March-April 2020
ISSN: 0193-4120 Page No. 7740 - 7748

Wrapper-based Feature Selection for Classifying Cued Speech Malay Syllables

Zaridah Mat Zain[1], Hariharan Muthusamy[2], Kushsairy Abd Kader[3], Zulkhairi Mohd Yusuf[4]
[1,3,4]Universiti Kuala Lumpur, British Malaysian Institute Km 8 Jalan Sungai Pusu, 53100 Gombak, Selangor Malaysia
[2]Temporary Campus-Govt. ITI, National Institute of Technology, Srinagar (Garhwal) – 246174, Uttarakhand India
zaridahmz@unikl.edu.my[1], hariharanm@nituk.ac.in[2], kushsairy@unikl.edu.my[3], zulkhairi@unikl.edu.my[4]

Article Info
Volume 83
Page Number: 7740 - 7748
Publication Issue:
March - April 2020

Article History
***Article Received**: 24 July 2019*
***Revised**: 12 September 2019*
***Accepted**: 15 February 2020*
***Publication**: 09 April 2020*

Abstract

Malay Cued Speech is adapted from English Cued Speech to assist hearing-impaired children in communicating using visual sign language. Malay Cued Speech is created to serve as a supplement of lipreading and allow complete access to spoken language in a purely visual form. Integrating computing technology to Malay Cued Speech offers excellent flexibility of learning and therapy. However, with significant variations of speech signals due to speaker variabilities such as gender, dialects and speaking style could make the task challenging. Without previous understanding on the acoustical properties of the speech, it is difficult to discover the relevant features of the dataset. Besides, irrelevant and redundant features might degrade the classification accuracy due to its large dimension of search space. In this paper, three wrapper-based feature selection using Genetic Algorithm (GA), Particle Swarm Optimization (PSO) and Binary PSO is proposed to discover which features are most useful. The acoustic feature set from 10 native children as configured in Interspeech 2010 Paralinguistic Challenge (IS10) is extracted. Extreme Learning Machine (ELM) is used to classify twenty-two Cued Malay Syllables. The best accuracy is achieved by GA-ELM (72.47%). The optimised features are then fed to a Heterogeneous Ensemble Classifiers (HCE) for further improvement. Radial Basis Function ELM, Polynomial Support Vector Machine and Linear Kernel ELM are constructed for the base classifiers. Multiple combination methods are tested to find diversity among the performances of each base classifiers to attain a significant improvement of the accuracy.

Keywords*; *Malay Cued Speech, Malay Syllables, Extreme Learning Machine, Ensemble Classifiers and Wrapper-Based Selection

I. INTRODUCTION

Cued Speech is known as a phonetic-based visual communication system for hearing-impaired. Cued Speech uses hand shapes to represent consonant phonemes and hand placements to represent vowel phonemes together with natural mouth movements to remove the ambiguity of lip-reading and clearly show spoken information through vision alone. The smallest unit of sound is phonemes and combining various phonemes make up word and sentence. This phenomenal seen phonemes as the building blocks of a language. With or without a hearing, hearing-impaired should see precisely what is being said. Cued Speech supports lip-reading skills and many Cued Speech users have developed strong lip-reading abilities. Also, the use of Cued Speech can be used to support the development of speech skills. Since Cued Speech is phonetically based, the hearing-impaired children are entirely aware of the sounds that make up words, which supports the articulation process. The advantage of cued speech over sign language system is that it can be learned within a short period. Learning the base of cues requires far less time than learning the thousands of symbol-like signals in a sign language system.

Cued Speech is now becoming more and more

Published by: The Mattingley Publishing Co., Inc.

TEST Engineering & Management

March-April 2020
ISSN: 0193-4120 Page No. 7740 - 7748

attended by the world and has been adapted to over sixty-five languages worldwide, including the Malay language. Mr Tan Chin Guan who was the Vice President of the National Society for the Deaf at that time adapted English Cued Speech to Malay Cued Speech in 1982 with the help of Dr Orin Cornett, the inventor of English Cued Speech. According to Dr Cornett, implementation of Malay Cued Speech is easier since the sound generated from the vowel 'a' is consistent in the Malay Language. Pusat Perututuran Kiu (PPK) was founded in 1988 as a teaching centre of Malay Language Cued Speech. PKK later is registered as a private school in 2002. The school is known as Sekolah Pendidikan Khas Pertuturan Kiu (SPKPKiu), located in Kampung Pandan, Kuala Lumpur.

