Contrasting the Branding Effect of Cost Per View vs Cost Per Click Campaigns

Chad Crowe

September 26, 2021

Abstract

Existing works have predicted user behavior on social media using either image or text data. This research takes a further step with a methodology to combine image and text data for predicting user behavior on social media. The methodology is applied to 350k Facebook firm-generated content (FGC) posts. This research demonstrates that the proposed methodology that utilizes both text and image data for predicting user behavior outperforms text-only or image-only models for predicting user likes, shares, comments, and comment sentiment. The results validate the need for using both image and text data for predicting user behavior on social media.

1. Introduction

1.1 Machine Learning on Social Media Data

Existing research recognizes the value of modeling user behavior on social media using machine learning models Li et al. (2015); Straton et al. (2017); Ohsawa & Matsuo (2013); Liu (2012); Li et al. (2015). This pattern of research reflects the desire to understand consumer behavior on social media Fisher (2009). These studies have focused on particular metrics, including user click-through rates Li et al. (2015), user interaction with Facebook posts Straton et al. (2017), predicting user tendency to follow pages Ohsawa & Matsuo (2013), and user sentiment Liu (2012); Wang et al. (2015). A large amount of research is concerned with better understanding user behavior on social media.

The existing research understands user behaviors through behavior models, which often include machine learning models Li et al. (2015); Straton et al. (2017); Ohsawa & Matsuo

(2013); Liu (2012); Li et al. (2015). Examples include statistical models Li et al. (2015), neural networks Straton et al. (2017), text idf models Ohsawa & Matsuo (2013), opinion mining Liu (2012), and sentiment analysis of images Wang et al. (2015). Each example demonstrate the common methodology of modeling user behavior using machine learning models in order to better understand social analytics and user behavior. Modeling user behaviors can provide beneficial research insights about social analytics.

Convolutional Neural Networks (CNNs) perform well at working with images and are used in correspondence with social media images. They have been used for gender classi-fication Hassner & Tal (2015), for visual sentiment Segalin et al. (2017); Xu et al. (2014), to detect sarcasm on Twitter Poria et al. (2016), for detecting stress in social media images Lin et al. (2014), to perform social media profiling Segalin et al. (2017), to predict social media popularity Gelli et al. (2015), and to predict which posts will receive the post clicks Khosla et al. (2014). CNNs are the standard in the realm of social media for image analysis Hassner & Tal (2015).

Models trained on text data exist to understand user behavior. Facebook likes have been predicted for hospital data using only post text data Straton et al. (2017). Research has used NLP to predict the likelihood a user will follow a Facebook page Ohsawa & Matsuo (2013). Other models have used text data for opinion mining Liu (2012). Text data is also frequently used for modeling user behavior on social media.

1.2   Gap

However, a gap exists in current research when modeling user behavior with social media data because researchers fail to incorporate multiple data types, despite the availability of both image and text data. An example is performing sentiment analysis of posts with images Wang et al. (2015) but failing to incorporate the post's text into the model. The same is true for using text-data to predict a post's CTR but ignoring its associated image data Li et al. (2015). The gap in research is a failure to utilize multiple available data types when modeling user behavior on social media.

The gap is demonstrated in the failure for image-based models to incorporate text data. Each of the CNN models fails to incorporate text data in their models for predicting gender classification, detecting sarcasm, profiling, and predicting social media popularity Hassner & Tal (2015); Poria et al. (2016); Segalin et al. (2017); Gelli et al. (2015). Image-based social media models fail to incorporate text-data in their methodology.

1.3   Applied Model for Advertising and Forecasting User Behavior

This research applies its methodology to a use case of predicting user behavior in response to advertising. We feel this use case is relevant to the Marketing Science Journal because of it demonstrates a method for improving models of user behavior on social media in response to advertising. Such topics might be utilized to provide advertisers with a competitive advantage, or alternatively used in future research to improve marketing models on social media. This paper provides marketing science about user behavior with regard to advertisements on social media.

Advertisers are most concerned with social media metrics, especially those that promote engagement Tiago & Verissimo (2014), which include click-through rate (CTR), brand awareness, and word-of-mouth buzz. Advertisers associate these with advertisement return on investment (ROI), which is known as the Holy Grail of social media Fisher (2009). However, advertisers calculate ROI, which often includes an increase in user interaction Romero (2011); Schacht et al. (2015). This study successfully models user engagement, which is of great interest to advertisers and social media platforms.

Advertisers want to impact future sales from the untapped market Guo et al. (2020). Their goals include creating brand stickiness, improving user relationship quality, creating unique visitors, increasing average time per visit to their website, get repeated visitors, and increase visit frequency Bhat et al. (2002). There are many ways to improve advertisement campaign performance, such as influencing both its content and content type Imsa & Irwansyah (2020). However, neither of these provides a direct forecast of the advertisement's performance. Given the cost of showing ads, quicker feedback mechanisms that can predict advertisement performance is useful in curating content and publishing on the platform with a great degree of confidence concerning the advertisement's performance Hu et al. (2016). Therefore, this study is helpful in that it provides improved mechanisms for forecasting advertisement performance on social media.

