دانشگاه علامه طباطبائی

Classification of Persian Product Reviews Using Neural Networks

The second Conference On Artificial Intelligence and Soft Computing in the Humanities

Feb 17th, 2022

Karim Akhavan Azari — Languages and Linguistics Center, Sharif University of Technology, Tehran, Iran. karim.akhavan@sharif.edu
Dr. Mohammad Bahrani — Department of Computer Science, Allameh Tabataba'i University, Tehran, Iran. bahrani@atu.ac.ir

List of Contents

Different Sentiments in Reviews

کالا در این رنج قیمت
محصولات نوا عالی
هستن حتما بخرید

Recommended

پلاستیکی هستش ولی
با توجه به قیمتش
شاید بیارزه

Neutral

نداشتن درپوش
پلاستیکی

Not Recommended

One Type of Challenge in Sentiment Analysis

⁉️

رنگش قرمز بود!

Neutral

⁉️

سایزش مثل عکس
نبود

Neutral

⁉️

اصن اینقدر خوب بود
که همون تو جعبه
میموند بهتر بود :)

Neutral

Types of Sentiment Analysis

- Document level
  - متاسفانه علیرغم جنس خوب ، سایز رویه بیرونی ۹ سانته که با مشخصات فنی موجود مطابقت نداره. خوب تو سینک فیت نمیشه و عملا بلا استفاده ست. بدرد سینکهای قدیمی میخوره. ارزش برگشت زدن هم نداره و باز هم آدم به این نتیجه میرسه که محصولی که فروشنده غیر خود دیجیکالا باشه خرید نشه
- Sentence level
  - خیلی خوشگله حتما همه بخرید
- Entity & Aspect level
  - صدای قابل قبولی دارد

Motivation Behind the Work

- Helps businesses to monitor product sentiment in customer feedback, and understand customer needs
    - More than 58 million internet users in Iran and around 46% of them (more than 26 million) make online transactions with many of them being online shopping[1]
- Helps customers in deciding on the best product to buy
    - Product ratings and comments impact the buyers' opinion on the shopping item[2]
- Analyzing Data at Scale
- Real-Time Analysis
    - Can be applied to conversations with customers, brand/product mentions (e.g. twitter), comments (e.g. Instagram, YouTube), etc.

1. DataReportal (2020), "DIGITAL 2020: IRAN", retrieved from https://datareportal.com/ reports/digital-2020-iran
2. von Helversen, Bettina, et al. "Influence of Consumer Reviews on Online Purchasing Decisions in Older and Younger Adults." Decision Support Systems, vol. 113, Sept. 2018, pp. 1–10. ScienceDirect, https://doi.org/10.1016/j.dss.2018.05.006.

Literature Review

- Machine learning algorithms such as Naive Bayes, Logistic Regression, Tree Classifiers, SVM
- Turney et al. (2003)
  - Classify words based on their polarity with an accuracy of 82.84% using LDA
- Catal et al. (2017)
  - Ensemble learning method with the combination of two SVM variants and a Naive Bayes algorithm. Classifies by majority voting, which on the average achieved an accuracy of 83.25% on three different datasets
- Abdi et al. (2019)
  - Deep learning combined feature vectors in terms of statistical, linguistic and sentiment knowledge, sentiment shifter rules, and word-embedding. Saw an improvement over other deep-learning-based sentiment classification researches
- (W ei, Y ., Lao, While clustering algorithms (Rehioui, H., & Idrissi, A., 2020; Mostafa, M.M., 2019)
  - Combination of clustering techniques which resulted in better clustering
- Roshanfekr et al. (2017)
  - Compares NBSVM with two a Bidirectional-LSTM and a Convolutional Neural Network. Although the NBSVM outperforms the other two (70% compared to 54.2 and 59.1 for BiLSTM and CNN in order), both of the deep learning methods had an overall much better performance considering their recall and F-score
- Zobeidi et al. (2019)
  - Classifies by feature extraction using a CNN and using a BiLSTM network on the Digikala Persian dataset on mobile and digital cameras resulting in an accuracy of 95% for two classes and 92% for multi-class classification

