types plain sentences into sentence role and context. In light of BERT framework, it can intelligently and accurately to pick up multiple meanings (the sentence's main topic) that supersedes conventional models. The method, which consists in assigning individual labels to both intents and slots, shows up once again the capability neural networks to track down rather complicated forms of language rules.

8.. "Nils Caldera’s" article "Sequence to Sequence Learning with Neural Networks" [7] shows.

The authors prepare a more sensitive method called an Adaptive Graph Interactive Framework that has a good accuracy in identifying intentions through a single sentence. The framework uses the bidirectional LSTM model to understand human similarities and difference and extract the required slot details from a single utterance.

9. "BERT models intended to identify intentions using a frequency cut-off strategy for developing domain-specific vocabulary — ‘Fine-Tuning the BERT Models for Intent Recognition utilizing a Frequency Cut-off Strategy for Domain-specific Vocabulary Expansion’ by Fernando Fernández-Martínez et al. [9]

The authors compare two approaches: both combining approaches i.e. joint mapping and mixture mapping. This implies that, in contrast to lieu of location embedding, that the term mixture mapping is more accurate when it comes to introducing the model to domain-specific vocabulary. This technology is aimed at reforming the models for the Word2Vec approach by tailoring them to match the context of a specified domain eliminating the need for large amounts of unsupervised pretraining

10. Importantly, the paper "Joint Learning Framework with BERT for Spoken Language Understanding" by Zhichang Zhang and others is referred to as [10].

This article tackles whether the previous and current text classification methods of natural language processing are traditional or cheery. On the contrary, the intuitive choices for feature extraction of traditional machine learning tools such as CNNs (Convolution Neural Networks) and LSTM (Long Short Term Memory) have been becoming more popular. The article kickstarts with the clarification of a conditional RNN encoder-decoder model comprising the two RNNs differently than other RNN methods which enables it to solve problems others have thrown down.

3. Proposed Solution

Our method puts together NLP (Natural Language Processing) technology, latest, with a specially designed neural network, for helping with the purposes that people express in the comments section of social media. With the BERT pretrain language model ,which has such as a wide experience when it comes to comprehension, incorporated in a neural network personalized architecture, we exploit this capability.

Our model utilizes BERT to circle around in the twisted comments of social media, thus allowing us to detect and analyze user intent with considerable accuracy. Our approach is principle nor does it involve emojis as an extra though; they already are! Emoticons, which are often used instead of words during online conversations, are significant as they imply emotional implications of the emotional expressions impossible to spell out words. Applying emojis into our model is a way of employing this handy non-verbal communication, thus leading to larger and better understanding the owner of the message mood.

.

3.1 Model Description

3.A transformer-based model know as BERT, which stands for Bidirectional Encoder Representations from Transformers that dominates in language understanding, is a starting point during decoding. It clarifies the relation of the sentence through all words together where it reaches an ample comprehension of the meaning in either way. In fact, she helps that BERT knows how to do the natural language comprehension and classification tasks very well because she uses already found in the most accurate representations.

*3.1.2 BertForSequenceClassification. This class enhances the BERT model for classifying sequences by: * Creating of new layer classification on top of model. * Implementing correctional neural weights and parameters towards near-optimal results for sequence. * Training the modified pre-trained BERT representations on the intent classification task by e.g. transfer learning.*

3.1.3 Data Preprocessing using BERT Tokenizer The BERT tokenizer makes data ready analyzes by BERT, a language model during processing. It change the text to the number series sequence that(s) BERT can understand. To ensure that the codes are running as expected and to test the reliability of the algorithms.both have the same length during tokenization process, equal amount of important parts are taken into consideration (truncation), and special symbols (attention mask, e.g.).

3. 1.4 Training Parameters Well training parameters as these comprise of learning rate, number of training iterations and data batch size, are changed with great care to get the model to train more speedily and effectively. The technique *applies the AdamW optimizer so as to improve the predicted values compared with the actual, while a linear schedule fits the learning rate during the training sessions. The update of optimizations is very relevant to the model's learning and training system.*

3.1.6Realization and Assessment of the model is the formal part which is evaluated with the help of set of different measurements. Cross entropy loss is being inferred to as the metric of evaluation which is 3.0752 at last epoch run no.20.

