A machine learning-based approach to forecasting alcoholic relapses

Conference Paper · October 2019

CITATIONS
4

READS
2,322

3 authors:

Nikola Katardjiev
Uppsala University
1 PUBLICATION **4** CITATIONS
SEE PROFILE

Steve McKeever
Uppsala University
74 PUBLICATIONS **663** CITATIONS
SEE PROFILE

Andreas Hamfelt
Uppsala University
49 PUBLICATIONS **345** CITATIONS
SEE PROFILE

A machine learning-based approach to forecasting alcoholic relapses

Nikola Katardjiev, Steve McKeever, and Andreas Hamfelt

Department of Informatics and Media, Uppsala University,
Kyrkogårdsgatan 10, 753 13 Uppsala, Sweden
{nikola.katardjiev,steve.mckeever,andreas.hamfelt}@uu.im.se
https://www.im.uu.se

Abstract. This research aims to explore alcoholic relapses by modelling four types of machine learning algorithms on clinical trial data of patients in an alcohol addiction treatment plan provided by an Uppsala-based company called Kontigo Care, with the goal of predicting downwards trends in the data that could indicate a relapse. The learning algorithms were support vector machine, random forest, decision trees, and a K-nearest neighbour regressor. Results indicated that using a random forest model, it is possible to forecast future data entries in the particular data set by beating the benchmark of a baseline predictor, suggesting that alcoholism could indeed be predicted, and thus future work leveraging larger data sets and deep learning methods could improve relapse predictions even further.

Keywords: Alcohol abuse, relapse, machine learning, time series forecasting.

1 Introduction

Irresponsible alcohol consumption is a wide-spread issue plaguing most modern countries; alcoholism is the third-leading cause of death in Russia [1], and it contributes to over 100,000 deaths in the United States alone [2]. Indeed, such is its effect that several countries, such as Sweden and Finland, have taken action to downsize and restrict the consumption of alcohol by the local population through restricting the sale to government-owned stores, Systembolaget and Alko, respectively [3].

However, individuals who try to restrict their alcohol consumption voluntarily often fail to do so [4], putting them at risk for a relapse following any temporary restriction they may have made. For those in recovery programmes and healthcare plans, a relapse following a substantial investment on the part of the caretaker can only be considered a failure and a sunk investment; it is estimated that relapse rates are usually found at around 60%, going as high as 90% in some cases [5]. Current research in addictive sciences is advising stronger patient monitoring and long-term treatment plans [6], which are argued to aid in

reducing relapse rates [7]. Solutions for monitoring these types of plans, however, are oftentimes rooted in more traditional techniques [8].

We seek to assess the efficacy of a machine learning approach to this domain. While healthcare at large has seen successfully applications of machine learning [9], the domain of addiction care has not been investigated thoroughly [10]. Given that machine learning has been applied to forecast other types of behaviours and trend changes, it follows that a careful application of those routines and appropriate transformations (given appropriate input data that corresponds to patient behaviour) should make it possible to forecast trend changes in alcohol abuse behaviour.

It is not unreasonable to then assume that there is a substantial knowledge gap in the development of heuristics for machine learning applications in a specialised field such as that of alcohol abuse treatment. This research aims to provide a proof-of-concept by modeling several common machine learning models against monitoring data of patients suffering from alcohol addiction. Due to the exploratory nature of our work, the data selection is taken from a single source, an Uppsala-based eHealth company called Kontigo Care AB [11].

The paper is structured as follows. In the next section we discuss past work in artificial intelligence being used in addiction monitoring and patient treatment. We then show our proposed implementations and explain the nature of the data set used in this research, before discussing the test results. Finally, we conclude and suggest various avenues for future work.

2 Related work

Only a handful of papers appear to have tackled the issue, such as the case of [8], which developed a classification-based approach to identifying relapses during ongoing treatment plans. They compared two models, a decision tree classifier and a Bayesian network, on 73 in-treatment patients, mapping outcomes as either successful or relapsing, with both models yielding a success rate higher than 50%. However, the research only modelled whether or not the patients would relapse following a completed treatment, whereas our research attempts to model relapses in ongoing treatment plans.

