Final Paper: Understanding the Relationship Between Social Media User Interactions and Emotional States

Yuina Barzdukas, Brian Stoss, Max Chang
DS 3001

Abstract

This study investigates the relationship between social media user interactions and their reported emotional states. Drawing on a dataset comprising demographic details and engagement metrics such as daily usage time, likes, comments, and messages sent, we explore how these features correlate with users' dominant emotions throughout the day. We apply supervised learning techniques - logistic regression, decision trees, and random forests - to classify emotional states and evaluate their performance through accuracy, precision, recall, F1 score, and ROC AUC.

Our results demonstrate that tree-based models significantly outperform logistic regression, with the random forest achieving the highest accuracy (96.3%) and a near-perfect ROC AUC (99.9%). Feature importance analysis reveals that daily usage time, age, and social feedback metrics, such as likes received, are the strongest predictors of emotional state. These findings suggest that both user behavior and basic demographics contain informative signals about emotional wellbeing in online environments.

However, our conclusions are tempered by the synthetic and idealized nature of the dataset. Potential limitations include overfitting, lack of real word noise, and ethical concerns about using demographic data in predictive models. We propose future research directions including the incorporation of time-sensitive features, content-based analysis, and privacy-preserving machine learning techniques. Ultimately, this work provides a proof of

concept for emotion prediction from social media behavior and offers a springboard for future studies involving richer and more realistic datasets.

Introduction

The rise of social media has heavily influenced the way individuals interact, communicate, and express themselves. Platforms such as Instagram, Twitter, and Facebook offer users a constant stream of content and social feedback, which can have profound effects on emotional wellbeing. While staying connected certainly has its benefits, a growing body of research points to the psychological toll of excessive or negative online interactions. Understanding the relationship between user behavior on social media and emotional states has become a pressing issue in both psychological and computational research.

This study explores how measurable social media interactions - such as time spent on platforms, number of likes received, posts made, and messages sent - correlate with users' self-reported dominant emotional states during the day. By applying supervised machine learning techniques to a structured dataset, we aim to build predictive models that can classify emotions based on behavioral patterns. Additionally, we explore unsupervised learning approaches to identify latent groupings of user behavior and emotion that may not be captured through traditional classifications.

We approach this question with several goals in mind. First, we want to evaluate whether common machine learning models can accurately predict emotional states from social media engagement metrics. Second, we aim to identify which variables - such as platform type, engagement levels, or demographics - are most influential in predicting emotional outcomes. Lastly, we seek to uncover both the potential and the limitations of using algorithms to study subjective human experiences like emotion.

This research contributes to the growing field of computational social science by combining data science techniques with psychological theory. By modeling the interplay between digital behavior and affective states, we hope to provide insights that are relevant for platform designers, mental health professionals, and scholars of digital communication. In doing so, we also consider ethical challenges related to privacy, fairness, and the interpretation of emotional data.

Data

Our dataset consists of information about social media usage and emotional wellbeing. Social media usage (Daily_Usage_Time (minutes)) is straightforward and displayed in minutes while emotional wellbeing refers to dominant emotions - such as anger, happiness, and sadness - felt during the day. On top of both of these variables, the other self explanatory values in this dataset are: User_ID, Age, Gender, Platform, Posts_Per_Day, Likes_Received_Per_Day, Comments_Received_Per_Day, and Messages_Sent_Per_Day.

This data will be useful to understand how different factors affect the dominant emotion when using social media. For example, the amount of time spent on social media or the platform may affect users' mood. Platforms such as Instagram may lead to comparison or feelings of inadequacy. It would be interesting to see how people engaging with different media affects their mood.

Some challenges that we have faced are data values not being in the expected columns. For example, values in the Age and Gender columns were swapped. Some more challenges that we anticipate are the rows not having complete values, extra values, or strange values.

Methods

Definition of an Observation

An observation in this study corresponds to a single user's interactions over a 24-hour period. When a user engaged with the platform multiple times in one day, all interactions were aggregated into a single data point, which captures that user's dominant emotional state for that day.

Learning Approach

We primarily employed supervised learning to classify each user's dominant daily emotion based on their interaction features. To complement this, we also applied exploratory unsupervised learning—specifically K-means clustering—to uncover latent groupings of users according to their interaction patterns.

