

RESEARCH

Open Access



# Context-aware prediction of active and passive user engagement: Evidence from a large online social platform

Heinrich Peters<sup>1\*</sup>, Yozen Liu<sup>2</sup>, Francesco Barbieri<sup>2</sup>, Raiyan Abdul Baten<sup>3</sup>, Sandra C. Matz<sup>1</sup> and Maarten W. Bos<sup>4</sup>

\*Correspondence:  
hp2500@columbia.edu

<sup>1</sup> Columbia University, New York, USA

<sup>2</sup> Snap Inc., Santa Monica, USA

<sup>3</sup> University of South Florida, Tampa, USA

<sup>4</sup> Carnegie Mellon University, Pittsburgh, USA

## Abstract

The success of online social platforms hinges on their ability to predict and understand user behavior at scale. Here, we present data suggesting that context-aware modeling approaches may offer a holistic yet lightweight and potentially privacy-preserving representation of user engagement on online social platforms. Leveraging deep LSTM neural networks to analyze more than 100 million Snapchat sessions from almost 80,000 users, we demonstrate that patterns of active and passive use are predictable from past behavior ( $R^2=0.345$ ) and that the integration of context features substantially improves predictive performance compared to the behavioral baseline model ( $R^2=0.522$ ). Features related to smartphone connectivity status, location, temporal context, and weather were found to capture non-redundant variance in user engagement relative to features derived from histories of in-app behaviors. Further, we show that a large proportion of variance can be accounted for with minimal behavioral histories if momentary context is considered ( $R^2=0.442$ ). These results indicate the potential of context-aware approaches for making models more efficient and privacy-preserving by reducing the need for long data histories. Finally, we employ model explainability techniques to glean preliminary insights into the underlying behavioral mechanisms. Our findings are consistent with the notion of context-contingent, habit-driven patterns of active and passive use, highlighting the value of contextualized representations of user behavior for predicting user engagement on online social platforms.

**Keywords:** Context-aware machine learning, User engagement, Active and passive use, LSTM, Snapchat

## Introduction

Few technologies have shaped the world as drastically as the advent of the internet and the rise of online social networks. As people's lives are increasingly mediated through online platforms competing for users' attention and striving to maximize user engagement, the ability to understand and predict user behavior has become vitally important. This capacity forms the basis of powerful personalization technologies that adapt to individual users' needs in order to provide a better user experience and ultimately generate revenue. The impact of such technologies can hardly be overestimated as they

are deeply embedded in the design—and will continue to be at the core of the business model—of online platforms that billions of people use on a daily basis.

The capability to predict user behavior at increasing levels of granularity, however, comes at a cost. Not only does it require increasingly sophisticated infrastructure and software, it can also compromise users' privacy [1–4] as the prevailing approach to modeling user behavior involves the creation of user-profiles and long histories of past behavior to predict future behavior [5], often across platforms [6]. While the design of predictive technologies tends to be agnostic to theoretical insights from the behavioral sciences, their application is closely aligned with the catchphrase that past behavior is the best predictor of future behavior [7, 8]. However, the availability of fine-grained behavioral user data also opens the door for modeling approaches aimed at deeper psychological constructs, such as personality [9–14] or habits [15–17].

In this paper, we propose that a context-aware approach to user modeling can increase the performance of predictive models while deepening our understanding of online social behavior. That is, in building contextualized, holistic representations of user engagement, predictive models can be made more efficient, more privacy-preserving, and more aligned with behavioral theory than current approaches. Specifically, we employ LSTM neural networks [18, 19], which excel at representing the recurrent behavioral patterns characteristic of media and technology habits [15–17, 20, 21], to explore how context-aware models incorporating smartphone connectivity status, location, temporal context, weather, and socio-demographic context can aid the prediction of active and passive user behaviors in online social platforms.

### **Background and related work**

Recognizing that user engagement is a multi-faceted phenomenon, past research has often differentiated between active and passive user behaviors. Active use includes behaviors that revolve around social interactions and the creation of content, while passive use includes behaviors that revolve around content consumption [22–26]. The distinction between active and passive use is of both theoretical and practical relevance. While prior work has examined the impact of active versus passive use on psychological outcomes such as well-being [22, 27], little is known about habitual, context-contingent patterns of active and passive use on social platforms. At the same time, the ability to predict active and passive use is important for businesses because the two modes of user engagement distinctively impact user experience and revenue. For example, it is important for online social platforms to host a variety of user-generated content in order to keep users socially engaged [25], but passive use has become increasingly relevant because ad revenue is closely tied to content consumption on all major platforms as content is interspersed with ads. Relatedly, the recent rise of short-form video content has strongly affected how people interact with online social platforms, and the context-contingent and potentially habitual nature of these novel forms of user behavior is currently not well understood.

The field of psychology has produced a rich body of work regarding the effects of context on human behavior. This includes research on ecological psychology [28–30], person-situation interactions [31–33], the conceptualization and measurement of situational cues [34–37], as well as media and technology habits [38–40]. With regard to

the prediction of user engagement, the literature on media and technology habits provides the most suitable theoretical framework connecting context and online behavior. Media and technology habits are learned routines that are contingent on contextual cues and emerge after repeated media use [20, 21, 38]. Habit formation is initially driven by reward learning, where the pleasure associated with content consumption, or the experience of social interaction and recognition, act as rewards. Once a habit is formed, the initial motivations become less important, and the learned response occurs automatically when triggered by associated contextual cues [40]. For example, a user might habitually scroll through social media when riding the bus to work or open an instant messaging app when they hear a notification sound. Generally, contextual cues can be conceptualized as specific context features that precede habitualized behaviors [40, 41]. Relevant cues include “technical cues” that are features of a medium or technology itself (e.g., notifications, buzzes, sounds) [39], but also more subtle situational triggers, such as the time of the day, properties of the environment, events and activities, or internal states [42–46].

Previous empirical research indicates that more than half of media consumption can be classified as habitual [47] and that repeated social media use is strongly habitual, with frequent users reporting higher degrees of automaticity [40]. Habitual use has also been found to be the most important predictor of media use for social sharing [48] and news consumption [49]. In line with these findings, recent research has shown that user activity on social media apps follows predictable patterns. For example, user engagement on Snapchat follows habitual patterns over time, enabling new neural network architectures specifically designed to represent cyclical patterns [15]. Similarly, fine-grained in-app action sequences follow predictable patterns that can be characterized as habitual [16]. However, with the exception of time, the role of context has largely been overlooked in studies linking habits and user engagement.

While there is a lack of previous research investigating how active and passive use relate to contextualized habits, the computer science literature on context-aware modeling offers valuable insights linking user activity and context. The notion of context-awareness - the “ability of a mobile user’s applications to discover and react to changes in the environment” [50] - shares its emphasis on context with the habits literature but focuses more on applications than on behavioral theory [e.g., 51, 52]. Recently, there has been a strong movement towards utilizing machine learning in context-aware technology to model user behavior and tailor applications to individual users’ needs [e.g., 53, 54, 52, 55–57]. For instance, novel methods have been developed to predict users’ activities [58, 59], locations [58, 60], and mental states [56, 61–63], as well as interactions with mobile applications [15, 52, 64–66]. Similarly, context-aware recommender systems have received attention in past research [54, 67]. A wide array of context features has proven useful for the purpose of context-aware modeling, including spatial and temporal context [64, 65], and device context [52], but also more indirect factors such as weather [68]. Despite these advancements, it has been pointed out that the collection, processing, and modeling of various context features remains a challenge, especially in mobile settings [69].