Forecasting user response to advertisements is important because advertisers view social media as a method for creating both tangible and intangible firm value that improves business performance Authors (2013). Tangible benefits include a decreased time needed for users to make a buying decision Authors (2013). Intangible benefits include how advertisements influence buyer decisions Authors (2013). In addition, with better forecasting, advertisers can improve planning their sales cycles and projected revenue Imsa & Irwansyah (2020). This paper provides details on improved user behavior forecasting, which is beneficial to advertising revenue.

Social media serves as a platform where brands can create and maintain an online presence Greenwood et al. (2016). Social media can create tangible value that improves business performance Authors (2013). The desire is that tangible user engagements result in faster user conversions Authors (2013). Social media can serve as a platform for influencing their target audiences and increase their bottom-lines.

1.4 Research Summary

Our research provides a method for combining text and image data for modeling user behavior on social media. We demonstrate the successful implementation of a model combining image and text data and demonstrate its improved performance over single-data type models. The chosen method makes use of an ensemble model whose input is a text-based NN and image-based CNN. The combined model outperforms the text and image models when predicting user click, share, comment, and comment sentiment. The results of the sixteen machine learning models are provided in the results section and the discussion provides insights concerning the model's performance and its application to advertising social media data. Future social media studies should adopt the combined model methodology when modeling user behavior on social media.

The remaining paper consists of five sections. The related works will cover existing studies that model user behavior on social media and will delineate studies relying on text data or image data. We also include methodologies for processing text and image data within the related works section. The methodology section will describe the creation of the sixteen machine learning models, four text-only, four image-only, and eight combined models with different architectures. The result section outlines the result of the combined model, juxtaposed with text-only and image-only models. The discussion delineates why the combined model produces an improved performance. The conclusion and future work outline ways future research can adopt these methods to better model user behavior on social media.

2. Related Work

2.0.1 Text-based Social Media Models

Text models exist to predict user interaction on Facebook Straton et al. (2017). The predicted user metrics include page likes, shares, and comment counts from this data. The analysis categorizes all posts into engagement categories, e.g., low, medium, and high. The Neural Network trains with on the text and time data. The model can accurately predict for

lower user engagement but fails to predict for higher levels of engagement. The study's sample size was 100k posts and did not incorporate images or comment text in its predictions. Nevertheless, the study found that text data can predict limited levels of user engagement.

A CTR study focuses on predictions based on user interests Li et al. (2015). The study is essential because it models the likelihood of user behavior based on user interests with advertiser data. The study performs its prediction by modeling the Twitter feed and the click rates for each type of user interest. As a result, the study successfully predicted user click-through rates based on how well user interests coincide with the advertisement's content.

Research exists that to measure user sentiment using either text or image data. Text data is useful for opinion mining Liu (2012), where opinion mining uses keywords as sentiment indicators. Fortunately, existing sentiment lexicons are available for predicting sentence sentiment Georgiou et al. (2015). In contrast, there is research that detects image sentiment by clustering images Wang et al. (2015). Methods exist that use either text or image data to predict user sentiment. However, there are no cases of using a combination of image and text data to predict user sentiment on social media.

2.0.2 Image-based Social Media Models

Many studies use Convolutional Neural Networks (CNN) for image analysis. The use cases are varied, and include: age and gender classification Hassner & Tal (2015); image polarity Poria et al. (2016); sarcasm detection Poria et al. (2016); and image popularity classification Khosla et al. (2014). Existing research has produced visual sentiment classifiers with CNNs Segalin et al. (2017); Xu et al. (2014) to identify stress within social media images, Lin et al. (2014), use supervised CNNs to performed social profiling to identify personality traits Segalin et al. (2017), perform sentiment analyses and estimated social media popularity with CNNs Gelli et al. (2015), use images to predict which types of images are popular on social media Gelli et al. (2015), and predict which posts will receive the most clicks Khosla et al. (2014). CNNs are frequently used in combination with images on social media for understanding user behavior.

2.1 Industry need for Modeling User Behavior

Companies calculate social media revenue return on investment (ROI) via their advertisement performance on the platform Fisher (2009). Therefore, ROI is the Holy Grail of social media Fisher (2009). When asked which social media metrics marketing managers care about most, they replied with brand awareness, word-of-mouth buzz, customer satisfac-

tion, user-generated content, and web analytics Tiago & Verissimo (2014). However, ROI is difficult to track Schacht et al. (2015). Most companies are unable to get revenue or cost savings from social media Romero (2011). Instead, ROI is measured via user consumption Schacht et al. (2015). The study performed a cross-platform analysis of ROI on Facebook, Twitter, and Foursquare. Schacht proved that tweets could predict rising Foursquare check-ins.

Users visit social media sites to gain information Fisher (2009)—for example, 34% of participants post products about opinions on blogs. Moreover, traffic to blogs keeps increasing 50% alone that year, compared to 17% at CNN, MSNBC, and the New York Times. 70% of consumers visit social media sites for information. 49% of the 70% buy based on social media content. 36% of participants better rate companies with blogs. 60% of users pass along social media data to other users. Persons use social media to learn and gain opinions about products and brands.

