An Advanced Computerized Artificial Intelligent System for Assisting Real Life Feature Extraction using AI-NOVA

Srisailapu D Vara Prasad [1]
GITAM Deemed to be University, Hyderabad, Telangana
sdvprasad554@gmail.com

J.Kavitha [2]
BVRIT HYDERABAD College of Engineering for Women, Hyderabad, Telangana
j.kavitha5555@gmail.com

G.Charles Babu[3]
Gokaraju Rangaraju Institute of Engineering and Technology(Autonomous), Hyderabad, Telangana
charlesbabu26@gmail.com

K. Venkata Murali Mohan [4]
TKR Engineering College, Hyderabad, Telangana
kvmuralimohan22@gmail.com

P.Gopala Krishna [5]
Gokaraju Rangaraju Institute of Engineering and Technology(Autonomous), Hyderabad, Telangana
dean_publicity@griet.ac.in

Abstract- **Artificial Intelligence (AI) technologies play a crucial role in human lives. Many of the tasks that humans perform are automated. Several day-to-day tasks are accomplished by using a voice command. This can be achieved by enabling a Voice Assistant. Inspired from Voice Assistants like Siri, Alexa, etc... this study has developed a new voice assistant named "NOVA", which provides all the necessary as well as useful features such as telling the date, time, weather report, a joke, playing games, searching for information on Wikipedia/Google, playing songs on YouTube, etc. Besides the above features, this study introduces a new and unique feature that helps users to perform Exploratory Data Analysis (EDA) on a given dataset and also, and the analyzed report will be saved in form of an html file.**

Keywords: Artificial intelligence, voice assistance, Exploratory Data Analysis

1. INTRODUCTION

With the technological developments, the use of Artificial Intelligence has come into the picture and people have shown quite a lot of interest in using Artificial Intelligence to make accessing things easier. Voice assistants became common and everyone is aware of them, they are used to provide us with our requirements just with a few instructions like -

a.   Tell us the time and date
b.   Set alarms and reminders
c.   Tell the weather conditions
d.   Play songs, tell jokes and even tell us about things that are not familiar with.

A voice assistant is a computer assistant that listens to specific vocal commands and responds with pertinent information or carries out particular tasks as asked by the user using speech recognition, language processing algorithms, and voice synthesis. Voice assistants can return pertinent information based on the user's spoken orders, also known as intentions, by listening for certain keywords and removing background noise.

This study has developed a voice assistant from scratch by using Python Programming Language and various Libraries provided by Python-like Pyaudio, speech recognition, gtts, and a few others. To make it stand out from the various other voice assistants like Alexa, Cortana, and Google assistant we gave our Voice assistant a unique feature, that is performing Exploratory Data Analysis.

2. OBJECTIVE

Our project aims to take the Voice Assistant to the next level by providing it with the features like performing exploratory data analysis on the data provided by the user. Computerized Artificial intelligence for Real Life Assistance-NOVA our Voice Assistant can generate reports on data that is given by the user with diverse types of graphs and plots. It helps us to summarize the main characteristics of the data given. It helps data scientists to discover patterns, spot anomalies, test a hypothesis, or check assumptions. With Python as the language used for coding the code is easy to understand and easy to execute.

3. EXISTING MODEL

The applications like SIRI on Apple devices, CORTANA on Microsoft devices, etc., are devices dedicated to providing us with basic assistance like playing songs, getting information on the internet, creating shopping lists, etc., but the existing voice assistants mentioned above are incapable of performing Exploratory Data Analysis.

Proposed System

To design a voice assistant which performs all the previously existing features and is capable of performing Exploratory Data Analysis and Report Generation. Through our project, we wish to provide everyone with their own personal assistant to improve users' privacy and its related issues.

Exploratory Data Analysis

It uses statistical graphics and other data visualization techniques to analyze data sets and highlight their key properties. Exploratory data analysis is used to discover information from the data that goes beyond formal modelling or hypothesis testing. We can even use various other types of graphs using the module named seaborn and matplotlib.

Several tools are useful for EDA, the typical graphical techniques used in EDA are:

a. Box plot
b. Histogram
c. Run chart
d. Pareto chart
e. Scatter plot

To perform Exploratory Data Analysis through Python Programming we use the library pandas_profiling and in the pandas_profiling library, we use the function profile_report to generate the EDA report on the data provided. This function helps us generate the EDA report with just a few lines of code and it is easy to code and easy to understand.

