International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10s
DOI: https://doi.org/10.17762/ijritcc.v11i10s.7608
***Article Received**: 30 May 2023 **Revised**: 29 July 2023 **Accepted**: 13 August 2023*

Blood Pressure Estimation from Speech Recordings: Exploring the Role of Voice-over Artists

Vaishali Rajput[1], Preeti Mulay[1], Rajeev Raje[2]
[1]Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India
vaishali.kalyankar.phd2019@sitpune.edu.in
preeti.mulay@sitpune.edu.in
[2]Indiana- Purdue University, Indiana, United States
rraje@iupui.edu

Abstract: Hypertension, a prevalent global health concern, is associated with cardiovascular diseases and significant morbidity and mortality. Accurate and prompt Blood Pressure monitoring is crucial for early detection and successful management. Traditional cuff-based methods can be inconvenient, leading to the exploration of non-invasive and continuous estimation methods. This research aims to bridge the gap between speech processing and health monitoring by investigating the relationship between speech recordings and Blood Pressure estimation. Speech recordings offer promise for non-invasive Blood Pressure estimation due to the potential link between vocal characteristics and physiological responses. In this study, we focus on the role of Voice-over Artists, known for their ability to convey emotions through voice. By exploring the expertise of Voice-over Artists in controlling speech and expressing emotions, we seek valuable insights into the potential correlation between speech characteristics and Blood Pressure. This research sheds light on presenting an innovative and convenient approach to health assessment. By unraveling the specific role of Voice-over Artists in this process, the study lays the foundation for future advancements in healthcare and human-robot interactions. Through the exploration of speech characteristics and emotional expression, this investigation offers valuable insights into the correlation between vocal features and Blood Pressure levels. By leveraging the expertise of Voice-over Artists in conveying emotions through voice, this study enriches our understanding of the intricate relationship between speech recordings and physiological responses, opening new avenues for the integration of voice-related factors in healthcare technologies.

Keywords: Blood Pressure, Voice-over Artists, speech recordings, emotion recognition.

I. Introduction

High Blood Pressure, also known as hypertension, is a prevalent and significant health concern globally. It is associated with various cardiovascular diseases and has been identified as a leading cause of morbidity and mortality [3-4]. To aid in early detection, opinion, and successful management of hypertension, accurate and prompt Blood Pressure monitoring is essential. Traditional Blood Pressure measurement methods involve the use of cuff-based devices, which can be inconvenient, uncomfortable, and require frequent measurements [3]. “Non-invasive” and “continuous Blood Pressure” estimation methods have gained attention as an alternative approach for monitoring Blood Pressure [3-4]. Speech recordings have emerged as a promising avenue for non-invasive Blood Pressure estimation due to the potential connection between vocal characteristics and physiological responses [1,16]. The motivation behind this research lies in exploring the potential of using speech recordings for Blood Pressure estimation and understanding the specific role of Voice-over Artists in this process. Voice-over Artists possess specialized skills in vocal modulation, emotional expression, and controlled speech performance [28]. They are trained to convey a wide range of emotions effectively through their voice and speech, making them an interesting group to study for emotion-related research [28].

II. Significance of the Study

The significance section of the research paper highlights the potential contributions and implications of the study. It explains the value and importance of the research in advancing scientific knowledge, addressing practical challenges, or making a broader impact. Here is an elaboration of the significance of the study:

A. Augmented Accessibility: Estimating Blood Pressure from speech recordings would eliminate the need for traditional cuff-based measurements, making monitoring more accessible, comfortable, and convenient for individuals [1]. The study of speech recordings as a means of Blood Pressure estimation offers the novel way of developing non-invasive, user-friendly, and continuous monitoring tools.

B. Enhancing Personalized Healthcare: By exploring the use of speech recordings for Blood Pressure estimation, the study may contribute to personalized healthcare approaches. It has the potential to provide individuals with the means to supervise their Blood Pressure in instantaneous-time, enabling proactive management and early detection of hypertension-related risks [1,12-13].

C. Bridging Disciplinary Gaps: The research strives to silent the knowledge variance between the fields of speech analysis, emotion detection, and cardiovascular health. By bringing together expertise from multiple disciplines, including psychology, computer science, and biomedical engineering, this study promotes interdisciplinary collaboration and knowledge integration. It encourages researchers and practitioners to explore new avenues for cross-disciplinary research and collaboration in the development of healthcare technologies.

D. Practical Applications: The outcomes of this study have practical applications in various domains. Speech-based Blood Pressure estimation can find applications in telemedicine, remote patient monitoring, and wearable devices [1,16]. It can enable continuous and unobtrusive monitoring of Blood Pressure, facilitating early detection of hypertension-related risks and timely intervention. Additionally, the research findings may have implications in fields such as affective computing, human-computer interaction, and voice-enabled technologies.

III. Voice-over Artist and Emotional Expression

Voice-over Artists play a crucial role in conveying emotions through their vocal performances [28]. They use their voice to bring characters, narratives, and messages to life, effectively communicating and evoking specific emotional responses in the audience [28-29]. Here are some key points to elaborate Voice-over Artists and emotional expression:

A. Vocal Techniques: Voice-over Artists utilize various vocal techniques to express emotions. They adjust their pitch, tone, volume, pacing, and emphasis to convey different emotional states [28]. For example, they may raise their pitch and increase volume to express excitement or use a slower pace and softer tone to convey sadness or calmness.

B. Emotional Authenticity: Skilled Voice-over Artists strive for emotional authenticity in their performances. They tap into their own emotional experiences, imagination, and understanding of the character or message to deliver a genuine and believable portrayal of emotions. They bring nuances, subtleties, and layers of emotional expression to their voice work [29].

