Using Natural Language Processing (NLP) Based Techniques for Handling Customer Relationship Management (CRM)

[1] Rohit Chitte, [2] Riddhi Mandal*, [3] Romir Mathur, [4] Adhyatma Sharma, [5] Dhananjay Bhagat

[1] [2] [3] [4] B.Tech., Department of Artificial Intelligence, G. H. Raisoni College of Engineering, Nagpur, India, India
[5] Assistant Professor, Department of Artificial Intelligence, G. H. Raisoni College of Engineering, Nagpur, India, India
Corresponding Author Email: [1] chitterohit2112@gmail.com, [2] riddhimandal1814@gmail.com, [3] mathurromir@gmail.com, [4] yash45430@gmail.com, [5] mailto:dhananjaybhagat84@gmail.com

Abstract— The paper talks about using NLP and its techniques to simplify CRM systems. The use of NLP techniques tries to remove human error in CRM systems. In CRM systems the entry of information has always found errors in regard to customer and service provider calls. Recently the data entry by the service provider has seen massive errors resulting in unsatisfactory engagement with the customer. The paper describes the use of NLP application of text to speech conversion and text summarization along with the use of NLP modules such as tokenization and stop word removal to manipulate the text to retrieve valuable information has improved the CRM system management. The paper shows flowcharts of working from input as entry of audio file using a GUI and producing the output in the form of a python dictionary containing all the required information. The paper shows the pre-processing of audio files where the NLP modules have been used. The paper illustrates the outputs and its working from start to end. The research talks about the future scope of CRM system and increasing its productivity by integrating factors of Artificial Intelligence.

Index Terms—Natural Language Processing (NLP), Customer Relationship Management(CRM), Graphic User Interface, Artificial Intelligence, Spell Checker, Speech-to-text.

I. INTRODUCTION

The enormous rise of the data due to digitalization and the applications of Machine learning (ML) algorithms to implement intelligent products and services. To handle the data, it's a very difficult task for human beings. In the growing world customer cognition is required [1] because working on a project which doesn't impact the customer will be useless. Natural Language processing (NLP) and Machine learning makes a tremendous approach in finding out the customers' needs. Businesses use CRM to hold customers and enhance the sales and marketing [2].

The system organizes the customer information at various platforms, at the point of contact between customer and company where the company may use live chat, website, social networks, direct mail, marketing materials and telephone. The system also gives detailed information on a customer's personal information, its purchasing and concern with the product and the past history the customer had made.

Fig.1. Audio to text Summarization

In the proposed work authors have used Machine Learning and Natural Language processing techniques to convert the audio conversation between the customer and customer service representative into the text format as an output in a summarised form. Further, these summarised texts are fed into a spell corrector module which makes corrections in the summarised text and then these corrected spelling in sentences make a complete script, which includes all the dialogue present in the audio which can be easily used for solving the queries of customers as shown in Fig.1.. This system leads to a helping model for customer service representatives, in which these outputs will be automatically stored and can give solutions according to the queries raised by the customer.

The study promotes the evolution in the domain of AI in CRM and conveys unique essential ways for further studies. The remaining part of the paper is arranged as follows: 2. Presents literature review 3. Methodology, 4. Result 5. Future Scope and 6. References
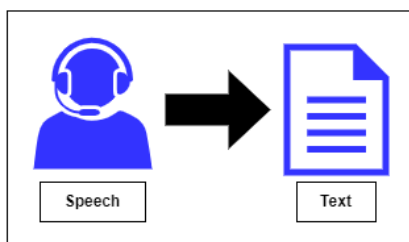
II. LITERATURE REVIEW

The usual approach towards using AI in customer relationship management involves the use of bibliometric analysis in machine learning algorithms. Bibliometric analysis[1] is nothing but the statistical and mathematical calculations of the bibliometric data. It is used to establish a relationship between keywords in a bibliometric data , this helps in generating an overall picture of the data and provide

possible trends and potential opportunities. This approach however cannot take speech data into account and requires article-based data or text data.

All the other approaches usually deal with either only the text data or the speech data[3]. Our approach is a full-fledged solution to the problem statement which involves converting speech data into text data first and then using data engineering and data cleaning to feed the cleaned data into a machine-learning model for a complete and useful summary.

