

# Baby Cry Classification Using Machine Learning

P.Ithaya Rani <sup>1</sup>, P.Pavan Kumar <sup>2</sup>, V.Moses Immanuel <sup>3</sup>, P.Tharun <sup>4</sup>, P.Rajesh <sup>5</sup>

Associate Professor, CSE , Koneru Lakshmaiah Education Foundation (KLEF)

Deemed to be University, Green Fields, Vaddeswaram, Guntur District, A.P., INDIA.

B.Tech Student, CSE , Koneru Lakshmaiah Education Foundation (KLEF) Deemed to be University,  
Green Fields, Vaddeswaram , Guntur District, A.P., INDIA.

**Abstract:- A Cry is a type of correspondence for kids to communicate their sentiments. Child cry can be portrayed by its regular occasional tone and the difference in voice. Through their child's cry discovery, guardians can screen their child somewhat just in significant conditions. Recognition of a child cry in discourse signals is a urgent advance in applications like remote child observing and it is likewise significant for researchers, who concentrate on the connection between child cry signal examples and other formative boundaries. This investigation of sound acknowledgment includes highlight extraction and arrangement by deciding the sound example. We use MFCC as an element extraction strategy and K-Nearest Neighbor (K-NN) for arrangement. K-Nearest Neighbor (KNN) is a characterization technique that is regularly utilized for sound information. The KNN classifier is displayed to yield extensively better outcomes contrasted with different classifiers.**

## I. INTRODUCTION

Cry signals or cry designs are under research examination for a long time. Researchers and examiners observed that the cry signs can give definite pictures about the physical and mental conditions of babies.

There are a few methodologies for this issue. One of the strategies to distinguish sleepiness is by utilizing facial highlights. This is a picture based framework in which the looks are considered by Feature extraction. For the most part, individuals shut their eyes or yawn when they are feeling tired. By thinking about these elements, we can anticipate the condition of the driver.

The dataset comprises of two arrangements of pictures, ready and tired. The framework configuration comprises of the fo signals or cry designs are under research examination for quite some time. Researchers and experts observed that the cry signs can give itemized pictures about the physical and mental conditions of babies.

From the exploration of WHO, consistently, almost 40% of newborn child passings are occurring in their underlying 30-50 days of life, 72% of baby passings occur inside the principal seven day stretch of their introduction to the world, and up to 2/3 of newborn child lives will be saved if the reason is seen significantly sooner.

The strategies which permit us to detect the past indications of baby wellbeing and cleanliness can assist us with diminishing newborn child passing rate. To be exact this can be the better objective of our proposition is than create or carry out a dependable framework that grants us to comprehend sicknesses dependent on cry sound assessment. Improvement of such a kind of framework at first notices the matter track down the solid cry parts or examples in an info waveform. The NCDS framework likely befuddles if the information discourse signal contains pointless commotions separated from the cry signal alone. Consequently, the best test in planning and carrying out a symptomatic framework is to execute a programmed distinguishing machine to precisely look through the inspiratory and expiratory pieces of a cry design. After bunches of examination on sicknesses and cry signals and their connection some valuable outcomes ended up creating programmed sound division of expiratory and inspiratory pieces of newborn child cries.

In case we are prepared to fragment sound cry flags and look at imperative pieces of a pre recorded sound sign, it very well may be extremely useful and less difficult to foster an absolutely programmed framework which helps in understanding sicknesses this strategy can be acclimated support our choices on understanding the newborn child cries. Through this we can decide the indications prior and make essential strides at a proficient and modest expense.

The new investigations on newborn child cries have shown that babies are weeping for quite some time upheld their prerequisite like yearning, weariness, awkward sentiments, torment, and different reasons. Scientists and understudies like pediatricians clinical specialists can relate between different types of baby cries and subsequently pre gauge the newborn child's prerequisite by utilizing the cry sounds, motions and other conduct delivered by the newborn child nonetheless, this can be a genuine time extensive issue (issue) for the fogeys not really talented to those that can't take appropriate consideration of the child. This task gives a programmed strategy to baby cry characterization which is prepared to utilize an informational index of 8 distinct newborn child cries.

Henceforth the most goal is to extricate valuable highlights from the cry sound sign i.e., the newborn child cry and test the obscure cry signal with the characterized mentor and know the significance of the baby cry, in this way taking Care for the baby in like manner.

