Grenze International Journal of Engineering and Technology, Jan Issue

An Audio Player using Voice Recognition

Kabilesh J[1], Karthik S[2],Krishna Bharath S[3,] and Rajesh Kumar S[4]
[1234]National Engineering College/ Computer Science and Engineering, Kovilpatti, India
[4] Assistant Professor,[National] Engineering College/ Computer Science and Engineering, Kovilpatti, India
Email:rajeshkumar_cse@nec.edu.in
[1-3]Email: {2017-18cse15, 2017-18cse52,2017-18cse17}@nec.edu.in

***Abstract*—This project is to develop an NLP-based audio player application for windows and android users, which allows us to control the audio playerusing our voices. The person who wants to hear music and has their wireless ear pods needs a hand-free method to control their music player while doing their other chores. There are other alternatives such as Alex, Google assistants to do the same, but they need an extra physical device that needs the internet connection to run. Thus, the product will help to solve these difficulties and helps the users for their easy access.**

***Index Terms*— Application, Voice Recognition, Natural Language Processing, Artificial Intelligence.**

I. INTRODUCTION

All of uslistento music, nowadays all the operating systems have at least one default media player. For example, Windows has Windows Media Player, Microsoft Movies &TV,Android has Google Play Music for audio.
In some scenarios like living in the Apartment, we can't be able to listen to music loudly, so we need to use earphones while doing some house chores and if we need to skip the next song or increase the volume, then we need to do it physically on the device, so we need a hand free method to control their music player while There are other alternatives such as Alex, Google Assistants are there can do the same, but they need an extra physical device with an internet connection to implement the same.

II. EXISTING WORKS

A. EMO Player

This player automatically plays songs based on the emotions of the user. It can able to recognize human facial emotions and play according to their current moods. The emotions can be able to identify using Machine Learning Algorithms.

B. Smart Music Player

This player will help the user to search the trending or their favorites songs and saves them in their cloud system and able to listen to music from the cloud instead of their physical device.

C. Spotify

This music player also gives access to songs from artists all over the interworld. There are also options to download the contents from them if they upgrade it to premium.

Grenze ID: 01.GIJET.7.1.603

III. Related Works

A. Alexa

Alexa is an Amazon'sassistance AI-first used in smart speakers called Echoes. It has capable of having interactions with the users, playing music, streaming videos, reading audiobooks, etc. Now the echoes have 360 sound and Dolby sounds.
Alexa plays songs depends upon the situation.For example, plays music to get relax at night, or some music while jogging.

IV. Proposed Systems

There are some devices in markets such as Alex, Google Assistants can be able to play the songs and able to control them using our voices, but we need to buy an extra physical device and a standard internet connection to run that device. But using our app we can be able to hear songs and able to control them using our voices for free of cost and need a small amount of internet to access data dictionary.

V. Algorithms

A. Automatic speech recognition (ASR)

This algorithm helps the user to speak to the machine as they were speaking to a fellow person. It is one of the most developing algorithms in Natural Language Processing. The accuracy of these algorithms is up to 95% and these can achieve through constant training and the most accurate results are given for the question type that gives Yes or No as an answer.

1. The user speaks to the machine.
2. The machine converts the user's voice commands into an audio format such as .wav.
3. The noise in your sounds will be removed.
4. Extract the root words from your commands.
5. If the machine understands your command, then it will responses accordingly.

B. Limited User Large Vocabulary

It is used for small-scale uses of NLP where the commands are not very complicated. The algorithm has good accuracy and has many vocabularies. The algorithms need to train so that it works best with some basic simple commands.

C. Speech To Text (STT)

STT lets the user control computer functionsby voice. The system is done in two parts,first, it processes acoustic audiosignals that are captured by a microphone and second itmaps the signal to words. Hidden Markov Model (HMM) is used for modeling each letter.Vector quantization trains the dataset.

VI. Modular Description

A. Song Picker

This is the First module where the user will choose the folder that has the audio files in it. Or the user can be able to enter the URL that has the audio links in it. Then the user can be able to set the songs as favorites then it will store them in the default locations that the use was present in the app. They so that they can be able to access it easily.

