Feature Selection of EEG signals for Emotional State Classification in Machine Learning

Zin War Oo[1] and Khine Khine Oo[2]
[1]No. 13, 87th Street, Mingalar Taung Nyunt Township, Yangon 11221, Myanmar,
Email: zinwaroo@ucsy.edu.mm
[2]Faculty of Information Science
University of Computer Studies, Yangon, Shwe Pyi Thar, 11411, Myanmar,
Email: khinekhineoo@ucsy.edu.mm

KEY WORDS: electro-encephalography (EEG), feature selection, embedded feature selection methods, Lasso and Ridge Regulation methods.

ABSTRACT: Emotional health is very important to build our life. It is equally as important as our physical health. Dealing with our emotions is a difficult task because we can't see them. To analysis emotional state is also the interested field of the researcher. We can test the brain wave to analysis the emotion using electro-encephalography (EEG) signals. There are many kinds of emotions. We propose to classify the human brain wave for happy, disgust, surprise, anger, sad and fear. To classify six types of human emotion needs data annotation, feature extraction, feature selection and classification methods by analyzing electroencephalography (EEG) signals. We propose the embedded feature selection methods to select the least number of features in order to increase accuracy and decrease the cost of data classification. In this propose model used the combination of the Lasso and Ridge Regulation methods to reduce the freedom of the model. The input dataset are from the Database for Emotion Analysis using Physiological Signals (DEAP). In this model propose the high accuracy by comparing with the previous method.

1. MANUSCRIPT

Emotion becomes from the strong feeling. That feeling derive from outside of our body which one's circumstances, mood, or relationships with others (Ekman, 1992). Many kinds of human emotions are effected to response the significant internal and external events. Most of the strong actions are influences by the emotion that we feel. So many kinds of emotions are happened in our daily lives. These emotion play a main role in how we feel inside our body and how we behave individually and socially to the environment. There are many psychologists try to identify the different types of emotions. Basically, emotions can be identified happiness, sadness, disgust, fear, surprise, and anger (Ekman, 1992).

Human's emotion is come out from the brain and deep affect to the soul because of the high thinking of the brain and feeling of the soul. All emotions are come out from the brain's limbic system. That is placed near brain stem and about the size of a walnut (Nolte, 2002). The limbic system controls emotion and other brain functions related to our instincts and memories.

There are so many kinds of methods to take the signals from the brain. To understand some of the processes of the brain like working memory, translating languages and sensory perception emotion to the dynamics in the brain can compute by studying the electricity of the brain electromagnetic waves. We can measure that waves using EEG electroencephalography that directly connect with scalp on the head. EEG measures ionic current voltage which are come out from within the neurons of the brain (Niedermeyer, 2004). EEG records the electrical activity of the brain's spontaneous over a definite of time and then record the wave signals of multiple electrodes placed on the scalp.

There are four steps to classify emotions state from the EEG signals: data annotation, feature extraction, feature selection, and classification. Data annotation can be labeled the type of emotions and the state of calm. The data analysis on the EEG signals is time-consuming or high computational analysis need to generated features for each window length are large in numbers. Based on the maximum relevance and the minimum redundancy relation of the features are selected for feature vector calculation. Feature selection is an area of the active research for the pattern recognition of the signals. Selection of the appropriate features is important for reduced computational cost, response time and performance of the system. To improve and simplify the quality of the dataset, it is needed to select the salient features of high-dimensional datasets. For this purpose Feature Selection (FS) is used in the machine learning technology. The spurious features are deleted from the original dataset of the features by using FS without sacrificing generalization performance. Normally the existence of the FS is the essential role in the real-world problems for the factors: abundance of noise, spurious information, irrelevant and redundant features in the original feature set (Harpale, Bairagi 2018).

In the feature selection methods, the redundant features are usually removed from dataset because of there is a subset of the other features. The noise features that do not provide any information about labels should also be removed because they will reduce the efficiency of the algorithm. The relevant features which consist of the significant information about given dataset of the features will remain. The methods for identifying diverse features, calculating relationships between features and selecting relevant features are needed through a very huge amount of data (Niedermeyer, 2004).

Embedded Filter models are computationally efficient, but don't take the biases of the learning algorithm. By comparing with the filter models, wrapper models obtain better predictive accuracy estimates. However, wrapper models are very computationally expensive. Embedded models are a trade-off between the two models by embedding the feature selection into the model construction. Thus, embedded models take the advantage of both filter models and wrapper models. Embedded models are far less computationally intensive than wrapper methods, since they don't need to run the learning models many times to evaluate the features, and they include the interaction with the learning model (Suhang, Jiliang, Huan, 2016). The accuracy of the propose method is higher comparing with the training data of predefine dataset.

