

Binary Classification of Physician's Online Reviews Using Deep Learning Techniques

Vanitha M

PG Scholar,

Dept. of Information Technology,
PSG College of Technology, Coimbatore, Tamil Nadu,
India.

Raja Mohana SP,

Assistant Professor,

Dept. of Information Technology,
PSG College of Technology, Coimbatore, Tamil Nadu,
India.

Abstract: - Online health forums are continuously becoming a popular platform for people to search health-related information. Patients are searching information not only about diseases but also about physicians and hospitals. Online health related search queries deal with medical related information, like performance reviews and ratings of physicians and hospitals. Physician reviews and ratings are qualitative feedback given by patients to their medical consultants. The purpose of this work is to find the type of review, positive or negative by which we can consecutively find the quality of services given by medical consultants to their patient. The review in online forums are considered and sentiments are analyzed to determine the type of review. The review data have been collected from RateMDs.com. Existing method [1] uses Convolution Neural Network with Convolutional Layer (CONV), ReLu Layer (RELU), Pooling Layer (POOL), and Fully-Connected Layer (FC) which uses the publicly available word2vec word vectors that were trained on 100 billion words from Google News where each word is represented as a 300-dimensional vector to train their model. The proposed with small changes, a method without pre-training word2vec using Convolution Neural Network with Embedding Layer, followed by a Convolutional, Max-pooling and Soft-Max Layer for classification improves the classification performance and accuracy.

INTRODUCTION:

People doesn't use online forums only to give review about the product they by online but also about the services they get. Now this is now been to medical departments as well, people are discussing about their health issues, drugs, services provided by doctors and hospitals. In a survey Fox and Duggan 2013, they found that 72% of internet user looks for health related information online. Also, based on Ellimoottil et al. 2012, survey with random 500 urologists, 80% of them have their reviews in online forums. Consumer's feedback is so important to any industry, but in this field a simple feedback or a single review plays a major role in their choice. There are multiple online websites to share the opinion such as RateMDS, Vitals and HealthGrades. Even the hospitals have feedback system in their own websites to know their positives and negatives, so they can improve and keep providing better services. The proposed system is to mine valuable information from reviews to understand a user's preferences. The overall objective of this system is to improve the accuracy and performance of review classification. Ranti's model [1] yields 93% for positive/negative classification using Convolution Neural Network with Convolutional Layer (CONV), ReLu Layer (RELU), Pooling

Layer (POOL), and Fully-Connected Layer (FC) which uses the publicly available word2vec word vectors that were trained on 100 billion words from Google News where each word is represented as a 300-dimensional vector to train their model. The proposed method without pre-training word2vec using Convolution Neural Network with Embedding Layer, followed by a Convolutional, Max-pooling and Soft-Max Layer for classification improves the classification performance and accuracy is increased to 97%.

Related Work:

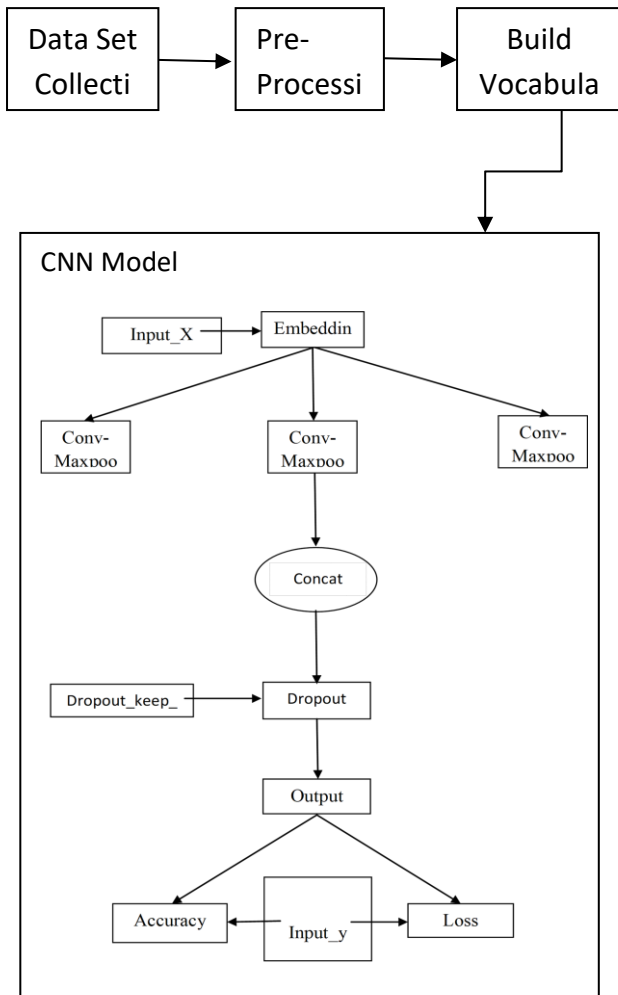
Paul et al. [5]. This system analyzes the user's satisfaction and dissatisfaction based on the reviews of doctors given by patients and predict their ratings. The novel method with convolutional neural network for optimizing functions got 93% accuracy rate for positive/negative binary classification of patient reviews and they got a mean absolute error as 0.525 for predicting rating out of 5 with error rate 0.71.

