The Complexity of Social Media Response: Statistical Evidence For One-Dimensional Engagement Signal in Twitter

Damian Konrad Kowalczyk
Microsoft Development, Kanalvej 7
2800 Kgs. Lyngby, Denmark
ORCID: 0000-0004-8882-8330
dakowalc@microsoft.com

Lars Kai Hansen
DTU Compute, Matematiktorvet 324
2800 Kgs. Lyngby, Denmark
ORCID: 0000-0003-0442-5877
lkai@dtu.dk

Abstract

Many years after online social networks exceeded our collective attention, social influence is still built on attention capital. Quality is not a prerequisite for viral spreading, yet the size of diffusion cascades is still the hallmark of social influence. Consequently, our exposure to low-quality content and questionable influence is expected to increase. Since the conception of influence maximization frameworks, multiple content performance metrics became available, albeit raising the complexity of influence analysis. In this paper, we examine and consolidate a diverse set of content engagement metrics. The correlations discovered, lead us to propose a new compound engagement signal. We then show it is more predictable than any individual influence predictor investigated before. Our proposed model achieves strong engagement ranking performance and is the first to explain half of the variance with features available at the time of engagement. We share the detailed numerical workflow to compute the new compound engagement signal. We consolidate a diverse set of content engagement metrics. The correlations discovered, lead us to propose a new compound engagement signal. We then show it is more predictable than any individual influence predictor investigated before. Our proposed model achieves strong engagement ranking performance and is the first to explain half of the variance with features available at the time of engagement. We share the detailed numerical workflow to compute the new compound engagement signal.

Extant work on influence analysis focuses on homogenous information networks and attributes the greatest influence to authors triggering the largest diffusion cascades (Franck 2019). When the author’s influence is modeled as the ability to maximize the expected spread of information in the network (Pezzoni et al. 2013), the most desirable user-generated content is the one propagated furthest, in Twitter measured by the number of retweets. We observe millions of attempts at viral spreading, only on Twitter today. The abundance of information to which we are exposed through online social networks is exceeding our capacity to consume it (Weng et al. 2012), let alone in a critical way. (Weng et al. 2012; Qiu et al. 2017) show that content quality is not a prerequisite for viral spreading, and (Lorenz-Spreen et al. 2019) show that the competition for our attention is growing, causing individual topics to receive even shorter intervals of collective attention. Consequently, our exposure to low-quality information and by extension low-quality influence, is increasing (Table 1). Still, few studies systematically investigate how to model the strength of influence in heterogeneous information networks, and the processes that drive popularity in our limited-attention world remain mostly unexplored (Franck 2019; Weng et al. 2012). Propagation metrics (retweet count in particular) do not capture, the average individual attention received. Retweet action does not inform, e.g., if the actor has actually read the content let alone consider the source, or whether that effort was left to the followers.

The digital footprint of an audience goes far beyond the retweet action. Platforms like Facebook and Twitter record an increasingly diverse set of user behaviors, including number of clicks, replies or favorites (likes). Since (Pezzoni et al. 2013), Twitter in 2015 has made many of these metrics available to the public, inviting a more comprehensive approach to influence modeling, albeit rising the complexity of all dependent tasks.

The four Tweets in Table 1 show that the mechanisms leading to high values of either retweet count, replies or favorites are complex. In the following work,

| Tweet (Body) | Retweets | Replies | Favorites |
|--------------|----------|---------|-----------|
| HELP ME PLEASE, A MAN NEEDS HIS NUGGS | 1.61M | 16K | 2.29M |
| No one is born hating another person because of the color of his skin or his background or his religion... " | 2.43M | 235K | 3.11M |
| "No one is born hating another person because of the color of his skin or his background or his religion... " | 1.04M | 63K | 1.44M |

Table 1: Four influential tweets ranked by the most popular influence predictor: size of diffusion triggered in the network

Copyright © 2019, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.
we investigate the multi-dimensional response of on-line audiences, to understand this complexity. We examine and consolidate multiple discrete engagement metrics towards a new compound engagement signal. While the new signal is statistically motivated, we next show the relevance of the signal for understanding engagement in multiple datasets. In particular we show that the new signal is more predictable than the individual metrics (e.g., diffusion size measured by retweet count) prevalent in literature. Our engagement model is the first to explain half of the variance with features available early, and to offer strong (Cohen 1988) ranking performance simultaneously. We share the computational of the new compound engagement signal to ensure reproducibility.

