Drug usage duration classification using Extreme Learning Machine based on personality features

by Nurul Hidayat

Submission date: 17-Mar-2021 02:21PM (UTC+0700)
Submission ID: 1535212556
File name: adinugroho2019.pdf (312.28K)
Word count: 3626
Character count: 17570

Drug usage duration classification using Extreme Learning Machine based on personality features

1st Sigit Adinugroho
Computer Vision Research Group
Faculty of Computer Science
Brawijaya University
Malang, Indonesia
sigit.adinu@ub.ac.id

2nd Yuita Arum Sari
Computer Vision Research Group
Faculty of Computer Science
Brawijaya University
Malang, Indonesia
yuita@ub.ac.id

3rd Nurul Hidayat
Informatics Department
Universitas Jenderal Soedirman
Purwokerto, Indonesia
nurul@unsoed.ac.id

Abstract—Determining the duration of drug consumption is essential for the success of treatment for drug abuse since the effectivity of such a program depends on the duration of the treatment. One promising set of features to identify the duration of drug consumption is personality features called Revised NEO Personality Inventory (NEO PI-R). In this paper, the Extreme Learning Machine model is employed to perform the classification. The model is trained and tested using 10-fold mechanism to verify the effectivity of the classification. The accuracy of the classifier differs, depending on the type of drug, with the maximum accuracy of 86.31% and the minimum one of 36.65%.

Index Terms—drug usage prediction, extreme learning machine, personality features

I. Introduction

Drug abuse has become major problem for countries around the globe. The production of some drug substances reached the highest level recorded. For instance, the opium production reached the staggering amount of 10,500 tons, an increase of 65% compared to the previous years and it is the highest number recorded by The United Nations Office on Drugs and Crime since the beginning of the twenty-first century. Since the nature of uncontrolled use of drugs that harm human health or even end up in death, their massive supply escalates the problem. In 2015 and 2016 drug abuse contributes to the decline of the US's life expectancy for the first time in 50 years. In general, it is estimated that 275 million people around the globe in 2016 consume drugs at least once in the past year, and 31 million of them experience drug use disorder. At this stage, a proper treatment is essential. [1].

Appropriate treatment to drugs addict has been proved to have positive impact, reducing the risk of users to consume them again, minimizing criminal behavior and other social problems [2]. The effectivity of such treatment depends on various factors. one of them is the duration of the remedy. A study by Institute of Medicine US revealed that there are positive correlation on treatment length and its result [2]. Another study suggested that the effect of a treatment is linear to its duration [3]. In addition, the duration of treatment is also related to the duration of drug use. Treatment period to reach one year or more of abstinence is significantly longer for people who used drug for a longer period [4]. Therefore, identifying the duration of drug use is crucial for determining treatment length, that in the ends determine the success of the treatment.

Various factors are known to prompt ones to consume the drug for the first time, including psychological, social, individual, environmental, and economic factors [5], [6], which correlates to personality characteristics [7]. So far, the most comprehensive measures for personality recognition is Revised NEO Personality Inventory (NEO PI-R) which measures Big Five personality traits, namely Neuroticism (N), Extraversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C) [8]. The Big Five personality traits are known to have relationship to drug consumption [9].

A number of attempts have been made to employ data mining and machine learning methods to detect drug use pattern based on personality traits. Alcohol user is predicted using two form backpropagation neural networks, ANN-D and ANN-C. The former predicts if a person is an alcohol user while the latter predict the consumption time. The accuracy of ANN-D is 98.7% and the accuracy of ANN-C is 49.1% [10]. The same method is also utilized for volatile substance abuse (VSA) detection [11]. Multiple algorithms, namely Bayesian Networks, Random Forest, Regression, SMO, J48, Bayesian Networks, Perceptron classifiers and Logistic regression, are applied to predict the frequency of drug use for four types of drugs: ketamine, heroin, crack and meth [12]. The most comprehensive study was conducted to classify user and non-user categories of 18 psychoactive drugs. Decision tree, random forest, k-nearest neighbors, linear discriminant analysis, Gaussian mixture, probability density function estimation, logistic regression and naive Bayes are involved to finish the task. The sensitivity and specificity for almost all drugs reached more than 70% for almost all drugs [9]. However, this research only performed binary classification from 7 available classes in the dataset.