Notably, past work has not always been consistent with regard to its definition of context. For example, the literature on context-aware computing tends to operate on a

broader notion of context compared to the behavioral science literature. The behavioral science literature usually defines context as the sum of environmental conditions or situational circumstances under which behavior occurs, explicitly juxtaposing context-characteristics and person-characteristics or inner states as distinct (and sometimes competing) determinants of behavior [e.g., 70]. This definition is aligned with the behavioral sciences' focus on individual behavior as the main object of study. The computer science literature, on the other hand, tends to include person characteristics such as preferences, psychological traits, and even biometric processes in their definitions of context [e.g., 71]. This perspective is aligned with a technology-centric view in which user characteristics are part of the context in which a technological system is studied. Since the present research aims to contribute not only to the technical literature on context-aware computing but also to the behavioral science literature, it employs the former definition of context, highlighting the dichotomy between environmental factors and person characteristics. Similarly, while past research has sometimes characterized preceding behavior as context [17, 72, 73], the present paper operates on a more narrow definition of context as the sum of environmental factors, focusing on geographic areas with certain socio-demographic properties, weather, time, and locations or places. This approach captures a comprehensive view of spatial context, including factors that are relatively constant and mostly outside users' control (socio-demographic context and weather), along with those that more closely represent users' decisions and mobility habits (location visits and connectivity status).

### **Present research**

The present research explores how context-aware models representing habitual patterns of behavior can aid the prediction of active and passive user behaviors. We analyze data from Snapchat, a major instant messaging and online social platform with close to 400 million daily active users [74]. In the US, the majority (65%) of 18–29 year-olds use Snapchat, most of them at least daily [75]. The app provides ample opportunities for active and passive user behaviors. A core functionality of the Snapchat app is messaging, allowing users to exchange written notes (chats), photos and videos (Snaps), or other multimedia content found on the app. Messaging is inherently social and therefore constitutes a prime example of active use. Other examples of active use include the Lens feature, which allows users to creatively edit photos and videos before sharing them, and the Stories feature allowing to post multimedia status updates that can be viewed by some or all of a user's friends. Examples of passive use include the Discover Feed and the Spotlight Feed, which provide a curated collection of multimedia content that users can engage with. The large user base, the high frequency of interaction, and the design of the platform allowing for a diverse set of active and passive user behaviors make Snapchat an ideal setting to study the prediction of user engagement.

Based on data from over 79,000 Snapchat users, we investigate the extent to which the integration of contextual information, including smartphone connectivity status, location, temporal context, weather, and socio-demographic context, can lead to more accurate, efficient, and privacy-preserving models of user engagement. Using deep LSTM neural networks [18, 19], we first show that - consistent with the idea of habit-driven user engagement—active and passive use follow predictable patterns over time

( $R^2=0.345$ ). Second, highlighting the context-contingent nature of media and technology habits, we demonstrate that the model can be improved by adding different types of context features ( $R^2=0.522$ ). Third, we show that very short sequences can lead to satisfactory predictive performance if momentary context is considered ( $R^2=0.442$ ). Finally, we identify the specific behavioral and contextual variables that are driving predictions in order to gain preliminary insights into the underlying behavioral mechanisms.

## Methods

### Sample and data collection

We utilized archival behavioral user data from a sample of frequent Snapchat users collected through the Snapchat app as part of regular business operations. Frequent users were defined as users who used the app every day for the six-month period before data collection. Among the users who met this criterion, we drew a random sample ( $N=100,000$ ;  $N=79,175$  after cleaning; details described below;  $M_{age} = 25.07$ ;  $SD_{age} = 7.24$ ; 56.31 % female) and obtained their behaviors on the Snapchat app during the 30-day period from July 6th to August 5th, 2021. The dataset contained 105,636,289 sessions (1334 sessions per user on average) which were distributed across 21,604,570 session hours (273 session hours per user on average). A session was defined as the interval between opening and closing the Snapchat app, and a session hour was defined as a full hour in which at least one session has occurred.

All behavioral user data were obtained from in-app event logs indicating that a user had interacted with a specific feature of the app. The set of behavioral features included the number of sessions, time spent on the app, and users' interactions with different features of the app (e.g., chats, Snaps, Stories, ads, Discover, Spotlight, creative tools, and Lenses), split by the type of interaction (view, send, create), and the source of the interaction (e.g., subscriptions, feed, reply; see SI 1). The total number of these behavioral features before further processing was 28 (see SI 2.1).

The set of context features included socio-demographic context, weather, temporal context, location visits, and connectivity status. This feature set was chosen to enable the models to learn a comprehensive representation of spatial context, including macro features that are relatively stable and largely beyond the users' control (e.g., socio-demographic features and weather at the ZIP code level), as well as more fine-grained features that reflect users' choices and mobility habits to a greater extent (e.g., location visits, connectivity status). In order to enrich the behavioral data with socio-demographic and weather data, the GPS coordinates of each session were mapped to the corresponding ZIP code [76]. Exact GPS coordinates were discarded in order to preserve users' privacy. We then used the ZIP code to collect weather data and socio-demographic census data for the area a user was located in during a given session. Weather data was collected from OpenWeatherMap [77]. We used OpenWeatherMap's historical weather API to collect hourly weather data for ten raw features for each ZIP code: temperature, perceived temperature, atmospheric pressure in hPa, humidity in percent, minimum temperature, maximum temperature, wind speed in meter/sec, wind direction in degrees, cloudiness in percent, and categorical weather descriptors (e.g., sunny, rain, snow, fog, extreme, etc.). Socio-demographic data were obtained from the official website of the US Census [78]. We captured 19 features belonging to several important socio-demographic

categories, including socio-economic status, racial composition, age distribution, gender distribution, and marital status at the ZIP code level. We also obtained location visitation data for users who agreed to share their location. To preserve users' privacy, the locations were represented as 11 high-level categories, including events, travel, education, nightlife, residence, food/beverage, shops/services, arts/entertainment, outdoors/recreation, other, and missing. Location features were operationalized as the maximum probability score produced by a location classifier for each location category in each given session hour. The exact location names were not used, making it impossible to infer individual users' exact location visits from the processed data. Additionally, we obtained connectivity status (Wi-Fi access, mobile data) as well as the temporal context (hour of the day, day of the week, day of the month, day of the year, time since the previous session) from the app usage logs (see SI 2.2).

The protocol was approved by the Columbia University IRB (AAAU2607). All methods were carried out in accordance with relevant guidelines and regulations. The study was exempt from the informed consent requirement by the Columbia University IRB because only anonymized archival data were used. While the research relies on proprietary data that cannot be shared openly, the code made available on this project's OSF page (<https://osf.io/nkfhz>) provides a detailed picture of the key properties of the dataset and data processing.

#### **Data preprocessing and operationalizations**

To facilitate training, hyperparameter tuning, and model evaluation, we split the data into three distinct datasets: a training set, a validation set, and a test set. The data was split at the person level, such that all records associated with any individual were assigned to only one of the three datasets. The training data consisted of 500,000 focal session hours, while the validation and test set each consisted of 50,000 focal session hours. Importantly, each focal session hour was associated with a history of up to 100 preceding session hours, each representing hundreds of sessions throughout the 30-day observation period. The target variable, user engagement, was operationalized as the ratio of active-use scores and passive-use scores. Active-use scores were calculated as weighted averages of event counts indicating behaviors associated with active use. This includes the number of chat messages sent, the number of direct Snaps created or sent, and the number of Stories posted by a user. The different input scores were weighted using min-max-transformations such that they equally contributed to the active-use score, irrespective of their absolute frequencies. This approach ensures that relatively rare actions that require higher levels of effort and engagement (e.g., posting a story) would not be drowned out by more frequent but effortless actions (e.g., sending a message). Passive-use scores were defined as the number of Stories viewed. This includes Stories posted by friends, as well as curated Stories viewed through the Discover and Spotlight feeds. These definitions map onto the distinction between social active use and passive use as presented in the Passive and Active Facebook Use Measure (PAUM) [79].