Contributions of this paper are summarized as follows:

1. analysis of three individual content performance signals, showing evidence of one-dimensional engagement signal on Twitter
2. new compound engagement formula, capturing over 75% of variance in available engagement signals
3. advancing feature representation of user generated language, and temporal descriptors, to predict the size of the audience, to understand this complexity. We examine and consolidate multiple discrete engagement metrics towards a new compound engagement signal. While the new signal is statistically motivated, we next show the relevance of the signal for understanding engagement in multiple datasets. In particular we show that the new signal is more predictable than the individual metrics (e.g., diffusion size measured by retweet count) prevalent in literature. Our engagement model is the first to explain half of the variance with features available early, and to offer strong (Cohen 1988) ranking performance simultaneously. We share the computational of the new compound engagement signal to ensure reproducibility.

Methodology

In this section we describe the application of unsupervised learning towards contributions (1,2,6), novel data collection and feature extraction approach towards contribution (1,3), and the chosen supervised learning method towards contributions (4,5,6).

Principal Engagement Component

We acquire the multivariate set of responses forming a vector

\[ \mathbf{e} = [e_{\text{retweets}}, e_{\text{replies}}, e_{\text{favorites}}]^T. \]  

(1)

Recent work on engagement modeling, e.g., [Lee, Hosanagar, and Nair 2018] defines any response as a sign of engagement, effectively reducing the multivariate response to a one-dimensional signal. However, to our knowledge the complexity of the engagement signal has not been explored more formally. While it appears credible that the population response signals,i.e., the dimensions of the of vector \( \mathbf{e} \), are highly correlated, we can test the effective dimension of the space populated by the vectors using so-called Parallel Analysis (PA) [Jorgensen and Hansen 2011]. In PA principal component analysis of the measured signals is compared with the distribution of the principal components of null data obtained by permutation under a (null) hypothesis that there is no dependency between the individual response signals. Consistent with this hypothesis we can permute the sequence of the signals for each observation separately. In particular we compute the upper 95% quantile for the distribution of the eigenvalues in the permuted data. Eigenvalues of the original unpermuted data set that reject the null hypothesis are considered "signal".

Principal components are computed on the response signals subject to a variance stabilization transformation,

\[ \hat{e} = \ln(e + 1), \]  

(2)

see e.g., [Can, Oktay, and Manmatha 2013; Kowalczyk and Larsen 2019].

Projection on the engagement component

Hypothesizing a one-dimensional engagement signal, we compute the value as the projection on the first principal component of the transformed data of dimension \( D = 3 \),

\[ E_1 = \sum_{i=1}^{D} w_i (\ln(e_i + 1) - \mu_i), \]  

(3)

where \( \mu_i = \frac{1}{N} \sum_{n=1}^{N} e_i,n \) is the \( i \)’th component of the \( D \)-dimensional mean vector for a sample of size \( N \), while \( w_i \) is the \( i \)’th component of the first principal component, computed on the same sample.

Gradient Boosted Regression Trees (GBRT)

We consider the problem of predicting audience engagement for a given tweet based on features available immediately after its delivery (Table 3). Features describing the author are used together with the content, language, and temporal descriptors, to predict the size of retweet cascade, number of likes, number of replies and the proposed compound engagement signal. GBRT is a tree ensemble algorithm which builds one regression tree at a time by fitting the residual of the trees that preceded it. The training process minimizing a chosen twice-differentiable loss function, can be described as

\[ \theta^* = \arg \min_{\theta} \sum_{i=1}^{N} L_{\text{SE}}(\hat{e}_i, e_i), \]  

(4)

where \( \theta \) contains all parameters of the proposed model, \( N \) is the number of examples, and \( L_{\text{SE}} \) is the squared error of an individual prediction,

\[ L_{\text{SE}}(e, \hat{e}) = (e - \hat{e})^2. \]  

(5)

We follow [Can, Oktay, and Manmatha 2013; Kowalczyk and Larsen 2019] to stabilize variance of all individual engagement signals via log-transformation as in Equation 2.
Gradient Boosting Framework We use Microsoft’s implementation of Gradient Boosted Decision Trees (Ke et al. 2017) for model training and tuning. LightGBM offers accurate handling of categorical features by applying (Fisher 1958), which limits the dimensionality of our tasks. The framework benefits from GPU acceleration, due to histogram-based algorithm for finding the best splits (Zhang, Si, and Hsieh 2017). However, what differentiates LightGBM the most in our experiments, is the efficient leaf-wise growth algorithm.

