Smoking lapses and cravings can be predicted with high accuracy using smokers' smartphones' movements.

Authors: Maryam Abo-Tabik[1]*†, Nicholas Costen[2] and Yael Benn[3]†
Affiliations:
1*Department of Computer Science, University of Central Lancashire, Preston, UK.
2Department of Computing and Maths, Manchester Metropolitan University, Manchester, UK.
3Department, of Psychology, Manchester Metropolitan University, Manchester, UK.
***Corresponding author(s). E-mail(s): mabo-tabik@uclan.ac.uk;**
Contributing authors: n.costen@mmu.ac.uk; y.benn@mmu.ac.uk;
†These authors contributed equally to this work.

Abstract: Decades of research on smoking-behavior had identified triggers that contribute to failed quitting attempts in order to develop effective interventions to support smokers. Triggers may be environmental (e.g., location), social, or internal (e.g. stress). Here, we show that movement data collected from smokers' smartphones' sensors (accelerometer, gyroscope and magnetometer) is a better predictor of smoking-behavior. Feeding the movement data into a Deep Learning (DL) model (1D-CNN-BiLSTM), smoking-behavior was predicted with 85% accuracy within a 5-minute window. This is compared to 63% accuracy when using traditional triggers such as time of the day. Crucially, movement data can be used to predict high-craving and lapses in the 3-months following quitting smoking with similarly high accuracy, even when predictions are made without any personal data (i.e., when the model is trained using only data from other smokers). Findings can be used to transform smoking-cessation smartphone apps, enabling the provision of just-in-time personalized support to those wishing to quit smoking. Importantly, the findings have implications beyond smoking-cessation applications, by revealing that human movements, largely overlooked to date, can be used for, early detection of, and intervention for, health (and other) behaviors, including those that are not genetic or typically characterized by changes in movement.

1. Introduction

Smoking is a public health emergency, as it increases the risks for multiple health conditions, which result in the death of around 8 million people worldwide annually [1, 2]. Despite many smokers trying to quit, quitting success rate remains low [3]. To support smokers who wish to quit, many mobile-based smoking cessation apps have been developed [4]. While there are several popular applications that are effective in improving quitting rate (e.g. the United Kingdom National Health Service quit smoking app [5]), these apps do not make use of the full range of current technological advances [6], hence provide less than optimal support.

Studies indicate that providing just-in-time targeted intervention could minimize relapse incidents [7, 8]. To enable the delivery of such timely intervention (e.g. via a mobile app), several attempts have been made to utilize Machine Learning (ML) algorithms for predicting smoking-behavior based on known smoking risk-factors such as high urges [9], geographical location [10] or social triggers [11]. However, these studies suffer from several limitations. Most importantly, the data they are based on largely relies on self-reporting from smokers (e.g. [9]), which is considered unreliable [12].

Multiple mobile sensors (e.g., accelerometer, Global Positioning System: GPS) are commonly embedded within smartphones and smart watches, and they offer a superior method for passively collecting real-time and objective data. Such data had been increasingly used within ML algorithms to analyze and understand complex behavioral patterns to improve insight into different behaviors and medical conditions [13]. For example, [4] attempted to predict smoking events using passively collected GPS and Accelerometer data from smokers' smartphones. The data was processed using a ML model that combined a 1-Dimention Convolutional Neural Networks (1D-CNN) with Control Theory. While the model achieved a 74% accuracy of predicting smoking events, this was only tested using 'regular' smoking data, i.e., before the smokers attempted to quit (hence, it is unclear whether the model would be able to predict lapses after quitting). The model suffers from several other limitations. First, it relies on GPS data, which is highly intrusive and as such, often unavailable to be collected continuously using modern mobile devices. Second, the complexity of the combined model would make it very difficult to be used as part of a smoking cessation app. Lastly, the model's performance, while high in accuracy, was only able to make prediction on smoking events within a 60-minute window, which is not sufficiently precise for 'just-in time' intervention.

More recently, studies have begun to use movement data alone, collected by digital sensors, to predict and understand behaviors in both humans (e.g. disruptive behavior among children [14]; classify stress or anxiety indicators [15]), and animals [e.g. [16]]. Furthermore, recent work had shown that some fine-movements, which even when detectable to the human eye often go un-noticed (e.g. pupil and head movement), could be effectively used to diagnose conditions such as Autism [17]. The method, currently referred to as 'Behavioral Phenotyping', has the potential to target a wide range of health behaviors [18–20], but to date have not been used for identifying or characterize non-genetic conditions. The integration of sensors technology with Artificial Intelligence (AI) has opened new opportunities for digital phenotyping and health monitoring, enabling the passive collection of sensors' data, call history, SMS patterns, and application usage, to be utilized for behavioral characterization and prediction [21].

Here, we report the outcome of two studies that investigated the capacity of sensors' data, collected from smokers' smartphones and processed using a Deep Learning (DL) neural network algorithm, to predict smoking-behavior before and after quitting. The outcome can demonstrate the potential of human movements, as detected by digital sensors (but perhaps not consciously registered by humans) to identify seemingly unrelated conditions. As such, the technology offers opportunities for characterising a wide-range of behaviors, and specifically for early detection and intervention for problematic health behaviors.