Modern research concerning time series forecasting can primarily be found in the field of machine learning; a paper by [12] details the development of an LSTM network capable of not only forecasting future time steps, but also performing anomaly detection on that time series. In other words, it can forecast a series and anticipate when aberrant behaviour will occur x time steps ahead.

Similarly, [13] used a recurrent ARIMA-like neural network to forecast several types of time series, ranging from Turkish stock prices to Australian beer consumption, generating a non-linear auto-regressive function as its output. This type of hybrid network has been developed in previous examples as well; [14] tested a case using a feed-forward neural network and an ARIMA model in tandem to forecast stock markets with results surpassing either independently.

There is evidence to suggest that hybrid ensemble models such as these can have improved performance on time series forecasting.

In the area of healthcare, machine learning can indeed be used to make predictions on patient data. Previous experiments in the domain have been used to predict myocardial infarctions [9], and machine learning was used to track the progression of Alzheimer's disease [15].

These cases are some strong samples of how machine learning is currently used to model predictions for time series forecasting, and we propose testing similar approaches in the domain of relapse forecasting.

3 Training and testing data

In 2018, a Sweden-based eHealth company called Kontigo Care published the results of a longitudal set of clinical trials done on 30 patients suffering from alcohol addiction [11]. The study was conducted over the course of an entire year, with the objective of monitoring, several times per day, alcohol levels in these patients, and this data forms the basis of our research. The clinical trial was undertaken in conjunction with Kontigo Care's Previct Alcohol product, a breathalyzer tool that patients can use to self-report their own behaviours, up to four times per day, each usage resulting in a unique measurement.

The breathalyzer results are calculated into a construct called Addiction Monitoring Index (AMI), an exponentially smoothed value scaling from 0 to 100 which acts as a metric for how well a patient is performing. An average AMI of 0 means that a patient has entered a full alcoholic relapse, while an AMI of 100 suggests that the patient has not missed any measurements, and each breathalyzer test has shown no traces of alcohol. The original research, from which the data stems, suggested that AMI is a valid method "for monitoring the recovery process and for early identification of lapse/relapse patterns" [11]. An additional study undertaken by the same company was conducted for a digital biomarker called Maximum Time Between Tests (MTBT), which is simply the difference in time between any one test and the previous. 54 patients were self-reporting using the same breathalyzer tool as in the previous clinical trial. The results showed that there was a correlation between MTBT and phosphatidyl ethanol (PEth) in the blood. The authors report the following: "The time-based digital biomarker maximum time between tests described here has the potential to become a generally useful metric for all scheduled measurement-based eHealth systems to monitor test behaviour and compliance, factors important for dosing of eHealth systems and for early prediction and interventions of lapse/relapse." [16]

MTBT is an indicator of a particular quirk of the dataset; a missed measurement (i.e., a null value) has some meaning with regards to how AMI is calculated. If a patient is expected to perform four tests, and one of them is skipped or otherwise not executed, it is nigh equivalent to having failed one breathalyzer test. In a sense, a null value does, in this dataset, still contain some information. While there is no guarantee that a patient has been drinking during the missed measurement, it is still a useful indicator that they could have done so.

Table 1: Example data of a fictive patient. AMI, on the far right, is the label, while all other columns sans "Entry" - are features.

Entry	**BAC**	**MTBT**	**PEth**	**AMI**
13	0.04	9.669	0.025	97.3
14	0.002	12.956	0.03	99.5
15	0.001	19.643	0.025	97.0
16	0.08	12.102	0.05	95.4
17	0.02	21.745	0.25	98.2
18	0.01	11.243	0.01	99.3
19	0.005	20.418	0.01	100.0

While MTBT can be a useful indicator of alcohol intake at a given moment, the aforementioned AMI is composed of several other key components. Phosphatidyl ethanol levels in the blood (PEth), blood alcohol content in permille reported in a measurement (BAC), and AMI in the previous measurements are all used in the calculation of the next measurement's AMI. For the purposes of this experiment, MTBT, PEth, BAC, and the previous AMI can be considered dataset features, while the next AMI measurement is the label.