Data Collection

| Dataset       | T2016-IMG | T2017-ML | T2018-ML |
|---------------|-----------|----------|----------|
| introduced    | Wang (2018) | Kowalczyk (2018) | now      |
| w/image only  | True      | False    | False    |
| languages     | English   | 18       | all      |
| months total  | 3         | 14       | 12       |
| month from    | 2016.10   | 2017.01  | 2018.01  |
| unique tweets | 2,848,892 | 9,719,264 | 28,883,324 |
| quoting       | 421,175   | 583,514  | 2,647,972 |
| retweets total| 5,929,850 | 11,361,699 | 42,919,158 |
| replies total | 717,644   | 3,576,976 | 12,414,907 |
| favorites total| 12,665,657 | 29,138,707 | 134,523,998 |
| no engagement | 1,547,829 | 5,929,850 | 14,813,772 |

Table 2: Datasets acquired

Unique Tweets We use Twitter Historical PowerTrack APIs to collect training and validation datasets described in Table 2. Retroactive filtering of Twitter archive allows us to rebuild datasets used in prior work e.g. (Wang, Bansal, and Frahm 2018; Kowalczyk and Larsen 2019) and to reduce topic bias via near-uniform sampling across long time-frames (Figure 1). Collecting a dataset similar to T2017-ML by sampling Twitter Firehose popular in prior work, would have taken 14 months.

Figure 1: T2017-ML volume per month: mitigating topic bias via extended time-frame sampling. Historical APIs allow near uniform distribution.

Engagement totals Three content engagement metrics are made publicly available by Twitter in 2014. We use Twitter’s Engagement Totals API to retrieve the number of retweets, replies and favorites ever registered for each tweet (even if removed later via unlike or account suspension). Use of the Engagement Totals API ensures 100% accuracy of our supervisory vector of response signals e.

Sentiment prediction (Hansen et al. 2011; Kowalczyk and Larsen 2019) show the impact of sentiment on tweet’s virality (retweetability). We reuse sentiment predictions from (Kowalczyk and Larsen 2019) for all tweets in the validation datasets to explore correlation with other engagement metrics and ensure fair comparison with previous results. The analysis was performed for tweets in 18 languages, using Text Analytics APIs from Microsoft Cognitive Services (Microsoft 2017).

Privacy respecting storage The data analyzed in this study is publicly available during collection. How much of it remains public, can change rapidly afterwards. We follow the architecture proposed by (Kowalczyk and Larsen 2019) to secure the data in a central highly scalable storage solution, exposed to applicable privacy requests from the Twitter’s Compliance Firehose API, and feature extraction requests from our Spark cluster.

Datasets Retroactively, we acquire three (Table 2) datasets:

1. **T2016-IMG** to enable fair comparison with the work of (Mazloom et al. 2016; McParlane, Moshfeghi, and Jose 2014; Khosla, Das Sarma, and Hamid 2014; Cappallo, Mensink, and Snoek 2015). Wang, Bansal, and Frahm 2018; Kowalczyk and Larsen 2019, validate our feature representation and training method. The dataset matches the same filters, as applied before (timeframe, language and the presence of an image attachment).

2. **T2017-ML** to evaluate the generalizability of our resulting models across seasons and languages (cultures) and comparison with the work of (Kowalczyk and Larsen 2019). This dataset represents a near-uniform sample of Twitter 2017 volume in all 18 languages supported by the sentiment analysis service (Microsoft 2017).

3. **T2018-ML** to evaluate the generalizability of our compound engagement signal across years. This dataset represents a near-uniform sample of entire Twitter 2018 volume in all known languages.