Because user engagement was analyzed at the hourly level, we aggregated the raw behavioral event logs to hourly count metrics for all 28 relevant behavioral features. Since the behavioral count data tended to be power-law distributed, we transformed the hourly count data in several ways: First, we removed data points in the top 0.001

quantiles of each feature distribution in order to remove the most extreme outliers, which are likely caused by technical glitches or bots. Second, we normalized the feature values by dividing hourly counts by hourly active time on the app while also retaining the original non-normalized feature space. This was done to not only measure user activity in absolute terms but also how active users were relative to the time they spent on the app, which can be interpreted as a measure of intensity. Third, we log-transformed the combined feature set to produce more balanced distributions—again, while retaining the original non-log-transformed feature space. Additionally, we generated 15 derived features, such as composite scores of active behaviors (including chats, Snaps, and Stories created or shared by a user), composite scores of passive behaviors (Stories consumed through the Discover, Spotlight, and friends' Story feeds), as well as several ratio and difference scores: the ratio and difference of chat messages sent and received, the ratio and difference of direct Snaps sent and received, the ratio and difference of Stories posted and viewed, as well as ratio and difference of active and passive use scores. Through this process, we obtained an overall space of 127 behavioral features (see SI 2.1).

Similar to behavioral features, context features were aggregated to the hourly level. For ZIP-code-level features (i.e., weather and socio-demographic census features), we calculated time-weighted averages (by active time on the app) if a user had visited more than one ZIP code within one hour. For location features, we cast each of the location categories as a feature and used the maximum probability score (per category, per hour) obtained from a location classifier. Missing values were imputed with zeros as the location classifier was not able to pick up on a relevant location in these cases. Connectivity status was derived as the ratio of the number of sessions involving mobile data usage and the number of sessions involving Wi-Fi access for each hour. It was not necessary to further aggregate the five temporal context features since none of them were captured at a more granular level than the hourly level. Overall, our model included 56 context features (see SI 2.2).

Finally, all numerical (interval and ratio scaled) features were normalized using min-max transformations, such that the final range for each feature was limited to values between 0 and 1 on the training set. Categorical (nominally and ordinally scaled) features were one-hot-encoded (dummy coded), increasing the number of weather features from 10 to 19. All transformations involving distribution parameters were fitted on the training set and then applied to the validation set and the test set using the distribution characteristics of the training set. The final feature set included 183 preprocessed predictors (see SI 2).

### **Modeling**

In distinction from much of the behavioral science literature on user engagement and media and technology habits, we employ a predictive modeling approach [80–84]. The prediction of user engagement was framed as a multivariate time series regression problem, where each feature represented a temporal sequence of behaviors or context events, and the target variable was a continuous variable representing active and passive use. A neural network architecture that is particularly well suited to represent recurrent patterns in multivariate time series is the LSTM architecture [18, 19]. LSTM neural network models are a type of recurrent neural network (RNN), that is able to preserve

information from previous steps more effectively by combining hidden states from previous time steps and current inputs before feeding them through a series of gates (input gate, forget gate, output gate), determining which information is passed on across time steps. In contrast to standard RNN models, which are limited to shorter time series, LSTM neural networks are able to represent the temporal order of widely separated events and temporally extended patterns in noisy input sequences and effectively extract information captured in the temporal distance between events in long sequences [19, 85]. In other words, LSTM models are designed to learn recurrent patterns over time and across long input sequences. LSTM models play an important role in a wide variety of applications, including time series prediction [86, 87], anomaly detection [88], speech recognition [89], and business process management [90, 91]. In the present project, we leverage the capacity of LSTM neural networks to make predictions from a large number of long time series.

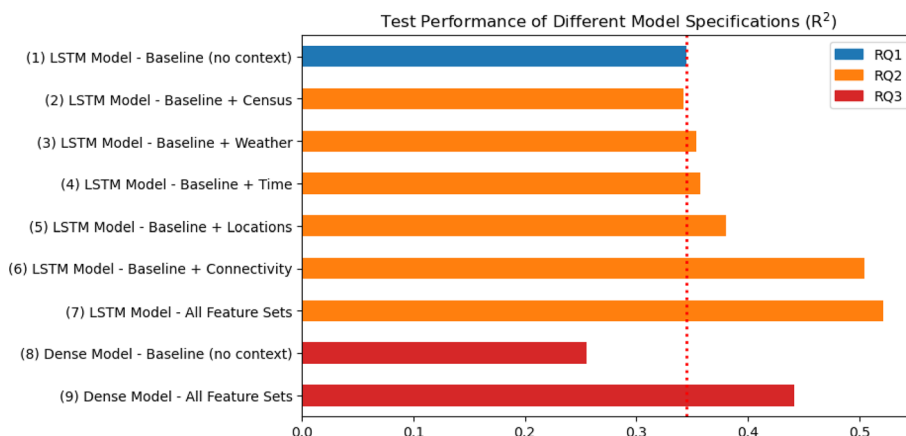
More concretely, we utilized a network architecture of stacked LSTM layers, where the output of each LSTM layer was fed into the next LSTM layer. The output of the last LSTM layer was fed into a final dense layer which was connected to a 1-neuron output layer with a linear activation function. All dense layers utilized rectified linear unit (relu) [92] activation functions, and all LSTM layers utilized hyperbolic tangent (tanh) [93] activation functions. We used Bayesian optimization [94] to tune the depth of the neural network and the dimensionality of the layers. We also tuned the learning rate and the regularization strength with dropout and recurrent dropout. The hyperparameter search space spanned LSTM branch depths of 1–4 layers, tapered LSTM layer dimensions [32, 64, 128, 256] (depending on the depth of the network), top layer dimensions [32, 64, 128], dropout and recurrent dropout values [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6], as well as learning rates [0.01, 0.001, 0.0001]. The batch size was set to a value of 2048 across all experiments. A full overview of the hyperparameter space can be found in Table SI 3. To ensure that performance differences between the different model specifications were indeed caused by differences in the information content of the feature sets and not merely by a mismatch between the dimensionality of the feature space and network configuration, we re-tuned the hyperparameters for each learning task. Hyperparameter tuning was performed with Bayesian optimization [94] using three randomly selected starting points and 100 search iterations per model. Models were trained over 50 epochs with early stopping and a patience parameter value of 5. After training, we selected the model state resulting in the lowest Root Mean Squared Error (RMSE) on the validation set. Generalized model performance was assessed by re-fitting the model with the best hyperparameter configuration and evaluating its predictions on the test set. To ensure the robustness of these performance estimates, we repeated the procedure ten times for each model. An overview of the hyperparameter settings for each model can be found in SI 4.

## Results

### Predicting active and passive use from histories of in-app user behaviors (RQ1)

To test whether active and passive use can be predicted from users' histories of in-app behaviors, we trained an LSTM model using all 127 behavioral feature sequences as inputs (Model 1). We found an  $R^2$  coefficient of 0.345, meaning that the model was able





**Fig. 1** Overview of the predictive performance of different model specifications. RQ1: Model trained on behavioral histories only. RQ2: Model specifications used to assess the performance increment due to different sets of context features, including socio-demographic context (Model 2), weather (Model 3), temporal context (Model 4), location visits (Model 5), network connectivity status (Model 6), and all context features (Model 7). RQ3: Predictive performance of dense neural network models trained on cross-sectional data. Model 8 was trained on only behavioral data from t-1. Model 9 was trained on behavioral data from t-1 and context features from t0. For more details, please see SI 5

to explain 34.5% of the variance in active-passive-use scores, which is considerably better than a naive baseline of  $R^2=0$ . The result indicates that patterns of past behavior are predictive of active and passive use at a given point in time, which is consistent with the idea of habit-driven patterns of engagement. A graphical representation of the results can be found in Fig. 1 and detailed values can be found in SI 5 (Model 1).

