Abstract

Drunk driving is a very widespread problem, causing many casualties and millions of dollars in insurance and damages per year. In 2020, despite the COVID19 pandemic greatly reducing road traffic, about 32 people died per *day* in the US alone from DUI (Driving under Influence)(nhtsa.gov). Existing solutions such as police testing drivers suspected to be under influence is simply impractical considering the number of possible intoxicated drivers. With more and more people gaining access to a vehicle, it is crucial that more effective strategies be developed to detect and combat drunk driving. The purpose of this study is to analyze whether Artificial Intelligence can be used to more effectively detect and prevent drunk driving and if a Machine Learning models like Logistic Regression and Decision Tree can be more accurate than a human police officer. To address this question, a dataset named drunkImagesWebp was used. Two machine learning algorithms, Logistic regression and decision tree were then trained on this dataset with facial image data of intoxicated people to accurately predict the sobriety of humans based on facial cues. After testing this model, it became clear that both Logistic regression and decision tree models can indeed accurately test a driver for signs of intoxication with well over 90% accuracy compared to human-administered tests with can only hit up to around 75% accuracy. By comparison, both the Logistic regression and Decision tree algorithms detected intoxication with 96% accuracy. This paper shows the potential AI can have in creating an automated solution to detecting and ultimately preventing drunk driving.

Introduction & Background

Drunk driving is one of today's principal issues and is reported to cause an estimated 10,000 deaths per year. Current safety measures intended to combat DUI include police stopping a suspicious or erratically moving vehicle and subjecting the driver to breath/coordination/blood tests to identify intoxicated drivers. However, obviously, with the sheer number of drivers on the road and the difficulty of choosing who to pull over and test, the existing safety measures are inefficient and mostly rely on drivers making the responsible decision themselves. Additionally, sobriety tests such as the Horizontal Gaze Nystagmus test and breathalyzer tests simply are not accurate enough to detect all drunk drivers and can lead to false positive cases as well. Conventional intoxication tests require cooperation from the subject the test is administered too, dragging out an already time-inefficient process. On the other hand, an automated model can detect intoxication in seconds, without the need for cooperation. This would allow an automated solution to address the widescale aspect of the issue and with little time wasted. The current model of DUI prevention is simply not sustainable and contributes to a serious potential threat on today's roads. Automating the detection and prevention of intoxicated drivers can allow more accurate solutions free of human error to truly address the widespread problem and make the roads a safer place. ~~This study explored multiple machine learning models to determine whether Artificial Intelligence can accurately detect whether a driver is DUI far more accurately than conventional methods as well as completely independently of human intervention. This was done so by choosing a dataset that contains hundreds of facial images of people in various states of intoxication and sobriety, then training several machine learning models on this data set, concluding with testing these models and scoring how accurately they could predict whether a person was intoxicated based on facial cues. The study compared different machine learning~~

~~methods against each other and against the accuracy scores of traditional human intoxication tests to test which was the most accurate at predicting intoxication.~~

Dataset (Preprocessing data)

To build a machine learning model to accurately identify intoxication based on facial features, a dataset is required of image data of facial features of people before and during intoxication. The dataset used is pulled from a public preexisting dataset named drunkImagesWebp created by SoBr, a company that builds touch-based alcohol detection devices, published on GitHub. The dataset contained images of 53 different people, which contained 4 quadrants with an image of the individual in various states of intoxication, equaling 212 samples in total. The 4 images for each subject were classified based on the level of intoxication, from fully sober to heavily intoxicated. To process the dataset, all the images were split into 4 separate images, which were classified into 4 lists (sober, drunk1, drunk2, and drunk 3) based on the intoxication level. These lists were reshaped into 2 dimensions, and another list was created as classification data for each item on the image data lists. These two datasets, X(image data containing the 4 lists) and Y(classification data), were then split into 75% training and 25% testing data.

Methods (Figures, equations)

In order to analyze whether artificial intelligence can effectively detect or prevent drunk driving, logistic regression and decision tree were used to test whether machine learning algorithms can detect intoxication and the general level of intoxication in an individual. A dataset containing image data of the facial expressions of a person when intoxicated vs sober was processed into two arrays containing image data and another containing the intoxication level label, which was split into training and testing data. Two machine learning algorithms, one logistic regression model and one decision tree model were trained on the dataset to successfully label the intoxication level of a person based on the image data.

