

Exploring the Time-efficient and Evolutionary-based Feature Selection Algorithms for Speech Data under Stressful Work Condition

Derry Pramono Adi¹, Lukman Junaedi¹, Frismanda¹,
Agustinus Bimo Gumelar^{1,2}, Andreas Agung Kristanto³

¹Fakultas Ilmu Komputer, Universitas Narotama

²Department of Electrical Engineering, Faculty of Intelligent Electrical and Informatics Technology (ELECTICS), Institut Teknologi Sepuluh Nopember

³Department of Psychology, Universitas Mulawarman

E-mail: derryalbertus@ieee.org, lukman.junaedi@narotama.ac.id, frismanda@fik.narotama.ac.id, *bimogumelar@ieee.org, andreasagungk@gmail.com

Received December 20, 2020; Revised January 21, 2021; Accepted February 23, 2021

Abstract

Initially, the goal of Machine Learning (ML) advancements is faster computation time and lower computation resources, while the curse of dimensionality burdens both computation time and resource. This paper describes the benefits of the Feature Selection Algorithms (FSA) for speech data under workload stress. FSA contributes to reducing both data dimension and computation time and simultaneously retains the speech information. We chose to use the robust Evolutionary Algorithm, Harmony Search, Principal Component Analysis, Genetic Algorithm, Particle Swarm Optimization, Ant Colony Optimization, and Bee Colony Optimization, which are then to be evaluated using the hierarchical machine learning models. These FSAs are explored with the conversational workload stress data of a Customer Service hotline, which has daily complaints that trigger stress in speaking. Furthermore, we employed precisely 223 acoustic-based features. Using Random Forest, our evaluation result showed computation time had improved 3.6 faster than the original 223 features employed. Evaluation using Support Vector Machine beat the record with 0.001 seconds of computation time.

Keywords: Feature Selection Algorithms, Curse of Dimensionality, Speech Data, Work Stress, Evolutionary Algorithm

1. INTRODUCTION

High dimension speech data in machine learning models are not to be taken lightly. In fact, any data are available in hundreds, and arguably thousands of acoustic features (1). According to values in each acoustic feature, indeed, not all of the features hold the relevancy threshold, thus having lower importance of being fed to the machine learning models (2).

If thousands of acoustic features recklessly become the input of a model, it will undoubtedly strain the computational time and computing resources (3). Figure 1 shows a simple visualization of handwritten digit image data from the Modified National Institute of Standards and Technology or MNIST, which has approximately 784 features (4). Figure 1 looked a little bit dense since PCA broke down 784 features into two principal components and nine eigenvectors (5).

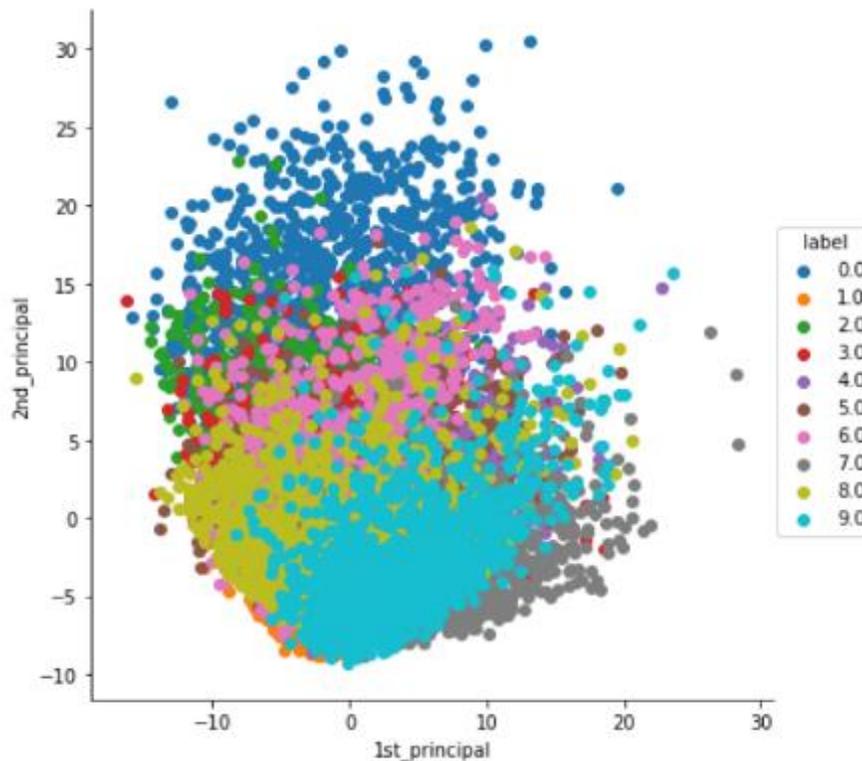


Figure 1. Illustration of High Dimension Data (5)

This research aims to determine the best feature selection algorithms suitable for speech data. Speech recognition in machine learning has become quite an exciting field of study. Typically, we would extract acoustic features from the speech data to determine the information's modality (of semiotics manner) and then feed it to the machine learning algorithms for the specific recognition task (6). Unfortunately, machine learning is a data-driven model. In order to make sense of the semiotics pattern, a large number of data is required, which poses a new problem: the curse of dimensionality (7). A massive load of data means high dimension data, while high dimension data leads to the curse of dimensionality. It is dubbed as a "curse" because having many data worsen the recognition stage instead of improving with new insight into the desired semiotics pattern. Assuming a hundred acoustic features for a hundred sound clips as the data, a simple multiplication will result in slower computation time and too high computing power. (3).

