Addictive Parameter In Ensemble Technique & Implementation Using Flask Server

G.Baby Saral[1*], R.Priya[2]

[1]Ph.D. Research Scholar, Department of Computer Application, VISTAS, Chennai.
[2]Professor, Department of Computer Application, VISTAS, Chennai
[1]online.babysaral@gmail.com, [2]priyaa.research@gmail.com

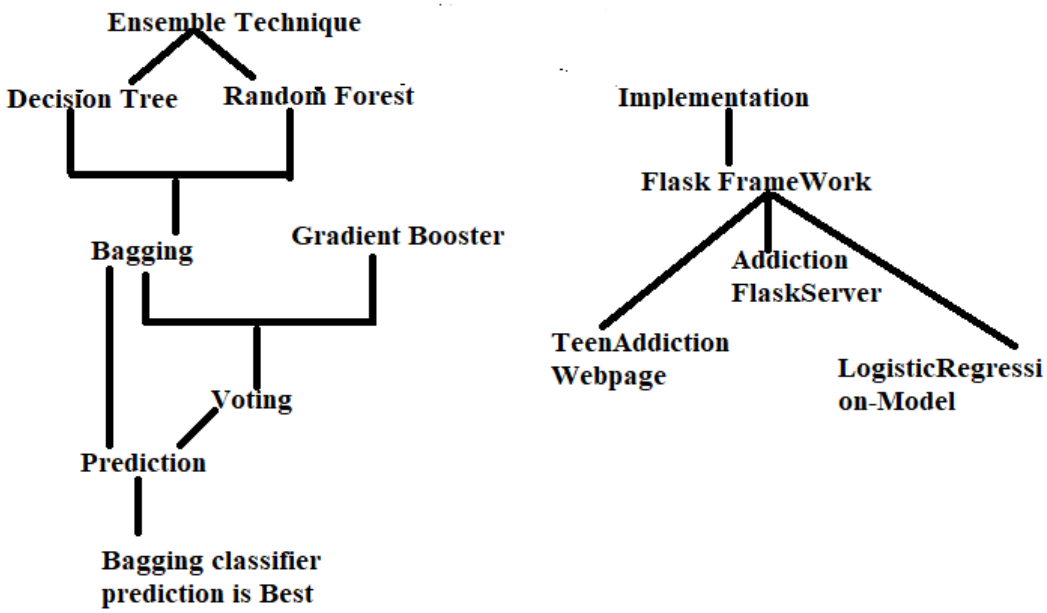
Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract:Exceeding use of screen has completely switched the teens routine life who are addicted with social media and other electronic accessories, which results in psychological disorder like anger, stress, depression, anxiety etc., The addiction data set applied with some algorithm like Random forest, decision tree and so on. The Ensemble Technique bagging and voting classifier are used for better prediction. The bagging classifier which combines the decision tree and random forest, the voting classifier which compares the bagging classifier and gradient booster. Among the two classifier bagging classifier gives the best prediction result. Implementation proceeded with webpage, Flask server and python model using logistic regression model. In future the app can be used as a standalone system, which now works in local server.

Keyword: Ensemble Technique, Flask Server, Machine Learning, Python.

