Natural Language Processing
Product Review Generation Using Generative Adversarial Network Based On Sentiments

26 June, 2021

1 Introduction

Roman Urdu has emerged as a popular language on online platforms especially in the Indo-Pak region. It has become a prevalent language for different forms of communication like reviews, opinions, discussions and suggestions; especially in E-commerce industry. There are several limitations linked to the use of this language in the given scenario, which are stated below.

1.1 Insufficiency of Consumer Reviews

User reviews and opinions of a certain product greatly impact how the product is perceived, while shopping online. According to [1], among young adults, a single positive or negative review overrides the significance of product features while making purchase decision. Large number of reviews either positive or negative act as a representative of user trust towards the brand [2]. They not only affect the consumer behavior towards the product but also help in boosting the business of the seller. The problem occurs when users do not give their feedback on the product due to different reasons. People usually shy away or do not bother to leave an opinion of the product. In such cases even good products are under shadowed. For this purpose review generation systems can be brought into play.

1.2 Roman Urdu a Resource Lacking Language

Roman Urdu is a resource lacking language. There are near to null data sets and techniques available for this language in the domain of Natural Language Processing. Thus generative modelling for this language is a challenging task.

1.3 Ineffective Review Generation

Realistic review generation in any language is altogether a difficult task. There are very few existing systems that generate product reviews in English language and the quality of their results is not impressive. These results can be easily distinguished as fake. Figure 1 taken from [3] is a running example of such system. It is provided certain product as seed value around which all resultant reviews are generated. As the figure shows the generated reviews are clearly distinguishable as unrealistic reviews and some of them are not even related to the product. In case of Roman Urdu there are no such systems that can generate realistic reviews for products.

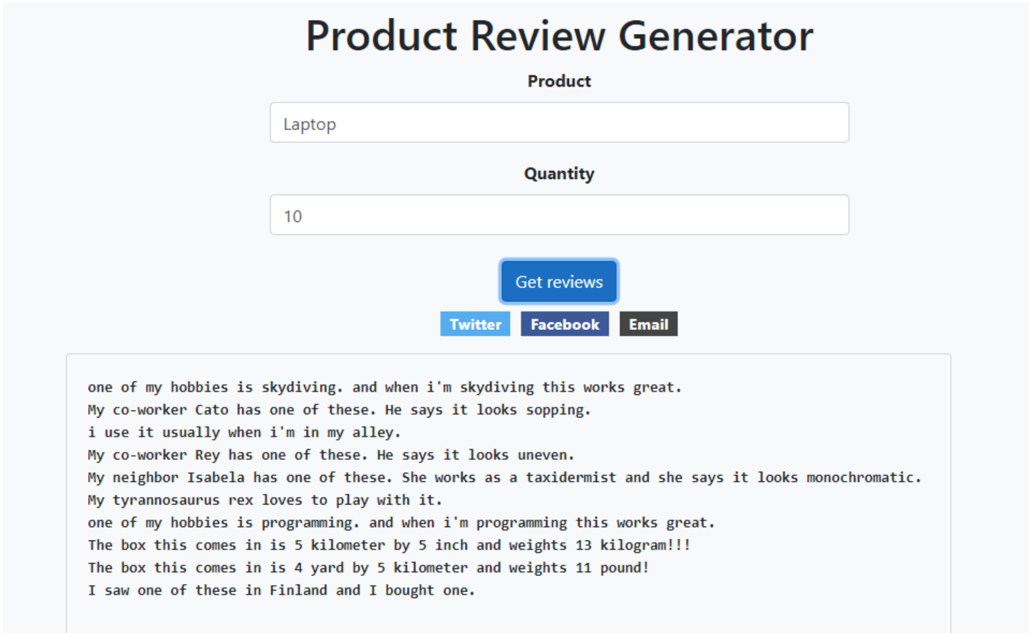


Figure 1: This is an image of existing review generating system.

2 Related Work

Research on Natural Language Generation (NLG) has shown extraordinary growth in recent years. NLG as we know of now is divided into three major approaches: traditional approach, statistical approach and deep learning approach [4]. Traditional approaches are mainly rule based which involve encoding human knowledge into sets of rules.[5] implements a rule based text generation system which generates product descriptions and journal articles in multiple languages. Rule based systems often have low precision and high recall, this means they can have high performance in specific use cases, but usually suffer performance degradation when generalized. Apart from that rule generation is a very time intensive and manual effort driven process.

In statistical approach instead of creating data that is missing in textual discourse, the idea is to leverage large amounts of already existing data for different problems. The crux of this approach is that if we can find enough number of examples of a given problem, we can potentially solve it using the power of induction. Statistical NLG makes heavy use of Machine Learning for developing models and deriving insights from labeled text. The drawback is that these approaches are compute intensive.