C. Interpretation and Intention: Voice-over Artists interpret the emotions depicted in the script or direction and bring their own artistic choices and intention to the performance. They analyse the context, underlying emotions, and character motivations to effectively convey the intended emotional message [28-30]. Their ability to understand and embody emotions enhances the impact of their voice performances [29].

D. Versatility: Voice-over Artists possess a wide range of emotional versatility. They can portray a diverse spectrum of emotions, ranging from joy, sadness, anger, fear, surprise, to more subtle emotions like curiosity, contentment, or anticipation [28-30]. They adapt their vocal delivery to suit different genres, styles, and characters, ensuring the emotional resonance of the content.

E. Impact on Audience: Voice-over Artists have the power to evoke emotional responses in the audience [29]. Their expressive voices can create a strong emotional connection, influencing how the audience perceives and engages with the content.

In summary, Voice-over Artists use their vocal skills, emotional authenticity, interpretation, and collaboration to effectively express a wide range of emotions in their performances. Their ability to convey emotions through voice plays a significant role in engaging and connecting with the audience, making them an essential element in various forms of media and communication [28-30].

IV.Existing Machine Learning algorithm for Emotion Recognition of Voice-over Artist

The choice of the best machine learning algorithm for emotion recognition of Voice-over Artists depends on various factors like the dataset, the complexity of the emotion recognition task, etc. Here are some commonly used machine learning algorithms for emotion recognition:

***A. Support Vector Machines*:** SVM is a popular algorithm for emotion recognition due to its capacity to handle “high-dimensional” feature spaces and its effectiveness in handling non-linear relationships. SVMs aim to find an optimal hyperplane that separates different emotion classes in the feature space [22-23].

B. Random Forest: An ensemble learning system called Random Forest mixes various decision trees to produce predictions. It can handle high-dimensional feature spaces and is robust against overfitting. Random Forests can capture complex relationships between features and emotions, making them suitable for emotion recognition [32].

C. Decision Tree: Decision Trees: Decision Trees are a straightforward yet effective method that classifies data using a hierarchical structure of decision nodes. They can capture intricate correlations between features and feelings and are interpretable [36].

D. K-Nearest neighbors: KNN is a straightforward and understandable algorithm that groups data depending on how similar its feature vectors are. New data points are given a class

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10s
DOI: https://doi.org/10.17762/ijritcc.v11i10s.7608
***Article Received**: 30 May 2023 **Revised**: 29 July 2023 **Accepted**: 13 August 2023*

label based on the labels of their k closest neighbours in the feature space [34].

***E. Naive Bayes*:** It is a probabilistic method that estimates a data point's likelihood of belonging to a certain class based on the values of its features. It bases predictions on the Bayes theorem and assumes that features are independent of one another [35].

In the proposed research work, SVM is selected for emotion recognition.

V. SVM for Emotion recognition of Voice-over Artist-Advantages and Rationale

Research in classifying and identifying emotions in speech based on acoustic characteristics of the individual voice is continuing. When it comes to Voice-over Artists, their vocal performances provide a unique opportunity to explore emotion recognition in a controlled and expressive context.

While SVM can have a higher computational complexity, it has shown excellent performance in various domains, including classification and regression tasks. SVM has several advantages, such as effective handling of high-dimensional data, robustness against overfitting, and the ability to handle nonlinear data through kernel functions [22-23,31,38]. With consideration of these benefits, the proposed research focuses on the use of SVM for emotion recognition. Fig.1 shows key points to recognize emotions of Voice-over Artists from speech.

The steps for performing Emotion Recognition of Voice-over Artists using SVM:

A. Data Collection: Gather a dataset of speech samples of Voice-over Artists, where each sample is labelled with the corresponding emotion category (e.g., happy, sad, angry, etc.)
B. Feature Extraction: Extract relevant features from the speech signals that can capture the emotional content. Commonly used features include "Mel-frequency cepstral coefficients (MFCCs)", prosodic features (e.g., pitch, intensity), and spectral features [15].
C. Data Pre-processing: Normalize the extracted features to ensure they are on a consistent scale. You can use techniques like mean normalization or standardization [37].
D. Dataset Split: From the dataset, create "training" and "testing" sets. The SVM model will be trained using the training set.
***E. SVM Model Training*:** Utilise the training data to train a SVM model. Set the SVM's hyperparameters, including the regularisation parameter, kernel type, and others [22].
***F. Model Evaluation*:** Utilise relevant working metrics, such as "accuracy", "precision", "recall", and "F1 score", to evaluate the trained SVM model against the testing data.
***G. Hyperparameter Tuning*:** Optimize the SVM hyperparameters to achieve better performance. This can be done through techniques like "grid search" or "random search" to find the optimal values [23].
***H. Model Deployment*:** After SVM model training and evaluated, it can be used to predict the emotions of unseen speech samples from Voice-over Artists.
***I. Performance Evaluation*:** The execution of the emotion recognition algorithm is typically weighed using system of measurement like "accuracy", "precision", "recall", and "F1 score". These measurements offer perceptions into the model's ability to correctly classify different emotions from voice recordings [22-23].

In summary, emotion recognition of Voice-over Artists from speech involves collecting a dataset, extracting acoustic features, training machine learning models, and evaluating their performance. This process allows for the automatic identification and classification of emotions expressed in the Voice-over Artists' recordings, enabling various applications in industries where emotional expression through speech is relevant.