3.1.7 Usage and Interaction in a Real-Time Environment Users will work with an interface having the features of the latest AI techniques to discover the upper level of a user's comments after their comments being instantly analyzed. It works as an edge that makes smooth the opening of extension ease applications where users are really active. The element demonstrates with slack by writing both the message and emoji. The interface can log in such a way, as the san to get idea of user's intentions. It allows for a mutual setting and lets the visitor to be more involved. Therefore, it assists the visitor to realize and respond to the artworks better.

The given model includes an integration of Tensorflow and Keras engine to ensure the adaptability be encapsulated by the range of structure of use and configurations. Apart from that, libraries such as torch, torchviz as well as transformers can facilitate the tokenization of data so as to help creators visualize the outcome and model development. This will improve its useability and facilitates the model's adoptability which makes the implementation process becomes flexible, especially in real-world scenarios.

3.2 Dataset Creation

3.2.1 Dataset Overview: This data which was built for this particular purpose drew its focus on nearly all major aspects of the factors which define the grassroots strategies.

social media usages became an instrument of the society's dialogues to identify and discuss different issues. On the other hand, the course characterizes another aspect of language interchange by precipitating the expansion of the vocabulary, grammar, and cultures of one language into another. That data encompasses

Tweaked so that its message could be transmitted via 9945 commediat to achieve the customer' wants by furnishing of the most desired kind of shoe among the 8 variations of shoes

work on models that not only smart go but also considered the contexts variation. Moreover, data arises from various social media platforms while generating data at the same time thus, it results in information improvement.

incorporating the topic as a whole and not just users' new habit of navigating within a single platform, but also provide new online communities.

3.2.2 Data Collection and Sources: The dataset is created by integrating the data from various sources like social media, online blogs, and user surveys.

website, blogging, social platforms, upload, and community contribution. Comments come from everywhere, whilst relying on training comments from the back and forth dialogue between the humans and the machine.

open gathering place for close communication, customer questions, idea exchange, and opinion giving. This blending of sources will not only ensure the involvement of real-life customer transactions from a diverse group, but will also provide a great deal of value and reliability to the dataset.

3.2.3Data Preprocessing: The next step would be to maintain consistency and quality of the dataset by removing comments and formatting them before adding them to the dataset. These steps are included in the cleanup: • Written comments are consistent and have strings format, so they will work well with other layers: To keep the process in consistency and work perfectly in further steps, comments are turned into stable strings.

• Paraphrased Text:Skipping Comments: The skipping process considers posts that carry irrelevant characters or emojis as unusable ones.The process then focuses its attention on the analysis of posts that carry a relevant intent to determine their subjectivity or objectivity. Encoding Intentions:Hence the intends that are sent with each comments is categorized as specific categories by means of techniques like one-hot encoding or label encoding. The purpose of the establishment is to involve in models training and assessment for the model's purposes of intention detection.

3.2.4 Sample Dataset Entries: Chart 1 is intended to describe a list of reviews in which there is a range of comments and labels showing if the comments are positive or negative. This data set symbolizes the types of messaging exchanged among users guased on their interacting habits which is showcased through examples in the data sample.

• "That's a really good point." (Intent: The previous communication pattern, which is usually having feedback mechanisms, avoids the case of extra revisions and files copying, so the process remains fast.

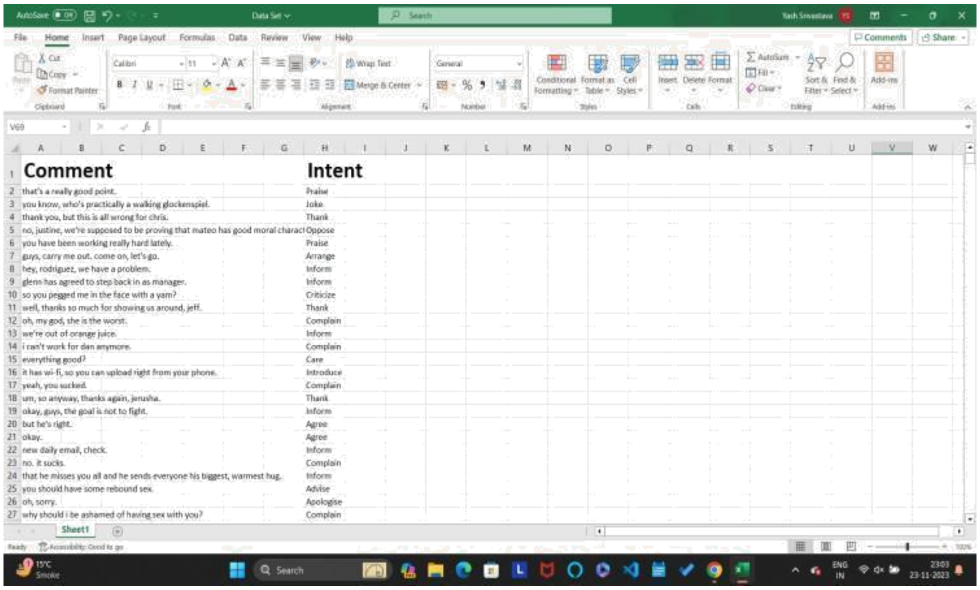
• "You know, who's practically a walking glockenspiel." (Intent: You are one of the factors that are behind my success, but I think you would be well placed in the Department of Earth Studies so as to give upon your focus on earth studies.

