Drug and Alcohol Dependence 252 (2023) 110964

Contents lists available at ScienceDirect

Drug and Alcohol Dependence

journal homepage: www.elsevier.com/locate/drugalcdep

Predicting quantity of cannabis smoked in daily life: An exploratory study using machine learning

Ching-Yun Yu [a,*], Yi Shang [a], Tionna M. Hough [b], Anthony L. Bokshan [b], Megan N. Fleming [b], Alison M. Haney [b], Timothy J. Trull [b,**]

[a] Department of Electrical Engineering and Computer Science, University of Missouri, USA
[b] Department of Psychological Sciences, University of Missouri, USA

ARTICLE INFO

Keywords:
Mobile survey
Machine learning
Digital health
Ecological momentary assessment
Cannabis

ABSTRACT

Background: Cannabis use is prevalent in the United States and is associated with a host of negative consequences. Importantly, a robust indicator of negative consequences is the amount of cannabis consumed.
Methods: Data were obtained from fifty-two adult, regular cannabis flower users (3+ times per week) recruited from the community; participants completed multiple ecological momentary assessment (EMA) surveys each day for 14 days. In this exploratory study, we used various machine learning algorithms to build models to predict the amount of cannabis smoked since participants' last report including forty-three EMA measures of mood, impulsivity, pain, alcohol use, cigarette use, craving, cannabis potency, cannabis use motivation, subjective effects of cannabis, social context, and location in daily life.
Results: Our best-fitting model (Gradient Boosted Trees; 71.15% accuracy, 72.46% precision) found that affects, subjective effects of cannabis, and cannabis use motives were among the best predictors of cannabis use amount in daily life. The social context of *being with others*, and particularly with a partner or friend, was moderately weighted in the final prediction model, but contextual items reflecting *location* were not strongly weighted in the final prediction model, the one exception being *not at work*.
Conclusions: Machine learning approaches can help identify additional environmental and psychological phenomena that may be clinically-relevant to cannabis use.

1. Introduction

Cannabis use is prevalent in the United States and is associated with numerous clinically-relevant psychological processes such as affect, impulsivity, and pain (Hasin and Walsh, 2020; Li et al., 2019; Trull et al., 2016; Volkow et al., 2014), as well as a host of negative consequences (e.g., impaired driving; poor work/school performance; Pearson, 2019). Importantly, a robust indicator of negative consequences is the amount of cannabis consumed (Callaghan et al., 2020). Therefore, research into the best predictors of the amount of cannabis consumed could inform future interventions aimed at mitigating negative consequences.

Because individual motives for and subjective experiences of cannabis use may vary across time and contexts, researchers seeking to understand how these processes unfold in daily life often use methods such as *Ecological Momentary Assessment* (EMA; Wycoff et al., 2018). Using smartphones for data collection multiple times per day, EMA allows researchers to assess cannabis use and other psychological phenomena in participants' natural environments, increasing ecological validity. EMA methods are particularly useful for studying substance use behaviors like cannabis use, as this approach mitigates retrospective recall bias by prompting participants to recall events over relatively shorter periods of time (e.g., in the past 15 minutes, past hour, and/or since the prior recording), thus reducing measurement error (Trull and Ebner-Priemer, 2013; Freeman et al., 2023; Mun et al., 2021; Solhan et al., 2009).

EMA research demonstrates that *subjective experiences of cannabis intoxication* may vary within-person based on method of cannabis consumption and various *qualities of the cannabis consumed* (e.g., potency, amount; Li et al., 2019; Okey et al., 2023; Trull et al., 2022). Furthermore, *motives for cannabis use* may differentially predict the subsequent quantity of cannabis one consumes (Bonar et al., 2017) as well as the *subjective experience* of cannabis use (Buckner et al., 2013, 2015;

* Correspondence to: 411 S. 6th Street, Columbia, MO 65201, USA.
** Correspondence to: University of Missouri, 210 McAlester Hall, Columbia, MO 65211, USA.
E-mail addresses: cytbm@missouri.edu (C.-Y. Yu), TrullT@missouri.edu (T.J. Trull).

https://doi.org/10.1016/j.drugalcdep.2023.110964
Received 2 July 2023; Received in revised form 2 September 2023; Accepted 10 September 2023
Available online 14 September 2023

Comulada et al., 2016; Ross et al., 2018; Shrier and Scherer, 2014; Votaw and Witkiewitz, 2021). Additionally, *momentary affect* may predict later cannabis use (Buckner et al., 2012a, 2012b, 2013, 2015; Chakroun et al., 2010) and fluctuate following cannabis use (Buckner et al., 2015; Henquet et al., 2010; Tournier et al., 2003; Tyler et al., 2015), though the direction of association across studies remains unclear. Although research is limited, individuals generally report elevated *impulsivity* on cannabis use days (Ansell et al., 2015; Trull et al., 2016), but the temporal precedence involved in this association is unknown. Finally, though *pain relief* is the most reported motive for medical cannabis use (Boehnke et al., 2019), the literature is inconclusive regarding whether elevated pain levels predict subsequent cannabis use and whether pain decreases following cannabis use (Kuerbis et al., 2019; Li et al., 2019).

The *context* of cannabis use in daily life has been relatively understudied (Denson et al., 2023). Most studies have sampled undergraduate or clinical participants in treatment for cannabis use disorder, only examined who the participant is with at the time of cannabis use, and few have considered the participant's location when using cannabis (Buckner et al., 2012a, 2012b, 2015, 2016; Henquet et al., 2010; Phillips et al., 2018; Sznitman et al., 2022; Treloar Padovano and Miranda, 2018; Tyler et al., 2015). Therefore, a more comprehensive evaluation of the associations between contexts and cannabis use seems warranted.