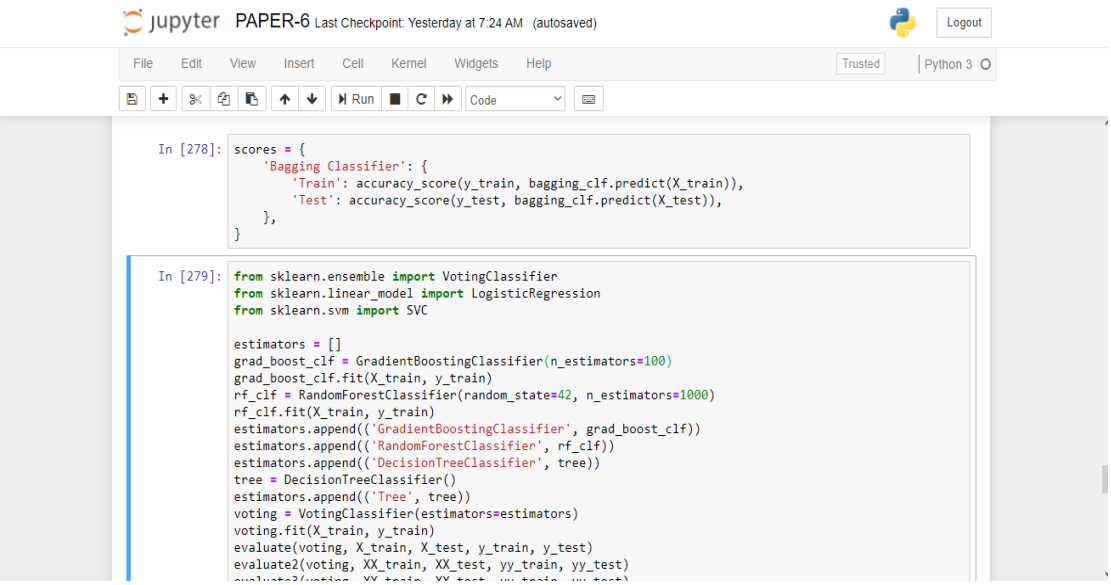
1. Introduction

Electric appliance abnormality is a symptom of neurological addiction influenced to use appliance such as the internet, game, Facebook, whatsapp and social media, so anxiety, depression, fear leads to health problem in the whole world and it affects all group all men, women and kids[5].All the outcome of anxiety and depression disorder results in weight loss or gain and psychological disorder like tension, fear(Much more fear avoid the situation that involve certain negative judgment), anger, low[6] concentration and many. A questionnaire study of students [1][2] forecasts the addictive diseases fatigue, depression, and anxiety[7] as a result of long-term screen use and addiction. The screen addictive disease details was interrelated and many algorithms are applied like random-forest, decision-tree etc for prediction. The Ensemble[10] technique which combine more one model for prediction. The most popular algorithms for combining the prediction from different models are bagging and voting building multiple models.



2. Bagging Classifier

The bagging involves taking multiple samples from data set with replacing that is choose the same value multiple times and train the model for each sample(sub model).Calculate the mean of each sub sample. Collect the mean of all the sub samples and among this sub sample calculate the average of all the sub sample, use the estimated mean for data. The final output prediction averaged across the prediction of all the sub models. The two bagging models involved are bagged decision tree and Random Forest.



The Bagged Decision Tree

The bagging decision tree achieve better with all the model which have high variance and decision tree are build-up without pruning. The bagging[8] which is an ensemble with meta-estimator fits each random subset model with initial data set corporate each and every specific prediction[11] either by averaging or voting to form final prediction. In this addictive prediction three model Gradient Booster, Random Forest and Decision Tee are applied training and testing score are estimated, the Random Forest and Decision Tee have the same accuracy score, which satisfies the Bagging classifier, hence the combination of two classifier gives exact prediction use ensemble technique and the scores are specified in the below table 1.The parameter of bagging classifier are base estimator fits on random subsets of dataset estimators ensemble the base estimators, max samples and max features draw samples & features from X to each base estimator, all these parameter are imported from sklearn ensemble bagging classifier.

S.No	CLASSIFIER	TRAINIG ACCURACY SCORE	TESTING ACCURACY SCORE
1.	GRADIENTBOOSTING	0.9799	0.8696
2.	RANDOM FOREST	0.9624	1.0000
3.	DECISION TREE	0.9624	1.0000

Random Forest

The Random Forest is meta estimator that fits number of decision tree on different small segment of data-set, the averaging technique balance and improves the prediction, accuracy and control over-fit. The small segment or sub sample is always same as the original or initial sample size but its drawn with replacement (same sample multiple times).Each sub sample is a tree in random forest, the maximum depth parameter calculates the maximum trees to estimate the depth, minimum samples split internal with required number of samples said to be as leaf node. The node will be split- ted if the split induces a decrease of the impurity greater than or equal to minimum impurity decreased value by using the parameter called minimum impurity decrease by applying ensemble random forest model[12][13].

Stochastic Gradient Boosting

Stochastic Gradient Boosting also called as Gradient Boosting is an ensemble technique to improve the performance using ensemble [3].Gradient-Boosting is an additional model which allows optimization of inconsistent differential loss functions. In each and every step n classes regression trees are fit on the negative gradient of the biform or multiform deviance loss function. Binary classification is a special case where only a single regression tree is induced.