## II. LITERATURE SURVEY

There are different strategies for recognizing laziness. A portion of the methodologies which are utilized in this space are talked about here.

### A. Parts of newborn child cry sound sign

The significant and key parts of a newborn child cry sound sign are motivation and termination leaves behind vocalization and hear-capable that can successfully look through INSV, EXP precisely inspiration (INSV) and lapse (EXP). The essential difficulties looked in this kind of framework is carrying out a technique inside an two essential and significant techniques: individual cry signal. The issue of cry identification is not quite the same as unvoiced, voiced division on the grounds that an average hear-capable newborn child cry sound sign contains each of the unvoiced and voiced parts[1][2].

### B. Voice Activity Detection (VAD)

The writing proposed by Kuo, 2010, has expressed that the principle issue is recognition of cry sound utilizing a framework recorded in a ton of uproarious homegrown nearby climate is too difficult to possibly be settled by VAD (Voice Activity Detection) modules, VAD manages the issue of looking or observing discourse designs from other hear-able dynamic districts of a considered sound sign. The other hear-able dynamic examples might be any sort like quiet, commotion or a doorbell cautioning. The Signal to Noise Ratio "SNR" is a vital boundary and it may bring about a ton of undesirable mistakes. VAD is essential in a few sound correspondence frameworks like programmed discourse acknowledgment, phones, other advanced assets, and transmission of discourse in genuine time[3]. A portion of the Common and broadly utilized VAD strategies contain Feature Extraction and Decision making. Highlights of a sign that permits calculation of energy, cepstral coefficients, ZCR, Marzinzik and Kollmeier, proposed range investigation in 2002, Wang and Tasi, proposed wavelet advertisement entropy changes in 2008 ; Juang et al., proposed choice guideline calculation dependent on outline by-outline and exceptionally straightforward principles for thresholding. In 2009, After applying natural VAD calculations from Rabiner-Sambur and G.729b technique in recognition of cry signal parts or fragments.

### C. Discoveries

The discoveries were:

- It is difficult to choose the edge settings in a loud homegrown climate.
- While information securing, the Traditional VAD module can't separate between EXP and [6] INSV (cry signal fragments) and recorded discourse signal sections.
- Traditional VAD modules can't recognize lapse (EXP) from motivation (INSV) portions of a cry sound signal[5]

Factual methodology is a decent answer for abstain from limiting the issue of changing edges. That is the reason due thought is given to factual model-based methodologies proposed by AbouAbbas in 2015b, 2015c modules.

### D. Existing framework

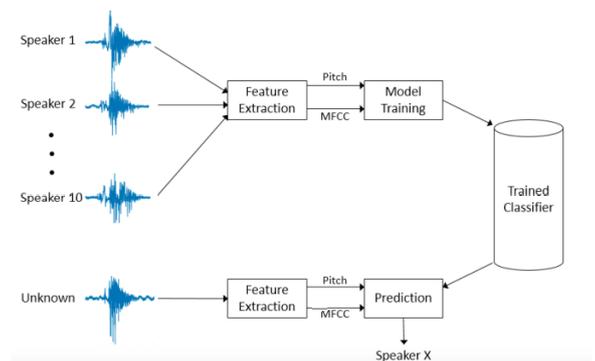
There are frameworks to identify whether or not a sound record gave is a Baby cry. The procedures utilized LFCC (Linear Frequency Cepstral Coefficients) for include extraction. There is likewise a framework to group the purposes behind child cry and in this framework different classifiers are utilized to order the reasons from the pre-characterized informational collection.

## III. PROPOSED METHODOLOGY

### A. The Dataset

The initial step is assortment of the dataset. We thought about a portion of the datasets for this framework. A dataset which contains sound records of child cries gathered by discourse research foundations specifically give a-cry-corpus has been gathered. This dataset contains sound examples of numerous newborn children caught under various circumstances more than a few circumstances. This dataset contains 8 arrangements of sound records i.e., Awake, Belly torment, Burping, Discomfort, Hug, Hungry, Sleepy and Tired.