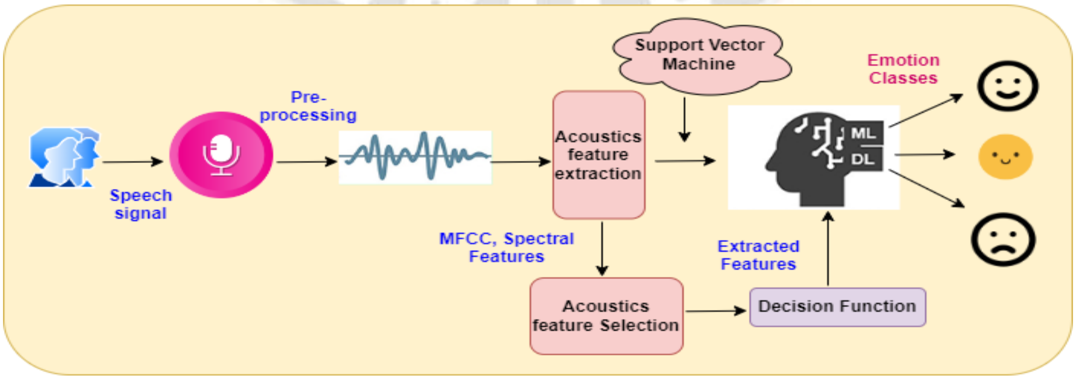

Fig.1 Emotion recognition using SVM.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10s
DOI: https://doi.org/10.17762/ijritcc.v11i10s.7608
***Article Received**: 30 May 2023 **Revised**: 29 July 2023 **Accepted**: 13 August 2023*

VI. Mathematical model of Support Vector Machine (SVM) for Emotion Recognition:

Given a training dataset consisting of N speech samples, each represented by a feature vector x_i of dimension D, and their corresponding emotion labels y_i (where y_i can be either +1 or -1 indicating the positive or negative emotion class), the goal of SVM is to find a hyperplane that separates the samples of different emotional classes with the largest margin.

The mathematical model can be represented as follows:

A. *Training Phase:*

Each voice sample in the annotated dataset used to train the SVM model has an emotion label attached to it (such as a positive or negative sentiment).

Objective function: minimize:

$$1/2 * ||w||^2 + C * \Sigma\xi_i \quad (1)$$

Subject to:

$$y_i * (w^T * \varphi(x_i) + b) >= 1 - \xi_i \quad (\xi_i >= 0) \quad (2)$$

Here, w is the weight vector perpendicular to the hyperplane, b is the bias term that controls the position of the hyperplane, ξ_i are slack variables that allow for some misclassification, and C is the regularization parameter that balances the trade-off between maximizing the margin and minimizing the classification errors [22-23].

In the objective function, the first term 1/2 * ||w||^2 represents the regularization term to control the complexity of the model and prevent overfitting. The second term C * Σξ_i is the slack penalty term that encourages the minimization of classification errors.

The constraints ensure that each sample is correctly classified (y_i * (w^T * φ(x_i) + b) >= 1) with a tolerance given by the slack variable ξ_i. The constraints also ensure that the slack variables are non-negative (ξ_i >= 0). The function φ (·) represents a non-linear mapping of the input feature vector x_i to a higher-dimensional feature space. This mapping allows SVM to handle non-linear decision boundaries by implicitly projecting the samples into a higher-dimensional space where linear separation is possible [22-23].

B. *Prediction Phase:*

During the prediction phase, the SVM model takes as input a new speech sample (represented by its feature

vector) and computes a decision function value. This decision function value can be interpreted as a confidence score or a measure of how likely the sample belongs to a particular emotion class. The decision function can be represented as:

$$f(x) = w^T * \varphi(x) + b \quad (3)$$

If f(x) >= 0, the sample is classified as the positive emotion class; otherwise, it is classified as the negative emotion class.

SVMs can be solved using optimization techniques such as quadratic programming or convex optimization algorithms to find the optimal values of w and b that minimize the objective function while satisfying the constraints [22-23,31].

To extend SVM for multi-class emotion classification with emotions such as sad, anxiety, surprise, happy, and neutral, etc we can use one of the following techniques:

One-vs-All (OvA) Approach:

In the One-vs-All approach, each emotion category is treated as a separate binary classification task. The method involves training C binary SVM classifiers, where C represents the number of emotion classes under consideration. For each class c, positive samples are drawn from that specific emotion category, while negative samples are selected from all other emotion classes. During the prediction phase, the algorithm calculates decision values for all C SVM classifiers and assigns the test sample to the class with the highest decision value. This approach enables the classification of the test sample into one of the emotion categories, based on the highest confidence score obtained from the individual binary classifiers [31].

***One-vs-One (OvO) Approach*:**

In the One-vs-One approach, a binary SVM classifier is trained for each pair of emotion classes. If there are C emotion classes, this strategy necessitates the creation of C * (C - 1) / 2 binary classifiers. During the prediction phase, the algorithm calculates decision values for all binary classifiers. To determine the final predicted emotion, a voting scheme is employed, where each binary classifier contributes its vote for the emotion class it supports. The emotion class with the highest number of votes is selected as the final predicted emotion for the test sample. This approach allows for the classification of the test sample into one of the emotion categories based on the majority vote from the individual binary classifiers [31].

