engineering proceedings

MDPI

Proceedings Paper

Feature Selection Based on Evolutionary Algorithms for Affective Computing and Stress Recognition †

Dilana Hazer-Rau [1,*], Ramona Arends [1], Lin Zhang [2] and Harald C. Traue [1]

[1] Section Medical Psychology, University of Ulm, Frauensteige 6, 89075 Ulm, Germany; ramona.arends@uni-ulm.de (R.A.); harald.traue@uni-ulm.de (H.C.T.)
[2] School of Psychology, Central China Normal University, No.152 Luoyu Road, Wuhan, China; lin_zhang@ccnu.edu.cn
* Correspondence: dilana.hazer@uni-ulm.de

† Presented at the 8th International Electronic Conference on Sensors and Applications, 1–15 November 2021; Available online: https://ecsa-8.sciforum.net.

Abstract: In the area of affective computing, machine learning is used to recognize patterns in datasets based on extracted features. Feature selection methods are used to select the most relevant features from the large number of extracted features. This paper presents a feature selection approach based on evolutionary algorithms using techniques inspired by natural evolution. Our proposed method consists of the steps Initialize, Mutate, Crossover, Evaluate and Select. First, an initial population consisting of a set of individuals is generated and the fitness of every individual is evaluated to select the fittest individuals for the next steps. Mutation sets used attributes to unused attributes and vice versa, while with crossover one part of one individual is crossed over with another part of the other individual. The performance of new individuals is evaluated by calculating the accuracy of the prediction and a selection scheme is adopted to select new individuals according to their fitness for the next generation. The feature selection method with evolutionary algorithms is integrated within our previously developed workflow for affective computing and stress recognition from biosignals and is evaluated using our uulmMAC database for machine learning applications. Our proposed approach is faster than the forward selection method and does not stop at a local optimum, allowing a promising feature selection alternative in the field of affective computing.

Keywords: feature selection; evolutionary algorithms; machine learning; affective computing; stress recognition; emotion recognition; biosignals; psychophysiology

Citation: Hazer-Rau, D.; Arends, R.; Zhang, L.; Traue, H.C. Feature Selection Based on Evolutionary Algorithms for Affective Computing and Stress Recognition. *Eng. Proc.* **2021**, *3*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Dilana Hazer-Rau

Published: 1 November 2021

1. Introduction

Machine learning enables the artificial generation of knowledge based on intelligent training of data. An artificial intelligent system learns from known (training) data and applies the gained knowledge to unknown (test) data. In the area of affective computing, machine learning is used to recognize emotion-related patterns in datasets based on specific features extracted from different modalities such as facial, speech, text or biosignal information. The feature extraction step is a main key task within the recognition process as it delivers significant information related to a specific affective state. However, the large number of features that can be extracted from specific data might also lead to inefficient classifications in terms of recognition rates and computation time [1]. Therefore, feature selection methods are used in the next step to select the most relevant and non-redundant features from the large number of extracted features. This step is crucial in the recognition workflow process to achieve optimal computations by enhancing the speed of the algorithms and increasing the rate of the recognition.

There are many feature selection methods available based on different strategies. In our previously developed processing workflow for affective computing [2], three feature selection methods were presented including: Forward Selection, Backward Elimination and Brute Force methods. In terms of computation time, it was mentioned that the Brute Force is usually the last option to be used, as it tries all possible combinations of features in order to select the ones leading to the highest performance. But also the Backward Elimination, which starts with the whole set of extracted features and subsequently removes features, as well as the Forward Selection, which begins with an empty selection of features and subsequently adds features, are associated with high computational time depending on the classification method used.

Therefore, in this paper, we present another feature selection approach based on evolutionary algorithms to further optimize the computational process within our recognition workflow. Evolutionary algorithms is a generic term for a number of different procedures that use Darwinian-like evolutionary processes to solve difficult computational problems. They are based on the Darwinian principle using techniques inspired by natural evolution, such as inheritance, mutation, selection and crossover [3]. In the 1960s scientists started to study evolutionary systems to solve optimization problems [4]. Genetic algorithms belong to the larger class of evolutionary algorithms being a search heuristic that mimics the process of natural evolution. Genetic Algorithms as a part of evolutionary algorithms were introduced by Holland to generate solutions to optimization problems [5]. Since then, evolutionary approaches have been adopted in many studies, for instance in the field of multimodal pain recognition [6] or for improved diagnostic ability of beat-to-beat variability analysis [7]. Studies comparing different feature selection methods were also conducted to investigate which strategy delivers the best classifications. It was shown that for four out of five datasets used, the best results were obtained using optimized selection with genetic algorithms [8].