• "Thank you, but this is all wrong for Chris." (Intent: During the process of online interactions such as work, education, and social contact which take up most of our daily living cyber attack landscape has seen exponential growth. Privacy and security challenges have since been cast into one complex network.

• "No, Justine, we're supposed to be proving that Mateo has good moral character." (Intent: Next, students should be shown how to examine and apply precise media materials.

3.2.5 Emojis and Their Significance: The inventory covers written significant emoji at the same time their corresponding meanings are visible in them and therefore is diverse in emojis. Emotion on the other hand is a face that brings harmony and peace through gestures and emotion of emojis which assist in their expression online. This dataset is represented in a list and the corresponding equivalent codes. This characteristic is one of the major language skills that help people quickly understand what the others mean whenever they are speaking through the social media platforms.

3.2.6Dataset Representation and Structure: It is a program where one section is commentary while the other side contains information commentary displayed on tables for side by side comparison. Each sentence is individualized so that is can portray just a statement by community member as well as the sentiment that determines his/ her feelings. The emojis might be mounted on the comment somehow with the help of the Unicode Characters, what is a more informative tool in data description and how the users express themselves using the comment text. These social intelligence data can be suitably utilized for creating models that understand beyond the words in the context; they comprehend not only the context itself in conversations related to social media, but its complexities too, which is imperative when dealing with social media conversations. Different from all other online translators that majorly produce machine translations, which unfortunately limits our understanding of communication from different real life situations, this is different and unique in a way that it has a number of actual conversations that are brought about by the users. As a result, it becomes a perfect tool to help us study online communications and language.



	Comment	Intent
2	that's a really good point.	Praise
3	you know, who's practically a walking glockenspiel.	Joke
4	thank you, but this is all wrong for chris.	Thank
5	no, justine, we're supposed to be proving that mateo has good moral charact	Oppose
6	you have been working really hard lately.	Praise
7	guys, carry me out. come on, let's go.	Arrange
8	hey, rodriguez, we have a problem.	Inform
9	glenn has agreed to step back in as manager.	Inform
10	so you pegged me in the face with a yam?	Criticize
11	well, thanks so much for showing us around, jeff.	Thank
12	oh, my god, she is the worst.	Complain
13	we're out of orange juice.	Inform
14	i can't work for dan anymore.	Complain
15	everything good?	Care
16	it has wi-fi, so you can upload right from your phone.	Introduce
17	yeah, you sucked.	Complain
18	um, so anyway, thanks again, jerusha.	Thank
19	okay, guys, the goal is not to fight.	Inform
20	but he's right.	Agree
21	okay.	Agree
22	new daily email, check.	Inform
23	no. it sucks.	Complain
24	that he misses you all and he sends everyone his biggest, warmest hug.	Inform
25	you should have some rebound sex.	Advise
26	oh, sorry.	Apologise
27	why should i be ashamed of having sex with you?	Complain

Fig.1) Data Set

3.3 Methodology & Architecture

The research work is shifted below with the help of text processing tools that involved a pretrained BERT-based uncased model that aids in processing of text information, classifying text content to categorize user comments in sought for intents. The approach illustrates a framework of mechanisms that are designed in a sequence to improve the performance of the system and further enhance the ability of the system to identify user's intent. Step 1: Data preparation

A.Import the comments and intents into Excel

B.Let the comments to be input, and the correct intent fare the output

C.Use label mapping as rule for assigning value for intents,

D.Use label encoding rule for intents value assignment.

E.Filter out the negatives emoticons or inappropriate symbols in the data set, because it will impact the quality of the data. Figure 1: Illustration of data preprocessing

Step 2: Feature engineering Features that have been transformed from the original intent and comment by using data preprocessing can be regarded as "features" in deep learning.