2. RELATED WORK

One of the research paper propose the feature selection for motor imagery EEG classification based on firefly algorithm to eliminate redundant features and to improve the classification accuracy. Firefly algorithm (FA) can consistently select the best subset of features but it is easily entrapped in a local optimum. To avoid this problem, used the combining method of the firefly algorithm (FA) and learning automata (LA) to optimize feature selection for motor imagery EEG. That paper compared with the genetic and adaptive weight particle swarm optimization algorithms to show the effectiveness of the propose method. At that paper, a real-time brain-computer interface system was implemented to prove the feasibility of the proposed methods being used in practical brain-computer interface systems (Aiming, 2017).

Another research paper proposed an adaptive method for feature selection and extraction for classification of epileptic EEG signal in significant states. That research paper identify pre-seizure state and seizure state of EEG signal using time and frequency features. It used fuzzy classifier to predict seizures. The methodology used the hypothetical testing to refine feature selection and pattern adapted wavelet transform to improve classification. The CHB-MIT EEG scalp dataset from Children's Hospital, Boston is used for experimentation to shows the high classification accuracy ((Harpale, Bairagi 2018)).

Another research paper proposed the feature selection model based on EEG signals for assessing the cognitive workload in drivers. That paper present a new feature selection model for the pattern recognition of EEG signals based on machine learning technique called GALoRIS. It combines Genetic Algorithms and Logistic Regression. That identify the data related to high and low cognitive workloads of subjects extracted the information from multiple EEG signals. That maximizing the model's predictive capacity, receiving a precision rate greater than 90% (Patricia, 2020).

3. PROPOSED METHODOLOGY

To classify the human emotion sate from the EEG signals, there will be so many kinds of signals. EEG detects the neutrons emitted from the human brain by directing scalp on the so many places on the brain. The signals of emotion are come out from the brain limbic system at the below back of the brain. In this paper, used the EEG signals come out from the place near the brain stem. To classify the human emotions, feature extraction need to do first. In this propose system will used the features that are the result of Logistic Regression classification algorithm for the input of feature selection methods.

In this proposed method used DEAP dataset (Database for Emotion Analysis using EEG Physiological Signals) (https://www.eecs.qmul.ac.uk) to show the consistency. It consists of two parts:

i . The rated results of the people who watch the video based on arousal, valence and dominance.
ii. The rating of the participant, physiological recordings and frontal face video of an experiment.

The participants watched the music videos and discrete 9-point scale for valence and arousal are rated. The strongest volunteer ratings and a small variation is selected to maximize the strength of elicited emotions. For each video *x* normalized arousal and valence score are calculated by mean rate divided by standard deviation (μ_x/σ_x) (Koelstra, 2011). The system flow diagram of our propose system is shown in Figure 1.

Logistic regression is the technique for classification problem by estimating the values of parameters coefficients. At the end of the training of the machine learning model, we got the function that best described the relationship between the input and the output values. The prediction of the output is transformed using a nonlinear function called the sigmoid function as well as the logit function developed by a statistician to describe probability of the result. Each object in the image is assigned a probability between 0 and 1. Logistic regression technique can be used for traditional statistics machine learning (Tolles, 2016).