Traditionally, researchers focus on a limited number of variables to predict important clinical outcomes. Although this approach has the advantage of limiting the number of predictors used in models to those that are believed to be most theoretically relevant to the outcome, in some cases not enough is known about a range of potential predictors of the outcome. Furthermore, initial discovery or exploratory research is often needed to identify promising predictors of the outcome that can later be tested in replication studies (Rubin and Donkin, 2022; Yarkoni and Westfall, 2017). Therefore, to evaluate the relations between a range of putative predictors and cannabis use, we opted to use machine learning algorithms to identify those features most associated with using more versus less cannabis flower during a given smoking episode.

In summary, increased cannabis use is robustly associated with more negative consequences, and daily life studies suggest a range of possible predictors for higher levels of cannabis use. The present study tested various machine learning algorithms to build models to predict the amount of cannabis smoked since participants' last report from momentary measures of mood, impulsivity, pain, alcohol/cigarette use, craving, cannabis potency, substance use motivation, and effects of cannabis. Because cannabis use and problems have been associated with contexts, we collected data on participants' social context and current location. Finally, we sampled adults in the community who were regular cannabis users, as these participants may be more representative of the general population of cannabis smokers.

2. Method

2.1. Participants

Fifty-two recreational cannabis users (mean age of 24.21 years; 49.06% women, 47.17% men, 3.77% other gender; 81.13% European American/White, 15.09% African American/Black, 7.55% Asian American) participated in this study (see Table 1). Inclusion criteria included being between 18 and 50 years old and smoking cannabis flower three or more times per week. Exclusion criteria included using cannabis *primarily* in any form other than combustible flower (e.g., edibles, concentrates), testing positive for using illicit substances other than cannabis (e.g., cocaine, opiates) during the initial visit, previous head trauma involving sustained impairment in mood, attention, or concentration, sought/intending to seek treatment for substance-related problems, or reporting active psychosis or suicidal ideation. Baseline data in Table 1 indicated that the sample included regular cannabis smokers who, on average, reported at least one day of heavy cannabis use (i.e., four grams or more) over the past six months.

Table 1
Descriptive Information.

	Mean	*SD*	Range
Age, gender, race/ethnicity	24.2	7.2	18–47
Participant age (years)	49.06%	0.38	3–4[2]
Percent women	81.13%	3.02	0.5–14.8
Percent European American/White	15.09%	0.58	0.05–2.67
Percent African American/Black	7.55%		
Percent European Asian American	3.83[1]		
Cannabis Use patterns (from baseline	4.17		
assessment)	0.65		
Average frequency of use over last 6 months (CUDIT)			
Maximum amount of cannabis smoked in one day over last 6 months (grams)			
EMA results (14 days)			
Mean grams on smoking days			
Mean number of use days per participant	9.21	3.26	2–15
MR compliance %	90.04	13.98	50–100
RP compliance %	72.69	20.32	9.26–98.08
Number of self-initiated reports	8.08	7.77	0–37

Note.
MR = Morning report. RP = random prompt. [1]Corresponds to consuming cannabis 4 or more times per week. [2]Corresponds to consuming cannabis 2–3 times per week over the past six months.

2.2. Procedure

Individuals were screened by phone for eligibility, and informed consent was obtained before study entry. Participants completed an initial visit, during which they learned to use the smartphone data collection app (*TigerAware*; Morrison et al., 2018, July) and completed baseline questionnaires (compensated $10 for completion). Participants then completed 14 days of EMA, including morning reports upon wakeup, four random prompts scheduled at quasi-random times (i.e., a random prompt within four, three-hour windows from12 noon-12 midnight), and self-initiated reports on their cannabis use. Participants were compensated up to $50 based on their survey compliance.

2.3. EMA measures

2.3.1. Affect

Participants reported their affect in the past 15 minutes using 17 items from the Positive and Negative Affect Schedule Extended Form (PANAS-X; Watson and Clark, 1994) using a Likert scale from 1 (*Very Slightly or Not at All*) to 5 (*Extremely*). Positive and Negative Affect items were averaged, respectively, into two composite predictors (PA and NA).

2.3.2. Impulsivity

Participants reported their impulsivity within the last 15 minutes using four items adapted from the UPPS Impulsive Behavior Scale (Whiteside and Lynam, 2001). One item was selected from each subscale of the UPPS, including Urgency (*acting on a strong impulse*), Lack of Premeditation (*doing something without really thinking it through*), Lack of Perseverance (*giving up easily*), and Sensation-seeking (*doing something for the thrill of it*). Items were rated on a Likert scale from 1 (*Very Slightly or Not at All*) to 5 (*Extremely*). These items were averaged into a composite predictor (IMP).

2.3.3. Alcohol and tobacco use

At each prompt, participants indicated whether they had consumed any alcohol or cigarettes since the last prompt.

2.3.4. Cannabis use

Participants reported whether they consumed cannabis since the last prompt in the self-initiated cannabis use survey and/or in the random reports (which contained a question about use since the last survey completed). If endorsed, follow-up questions were asked, such as how many times they consumed cannabis and *how much cannabis they consumed in grams since the last prompt*. At the orientation session, to help train participants to provide accurate estimates of cannabis flower quantity, participants were presented with several different visual references as aids to estimate grams (or portions of a gram) of cannabis flower. Participants also reported when they finished using cannabis, with the options of “just now,” “15 minutes ago,” “30 minutes ago,” “45 minutes ago,” “1 hour ago,” and “more than 1 hour ago.” Finally, participants reported how they would rate the “strength or potency” of the cannabis they consumed since the last survey, with options of “less strong than,” “similar in strength to,” or “stronger than” what they typically consume.

2.3.5. Motives for cannabis use

When they endorsed cannabis use, participants were asked about their motives for consumption using 12 items from the Marijuana Motives Measure (MMM; Simons et al., 1998), using a Likert scale from 1 (*Strongly Disagree*) to 4 (*Strongly Agree)*.