With our approach of speech-to-text conversion and then using machine learning on the data, we have achieved a good accuracy[4]. The data generated from various sources and customers when used in an efficient manner can help a company deduce important insights about the market or their product. The quality of data being high as the source is directly the customers thus it will also help in updating or improving the quality of the product, providing higher quality product leading to customer satisfaction.

Ben Dickson (2017) -The founder of tech talks who blogs and writes about upcoming technology has described how CRM is the new industry in which AI is being introduced. He mentions how SalesForce, the leader in the CRM industry, announced Einstein, an Artificial intelligence assistant that when launched[5], will work across all platforms and work relentlessly to study the flood of data that Sales Forces collects from sales, e-commerce activity, emails, etc. It helps the company by highlighting the main feature and engaging in an analysis of the weak points of the company.

Genesis Rivas (2017) - Who studied intelligent data analysis and automated entry system claims, firms and companies can put up an overall profile of a customer or user that can help[6] in predicting the customer's response to a product. This can further enhance the firm's data-collecting procedure and bring about positive growth. . Furthermore AI can also be used for customer engagement and improve the understanding[7] and analysis of customer emotions. With the use of AI techniques like facial recognition and other biometric techniques, a sense of trust can be created among the customers. The evolution towards a more efficient CRM allows cost and time saving and can help in the overall sales of a company.

AI in CRM is already being used by vendors around the globe for different reasons, to increase business or to reduce the time required for sales and marketing. Some of the prominent vendors that use AI in CRM are[8]:

1. Adobe Sensei
2. IBM Watson
3. Microsoft Cognitive Services and Dynamics
4. Oracle AI
5. Salesforce Einstein
6. SAP Leonardo
7. Sugar CRM
8. Zoho

III. METHODOLOGY

The section illustrates the complete proposed model consists of several steps as shown in Fig.2 Represents the whole steps involved for the system. In the proposed technique, the system is divided into three different segments named Speech to Text, Spelling Corrector, and Keyword spot, where the audio file come across all the step and gives an output in a keyword. The recorded audio clips dataset is summarised and extracted the features into the text format using Google Speech recognition Application Programming Interface (API)[10]. The project expects audio files as an input e.g ".wav" format. After uploading an audio file it is broken into multiple chunks based on pauses in the audio file the audio[11] files are stored at a specific location the speech-to-text module generates text one by one from these chunks. This text is fed to the spell corrector module for checking and rectifying spelling errors. Further keyword spotting is applied on rectified text then the final output is shown.
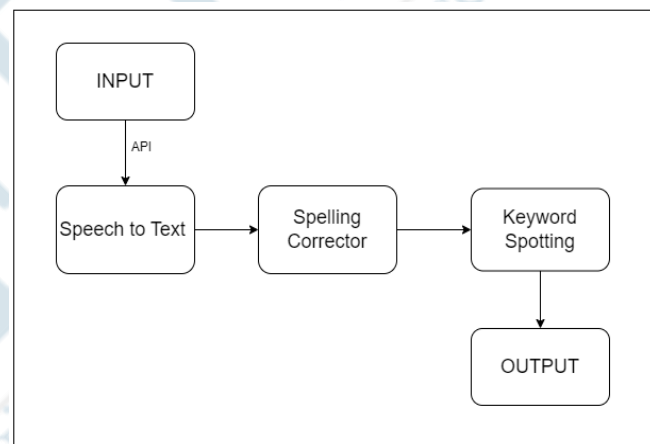

Fig.2. Flowchart of steps involved in the system

ClientService.py is the main application file. Execution of the project starts from a point in this file. Necessary imports have been done starting from different modules defined in the project. A detailed explanation of each module is given in respective sections. As the project is deployed as a web application, the Flask module has been used to serve the purpose of creating a web application as shown in fig.3. The components of ClientService.py are as follows the First one is Class ClientService Client Service class has necessary object variables and methods. Creates chunks of input audio file, Stores at the given path converts speech to text and extracts keys from the predefined dictionary. The next component[12] is Object Variables which includes Folder Path i.e Path of the folder where input audio files are stored, Filelist List of all the files in FolderPath, and SeparatedOutputFiles: Path of the folder to save separated Audio Files. , OutputText: Python Dictionary to store Output Text after converting Speech to Text. The next crucial component is Methods out of which the first one is