B. Song Player

Here the music and audio files that are in the files will be played. There the user can be able to use their voice commands to play as they wanted to do. Then the user Commands will be executed as per the Trained modules. If the user wants to use the buttons instead of using the voice commands.

C. Favorite Section

If the user likes a song, then he can be able to set that as Favorites, and when he asks for his favorites then the player will play the songs that are marked as favorite. While playing the music the AI must read the surrounding environment and if there is noise around then the volume has to be increased. If the user is speaking to someone then the audio player has to pause the song and continue when the conversation is over. The player must read the

type of song it was playing and automatically reads its type such as melody, rock, jazz, etc. Each of the song types has its own best amplifier tone to get the best out of it. The AI will recognize the type of the song and plays accordingly.

VII. Conclusions

It will provide a hand-free environment for the users to use their audio player while they are doing other chores. These will help them use it hand-free. It will also help the differently-abled people and the users whose hands are doing some other jobs. The users with wireless ear pods benefit the most by control their music player while doing using their commands.

References

[1] ZhangLandLiuB.Aspectandentityextractionforopinionmining.In:ChuWW(ed.)*Dataminingand knowledge discoveryfor big data.NewYork: Springer*, 2014, pp. 1–40. https://doi.org/10.1007/978-3-642-40837-3_1

[2] Jime ´nez-Zafra SM, Martı ´n-Valdivia MT, Martı ´nez-Ca ´mara E et al. Combining resources to improve unsupervised sentiment analysis at aspect-level. J Inf Sci 2016; 42: 213–229.

[3] https://doi.org/10.1177/0165551515593686

[4] Muangon A, Thammaboosadee S and Haruechaiyasak C. A lexiconizing framework of feature-based opinion mining in tourism industry. In: Proceedings of the fourth international conference on digital information and communication technology and it's applications (DICTAP), Bangkok, Thailand, 6–8 May 2014, pp. 169–173. New York: IEEE. https://doi.org/10.1109/DICTAP.2014.6821677

[5] Wang W, Wang H and Song Y. Ranking product aspects through sentiment analysis of online reviews. J Exp Theor Artif In 2017; 29: 227–246. https://doi.org/10.1080/0952813X.2015.1132270

[6] Marrese-Taylor E, Vela´squez JD and Bravo- Marquez F. Opinion zoom: a modular tool to explore tourism opinions on the web.In: Proceedings of the 2013

[7] IEEE/WIC/ACM international joint conferences on web intelligence (WI) and intelligent agent technologies (IAT), vol.3, Atlanta, GA, 17–20 November 2013, pp.261–264. New York: IEEE. https://doi.org/10.1109/WI-IAT.2013.193

[8] Marrese-Taylor E, Vela´squez JD and Bravo- Marquez F. A novel deterministic approach for aspect-based opinion mining in tourism products reviews. Expert Sys Appl 2014; 41: 7764–7775. https://doi.org/10.1016/j.eswa.2014.05.045

[9] Colhon M, Ba˘ dica˘ C and Sxendre A. Relating the opinion holder and the review accuracy in sentiment analysis of tourist reviews. In: Proceedings of the international conference on knowledge science, engineering and management, Melbourne, VIC, Australia, 19–20 August 2014, pp. 246–257. New York: Springer. https://doi.org/10.1007/978-3-319-12096-6_22

[10] Mukherjee A and Liu B. Aspect extraction through semi-supervised modeling. In: Proceedings of the 50th annual meeting of the association for computational linguistics: long papers, vol. 1, Jeju Island, Korea, 8–14 July 2012, pp. 339–348. New York:ACM.