**Effects of context-awareness on model performance (RQ2)**

In order to assess whether context-awareness can contribute to the prediction of user engagement, we employed LSTM neural networks in conjunction with ablation study techniques. In particular, we defined six additional model specifications that gave the model access to different subsets of context features and allowed us to observe how model performance changed in response to the addition/removal of context features. We used the previous model consisting of only behavioral sequences as a baseline. We then added each of the five individual brackets of context features (socio-demographic context, weather, temporal context, location visitations, connectivity status) to the baseline model in order to assess the increase in predictive performance associated with each of the different types of context features independently. Finally, we added all context features at once in order to assess the overall increase in predictive performance. An overview of the learning tasks, feature sets, and results can be found in Fig. 1 and SI 5.

The results show that weather (Model 3,  $R^2=0.354$ ) and temporal context (Model 4,  $R^2=0.357$ ) were associated with small performance improvements (2.5% and 3.5%, respectively), while location information (Model 5,  $R^2=0.380$ ) was associated with a moderate performance improvement of 10% over the baseline (Model 1). With a performance improvement of 46%, connectivity status (Model 6,  $R^2=0.504$ ) added by far the most incremental explained variance. Adding census data to the baseline model did not result in improved performance (Model 2,  $R^2=0.342$ ). Adding all features to the

baseline simultaneously (Model 7,  $R^2=0.522$ ) resulted in the best model and the largest performance improvement (51%), indicating that the different feature sets capture some non-redundant information. The model also dramatically outperformed models trained exclusively on context features (see Section 3.3). These findings suggest that complementing behavioral features with context features, especially connectivity status, can substantially improve the performance of models predicting user engagement. The findings are consistent with the idea of context-contingent patterns of engagement.

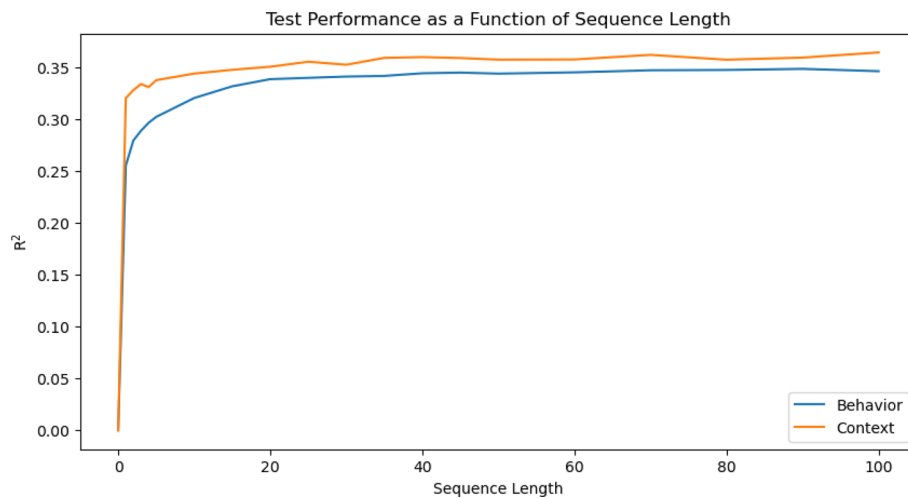
### Effects of context in models trained on truncated sequences (RQ3)

In order to assess the history-dependence of the predictive models, we conducted an ablation study in which we manipulated the length of the sequence that was accessible to the LSTM model. Sequence lengths ranged from only 1–100 time steps, representing anything from just one session up to 100 session hours distributed across 28 days. We first conducted this analysis for behavioral sequences and then separately for sequences of context data. In the case of behavioral histories, the longest sequence ranged from  $t-101$  to  $t-1$ . No information about  $t_0$  was included in order to ensure that no information regarding user activity in the target session leaked into the predictor set. In the case of context histories, the longest sequence ranged from  $t-100$  to  $t_0$ , such that the last data point in the sequence represented the momentary context in  $t_0$ . Model performances are depicted in Fig. 2.

As expected, longer time series were generally associated with higher model performance. This effect was more pronounced in the case of behavioral histories compared to context histories. For example, a model trained on behavioral histories of sequence length 1 captured 73% of the maximum performance, while a model trained on context histories of sequence length 1 captured 88% of the maximum performance. Models utilizing a sequence length of 50 yielded at least 99% of the performance reached by models trained on the maximum length sequences with 100 time steps. The results indicate that temporal histories capture information about user engagement and that a relatively large share of variance in user engagement can be explained with very short sequences, particularly in the case of context predictors.

The previously trained models rely on sequences of historical data to predict user engagement. While the findings suggest that even short sequences of behavioral and contextual data offer moderate levels of predictive performance, we designed an additional experiment to test the incremental predictive power of context data at the extreme. Specifically, we trained a series of models based on minimal historical data, including only cross-sectional behavioral features from  $t-1$  and momentary context features from  $t_0$ . Given that this procedure eliminates the sequential structure of the input data, we shifted the analytical strategy from the LSTM architecture to a feed-forward neural network architecture.

The results show a substantial drop in performance when only behavioral features from  $t-1$  are provided to the model. With an  $R^2=0.256$ , the performance of the model incorporating cross-sectional behavioral data (Model 8) was 26% lower compared to the baseline model incorporating behavioral histories of sequence length 50 and 51% lower compared to the best model incorporating the full feature set at a sequence length of 50. The cross-sectional model trained on behavioral and context features (Model 9), on the



**Fig. 2** Model performance as a function of sequence length for behavioral histories (blue) and context histories (orange). The curves show diminishing returns to adding additional time steps to the model. Models built on behavioral histories benefit more from long time series compared to models built on context histories

other hand, showed a predictive performance of  $R^2=0.442$ . This was 15% lower than that of the best model from the previous analysis but 28% higher than the performance of the baseline model with full access to behavioral histories and no context features. Most importantly, predictive performance was 73% higher than the performance of the model built on only cross-sectional behavioral data. The results demonstrate that the prediction of user engagement - even in the extreme case of fully truncated time series - works reasonably well if context is considered. For a visual comparison, see Fig. 1.

#### Model explainability analyses (RQ4)

While deep neural networks excel at making predictions from complex data, they do not provide much insight regarding the strength and directionality of specific feature-target associations by default. In order to open up the black box and explain how the models arrive at their predictions, we employed Shapley values, a popular model explainability technique [95]. Shapley values are additive feature importance measures that quantify the contribution of each feature with respect to each individual predicted value. Intuitively, this is achieved by calculating the average marginal contribution of each feature across all possible subsets of features. The marginal contribution of a feature is defined as the difference in predicted values between a model including the focal feature and a model withholding the feature (e.g., by replacing feature values with unrelated values sampled from a background dataset). Because the effect of withholding a feature is contingent on other features, the marginal effects need to be computed and averaged for all possible subsets of features. This procedure is repeated for each feature in the overall feature set. In the present project, we used the Shapley Additive Explanations (SHAP) [95] Python library. Due to their computational complexity, SHAP values were generated for the dense neural network model incorporating the full feature set (Model 9) on a subsample of 20,000 data points from the test set.

In order to estimate general feature importance scores, we computed the arithmetic mean of the absolute Shapley values for each feature across all data points in the test sample. To analyze the directionality of the association, we analyzed the relationship between feature values and Shapley values. For comparison, we also analyzed the relationships between feature values and actual target values on the test set using Pearson's correlation coefficients for metric features and point-biserial correlation coefficients for one-hot encoded categorical features. A positive correlation between feature values and Shapley values indicates that higher feature values are, on average, associated with higher predicted values in the model. A positive correlation between feature values and actual target values would indicate linear feature-target associations irrespective of their representations in the model. The results are shown in Fig. 3.

