# Exploring the potential of an audio application for teaching AI-based classification methods to a wider audience

## Explorando el potencial de una aplicación de audio para enseñar métodos de clasificación basados en IA a una audiencia más amplia

Gabriel Coto-Fernández[1], Marvin Coto-Jiménez[2]

Coto-Fernández, G.; Coto-Jiménez, M. Exploring the potential of an audio application for teaching ai-based classification methods to a wider audience. *Tecnología en Marcha*. Vol. 35, special issue. November, 2022. International Work Conference on Bioinspired Intelligence. Pág. 33-41.

https://doi.org/10.18845/tm.v35i8.6444

1 Saint Joseph High School. Costa Rica
2 PRIS-Lab, Electrical Engineering Department. University of Costa Rica. Costa Rica. E-mail: marvin.coto@ucr.ac.cr
https://orcid.org/0000-0002-6833-9938

## Keywords

Audio; artificial intelligence; classification; teaching.

## Abstract

Knowledge about artificial intelligence (AI) is becoming increasingly important for many careers, especially those based in science and engineering. Besides formal education, the impact of AI on society lead to consider educational projects for teaching the fundamental concepts of AI at wider audiences, including high school levels. This can help more general audiences to better understand how AI works, with the hope that also parents and educators can help students develop a healthy appreciation for implications and limitations, along with an appropriate relationship and deeper interest on it. In this paper, we present a pilot project for teaching an AI-based classification method that is empirically evaluated with real data of a real problem, which can be understood and tackled with basic mathematical tools and activities suitable for high school students. With this proposal, we aim to show how audio and speech applications can inform a wider audience about advances in AI, its characteristics, and its future impact on society. Results and lessons learned from this project can form the basis for further projects using different tools and data, according to students' interests and initiative.

## Palabras clave

Audio; inteligencia artificial; clasificación; docencia.

## Resumen

El conocimiento sobre inteligencia artificial (IA) se está volviendo cada vez más importante para muchas carreras, especialmente las basadas en la ciencia y la ingeniería. Además de la educación formal, el impacto de la IA en la sociedad lleva a considerar proyectos educativos para enseñar los conceptos fundamentales de la IA a un público más amplio, incluidos los niveles de secundaria. Esto puede ayudar a un público más general a comprender mejor cómo funciona la IA, con la esperanza de que también los padres y educadores puedan ayudar a los estudiantes a desarrollar una apreciación saludable de las implicaciones y limitaciones, junto con una relación adecuada y un interés más profundo en ella. En este artículo presentamos un proyecto piloto para la enseñanza de un método de clasificación basado en IA que se evalúa empíricamente con datos reales de un problema real, que se puede entender y abordar con herramientas y actividades matemáticas básicas adecuadas para estudiantes de secundaria. Con esta propuesta, nuestro objetivo es mostrar cómo las aplicaciones de audio y voz pueden informar a una audiencia más amplia sobre los avances en IA, sus características y su impacto futuro en la sociedad. Los resultados y las lecciones aprendidas de este proyecto pueden formar la base para otros proyectos utilizando diferentes herramientas y datos, de acuerdo con los intereses y la iniciativa de los estudiantes.

## Introduction

Artificial intelligence (AI) has become part of our everyday life: its existence, recent applications and potential impact are constantly discussed in mass media. Children now interact with tablets, toys and video games that implement AI algorithms, all of which contribute to their understanding of these devices. In this way, their knowledge about AI becomes more deeper and more nuanced [1]. Despite the importance of AI, few people know about the technology and its requirements [2]. This is one reason it is important to familiarize young people in their school years with the technical details and the underlying concepts of AI.

The concept of AI literacy has been well discussed in popular media outlets. It is expected that literacy in AI will become a major issue in the future. With

AI literacy, students can receive a solid preparation for subsequent studies at the university level and in their future careers.

