DETECTING SEVERITY OF NICOTINE ADDICTION VIA EYE TRACKING

A PREPRINT

Mayank Sharan
Computer Science and Engineering
Indian Institute of Technology
Delhi
mayanksharan96@gmail.com

Rahul Garg
Computer Science and Engineering
Indian Institute of Technology
Delhi
rahulgarg@cse.iitd.ac.in

December 21, 2019

ABSTRACT

Smoking addiction is a social epidemic affecting about 1 in 5 adults in the world directly and many more through passive smoking. Diagnostic tools to determine the severity of the addiction of nicotine on the other hand are either not very reliable or are very restrictive in when, where and by whom can they be used. This makes the procedure of identification and rehabilitation less accessible to those who need it. This paper explores an effort to see if we can create a computer based test to evaluate the severity of nicotine addiction using a simple eye tracking test.

1 Introduction

WHO [1] estimates say that 20.2% of the world's population aged over 15 were smokers in 2015. The estimates look especially bad in the European Region where the prevalence was estimated to be 38.7% in males and 21% in females in 2015. Tobacco smoking is well known to cause a plethora of health problems all the way up to lung cancer. Important characteristics of smoking that make it very dangerous are it is addictive and second hand smoking can cause health problems for people who do not smoke themselves.

The treatment of of any addiction adheres to the tangible outcomes of helping the person stopping the usage, stay free of the drug and be productive in the society. Depending on the severity of the addiction this treatment can range from use of nicotine patches and gum and taking help from support groups to proper rehabilitation at a medical facility with medicine, oversight and regular follow-ups. In a scenario with limited medical resources identification of what measures to administer to whom becomes a major requirement to be able to tackle addiction as a social epidemic.

The severity of nicotine is usually measured by a test with multiple question items designed to assess several aspects of the addiction. Available evaluation methods [2] [3] of this nature include Addiction Severity Instrument (ASI), Cigarette Dependence Scale (CDS), Chemical Use Abuse and Dependence Scale (CUAD), Fagerstrom Test for Nicotine Dependence (FTND) [4], Nicotine Dependence Syndrome Scale (NDSS), Severity of Dependence Scale (SDS), Substance Use Involvement Index (SUII), Tobacco Dependence Screener (TDS) and Wisconsin Inventory of Smoking Dependence Motives (WISDM).

The other way to detect the amount of recent use of nicotine is by detecting one of the numerous biomarkers [5] that result from smoking tobacco. The most common methods being measuring breath carbon monoxide, testing saliva, blood, urine and hair. The markers are Nicotine, Cotinine, Antabine, exhaled CO, Carboxyhaemoglobin, Acetonitrile, Thiocyanate etc.

The above summarized measures come with limitations of their own. The subjective measures are human administered tests that inherently makes them prone to bias and manipulation. Also some of the tests requires highly trained professionals to administer them, limiting the environments and settings they can be used in. The objective measures are invasive in nature and take some time to process. The resources required to perform these tests again limits the outreach and availability of these tests.

FTND is one of the most widely used scale to measure nicotine dependence severity [6]. It is one of the most objective human administered questionnaires for addiction severity. It assigns a score from 0-10 based on the answers given to 6 multiple choice questions. The simplicity and the objective nature of the test are why we are going to try and predict FTND scores using the eye tracking measurements.

The theory of incentive salience [7] postulates that addictive drugs enhance the pleasure rewarding dopamine pathways. These pathways attach "incentive salience" as an attribute to related stimuli - Incentive Salience refers to the want experienced for a rewarding stimulus. Repeated usage increases the aforementioned effect perhaps to permanency. These effects can happen independent of the subjective "pleasure" associated with the drugs. Attentional bias [8] is the tendency of perception to be affected by recurring thoughts.

In this paper motivated by the theories of incentive salience and attentional bias, we designed a few paradigms that extract semantic features and using those features we learn machine learning models to assign weights to those features so that they fit to the FTND scores. The test includes four tasks.

The first being an anti saccade measurement where an image is displayed at one of four positions on the screen and the subject is instructed to look away from the image. The second task asks the subject to focus on a cross traversing a circular trajectory on the screen while an image cue is displayed at the center of the screen. The third task entails the subject focusing on a cross at the center of the screen while an image follows a circular path. The fourth and the final task is to memorize a sequence of 6 digits that appear on the screen while a video is playing and enter the sequence at the end.