[11] Kayaalp NF, Weckman GR, Ii WAY et al. Extracting customer opinions associated with an aspect by using a heuristic based sentence segmentation approach. Int J Bus Inf Syst 2017; 26: 236–260. https://doi.org/10.1504/IJBIS.2017.086335

[12] Wu Y and Ester M. Flame: a probabilistic model combining aspect based opinion mining and collaborative filtering. In:Proceedings of the eighth ACM international conference on web search and data mining, Shanghai, China, 2–6 February 2015, pp. 199–208. New York: ACM. https://doi.org/10.1145/2684822.2685291

[13] Shams M and Baraani-Dastjerdi A. Enriched LDA (ELDA): combination of latent Dirichlet allocation with word co-

[14] occurrence analysis for aspect extraction. Expert Syst Appl 2017; 80: 136–146. https://doi.org/10.1016/j.eswa.2017.02.038

[15] Mubarok MS and Adiwijaya Aldhi MD. Aspect- based sentiment analysis to review products using naïve Bayes. AIP Conf Proc 2017; 1867: 020060. https://doi.org/10.1063/1.4994463

[16] Xu X, Cheng X, Tan S et al. Aspect-level opinion mining of online customer reviews. China Commun 2013; 10: 25–41. https://doi.org/10.1109/CC.2013.6488828

[17] Catal C and Nangir M. A sentiment classification model based on multiple classifiers. Appl Soft Comput 2017; 50: 135–141. https://doi.org/10.1016/j.asoc.2016.11.022

[18] Pontiki M, Galanis D, Papageorgiou H et al. SemEval-2016 task 5: aspect based sentiment analysis. In: Proceedings of the 10th international workshop on semantic evaluation (Semeval-2016), San Diego, CA, 16–17 June 2016, pp. 19–30. Stroudsburg, PA:Association for Computational Linguistics. https://doi.org/10.18653/v1/S16- 1002 https://doi.org/10.18653/v1/S15-2082

[19] Akhtar MS, Gupta D, Ekbal A et al. Feature selection and ensemble construction: a two-step method for aspect based sentiment analysis. Knowl-Based Syst 2017; 125: 116–135. https://doi.org/10.1016/j.knosys.2017.03.020

[20] Ali F, Kwak KS and Kim YG. Opinion mining based on fuzzy domain ontology and support vector machine: a proposal to automate online review classification. Appl Soft Comput 2016; 47: 235–250. https://doi.org/10.1016/j.asoc.2016.06.003
[21] T. A. Rana, Y.-N. Cheah. Aspect extraction in sentiment analysis: comparative analysis and survey. Springer Science+Business Media Dordrecht 2016, Artif Intell Rev (2016) 46:459– 483. https://doi.org/10.1007/s10462-016-9472-z
[22] W. Medhat et al. Sentiment analysis algorithms and applications: A survey, 2090-4479Ó2014 Production and hosting by Elsevier B.V. on behalf of Ain Shams University.
[23] Lakshmana Phaneendra Maguluri et al. A New sentiment score based improved Bayesian networks for real-time intraday stock trend classification. International Journal of Advanced Trends in Computer Science and Engineering vol.8 no.4, pp.1045-1055, July-Aug 2019. https://doi.org/10.30534/ijatcse/2019/10842019
[24] Maguluri, L.P. et al. Comparative analysis of machine learning algorithms on social media test. International Journal of Engineering and Technology(UAE).2018
[25] Joshua Goodman: A bit of progress in language modeling, MSR Technical Report, 2001.
[26] Yoon Kim, et al.: Character-Aware Neural Language Models, 2015.
[27] Kishore Papionine: A Method for Automatic Evaluation of Machine Translation, ACL, 2002
[28] Matthew E. Peters: Deep Contextualized word representations, 2018
[29] Alec Radford, et al.: Language Models are Unsupervised Multitask Learners,2018.
[30] Donald Hindle and Mats Rooth. Structural Ambiguity and Lexical Relations, Computational Linguistics, 2018.
[31] Daniel Andor et al.: Globally Normalized-Based Neural Networks, 2016.
[32] Oriol Vinyals, et al.: Grammar as a Foreign Language, 2015.
[33] Zhiheng Huang et al.: Bidirectional LSTMCRF Models for Sequence Tagging, 2015.
[34] Xuezhe Ma and Eduard Hovy: End-to-end Sequence Labeling via Bi-directional LSTM- CNNsCRF, ACL 2016
[35] https://github.com/watson-developer-cloud/unity- sdk
[36] https://developer.ibm.com/technologies/artificialint elligence/projects/watsondeveloper-cloud-unity- sdk/