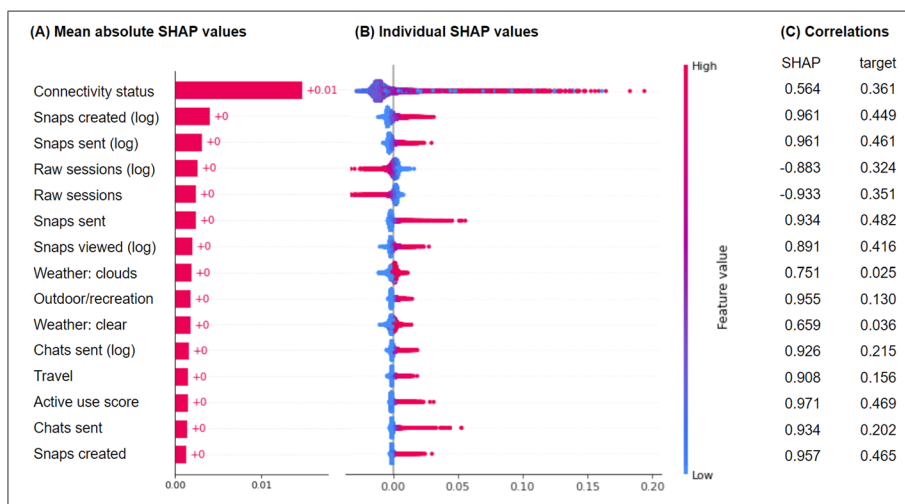
Connectivity status and location visits in  $t_0$ , along with in-app behaviors related to active use in  $t-1$ , were particularly important predictors. Connectivity status showed by far the greatest average SHAP value. Mobile data usage was overall associated with higher predictions of active use. A closer inspection of the SHAP values revealed that, in some cases, feature values for high mobile data usage were associated with negative SHAP values. This indicates that contrary to the general trend, mobile data usage sometimes decreases predictions of active use, meaning the model represents interaction effects or nonlinearities. The idea of interactive relationships is also reflected in another finding: despite its high overall impact on predicted values, the correlations between connectivity status and SHAP values, as well as actual target values, are lower than those of the behavioral predictors. This is likely the case because correlations, as opposed to SHAP values, only represent linear relationships.

## Discussion

### Interpretation of results

The results show that user engagement is highly predictable and that context-awareness can substantially improve the performance of models predicting patterns of active versus passive use compared to models trained on behavioral histories alone, especially when minimal temporal histories are considered. While the baseline model explained 34.5% of the variation in active-passive use scores, the context-aware model, including all context features, explained 52.2%, amounting to a performance improvement of 51%. The performance improvement was even more pronounced in the models trained on truncated sequences, where the context-aware model scored 73% higher than the behavioral baseline model. These findings underscore the value of contextualized representations of user behavior for predicting user engagement on social platforms. Additionally, our findings are aligned with previous research suggesting that user behavior on Snapchat follows habitual patterns over time [15, 16]. More broadly, our findings are consistent with the notion that social sharing [48, 96] and media consumption [17, 40, 46, 47, 49] are habit-driven.

In line with the idea that specific behavioral outcomes are best predicted by constructs that match their level of analysis, we found that connectivity status (specific to the individual user's device) and location (an individual user's immediate surroundings) were associated with the greatest uptick in predictive performance, while less specific context factors operating at larger geographic scales - such as weather or



**Fig. 3** Mean absolute SHAP values (panel A), distribution of individual SHAP values colored by underlying feature values (panel B), and correlations between feature values and SHAP values, as well as feature values and target values (panel C). Connectivity status, online behaviors in the previous session, location visits, and weather features showed the largest impact on predictions

socio-economic context - were of little help. The small effect of weather on predictive performance is consistent with previous work suggesting that weather has only small effects on psychological states and behavioral outcomes [97].

The finding that behavioral and contextual histories from the recent past were more predictive than information from the distant past is plausible, considering that the effects of predictors captured at a more recent point in time are more direct and less likely to be washed out by noise. Additionally, more recent predictor states contain information about previous states, such that adding previous states to a model will only marginally improve model performance insofar as they contain non-redundant information. At the same time, the finding reflects a general limitation of LSTM models, which place more emphasis on more recent data points due to information decay [98]. The fact that the effect was more pronounced in the case of behavioral histories, however, speaks to the relevance of longer-term behavioral patterns, contrasted with more temporally bounded context effects.

The results of the model explainability analysis using SHAP values are aligned with the findings of the previous ablation study, indicating that connectivity status and location features are particularly important context factors. Mobile data usage was generally associated with active use. This is plausible given that people might limit passive use in order to save mobile data, but more importantly, people likely encounter more situations that are worth sharing when they are out and about as compared to when they are at home or at work, where they would typically have Wi-Fi access. The latter point is especially plausible given the image-based nature of Snapchat. The finding also ties in with previous work suggesting that social sharing is habitual [48]. This interpretation is aligned with the finding that locations related to outdoor/recreation and travel are also associated with higher levels of active use in the present study. The fact that high feature values were not uniformly associated with high SHAP

values is consistent with the idea that the neural network model represents interactive relationships and that those interactive relationships are of particular importance for the prediction of user engagement.

Taken together, our research extends previous work in several ways: First, we add nuance to the study of user engagement by focusing on the distinction between active and passive use, which had not been studied through the lens of media and technology habits before [22–26]. Second, we use highly granular large-scale user data to explore the effects of a wide range of objectively captured context factors, contrasting their predictive contributions against each other and against behavioral features. In doing so, our study pushes the boundaries of both research on media and technology habits and research on context-aware modeling of user engagement. Finally, our study demonstrates how integrating objective sensing and location-tracking can enrich behavioral research. By taking research out of the lab and into the real world, we glean insights that advance behavioral theory while also speaking to people's naturalistic experiences in everyday life, thus enhancing the ecological validity and applicability of the findings.

#### **Limitations and directions for future research**

Our work has several limitations that have the potential to stimulate future research. First, the present study squarely follows a predictive modeling approach. This approach contrasts with explanatory methods, typically used in the behavioral and social sciences, that employ statistical inference to uncover causal relationships between variables [80–84]. While our methodology allows us to investigate the predictive contributions across a range of theoretically informed feature sets based on highly granular longitudinal user data, it inherently prioritizes prediction at the expense of causal explanation. Recognizing this distinction, future research could build on our findings to derive more targeted research questions and focus on specific causal mechanisms.

Second, while we interpret the presence of recurrent predictable patterns of user engagement through the lens of media and technology habits, our findings also relate to other theoretical orientations such as ecological psychology [28–30] and person-situation interactions [31–33]. Although our research was not designed as an explicit test of theoretical predictions from either framework, our findings are aligned with these theoretical traditions. Specifically, they add to the existing literature by examining the predictive power of context in the setting of online social platforms and by focusing on the prediction of user engagement from highly granular behavioral data based on recurrent patterns. By offering empirical support for the existence of interactive relationships between user engagement and certain context features, our findings corroborate the main premise of person-situation interactions. However, additional research is needed to disentangle these relationships in a more explicit manner.

Third, while our study leverages an exceptionally granular and comprehensive feature set, we encountered several tradeoffs with respect to the operationalization of context. For instance, integrating network-related features to capture user relationships [60] could have enriched our analysis by offering deeper insights into the predictive utility of social contexts. Additionally, the location visits, a critical component of our analysis,

were inferred from noisy data. This limitation was compounded by our decision to use broad location categories in order to safeguard user privacy, potentially at the expense of achieving maximum predictive performance. Furthermore, past research has highlighted the value of incorporating self-reported measures to capture the psychological experience of context [34–36], a methodological approach we could not employ due to our reliance on objective behavioral data. Future work should continue to integrate objectively measured cues with self-report measures to gain deeper insights into psychological mechanisms.

Fourth, we computed active-passive use scores as the ratio of active and passive behaviors within a given session hour. This decision was necessary given that active and passive use in isolation are highly correlated with overall app use, rendering the prediction task trivial and less interesting from a theoretical perspective. Our goal was to create a measure that captures users' relative tendency towards either of the usage modes. Future work could avoid this tradeoff by shifting the level of analysis to that of individual user behaviors.