2.3.6. Cravings for cannabis use

Participants rated (1) how much they craved cannabis and (2) how hard it was to stop thinking about using cannabis in the past 15 minutes using a Likert scale from 1 (*Very Slightly or Not at All*) to 5 (*Extremely*).

2.3.7. Subjective effects of cannabis

Upon endorsing cannabis use, participants reported their subjective experience of intoxication, indicating the extent to which 14 options applied to them (Chait et al., 1988) on a Likert scale from 1 (*Very Slightly or Not at All)* to 5 (*Extremely*).

2.3.8. Pain

Participants reported how much pain they felt in the past 15 minutes on a Likert scale from 1 (*No Pain*) to 10 (*The Worst Possible Pain*).

2.3.9. Location

Participants reported whether their current location was at home, work, a bar/restaurant, outside, another public space, or another kind of location.

2.3.10. Social contexts

Participants reported whether they were with anyone in the past 15 minutes. If they reported they were with someone, they specified whether they were with a partner, a friend, or someone else.

3. Analytic plan

Given the literature on correlates of cannabis use, we selected forty-three attributes/predictors from the Cannabis Use and Random Prompt survey data that focus on mood, impulsivity, pain, alcohol use, cigarette use, craving, cannabis potency, cannabis use motivations, effects of cannabis, and where/with whom participants were (see Table 2). Two ordinal classes for cannabis use were defined based on an examination of cannabis quantity score distributions (0.2 g or less [~50%] and greater than 0.2 g [~50%]) for cannabis used by participants. Our dependent variable represented whether reported cannabis quantity fell within the range of one ordinal class or the other.

The raw data set contained 852 rows and 43 attributes/predictors from 52 participants; each row represents a single EMA observation/survey response in which a participant endorsed having consumed cannabis since the last survey they answered. These data were split into labeled data (813 rows) and unlabeled data (39 rows); unlabeled data were those endorsements of cannabis use without information on the amount used (~5% of data rows). Labeled data were divided into the training and validation sets in the ratio of 6:4. At this stage, we set a random seed to ensure that the data split according to this ratio were consistent in each run. The training set contained 467 rows, while the test set contained 346 rows.

Table 2
Forty-three attributes serving as predictors of the amount of cannabis smoked in daily life, and each attribute’s correlation with amount of cannabis smoked (dichotomous; 0.2 g or less and greater than 0.2 g; **bolded**).

Affect, Pain, and Impulsivity (last 15 minutes)
Positive Affect (*PA;* ***0.156***); Negative Affect (*NA;* ***0.068***); Physical Pain (*Pain;* ***0.190***); Impulsivity (*IMP;* ***0.08***)
Cannabis craving (last 15 minutes)
How much did you crave cannabis? (*Craving1;* ***-0.095***); How hard was it to stop thinking about using cannabis? (*Craving2;* ***-0.176***)
Situation (last 15 minutes) (Yes/No)
Have you been with anyone? (*anyoneYN;* ***0.051***); Who have you been with–partner/spouse? (*with_SOYN;* ***0.172***); Who have you been with–friend/acquaintance? (*with_friend;* ***0.155***)
Location (current) (Yes/No)
Home (*loc_homeYN;* ***0.028***); Work (*loc_workYN;* ***-0.153***); bar/restaurant (*loc_barYN;* ***-0.01***); Outside (*loc_outsideYN;* ***0.030***); Other public place (*loc_otherpubYN;* ***0.013***); Other (*loc_otherYN;* ***0.030***)
Recent other substance use (Yes/No)
Alcohol (*alcYN;* ***0.031***); Cigarettes (*cigYN;* ***0.038***)
Potency of the cannabis you consumed
(*mj_potency*; **0.141**)
Less strong than what I typically consume (0); Similar in strength to what I typically consume (1); Stronger than what I typically consume (2).
Motivations for cannabis use (I used cannabis....) ***[1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree])***
Because I like the feeling (*cm.feeling;* ***0.020***); Because it makes social gatherings more fun (*cm.social;* ***0.140***); To be liked by others (*cm.liked;* ***0.030***); Because it helps me when I feel depressed or nervous (*cm.depressed;* ***0.017***); To understand things differently (*cm.understand;* ***0.036***); Because it’s fun (*cm.fun;* ***-0.021***); Because it improves parties and celebrations (*cm.parties;* ***0.075***); So I won’t feel left out (*cm.left.out;* ***0.073***); To forget about my problems (*cm.problems;* ***0.151***); To expand my awareness (*cm.expand;* ***0.056***); Because it gives me a pleasant feeling (*cm.pleasant;* ***0.022***); To get high (*cm.high;* ***0.084***)
Effect/Subjective response (Since using cannabis, how much have you felt....) ***(1 = not at all, 5 = extremely)***
Things around me seem more pleasing than usual (*ce.pleasing;* ***0.054***); I feel as if something pleasant had just happened to me (*ce.pleasant;* ***-0.094***); I have difficulty in remembering (*ce.difficulty;* ***-0.154***); I feel a very pleasant emptiness (*ce.emptiness;* ***-0.127***); My mouth seems very dry (*ce.dry;* ***0.00***); Some parts of my body are tingling (*ce.tingling;* ***-0.020***); My movements seem slower than usual (*ce.slower;* ***-0.013***); I notice that my heart is beating faster (*ce.heart;* ***0.012***); My thoughts seem to come and go (*ce.thoughts;* ***-0.096***); I notice my hands are shaking (*ce.hands;* ***0.014***); I have an increasing awareness of my bodily sensations (*ce.sensations;* ***0.038***); I am more suspicious of others (*ce.suspicious;* ***0.054***); high (*ce.high;* ***0.175***);

The mean values for each attribute/predictor in the training and test set were calculated for each participant, and any missing values in both the training set and the test set were replaced with the mean values (i.e., mean imputation). For both the training set and the test set, each participant’s rows were up-sampled, if necessary, to match the number of rows for the participant with the highest count, ensuring all users had an equal number of rows and that no participants’ data was weighted more heavily than others. The participant with the highest number of rows in the training set had 35, while the participant with the highest count in the test set had 24 rows. A random seed was set to ensure that randomly selected rows were added for those participants with fewer rows of data. After up-sampling, the training set included 52 participants multiplied by 35 rows (total of 1820 rows), and the test set contained 52 participants multiplied by 24 rows (total of 1248 rows).