2.1.1 Research Questions

Research_Question: Can a combination of text and image data better predict user engagement on social media using machine learning?

1. When predicting user engagement on social media, models trained using both text and image data outperform either text or image models.

The question explores the existing gap in using a machine learning architecture that digests both image and text data to predict user engagement. Such an architecture might include text-based NN, CNNs, and popular models like decision trees. The predicted user engagement consists of the count of likes, comments, shares, and comment sentiment. Models that predict numbers use regression and mean-squared error (MSE) as their loss function. This research explores whether model architectures that combine text and image better produce a model with a lower loss than their text and image counterparts.

3.   Methods

Our proposed method combines both text and image data as features in a machine-learning algorithm to predict user interaction with page content. We did this by examining user interaction with Facebook posts, representing both images and text as features within a machine learning model.
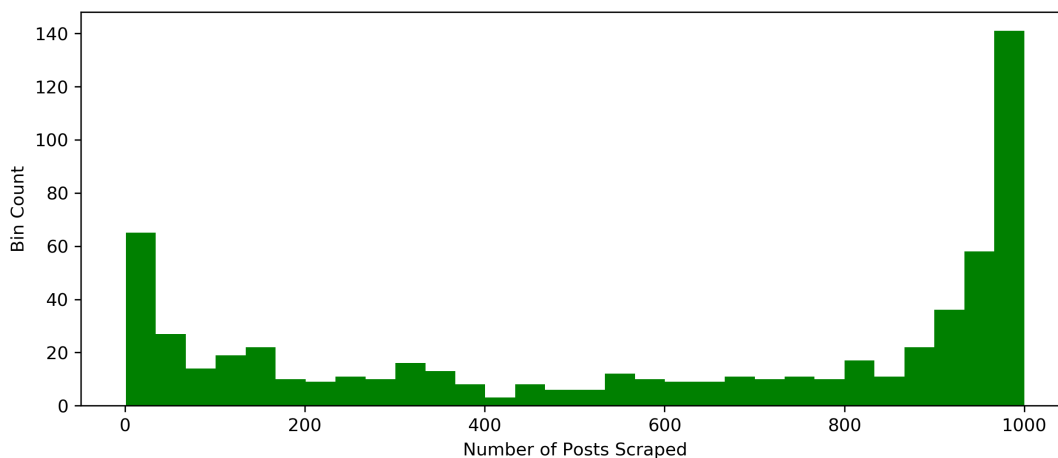
Figure 1: Histogram of the Number of Posts Scraped Per Facebook Page

3.1 Data Context

The research obtains the Facebook post URLs from AdEspresso. Hootsuite owns the web-
site, a social media management platform created in 2008 (Hootsuite). The company man-
ages content for many platforms, including Twitter, Facebook, Instagram, LinkedIn, Google+,
and YouTube (Hootsuite). The AdEspresso website provides a management platform for so-
cial media advertising. It allows marketers to create and promote advertisements in a single
place. The website features over one-hundred-thousand demo Facebook advertisements.
This paper scraped the links of 281,090 available advertiser pages linked on this website.

3.1.1 Data Origin

3.1.2 Data Collection

The researchers created a Python web scraper to crawl Facebook pages via the Facebook
Graph API, a publicly available API. The top 1000 posts are scraped from each Facebook
page, though most pages contain fewer than 100 posts. Figure-1 shows a histogram of the
number of posts scraped from each Facebook page. In total, the study collected 366,415
Facebook posts and 1,305,375 million comments.

The study also analyzed comment sentiment scores for each post. This study made use
of VADER, a parsimonious rule-based model for Sentiment Analysis of Social Media Text
Gilbert (n.d.) for its sentiment analysis. The sentiment analysis output was a score between
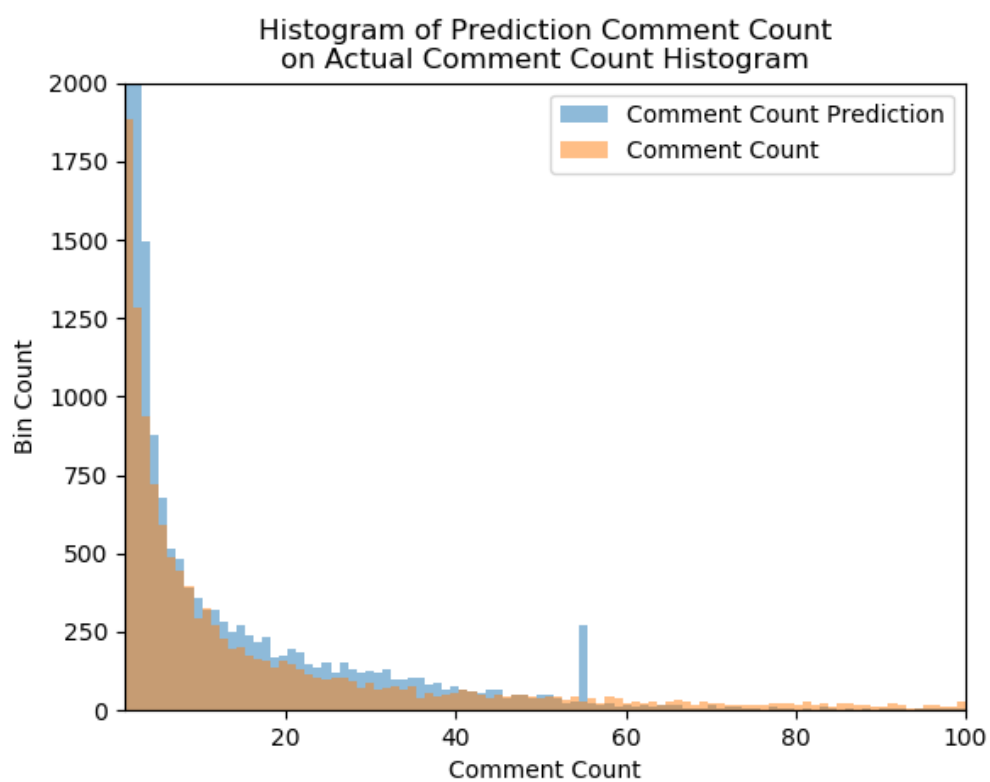-1 and 1, where 0 denotes neutral sentiment, -1 is very negative, and +1 a very positive