Developing a smoking-behavior identification model that can be effectively used within smart smoking cessation apps requires training the model using ecologically valid data. Data from N=17 smokers was collected using a simple smartphone app. In the first phase of data collection (henceforth, Phase-1), smokers were asked to report in real time, by pressing a button on their app, every cigarette they smoked over a two weeks period. At the end of the two weeks, participants were given a 5-days window to quit smoking, before progressing to the second phase of data collection (Phase-2). In Phase-2, participants were asked to report, over a three months post-quitting period, any smoking lapses. They were also asked to report their craving level on a scale from 1 (very low) to 5 (very high), every time that they experienced high craving, or at least once a day if they did not experience high cravings (whenever they remembered to do so). This was important as a measure of engagement with the app. Smartphoness' sensors' data (Accelerometer: ACC; Gyroscope: GYR; Magnetometer: MAG; Light: L; Time of the Day: T and GPS) were continuously recorded every minute throughout both phases of data collection. The collected data was pre-processed and then fed into four different DL models. Each model was assessed for its effectiveness in predicting smoking events during the pre-quit period (Phase-1) and instances of smoking lapses and cravings after quitting (Phase-2).

To develop a DL prediction model with a 5-minutes prediction period, the collected data was converted to non-overlapped samples. This is common practice in action prediction ML models when training uses samples that are derived from raw signal data [22, 23]. All data samples with majority of missing samples where removed (as these are not useful for modeling), and the remaining samples were assigned either a (0) value for non-smoking samples or (1) for smoking events pre-quitting, and reported cravings and lapses post-quitting. Finally, to avoid the unbalanced data problem [24], the data was randomly down-sampled with noise removal [25, 26].

In the each data sample included 25 rows of smartphone sensors data, with each row consisting of 5 columns (i.e. ACC, GYR, MAG, L, & T). GPS data was not used in the model, as due to the sensitivity of the data, modern smartphones only enable GPS data collection when the app is open- i.e., when smokers reported an event on the app, making the data highly biased (it was only available for smoking and craving samples).

2. Using each smoker's Phase-1 data, to predict (in Phase-1) smoking-behavior, and lapses and cravings (in Phase-2)

At the first stage, we evaluated the performance of four DL models: 1D-CNN, Long Short Term Memory (LSTM), Bidirectional LSTM (BiLSTM), and 1D-CNN-BiLSTM. The models were tested for their ability to predict smoking events based on smokers' Phase-1 data (pre-quit data) served as the training dataset for the different DL models, utilizing 90% of smokers' collected data to predict smoking events in the remaining 10% of the Phase-1 data.

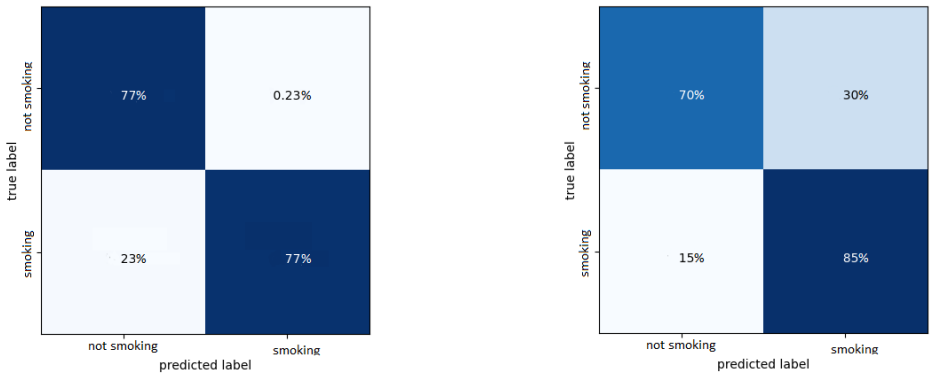
Table 1(A) displays the prediction accuracy of the different models using different sensors' data as input. The table reveals that the combined 1D-CNN-BilSTM can predict smoking events with 85% accuracy within a 5-minutes window, using the phone's movement data alone (ACC, GYR and MAG).

Subsequently, we employed the same DL models, trained with 100% smokers' Phase-1 data, to test their capacity in predicting smokers' Phase-2 (12-week post-quit period) smoking lapses and reported cravings (Table 1(B)). Again, the 1D-CNN-BilSTM was superior, predicting behavior with 78% accuracy within a 5-minutes window.

Table 1: Comparison of models' accuracy in predicting smoking events pre-quitting (Phase-1), and lapses and cravings post-quitting (Phase-2).

Input type	LSTM	1D-CNN	BiLSTM	1D- CNN-BiLSTM
A. smokers' Phase-1 data, used for predicting smoking events pre-quit				
Accelerometer (ACC)	0.833	0.818	0.841	0.847
Gyroscope (GYR)	0.829	0.818	0.826	0.826
Magnetometer (MAG)	0.550	0.771	0.585	0.795
Light (L)	0.702	0.756	0.698	0.752
Time of the Day (T)	0.632	0.651	0.632	0.632
ACC-GYR-MAG	0.574	0.798	0.554	**0.854**
B. smokers' Phase-1 data, used for predicting lapsing and craving at Phase-2 'post-quit'				
Accelerometer (ACC)	0.767	0.743	0.735	0.772
Gyroscope (GYR)	0.755	0.753	0.769	0.759
Magnetometer (MAG)	0.459	0.717	0.503	0.734
Light (L)	0.670	0.665	0.667	0.683
Time of the Day (T)	0.603	0.596	0.603	0.602
ACC-GYR-MAG	0.509	0.713	0.517	**0.775**

