

Assessing the Feasibility of Machine Learning To Predict Chronic Pain in Adolescence

ABSTRACT

"Chronic pain affects between 15 to 40% of adolescents worldwide. The impact and prevalence of chronic pain can be felt every day in terms of missed school days, strained familial relationships, and financial stress. While rehabilitation programs specifically designed for chronic pain management exist, they cannot always adapt to the idiosyncratic nature of chronic pain. Machine learning presents a framework to use diary data from individuals in pain and make predictions about the trajectories of their pain and related functioning. This study's goal is to assess the feasibility of using machine learning to predict pain and functioning by constructing, training, and evaluating multiple models that take a variable-centered approach to chronic pain."

INTRODUCTION

IMPETUS/CP BACKGROUND

Chronic pain is defined as pain that persists for over six months. Chronic pain differs from acute pain in that it is not caused by physical tissue damage, but rather a failure of the nervous system. Chronic pain can even occur in the absence of precipitating tissue damage. A problem facing rehabilitation efforts is that chronic pain is that it cannot be effectively managed with medication. This understanding led to many chronic pain rehabilitation programs to focus on an alternate method of treatment - lifestyle changes. The reality for many in chronic pain is that their pain may never go away. In these circumstances, the treatment priority changes from curing the condition to managing it. Lifestyle based chronic pain rehabilitation focuses on a wide range of targets such as sleep quality, proper diet and hydration, and daily habits. Chronic pain rehabilitation focuses on restoring function, which serves two purposes. Increasing functioning reduces functional decline due to inactivity and long term bed rest and retrains the nervous system to no longer interpret normal stimuli as pain. Current research shows that restoring function precedes a reduction in pain, rather than vice versa (Benore et al, 2018).

ADOLESCENTS IN CHRONIC PAIN

It is estimated that five percent of adolescents are living with life-changing chronic pain. The initial surveys of chronic pain prevalence in adolescents demonstrated strong relationships between prevalence and both age and sex. Girls were more likely to be in chronic pain than boys, and increasing age was associated with higher prevalence rates (Perquin et al., 2000). These studies also examined the prevalence across different chronic pain conditions. While migraine headaches are the most common chronic pain condition, there were also high incidence rates for abdominal pain, back pain, and other musculoskeletal pain conditions. Comorbidity of chronic pain conditions is standard, with a large portion of those in pain having multiple diagnoses (King et al., 2011).

Adolescents are particularly susceptible to the debilitating effects of chronic pain, primarily due to psychosocial and neurological development that occurs during the adolescent developmental stage. A 2001 study suggested that heightened neuroticism, fear of failure, and desire for social acceptance leaves adolescents particularly vulnerable to the effects of chronic pain on daily functioning (Merlijn et al., 2001).

IMPACT OF CP ON ADOLESCENCE

Adolescents in chronic pain feel the effects of their pain far beyond any physical sensation. The detrimental effects of pain can be seen in school, at home, and in the family dynamics of those in pain. A 2008 sample of adolescents in pain missed an average of five of the past 20 days of school. 47% of the sample showed a reduced academic performance that worsened over time (Logan et al., 2008). The snowball effect of short term school absences can result in missing entire school years or potentially even a complete withdrawal from school for some. The effects of chronic pain in adolescents expands beyond the teen into their families. A meta-analysis of families with adolescent children in pain revealed a strong association between family functioning indices and pain-related disability rather than the pain itself (Lewandowski et al, 2010). Recently, chronic pain rehabilitation programs have begun to focus on mitigating the impact of pain on these factors (Benore et al, 2018).

PROBLEM STATEMENT

Designing lifestyle change based treatments for adolescents is not an easy process. Adolescence is a developmental stage where making lifestyle changes now in the hope of managing pain later is a particularly tough sell. The dual systems development model proposed by Laurence Steinberg shows that differences in neuronal growth rates between reward-seeking and cognitive control areas of the brain predispose teens to desire short term gratification (Steinberg, 2010, Albert & Steinberg, 2011). In terms of chronic pain management, it is particularly difficult for teens to see a connection between decisions made in the short term and their long term results. One can imagine the difficulty of trying to convince a teenager to exercise six days a week by explaining that they may feel better three months later.

Albert Bandura's Social Learning theory provides a useful framework for understanding how to convince a teen to achieve a long term goal. The theoretical model explains that in order to reach a long term outcome, the teen must want that outcome, believe that they want that outcome, see that the outcome works for them, and see that the outcome is achievable (Bandura, 1977). Simply put, claiming that immediate lifestyle changes will result in future improvements in functioning is not a convincing argument for many adolescents living with chronic pain.

While rehabilitation programs geared to adolescent chronic pain exist, they are not always capable of taking a teen through the stages of Bandura's model. The main difficulty lies in the stages where the teen sees the outcome works for them and is achievable. Simply explaining to teens that lifestyle changes will result in future reductions in pain does not demonstrate the value of making those changes in the short term. Some inpatient rehabilitation programs exist that help teens make changes

gradually and then examine pain severity reports after discharge, but these programs are limited by their immense costs (over \$40,000 for three-week programs) and their lack of follow up programming. Additionally, chronic pain is highly idiosyncratic, and programs that use group paradigms may not be able to utilize individual information to its fullest extent.

MACHINE LEARNING AS A SOLUTION

Designing a solution that takes a teen from wanting a goal to achieving it requires the use of large amounts of individual data collected regularly in order to make predictions of trends of pain and related functioning. Machine learning provides an efficient framework for drawing on large quantities of individual data in order to make sophisticated predictions about each individuals' future.

Machine learning is a field combining statistical modeling, mathematics, and computer science to complete prediction and inference tasks on large scale datasets. Machine learning differs from traditional statistical modeling in terms of the machinery employed to test hypotheses. Traditional statistical modeling tests a single hypothesis by fitting a given model to a set of data, while machine learning constantly iterates through models and eventually converges to a final model.

This study employs a form of machine learning called *supervised machine learning*. In supervised machine learning, a dataset is collected that contains a dependent variable (called a *label*) and independent variables (called *attributes* or *features*). The dataset is then split into *training* and *test* sets. The training set is supplied to the model in order to learn a prediction function that best fits that training data. The model can then be evaluated by supplying the test set and generating predictions. The percentage of correct predictions across the test set is the accuracy of the model.

In order to design a machine learning model to answer a question, three key components are required: a task, a set of metrics, and experience. The task to complete determines the nature of the machine learning model. Many supervised machine learning tasks fall under two categories: classification and regression. In this study, prediction of pain severity ratings (on a 1 to 5 scale) is the primary goal. We want a model that is capable of classifying pain ratings given a set of diary data as accurately as possible, so we will utilize the classification framework as it measures accuracy in terms of whether the prediction is exactly correct or not. This type of task is referred to as a multinomial classification task, where the model will generate a prediction of either 1,2,3,4 or 5 for each input. Performance metrics allow for the evaluation of a model's performance on a task. In this case, classification accuracy (the number of correct predictions made by the model out of total predictions made) given various conditions serves as the primary metric. Finally, the experience required is a large scale dataset of individual-level pain and related functioning data collected regularly.

RESEARCH QUESTIONS

The goal of this study is to design a machine learning model that uses regularly collected pain and related function data to make predictions about future trends in pain severity ratings. More specifically,

this study aims to use a machine learning model to examine the relationship between the amount of past data input into the model and the prediction accuracy of the model at various points in time.

To quantify the past and future, we denote the number of past days of data input into the model as k and the number of days into the future to predict as N . For example, if $k = 2$ and $N = 0$, the question is, "what is the accuracy of the model predicting today's pain rating from 2 days of past information?". By varying k and N , a table of prediction accuracies arises that can then be analyzed in order to determine trends.

Before analyzing any longitudinal relationship within the data, the data needs to be tested for fitness to the task. If the data on a given day does not predict the concurrent pain rating, the remainder of the analysis would not yield any useful information. If the data is predictive of concurrent ratings, there are two immediate longitudinal relationships that can be tested. The predictive power of a single day of diary data can be tested by holding k constant and increasing N , while the predictive power of additional past information to predict a current day can be tested by holding N constant while increasing k .

PLAN OF ANALYSIS

With a defined task (classifying pain ratings) and a metric (prediction accuracy), only one component remains - experience. In order to design a dataset for this task, we need pain severity ratings and several indices of functioning. A review of the literature on the detrimental effects of chronic pain on adolescents suggests that sleep (Tham et al, 2019), diet and hydration (Brain et al, 2019, Bear et al, 2016), exercise (Kichline & Cushing, 2019), stress reduction (Benore et al, 2015), and daily functioning such as getting out of bed or taking a shower (Benore et al, 2018) are all impacted. By collecting data on these functioning types and pain ratings on a daily basis for a period of 28 days, we construct a dataset that fits the parameters for supervised machine learning.

The reason for collecting our own data is that there is no dataset currently available on adolescents in pain with the parameters described above. While it was not possible to collect thousands of adolescents' data over a 28 day period, it is possible to use the collected data as a seed for the generation of a synthetic dataset. By employing joint distribution roulette wheel sampling, a synthetic dataset can be generated with a much larger number of participants while preserving the longitudinal nature of the original seed data.

After collecting the seed data and generating the synthetic dataset, an algorithm will then constantly reconstruct the synthetic data to test different pairings of k and N using several supervised machine learning architectures. These architectures were selected as they span the performance-interpretability tradeoff in machine learning, with some architectures having very human interpretable output but weaker raw performance, while other architectures have unintelligible output while demonstrating very strong classification performance. Key features of the model architectures are summarized in table 1.

Table 1
Summary of Model Architectures

Architecture	Method	Assumes Independence of Attributes	Assumes Equal Importance of Attributes
Naive Bayes	Bayesian Estimation	Yes	Yes
Logistic Regression	Optimal Weights for $y_{\text{pred}} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \dots$	No	Yes
Decision Tree	Rule Based Learning	No	No
Neural Network	Optimal Weights for $y_{\text{pred}} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \dots$	No	Yes
Random Forest	Ensemble Rule Based Learning	No	No

The models described in table 1 use different methodologies to arrive at their predictions. Naive Bayes uses Bayesian estimation to find the most likely label given a set of attributes. This algorithm assumes that every attribute in the dataset is independent of all other attributes and that all variables are equally important in predicting the label.

Logistic Regression serves as a baseline against all other models, as it is most similar to standard statistical modeling. This study employs multinomial logistic regression, different from standard binary logistic regression as the label has 5 classes rather than 2. The key difference is the use of a function known as the softmax that outputs a vector of probabilities for each label based on the model parameters and then selects the highest probability as the final prediction.

Artificial Neural Networks (ANNs) are a natural expansion of the logistic regression algorithm, with the advantage of being able to plot nonlinear decision boundaries. Logistic regression is the simplest case of a neural network, and the ability for many logistic regression “units” to be aggregated together allows for the plotting of more complex decision boundaries. Additionally, different hyperparameter tweaks (changes to dimensions of the network, learning rate, epochs of training) can vastly affect the performance of the network.

Decision trees fall under the higher interpretability section of the performance interpretability tradeoff, primarily as they employ rule-based learning. A classic decision tree outputs a set of rules that, when followed for an instance, produce a prediction for the label. Decision trees also allow for individual importance ratings to be attached to each attribute, rather than assuming equal importance among all attributes.

Random Forest is unique among these models as it is the only one to employ ensemble-based learning. Ensemble methods use multiple models and aggregate the information in order to make a final prediction. In this case, the random forest is constructed of 100 decision trees that are all initialized with random starting values. Each tree generates a prediction for the given test example, and then the mode of the distribution of decisions is taken as the final prediction. This construction defends the model from overfitting to their training data and having predictions based on a single process. Like the single decision tree, the random forest assigns individual importances to each attribute in the dataset.

METHODS

SURVEY DESIGN

Two Qualtrics surveys formed the basis for data collection. The intake survey collected both demographic information and brief pain history. The demographic section of the intake survey asked for information on participant age, gender identity, racial background, and diagnoses of chronic pain conditions, including migraines, fibromyalgia, and other sources. The goal of the pain history was to establish a baseline level of both pain intensity and functioning, specifically targeting non-medicinal aspects of rehabilitation. Questions in this section assessed the severity of disruption to sleep, diet, exercise, mood, and daily routine from pain.

