ISSN PRINT 2319 1775 Online 2320 7876

Research paper **UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022**

To Study And Analyze The Use Of Speech Recognition Systems And Its Benefits Across Various Demographics

Dr. Kavita Khadse

Assistant Professor, CRKIMR, Mumbai

Abstract:

Communication among the human being is dominated by spoken language, therefore it is natural for people to expect speech interfaces with computer .computer which can speak and recognize speech in native language. Machine recognition of speech involves generating a sequence of words best matches the given speech signal. Some of known applications include virtual reality, multimedia searches, auto-attendants, travel information and reservation, translators, natural language understanding and many more applications.

INTRODUCTION

The speech is primary mode of communication among human being and also the most natural and efficient form of exchanging information among human in speech. So, it is only logical that the next technological development to be natural language speech recognition. Speech recognition can be defined as the process of converting speech signal to a sequence of words by means algorithm implemented as a computer program. Speech processing is one of the exciting areas of signal processing. The goal of speech recognition area is to developed technique and system to developed for speech input to machine based on major advanced in statically modeling of speech, automatic speech recognition today find widespread application in task that require human machine interface such as automatic call processing. Since the 1960s computer scientists have been researching ways and means to make computers able to record interpret and understand human speech. Throughout the decades this has been a daunting task. Even the most rudimentary problem such as digitalizing (sampling) voice was a huge challenge in the early years. It took until the 1980s before the first systems arrived which could actually decipher speech. Off course these early systems were very limited in scope and power.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

1.1. Type of speech

Speech recognition system can be separated in different classes by describing what type of utterances they can recognize.

1.1.1 Isolated Word

Isolated word recognizes attain usually require each utterance to have quiet on both side of sample windows. It accepts single words or single utterances at a time .this is having "listen and non listen state". Isolated utterance might be better name of this class.

1.1.2 Connected word

Connected word system are similar to isolated words but allow separate utterance to be "run" together minimum pause between them.

1.1.3 Continuous Speech

Continuous speech recognizers allows user to speak almost naturally, while the computer determine the content recognizer with continuous speech capabilities are some of the most difficult to create because they utilize special method to determine utterance boundaries.

1.1.4 Spontaneous Speech

At a basic level, it can be thought of as speech that is natural sounding and not rehearsed. An ASR system with spontaneous speech ability should be able to handle a variety of natural speech feature such as words being run together.

1.2 Overview of speech recognition

1.2.1 Between years 1920 to 1960s

In the early 1920s machine recognition came in to an existence. The first machine to recognize speech to any important step commercially named, radio rex (toy) was manufactured in 1920. Research into the concepts of speech technology started in the year 1936 at bell laboratory. In 1939, bell labs showed a speech synthesis machine at world fair in New York before they later abandoned struggle to develop speech simulated listening and recognition based on an

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022**

inappropriate conclusion that artificial intelligence would eventually be necessary for accomplishment.

In 1950s, earliest attempt to device system for automatics speech recognition by machine were made when several researchers tried to exploit the fundamental concepts of acoustic phonetics. During this period, most of the speech recognition system investigated special resonances during the vowel system of each utterance which were extracted from output region of each utterance which were extracted from output signals of an analogue of filter bank and logic circuits.

In 1952, at bell laboratories, David, Biddulph and Balashek developed a system for isolated digit recognition for a single speaker. The system depends deeply on measuring spectral resonances during the vowel region of each digit. In 1956, at RCA laboratory, Olson and Belar tried to recognize 10 distinct syllables of a single talker, as personified in 10 monosyllabic words. In another effort of fry and denes at University College in England, in 1959, a phoneme recognizer that recognizes four vowels and nine consonants were developed. They used a spectrum analyzer and a pattern matcher to make the recognition decision.

Again during 1959 period vowel recognizer of Forgie and Forgie built at MIT Licoin laboratory in which 10 vowels embedded in a /b/ vowel /t/ format were recognized in a speaker independent mode.

1.2.2 Between years 1960 to 1980s

In the 1960s at Japanese labor Atory Suzuks and Nakata started their research in the Speech recognition field and developed special purpose hardware as part of their system due to computation that were not fast enough then. In 1962s, another hardware effort in japan was the work of Sakari and Doshita of Kyoto University, who developed hardware phoneme recognizer. The third Japanese work was the digit recognizer hardware of negate and coworkers at NEC laboratory in 1963.

In the separate effort of Japan Sakoe and Chiba at NEC laboratories dynamic programming technique was used to solve the non-uniformity problems the final achievement of annotation in the late 1960s was pioneering research of Reddy in the area of continuous speech recognition by dynamic tracking of phonemes.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022**

The field of isolated word or discrete utterance recognition became a possible and functional technology in 1970s through the fundamental studies by Velichko and Zagoruyko in Russia, Itakura in United State, Cakoe and Chiba in Japan. During this decade, striving speech understanding project was funded by defense advanced research projects agencies (DARPA), which lead to various seminal system and technology. One of the demonstration of speech understanding was achieved by CMU in 1973 and heresay-i system was able to used semantic data to significantly moderate the number of alternatives considered by the recognizer. CMU's Harphy System was displayed to be able to recognize speech using a vocabulary of 1, 011 words with the judicious accuracy.

Another success of research in the 1970s was the beginning of a longstanding, extremely successful effort of group in large vocabulary speech recognition at IBM in which researchers studied three different tasks over a period of almost two decades, namely the New Raleigh language for simple database queries, laser potent text language for transcribing laser potent and lastly the office correspondent tasks called Tangora for dictation of simple communications.

1.2.3 Between years 1980 to 200s

Research in field of speech recognition in 1980s was characterized by a shift in technology from template based approach to statistical modeling method, most especially the Hidden Markov Model (HMM) approach.

The approach of Hidden Markov Model (HMM) was well known and understands in a few laboratories like primary IBM, Institute for Defense Analysis (IDA) and Dargon systems but it became extensively used in the middle of 1980s. Another innovative technology that came into existed in the late 1980s was the method of applying neural network to problem of speech recognition. The approach was first introduced in the 1950s, but they did not prove useful initially because they had many practical problems.

Era of 1980s was decade in which a major motivation was given to large vocabulary and continuous speech recognition system by the defences advanced research project agency (DARPA) community, sponsored a large research program aimed at accomplishing high word accuracy for 1000 word continuous speech recognition, database management task. Major research contributions resulted from effort at CMU, AT&T Bell labs, Lincoln Labs and SRI. The

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

CMU (also known as sphinx system) successfully integrated the statistical method of hmm with the network search strength of the earlier harpy system.