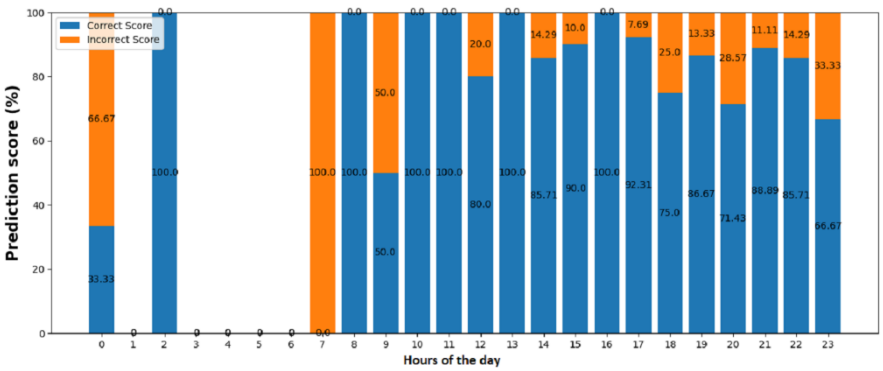
The lower prediction accuracy in Phase-2 is expected due to the change in smoking behavior during this period. It is particularly noted that data from sensors previously not considered by scientists for characterizing smoking-behavior, i.e., movement data, have very high predictive value (e.g., ACC, 0.772). This is in contrast to data such as, 'time of the day', which was previously considered a major predictive value for smoking-behavior [27], but in fact provides the lowest single predictive value (0.602). However, the relatively low predictability of time and light sensors could be due to the nature of the data, which can cause class overlapping.
In both the above studies the overall accuracy of the integrated 1D-CNN & BiLSTM model shows minimal difference in performance when using ACC data alone as input compared to using the combined input from ACC, GYR, and MAG sensors. Still, in both studies the 1D-CNN-BiLSTM model demonstrates enhanced ability in predicting smoking events, lapses and reported cravings when data from all three movement sensor types are combined.
This is clearly demonstrated by the confusion matrix (Figure 1), showing 0.85 accuracy when using all three sensors compared to 0.77 when using the ACC alone. This difference is critical when implementing smoking interventions, to avoid sending false intervention-messages to smokers, which may instead of helping, act as a reminder of their smoking habit, with negative impact on their quitting attempt as reported in [10].



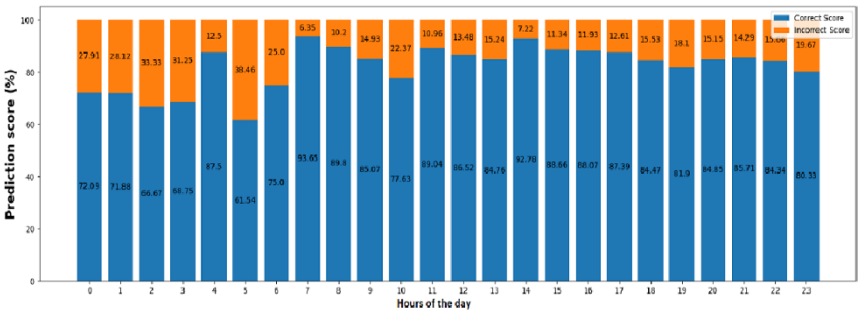
(a) Prediction using only ACC data (b) Prediction using ACC, GYR, and MAG

Fig. 1: 1D-CNN-BiLSTM confusion matrix.

To validate the efficacy of the algorithm for use within smart smoking cessation apps, we conducted a more in-depth analysis of the model's ability to differentiate between lapses and craving events. Phase-2 labels were reclassified into either 'smoking' (actual lapse) or 'craving' (every instance when the smoker reported a craving level ranging from 1 to 5). The rationale is that even if reported craving is low, the act of reporting indicates that the smoker had thought about their smoking habit at that moment.
Figures 2 shows the efficacy of the combined ACC, GYR and MAG inputs in predicting moments of either smoking lapse (a) or craving-reporting (b) at different time of the day.



(a) 1D-CNN-BiLSTM to differentiate smoking lapses using ACC, GYR, and MAG data



(b) 1D-CNN-BiLSTM performance in predicting craving events using ACC, GYR, and MAG data

Fig. 2: 1D-CNN-BiLSTM prediction score for smoking lapses and craving during different hours of the day.

Lastly, Figure 3 shows the model's prediction accuracy for cravings at various reported levels (1, 2, 3, 4, or 5), along with smoking lapses. The results clearly demonstrate the ability of the model to recognize and differentiate lapse incidents and reported craving events. This highlights the potential of the model for supporting individuals through the challenges of quitting their smoking addiction by effectively intervening in timely manner during high-risk scenarios associated with lapsing.