### B. System Architecture



### C. Feature Extraction

From the get go, the information is taken. The info is normally a sound document which is to be handled further. The discourse signal contains a larger than average number of information which mirrors the enthusiastic attributes, sex grouping and hence the speaker's personality. Each discourse and speaker has extraordinary individual attributes which are implanted in their discourse expressions. Element extraction from the discourse signal is the sign handling forepart process which changes over the human discourse into some valuable parametric portrayal. Element extraction assumes a significant part inside the general presentation of a discourse acknowledgment besides as a speaker acknowledgment framework. a legitimate element extraction strategy should catch the significant qualities of the sign and may likewise dispose of some superfluous properties. Element extraction is the method involved with keeping helpful data from the sign while disposing of excess and undesirable data . Component boundaries assume a major part to separate discourses moreover as speakers from each other. These boundaries are helpful for examination in different discourse united applications like discourse acknowledgment, speaker acknowledgment, discourse combination and discourse

coding. In this way, we attempted to remove some conspicuous qualities from the discourse to fulfill the presentation of our proposed framework.

### C. Mel recurrence cepstral coefficients(MFCC)

```
[[ -83.377686  33.08731  -42.55658  22.141876  -12.549919
  23.294508  -22.59206  14.99546  1.4717332  4.209373
  -0.74977094 -14.92513  0.9372652  8.224226  -0.12192553
  13.989015  6.5319715  1.9878587  -8.569176  -6.5270405
  0.10054475  4.7355175  -4.119027  -3.9245353  -5.563842
  3.058156  -2.8110352  -2.5545697  -0.23786436  7.003273
  3.155918  -1.0060205  -2.7972405  4.107716  -0.52387434
  -1.1573135  -1.2756749  1.5718873  -3.6838982  -0.8891317 ]]
```

MFCC is one in all of the principal famous component extraction strategies utilized in programmed discourse or speaker acknowledgment frameworks utilizing the Mel scale which depends on the human ear scale[11]. It's upheld the nonlinear human impression of the recurrence of sounds. These coefficients address sound upheld discernment. They're gotten from the Mel recurrence cepstrum.[14] The otherworldly data can after that be changed over to MFCC by going the signs through band pass channels where higher frequencies are misleadingly supported, so applying a backwards Fast Fourier Transform (FFT) consequently. It consolidates the advantages of the cepstrum examination with a perceptual recurrence scale upheld basic bands.[17] thus, the upper frequencies are turning out to be more conspicuous. Since the Mel recurrence cepstrum can address an audience's reaction framework obviously, hence MFCC is normally viewed as the least complex accessible estimation of the human ear.

## IV. ALGORITHM USED

### A. K-Nearest Neighbor(KNN)

K-Nearest Neighbor is one in everything about best Machine Learning calculations upheld Supervised Learning technique. The K-NN calculation accepts the similitude between the new case/information and accessible cases and places the new case into the classification that is generally similar to the accessible categories. K-NN calculation stores every one of the accessible information and groups a spic and span data upheld the likeness. This infers when new information shows up then it is effectively arranged into an appropriate class by utilizing K-NN calculation. The K-NN calculation will be utilized for Regression likewise concerning Classification however generally it's utilized for the Classification issues. K-NN could be a non-parametric calculation, which infers it doesn't make any presumption on hidden information. It is moreover called a lethargic student calculation since it doesn't gain from the preparation set promptly rather it stores the dataset and at the hour of characterization, it plays out an activity on the dataset.

After the cepstral coefficients are separated from the sound example, they are put away in a cluster and its mean is determined. At the point when the model is applied on the information, as indicated by the worth of k, the arrangement classification is chosen and is given as yield.

### B. Innocent Bayes

The Naive Bayes calculation could be a managed learning calculation, which is predicated on Bayes hypothesis and utilized for tackling grouping issues. It's primarily utilized in text arrangement that has a high-dimensional preparing dataset. Innocent Bayes Classifier is one among the simple and best Classification calculations which helps in building the quick AI models which will make speedy predictions. It is a probabilistic classifier, which suggests it predicts on the possibility of the likelihood of an item. Some well known examples of Naive Bayes Algorithm are spam filtration, Sentimental examination, and arranging articles.

### C. Backing Vector Machine

Backing Vector Machine or SVM is one among the principal well known Supervised Learning calculations, which is utilized for Classification furthermore as Regression issues. In any case, fundamentally, it's utilized for Classification issues in Machine Learning. The objective of the SVM calculation is to shape the best line or choice limit which will isolate n-dimensional space into classes so we can undoubtedly put the new data inside the right classification inside what's to come. This best choice limit is named a hyperplane. SVM picks the extreme focuses/vectors that assistance in making the hyperplane. These outrageous cases are called support vectors.