3.1. Model training steps

All model building, training, and evaluation was implemented using the ***RapidMiner*** software platform (Mierswa and Klinkenberg, 2018). We chose five frequently used algorithms: *Decision Trees, Random Forest, Gradient Boosted Trees, Support Vector Machine, and H2O’s Deep Learning*. For each, the training process consisted of two phases. In the *first phase*, the optimal parameters were determined. In the *second phase*, during model learning, it was observed that most confidence values did not resemble probabilities and could sometimes be severely skewed. To address this, an improved Platt scaling method was used to rescale the confidence produced by the model (Platt, 1999). This required a labeled calibration dataset that the model had not been trained on. Hence, the training set was divided into a 9:1 ratio before training to create the calibration dataset. In finding the optimal parameters, cross-validation with 10 folds was used. The cross-validation involved a training section and a testing section, and the improved Platt scaling method was employed to rescale the confidence generated by the model. The trained model was then used in the testing process to assess its performance. As there was no need for the process of finding the optimal parameters, the improved Platt scaling method was used in both the first and second phases.

To evaluate performance, we calculated the classification accuracy and produced the confusion matrix for each model. After inputting the trained model, training data, and validation data, the explain operator yielded the explained prediction results and the weights of the attributes in the output. The explained prediction results were displayed as a table highlighting the attributes that most strongly support each prediction result, while the local attribute weights were identified by the correlation of a neighboring set of data points generated by the operator. Although the relationship between attributes and predictions may be highly nonlinear globally, the local linear relationship was sufficient to explain the prediction results. In addition, this operator could compute model-specific but model agnostic global attribute weights. If the true labels of the test data are known, then all supporting and contradictory local explanations positively affect the weights of correct and incorrect predictions, respectively. In contrast, if the true labels are unknown, the global weights use only the supporting local weights.

4. Results

Table 2 presents all 43 features that were included as well as each feature's bivariate correlation with the ordinal class of amount of cannabis used in parentheses. Negative associations suggested a feature was associated with cannabis use at lower quantities, while positive associations suggested a feature was associated with higher quantities of use. These results index the strength of association of each feature with the ordinal class and the valence of the association. The five features with the ***highest, positive association*** with cannabis amount were: physical pain (0.19); felt high (0.175); with a partner (0.172); PA (0.156); and with a friend (0.155). The five features with the ***highest, negative association*** with cannabis amount were: difficulty stopping thinking about cannabis (−0.176); difficulty in remembering (−0.154); at work (−0.153); feel a very pleasant emptiness (−0.127); and thoughts seem to come and go (−0.096).

Table 3 shows the *test accuracy, true positive rate (sensitivity), false positive rate, false negative rate, true negative rate (specificity), recall, and precision* for all five models. The test accuracy score is an index that measures the number of correct predictions made by a model in relation to the total number of predictions made, while precision refers to the likelihood that a positive prediction is indeed correct. In terms of performance, the highest test accuracy we obtained came from Gradient Boosted Trees at 71.15%, followed by Support Vector Machine at 70.27%. In general, test accuracy above 70% is considered good. The Gradient Boosted Trees algorithm also produced the highest precision (72.46%).

Fig. 1 displays attributes that have the greatest impact on model predictions. The *top ten attributes* were: (1) I have difficulty in remembering (*ce.difficulty*); (2) pain; (3) negative affect (*NA*); (4) positive affect (*PA*); (5) Because it makes social gatherings more fun (*cm.social*); (6) Because I like the feeling (*cm.feeling*); (7) cannabis potency; (8) To be liked by others (*cm.liked*); (9) I feel as if something pleasant had just happened to me (*ce.pleasant*); and (10) I feel a very pleasant emptiness (*ce.emptiness*).

The *weakest ten attributes* were: (1) At Other location (*loc_otherYN*); (2) At a bar/restaurant (*loc_barYN*); (3) At Other public place (*loc_otherpubYN*); (4) My mouth seems very dry (*ce.dry*); (5) Outside (*loc_outsideYN*); (6) I am more suspicious of others (*ce.suspicious*); (7) Alcohol (*alcYN*); (8) Because it's fun (*cm.fun*); (9) Some parts of my body are tingling (*ce.tingling*); and (10) Cigarettes (*cigYN*).

5. Discussion

We sought to identify the strongest correlates of higher quantity of cannabis use in daily life among 52 regular cannabis smokers. Based on earlier investigations, we pre-selected 43 attributes from our smartphone survey data that focused on affect, impulsivity, pain, alcohol use, cigarette use, craving, cannabis potency, substance use motivation, effects of cannabis, and context (where/with whom participants were). We then trained five models on the dataset, and our best fitting model, Gradient Boosted Trees, correctly predicted the class of amount of cannabis smoked in approximately 71% of cases, which is appreciably better than classification by the baseline model (Decision Tree; ~53% accuracy).

As presented in Fig. 1, among the attributes that had the greatest impact on model predictions were affective state (PA/NA), pain, endorsement of a range of cannabis effects (e.g., difficulty in remembering; feel as if something pleasant had just happened to me; feel a very pleasant emptiness), and motivations to use cannabis (e.g., it makes social gatherings more fun; because I like the feeling; to be liked by others).