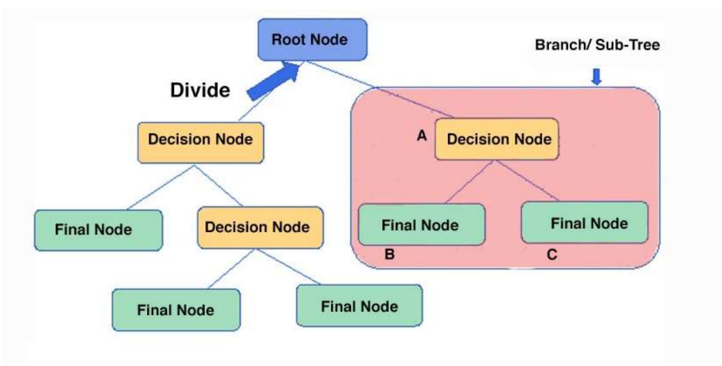
Classification using Logistic Regression

The logistic regression model predicts whether or not an individual is sober or drunk by the similarity in facial features to image data given for each of the intoxication levels. Depending on the number of features present the model can detect the level of sobriety of the individual. As such the dataset can be fit into a logistic regression curve based on the individual's identified facial features and level of intoxication to accurately label their condition. The model uses the equation below to predict the probability of subject's intoxication status.

$$P = \frac{e^{a+bX}}{1+e^{a+bX}}$$

P in this equation represents the probability, a and b are the parameters of the model, and X is the independent variable (facial data features). Logistic regression essentially calculates the probability of an item being either 1 or 0 (true/false), and e = 2.718 or the base of a natural logarithm. Since the function is logistic, we are calculating the logarithm of the odds, or P/1-P. P is equation the proportions of 1s in the sample, which is also the same as the probability of a 1 occurring within the sample. Since the probability of the 1s and 0s occurring must equal 1 when added, we can derive the odds function as the probability of a 0 occurring is equal to 1-P. When we are using the logistic regression model to classify data, the percent probability can be used to sort the data as either a 1 or a 0. It should be noted that since we are using logistic regression, X is not directly related to P. In our case, the logistic regression model can sort each image into one of the 4 classes (sober, drunk 1, drunk 2, etc) by using this equation to calculate the probability of the image either being in that class or not.

Classification using Decision Tree Classifier



The decision tree model predicts whether or not an individual is sober or drunk by evaluating the image data for certain existent tells or features (decision node), then either makes a decision on the classification of the image (final node) or makes another decision based on another feature (decision node). Each split of the decision tree is called a branch. Additionally, depending on the presence of certain facial features/expressions during a decision node, the model can classify the level of sobriety of the subject and sort it into one of the 4 classes (sober, drunk1, drunk2, etc) (final node). The accuracy of each of the branches of the decision trees was scored as seen in Figures 2 and 3. While experimenting with a decision tree model with 9 vs 5 branches to search for evidence of overfitting, as the dataset is quite small, and a more complicated algorithm could cause higher accuracy on training data but lower accuracy on testing data, accuracy score results for both training and testing data pointed to the absence of overfitting, and indeed the decision tree algorithm with more depth returned higher accuracy score for both training and test data.

Results

```
[ ] from sklearn.metrics import accuracy_score
    from sklearn.linear_model import LogisticRegression
    lr_model = LogisticRegression()
    lr_model.fit(X_train, y_train)
    y_pred = lr_model.predict(X_test)
    print(accuracy_score(y_test, y_pred))

    0.9622641509433962
```

Figure 1. Logistic regression model accuracy scores

```
cvs = cross_val_score(clf, X_train, y_train, cv=9)
cvs1 = cross_val_score(clf, X_test, y_test, cv=9)
print(cvs)
print(cvs1)

[1.         0.88888889 1.         0.94444444 0.94444444 1.
 0.94117647 0.82352941 1.        ]
[1.         0.83333333 1.         1.         1.         1.
 0.83333333 1.         1.        ]

avg_cvs = sum(cvs)/len(cvs)
print(avg_cvs)
avg_cvs1 = sum(cvs1)/len(cvs1)
print(avg_cvs1)

0.9491648511256354
0.9629629629629631
```

Figure 2. Decision tree classifier model accuracy score with 9 branches

```
[13] from sklearn.tree import DecisionTreeClassifier
     from sklearn.model_selection import cross_val_score
     clf = DecisionTreeClassifier(random_state=0)
     cvs = cross_val_score(clf, X_train, y_train, cv=5)
     cvs1 = cross_val_score(clf, X_test, y_test, cv=5)
     print(cvs)
     print(cvs1)

     [0.9375     0.96875    0.9375     0.875      0.83870968]
     [0.90909091 1.         0.90909091 1.         0.9       ]

     avg_cvs = sum(cvs)/len(cvs)
     print(avg_cvs)
     avg_cvs1 = sum(cvs1)/len(cvs1)
     print(avg_cvs1)

     0.911491935483871
     0.9436363636363637
```

Figure 3. Decision tree classifier model accuracy score with 5 branches

As theorized, using machine learning models for image classification allowed an automated approach to successfully detect intoxication with high accuracy scores. Both logistic regression and the decision tree algorithms had roughly the same accuracy score of over 96%. Because the dataset is so small, this model could experience issues due to overfitting. To increase the accuracy score and solve this issue, the number of cross-validation (cv) folds were changed. Due to a smaller dataset, more folds should have returned with less overfitting, which is seen from the model with 9 cv folds returned a higher accuracy score than a model with fewer folds. This shows that with a larger number of cross-validation folds the decision tree model produces better results since the training data is larger and because the model is being trained more times to reduce errors, meaning the model can more accurately predict intoxication and returns a higher accuracy score. A bigger dataset with more variety in facial features, gender, race, and other identifying features would allow the model to produce better results as well and reduce overfitting because the larger number of subjects would enable the model to be more accurate as the dataset would be more reflective of real-life subjects.