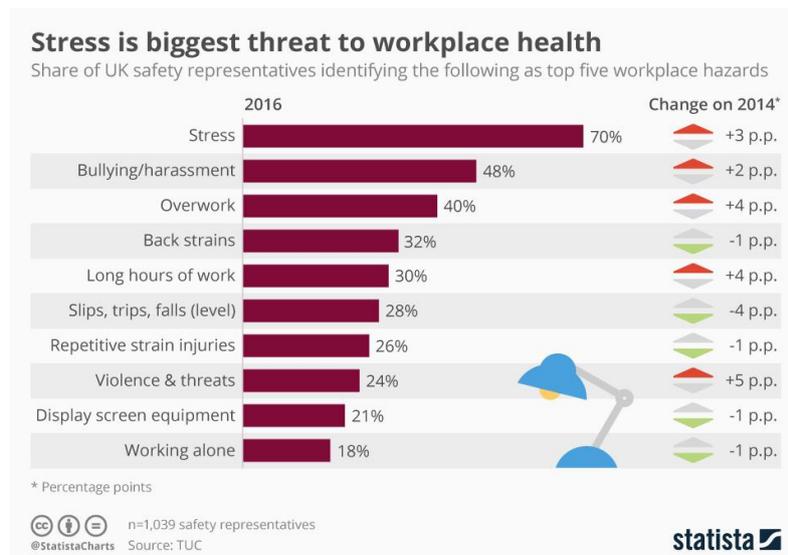


Figure 2. A 2016 Workplace Stress Survey by Statista in the United Kingdom (8)

Many studies showed excellent semiotic pattern recognition performance for the specific case of speech data under stressful work conditions (9). As speech is the fastest form of communication, humans often use it to imply much information (10). This information may vary, but one of them is interesting: stressful speech. A survey in 2016 (shown in Figure 2) says that the everyday and mundane nine-to-five work is stressful for 70% of workers in the United Kingdom (8). Furthermore, in 2017, a total of 65% of 1,210 respondents in a similar survey for interns residing in Scotland would say that they are in a moderately stressed condition (11). The high prevalence of work stress encourages us to dive deeper into understanding stress through the fastest human communication form.

We selected Customer Service hotline records as they deal with day-to-day complaints and are, therefore, in highly stressful conditions. Although the conversation (in a contextual manner) signifies little evidence to stress, however, under stressful situations, humans involuntarily accommodate certain sound qualities as an indication that they are experiencing work stress (12), hence the adoption of speech data and acoustic features in this experiment.

To improve the stress semiotic pattern recognition, we choose to reduce the number of acoustic features in stressful speech data using FSA. In this work, we use the Evolutionary Algorithm (hereinafter be referred to as Evo), Harmony Search (hereinafter be referred to as Harmony), Principal Component Analysis (PCA), Genetic Algorithm (hereinafter be referred to as Gen), and Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Bee Colony Optimization (BCO). These FSAs are believed not to eliminate the insight of stress itself but instead leave significant acoustic features for later classification tasks using the hierarchical models of ML. The mentioned

hierarchical models of ML are Random Forest (RF) and Support Vector Machine (SVM).

The remainder of this work is disclosed as follows: Section 1 describes the importance of FSA in alleviating the curse of dimensionality and the speech data under stressful work conditions. Past results of speech recognition and FSAs, which nearly drive this experiment, are written in Section 2, while Section 3 constitutes this paper's originality. Section 4 elaborates on the grand experiment design in FSA of stressful speech data, as the experiment result is shown in Section 5. Subsequently, Section 6 brings us to the conclusion of the experiment.

2. RELATED WORKS

With ML advancements, FSAs are also getting the spotlight by progressively improved through a meta-model and a new hybrid model. FSA played a significant role in choosing the best features and subset of features while still retaining the information.

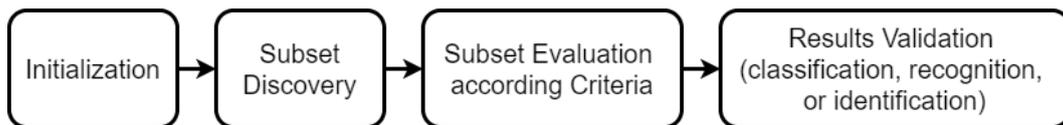


Figure 3. General Process of Feature Selection Algorithms

Figure 3 displayed the general process of FSA. Every FSA employed in this experiment works to find the best subset according to criteria and then have its result validated using ML models. We also showed the ability of hierarchical models of ML in the speech classification task.