Step 3: Model learning During this step of training the deep learning model, the machine will adjust the parameter according to the gradients. At this time, the model players the role of adjusting the weight of each parameter in their learning.

Step 4: Model evaluation Here is an example of testing on a machine whether it can correctly classify a person's intention. For example, as shown in Figure 6 below, type the input inside the box and click the appropriate intent to check if the model can correctly classify the intention. The smaller the distance between the input and the correct label, the higher the evaluation score. If the prediction result is "Indirect expressions of negative emotions towards the target", the evaluation score will be marked as satisfactory.

Model Training: Finally, we inform a specified BERT model for the classification. To train our dataset we'll split it into training (80%) and testing (20%) sets. It comes to this stage, we can assess the performance of the model. Our project extends BERT with classifying capabilities, using architectural improvements. We are going to lay a train on it by running it through many epochs with AdamW optimization. We'll also adjust learning rate according to the timeline to ensure the model will be as good as possible. The model takes both sentence by sentence coursework, sometimes complex and others easy. This ability to generalize makes it possible for the algorithm to learn from small data sets and similar scenarios.

Model Evaluation: To make a good evaluation of the model, we look to how good the model is at the data testing. Precision, recall and F1-Score scores are used to evaluate the quality of entities identified. These scores will provide a picture showing where the entity identification model excels and fails. It is extremely useful for making adjustments and correction, which does a fine job in driving the model smartness.

Prediction Function: To do that, a prediction function is built that shows predicted intents and confidence metrics for each user input at the moment the input is posted. Thus, it prevents user comments from getting the some of the emojis or other special characters which could spoil the prediction accuracy and go on with the prediction. When the data comes in, the function performs tokenization and preprocessing as an attempt to keep the comment in full text form Therefore, it traps the model using that which has been trained to detect the unlabeled intent level and the confidence level for the said statement.

User Interaction: User interaction is produced by the interactive spiral of user and computer. This implies the possibility to check the sentiment and the strength of comments and automatically predicting intention and confidence levels. The cycle will keep on running whenever the loop detects 'exit' command, so that there should be good synergy between the human and machine that makes the system more user friendly and interactive.

In brief, the approach which we addressed unites the most effective ways to find out about different people attitude to social media. It applies most modern tools and methods of a non-human intent candidacy in online chats and posts to be punctual and accurate. Some key steps make this effective: repair raw data is expected, importantly removing symbol from the sentence. After that, using more advanced text processing tools, it identifies the significant characteristics of a text, such as keywords and entities.

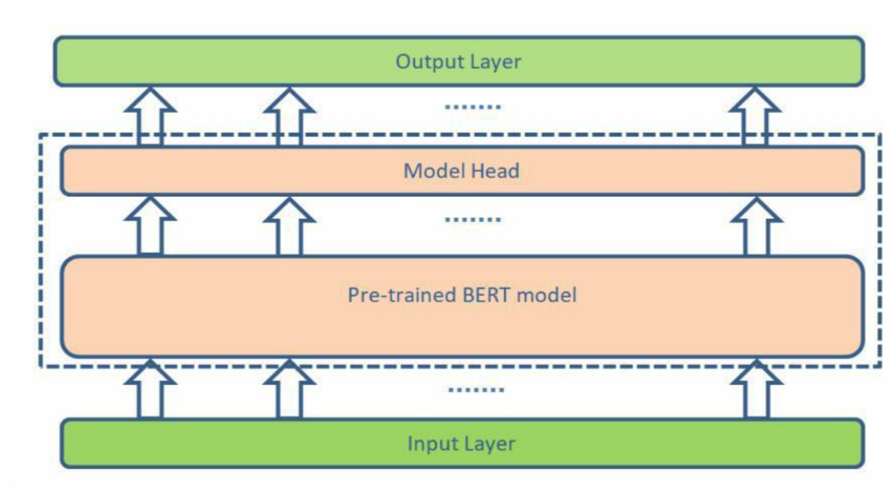


Fig 2.) BERT - Masked Language Modelling Diagram[10]