RECRUITMENT/COMPENSATION

The criterion for inclusion in the study was college-aged (18-25) and living with chronic pain for more than six months. Recruitment was conducted at Oberlin College, targeting undergraduate students living with chronic pain. Each participant was contacted by the principal investigator through email and asked if they would participate in a data collection study regarding chronic pain and adolescence. The students were then informed that the base compensation for participating in the study was \$35, and they stood to make a \$5 bonus each week if they completed all seven surveys given that week. The maximum potential compensation for each participant was precisely \$50. Those students who agreed to participate ($N=8$) were then added to an anonymous email list to receive the intake and daily surveys.

DEMOGRAPHICS

All participants were between 20 and 25 years old ($M = 21.375$, $SD = 1.3$), with 37.5% self-identifying as male, 37.% as female, and 25% as gender non conforming. 87.5% of participants identified as White, while 12.5% of participants identified as Hispanic/LatinX. All participants reported an average pain level between 2 and 6 out of a possible 10 ($M = 4.75$, $SD = 1.48$). 62.5% of participants supplied a diagnosis for at least one chronic pain condition, with 37.5% indicating diagnoses of multiple conditions.

All participants reported an impact of chronic pain on their ability to engage in leisure, family, and work activities due to their pain. Multiple participants noted that remaining still for extended periods or specific activities, including eating and walking, caused them the most pain. 75% of participants reported fair sleep quality, while the remaining 25% reported poor sleep quality. 75% of participants also reported feeling fatigued regularly over the past month. 87.5% of participants reported some disruption to their day to day activities, and 75% reported some disruption to their social activities. All participants reported at least an 8 out of 10 when asked to indicate their ability to tolerate moderate pain.

COLLECTION PROCEDURE

Each participant was emailed a link to the intake survey where they read and electronically signed an informed consent form. Those who consented then provided an email address for receiving the daily surveys and were given a unique identification code generated by Qualtrics for use in place of personal information for the remaining surveys. For the next 28 days after filling out the intake survey, each participant was emailed a link to a daily survey at 3:30 pm EDT and asked to complete the survey. No pain severity readings were recorded until seven days in data collection due to a data entry error during the survey creation process. After the 28 day collection period, participants were compensated through the TANGO platform for amounts between \$35 and \$50 based on their completion rate. After destroying all identifying information, the data is then exported from Qualtrics and loaded into Rstudio to serve as the seed for generating the synthetic dataset.

DATASET SYNTHESIS

Using RStudio, a Joint distribution roulette wheel sampling of the seed data produced the synthetic dataset. The first step was to remove metadata and identify columns in the dataset that corresponded to each scale assessed in the surveys. The dataset generation function took a number of participants as an argument and generated 28 days of data for each participant.

In roulette wheel sampling, a "wheel" is created from the proportion of each response to a given question. For example, if 25% of participants answered yes to a question, 35% answered no, and the remaining 40% refused to answer. In this case, the values 0-0.25 represent yes, 0.25-0.6 represent no, and 0.6-1.0 represents a refusal to answer. When the algorithm needs to synthesize a value for a given instance, the function draws a random uniform value and tests it against the wheel. If the value falls within one of these ranges, then that response is selected for the synthetic instance.

The resulting dataset had an intraclass correlation coefficient of 0.789 for pain ratings (see table 2). This demonstrates that 80% of the variance in pain ratings are attributable to between-person differences, while the remaining 20% is attributable to within-person differences and error. This demonstrates two fundamental features of the synthetic dataset: that attributes are not independent of one another, and that the predictive power of our models will be limited as we are examining the within-person component of the variance in pain ratings. It is also worthy of note that there is a recency effect present in the inter-item correlation matrix (see appendix). While each day's pain is correlated strongly with the next and previous days' ratings, the correlations drop off quickly after several days and begin to vary about 0 after 12 days.

Table 2
Intraclass Correlation Analysis on Pain Ratings

ICC	N of items
0.789	28

ML ANALYSIS

At this point, the synthetic data is read into a python script that handles the actual machine learning model construction. The first function in the script solves the problem of testing different values of k and N from the same data. Normally, the instances are not modified in this manner during a supervised machine learning analysis, but the data is not presently in a format to answer questions like “how accurately can I predict N days into the future given k days of past data?”. For example, in order to test if five days of past data is enough to predict the next day's pain rating, an instance needs to use the next day's pain rating as a label and the entire instance of the past five days for a given participant as the factors. A python function that takes in a value k for past days and a value N for the prediction day accomplishes this task. At this point, the script splits the prepared synthetic dataset into a training set and a test set based on a given proportion.

After splitting the training and test set, the script runs through a loop that tests every pairing of k and N up to $k = 7$, $N = 7$. Each time through the loop, the script reorganizes the dataset to test a specific k N pairing and then runs all 5 models on the data. Each model is trained on the training set, tested on the test set, and then produces an accuracy rating for that pairing, alongside other performance metrics. During this iterative process, the accuracies for each model, and each pairing are stored in tables. These tables are exported in .CSV format at the conclusion of the script alongside confusion matrices for each model and each pairing (see appendix).

Many of the models employed require hyperparameters to be set before conducting any analyses. For many of the models, a random seed was required, which was set as a command-line argument to the python script. The neural network's hyperparameters were tuned via grid search to arrive at a construction with two hidden layers of 150 neurons each, the ADAM optimization algorithm, a learning rate of 0.001, and 5000 epochs of training time. The random forest model utilized 100 estimators that were each initialized randomly based on the supplied seed.

RESULTS

Predicting Pain from Diary Data

Our first step was to validate that the survey data could accurately predict concurrent pain ratings using traditional statistical methods. Logistic regression serves as a baseline statistical model both for its similarity to non-machine learning statistical modeling and ease of interpretability. Analyzing the output of the regression when $k = 1$ and $N = 0$ gives us critical insight: some variables are more predictive of pain than others. (It should be noted that the logistic regression assumes that each day's data is independent of the others, which is violated by the nested nature of the dataset.) The subsection of table 3 summarizes the output of the regression model. The entries in each row represent the log odds ratio for the given attribute. Each entry corresponds to how much more likely each label value is than 1 given a unit increase in the attribute. It is clear from the table that some attributes, like reporting feeling distressed, have much greater influence on these odds ratios than others (like whether or not someone

played music). Overall, the regression model's 90.4% accuracy demonstrates that standard statistical methods can predict pain from the given data.

Table 3

Snippet of Logistic Regression Model Output for $k = 1, N = 0$

	<i>Dependent variable:</i>			
	2 (1)	3 (2)	4 (3)	5 (4)
'Q4_1: Indicate if you felt each of the following since your last survey - Interested'	819.296*** (9.842)	-77.180*** (10.182)	84.585*** (5.879)	61.303*** (6.803)
'Q4_2: Indicate if you felt each of the following since your last survey - Distressed'	88.013*** (26.839)	410.003*** (69.108)	492.635*** (33.987)	270.437*** (72.141)
'Q8_1: Indicate if you did any of the following since your last survey - Read a book'	-2.923 (18.471)	32.781 (24.906)	-3.186 (18.564)	-0.214 (18.546)
'Q8_2: Indicate if you did any of the following since your last survey - Created art'	-2.797 (16.244)	-11.735 (16.998)	10.916 (16.732)	11.620 (16.881)
'Q8_3: Indicate if you did any of the following since your last survey - Played a game'	-43.121 (44.989)	-55.398 (45.612)	-58.083 (45.813)	-58.947 (45.927)

Note: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

From here, the question becomes whether or not more sophisticated machine learning models can outperform the baseline logistic regression in terms of classification performance. Figure 1 displays the prediction accuracies of each model architecture tested when k is 1 (only the current day's pain data is supplied), and N is 0 (predicting the label attached to that day's data). The figure demonstrates that some models outperform the baseline set by the regression while other architectures fall flat. This observation demonstrates not only that machine learning is a potentially suitable framework for the problem, but also that there is added value in selecting an appropriate model for a given task.

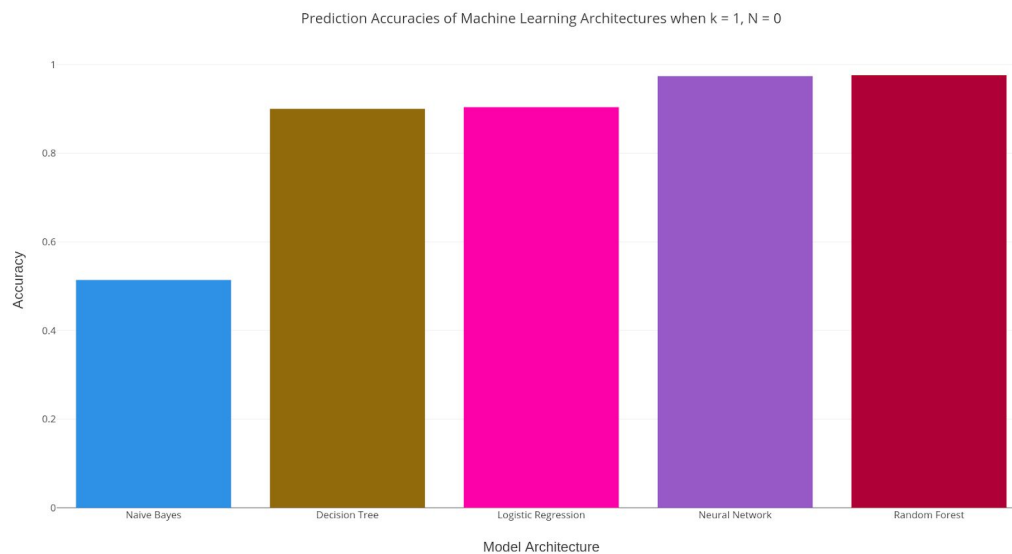


Figure 1. Bar graph of prediction accuracies of models given $k = 1, N = 0$.

Organization

The model results are presented from the simplest to the most complex. The naive Bayes classifier serves as the most simple example of machine learning based classification, followed by an artificial neural network (a natural extension of the logistic regression model). From there, a single decision tree shows a single permutation of rule-based learning. Finally, the random forest model demonstrates ensemble rule-based learning. The overall performance on all k N pairings for each architecture is summarized in annotated heatmaps (see appendix).

Naive Bayes

The naive Bayes classifier had by far the worst performance of all models tested, reaching an accuracy of only 51.4% when $k = 1$, $N = 0$. The explanation for the sub-par performance is that a fundamental assumption of naive Bayes is violated in this dataset. Naive Bayes rests on the naive assumption - the assumption of complete independence of attributes. Naive Bayes assumes that all of the predictor variables are independent of one another, and assigns equal weight to each attribute when making a prediction. This assumption is contrary to the entire methodology of this study, as the dataset is constructed from scales of chronic pain-related function, each in turn composed of related questions. These relationships between predictors are demonstrated in the poor performance of naive Bayes relative to the other models. If the Naive classifier had performed better than other architectures, it would demonstrate a fundamental flaw in the survey design.

Neural Network

The neural network performed slightly better than the logistic regression baseline, reaching an accuracy of 97.3% for $k = 1$, $N = 0$. The primary concerns with the neural network model are both the assumption of equal importance among attributes as well as a significant class imbalance in the training data. The class percentages - the percentage of each pain rating label out of total instances - are summarized in table 4. The overwhelming majority of instances have either 2 or 4 as the label. This imbalance is a bias in the training data that will affect the model's classification performance. The confusion matrices (see appendix) for the neural network models show a clear bias in favor of predicting 2 and 4. Additional training data may alleviate this bias and, in turn, improve the network's classification performance.

Table 4

Frequencies of instance labels in training data

label	Frequency	Percent	Valid Percent	Cumulative Percent
1	427	15.250	15.250	15.250
2	1042	37.214	37.214	52.464
3	215	7.679	7.679	60.143
4	806	28.786	28.786	88.929
5	310	11.071	11.071	100.000
Missing	0	0.000		
Total	2800	100.000		

Decision Tree

The single decision tree performed similarly to the baseline, reaching an accuracy of 90%. While the decision tree has a natural advantage in the ability to assign importance to each attribute in the dataset, a single decision tree is not always capable of defining a set of rules that capture every nuance present in the training set. The fact that the single tree is outperformed by the random forest model suggests that rule based learning seems effective, though the additional power from ensemble learning provides additional benefit on this task.