2.1. Acoustic Speech Features

For this experiment, we used the jAudio and openSMILE as the feature extractor tool. The jAudio is a comprehensive acoustic and music information retrieval-related feature extractor, while openSMILE is the engine of two vital and glottal-related speech feature extraction: jitter and shimmer.

Our baseline features are the Mel-Frequency Cepstral Coefficient (MFCC) and Linear Predictive Coding (LPC). In this experiment, MFCC and LPC are used as baseline functions, taking into account their success in expressing specific vocal characteristics (13). MFCC calculation formula is depicted in Equation (1), where f is the frequency being converted to Mel scale. On the other hand, the LPC estimation formula is depicted in Equation (2), where \hat{s} is the predicted sample, s is the speech sample, p is the predictor coefficients.

$$Mel(f) = 1125 \ln \left(1 + \frac{f}{700} \right) \quad (1)$$

$$\hat{s}(n) = \sum_{k=1}^p a_k s(n-k) \quad (2)$$

MFCC computation is the replication of a human hearing device in considering an artificial application of the operating theory of the ear, along with the presumption that an accurate speaker recognizer can be correlated with the human ear (14). By comparison, LPC is a common method of voice research. It resembles the vocal tract in humans and represents a robust characteristic of speech. It also analyzes the speech signal by entering the formants, properly disposing of its effects from the speech signal, and is supposed to leave behind a concentration and frequency residue (15).

$$Jitter = \frac{1}{N-1} |t_i - t_{i+1}| \quad (3)$$

Some useful speech characteristics to consider are jitter, shimmer, Harmonic-to-Noise Ratio (HNR), the fundamental frequency (F0), pitch, intensity, loudness, and spectral variance. Jitter can be characterized in terms of absolute and Relative Average Perturbation (RAP) values by calculating fluctuations in the fundamental frequency from one cycle to another (16). Jitter calculation formula is depicted in Equation (3), where N is the total of extracted fundamental frequency, t_i is the time instant at the fundamental frequency period lengths.

$$Shimmer(dB) = \frac{1}{N-1} \sum \left| 20 \log \left(\frac{A_{i+1}}{A_i} \right) \right| \quad (4)$$

Shimmer is a calculation of amplitude variations from peak-to-peak at moments of fundamental frequency (16). Shimmer calculation formula is depicted in Equation (4), where A_i is the peak-to-peak amplitude at the fundamental frequency period.

These features are among the tools used to represent emotions in voice and expression recognizers as well. These two particular features are difficult to obtain, since extraction relies entirely on how exactly the fundamental frequency has been accomplished (17). Essentially speaking, a piece of signature evidence such as jitters, stutter, and perturbation that may come up in a stressful speech can be captured via Jitter and Shimmer features (hence the name). Furthermore, Table 1 showed the list of overall acoustic features for the FSA experiment and the total number of derivation by each features.

Table 1. Overall Acoustic Features and Derivation Number

No	Extracted by	Feature Name	Total Features w/ Derivation(s)
1	openSMILE	Harmonic-to-Noise Ratio	15
2	openSMILE	Fundamental Frequency	15
3	openSMILE	Intensity	15
4	openSMILE	Loudness	15
5	openSMILE	Jitter	30
6	openSMILE	Shimmer	15
7	openSMILE	Zero Crossing Rate	15

No	Extracted by	Feature Name	Total Features w/ Derivation(s)
8	jAudio	Spectral Centroid	6
9	jAudio	Spectral Rolloff	6
10	jAudio	Spectral Flux	6
11	jAudio	Spectral Variability	6
12	jAudio	Root Mean Square Error	6
13	jAudio	Mel-Frequency Cepstral Coefficient	30
14	jAudio	Linear Predictive Coding	24
15	jAudio	Partial-based Spectral Centroid	6
16	jAudio	Partial-based Spectral Flux	6
17	jAudio	Peak-based Spectral Smoothness	6
Total			223

2.2. Feature Selection Algorithms: Evolutionary

In this section, we elaborate on the FSAs in previous research work and their successful combination. We also showed the ability of hierarchical models of ML in the speech classification task. Figure 4 illustrated the general categories of Evo as FSA.

Paradigms: According to Sharma and Kaur, specific nature-inspired metaheuristic algorithms may classify function subsets with special characteristics. They indicated that it is worth exploring if these particular characteristics could be used to create a series of higher quality features (18). Evo and Gen is by far the most natural representation of binary classification, where selected features are represented as 1, and 0 is for the not-selected features (19).