VII. Multiclass Support Vector Machine (SVM) for Emotion Recognition:

To classify emotions such as sad, anxiety, surprise, happy, and neutral using SVM, we have modified the problem as a multi-class classification task. Here is a mathematical formulation of the SVM model for emotion classification: Suppose we have a training dataset consisting of N samples, each with D-dimensional feature vectors and their corresponding emotion labels.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10s
DOI: https://doi.org/10.17762/ijritcc.v11i10s.7608
***Article Received**: 30 May 2023 **Revised**: 29 July 2023 **Accepted**: 13 August 2023*

A. Data Preparation:

Let X be the input feature matrix of size N x D, where each row represents a feature vector. Let y be the emotion label vector of size N x 1, where each element represents the corresponding emotion label (sad, anxiety, surprise, happy, neutral).

B. Label Encoding:

Convert the emotion labels (sad, anxiety, surprise, happy, neutral) into numerical labels, such as 0, 1, 2, 3, 4, respectively.

C. SVM Model Training:

Train a multi-class SVM model using the training dataset (X, y). Choose an appropriate kernel function and set other hyperparameters. Use the "One-vs-One (OvO)" or "One-vs-All (OvA)" approach to handle multi-class classification. The SVM method identifies

the ideal hyperplane(s) that maximally separates the various emotions in the feature space during training.

D. SVM Model Prediction: Given a new test sample with a feature vector x_test, apply the trained SVM model to predict its emotion label.

Compute the decision values or probabilities for each emotion class using the trained SVM model. Choose the emotion label with the highest decision value or probability as the predicted emotion for the test sample.

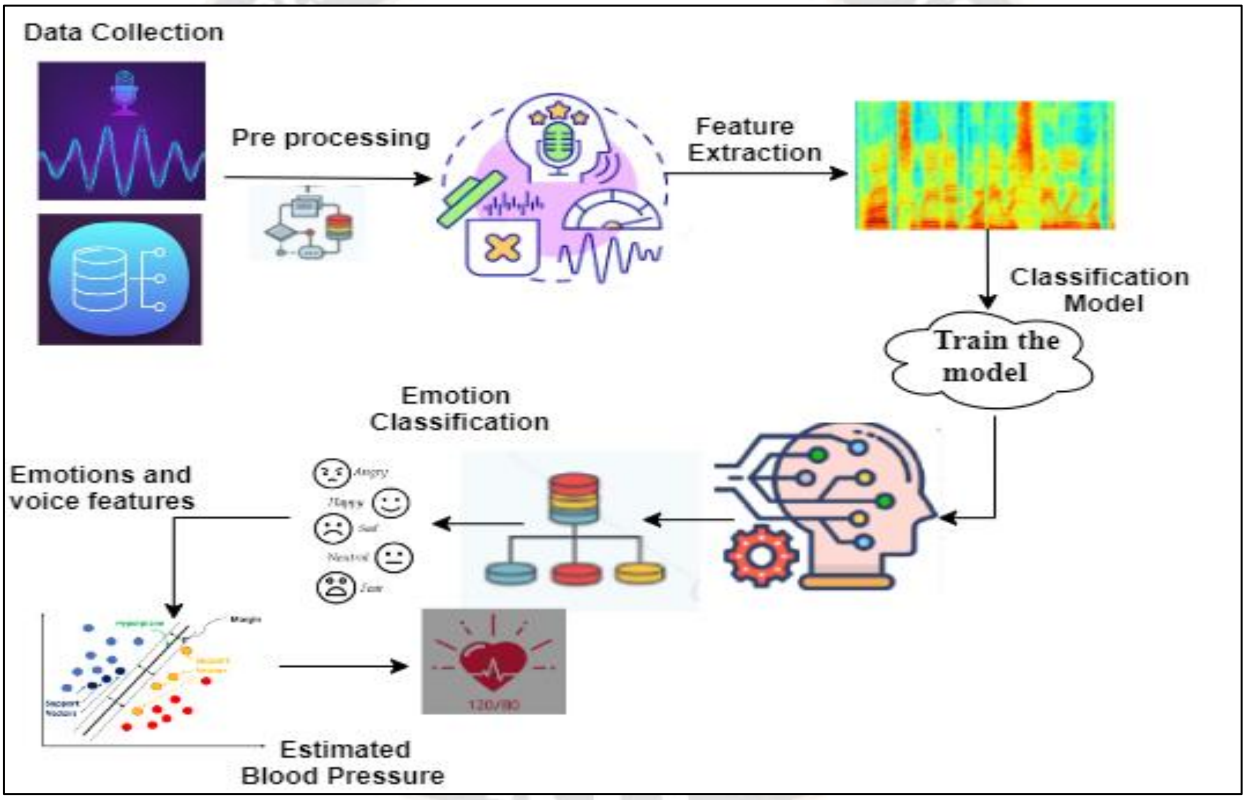

Fig.2 Blood Pressure estimation from the Emotions

VIII. Mathematical Model for Emotion based Blood Pressure Estimation:

The emotion classification step assigns a label (e.g., sad, anxiety, surprise, happy, neutral) to each voice sample. Let E be the set of recognized emotions, where E = {e1, e2, ..., en}. Each voice sample is associated with an emotion label, denoted as Emotion(i), where i represents the sample index. After the emotion classification step, the classified emotions can be utilized to estimate Blood Pressure (BP) levels as shown in Fig.2.

To estimate Blood Pressure (BP) levels after emotion classification, a mathematical model can be developed that relates the recognized emotions to corresponding

BP values [17-18]. Here is a detailed mathematical model for BP estimation:

A. Model Training:

The emotion-to-BP mapping function converts the recognized emotions into estimated BP values [6,10-11]. Let BP(i) represent the estimated Blood Pressure for voice sample i. The

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10s
DOI: https://doi.org/10.17762/ijritcc.v11i10s.7608
***Article Received**: 30 May 2023 **Revised**: 29 July 2023 **Accepted**: 13 August 2023*

mapping function can be represented as BP(i) = f(Emotion(i)), where f() is the mapping function.