Generative models in NLG utilizing the power of deep learning has shown results that have outperformed those achieved by statistical approaches [6]. In the domain of deep generative models, the concept of distributed representation of input is of fundamental importance. Word2Vec is one of such approach. It uses feed-forward fully connected ANN architecture [7]. The limitation of Word2Vec is that it only considers the local context of language. Recurrent Neural Network has also been tried on NL tasks. It has internal memory due to which it can remember previous results as well, they are used in sequential data generation models. [8] uses RNN for generating text word by word. The limitation of RNN is that they are difficult to train optimally, they also face problem of vanishing and exploding gradients thus they have not been pursued further for NL tasks. LSTM is also a variation of RNN without the hidden states. The benefit of LSTM is that it only learns to keep relevant information to make predictions and forgets all the irrelevant information , they also do not face the problem of vanishing and exploding gradients like that in RNNs.[9] uses LSTM for context based text generation.[10] [11] also use LSTM. The major drawback of LSTM is that it can take ages to train it. The above mentioned models usually rely on labelled data, but most of the time labelled data is not available. Variation Auto Encoder [**kingma2013auto**] is one of the generative models that works on unlabeled data. In recent years Generative Adversarial Network has emerged as a state of the art approach to deep generative models.

3 Design Methodology

Online reviews have become the most important aspect for sellers these days. Users rely heavily on these opinions to make decisions about which product to purchase and from which vendor. Therefore it has become imperative for businesses to receive and maintain positive reviews from and avoid negative reviews at all cost as it could potentially damage their business.

This has led many companies to acquire systems that can generate reviews on behalf of customers regarding their products. However these techniques are still not good enough to convince humans into believing they are real reviews and can easily be detected as not legitimate. This is because the system lacks finesse and fails to generate reviews that closely resemble human language. Which is why we have proposed a system which uses real human reviews in order to understand the language and then generate similar reviews based on that. We also target sentiment in order to ensure that the proposed system generates only positive reviews by training the system first to understand the difference between the two.

There are different methods of detecting fake reviews online, one of which is a language-based method which is related to linguistic characteristics, it focuses on the features of the language that is mostly used in these fake reviews. In order to get past such detection techniques the system needs to be robust enough to fine tune its language by copying real reviews. The proposed system uses GPT-2 to generate meaningful and fluent sentences that closely resemble the original reviews sentiments and then uses RUBERT [12] to filter out the reviews that show undesired sentiments.

4 Proposed Work

The proposed system follows three step process to generate reviews, first is preprocessing the review text, second is review discrimination and third is review generation.

4.1 Data Collection

We scrape the data from *Daraz.pk* with the help of crawler.

4.2 Data Preprocessing

The preprocessing step is further divided into three phases:

1. **Removal of Non Meaningful Sentences**

   In this phase we filter out text that does not contain product name as they are vague and do not provide much information about the product. For example we have two different reviews for the same product the decision to retain or discard the review will be as follows:

User Review	Decision
yeh earphones bohat ache quality kay hai	review has explicit mention of product, retain it
yeh bohat acha hai	review does not mention the product, discard it

2. **Canonicalization**

   In this phase we remove all digits and punctuations from text.

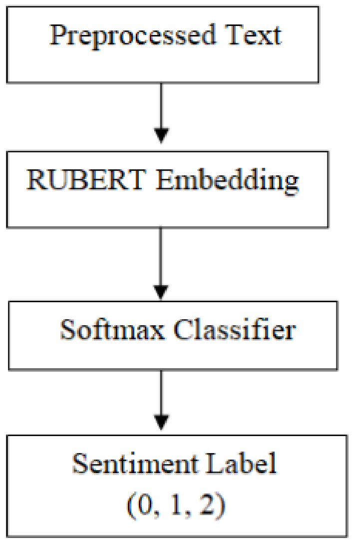
3. **Add Special Tokens**

   In the final phase of preprocessing we add [CLS] and [SEP] tokens at appropriate positions.

Performing preprocessing in this way ensures the system is trained on reviews that are meaningful in terms of information which in turn will help to generate reviews that deliver that same context of information.

4.3 Review Discriminator in GAN

In this step we train RUBERT [12] using softmax classification. We feed RUBERT the preprocessed text and it calculates the probabilities of input which belong to different class labels related to sentiments (positive or negative). These probabilities sum up to 1.



The softmax function is defined as:

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{i=1}^{K} e^{z_i}} \quad \text{for } i = 1, \ldots.., K$$

Where z is the output of the softmax layer. The output node with highest probability is selected as a predicted label.