GROUPING DATA

The group by measure in pandas is a useful tool for determining how various categorical qualities affect other data variables.

ANOVA

Analysis of Variance is referred to as ANOVA. It is carried out to determine the relationship between the various categorical data groupings. As a result of the ANOVA,

we have two measures: –F-test score: which shows the variation of group mean over variation

p-valueIt demonstrates how significant the outcome is. Using the Python module scipy method name, this may be done. f_oneway()
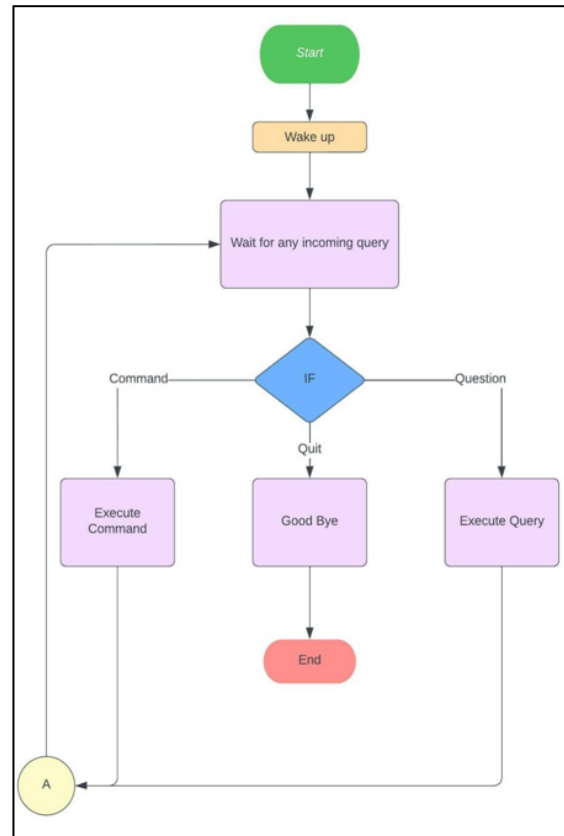
Flow Chart

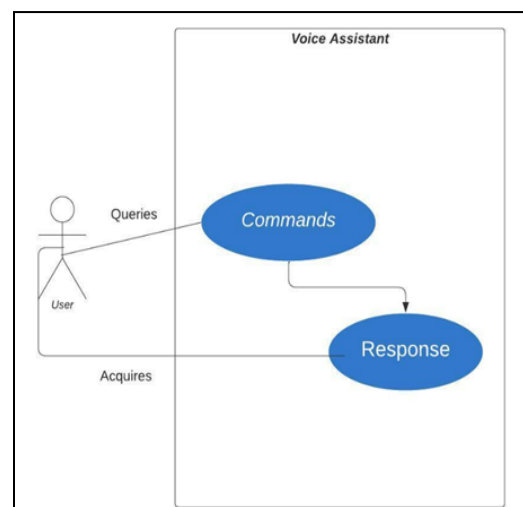**Figure 1: Flow chart describing the working of Virtual Assistant**

Figure 2: Diagram representing the Use Case

A voice assistant is a computer assistant that listens to specific vocal commands and responds with pertinent information or carries out particular tasks as asked by the user using speech recognition, language processing algorithms, and voice synthesis.

Artificial intelligence (AI) and voice recognition are used by voice assistants to accurately and quickly deliver the desired outcome. While asking a computer to set a timer might seem straightforward, the technology is fascinating.

Voice Recognition

Voice recognition functions by converting the analogue signal produced by the user's voice into a digital signal. After doing this, the computer attempts to discern the user's intent by matching the digital signal to words and phrases. For the computer to be able to accurately match the digital signal, it needs a database of already-existing words and syllables in the target language. Pattern recognition, which is the main mechanism underpinning voice recognition, involves comparing the incoming signal with this database.

Artificial Intelligence

Machines are used to mimic and duplicate human intelligence through artificial intelligence. The first time the question "Can computers think?" was posed was in Alan Turing's 1950 paper "Computing Machinery and Intelligence," which bears his name. Later, Alan Turing created the Turing Test, a technique for judging a computer's aptitude for human-like thought.

