

Research Article

An Efficient Sentence-based Sentiment Analysis for Expressive Text-to-speech using Fuzzy Neural Network

B. Sudhakar and R. Bensraj

Department of Electrical Engineering, Annamalai University, College Rd, Annamalai Nagar, Chidambaram, Tamil Nadu 608002, India

Abstract: In recent years, speech processing has become an active research area in the field of signal processing due to the usage of automated systems for spoken language interface. In developed countries, the customer service with automated system in speech synthesis has been the recent trend. The existing automated speech synthesis systems have certain problems during the real time implementation such as lack of naturalness in output speech, lack of emotions and so on. In this study, the novel Text to Speech system is introduced along with the sentiment analysis in Tamil language. The input text is first classified into the positive, negative and neutral based on the emotions in the sentence then the text is converted into speech with emotions during TTS conversion. Existing approaches used neural network based classifiers for classification. But, neural networks have certain drawbacks in real time training. So, this research study uses Fuzzy Neural Network (FNN) to classify the sentence based on the emotions. The text to speech with sentiment analysis effective scheme which is evaluated using Doordarshan news Tamil dataset. The proposed scheme is implemented using MATLAB. This TTS system has several social applications, especially in railway stations where the announcements can be made through expressive speech.

Keywords: Another Tool for Language Recognition (ANTLR), Natural Language Processor (NLP), Text to Speech (TTS)

INTRODUCTION

Speech processing has become one of the most important research areas due to its importance in various applications. In recent years, TTS systems have been extensively used in many real time applications for automated speech synthesis around the world. There has been a tremendous growth and development in the area of Text to speech in last two decades by introducing many efficient methods to improve communication in efficient manner (Bing, 2012; Rudy and Mike, 2009). Sentiment analysis has been a recent trend in the area of speech processing and has attracted various researchers.

Sentiment analysis has become one of the essential fields for research in the computational linguistics and is used for analyzing the people's expression in either speech or text. It is also called as opinion mining as it analyzes the people's opinions, attitudes and emotions. In recent years, sentiment analysis has been extensively used for analyzing the comments, feedbacks and critiques in TTS synthesis. Generally, sentiment analysis falls under the category of classification task.

Sentiment analysis process the input text by grouping positive, negative and neutral sentences. TTS system has several real time societal applications such

as retail banking, railway stations, colleges, universities, etc. In retail banking, this system can be used for analyzing the opinion about the new offerings, customer feedbacks as it helps to improve the data in a more expressive way. It is very useful to the politicians and policy makers for analyzing the public opinion about policies and political issues during electoral campaign. This study can also be effectively applied in railway stations for announcements which can be made in a more expressive way.

Three types of analysis are possible in sentimental analysis based on the documents. They are explained in detailed manner.

Document level: In this method, the whole document is considered as input for analyzing and it classifies the whole document as positive or negative based on the opinion (Pang *et al.*, 2002; Turney, 2002). It is called document level sentiment analysis. The analyzer classifies the document by assuming that the entire document expresses the opinion about the single product. Therefore, it is not suited for the documents which compare multiple products opinions.

Sentence level: In this method, single sentence is considered as input to analyzer and it is classified based

Corresponding Author: B. Sudhakar, Department of Electrical Engineering, Annamalai University, College Rd, Annamalai Nagar, Chidambaram, Tamil Nadu 608002, India

This work is licensed under a Creative Commons Attribution 4.0 International License (URL: <http://creativecommons.org/licenses/by/4.0/>).

on the opinion expressed in the sentence. The sentences are normally grouped into positive, negative and neutral. It is similar to the subjectivity classification which classifies the sentence based on the opinion and subjective views. Although the sentence level classification gives better results than the document level still it clauses are not efficient to obtain the high classification accuracy.

Entity and aspect level: The sentence and document level analysis have certain limitations such as poor classification accuracy. Since both analyzers did not find user's opinion clearly, this method does not give an efficient output (Hu and Bing, 2004). This method classifies the document based on the opinion instead of language constructs such as documents, paragraphs, sentences, clauses or phrases. The main aim of this analyzer is to discover the sentiments on entities or aspects.

This study focuses on TTS conversion along with sentence level sentiment analysis, since earlier works on TTS analysis did not concentrate on sentiment analysis. If text is given as input to the system, then it is classified based on positive, negative and neutral based text. Based on text type or category, original text is converted into speech with the emotions of the user. The proposed system can be used in real time applications such as railway stations for announcing the train information, retail banking industries, automated telephone system in customer services.

LITERATURE REVIEW

Tiomkin *et al.* (2010) presented a segment-wise model for enhancing the baseline TTS system. This system is used to provide additional degrees of freedom in speech feature determination. The author utilizes the degree of freedom to increase the speech feature vector norm constraint and it is compared with the original norm constraint. The obtained speech from the method is less over smoothed and it sounds more natural.

Chalamandaris *et al.* (2010) proposed a design and implementation approach for the efficient integration of this technology into screen reading environments. This study mainly focused on the issues of natural language processing, speed optimization, multilingual design and overall quality optimization. The system is evaluated by series of test and the performance qualities of the approach are compared with the conventional concepts.