While the potential of Malay Cued Speech in improving literacy is significant potential, the speech learning centre is only provided at SPKPKiu and with little evidence of success (Yasin, Bari, & Hassan (2013) and (Mohid & Zin, 2011). Malay Cued Speech is currently lack of popularity, and sign language is still dominant among hearing-impaired children in Malaysia (Miles, Khairuddin & McCracken, 2018). Parents may not be aware of the existing of such an excellent cue system and technology is not fully utilised yet in Malay Cued Speech. People outside of Klang Valley may find difficulties to find a Malay Cued Speech instructor, and if they could attend a short course of Malay Cued Speech at SPKPKiu, fluency of Malay language may be hard to develop since the learning centre is hard to reach. Moreover, language learning progress of Malay Cued Speech should be consistently practised and monitored. According to Miles, Khairuddin & McCracken (2018), currently sign language is not available in mainstream schools, and being able to speak is an essential to be able to participate in the classroom activities and the limitations of the high cost of hearing aids and cochlear implants limits the number of deaf children to benefit from the technology. Lack of access and limited curriculum without reliable assistive technology and specialist support shows a failure of technology and appropriate communication support among these children is challenging to establish.

Looking at the potential of Computer Assisted Language Learning (CALL) in Malay Language as suggested in (Abdullah, Hisham, & Parumo (2009), Mohid & Zin (2010), Rosdi, Mustafa, Salim & Hamid (2017) to be useful and fully utilized by hearing-impaired children and the most importantly for these children to learn to speak, this paper proposed a development of teaching assistant software based on Automatic Recognition System (ASR), explicitly designed for Malay Cued Speech children at SPKPKiu.

II. RESEARCH METHODOLOGY

This section introduces the research methods involved in the machine learning design for the development of the speech engine. There are six main procedures involved in the process. It started with the audio recordings of normal-hearing children, followed by data pre-processing. In the next part, the features are extracted based on IS10 configurations to select the best features that represent the targeted Malay syllables using three different wrapper-based feature selection methods. Once the best features are obtained, the classification process of Heterogeneous Ensemble Classifiers (HEC) will be performed to classify the target Malay syllables. Figure. I. shows the overall framework of the speech engine development of the proposed teaching assistant software.

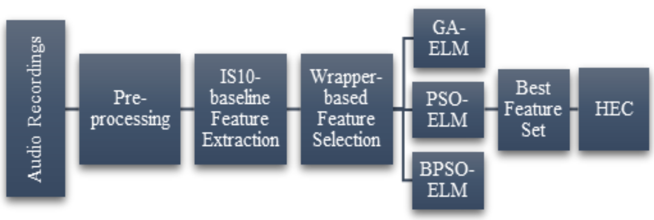

Figure. I. Block diagram of the proposed method

III. AUDIO RECORDINGS

A ten-speakers dataset of twenty-two Malay CV

March-April 2020
ISSN: 0193-4120 Page No. 7740 - 7748

(consonant-vowel) syllables including /ba, ca, da, fa, ga, ha, ja, ka, la, ma, na, pa, ra, sa, ta, va, wa, ya, za, nga, nya, sya/ is collected for the experiment. The selected syllables are aligned with the phonetics exercises in SPKPKiu. The speakers selected for the recordings are 10 native children aged between 7 to 12 years old. All the speakers use Malay Language as their first language. The recordings are done in a typical quiet room using Audacity software with a noise-cancellation microphone with the consent from the school principal. The speakers are asked to sit with the body upright and mouth facing the microphone. Each speaker produced the twenty-two Malay syllables sequentially at a regular speaking rate. The procedure is repeated 25 times, generating 25 productions of each syllable. A total of 5500 samples are acquired for data pre-processing.

IV. PRE-PROCESSING

For the pre-processing, all the recorded samples are downsampled to 8 KHz. The unvoiced portion has been removed from the signal based on the energy present in the frame, and the remaining voiced portion is concatenated into a single frame. The signals are then filtered by first order low pass filter to minimise the spectral distortion and signal discontinuity in each frame. The typical value of pre-emphasis coefficient was selected for the speech processing. The first-order pre-emphasis filter is defined as the following equation:

$$H(z)=1-a*z^{(-1)} \quad (1)$$

Where a is the pre-emphasis coefficient, and a typical value of 0.95 is selected (Jaafar and Ramli, 2013). The filtered signal is then segmented into 25ms frames with 50% overlap.