Later, four definitions of AI, including "thinking humanly/rationally" and "doing humanly/rationally," were developed. The second two deal with real behaviour, while the first two deal with rationale. Modern artificial intelligence is often seen as a computer programme created to carry out tasks that typically call for human interaction. The use of machine learning allows these systems to develop upon themselves.

Machine Learning

Machine learning is a subtype of artificial intelligence in which programs are produced automatically rather than manually by human programmers. Programmers offer the AI "patterns" to identify and learn from, as well as vast volumes of data to comb through and analyze, rather than coding the entire programme from scratch. Therefore, the AI looks for patterns within this data and uses them to improve its already existing functions rather than having explicit rules to follow.

Voice A can benefit from machine learning by giving the algorithm hours of speech from diverse dialects and accents. Machine learning tools take an input and an output and utilize those to generate the programme itself, unlike traditional programs that need input and rules to develop an output. Unsupervised learning and supervised learning are the two methods used in machine learning. In supervised learning, the model receives data that has already been partially labelled; as a result, some of the input data will already have the correct answer associated with it.

This gives the model direction as it develops the right method and classifies the remaining data. Since there are no labels on any of the data in unsupervised learning, the model must accurately identify the pattern. This helps the model identify patterns that it might not have otherwise discovered because the data is much more unpredictable, which is one of the reasons it is so helpful.

4. RESULTS

The following pictures depict the working process of our voice assistant.
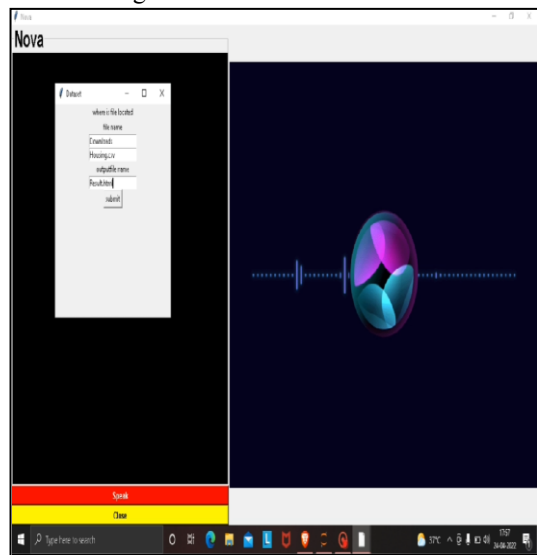
EDA through NOVA:

Figure 3: Accessing a CSV file through NOVA

Figure 4: Description of the basic Overview of data

After giving the command "perform data analysis" to NOVA it displays a frame as shown in Fig 4. after entering the name of the CSV file and location in the given frame as shown in Fig 5.
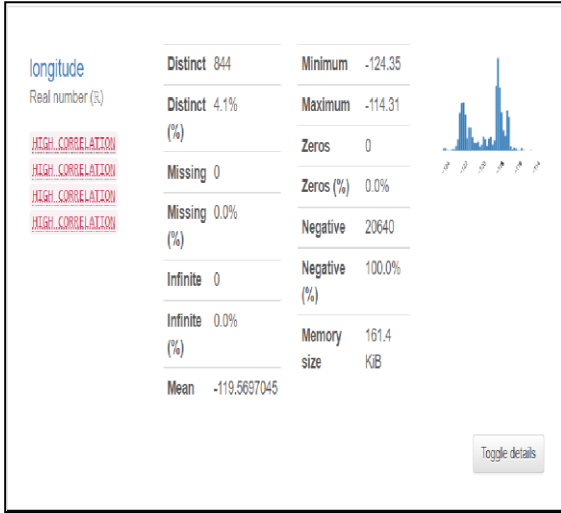

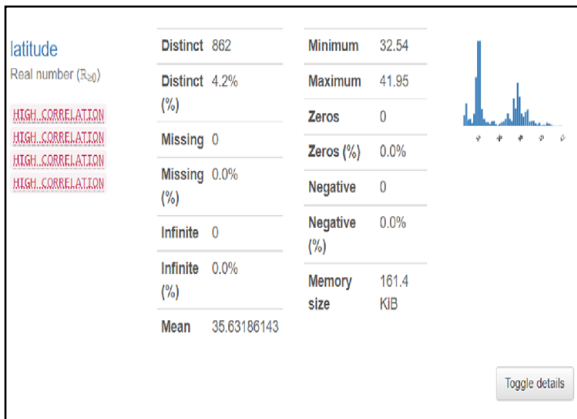**Figure 5: Description of a particular Attribute in the data-I**