Datasets T2016-IMG and T2017-ML are split into 70% training, 20% test and 10% validation sets. In this study, T2018-ML dataset is used in unsupervised experiments only. To aid reproducibility, we share unique ID’s of acquired tweets along with sentiment predictions.

Feature extraction Table 3 describes features collected for each tweet. To ensure scalability in production, only information available at the time of engagement is considered. Since their introduction in 2015 by Twitter, quote tweets have gained on popularity. Over 3.5 million tweets collected for this study quote another (Table 2). We have not seen the original (quoted) tweets considered in literature before. In our study
### Results

We begin with examining all available content performance signals (count of retweets, replies and favorites) in the extended time-frame datasets. We look for potential correlations that could enable reducing the dimension of engagement using Parallel Analysis. In the supervised experiments, first we evaluate our methodology and feature representation against previous state-of-the-art methods, by modelling the individual influence metrics (e.g. virality) and the compound engagement on the benchmark dataset T2016-IMG. Finally we evaluate the generalizability of our method across topics and cultures, modeling engagement on the multilingual extended-timeframe dataset T2017-ML.

#### Evidence for a one-dimensional engagement signal

We perform Parallel Analysis and compute the principal components and their associated projected variances for the log-transformed data as well as for Q = 100 permutations of the data assuming the no correlation null. The one-sided upper 95% quantile is computed from the permuted samples. Variances of the unpermuted signals and the 95% quantiles for the three eigenvalues of the permuted data are shown in figure 2. Very similar results are obtained for the 2018 data set (not shown).

#### The engagement signal

We perform principal component analysis of the two data sets keeping a single principal component. The mean vectors and projections are found in Table 4. The variance explained by the first components in the three analyses were: 2016 : 83%, 2017 : 72%, 2018 : 77%.

![Figure 2: Parallel Analyses of the response signals for the 2017 data set provide evidence for a one-dimensional engagement signal: Only the first component ('1'- red dotted line) exceeds the 95% quantile of the corresponding eigenvalue in the null hypothesis (blue dashed line).](image)

### Predicting Engagement

First round of our supervised experiments focus on evaluating our user generated content feature representation and GBRT approach against previous state-of-the-art methods, in modeling the established engagement signals, like the size of diffusion (e.g. retweet count), response (i.e. number of replies) and popularity (i.e. number of favorites/likes), before attempting to predict the compound engagement.

#### Metrics

The absolute measure of fit (RMSE) is chosen as an objective of optimization, to penalize large errors. We compute the Spearman ρ ranking coefficients, to measure each models ability to rank the content depending on the definition of engagement. We compute the relative measure of fit $R^2$ to compare the variance explained by the compound engagement vs individual engagement signals. Interpretation of $R^2$ and Spearman ρ is domain specific, with guidelines for social and behavioral sciences proposed by (Cohen 1988). SciPy version 1.3.1 is used to ensure ρ tie handling. The p-value for all reported ρ results is $p < 0.001$. Each metric is an average from 3-fold cross-validation.

|          | retweets | replies | favorites |
|----------|----------|---------|-----------|
| T2017-ML | $\mu_1$  | $\mu_2$ | $\mu_3$   |
|          | 0.451    | 0.049   | 0.145     |
|          | 0.082    | 0.880   | 0.148     |
| T2018-ML | 0.450    | 0.066   | 0.188     |
|          | 0.080    | 0.872   | 0.205     |

Table 4: First principal components of the extended time-frame engagement signals
Table 5: Method evaluation on the T2016-IMG dataset.

| Method                                      | $R^2$ | $\rho$ | RMSE |
|---------------------------------------------|-------|--------|------|
| McParlane, Moshfeghi, and Jose 2014         | -     | 0.257  | -    |
| Khosla, Das Sarma, and Hamid 2014           | -     | 0.254  | -    |
| Cappallo, Messink, and Snoek 2013           | -     | 0.258  | -    |
| Mazloom et al. 2016                        | -     | 0.262  | -    |
| Wang, Bansal, and Frahm 2018†              | 0.391 | 0.504  | 0.555|

Table 6: Engagement prediction performance on T2017-ML dataset.