## V. RESULT

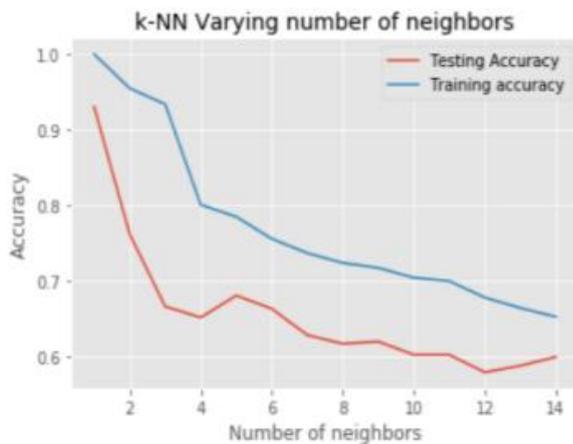
### A. Mel Frequency cepstral Coefficients(MFCC)

These are the cepstral coefficients gotten when an arbitrary sound record is utilized for testing. There are 40 cepstral coefficients altogether and each worth is unique. There are positive and negative values.[14] Positive worth of the cepstral coefficient suggests that the main part of ghastly energy is amassed in low recurrence areas. Negative worth of the cepstral coefficient implies the otherworldly energy is packed in high recurrence regions.[15][17] Here, the measure of cepstral coefficients is 40 since it yields better outcomes.

### B. Correlation among KNN, Naive Bayes and SVM

The precision got when Naive Bayes is applied on the information is 45% and on account of SVM, the exactness is 42% and the most noteworthy precision that is acquired is 76.16% when KNN is utilized for grouping.

### C. K worth in KNN process



The correlation aftereffect of the K incentive for the grouping system is displayed in fig.3. The most elevated precision acquired is 76.16% when utilizing K worth is 2. In addition, the littlest precision acquired is in the middle of half and 60% when the K worth is in the middle of 4 and 10. This can be impacted by the preparation information and testing information. For this framework, the most elevated noticed precision is 76.16% and is seen when the K worth is 2.

## VI. RELATED WORK

It is a crucial task to discriminate the infant cries, so in this work (6), dealt with K-NN classifier with features such as short-time energy, harmonic to average power ratio (HAPR), Mel frequency coefficient, and harmonicity factor (HF) to recognize the infant cry sounds. In this work (7), convolutional restricted Boltzmann machine was used to analyze the unsupervised auditory filter banks. The network consists of the visible and hidden layers, and the weights were shared between those layers. The non-linear activation of Noisy Leaky Rectifier Linear Unit (NLReLU) was used. The parameters of the network were optimized by using the Adam optimization method. Convolutional restricted Boltzmann machine and discrete cosine transform were applied to reduce the feature dimensions. Those features were compared with MFCC features, and it was found that CNBM-based feature performs well in the discrimination of healthy and pathological auditory cries. In this case (8), they employed a convolutional neural network in infant cry vocalizations. The cry segments were manually extracted from the audio signal and segmented into a 4–8-s duration of segments. Audio signals were represented as spectrogram through short-time Fourier transform, which is based on Fourier transform. The spectrogram is the input for convolutional neural network. The convolution layer can obtain the features from the spectrogram, and the network can successfully discriminate the baby cry vocalizations. or removing essential data without lawful authorization.

## VII. CONCLUSION

The main aim is to take an audio file as an input document as an info and arrange the justification for the cry from the eight diverse order marks previously characterized. The sound documents are recorded under loud conditions and from various cases. Through this task, we took in the significance of preprocessing for highlight extraction and distinguish the importance of newborn child cries with the goal that it very well may be a contributing variable to deal with the babies appropriately. By utilizing KNN classifier we acquired 76% precision at k=2 which characterizes and gives the right explanation in larger part of the occasions. Python above 3.6 was utilized to implement. This proposed framework takes a sound document as an information and distinguishes the justification for cry. This same thought can be utilized in fields like remote child controlling and in clinical applications to comprehend the justification behind child cry with the goal that they could deal with them as needs be.