1.2.4 Between years 1990 to 2000

The year 1990s was a decade in which a number of innovations took place in the area of pattern recognition. The problem of pattern recognition which traditionally followed the framework of Bayes and required estimation of distributions for the data was changed into an optimization problem resulting to the reduction of the empirical recognition error.

During this decade, a key issue in design and implementation of speech recognition system was how to appropriately select the speech material used to train the recognition algorithm. A number of human language technology projects funded by DARPA in the 1980s and 1990s further enhanced the progress, as showed by many papers published in the proceedings of the DARPA speech and natural language/ human language workshop. The research describes the development of accomplishments for speech recognition that were conducted in the 1990s, at Fajitsu laboratories limited.

1.2.5 Between years 2000 till date

In year 2000, Variational Bayesian (VB) estimation and clustering technique were developed. This VB method was based on a subsequent distribution of parameters. Giuseppe Richardi developed this technique to solve the problem of adaptive learning in automatic speech recognition and also proposed active learning algorithm for automatic speech recognition in the year 2005. Some enhancements have been worked out on large vocabulary continuous speech recognition system on performance improvement.

Sadaoki and Furui investigated SR technique in 2005 that can adapt to solve speech variation using a large number of model trained based on clustering technique. De-Waehter, attempted to overcome the time dependencies problems in speech recognition (SR) by using straight forward template matching technique. In 2008, the authors explained the application of corresponding frame level of phone and phonological attributed classes. In the recent work carried out by Vrinda and Chander, suitable speech recognition was developed for hindi language, for the people that are physically challenges and cannot able to operate the computer through keyboard

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

and mouse, using Hidden Markov Model (HMM) to recognize speech sample to give admirable result for isolated words

1.3 Speech Databases

Speech databases have an extensive uses in Automatics Speech Recognition (ASR). It is also used in other key applications such as automatic speech synthesis, coding and analysis, speaker language identification and verification. These applications required large amounts of recoded database. Speech databases are most generally classified into multi-session and single-session databases.

Multi-session databases allow estimation of temporal intra-speaker variability. Based on acoustic environment, databases are recorded either in noise free environment, such as office or home. Also, according to the purpose of the databases, some corporations are designed for mounting and evaluating speech recognition.

While convenient, speech recognition technology still has a few issues to work through, as it is continuously developed. The pros of speech recognition software are it is easy to use and readily available. Speech recognition software is now frequently installed in computers and mobile devices, allowing for easy access.

The downside of speech recognition includes its inability to capture words due to variations of pronunciation, its lack of support for most languages outside of English and its inability to sort through background noise. These factors can lead to inaccuracies.

Speech recognition performance is measured by accuracy and speed. Accuracy is measured with word error rate. WER works at the word level and identifies inaccuracies in transcription, although it cannot identify how the error occurred. Speed is measured with the real-time factor. A variety of factors can affect computer speech recognition performance, including pronunciation, accent, pitch, volume and background noise.

The general difficulty of measuring performance lies in the fact that the Recognized word sequence can have a different length from the reference word sequence. The WER is derived

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

from the Levenshtein distance, working at the word level instead of the phoneme level. This problem is solved by first aligning the recognized word sequence with the reference word sequence using dynamic string alignment.

1.4 Current Scenario

Speech recognition technology might be becoming standard in new gadgets, but its accuracy will be what determines whether it really becomes a can't-live-without feature.

That's one of the messages delivered by Silicon valley venture capitalist Mary Meeker in her annual internet trends report. Meeker points out that voice input has the potential to be the most efficient form of computing: humans can speak 150 words per minute on average, but can only type 40. Now is the time for voice recognition to take over, too, since the technology is a logical fit with internet of things-connected devices, such as amazon echo or the apple watch.

What's kept speech recognition from becoming a dominant form of computing is its unreliability. Regional accents and speech impediments can throw off word recognition platforms, and background noise can be difficult to penetrate. And simply recognizing sounds isn't enough--to have any level of effectiveness, systems need to be able to distinguish between homophones (words with the same pronunciation but different meanings) and learn new words and proper names.

But it's getting closer. Meeker's presentation cited Andrew Ng, former Stanford professor and current chief scientist at Chinese search engine Baidu, as saying that 99 percent is the key metric: as accuracy in low-noise environments rises from 95 to 99 percent, voice recognition technology will expand from limited usage to massive adoption.

As recently as 2010, Meeker's presentation says, industry leaders were hovering around 70 percent accuracy. Now, some are approaching that key 99 percent threshold. Here are some of the best, in order of accuracy.

1. Baidu

The "google of china" is the country's biggest search engine, and at 96 percent, its voice recognition is better than most humans at identifying spoken words. The software it uses, deep speech 2, was developed in silicon valley and learned to understand words by listening to

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

thousands of hours of recordings while simultaneously reading their transcriptions. The system understands both english and mandarin, and it's growing in popularity in china, where voice commands are more popular due to the time it takes to type with the massive mandarin alphabet--and, of course, where google is blocked by the communist government.

2. Hound

The hound app, silicon valley company soundhound's flagship product, is a digital assistant that launched in march. It answers verbal questions and completes tasks like calculations, correctly identifying 95 percent of words in the process. A product nine years in the making, the app has a shazam-like feature that identifies songs--including, in some cases, ones hummed into it. Founder keyvan mohajer told tech crunch that his company started working on the technology before industry leaders like apple did, which has given it a head start in creating some of the best voice recognition technology there is.

3. Siri

Apple's siri might frustrate when it comes to finding answers, but as far as voice recognition goes, america's most-used personal assistant is near the top. At 95 percent accuracy, siri outpaces all its fellow silicon valley giants. And as for those faulty or nonsensical answers, the company hired a team of speech recognition experts trained in deep learning in 2014. The assistant's accuracy and intelligence should keep improving, which should make it less likely that siri responds to your request for help with a gambling problem with a list of casinos.

4. Google now

Google's voice search is 92 percent accurate, and can be used via the google app or for voice diction on android phones. Baidu's ng, who used to work at google, predicted that 50 percent of web searches will be performed using speech or images by 2019--and you can fully expect google to lead that charge. Google has done more work lately to improve accuracy in loud places, a feature that could help put it over the top.