Sharma et al. [1], This system analyzes the patient's sentiments using reviews of doctors and predict doctor's ratings based on categories such as Knowledge, Staff and Helpfulness. Convolutional neural network with pre-trained word vectors using Adadelta optimizer, categorical cross-entropy for loss function and dropout layer to regularize. Here, they have increased the performance of Paul et al. [5] and got better result for both binary and 5-point rating classification problems.

Monett et al. [8] The models mainly based on ratings from customer reviews of mobile apps that are collected from Google Play Store. The model makes predictions based on the review polarity. Predicting ratings is important for sentiment analysis because it can better be understood how customers choose products. Text based rating prediction is performed better even when only phrase-level sentiment classification is available.

Lopez et al. in [12], This model rate physicians based on patient reviews by qualitative content analysis of 712 online reviews from two different rating websites. They sampled reviews of 445 primary care doctors from four different U.S. urban locations. They found that most of internet reviews of primary care physicians are positive. Their findings confirmed that the patient's satisfaction is also based on staff, access, and convenience. In addition, negative interpersonal

reviews underscore the importance of well-perceived bedside manner for a successful patient–physician relationship. Proposed Method:



Data Collection: The doctor review dataset is collected from the website RateMDs.com It consists of 54017 individual reviews which are 38847 positive and 15170 negative reviews. Example of review: “DR Carter is an outstanding doctor. He’s an example of an utmost professional...doing his best to treat see as many patients as possible. He doesn’t manage a social club or a therapist/counselor!!! LoL”, “Although she was a nice person, I wouldn’t trust her to handle medical issues.”

Pre-processing: Data-collection is complex process and it’s not controlled, which results in invalid or non-useful data like out-of-range values, impossible data combinations, missing values, etc. Using data which are not cleaned can lead to wrong results. As noisy data keeps on increasing, training will be more difficult and inaccurate. Thus, quality of data is more important.

Build vocabulary: The training dataset consists of two csv file, positive.pos and negative.neg which has positive negative reviews respectively. It is converted from text input column to numeric value. There are 54,017 unique reviews in the dataset with 38847 positive reviews and 15170 negative

reviews. Vocabulary is built using built-in function `tflearn.data_utils.VocabularyProcessor(max_document_length,min_frequency,vocabulary, tokenizer_fn)` which maps each word to an integer. Each sentence becomes a vector of integers.

a	1
caring	2
doctor	3
who	4
cares	5
about	6
his	7
patients	8
he	9
has	10
great	11
christian	12
man	13

[[1	2	3	.	.	.	,	0	0	0]	
[19	2	21	.	.	.	,	0	0	0]	
[53	54	3	.	.	.	,	0	0	0]	
.	.	,	.	.	.	,	0	0	0]	
[9	28	58	.	.	.	,	0	0	0]	
[132	7	41	109	.	.	.	,	0	0	0]

The Model:

The first layers embed words into low-dimensional vectors. The second layer does convolutions over the embedded word vectors using multiple filter sizes, sliding over 3, 4 and 5 words for each iteration. Next, max-pool the output of the convolutional layer to long feature vector, then dropout regularization, and classification of result using a softmax layer.

Embedding Layer: First layer is embedding layer, it maps vocabulary word index into low-dimensional vector representation.

`tf.nn.embedding_lookup(none,sequence_length, embedding_size,1)` is used to build vector. TensorFlow’s convolutional `conv2d` function uses a 4-dimensional tensor as input which are batch, width, height and channel. The result of embedding layer doesn’t have the channel dimension, so add it manually, leaving with a layer of shape.

Convolution and Max-Pooling Layers: The filters of different sizes are used. Convolution creates tensors of different shapes and iterated through them, create a layer for each of them, and then the result is merged to get single feature vector.

Each filter slides over the entire embedding, but differs in number of words it covers. Narrow convolution gives an output as [1, sequenceLength - filterSize + 1, 1, 1]. Performing max-pooling over the output of a specific filter size leaves us with a tensor of shape [batchSize, 1, 1, num_filters]. The feature vector, in which the last dimension corresponds to our features. The pooled output tensors from every filter size is combined to single long feature vector of shape [batch Size, num Filters Total]. Using -1 in tf.reshape tells TensorFlow to flatten the dimension when possible.

Dropout Layer: CNN is regularized using dropout layer. It disables some neurons which avoid them from learning particular feature. This is set to 0.5 during training, and to 1 during evaluation to disable the dropout.