These features are somewhat consistent with existing understandings of the factors that influence cannabis use. Concerning affective state and pain, several motivational theories of substance use are relevant. For example, the affective-motivational model of drug addiction suggests that individuals may use more of a substance to deal with painful or aversive states (Baker et al., 2004). Ostensibly, by using more cannabis when these states are present, negative affect and pain may be reduced, serving to negatively reinforce higher amounts of use. On the other hand, cannabis may be used more to enhance positive affect, and there is evidence that positive affect may be elevated in anticipation of the effects of cannabis use (Buckner et al., 2015; Chakroun et al., 2010).

Concerning cannabis motives, our results converge with the evidence that cannabis use may be used to improve social interactions and to conform to others using cannabis or increase the probability of fitting in. Regarding subjective effects, the most highly weighted predictor of use of higher cannabis amounts was *lower scores* on the item, "I have difficulty in remembering." This may suggest individuals are more likely to use greater amounts when their memory is not impaired by cannabis use, whereas individuals who report some degree of memory problems

Table 3
Performance Comparison of Machine Learning Models in Predicting Larger Amount of Cannabis Smoked (i.e., > 0.20 g).

MODEL	Accuracy	True Positive Rate (SENS)	False Positive Rate	False Negative Rate	True Negative Rate (SPEC)	Recall	Precision
Decision Tree	53.29%	83.05%	73.40%	16.95%	26.60%	83.05%	50.36%
Random Forest	61.10%	65.93%	37.54%	34.07%	62.46%	65.93%	61.16%
Deep Learning (H2O)	67.31%	65.08%	30.70%	35.42%	69.30%	65.08%	65.53%
Support Vector Machine	70.27%	71.19%	30.55%	28.81%	69.45%	71.19%	67.63%
Gradient Boosted Trees	71.15%	62.88%	21.43%	37.12%	78.57%	62.88%	72.46%

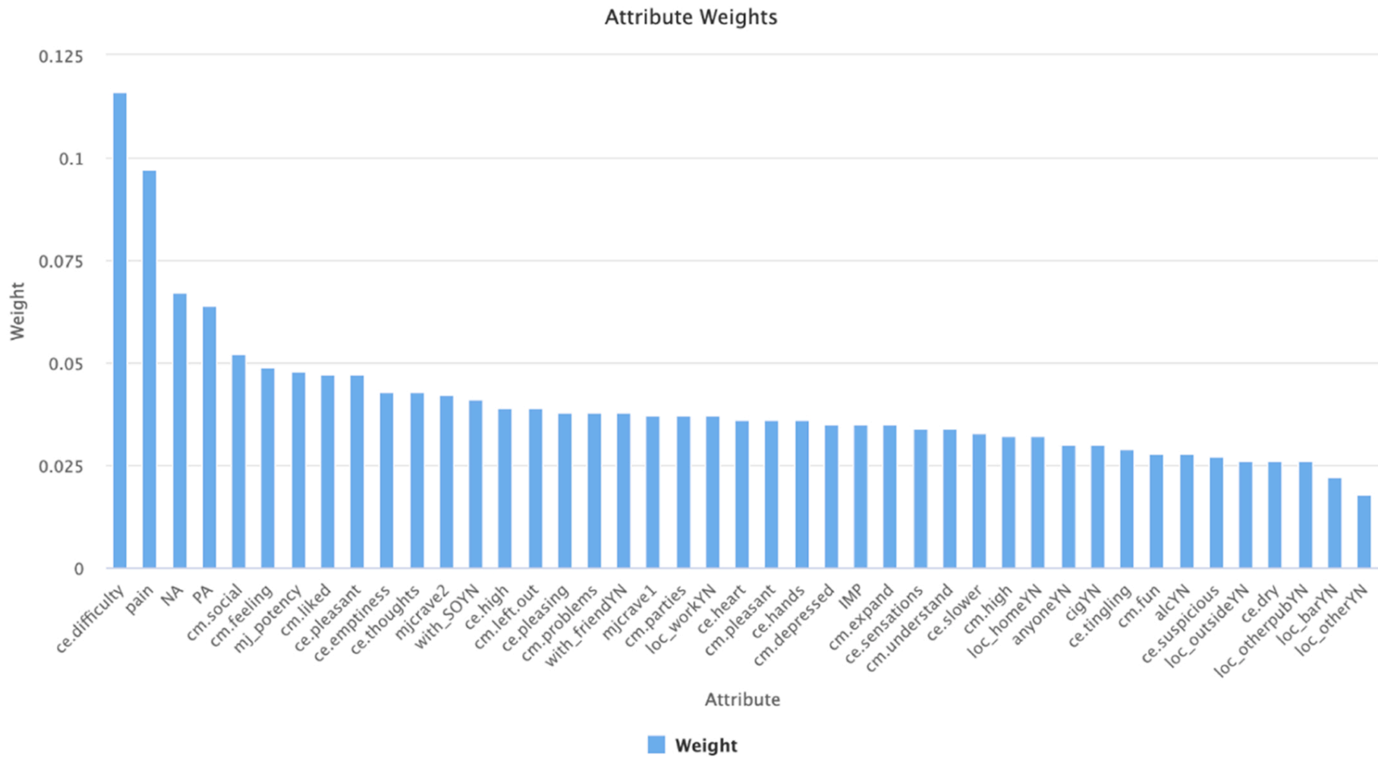

Fig. 1. The attribute weights for Gradient Boosted Trees. **Note:** see Table 2 for definition of abbreviations.

with lower levels of cannabis use may abstain from heavier use to avoid worsening its effects. Other subjective effect predictors were also related to *decreased amounts of cannabis used*, including pleasant subjective effects (I feel as if something pleasant had just happened to me; I feel a very pleasant emptiness), suggesting that individuals were more likely to report pleasant effects when using lower amounts of cannabis. This could be seen as evidence for moderation of larger amounts of cannabis if one is already in a relatively positive state.

