Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 27 (2024), No. 4, pp. 1345–1355
DOI : 10.47974/JDMSC-1989

Teachable machine : A web based machine learning tool for user voice biometric authentication system

Vinod Chhipa *
Rakesh Poonia †
Department of Computer Application
Engineering College Bikaner
Bikaner
Rajasthan
India

Abstract

Now a days field of Artificial Intelligence and Machine Learning growing faster in this digital era. Teachable Machine is a web based; online open source Machine Learning classification model tool provides GUI (Graphical User Interface) and user friendly environment for end users. Teachable Machine, a web-based artificial intelligence tool service, was launched by Google. for teachers, students, data scientists and researchers without technical expertise in coding. Teachable Machine is an ML and DL model development tool without using any programming language. Teachable machine basically provides model training such as image, audio and pose identification. In this paper we are training user voice samples to recognize user voice to authenticate authorized user using his/her voice. Teachable Machine gets background noise and other multiple class audio samples to train our model. Teachable Machine can be utilized in various online transaction based application to access or complete transactions. User voice can be recognized in real time systems. It uses tenserflow.js library for training voice model with less response time and better accuracy rate as compare to other existing techniques. The customized voice model can be export by user, to utilize model in object detection. Exported model can be integrated with other programming languages such as Android, JavaScript and P5js. It helps users to develop AI and ML based applications for voice authentications as well as voice recognition. It can easily and quickly classify the datasets given to it.

Subject Classification: *68T10, 68T05.*

Keywords: *Machine learning, Voice, Authentication, Biometric, Tool, Deep learning, Artificial intelligence.*

* *E-mail:* **vinodchhipa55@gmail.com (Corresponding Author)**
† *E-mail:* **rakesh.ecb98@gmail.com**

© Taru Publications

1. Introduction

Google was motivated by Wekinator (Mac App) founded by Rebecca Fiebrink in 2009. From Past few years it was noticed that the field of Artificial Intelligence important for students and researchers. Voice Model can be created free and easily with the help of teachable machine web tool [1]. The voice biometric authentication model that can use unique voice characteristics of individuals. It is using a tenserflow libraries to train and test models. First it train the voice samples as well as background noise of approx. 20 seconds. Then it stores voice samples of multiple classes, these sampled are stored in database. The proposed system getting 0 to 9 digits as voice samples. Each digit required to create a separate class. These digits are labeled by digit value for class name. For the better accuracy rate we have to input voice sample through headphone or microphone. It can be used to recognize and authenticate real time audio. It doesn't require any account for users for creating model [2].

2. Artificial Intelligence Methodology

Artificial Intelligence falls under the umbrella of Science. and develops machines to find out better solutions of complex and critical problems in human like fashion. Artificial Intelligence based technologies are the future of upcoming next generation. It is uses in various fields like engineering, medical, finance, research and education etc. AI is a trained machine, which behave or work like a human brain in decision making and critical thinking Ex Alexa. AI based technology is present and future generation of computers which are more helpful in innovations and discoveries such as detecting disease and designing driverless cars [3].

Machine Learning : ML is a subset of AI Technology. Machine Learning plays an important role to achieve Artificial Intelligence Ex. Spam Detection [4].

3. Existing System

Existing voice biometric authentication technique uses MFCC (Mel Frequency Cepstral Coefficients) and LPCC (Linder Predictive Cepstral Coefficients) for feature extraction. training voice model. Gaussian Mixture Model(GMM) model is used for training voice model.

4. Machine Learning Voice Model Framework

We can design a audio model using machine learning algorithm. The Machine Learning voice model framework is illustrated in Figure 1.