5. Microsoft cortana

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

Cortana, the microsoft phone assistant now built into windows 10, composes messages, performs searches, and sets calendar events by way of voice commands. It's been measured above 90 percent accuracy--quite an improvement considering windows 95 had an error rate of close to 100 percent.

6. Amazon alexa

The amazon echo can do a lot--play music, adjust lighting, read recipes--without needing a screen or any manual activation. While the company won't reveal its internal word error rates, many users have pegged its word recognition as being a shade behind other voice platforms. The good news, though, is that alexa adapts to your voice over time, helping offset any issues it has with your particular dialect. And while others require the speaker to be within a few feet of its microphones, alexa operates from the next room.

LITERATURE REVIEW

1. **A Review on Speech Recognition Technique by Santosh Gaikwad, Bharti Gawali, Pravin Yannawar (Nov 2010)**

The paper discusses the techniques developed in each stage of speech recognition system. It also presents the list of techniques with their properties for Feature extraction. Through this information the authors have found that MFCC is used widely for feature extraction of speech. The paper is descriptive in nature and does not have any facts and figures to back the findings.

2. **Automatic Speech Recognition: A Review by Preeti Saini, Parneet Kaur (2013)**

The paper gives us the journey that Speech Recognition Systems have made over the years. Although the paper provided the basic understanding of the concept, there was no definitive conclusions that had been drawn by the authors.

3. **A Survey On Feature Extraction And Classification Techniques For Speech Recognition By Sanjay A. Valaki, Harikrishna B. Jethva**

Speech processing is really essential research spot where speaker recognition, speech synthesis, speech codec, speech noise reduction are some of the research areas. Many of the languages have

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022**

different speaking styles called accents or dialects. Speech recognition system is a way for the interface of human to machine. Automatic speech recognition is advance way to operate computer without much efforts through speech only. In this paper we have discuss lpc and mfcc techniques for future extraction and some classification methods to classify after recognition of speech word like hmm and ann. This paper is concludes with the decision on feature direction for developing technique in human computer interface system using gujarati language.

4. **Automatic Speech Recognition And Verification Using Lpc, Mfcc And Svm By Aaron M. Oirere, Ganesh Bapurao Janvale, Ratnadeep R.**

Speech has much capability as an interface between human and computer which comes under the human computer interaction (hci). The major challenge has been the nature of voice is ever varying speech signal. The paper presents the development of the speech recognition system using swahili speech database which was collected in three sets: digits, isolated words and sentences from both native and non native speakers of swahili language. Different feature extraction techniques deployed in the system are: linear prediction coding (lpc) and mel-frequency coefficients (mfcc). We have used the 12 coefficient features from mfcc and 20 coefficients features from lpc. All these features extracted techniques are applied and tested for the own developed swahili speech database. Recognition and verification were done using confusion matrix and support vector machine (svm) as a classifier for the classification purpose. Lda was tested for the entire dataset for the dimension reduction. Lda gave a good clustering. The performance of the system was checked on basis of their accuracy; confusion with mfcc 50.9%, confusion with lpc 50.1%, the higher recognition rate in each data set were as follows numeric data: mfcc: 75%, lcp:72% , isolated word data: mfcc: 65.2% lpc: 66.67%, sentence data mfcc: 63.8%, lpc: 59.6.

5. **Analysis And Optimization Of Telephone Speech Command Recognition System Performance In Noisy Environment**

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022**

This paper deals with the analysis and optimization of a speech command recognition system (scrs) trained on czech telephone database speechdat (e) for use in a selected noisy environment. The scrs is based on hidden markov models of context.

6. **Speech Recognition By Machine: A Review By M.A.Anusuya, S.K.Katti**

This paper presents a brief survey on automatic speech recognition and discusses the major themes and advances made in the past 60 years of research, so as to provide a technological perspective and an appreciation of the fundamental progress that has been accomplished in this important area of speech communication. After years of research and development the accuracy of automatic speech recognition remains one of the important research challenges (eg., variations of the context, speakers, and environment).the design of speech recognition system requires careful attentions to the following issues: definition of various types of speech classes, speech representation, feature extraction techniques, speech classifiers, database and performance evaluation. The problems that are existing in asr and the various techniques to solve these problems constructed by various research workers have been presented in a chronological order. Hence authors hope that this work shall be a contribution in the area of speech recognition. The objective of this review paper is to summarize and compare some of the well known methods used in various stages of speech recognition system and identify research topic and applications which are at the forefront of this exciting and challenging field.

7. **A Model And A System For Machine Recognition O F Speech By D. Raj Reddy, Lee D. Erman, And Richard B. Neely**

This paper presents a model for machine recognition of connected speech and the details of a specific implementation of the model, the hearsay system. The model consists of a small set of cooperating independent parallel processes that are capable of helping in the decoding of a spoken utterance either individually or collectively. The processes use the 4 'hy pothesize-and-test" paradigm. The structure of hearsay is illustrated by considering its operation in a particular task situation: voice-chess. The task is to recognize a spoken move in a given board position. Procedures for determination of parameters, segmentation, and phonetic descriptions are outlined. The use of semantic, syntactic, lexical, and phonological sources of

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

knowledge in the generation and verification of hypotheses is described. Preliminary results of recognition of some utterances are given.

8. The Impact Of Speech Recognition On Speech Synthesis By Mari Ostendorf, Ivan Bulyko

Speech synthesis has changed dramatically in the past few years to have a corpus-based focus, borrowing heavily from advances in automatic speech recognition. In this paper, we survey technology in speech recognition systems and how it translates (or doesn't translate) to speech synthesis systems. We further speculate on future areas where asr may impact synthesis and vice versa.