ProcessAudioFile it's objective is to Take audio files from FolderPath, Break the audio file into chunks based on Frequency(Silence) Internally uses the GenerateTranscript Function from file ai speech to text.transcript generator Module as shown in fig.4, Generated Text from Speech is stored in inputFileTranscripted and printed[13]. Applies Spelling Corrector on generated text internally uses spell corrected from file ai spellingcorrector. spellcorrector module for spelling correction. And then prints input text and corrected text. Further uses addkey function from class AddMultiKeywords from module file ai keywordspotter.keywordSpotter , the task of addkey function is to extract keys if the values of those specific keys are encountered in text. The project is controlled over distinct Flask Routes which are as follows, Homepage("/"): after hitting the URL index.html page is loaded to display UI. Getinputfile("/upload"): Uploading the file after pressing upload button at UI.processInputFiles("/Startprocessing"): Starts processing the audio file i.e breaking the audio file in chunks based on silence, storing it at a specific location converting speech to text, applying spelling corrector, finding keys as shown in fig.5.

Fig.3. Speech to text

The project includes some of the important files and folders that have specific python modules or are used to store intermediate results. These are as follows, ArchivedInputFiles[Folder]This Folder stores the previous input files in the Input Files Folder on which processing has been already done. This is to accumulate the already processed data[14]. File ai keywordspotter[Python Package user defined] File ai keywordspotter is a python package, it contains a python module named keywordSpotter.py and in it, there is a class named AddMultiKeyword, This class has two object variables and one method. Variables used in this module are Text which is used to receive after converting speech to text from individual chunks of audio. Keyword dict stores thedictionary of keywords that are to be spotted in the sentence stored in the text variable. Addkey method uses an instance of KeywordProcessor class from the flashtext module of python two methods of keywordprocessor class have been used in this method first one is add keywords from dict()[] which accepts dictionary checks if values are instances of list object in python and if yes then adds the key-value pairs to the dictionary.
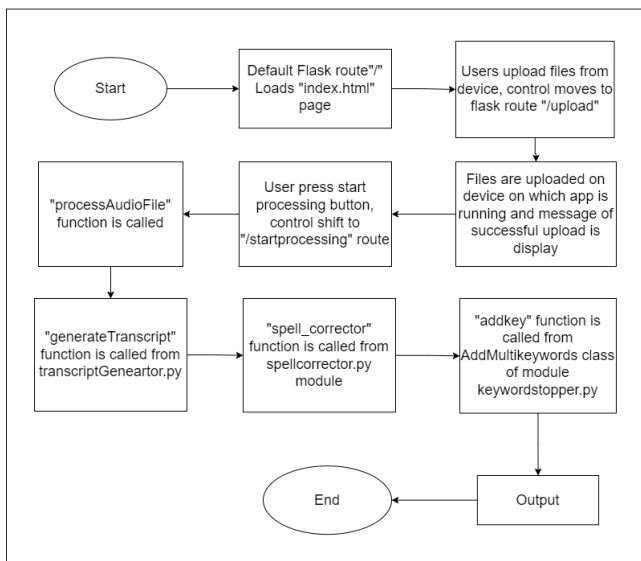

Fig.4. Spelling Corrector

Fig.5. Keyword Spotting

The second one is extract keyword which accepts a sentence/Text and extracts the keys if the values corresponding to those keys are matched/found in the sentences. e.g ["football" : ['messi","ronaldo",sunil chetri","kaka"] extract keyword("I love the game that is played by a wonderful player sunil chetri") above function will return "football" because above sentence has a value sunil chetri that corresponds to a key football. file ai speech to text[Python Package user defined]. This package contains a module: transcriptGenerator.py, the primary function of this module is to break down the audio file[9][15] into multiple chunks based on frequency change(silence) then implement speech-to-text on those chunks and generate a transcript. Modules imported are Speech recognition, Os, and Pydub. Only one function i.e generateTranscript has been defined in this module which takes care of the entire functionality. Input parameters are path(path of input audio files) and folderpath(path of folder to store the chunks of audio) and return a dictionary where keys are names of audio file chunks stored in folderpath variable and values are the text i.e obtained after applying speech to text[16] on those chunks. file ai spellingcorrector [python package user defined] spellcorrector.py is a python module basic functionality of this module is to rectify the spelling error in the text given as a parameter to the function[17]. Input is incorrect sentence and output is corrected sentence. Modules imported are symspellpy, pkg resources as shown in fig.6.