V. INTERSPEECH 2010 PARALINGUISTIC CHALLENGE (IS10) FEATURE EXTRACTION

The INTERSPEECH 2010 Paralinguistics Challenge (IS10) based features were extracted by using the openSMILE toolbox. A total of 1582 features, as shown in Table I., were extracted to gain a more in-depth insight into which features are of importance for the task. Bipolar normalisation or commonly known as minmax normalisation (between -1 to 1) is applied to the IS10 feature set to reduce data redundancy.

Table I. IS10 Descriptions

Group	LLD	Functional	Total Features
A	PCM Loudness	Max and min position	42
	MFCC 0-14	Arithmetic mean, standard deviation	630
	Log. Mel Frequency Band 1–8	Skewness, kurtosis	336
	LSP 0–7	Linear regression slope, offset	336
	F_0 envelope	Quartile 1, 2 and 3	42
	Probability of voicing	Quartile range 2–1 / 3–2 / 3–1	42
B	F_0 via Sub Harmonic Sum	Linear regression Error Quadratic, Absolute	38
	Local jitter	Percentile 1/99 (min or max)	38
	Delta jitter	Percentile range 1-99	38
	Shimmer	Up-level time 75% and 90%	38
	Speaking rate		1
	Segment Length		1
Total Extracted Features			**1582**

VI. WRAPPER-BASED FEATURE SELECTION

Total extracted features from the IS10 baseline feature set, resulting in a high dimensional of feature set (1582 features x 5500 audio samples). Data with high dimensionality has presented severe challenges to existing learning methods, i.e., the curse of dimensionality and overfitting, resulting in low recognition accuracies in return. Moreover, the complexity grows exponentially with the increasing number of features making the search of all possible

TEST Engineering & Management
March-April 2020
ISSN: 0193-4120 Page No. 7740 - 7748

spaces infeasible.

There are two major categories in feature selection; filter-based and wrapper-based. Wrapper-based feature selection combines machine learning classifier to develop a heuristic mechanism to select a subset of significant features that gives the best classifier's performance while filter-based feature selection evaluates each feature statistically. Wrapper-based feature selection working on the whole features at the same time, contras from filter-based which working on a single feature at one time making the sum of the information gathered may not be relevant for all domain. It would be beneficial to select as many as possible features from the large dataset to discover more information among those features while the recognition accuracy is monitored closely. Based on these arguments, Genetic Algorithm, Particle Swarm Optimization and Binary Particle Swarm Optimization wrapper-based feature selection are proposed to work with Extreme Learning Machine classifier to select the best feature set of the Malay syllables.

Speaker Dependent (SD) and Speaker Independent (SI) experiments are tested to analyse the performances of the extracted features. Different data partitioning methods are employed for both experiments. For the SD experiments, the training and test partitions are obtained by stratifying randomly of the speakers with a 70% - 30% split. In simple random sampling, every observation in the primary data set has an equal probability of being selected for the partition data set. In other words, each observation has a 70% chance of being selected. The more testing data, the less variance you can expect in your results.

For the SI partitioning, the following procedures are followed:

```
for fold_N = 1:N_folds
    indices = fold_N;
    train_indices = (Patterns(:,1)~=indices);
    test_indices = (Patterns(:,1)==indices);
    train_speaker = find(train_indices);
    test_speaker = find(test_indices);
    TrainPatterns = Patterns(train_speaker,2:end-1);
    TestPatterns = Patterns(test_speaker,2:end-1);
    TrainTargets = Targets(train_speaker,end);
    TestTargets = Targets(test_speaker,end);
end
```

Where N_folds is fixed to 10 equals to the number of subjects, patterns are the feature set, and targets are the labels (22 Malay syllables). For each run, one subject is used for the testing set (550 audio samples x 1582 features) and the remaining nine subjects (4950 audio samples x 1582 features) are used for the training set. The above procedures are followed to avoid the overlap of the testing data. The resultant differences are again assumed to be an independently drawn sample from an approximately normal distribution. The testing sets are independent, and the size is small, which entails a high variance of the estimates.