Table 2: Descriptive statistics of relevant columns in the data set

	BAC	**MTBT**	**PEth**	**AMI**
Count	13959	13959	13959	13959
Mean	0.035918	0.0042	0.03	62.29
Std	330.69	0.0773	3.230	32.30
Min	0.00	0.000	0.000	0.012
25%	0.02	13.000	0.000	39.34
50%	0.06	15.000	0.000	72.51
75%	0.01	33.000	0.000	90.28
Max	4.822	3648	4.100	100.0

See Table 1 for an example of how the data can be structured. Note that the data is fictional, and does not represent a real patient in the clinical trials. The data set used for this research consists of the original research data for the clinical trials of 30 patients. The descriptive statistics of MTBT, AMI, PEth, and BAC are illustrated in Table 2.

Figure 1 shows the AMI curve of a single patient across approximately 200 days treatment. AMI is an exponentially smoothed value, meaning as it approaches its upper and lower bounds, it becomes increasingly difficult to 'stay' in those values; a missed measurement at a high value with cause a big drop, for example. In this research, AMI can be considered an output value ranging from 0 to 100, with a maximum change between any two time steps of 21%. Figure 1 provides an example of how AMI is smoothed towards the upper bounds.

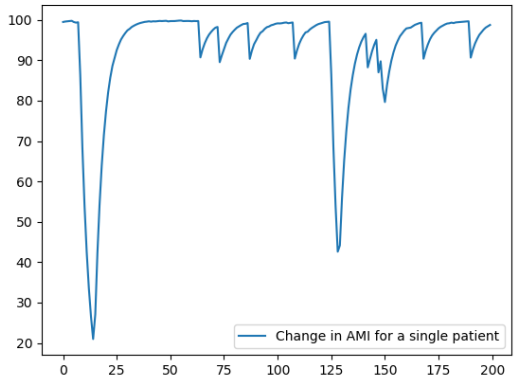


Fig. 1: Exponential smoothing in AMI of a single patient

The fact that the data is exponentially smoothed is particularly relevant, as maintaining a high AMI becomes progressively harder the closer one is to 100; in order to maintain that level, all breathalyzer tests must pass and none may be missed. Similarly, in order to stay at 0, a patient must be constantly failing or skipping all their tests. Hence, if the behaviour of patients was entirely random, the distribution of AMI would be expected to be primarily in the middle values, with 0 and 100 being particularly rare. A histogram of AMI values, Figure 2, shows that this is not the case:

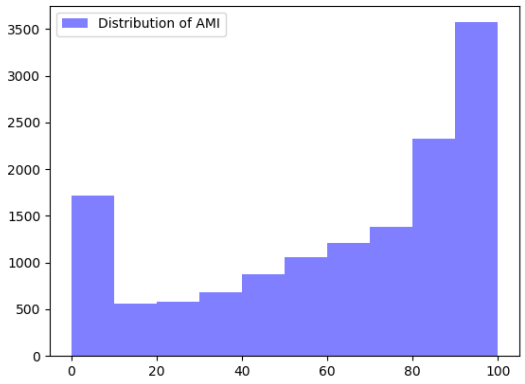


Fig. 2: A histogram showing the spread of AMI across the patients

Indeed, the histogram reveals that 0 and 100 are the most common values in the data set. This is a strong indicator that once patients fall into a lapse/relapse pattern, they tend to stay there for long periods of time. Similarly, a patient who is performing very well is likely to continue performing well. Hence, it becomes increasingly important to attempt to forecast increased relapse risks, as previously well-performing patients could enter a lapse/relapse pattern and subsequently stay in that pattern for a substantial time.

The model fitting will not be done on the data set only as is; a relapse is not determined only by whether or not a person has been consuming alcohol in the last measurement, but rather by a longer time sequence. It stands to reason, that past measurements can serve as indicators of a relapse occurring. As an abstract example, whether or not a day is a weekend could act as an indicator of cyclical behaviour; patients who maintain working lives or other weekly responsibilities may be consuming alcohol only on weekends, and by only considering the previous measurement in the data, such an occurrence would be missed. As expressed in To accommodate this, the data set is time shifted an arbitrary amount of steps in the past, like so: $y'_t = prediction(y_{t-3}, y_{t-4}, ..., y_{t-n})$.
Note that y'_t is the predicted AMI 3 days in advance. The data is time lagged to minimise the impact of the limited change in AMI from time step to time step; as the data is exponentially smoothed, any prediction of y'_t given y_{t-1} is flawed, as whatever number is provided, it can only change by at most a few percentage points, so any prediction within that range is decent. By lagging the series with two additional days, the amount of possible outcomes increases drastically.