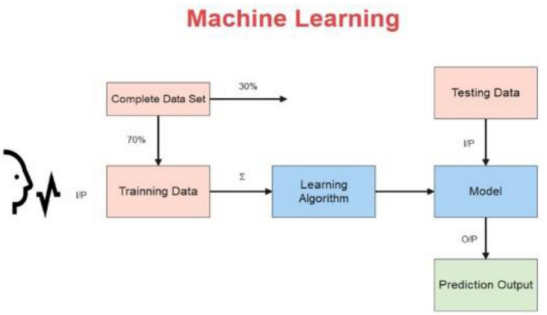


Figure 1

Machine Learning voice model framework

Machine Learning v/s Deep Learning v/s CNN

In below given Figure 2 Both the techniques are used in object detection but Machine Learning framework contains separate steps for feature extraction and classification on the other hand Deep Learning process having middle step which is combination of both Feature extraction and classification of object[5][6]. In CNN there are five main stages to train a model but in Machine Learning model can be trained in three main stages[7].

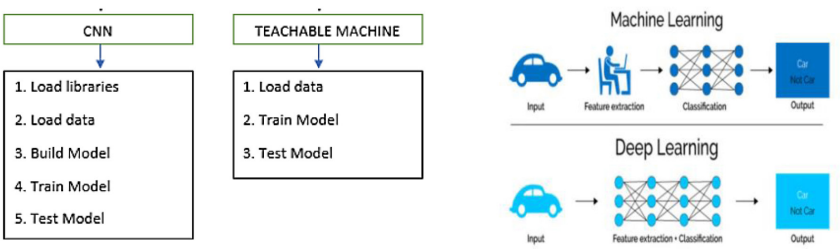


Figure 2

ML v/s CNN v/s Deep Learning

Note: (Source URL : https://www.softwaretestinghelp.com/data-mining-vs-machine-learning-vs-ai/)

5. Research Methodology

The proposed Teachable Machine Learning Audio Model is divided into three main categories

A. Creating New Project
B. Input Voice Data
C. Training Voice Model
D. Save and Export Voice Model
E. Download Audio Model
F. Integrate Voice Model on Web Page
G. System Performance Analysis

A. Creating New Project: To create new project in teachable machine open URL (https://teachablemachine.withgoogle.com/) in web browser. On the left top menu click on new project option, it shows three types of project Image project, Audio project and Pose project. We will start with an Audio Project as shown in Figure 3.

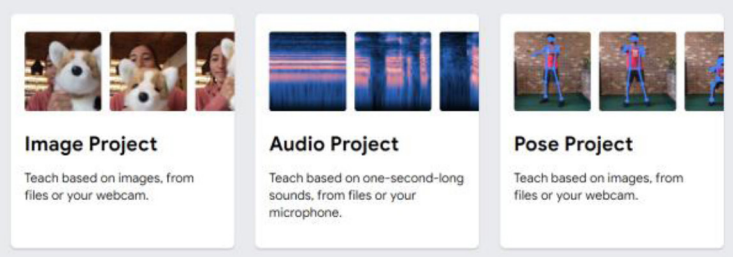


Figure 3

Machine Learning model categories

Note : (Source URL : https://teachablemachine.withgoogle.com/)

B. Input Voice Data: In this stage user speaks digits ie 0,1,2,3,4,5,6,7,8,9 each class for each digits. An audio model required at least 20 seconds of background noise. We cannot delete or disable this background class. Other than background class remaining classes required at least 8 second of voice sample. Click on record button to record user voice samples. Once the voice is recorded click on extract sample button to extract voice samples. Click on Add Class button to add new class in the project. Enable your microphone to access on this website. Below Figure 4 shows voice samples.