Random Forest

Of all architectures tested, random forest had the most impressive classification performance, reaching an accuracy of 97.6%. Random forest has both the ability to assign importance to each predictor variable or attribute like a decision tree, but also benefits from its nature as an ensemble learning method. All other architectures tested utilize a single converged model to make predictions, but random forest utilizes multiple decision trees to generate a distribution of predictions, and then selects the mode of that distribution to be the final prediction. This provides additional robustness to the predictions generated. In terms of assigning importance, random forest generates an importance rating (see table 5) to each attribute. These ratings suggest that some scales, such as the mood scale (Q_4), were more predictive of pain ratings than others.

Table 5

Importance Ratings of Attributes Assigned by Random Forest

Attribute	importance
Q4_2: Indicate if you felt each of the following since your last survey - Distressed	0.04331892
Q4_6: Indicate if you felt each of the following since your last survey - Guilty	0.03689005
Q4_7: Indicate if you felt each of the following since your last survey - Scared	0.0351594
Q4_3: Indicate if you felt each of the following since your last survey - Excited	0.03294273
Q7_5: Indicate if you did any of the following since your last survey - Prepared food	0.03138055
Q4_5: Indicate if you felt each of the following since your last survey - Strong	0.03049409
Q12_3: Indicate if you did any of the following since your last survey - Meditated	0.02993357
Q4_4: Indicate if you felt each of the following since your last survey - Upset	0.02967443
Q9_4: Indicate if you did any of the following since your last survey - Volunteered	0.02912034
Q4_14: Indicate if you felt each of the following since your last survey - Inspired	0.02849745
Q8_8: Indicate if you did any of the following since your last survey - Enjoyed my hobbies	0.02722411
Q9_3: Indicate if you did any of the following since your last survey - Helped out family/friends	0.02670182
Q7_7: Indicate if you did any of the following since your last survey - Showered	0.02667582
Q4_10: Indicate if you felt each of the following since your last survey - Proud	0.0251548
Q4_18: Indicate if you felt each of the following since your last survey - Jittery	0.02433615
Q9_5: Indicate if you did any of the following since your last survey - Fed pets	0.02415844
Q4_11: Indicate if you felt each of the following since your last survey - Irritable	0.02389813
Q8_3: Indicate if you did any of the following since your last survey - Played a game	0.02384421
Q4_17: Indicate if you felt each of the following since your last survey - Attentive	0.02225258
Q12_2: Indicate if you did any of the following since your last survey - Mindfulness	0.0217035
Q8_7: Indicate if you did any of the following since your last survey - Played with pets	0.02028188
Q4_8: Indicate if you felt each of the following since your last survey - Hostile	0.018619
Q4_20: Indicate if you felt each of the following since your last survey - Afraid	0.01858849
Q11_3: Indicate if you did any of the following since your last survey - Got my heart rate up	0.0179688
Q4_19: Indicate if you felt each of the following since your last survey - Active	0.01773767
Q4_15: Indicate if you felt each of the following since your last survey - Nervous	0.01688955
Q4_16: Indicate if you felt each of the following since your last survey - Determined	0.01660076
Q7_4: Indicate if you did any of the following since your last survey - Ran errands	0.01658782
Q8_4: Indicate if you did any of the following since your last survey - Watched TV	0.01631227
Q10_1: Indicate if you did any of the following since your last survey - Ate healthy	0.01584844
Q4_9: Indicate if you felt each of the following since your last survey - Enthusiastic	0.01568096

Q8_2: Indicate if you did any of the following since your last survey - Created art	0.01558763
Q10_2: Indicate if you did any of the following since your last survey - Followed diet	0.01517544
Q8_6: Indicate if you did any of the following since your last survey - Played music	0.01395268
Q4_13: Indicate if you felt each of the following since your last survey - Ashamed	0.01364892
Q9_2: Indicate if you did any of the following since your last survey - Did homework	0.01262218
Q4_12: Indicate if you felt each of the following since your last survey - Alert	0.01248701
Q13_2: Indicate if you did any of the following since your last survey - Rested during the day	0.01209593
Q13_3: Indicate if you did any of the following since your last survey - Got up on time	0.01182908
Q11_2: Indicate if you did any of the following since your last survey - Took a stroll	0.01011041
Q8_1: Indicate if you did any of the following since your last survey - Read a book	0.01003662
Q8_5: Indicate if you did any of the following since your last survey - Wrote in a journal	0.00979411
Q7_3: Indicate if you did any of the following since your last survey - Left for school/class	0.00967585
Q10_3: Indicate if you did any of the following since your last survey - Stayed hydrated	0.00942652
Q11_1: Indicate if you did any of the following since your last survey - Exercised	0.0085884
Q13_4: Indicate if you did any of the following since your last survey - Got to sleep	0.00857236
Q13_1: Indicate if you did any of the following since your last survey - Went to bed on time	0.00848681
Q7_8: Indicate if you did any of the following since your last survey - Brushed hair	0.00798674
Q12_1: Indicate if you did any of the following since your last survey - Was thankful	0.00718257
Q12_4: Indicate if you did any of the following since your last survey - Yoga/Tai Chi	0.00653078
Q9_1: Indicate if you did any of the following since your last survey - Was supportive to a friend	0.00596623
Q7_9: Indicate if you did any of the following since your last survey - Took medications	0.00545064
Q7_6: Indicate if you did any of the following since your last survey - Brushed teeth	0.0053845
Q8_9: Indicate if you did any of the following since your last survey - Played Sports	0.0045071
Q7_10: Indicate if you did any of the following since your last survey - Washed up	0.00302055
Q12_6: Indicate if you did any of the following since your last survey - Biofeedback	0.00257781
Q12_5: Indicate if you did any of the following since your last survey - Prayed	0.00252744
Q4_1: Indicate if you felt each of the following since your last survey - Interested	0.00229894
Q7_2: Indicate if you did any of the following since your last survey - Got dressed	0
Q7_1: Indicate if you did any of the following since your last survey - Got out of bed	0

Note: importance table generated with $k = 1$, $N = 0$

Using Today to Predict the Future

The k N pairing framework allows for the examination of the predictive power of a single day of diary data. Determining the predictive power of today's diary data involves testing what happens to the predictive accuracy, holding $k = 1$ as N increases. Figure 2 demonstrates the predictive accuracy of each model holding $k = 1$.

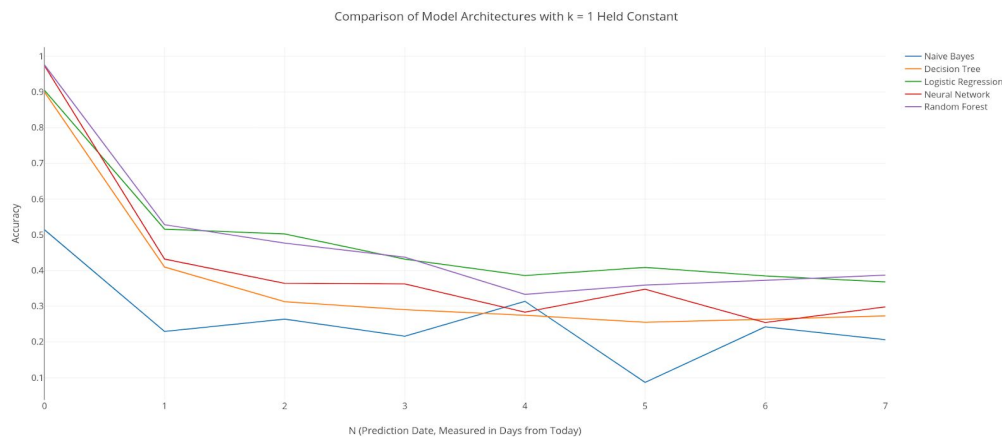


Figure 2. Prediction Accuracies of Model Architectures at $k = 1$, N increases from 0 to 7.

Across all models, we see a precipitous drop in classification performance between $N = 0$ and $N = 1$, which only drops further out until $N = 7$. This effect is most likely due to the dataset's relatively weak signal to noise ratio. Holding k constant and increasing N attempts to use a tiny segment of time to make accurate predictions about distant outcomes. This pattern of results indicates that today's diary data is more predictive of today's pain ratings than later pain ratings.

It is worthy of note that despite decoupling the prediction day and the diary data, the random forest and logistic regression models still manage to perform better than chance accuracy (20%). The random forest managed to maintain an accuracy of over 33% at all values of N tested. The robustness of random forest against irrelevant attributes and the power of ensemble learning are likely causes for the ability to extract more meaningful information out of a relatively small dataset.

Augmenting Predictions with Past Data

The longitudinal nature of the dataset naturally begs the question of whether or not a prediction can be improved given additional past data. This question can be answered by examining the behavior of a fixed N and increasing k . Figure 3 compares the various models with a fixed N of 0 and increasing the value of k .

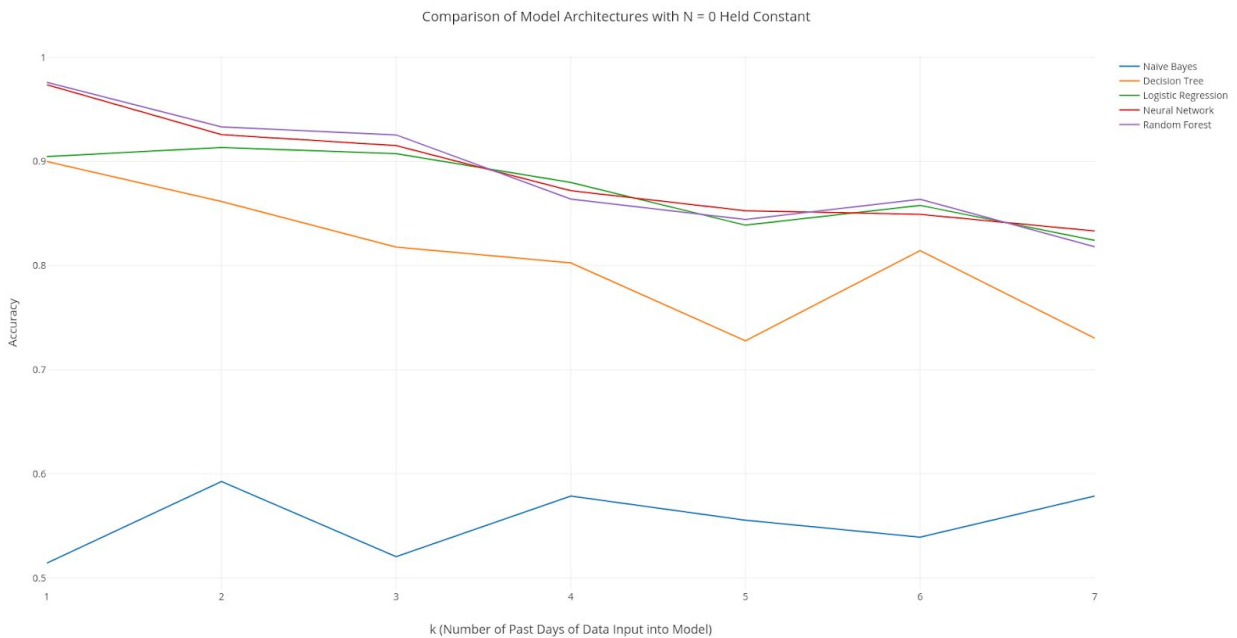


Figure 3. Prediction Accuracies of Model Architectures at $N = 0$, k increases from 1 to 7.

Across all models, the effect of increasing k was a significant drop in classification accuracy. This effect is the net result of irrelevant attributes drowning the signal in increasing amounts of noise as k increases. When several questions are not predictive of pain on the same day, aggregating those same questions over multiple days results in a large proportion of attributes not having much predictive power. For architectures that do not assign importance to attributes, this results in extreme performance

degradation. Even architectures that can distinguish irrelevant attributes demonstrate appreciable losses, as their robustness has a threshold. The aggregation of noise is exemplified in analyzing the differences in performance between the models at $k = 1$ $N = 0$ and $k = 7$ $N = 0$ (see figure 4). The only model to not have a marked drop in classification performance is naive Bayes, which is already performing 50% worse than all other architectures. The assumption of independence shields naive Bayes from the noise generated by additional past data but does not improve prediction accuracy in any meaningful way.