Figure 6: Description of a particular Attribute in the data-II

The breakdown of the profile report given by NOVA is represented in the above figures, these figures show the description of a particular attribute, also the total number of missing values belonging to that attribute, the maximum and minimum values in that attribute, and their correlation with the other attributes.
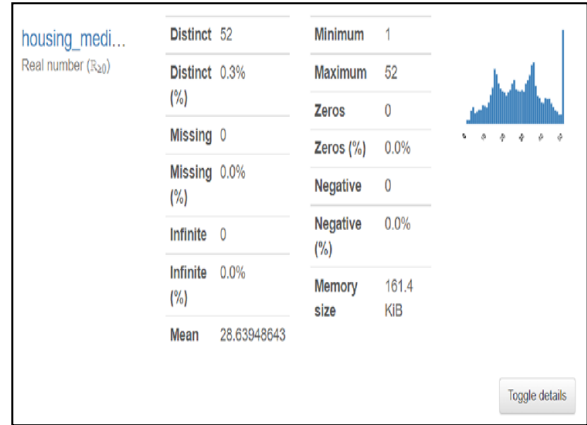
Figure 7: Description of a particular Attribute in the data-III
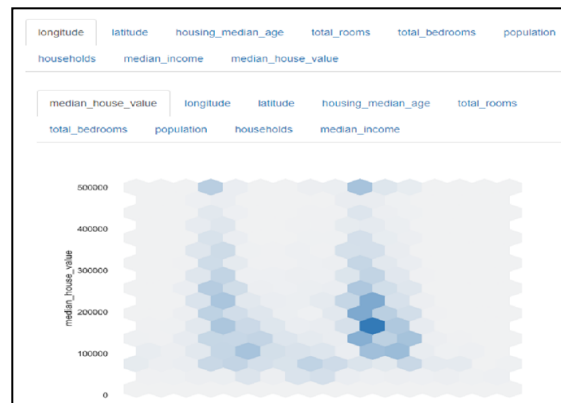

Figure 8: Hexbin graph representing the interaction between two variables

Figure 8. shows the Hexbin graph representing the interactions between any two variables of the data. Here, we can change the variables and check for different relations.
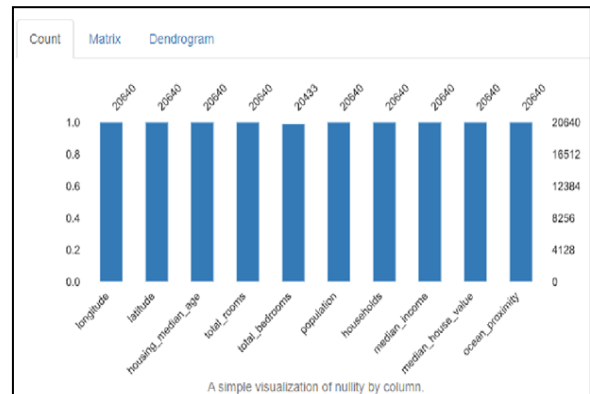

Figure 9: Bar graph representing the missing values in the data
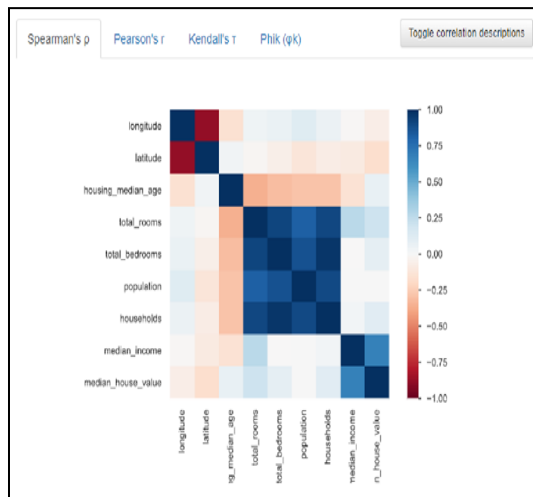
Figure 10: Heat map representing the various correlations in the data