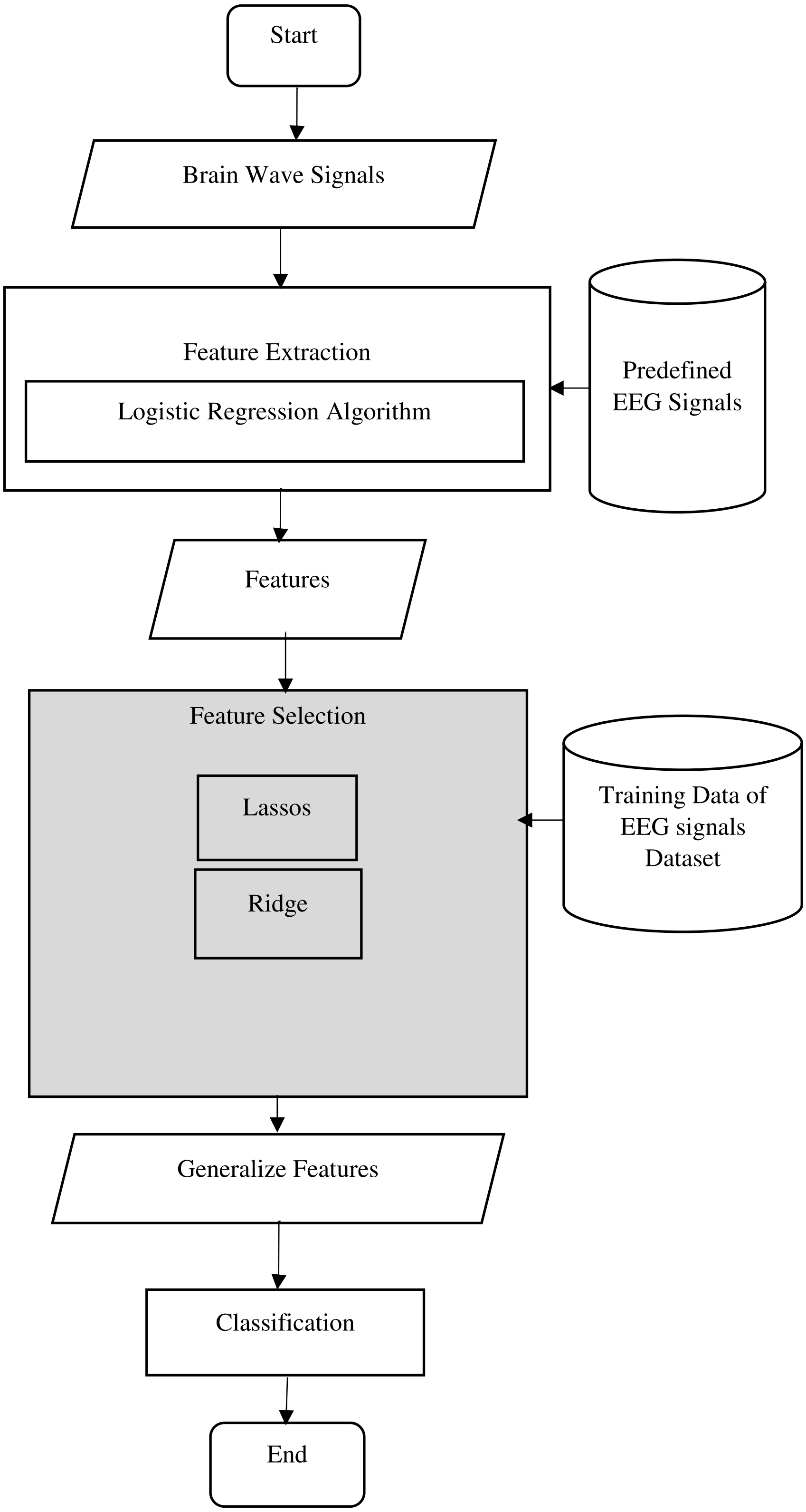

Figure 1. System flow diagram for feature selection of EEG signals

Feature selection is one of the main part. By using the feature selection methods that can be shorter training time, easier to interpret the model, reduce overfitting and improves accuracy. There are so many methods to do feature selections such as filter methods, wrapper methods and embedded methods for machine learning. In this propose model based on the embedded methods because it is faster than wrapper methods and more accurate than filter method. Below the embedded methods there were included regularized regression models that is a technique that regularize the estimates or shrink the co-efficient towards zero and slight modification to the least square estimation. (Tolles, 2016). Regularization adds the different parameters of the model to reduce the freedom of the model. The model will be less likely to fit the noise of the training data and it will be improve the generalization abilities of the model.

There are basically three types of embedded methods:

- L1 regularization (Lasso - Least Absolute Shrinkage and Selection Operator)
- L2 regularization (Ridge)
- L1/L2 regularization (Elastic net)

$$\text{RSS} = \sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{P} \beta_i\, x_{ij}\right)^2 \qquad (1.1)$$

RSS = Regularized regression
β co-efficient ≅ 0,
y_i = target,
P = number of predictors bit,
x_{ij} = set of features or fill top predictors,
β_0, β_i = set of estimates

We have assumed in the equation above the data-set has M instances and P features. The training data set of the DEAP are used whether the model is suffering from over-fitting or under-fitting. If there have very few features on a data set the score is poor for training set then there will be a problem of under-fitting. On the side if we have large number of features and training score is relatively poor than it's the over-fitting. To prevent over-fitting and to reduce model complexity we propose the simple techniques Lasso and Ridge regression.