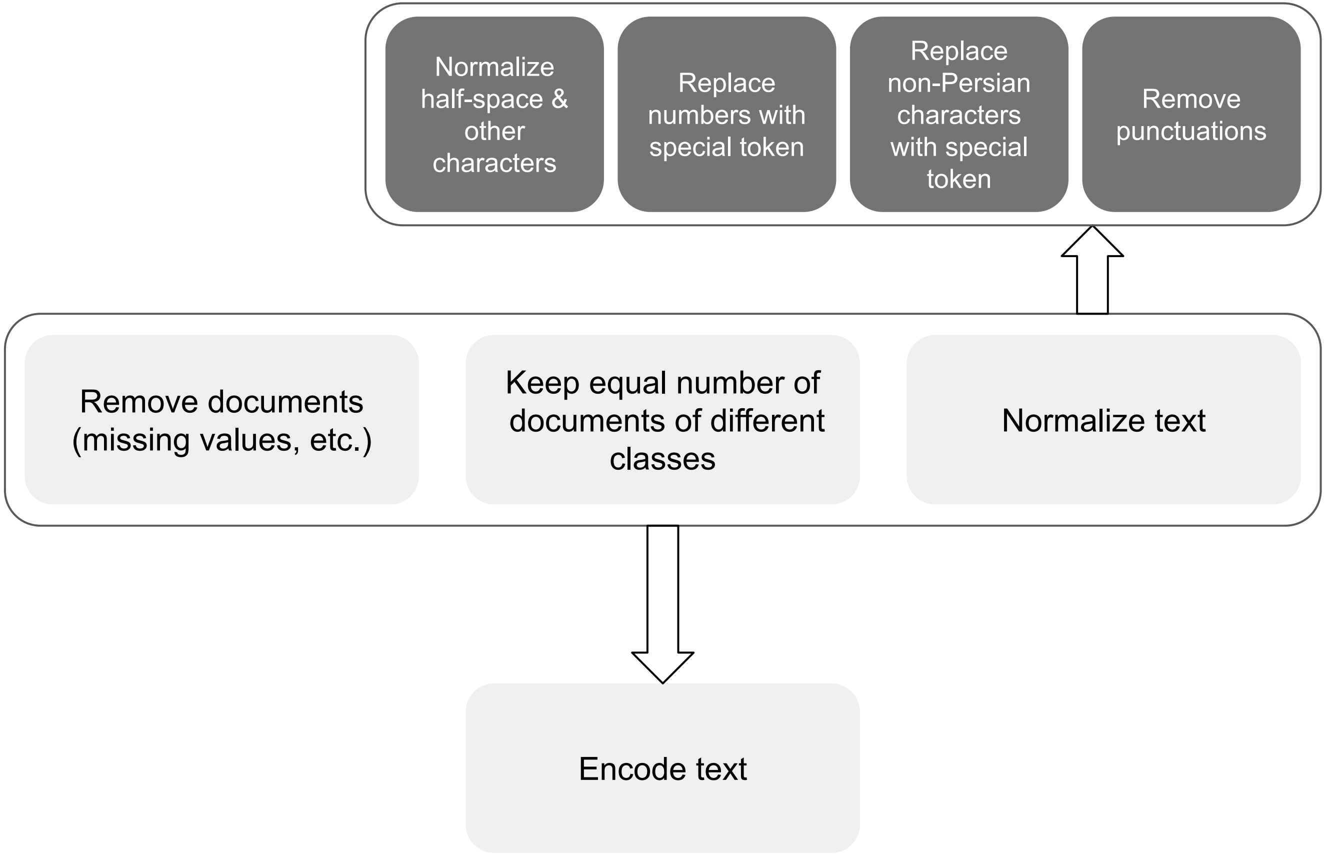
Dataset[1]

Size of dataset: 100000 documents
Labels: “recommended”, “Not_recommended”, or “no_idea”
Size of dataset after preprocessing: 62131 documents
Train/evaluation ratio: 90% / 10%

product_id	product_title	recommend	comment
1	ساعت دیواری آرام مدل هنرکار SD1001	recommended	خیلی ساعت قشنگیه تنها ایرادی که داره اینه که رو دیوار مسطح وای نمی ایسته...
2	اتو مو مک استایلر مدل MC-2017	not_recommended	دختر عمه ی من ارایشگره و این اتو رو خریده من بیار استفاده کردم عالی بود...
3	ساعت مچی عقربه ای مردانه کاسیو جی-شاک GA-100-1A1DR	recommended	خیلی کاربردی و خوبه.
4	دوربین دوچشمی سلسترون مدل 12x25 Focus View	recommended	اگر در پیشنهاد ویژه خریداری شود به نسبت قیمت، دوربین مناسبیست....

1. https://www.digikala.com/opendata/

Pipeline: Preprocessing



Pipeline: Network Architecture

Three different models:

- RNN (baseline)
- LSTM
- BiLSTM

Pipeline: Network Architecture

The embedding layer:

- Feed 1000 selected features as the input to this layer
- Project the input features into a 300 dimension layer

The hidden layer:

Each has a layer with 128 neurons

- RNN
- LSTM
  - Has another second layer, similar to the first one
  - The first layer returns a dense representation of input for each time step
- BiLSTM

Dense layer:

- Three neurons (3 classes)
- Softmax

Pipeline: Network Architecture

The embedding layer:

- Feed 1000 selected features as the input to this layer
- Project the input features into a 300 dimension layer

The hidden layer:

Each has a layer with 128 neurons

- RNN
- LSTM
    - Has another second layer, similar to the first one
    - The first layer returns a dense representation of input for each time step
- BiLSTM

Dense layer:

- Three neurons (3 classes)
- Softmax

Pipeline: Network Architecture

The embedding layer:

- Feed 1000 selected features as the input to this layer
- Project the input features into a 300 dimension layer

The hidden layer:

Each has a layer with 128 neurons

- RNN
- LSTM
  - Has another second layer, similar to the first one
  - The first layer returns a dense representation of input for each time step
- BiLSTM

Dense layer:

- Three neurons (3 classes)
- Softmax

Pipeline: Network Architecture

Config:

- Batch size: 64
- LR: 0.002
- Dropout: 0.4
- Optimizer: Adam
- Loss function: Categorical cross-entropy

Evaluation

Performance of each model:

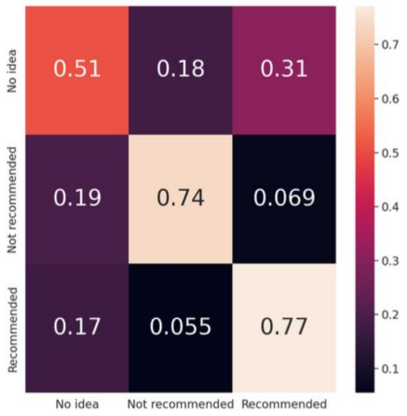
Model	Accuracy[%]	Precision[%]	Recall[%]	F1 Score[%]
RNN*	66.02	76.66	72.4	74.03
LSTM*	68.86	78.97	73.7	75.53
BiLSTM*	67.1	78.67	76.1	77.15
NBSVM-bi[1]	-	70.7	31.9	44.0
BiLSTM[1]	-	54.2	52.2	53.2
CNN[1]	-	59.1	52.2	55.4

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i} \qquad \text{Recall}_i = \frac{TP_i}{TP_i + FN_i}$$

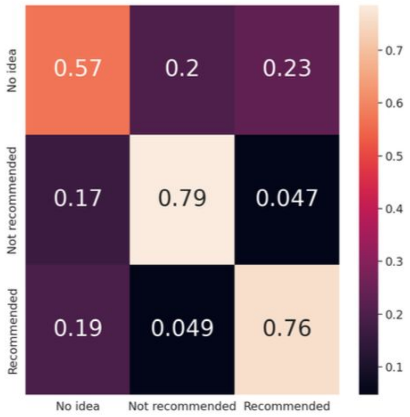
1. Roshanfekr, B., Khadivi, S., & Rahmati, M. (2017). Sentiment analysis using deep learning on Persian texts. 2017 Iranian Conference on Electrical Engineering (ICEE), 1503-1508.

Evaluation

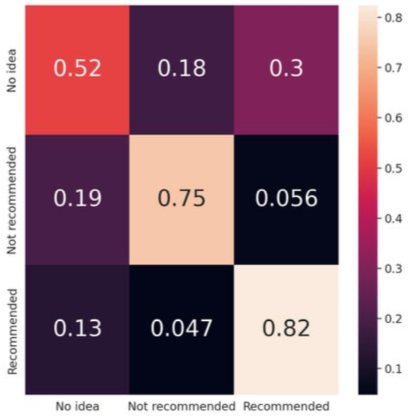
Confusion matrix of each model:



RNN Model



LSTM Model



BiLSTM Model

Conclusion & Summary

- Sentiment analysis are of great help to both businesses and users alike
- Seq2Seq networks like BiLSTM are both effective & fast to train
    - Increasing the data could help even more
- Lack of data in Persian
- Use of transformer networks such as Bert
    - Might be more costly to train
    - Might need more data

Further Research

- Working with other types of sentiment analysis
  - Graded Sentiment Analysis
  - Emotion detection
  - Aspect-based Sentiment Analysis
- Collecting larger datasets
  - Translating Yelp dataset or Amazon Reviews Corpus

References

- Abdi, A., Shamsuddin, S.M., Hasan, S., & Piran, J. (2019). Deep learning- based sentiment classification of evaluative text based on Multi-feature fusion. Inf. Process. Manag., 56, 1245-1259.
- Catal, C., & Nangir, M. (2017). A sentiment classification model based on multiple classifiers. Applied Soft Computing, 50, 135–141. https://doi.org/10.1016/j.asoc.2016.11.022
- Duyu Tang, Bing Qin, and Ting Liu. 2015b. Learning semantic representations of users and products for document level sentiment classification. In Proc. ACL.
- Long, Y., Ma, M.D., Lu, Q., Xiang, R., & Huang, C. (2018). Dual Memory Network Model for Biased Product Review Classification. WASSA@EMNLP.
- Turney, P. D., & Littman, M. L. (2003). Measuring praise and criticism: Inference of semantic orientation from association. ACM Transactions on Information Systems, 21(4), 315–346.
- Wei, Y., Lao, Y., Sato, Y., & Han, D. (2019). Product-review Classification Combining Multiple Clustering Algorithms. ICFET 2019.
- Zobeidi, S., Naderan, M., & Alavi, S.E. (2019). Opinion mining in Persian language using a hybrid feature extraction approach based on convolutional neural network. Multimedia Tools and Applications, 78, 32357 - 32378.

Thanks for Your Attention!

erfan226@sharif.edu

erfan226

erfan226.github.io