| Method                                      | $R^2$ | $\rho$ | RMSE |
|---------------------------------------------|-------|--------|------|
| Kowalczyk and Larsen 2019                  | 0.402 | 0.369  | 0.336|
| virality (retweets)                         | 0.425 | 0.371  | 0.329|
| response (replies)                         | 0.302 | 0.512  | 0.292|
| popularity (favorites)                      | 0.493 | 0.526  | 0.484|
| engagement (compound)                       | 0.507 | 0.529  | 0.228|

T2017-ML dataset with Nvidia RTX 2080Ti GPU took 4 minutes. Computing predictions for the 3 million unique tweets in the validation set, took another 55 seconds. This implies throughput of over 54,000 vectorized tweets / second, with a single GPU.

Figure 3 offers a comparison of feature importance between all models trained on T2017-ML dataset. Feature importance is calculated for every model by aggregating purity gain per feature, before averaging across 3-folds and normalizing to compare across models. The uncertainty for virality features does not exceed 6%. When
Table 7: Four influential tweets, ranked by the compound engagement

| Tweet (body)                                                                 | Engagement |
|------------------------------------------------------------------------------|------------|
| “No one is born hating another person because of the color of his skin or his background or his religion...” | 9.283      |
| “If only Bradley’s arm was longer. Best photo ever. #oscars”                  | 9.266      |
| ZOZOTOWN新春セール史上最大で取引100を先は(...)                                     | 9.158      |
| “HELP ME PLEASE. A MAN NEEDS HIS NUGGS”                                     | 8.822      |

predicting response (i.e. number of replies) we find the number of users mentioned to have the highest predictive value, while the number of image attachments (i.e. mediaCount) to have virtually none. Number of followers, the most popular feature in all prior work on virality prediction is only 4th when predicting compound engagement. Average number of followers received with each status (followersStatusRatio) or number of times the author liked another tweet are far more predictive of engagement.

Conclusion
In this study we have analysed the complexity of the multivariate response of users engaging with social media. We have examined and consolidated various engagement metrics available today from Twitter.

The significant correlation found between individual response signals leads us to propose a new one-dimensional compound engagement signal. We showed on multiple benchmark datasets, that compound engagement is more predictable than any individual engagement signal, significantly outperforming the size of diffusion cascade, predominant in influence maximization frameworks [Franck 2019; Eshgi et al. 2019].

Our engagement model is the first to explain half of the variance with features available early, and to offer strong (Cohen 1988) ranking performance simultaneously. The model is ready for production with immediate application to social media monitoring, campaign engagement forecasting, influence prediction and maximisation. We share the compound engagement workflow and parameters (Eq. (3) and Table (4)) to ensure reproducibility and inspire work on engagement modeling. We propose the ability to engage the audience, as a new baseline for social influence modeling.

Acknowledgements
This project is supported by Microsoft Development Center Copenhagen and the Danish Innovation Fund, Case No. 5189-00089B. We would like to thank Sandeep Aparajit, Pushpraj Shukla, Jörg Derungs, Ralf Gautschi, Walter Sun and Charlotte Mark. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the sponsors.
References

Beyersdorf 2019] Beyersdorf, B. 2019. Regulating the Most Accessible Marketplace of Ideas in History: Disclosure Requirements in Online Political Advertisements after the 2016 Election. California Law Review 107.

Bueno 2016] Bueno, C. 2016. The attention economy: Labour, time and power in cognitive capitalism.

Bybee and Jenkins 2019] Bybee, K. J., and Jenkins, L. 2019. Free Speech, Free Press, and Fake News: What If the Marketplace of Ideas Isn’t About Identifying Truth? SSRN Electronic Journal.

Can, Oktay, and Manmatha 2013] Can, E. F.; Oktay, H.; and Manmatha, R. 2013. Predicting retweet count using visual cues. In Proceedings of the 22nd ACM international conference on Conference on information & knowledge management - CIKM ’13.

Cappallo, Mensink, and Snoek 2015] Cappallo, S.; Mensink, T.; and Snoek, C. G. 2015. Latent Factors of Visual Popularity Prediction. In Proceedings of the 5th ACM on International Conference on Multimedia Retrieval - ICMR ’15.

Cohen 1988] Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates.