In each of these tasks we have smoking and non smoking cues. To the best of our knowledge no prior study tries to use ML approaches on eye tracking data to predict the severity of addiction scores. There is one that comes close which tries to correlate smoking addiction severity with eye tracking and FMRI metrics [9]. Similar studies have tried to do the same thing for different drugs like cocaine, alcohol, morphine and methamphetamine.

Our initial results are very promising. Even though any of the metrics on an individual level does not display a strong correlation with the FTND scores, a combination of them gives us a reasonable model to compute FTND scores with good confidence. We will go into details of the same in the results section.

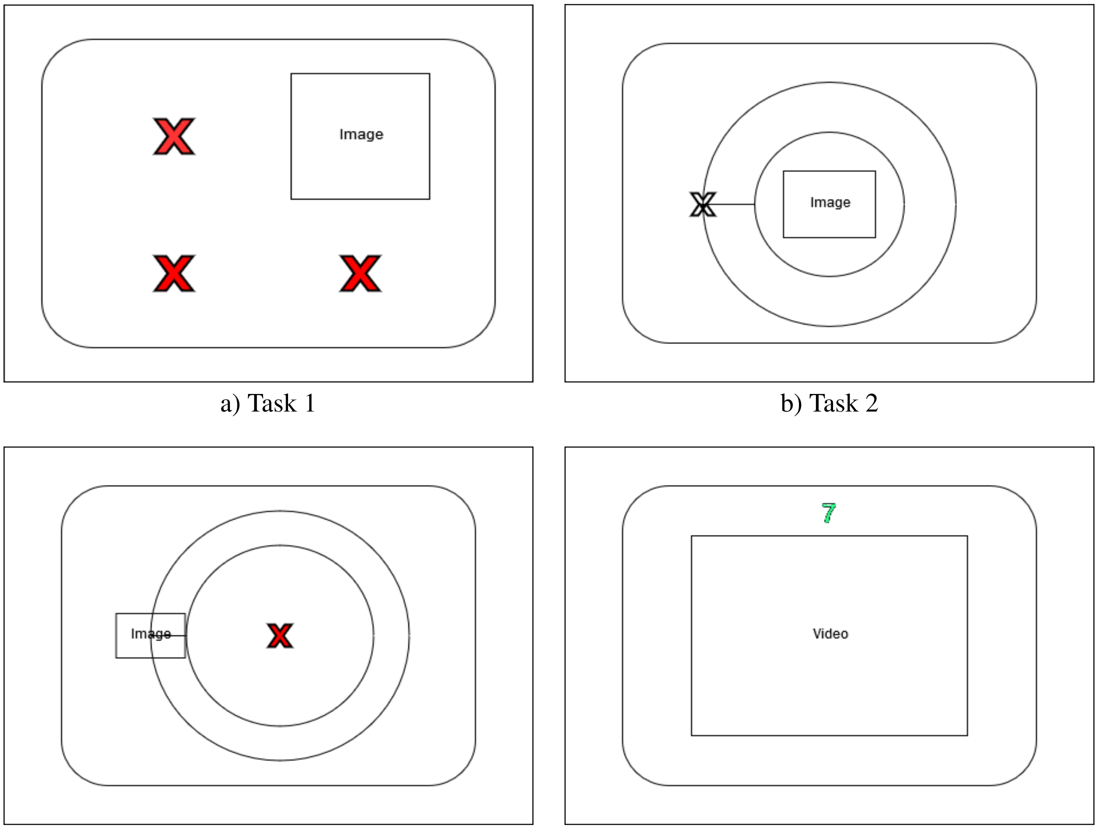
The paper will further go into the depth of the design of the task, then we will explain how the study was performed and the data collected. The next section then highlights the analysis that was performed on the data collected and a summary of the results that we got along with the interpretation is what follows. Towards the end we bind this together with inferences that are generated from the study and future work that can be done building on this.

2 Materials and Methods

Incentive salience as a theory has gone under iterations and further research since it was first proposed in 1993 by T. E. Robinson and K. C. Berridge [7]. It is one of the fundamental motivations behind the design of this task. This theory set out to explain the psychological and neurobiological basis of drug craving. Under this it posits that addictive drugs share the ability of enhance the mesotelencephalic dopamine neurotransmission, one of the functions of which is to attribute 'incentive salience' to perception and mental representation of stimuli associated with activation of the system. Simplified this means that when a person is presented with a stimulus that can activate this system the psychological process of incentive salience transforms the mental translation of it imbuing them with 'salience' i.e. making them attractive, 'wanted' stimulus. The takeaway here is that the brain responds differently to stimuli based on the craving associated with them, especially drugs.