For our research, in order to account for patient history, the usage of time lag and a time window is employed; the experiment is rendered once with a time lag of 30 and a window of 5 (in other words, providing inputs for days 1-30 to predict the output of day 35). The window needs to be sufficiently large such that exponential smoothing does not become a relevant factor in predictions; given that the smoothing factor limits how much the data can change day-by day, so by increasing the window we reduce the amount of variance that can be explained by the smoothing. Based on this, if the window was set to only 1, well over 90% of the variance could be explained by exponential smoothing alone.

By extending the time lag to a state where the entire range of 0-100 is included, it becomes significantly more difficult to make accurate predictions by repeating the previous measurement's AMI with some random noise added.

The models will be fitted and trained on this transformed data set. We increase the range to overlap the past month of measurements; this causes significant overlap between entries; much of the data between entries y_t and y_{t-1} will simply be repeated data. We argue that this redundancy of data is necessary to capture the behaviour of the patients in a representative fashion.

4 Implementation and approach

This research serves as a proof-of-concept that alcohol abuse patterns are to some degree predictable several days in advance, thereby making them preventable.

The objective is to fit a data set that in some way models alcohol usage over time onto machine learning models, and to show that it is possible to outperform baseline predictors. We compare four common machine learning algorithms – decision trees, random forests, support vector machines, and K-nearest neighbours – against a baseline predictor and compare highest degree of explained variance.

4.1 Model development, training process, and data collection

Each model was developed and tested independently, but the process of data collection (having developed and ensured the stability of each respective algorithm) followed the same three-step approach:

1. Train the algorithm using the split training data.
2. Test the resulting model with respect to the unseen test data.
3. Perform optimisations and validations to examine algorithm performance.

In this case, data collection is performed through the form of the *root mean squared error* (RMSE) and the *explained variance* (r^2 of the model predictions against real observations). Data collection is also undertaken by examining predicted trend changes in each model to examine their capabilities as forecasting tools. As the data set is rather small (13,959 entries, as shown later on), we use an 80/20 split of training/test data, using K-fold cross validation to ignore the use of validation data.

In order to provide baseline comparisons, which add relevancy to the models rating, we include a baseline predictor that simply returns the last known value in the time series; we also test a predictor that outputs the mean difference in AMI. This is a simple yet effective strategy that has seen usage in time series forecasting [17]. The mean difference is calculated only from the same training sets that the machine learning models are exposed to; this is to ensure that the testing data remains completely unseen. We define the baseline predictor as such: $y'_t = y_{t-5} + \textit{mean difference}\ (\textit{training data})$.
The mean difference is calculated exclusively from the training data, so as not to give it an unfair advantage on the testing set. It is assumed that a model that can outperform the baseline predictor should contain some power in predicting alcoholic relapses in the modeled data.

4.2 Algorithms

We use a single *decision tree* as our first tested model. They have a high degree of transparency, and if the data can be modeled and predicted in a simple enough manner, it may be sufficient to use a single tree for our purposes. Conversely, if the problem cannot be modeled simply enough, we expect the decision tree to overfit onto the training set. The second predictor is a *random forest* [18]. Studies [19] appear to indicate that they show more promise in solving regression problems with higher levels of accuracy. We argue that our inclusion of a random forest is predicated on the case that decision trees do not have sufficient

explanative power to cover a problem on this nature. Finally, we use a *support vector machine* (SVM) fitted with a *radial basis function* (RBF) kernel to model the data in the same manner as the previous two models, as well as a K-nearest neighbour regressor. This is to help show if models perform differently on the data set, helping suggest that they are not forecasting the same thing.

The models were trained on 80% of the samples, selected semi-randomly; for any given patient, all entries of that patient needed to be in either the training set or the testing set. Splitting data from the same patients into both data sets would give the models an unfair advantage on predicting on patients it's already been trained on, which would cause the test set to overlap with the training set and make it less reusable for new patients. If sample Y_t is found in the training set, it does not make sense to add Y_{t-1} to the test set, as a significant amount of data is repeated, only in different columns.