In the following, we present the implementation of a forward selection method based on evolutionary algorithms and describe its integration within our previously developed workflow for affective computing and stress recognition [2]. Then, we evaluate this approach using biosignal data from our uulmMAC dataset [9] and finally discuss some options for future optimizations.

2. Materials and Methods

Our feature selection method with evolutionary algorithms is based on a genetic algorithm that uses techniques inspired by natural evolution, such as mutation, crossover and selection. In the context of feature selection, *mutation* denotes switching features on and off, while *crossover* denotes interchanging used features. *Selection* is achieved using a specified selection scheme parameter [10].

Given a clearly defined problem to be solved and a bit string representation for candidate solutions, a simple genetic algorithm works as follows [11]:

- Start with a randomly generated population of n parent individuals, where each individual represents a solution to a problem.
- Calculate the *fitness* (accuracy of the prediction, stating how good the individual solves the problem) of each parent individual in the population.
- Repeat a set of steps including mutation, crossover, evaluation and selection, until n *offspring* (mutated and/or recombined version of the parent individuals, also synonym for all generated child individuals) has been created.

Each iteration of this process is called a generation. A genetic algorithm is typically iterated for 50 to 500 or more generations.

The proposed method is realized using the "Optimize Selection (Evolutionary)" operator from RapidMiner, which uses a genetic algorithm to select the most relevant features of a given dataset. It consists of the steps *Initialize*, *Mutate*, *Crossover*, *Evaluate* and *Select* and is implemented as follow:

Initialize: First an initial population consisting of p individuals is generated, in which every individual is a vector of a randomized set of attributes (features). In our example the population size parameter *p* is set to 20 and each individual has a minimum and maximum size of attributes of 3 and 10, respectively. Each attribute is switched on with probability defined with the p-initialize parameter, set in our example to *p_i* = 0.5.

Mutate: For all individuals in the population, mutation is performed by setting used attributes to unused with probability *p_m* and vice versa. The probability *p_m* is defined by the p-mutation parameter, given typically as a very small rate [11]. In our case, we set the mutation rate to *p_m* = -1.0, which is equivalent to a probability of $1/n$ where n is the total number of attributes. Mutation allows adding new child individual information while slightly changing the parent individual.

Crossover: Crossover for interchanging used features is performed on two individuals chosen from the population, with probability *p_c*. The probability *p_c* is defined by the p-crossover parameter and is set to *p_c* = 0.5. The type of crossover is defined by the crossover type parameter and is set to uniform. In uniform crossover, we select two individual for crossover and assign *heads* to one parent and *tails* to the other. We then flip a coin for each position for the first child and make an inverse copy for the second child. The uniform operator has the property that the elements of an individual are position independent [12].

Evaluate: In the next step, the fitness of all individuals generated with mutation and crossover is evaluated. Therefore, the accuracy of the prediction is calculated using a given classification algorithm. In this paper, we use the k-Nearest-Neighbor and the Random Forest classifiers to evaluate the fitness of an individual by computing the accuracy of the correct predicted emotional state. The higher the fitness of an individual is, the more likely it is selected for the next generation.

Select: Finally, a selection scheme is adopted to map all individuals according to their fitness and draw *p* individuals at random according to their probability for the next generation, where *p* is again the population size parameter. In this paper, we use the Roulette Wheel selection scheme in which the number of times an individual is expected to be selected for the next generation is equal to its fitness divided by the average fitness in the population [11].

This process is repeated as long as the stopping criterion is not yet reached. The stopping criterion is set after a maximum of 50 generations or after 2 generations without improval. The described parameters are illustrated in Figure 1. These can be adjusted independent on the used classification algorithm. A detailed description of the different parameters as well as other available options can be found in the documentation section of RapidMiner [10].

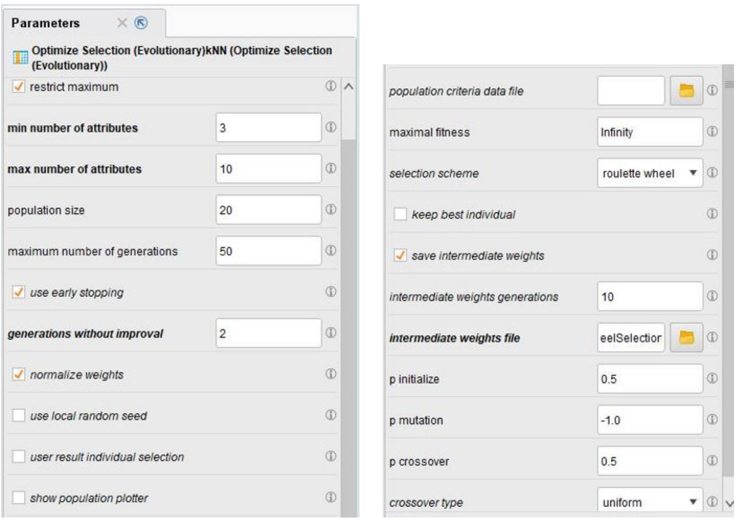

Figure 1. Parameters related to the feature selection method based on evolutionary algorithms. They can be adjusted independent on the used classification algorithm.