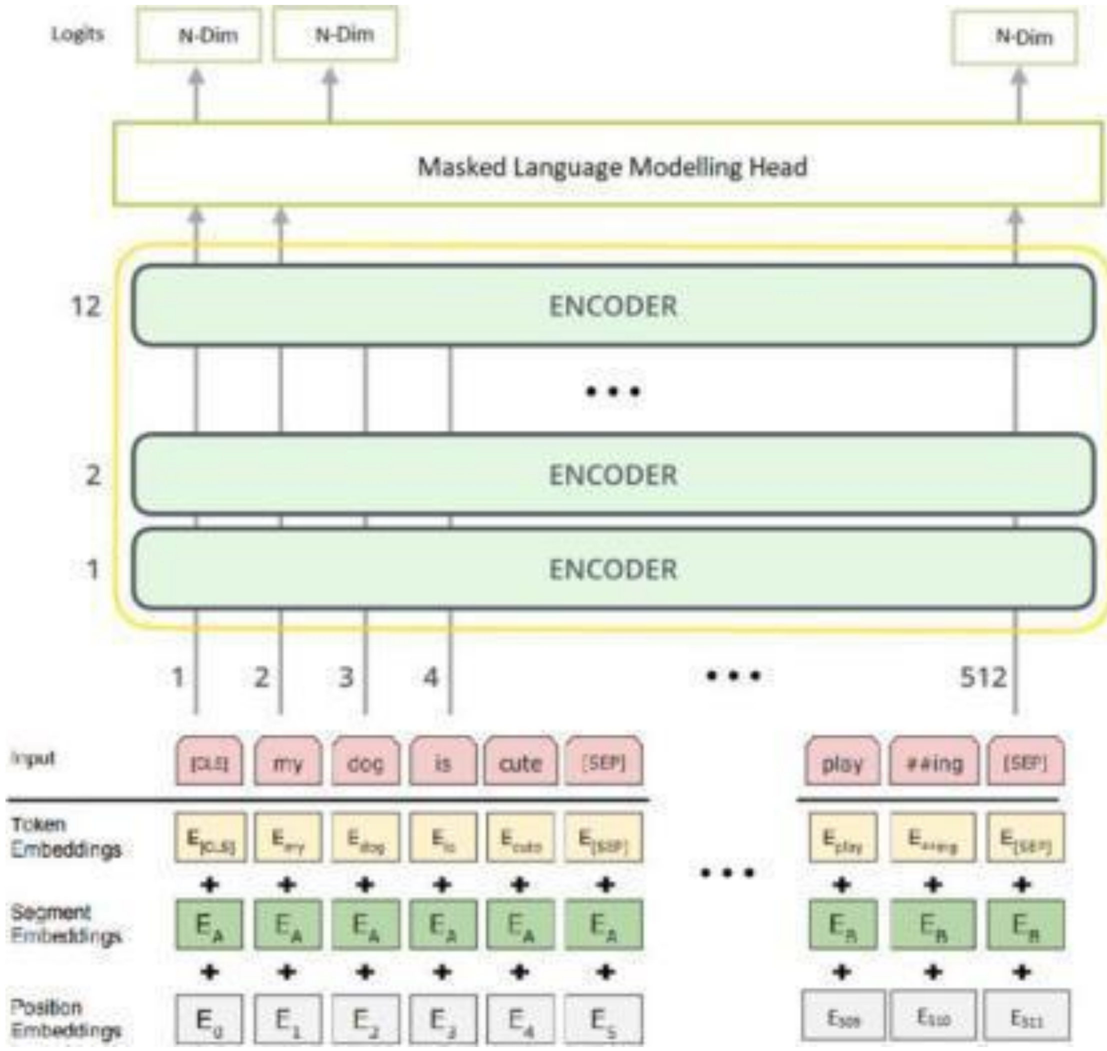


Fig 3.) BERT Base Model - Masked Language Modelling Diagram[10]

The model's input has in total three components [10]:

Positional Embedding: this represents, index number of the input token.
Segment Embedding: this indicates the sentence's position’s in the sequence of sentences.
Token Embedding: it contains the token/tokens generated by the tokenizer for a given word/words.

2. Results

Our proposed approach to how users can be understood in comments have already been confirmed during our experiments. The BERT based system covers intent recognition. The system has shown potential results as opposed to training processing. The average loss chart continues to show the gradual decline of the loss values from the initial 3.2160 to the final 3.0752 during training. This demonstrates that the model receives training, and classifies user intention from Facebook comments after repeated cycles of training, very accurately. Consistent improvement in preciseness suggests that the manner performs well at optimizing parameters and increases with time. Through the means of the training loss trend analysis, we can check the model's architecture and training settings for their accuracy thus obtaining a better performance for the model in the intent recognition process.