Alias *et al.* (2008) introduced a novel TTS strategy is called multi-domain TTS for synthesizing speech among different domains. It is widely used in spoken language systems. Many works have been carried out to improve the text to speech system. Many algorithms are introduced for improving the system. In first stage, text classifier is introduced in the traditional TTS in order to classify the text by selecting the appropriate domain.

The text classifier used in this method is to classify the text based on the content and its structure. In second stage, the text modeling scheme is also introduced to improve TTS. The text modeling is used to represent the text as a directional weighted word based graph. It is mainly based on the associative relational network.

Bellegarda (2011) presented a general framework based on the latent affective mapping and also proposed to utilize the two levels of semantic information. The first level of semantic information is mainly considered the foundations of domain and second level considered the fabric of the language. The connections between these two semantic levels are used to improve the classification process. This method is mainly used for improving the automatic emotion analysis in text to speech systems. The author compared this method with other existing systems in order to show the effective results. Although the obtained speech is quite natural, its emotion analysis performance is not effective.

Alexandre and Francesc (2009) described a text classifier for automatically tagging the sentiment of input text according to the emotion that is being conveyed. The author used pipelined framework which is composed of Natural Language Processing modules for feature extraction. The binary classifier is used for decision making between positive and negative sentences. The author evaluated the performance of the presented method by training the classifier with the help of semeval 2007 dataset. The dataset consist of sentence based on different emotions.

Alm *et al.* (2005) proposed a framework called SNoW learning architecture which is mainly based on the emotion prediction problem. The author mainly aimed to classify sentence based on the emotion during the story narration for children's in text to speech synthesis. The author evaluated the Snow learning architecture with 22 fairy tale shows and classifies the results based emotions in the content.

METHODOLOGY

Proposed TTS framework: The proposed work focuses on TTS systems based on the sentiment analysis. The proposed framework is mainly focused on reviewing and gathering a set of features which is used for sentiment classification. The proposed framework is constructed based on the framework of EmoLib and it is particularly developed for south Indian Dravidian languages. The framework of the EmoLib is presented in Fig. 1. The EmoLib framework is used to analysis the input text based on the features extracted from the text and it uses the different types of classifiers for problems. Recently, TTS systems are widely used in the railway stations for announcing the train information. The proposed framework is used to improve the TTS systems used in railway stations. It is improved by adding the features of sentiment analysis in TTS (Trilla *et al.*, 2010). The framework of the proposed system is clearly explained and it is shown in the Fig. 2.