Building on this is the theory of attentional bias [8]. Attentional bias is the tendency of perception to be affected by recurring thoughts. Attentional bias and craving have a mutually excitatory relationship thus giving basis for correlation to be seen in the resulting behavior. This background support forms the basis of the tasks being designed to quantify the response to visual stimuli that elicit craving.

The study works in a simple sequence of steps. The test is taken by the subject where the eye tracking mea-



a) Task 1 b) Task 2 c) Task 3 d) Task 4

Figure 1: Illustrations of the Tasks

surements are stored specifically the x and y coordinates of the gaze on the screen. After that the raw data is processed to generate semantically relevant features for each task. These features are then passed into a machine learning model that generates a score that is an estimate of where the subject lies on the FTND scale.

In the first task an image (smoking related or neutral) is displayed randomly in one of four locations on the screen with the subject instructed to look away from the image and at a red cross present at each of the three positions the image is not displayed (Figure 1a). This task aims to measure anti saccade response of the subject i.e. the reaction time of the subject to execute the instruction of detaching his/her gaze from the stimulus.

In the second task an image (smoking related or neutral) is displayed at the center of the screen while a pointer (red cross) traverses a circular trajectory where the radius changes after every completed circle around the cue (smoking related or neutral) (Figure 1b). The subject is instructed to follow the motion of the pointer. This task aims to measure the pro saccade response from the subject, basically how many times does the subject deviate from following the pointer and has a look at the image instead and also the duration of the same.

In the third task a red cross is displayed at the center of the screen while an image (smoking related or neutral) traverses a circular trajectory around the cross where the radius changes after every completed circle (Figure 1c). The subject is instructed to fix their gaze at the centre of the screen. This task is also a pro saccadic measurement but in a different flavor from the previous one. This is a non stationary stimulus and the task of focusing on the center of the screen is less engaging than the instruction to follow a pointer in a trajectory. Similar measurements as the previous task are made.

The fourth task is a modified version of the n-Back test. A video containing smoking related stimulus and one containing a neutral stimulus is played in a size lesser than that of the screen leaving space above and below the video (Figure 1d). During the video 6 digits appear one by one at random above or below the video . At the end of the

Table 1: Participant Demographic Data

FTND Score	Number of Participants	Age	Number of Participants
0	15	18	3
1	4	19	5
2	1	20	7
3	2	21	5
4	3	22	7
5	0	23	1
6	3	24	0
7	2	25	2

Table 2: Features

Task	Feature Name	Feature Description
Task 1	Reaction time (S)	Average time taken to look away from a smoking cue
Task 1	Number of returns (S)	Average number of times the gaze returns to smoking cue
Task 1	Time of return (S)	Average time for which the gaze returns to smoking cue
Task 1	Reaction time (N)	Average time taken to look away from a neutral cue
Task 1	Number of returns (N)	Average number of times the gaze returns to neutral cue
Task 1	Time of return (N)	Average time for which the gaze returns to neutral cue
Task 2,3	Time on Cross (S)	Average time with gaze on cross for smoking cue
Task 2,3	Time on Image (S)	Average time with gaze on image for smoking cue
Task 2,3	Time on Cross (N)	Average time with gaze on cross for neutral cue
Task 2,3	Time on Cross (N)	Average time with gaze on image for neutral cue
Task 4	Time on Video (S)	Average time with gaze on video for smoking cue
Task 4	Time above video (S)	Average time with gaze above video for smoking cue
Task 4	Time below video (S)	Average time with gaze below video for smoking cue
Task 4	Percentage match (S)	Percentage of match of the sequence entered after smoking cue
Task 4	Time on Video (N)	Average time with gaze on video for neutral cue
Task 4	Time above video (N)	Average time with gaze above video for neutral cue
Task 4	Time below video (N)	Average time with gaze below video for neutral cue
Task 4	Percentage match (N)	Percentage of match of the sequence entered after neutral cue

video the subject is asked to reproduce the sequence. This task measures the attentiveness via the measure of how much of the sequence the subject gets right and we also compute the duration of the gaze on the video that plays and duration of the gaze being elsewhere.