The mapping function f() needs to be trained using a dataset that includes paired samples of recognized emotions and corresponding actual BP measurements.

The training dataset consists of N samples, denoted as {(Emotion(1), BP(1)), (Emotion(2), BP(2)), ..., (Emotion(N), BP(N))}.

The goal is to learn the mapping function f() that can accurately predict BP values based on recognized emotions.

B. Regression Model:

Several regression algorithms, including as linear regression, support vector regression (SVR), and artificial neural networks, can be used to train the mapping function [7-8,24,27].

Let X be the feature vector representing the recognized emotions, X = [x1, x2, ..., xm].

Each recognized emotion can be encoded as a feature, where xi represents the presence or intensity of a specific emotion. The regression model aims to learn the relationship between the feature vector X and the corresponding BP values.

C. Training and Optimization:

The training process involves fitting the regression model to the training dataset to minimize the prediction error between the estimated BP values and the actual BP measurements. The optimal parameters of the regression model can be found using optimisation methods such as gradient descent or stochastic gradient descent.

D. BP Estimation:

Once the mapping function f () is trained, it can be used to estimate the BP values for voice samples with recognized emotions. Given a new voice sample with recognized emotions, the feature vector X is constructed based on the presence or intensity of each recognized emotion. The mapping function f() is then applied to X to predict the corresponding BP value.

The mapping function f() in the context of Blood Pressure (BP) estimation from recognized emotions is responsible for transforming the recognized emotions into estimated BP values. It captures the relationship between the input (recognized emotions) and the output (estimated BP) through a mathematical model [6,14]. The specific formulation of f() will depend on the chosen regression algorithm and the features extracted from the recognized emotions.

We have five recognized emotions: sad, anxiety, surprise, happy, and neutral. We represent these emotions as a feature vector X = [x1, x2, x3, x4, x5], where xi represents the intensity or presence of the corresponding emotion.

The mapping function f() can be formulated using a regression model. Here's an example of how f() can be represented using a linear regression model:

$$f(X) = b0 + b1x1 + b2x2 + b3x3 + b4x4 + b5*x5 \;.. (4)$$

In this equation, b0 is the intercept term, and b1, b2, b3, b4, b5 are the coefficients that capture the relationship between each emotion feature and the estimated BP value.

Alternatively, if we use SVR, the mapping function can be expressed as:

$$f(X) = \Sigma\, (ai * K(X, Xi)) + b \qquad (5)$$

In this equation, ai represents the weights assigned to support vectors, Xi represents the feature vectors of the support vectors, K(X, Xi) represents the kernel function that measures the similarity between X and Xi, and b is the bias term.

During the training phase, the regression model will learn the optimal values for the coefficients (b0, b1, ..., bn) or the support vectors (ai) and bias (b) to minimize the prediction error between the estimated BP values and the actual BP measurements in the training dataset.

X. Experimental Results:

The performance of different algorithms for speech-emotion recognition was evaluated using a dataset of 500 audio clips of Voice-over Artists encompassing five different emotions: Sad, Anxiety, Surprise, Happy, and Neutral. The algorithms considered for analysis were SVM, Decision Tree, Naïve Bayes, Random Forest, and KNN. In terms of accuracy, SVM achieved the highest accuracy of 85%, followed closely by Random Forest with an accuracy of 84%. Precision, recall, and F1-score were also evaluated for each algorithm. SVM demonstrated the highest precision and F1-score among the algorithms, indicating its ability to accurately classify the different emotions.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10s
DOI: https://doi.org/10.17762/ijritcc.v11i10s.7608
***Article Received**: 30 May 2023 **Revised**: 29 July 2023 **Accepted**: 13 August 2023*

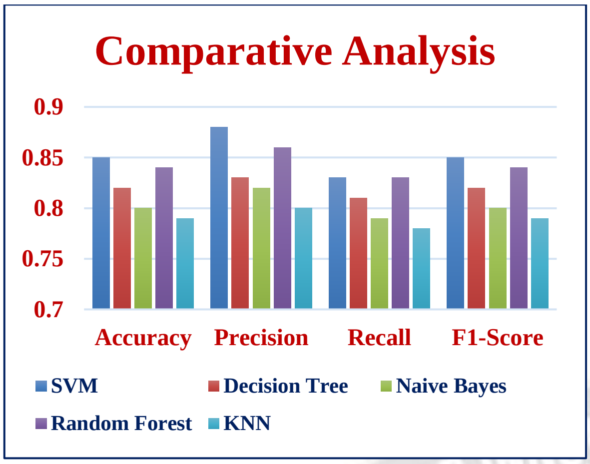

Fig.3. Comparative Analysis of Different Algorithms for Speech-Emotion Recognition

Fig.3. shows the comparative analysis of different algorithms for Speech-Emotion Recognition implemented using different algorithms. Confusion matrices were generated for each algorithm to provide insights into the classification performance for individual emotions as shown from Table 1 to Table 5. The confusion matrices depicted the number of correct and incorrect predictions for each emotion category.

The analysis indicates that certain emotions, such as stress and anxiety, are associated with higher Blood Pressure levels, while emotions like neutral and happiness are linked to normal Blood Pressure levels. Also, sad emotion may result into hypotension in some cases as shown in Table 6.