Extreme Learning Machine (ELM) is used as the classifier to find fitness function in wrapper-based GA, PSO and BPSO. ELM was first introduced by Huang et al. (Huang, Zhu & Siew, 2006). ELM is a learning algorithm for a single hidden layer feedforward neural network. Compared with the conventional neural network learning algorithm, ELM overcomes the slow training speed. ELM is based on empirical risk minimisation theory, and its learning process needs only a single iteration. The algorithm avoids multiple iterations and local minimisation. Based on Mirjalili and Lewis (2014), they stated that ELM is very efficient for classifications problem and supported by Huang, Huang, Song, & You (2015) and Deng, Huang & Tang (2015) mention that compared to traditional Feed-forward Neural Network learning methods,

TEST Engineering & Management

March-April 2020
ISSN: 0193-4120 Page No. 7740 - 7748

ELM is remarkably efficient and tends to reach a global optimum.

VII. WRAPPER-BASED FEATURE SELECTION USING GENETIC ALGORITHM EXTREME LEARNING MACHINE

GA procedure is begin with a set of solutions represented by a chromosome or a population. Solutions from the initial population are formed a new population. Solutions are selected according to their fitness function which is for this case is the recognition accuracy of the classifier. The higher accuracy obtained from the solution, the more chances for the population to reproduce. The fitness function is the accuracy obtains from the ELM classifier. Two different kernel functions are used to evaluate the performance of the selected subset; Radial Basis Function (RBF) and linear (LIN) kernel. Regularisation coefficient and kernel parameter are optimised and fixed to 40 and 0.1, respectively. This is repeated until the maximum number of generations is satisfied. The best subset is used for improvement using ensemble classifiers. The following steps summarise the procedures of GA-ELM.

1. Generate a random population of n chromosomes
2. Evaluate the fitness function, f(x) from the ELM classifier of each chromosome x in the population.
3. Reproduce a new population by repeating the following steps until the new population is complete:

a. Pick two parents of the chromosomes based on their fitness function

b. Cross over the parents of the chromosomes to form a new children. If no cross over is performed, offspring is the exact copy of the parents.

c. With a mutation, probability mutates new offspring at each locus (position in chromosome).

d. Place new offspring in a new population

e. Use the newly generated population for the next loop

If the maximum number of generations reached, the program will stop and return the best solution in the current population

VIII. WRAPPER-BASED FEATURE SELECTION USING PARTICLE SWARM OPTIMIZATION EXTREME LEARNING MACHINE

Particle Swarm Optimization (PSO) is a population-based search algorithm developed by Eberhart and Kennedy in 1995, inspired by the social behaviour of birds flocking and share many similarities with GA. The fitness evaluation procedure determines the PSO based feature selection. In this paper, the classification accuracy of ELM is used as the fitness function. The following pseudocode illustrates the PSO algorithm:

Step 1: read the feature set from IS10

Step 2: initialise the PSO

Step 3: initialise the number of population

Step 4: calculate the fitness function

Step 5: for loop of the iteration

Step 6: calculate the velocity

Step 7: calculate the fitness function with the updated velocity

Step 8: if fitness is less than the pBest value

pBest is fitness function

Step 9: if pBest is less than gBest

gBest is pBest

Step 10: the best fitness function is gBest

Step 11: end

In a PSO algorithm, the population is initiated randomly with particles and evaluated to compute fitness together. The updated velocity then

Published by: The Mattingley Publishing Co., Inc.

March-April 2020
ISSN: 0193-4120 Page No. 7740 - 7748

determines the particle best value of the individual (pBest) and the best particle in the whole swarm (gBest). Evaluation is again performed to compute the fitness function of the particle swarm optimization. The loop is terminated once the stop criteria is met.

IX. BINARY PARTICLE SWARM OPTIMIZATION EXTREME LEARNING MACHINE

Binary Particle Swarm Optimization (BPSO) is constructed based on Mirjalili & Lewis (2013). The binary version of the PSO is introduced to find the best position by taking the value of binary numbers (1 or 0) with the probability of 0.5. The procedures of PSO is followed in BPSO.

X. ENSEMBLE CLASSIFIERS

Ensemble classifiers is a meta-algorithm that combine several classifiers (usually in odd numbers) with improving the prediction accuracy. There are two ways of ensemble methods, one is by combining learners of the same type leading to homogeneous ensemble classifiers, and the other one is combining several different types of classifiers leading to heterogeneous ensemble classifiers.

The purpose of having an ensemble classifier is to find an agreement between a set of classifiers. The idea behind all ensemble-based systems is that if individual classifiers are diverse, then they can make different errors, and combining these classifiers can reduce the error through averaging (Stepenosky, Green, Kounios, Clark, & Polikar, 2006).