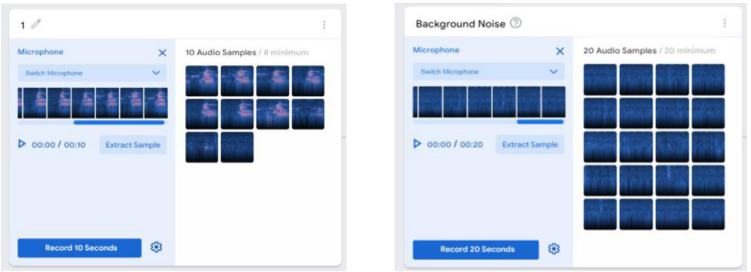


Figure 4

Voice samples

C. **Training Voice Model:** Training samples (85.0% of the samples) are used to train the model how to correctly classify new samples into the classes you have made. Test Samples (15.0% of the samples) are never used to train the model, so after the model has been trained on the training samples they are used to check the performance of model. Now click on train model button. This may take a while to complete the training process depend on epochs rate. By default epochs rate is 50 in audio model. If your epochs is 60 , it means the model you training will work throughout the entire training dataset 60 times. Teachable Machine mainly uses Tensorflow.js, a library for machine learning in JavaScript, to train and run the models you make in your web browser. The value epochs defines how many times the dataset iterate through the trained model[8]. The total training time depends on number of classes used in model. The permutation and combinational technique is used at the time of training phase of model. Below Figure 5 shows various model training options.

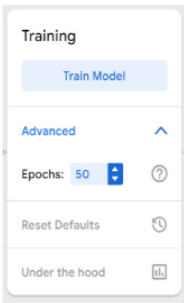


Figure 5

Train model options

D. Save & Export Voice Model: We can save and export our voice model to user in other projects. Following are the ways to save our project

(a) Save Audio Project on Google Drive : We can save our audio project on Google Drive account which containing samples. When we save project on drive it creates .zip file format. Later on we can access this file from drive

(b) Download Audio Samples : We can download the class samples for further access the data of same class in other projects.

(c) Download Audio Model : We can download audio model to use that model in other applications and noting is saved on browser or any web servers [4]. Figure 6 presents the sample code to access model.

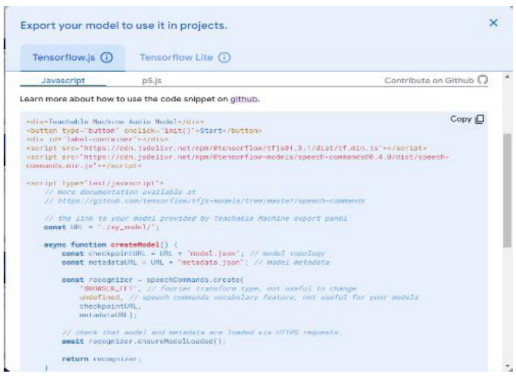


Figure 6

JavaScript code to access model

E. Download Audio Model : When we download audio model it download into zip format, which contains three main files:

- **weights.bin :** This bin file contains weights of trained model[4].
- **model.json :** This JSON file contains model topology[4].
- **metadata.json :** This file is a JSON file which contains used libraries, version, metadata, list of used label names, model number and image size to train the model[4].

F. Integrate Voice Model on Web Page : We can export our model file as tenserflow.js and can host on teachable machine server, so we can access the hosted model through shareable URL in websites or applications[4]. Other option to download in Tenwerflow (Javascript

& P5.js) or Tenserflow Lite (Android). Although we can export our model to use in other platforms and programming environments.

G. System Performance Analysis : Here are few graphs that can help you understand how well your model is working

a) Accuracy : Precision refers to the proportion of correct classifications made by a model during its training process. For instance, if a model correctly classifies 80 out of 100 samples, the precision would be calculated as 80/100 = 0.8. A perfect prediction by the model would result in a precision value of 1, while any other prediction would yield a value less than 1. [1] [9][10]. Below given Figure 7 shows the accuracy graph.

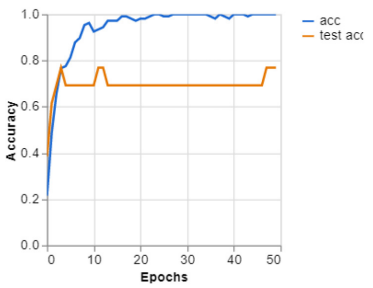


Figure 7

Show accuracy graph

b) Loss : Loss is a metric used to assess the effectiveness of a model in accurately predicting the correct classification for a given set of samples. If the model's predictions are flawless, the loss will be zero (0); otherwise, the loss will be greater than zero (>0)[1] [9][10]. Below given Figure 8 shows the loss graph.