We expected that contextual factors might be among the strongest predictors of cannabis consumed. On the one hand, items reflecting *being with others*, and particularly with a partner or friend, were moderately weighted in the final prediction model. This finding is consistent with several previous studies that have found more cannabis use occurs when in the presence of others (e.g., Buckner at al., 2013; Phillips et al., 2018). However, contextual items reflecting *location* were not strongly weighted in the final prediction model, the one exception being *not at work*. Previous studies have suggested that being at home or at a friend/relative's house was significantly associated with both craving and cannabis use (e.g., Gunn et al., 2021; Shrier and Rhoads, Burke, et al., 2014). One possible interpretation for these results is that *being with others* is more important than *where you are with others* in terms of using more cannabis.

5.1. Limitations and implications

As mentioned earlier, a machine learning approach to classifying individuals has the advantages of evaluating a range of predictors. These approaches not only allow researchers to evaluate predictors using more focal, hypothesis-driven approaches, but also enable the discovery of less-studied but, perhaps, important correlates. In addition, these approaches can analyze large data sets efficiently and can be improved over time to increase accuracy of classification. However, these approaches' results are influenced by the representativeness of the training data and the sample itself.

Although our results revealed the strongest predictors of amount of cannabis used in daily life, there are limitations. First, our sample was relatively small and consisted of regular cannabis smokers (i.e., 3+ times per week) from the community. It is possible that a different pattern of results would be obtained from undergraduate or clinical samples, lighter cannabis smokers, or those who use cannabis via different delivery methods (e.g., concentrates, edibles). Secondly, because we were focusing on the total amount of cannabis smoked in each episode, our analyses essentially focus on event-level associations between cannabis amounts and our predictors. Therefore, we are unable to determine the temporal precedence of cannabis smoking and these predictors/features. These relations are best interpreted as concurrent associations; we cannot assume that cannabis smoking led to changes in these features or that these features influenced the amount consumed. Future research might also examine feature ratings endorsed before cannabis smoking episodes.

In addition, it is worth noting that there were multiple episodes of cannabis use in approximately 44% of the days analyzed for this study. There was a small but significant negative correlation ($r = -0.17$, $p = 0.002$) between lagged grams and concurrent grams consumed on these days. This may suggest that individuals may have "self-regulated" on these days by smoking less cannabis if they smoked a relatively greater amount in a previous session. It may be possible that this influenced some results indicating negative associations between features and cannabis amounts.

Our best performing model improved upon the baseline model; however about 29% of smoking sessions were misclassified in terms of the amount of cannabis smoked. It is possible that the accuracy of models could be improved with additional predictors/features that were not included in this study. For example, while we assessed whether the participant was alone or with others, we did not assess whether anyone in the participants' presence was smoking cannabis. Another limitation is that we relied on participants' reports of how much cannabis they consumed. Research suggests that participants' estimates of the amount of cannabis used may be unreliable (e.g., Prince et al., 2018).

Despite these limitations, we briefly highlight several possible implications for reducing the amount of cannabis smoked per day. On average, more cannabis was consumed when participants reported relatively higher levels of NA and pain. When indicated, interventions might include training in alternative coping skills or the use of other medical interventions, as opposed to using cannabis. Another set of significant predictors of increased cannabis use include motives, specifically using cannabis to make social gatherings more fun, to be liked by others, and because of the feeling. Cognitive behavior therapy skills aimed at identifying/avoiding triggering situations, problem solving, and refusal skills may be helpful in mitigating the influence of these motives for cannabis use.

6. Conclusions

Our best-fitting model found that affect, pain, subjective effects of cannabis, and cannabis use motives were among the best predictors of cannabis use in daily life. This is consistent with research findings that individuals may use cannabis for motives or reasons, such as experiencing a subjective or affective change. However, additional research is needed to understand how other individual and contextual differences may directly or indirectly influence the use of cannabis.

Role of funding source

Funding support received from the National Institute of Health: R01 AA027824 (Trull) and T32 AA013526 (Haney, Hough).

CRediT authorship contribution statement

YS and TT contributed to the design of the study and the collection of data, and CYY processed the data and implemented the analyses. CYY produced the first draft of the manuscript, edited by YS, TH, AB, MF, AH, and TT, and approved by all authors.

Declaration of Competing Interest

YS and TT are co-founders of TigerAware LLC who created the software platform used to collect data, but they do not receive any compensation from the company. Other authors have no conflicts of interests to disclose.

References

Ansell, E.B., Laws, H.B., Roche, M.J., Sinha, R., 2015. Effects of marijuana use on impulsivity and hostility in daily life. Drug Alcohol Depend. *148*, 136–142. https://doi.org/10.1016/j.drugalcdep.2014.12.029.

Baker, T.B., Piper, M.E., McCarthy, D.E., Majeskie, M.R., Fiore, M.C., 2004. Addiction motivation reformulated: An affective processing model of negative reinforcement. Psychol. Rev. *111* (1), 33–51. https://doi.org/10.1037/0033-295X.111.1.33.

Boehnke, K.F., Gangopadhyay, S., Clauw, D.J., Haffajee, R.L., 2019. Qualifying conditions of medical cannabis license holders in the United States. Health Aff. *38* (2), 295–302. https://doi.org/10.1377/hlthaff.2018.05266.

Bonar, E.E., Goldstick, J.E., Collins, R.L., Cranford, J.A., Cunningham, R.M., Chermack, S.T., Blow, F.C., Walton, M.A., 2017. Daily associations between cannabis motives and consumption in emerging adults. Drug Alcohol Depend. *178*, 136–142. https://doi.org/10.1016/j.drugalcdep.2017.05.006.