According to reported studies, children can be too trusting of technological devices and can be influenced by them [3]. Better understanding of the principles and implications of AI in high schools and by parents and educators can help them teach children to gain an appreciation for AI's abilities and limitations, as well as build a proper relationship with it.

Some of the limitations that exist for the inclusion of teaching AI concepts in high schools and other wider audiences have been due to the complexity of the mathematical concept of algorithms, the computing and programming background needed to understand AI concepts, the many difficulties in the implementation of projects with a proper understanding of the implications and the environment where real applications should be developed.

Given those difficulties, it becomes very important to carefully select situations and teaching strategies that motivate new audiences and provide them an adequate level of understanding and appreciation for AI. Based on previous experience, some themes have been shown as key to teaching AI to children and wider audiences [4, 5]: 1) AI systems are based on knowledge acquired from hu- man beings through examples. 2) AI systems do not know everything and make mistakes. 3) AI systems are corrected and improved by human beings.

Considering these key themes and the importance of real-world applications that can be understood and addressed by younger learners and general audiences, we propose the exploration of audio and speech in teaching strategies on AI concepts, especially those based on classification methods. For this purpose, we selected a real-world problem that arose from the need to help people with disabilities, and the opportunity to provide them with alternative and augmentative communication systems [6].

This problem led us to design an educational experience to explore the possibilities for teaching basic classification methods, including the complete process of building a data-set, presenting and explaining feature extraction, and visually applying the classification concepts.

## Related Work

Several researchers have reported successful strategies for teaching AI concepts to children and other general audiences. For example, in [7], an approach for an agriculturally-based AI challenge that helped students learn the process of creating machine learning models was presented. The study found that machine learning can be used as a tool to conduct interdisciplinary education at the middle school level.

With a historical perspective, the authors in [8] discuss the Turing Test as an educational activity for undergraduate students. Some other long-term courses also introduce basic concepts, vocabulary and history [9] at appropriate levels, including topics presented in important books commonly used for graduate students, but adapted for younger students [2].

Associations such as the Association for the Advancement of Artificial Intelligence (AAAI) and the Computer Science Teachers Association (CSTA) in United States of America have formed working groups to develop national guidelines for teaching AI, machine learning and robotics to K-12 students. One of the main objectives of the projects is to invite the AI research community to reflect their ideas and developments of AI in a way that every K-12 student should know, and how communicate with the general public about advances in AI and their future impact on society [10].

In Latin America, the idea of presenting AI concepts to a wider audience have been discussed in [4], with the particular aim of an exhibition in a science museum.

All this initiatives apply different strategies to particular age groups, and focus on concepts of AI that suit some of the vast possibilities of education. Our approach take advantage of these experiences in term of do's and don'ts, situations to avoid and to consider, but with a new application based on audio and speech processing for helping people with disabilities.

The rest of this paper is organized in the following way: Section II presents the motivation that lead the proposal, in terms of developing communication solutions to people with disabilities. Section III shows the experimental results of the experience, Section IV the results and finally Section V present the conclusions.

## Experimental and teaching framework

### Motivation

The motivation for the application of speech in this proposal arose from the fact that speech is the most common means of communication between people. However, some disability conditions may limit the abilities of people of different ages, genders and others conditions to speak. Within this framework, an AI-based application of speech technologies can provide opportunities for communication, cognitive support and functional independence for many users.

For example, some persons with brain paralysis cannot express themselves verbally, but are capable of producing sounds that express emotional states or positive/negative answers to questions. A simple device that incorporates a piezo-electric sensor and sound amplifier can be built with the appropriate characteristics for its implementation by taking audible vocalizations directly from the user's throat, and recording them with a computer or embedded system.

The AI-based classification algorithm can label new sounds according to the labeled sounds in a database. In the final stage, speech synthesis can pronounce the word in a a manner that can be understood by most people.

This motivation can be presented to high school students that can also participate in the making of the simple sensor, as shown in Figure 1.