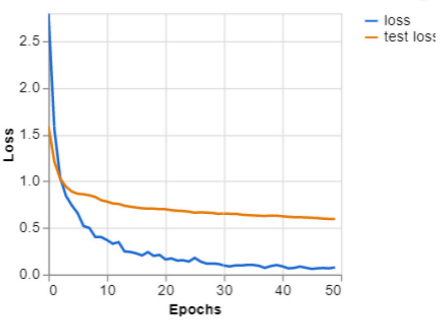


Figure 8

Show loss graph

To gain a better understanding of this concept, consider two models: A and B. Model A correctly predicts the classification for a sample but with only 60% confidence in that prediction. On the other hand, Model B also correctly predicts the classification for the same sample but with 90% confidence. Both models have the same level of accuracy, but Model B exhibits a lower loss value. [10].

c) **Underfit** : When a model fails to accurately classify data, it is considered underfit, signifying its inability to capture the intricate complexities inherent in the training samples. [10].

d) **Overfit** : The phenomenon of over fitting arises when a model becomes too fixated on accurately classifying training samples, resulting in its inability to make accurate classifications on test samples. [10].

H. Results : The below given Figure 9 shows the result found when user speaking specific OTP digit in two different classes.

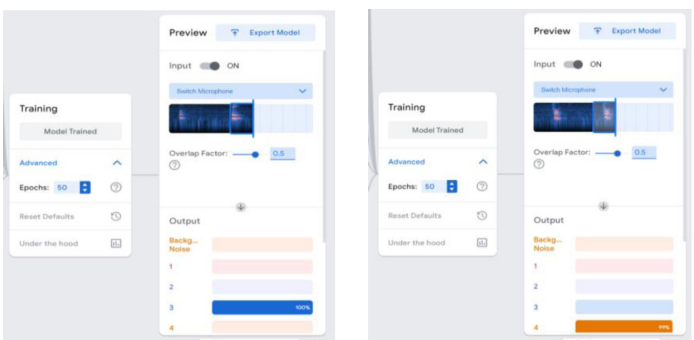


Figure 9

Class accuracy view real time

6. Research Flow Chart System

The below given Figure 10 displays the actual flow of complete system in various symbols. The flow chart helps programmer to check the accuracy and error in program. In this work the user will login with their login credentials ie. Username and Password. If the login success the transaction can be made if the user is registered or already enrolled on the portal with voice samples[7].

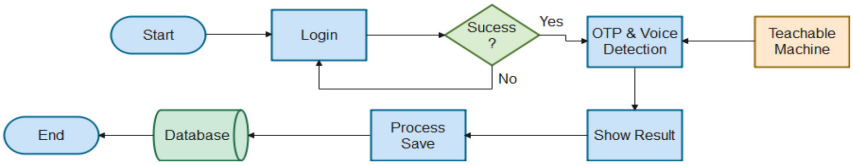


Figure 10

Flow chart

7. Discussion and Conclusion

The proposed voice authentication machine learning tool can be utilized to authenticate user voice samples based on various voice features. The system is validating user voice based on randomly generated OTP digits. In this paper we have created separate classes for digits to recognized voice as well as digit value from user. This system is based on text dependent voice authentication system. The system is first complete the enrollment of user on the basis of voice samples. User will get 6 digit randomly generated OTP on registered mobile that can be validate through user voice. The voice accuracy may be affected due to environmental condition of health issues.

References

[1] Michelle Carney, Barron Webster. Irene Alvarado and Kyle Phillips," Teachable Machine: Approachable Web-Based Tool for Exploring Machine Learning Classification"

[2] Hitesh kumar Babubhai Vora, Hardik Anilbhai Mirani, Vraj Bhatt, "Traditional Machine Learning and No-Code Machine Learning with its Features and Application", *International Journal of Trend in Scientific Research and Development (IJTSRD)*, Vol 5, Issue 2, pp. 29-32, Jan-Feb (2021).