Table 1. Confusion Matrix using Decision Tree

	Sad	Anxiety	Surprise	Happy	Neutral
Sad	**95**	4	1	0	0
Anxiety	6	**88**	4	1	1
Surprise	2	7	**90**	0	1
Happy	0	1	0	**92**	2
Neutral	0	2	2	1	**91**

Table 2. Confusion Matrix using Naïve bayes

	Sad	Anxiety	Surprise	Happy	Neutral
Sad	**89**	6	1	1	2
Anxiety	8	**82**	7	1	1
Surprise	1	8	**84**	1	1
Happy	1	2	0	**89**	2
Neutral	1	4	1	0	**90**

Table 3. Confusion Matrix using Random Forest

	Sad	Anxiety	Surprise	Happy	Neutral
Sad	**90**	6	1	1	2
Anxiety	8	**83**	7	1	1
Surprise	1	8	**82**	1	1
Happy	1	2	0	**90**	2
Neutral	1	4	1	0	**92**

Table 4. Confusion Matrix for KNN

	Sad	Anxiety	Surprise	Happy	Neutral
Sad	**93**	5	0	1	1
Anxiety	6	**87**	7	0	0
Surprise	3	10	**87**	0	2
Happy	1	2	0	**96**	1
Neutral	1	4	1	0	**92**

Table 5. Confusion Matrix for SVM

	Sad	Anxiety	Surprise	Happy	Neutral
Sad	97	5	0	1	1
Anxiety	6	92	7	0	0
Surprise	3	10	94	0	2
Happy	1	2	0	96	1
Neutral	1	3	2	0	94

Table 6. Estimated SBP and DBP values

Emotion	Estimated SBP	Estimated DBP
Anxiety	138.37	89.87
Happy	118.94	80.54
Neutral	117.02	78.20
Sad	98.89	68.25
Surprised	115.11	82.10

Table 7 gives comparative analysis of existing algorithms used for emotion-based Blood Pressure estimation.

Table 7. Comparative analysis of existing algorithms used for emotion-based Blood Pressure estimation.

Source: [22-23,25-26,32-36]

Algorithm	Advantages	Disadvantages
SVM	- High accuracy in classification	- Computationally intensive and may have longer training time
	- Effective in handling high-dimensional data	- Sensitive to parameter settings, requires careful tuning
	- Works well with small to medium-sized datasets	- May not perform well with imbalanced datasets
	- Can handle variable-length input sequences	- Difficulty in interpreting the model due to its complex nature
Naive Bayes	Fast and efficient	Assumes feature independence
	Works well with high-dimensional data	Can be affected by irrelevant features
Decision Tree	- Creates strong predictive models	- Prone to overfitting
	- Handles numerical and categorical features	
	- Performs well with imbalanced data	
Random Forest	- Ensemble method that combines multiple decision trees for improved performance	- Can be prone to overfitting with excessive trees
	- Robust to noise and outliers	- Increased complexity and computational requirements with large datasets
KNN	- Simple and intuitive algorithm	- Computationally expensive during inference for large datasets
	- Handles both numerical and categorical data	- Not suitable for high-dimensional data

The findings suggest that emotions play a crucial role in influencing the physiological state of Voice-over Artists, particularly their Blood Pressure. This implies that emotional well-being and management are important factors to consider in maintaining a healthy cardiovascular system for individuals in the Voice-over profession.

XI. Conclusion:

In conclusion, the correlation between “emotions” and “Blood Pressure” is complex and multifaceted. Various emotions, such as anger, stress, anxiety, fear, sadness, happiness, surprise, and neutral states, can have different effects on Blood Pressure. Anger, stress, anxiety, and fear tend to increase Blood Pressure. On the other hand, sadness may lead to a temporary decrease in Blood Pressure, while happiness and neutral emotions generally have minimal impact. Understanding the impact of emotions on Blood Pressure is important as it provides insights into the physiological responses associated with different emotional states. It highlights the role of emotions in cardiovascular health and the potential influence on the risk of hypertension and other cardiovascular conditions.

Moreover, recognizing the relationship between emotions and Blood Pressure has implications for various fields, including psychology, neuroscience, and cardiovascular health. It underscores the importance of managing emotions, stress, and promoting emotional well-being as part of a comprehensive approach to maintaining optimal cardiovascular health. Overall, the study highlights the impact of emotions on Blood Pressure and emphasizes the need for further research and awareness regarding the emotional health of Voice-over Artists to promote their overall well-being.

Further research in this area can contribute to a deeper understanding of the mechanisms underlying the emotional regulation of Blood Pressure and facilitate the development of more targeted interventions for individuals at risk of hypertension or other cardiovascular issues.

Acknowledgement

I would like to extend my heartfelt gratitude to **Dr. Sapana and Dr. Shriananta Pardeshi,** for their invaluable support and guidance throughout this research endeavour. Their expertise and assistance in the Blood Pressure study, have been instrumental in the success of this study. I am truly grateful for their unwavering support and dedication, which has been a driving force behind the successful completion of this research work.

Conflict of Interests

The authors declare that they have no conflict of interest.

Author Contributions

V.R.: Conceptualization , Data curation, Methodology, Software , Validation, Writing – Original draft preparation.

P.M.: Conceptualization, Investigation, Visualization, Validation, Supervision, Writing – Reviewing and Editing.

R.R. : Supervision, Writing – Reviewing and Editing.