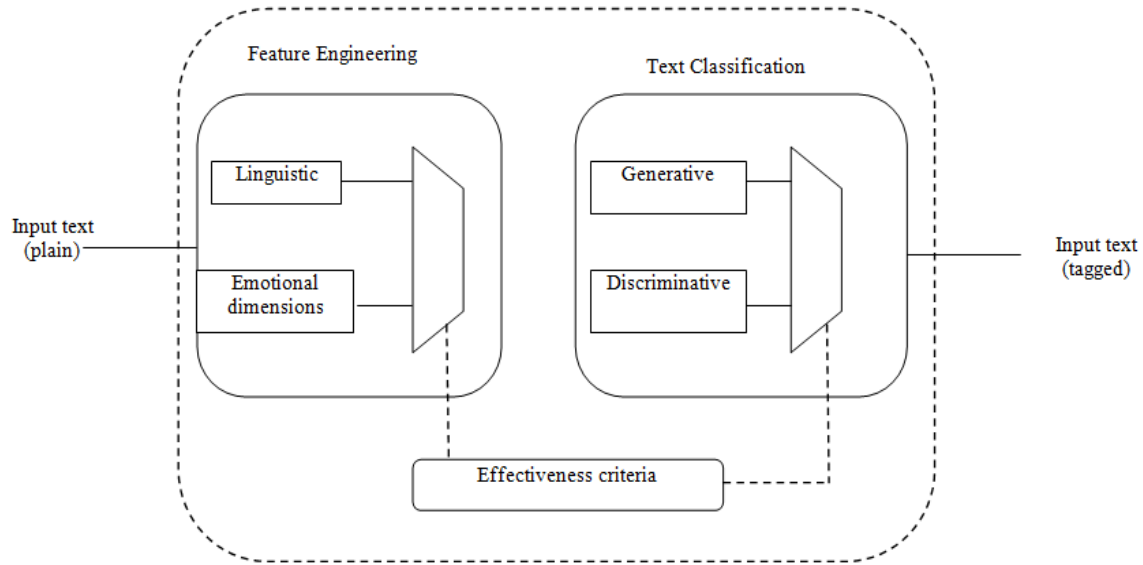


Fig. 1: Overview of sentiment analysis framework

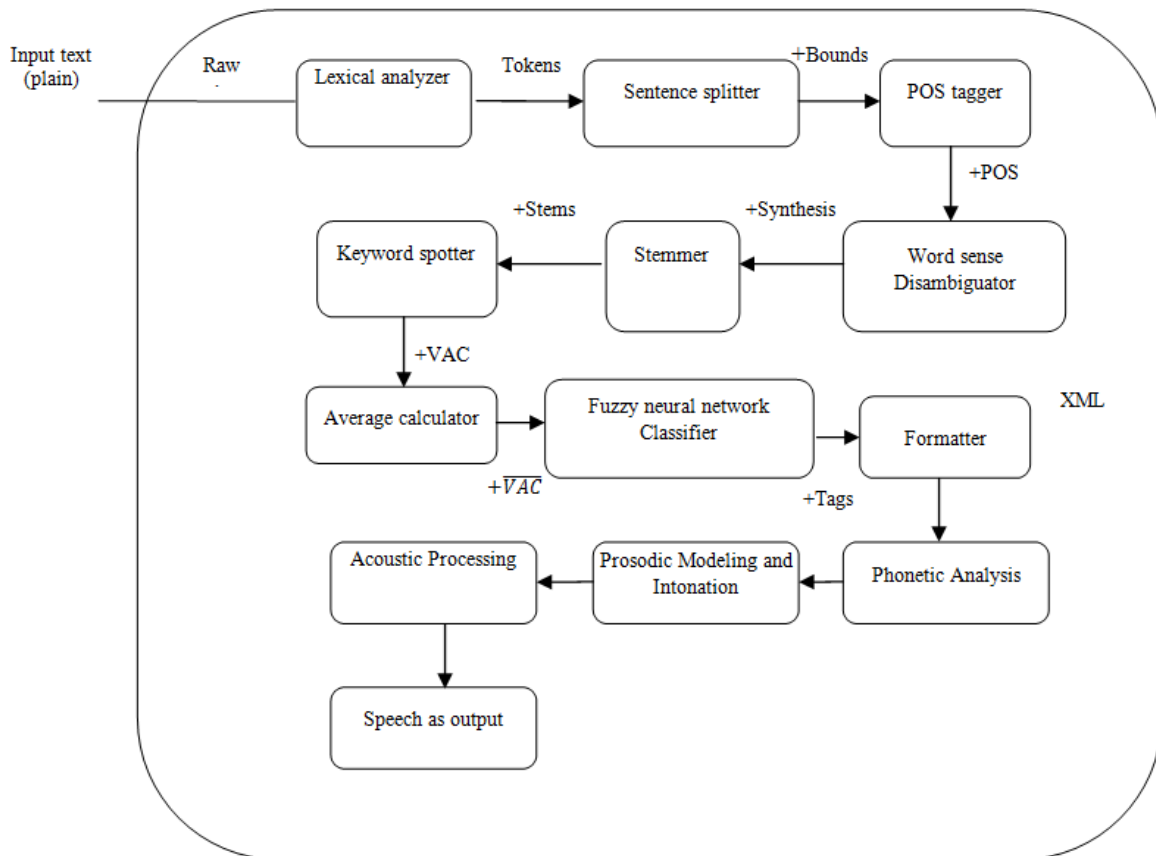


Fig. 2: Proposed framework of text to speech system

Framework:

Lexical analyzer: The lexical analyzer is used for converting the plain input Tamil text to individual units is as called tokens. These tokens are expressed in regular patterns as established by the grammar of the

language. It is the INPUTTER of the pipeline. ANTLR (ANother Tool for Language Recognition is powerful parser generator) is used in this framework for reading, processing, executing, or translating structured text or binary files. In addition to this, it is used to determine

the content words, negation words and intensifiers in the text (Pang and Lee, 2008). ANTLR is also used for filtering the stop words in the sentences (Sebastiani, 2002). Punctuations and special characters in the sentences are also removed using ANTLR.

Sentence splitter: The sentence splitter splits the whole document into paragraph and sentences. The sentence splitter uses the binary decision for delimiting the sentences. In general, periods, uppercase letters, exclamation points and question marks are good indicators of sentence boundaries. The example for sentence splitter is example:

கோவிலில் ஆறு அடி உயரமான மணி உள்ளது
கோவிலில் ஆறு அடி உயரமான மணி
உள்ளது

POS tagger: POS Tagger is used to determine the functions of noun, verbs and adjectives (classes of words with a possible affective content) within the sentence.

SVMTool for Tamil POS tagger: The SVMTool is mainly composed of three modules namely the model learner (SVMTlearn), the tagger (SVMTagger) and the evaluator (SVMTeval). Tamil language is one of the important languages when compared to the other languages and it is inflected with more grammatical features which makes the POS tagger system complex. In this study automatic parts of speech tagger is constructed for determining the grammatical features of Tamil languages. POS tagger is constructed using the SVM tools.

Word-sense disambiguator: It resolves the meaning of affective words (i.e., nouns, adjectives and verbs) according to their context. In this study, the word disambiguator uses the semantic similarity measure to score the senses of an affective word with the context words. The word disambiguator algorithm uses Supervised learning Approach of word sense disambiguation. Additionally, the module retrieves the set of synonyms for the resulting sense in order to expand the feature space (Manning *et al.*, 2008).