Voluntary students or participants themselves can participate in the recording of the database, consisting on isolated gesticulations to express "yes" or "no", as described in the next section.

(a) Piezo-electric sensor and amplifier

(b) Recording sessions

**Figure 1.** Illustration of the proposed device.

## Database description

Our proposal is to present the topic of AI-based classification methods with the necklace and the piezoelectric sensor describe in II.A, for the purpose of producing an alternative communication device for people with disabilities. To test this device and its ability to classify sounds with positive or negative meaning, the proposal is to record at least two volunteers producing three hundred and sixty sounds, with the following characteristics:

1. Thirty sounds indicating affirmation and thirty denial, with the head facing forward.
2. Thirty sounds indicating affirmation and thirty denial, with the head turned as far as possible to the right.
3. Thirty sounds indicating affirmation and thirty denial, with the head turned as far as possible to the left.

According to our experience, the recording sessions for this database can last about 20 minutes per volunteer, which represents a proper time for an experience suited for a wider audience. The head turns are proposed to verify the robustness of the device against involuntary movements, or the movements of continuous use in a realistic environment. During process, the feedback of the students or users can be of great benefit and can lead to discuss and record new data or to systematize the sessions.

For this work, we perform a recording session with volunteers, as a mean to properly measure the difficulties of the process and to validate the proposal. Feature extraction and results are reported to set expected values.

## Feature extraction

The process of production of speech or sounds with the sounding apparatus is complex, since it involves a series of organs, from the lungs, the larynx, the vocal cords, the nose and the tongue, among others [11].This complexity is illustrative in explaining the students the complexity of the speech production process and the subtle differences that let human beings differentiate voices and words.

This complexity also allows the introduction of how numerically characterize the sounds of the database. Whilst the parametrization of speech production is a complex subject for high school or more general audiences, the understanding of pitch, related to deep and sharp voice, or the melody produce in talking, can be shared and applied in this experience.

To maintain the necessary simplicity in the process, the classification of the sounds in the database is proposed using basic statistics of a single parameter: the fundamental frequency. This decision is based on the fact that there are numerous implementations for the extraction of the parameter, in addition to the simplicity of the process to properly implement a recognition system for the application of assistive technologies. There are also open research lines to detect it in adverse conditions [12], which is useful for real applications.

For each of the sounds in the database, the fundamental frequency and four characteristics are obtained:

- Min (minimum value of $f_0$ of the while file)
- Max (maximum value of $f_0$ of the while file)
- Mean $f_0$ value of $f_0$ of the while file)
- Median $f_0$ value of $f_0$ of the while file)

For evaluation purposes, we propose a subset of about 20 recordings (with the corresponding four statistical characteristics) to test the algorithm applied.

### Algorithm and exploration

For a proper visualization and understanding of the AI-based classification scheme, we intend to apply the parameters in pairs, for a 2D representation of each file in the plane. One of the first questions that arose in real-life experimentation with classification problems is: What features should we use? This question can motivate discussion and experimentation with students.

For example, we can choose some of the pairs of values for a manual application of an algorithm such as KNN, where only distances between points need to be measured or estimated from a planar representation. In particular, the case of 1-NN can be presented and explained with ease at the high school level, where the notions plane, points and coordinates are well known.

The following steps can be applied in an teaching experience, after the intro- duction of the problem and an explanation of the 1-NN algorithm:

1. Plot the pairs of values of each possible combination of features in a plane, with corresponding labels for the negative or positive message of the audio.
2. Ask the students or participants to manually label each of the points of the test set according to its nearest neighbour.
3. Keep a second plot of each pair of features with the real labels of the points in the test set, for comparison purposes.
4. Compute the errors manually and decide on the best pair of features.
5. Compare the results with a computer version of the algorithm, where further experimentation can be done; for example, 2-NN, 3-NN and 4-NN.