3. Results and Discussion

The feature selection method based on evolutionary algorithms was first designed in RapidMiner as described in the previous section. Figure 2 illustrates the implementation of this method using the "Optimize Selection (Evolutionary)" operator. It is integrated within the feature selection subprocess of our previously developed processing workflow for affective computing and stress recognition [2].

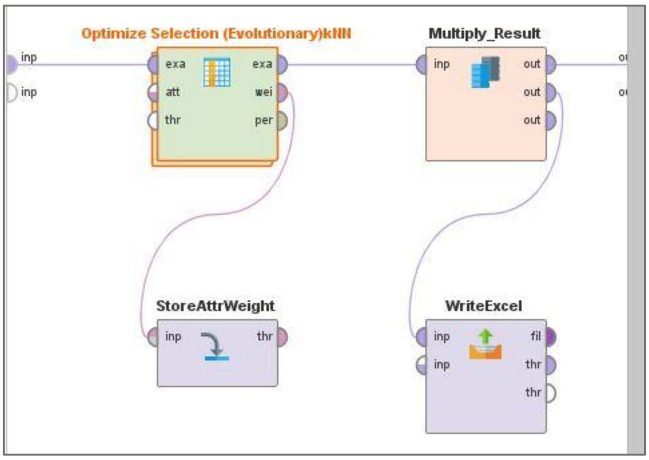

Figure 2. Implementation of the "Optimize Selection (Evolutionary)" operator, integrated within the forward selection subprocess of the affective computing workflow.

The proposed method was then evaluated using biosignal data from our uulmMAC database for affective computing and machine learning applications [9]. For the evaluation, we applied our processing workflow using both the evolutionary algorithms and the Forward Selection method. The latter was chosen for comparison being the fastest among the other two approaches of Backward Elimination and Brute Force. As classifiers, we applied the k-Nearest-Neighbor and the Random Forest algorithms to compute the accuracy of the prediction. As for the validation, we used the cross validation method.

A total of 162 different features were extracted from the biosignal data, including category-based features for the respiration, skin conductance level, temperature and electromyography channels and signal-specific features for the electrocardiogram channel. Among the six different sequences available in the dataset (Interest, Overload, Normal, Easy, Underload, Frustration), we evaluated a two-class problem by computing the recognition rates for the states Overload and Underload.

Our results show that the proposed feature selection approach based on evolutionary algorithms has a much faster runtime compared to the Forward Selection method and does not stop at a local optimum, allowing a promising feature selection alternative in the field of affective computing. Preliminary classifications resulted in a fivefold enhancement of the computation time when using the Random Forest classifier at the same recognition rate of 87%, while a threefold enhancement of computation time was obtained using the k-Nearest-Neighbor classifier, however with a decrease of performance from 69% to 83%.

We are currently investigating some other options to further optimize the results. For instance, by increasing the stopping criteria to more generations and evaluating the effect on the classification rates relative to the increase of computation time. We will conduct further computations using multi-class problems to classify the other affective states from the uulmMAC dataset and extend our preliminary classifications [13]. The results will be evaluated with different classifiers and validation methods as previously adapted by our performance study [14]. Further, in the present work we use the Roulette Wheel selection scheme to select the fittest individuals for the next generation. In the Roulette Wheel selection, the survival probability of each individual is proportional to its relative fitness.

Further selection schemes could be investigated such as the Tournament Selection, in which a randomly selected number of individuals is first selected to take part in a tournament, and the individuals with the highest fitness of this tournament are subsequently selected into the next generation until a predefined number of individuals is reached in the new generation [15].

4. Conclusions

This paper presents a feature selection method based on evolutionary algorithms to optimize the computational efficiency for machine learning applications. It is implemented within our workflow for affective computing and stress recognition using psychophysiological data. We first discussed the importance of feature selection for the recognition process, then we introduced our method based on genetic algorithms and described the implementation as well as the results and next steps. The present work is a valuable effort to enhance the classifications especially in terms of computation time, which is essential for real-life machine learning applications in the field of emotion and stress recognition.