Three classifiers were selected as the base classifiers, Radial Basis Function (RBF) Kernel Extreme Learning Machine (RBF_ELM), Polynomial Support Vector Machine (Poly_SVM) and Linear Kernel Extreme Learning Machine (Lin_ELM). These three classifiers were then combined using eight different ensemble combination techniques: majority voting, maximum, summation, minimum, mean, product, decision templates and Dempster-Shafer.

XI. RESULTS AND DISCUSSIONS

To show the effectiveness of the proposed GA-ELM learning algorithm for RBF and LIN networks, GA-ELM is compared with the Particle Swarm Optimization (PSO) and Binary PSO (BPSO).

Particle Swarm Optimization (PSO) is a relatively new heuristic search method whose mechanics are inspired by the swarming or collaborative behaviour of a biological population. PSO is a population-based search method and it is build up similarly to the Genetic Algorithm (GA). In other words, PSO and the GA move from a set of population to another population in a single iteration with possible improvement using a combination of deterministic and probabilistic rules (Hassan, Cohanim, De Weck & Venter, 2005).

This section attempts to examine whether GA has the same effectiveness as PSO in recognition accuracy. Table II. shows twelve independent simulations using MATLAB platform for GA-ELM, PSO-ELM and BPSO-ELM.

Table II. Comparison Between Ga-elm, pso-elm and bpso-elm

Classifier	No of Pop	Max Gen	No. Of Selected Features	Elite Accuracy (%)
GA-RBF_ELM	20	10	758	65.35
GA-Lin_ELM	20	10	800	67.45
PSO-ELM	20	10	767	64.58
BPSO-ELM	20	10	785	64.16
GA-RBF_ELM	20	30	765	68.65
GA-Lin_ELM	20	30	778	69.76
PSO-ELM	20	30	779	65.29
BPSO_ELM	20	30	808	64.16
GA-RBF_ELM	20	50	702	71.04
GA-Lin_ELM	20	50	771	72.47
PSO-ELM	20	50	788	67.45
BPSO-ELM	20	50	794	65.58

Several populations are fixed to 20 for all runs. Several generations are varied from 10 to 50

Published by: The Mattingley Publishing Co., Inc.

March-April 2020
ISSN: 0193-4120 Page No. 7740 - 7748

generations. It can be observed that the best subset selected is from GA-Lin_ELM with 72.47% of accuracy.

The obtained results by PSO and BPSO show that the fitness values are improved, but as the algorithm continues to diverge from the optimal solution, they may trap in a local optimum. The reason for the divergence can be found in Nezamabadi-pour, Rostami-Shahrbabaki & Maghfoori-Farsangi, (2008). When the algorithm is reached to the optimum solution, the probability of changing the position of the particles is almost zero, while at this point, the position will change by taking the position values with the probability of 0.5. This causes the algorithm not to converge well.

It can be concluded that GA is superior compared to PSO and BPSO. However, the best subset is selected for 50 generations. Considering the execution times of 50 generations, GA and ELM are not recommended for online optimisation of ensemble classifiers. Thus, the single classifier is chosen for offline feature selection, and the best subset is then fed as input to ensemble classifiers for further improvement.

Table III. Shows the heterogeneous ensemble classifiers simulation results. Different feature sets have been employed to investigate the performance of the ensemble classifiers.

Table III Ensemble Classifiers Results

Feature Set	Max Gen	Base Classifier	SD Base Accuracy (%)	SI Base Accuracy (%)
IS10	-	LIN_ELM	97.19	63.53
		RBF_ELM	98.32	56.53
		POLY_SVM	98.59	67.13
RBF_ELM	10	LIN_ELM	95.88	63.91
		RBF_ELM	99.05	65.35
		POLY_SVM	98.39	68.91
LIN_ELM	10	LIN_ELM	96.58	67.85
		RBF_ELM	98.94	62.11
		POLY_SVM	98.43	65.60
RBF_ELM	30	LIN_ELM	96.45	65.04
		RBF_ELM	99.07	68.65
		POLY_SVM	98.41	71.31
LIN_ELM	30	LIN_ELM	96.40	69.76
		RBF_ELM	99.01	62.84
		POLY_SVM	98.54	66.49
RBF_ELM	50	LIN_ELM	95.99	65.80
		RBF_ELM	99.21	71.04
		POLY_SVM	98.33	70.35
LIN_ELM	50	LIN_ELM	96.53	72.47
		RBF_ELM	99.05	63.64
		POLY_SVM	98.38	67.16

Table IV. and V. show the improvements made by the ensemble classifiers using eight combination methods in SD environment. Experimental results obtained in both Table IV. and V. demonstrate overall results on SD show consistency in the recognition accuracy. The consistency proved that the database and methods used are appropriate.