Fifth, past research has sometimes treated preceding behaviors or activities as a specific type of context [34, 45]. We fundamentally agree with this perspective but decided to emphasize the dichotomy of spatial context and behavioral histories as the distinction maps more cleanly onto current approaches in user modeling, which has often relied exclusively on past behavior [99, 100].

Finally, our research was conducted on a single online social platform with unique affordances [101] and in a specific cultural context [102]. Additionally, we only included data from adult users, which excludes a substantial share of Snapchat's user base. As Snapchat shares important features, such as the specific modalities for active and passive use (Chat, Stories, Spotlight, Discover) with other online social platforms, we believe it is likely that our findings would apply to related settings. However, the extent of generalizability would have to be examined in future research.

### **Implications**

Our findings indicate that active and passive use follow predictable patterns over time and that context-awareness can substantially improve model performance. This has implications for the use of predictive models and their integration into the design of online platforms. Being able to accurately predict user behavior through context-aware modeling could facilitate the dynamic allocation of computational resources to adaptively support different modes of user engagement. For example, knowing when an app user likes to consume rather than create content would enable a platform to allocate resources to a recommender system and preload relevant content. Similarly, knowing when a user is likely to actively create and share content would allow them to pre-allocate resources to the camera process and load the user's favorite filters. At the same time, a better understanding of the relationships between context factors and user engagement could facilitate the design of features that encourage active or passive use by making specific context factors salient to the user. For example, a system that draws users' attention to the social affordances of their environment might encourage active use and discourage passive use in certain settings. While the present paper focuses on active and

passive use, psychologically informed context-aware models can be used to personalize a wide array of features related to users' momentary states, including physiological (e.g., being hungry or tired), social (e.g., being with family, friends, or work colleagues), or psychological states (e.g., mood and affect). Such personalizations that take into account context-contingent states will ultimately improve user experience and create new opportunities for online platforms to generate revenue.

Importantly, our findings can help decision-makers navigate resource and privacy tradeoffs. The ablation study that was used to analyze performance as a function of the length of data histories shows strongly diminishing returns to longer time series, especially in excess of sequence lengths of 50 time steps. While the specific relationship is highly dependent on the predictive task and the model at hand, our procedure constitutes a simple method to assess a model's dependence on data histories that can easily be used to inform decisions in practice to preserve resources and limit privacy concerns. Additionally, our results show that the integration of context data can compensate for the truncation of behavioral histories, offering an opportunity to fade out models relying on historical data in favor of more ephemeral modeling approaches representing momentary user states. Accordingly, in situations where maximizing predictive accuracy is not an absolute imperative, it might be recommended to use short-term behavioral data in conjunction with easily obtainable context features, such as connectivity status, to build models that are privacy-preserving and cheap to run. Smaller, less computationally expensive models can, in turn, unlock additional privacy benefits—for example, it would be possible to run models locally on a mobile device without sharing sensitive data.

Taken together, our findings challenge the narrative that long-term user data are needed in order to make accurate predictions of user engagement and provide personalized services [100, 103, 104]. Instead, our results show that it is possible to trade a relatively small reduction in predictive performance for a considerable reduction in data requirements and computational resources. Practitioners should be aware of this fact in order to make informed decisions and balance considerations of predictive performance, resources, and privacy when implementing predictive models in real-world settings.

## Conclusion

In conclusion, our findings show that active and passive use follow predictable patterns over time and emphasize the importance of context in predicting user engagement on social platforms. Consistent with theories on media and technology habits, the combination of behavioral data and context information leads to a substantial increase in the predictive performance of user engagement. The present paper presents the first rigorous large-scale assessment of the predictive power of a wide array of context factors with respect to active and passive use on mobile social platforms. Our findings have potential implications for managerial, engineering, and design decisions - specifically with respect to the dynamic allocation of computational resources to adaptively support different modes of user engagement, privacy-preserving modeling of user behavior, and the design of features that encourage active or passive use. Furthermore, our work provides a starting point for a broader research program investigating additional context factors, different types of user engagement, and underlying psychological mechanisms.



## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40537-024-00955-0>.

Supplementary file 1.

### Acknowledgements

We thank Ron Dotsch and Joseph B. Bayer for their thoughtful feedback.

### Author contributions

HP, YL, and MWB developed the research idea and the research design. HP and YL collected and analyzed the data. HP wrote the paper. YL, FB, RAB, SCM, and MWB provided feedback on research design, analyses, and writing.

### Funding

The research was funded by Snap Inc.

### Availability of data and materials

The datasets generated and/or analyzed during the current study are not publicly available due to their highly sensitive, proprietary nature. Requests can be granted upon approval by Snap Inc.

## Declarations

### Ethics approval and consent to participate

The data collection was approved by the Columbia University IRB (AAAU2607). All methods were carried out in accordance with relevant guidelines and regulations. The study was exempt from the informed consent requirement by the Columbia University IRB because only anonymized archival data was used.

### Consent for publication

Not applicable.

### Competing interests

YL, FB, and MWB are employe by Snap Inc. The research was funded by Snap Inc. HP, RAB, and SCM declare no potential competing interests.

Received: 18 August 2023 Accepted: 11 July 2024

Published online: 08 August 2024

## References

1. Casadesus-Masanell R, Hervas-Drane A. Strategies for managing the privacy landscape. *Long Range Plan.* 2020;53(4):101949.
2. Chellappa RK, Sin RG. Personalization versus privacy: an empirical examination of the online consumer's dilemma. *Inform Technol Manag.* 2005;6(2):181–202.
3. Cloarec J. The personalization-privacy paradox in the attention economy. *Technol Forecast.* 2020. <https://doi.org/10.1016/j.techfore.2020.120299>.
4. Gal-Or E, Gal-Or R, Penmetsa N. The role of user privacy concerns in shaping competition among platforms. *Inform Syst Res.* 2018;29(3):698–722.
5. Eke CI, Norman AA, Shuib L, Nweke HF. A survey of user profiling: state-of-the-art, challenges, and solutions. *IEEE Access.* 2019;7:144907–24.
6. Veiga MH, Eickhoff C. A Cross-Platform Collection of Social Network Profiles. In: Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, New York, NY: ACM, 2016. <https://doi.org/10.1145/2911451.2914666>.
7. Aarts H, Verplanken B, van Knippenberg A. Predicting behavior from actions in the past: repeated decision making or a matter of habit? *J Appl Soc Psychol.* 1998;28(15):1355–74.
8. Ouellette JA, Wood W. Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. *Psychol Bull.* 1998;124(1):54–74.
9. Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. *Proc Natl Acad Sci.* 2013;110(15):5802–5.
10. Kulkarni V, et al. Latent human traits in the language of social media: an open-vocabulary approach. *PLoS ONE.* 2018;13(11): e0201703.
11. Matz SC, Kosinski M, Nave G, Stillwell DJ. Psychological targeting as an effective approach to digital mass persuasion. *Proc Natl Acad Sci.* 2017;114(48):12714–9.
12. Youyou W, Kosinski M, Stillwell D. Computer-based personality judgments are more accurate than those made by humans. *Proc Natl Acad Sci USA.* 2015;112(4):1036–40.
13. Peters H, Matz S. Large language models can infer psychological dispositions of social media users. arXiv preprint, 2023. <https://doi.org/10.48550/arXiv.2309.08631>.
14. Peters H, Cerf M, Matz S. Large language models can infer personality from free-form user interactions. arXiv preprint. 2024. <https://doi.org/10.48550/arXiv.2405.13052>.
15. Chowdhury FA, et al. CEAM: the effectiveness of cyclic and ephemeral attention models of user behavior on social platforms. *Proc Int AAAI Conf Web Soc Media.* 2021;15:117–28.