This paper serves an experimental result on predicting the last time a respondent uses a specific type of drug using extreme learning machine. It also aims to find the

978-1-7281-3880-0/19/$31.00 ©2019 IEEE

proper ELM architecture that fits the problem as well as to figure out if personality features are sufficient for drug usage prediction. Our contribution is in predicting drug use period for 18 legal and illegal drugs, which has never been conducted before.

The rest of the paper is organized as follow. Section II describes the data and the proposed methods in detail while the result is presented in Section III

II. Methods and Data

The dataset used in this research is explained in Section II-A. The detail of the method is delineated in II-B.

A. Description of Dataset

The dataset for this research comes from Drug consumption (quantified) Data Set available at UCI Machine Learning Repository [9]. The data contains 1885 data points as a result of an online survey. Each respondent is required to answer question about the last time he used a specific drug by selecting one of seven available categories : never used (CL0), used over a decade ago (CL1), used in last decade (CL2), used in last year (CL3), used in last month (CL4), used in week (CL5), used in last day (CL6). There are 18 central nervous system psychoactive drugs are asked during the survey as shown in Table I. Along with drug consumption time, 12 attributes are also collected from respondents. Five attributes represent personality, which are neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness, as well as level of education, age, gender, country of residence, ethnicity, impulsivity, and sensation seeking. In this research, country, gender, and ethnicity features are omitted since they do not provide strong information for binary classification [9]. Therefore, in the subsequent discussion the term all features refers to 9 features kept from the dataset. The features in the published dataset have been quantified from their original categorical values.

TABLE I
Types of Drugs in the dataset

Name of Drugs		
alcohol	amphetamines	amyl nitrite
benzodiazepine	cannabis	chocolate
cocaine	caffeine	crack
ecstasy	heroin	ketamine
legal highs	LSD	methadone
mushrooms	nicotine	volatile substance abuse

B. Extreme Learning Machine as a classifier

Extreme Learning Machine (ELM) [13] is a new concept in artificial neural network that advanced the well-known training algorithm for multilayer feedforward neural network. The architectural design of ELM follows the standard feedforward neural network with single hidden layer as depicted in Fig. 1. It consists of single input layer in the left, single hidden layer in the middle and an output layer in the right. The number of neurons in input layer must match the number of features in the data while the number of neurons in the output layer normally corresponds to the number of classes in the classification problem and only one neuron for regression purpose. Each layer is fully connected and weighted to the one in front of it. It is worth to mention that there exist an activation function in the hidden layer. Various activation functions are available, however this paper implements the sigmoid function as in (1).

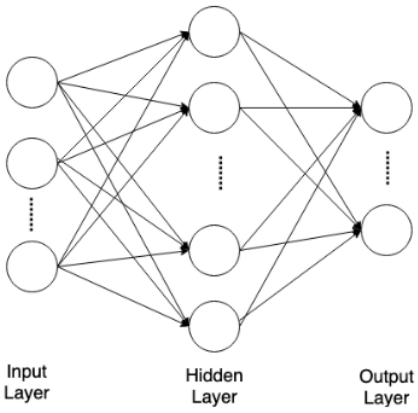

Fig. 1. ELM architecture

$$f(x) = \frac{1}{1+e^{-x}} \tag{1}$$

ELM implements a least-square solution for solving a linear system. The way it solves a linear system differs from backpropagation training algorithm which makes use of gradient descent. The one way least-square calculation makes ELM training time is significantly faster compared to backpropagation since it does not require multiple epoch looping that is required by backpropagation. In theory, ELM is able to resemble the approximation capability of backpropagation [14].

The first stage of ELM training works by generating random weights and bias that connect input layer to hidden layer w_j and b_j. The output $\hat{y}$ is calculated by (2).

$$\hat{y}_i = \sum_{j=1}^{m} \beta f(w_j x_i + b_j), 1 \leq i \leq n \tag{2}$$

where n is the number of training-target pairs (x_i, y_i), m denotes the number of hidden neurons, β_j is the jth output weight, $f(x)$ is the activation function, w_j is input weight and b_j is the bias.