Table IV Ensemble Classifiers Results (SD)

FEATURE SET	COMBINATION METHOD (AVERAGE FOLD ACCURACY %)			
	VOTE	MAX	SUM	MIN
IS10	98.92	97.80	99.09	98.66
GA-RBF_ELM_10	98.90	97.52	99.14	98.43
GA-LIN_ELM_10	99.02	97.95	99.19	98.47
GA-RBF_ELM_30	99.11	97.94	99.13	98.44
GA-LIN_ELM_30	99.09	97.90	99.15	98.60
GA-RBF_ELM_50	99.08	97.87	99.17	98.38
GA-LIN_ELM_50	98.98	97.79	99.07	98.44

Table V Ensemble Classifiers Result (sd) Cont.

FEATURE SET	COMBINATION METHOD (AVERAGE FOLD ACCURACY %)			
	MEAN	PROD	DEC	DEMP
IS10	99.09	99.05	99.10	99.08
GA-RBF_ELM_10	99.14	98.95	99.16	99.05
GA-LIN_ELM_10	99.19	99.07	99.22	99.13

Published by: The Mattingley Publishing Co., Inc.

March-April 2020
ISSN: 0193-4120 Page No. 7740 - 7748

GA-RBF_ELM_30	99.13	98.98	99.21	99.09
GA-LIN_ELM_30	99.15	99.04	99.22	99.13
GA-RBF_ELM_50	99.17	98.88	99.21	99.02
GA-LIN_ELM_50	99.07	98.93	99.16	99.00

Table VI. and VII. show the improvements made by the ensemble classifiers using eight combination methods in SI environment. Experimental results obtained in both Table VI. and VII. demonstrate that ensemble classifiers significantly improved the base classifier's accuracy by a maximum of three per cent. The best recognition improvement is obtained from GA-RBF_ELM_50 using decision template combination method.

Table VI Ensemble Classifiers Results (SI)

FEATURE SET	**COMBINATION METHOD (AVERAGE FOLD ACCURACY %)**			
	VOTE	**MAX**	**SUM**	**MIN**
IS10	67.29	60.95	67.96	67.33
GA-RBF_ELM_10	70.40	66.62	71.58	69.16
GA-LIN_ELM_10	69.38	66.23	70.09	65.82
GA-RBF_ELM_30	73.02	67.80	73.85	71.35
GA-LIN_ELM_30	69.80	67.36	70.47	66.56
GA-RBF_ELM_50	73.82	69.55	74.36	70.65
GA-LIN_ELM_50	70.89	69.53	71.71	67.35

Table VII Ensemble Classifiers Result (SI) Cont.

FEATURE SET	**COMBINATION METHOD (AVERAGE FOLD ACCURACY %)**			
	MEAN	**PROD**	**DEC**	**DEMP**
IS10	67.96	68.56	68.33	68.73
GA-RBF_ELM_10	71.58	70.87	71.47	71.55
GA-LIN_ELM_10	70.09	69.13	70.18	69.73
GA-RBF_ELM_30	73.85	73.56	74.44	74.13
GA-LIN_ELM_30	70.47	69.04	70.36	70.02
GA-RBF_ELM_50	74.36	73.53	75.13	74.07
GA-LIN_ELM_50	71.71	70.51	71.25	71.36

The combination of all three classifiers achieves better classification accuracy and the generalisation performance of the ensemble increases. From the overall result, we obtained 10% improvement of GA-RBF_ELM feature set using ensemble classifiers over the baseline features of IS10, testing on independent speaker environment.

XII. Conclusion

In this paper, wrapper-based GA, PSO and BPSO feature selection with ELM and ensemble classifiers are analysed. Three different type of classifiers are proposed to model a heterogeneous ensemble classifiers. More specifically, we adopt three types of classifiers, namely Radial Basis Function (RBF) Kernel Extreme Learning Machine (RBF_ELM), Polynomial Support Vector Machine (Poly_SVM) and Linear Kernel Extreme Learning Machine (Lin_ELM).