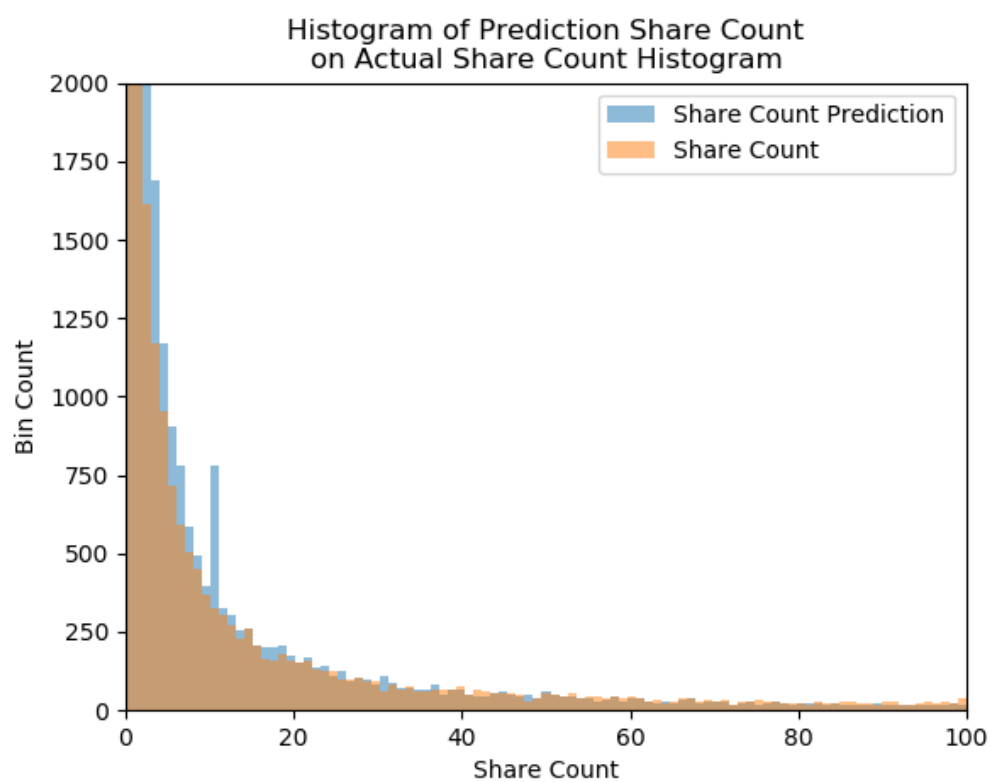Figure 2: Actual vs Predicted Comment Count Histogram

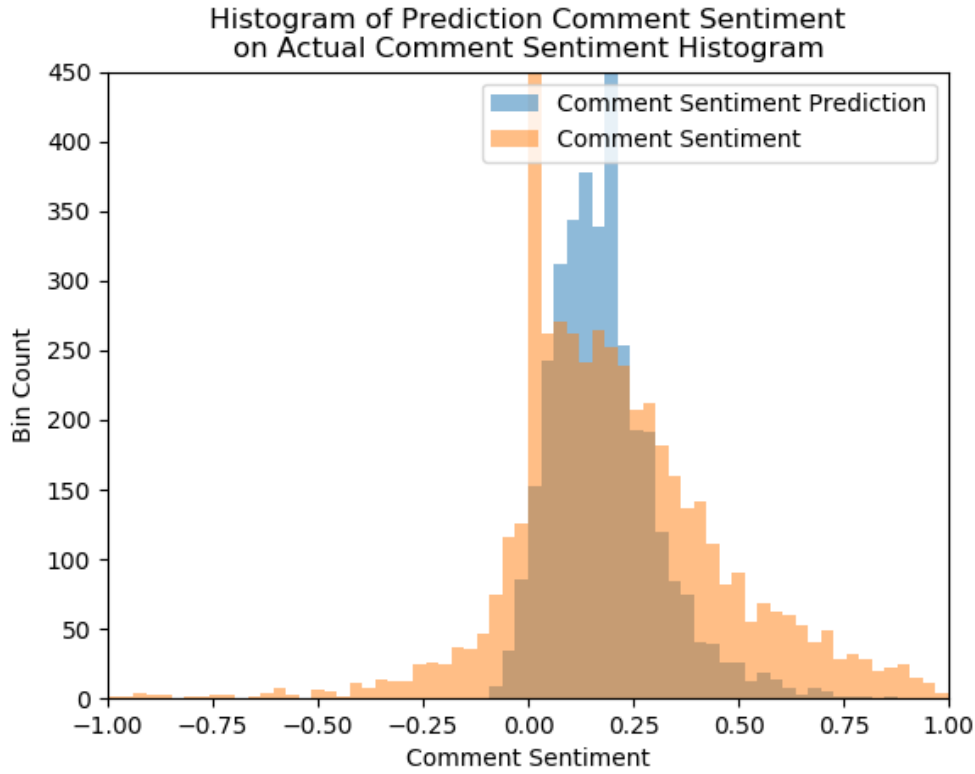Figure 3: Actual vs Predicted Share Count Histogram