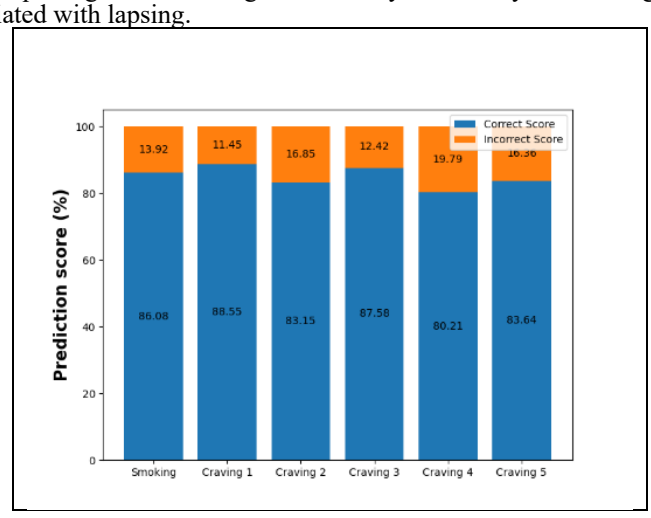


Fig. 3: Smoking lapses and self-reported craving level (on a scale from 1-5) predictions as made by the 1D-CNN-BiLSTM model with input from ACC, GYR, and MAG sensors.

3. Using other smokers' pre-quit data, to predict post-quit lapses and cravings of a new smoker.

If the model is able to predict lapsing behavior of any smoker using a model trained by the smoking and lapses data from other smokers, it would mean that an app can be downloaded by smokers and instantly used to support their quitting, without having to go through the two-weeks 'training' period to learn the individual's pattern. It would also mean that smokers' movements, as recognized by their smartphone's sensors, are universal in the identification of smoking lapses. This would suggest that those movements, which have not to date been identified by scientists despite decades of research into smoking behavior, may offer a window to the diagnosis or intervention of other conditions or behaviors.

To test the accuracy of the DL model in making such predictions, each participant's Phase-1 data was, in turn, held back from the training set, and the model was tested using the Phase-2 data of that participant. In this test only 14 out of the 17 participants' data was used because three participants did not report any post-quitting data. Only the 1D-CNN-BiLSTM with 9 vectors inputs from the ACC (x,y,z), GYR (x,y,z) and MAG (x,y,z), was used, given the superiority of the model in predicting smoking and lapses at the previous analysis.

Figure 4 shows the Receiver Operating Characteristic (ROC) curve for the model, which is a performance measure used in ML to understand the capability of a model to distinguish between different classes. While the Area Under the ROC Curve (AUC) values range between 0.68 to 0.87, the 1D-CNN-BiLSTM achieves an average AUC = 0.8 within a 5-minutes window. The degradation in the model's capacity to achieve accurate prediction among certain participants lacks a specific indicator, suggesting that there is no specific pattern of their behavior that could help the model improve its prediction rate.

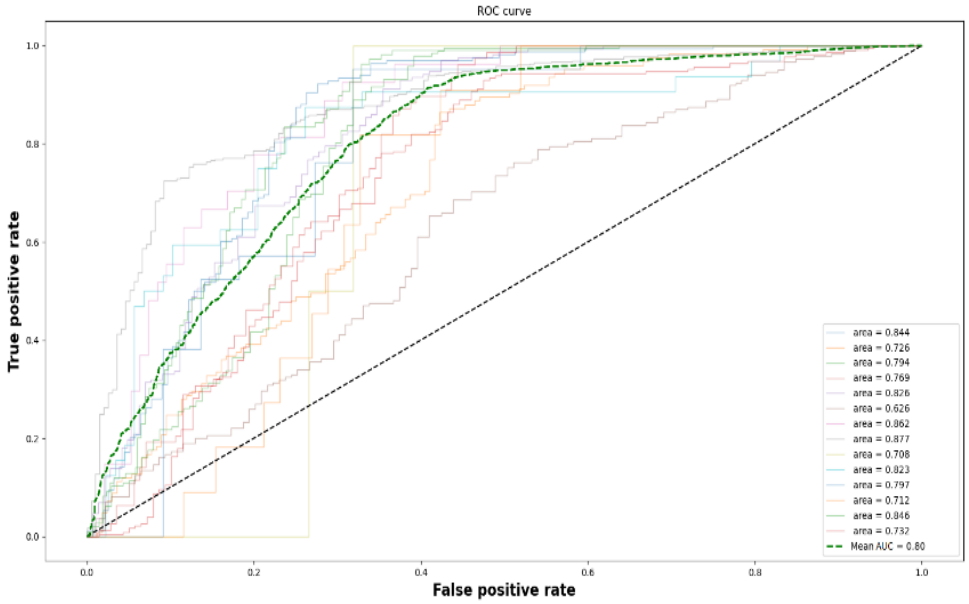
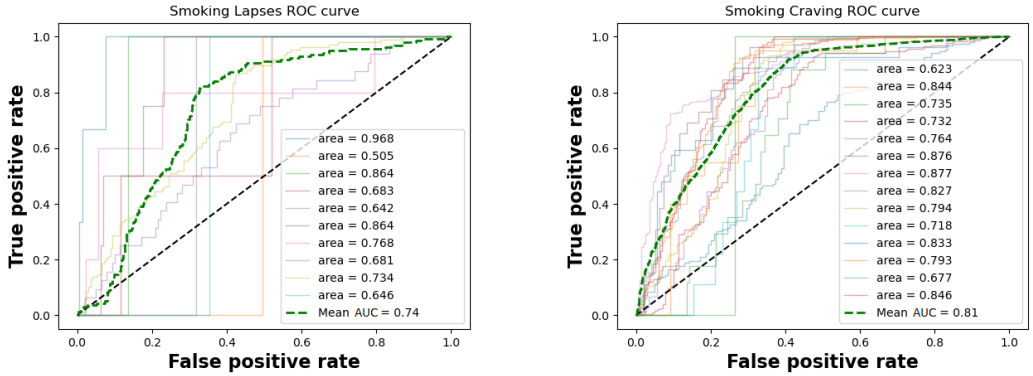


Fig. 4: The 1D-CNN-BiLSTM ROC curve, trained using Phase-1 data of 13 smokers, to predict the Phase-2 smoking patterns of the 14th smoker. Light colored lines represent each of the 14 participants while the green dotted line represents the average of all participants.