9. Advances In Artificial Intelligence Using Speech Recognition By Khaled M. Alhawiti

This research study aims to present a retrospective study about speech recognition systems and artificial intelligence. Speech recognition has become one of the widely used technologies, as it offers great opportunity to interact and communicate with automated machines. Precisely, it can be affirmed that speech recognition facilitates its users and helps them to perform their daily routine tasks, in a more convenient and effective manner. This research intends to present the illustration of recent technological advancements, which are associated with artificial intelligence. Recent researches have revealed the fact that speech recognition is found to be the utmost issue, which affects the decoding of speech. In order to overcome these issues, different statistical models were developed by the researchers. Some of the most prominent statistical models include acoustic model (am), language model (lm), lexicon model, and hidden markov models (hmm). The research will help in understanding all of these statistical models of speech recognition. Researchers have also formulated different decoding methods, which are being utilized for realistic decoding tasks and constrained artificial languages. These decoding methods include pattern recognition, acoustic phonetic, and artificial intelligence. It has been recognized that artificial intelligence is the most efficient and reliable methods, which are being used in speech recognition.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

When we talk about online learning or e-education, we frequently talk about how technology is changing and how the education sector is changing quickly. The significant changes occurred around the internet learning were around Coronavirus while the development of web based learning returns to 90's. When talking about online learning, technology and the pedagogy used by various teachers and facilitators are discussed most frequently. However, motivation is the only factor that has an effect on students' learning patterns, efficiency, and engagement. (Dr. Anju Mahendru, 2022)

Nearly every industry will benefit from Artificial Intelligence (AI) as a result of Industry Revolution 4.0. Many industries are already reaping the benefits of AI, which has been continuously enhancing students' learning abilities and performance. It has likewise gained a vital headway in schooling area. the widespread use of voice-assistance AI tools by a significant number of management students. It makes sure that there is no variety in the mindfulness level of the understudies in view of segment characterization. (Khadse D. K., Awareness & Applications of Artificial Intelligence Tools for Management Students, (2020))

In point of fact, virtual and expanded truths are increasingly making their way into our newsfeeds on account of significant bets placed on high-value returns by colossal corporations such as Facebook, Google, Samsung, and a number of others. Expanded reality (AR) has turned into one more well-known articulation in the high level world, and difficult to find someone's rarely known about this very front development. Currently, augmented reality is used in a variety of industries, including online gaming and education and medical services. (Khadse D. K., Exploratory study of Augmented Reality SDK'S & Virtual Reality SDK'S, 2022)

An enormous piece of the new school graduates these days are occupied with Person to person communication Locales. In this paper, we have done a student study for Person to person communication Locales (Sns') utilization and its relating influences and besides examined this data. (Khadse D. K., 2020)

Research problem

To study and analyze the use of speech recognition systems and its benefits across various demographics

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

Research objective

Humans can speak 150 words per minute on average, but can only type 40. So using speech recognition we can get the work done almost 4 times faster. As the industry makes this shift from text to speech, through this project we want to determine the use of this technology among people of different demographics.

METHODOLOGY

3.1 Sampling design

Stratified or mixed sampling where the population embraces a number of distinct categories, the frame can be organized by these categories into separate "strata." each stratum is then sampled as an independent sub-population, out of which individual elements can be randomly selected .In statistics, stratified sampling is a method of sampling from a population. We have tried to include units representing a fair share from all the above mentioned demographic factors. The sample of population understudy varies from the ages 8 to 56. They have been categorized in 4 categories.

3.2 Data analysis

We have used SPSS software to analyze the data that has been collected. SPSS is a widely used program for statistical analysis in social science. It is also used by market researchers, health researchers, survey companies, government, education researchers, marketing organizations, data miners, and others. The original SPSS manual (nie, bent & hull, 1970) has been described as one of "sociology's most influential books" for allowing ordinary researchers to do their own statistical analysis. In addition to statistical analysis, data management (case selection, file reshaping, creating derived data) and data documentation (a metadata dictionary was stored in the datafile) are features of the base software.

Statistics included in the base software:

- Prediction for numerical outcomes: linear regression

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

- Prediction for identifying groups: factor analysis, cluster analysis (two-step, k-means, hierarchical), discriminant
- Geo spatial analysis, simulation
- R extension(gui)

3.3 HYPOTHESIS

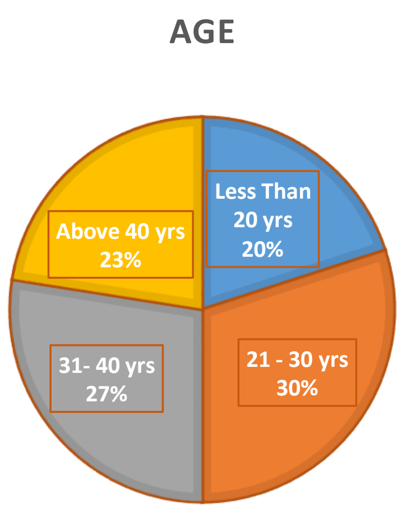
First we have performed one way ANOVA to see if education have any effect on use of different speech recognition system softwares. The hypothesis for the above statement is

H0: education doesn't have any effect on use of different speech recognition system