5 Results

Once trained on the appropriate data, the models were tasked with making predictions on both the training and testing sets, in order to show not only how well they performed compared to each other, but also to see to what extent overfitting may have played a part in their performances, as shown in Table 3.

Table 3: Summary of algorithm performances. We show root mean square error of train and test data, as well as the explained variance for each model.

Model	**Train RMSE**	**Train Variance**	**Test RMSE**	**Test Variance**
Decision Tree	0.077	1.000	18.091	0.646
Random Forest	4.908	0.977	12.124	0.841
Support Vector regressor	23.344	0.489	28.183	0.142
K-nearest Neighbour regressor	12.780	0.847	14.237	0.781
Baseline Predictor	16.529	0.744	15.062	0.755

The baseline predictor is a relatively accurate indication of what extent exponential smoothing (as shown in Figure 1) can be used to explain the variance in the data set. Looking at the results, the baseline predictor can explain 75.5% of the variance in the data. It is not unreasonable, then, to suggest that exponential smoothing can explain most of that variance. Hence, we argue that a model that can explain more of the variance must be forecasting better than the baseline predictor.

We find that a majority of the models perform extremely well on the training set; in three out of four cases, the models are able to explain over 80% of the variance. Of course, this is on the training set, so we expect an extremely high correlation here. Indeed, a single decision tree can explain 100% of the variance

in the training set. An impressive number, but once exposed to previously unseen data in the testing set, the decision tree can only explain 64.6% of the variance: a clear warning sign of overfitting.

In general, all models saw increased RMSE values and decreased r^2-values as predictions were moved onto the test sets, as one would expect, given that they have been exposed to training sets previously. It is notable that the support vector regressor performed extremely poorly on the test set, explaining only 14.2% of the variance in the data set. As illustrated earlier, simply using exponential smoothing can explain significantly more of the variance.

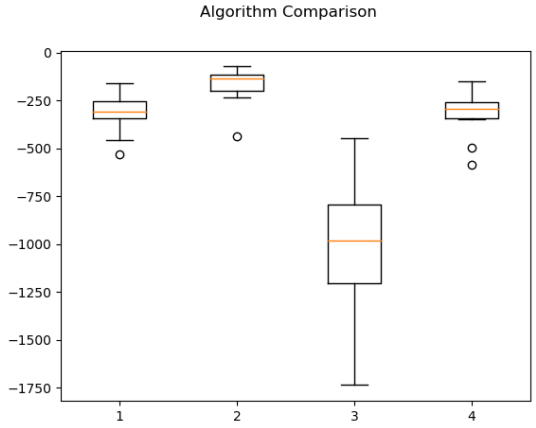


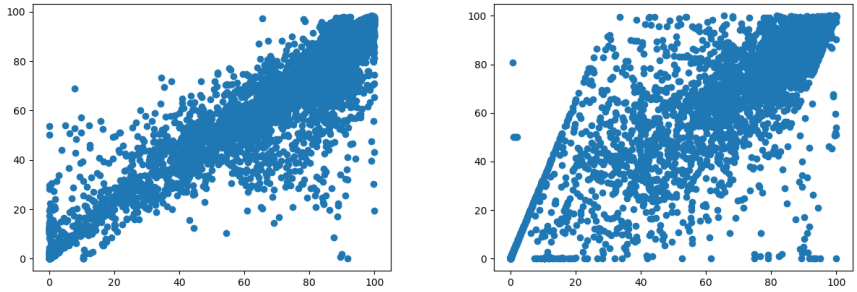
Fig. 3: A box-whisker plot comparing the four different machine learning models on negative square mean error. From left to right: decision tree, random forest, support vector regressor, and K-nearest neighbour.

In Figure 3 we find a much larger spread in predictions in one model than the other on the testing set; the support vector machine had a range of prediction errors from 22.5 AMI points to 42.4; this contrasts with the decision tree, that, despite overfitting severely onto the training data, produced its largest erroneous prediction at 22.3 AMI points off, while its smallest error was in the single digit.