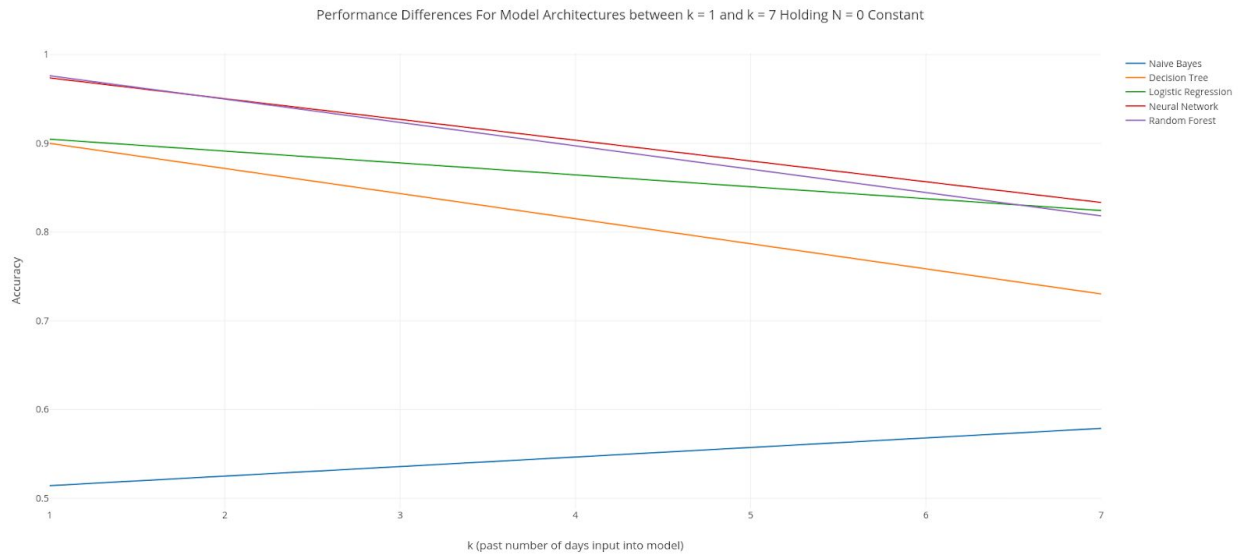


Figure 4. Performance Differences For All Models Between $k = 1$ and $k = 7$ ($N = 0$ constant)

CONCLUSION

DISCUSSION/LIMITATIONS

This study aimed to use different machine learning models to examine relationships between the amount of past data input into the model and prediction accuracies at various time points. Previously, machine learning has been used to examine acute pain, primarily in hospital contexts (Löttsch & Ultsch, 2018). This study is the first to use lifestyle based rehabilitation data to predict pain severity. Overall, the analysis determined that machine learning is a useful framework in which to analyze regularly collected diary data in order to predict pain severity ratings in the short term. However, the limitations of the dataset demonstrate a need for additional research to examine long term trends in pain using machine learning.

The initial testing determined that functioning data can be effectively utilized to predict self-report pain severity ratings. First, the results of the naive Bayes model supported the assumptions of non independence in the dataset. Second, the fact that random forest outperformed the logistic regression

baseline demonstrated that machine learning could be a more effective framework for this prediction task than traditional statistical learning methods. The importance table generated by the random forest model demonstrated a large number of irrelevant attributes, but the robustness of random forest to irrelevant attributes and the ensemble prediction method resulted in strong classification performance. While the neural network and decision tree models performed on par with logistic regression, these models will likely benefit from additional training data and tweaks to the hyperparameters.

The classification results also suggested that today's diary data is predictive of future days pain ratings to a certain degree. While all model architectures tested demonstrated reduced performance as N increased, the logistic regression and random forest models still predicted pain above chance accuracy up to 7 days in the future. It is worthy of note that this does not necessarily mean that the models have successfully determined a long term trend from a single day of data. It is clear that the weak signal to noise ratio of the data has a detrimental effect on classification performance, and the higher than chance accuracies of the models may suggest that they are detecting a "baseline" - a stable state of pain that does not necessarily reflect changes in pain trajectories. This question can be examined in a further study using a larger dataset and broader ranges of k and N .

Finally, the data suggest that simple models are not able to improve a given day's prediction based on additional past data. While this may seem disheartening, it is crucial to understand that the qualities of the dataset are such that additional steps are required to achieve maximum performance from these models. A standard stage in the machine learning pipeline after collecting data but before analysis is feature selection. Feature selection in the context of this thesis is the process of attempting a priori to determine which attributes are most likely to be predictive of pain. Feature selection is usually accomplished by correlating attributes with either the pain severity ratings directly or by the importance metric discussed during the random forest analysis. The process can increase classification accuracy and reduce model overfit.

IMPLICATIONS/FUTURE DIRECTIONS

The results of the analysis show promise for machine learning as the solution to analyzing diary data in order to make predictions about future pain levels. While far from a finished product that adolescents in chronic pain can utilize directly, these models have the potential to demonstrate that what one does now affects how they will feel in the future. The current infrastructure for caring for adolescents in pain is costly, both in terms of dollars lost and emotional cost to teens and their families. These models could form the basis for a new rehabilitation aid that does not require extended hospital stays or other costly measures.

Further research on using machine learning to examine diary data and predicting pain ratings should utilize a more extensive dataset collected over a more extended period. With sufficient signal to counteract the noise, it may be possible to determine long term trends from sufficient past data. If this is the case, it may be that an optimal k N pairing exists such that k number of previous days consistently produces highly accurate predictions N days into the future. If that optimal pairing is discovered, the model could serve as a basis for educational tools that show adolescents in pain that making changes over

some time will have tangible changes in their future pain trends. While it may not instantly solve the problems of teens refusing to make lifestyle changes, these models could lend weight to the arguments in favor of lifestyle change as a method for chronic pain rehabilitation.

ACKNOWLEDGMENTS

I would like to thank the Psychology Department for all the support they have given me over the past 4 years at Oberlin. I would like to personally thank my thesis committee (Dr. Nancy Darling, Dr. Adam Eck, Dr. Paul Thibodeau) for their mentorship, understanding, and time. I would also like to thank Joan Gleason for all of her assistance in scheduling and administrative help. This thesis would not exist without all of your dedication and support.

This thesis was supported by a Jerome Davis Research Grant.

REFERENCES

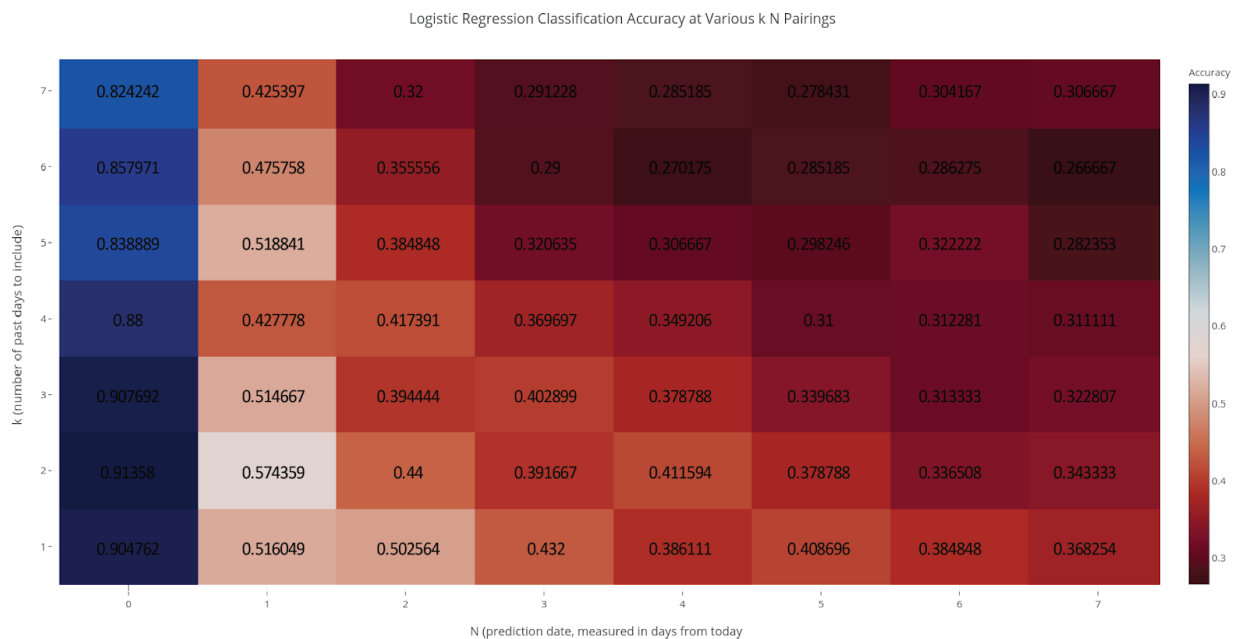
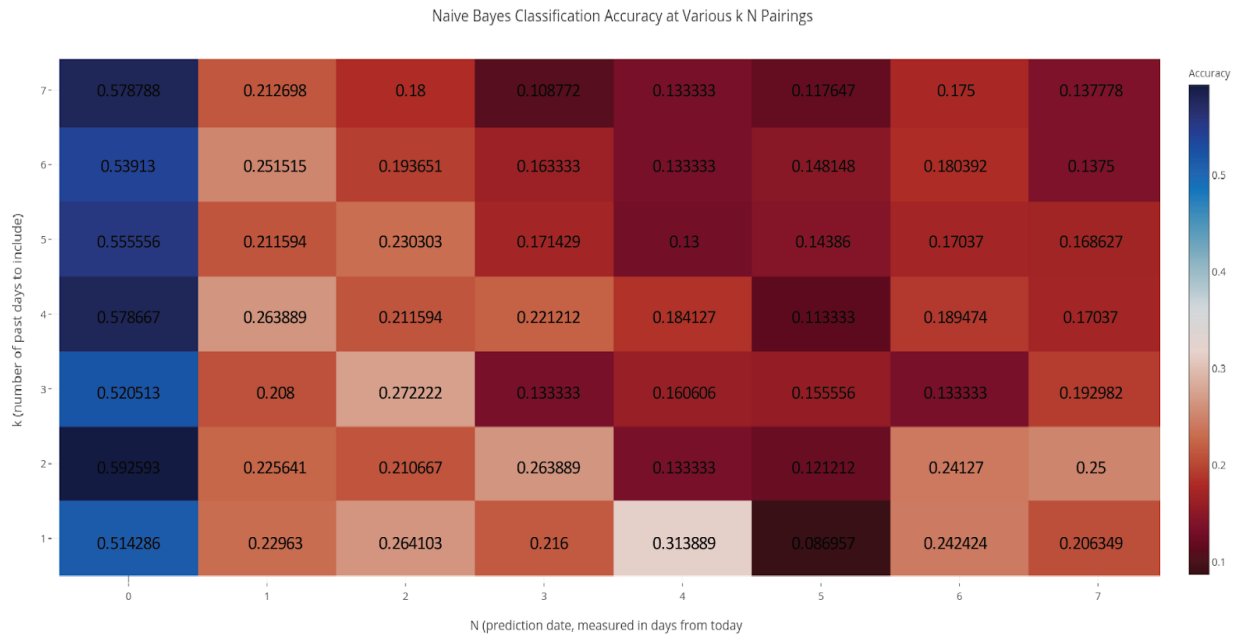
- Albert, D., & Steinberg, L. (2011). Judgment and decision making in adolescence. *Journal of Research on Adolescence, 21*(1), 211-224. doi:10.1111/j.1532-7795.2010.00724.x
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. doi:10.1037/0033-295X.84.2.191
- Bear, T., Philipp, M., Hill, S., & Mündel, T. (2016). A preliminary study on how hypohydration affects pain perception. *Psychophysiology, 53*(5), 605-610. doi:10.1111/psyp.12610
- Benore, E., D, Auria, A Banez, G. A., Worley, S, & Tang, A. N. (2015). The influence of anxiety reduction on clinical response to pediatric chronic pain rehabilitation. *Clinical Journal of Pain,*
- Benore, E., Brenner, A., Banez, G. A., Wang, L., & Worley, S. (2018). It takes two: Parent functioning within the pediatric chronic pain experience and interdisciplinary rehabilitation treatment. *Rehabilitation Psychology, 63*(2), 205-214. doi:10.1037/rep0000177
- Benore, E., Webster, E. E., Wang, L., & Banez, G. (2018). Longitudinal analysis of patient-reported outcomes from an interdisciplinary pediatric pain rehabilitation program for children with chronic migraine and headache. *Headache, 58*(10), 1556-1567. doi:10.1111/head.13389
- Brain, K., Burrows, T. L., Rollo, M. E., Chai, L. K., Clarke, E. D., Hayes, C., . . . Collins, C. E. (2019). A systematic review and meta-analysis of nutrition interventions for chronic noncancer pain. *Journal of Human Nutrition and Dietetics, 32*(2), 198-225. doi:10.1111/jhn.12601
- Kichline, T., & Cushing, C. C. (2019a). A systematic review and quantitative analysis on the impact of aerobic exercise on pain intensity in children with chronic pain. *Childrens Health Care, 48*(2), 244-261. doi:10.1080/02739615.2018.1531756
- King, S., Chambers, C. T., Huguet, A., Macnevin, R. C., Mcgrath, P. J., Parker, L., & Macdonald, A. J. (2011). The epidemiology of chronic pain in children and adolescents revisited: A systematic review. *Pain, 152*(12), 2729–2738. doi: 10.1016/j.pain.2011.07.016
- La Buissonniere-Ariza, V., Hart, D., Schneider, S. C., McBride, N. M., Cepeda, S. L., Haney, B., . . . Storch, E. A. (2018). Quality and correlates of peer relationships in youths with chronic pain. *Child Psychiatry & Human Development, 49*(6), 865-874. doi:10.1007/s10578-018-0802-z
- Lewandowski, A. S., Palermo, T. M., Stinson, J., Handley, S., & Chambers, C. T. (2010). Systematic Review of Family Functioning in Families of Children and Adolescents With Chronic Pain. *The Journal of Pain, 11*(11), 1027–1038. doi: 10.1016/j.jpain.2010.04.005
- Logan, D. E., Simons, L. E., Stein, M. J., & Chastain, L. (2008). School Impairment in Adolescents With Chronic Pain. *The Journal of Pain, 9*(5), 407–416. doi: 10.1016/j.jpain.2007.12.003
- Lötsch, J., & Ultsch, A. (2018). Machine Learning in Pain Research. *Pain, 159*(4), 623-630. doi: 10.1097/j.pain.0000000000001118
- Merlijn, V. P., Hunfeld, J. A., Wouden, J. C. V. D., Hazebroek-Kampschreur, A. A., Koes, B. W., & Passchier, J. (2003). Psychosocial factors associated with chronic pain in adolescents. *Pain, 101*(1), 33–43. doi: 10.1016/s0304-3959(02)00289-0
- Odell, S., & Logan, D. E. (2013). Pediatric pain management: The multidisciplinary approach. *Journal of Pain Research,*