Figure 4: Actual vs Predicted Comment Sentiment Histogram

sentiment. Posts with comments were given a sentiment score by averaging the sentiment score from all its comments. Figure-4 shows a histogram of comment sentiments for the Facebook posts.

3.1.3 Image Processing

Images are inherently highly dimensional. In order to reduce the number of features created by using images as part of our machine learning algorithm, we took several steps to reduce each image's dimensionality. First, we used principal component analysis. The number of dimensions to keep was set at 20 since the image variance is lower after denoising, reducing noise and data size. Image denoising can further reduce noise. Second, the research applied a Gaussian blur with a standard deviation of five to each image, which applies image blurring and emphasizes edges. Dilation and erosion are also applied to the image to remove noise in the edge space.

3.2 Text Processing

The method needs to transform text strings in order to create features for a neural network. Our process transformed text data into word vectors. First, the program split data into word tokens using whitespace as a delimiter. The program grouped these tokens into sentences, and both lowercased and removed common English stopwords and words with three or fewer characters. A port stemmer creates stems for all the words. A POS tag library performed parts-of-speech tagging. The program then extracted stems with a word lemmatizer, which takes the stem and the POS tag as input. Finally, the program fed the lemmatized text sentences to a td-IDF vectorizer. This created word vectors for training the neural network.

3.3 Methodology

The collected data consists of two types of data: text and image data. The text consists of Facebook posts and comments, and image data are any images attached to the post. These train and test this study's machine learning models.

The machine learning models are separated into three groups, corresponding to the types of input data. The three types of input data were text-only data, image data, and input consisting of text and image data. Each input type had a corresponding machine learning architecture for training. For example, text data train a NN, images train on a CNN, and the combination of text and images train on a combined NN and CNN model. Thus, three input types for three types of models make a total of nine models.

4. Results

Our machine learning algorithm's output was training against each post's comment count, share count, and comment sentiment. The experiments use different sized data sets. Comment and share count metrics utilized 366,415 posts in training and testing. Comment sentiment experiments included 201,215 posts. The comment sentiment data set was smaller due to limitations in the Facebook API.

The Python library Keras trained the machine learning models. The models predict continuous variables, which requires the model to perform regression. Keras provides many loss functions for regression. The research's ultimate goal is to differentiate between ad performance, so the researchers chose a loss function that more heavily penalizes significant differences in user interaction, which is accomplished by choosing the Mean Squared Error

Metrics / Model	Text-Based NN	Image-Based CNN	Combined Decision Tree	Combined NN
Share Count	3.44	1.01	2.58	1.00
Comment Count	1.02	1.01	1.29	1.00
Comment Sentiment	1.42	1.14	4.20	1.00

Table 1: Model Mean Squared Error Reported as a Ratio to the Best Model's Performance

(MSE) loss function since it has a squared loss penalty for any differences. The combined model outperformed both the text-only and the image-only models for predicting all social media metrics. Figure-**??** displays results as a ratio to the combined model's loss.

4.0.1   Model Performance

The combined model best predicted all user behavior metrics. Each machine learning model had its lowest MSE predicting for share count. The CNNs achieved a lower MSE than the NN on all metrics. Moreover, the CNN performance is best on data exhibiting a higher variance. Figure 3 and 4 show the predicted vs actual distribution for the combined model.

4.1   Linear Models of Scraped Data

This research created linear models to better understand the data before creating machine learning models. R scripts generated the linear models. The reported numbers are the coefficient of determination of the linear models. The linear models found an R^2 of 0.22 between talking about count and share count. The analysis found an R^2 of 0.38 between the fan count and talking about the count. The most significant result was an R^2 of 0.44 between comment sentiment and comment count. An R^2 of 0.38 and 0.44 are significant enough to create multicollinearity issues in machine learning features, which means that best practice avoids including two strongly correlated features. In addition, it is best practice to avoid correlated variables because they can negatively affect model weights by giving double the signal due to variable correlation.