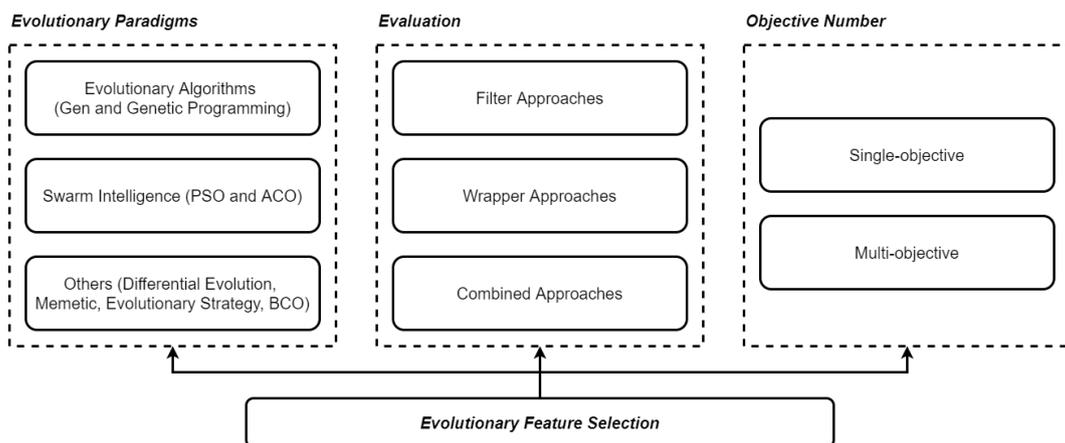


Figure 4. Evo-FSA Diagram of Categories

The PSO is continuously improving the potential approach for consistency estimation, i.e., growing predictive accuracy values. The PSO depicts the population of candidate solutions to the situation. This solution is

called "particles." The GPSO is used in Alba et al. to select a DNA microarray data element (20). Their study incorporates breast cancer type classification using SVM, eventually yield a result of 95.86%.

Chen et al. used the rough set method of ACO in their FSA trial. In truth, shared knowledge is used as a piece of heuristic information. They looked at the Mushroom dataset at the University of California, Irvine (UCI) Repository with the updated ACO classifier and obtained 100% classification accuracy, sadly with 51.39 seconds of computing time (21).

Criteria Evaluation/Assessment: Measurements from various fields, including information-based statistical measures, correlation measures, distance measures, and accuracy measures, have been applied to the filter approaches. In contrast, the chosen features' classification efficiency is used as an assessment criterion for wrapper feature selection approaches. The most common classification algorithms, such as DTs, SVMs, and Neural Networks (NNs), have been used for feature selection wrapper approaches in FSA (22). All filter and wrapper methods will also provide a separate category of approaches. This method means that the assessment process involves all filter and wrapper steps, creating a new framework which is often referred to as hybrid approaches (22).

Objective Number: Depending on the number of objectives, Evo-FSA approaches are split into single-objective and multi-objective approaches. The one that combines the number of features and the classification output of a single fitness function is known as single-objective FSA (23). Many of the current feature selection approaches seek to optimize the classification output either during the initial search or aggregate the classification results and attributes into a single objective attribute. All the multi-objective feature selection algorithms to date are considered the population-based process was generating alternative routes of solutions in a single session (24). In this experiment, we only use the single-objective approach because the "exploring" mission only stops at one speech classification task.

2.3. Feature Selection Algorithms: Harmony and PCA

Harmony is a global evolutionary algorithm inspired by the musicians' method of improvisation. The harmony that is any solution of this algorithm will be preserved in promising solutions called Harmony Memory (HM). The main idea mapping with an illustration was presented to explain how selecting features can be converted into optimization problems and further discussed by the HS algorithm by Diao et al. (25). In Harmony, each musician can vote for a feature to be included in the feature subset when an evolving subset is invented. Harmony is then paired with the vote of all artists, showing the features are being nominated. The whole collection of initial features contains a selection of notes accessible to each musician (26). Many artists are allowed to select the same attribute, and they can choose no attribute at all. The fitness

function used would become a feature subset assessment tool that analyzes and values of each of the new subsets discovered (26).

PCA is an unsupervised type of FSA that extends variables orthogonally into a new space, based on their variance (27). The lower variance characteristics are overlooked. PCA analyzes a data table describing observations represented by various dependent variables, which are typically inter-correlated. Its purpose is to derive valuable information from data that can be used and to convey this information as a series of different orthogonal variables called principal components (27). PCA also reflects the pattern of correlation between measurements and variables by representing them as data points.

2.8. Hierarchical Model

SVM and RF have become the most popular classification algorithms due to the maximum performance and excellent semiotic pattern recognition. Both methods follow hierarchical notion, where the usual DT might have propagates in a top-down fashion, but not for SVM and RF. A hierarchical notion follows the fundamental theory of multi-level modeling systems. Usually, it uses hierarchical structures that represent the generative mechanism being modeled (28).

RF is a baggage approach that treats weak classifiers in a particular manner (29). This simple method uses weak learners, and the inclusion of new weak learners is the same. However, when introducing a new weak classifier, the bagging process of RF searches for the best attribute in a random subset of data instead of unique data that is problematic for the existing collection of weak classifiers (30). In other words, several random-generated yet separate decision trees are combined into one learner. Ironically, the mixture of several randomly sampled data-based DT performs well in practice because randomness allows the model to prevent data overfitting.