Result

Epoch	Avg. Training Loss	Epoch	Avg. Training Loss
1	3.2160	2	3.1336
3	3.0643	4	3.1033
5	3.1618	6	3.1341
7	3.0805	8	3.1238
9	3.1468	10	3.1189
11	2.9716	12	3.1117
13	3.1765	14	3.1466
15	3.1269	16	3.0879
17	3.1142	18	3.0828
19	3.0661	20	3.0752

Fig 4.) Epochs & Avg. Training Loss

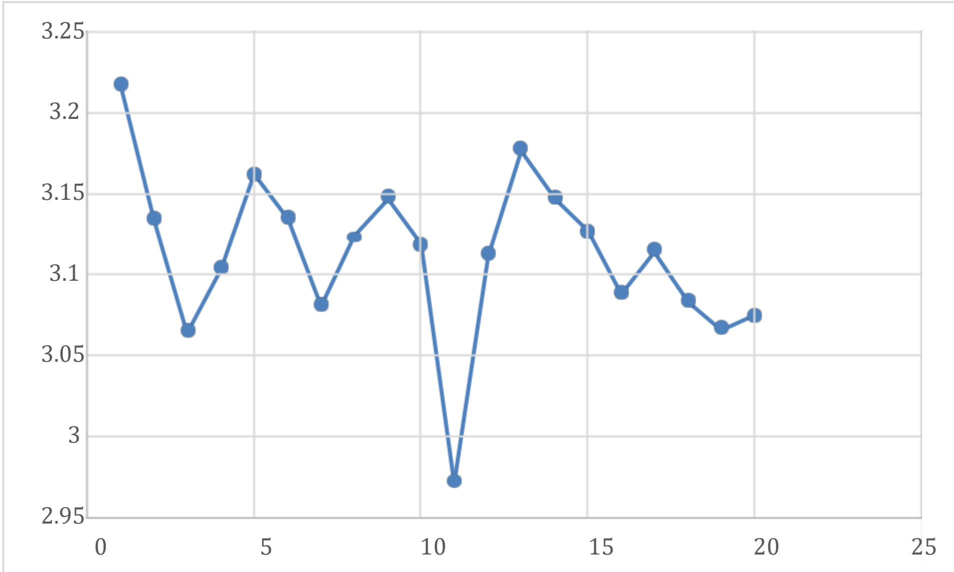


Fig 5.) Average Training Loss

5. Conclusion

The BERT and NLP technology which have been developed can be used to understand what people are trying to among others on social networks. This way of collecting data enables us to track user behavior online, and we can take more informed decisions, thus, delivering better internet experience eventually. It exactly puts the people's intentions in categories and make use of this fact to find the hidden pictures and guide in the development of online interactions.

For the AI system also manages to improve social media interactions, but beyond that it shows much effect in other fields such as health care, education, and government. It is introduced a brand new topic which can make us exchange our behaviors of interacting and decision-making. This system was created to discern the intentions of people and through BERT, it can definitely bring change to our online activities and shopping That it shows how multifaceted and influential AI can be in determining ways people relate one on one and can affect the shape of the future internet.

6. References

[1] Karimi, A., Rossi, L., & Prati, A. (Year). "Improving BERT Performance for Aspect-Based Sentiment Analysis." University of Parma, Department of Engineering and Architecture.

[2] Win, Y. (Year). "Classification using Support Vector Machine to Detect Cyberbullying in Social Media for Myanmar Language." In Proceedings of the 2019 IEEE International Conference on Consumer Electronics - Asia (ICCE- Asia).

[3] Devlin, J., Chang, M. W., Lee, K., & Toutanova, (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL).

[4] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (NeurIPS).

[5] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Agarwal, S. (2020). Language models are few-shot learners. In Advances in neural information processing systems (NeurIPS).

[6] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Kudlur, M. (2016). TensorFlow: A system for large-scale machine learning. In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI).

[7] Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP).

[8] Hugging Face (2022). Transformers Library Documentation. https://huggingface.co/transformers/

[9] Wijeratne, S., Balasuriya, L., Sheth, A., Doran, D., & Welles, B. (2017). Sarcasm detection on Twitter: A behavioral modeling approach. Decision Support Systems, 95, 93-101.

[10] TY - BOOK AU - Meem, Raisa AU - Hasan, Khandaker PY - 2023/12/12 SP - T1 - Improving Sentiment Analysis in Online Course Reviews with BERT and Transformer Attention Mechanism VL - DO - 10.21203/rs.3.rs-3741963/v1 ER.