- Perquin, C. W., Hazebroek-Kampschreur, A. A., Hunfeld, J. A., Bohnen, A. M., Suijlekom-Smit, L. W. V., Passchier, J., & Wouden, J. C. V. D. (2000). Pain in children and adolescents: a common experience. *Pain*, 87, 51-58. doi: 10.1016/s0304-3959(00)00269-4
- Ruskin, D., Harris, L., Stinson, J., Kohut, S. A., Walker, K., & McCarthy, E. (2017). "I learned to let go of my pain". the effects of mindfulness meditation on adolescents with chronic pain: An analysis of participants' treatment experience. *Children-Basel*, 4(12), 110. doi:10.3390/children4120110
- Steinberg, L. (2010). A dual systems model of adolescent risk-taking. *Developmental Psychobiology*. doi: 10.1002/dev.20445
- Tham, S., Daley, L., & Palermo, T. (2019). Sleep deficiency and disability in adolescents with chronic abdominal pain. *Journal of Pain*, 20(4), S5. doi:10.1016/j.jpain.2019.01.032

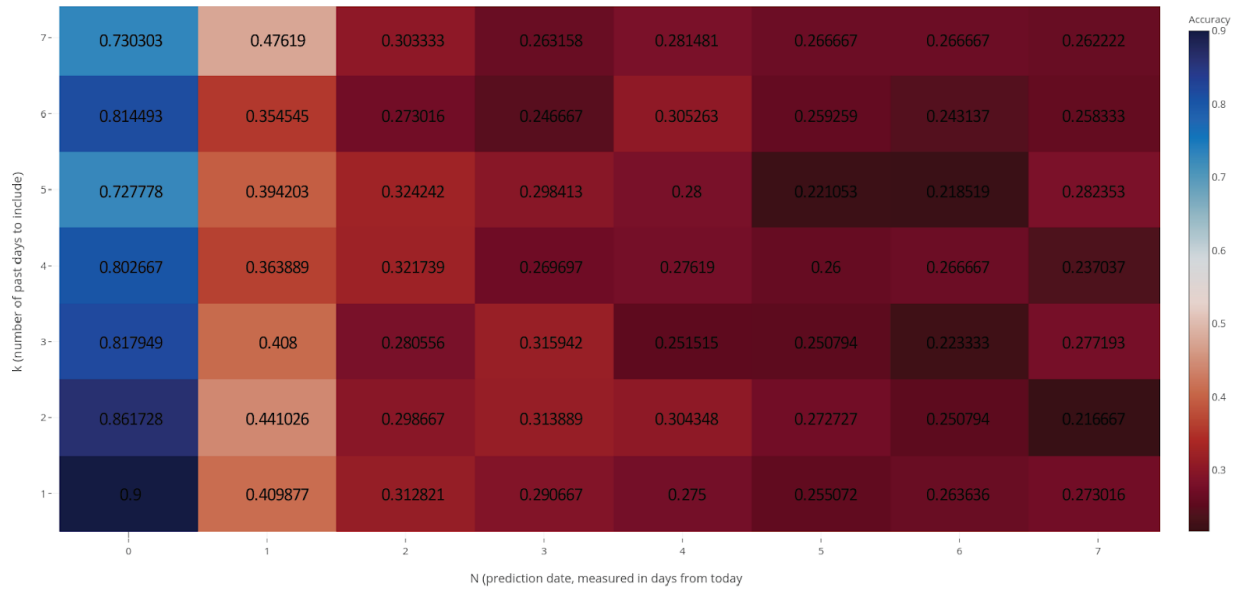
APPENDIX

GitHub Repository (Contains all analysis materials): <https://github.com/MkramerPsych/Honors>

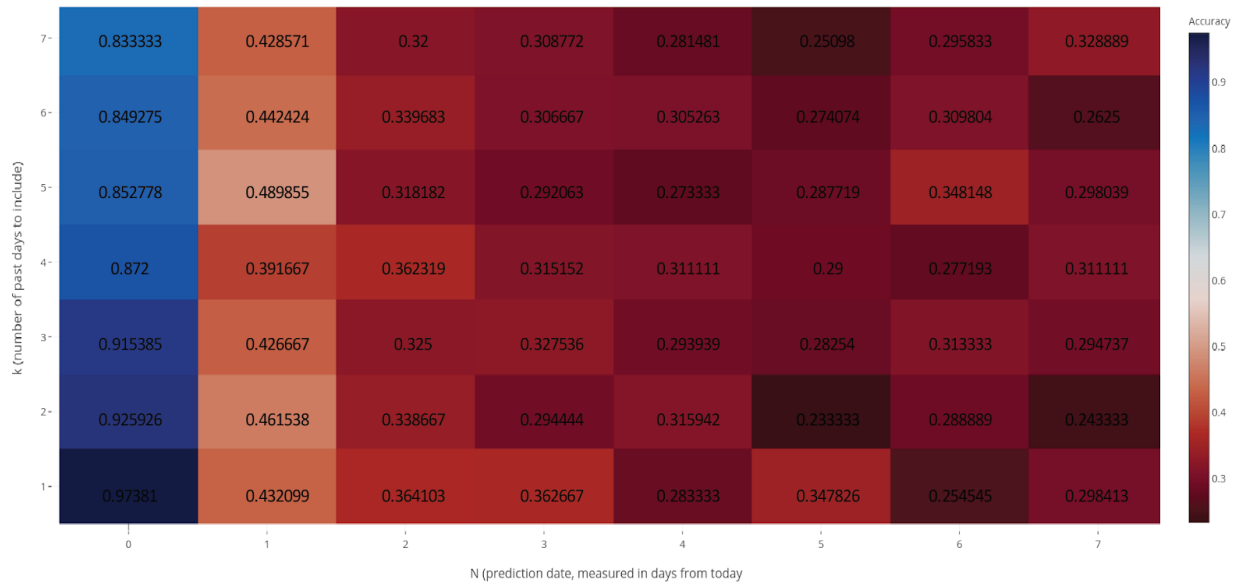
Prediction Accuracy Heatmaps for all Model Architectures Tested

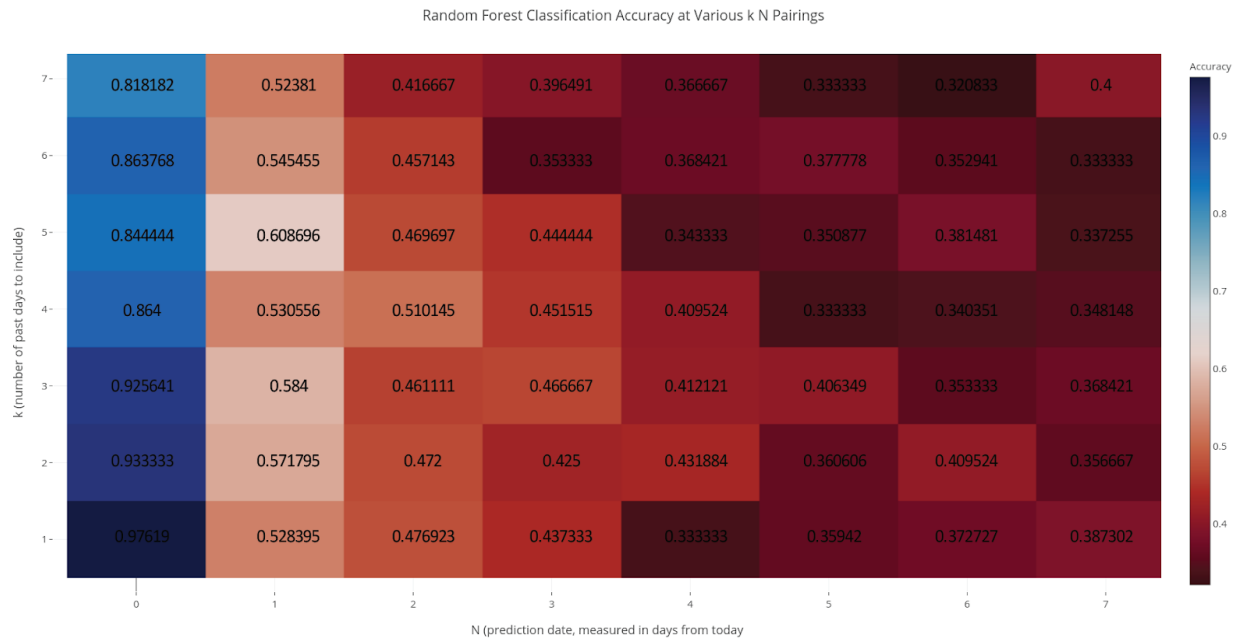


Decision Tree Classification Accuracy at Various k N Pairings



Neural Network Classification Accuracy at Various k N Pairings





Consent Form

Thank you for visiting!

This is a study to collect initial data for a project that aims to create a model of chronic pain in adolescents. We hope to use the data you and others provide to help develop a model that can aid in identifying more effective methods of treating adolescents in chronic pain,

What will you be asked to do?