Stemmer: Stemmer groups the words based on the meaning of the words for efficient information retrieval performance. It removes the inflection of words for indexing purposes using the Iterative Stemmer for Tamil (Baskaran and Vijay-Shanker, 2003). Semantically related words should map to the same stem, base or root form and this should compensate for data sparseness.

Keyword spotter: This module provides the emotional dimensions to the emotional words using the dictionary of affect for Tamil (Bradley and Lang, 1999). It considers both the stems of the lexical instances as well as their POS tags. Words are mapped into a tri-dimensional space (Francisco *et al.*, 2011), also known as the circumflex, which defines the Valence (positive/negative evaluation), Activation (stimulation of activity) and Control (submissiveness) features (VAC) of the affect that they convey.

Average calculator: This section is used to calculate the average emotional dimensions for each input text. In the current work, this is the arithmetic mean of the dimensions at the sentence level (i.e., a centroid-based approach).

Classifier: The input text is classified based on the emotions derived from the sentence. The classifier assigns labels to each input text with suitable emotion based on the affective attributes. It predicts the most appropriate sentiment label according to the features extracted from the terms observed in the text, which is usually taken for a bag of words.

Formatter: It presents the results in a usable form, which follows a XML specification (Schröder *et al.*, 2011), ready to be used by the TTS system that follows. For example, the module can use the Speech Synthesis Markup Language or the Emotion Markup Language (Baggia *et al.*, 2010).

Text analysis: Texts written in the Tamil font using the software are converted into the ISCII format. Then it is given as input to the phonetic analysis. Tamil language rules are applied to convert the text to sound waves.

Phonetic analysis: Phonetic Analysis is used to convert the orthographical symbols into phonological ones using a phonetic alphabet. A collection of phones that constitute minimal distinctive phonetic units are called Phoneme. In this study grapheme to phoneme converter is used to convert the ISCII format text into the sound waves. It used syntax and Letter to Sound rules (LTS) for converting into the Tamil phonetic format. These phones are the one which are defined in the Tamil phone set.

Prosodic modeling and intonation: The prosody is defined as the combination of stress pattern, rhythm and intonation in a speech. The prosodic modeling is mainly used to describe the speaker's emotion through pitch variations. As a result, the detailed analysis of the text is needed in this section by analyzing the pitch variation in order to obtain the high quality of the speech (Table 1 and 2).

Intonation is simply a variation of speech while speaking. Intonation is an important task in TTS synthesis systems because it affects naturalness of the

Table 1: Vowels and its pronunciation

Vowels	Pronunciation
అ	a
ఆ	aa/A
ఇ	i
ఈ	ii/I
ఉ	u
ఊ	uu/U
ఎ	e
ఏ	ee/E
ఐ	ai
ఓ	o
ఔ	oo/O
ఌ	au

Table 2: Consonants and its pronunciation

Consonants	Pronunciation
క	k
గ	ng
చ	ch
జ	nj
ట	t
థ	N
ద	th
ధ	nth
ప	p
మ	m
య	y
ర	r
ల	l
వ	v
శ	zh
ఱ	L
ఱ	tR
న	n

speech. In order to obtain the quality of speech conversion in TTS, better intonation method is needed. In this study Classification and Regression Tree (CART) based tree intonation method is used for obtaining the high quality speech.

Classifier:

Dimensionality reduction and weighting: In sentence-level Text Classification, the text feature plays an important role in the classification of sentence. If all the features are used for classification purpose then it increases feature space size. It affects the overall performance of the system. As a result, the important features are selected and the most relevant features raise the discriminating properties of the data thus improving the classification effectiveness.

These methods are described as follows:

Term selection: In this section, the most relevant text features are selected in order to reduce the dimensionality of the feature space (Sebastiani, 2005; Dang *et al.*, 2010). This improves both the effectiveness of the classifier as well as its computational performance given the fewer number of features to process. In this study, modified Local Context Analysis (LCA) is used as term selection of features to improve the classifier performance.

Term weighting: In this section, the features power is discriminated without reducing the dimensionality of the feature space. The term weighting process is better than term selection process because of its criteria used (Manning *et al.*, 2008). The text features are selected based on the frequency of terms which is naturally encoded. The binary term is used to denote the presence or absence of term and it is comparatively better than others (Pang and Lee, 2008). In this study, BO1 is used for term weighting process. The Bose-einstein statistics (Bo1) uses the Divergence from Randomness (DFR) term weighting model (Dipasree *et al.*, 2013; Kosko, 1992) for effective selection of text features. The term *t* is measured from the top ranking documents. The term weighting is calculated using the term *t* is given by:

$$S(t) = (\sum_{d \in PRD} tf(t, d)) * \log_2 \left(\frac{1 + f_{avg}(t, C)}{f_{avg}(t, C)} \right) + \log_2 (1 + f_{avg}(t, C))$$

where,

$$f_{avg}(t, C) = \sum_{d \in C} \frac{tf(t, d)}{N}$$

Fuzzy neural network classifier: Fuzzy Neural Network (FNN) classifies the input text based on the emotions derived from the sentence. The classifier assigns labels to each input text with suitable emotion based on the effective attributes. It predicts the most appropriate sentiment label according to the features extracted from the terms observed in the text, which is usually taken for a bag of words. FNN is used as a classifier to classify the text based on the emotions.