Buckner, J.D., Crosby, R.D., Wonderlich, S.A., Schmidt, N.B., 2012b. Social anxiety and cannabis use: An analysis from ecological momentary assessment. J. Anxiety Disord. *26* (2), 297–304. https://doi.org/10.1016/j.janxdis.2011.12.006.

Buckner, J.D., Crosby, R.D., Silgado, J., Wonderlich, S.A., Schmidt, N.B., 2012a. Immediate antecedents of marijuana use: An analysis from ecological momentary assessment. J. Behav. Ther. Exp. Psychiatry *43* (1), 647. https://doi.org/10.1016/j.jbtep.2011.09.010.

Buckner, J.D., Shah, S.M., Dean, K.E., Zvolensky, M.J., 2016. Cannabis use frequency and use-related impairment among African-American and White users: the impact of cannabis use motives. Ethn. Health *21* (3), 318–331.

Buckner, J.D., Zvolensky, M.J., Crosby, R.D., Wonderlich, S.A., Ecker, A.H., Richter, A., 2015. Antecedents and consequences of cannabis use among racially diverse cannabis users: an analysis from ecological momentary assessment. Drug Alcohol Depend. *147*, 20–25. https://doi.org/10.1016/j.drugalcdep.2014.12.022.

Buckner, J.D., Zvolensky, M.J., Ecker, A.H., 2013. Cannabis use during a voluntary quit attempt: an analysis from ecological momentary assessment. Drug Alcohol Depend. *132* (3), 610–616. https://doi.org/10.1016/j.drugalcdep.2013.04.013.

Callaghan, R.C., Sanches, M., Kish, S.J., 2020. Quantity and frequency of cannabis use in relation to cannabis-use disorder and cannabis-related problems. Drug Alcohol Depend. *217*, 108271 https://doi.org/10.1016/j.drugalcdep.2020.108271.

Chait, L.D., Evans, S.M., Grant, K.A., Kamien, J.B., Johanson, C.E., Schuster, C.R., 1988. Discriminative stimulus and subjective effects of smoked marijuana in humans. Psychopharmacology *94* (2), 206–212.

Chakroun, N., Johnson, E.I., Swendsen, J., 2010. Mood and personality-based models of substance use. Psychol. Addict. Behav. *24* (1), 129–136. https://doi.org/10.1037/a0018184.

Comulada, W.S., Swendeman, D., Wu, N., 2016. Cell phone-based ecological momentary assessment of substance use context for Latino youth in outpatient treatment: who, what, when and where. Drug Alcohol Depend. *167*, 207–213. https://doi.org/10.1016/j.drugalcdep.2016.08.623.

Denson, R.K., Hedeker, D., Mermelstein, R.J., 2023. Association between affect and cannabis use varies by social context. Drug Alcohol Depend. *243*, 109750 https://doi.org/10.1016/j.drugalcdep.2022.109750.

Freeman, L.K., Haney, A.M., Griffin, S.A., Fleming, M.N., Vebares, T.J., Motschman, C.A., Trull, T.J., 2023. Agreement between momentary and retrospective reports of cannabis use and alcohol use: comparison of ecological momentary assessment and timeline follow back indices. Psychol. Addict. Behav. *37* (4), 606–615. https://doi.org/10.1037/adb0000897.

Gunn, R.L., Sokolovsky, A., Stevens, A.K., Hayes, K., Fitzpatrick, S., White, H.R., Jackson, K.M., 2021. Contextual influences on simultaneous alcohol and cannabis use in a predominately white sample of college students. Psychol. Addict. Behav. *35* (6), 691–697. https://doi.org/10.1037/adb0000739.

Hasin, D., Walsh, C., 2020. Cannabis use, cannabis use disorder, and comorbid psychiatric illness: a narrative review. J. Clin. Med *10* (1), 15. https://doi.org/10.3390/jcm10010015.

Henquet, C., van Os, J., Kuepper, R., Delespaul, P., Smits, M., Campo, J.À., Myin-Germeys, I., 2010. Psychosis reactivity to cannabis use in daily life: an experience sampling study. Br. J. Psychiatry *196* (6), 447–453. https://doi.org/10.1192/bjp.bp.109.072249.

Kuerbis, A., Reid, M.C., Lake, J.E., Glasner-Edwards, S., Jenkins, J., Liao, D., Candelario, J., Moore, A.A., 2019. Daily factors driving daily substance use and chronic pain among older adults with HIV: an exploratory study using ecological momentary assessment. Alcohol *77*, 31–39. https://doi.org/10.1016/j.alcohol.2018.10.003.

Li, X., Vigil, J.M., Stith, S.S., Brockelman, F., Keeling, K., Hall, B., 2019. The effectiveness of self-directed medical cannabis treatment for pain. Complement Ther. Med *46*, 123–130. https://doi.org/10.1016/j.ctim.2019.07.022.

Mierswa, I., & Klinkenberg, R. (2018). *RapidMiner Studio* (Version 9.1). Altair Engineering Company. ⟨https://rapidminer.com/⟩

Morrison, W., Guerdan, L., Kanugo, J., Trull, T., Shang, Y., 2018, July. Tigeraware: an innovative mobile survey and sensor data collection and analytics system. Pap. Presente 2018 IEEE Third Int. Conf. Data Sci. Cyber (DSC), Guangzhou, China.

Mun, E.Y., Li, X., Businelle, M.S., Hébert, E.T., Tan, Z., Barnett, N.P., Walters, S.T., 2021. Ecological momentary assessment of alcohol consumption and its concordance with transdermal alcohol detection and timeline follow-back self-report among adults experiencing homelessness. Alcohol Clin. Exp. Res *45* (4), 864–876. https://doi.org/10.1111/acer.14571.