References

[1] Argha, A., Celler, B. G., & Lovell, N. H. (2021). A Novel Automated Blood Pressure Estimation Algorithm Using Sequences of Korotkoff Sounds. IEEE Journal of Biomedical and Health Informatics, 25(4), 1257–1264. https://doi.org/10.1109/JBHI.2020.3012567

[2] Badshah, A. M., Rahim, N., Ullah, N., Ahmad, J., Muhammad, K., Lee, M. Y., Kwon, S., & Baik, S. W. (2019). Deep features-based speech emotion recognition for smart affective services. Multimedia Tools and Applications, 78(5), 5571–5589. https://doi.org/10.1007/s11042-017-5292-7

[3] Farki, A., Kazemzadeh, R. B., & Noughabi, E. A. (2021). A Novel Clustering-Based Algorithm for Continuous and Non-invasive Cuff-Less Blood Pressure Estimation. In arXiv physics.med-ph]. arXiv. http://arxiv.org/abs/2110.06996

[4] Harfiya, L. N., Chang, C.-C., & Li, Y.-H. (2021). Continuous Blood Pressure Estimation Using Exclusively Photopletysmography by LSTM-Based Signal-to-Signal Translation. Sensors , 21(9). https://doi.org/10.3390/s21092952

[5] Yang, S., Zhang, Y., Cho, S.-Y., Correia, R., & Morgan, S. P. (2021). Non-invasive cuff-less Blood Pressure estimation using a hybrid deep learning model. In Optical and Quantum Electronics (Vol. 53, Issue 2). https://doi.org/10.1007/s11082-020-02667-0

[6] Mottaghi, S., Moradi, M. H., & Moghavvemi, M. (2014). Neuro-fuzzy Indirect Blood Pressure Estimation during Bruce Stress Test.

[7] Guizzo, E., Weyde, T., & Leveson, J. B. (2020). Multi-Time-Scale Convolution for Emotion Recognition from Speech Audio Signals. ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 6489–6493. https://doi.org/10.1109/ICASSP40776.2020.9053727

[8] Mustaqeem, Mustaqeem, Sajjad, M., & Kwon, S. (2020). Clustering-Based Speech Emotion Recognition by Incorporating Learned Features and Deep BiLSTM. In IEEE Access (Vol. 8, pp. 79861–79875). https://doi.org/10.1109/access.2020.2990405

[9] Martínez, G., Howard, N., Abbott, D., Lim, K., Ward, R., & Elgendi, M. (2018). Can Photoplethysmography Replace Arterial Blood Pressure in the Assessment of Blood Pressure? Journal of Clinical Medicine Research, 7(10). https://doi.org/10.3390/jcm7100316

[10] Sakai, M. (2015a). Case study on analysis of vocal frequency to estimate Blood Pressure. 2015 IEEE Congress on Evolutionary Computation, CEC 2015 - Proceedings, 00(c), 2335–2340. https://doi.org/10.1109/CEC.2015.7257173

[11] Sakai, M. (2015b). Modeling the Relationship between Heart Rate and Features of Vocal Frequency. International Journal of Computer Applications, 120(6), 32–37. https://doi.org/10.5120/21233-3986

[12] Ogedegbe, G., & Pickering, T. (2010). Principles and Techniques of Blood Pressure Measurement. Cardiology Clinics, 28(4), 571–586. https://doi.org/10.1016/j.ccl.2010.07.006

[13] Pechprasarn, S., Sukkasem, C., Sasivimolkul, S., & Suvarnaphaet, P. (2019). Machine Learning to identify factors that affect Human Systolic Blood Pressure. BMEiCON 2019 - 12th Biomedical Engineering International Conference. https://doi.org/10.1109/BMEiCON47515.2019.8990356

[14] Ryskaliyev, A., Askaruly, S., & James, A. P. (2016). Speech signal analysis for the estimation of heart rates under different emotional states. 2016 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2016, 1160–1165. https://doi.org/10.1109/ICACCI.2016.7732201

[15] Krishna Kishore, K. V., & Krishna Satish, P. (2013). Emotion recognition in speech using MFCC and wavelet features. Proceedings of the 2013 3rd IEEE International Advance Computing Conference, IACC 2013, 842–847. https://doi.org/10.1109/IAdCC.2013.6514336

[16] Kachuee, M., Kiani, M. M., Mohammadzade, H., & Shabany, M. (2017). Cuffless Blood Pressure Estimation Algorithms for Continuous Health-Care Monitoring. IEEE Transactions on Biomedical Engineering, 64(4), 859–869. https://doi.org/10.1109/TBME.2016.2580904

[17] El Ayadi, M., Kamel, M. S., & Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognition, 44(3), 572–587. https://doi.org/10.1016/j.patcog.2010.09.020

[18] Aich, S., Kim, H. -., Younga, K., Hui, K. L., Al-Absi, A. A., & Sain, M. (2019). A supervised machine learning approach using different feature selection techniques on voice datasets for prediction of parkinson's disease. Paper presented at the International Conference on Advanced Communication Technology, ICACT, 2019-February 1116-1121. doi:10.23919/ICACT.2019.8701961 Retrieved from www.scopus.com

[19] Mustaqeem, Mustaqeem, & Kwon, S. (2021). MLT-DNet: Speech emotion recognition using 1D dilated CNN based on multi-learning trick approach. In Expert Systems with Applications (Vol. 167, p. 114177). https://doi.org/10.1016/j.eswa.2020.114177

[20] Nie, W., Ren, M., Nie, J., & Zhao, S. (2021). C-GCN: Correlation Based Graph Convolutional Network for Audio-Video Emotion Recognition. IEEE Transactions on Multimedia, 23, 3793–3804. https://doi.org/10.1109/TMM.2020.3032037

[21] Mustaqeem, & Kwon, S. (2019). A CNN-Assisted Enhanced Audio Signal Processing for Speech Emotion Recognition. Sensors , 20(1). https://doi.org/10.3390/s20010183