Fig.6. Flowchart of the process of the model

The next folder is Static [Folder] Contains the js subdirectory which consists of the index.js file which is responsible for UI. Template [Folder]Consist of index.html file which is responsible for UI. Symspellpy[Python Package imported] Symspellpy package is imported from the Pypy distributor this package is used in spellcorrector module to rectify spelling mistakes.

IV. RESULT

The resulting output is broken down into three steps A. Speech to text B. Spelling Corrector C. Keyword spotting.

A. Speech to text

The output for this module shows all the sentences spoken on an audio file as broken-down sentences as part of the conversion, the audio file is broken down with the help of a frequency modulator. The final statements are shown in Fig.7 in terms of chunks for clarity.

B. Spelling Corrector

The chunks of the audio file are later pre-processed to remove spelling mistakes from the speech-to-text converted data. It is a very predominant issue of speech-to-text conversion that it misspells some words resulting in lower accuracy of the fetched data, hence to rectify this we use a spelling correctorfor producing a refined array of words as shown in Fig.8. The output shows the words which were previously wrongly written are now corrected.

Output

Speech to Text

{'chunk 1': 'I come from england. ', 'chunk 2': 'You cancel england. ', 'chunk 3': 'Do you like manchester united. ', 'chunk 4': 'Ipl light football. ', 'chunk 5': "You don't like football. ", 'chunk 7': 'Ic. ', 'chunk 8': 'Why did you come here. ', 'chunk 9': 'I think the vacation. ', 'chunk 10': 'Fifa vacation are you having fun. ', 'chunk 11': 'Yes i'm having a great time. ", 'chunk 12': 'What do you like the best here. ', 'chunk 13': 'What do i like the best. ', 'chunk 14': 'People very kind. ', 'chunk 16': 'The people here are time. ', 'chunk 17': 'You like first wizard. ', 'chunk 18': 'Yeah where is great. ', 'chunk 19': 'Is the field to spicy food. ', 'chunk 20': 'Now i love bicycle. ', 'chunk 21': 'Eat fit everyday. ', 'chunk 22': 'How much longer will you stay. ', 'chunk 23': 'State two days. ', 'chunk 24': 'When you go. '}

Fig.7. Output of Speech-to-text

Spelling Corrector

{'chunk 1': 'i come from england', 'chunk 2': 'you cancel england', 'chunk 3': 'do you like manchester united', 'chunk 4': 'ill light football', 'chunk 5': "you don't like football", 'chunk 7': 'in', 'chunk 8': 'why did you come here', 'chunk 9': 'i think the vacation', 'chunk 10': 'fife vacation are you having fun', 'chunk 11': "yes i'm having a great time", 'chunk 12': 'what do you like the best here', 'chunk 13': 'what do i like the best', 'chunk 14': 'people very kind', 'chunk 16': 'the people here are time', 'chunk 17': 'you like first wizard', 'chunk 18': 'yeah where is great', 'chunk 19': 'is the field to spicy food', 'chunk 20': 'now i love bicycle', 'chunk 21': 'eat fit everyday', 'chunk 22': 'how much longer will you stay', 'chunk 23': 'state two days', 'chunk 24': 'when you go'}

Fig.8. Output of Spelling Corrector

C. Keyword spotting

The end result shows what were the post-required data for the company and is presented in terms of a sequence of words. keyword spotting is done to get the most valuable parts of the conversations.the data fetched from the conversation highlights the name of the issue and the required other details which need to be used for generating tokens as shown in fig.9 output of keyword spotting.

Keyword Spotting with sentence

{'chunk 1': {'key': 'place', 'value': 'england', 'text': 'I come from england. '}, 'chunk 2': {'key': 'place', 'value': 'england', 'text': 'You cancel england. '}, 'chunk 3': {'key': 'team', 'value': 'manchester united', 'text': 'Do you like manchester united. '}, 'chunk 4': {'key': 'game', 'value': 'football', 'text': 'Ipl light football. '}, 'chunk 5': {'key': 'game', 'value': 'football', 'text': "You don't like football. "}}

Fig.9. Output of Keyword Spotting

V. FUTURE SCOPE

The age of AI has grown massively and has contributed to many fields. CRM systems with the enablement of AI have seen sound development. The paper proposes that in the future an automatic system is developed from the service provider end which asks questions and records data and it will also work in different languages. The system uses frequency-enabled methods to differentiate voice and improves the accuracy of the data recorded. The system

should automatically generate tickets for issues raised and promptly manage customers and handle their request.