ELM predicts the target y_i by $\hat{y}$. In ideal condition, ELM perfectly predict $\hat{y}$ equals to y_i as presented in

(3). The solution of (3) can be done by least-square approach in form of $\boldsymbol{H\beta = y}$. H is calculated in (4), $\beta = (\beta_1, \beta_2, \ldots, \beta_m)^T$, and $y = (y_1, y_2, \ldots, y_m)^T$.

$$\sum_{j=1}^{m} \beta f(w_j x_i + b_j) = y_i, 1 \leq i \leq n \tag{3}$$

$$H = \begin{bmatrix} f(w_1 x_1 + b_1) & \cdots & f(w_m x_1 + b_m) \\ \vdots & \ddots & \vdots \\ f(w_1 x_n + b_1) & \cdots & f(w_m x_n + b_m) \end{bmatrix} \tag{4}$$

After input weight and bias random generation, output weight β is calculated by (5)

$$\beta = H^{\dagger} y \tag{5}$$

where $H^{\dagger}$ is Moore–Penrose pseudo-inverse of $H \cdot B$.

III. Experimental Result

In order to assess the performance of the classification method, all data points from the dataset are used as training and testing data. The partition of training data and test data is conducted via k-fold mechanism and in this research, the k parameter is set to 10. Each fold alternately acts as testing data while the remaining 9 folds become training data. Accuracy for each training and testing session is measured and the result of 10 pairs of them are averaged to form overall accuracy.

First experiment is intended to understand the effect of the number of hidden neurons to the accuracy. In this scenario, ELMs with the various number of hidden neurons, ranging from 1 to 7, are tested. All 9 features are included for the test.

Table II shows all accuracy data for each drug type and number of the hidden neuron. The maximum accuracy achieved is shaded in red. In 38.89% types of drugs, the number of hidden neuron does not have any impact on accuracy. Those drugs are Amyl nitrite, Caffeine, Crack, Heroin, Ketamine, Methadone, and VSA. The accuracy of 27.78% of all drugs reaches the maximum value when using only one hidden neuron. The equal proportion of the drugs reaches their peak classification accuracy by using the maximum number of hidden neurons. The remaining 5.55% has its maximum accuracy by using 6 hidden neurons.

In order to understand the magnitude of hidden neuron selection to the performance of the ELM classifier, we calculate the difference between maximum accuracy and minimum accuracy in percentage for each drug. In the Table IV, it is shown that the difference of accuracy is less than 4% for the vast majority of drugs. There is exception for nicotine and cannabis in which different selection of hidden neurons yields 23.25% and 54.39% of accuracy change, respectively. In general, difference in number of neuron does not significantly affect the accuracy of the classification using ELM.

TABLE IV
Accuracy difference caused by the number of hidden neurons

Drug	Accuracy difference
Alcohol	1.25%
Amphetamines	1.97%
Amyl nitrite	0.00%
Benzodiazepine	0.95%
Caffeine	0.00%
Cannabis	54.39%
Chocolate	1.42%
Cocaine	0.92%
Crack	0.00%
Ecstasy	1.30%
Heroin	0.00%
Ketamine	0.00%
Legal highs	3.98%
LSD	1.42%
Methadone	0.00%
Magic mushrooms	1.17%
Nicotine	23.25%
VSA	0.00%

The second experiment is performed to investigate if 5

TABLE II
Prediction accuracy for all features

	Accuracy																
N hidden	Alcohol	Amphetamines	Amyl nitrite	Benzodiazepine	Caffeine	Cannabis	Chocolate	Cocaine	Crack	Ecstasy	Heroin	Ketamine	Legal highs	LSD	Methadone	Magic mushrooms	Nicotine	VSA
1	0.403	0.518	0.692	0.531	0.735	0.239	0.428	0.551	0.863	0.542	0.852	0.791	0.580	0.567	0.758	0.521	0.314	0.772
2	0.402	0.516	0.692	0.529	0.735	0.304	0.427	0.550	0.863	0.537	0.852	0.791	0.585	0.566	0.758	0.519	0.336	0.772
3	0.402	0.515	0.692	0.530	0.735	0.329	0.426	0.549	0.863	0.537	0.852	0.791	0.578	0.563	0.758	0.517	0.343	0.772
4	0.402	0.513	0.692	0.528	0.735	0.346	0.426	0.549	0.863	0.539	0.852	0.791	0.586	0.565	0.758	0.515	0.362	0.772
5	0.400	0.512	0.692	0.528	0.735	0.366	0.423	0.547	0.863	0.541	0.852	0.791	0.589	0.565	0.758	0.517	0.373	0.772
6	0.400	0.512	0.692	0.526	0.735	0.367	0.422	0.546	0.863	0.540	0.852	0.791	0.594	0.571	0.758	0.520	0.379	0.772
7	0.405	0.508	0.692	0.528	0.735	0.369	0.423	0.548	0.863	0.544	0.852	0.791	0.601	0.568	0.758	0.520	0.387	0.772