H1: education has effect on use of different speech recognition system

DATA PRESENTATION & ANALYSIS

1. Age

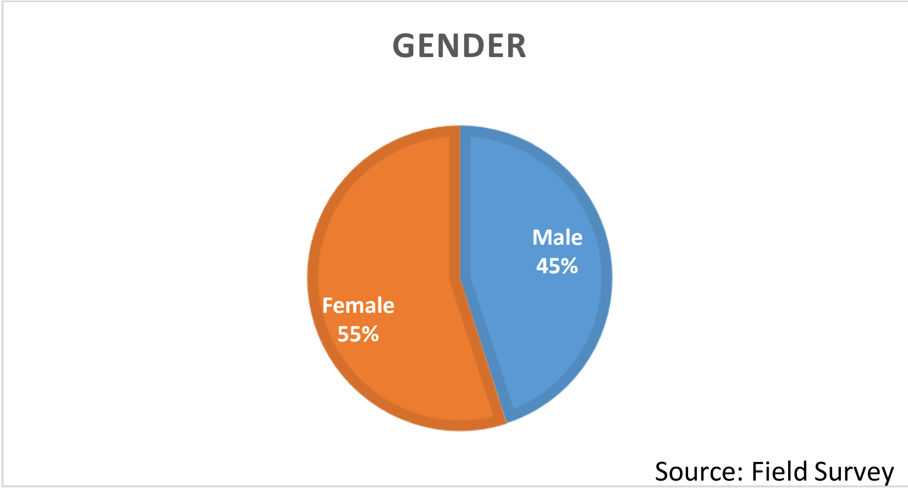


Source: Field Survey

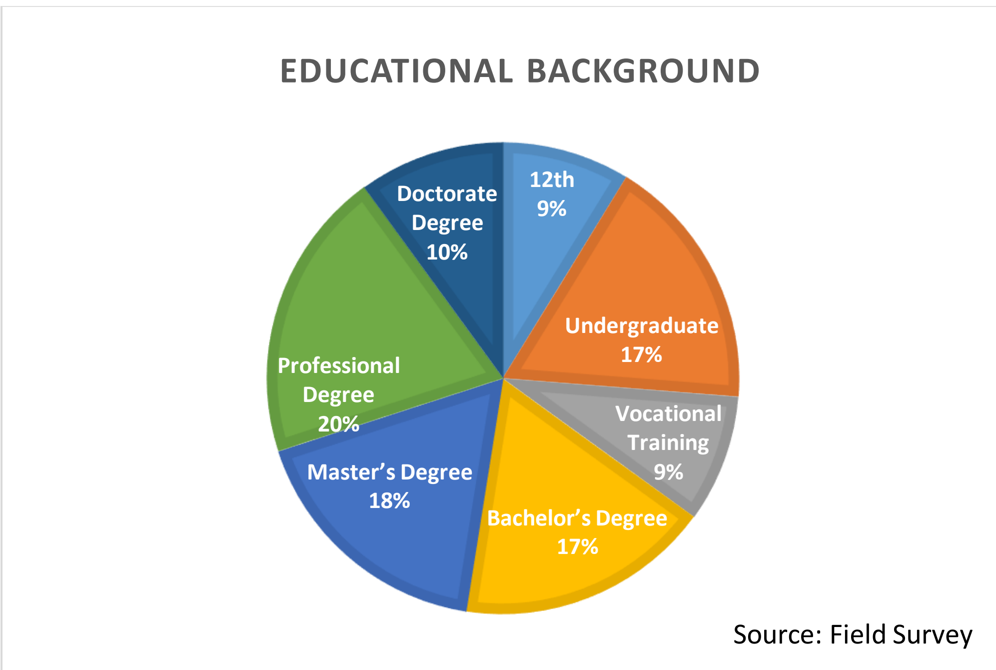
2. Gender

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022**



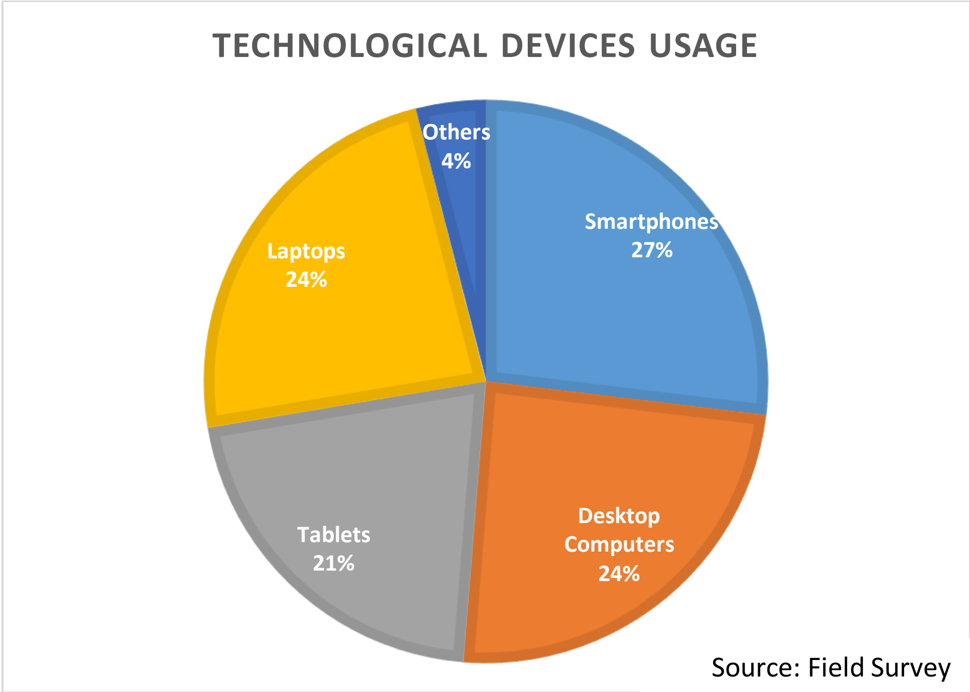
3. Educational background



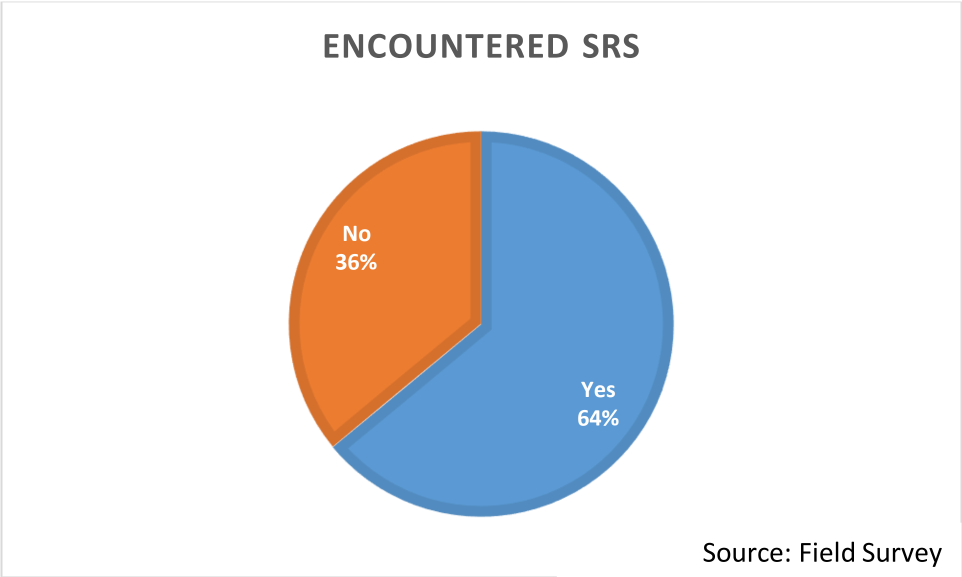
4. Technological device/s Utilization:

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022



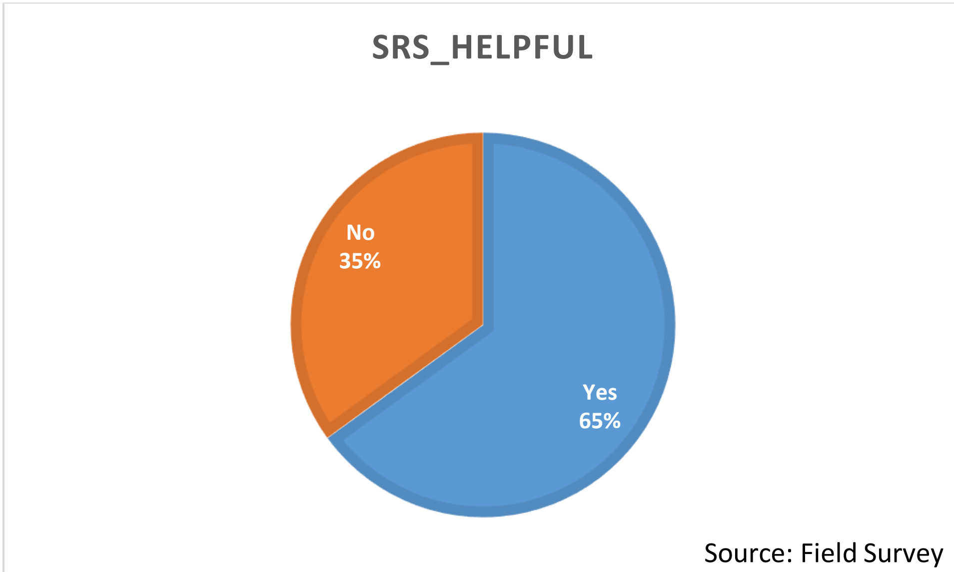
5. Speech Recognition System Encountered:



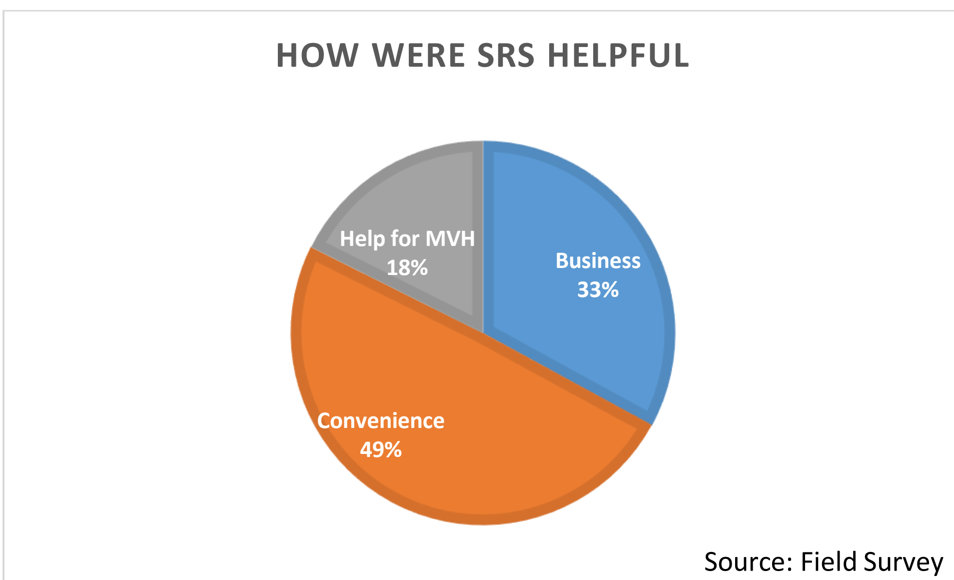
ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022**

6. Speech Recognition Systems helpful or not:



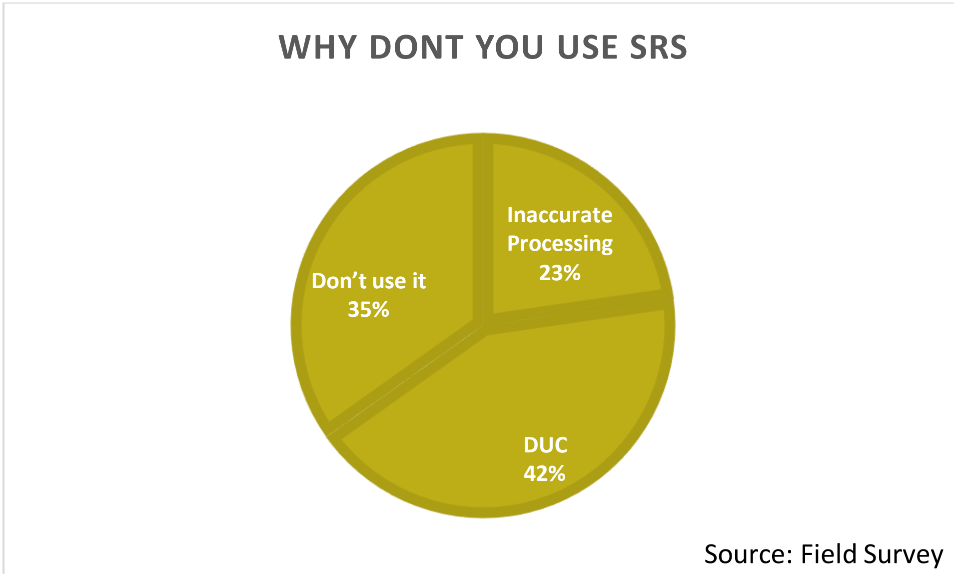
7a. Whether SRS helpful if yes and what way?



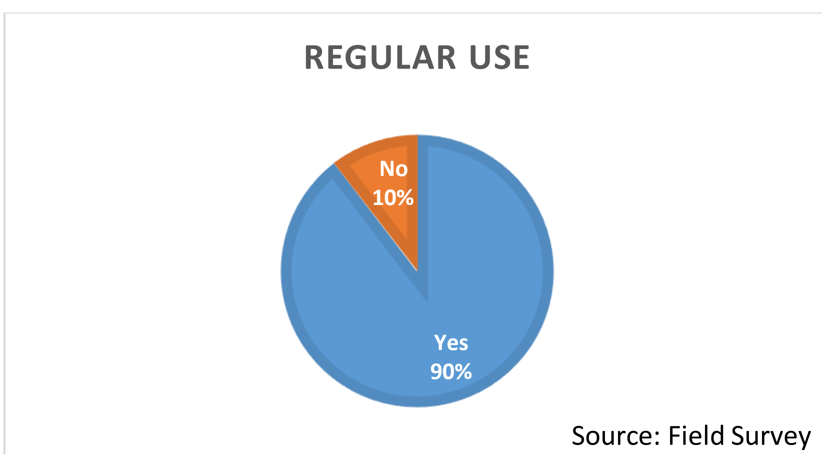
ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

7b. If 'No', why not?