SVM is one of the most popular and commonly used supervised learning processes. The basic definition of SVM is based on the case of binary classification. SVM determines the margin that best separates the feature space with the data points dispersed in the feature space (30). The linear division of space is the easiest method for the division of space.

3. ORIGINALITY

In this work, we employed many feature selection algorithms. We used Evo, Harmony, Gen, PCA, PSO, ACO, and BCO. The dataset involved is Customer Service conversation recordings during work hours. As there are massive experiments on implementing FSA in many cases, none is considered the best as it depends on the case and the concern. We specifically employ a wide range of FSAs with the hierarchical classification model. Moreover, the available dataset is considered natural, low-noise, and the classes are company-graded.

Our primary concern is of computation time. ML is a data-hungry advancement and therefore causing high dimension data. Processing raw and high dimension data will inevitably impact the computation time. As well as accuracy, we also compare each of the FSA methods' time to determine the most time-efficient algorithm. Accuracy represents the relevant-still data even with reduced dimension and features mapped into new hyperspace (specifically only done by the PCA). At the same time, faster computation time is inversely proportional to the computation resource.

In short, our experiment can define clear-cut contribution as advancing ML by considering how many features applicable in the one-time running. There is no question of using as many features as possible, but as the computation time and resources are stretched pretty thin, we can look at the advantages of using FSA methods. Moreover, we used under-recognized research in the various information in speech recognition, including stressful speech. The acoustic-based features selected from FSA methods in question can remarkably improve computation time and resources. Figure 5 shown what FSA methods are being used in this experiment and what category our experiment falls into.

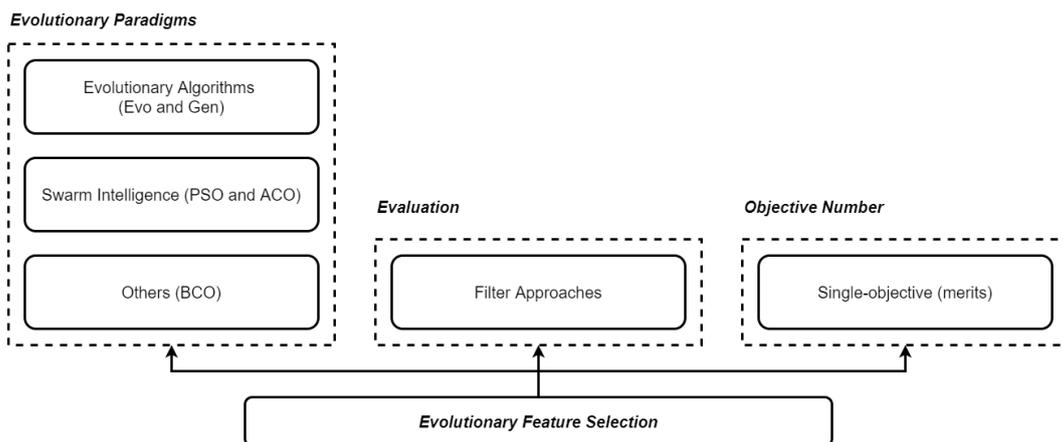


Figure 5. Proposed Work of Evo-FSA

4. SYSTEM DESIGN

Studies have shown the superiority of the hierarchical model, as well as the competency of FSA. Evo is an ultimately sought out FSA, while PCA is the most robust FSA there is. One of the purposes is to form a better-optimized model with the reduced data dimension as the input. In this experiment, the selected acoustic features from the dataset in hand were carried out using various FSA methods. Furthermore, Table 2 showed the on-budget hardware for this particular experiment. There is no need to discuss minimal specifications by showing Table 2, but we experimented on many PCs, in which Table 2 is the lowest hardware specifications. Hence, we are proving the hypothesis that even the lowest hardware has the capability to compute

complex speech recognition problems by applying FSA. In short, our computing resource is relatively low yet up-to-date with the current requirement of basic ML models, thereby strengthening the idea of having the FSA method in the experiment.

Table 2. Hardware Specification

Part	Specification
CPU	AMD Ryzen 7 3700X
GPU	Gigabyte GeForce GTX 1650 Super 4GB OC
RAM	Corsair Vengeance LPX 32GB DDR4
Storage	ADATA SX8200 Pro NVMe M.2 1TB

Moreover, we used a training-testing sample on the standard scale of 60:40. This sampling means we used 60% of the data for the training process and the remaining 40% of the testing process data. With 1,000 data, the training stage has 600 data, while the testing stage has 400 data. The sampling and file access in approximate use of WEKA is depicted in Figure 6. Speaking of data, although we used two different feature extraction software, we have since manually merged them into one Comma Separated Values (*.csv) file. The software openSMILE provides Jitter and Shimmer feature extraction features. Meanwhile, jAudio provides a much more fundamental feature extraction function for MFCC, LPC, and Spectral variance.