TABLE III
Prediction accuracy for 5 personality features

	Accuracy																
N hidden	Alcohol	Amphetamines	Amyl nitrite	Benzodiazepine	Caffeine	Cannabis	Chocolate	Cocaine	Crack	Ecstasy	Heroin	Ketamine	Legal highs	LSD	Methadone	Magic mushrooms	Nicotine	VSA
1	0.403	0.518	0.692	0.531	0.735	0.240	0.428	0.551	0.863	0.542	0.852	0.790	0.580	0.567	0.758	0.521	0.324	0.772
2	0.401	0.518	0.692	0.529	0.735	0.285	0.427	0.549	0.863	0.541	0.852	0.790	0.579	0.567	0.758	0.521	0.332	0.772
3	0.401	0.516	0.692	0.528	0.735	0.331	0.425	0.549	0.863	0.539	0.852	0.790	0.581	0.566	0.758	0.516	0.354	0.772
4	0.400	0.513	0.692	0.528	0.735	0.351	0.423	0.547	0.863	0.538	0.852	0.790	0.580	0.564	0.758	0.519	0.355	0.772
5	0.400	0.513	0.692	0.526	0.735	0.358	0.423	0.545	0.863	0.536	0.852	0.790	0.583	0.566	0.758	0.516	0.361	0.772
6	0.400	0.511	0.692	0.525	0.735	0.366	0.423	0.541	0.863	0.536	0.852	0.790	0.584	0.565	0.758	0.515	0.370	0.772
7	0.399	0.510	0.692	0.525	0.735	0.366	0.419	0.543	0.863	0.536	0.852	0.790	0.585	0.564	0.758	0.515	0.370	0.772

personality features (neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness) are sufficient to classify drug usage duration. Table III reveals accuracy for all drugs using 5 personality features only. The maximum accuracy is also shaded in red. There are some similarities found, first, the pattern of the number of hidden neurons used is similar. Second, the maximum accuracy achieved for each drug does not deviate significantly from the experiment using all features. Further evidence is shown in Table V which measures the difference of maximum accuracy attained by using 5 features and all features in percent. It is clear that feature reduction diminishes the prediction accuracy, indicated by minus values. However, as we look the table more closely, only accuracy of 7 drugs or 38.89% of all drugs are negatively affected by feature trimming. The degree of accuracy change varies, 5 drugs exhibit accuracy drop by less than -1%, while the remaining two have accuracy down by -2.66% and -4.39% respectively. The result for multi class classification is in line with that of binary class in which personality features are strongly associated with user or non-user groups.

TABLE V
accuracy difference for 5 features, compared to all features

Drug	Accuracy difference
Alcohol	-0.49%
Amphetamines	0.00%
Amyl nitrite	0.00%
Benzodiazepine	0.00%
Caffeine	0.00%
Cannabis	-0.81%
Chocolate	0.00%
Cocaine	0.00%
Crack	0.00%
Ecstasy	-0.37%
Heroin	0.00%
Ketamine	-0.13%
Legal highs	-2.66%
LSD	-0.70%
Methadone	0.00%
Magic mushrooms	0.00%
Nicotine	-4.39%
VSA	0.00%

Overall classification result based on personality information only is presented at Fig. 2. Two types of drugs, namely crack and heroin, are successfully classified with high degree of accuracy by more than 80%. 12 other drugs (Amphetamines, Amyl nitrite, Benzodiazepine, Caffeine, Cocaine, Ecstasy, Ketamine, Legal highs, LSD, Methadone, Magic mushrooms and VSA) are correctly recognized by moderately high level of accuracy between 50% and 80%. The remaining 4 types of drugs are hardly recognized by personality features only, indicated by low accuracy below 50%.