Comparing Figure 3 and Table 3, the random algorithm performed the best, dropping only a relatively small amount in predictive power when moving onto the test set, being able to explain 84.1% of the variance in the data, and with the lowest RMSE recorded on the test set. Similarly, the box-whisker plot shows how reliable its predictions are, with only a single outlier, and very low spread in the errors of its predictions. However, it should be noted that the random forest is only a somewhat better predictor than the K-nearest neighbour regressor, which also outperforms the baseline predictor, with both lower RMSE and more variance explained (14.237 and 0.781 compared to 15.062 and 0.755, respectively).

6 Analysis

While the K-nearest neighbour model does boast better statistics than the baseline predictor, the difference is rather miniscule, and could be tallied up to being purely coincidental. Decision trees are very prone to overfitting to the training data for a problem of this nature, thus they can easily be dismissed. We suspect the support vector regressor performed so poorly due to the difficulty in optimising them. As the results clearly indicated, the random forest model outperformed our baseline predictor, and to better illustrate how its predictions compared to the baseline, Figure 4a and 4b plot predicted and observed AMI-values for the two models, respectively.



(a) For a random forest algorithm (b) For a baseline predictor

Fig. 4: A scatterplot showing the spread of predicted AMI on the X-axis against observed AMI on the Y-axis.

The closer the data points are arranged in a straight line along the axis $y = x$, the more accurate the model predictions are. In Figure 4a, we can see that the data is approximately arranged in a line, though there appears to be a case of excessive warning in the high end of the spectrum; that is, the model predicts low values, indicating relapses where there were none in actuality, so the model is predicting a *false positive*. The same but in reverse happens on the low end of the spectrum. This is particularly disturbing, as it is indicative of the model suggesting that a patient will do well, when, actually, a relapse has occurred; a *false negative* has occurred.

However, as Figure 4b shows, the baseline predictor's forecasted AMI-values have a significantly larger spread; it does not warn in case of relapses nearly enough, and misses a significant amount. This is clearly visible to the naked eye; the aforementioned target, $y = x$, is noticeably less present. We also see a skewed 'grid'-like feature in the data set; this is because the majority of values are oftentimes just the same value shifted some days forward or backwards. As

it was shown earlier that the baseline predictor was more likely to perform both false positives and false negatives, it can be argued that the random forest was predicting more accurately.

7 Conclusion

Our findings suggest that machine learning models can make relatively meaningful predictions on alcoholic relapses five days in advance, as they can beat baseline predictors, both in terms of RMSE and explained variance. Crucially, one model was able to outperform a mean difference predictor. There is a gap in the domain-specific applications of machine learning in this field, and this paper serves as a proof-of-concept that shows that simple, commonplace models can indeed be used to forecast relapses in alcohol abuse. Hence, we argue that further research in this area is highly advised.

There are some caveats that must be highlighted. The study lacks a certain level of generalizability, as the methodology for this paper was built on a single company's own construct for measuring sobriety. This implies that the approach taken is not necessarily generalizable to other data sets dealing with the same domain. In order to examine the subject in a more general sense, one would have to incorporate aspects touched upon by [8], requiring transformations of output variables. Hence, this research can only serve as a proof-of-concept that relapse forecasting is indeed possible, given sufficient data.

There were other factors neglected during this research. Due to availability concerns, only data pertaining to the measurements completed in the clinical trials performed by Kontigo Care [11] were used, but there are other potential indicators of substance abuse. As mentioned in the data set discussion, weekdays, whether or not a salary day has passed, seasons, or vacation periods could be used to make more accurate predictions.

The reason why more modern techniques, such as LSTM networks [12], were not tried is because of the limited data set size. Neural networks and deep learning demand extremely large quantities of data, well beyond the scope of the initial clinical study. However, as this paper shows, relapse forecasting is a possibility, and given the costs associated with treatment plans and eventual relapses [2], further research is beneficial. With larger data sets, deep learning could be used to make extremely accurate long- and short-term predictions alike.