4.2   Text-based NN Models

This research utilized Keras to create, train, and test this model. Initial training experimented with 2-8 hidden layers. The final model utilized two hidden layers, as two layers performed similarly to eight hidden layers. This research initially experimented with word

vector sizes from 1k-400k. Good performance and fast training occurred with a word vector size of 10k. The NN models used 10k as the word vector size.

4.3 Image-based CNN Models

Related work includes sections on keypoint descriptors and neural networks. CNNs detect image features and key points during repeated pooling and filter stages. This replaces the usefulness of image descriptors. Applied research uses CNNs in favor of image descriptors. Moreover, CNNs were faster to train than comparing image descriptors between images. This research trained with as many as eight layers and as few as one hidden layer. Model depths above three trained to a similar degree of accuracy. The models experimented with different convolutional size sequences. The experiments include doubling in size, usually from 64 to 128, with a 2x2 filter. The alternative was reduced by size in half with a 2x2 filter. Models with a 2x2 filter that decreased their size by half produced the highest accuracies. The final model contained four hidden layers. This research used this CNN configuration for all CNN model training. The final image-based model performed with a better loss than the text-based NN model on all metrics.

4.4 Combined Models

The combined model concatenates both text-based NN and image-based CNN. The Keras API includes the ability to combine models. The API concatenated each model after the CNN flattened. At this point, the CNN portion has finished, and the remaining eight nodes form a NN. At the point of model concatenation, the NN also decreased to eight nodes. The result was a NN layer of sixteen densely connected nodes—the sixteen-node NN connected to a final output node.

4.5 Application

The research applied the models in a real-world application. The application explored a use case to demonstrate the model's ability to choose which two advertisements would have the best performance. The research only compares advertisements with a significant value difference, where a significant difference is anything more than one standard deviation, as defined by the data distribution. The research found that models are unlikely able to differentiate between advertisements with similar performances. The combined model performs all predictions since it performed best across all metrics. Moreover, the application is the

most useful if it can detect poor-performing and best-performing advertisements.

The scenario is predicting user engagement for advertisements. The scenario pairs Facebook posts together. The model predicts user engagement for all posts. The program scores how often the model correctly predicted for greater user engagement. The program reports the result for each metric. The score is the number of correct predictions over incorrect predictions. This research compared the score with a random model. The random model produced the correct answer 50% of the time. The combined model scored 93% for comment sentiment, 65% for comment count, and 63% for share count.

The model performance shows its applicability in the real world. Platforms could employ the models to aid advertisers in choosing the best performing advertisement before paying to advertise it on social media. The model can tell advertisers which ads would perform best on the platform, which allows advertisers have their ads vetted. The vetting could prevent advertisers from spending large amounts of money showing worse ads. Moreover, the vetting would allow advertisers to only show ads that will perform best. In the context of billions, 93%, 65%, and 63% accuracy is a substantial monetary difference.

4.6   Linear Models

The data correlations were sensible. The metrics for comment count and share count are correlated. This correlation is sensible, considering that users talk about shared content. Moreover, the shared content is interesting. Shared content also creates more interesting content. Linear models also found a correlation between fan count and share count. The more fans a page has, the more opportunities exist for people to share the content. Alternatively, people are fans of the page because they like the content and if people like content, they are more likely to share it.

Linear models also discovered a correlation between comment count and comment sentiment. Posts that receive many comments are likely to receive positive comments. We can infer that happier users comment at a higher rate, or positive content elicits more comments, which are also likely to be positive. Interestingly, comment count was not correlated to share count. One might infer that just because a post draws attention and comments does not mean they are likely to get shared. It might also mean that happier content, which receives more comments, does not cause users to share that content. Page metrics are also not correlated with comment count, i.e., a famous or less popular page did not necessarily generate more or fewer comments. The study uses the lack of correlation to justify comparing posts across Facebook pages.

4.7   Metric Prediction

The data correlations were sensible. The metrics for comment count and share count are correlated. This correlation is sensible, considering that users talk about shared content. Moreover, the shared content is interesting. Shared content also creates more interesting content. Linear models also found a correlation between fan count and share count. The more fans a page has, the more opportunities exist for people to share the content. Alternatively, people are fans of the page because they like the content and if people like content, they are more likely to share it.

Linear models also discovered a correlation between comment count and comment sentiment. Posts that receive many comments are likely to receive positive comments. We can infer that happier users comment at a higher rate, or positive content elicits more comments, which are also likely to be positive. Interestingly, comment count was not correlated to share count. One might infer that just because a post draws attention and comments does not mean they are likely to get shared. It might also mean that happier content, which receives more comments, does not cause users to share that content. Page metrics are also not correlated with comment count, i.e., a famous or less popular page did not necessarily generate more or fewer comments. The study uses the lack of correlation to justify comparing posts across Facebook pages. The machine learning models were quick to train. The large batch sizes of 256 helped initialize model weights and prevent overfitting. Employing the Adam optimizer also helped reduce training time. This research scaled the data using min-max scaling. Each type of model required a similar quantity of training time.

Models with a higher variance achieved the worst overall loss/variance ratio. It seems that the larger the data variance, the higher the resulting mode's MSE. Share count had the highest variance of all the measured metrics, 1000x more than the comment count. The high MSE likely reflects that share counts are less related to the image and text data. Share counts might be a factor of other features, like page popularity.