Loss and Accuracy: The loss is used to measure the error in network created, and it should be minimized. The standard loss function, cross-entropy loss is used. tf.nn.softmax_cross_entropy_with_logits is used to calculate loss. Cross entropy is given by,

$$H(p, q) = - \sum_x p(x) \log q(x).$$

Scores and Predictions:

Using the feature vector from max-pooling (with dropout applied) can generate predictions by doing a matrix multiplication and picking the class with the highest score. Softmax function to convert raw scores into normalized probabilities, but that wouldn't change the final predictions.

Result Analysis:

This section, briefs about the results obtained using Convolution Neural Network with Embedding Layer, followed by a Convolutional layer, Max-pooling layer and Softmax Layer to classify doctor reviews is applied to doctor review dataset collected from ratemds.com.

Method	Accuracy
Paul et al. [5]	73%
Ranti et al [1]	93%
Proposed System	97%

The proposed CNN model provides 97% accuracy in classifying users reviews into positive and negative which is higher than accuracy reported in other CNN model given by Paul et al. [5] and Ranti et al [1]. Paul et al. [5] novel method with convolutional neural network for optimizing functions got 76% accuracy rate for positive/negative binary classification of patient reviews and they got a mean absolute error as 0.525 for predicting rating out of 5 with error rate 0.71. Ranti et al. [1] system analyzes the patient's sentiments using reviews of doctors and predict ratings on three categories as Knowledge, Staff and Helpfulness. Convolutional neural network with pre-trained word vectors

using Adadelta optimizer, categorical cross-entropy for loss function and dropout layer to regularize and got better result as 93% for binary classification.

Effect of Hyper Parameters:

Training epochs -The accuracy increases as number of training epochs increases, it converges at 220. Default dimension - 128 is used for character embedding. Window size - As window size increases, the training takes longer time and no significant increase in accuracy is found, thus, used multiple window sizes, 3,4,5 and 128 filters each.

Conclusion and Future Work:

Thus, the proposed system using Convolution Neural Network with Embedding Layer, followed by a Convolutional, Max-pooling and Softmax Layer to classify doctor reviews is applied to doctor review dataset collected from ratemds.com. The result obtained using CNN method shows accuracy increase when compared to existing methods. The proposed method does the binary classification as positive, negative review for doctor dataset. For future work, the system can be extended to predict five point ratings.

REFERENCES:

- [1] Ranti D. Sharma, Samarth Tripathi, Sunil K. Sahu, Sudhanshu Mittal, and AshishAnand, "Predicting Online Doctor Ratings from User Reviews Using Convolution Neural Networks" *International Journal of Machine Learning and Computing*, Vol. 6, No. 2, April 2016.
- [2] HemaSagarVakati, Jebakumar R, "Predicting Ratings for User Reviews and Opinion Mining Analyze For Physicians Hospitals" *Asian Journal of Pharmaceutical and Clinical Research*, Vol 10, Issue 3, 2017
- [3] Alodadi, Nujood, and Lina Zhou, "Predicting the Helpfulness of Online Physician Reviews." *Healthcare Informatics (ICHI)*, 2016 *IEEE International Conference on. IEEE*, 2016.
- [4] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. J Mach "Dropout: A simple way to prevent neural networks from over fitting." *Learn Res* 2014;15(1):1929-58.
- [5] Paul MJ, Wallace BC, Dredze M. "What Affects Patient (Dis)satisfaction? Analyzing Online Doctor Ratings with a Joint Topic-Sentiment Model." Copyright ©2013.
- [6] Ahire SB, Khanuja HK. "A personalized framework for healthcare recommendation". *Int J ComputAppl* 2015;110(1):89-92.
- [7] M. M. Galizzi, M. Miraldo, C. Stavropoulou, M. Desai, W. Jayatunga, M. Joshi, and S. Parikh, *BMJ* "Who is more likely to use doctor-rating websites, and why? A cross-sectional study in London," *Open*, vol. 2, no. 6, e001493, 2012.
- [8] Monett D, StolteH. "Predicting star ratings based on annotated reviews of mobile Apps" by. [dx.doi.org/10.15439/2016F141](https://doi.org/10.15439/2016F141) 24 October 2016.
- [9] B. Kadry, L. F. Chu, B. Kadry, D. Gammas, and A. Macario, "Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating," *Journal of Medical Internet Research*, vol. 13, no. 4, 2011.
- [10] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. International Conference on Artificial Intelligence and Statistics*, pp. 315–323
- [11] Kim, Yoon. (2014). "Convolutional neural networks for sentence classification" (2014).
- [12] Xiang Ma, Xiaojiang Lei, Guoshuai Zhao, Xueming Qian, "Rating prediction by exploring user's preference and sentiment", *Multimedia Tools and Applications*, pp. , 2017, ISSN 1380-7501.
- [13] X. Lei, X. Qian, G. Zhao, "Rating prediction based on social sentiment from textual reviews", *IEEE Trans. Multimedia*, vol. 18, no. 9, pp. 1910-1921, Sep. 2016.