2.1 Data Collection

Given that the study collected data from humans due procedure was followed and approvals were taken. The study was approved by the Institute Ethics Committee of IIT Delhi for a participant pool size of 30. The study had strict measures to ensure that the data collection was done ethically. The subjects were students from IIT Delhi. All subjects were male. Complete confidentiality was ensured to all subjects as in each subject was assigned a study identifier which was the only tag associated with any of the data collected. Informed consent was obtained from each subject and they were provided with written details of the task which was explained to them with the contact details of the investigators conducting the study.

FTND [4] [10] was administered to the subjects to obtain smoking addiction severity score and a modified version of MINI was used to evaluate if the subject satisfied any of the exclusion criterion based on psychiatric disorders. Participants with physical problems obstructing participation in the study and ones with history of learning disability or traumatic brain injury were also excluded from the study. Participants scoring a 0 on the FTND were classified into the control group and those scoring higher than 0 formed the nicotine dependent group. There were 15 subjects in each group. Measures were taken to ensure that the distribution of the nicotine dependent group was as varied on the FTND score scale as possible. The task was performed by the subjects in an isolated light adjusted environment with standardized position of the head w.r.t. the screen and the eye tracker.

Table 3: Lasso Regression Results

Metric	Task 1	Task 2	Task 3	Task 4	Total
Mean Squared Error	5.50	5.06	5.32	4.91	2.11
Mean Absolute Error	2.02	1.83	1.90	1.95	1.09

2.2 Feature Selection and Model Building

All features that have been extracted from the ask for the purposes of further analysis are based on the semantic sense they make. It is definitely possible to go ahead and dig much deeper into the data to look for features at a more granular level but that comes with the risk of capturing artifacts which aggregate measures usually even out making them more reliable. In each task each feature mentioned is calculated as aggregate for all smoking cues separately and for all neutral cues separately.

In task 1 we measure the initial reaction time taken to look away from the image, number of times the gaze returns to the image and the average time spent with gaze on the image on a return after the initial reaction. In task 2 and 3 we measure the time spent on the cross and the time spent on the image. In task 4 we calculate the time on the video, time above the video, time below the video and the percentage match of the sequence of 6 digits entered with the sequence entered. The features are summarized in table 3.

The features that have been extracted from the data for each task have been subjectively chosen. The features are designed carefully to strike a balance between semantics and the level of granularity we can extract the features at without it becoming too small a sample size for the features to not be outliers.

In the models other than just the features as it is we also feed in the products of those features to see if they are picked up for modeling. The product based features were assigned very low weights and the models with just the linear features performed better.

We fit different regression models to take in these features and output a score which is supposed to be a prediction of the FTND score. We fit models for features from each task separately and models for all features combined from the 4 tasks. All features are normalized to values between 0 and 1 before the models were trained.

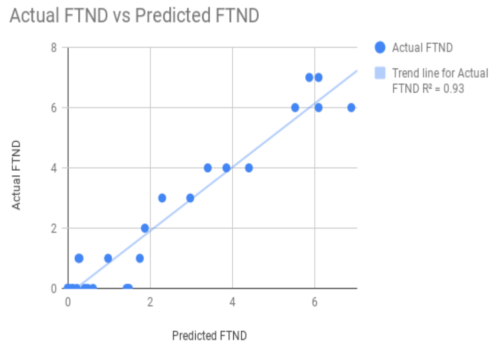
We tried lasso regression, SVM regression, ridge loss regression and random forest regression[11]. Random forest regression was set to 3 estimators and a maximum depth of 10 for individual tasks and 10 estimators and maximum depth of 35 for the analysis when using all the features. The ridge loss model had the penalty factor set at 1.0 since that gave the best results in the search space. The results for each of them are summarized in the tables 4-7. As the evaluation metric we run leave one out cross validation and based on the predicted FTND scores we calculate mean squared error and mean absolute error across the 30 samples.

3 Results

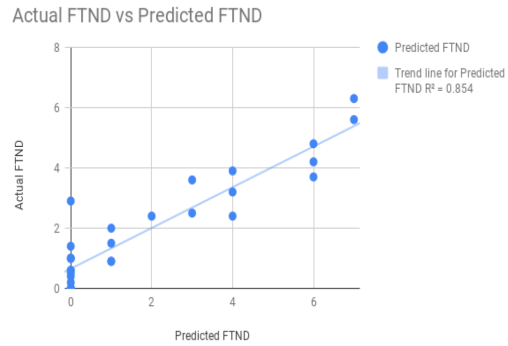
The results are reasonably good for models that are not trying to overfit and have data as small as a sample size of 30. Given that these 30 might also have some outliers the error metrics are suggest that the models are doing pretty good. The best performing model across the board is the random forest, followed by the SVM and then the ridge loss and the lasso. Comparing models on the combined data results ridge loss and the random forest are the best while Lasso and SVM do not perform as well.