## REFERENCES

- [1]. J. Saraswathy, M. Hariharan, Sazali Yaacob and Wan Khairunizam, "Automatic Classification of Infant Cry:A Review", School of Mechatronic Engineering, Universiti Malaysia Perlis (UniMAP),02600 Perlis, Malaysia,2011. Kevin Kuo, "Feature extraction and recognition of infant cries". 2018.
- [2]. Jose Orozco Garcia, Carlos A. Reyes Garcia," Mel-Frequency Cepstrum Coefficients Extraction from Infant Cry for Classification of Normal and Pathological Cry with Feed-forward Neural Networks", Instituto Nacional de Astrofisica Optica y Electronica (INAOE) Luis Enrique Err 23 Page 11-25 © MANTECH PUBLICATIONS 2019. All Rights Reserved Journal of Artificial Intelligence, Machine Learning and Soft Computing Volume 4 Issue 1 I,Tonantzintla, Puebla, Mexico.
- [3]. Jose Orozco, Carlos A. ReyesGarcia," Implementation and Analysis of Training Algorithms for the Classification of Infant Cry with Feed-forward Neural Networks", Instituto Nacional de Astrofisica Optica y Electronica (INAOE) Luis Enrique Err # 1,Tonantzintla, Puebla, Mexico,2018.
- [4]. Pritam Pal, Ananth N. Iyer and Robert E. Yantorno, "Emotion Detection From Infant Facial Expressions And Cries" Electrical & Computer Engineering Department, Temple University 12th & Norris Streets, Philadelphia, 2019.
- [5]. Ioana-Alina Bănică, Horia Cucu, Andi Buzo, Dragoş Burileanu, and Corneliu Burileanu," Baby Cry Recognition in Real-World Conditions", Speech and Dialogue Research Laboratory, University Politehnica of Bucharest, Romania, 2019.
- [6]. Andi Buzo, Dragos Burileanu, Corneliu Burileanu," Automatic Methods for infants cry classification", Speech and Dialogue Research Laboratory, University Politehnica of Bucharest, Bucharest, Romania 2019.

- [7]. Chuan-Yu Chang, Jia-Jing Li, “Application of Deep Learning for Recognizing Infant cries”, Chuan-Yu Chang, Jia-Jing Li National Yunlin University of Science & Technology, Taiwan, 2018.
- [8]. J.Saraswathi, N.Thiyagar, M.Hariharan, “Infant Cry Classification: Time Frequency Analysis”, International conference on control system, computing and engineering, Malaysia, 2020.
- [9]. Shubham Asthana, Naman Varma, Vinay Kumar Mittal, “Preliminary Analysis of causes of Infant Cry”, Electronics and Communication Engineering 1, 2 The LNM Institute of Information Technology 1, 2 Rupa Ki Nangal, Post-Sumel, Via Jamdoli, Jaipur 302 031, Rajasthan, India 3 Indian Institute of Information Technology 3 Sri City, Satyavedu Mandal, Chittoor District 517 588, Andhra Pradesh, India, 2018.
- [10]. Mrs. Sameena Bano, Dr. K.M. RaviKumar, “Decoding Baby Talk: A Novel Approach for Normal Infant Cry Signal Classification”, 2015 International Conference on Soft-Computing and Network Security (ICSNS - 2015), Feb. 25 – 27, Coimbatore, INDIA, 2019.
- [11]. Suryakanth V.Gangashetty, Pruthvi Raj Myakala, Rajashree Nalumachu, Shivam Sharma, VK Mithal, “Acoustic Analysis of Infant Cry Signal Toward Automatic Detection of the cause of Crying”, Indian Institute of Information Technology Chittoor, Sri City, A.P., India 1, 2, 3 International Institute of Information Technology, Hyderabad, Telangana, India 4 CEO, Ritwik Software Technologies Pvt. Ltd., Hyderabad, India 5, 2020.
- [12]. Pruthvi Raj Myakala, Rajashree Nalumachu, Shivam Sharma, VK Mithal, “An Intelligent System for Infant Cry detection and information in real time”, Speech and Dialogue Research Laboratory, University Politehnica of Bucharest, Bucharest, Romania, 2020.
- [13]. Aomar Osmani, Massinissa Hamidi, Abdelghani Chibani, “Machine Learning Approach for Infant Cry Interpretation”, International Conference on Tools with Artificial Intelligence, 2019.
- [14]. Rodica Ileana Tuduce, Horia Cucu, and Corneliu Burileanu, “Why is my Baby Crying? An in-depth Analysis of Paralinguistic Features and Classical Machine Learning Algorithms for Baby Cry Classification”, Speech and Dialogue Research Laboratory, University Politehnica of Bucharest, Romania, 2018.