Lindsay Anne D. Peñaranda

2018-01719

1. **Introduction**
    Social media is a tool that has only increased in usage over time. The convenience of finding news, getting entertainment, and communicating with our friends and family in a single, free-to-use website or mobile application is not something to pass on. Especially with the COVID-19 pandemic, more people have turned to social media to promote their business, to engage in discourse about anything under the sun, and so on.

    Trending is a term used to label something that is currently popular or widely discussed online, especially on social media ("Trending", n.d.). On Facebook, a specific social media site, trending topics are personalized based on the user's location and social behavior, as well as what's popular in general (Myers, 2016). For Twitter, a trending hashtag is determined by sharp spikes of discourse on the website rather than gradual, sustained growth.

    This paper aims to predict whether a Facebook post will trend or not. Trending is defined if there are at least 100 comments within 24 hours of posting. The neural network will be helpful for people or small businesses aiming to boost their products with the use of social media.
    a. **Inputs**
        The input variables used in this project are the following:
        i.   Page likes – the number of likes of a given page
        ii. Page check-ins – the number of individuals that visited a given page
        iii. Page talking about – the daily interest of individuals to the page
        iv. Page category – the category of the page
        v.   C1 – the total number of comments before the given base time
        vi. C2 – the number of comments in the last 24 hours, relative to base time
        vii. Base time – selected time in order to simulate the scenario
        viii. Post length – the number of characters in a post
        ix. Post share count – the number of shares in a post
        x.   Post promotion – a binary value determining if the post is promoted
        xi. Day of post publication – binary encoded features that represent the day on which the post was published
        xii. Trending – the target variable, a binary value determining if the post trended or not
    b. **Output**
        The output variable will be binary, signifying if a post will trend (1) or not (0).

2. **Review of Related Literature**
    A study by Miao et al. (2016) makes use of a cost-effective framework with algorithms that can select a subset of representative users of social media, and then monitor these users to detect the overall trending topics and predict their future popularity in the social media site. The researchers collected 1.6 million microposts in Weibo, a social media site, and the proposed system detected 92% of the trending topics that are published in Weibo official trends. This study

was able to detect and predict the future trending topics of Weibo, however it does not account for predicting if an individual post will trend or not.

A study by de Oliveira and Goussevskaia (2020) focuses on how engagement on an Instagram post is correlated with the usage of topics and global topic trends in the hashtags. The researchers used a large-scale dataset comprised of over 3 million posts by over 7000 influencers that posted at least one sponsored post. The researchers found that nano and micro-influencers' post engagement is more affected by global trends compared to celebrities, yet have more potential to exploit subject-specific hashtags to boost engagement even more. This study accounts for users in Instagram, which has a completely different basis for trending posts compared to Facebook. Users in Instagram also have different usage habits and culture compared to users in Facebook.

A related study by Xie et al. (2017) focuses on predicting the popularity of mixed-media items on Tumblr using the key traits of early adopters, individual classifiers, and importance of various features. The early adopters of a post are defined as the original poster of the content and the first n-1 resharers. It was found that network features (a reblog network was constructed with Tumblr in mind) play a more important role compared to the content and temporal features of a post. Similar to de Oliveira and Goussevskaia's study (2020), this only took Tumblr into consideration, which has a different user culture compared to Facebook.

3. **Methods**
   a. **ANN**
      The ANN used for this project is a three-layer neural network – two layers for the input and output and one hidden layer. The input layer has 17 nodes, which accepts each feature as input. Each hidden layer has 13 nodes. The output layer has 1 node, which signifies if a post will potentially trend or not. The activation function used in every node is a sigmoid function.
   b. **Datasets**
      The dataset is split on 80-20 for training and validation. A random number generator is used to randomly select rows of data for the training and validation datasets.
   c. **Quality Assurance (QA)**
      From training and validation, the neural network's accuracy is computed every epoch by comparing the predicted value of the target to the true value of the target.

4. **Results**
   After training the neural network, the accuracy yielded from validation was 94% after fifteen epochs. The weights of each node are as follows:

A0,1	2.8269750e-01
A0,2	2.3136739e-01
A0,3	1.4474218e-04
A0,4	1.1624871e-01
A0,5	3.1280681e-01
A0,6	2.6946884e-01

A0,7	3.0724269e-01
A0,8	3.1538102e-01
A0,9	3.8938808e-01
A0,10	7.6813167e-01
A0,11	1.8282229e-01
A0,12	9.8515458e-02
A0,13	3.7088516e-01
A1,1	-0.05460743
B0,1	0.01272895
B0,2	0.58970004
B0,3	-1.0714889
B0,4	-1.1514671
B0,5	-1.1012862
B0,6	0.54498
B0,7	0.4507158
B0,8	-1.2711614
B0,9	-1.1824149
B0,10	0.03475037
B0,11	0.3020023
B0,12	0.2674069
B0,13	-1.0231405

After training and validation, the neural network has an accuracy of 94%.

5. References

Miao, Z., Chen, K., Fang, Y., He, J., Zhou, Y., Zhang, W., & Zha, H. (2017). Cost-effective online trending topic detection and popularity prediction in microblogging. ACM Transactions on Information Systems, 35(3), 1–36. https://doi.org/10.1145/3001833

Myers, E. (2016). What's trending on Facebook and Twitter? it could be your business. OutboundEngine. Retrieved January 5, 2022, from https://www.outboundengine.com/blog/whats-trending-on-facebook-and-twitter/

Oliveira, L. M., & Goussevskaia, O. (2020). Topic trends and user engagement on Instagram. 2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT). https://doi.org/10.1109/wiiat50758.2020.00073

Trending. (n.d.). Oxford Languages.

Xie, D., Xu, J., & Lu, T.-C. (2017). What's trending tomorrow, today: Using early adopters to discover popular posts on Tumblr. 2017 IEEE International Conference on Big Data (Big Data). https://doi.org/10.1109/bigdata.2017.8258165