The metrics are in compiled in tables 4-7 for each of the models. To compute these metrics with the leave one out cross validation technique, we train on 29 inputs and predict the 30th and compute square and absolute values of the difference between the actual and predicted FTND. We take the average of these values across the 30 inputs being predicted and call them mean squared and mean absolute error respectively. Task i refers to the metrics for when model is fit and evaluated using only the features extracted from task 1. Total refers to the metrics from the model trained and evaluated using features from all 4 tasks.

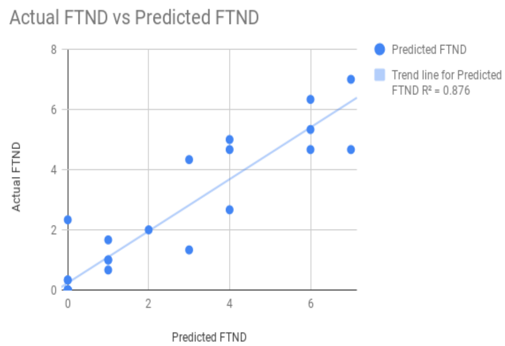
Another interesting observation is that if we set all negative prediction values to zero for the combined data ridge loss model the mean absolute error becomes 0.47 and the mean squared error becomes 0.43 (Figure 2a), making it the best model in that scenario. The insight we get from this is that the models are having trouble accommodating the variation in data amongst the control group and mapping them to the same FTND score of zero.



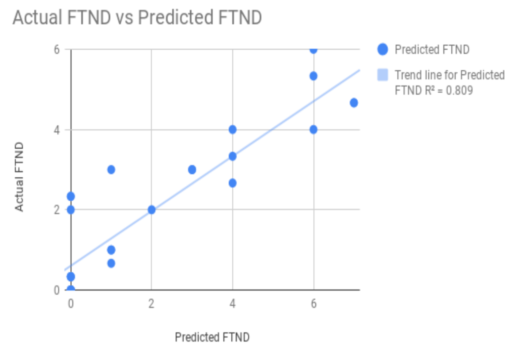
a) Ridge loss Total scatter plot



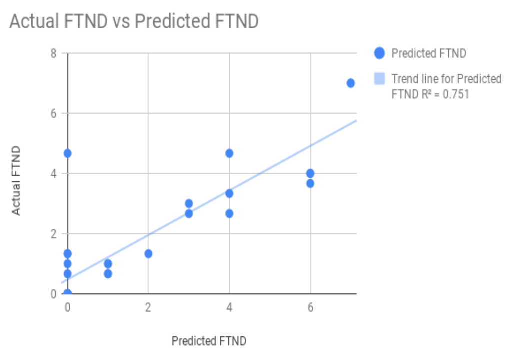
b) Random Forest Total scatter plot



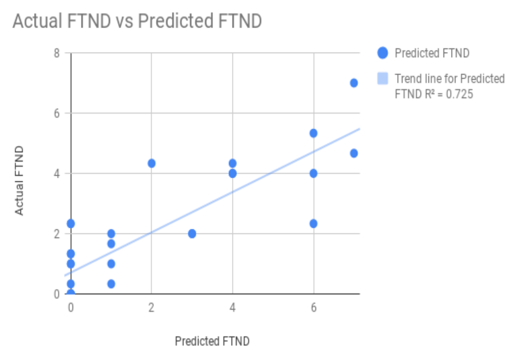
c) Random Forest Task 1 scatter plot



d) Random Forest Task 2 scatter plot



e) Random Forest Task 3 scatter plot



f) Random Forest Task 4 scatter plot

Figure 2: Scatter Plots

Table 4: SVM Regression Results

Metric	Task 1	Task 2	Task 3	Task 4	Total
Mean Squared Error	3.99	3.98	4.16	3.98	3.99
Mean Absolute Error	1.13	1.14	1.17	1.13	1.13

Table 5: Ridge Loss Regression Results

Metric	Task 1	Task 2	Task 3	Task 4	Total
Mean Squared Error	5.38	5.06	5.31	4.42	1.09
Mean Absolute Error	1.99	1.82	1.89	1.82	0.70