As before, labels were re-classified to separate smoking lapses and reported craving level, to enable the analysis of ROC for both incidents separately. Figure 5 shows that the 1D-CNN-BiLSTM has overall good prediction, with higher ability to predict cravings (AUC = 0.81) than reported smoking lapses (AUC = 0.74).



(a) Reported smoking lapses (b) Reported smoking cravings

Fig. 5: ROC curve after re-classified Phase-2 data to smoking lapses and reported cravings. Note that the AUC value for smoking lapses is for only 10 participants, as the other 4 did not report lapses during the quit period.

4. Discussion

The research demonstrated that human movements, as detectable by smartphones' sensors, can be used to predict health behaviors such as smoking lapses and cravings with higher accuracy than that achieved by human-observable patterns (e.g., locations, time of the day). Despite significant advancements in the field of digital health and behavioral predictions and phenotyping, there exists a notable gap in the processing of smartphone-collected data using DL algorithms for developing effective characterisation, diagnostic and interventions tools for smoking-cessation and other (health) behaviors.

To date, applications that are designed to support smoking identification and/or support for cessation, if using ML algorithms at all, have primarily focused on motion sensors from wearable devices to identify specific actions (e.g., arm movement to identify smoking action [29–31]) or self-reported smoking triggers (e.g. locations, [10]) neglecting the wealth of information that can be collected from the multifaceted sensors embedded in smartphones. There is therefore a need to explore how DL models can be used to effectively predict behaviors, and capture temporal patterns, and accurately identify the precise time-window in which support or interventions are needed. Such support for smokers may be particularly valuable during the first few months after quitting [32, 33].

Overall, the research posed the null hypothesis that passively collected data from smartphones, analyzed using DL algorithm, will have no significant effect on the prediction of smoking behavior. To test this hypothesis, the Friedman test [28] was used to compare multiple classifiers with various data types: ACC, GYR, MAG, L, T, and a combined ACC-GYR-MAG input. In the ensuing analysis, each of the four DL models (LSTM, CNN, BiLSTM, and 1D-CNN-BiLSTM), were subjected to two significant levels, $\alpha = 0.05$ and $\alpha = 0.1$. The evaluation criterion was based on the generation of p-values, where a p-value below the chosen 0.01 signifies a statistically significant difference. To this end, the outcome of the Friedman test for all four DL models rejects the null hypothesis (Table 2). These findings show the substantial impact of passively collected smartphones' sensors data on the disparities in performance of the evaluated DL models in predicting smoking events.

Table 2: Summary of Friedman test Results

Model	Q-Statistic	Degrees of Freedom	p-Value
LSTM	27.489796	5.0	0.000046
CNN	15.826446	5.0	0.007357
BiLSTM	28.106996	5.0	0.000035
1D-CNN-BiLSTM	19.016736	5.0	0.001908

Despite the many advances achieved by this work, it is not without limitations. First, the collected data was geographically restricted to the United Kingdom. This limitation raises questions about the generalizability of the findings to diverse demographics and cultural contexts. As the model relies on movements for which we do not know if there are cultural or regional differences, it would be good to validate this with data from other cultures. This is of particular interest given that smoking triggers vary across different populations and regions (e.g., alcohol consumption is only a trigger in cultures that consume alcohol). Moreover, while the study demonstrates the effectiveness of the DL models in predicting smoking behavior within the studied population, further validation in a diverse populations (e.g. individuals with limited mobility) is necessary to ensure the robustness and applicability of the models in universally supporting smokers through the quitting process. Expanding the dataset to encompass diverse populations and providing validation within real-world applications is crucial for ensuring the robustness and generalizability in the complex landscape of providing effective and timely intervention for all smokers.

Despite the above limitation, the outcome of this research is promising, as the DL model was trained using data collected in an uncontrolled environment, with participants going about their usual daily routines without any

changes except for pressing a button to report smoking events in real time during a 2-weeks pre-quit period, and reporting smoking lapses and craving during a 3-months post-quit period. These may not be needed when the model is implemented into a smoking cessation app. Future work should explore to what extent data from these sensors can be used to predict or characterised other human behaviors that are not considered genetic or those that are specifically associated with visible changes in movement.

5. Methods

Seventeen smokers (10 Females and 7 males, Mean age= 37.18, minimum 5 cigarettes a day for a minimum of 6 months) were recruited for the study and compensated. The study received full ethical approval (see details 5). Before condensing and cleaning the data as described above, the entire dataset included 404478 5-minute samples (Phase-1:81837, Phase-2:322641). The final data set used for the training and validation process included a total of 2572 Phase-1 and 3616 Phase-2 5-minutes data period samples (overall smoking lapses: 158, craving reporting: 1650, and 4337 no smoking events).