SPEECH PROJECTS

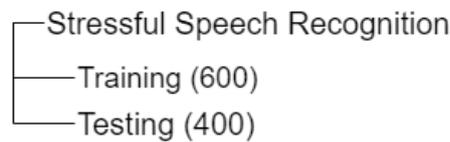


Figure 6. File Access for the Stressful Speech Data

5. EXPERIMENT AND ANALYSIS

Figure 7 depicts our experiment design flow. We firstly collect all sound clips data from the Customer Service department, where all 1,000 sound clips are already graded as either stress or non-stress by the managerial standard. All sound clips are at a sample rate of 8,000 Hz and approximately 1-5 seconds in length.

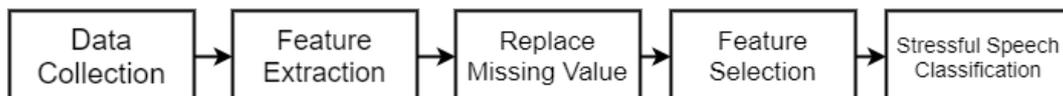


Figure 7. Experiment Design

Then, we extract all features using openSMILE and jAudio. All features have resulted in the *.csv file. The file is then fed to the WEKA. We quickly realized that not all of the feature extraction process yields the desired value.

In fact, some of them did not produce any value. This finding would lead to data imbalance. Therefore, we applied the Replace Missing Value built-in package as soon as the data is inputted.

After all these steps, we finally get into feature selection. All of the FSA methods are tested using the same dataset using data composition of 60:40, in which all result buffer is saved in the form of *.arff files. The files are later loaded accordingly in the classification task. RF has a specific batch task of a hundred, while SVM employs the Polynomial kernel.

5.1 Selected Features

Table 3 shown the selected features from the process of each FSAs. It is ranked from the highest total number of features and derivations selected. Moreover, PCA ranked highest with 51 features selected for further process, while Harmony Search only needs one MFCC feature to move to the classification task.

From Table 3, we can see that four FSA methods, namely Gen, ACO, PSO, and Evo, selected Jitter and Shimmer (and its derivations), which proves our hypothesis, that the values of stressful speech perturbation extracted using Jitter and Shimmer is necessary. All FSA methods, except PCA, also selected at least the MFCC feature, which proves that MFCC is worth extraction for speech recognition problems.

Table 3. Selected Features ranked from Highest Number

FSA	Selected Features	Feature Numbers
PCA	Eigenvectors on its own	51
Gen	HNR, F0, Intensity, Jitter, Shimmer, Spectral Flux, MFCC, LPC, Partial-based Spectral Centroid, Peak-based Smoothness	32
ACO	HNR, Intensity, Loudness, Jitter, Shimmer, ZCR, Spectral Flux, Spectral Variability, MFCC, LPC, Partial-based Spectral Centroid	27
PSO	HNR, F0, Intensity, Jitter, Shimmer, Spectral Flux, MFCC, LPC, Partial-based Spectral Centroid, Peak-based Smoothness	23
Evo	HNR, Intensity, Loudness, Jitter, Shimmer, Spectral Flux, MFCC	16
BCO	Spectral Rolloff, MFCC, LPC	3
Harmony	MFCC	1

5.2 Classification Performance

Based on-specified subset of functions, FSA is used to estimate the efficiency of the classification. Accuracy is the output criterion used to compare the various RF and SVM outcomes with the various reduced features using different FSAs. As time is of the essence, we also considered computation time. Figure 8 showed the classification performance from each FSAs using RF, while Figure 9 showed the classification performance from each FSAs using SVM. From both figures, it is clear that Gen-RF has the highest accuracy of stressful speech classification tasks with 92.7% accuracy. Harmony-SVM has the most time-efficient result of classification, which took 0.001 seconds. In the chart, we have since rounded up the computation time number into 0 for graphical clarity reasons.

Clearly, all result shows excellent performance. The FSA methods greatly improve accuracy and time, but only if the Machine Learning method matched the necessary hyperparameter settings. In Figure 9, although the computation time is considered acceptable, the shown accuracy is a little bit chaotic and underperformed. The Polynomial kernel of the SVM may have caused this.

Also, Figure 8 shown a similar accuracy result, which leads us to believe that the acoustic-based features incorporated in this study have a high similarity value. These findings may lead to increased accuracy but can also indicate underfit data. However, this is not the case for PCA, as it works with reforming features into eigenvectors, ultimately scaling up the data variance in a small number of attributes. Moreover, the RF applying a randomness function seems to expect both underfit and overfit data, therefore automatically handling these data challenges in an instant. Based on the result, our Polynomial SVM seems to lack data handling ability, leaving (suspected) underfit data as is.