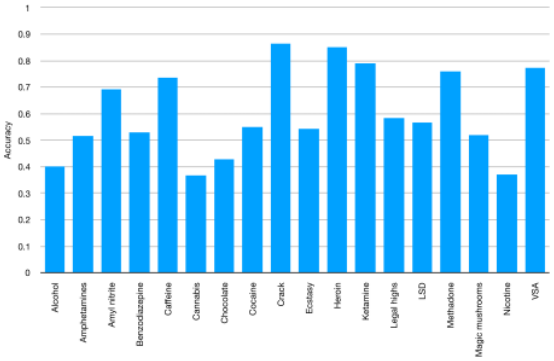

Fig. 2. Classification accuracy for each drug using 5 personality features

IV. Conclusion

This paper presents an approach to classify drugs usage duration based on personality features using extreme learning machine. Personality features only are able to perform accurate classification, compared to all available features. The performance of the classifier is not affected by the number of hidden neurons for almost all types of drugs. However, the accuracy of ELM classifier varies, depending on the type of drug. The highest accuracy achieved is 86.31% while the lowest is 36.65%. More work should be directed to identify the most appropriate features to classify a certain type of drug.

References

[1] United Nations and U. N. O. on Drugs and Crime, World drug report 2018, Vienna, 2018.
[2] Institute of Medicine (US) Committee for the Substance Abuse Coverage Study, Treating Drug Problems: Volume 1: A Study of the Evolution, Effectiveness, and Financing of Public and Private Drug Treatment Systems, D. R. Gerstein and H. J. Harwood, Eds. Washington (DC): National Academies Press (US), 1990.
[3] C. E. Grella, Y.-I. Hser, V. Joshi, and M. Douglas Anglin, "Patient histories, retention, and outcome models for younger and older adults in DATOS," Drug and Alcohol Dependence, vol. 57, no. 2, pp. 151–166, Dec. 1999.
[4] M. L. Dennis, C. K. Scott, R. Funk, and M. A. Foss, "The duration and correlates of addiction and treatment careers," Journal of Substance Abuse Treatment, vol. 28, no. 2, Supplement, pp. S51–S62, Mar. 2005.
[5] C. A. A. Ventura, J. d. Souza, M. Hayashida, and P. S. Ferreira, "Risk factors for involvement with illegal drugs: opinion of family members or significant others," Journal of Substance Use, vol. 20, no. 2, pp. 136–142, Mar. 2015.
[6] M. J. Cleveland, M. E. Feinberg, D. E. Bontempo, and M. T. Greenberg, "The Role of Risk and Protective Factors in Substance Use Across Adolescence," Journal of Adolescent Health, vol. 43, no. 2, pp. 157–164, Aug. 2008.
[7] C. Dube, M. Arora, S. Gupta, and B. Kumar, "Five Factor correlates: A comparison of substance abusers and non-substance abusers," Journal of the Indian Academy of Applied Psychology, vol. 36, no. 1, pp. 107–114, 2010.
[8] P. Costa, R. McCrae, and I. Psychological Assessment Resources, Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI). Psychological Assessment Resources, 1992.

[9] E. Fehrman, A. K. Muhammad, E. M. Mirkes, V. Egan, and A. N. Gorban, "The Five Factor Model of personality and evaluation of drug consumption risk," arXiv:1506.06297 [stat], Jun. 2015.
[10] D. Kumari, S. Kilam, P. Nath, and A. Swetapadma, "Prediction of alcohol abused individuals using artificial neural network," International Journal of Information Technology, vol. 10, no. 2, pp. 233–237, Jun. 2018.
[11] P. Nath, S. Kilam, and A. Swetapadma, "A machine learning approach to predict volatile substance abuse for drug risk analysis," in 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Nov 2017, pp. 255–258.
[12] S. G. Jacob, S. Sridhar, and N. Murugavel, "A Comparative Study on the Performance of Classifiers in Predicting Frequency of Drug Intake: A case study with Ketamine, Heroin, Crack and Meth," International Journal of Engineering, Technology, Science and Research, vol. 4, no. 9, pp. 378–383, 2017.
[13] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), vol. 2, Jul. 2004, pp. 985–990 vol.2.
[14] Guang-Bin Huang, Qin-Yu Zhu, K. Z. Mao, Chee-Kheong Siew, P. Saratchandran, and N. Sundararajan, "Can threshold networks be trained directly?" IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 53, no. 3, pp. 187–191, March 2006.

Drug usage duration classification using Extreme Learning Machine based on personality features

ORIGINALITY REPORT

6%	6%	9%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Sigit Adinugroho, Yuita Arum Sari. "Leaves classification using neural network based on ensemble features", 2018 5th International Conference on Electrical and Electronic Engineering (ICEEE), 2018 Publication	3%
2	arxiv.org Internet Source	3%

Exclude quotes	Off	Exclude matches	< 3%
Exclude bibliography	Off		