[3] Nataliia V. Valko, Tatiana L. Goncharenko, Nataliya O. Kushnir and Viacheslav V. Osadchyi," Cloud technologies for basics of artificial intelligence study in school", Vol 3085, pp. 170-183 (2021).

[4] Fabiano Pereira de Oliveira, Christiane Gresse von Wangenheim, Jean C. R. Hauck , "TMIC: App Inventor Extension for the Deployment of Image Classification Models Exported from Teachable Machine"

[5] P. Yogendra Prasad, Dr. Dumpa Prasad, Dr.D Naga Malleswari," Implementation of Machine Learning Based Google Teachable Machine in Early Childhood Education", *International Journal of Early Childhood Special Education (INT-JECSE)*, Vol 14, pp. 4132-4148 (March 2022).

[6] M. Aqil, F. Tabri, N. N. Andayani, S. Panikkai, "Integration of smartphone technology for maize recognition", IOP Conf. Series: Earth and Environmental Science, pp. 1-6 (2021), doi: 10.1088/1755-1315/911/1/012037.

[7] Edwin Ariesto Umbu Malahina1, Ryan Peterzon Hadjon, Franki Yusuf Bisilisin," Teachable Machine: Real-Time Attendance of Students Based on Open Source System", *The IJICS (International Journal of Informatics and Computer Science)*, Vol 6, No 3, pp. 140–146, November (2022).

[8] Hyunja Jeong," Feasibility Study of Google's Teachable Machine in Diagnosis of Tooth-Marked Tongue", Vol 20, No 4, pp. 206-212 (2020), doi: https://doi.org/10.17135/jdhs.2020.20.4.206.

[9] Ashwini Kalshetty, Sutapa Rakshit, "Use case of no code machine learning tools for medical image classification", Research Suare, pp. 1-10, doi: https://doi.org/10.21203/rs.3.rs-498907/v1.

[10] Kumar, Ankit, et al. "An improved quantum key distribution protocol for verification." *Journal of Discrete Mathematical Sciences and Cryptography* 22.4 : 491-498 (2019).

[11] Kumar, Ankit, et al. "An enhanced quantum key distribution protocol for security authentication." *Journal of Discrete Mathematical Sciences and Cryptography* 22.4 : 499-507 (2019).

[12] J. A. Jenoshan, "Improved voice authentication system for hands-free computer interaction" , pp. 1-6, July (2022).

[13] https://levity.ai/blog/difference-machine-learning-deep-learning

[14] Indrajit De, Ambuj Kumar Agarwal, Bharat Bhushan, Aarti Kalnawat, Piyush Mathurkar & Amit Garg ,"Implementing multi-factor authentication (MFA) for robust network access security", *Journal of Discrete Mathematical Sciences and Cryptography*, Vol 27, No 2-B, pp. 821–832 (2024), doi: 10.47974/JDMSC-1958.

[15] Suman Devi & Avadhesh Kumar, " An artificial intelligence based authentication mechanism for wireless sensor networks using blockchain", *Journal of Information and Optimization Sciences,* Vol 45, No 2, pp. 581–594 (2024), doi: 10.47974/JIOS-1596.

[16] Sandeep Kumar Sharma, Anil Kumar, Rashmi Ashtagi & Rekha Jain," OCA: An intelligent model for improving security breach of biometric based authentication systems", *Journal of Discrete Mathematical Sciences and Cryptography*, Vol. 26, No 5, pp. 1415–1425 (2023), doi: 10.47974/JDMSC-1765.

[17] Priti Golar & Rika Sharma, "A secured image based three factor user authentication system", *Journal of Statistics and Management Systems*, Vol 27, No 2, pp. 295–302 (2024), doi: 10.47974/JSMS-1255.