The combined model outperformed all other models for each metric. The result is even more surprising in light of the decision tree performance. The decision tree also took as input data from both the text-based NN and the image-based CNN. However, the decision tree performed far worse than the combined model. While the exact reason is unknown, there are a few differences between the two models. First, the combined model closely integrates with the text-based NN and image-based CNN. Second, when the combined model trains, it also trains its two-parent models. As a result, the decision tree did not learn alongside its inputs. Likely, parent models compensate for mistakes by training together.

The model concatenation occurs in the later stages of each model. The goal was to combine these dense networks before the final output and after the CNN. The goal was a combined model that kept parent model training more-so independent. Combining models at the end keeps both models somewhat separate. This research plotted prediction vs. test data for all models. The combined model best resembled the test data's distribution. The text-based NN and image-based CNN tended to skew towards zero due to spread predictions.

Image-based models performed better than text-based models on all metrics, likely because it emphasizes images' importance for predicting user interaction on social media. Text-based models were poor predictors for many metrics. Text-based models did especially badly in prediction share counts. One may guess that users are sharing content they consider worthy of sharing.

4.8   Model Loss and Outperforming Random Guessing

In light of the newly created data, there are no other baseline measures for what constitutes good model performance. One goal was to produce a model that performed better than random guessing. Random guessing alone is not representative of the data. A better guess is based on the input's distribution. A good random guess would consider the data's value at each point along with its distribution. Such a distribution would weigh each value by the frequency of data. This calculation is the expected value. For a normal distribution, the expected value is equal to the mean. Also, each output in the research resembles the normal distribution after a log transformation. If the model always predicts the mean, the MSE is equal to the data's variance. The variance is a squared order of the data's distance from the mean. The MSE is also a squared order of the model's prediction from the actual value. Producing a model whose loss is below the variance is a general measure of demonstrating the model is doing a significant amount of learning. Fortunately, the combined model's performance was always much lower than each output data's variance.

4.9   Study Contributions

The study demonstrates the improved model performance by combining image and text data when predicting user behavior on social media. The research found that machine learning with both image and text data results in enormous improvements for predicting comment sentiment. This serves as a benchmark for predicting user sentiments on social media. It also provides a comparison between text-based, image-based, and combined text

and image model performance. The code is available on GitHub for future studies.

The research serves as a case study for combining text and image data. The study provides performant model hyperparameters in model depth, dropout rate, and word vector sizes. The expected results will be generalizable to other studies and encourage others to explore understanding advertisement performance on social media.

4.10 Addressing Research Questions

This research question asks if a machine learning model, trained with text and image data, can predict user interaction on social media. The research found that models utilizing image and text data outperformed their single-data type counterparts. The research questions also asked how well the models can predict user interaction in terms of MSE and variance. Each machine learning model achieved a MSE below the data variance, demonstrating a measure of model success learning user behavior. In addition, the combined model predicts which of two advertisements would evoke the more positive comment sentiment 93% of the time, further confirming the model's successful performance.

4.11 Limitations

Model training time was a limitation. There were twelve models. Each model requires many hours of training time. Machines with GPUs can only train one model at a time, so this process was very time-consuming. Time was not sufficient to train all models thoroughly. With more time, the models could have reached lower losses. One observation is that the models generally begin to overtrain after 30 epochs. More regularization might mitigate this problem. Such mitigations would likely lead to increased model performance. Heuristics dictated for hyperparameters and batch sizes.

4.12 Future Work

Given that share count and comment-sentiment have a coefficient of determination over 0.4, the share count model is likely a good input for predicting comment sentiment. Future research can incorporate this input to improve future models when predicting user behavior. In addition, future research could overcome existing Facebook API scraping problems. Another avenue for research is collecting comment sentiments for the 350k Facebook posts and using the new data to train better models. Moreover, future work can incorporate the combined text and image model with the image-based CNN and text-based NN into an en-

semble model to see if it improves performance.

The models created from this research could generate data to train a generative model. The generative model could transform images and text into advertisements that should generate more user interaction. The transformed and original advertisements could both be shown on social media and their user interactions compared. Thus, the study might demonstrate that generative models generate and improve existing advertising content.

References

Authors, F. (2013). Do users look at banner ads on Facebook? *Journal of Research in Interactive Marketing*, *7*(2), 119–139.

Bhat, S., Bevans, M., & Sengupta, S. (2002). Measuring users' web activity to evaluate and enhance advertising effectiveness. *Journal of Advertising*, *31*(3), 97–106. doi: 10.1080/00913367.2002.10673679

Fisher, T. (2009). Roi in social media: A look at the arguments. *Journal of Database Marketing and Customer Strategy Management*, *16*(3), 189–195. doi: 10.1057/dbm.2009.16

Gelli, F., Uricchio, T., Bertini, M., Del Bimbo, A., & Chang, S.-F. (2015). Image Popularity Prediction in Social Media Using Sentiment and Context Features. *Proceedings of the 23rd ACM international conference on Multimedia - MM '15*, 907–910. Retrieved from http://dl.acm.org/citation.cfm?doid=2733373.2806361 doi: 10.1145/2733373.2806361

Georgiou, D., MacFarlane, A., & Russell-Rose, T. (2015). Extracting sentiment from healthcare survey data: An evaluation of sentiment analysis tools. *Proceedings of the 2015 Science and Information Conference, SAI 2015*, 352–361. doi: 10.1109/SAI.2015.7237168

Gilbert, E. (n.d.). VADER : A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text.