Figure 9, shows a bar graph representing the missing values, here we can even select the option Dendrogram to get a dendrogram representing the missing values. Fig.10 shows the different correlations in the data. The Toggle description button gives the definitions of the correlations used.
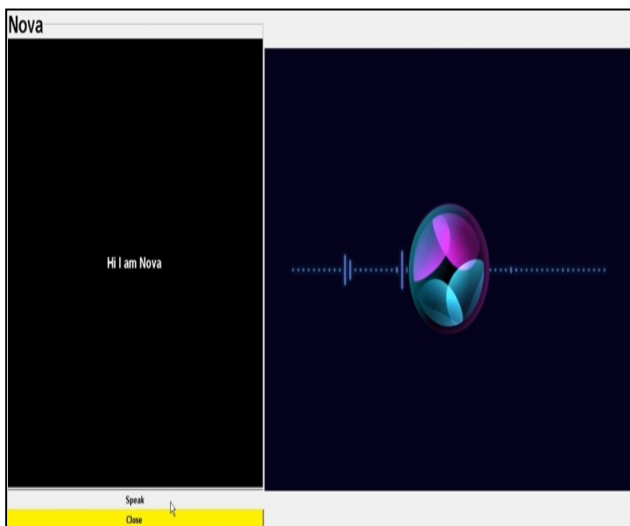
Other Features of NOVA
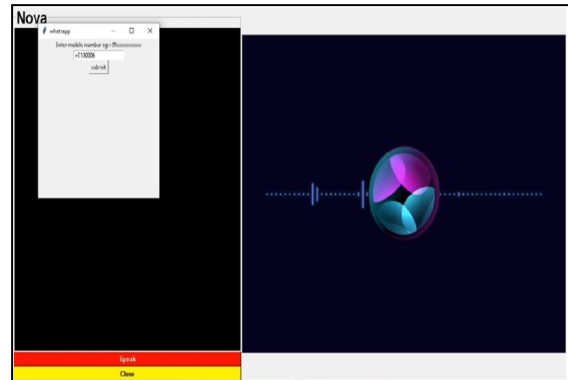

Figure 11: The Graphical User Interface of Nova

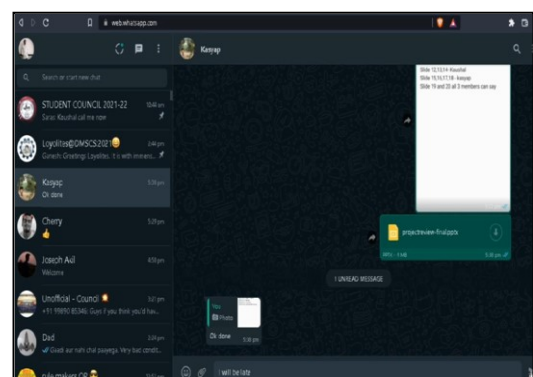**Figure 12: Sending Messages on WhatsApp – I**