4.4 Review Generator in GAN

In this step we take this preprocessed review and feed it to the GPT-2 model which generates corresponding fake reviews. GPT-2 generates these reviews using probability distribution of the corpus for example if input tokens are given in sequence of x = (x1,. . . .xn) then the probability of this input can be calculated as:

$$P(x) = \prod_{i=1}^{n} P(x_i | x_1, \ldots\ldots, x_{n-1})$$

This factorized probability of sequence is calculated by using conditional probability of each token, provided number of n-context tokens in neural networks. The cost function maximizes the sum of log of conditional probabilities over these sequences of tokens:

$$\Theta^* = \underset{\Theta}{\operatorname{argmax}} \sum \log P\,(x_i | x_1, \ldots\ldots, x_{n-1}; \Theta)$$

Where θ is the neural network parameters and learnt using feed forward neural network, vanilla RNN and transformer architecture.

Fig 2 is the bigger picture of the model.

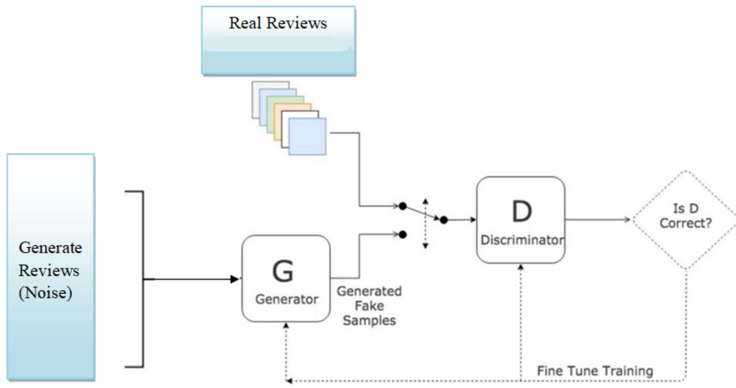


Figure 2: This is the overall architecture of the model.

5 Evaluation Methodology

We use two techniques to evaluate our results. First one is human inspection, in which real and model generated reviews are given to testers. They label the reviews(machine generated,real).These labels are checked against the actual. This gives us the quality of generated the reviews. If the reviews are good enough to be passed as a human entered review then the performance of our model is good.

We also use detector softwares, that detect machine generated reviews. If the reviews are undetectable then the model has good performance.

References

[1] Bettina von Helversen et al. "Influence of consumer reviews on online purchasing decisions in older and younger adults". In: *Decision Support Systems* 113 (2018), pp. 1–10.

[2] Jinghuan Zhang, Wenfeng Zheng, and Shan Wang. "The study of the effect of online review on purchase behavior". In: *International Journal of Crowd Science* 4 (2020), pp. 73–86.

[3] *Product Review Generator*. URL: `https://randommer.io/Text/Review`.

[4] Sashank Santhanam and Samira Shaikh. *A Survey of Natural Language Generation Techniques with a Focus on Dialogue Systems - Past, Present and Future Directions*. 2019.

[5] Armin Bauer, Niki Hoedoro, and Adela Schneider. "Rule-based Approach to Text Generation in Natural Language - Automated Text Markup Language (ATML3)". In: *Proceedings of the RuleML 2015 Challenge, the Special Track on Rule-based Recommender Systems for the Web of Data, the Special Industry Track and the RuleML 2015 Doctoral Consortium hosted by the 9th International Web Rule Symposium (RuleML 2015), Berlin, Germany, August 2-5, 2015*. Vol. 1417. CEUR Workshop Proceedings. 2015.

[6] Touseef Iqbal and Shaima Qureshi. "The survey: Text generation models in deep learning". In: *Journal of King Saud University - Computer and Information Sciences* (2020).

[7] Tomas Mikolov et al. *Efficient Estimation of Word Representations in Vector Space*. 2013.

[8] Ilya Sutskever, James Martens, and Geoffrey Hinton. "Generating Text with Recurrent Neural Networks". In: 2011.

[9] Sivasurya Santhanam. "Context based Text-generation using LSTM networks". In: *CoRR* abs/2005.00048 (2020).

[10] Xingyuan Chen et al. *Adversarial Sub-sequence for Text Generation*. May 2019.

[11] Wenlin Wang et al. *Topic-Guided Variational Autoencoders for Text Generation*. 2019.

[12] Usama Khalid, Mirza Omer Beg, and Muhammad Umair Arshad. *RUBERT: A Bilingual Roman Urdu BERT Using Cross Lingual Transfer Learning*. 2021.