On the intake questionnaire, you will be asked to enter your email (and phone number if you would like text reminders) as well as a few demographic questions. After that, you will be asked to provide some background on your pain history.

After the initial questionnaire, you will be asked to complete a brief survey every day for a period of 28 days. The questionnaire should only take approximately 5 minutes per day. The daily questionnaire asks questions about mood, pain, and functioning.

The whole study will occur over a 28 day period.

What are the risks and benefits of participating?

This study is classified as having 'minimal risk' – in other words, it is no more risky than things you might experience in everyday life.

You will receive \$7.50 for each week you remain in the study. Every week in which you complete a survey every day will earn you a \$5 bonus. You stand to make \$50 for completing 28 surveys.

Your end reward will be sent via Tango, where you have the option of having a check mailed to your address, certain gift cards, or a donation to charity.

Beyond the financial compensation, your data can help in the creation of models that will help identify the best courses of treatment for adolescents in chronic pain.

Is my information confidential?

Yes. Your ID codes are generated by Qualtrics and your emails/phone numbers will be stored on a secure server.

What will happen to my data?

At the conclusion of the study, all identifying information about you (email/cellphone number) will be destroyed. While your data will become part of a validation set for a model, you will have the option to submit or withhold data during the final survey.

Still interested? Thank you!

By clicking here you are affirming that (a) you are 18 years old or older and (b) you have read the study description above and (c) you are voluntarily participating in this study.

More Questions?

If you have questions about the study, please contact Max Kramer (mkramer@oberlin.edu or 773-318-5225)

If you have questions about your rights as a research participant, please contact Associate Dean Daphne John, Office of the Dean of Arts and Sciences, Cox 101 (djohn@oberlin.edu or 440-775-8410).

-
-
-

- YES! I affirm I am 18 years old or older, have read the study description, and am participating voluntarily.
- No, I don't want to participate.

Default Question Block

Welcome to the intake questionnaire! Over the next four weeks, you will be asked to take a short survey each day.

At the end of each week, if you complete a survey every day of that week you will get a \$5 bonus!

If you enter your cellphone number, you can get text reminders and take the survey straight from your phone! Otherwise, you will receive notifications via email.

Block 1

Enter your cellphone number (or leave blank to only get email alerts).

Block 2

Please enter your email address below.

Block 3

Demographic Information

How old are you?

What is your gender identity? You may select more than one option.

- Female
- Male
- Gender non-conforming
- Transgender
- Other: I identify as
- Prefer not to respond

How would you describe your ethnicity and/or racial background? You may select more than one option.

- American Indian or Alaska Native (e.g., Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Nome Eskimo Community, etc.)
- Asian (e.g., Chinese, Filipino, Asian Indian, Vietnamese, Korean, etc.)
- Black (e.g. African American, Jamaican, Haitian, Nigerian, Ethiopian, Somalian, etc.)
- Latinx, Hispanic, or Spanish origin (e.g., Mexican, Mexican American, Puerto Rican, Cuban, Salvadorian, Dominican, Colombian, etc.)
- Middle Eastern or North African (e.g., Lebanese, Iranian, Egyptian, Syrian, Moroccan, Algerian, etc.)
- Native Hawaiian or Pacific Islander (e.g., Samoan, Chamorro, Tongan, Fijan, Marshallese, etc.)
- White (e.g., German, Irish, English, Italian, Polish, French, etc.)
- Other: I identify as
- Prefer not to respond

Pain History

In the past month, how would you rate your pain on average?



Last semester, about how many times did you experience:

	Never	Rarely	Once or Twice a Month	Almost Every Week	Almost Every Day	I choose note to answer this question
Mild Pain	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Moderate Pain	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Severe Pain	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How often did you experience pain due to behaviors like engaging in sports, playing an instrument, hangovers, or another specific activity you enjoyed?

	Never	Rarely	Once or Twice a Month	Almost Every Week	Almost Every Day	I choose not to answer this question
Mild Pain	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Moderate Pain	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Severe Pain	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How often did you experience pain due to a chronic condition such as migraines, fibromyalgia, PID, arthritis, a back injury, or joint inflammation?

	Never	Rarely	Once or Twice a Month	Almost Every Week	Almost Every Day	I choose not to answer this question
Migranes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fibromyalgia	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
PID	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Arthritis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Back Injury	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Joint Inflammation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please respond to each question or statement by marking one box per row.

	Without any difficulty	With a little difficulty	With some difficulty	With much difficulty	Unable to do	I choose not to answer this question
Are you able to do chores such as vacuuming or yard work?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Are you able to go up and down stairs at a normal pace?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Are you able to go for a walk of at least 15 minutes?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Are you able to run errands and shop?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please respond to each question or statement by marking one box per row.

	Never	Rarely	Sometimes	Usually	Always	I choose not to answer this question
I have trouble doing all of my regular leisure activities with others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have trouble doing all of the family activities that I want to do	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have trouble doing all of my usual work (include work at home)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Never	Rarely	Sometimes	Usually	Always	I choose not to answer this question
I have trouble doing all of the activities with friends that I want to do	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

What activity caused you pain most frequently?

In the past month...

	Very poor	Poor	Fair	Good	Very good	I choose not to answer this question
My sleep quality was	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please respond to each question or statement by marking one box per row.

In the past month...

	Not at all	A little Bit	Somewhat	Quite a bit	Very much	I choose not to answer this question
I feel fatigued	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have trouble starting things because I am tired	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How run-down did you feel on average?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How fatigued were you on average?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please respond to each question or statement by marking one box per row.

In the past month...

	Never	Rarely	Sometimes	Often	Always	I choose not to answer this question
I felt fearful	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found it hard to focus on anything other than my anxiety	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My worries overwhelmed me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt uneasy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please respond to each question or statement by marking one box per row.

In the past month...

	Never	Rarely	Sometimes	Often	Always	I choose not to answer this question
I felt worthless	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt helpless	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt depressed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt hopeless	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please respond to each question or statement by marking one box per row.

In the past month...

	Not at all	A little bit	Somewhat	Quite a bit	Very much	I choose not to answer this question
My sleep was refreshing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I had a problem with my sleep	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I had difficulty falling asleep	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please respond to each question or statement by marking one box per row.

In the past month...

	Not at all	A little bit	Somewhat	Quite a bit	Very much	I choose not to answer this question
How much did pain interfere with your day to day activities?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How much did pain interfere with work around home?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How much did pain interfere with your ability to participate in social activities?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How much did pain interfere with your household chores?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please respond to each question or statement by marking one box per row.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	I choose not to answer this question
No one's been able to tell me exactly why I'm in pain.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My pain is confusing me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I don't know enough about my pain.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	I choose not to answer this question
I can't figure out why I am in pain.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In general, how would you rate your ability to tolerate moderate pain?

Worse than most people Better than most people

0 1 2 3 4 5 6 7 8 9 10

Default Question Block

Welcome Back!

This is the daily questionnaire for 4/27/2020 !

You will first be asked to enter the ID you were provided by email, then you will be asked to complete a short survey about your mood, pain intensity, and other factors.

Voluntary Participation.

Your participation is completely voluntary and you are free to withdraw at any time. If you choose to participate, you may skip any other questions without penalty. In addition, on the last day of the survey, we will ask you to confirm that we may use your data.

Questions?

If you have questions about the study, please contact Nancy Darling (ndarling@oberlin.edu or 440-775-8363) Max Kramer (mkramer@oberlin.edu or 773-318-5225).

If you have questions about your rights as a research participant, please contact Associate Dean Daphne John, Office of the Dean of Arts and Sciences, Cox 101 (ocirb@oberlin.edu. or 440-775-8410).

Block 9

Please Enter your ID code from the Intake Survey

Mood Scale

Indicate if you felt each of the following since your last survey

	No	Yes
Interested	<input type="radio"/>	<input type="radio"/>
Distressed	<input type="radio"/>	<input type="radio"/>
Excited	<input type="radio"/>	<input type="radio"/>
Upset	<input type="radio"/>	<input type="radio"/>
Strong	<input type="radio"/>	<input type="radio"/>
Guilty	<input type="radio"/>	<input type="radio"/>
Scared	<input type="radio"/>	<input type="radio"/>
Hostile	<input type="radio"/>	<input type="radio"/>
Enthusiastic	<input type="radio"/>	<input type="radio"/>
Proud	<input type="radio"/>	<input type="radio"/>

	No	Yes
Irritable	<input type="radio"/>	<input type="radio"/>
Alert	<input type="radio"/>	<input type="radio"/>
Ashamed	<input type="radio"/>	<input type="radio"/>
Inspired	<input type="radio"/>	<input type="radio"/>
Nervous	<input type="radio"/>	<input type="radio"/>
Determined	<input type="radio"/>	<input type="radio"/>
Attentive	<input type="radio"/>	<input type="radio"/>
Jittery	<input type="radio"/>	<input type="radio"/>
Active	<input type="radio"/>	<input type="radio"/>
Afraid	<input type="radio"/>	<input type="radio"/>

Morning routine

Indicate if you did any of the following since your last survey

	No	Yes
Got out of bed	<input type="radio"/>	<input type="radio"/>
Got dressed	<input type="radio"/>	<input type="radio"/>
Left for school/class	<input type="radio"/>	<input type="radio"/>
Ran errands	<input type="radio"/>	<input type="radio"/>
Prepared food	<input type="radio"/>	<input type="radio"/>
Brushed teeth	<input type="radio"/>	<input type="radio"/>
Showered	<input type="radio"/>	<input type="radio"/>
Brushed hair	<input type="radio"/>	<input type="radio"/>
Took medications	<input type="radio"/>	<input type="radio"/>
Washed up	<input type="radio"/>	<input type="radio"/>

Activity checklist

Indicate if you did any of the following since your last survey

	No	Yes
Read a book	<input type="radio"/>	<input type="radio"/>
Created art	<input type="radio"/>	<input type="radio"/>
Played a game	<input type="radio"/>	<input type="radio"/>
Watched TV	<input type="radio"/>	<input type="radio"/>
Wrote in a journal	<input type="radio"/>	<input type="radio"/>
Played music	<input type="radio"/>	<input type="radio"/>
Played with pets	<input type="radio"/>	<input type="radio"/>

	No	Yes
Enjoyed my hobbies	<input type="radio"/>	<input type="radio"/>
Played Sports	<input type="radio"/>	<input type="radio"/>

Helping others

Indicate if you did any of the following since your last survey

	No	Yes
Was supportive to a friend	<input type="radio"/>	<input type="radio"/>
Did homework	<input type="radio"/>	<input type="radio"/>
Helped out family/friends	<input type="radio"/>	<input type="radio"/>
Volunteered	<input type="radio"/>	<input type="radio"/>
Fed pets	<input type="radio"/>	<input type="radio"/>

Diet/Hydration

Indicate if you did any of the following since your last survey

	No	Yes
Ate healthy	<input type="radio"/>	<input type="radio"/>
Followed diet	<input type="radio"/>	<input type="radio"/>
Stayed hydrated	<input type="radio"/>	<input type="radio"/>

Exercise

Indicate if you did any of the following since your last survey

	No	Yes
Exercised	<input type="radio"/>	<input type="radio"/>
Took a stroll	<input type="radio"/>	<input type="radio"/>
Got my heart rate up	<input type="radio"/>	<input type="radio"/>

Mindfulness/Stress Reductions

Indicate if you did any of the following since your last survey

	No	Yes
Was thankful	<input type="radio"/>	<input type="radio"/>

	No	Yes
Mindfulness	<input type="radio"/>	<input type="radio"/>
Meditated	<input type="radio"/>	<input type="radio"/>
Yoga/Tai Chi	<input type="radio"/>	<input type="radio"/>
Prayed	<input type="radio"/>	<input type="radio"/>
Biofeedback	<input type="radio"/>	<input type="radio"/>

Sleep

Indicate if you did any of the following since your last survey

	No	Yes
Went to bed on time	<input type="radio"/>	<input type="radio"/>
Rested during the day	<input type="radio"/>	<input type="radio"/>
Got up on time	<input type="radio"/>	<input type="radio"/>
Got to sleep	<input type="radio"/>	<input type="radio"/>

Pain

How is your pain today?



Contact

More Questions?