The research utilized a stacked DL model that combines 1D-CNN [34] and BiLSTM [35, 36] DL models. In the combined DL model, the 1D-CNN component can learn local patterns from the input data but it cannot learn sequential correlations, whereas the BiLSTM component is specialized for sequential modeling, and can extract correlated patterns. Therefore, the combined model has the potential to improve the prediction accuracy of smoking (and craving) events, as this requires processing of large environmental datasets produced by phone sensors (input), while correlating it with smoking related events in the past and the future (sequential correlations). This approach was shown to be effective in several domains, including time series prediction and health applications [34,37,38].

The input layer accepts x-input vector (x is the number of inputs that changes based on the used data, ACC, GYR, MAG, T, & L) each with 25-row data (25-rows represent 5 samples of data for each minute, resulting in a total (5 x 5) durations of 5-minutes). This is passed to the convolutional layer, with a filter-size = 128; the convolutional operation takes the form of

$$C_i = h(W^T \otimes x_{i-10:i} + b_i) \quad \text{..... (1)}$$

The $\otimes$ is convolution operator, w^T the network weights, b is the bias, and h is the non-linear ReLU activation function with l_2 weight regularization. The feature map output of this layer is then passed to a process of batch normalization, which is followed by a max-pooling layer that reduces the features' variance by using a max function as follows:

$$P_{i,k} = Max\left(C_{(i,k)}\ U_{3,1}\right) \quad \text{..... (2)}$$

k is the filter number, and $U_{3,1}$ is sliding max window of size (3x1). The second level of the DL model is the BiLSTM network (combining forward and backward LSTM models). Each LSTM consists of several multiplicative memory cells, each with three gates: input, output, and forget gate. These gates control the output from the cell, which is either 'keep', 'release', or 'reset'. The following equations show the functionality of the BiLSTM model:

$$f_t^{\rightarrow} = \sigma\left(w_f^{\rightarrow} \bullet [h_{t-1}^{\rightarrow}, x_t^{\rightarrow}] + b_t^{\rightarrow}\right) \quad \text{..... (3)}$$
$$i_t^{\rightarrow} = \sigma(w_i^{\rightarrow} \bullet [h_{t-1}^{\rightarrow}, x_t^{\rightarrow}] + b_i^{\rightarrow}) \quad \text{..... (4)}$$
$$\tilde{c}_t^{\rightarrow} = tanh(w_c^{\rightarrow} \bullet [h_{t-1}^{\rightarrow}, x_t^{\rightarrow}] + b_c^{\rightarrow}) \quad \text{..... (5)}$$
$$c_t^{\rightarrow} = f_t^{\rightarrow} * \tilde{c}_{t-1}^{\rightarrow} + i_t^{\rightarrow} * \tilde{c}_t^{\rightarrow} \quad \text{..... (6)}$$
$$o_t^{\rightarrow} = \sigma(w_o^{\rightarrow} \bullet [h_{t-1}^{\rightarrow}, x_t^{\rightarrow}] + b_o^{\rightarrow}) \quad \text{..... (7)}$$
$$h_t^{\rightarrow} = o_t^{\rightarrow} * tanh(c_t^{\rightarrow}) \quad \text{..... (8)}$$

In the BiLSTM fi,it, and ot are calculated for both forward and backward networks; for the forward LSTM network is calculated using the current t and the previous t−1, While the for the backward LSTM network is calculated using the current t and the next t + 1 as

$$f_t^{\leftarrow} = \sigma\left(w_f^{\leftarrow} \bullet [h_{t+1}^{\leftarrow}, x_t^{\leftarrow}] + b_t^{\leftarrow}\right) \quad \text{..... (9)}$$
$$i_t^{\leftarrow} = \sigma(w_i^{\leftarrow} \bullet [h_{t+1}^{\leftarrow}, x_t^{\leftarrow}] + b_i^{\leftarrow}) \quad \text{..... (10)}$$
$$\tilde{c}_t^{\leftarrow} = tanh(w_c^{\leftarrow} \bullet [h_{t+1}^{\leftarrow}, x_t^{\leftarrow}] + b_c^{\leftarrow}) \quad \text{..... (10)}$$
$$c_t^{\leftarrow} = f_t^{\leftarrow} * \tilde{c}_{t+1}^{\leftarrow} + i_t^{\leftarrow} * \tilde{c}_t^{\leftarrow} \quad \text{..... (12)}$$
$$o_t^{\leftarrow} = \sigma(w_o^{\leftarrow} \bullet [h_{t+1}^{\leftarrow}, x_t^{\leftarrow}] + b_o^{\leftarrow}) \quad \text{..... (13)}$$
$$h_t^{\leftarrow} = o_t^{\leftarrow} * tanh(c_t^{\leftarrow}) \quad \text{..... (14)}$$

The final h_t will be is the concatenated vector of both of and as follows,

$$h_t = h_t^{\rightarrow} \otimes h_t^{\leftarrow} \quad \text{..... (15)}$$

The output from the BiLSTM is then passed to the final level of the model, which is made of three fully connected neural networks. These are used for generating one hot-encoded output that represents a prediction of either a smoking or a not-smoking session within a given 5-minute slot.