[22] Fahad Nasser Alhazmi. (2023). Self-Efficacy and Personal Innovation: Conceptual Model Effects on Patients' Perceptions of PHR Use in Saudi Arabia. International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 369–376. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2676

[23] Alkan, A., & Günay, M. (2012). Identification of EMG signals using discriminant analysis and SVM classifier. Expert Systems with Applications, 39(1), 44–47. https://doi.org/10.1016/j.eswa.2011.06.043

[24] Pan, Y., Shen, P., & Shen, L. (2012). Speech emotion recognition using support vector machine. International Journal of Smart Home, 6(2), 101–108. https://doi.org/10.1109/kst.2013.6512793

[25] Mustaqeem, Mustaqeem, & Kwon, S. (2020). CLSTM: Deep Feature-Based Speech Emotion Recognition Using the Hierarchical ConvLSTM Network. In Mathematics (Vol. 8, Issue 12, p. 2133). https://doi.org/10.3390/math8122133

[26] Song, K., Park, T.-J., & Chang, J.-H. (2021). Novel Data Augmentation Employing Multivariate Gaussian Distribution for Neural Network-Based Blood Pressure Estimation. NATO Advanced Science Institutes Series E: Applied Sciences, 11(9), 3923. https://doi.org/10.3390/app11093923

[27] Y., A., A., K., A., Y., & Pandya, N. (2017). Survey paper on Different Speech Recognition Algorithm: Challenges and Techniques. International Journal on Speech Communications, 175(1), 31–36. https://doi.org/10.5120/ijca2017915472

[28] Shakti P.Rath(2017). Scalable algorithms for unsupervised clustering of acoustic data for speech recognition. Computer Speech & Language. Volume 46, November 2017, Pages 233-248. https://doi.org/10.1016/j.csl.2017.06.001

[29] T. B. Amin, P. Marziliano and J. S. German, "Nine Voices, One Artist: Linguistic and Acoustic Analysis," 2012 IEEE International Conference on Multimedia and Expo, Melbourne, VIC, Australia, 2012, pp. 450-454, doi: 10.1109/ICME.2012.142.

[30] N. Obin and A. Roebel, "Similarity Search of Acted Voices for Automatic Voice Casting," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 24, no. 9, pp. 1642-1651, Sept. 2016, doi: 10.1109/TASLP.2016.2580302.

[31] Phalguni Phatak, Sonia Shaikh, Neha Jamdhade & Pallavi Sovani Kelkar (2021) Do Voice-Over Artists Convey Emotion Better Than Untrained Voice Users?, Voice and Speech Review, 15:3, 315-
329, DOI: 10.1080/23268263.2021.1882751

[32] Duan, KB., Rajapakse, J.C., Nguyen, M.N. (2007). One-Versus-One and One-Versus-All Multiclass SVM-RFE for Gene Selection in Cancer Classification. In: Marchiori, E., Moore, J.H., Rajapakse, J.C. (eds) Evolutionary Computation,Machine Learning and Data Mining in Bioinformatics. EvoBIO 2007. Lecture Notes in Computer Science, vol 4447. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-71783-6_5

[33] Dhabliya, M. D. . (2021). Cloud Computing Security Optimization via Algorithm Implementation. International Journal of New Practices in Management and Engineering, 10(01), 22–24. https://doi.org/10.17762/ijnpme.v10i01.99

[34] Hong Feng and Xunbing Shen. 2022. A random forest algorithm-based emotion recognition model for eye features. In Proceedings of the 3rd International Symposium on Artificial Intelligence for Medicine Sciences (ISAIMS '22). Association for Computing Machinery, New York, NY, USA, 148–152. https://doi.org/10.1145/3570773.3570851

[35] Hariharan Muthusamy, Kemal Polat, Sazali Yaacob, "Improved Emotion Recognition Using Gaussian Mixture Model and Extreme Learning Machine in Speech and Glottal Signals", Mathematical Problems in Engineering, vol. 2015, Article ID 394083, 13 pages, 2015. https://doi.org/10.1155/2015/394083

[36] Xing, W., & Bei, Y. (2020). Medical Health Big Data Classification Based on KNN Classification Algorithm. IEEE Access, 8, 28808–28819. https://doi.org/10.1109/access.2019.2955754

[37] An, Y., Sun, S., & Wang, S. (2017). Naive Bayes classifiers for music emotion classification based on lyrics. https://doi.org/10.1109/icis.2017.7960070

[38] Mark White, Thomas Wood, Carlos Rodríguez, Pekka Koskinen, Jónsson Ólafur. Exploring Natural Language Processing in Educational Applications. Kuwait Journal of Machine Learning, 2(1). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/168

[39] Liu, Z., Wu, M., Cao, W., Mao, J., Xu, J., & Tan, G. (2018). Speech emotion recognition based on feature selection and extreme learning machine decision tree. Neurocomputing, 273, 271–280. https://doi.org/10.1016/j.neucom.2017.07.050

[40] Zhang, W. ., & Chung Ee, J. Y. . (2023). AnIntelligent Knowledge Graph-Based Directional Data Clustering and Feature Selection for Improved Education. International Journal on Recent and Innovation Trends in Computing and Communication, 11(6s), 22–33. https://doi.org/10.17762/ijritcc.v11i6s.6807

[41] 38. Velunachiyar, S., & Sivakumar, K. (2023). Some Clustering Methods, Algorithms and their Applications. International Journal on Recent and Innovation Trends in Computing and Communication, 11(6s),401–410. https://doi.org/10.17762/ijritcc.v11i6s.6946