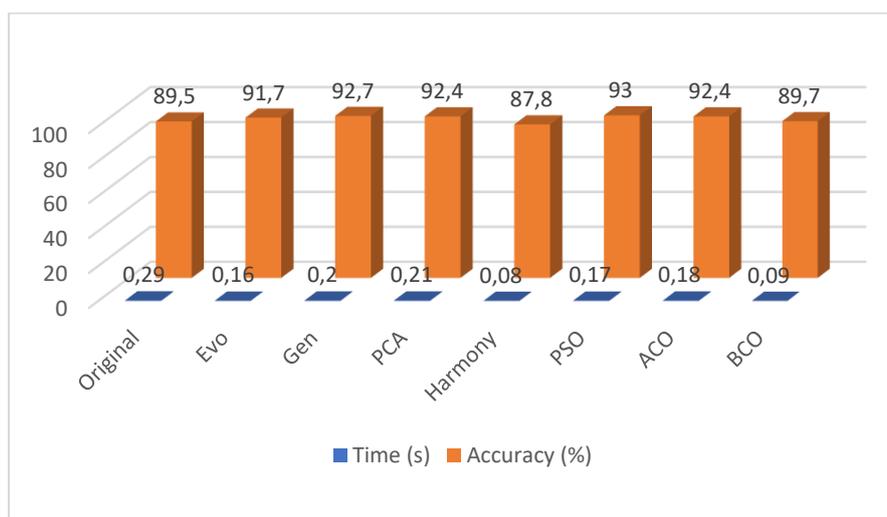


Figure 8. FSAs Evaluation using RF

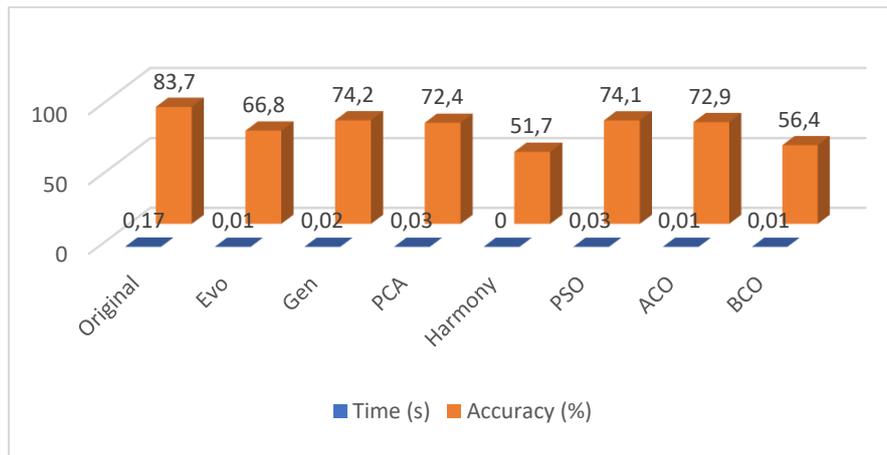


Figure 9. FSAs Evaluation using SVM

6. CONCLUSION

In this work, we used the Evolutionary Algorithm, Genetic Search, Principal Component Analysis, Harmony Search, Particle Swarm Optimization, Ant Colony Optimization, and Bee Colony Optimization as the Feature Selection Algorithm (FSA) in the case of stressful speech data. The main goal of FSA is to alleviate the curse of dimensionality and select the most relevant features. It dramatically improves computation time and resources. This experiment proves that FSA did improve computation time. Evolutionary-based FSA, however, showed the best result using Random Forest. Genetic Search yields 92.7% of stressful speech classification, with 0.2 seconds of computation time. All FSAs under Random Forest have improved classification accuracy. Furthermore, the majority of FSAs have proved the importance of employing Jitter and Shimmer acoustic features. But in all fairness, all FSA (except PCA on using its eigenvectors) chose to employ at least one MFCC.

Unfortunately, although all results showed improved computation time, FSAs still show inadequate accuracy, many of which are resulted from SVM classification. The computation time did not necessarily in the play of improving accuracy. This result encourages us to tinker further with the SVM kernel and hyperparameters to choose the best settings. In the future, we also would like to use Evo-FSA with other approaches such as the wrapper approach and multi-objective Evolutionary algorithm.

REFERENCES

- [1] Altman N, Krzywinski M. **The Curse(s) of Dimensionality**. *Nat Methods*. 2018;15(6):399–400.
- [2] Syarif I. **Feature Selection of Network Intrusion Data using Genetic Algorithm and Particle Swarm Optimization**. *Emit Int J Eng Technol*. 2016 Dec 15;4(2). Available from: <https://emitter.pens.ac.id/index.php/emitter/article/view/149>