8. Do you use the Speech Recognition Systems of your devices regularly?



ISSN PRINT 2319 1775 Online 2320 7876
Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

The result obtained for the stated hypothesis is as below:

Test of homogeneity of variances

	Levene statistic	Df1	Df2	Sig.
Apples_siri	.829	6	73	.551
Google_now	1.574	6	73	.167
Microsoft_cortana	2.090	6	73	.065
Amazon_alexa	2.350	6	73	.039
Hound	3.247	6	73	.007

Anova

		Sum of squares	Df	Mean square	F	Sig.
Apples_siri	Between groups	7.152	6	1.192	.940	.472
	Within groups	92.536	73	1.268		
	Total	99.688	79			
Google_now	Between groups	5.366	6	.894	.747	.614
	Within groups	87.384	73	1.197		

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

	Total	92.750	79			
Microsoft_cortana	Between groups	10.571	6	1.762	1.068	.390
	Within groups	120.429	73	1.650		
	Total	131.000	79			
Amazon_alexa	Between groups	14.736	6	2.456	1.069	.389
	Within groups	167.652	73	2.297		
	Total	182.388	79			
Hound	Between groups	27.888	6	4.648	1.448	.208
	Within groups	234.312	73	3.210		
	Total	262.200	79			

The test of homogeneity of variances is done using levene's test. The procedure to perform levene's is test:

1. Calculate each $z_{ij} = jy_{ij} \quad \bar{y}_i j$
2. Run an anova on the set of z_{ij} values.
3. If p-value, reject h_o and conclude the variances are not all equal

As we can see from the test of homogeneity of variances table there is a significant relation between education and use of Microsoft Cortana, Amazon Alexa & Hound while Apple's Siri and Google now are independent.

Secondly, we check if there is any correlation between age and use of SRS for various parameters

Correlations

		Age	Business	Convenience	Help_for_mvh	Others_used
Age	Pearson correlation	1	.157	-.455**	.126	.[a]
	Sig. (2-tailed)		.163	.000	.265	.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

	N	80	80	80	80	80
Business	Pearson correlation	.157	1	.173	.185	.[a]
	Sig. (2-tailed)	.163		.124	.101	.
	N	80	80	80	80	80
Convenience	Pearson correlation	-.455**	.173	1	.200	.[a]
	Sig. (2-tailed)	.000	.124		.075	.
	N	80	80	80	80	80
Help_for_mvh	Pearson correlation	.126	.185	.200	1	.[a]
	Sig. (2-tailed)	.265	.101	.075		.
	N	80	80	80	80	80
Others_used	Pearson correlation	.[a]	.[a]	.[a]	.[a]	.[a]
	Sig. (2-tailed)	.	.	.	.	
	N	80	80	80	80	80

**. Correlation is significant at the 0.01 level (2-tailed).

As we can see age is highly correlated with use for help for motor, verbal and/or hearing difficulties and has slightly less correlation with use for business.

CONCLUSION

Analyzing the field survey data we understand that Speech Recognition Systems are gaining popularity across all demographics. It is seen that as the technology is progressing, SRS will be the way ahead as it proves to be an easier alternative than texting and typing.

Bibliography

Agarwal, D. K. (2020). To Study Role and Applications of Natural Language Processing in Business and Education. *Quest journal of Management Research, CRKIMR, volume XI, Issue II, September 2020*, ISSN 0926-2000, Indexed in ProQuest.

Alhawiti, K. (2015). Advances in Artificial Intelligence Using Speech Recognition. *World Academy of Science, Engineering and Technology International Journal of Computer and Information*

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022

Engineering , World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering Vol:9.

Anusuya, M. (2009). Speech Recognition by Machine. *(IJCSIS) International Journal of Computer Science and Information Security*, Vol. 6, Mysore, India.

Bulyko, O. &. (n.d.). The Impact of Speech Recognition on Speech Synthesis. *Washington, WA*.

Dr Kavita Khadse & Manjiri Raut, H. D. (23rd February 2020). Nation Branding: Building Brand Image of ASEAN Countries through Tourism . *International Conference on "Harnessing India's Resources to make India Self-Reliant in collaboration with ASEAN countries*, Quest Journal of Management Research, ISSN 0976-2000. , Volume XII Issue I, CRKIMR, Indexed in ProQuest.

Dr. Anju Mahendru, D. M. (2022). The Impact of Self efficacy on students engagement in online learning mediating the Role of Motivation. *UGC Care Group II, Scopus Indexed journal, Journal of Positive school of Psychology*, ISSN- 2717-7564, Vol.6, No.6(2022).

Dr. Kavita Khadse & Aditya Nijap, U. U. (19th January 2019). "Business Agility: Industries Adapting to Plastic Phase Out" . *Paper was presented in International Conference on "Business Agility: Capabilities and Insights", 19th January 2019. Published in Journal of Management & Research*, ISSN No: 0976-0628, Volume 11 Issue I March 2019, CIMR, Indexed in ProQuest.

Dr. Kavita Khadse & Hardik Mundhada, T. R. (2020). Virtual Water Trade: An assessment of Implementation feasibility in India. *International Conference on "Water secure world", 18th January 2020, Published in Journal of Management & Research*, ISSN No: 0976-0628, Volume II Issue II September 2020, CIMR, Indexed in ProQuest.

Dr. Kavita Khadse & Pranav Pai, A. D. (2019). Business Agility: Artificial Intelligence in Management Education. *Paper was presented in International Conference on "Business Agility: Capabilities and Insights", 19th January 2019. Published in Journal of Management & Research*, ISSN No: 0976-0628, Volume 11 Issue I March 2019, CIMR. .

Dr. Kavita Khadse, A. M. (2022). A study of Implementation, Challenges of Bitris2424 for CRM. *Quest Journal of Management Research, CRKIMR, Indexed in ProQuest*, Pg. No.27-33, published in Volume XIII Issue 1I, December 2022, ISSN 0976-2000.

Dr. Kavita Khadse, D. U. (2020). Indian FMCG: Planning Route to Grow Healthier. *New trends in Research and Innovation Technology, Journal of Research and development, Multidisciplinary International Level , Referred Journal*, ISSN: 2230-9578, Volume 10, Issue 13 with Impact factor of 5.13 .