Davenport and Beck 2001] Davenport, T., and Beck, J. 2001. The attention economy: Understanding the new currency of business.

Eshgi et al. 2019] Eshgi, S.; Maghsudi, S.; Restocchi, V.; Stein, S.; and Tassiulas, L. 2019. Efficient influence maximization under network uncertainty. In INFOCOM 2019 Workshop proceedings.

Fisher 1958] Fisher, W. D. 1958. On Grouping For Maximum Homogeneity. American Statistical Association Journal.

Franck 2019] Franck, G. 2019. The economy of attention. Journal of Sociology 55(1):8–19.

Hansen et al. 2011] Hansen, L. K.; Arvidsson, A.; Nielsen, F. A.; Colleoni, E.; and Etter, M. 2011. Good friends, bad news - Affect and virality in twitter. In Communications in Computer and Information Science.

Horn 1965] Horn, J. L. 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30(2):179–185.

Jørgensen and Hansen 2011] Jørgensen, K. W., and Hansen, L. K. 2011. Model selection for gaussian kernel pca denoising. IEEE transactions on neural networks and learning systems 23(1):163–168.

Ke et al. 2017] Ke, G.; Meng, Q.; Wang, T.; Chen, W.; Ma, W.; Liu, T.-Y.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; and Liu, T.-Y. 2017. LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems.

Khosla, Das Sarma, and Hamid 2014] Khosla, A.; Das Sarma, A.; and Hamid, R. 2014. What makes an image popular? In Proceedings of the 23rd international conference on World wide web - WWW ’14.

Kowalczyk and Larsen 2019] Kowalczyk, D. K., and Larsen, J. 2019. Scalable Privacy-Compliant Virality Prediction on Twitter. In Proceedings of AffCon 2019 @ AAAI, volume 2328, 12–27.

Lee, Hosanagar, and Nair 2018] Lee, D.; Hosanagar, K.; and Nair, H. S. 2018. Advertising content and consumer engagement on social media: evidence from facebook. Management Science 64(11):5105–5131.

Lorenz-Spreen et al. 2019] Lorentz-Spreen, P.; Monsted, B. M.; Hövel, P.; and Lehmann, S. 2019. Accelerating dynamics of collective attention. Nature Communications 10(1):1759.

Mazloom et al. 2016] Mazloom, M.; Rietveld, R.; Rudinac, S.; Worrting, M.; and van Dolen, W. 2016. Multimodal Popularity Prediction of Brand-related Social Media Posts. In Proceedings of the 2016 ACM on Multimedia Conference - MM ’16.

McParlane, Moshtaghi, and Jose 2014] McParlane, P. J.; Moshtaghi, Y.; and Jose, J. M. 2014. "Nobody comes here anymore, it’s too crowded": Predicting Image Popularity on Flickr. Proceedings of International Conference on Multimedia Retrieval - ICMR ’14.

Microsoft 2017] Microsoft. 2017. Cognitive Services APIs reference. https://westus.dev.cognitive.microsoft.com/docs/services/]

Pezzoni et al. 2013] Pezzoni, F.; An, J.; Passarella, A.; Crowcroft, J.; and Conti, M. 2013. Why do I retweet it? An information propagation model for microblogs. In Lecture Notes in Computer Science (including sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).

Qiu et al. 2017] Qiu, X.; F. M. Oliveira, D.; Sahami Shirazi, A.; Flammini, A.; and Menczer, F. 2017. Limited individual attention and online virality of low-quality information. Nature Human Behaviour 1(7):0132.

Wang, Bansal, and Frahm 2018] Wang, K.; Bansal, M.; and Frahm, J. M. 2018. Retweet wars: Tweet popularity prediction via dynamic multimodal regression. In Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018.

Weng et al. 2012] Weng, L.; Flammini, A.; Vespignani, A.; and Menczer, F. 2012. Competition among memes in a world with limited attention. Scientific reports 2:335.

Wu 2017] Wu, T. 2017. The attention merchants: The epic scramble to get inside our heads.

Zhang, Si, and Hsieh 2017] Zhang, H.; Si, S.; and Hsieh, C.-J. 2017. GPU-acceleration for Large-scale Tree Boosting.