Gradient Boosting Parameters

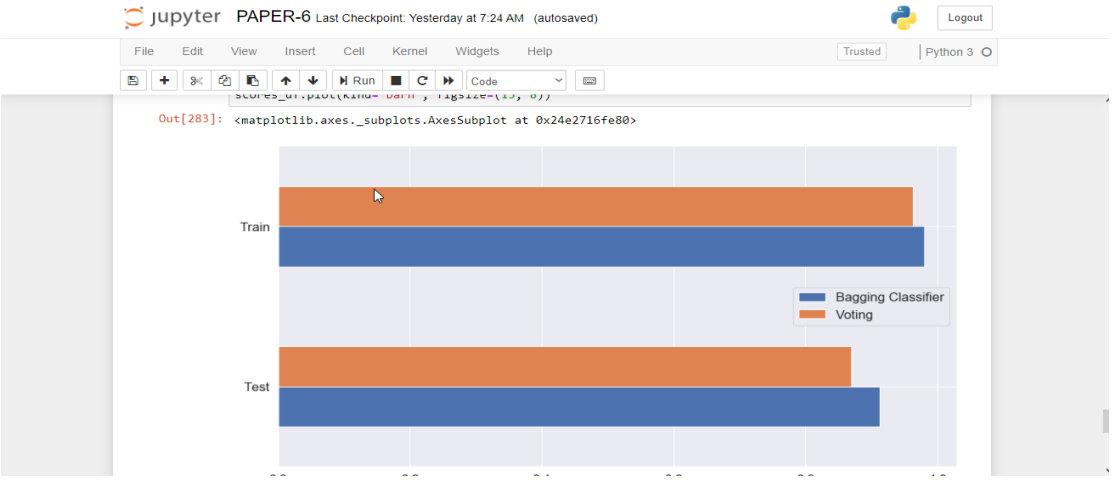
Learning rate: Learning rate shrinks the contribution of each tree by learning rate. There is a trade-off between learning rate and n estimators.

N estimators: The number of boosting stages to perform. Gradient boosting is purely robust to over-fitting, hence big number results in best performance.

Sub sample:The fraction of samples to be used for fitting the individual base learners, if smaller than 1.0 this results in stochastic gradient boosting. subsample interacts with the parameter n estimators. Choosing sub-sample <1.0 leads to a reduction of variance and an increase in bias.

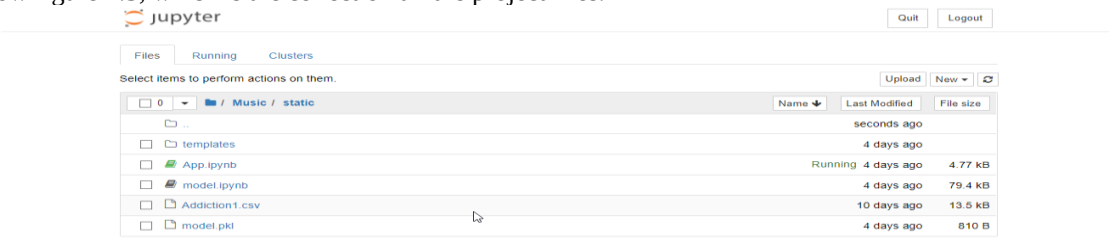
Criterion: The function to measure the quality of a split. Supported criteria are mean squared error with improvement score by Friedman, “mse” for mean squared error and “mae” for mean absolute error. The default value of “friedman mse” is generally the best as it can provide a better approximation is some cases.

Min samples split: The minimum number of samples required to split an internal node.



3. Implementation Using Flask

Flask is a web development framework which has the project structure with static folder with two python file, one model file and one data-set that is csv file, one folder called templates which has HTML, style sheet file for designing the page. Flask is a web-app written in python and deploys the app without the need of protocol and thread management, many tools and libraries to create the addictive prediction app.pip install flask command install the flask in anaconda of python. Import the flask, request, render template and pickle. The project structure in below figure 1.3, which is the collection all the project files.