16. Liu Y, Shi X, Pierce L, Ren X. Characterizing and forecasting user engagement with in-app action graph: a case study of snapchat. arXiv preprint, 2019. <https://doi.org/10.48550/arXiv.1906.00355>.
17. Peters H, Bayer JB, Matz SC, Chi Y, Vaid SS, Harari GM. Social media use is predictable from app sequences: using LSTM and transformer neural networks to model habitual behavior. arXiv preprint, 2024. <https://doi.org/10.48550/arXiv.2404.16066>.
18. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw.* 2015;61:85–117.
19. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* 1997;9(8):1735–80.
20. Schnauber-Stockmann A, Naab TK. The process of forming a mobile media habit: results of a longitudinal study in a real-world setting. *Media Psychol.* 2019;22(5):714–42.
21. Tokunaga RS. Media use as habit, in the international encyclopedia of media psychology. Hoboken: John Wiley & Sons Ltd; 2020. p. 1–5.
22. Escobar-Viera CG, et al. Passive and active social media use and depressive symptoms among United States adults. *Cyberpsychol, Behav, Soc Netw.* 2018;21(7):437–43.
23. Hemmings-Jarrett K, Jarrett J, Blake MB. Evaluation of user engagement on social media to leverage active and passive communication. In: 2017 IEEE International Conference on Cognitive Computing (ICCC), Honolulu, HI, 2017. <https://doi.org/10.1109/IEEE.ICCC.2017.24>.
24. Khan ML. Social media engagement: what motivates user participation and consumption on YouTube? *Comput Human Behav.* 2017;66:236–47.
25. Pagani M, Hofacker CF, Goldsmith RE. The influence of personality on active and passive use of social networking sites. *Psychol Mark.* 2011;28(5):441–56.
26. Trifiro BM, Gerson J. Social media usage patterns: research note regarding the lack of universal validated measures for active and passive use. *Soc Media Soc.* 2019. <https://doi.org/10.1177/2056305119848743>.
27. Verduyn P, Gugushvili N, Kross E. Do social networking sites influence well-being? The extended active-passive model. *Curr Dir Psychol Sci.* 2022;31(1):62–8.
28. Barker RG. Explorations in ecological psychology. *Am Psychol.* 1965;20(1):1–14.
29. Heft H. Ecological psychology in context: James Gibson, Roger Barker, and the legacy of William James's radical empiricism. New York: Psychology Press; 2016.
30. Lobo L, Heras-Escribano M, Travieso D. The history and philosophy of ecological psychology. *Front Psychol.* 2018;9:2228.
31. Rauthmann JF. Person-situation interactions. In: Carducci BJ, C. Nave S., and R. E. Riggio, editors. *The wiley encyclopedia of personality and individual differences*. Hoboken: John Wiley & Sons; 2020.
32. D. C. Funder. Persons, behaviors and situations: an agenda for personality psychology in the postwar era. *J Res Pers.* 2009;43(2):120–6.
33. Fleeson W, Nofhle E. The end of the person-situation debate: an emerging synthesis in the answer to the consistency question. *Soc Pers Psychol Compass.* 2008;2(4):1667–84.
34. Rauthmann JF, Sherman RA. Conceptualizing and measuring the psychological situation. In: Wood Dustin, Read Stephen J, Harms PD, Slaughter Andrew, editors. *Measuring and modeling persons and situations*. San Diego: Elsevier Academic Press; 2021. p. 427–63.
35. Rauthmann JF, et al. The situational eight DIAMONDS: a taxonomy of major dimensions of situation characteristics. *J Pers Soc Psychol.* 2014;107(4):677–718.
36. Schoedel R, Kunz F, Bergmann M, Bemmman F, B-hner M, Sust L. Snapshots of daily life: situations investigated through the lens of smartphone sensing. *J Pers Soc Psychol.* 2023;125(6):1442–71.
37. D. C. Funder. Taking situations seriously: the situation construal model and the riverside situational Q-sort. *Curr Dir Psychol Sci.* 2016;25(3):203–8.
38. Bayer J, Anderson I, Tokunaga R. Building and breaking social media habits. *Curr Opin Psychol.* 2022;45:279–88.
39. Bayer JB, LaRose R. Technology habits: progress, problems, and prospects. In: Verplanken B, editor. *The psychology of habit*. Cham: Springer International Publishing; 2018. p. 111–30.
40. Anderson IA, Wood W. Habits and the electronic herd: the psychology behind social media's successes and failures. *Consum Psychol Rev.* 2021;4(1):83–99.
41. Gardner B, Rebar AL. Habit formation and behavior change. In: Braddick Oliver, editor. *Oxford research encyclopedia of psychology*. Oxford: Oxford University Press; 2019.
42. LaRose R. The problem of media habits. *Commun Theory.* 2010;20(2):194–222.
43. Naab TK, Schnauber A. Habitual initiation of media use and a response-frequency measure for its examination. *Media Psychol.* 2016;19(1):126–55.
44. Verplanken B, Wood W. Interventions to break and create consumer habits. *J Public Policy Mark.* 2006;25(1):90–103.
45. Wood W, Neal DT. A new look at habits and the habit-goal interface. *Psychol Rev.* 2007;114(4):843.
46. Bayer JB, Campbell SW, Ling R. Connection cues: activating the norms and habits of social connectedness. *Commun Theory.* 2016;26(2):128–49.
47. Wood W, Quinn JM, Kashy DA. Habits in everyday life: thought, emotion, and action. *J Pers Soc Psychol.* 2002;83(6):1281.
48. Choi M, Toma CL. Understanding mechanisms of media use for the social sharing of emotion: the role of media affordances and habitual media use. *J Media Psychol Theor Methods Appl.* 2021. <https://doi.org/10.1027/1864-1105/a000301>.
49. Diddi A, LaRose R. Getting hooked on news: uses and gratifications and the formation of news habits among college students in an internet environment. *J Broadcast Electron Media.* 2006;50(2):193–210.
50. Schilit B, Theimer M. Disseminating active map information to mobile hosts. *IEEE Netw.* 1994;8(5):22–32.
51. Yurur O, Liu CH, Sheng Z, Leung VCM, Moreno W, Leung KK. Context-awareness for mobile sensing: a survey and future directions. *IEEE Commun Surv Tutor.* 2016;18(1):68–93.
52. Sarker I, Colman A, Han J, Watters P. Context-aware machine learning and mobile data analytics: automated rule-based services with intelligent decision-making. Cham: Springer International Publishing; 2021.