FNN is the combination of fuzzy theory and neural network (Lin *et al.*, 1999; Chen and Teng, 1995). Since the both models are integrated, it could have better results. Both models are created based on the functions of human brain, psychological reasoning and mental status of the humans.

Fuzzy system: The input data is converted into the set of fuzzy data. It is called as fuzzification. The fuzzy systems use the different types of membership function for the effective output. In this study, triangular membership functions are used.

The fuzzy inference mechanism is the core of the fuzzy system (Fig. 3). The classifier has two inputs and the fuzzy rules expressed as follows:

$$R^j: \text{if } x_1 \text{ is } A_1^j \text{ and } x_2 \text{ is } A_2^j \text{ then } x \text{ belongs to class } y_j \text{ with } CF = CF_j$$

R^j denotes the *j*-th fuzzy rule. A_1^j and A_2^j denotes the fuzzy sets where $j = 1$ to n input text with n rules. $y_j \in y \{1, 2, 3 \dots M\}$ is the output text of the *j*-th fuzzy

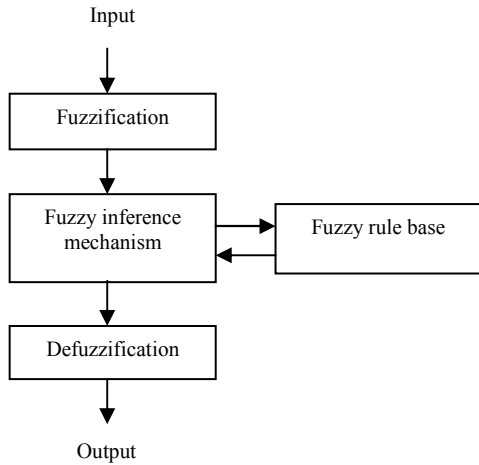


Fig. 3: Fuzzy inference system

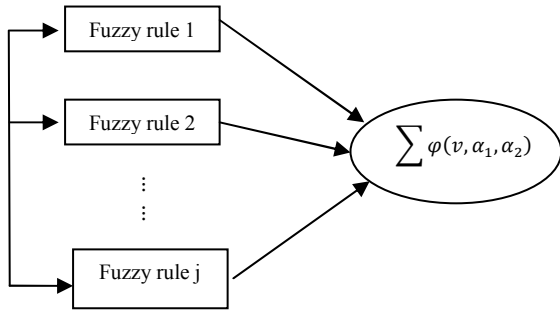


Fig. 4: The artificial neural network classifier combined with single neuron

rule, which is one for the M labels, CF_j denotes the reliability of the fuzzy rule R^j . The Generalized Modus Ponens is applied to obtain the effective results and Max-Min composition operation and the fuzzy inference outputs expressed as follows (Trilla *et al.*, 2010):

$$\mu_y(y) = \max_j \{ \mu_{y,j} \} | y_j = y$$

where,

$$\mu_{y,j} = \mu_{m_{A,j}}(x) \cdot CF_j$$

$$\mu_{m_{A,j}}(x) = \min \{ \mu_{A_1}(x_1), \mu_{A_2}(x_1) \}$$

The process of converting the linguistic variable into a crisp value is called defuzzification. Formally, the outputs from the fuzzy inference may be the fuzzy sets or specific values. If the result obtained from the inference system is a fuzzy set, then median method and the center area method are applied to convert the fuzzy set to crisp outputs. In this study, an artificial neural network based on single neuron is used to

Table 3: Rule base

Word	VN	N	Ne	P	VP
VN	NM	NM	NM	NS	ZO
N	NM	NM	NS	ZO	PS
NE	NM	NS	ZO	PS	PM
P	NS	ZO	PS	PM	PM
VP	ZO	PS	PM	PM	PM

VN: Very negative; N: Negative; NE: Neutral; P: Positive; VP: Very positive; NM: Negative maximum; NS: Negative small; ZO: Zero; PS: Positive small; PM: Positive maximum

defuzzify the fuzzy set. Each data in the fuzzy set are connected to the single neuron and it is converted into the crisp outputs i.e., classified text with emotions. The single neurons accept the p inputs from the fuzzy set and calculate the bond value for each data using the transfer function:

$$y = \varphi(v)$$

The output is the classification inference probability of each text. The formula for classification:

$$v = \sum_{i=0}^p w_i u_i$$

$$y = \varphi(v) = \frac{\alpha_2}{1 + e^{-\alpha_1 v}}$$

where, α_1 controls the shape of the function, α_2 adjusts the size of scale. Figure 4 shows the combination of the fuzzy system and the single neuron.

Based on these rules, the words are categorized and the particular frequency is allotted for the words. For example, if a positive word ‘P’ like ‘happy’ is given as input, the matching frequency is checked in the rule base Table 3 and the corresponding output ‘PM’ would be obtained which results in the corresponding expressive speech (Fig. 5).