The necessity of AI in this problem can be presented in terms of the difficulties that arise when elements that are always different (such as words and human sounds) need to be classified according to preset categories.

## Results and discussion

This proposal can be useful in a teaching experience if the results represent a meaningful challenge for the participants and the error rates are close to those of a real application. For example, if the error rate is as high as 50%, the whole experience cannot be considered successful, nor will it lead to appreciation of the classification method and AI methodologies.

To verify the possibilities of the data-set to achieve good results, as well as manual labeling of the plots, in Figures 2 and 3 two pairs of features are shown, with the corresponding training and testing data, with spaces for manual labeling presented with lines.

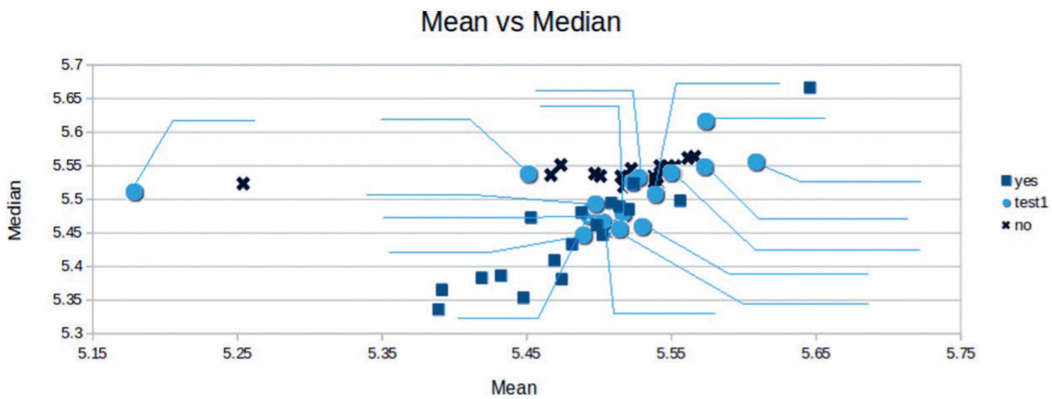


**Figure 2.** Set of points of the train and test data for Mean and Median features.

The two plots also show how different the classification method can be in terms of how points are distributed and separated in the training set. Some of the points of the test set are very close in both cases, and proximity to its nearest neighbor can foster collaborative work and discussion from participants.

Table 1 shows the results of a preliminary manual labeling and comparison with the KNN algorithm implemented in the Weka system, which can be very useful for demonstrations in teaching, due to its flexibility and user-friendly interface.

Many questions can be asked about the results shown in the plots and the table; for example: can human labeling surpass computer results? How much time does it take to prepare the files (such as the ARFF format for the Weka system) in comparison with manual labeling of results? What other features can be extracted from the dataset (such as range or quartiles)? Is it easy to find the best parameters for a classification method? Can computers make mistakes in classification or other AI-related tasks? How can these errors be minimized?

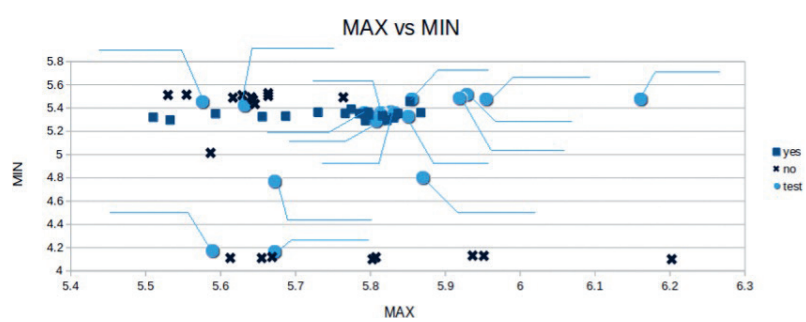


**Figure 3.** of points of the train and test data for Max and Min features.

**Table 2.** t -test results comparing ex-combatants and civilians on SCB scores.

| Features | 1-NN (Human) | 1-NN | 2-NN | 3-NN | 4-NN |
|---|---|---|---|---|---|
| Max *vs* Min | 20 | 25 | 30 | 30 | 30 |
| Mean *vs* Min | 10 | 10 | 15 | 10 | 10 |
| Mean *vs* Max | 40 | 40 | 30 | 30 | 40 |
| Mean *vs* Median | 10 | 20 | 25 | 10 | 15 |