Declarations

The study was supported by internal research funding from Manchester Metropolitan University, UK (MMU pump funding from the faculty of engineering, 2018 awarded to the 2nd and 3rd authors). The authors declare no conflict of interest. The work received ethical approval from MMU, in line with the British Psychological Society (Ethos Numbers: 0441, 17913). All participants provided informed written consent to taking part, and

for the data to be published. Data and code are available by request from the 1st author- this will be free for research and education, and may require a fee for commercial application.

Author contribution

MAT: Data curation, Formal Analysis, Investigation, Methodology, Project administration, Visualization, Writing - original draft. NC: Formal Analysis, Funding acquisition, Methodology, Supervision, Validation, Writing - review & editing YB: Conceptualization, Funding acquisition, Methodology, Supervision, Validation, Writing - review & editing

References

[1] Our World in Data. *Smoking is responsible for about 8 million premature deaths each year*. Available online: https://ourworldindata.org/smoking (Accessed on December 14, 2023.).

[2] World Health Organization. *Tobacco*. Available online: https://www.who.int/ news-room/fact-sheets/detail/tobacco. Accessed on December 14, 2023.

[3] Centers for Disease Control and Prevention. *Why Quitting Smoking is Hard*. Available online: https://www.cdc.gov/tobacco/campaign/tips/quit-smoking/. Accessed on December 14, 2023.

[4] Abo-Tabik, M., Costen, N., Darby, J., & Benn, Y. (2020). Towards a smart smoking cessation app: A 1D-CNN model predicting smoking events. *Sensors*, 20(4), 1099.

[5] NHS. *Quit Smoking*. https://www.nhs.uk/better-health/quit-smoking/

[6] M. Abo-Tabik, Y. Benn, N. Costen. *Are Machine Learning Methods the Future for Smoking Cessation Apps'* Sensors, 2021.

[7] Hebert, E.T., Ra, C.K., Alexander, A.C., Helt, A., Moisiuc, R., Kendzor, D.E., Vidrine, D.J., Funk-Lawler, R.K., and Businelle, M.S. *A Mobile Just-In-Time Adaptive Intervention for Smoking Cessation: Pilot Randomized Controlled Trial*. Journal of Medical Internet Research, 22(3), e16907, 2020.

[8] Hebert, E. T., Stevens, E. M., Frank, S. G., Kendzor, D. E., Wetter, D. W., Zvolensky, M. J., Buckner, J. D., & Businelle, M. S. (2018). *An ecological momentary intervention for smoking cessation: the associations of just-in-time, tailored messages with lapse risk factors.* Addictive Behaviors, 78, 30-35.

[9] Dumortier, A., Beckjord, E., Shiffman, S., & Sejdi', E. (2016). *Classifying smoking urges via machine learning*. Computer Methods and Programs in Biomedicine, 137, 203-213.

[10] Schick, R.S., Kelsey, T.W., Marston, J., Samson, K., & Humphris, G.W. *MapMySmoke: Feasibility of a new quit cigarette smoking mobile phone application using integrated geo-positioning technology, and motivational messaging within a primary care setting. Pilot and Feasibility Studies*, 4(1), 1-10 (2018).

[11] Franzwa, F., Harper, L.A., and Anderson, K.G. (2022). *Examination of social smoking classifications using a machine learning approach*. *Addictive Behaviors*, 126, 107175.

[12] Gerritsen, M., Berndt, N., Lechner, L., de Vries, H., Mudde, A., & Bolman, C. (2015). Self-reporting of smoking cessation in cardiac patients: how reliable is it and is reliability associated with patient characteristics' *Journal of Addiction Medicine*, 9(4), 308-316.

[13] Baumgartner, R. *Precision medicine and digital phenotyping: digital medicine's way from more data to better health*. *Big Data And Society*, 8(2), 2021.

[14] Romanowicz, M., Croarkin, K.S., Elmaghraby, R., Skime, M., Mayo Clinic Inpatient Child and Adolescent Psychiatry Registered Nurse Staff, Croarkin, P.E., Vande Voort, J.L., Shekunov, J., and Athreya, A.P. *Machine Learning Identifies Smartwatch-Based Physiological Biomarker for Predicting Disruptive*

Behavior in Children: A Feasibility Study. *Journal of Child and Adolescent Psychopharmacology*, 33(9), pp.387-392, 2023.

[15] Razavi, M., McDonald, A., Mehta, R., & Sasangohar, F. (2023). Evaluating Mental Stress Among College Students Using Heart Rate and Hand Acceleration Data Collected from Wearable Sensors. arXiv preprint arXiv:2309.11097.

[16] Weissbrod, A., Shapiro, A., Vasserman, G., Edry, L., Dayan, M., Yitzhaky, A., Hertzberg, L., Feinerman, O., & Kimchi, T. (2013). Automated long-term tracking and social behavioural phenotyping of animal colonies within a semi-natural environment. *Nature communications*, 4(1), 2018.

[17] Perochon, Sam, J. Matias Di Martino, Kimberly LH Carpenter, Scott Compton, Naomi Davis, Brian Eichner, Steven Espinosa, et al. *Early detection of autism using digital behavioral phenotyping*. Nature Medicine, 2023, pp. 1-9.