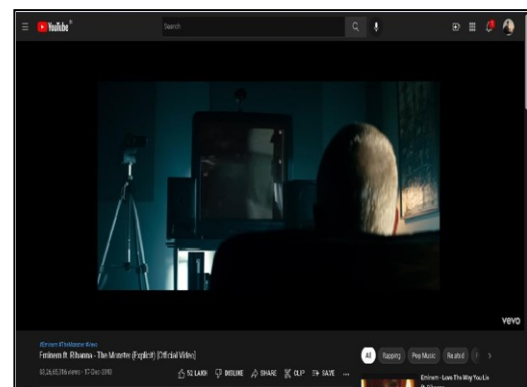

Figure 13: Sending Message on WhatsApp – II

Figure 14: Automating YouTube through Nova via voice

When the user gives the command "let's play a game" then NOVA opens a window for the game ROCK-PAPER-SCISSORS here, NOVA generates a random option so depending on the user's choice it shows a small message whether it's a win or loss as shown in the Fig 14.
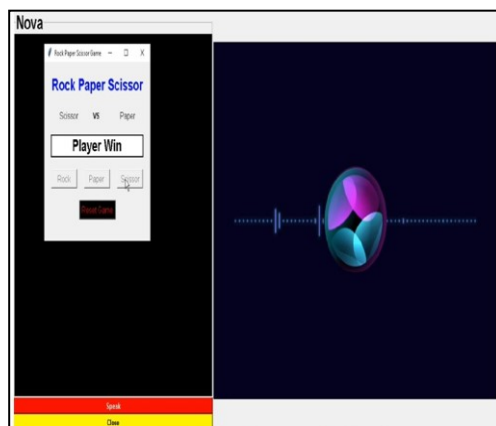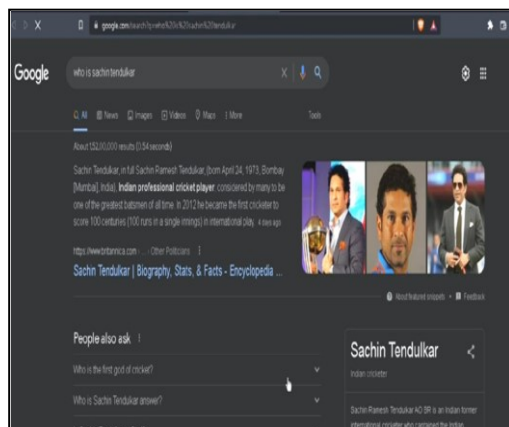
Figure 15: Playing Games with Nova – II

Figure 16: Searching for information using Nova

5. CONCLUSION

This article consists of a brief explanation of the proposed execution. The main objective of this research work is to develop Voice Assistant - Computerized AI for Real Life Assistance-NOVA a special feature which is the ability to perform exploratory data analysis on any dataset provided by the user. This research work has developed a personal voice assistant NOVA using libraries present in the Python Programming Language, through the Jupyter notebook. This is an era where the applications of Artificial Intelligence (AI) are observed in every field of work, every task performed is mostly automated. Several day-to-day tasks are achieved through a voice command. This provided us with the inspiration to create voice assistant with a special feature that is performing Exploratory Data Analysis (EDA). This research work helps to easily analyze a particular dataset with just a voice command. This project can be further expanded to android devices as it is currently only a system-based GUI

and we can also try to provide the voice assistant with access to the cloud storage as it will be convenient and can store very huge amounts of data.

REFERENCES

[1] Shaughnessy, IEEE, "Interacting with Computers by Voice: Automatic Speech Recognition and Synthesis" proceedings of the IEEE, vol. 91, no. 9, september 2003.

[2] Patrick Nguyen, Georg Heigold, Geoffrey Zweig, "Speech Recognition with Flat Direct Models", IEEE Journal of Selected Topics in Signal Processing, 2010.

[3] Mackworth (2019-2020), Python code for voice assistant: Foundations of Computational Agents- David L. Poole and Alan K. Mackworth.

[4] Nil Goksel, CanbekMehmet ,EminMutlu, "On the track of Artificial Intelligence: Learning with Intelligent Personal Assistant", proceedings of International Journal of Human Sciences, 2016.

[5] Keerthana S, Meghana H, Priyanka K, Sahana V. Rao, Ashwini B "Smart Home Using Internet of Things", proceedings of Perspectives in Communication , Embedded -systems and signal processing, 2017.

[6] Sutar Shekhar, P. Sameer, Kamad Neha, Prof. Devkate Laxman, "An Intelligent Voice Assistant Using Android Platform", IJARCSMS, ISSN: 232-7782, 2017.

[7] Rishabh Shah, Siddhant Lahoti, Prof. Lavanya. K, "An Intelligent Chatbot using Natural Language Processing", International Journal of Engineering Research , Vol.6 , pp.281-286, 2017.

[8] Luis Javier Rodríguez-Fuentes, Mikel Peñagarikano, Aparo Varona, Germán Bordel, "GTTS-EHU Systems for the Albayzin 2018 Search on Speech Evaluation", proceedings of IberSPEECH, Barcelona, Spain, 2018.

[9] Ravivanshikumar ,Sangpal,Tanvee ,Gawand,Sahil Vaykar, "JARVIS: An interpretation of AIML with integration of gT T S and Python", proceedings of the 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kanpur, 2019.

[10] Luis Javier Rodríguez-Fuentes, Mikel Peñagarikano, Aparo Varona, Germán Bordel, "GTTS-EHU Systems for the Albayzin 2018 Search on Speech Evaluation", proceedings of IberSPEECH, Barcelona, Spain, 2018McKinney, W. (2012). Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. " O'Reilly Media, Inc.".