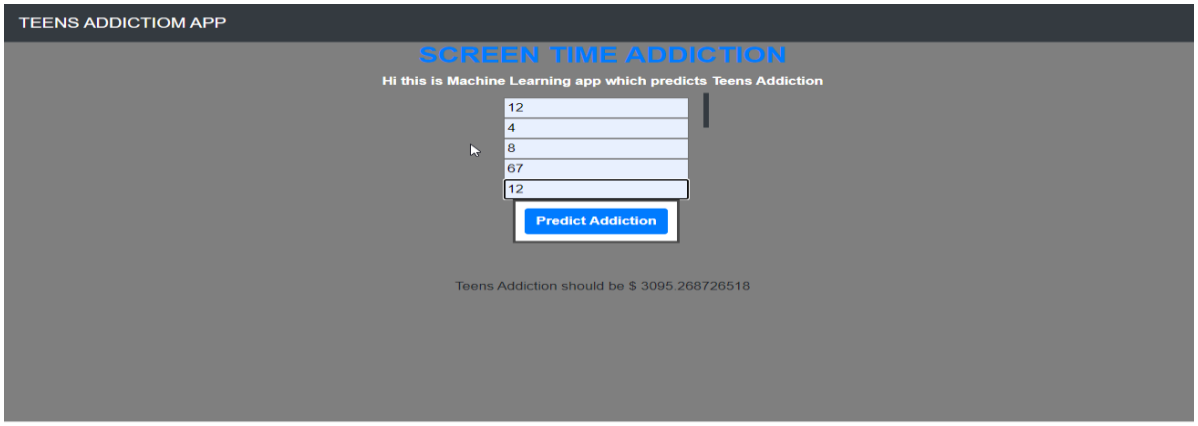


Teen Addiction Web-page

The teens addiction page designed with four attribute text box like spending, lost, within , upset and a submit button called prediction addiction, which gives the predicted value with the help of logistic regression model[14][15],cascading style sheet outer layer of teen addiction webpage control the size, color, text, background color and positioning the element.

<form action="{{ url_for('predict')}}"method="post">
<input type="text" name="Spending" placeholder="Spending" required="required" />
<input type="text" name="Lost" placeholder="Lost" required="required" />

```
<input type="text" name="within" placeholder="within" required="required" /><br>
<input type="text" name="Upset" placeholder="Upset" required="required" /><br>
<input type="text" name="AddictedNot Addicted" placeholder="AddictedNot Addicted"
required="required" /><br>
</div>
<!-- Show button -->
<div class="button_cont" align="center"><a class="button_css" href="" target="_blank" rel="nofollow
noopener">
<button type="submit" class="btn btn-primary btn-block btn-large"><strong>Predict
Addiction</strong></button></a>
</div>
</form>
```



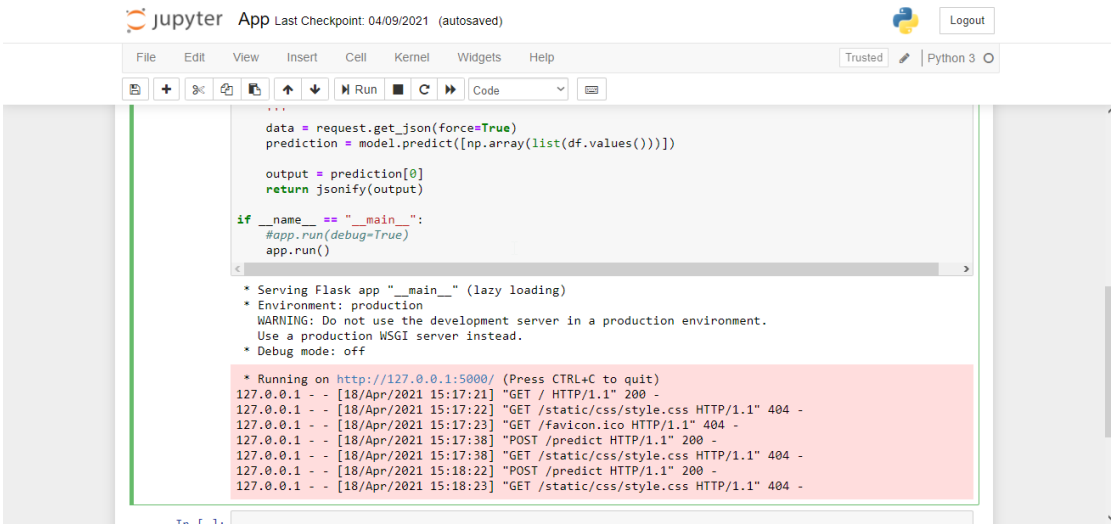
4. Addiction Flaskserver

The Teens addicted API receives addicted attribute values through graphical user interface calculate the predicted addiction value based on logistic regression model and to get the output prediction values, the input should be de-serialize with the help of pickle model in the form of python object [16][17].The main page is index.html, on predicting(submitting) form values the post, render template and request ,the final output prediction is displayed.