[18] Burgermaster, M., & Rodriguez, V. A. (2022). *Psychosocial-behavioral phenotyping: a novel precision health approach to modeling behavioral, psychological, and social determinants of health using machine learning*. *Annals of Behavioral Medicine, 56*(12), 1258-1271.

[19] Mansoor, H., Gerych, W., Alajaji, A., Buquicchio, L., Chandrasekaran, K., Agu, E., & Rundensteiner, E.A. *Visual Analytics of Smartphone-Sensed Human Behavior and Health*. *IEEE Computer Graphics and Applications*, 41(3), 96-104, 2021.

[20] Ribeiro Filho, J.D.P., Teles, A.S., Silva, F.J., and Coutinho, L.R. *Towards clustering human behavioral patterns based on digital phenotyping*. In *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 271-276, June 2021. IEEE.

[21] Perez-Pozuelo, I., Spathis, D., Clifton, E.A., and Mascolo, C. *Wearables, smartphones, and artificial intelligence for digital phenotyping and health*. In *Digital Health*, Elsevier, 2021, pp. 33-54.

[22] Jordao, A., Nazare Jr, A.C., Sena, J. and Schwartz, W.R. *Human activity recognition based on wearable sensor data: A standardization of the state-of-the-art*. arXiv preprint arXiv:1806.05226, 2018.

[23] Mekruksavanich, S., Hnoohom, N., and Jitpattanakul, A. *A hybrid deep residual network for efficient transitional activity recognition based on wearable sensors*. Applied Sciences, 12(10), 4988, 2022.

[24] Bourel, M., Segura, A. M., Crisci, C., Lopez, G., Sampognaro, L., Vidal, V., Kruk, C., Piccini, C., & Perera, G. *Machine learning methods for imbalanced data set for prediction of faecal contamination in beach waters*. *Water Research*, 202, p. 117450. 2021.

[25] Liu, Y., Chawla, N. V., Harper, M. P., Shriberg, E., & Stolcke, A. (2006). *A study in machine learning from imbalanced data for sentence boundary detection in speech*. Computer Speech & Language, 20(4), 468-494.

[26] Dogra, V., Verma, S., Verma, K., Jhanjhi, N.Z., Ghosh, U., & Le, D.N. (2022). *A comparative analysis of machine learning models for banking news extraction by multiclass classification with imbalanced datasets of financial news: challenges and solutions*. Journal of Finance and Information Management, 19(3), 123-145.

[27] Grainge, M.J., Shahab, L., Hammond, D., O'Connor, R.J., and McNeill, A. *First cigarette on waking and time of day as predictors of puffing behaviour in UK adult smokers*. Drug and Alcohol Dependence, 101(3), pp. 191-195, 2009.

[28] Hoffman, J.I. (2015) *Analysis of Variance II. More Complex Forms*, Biostatistics for Medical and Biomedical Practitioners. London, England: Academic Press, pp. 421–447. Available at: https://doi.org/10.1016/B978-0-12-802387-7.00026-3.

[29] Benouis, M., Abo-Tabik, M., Benn, Y., Salmon, O., Barret-Chapman, A., Costen, N., *Behavioural smoking identification via Hand-Movement dynamics*. In: *2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, August 2019, pp. 1734-1739. IEEE.

[30] Mukhopadhyay, S., Dey, S. and Ghose, A. *TinyPuff: Automated design of Tiny Smoking Puff Classifiers for Body Worn Devices*. In Proceedings of the 8th Workshop on Body-Centric Computing Systems, June 2023, pp. 7-12.

[31] Hnoohom, N., Mekruksavanich, S., & Jitpattanakul, A. (2023). An Efficient ResNetSE Architecture for Smoking Activity Recognition from Smartwatch. *Intelligent Automation & Soft Computing*, 35(1).

[32] Brown RA, Lejuez CW, Kahler CW, Strong DR, Zvolensky MJ. *Distress tolerance and early smoking lapse.* Clinical Psychology Review. 2005 Sep;25(6):713-33. doi: 10.1016/j.cpr.2005.05.003. PMID: 16023275; PMCID: PMC1839854.

[33] Borland, R., Partos, T. R., Yong, H. H., Cummings, K. M., & Hyland, A. (2012). How much unsuccessful quitting activity is going on among adult smokers' *Addiction*, 107(3), 673-682.

[34] Lu, Y., & Li, K. (2023). Multistation collaborative prediction of air pollutants based on the CNN-BiLSTM model. Environmental Science and Pollution Research, 30(40), 92417-92435

[35] Yang, M. and Wang, J. (2022). Adaptability of financial time series prediction based on BiLSTM. Procedia Computer Science, 199, pp. 18-25.

[36] Kulshrestha, A., Krishnaswamy, V., & Sharma, M. (2020). Bayesian BILSTM approach for tourism demand forecasting. *Annals of Tourism Research, 83*, 102925

[37] Kavianpour, P., Kavianpour, M., Jahani, E., & Ramezani, A. (2023). A CNNBiLSTM model with attention mechanism for earthquake prediction. *The Journal of Supercomputing*, pp. 1-33

[38] Alkhodari, M. and Khandoker, A.H. *Detection of COVID-19 in smartphonebased breathing recordings using CNN-BiLSTM: a pre-screening deep learning tool*. medRxiv, pp.2021-09, 2021