Dr. Kavita Khadse, K. S. (2022). Understanding the Consumer Behaviour for Online Food Delivery during COVID-19. *Quest Journal of Management Research, CRKIMR Indexed in ProQuest*, Pg. No. 41 – 49, published in Volume XIII Issue 1.

Dr. Kavita Khadse, M. M. (2022). A study on Management of IT Assets of Employees for Vedang Cellular Services Pvt. Ltd. *Quest Journal of Management Research, CRKIMR, Indexed in ProQuest*, Pg. No. 01-15, published in Volume XIII Issue 1I, December 2022, ISSN 0976-2000.

Dr. Kavita Khadse, M. P. (2020). Cloud computing Awareness, Adoption and Usage among Management Students. *International Journal of Concerns, Complexities and Dialogue, Double Blind Peer Reviewed, Multidisciplinary E journal*, Volume I, Issue I , Jan –March 2021.

Dr. Kavita Khadse, S. J. (2022). Business Resilience: Study of Management Student's Perception about Online Trading Platform. *Quest Journal of Management Research*, ISSN 0976-2000 Volume XIII Issue 1, CRKIMR, and Indexed in ProQuest.

Dr. Kavita Khadse, T. J. (2021). Applications of Artificial Intelligence in Digital Marketing for Various Sectors like E-commerce, IT and Food Chain with Reference to ASEAN Countries as India and Singapore . *Paper was presented in International Conference on "Harnessing India's Resources to make India Self-Reliant", on 23rd February 2021,", Published in Journal of Management & Research*, ISSN No: 0976-0628, Volume XIII Issue I March 2020, CIMR, like 40 – 55, Indexed in ProQuest.

Dr. Makarand Upadhyaya, D. K. (2022). Effect of Covid-19 Lockdown on Employees and Environment. *UGC Care Group II, Scopus Indexed journal , INTERNATIONAL JOURNAL OF SPECIAL EDUCATION*, ISSN 0827-3383, Vol.37, No.3, 2022.

Dr. Ranit Kishore, D. K. (2022). Impact of Employee Engagement on Turnover Intention in the Context of Hospitality Industry . *UGC Care Group II, Scopus Indexed journal, INTERNATIONAL JOURNAL OF SPECIAL EDUCATION*, ISSN 0827-3383, Vol.37, No.3.

Gaikwad, G. &. (2010). A Review on Speech Recognition Technique,. *International Journal of Computer Applications (0975 – 8887) Volume 10, Aurangabad, India*, (0975 – 8887) Volume 10, Aurangabad, India.

Janvale, O. &. (2015). Automatic Speech Recognition and Verification using LPC, MFCC and SVM. *International Journal of Computer Applications* , (0975 – 8887) Volume 127, Aurangabad, India.

Jethva, V. &. (2016). A Survey on Feature Extraction and Classification Techniques for Speech Recognition. *IJARIIE*, ISSN(O) Vol-2, Ahmedabad, India.

Kaur, S. &. (2013). Automatic Speech Recognition. *International Journal of Engineering Trends and Technology*, Volume 4, Haryana, India.

Kavita Khadse, D. R. (2014). Sustainable Inclusive Growth in Management Education through Adoption of Instructional Technologies. *National conference on "Sustainable Inclusive Growth through Socially Responsible Enterprises", 22 March 2014, published in Volume VI, Issue 9, September 2014*, ISSN 0976-2000, Quest Journal of Management Research, CRKIMR.

Kavita Khadse, D. R. (2015). India Vision 2020-How the Higher Education in India Will Change? *National Conference on "India Vision 2020: Entrepreneurial Opportunities and Management Challenges*, 20 March 2015, Published in volume VI, Issue I, March 2015, ISSN 0926-2000, Quest journal of Management Research, CRKIMR.

Kavita Khadse, D. R. (2016). Flipped Classrooms Leveraged Management Education in India. *International Conference on "Diversity: Leveraging the Differences", 19th March 2016, Published*

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 10, 2022**

in volume VIII, Issue II, September 2016, ISSN 097676-0628, Journal of Management Research, CIMR .

Khadse, D. K. ((2020)). Awareness & Applications of Artificial Intelligence Tools for Management Students. *hycology & Education Journal, ISSN 00333077, UGC Care Group II, Scopus Indexed journal, Volume 57 No. 9* , UGC Care Group II, Scopus Indexed journal, Volume 57 No. 9 .

Khadse, D. K. (2020). To Explore the Effect of Social Networking Sites on Students Academic Performce. *UGC Care Group II, Scopus Indexed journal, Turkish Journal of Computer and Mathematics Education*, ISSN: 1309-4653,Vol.11No.3(2020), 856-866.

Khadse, D. K. (2022). Exploratory study of Augmented Reality SDK'S & Virtual Reality SDK'S. *International Conference on "Unlocking the Potential of Post-Covid Transformations in Commerce & Management for Economic Development and Sustainability" on 24th April 2021*, Published in UGC Care Group II, Scopus Indexed journal, Palarch's Journal of Archaeology of Egypt/Egyptology, ISSN 1567-214X, Volume -18, Issue - 7, 2021, pg 2208-2222.

Khadse, K. (2017). Applications of Big Data in Management Education. *International conference on "Business & Society: Value Creation through Analytics", 19th December 2017, Published in Volume IX, Issue 12, December 2017*, ISSN 0976-0628. Journal of Management Research, CIMR. .

Novotny, S. &. (2004). Analysis and Optimization of Telephone Speech Command Recognition System Performance in Noisy Environment. *RADIOENGINEERING*, VOL. 13, Praha, Czech Republic.

Patil, D. K. (2023). A STUDY ON FORECASTING CPI INFLATION IN INDIA USING MACHINE LEARNING ALGORITHM MODEL. *UGC-CARE List Group I, JOURNAL OF THE ASIATIC SOCIETY OF MUMBAI*, ISSN: 0972-0766, Vol. XCVI, No.20, 2023, Page No 79-83.

Reddy, E. &. (1973). A Model and a System for Machine Recognition of Speech. *IEEE Vol. AU-21, Pittsburgh, PA.*, IEEE Vol. AU-21, Pittsburgh, PA.

Sharma, D. K. (2019). Study of factors influencing Management Faculties adoption of Instructional Technology and Designing Conceptual Model. *National Conference of Maharashtra state Commerce Association on "Commerce and Management in 21st Century, Published in International Research Fellow Associations, Research Journey, UGC approved Journal, , ISSN - 2348-7143*, January 2019.

****