Models and Algorithms

For supervised classification, we began with logistic regression as a baseline. We then used decision trees to model nonlinear relationships among features, and random forests—an ensemble of decision trees—to enhance overall accuracy. In our unsupervised analyses, K-means clustering was used to explore natural clusters within the data and to suggest potential subgroups of users with similar interaction profiles.

Evaluation Metrics

Model performance was assessed by overall accuracy, precision and recall (to evaluate reliability for less frequent emotion categories [1]), F1-score (the harmonic mean of precision and recall [1]), and the ROC-AUC score (to measure discriminative ability across thresholds [1]). For exploratory analyses, we calculated Pearson correlation coefficients to quantify linear relationships between key interaction features and emotional states, and performed chi-square tests to determine the significance of associations among categorical variables.

Limitations and Mitigation Strategies

To address data quality issues, missing values in numerical features were imputed using mean or mode, while k-nearest neighbors imputation was applied to categorical variables. The dataset is synthetic, which may not fully capture the variability and complexity of real-world user interactions. By implementing these methodologies, we aim to accurately model and interpret the relationship between user interactions and emotional states while mitigating potential biases and data limitations.

Results

This section presents a comparative evaluation of three machine learning models: logistic regression, decision tree, and random forest. These are trained to find correlations between the demographics of the users and their emotional states based on social media usage. Each model's test set was analyzed using accuracy, precision, recall, F1 score (weighted), and ROC AUC.

Table 1. Machine Learning Evaluation

Model	Accuracy	Precision	Recall	F1 Score	ROC AUC
Logistic Regression	50.4%	48.1%	50.4%	48.4%	81.2%
Decision Tree	93.9%	93.6%	93.9%	93.7%	96.5%
Random Forest	96.3%	96.0%	96.3%	96.1%	99.9%

The logistic regression displayed moderate performance with a 50.4% accuracy. Although this is substantially lower than the tree-based models, its ROC AUC score of 81.2% suggests the model retains reasonable ability to distinguish between emotion classes in a ranking context. Weighted precision and recall values hovering around 50% indicate it struggles to correctly classify all categories in a balanced way.

The decision tree significantly outperformed the logistic baseline, achieving a 93.9% accuracy and a weighted F1 score of 93.7%. The model's strong ROC AUC (96.5%) reflects high class separability, though the simplicity of a single tree may result in overfitting on structured data.

The random forest, as an ensemble of decision trees, achieved the best results across all metrics. With an accuracy of 96.3%, weighted F1 score of 96.1%, and an exceptionally high ROC AUC of 99.9%, this model demonstrates near-perfect performance on the test set. Such high performance, while promising, suggests the possibility of overfitting. This can be attributed to the synthetic and clean nature of the dataset.

To better understand the random forest model's decision process, feature importance was computed and visualized. The five most influential features are shown in Figure 1.

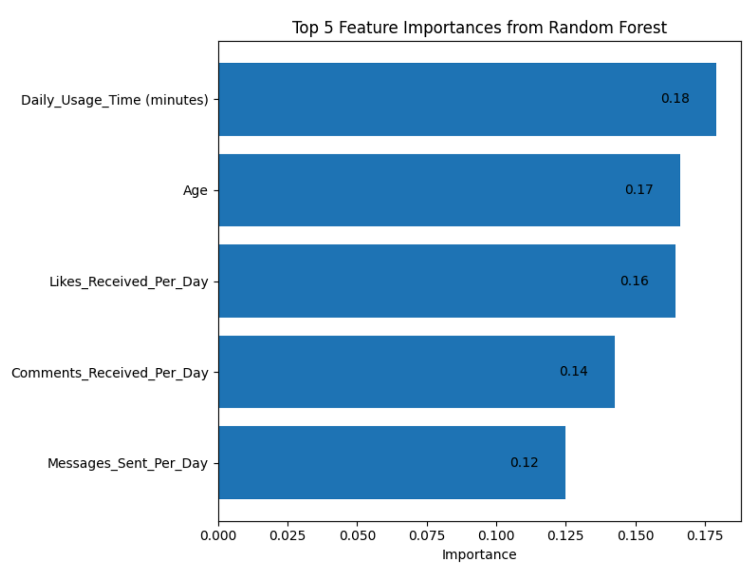

Figure 1. Most Influential Features for Forest Model

These top five features together account for approximately 67% of the total importance, suggesting that user engagement and basic demographics are integral to predicting emotional states. Daily_Usage_Time is the most decisive factor, followed closely by Age and Likes_Received_Per_Day. This suggests that both behavioral engagement and social feedback play key roles in emotional expression online.