53. Nascimento N, Alencar P, Lucena C, Cowan D. A context-aware machine learning-based approach. In: Proceedings of the 28th Annual International Conference on Computer Science and Software Engineering: IBM Corp, Riverton, NJ, 2018; pp 40–47.
54. Raza S, Ding C. Progress in context-aware recommender systems-an overview. *Comput Sci Rev.* 2019;31(C):84–97.
55. Mijnsbrugge DV, Ongenaes F, Van Hoecke S. Context-aware deep learning with dynamically assembled weight matrices. *Inform Fusion.* 2023. <https://doi.org/10.1016/j.inffus.2023.101908>.
56. Salido Ortega MG, Rodr-guez L-F, Gutierrez-Garcia JO. Towards emotion recognition from contextual information using machine learning. *J Ambient Intell Human Comput.* 2020;11(8):3187–207.
57. Miranda L, Viterbo J, Bernardini F. A survey on the use of machine learning methods in context-aware middlewares for human activity recognition. *Artif Intell Rev.* 2022;55(4):3369–400.
58. Liao D, Liu W, Zhong Y, Li J, Wang G. Predicting Activity and Location with Multi-task Context Aware Recurrent Neural Network, in Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, Stockholm, Sweden: International Joint Conferences on Artificial Intelligence Organization, 2018;3435–3441.
59. Hasan M, Roy-Chowdhury AK. Context aware active learning of activity recognition models. Piscataway, NJ: IEEE; 2015. p. 4543–51.
60. Baten RA, et al. Predicting future location categories of users in a large social platform. *Proc Int AAAI Conf Web Soc Media.* 2023;17:47–58.
61. Lam G, Dongyan H, Lin W. Context-aware deep learning for multi-modal depression detection. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, United Kingdom: IEEE, 2019; pp. 3946–3950.
62. Müller SR, Chen X, Peters H, Chaintreau A, Matz SC. Depression predictions from GPS-based mobility do not generalize well to large demographically heterogeneous samples. *Sci Rep.* 2021;11:14 007.
63. Onim MSH, Rhodus E, Thapliyal H. A review of context-aware machine learning for stress detection. *IEEE Consum Electron Mag.* 2023. <https://doi.org/10.1109/MCE.2023.3278076>.
64. Fan Y, et al. Personalized context-aware collaborative online activity prediction. *Proc ACM Interact, Mobile, Wear Ubiquitous Technol.* 2019;3(4):132:1–132:28.
65. Xia T, et al. DeepApp: predicting personalized smartphone app usage via context-aware multi-task learning. *ACM Trans Intell Syst Technol.* 2020;11(6):1–12.
66. Huang K, Zhang C, Ma X, Chen G. Predicting mobile application usage using contextual information, in Proceedings of the 2012 ACM Conference on Ubiquitous Computing, New York, NY, USA: ACM, 2012;1059–1065.
67. Chen A. Context-aware collaborative filtering system: predicting the user-s preference in the ubiquitous computing environment. In: Strang T, Linnhoff-Popien C, editors. Location- and context-awareness. Berlin, Heidelberg: Springer; 2005. p. 244–53.
68. Nawshin S, Mukta MSH, Ali ME, Islam AKMN. Modeling weather-aware prediction of user activities and future visits. *IEEE Access.* 2020;8:105127–38.
69. Sarker IH, Hoque MM, Uddin MK, Alsanoosy T. Mobile data science and intelligent apps: concepts, AI-based modeling and research directions. *Mobile Netw Appl.* 2021;26(1):285–303.
70. Sansone C, Morf CC, Panter AT. The Sage handbook of methods in social psychology. Thousand Oaks, CA: SAGE; 2004.
71. Gajjar MJ. Context-aware computing. In: Gajjar Manish J, editor. Sensors mobile, computing context-aware. Amsterdam: Elsevier; 2017. p. 17–35.
72. Wood W. Habit in personality and social psychology. *Pers Soc Psychol Rev.* 2017;21(4):389–403.
73. Buyalskaya A, Ho H, Milkman KL, Li X, Duckworth AL, Camerer C. What can machine learning teach us about habit formation? Evidence from exercise and hygiene. *Proc Natl Acad Sci.* 2023;120(17): e2216115120.
74. S. Inc., Snap Inc. Announces Second Quarter 2023 Financial Results, 2023. <https://www.sec.gov/Archives/edgar/data/1564408/000156440823000036/snap-2023725xexx991pressre.htm>.
75. Auxier B, Anderson M. Social Media Use in 2021, 2021. <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/>.
76. Peters H, Marrero Z, Gosling SD. The big data toolkit for psychologists: data sources and methodologies. In: Matz Sandra, Matz Sandra C, editors. The psychology of technology: social science research in the age of big data. Washington, DC: American Psychological Association; 2022. p. 87–124.
77. OpenWeatherMap, 2021. <https://openweathermap.org/history>
78. Census Bureau - Advanced Search, 2021. <https://data.census.gov/advanced>.
79. Gerson J, Plagnol AC, Corr PJ. Passive and active facebook use measure (PAUM): validation and relationship to the reinforcement sensitivity theory. *Pers Individ Differ.* 2017;117:81–90.
80. Hofman JM, Sharma A, Watts DJ. Prediction and explanation in social systems. *Science.* 2017;355(6324):486–8.
81. Hofman JM, et al. Integrating explanation and prediction in computational social science. *Nature.* 2021;595(7866):181–8.
82. Yarkoni T, Westfall J. Choosing prediction over explanation in psychology: lessons from machine learning. *Perspect Psychol Sci.* 2017;12(6):1100–22.
83. Breiman L. Statistical modeling: the two cultures. *Stat Sci.* 2001;16(3):199–231.
84. Shmueli G. To explain or to predict? *Stat Sci.* 2010;25(3):289–310.
85. Gers FA, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM. *Neural Comput.* 2000;12(10):2451–71.
86. Hua Y, Zhao Z, Li R, Chen X, Liu Z, Zhang H. Deep learning with long short-term memory for time series prediction. *IEEE Commun Mag.* 2019;57(6):114–9.
87. Siami-Namini S, Tavakoli N, Siami Namin A. A comparison of ARIMA and LSTM in forecasting time series. In: Proceedings of the 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL: IEEE, 2018; pp. 1394–1401.
88. Malhotra P, Vig L, Shroff G, Agarwal P. Long Short Term Memory Networks for Anomaly Detection in Time Series, *Computational Intelligence*, 2015:7

89. Graves A, Jaitly N, Mohamed A-R. Hybrid speech recognition with Deep Bidirectional LSTM. In: IEEE Workshop on Automatic Speech Recognition and Understanding. Olomouc, Czech Republic: IEEE. 2013;2013: pp. 273–8.
90. Camargo M, Dumas M, Gonz-lez-Rojas O. Learning accurate LSTM models of business processes. In: Hildebrandt T, van Dongen BF, R-glinger M, Mendling J, editors. Business process management. Cham: Springer International Publishing; 2019. p. 286–302.
91. Tax N, Verenich I, La Rosa M, Dumas M. Predictive business process monitoring with LSTM neural networks. arXiv preprint. 2017. <https://doi.org/10.48550/arXiv.1612.02130>.
92. Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines, in Proceedings of the 27th International Conference on International Conference on Machine Learning, Madison, WI, USA: Omnipress, 2010;807–814.
93. Kalman B, Kwasny S. Why tanh: Choosing a sigmoidal function. In: Proceedings of the International Joint Conference on Neural Networks, Baltimore, MD: IEEE, 1992; pp. 578–581.
94. Falkner S, Klein A, Hutter F. BOHB: Robust and efficient hyperparameter optimization at scale. In: Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden: PMLR, 2018; pp. 1437–46.
95. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, Red Hook, NY: Curran Associates Inc., 2017; pp. 4768–77.
96. Bayer JB, Campbell SW. Texting while driving on automatic: considering the frequency-independent side of habit. *Comput Human Behav.* 2012;28(6):2083–90.
97. Harley T. *The psychology of weather*. 1st ed. New York, NY: Routledge; 2018.
98. Chien H-YS, Turek JS, Beckage N, Vo VA, Honey CJ, Wilke TL. Slower is better: revisiting the forgetting mechanism in LSTM for slower information decay. arXiv preprint. 2021. <https://doi.org/10.4855/arXiv.2105.05944>.
99. Purificato E, Boratto L, De Luca EW. User modeling and user profiling: a comprehensive survey. arXiv preprint. 2024. <https://doi.org/10.4855/arXiv.2402.09660>.
100. Webb GI, Pazzani MJ, Billsus D. Machine learning for user modeling. *User Model User-Adapt Interact.* 2001;11(1):19–29.
101. Bucher T, Helmond A. The affordances of social media platforms. In: Bruges J, Marwick A, Poell T, editors. *The SAGE handbook of social media*. Thousand Oaks, CA: Sage Publications; 2018. p. 233–53.
102. Alsaleh DA, Elliott MT, Fu FQ, Thakur R. Cross-cultural differences in the adoption of social media. *J Res Interact Market.* 2019;13(1):119–40.
103. Jannach D, Resnick P, Tuzhilin A, Zanker M. Recommender systems-beyond matrix completion. *Commun ACM.* 2016;59(11):94–102.
104. Najafabadi MK, Mohamed AH, Mahrin MN. A survey on data mining techniques in recommender systems. *Soft Comput.* 2019;23(2):627–54.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.