References

1. Baltagi, B.H., Geishecker, I.: Rational Alcohol Addiction: Evidence from the Russian Longitudinal Monitoring Survey. Health Economics **15**(9) (2006) 893 – 914
2. Miller, N.S., Gold, M.S.: Comorbid Cigarette and Alcohol Addiction. Journal of Addictive Diseases **17**(1) (1998) 55 – 66 PMID: 9549603.
3. Örnberg, J.C., Ólafsdóttir, H.: How to sell alcohol? nordic alcohol monopolies in a changing epoch. Nordic studies on alcohol and Drugs **25**(2) (2008) 129–153
4. Moos, R.H., Moos, B.S.: Rates and Predictors of Relapse after Natural and Treated Remission from Alcohol Use Disorders. Addiction **101**(2) (2006) 212–222

5. Stout, R.L., Rubin, A., Zwick, W., Zywiak, W., Bellino, L.: Optimizing the cost-effectiveness of alcohol treatment: A rationale for extended case monitoring. Addictive Behaviors **24**(1) (1999) 17–35
6. Thylstrup, B.: Numbers and Narratives. Relations between Patient Satisfaction, Retention, Outcome and Program Factors in Outpatient Substance Abuse Treatment. Nordic Studies on Alcohol and Drugs **28**(5-6) (2011) 471 – 486
7. McLellan, A.T., McKay, J.R., Forman, R., Cacciola, J., Kemp, J.: Reconsidering the Evaluation of Addiction Treatment: from Retrospective Follow-up to Concurrent Recovery Monitoring. Addiction **100**(4) (2005) 447 – 458
8. Connor, J.P., Symons, M., Feeney, G.F.X., Young, R.M., Wiles, J.: The Application of Machine Learning Techniques as an Adjunct to Clinical Decision Making in Alcohol Dependence Treatment. Substance Use and Misuse **42**(14) (2007) 2193–2206
9. Weiss, J.C., Natarajan, S., Peissig, P.L., McCarty, C.A., Page, D.: Machine learning for personalized medicine: Predicting primary myocardial infarction from electronic health records. AI Magazine **33**(4) (2012) 33
10. Cruz, J.A., Wishart, D.S.: Applications of Machine Learning in Cancer Prediction and Prognosis. Cancer Informatics **2** (2006) 59 – 78
11. Hämäläinen, M.D., Zetterström, A., Winkvist, M., Söderquist, M., Karlberg, E., Öhagen, P., Andersson, K., Nyberg, F.: Real-time monitoring using a breathalyzer-based ehealth system can identify lapse/relapse patterns in alcohol use disorder patients. Alcohol and Alcoholism **53**(4) (2018) 368–375
12. Lee, T.J., Gottschlich, J., Tatbul, N., Metcalf, E., Zdonik, S.: Greenhouse: A Zero-Positive Machine Learning System for Time-Series Anomaly Detection. (01 2018)
13. Akdeniz, E., Egrioglu, E., Bas, E., Yolcu, U.: An ARMA Type Pi-Sigma Artificial Neural Network for Nonlinear Time Series Forecasting. Journal of Artificial Intelligence and Soft Computing Research **8**(2) (2018) 121 – 132
14. Zhang, G.P.: Time Series Forecasting using a Hybrid ARIMA and Neural Network Model. Neurocomputing **50** (2003) 159 – 175
15. Xu, J., Deng, C., Gao, X., Shen, D., Huang, H.: Predicting alzheimers disease cognitive assessment via robust low-rank structured sparse model. **2017** (2017) 3880
16. Zetterström, A., Hämäläinen, M.D., Karlberg, E., Winkvist, M., Söderquist, M., Öhagen, P., Andersson, K., Nyberg, F.: Maximum Time Between Tests: A Digital Biomarker to Detect Therapy Compliance and Assess Schedule Quality in Measurement-Based eHealth Systems for Alcohol Use Disorder. Alcohol and Alcoholism **54**(1) (12 2018) 70–72
17. Kilian, L., Taylor, M.P.: Why is it so difficult to beat the random walk forecast of exchange rates? Journal of International Economics **60**(1) (2003) 85–107
18. Liaw, A., Wiener, M.: Classification and Regression by RandomForest. R News **2/3** (12 2002) 18 – 22
19. Segal, M.R.: Machine Learning Benchmarks and Random Forest Regression. UCSF: Center for Bioinformatics and Molecular Biostatistics (2004)