Addiction Serial and De-serial

The pickle model dump the regression algorithm and written using write mode as a python object on the disk can be transferred as de-serialized by loading read mode in App. ipynb file.

```
pickle.dump(regressor, open('model.pkl','wb'))
app = Flask(__name__)
model = pickle.load(open('model.pkl', 'rb'))
```



5. Logistic Regression Model

The numpy, matplotlibrary, pandas and pickle files are imported, the dataset Addiction1.csv file is read, the NaN value is checked for all independent value and the target value with mean, the in place value is true, the training data[8][9] in fit in the model. Dump the model an predict the addictive attribute value.

```
# Saving Teens Addiction model to disk
pickle.dump(regressor, open('model.pkl','wb'))
In [15]:

# Loading Teens Addiction model to compare the results
model = pickle.load(open('model.pkl','rb'))
print(model.predict([[ 1, 3, 2, 9, 6]]))
[4.64568606]
```

6. Conclusion

Increase in Digital device usage is a indication of Psychological disorder Fear, Anger, s tress, depression, anxiety are predicted, using Ensemble technique known as bagging and voting classifier, bagging associate the Random Forest and Decision Tree ,voting bind the bagging and Gradient Booster, final best prediction results by bagging classifier.The implementation by using FlaskServer,which works in local server,in future used as independent Addiction App.

References

1. “Electronic screen addiction in digital world “, ,IJICS,Volume 6, Issue 6, June 2019,http://ijics.com.
2. Screen Time Addiction and psychological metric analysisamong teens using machine learning techniques”,G.Baby Saral, R.Priya,JARDCS.Vol. 11,10-Special Issue 2019.
3. XGBoost:A Scalable Tree Boosting System”,Tianqi Chen,Carlos Guestrin,Springer,August,2016,https://doi.org/10.1145/2939672.2939785.
4. Asp vision:optically computing the first layer of convolutional neural networks using angle sensitive pixels”,H.G.Chen,S.Jayasuriya,J.Yang,j.Stephen,IEEE,2016
5. Performance Analysis of machine learning classifier for non-technical loss detection”,khawaja MoyeezUllah Ghori,Muhammad Imran,Asad Nawaz,Springer,January 18 2020,https://doi.org/10.1007/s12652-019-01649-9.
6. EEEG-Based Identification of Latent Emotional Disorder Using the Machine Learning Approach”,Yaling,Fan Wul,Lei Du,IEEE 2019,978-1-5386-6243-4/19.
7. Clustering based semi-supervised machine learning for DDoS attack classification”,Muhammad Aamir,Syed Mustafa Ali Zaidi,ElSEVIER,April 2019,https://doi.org/10.1016/j.jksuci.2019.02.003.
8. Online power quality disturbance detection by support vector machine insmart meter”,Imtiaz PARVEZ,Maryamossadat AGHILI,Arif I.SARWAT,Springer 2019,https://doi.org/10.1007/s40565-018-0488-z.

9. Black-Box vs. White-Box: Understanding Their Advantages and Weaknessess from a Practical of View",octavio,Loyala,Gonzalez,IEEE,2019,10.1109/ACCESS.2019.2949286.
10. ".A mixed approach of deep learning method and rule-based method to improve aspect level sentiment analysis",P.Ray,A.Chakrabarti,Sciencedirect,March 2019,http://www.sciencedirect.com/science/article/pii/S2210832718303156.
11. A comprehensive evaluation of ensemble learning for stock-market prediction",Isaac kofi Nti,Adebayo Felix Adekoya,Springer,March 2020,https://doi.org/10.1186/s40537-020-00299-5.
12. 12."On the economic significance of stock market prediction and the no free lunch theorem",Bousono-Calzon C,Bustarviejo-Munoz J,Aceituno-Aceituno P,IEEE,March,2019,https://doi.org/10.1109/ACCESS.2019.2921092.
13. Random forest based feature selection of macroeconomic variables for stock market prediction",Adekoya AF,Weyori BA,Science publication,March 2019,https://doi.org/10.3844/ajassp.2019.200.212.
14. Towards comparing and using machine learning techniques for detecting and predicting heart attack and diseases ",IEEE, April 2019,10.1109/BigData47090.2019.9005488.
15. Smart mobile application to recognize tomato leaf diseases using Convolution Neural Networks",Azeddine Elhassouny,Florentin Smarandache,IEEE,2019,doi:10.1109/iccsre.2019.8807737.
16. An Introduction to Machine Learning for Panel Data",James Ming,chen,Springer March 2021,https://doi.org/10.1007/s11294-021-09815-6.
17. An app to assist farmers in the identification of diseases and pests of cofee leaves using deep learning",Jose G.M.Esgario Pedro B.C. de Castro,Elsevier,Feb,2021,https://doi.org/10.1016/j.inpa.2021.01.004.