In Lasso regularization method adds a penalty equal to the sum of the absolute value of the coefficients used for variable selection to find some of the β exactly zero for unimportant variables. Lasso embedded method extern the RSS equation like:

$$\text{RSS} + \lambda \sum_{j=1}^{P} |\beta_j| \qquad (1.2)$$

The λ parameter is the regularization penalty. In the Ridge embedded method add the RSS equation with another term and minimize the whole things to ensure β value to be low value.

$$\text{RSS} + \lambda \sum_{j=1}^{P} \beta_j^2 \qquad (1.3)$$

In this propose model, combine the results of Lasso and Ridge regularization method to get the more accurate result for the classification.

4. RESULT

The five different rhythms from the EEG signals is detected by segmenting the frequency. The signal becomes 5 chunk signals. The training dataset of DEAP datasets (Pan, 2019) contains all together 40 channels recorded including with 32 channels for EEG. Ridge regression shrinks the coefficients to reduce the model complexity and collinearity. Figure 2 show the result of the feature from Lasso regression and the Lasso-Ridge regression methods.

There are so many feature selection methods. Among them L1-SVM (L1-based linear support vector machine), EN (Elastic Net), L1-LGR (L1- based logistic regression), and GBDT (Gradient Boosting Decision Tree) are the most famous and most of the researcher used these methods for their system (Pan, 2019). Figure 3 show the comparison of the different feature selection methods.