If you have questions about the study, please contact Nancy Darling (ndarling@oberlin.edu or 440-775-8363) Max Kramer (mkramer@oberlin.edu or 773-318-5225).

If you have questions about your rights as a research participant, please contact Associate Dean Daphne John, Office of the Dean of Arts and Sciences, Cox 101 (djohn@oberlin.edu or 440-775-8410).

Closing

MK Honors Thesis Dataset Generation

1.0 Setup

In order to synthesize an artificial dataset, we must first load our seed data into R.

```
require(readxl)
```

```
## Loading required package: readxl
```

```
library(readxl)  
library(descr)
```

```
## Warning: package 'descr' was built under R version 3.6.3
```

```
Intake <- read_excel("Intake+Questionnaire+for+MK+Honors+Thesis_December+3,+2019_12.02.xlsx")  
Daily <- read_excel("Daily+Questionnaire+for+MK+Honors+Thesis_December+3,+2019_11.54.xlsx")
```

The data are read in as character vectors by default rather than categorical factors. We must convert the data typing of each variable to factor before proceeding.

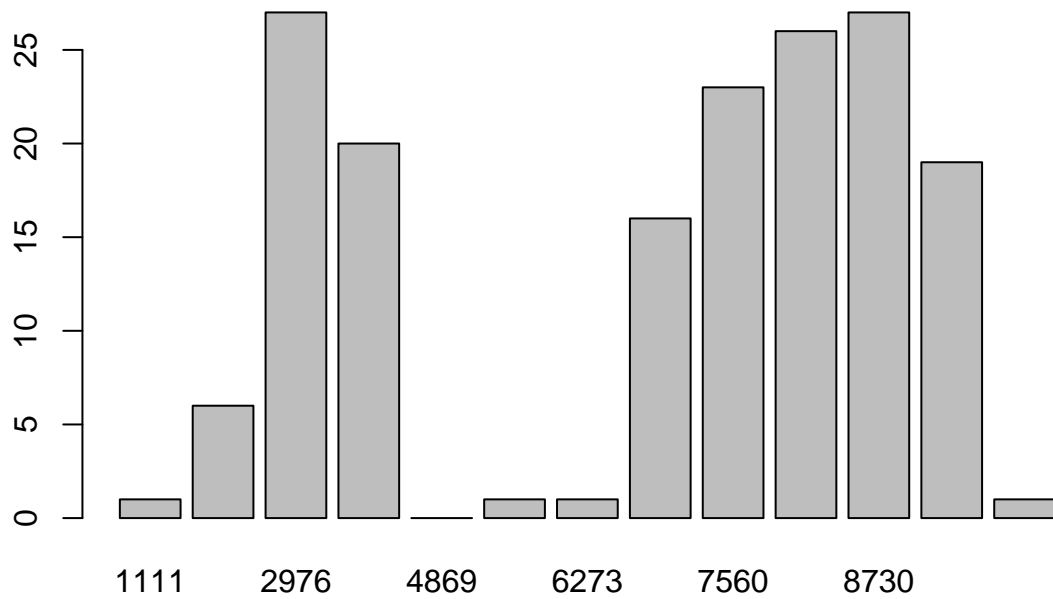
```
Intake2 <- as.data.frame(unclass(Intake))  
Daily2 <- as.data.frame(unclass(Daily))
```

Finally, labels and metadata should be stripped from the dataframe. For convenience, we also relabel the Intake2 and Daily2 datasets to i and d respectively.

```
i <- Intake2[-1,] # renaming and parsing factor labels  
d <- Daily2[-1,] # renaming and parsing factor labels
```

```
Daily2$Q15[Daily2$Q15 == 4869] <- 4689
```

```
freqlist <- freq(Daily2$Q15)
```

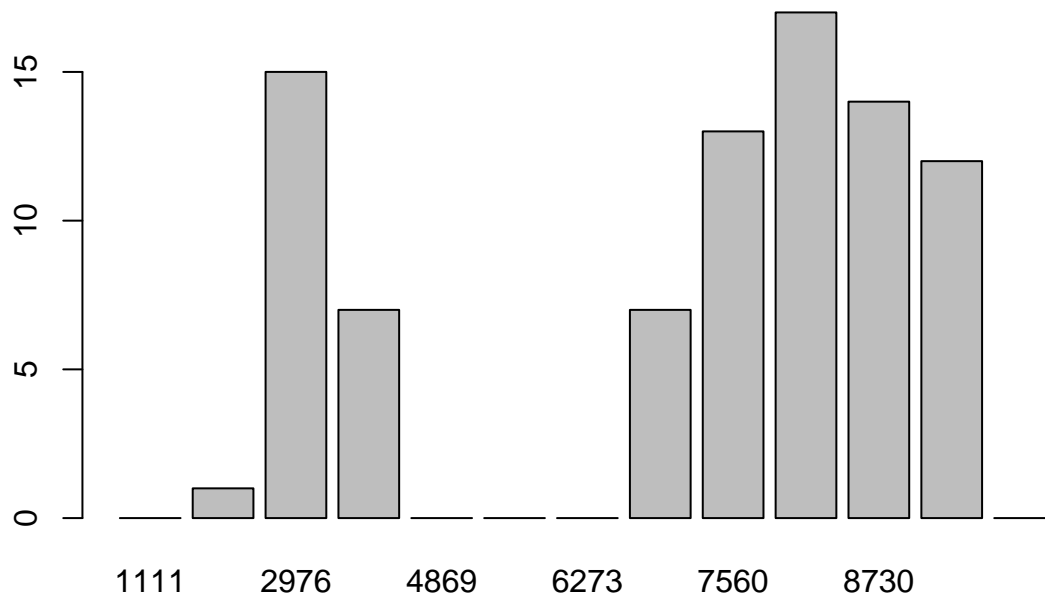


```
f <- which(freqlist[,1] > 1)
f <- names(f)[1:(length(f)-2)]

d <- Daily2[as.character(Daily2$Q15) %in% f,]
d <- d[d$Finished == "True",]

d_drop <- d[!is.na(d$Q17),] # removing NA values

freqlist <- freq(d_drop$Q15)
```



```
f <- which(freqlist[,1] > 1)
f <- names(f)[1:(length(f)-1)]
d_drop <- d_drop[as.character(d_drop$Q15) %in% f,]
d_main <- d_drop[,19:79] # parse metadata
d_main <- na.omit(d_main) ## THIS LINE REMOVES ALL INCOMPLETE CASES ##
d_label <- d_main[,61] # label vector

l_counts <- matrix(0,ncol=5,nrow=5)
for (i in f){
  d_sub <- d_drop[d_drop$Q15 == i,]
  for (row in 2:nrow(d_sub)){
    pastL <- d_sub$Q17[row-1]
    L <- d_sub$Q17[row]
    l_counts[pastL,L] <- l_counts[pastL,L] + 1
    #print(paste(i,row,pastL,L,nrow(d_sub)))
  }
}
roulette_l <- prop.table(l_counts,1)
```

1.1 Artificial Dataset Synthesis

The artificial dataset will be generated using joint distribution roulette wheel sampling.

Variables on daily: 19:78

MOOD: 1 - 20 DAILY FXN: 21 - 30 ACTIVITY FXN: 31 - 39 RELIABILITY FXN: 40 - 44 DIET/HYD FXN: 45 - 47 EXERCISE FXN: 48 - 50 MINDFULNESS/BIOFEED FXN: 51 - 56 SLEEP FXN: 57 - 60 PAIN LVL: 61

```
sampleLabel <- function(day,pastL,d_label,roulette_l){
  if (day == 1){
    rand <- sample(1:length(d_label),1) # randomly sample a label from labels
    L <- d_label[rand] # sample that entry of d_label
    return(L)
  }
  else{
    probs <- roulette_l[pastL,]
    rand <- runif(1)
    for (L in 1:length(probs)){
      if (probs[L] >= rand){
        return(L)
      }
      else{
        rand <- rand - probs[L]
      }
    }
    return(length(probs))
  }
}

createDay <- function(L, d_main, scales, nscales) {
  inst <- c(L) # create new artificial instance with label L
  d_subset <- d_main[d_main$Q17 == L,] # only select cases with label
  for (s_i in 1:nscales){ # instance generation
    start = scales[2*s_i-1]
    end = scales[2*s_i]
    rand <- sample(1:nrow(d_subset),1)
    for(att in start:end){
      inst <- c(inst,as.integer(d_subset[rand,att]) - 2)
    }
  }
  return(inst)
}

rouletteWheelSampling <- function(n)
{
  #Variables
  art <- data.frame()
  scales <- c(c(1,20),c(21,30),c(31,39),c(40,44),c(45,47),c(48,50),c(51,56),c(57,60))
  nscales <- length(scales)/2

  paste('Now Generating',n,'samples',sep=" ")

  for (person in 1:n) {
    L <- 1 # arbitrary
    for (day in 1:28) {
```

```

L <- sampleLabel(day,L,d_label,roulette_1)
meas <- createDay(L, d_main, scales, nscales)
inst <- c(person, day, meas)
art <- rbind(art, inst)
}
}
my_names <- names(Daily)

for (i in 1:ncol(Daily)) {
  my_names[i] <- paste0(my_names[i], ": ", Daily[1, i])
}
my_names <- my_names[19:79]

names(art) <- c("participant","day","label",my_names[1:60])
return(art)
}

art <- rouletteWheelSampling(100)
write.csv(art,'HonorsData.csv',row.names = FALSE)

```

"""
*Developing a Machine Learning Algorithm to Predict Daily Functioning in a Population
of Adolescents Living With Chronic Pain*

*Psychology Honors Thesis
Max Kramer, Oberlin College Class of 2020*

This code is designed to be run after the R markdown file generates the dataset
"""

```
#  
# EXTERNAL LIBRARY IMPORTATION  
#  
import sys  
import csv  
from math import sqrt  
import pandas as pd  
import numpy as np  
import matplotlib.pyplot  
from sklearn import preprocessing, metrics, model_selection  
  
#  
# MODELS  
#  
from sklearn.naive_bayes import GaussianNB  
from sklearn.neural_network import MLPClassifier  
from sklearn.neural_network import MLPRegressor  
from sklearn.tree import DecisionTreeClassifier  
from sklearn.tree import DecisionTreeRegressor  
from sklearn.ensemble import RandomForestClassifier  
from sklearn.ensemble import RandomForestRegressor  
from sklearn.linear_model import LogisticRegression  
from sklearn.linear_model import LinearRegression  
  
#  
# Helper Functions  
#  
def readin(path, verbose=False): # takes verbose flag  
    dataset = pd.read_csv(path) # read csv from supplied filepath  
    if verbose: # for diagnostics  
        print('dataset contains {} instances and {} attributes'.format(dataset.shape[0], dataset.shape[1] - 1))  
    return dataset  
  
def instanceFormat(dataset,k,n):  
    X = []  
    y = []  
    participants = np.arange(1,101)  
    for participant in participants:  
        for label_day in range(k+n,29):  
            label = (dataset.label)[(dataset.participant == participant) & (dataset.day == label_day)].to_numpy()[0]  
            y.append(label)  
            inst = []  
            for day_offset in range(k-1,-1,-1):  
                attribute_day = label_day - n - day_offset  
                original_row = dataset[(dataset.participant == participant) & (dataset.day == attribute_day)].to_numpy()[0].tolist()  
                if day_offset == 0 and n == 0:  
                    inst.extend(original_row[3:])  
                else:  
                    inst.extend(original_row[2:])  
            X.append(inst)  
    return np.asarray(X),np.asarray(y)
```

```

def split(X, y, train_percent, seed, verbose=False): # Split dataset
    X_train, X_test, y_train, y_test = model_selection.train_test_split(X, y, train_size=train_percent,
                                                                    random_state=seed) # creates test and train set

    if verbose: # for diagnostics
        print('training set contains {} instances'.format(X_train.shape[0]))
        print('test set contains {} instances'.format(X_test.shape[0]))
        print('split complete')
    return X_test.shape[0], X_train, X_test, y_train, y_test

def ConfidenceInterval(acc, testset_size): # generate 95% CI with Bonferroni Correction for 4 comparisons per dataset
    CI = 2.39 * sqrt((acc * (1 - acc)) / testset_size)
    return CI

##### CLASSIFICATION MODELS #####

#
# ASSUMPTIONS TEST: Naive Bayes
#

def NaiveBayes(k,N,testset_size, dataset, X_train, X_test, y_train, y_test):
    clf = GaussianNB()
    clf.fit(X_train, y_train)
    acc = clf.score(X_test, y_test)
    predicted = clf.predict(X_test)
    summary = metrics.classification_report(y_test, predicted)
    conmat = metrics.confusion_matrix(y_test, predicted)
    with open('results_NB_{}_{}.csv'.format(k,N), mode='w') as csvout:
        writer = csv.writer(csvout, delimiter=',')
        writer.writerows(conmat)
    CI = ConfidenceInterval(acc, testset_size)
    print('Naive Bayes: {} accuracy 95% CI : [{} , {}]'.format("%.3f" % acc, "%.3f" % (acc - CI), "%.3f" % (acc + CI)))
    print()
    print(summary)
    return acc

#
# BASELINE: Logistic Regression
#

def LR(k,N,testset_size, dataset, X_train, X_test, y_train, y_test,seed):
    clf = LogisticRegression(random_state=seed)
    clf.fit(X_train, y_train)
    acc = clf.score(X_test, y_test)
    predicted = clf.predict(X_test)
    summary = metrics.classification_report(y_test, predicted)
    conmat = metrics.confusion_matrix(y_test, predicted)
    with open('results_LR_{}_{}.csv'.format(k,N), mode='w') as csvout:
        writer = csv.writer(csvout, delimiter=',')
        writer.writerows(conmat)
    CI = ConfidenceInterval(acc, testset_size)
    print('LR: {} accuracy 95% CI : [{} , {}]'.format("%.3f" % acc, "%.3f" % (acc - CI), "%.3f" % (acc + CI)))
    print()
    #debug = 1
    #if debug:
        #print(clf.intercept_)
        #print(clf.coef_[0].size)
        #print(clf.coef_)
        #print()
    print(summary)
    return acc

#
# Decision Tree
#

def DecisionTree(k,N,testset_size, dataset, X_train, X_test, y_train, y_test, seed):