Limitations

The data used identifies sobriety based on facial cues; however, the dataset is limited, and the algorithms need to be tested on larger datasets as facial cues are not universal and may manifest differently in different people. While the dataset used for the algorithm contained somewhat diverse images, it was only 212 samples large, so a bigger dataset might be necessary to create a more practical model for real-world application. Both the decision tree and logistic regression using only 4 images per person for the image data, instead of video data, which could lead specifically to the decision tree algorithm struggling with comparing larger data sets with changing facial expressions and creating non-existent patterns, leading to false positives and higher inaccuracy.

Future Work

The algorithm only labeled images based on 4 classifications of sobriety, and future work could try and have the model predict the actual Blood Alcohol Content (BAC) of the individual. Future improvements could also consist of testing the algorithm on larger datasets with a greater variety of facial cues, as well as testing different machine learning methods such as KNNs or CNNs on these datasets. Hardware and software can be built around these AI models, such as hardware like cameras installed into a car that detect intoxication as a driver steps into the vehicle and starts the engine, which will feed the image data to the algorithm to process. This can also be combined with handheld devices on police officers, allowing them to analyze a person's facial cues for intoxication if they did need to pull the subject over.

Conclusion

The research conducted proves that image data of individual facial features can be analyzed by AI models such as Logistic regression and decision tree and to accurately label not only on whether the subject is drunk or sober but also on 4 levels of intoxication. Using these AI models, it is possible to build a completely automated sobriety test, that requires no physical cooperation from an individual or presence from a tester such as a police officer but returns a much higher accuracy score of 96% when predicting intoxication unlike a human-administered test which has a high margin of error and only 50-75% accuracy. Such a model can also be improved upon by using new datasets that relate facial features with Blood Alcohol Content (BAC) to predict the exact percentage BAC of an individual, allowing for more precise detection of their level of intoxication. Using AI to automate the process of sobriety tests is revolutionary in increasing the accuracy of the tests, cutting down the workload and time for police officers, and as hardware applications can work both with and without the presence of a police officer, this solution can be scalable to the massive number of drivers worldwide. Automated sobriety tests are already a heavy improvement to manual sobriety tests, and with many potential improvements, represent the future of safety on roads and the elimination of the deadly issue of drunk driving. Future work such as hardware or more diverse datasets can also be created to complement the model and allow for more practical, real-time applications.

Acknowledgements

This research paper was created under the Inspirit AI 1-on-1 AI project mentorship program under the mentorship of Linda Banh (lbanh@alumni.stanford.edu). The dataset used is compiled from a published GitHub link of SoBr, an intoxication detection program, cited below in references.

References

"Drunk driving," *NHTSA*. [Online]. Available: https://www.nhtsa.gov/risky-driving/drunk-driving#:~:text=Every%20day%2C%20about%2032%20people,one%20person%20every%2045%20minutes. [Accessed: 13-Jan-2023].

J. Chang, "JASONCHANG0/SOBR," *GitHub*. [Online]. Available: https://github.com/jasonchang0/SoBr. [Accessed: 13-Jan-2023].

M. Skinner, "Are field sobriety tests accurate? - west chester criminal defense attorney: Skinner Law Firm," *West Chester Criminal Defense Attorney | Skinner Law Firm*, 31-Mar-2022. [Online]. Available: https://www.skinnerlawfirm.net/blog/are-field-sobriety-tests-accurate/. [Accessed: 13-Jan-2023].

"Are field sobriety tests accurate?" *Gorelick Law*. [Online]. Available: https://www.gorelick-law.com/are-field-sobriety-tests-accurate. [Accessed: 13-Jan-2023].

FieldSobrietyTests.org, "Accuracy of Field Sobriety Tests," *Accuracy of field sobriety tests*. [Online]. Available: http://www.fieldsobrietytests.org/accuracyoffieldsobrietytests.html. [Accessed: 13-Jan-2023].

Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.

"Decision tree algorithm in Machine Learning - Javatpoint," *www.javatpoint.com*. [Online]. Available: https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm. [Accessed: 25-Feb-2023].

M. Brannick, "Logistic Regression," *Logistic regression*. [Online]. Available: http://faculty.cas.usf.edu/mbrannick/regression/Logistic.html. [Accessed: 25-Feb-2023].