VI. ACKNOWLEDGMENT

REFERENCES

[1] Ledro, C., Nosella, A. and Vinelli, A. (2022), "Artificial intelligence in customer relationship management: literature review and future research directions", Journal of Business and Industrial Marketing, Vol. 37 No. 13, pp. 48-63.

[2] Al Smadi, T., Al Issa, H.A., Trad, E. and Al Smadi, K.A. (2015) Artificial Intelligence for Speech Recognition Based on Neural Networks. Journal of Signal and Information Processing, 6, 66-72.

[3] Changhan Wang , Yun Tang, Xutai Ma1, Anne Wu , Sravya Popuri , Dmytro Okhonko , Juan Pino "FAIRSEQ S2T: Fast Speech-to-Text Modeling with FAIRSEQ" ,arXiv:2010.05171v2 [cs.CL] 14 Jun 2022

[4] Ye Jia Yu Zhang Ron J. Weiss Quan Wang Jonathan Shen Fei Ren Zhifeng Chen Patrick Nguyen Ruoming Pang Ignacio Lopez Moreno Yonghui Wu "Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis", arXiv:1806.04558v4 [cs.CL] 2 Jan 2019

[5] Xutai Ma, Juan Pino, Philipp Koehn, "SimulMT to SimulST: Adapting Simultaneous Text Translation to End-to-End Simultaneous Speech Translation", arXiv:2011.02048v1 [cs.CL] 3 Nov 2020

[6] Tahir Iqbal Md Nazmul Khan, "The Impact of Artificial Intelligence (AI) on CRM and Role of Marketing Managers", January, 2021

[7] Mishra, Niharika and Mukherjee, Sushanta, Effect of Artificial Intelligence on Customer Relationship Management of Amazon in Bangalore (2019). International Journal of Management, 10 (4), 2019, pp. 168–172.

[8] Kavitha Natarajan, "A Study on Impact of Artificial Intelligence in Customer Relationship Management", Volume 8, Issue 8

[9] S. Kumar Deb, R. Jain and V. Deb, "Artificial Intelligence—Creating Automated Insights for Customer Relationship Management," 2018 8th International Conference on Cloud Computing, Data Science Engineering (Confluence), 2018, pp. 758-764, doi: 10.1109/CONFLUENCE.2018.8442900

[10] Hopkinson, P, Perez-Vega, R and Singhal, A 2018, Exploring the Use of AI to Manage Customer Relationship. in Proceedings of the 51st Academy of Marketing Conference. 51st Academy of Marketing Conference 2018, Stirling, United Kingdom, 2/07/18.

[11] R. Boorugu and G. Ramesh, "A Survey on NLP based Text Summarization for Summarizing Product Reviews," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), 2020, pp. 352-356, doi: 10.1109/ICIRCA48905.2020.9183355.

[12] J. Doshi, "Chatbot User Interface for Customer Relationship Management using NLP models," 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), 2021, pp. 1-4, doi: 10.1109/AIMV53313.2021. 9670914.

[13] Y. -C. Guo and D. -X. Niu, "A Knowledge-Based Intelligent System for Power Customer Service Management," 2007 International Conference on Machine Learning and Cybernetics, 2007, pp. 2925-2930, doi: 10.1109/ICMLC.2007.4370648.

[14] S. Kumar Deb, R. Jain and V. Deb, "Artificial Intelligence—Creating Automated Insights for Customer Relationship Management," 2018 8th International Conference on Cloud Computing, Data Science and Engineering (Confluence), 2018, pp. 758-764, doi: 10.1109/CONFLUENCE.2018.8442900.

[15] A. Nguyen et al., "System Design for a Data-Driven and Explainable Customer Sentiment Monitor Using IoT and Enterprise Data," in IEEE Access, vol. 9, pp. 117140-117152, 2021, doi: 10.1109/ACCESS.2021.3106791.

[16] C. -S. Lee et al., "Model-based Customer-Relationship Management System and Strategic Board Game: Analogical Training," 2021 3rd International Workshop on Artificial Intelligence and Education (WAIE), 2021, pp. 69-73, doi: 10.1109/WAIE54146.2021.00022.

[17] Sujata D More, Dr. M. B. CHADAK Mangala S, Madankar, "Data Extraction by Using Natural Language Processing Tool", 2018/1/1 Journal Volume 8 Issue 5.