- [3] Mohammadi FG, Amini MH, Arabnia HR. **Evolutionary Computation, Optimization, and Learning Algorithms for Data Science**. Optim Learn Control Interdependent Complex Networks Adv Intell Syst Comput. 2020;1123:37–65.
- [4] LeCun Y, Cortes C. **MNIST handwritten digit database**. 2010; Available from: <http://yann.lecun.com/exdb/mnist/>
- [5] Arora M. **Feature Extraction-Principal Component Analysis**. Medium. 2019.
- [6] Gharavian D, Sheikhan M, Ghasemi SS. **Combined Classification Method for Prosodic Stress Recognition in Farsi Language**. Int J Speech Technol. 2018;21(2):333–41.
- [7] Bäck T, Fogel DB, Michalewicz Z. **Evolutionary computation 1: Basic algorithms and operators**. CRC press; 2018.
- [8] Armstrong M. **Stress Is Biggest Threat To Workplace Health**. Stress and burnout - Statista. 2016.
- [9] Ahmad J, Sajjad M, Rho S, Kwon S il, Lee MY, Baik SW. **Determining Speaker Attributes from Stress-affected Speech in Emergency Situations with Hybrid SVM-DNN Architecture**. Multimed Tools Appl. 2018;77(4):4883–907.
- [10] Adi DP, Gumelar AB, Arta Meisa RP. **Interlanguage of Automatic Speech Recognition**. In: 2019 International Seminar on Application for Technology of Information and Communication (iSemantic). IEEE; 2019. p. 88–93.
- [11] Stewart C. **Distribution of stress levels at work by those in paid employment and training programs in Scotland in 2017, by gender**. State of Health - Statista. 2020.
- [12] Holland P, Collins AM. **“Whenever I can I push myself to go to work”: a qualitative study of experiences of sickness presenteeism among workers with rheumatoid arthritis**. Disabil Rehabil. 2018 Feb;40(4):404–13.
- [13] Venkataramanan K, Rajamohan HR. **Emotion Recognition from Speech**. arXiv Prepr arXiv191210458. 2019 Dec;
- [14] Ahmad J, Fiaz M, Kwon S, Sodanil M, Vo B, Baik SW. **Gender Identification using MFCC for Telephone Applications - A Comparative Study**. 2016;3(5).
- [15] Deb S, Dandapat S. **Classification of Speech Under Stress Using Harmonic Peak to Energy Ratio**. Comput Electr Eng. 2016;55:12–23.
- [16] Chadha AN, Zaveri MA, Sarvaiya JN. **Optimal feature extraction and selection techniques for speech processing: A review**. In: 2016 International Conference on Communication and Signal Processing (ICCSP). IEEE; 2016. p. 1669–73.
- [17] Slavich GM, Taylor S, Picard RW. **Stress Measurement Using Speech: Recent Advancements, Validation Issues, and Ethical and Privacy Considerations**. Stress. 2019;22(4):408–13.

- [18] Sharma M, Kaur P. **A Comprehensive Analysis of Nature-Inspired Meta-Heuristic Techniques for Feature Selection Problem.** Arch Comput Methods Eng. 2020 Feb;
- [19] Chen X, Bai R, Qu R, Dong H, Chen J. **A Data-Driven Genetic Programming Heuristic for Real-World Dynamic Seaport Container Terminal Truck Dispatching.** In: 2020 IEEE Congress on Evolutionary Computation (CEC). IEEE; 2020. p. 1–8.
- [20] Talbi E-G, Jourdan L, Garcia-Nieto J, Alba E. **Comparison of Population based Metaheuristics for Feature Selection: Application to Microarray Data Classification.** In: 2008 IEEE/ACS International Conference on Computer Systems and Applications. IEEE; 2008. p. 45–52.
- [21] Chen Y, Miao D, Wang R. **A Rough Set Approach to Feature Selection based on Ant Colony Optimization.** Pattern Recognit Lett. 2010 Feb;31(3):226–33.
- [22] Rodriguez-Galiano VF, Luque-Espinar JA, Chica-Olmo M, Mendes MP. **Feature selection approaches for predictive modelling of groundwater nitrate pollution: An evaluation of filters, embedded and wrapper methods.** Sci Total Environ. 2018 May;624:661–72.
- [23] Yuxin Z, Shenghong L, Feng J. **Overlapping community detection in complex networks using multi-objective evolutionary algorithm.** Comput Appl Math. 2017 Mar;36(1):749–68.
- [24] Sarkar S, Das S, Chaudhuri SS. **Multi-level thresholding with a decomposition-based multi-objective evolutionary algorithm for segmenting natural and medical images.** Appl Soft Comput. 2017 Jan;50:142–57.
- [25] Diao R, Shen Q. **Feature selection with harmony search.** IEEE Trans Syst Man, Cybern Part B. 2012;42(6):1509–23.
- [26] Yusup N, Zain AM, Latib AA. **A review of Harmony Search algorithm-based feature selection method for classification.** J Phys Conf Ser. 2019 Mar;1192:12038.
- [27] Jasim YA, Al-Ani AA, Al-Ani LA. **Iris recognition using principal component analysis.** Proc - 2018 1st Annu Int Conf Inf Sci AiCIS 2018. 2019;5(5):89–95.
- [28] Shrestha A, Mahmood A. **Review of Deep Learning Algorithms and Architectures.** IEEE Access. 2019;7:53040–65.
- [29] Couronné R, Probst P, Boulesteix A-L. **Random forest versus logistic regression: a large-scale benchmark experiment.** BMC Bioinformatics. 2018 Dec;19(1):270.
- [30] Cho G, Yim J, Choi Y, Ko J, Lee S-H. **Review of Machine Learning Algorithms for Diagnosing Mental Illness.** Psychiatry Investig. 2019 Apr;16(4):262–9.