Speaking range: One hundred and fifty for sad voice to 197 for happy voice and default value for neutral voice are 160.

F0 range: Sad voice has smallest range and happy voice has highest one.

Mean pitch level: Happiest voice has highest mean and sad voice has lowest value.

F0 slope: Happy voice has high descendent slope and sad voice has flat one.

Spectral tilt: A lower tilt value increases the high frequency contents of the voice source, producing clearer voices; it is a specially used for happy voices.

Additive noise: Pitch-synchronous noise added to the voice source; used for sadness.

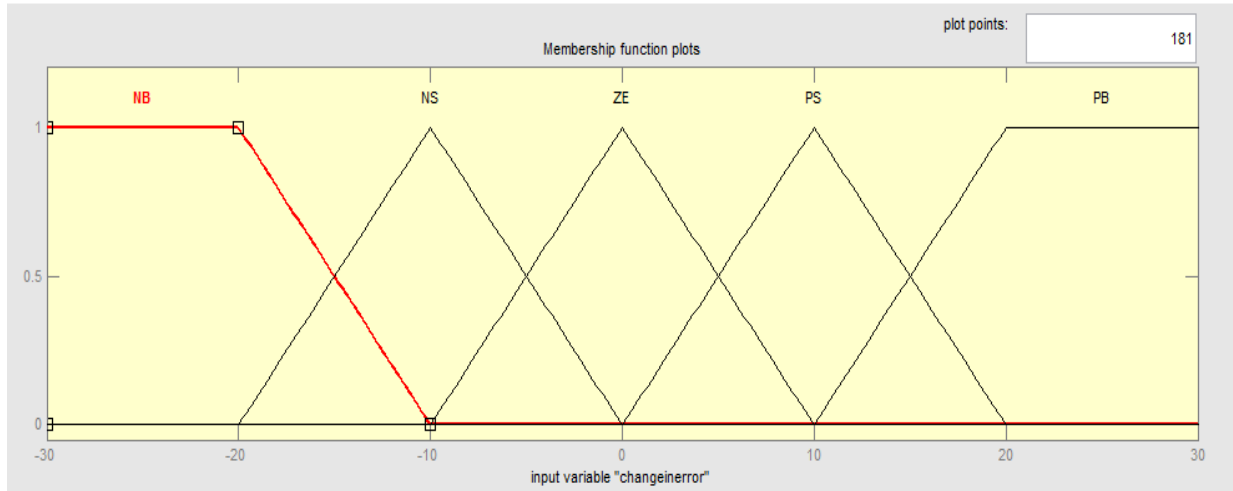


Fig. 5: Membership function for fuzzy rules customizable parameters of emotional synthesis

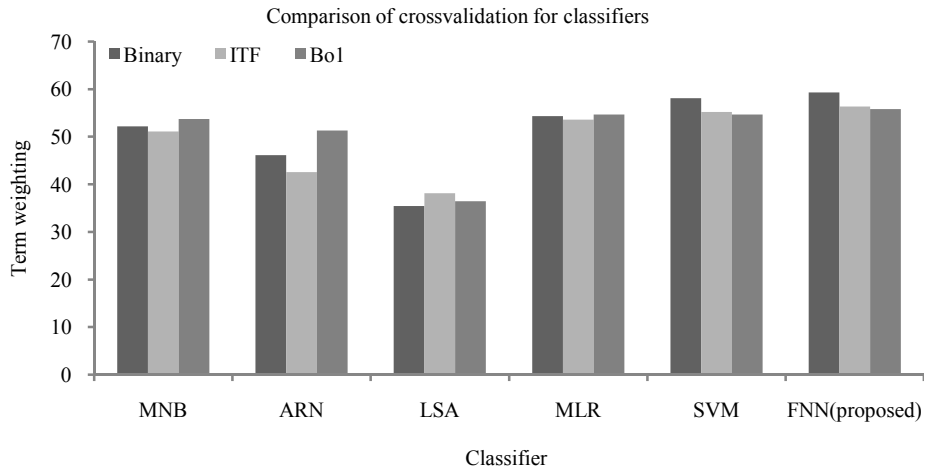


Fig. 6: Comparison of cross validation using term weighting for different classifiers

Type of emotion: Special rules are applied for happy and neutral voice, changing the default intonation contour characteristics.

SIMULATION RESULTS AND DISCUSSION

In this study, the high quality sentiment based text to speech analysis is proposed. The results obtained from the proposed system have naturalness and high quality. The proposed system is simulated in MATLAB. The dataset used in the proposed system is doordarshan news bulletins. The text is given as input to the microcontroller through the keyboard and it is converted into speech sound with emotions. The properties of input text features in terms of instance and feature counts are explained in Table 1. Table 2 and Fig. 6 shows the comparison the different features using the term weighting method. Exceptionally, the Fuzzy neural network classifier shows a significant improvement for Bo1 with respect to Binary and

especially to Inverse Term Frequency (ITF). The proposed classifier is better than the existing classifier such Associative Relational Network-Reduced (ARN-R), Multinomial Logistic Regression (MLR), Multinomial Naive Bayes (MNB), Support Vector Machine (SVM), Latent Semantic Analysis (LSA) (Table 4).