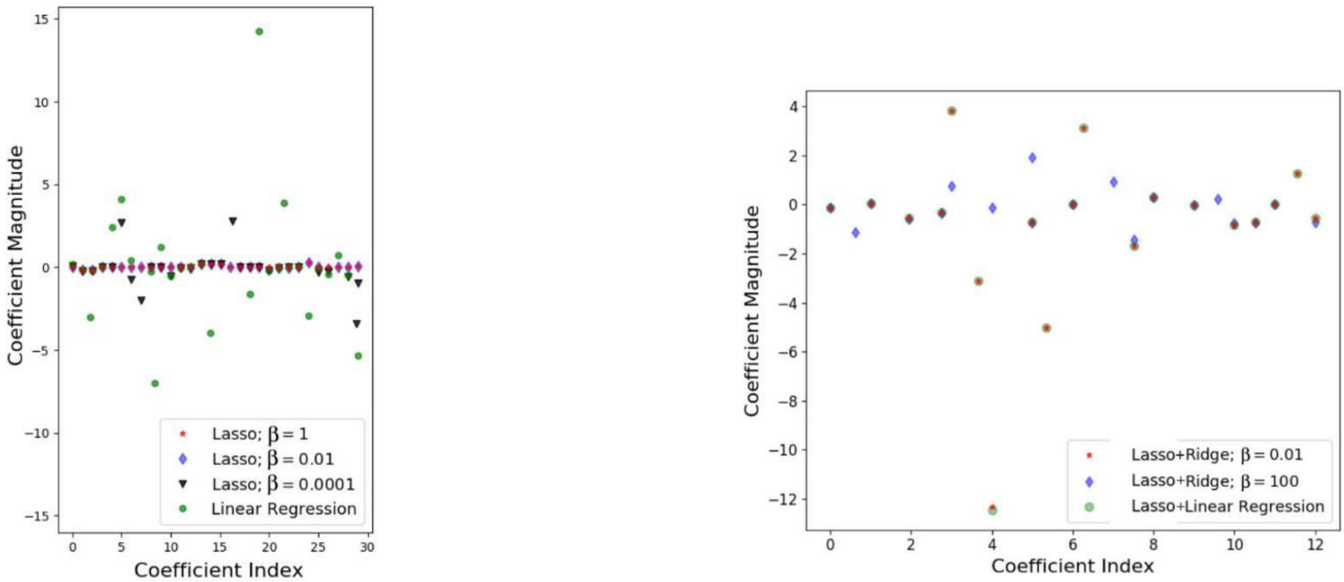

Figure 2: The result features from the Lasso and Lasso+Ridge regression method

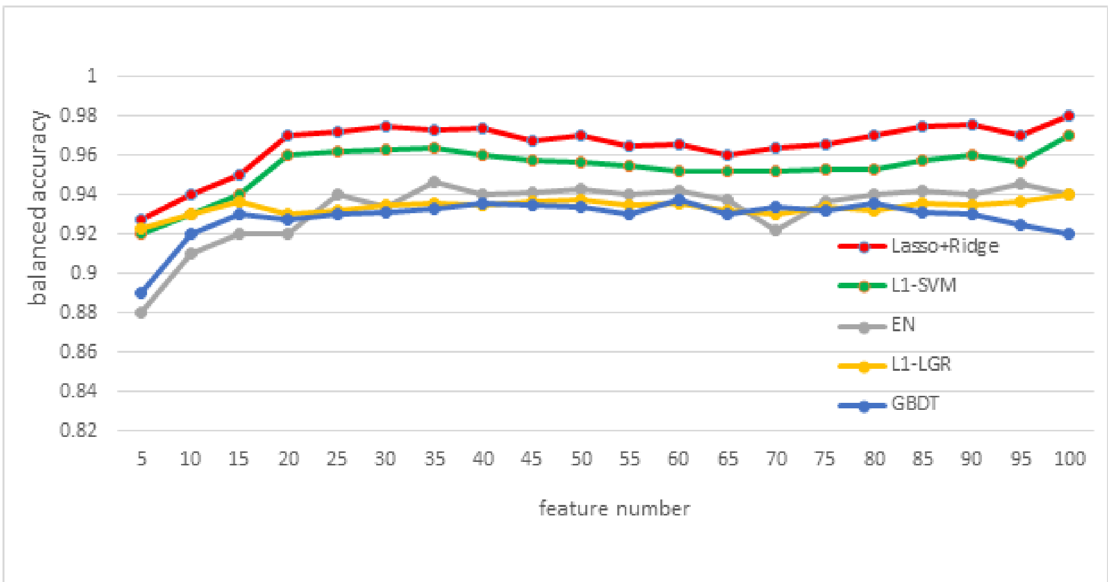

Figure 3: Comparison of the different feature selection methods with different feature numbers

5. CONCLUSION

To know the state of human emotion is one of the role to treat the patients. This research is intended to achieve higher classification accuracy using machine learning algorithm. In this paper, we proposed the feature selection algorithm by combining Lassos and Ridge regression methods. Our propose system output may be the effective machine-learning methods for feature selection to identify the EEG signal. Our study show the higher accuracy by comparing the feature selection methods. In future work, we will improve the feature classification method to classify the feature and to output the type of emotions. In this research will demonstrate that the best selecting for the low features to classify the right emotional state of the human.

References from Journals:

Aiming, L., Kun, C., Quan, L., Qingsong, A., Yi Xie and Anqi, C., 2017. Feature Selection for Motor Imagery EEG Classification Based on Firefly Algorithm and Learning Automata. www.mdpi.com/journal/sensors.

Harpale, V., Bairagi, V., 30 April 2018. An adaptive method for feature selection and extraction for classification of epileptic EEG signal in significant states. Journal of King Saud University, Computer and Information Sciences.

Patricia Becerra-Sánchez , Angelica Reyes-Munoz and Antonio Guerrero-Ibañez, 17 October 2020. Feature Selection Model Based on EEG Signals for Assessing the Cognitive Workload in Drivers. www.mdpi.com/journal/sensors.

Saito, T., Rehmsmeier, M., 2015. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. PLoS One.

References from Books:

Ekman, P., 1992. Are there basic emotions? Psychological Review, 99(3), 550-553.

Niedermeyer, E., Da Silva, F.L., 2004. Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincott Williams & Wilkins. ISBN 978-0-7817-5126-1.

Nolte, J., 2002. The Human Brain. 5th ed. St. Louis: Mosby.

References from websites:

Koelstra, S., Muhl, C., Soleymani, M., Jong-Seok, L., December 2011. DEAP: A Database for Emotion Analysis using Physiological Signals. IEEE Transactions on Afective Computing.

Ozerdem MS, Polat H., 2017. Emotion recognition based on EEG features in movie clips with channel selection. Springer Brain Informatics.

Pan Sun, (Student Member, Ieee), Defeng Wang, Vincent Ct Mok, And Lin Shi, August 9, 2019. Comparison of Feature Selection Methods and Machine Learning Classifiers for Radiomics Analysis in Glioma Grading. IEEE.

Suhang Wang, Jiliang Tang, and Huan Li, 2016. Feature Selection. Arizona State University, Tempe, AZ, USA, Michigan State University, East Lansing, MI,USA, Springer Science+Business Media New York.

Tolles, Juliana, Meurer, William, J., 2016. Logistic Regression Relating Patient Characteristics to Outcomes.