Table 6: Random Forest Regression Results

Metric	Task 1	Task 2	Task 3	Task 4	Total
Mean Squared Error	0.72	1.23	1.45	1.64	1.07
Mean Absolute Error	0.49	0.65	0.65	0.84	0.76

Table 7: Trendline fit and R squared results

Model	R^2	Trendline Slope	Trendline Intercept
Random Forest Task 1	0.876	0.861	0.238
Random Forest Task 2	0.809	0.683	0.603
Random Forest Task 3	0.751	0.739	0.475
Random Forest Task 4	0.725	0.668	0.709
Random Forest Total	0.854	0.676	0.655
Ridge Loss Total	0.930	1.06	-0.207

The chosen model is the Random forest regression model because that allows for more flexibility to accommodate for variations in data of the control group. The performance of the random forest regression is best using only task 1 features where the mean absolute error is 0.49 and the mean squared error is 0.72. The scatter plots are presented in figure 2.

Task 1 model assigned significant weights to all features except average number of returns in smoking stimuli. Task 2 model assigns almost equal weightage to the four features. Task 3 model also selects all features like task 2 but gives higher weightage to the features from the neutral stimuli than the features from the smoking stimuli. Task4 model assigns low weightage to the percentage match of sequence from both the smoking stimulus as well as the neutral stimulus.

The combined model gives low weightage to task 1 features of average number of returns for smoking stimuli, the average duration of return for smoking stimuli, task 2 features of average time on image for smoking stimuli, average time on cross for neutral stimuli and the task 4 feature of percentage match of sequence for the neutral stimulus.

4 Discussion and Future Work

The current results are encouraging and definitely suggest that further investigation be done into this. Replacing FTND as an objective measure is definitely possible once this result has been verified and tested thoroughly by other studies. Eye tracking would be a preferred alternative to FTND since it is much harder to manipulate.

In terms of future direction this protocol needs to be tested on more subjects and be verified. This protocol also needs to be cross checked with other subjective and objective measures to validate the accuracy and usability of this paradigm.

References

[1] WHO. Who global report on trends in prevalence of tobacco smoking 2000-2025, second edition. May 2018.

[2] Kevin P. Conway, Janet Levy, Michael Vanyukov, Redonna Chandler, Joni Rutter, Gary E. Swan, and Michael Neale. Measuring addiction propensity and severity: The need for a new instrument. *Drug Alcohol Depend*, 111((1-2)):4–12, September 2010.

[3] Naomi Sato, Tomonori Sato, Akiko Nozawa, and Haruhiko Sugimura. Assessment scales for nicotine addiction. *Journal of Addiction Research and Therapy*, S1:008, January 2012.

[4] Heatherton TF, Kozlowski LT, Frecker RC, and Fagerstrom KO. The fagerstrom test for nicotine dependence: a revision of the fagerstrom tolerance questionnaire. *Br J Addict*, 86, 1991.

[5] Neal L. Benowitz, Janne Hukkanen, and III Peyton Jacob. Nicotine chemistry, metabolism, kinetics and biomarkers. *Handb Exp Pharmacol*, 192:29–60, October 2010.

[6] Joseph R. DiFranza, Robert J. Wellman, Judith A. Savageau, Ariel Beccia, W. W. Sanouri A. Ursprung, , and Robert McMillen. What aspect of dependence does the fagerström test for nicotine dependence measure? *Hindawi*, 2013:8, October 2012.

[7] Robinson TE and Berridge KC. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research Review*, 18(3):247–291, September 1993.

[8] Field M and Cox WM. Attentional bias in addictive behaviors: a review of its development, causes, and consequences. *Drug Alcohol Depend*, 97(1-2):1–20, September 2008.

[9] O-Seok Kang, Dong-Seon Chang, Geon-Ho Jahng, Song-Yi Kim, Hackjin Kim, Jong-Woo Kim, Sun-Yong Chung, Seung-In Yang, Hi-Joon Park, Hyejung Lee, and Younbyoung Chae. Individual differences in smoking-related cue reactivity in smokers: An eye-tracking and fmri study. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 38:285–293, April 2012.

[10] Pomerleau C S, Majchrezak MI, and Pomerleau OF. Nicotine dependence and the fagerstrom tolerance questionnaire: a brief review. *J Substance Abuse*, 1, 1989.

[11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.