Dataset: The data corpus of doordarshan news bulletins are taken to analyze the proposed system for tamil

Table 4: Properties of the dataset in terms of instance and feature counts

Instant properties	Counts
Total sentence	1250
Positive	174
Neutral	764
Negative	312
With repeat words	46
Without stop words	4
Average length	7.53

Table 5: Average results with the whole set of features using 10-fold cross-validation (mean std)

Classifier	Term weighting		
	Binary	ITF	Bo1
MNB	52.20±4.30	51.09±3.93	53.72±6.54
ARN-R	46.14±6.33	42.56±5.10	51.28±3.82
LSA	35.45±6.64	38.10±6.24	36.44±8.28
MLR	54.32±6.43	53.58±6.72	54.66±5.14
SVM	58.12±4.15	55.20±5.16	54.67±5.13
Fuzzy neural network	59.23±5.25	56.33±6.12	55.78±6.85



Fig. 7: Simulated output for happy voice

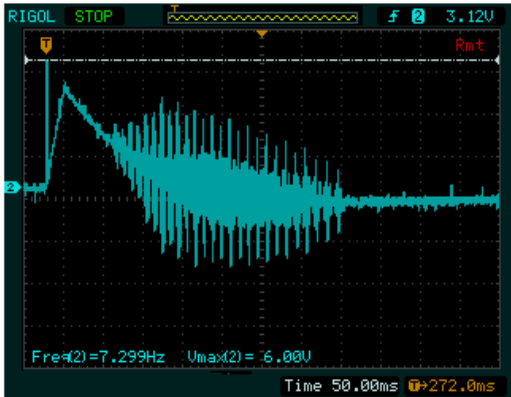


Fig. 8: Simulated output for neutral voice

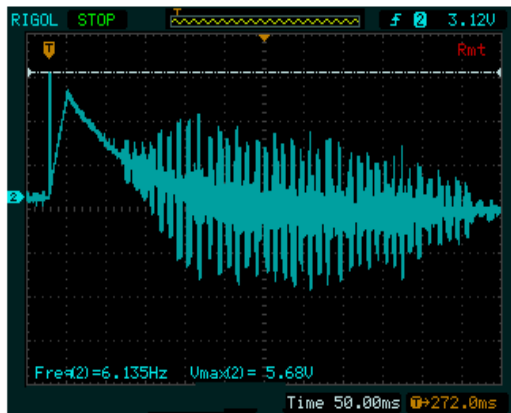


Fig. 9: Simulated output for sad voice

language. The speaking style in this corpus is news reading style. The fundamental frequency variation (F0 contour) in each syllable was analyzed (Table 5).

The following are the sample outputs of the proposed text to speech:

அவள் இன்று மிகவும் சந்தோஷமாக இருக்கிறாள்

Input text: Aval indru megavum santhosamaga irukiral.

Figure 7 shows the output sound wave form for positive sentences. The given input text is first classified into the positive sentence using classifier and then it is converted into the positive wave form based on Tamil phonetic language rules:

இந்தியா இன்று இலங்கையிடம் படுதோல்வியுற்றது ஆனால் சிறந்த விராக சச்சின் தாந்தேடுக்கப்பட்டார்

Input text: India indru ilagaiedam padutholviuttrathu aanal siratha veeraraka therenthadukapattar.

Figure 8 shows the output sound wave form for neutral sentences. The given input text is first classified into the neutral sentence using classifier and then it is converted into the neutral wave form based on Tamil phonetic language rules:

இந்தியா இன்று இலங்கையிடம் படுதோல்வியுற்றது

Input text: India indru ilagaiedam padutholviuttrathu.

Figure 9 shows the output sound wave form for negative sentences. The given input text is first classified into the negative sentence using classifier and then it is converted into the negative wave form based on Tamil phonetic language rules.

CONCLUSION

This research study proposes a novel phrase-level sentiment analysis designed and implemented for Tamil Language. A database has been created from the various domain words and syllables using doordarshan news bulletins. The given input text is analyzed and classified based on the emotions and opinion expressed in the sentence. The desired speech is produced by the concatenative speech synthesis approach. FNN has been used in this study for training the words. FNN has shown significant results with its training and testing capabilities. The proposed system is implemented using MATLAB and emotions markup language is used for obtaining efficient output. The emotions analyzed in this study are Happy, Sad and Neutral. The output speech clearly depicts the emotions of the user which is the main feature of this study. The proposed system can be implemented in real time applications such as automated train information announcement in railway stations, automated telephone in retail banking customer service.