Greenwood, S., Perrin, A., & Duggan, M. (2016). *Social Media Update 2016.* Pew Research Center.

Guo, L., Lu, R., Zhang, H., Jin, J., Zheng, Z., Wu, F., … Gai, K. (2020). A Deep Prediction Network for Understanding Advertiser Intent and Satisfaction. In *International conference on information and knowledge management, proceedings* (pp. 2501–2508). doi: 10.1145/3340531.3412681

Hassner, G. L., & Tal. (2015). Age and Gender Classification using Convolutional Neural Networks. *2008 8th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2008*, *24*(3), 2622–2629. Retrieved from http:// ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4761364%5Cnhttp:// portal.acm.org/citation.cfm?doid=1180639.1180711%5Cnhttp://www.ncbi.nlm.nih .gov/pubmed/21193381%5Cnhttp://dx.doi.org/10.1016/j.patcog.2012.09.011 doi: 10.1109/AFGR.2008.4813314

Hu, Y., Shin, J., & Tang, Z. (2016). Incentive problems in performance-based online advertising pricing: Cost per click vs. cost per action. *Management Science*, *62*(7), 2022–2038. doi: 10.1287/mnsc.2015.2223

Imsa, M. A., & Irwansyah. (2020). Online Advertising Effectiveness for Advertiser and User. , *459*(Jcc), 216–221. doi: 10.2991/assehr.k.200818.050

Khosla, A., Das Sarma, A., & Hamid, R. (2014). What Makes an Image Popular ? *Proceedings of the 23rd International Conference on World Wide Web*, 867—-876. Retrieved from http://dl.acm.org/citation.cfm?id=2567996 doi: 10.1145/2566486.2567996

Li, C., Lu, Y., Mei, Q., Wang, D., & Pandey, S. (2015). Click-through Prediction for Advertising in Twitter Timeline. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '15*, 1959–1968. Retrieved from http:// dl.acm.org/citation.cfm?doid=2783258.2788582 doi: 10.1145/2783258.2788582

Lin, H., Jia, J., Guo, Q., Xue, Y., Li, Q., Huang, J., . . . Feng, L. (2014). User-level psychological stress detection from social media using deep neural network. *Proceedings of the ACM International Conference on Multimedia - MM '14*, 507–516. Retrieved from http://dl .acm.org/citation.cfm?doid=2647868.2654945 doi: 10.1145/2647868.2654945

Liu, B. (2012). Sentiment analysis and opinion mining. (May), 1–108. doi: 10.2200/ S00416ED1V01Y201204HLT016

Ohsawa, S., & Matsuo, Y. (2013). Like Prediction: Modeling Like Counts by Bridging Facebook Pages with Linked Data. *Proceedings of the 22Nd International Conference on World Wide Web Companion*, 541–548. Retrieved from http://dl.acm.org/citation.cfm?id= 2487788.2487992 doi: 10.1145/2487788.2487992

Poria, S., Cambria, E., Hazarika, D., & Vij, P. (2016). A deeper look into sarcastic tweets using deep convolutional neural networks.

Romero, N. L. (2011). ROI. Measuring the social media return on investment in a library. *Bottom Line*, *24*(2). doi: 10.1108/08880451111169223

Schacht, J., Hall, M., & Chorley, M. (2015). Tweet if you will - The real question is, who do you influence? *Proceedings of the 2015 ACM Web Science Conference*(June). doi: 10.1145/2786451.2786923

Segalin, C., Cheng, D. S., & Cristani, M. (2017, mar). Social profiling through image understanding: Personality inference using convolutional neural networks. *Computer Vision and Image Understanding*, *156*, 34–50. Retrieved from https://www.sciencedirect.com/science/article/pii/S1077314216301679 doi: 10.1016/J.CVIU.2016.10.013

Straton, N., Mukkamala, R. R., & Vatrapu, R. (2017). Big social data analytics for public health: Predicting facebook post performance using artificial neural networks and deep learning. In *2017 ieee international congress on big data (bigdata congress)* (p. 89-96). doi: 10.1109/BigDataCongress.2017.21

Tiago, M. T. P. M. B., & Verissimo, J. M. C. (2014). Digital marketing and social media: Why bother? *Business Horizons*, *57*(6), 703–708. doi: 10.1016/j.bushor.2014.07.002

Wang, Y., Wang, S., Tang, J., Liu, H., & Li, B. (2015). Unsupervised sentiment analysis for social media images. *Proceedings of the 24th International Conference on Artificial Intelligence*, 2378–2379. Retrieved from https://arxiv.org/abs/1604.03489

Xu, C., Cetintas, S., Lee, K.-C., & Li, L.-J. (2014). Visual sentiment prediction with deep convolutional neural networks.