Okey, S.A., Waddell, J.T., Shah, R.V., Kennedy, G.M., Frangos, M.P., Corbin, W.R., 2023. An ecological examination of indica versus sativa and primary terpenes on the subjective effects of smoked cannabis: A preliminary investigation. Cannabis Cannabinoid Res. Advance online publication,. https://doi.org/10.1089/can.2022.0213.

Pearson, M.R., 2019. A meta-analytic investigation of the associations between cannabis use and cannabis-related negative consequences. Psychol. Addict. Behav. *33* (3), 190–196. https://doi.org/10.1037/adb0000452.

Phillips, K.T., Phillips, M.M., Lalonde, T.L., Prince, M.A., 2018. Does social context matter? an ecological momentary assessment study of marijuana use among college students. Addict. Behav. *83*, 154–159. https://doi.org/10.1016/j.addbeh.2018.01.004.

Platt, J., 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Adv. Large Margin Classif. *10* (3), 61–74.

Prince, M.A., Conner, B.T., Pearson, M.R., 2018. Quantifying cannabis: a field study of marijuana quantity estimation. Psychol. Addict. Behav. *32* (4), 426–433. https://doi.org/10.1037/adb0000370.

Ross, C.S., Brooks, D.R., Aschengrau, A., Siegel, M.B., Weinberg, J., Shrier, L.A., 2018. Positive and negative affect following marijuana use in naturalistic settings: an ecological momentary assessment study. Addict. Behav. *76*, 61–67. https://doi.org/10.1016/j.addbeh.2017.07.020.

Rubin, M., Donkin, C., 2022. Exploratory hypothesis tests can be more compelling than confirmatory hypothesis tests. Philos. Psychol. 1–29. https://doi.org/10.1080/09515089.2022.2113771.

Shrier, L.A., Scherer, E.B., 2014. It depends on when you ask: motives for using marijuana assessed before versus after a marijuana use event. Addict. Behav. *39* (12), 1759–1765. https://doi.org/10.1016/j.addbeh.2014.07.018.

Shrier, L.A., Rhoads, A., Burke, P., Walls, C., Blood, E.A., 2014. Real-time, contextual intervention using mobile technology to reduce marijuana use among youth: a pilot study. Addict. Behav. *39* (1), 173–180. https://doi.org/10.1016/j.addbeh.2013.09.028.

Simons, J., Correia, C.J., Carey, K.B., Borsari, B.E., 1998. Validating a five-factor marijuana motives measure: Relations with use, problems, and alcohol motives. J. Couns. Psychol. *45* (3), 265–273. https://doi.org/10.1037/0022-0167.45.3.265.

Sznitman, S.R., van Rijswijk, L., Cousijn, J., 2022. Cannabis use as a predictor and outcome of positive and negative affect in college students: an ecological momentary assessment study. Addict. Behav. *128*, 107221.

Solhan, M.B., Trull, T.J., Jahng, S., Wood, P.K., 2009. Clinical assessment of affective instability: comparing EMA indices, questionnaire reports, and retrospective recall. Psychol. Assess. *21* (3), 425–436. https://doi.org/10.1037/a0016869.

Treloar Padovano, H., Miranda, R., 2018. Subjective cannabis effects as part of a developing disorder in adolescents and emerging adults. J. Abnorm Psychol. *127* (3), 282–293. https://doi.org/10.1037/abn0000342.

Tournier, M., Sorbara, F., Gindre, C., Swendsen, J.D., Verdoux, H., 2003. Cannabis use and anxiety in daily life: a naturalistic investigation in a non-clinical population. Psychiatry Res *118* (1), 1–8. https://doi.org/10.1016/S0165-1781(03)00052-0.

Trull, T.J., Ebner-Priemer, U., 2013. Ambulatory assessment. Annu Rev. Clin. Psychol. *9*, 151–176. https://doi.org/10.1146/annurev-clinpsy-050212-185510.

Trull, T.J., Wycoff, A.M., Lane, S.P., Carpenter, R.W., Brown, W.C., 2016. Cannabis and alcohol use, affect and impulsivity in psychiatric out-patients' daily lives. Addiction *111* (11), 2052–2059. https://doi.org/10.1111/add.13471.

Tyler, E., Jones, S., Black, N., Carter, L.A., Barrowclough, C., 2015. The relationship between bipolar disorder and cannabis use in daily life: an experience sampling study. PLoS ONE 10 (3), e0118916. https://doi.org/10.1371/journal.pone.0118916.

Volkow, N.D., Baler, R.D., Compton, W.M., Weiss, S.R., 2014. Adverse health effects of marijuana use. N. Engl. J. Med *370* (23), 2219–2227. https://doi.org/10.1056/NEJMra1402309.

Votaw, V.R., Witkiewitz, K., 2021. Motives for substance use in daily life: a systematic review of studies using ecological momentary assessment. Clin. Psychol. Sci. *9* (4), 535–562. https://doi.org/10.1177/2167702620978614.

Watson, D., Clark, L.A., 1994. The PANAS-X: Manual for the positive and negative affect schedule-expanded form. University of Iowa,. https://doi.org/10.17077/48vt-m4t2.

Whiteside, S.P., Lynam, D.R., 2001. The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. Pers. Individ Dif. *30* (4), 669–689. https://doi.org/10.1016/S0191-8869(00)00064-7.

Wycoff, A.M., Metrik, J., Trull, T.J., 2018. Affect and cannabis use in daily life: a review and recommendations for future research. Drug Alcohol Depend. *191*, 223–233. https://doi.org/10.1016/j.drugalcdep.2018.07.001.

Yarkoni, T., Westfall, J., 2017. Choosing prediction over explanation in psychology: lessons from machine learning. Perspect. Psychol. Sci. *12* (6), 1100–1122. https://doi.org/10.1177/1745691617693393.