REFERENCES

- Alexandre, T. and A. Francesc, 2009. Sentiment classification in English from sentence-level annotations of emotions regarding models of affect. *Proceeding of the 10th Annual Conference of the International Speech Communication Association (INTERSPEECH, 2009)*. Brighton, United Kingdom, pp: 516-519.
- Alias, F., X. Sevilano, J.C. Socoro and X. Gonzalvo, 2008. Towards high-quality next-generation text-to-speech synthesis: A multidomain approach by automatic domain classification. *IEEE T. Audio Speech*, 16(7): 1340-1354.
- Alm, C.O., D. Roth and R. Sproat, 2005. Emotions from text: Machine learning for text-based emotion prediction. *Proceedings of the Joint Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp: 579-586.
- Baggia, P., P. Bagshaw, M. Bodell, D.Z. Huang, L. Xiaoyan, S. McGlashan, J. Tao, Y. Jun, H. Fang, Y. Kang, H. Meng, W. Xia, X. Hairong and Z. Wu, 2010. *Speech Synthesis Markup Language (SSML) Version 1.1*, W3C, Tech. Retrieved from: <http://www.w3.org/TR/speech-synthesis11/>.
- Baskaran, S. and K. Vijay-Shanker, 2003. Influence of morphology in word sense disambiguation for Tamil. *Proceeding of the International Conference on Natural Language Processing, Mysore*.
- Bellegarda, J.R., 2011. A data-driven affective analysis framework toward naturally expressive speech synthesis. *IEEE T. Audio Speech*, 19(5): 1113-1122.
- Bing, L., 2012. *Sentiment Analysis and Opinion Mining*. Morgan and Claypool Publishers, San Rafael, California.
- Bradley, M.M. and P.J. Lang, 1999. *Affective Norms for English Words (ANEW): Instruction manual and affective ratings*. Technical Report C-1, Center for Research in Psychophysiology, University of Florida, Gainesville, FL, pp: 1-45.
- Chalamandaris, A., S. Karabetos, P. Tsiakoulis and S. Raptis, 2010. A unit selection text-to-speech synthesis system optimized for use with screen readers. *IEEE T. Consum. Electr.*, 56(3): 1890-1897.
- Chen, Y.C. and C.C. Teng, 1995. A model reference control structure using a fuzzy neural network. *Fuzzy Set. Syst.*, 73: 291-312.
- Dang, Y., Y. Zhang and H. Chen, 2010. A lexicon-enhanced method for sentiment classification: An experiment on online product reviews. *IEEE Intell. Syst.*, 25(4): 46-53.
- Dipasree, P., M. Mandar and D. Kalyankumar, 2013. Query expansion using term distribution and term association. *Inform. Retrieval*, 1: 1-19.
- Francisco, V., R. Hervás, F. Peinado and P. Gervs, 2011. EmoTales: Creating a corpus of folk tales with emotional annotations. *Lang. Resour. Eval.*, 45: 1-41.
- Hu, M. and L. Bing, 2004. Mining and summarizing customer reviews. *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD, 2004)*, pp: 168-177.
- Kosko, B., 1992. *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice Hall, Englewood Cliffs, NJ.
- Lin, F.J., W.J. Hwang and R.J. Wai, 1999. A supervisory fuzzy neural network control system for tracking periodic inputs. *IEEE T. Fuzzy Syst.*, 7(1): 41-52.
- Manning, C.D., P. Raghavan and H. Schütze, 2008. *Introduction to Information Retrieval*. Cambridge University Press, Cambridge, MA.
- Pang, B. and L. Lee, 2008. Opinion mining and sentiment analysis. *Found. Trends Inf. Retrieval*, 2(1/2): 1-135.
- Pang, B., L. Lee and S. Vaithyanathan, 2002. Thumbs up? Sentiment classification using machine learning techniques. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP, 2002)*, pp: 79-86.
- Rudy, P. and T. Mike, 2009. Sentiment analysis: A combined approach. *J. Informetr.*, 3(2): 143-157.
- Schröder, M., H. Pirker, M. Lamolle, F. Burkhardt, C. Peter and E. Zovato, 2011. Representing Emotions and Related States in Technological Systems. In: Petta, P., R. Cowie and C. Pelachaud (Eds.), *Emotion-Oriented Systems-the Humaine Handbook*. Cognitive Technologies, Springer, Heidelberg, pp: 367-386.
- Sebastiani, F., 2002. Machine learning in automated text categorization. *ACM Comput. Surv.*, 34: 1-47.
- Sebastiani, F., 2005. Text Categorization. In: Zanas, A. (Ed.), *Text Mining and its Applications to Intelligence, CRM and Knowledge Management*. WIT Press, Southampton, UK, pp: 109-129.
- Tiomkin, S., D. Malah and S. Shechtman, 2010. Statistical text-to-speech synthesis based on segment-wise representation with a norm constraint. *IEEE T. Audio Speech*, 18(5): 1077-1082.
- Trilla, A., F. Alias and I. Lozano, 2010. Text classification of domain-styled text and sentiment-styled text for expressive speech synthesis. *Proceeding of the 6th Jornadas en Tecnología del Habla (FALA, 2010)*, pp: 75-78.
- Turney, P.D., 2002. Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews. *Proceeding of the 40th Annual Meeting of the Association for Computational Linguistics (ACL, 2002)*, pp: 417-424.