Further exploration with the algorithms, beyond what is shown in Table 1, can be easily done with the software. For example, the application of more than two features at the same time, and even the results from other, unknown algorithms that can surpass the obtained results, can show the wide variety and the importance of more exploration and study of AI.

## Conclusions

In the present work, we explored the application of a real-life problem, related to a communication device for people with disabilities, in order to present a teaching experience of AI-based classification methods to a wider audience, who has some basic statistic knowledge and understanding of a few mathematical concepts, such as those studied in the first years of high school. For this purpose, we show a low-cost device that can be understood or even built for the participants, and the methodology to present and contextualize the problem, as well as the necessity of AI for the classification of sound recordings through the device.

Almost all of the problems of a real-life application of AI can be explored with the proposal, including recording of the data-set, parametrization, feature extraction and selection of the algorithm, all adapted and simplified for a wider audience. The results of a preliminary experience show the impact of feature selection on the results, and the meaningful experience that a manual application of the algorithm and the comparison to the results and the computer represent. For future work, we aim to develop high quality materials (such as large scale plots and presentations) to implement the proposal with a group of high school students. There is a wide margin for experimentation, from the materials and characteristics of the device, as well as the development of the database and extended conditions of use that can also be explored.

## References

[1] Williams, Randi, et al. “Popbots: Designing an artificial intelligence curriculum for early childhood education.” Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. 2019.

[2] Burgsteiner, Harald, Martin Kandlhofer, and Gerald Steinbauer. “Irobot: Teaching the basics of artificial intelligence in high schools.” Thirtieth AAAI Conference on Artificial Intelligence. 2016.

[3] Williams, Randi, et al. “My doll says it’s ok” a study of children’s conformity to a talking doll.” Proceedings of the 17th ACM Conference on Interaction Design and Children. 2018.

[4] Candello, Heloisa, et al. “30 Minutes to Introduce AI to Kids.” (2019).

[5] Kandlhofer, Martin, et al. “Artificial intelligence and computer science in education: From kindergarten to university.” 2016 IEEE Frontiers in Education Conference (FIE). IEEE, 2016.

[6] González-Salazar, Astryd, Michelle Gutiérrez-Muñoz, and Marvin Coto-Jiménez. “Enhancing Speech Recorded from a Wearable Sensor Using a Collection of Autoencoders.” Latin American High Performance Computing Conference. Springer, Cham, 2019.

[7] Sakulkueakulsuk, Bawornsak, et al. “Kids making AI: Integrating Machine Learning, Gamification, and Social Context in STEM Education.” 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE). IEEE, 2018.

[8] Torrey, Lisa, et al. "The Turing Test in the classroom." Thirtieth AAAI Conference on Artificial Intelligence. 2016.

[9] Heinze, Clint Andrew, Janet Haase, and Helen Higgins. "An action research report from a multi-year approach to teaching artificial intelligence at the k-6 level." First AAAI Symposium on Educational Advances in Artificial Intelligence. 2010.

[10] Touretzky, David, et al. "Envisioning AI for K-12: What Should Every Child Know about AI?." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. 2019.

[11] Holmes, Wendy. Speech synthesis and recognition. CRC press, 2001.

[12] Coto-Jiménez, Marvin, John Goddard-Close, and Fabiola Martínez-Licona. "Im- proving automatic speech recognition containing additive noise using deep denoising autoencoders of LSTM networks." International Conference on Speech and Computer. Springer, Cham, 2016.