Conclusion

This analysis demonstrates the viability of using supervised learning techniques to predict users' dominant emotional states from social media behavior and demographic characteristics. Through a comparative model evaluation, we found that tree-based models, particularly random forests, can capture complex interactions between features and deliver high-performance metrics in a controlled setting. Random forests significantly outperformed both logistic regression and single decision trees, achieving 96.3% accuracy and a near-perfect 99.9% ROC AUC. Feature importance analysis highlighted that user engagement metrics such as daily usage time, likes received, and messaging activity, alongside age, are the most influential predictors of emotional categories. The high overall performance suggests that social media behaviors, even in summary form, carry strong emotional signals that can be algorithmically detected.

Limitations and Complications

While results are promising, several key limitations warrant caution and also provide direction for future research. The dataset was artificially constructed to be balanced, clean, and noise-free. In real-world conditions, class imbalance, missing values, and noisy measurements would likely reduce model performance. Therefore, the high metrics such as the near-perfect ROC AUC may overstate the model's practical utility. Future work should validate findings using real, anonymized user data. It should introduce controlled noise and imbalance to simulate real-world complexity. It should also apply models to external datasets to assess generalization.

There also may be some overfitting risks. The near-perfect evaluation scores for the random forest model are symptomatic of potential overfitting. Although ensemble models are generally more robust than single estimators, the complexity of the forest (many deep trees) may lead it to memorize patterns in synthetic data that do not generalize. To address this, use cross-validation and regularization (e.g., limiting tree depth). We could also tune hyperparameters like max features per split and introduce dropout-style randomness or feature bagging to improve robustness.

There are some restraints with the data. The dataset was limited to static user traits (age, gender) and aggregate engagement statistics (likes, comments, messages). This restricts the model's ability to capture dynamic emotional patterns that are affected by context, time, and content. Future work could focus on temporal features such as the time of day the user is on social media, session duration, or behavior changes over time. Moreover, by analyzing the content that the user is consuming such as the sentiment or topic on messages and comments, there could be deeper understanding of how the subject matter can influence emotional states. Network metrics could also be included, which would include social graph features like friend count, reciprocity, and clustering. Incorporating richer features could enable more nuanced emotional prediction and reveal temporal or causal dynamics in emotional expression.

Finally, there are ethical considerations. Including demographic features like age and gender in predictive models risks amplifying societal biases or infringing on privacy. Even in hypothetical datasets, it is important to recognize the potential for harm if these models were applied in real scenarios. Future iterations of this work must conduct fairness audits to assess bias in predictions across demographic subgroups and adopt privacy-preserving techniques such as differential privacy or federated learning.

Future Work

Several research extensions arise from the limitations above, specifically real data collection. By partnering with platforms or academic institutions to collect anonymized emotional self-reports alongside usage data, meaningful data can be extracted. Combining behavioral, textual, visual, and physiological data (e.g., wearable sensors) may yield more holistic emotional profiling. Rather than one-size-fits-all classifiers, adaptive models that learn a user's emotional baseline and deviations over time could increase accuracy and utility. Tracking emotional shifts over time such as changes after life events or platform usage breaks could offer predictive insights for mental health monitoring or content recommendation.

This project demonstrates that machine learning models can detect patterns between online behavior and emotional expression with high fidelity at least under idealized conditions. The performance of random forests in particular reveals a rich and learnable structure in behavioral data. Although we were able to produce promising results on synthetic data, it is eminent that real-world application requires careful attention to data quality, fairness, and privacy. We encourage future researchers to build on this work with a focus on ethical practices and practical deployment for user well-being.

References

Frost, Jim. "Principal Component Analysis Guide & Example." *Statistics By Jim*, 29 Jan. 2023, statisticsbyjim.com/basics/principal-component-analysis/.

Olumide, Shittu. "A Complete Guide to Model Evaluation Metrics." *Statology*, 23 Jan. 2025, www.statology.org/complete-guide-model-evaluation-metrics/.