To improve the ensemble recognition accuracies, eight different combination methods have been explored. These methods vary in their approach to treat the training data, the type of algorithms used, and the combination methods followed. Diversity among the performance of every single classifier in the ensemble is essential for combining the predictions from several base classifiers.

Different experiments are followed to introduce diversity among member classifiers. Speaker Dependent (SD) and Speaker Independent (SI) has been tested to validate each base classifier before adding it to the ensemble is used. Ensemble methods have been successfully proven in improving recognition performance on the selected feature set. The overall results show improvements of 10% from the base feature set, IS10. It is well known that the SI system often gives lower accuracy compared to

TEST Engineering & Management
March-April 2020
ISSN: 0193-4120 Page No. 7740 - 7748

the SD system, which shows very consistent results with over 95.00% of accuracy. This is a positive result, and there is room for improvements. Future study may evaluate using more substantial and more natural corpora. Natural Language Processing (NLP) can be implemented to enable real-time applications.

REFERENCES

[1] Goh, B. S., Fadzilah, N., Abdullah, A., Othman, B. F., & Umat, C. (2018). Long-Term Outcomes of Universiti Kebangsaan Malaysia Cochlear Implant Program Among Pediatric Implantees. International Journal of Pediatric Otorhinolaryngology, 105, 27-32.

[2] Miles, S., Khairuddin, K. F., & McCracken, W. (2018). Deaf Learners' Experiences in Malaysian Schools: Access, Equality and Communication. Social Inclusion, 6(2), 46-55.

[3] Alias, A., Azahari, M. H., Kamaruddin, N., Zahari, F., & Razak, F. A. (2019). Images as an Indicator for Deaf Students' Teaching and Learning: An Exploratory Study. Regional Conference on Science, Technology and Social Sciences, 47-53.

[4] Stepenosky, N., Green, D., Kounios, J., Clark, C. M., & Polikar, R. (2006). Majority Vote And Decision Template Based Ensemble Classifiers Trained on Event-Related Potentials for Early Diagnosis of Alzheimer's Disease. International Conference on Acoustics, Speech and Signal Processing (Vol. 5, pp.V-V).

[5] Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: theory and applications. Neurocomputing, 70(1-3), 489-501.

[6] Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Let a Biogeography-Based Optimizer Train Your Multi-Layer Perceptron. Information Sciences, 269,pp.188-209.

[7] Huang, G., Huang, G. B., Song, S., & You, K. (2015). Trends in Extreme Learning Machines: A Review. Neural Networks, 61, 32-48.

[8] Deng, C., Huang, G., Xu, J., & Tang, J. (2015). Extreme Learning Machines: New Trends and Applications. Science China Information Sciences,58(2),p.p1-16.

[9] Hassan, R., Cohanim, B., De Weck, O., & Venter, G. (2005). A Comparison of Particle Swarm Optimization and the Genetic Algorithm. 46th Structures, Structural Dynamics and Materials Conference, p-p.1897.

[10] Abdullah, M. H. L., Hisham, S., & Parumo, S. (2009). MyLexics: assistive courseware for Dyslexic children to learn primary Malay language. ACM SIGACCESS Accessibility and Computing, (95), pp. 3-9.

[11] Mohid, S. Z., & Zin, N. A. M. (2010). Courseware accessibility for hearing impaired. In Information Technology (ITSim), 2010 International Symposium in (Vol. 1, pp. 1-5). IEEE.

[12] Rosdi, F., Mustafa, M. B., Salim, S. S., & Hamid, B. A. (2017). The Effect of Changes in Speech Features on the Recognition Accuracy of ASR System: A Study on the Malay Speech Impaired Children. Malaysian Journal of Computer Science, 30(1), 48-62.

[13] Mirjalili, S., & Lewis, A. (2013). S-shaped versus V-shaped transfer functions for binary particle swarm optimisation. Swarm and Evolutionary Computation, 9, 1-14.

[14] Nezamabadi-pour, H., Rostami-Shahrbabaki, M., & Maghfoori-Farsangi, M. (2008). Binary particle swarm optimization: challenges and new solutions. CSI J Comput Sci Eng, 6(1-A), 21-32.

[15] Jaafar, H., & Ramli, D. A. (2013). Automatic syllables segmentation for frog identification system. In 2013 IEEE 9th International Colloquium on Signal Processing and its Applications (pp. 224-228). IEEE.