```

```

clf = DecisionTreeClassifier(random_state=seed)
clf.fit(X_train, y_train) # fit model to data
acc = clf.score(X_test, y_test)
predicted = clf.predict(X_test)
summary = metrics.classification_report(y_test, predicted)
conmat = metrics.confusion_matrix(y_test, predicted)
with open('results_DT_{k}_{N}.csv'.format(k,N), mode='w') as csvout:
    writer = csv.writer(csvout, delimiter=',')
    writer.writerows(conmat)
CI = ConfidenceInterval(acc, testset_size)
print(
    'Decision Tree: {} accuracy 95% CI : [{} , {}]'.format('%0.3f' % acc, '%0.3f' % (acc - CI), '%0.3f' % (acc + CI)))
print()
print(summary)
return acc

```

```

#
# Random Forest
#

```

```

def RandomForest(k,N,testset_size, D, dataset, X_train, X_test, y_train, y_test, seed):
    clf = RandomForestClassifier(n_estimators=100, random_state=seed)
    clf.fit(X_train, y_train) # fit model to data
    acc = clf.score(X_test, y_test)
    predicted = clf.predict(X_test)
    summary = metrics.classification_report(y_test, predicted)
    conmat = metrics.confusion_matrix(y_test, predicted)
    with open('results_Forest_{k}_{N}.csv'.format(k,N), mode='w') as csvout:
        writer = csv.writer(csvout, delimiter=',')
        writer.writerows(conmat)
    CI = ConfidenceInterval(acc, testset_size)
    print(
        'Random Forest: {} accuracy 95% CI : [{} , {}]'.format('%0.3f' % acc, '%0.3f' % (acc - CI), '%0.3f' % (acc + CI)))
    if k == 1 and N == 0:
        importance = pd.DataFrame(clf.feature_importances_, index = D.columns[3:],
            columns=['importance']).sort_values('importance', ascending=False)
        #debug = 1
        #if debug:
            #print(importance)
        importance.to_csv('importance_RF.csv')
        print()
    print(summary)
    return acc

```

```

#
# Neural Network
#

```

```

def shallowNN(k,N,testset_size, dataset, X_train, X_test, y_train, y_test, seed):
    clf = MLPClassifier(hidden_layer_sizes=(150,150),solver='adam', max_iter=5000, learning_rate_init=0.001,
        random_state=seed)
    clf.fit(X_train, y_train)
    acc = clf.score(X_test, y_test)
    predicted = clf.predict(X_test)
    summary = metrics.classification_report(y_test, predicted)
    conmat = metrics.confusion_matrix(y_test, predicted)
    with open('results_NN_{k}_{N}.csv'.format(k,N), mode='w') as csvout:
        writer = csv.writer(csvout, delimiter=',')
        writer.writerows(conmat)
    CI = ConfidenceInterval(acc, testset_size)
    print('Shallow NN: {} accuracy 95% CI : [{} , {}]'.format('%0.3f' % acc, '%0.3f' % (acc - CI), '%0.3f' % (acc + CI)))
    print()
    print(summary)
    return acc

```

```

##### REGRESSION MODELS #####

```



```

#
# BASELINE: Linear Regression
#

def LinReg(testset_size, dataset, X_train, X_test, y_train, y_test):
    reg = LinearRegression()
    reg.fit(X_train, y_train)
    R2 = reg.score(X_test, y_test)
    predicted = reg.predict(X_test)
    MSE = metrics.mean_squared_error(y_test, predicted)
    #CI = ConfidenceInterval(acc, testset_size)
    print('Linear Regression: R^2 = {} MSE = {}'.format('%0.3f' % R2, '%0.3f' % MSE))
    print()
    return R2, MSE

#
# Decision Tree Regressor
#

def DecisionTreereg(testset_size, dataset, X_train, X_test, y_train, y_test, seed):
    reg = DecisionTreeRegressor(random_state=seed)
    reg.fit(X_train, y_train) # fit model to data
    R2 = reg.score(X_test, y_test)
    predicted = reg.predict(X_test)
    MSE = metrics.mean_squared_error(y_test, predicted)
    #CI = ConfidenceInterval(MSE, testset_size)
    print(
        'Decision Tree: R^2 = {} MSE = {}'.format('%0.3f' % R2, '%0.3f' % MSE))
    print()
    return R2, MSE

#
# Random Forest Regressor
#

def RandomForestreg(testset_size, dataset, X_train, X_test, y_train, y_test, seed):
    reg = RandomForestRegressor(n_estimators=100, random_state=seed)
    reg.fit(X_train, y_train) # fit model to data
    R2 = reg.score(X_test, y_test)
    predicted = reg.predict(X_test)
    MSE = metrics.mean_squared_error(y_test, predicted)
    #CI = ConfidenceInterval(acc, testset_size)
    print(
        'Random Forest: R^2 = {} MSE = {}'.format('%0.3f' % R2, '%0.3f' % MSE))
    print()
    return R2, MSE

#
# Neural Network Regressor
#

def shallowNNreg(testset_size, dataset, X_train, X_test, y_train, y_test, seed):
    reg = MLPRegressor(solver='adam', max_iter=1000, alpha=1e-3, random_state=seed)
    reg.fit(X_train, y_train)
    R2 = reg.score(X_test, y_test)
    predicted = reg.predict(X_test)
    MSE = metrics.mean_squared_error(y_test, predicted)
    #CI = ConfidenceInterval(acc, testset_size)
    print('Shallow NN: R^2 = {} MSE = {}'.format('%0.3f' % R2, '%0.3f' % MSE))
    print()
    return R2, MSE

##### MAIN #####

def main():

```

```

seed = int(sys.argv[1])
LR_table = np.zeros((7,8))
NB_table = np.zeros((7,8))
DT_table = np.zeros((7,8))
RF_table = np.zeros((7,8))
NN_table = np.zeros((7,8))
# Linreg_R2_table = np.zeros((7,8))
# DTreg_R2_table = np.zeros((7,8))
# RFreg_R2_table = np.zeros((7,8))
# NNreg_R2_table = np.zeros((7,8))
# Linreg_MSE_table = np.zeros((7,8))
# DTreg_MSE_table = np.zeros((7,8))
# RFreg_MSE_table = np.zeros((7,8))
# NNreg_MSE_table = np.zeros((7,8))
Dataset = 'Honors'
dataset = readin('./HonorsData.csv')
for k in range(1,8):
    for N in range(8):
        X , y = instanceFormat(dataset,k,N)
        testset_size, X_train, X_test, y_train, y_test = split(X,y,0.85,seed)
        # print('K is {}, N is {}'.format(k,N))
        # R2_LR, MSE_LR = LinReg(testset_size, dataset, X_train, X_test, y_train, y_test)
        # Linreg_R2_table[k-1,N] = R2_LR
        # Linreg_MSE_table[k-1,N] = MSE_LR
        # print('K is {}, N is {}'.format(k,N))
        # R2_DT, MSE_DT = DecisionTreereg(testset_size, Dataset, X_train, X_test, y_train, y_test, seed)
        # DTreg_R2_table[k-1,N] = R2_DT
        # DTreg_MSE_table[k-1,N] = MSE_DT
        # print('K is {}, N is {}'.format(k,N))
        # R2_RF, MSE_RF = RandomForestreg(testset_size, dataset, X_train, X_test, y_train, y_test, seed)
        # RFreg_R2_table[k-1,N] = R2_RF
        # RFreg_MSE_table[k-1,N] = MSE_RF
        # print('K is {}, N is {}'.format(k,N))
        # R2_NN, MSE_NN = shallowNNreg(testset_size, dataset, X_train, X_test, y_train, y_test, seed)
        # NNreg_R2_table[k-1,N] = R2_NN
        # NNreg_MSE_table[k-1,N] = MSE_NN

        print('K is {}, N is {}'.format(k,N))
        LR_acc = LR(k,N,testset_size, Dataset, X_train, X_test, y_train, y_test,seed)
        LR_table[k-1,N] = LR_acc
        print('K is {}, N is {}'.format(k,N))
        NB_acc = NaiveBayes(k,N,testset_size, Dataset, X_train, X_test, y_train, y_test)
        NB_table[k-1,N] = NB_acc
        print('K is {}, N is {}'.format(k,N))
        DT_acc = DecisionTree(k,N,testset_size, Dataset, X_train, X_test, y_train, y_test, seed)
        DT_table[k-1,N] = DT_acc
        print('K is {}, N is {}'.format(k,N))
        RF_acc = RandomForest(k,N,testset_size, dataset, Dataset, X_train, X_test, y_train, y_test, seed)
        RF_table[k-1,N] = RF_acc
        print('K is {}, N is {}'.format(k,N))
        NN_acc = shallowNN(k,N,testset_size, Dataset, X_train, X_test, y_train, y_test, seed)
        NN_table[k-1,N] = NN_acc
        print("-----")

# np.savetxt('Linreg_R2.csv',Linreg_R2_table,delimiter=',',fmt='%f')
# np.savetxt('Linreg_MSE.csv',Linreg_MSE_table,delimiter=',',fmt='%f')
# np.savetxt('DTreg_R2.csv',DTreg_R2_table,delimiter=',',fmt='%f')
# np.savetxt('DTreg_MSE.csv',DTreg_MSE_table,delimiter=',',fmt='%f')
# np.savetxt('RFreg_R2.csv',RFreg_R2_table,delimiter=',',fmt='%f')
# np.savetxt('RFreg_MSE.csv',RFreg_MSE_table,delimiter=',',fmt='%f')
# np.savetxt('NNreg_R2.csv',NNreg_R2_table,delimiter=',',fmt='%f')
# np.savetxt('NNreg_MSE.csv',NNreg_MSE_table,delimiter=',',fmt='%f')

np.savetxt('LR.csv',LR_table,delimiter=',',fmt='%f')
np.savetxt('NB.csv',NB_table,delimiter=',',fmt='%f')
np.savetxt('DT.csv',DT_table,delimiter=',',fmt='%f')

```

```
np.savetxt('RF.csv',RF_table,delimiter=',',fmt='%f')  
np.savetxt('NN.csv',NN_table,delimiter=',',fmt='%f')
```

```
if __name__ == '__main__':  
    main()
```