Author Contributions: Conceptualization: D.H.-R.; methodology: D.H.-R., R.A. and L.Z.; validation: D.H.-R., R.A. and L.Z.; formal analysis: D.H.-R., R.A. and L.Z.; investigation: D.H.-R.; resources: D.H.-R. and H.C.T.; data curation: D.H.-R.; writing—original draft preparation: D.H.-R. and R.A.; writing—review and editing: D.H.-R. and H.C.T; visualization: D.H.-R. and R.A.; supervision: D.H.-R. and H.C.T.; project administration: D.H.-R.; funding acquisition: D.H.-R. and H.C.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a Margarete von Wrangell (MvW) habilitation scholarship funded by the Ministry of Science, Research and Arts (MWK) of the state of Baden-Württemberg for Dilana Hazer-Rau.

Institutional Review Board Statement:

Informed Consent Statement:

Data Availability Statement:

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
2. Hazer-Rau, D.; Zhang, L.; Traue, H.C. A Workflow for Affective Computing and Stress Recognition from Biosignals. *Eng. Proc.* **2020,** *2*, 85, doi:10.3390/ecsa-7-08227.
3. De Jong, K. *Evolutionary Computation—A Unified Approach*; MIT Press: Cambridge, UK, 2008; Chapter 1.
4. Rechenberg, I. *Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution*; Frommann-Holzboog Verlag: Stuttgart, Germany, 1973.
5. Holland, J.H. *Adaptation in Natural and Artificial Systems*, 2nd ed.; University of Michigan Press: Ann Arbor, MI, USA, 1975; MIT Press: Cambridge, UK, 1992.
6. Mamontov, D.; Polonskaia, I.; Skorokhod, A.; Semenkin, E.; Kessler, V.; Schwenker, F. Evolutionary Algorithms for the Design of Neural Network Classifiers for the Classification of Pain Intensity. In *Multimodal Pattern Recognition of Social Signals in Human-Computer-Interaction (MPRSS 2018), Lecture Notes in Computer Science (LNCS)*; Schwenker, F., Scherer, S., Eds.; Springer: Cham, Switzerland, 2019; Volume 11377, pp. 84–100, doi:10.1007/978-3-030-20984-1_8.
7. Kestler, H.A.; Haschka, M.; Muller, A.; Schwenker, F.; Palm, G.; Hoher, M. Evolutionary optimization of a wavelet classifier for the categorization of beat-to-beat variability signals. In Proceedings of the Computers in Cardiology 2000, Vol.27 (Cat. 00CH37163), Cambridge, MA, USA, 24–27 September 2000; pp. 715–718, doi:10.1109/CIC.2000.898624.
8. Arunadevi, J.; Nithya, M.J. Comparison of Feature Selection Strategies for Classification using Rapid Miner. *Int. J. Innov. Res. Comput. Commun. Eng.* **2016**, *4*, 13556–13563, doi:10.15680/IJIRCCE.2016.0407088.
9. Hazer-Rau, D.; Meudt, S.; Daucher, A.; Spohrs, J.; Hoffmann, H.; Schwenker, F.; Traue, H.C. The uulmMAC Database—A Multimodal Affective Corpus for Affective Computing in Human-Computer Interaction. *Sensors* **2020**, *20*, 2308, doi:10.3390/s20082308.
10. RapidMiner Documentation—Operator Reference Guide. Available online: https://docs.rapidminer.com/latest/studio/operators/modeling/optimization/feature_selection/optimize_selection_evolutionary.html (accessed on 27 September 2021).
11. Mitchell, M. *An Introduction to Genetic Algorithms*; MIT Press: Cambridge, MA, USA, 1996; Chapter 1.6, pp. 8–12.

12. Syswerda, G. Simulated Crossover in Genetic Algorithms. In Proceedings of the Second Workshop of Foundations of Genetic Algorithms, Vail, CO, USA, 26–29 July 1992; Whitley, D., Ed.; Morgan Kaufmann Publishers: San Mateo, CA, USA, 1993; pp. 239–255, doi:10.1016/B978-0-08-094832-4.50021-0.
13. Daucher, A.; Gruss, S.; Jerg-Bretzke, L.; Walter, S.; Hazer-Rau, D. Preliminary classification of cognitive load states in a human machine interaction scenario. In Proceedings of the International Conference on Companion Technology (ICCT'17), Ulm, Germany, 11–13 September 2017; pp. 1–5.
14. Hazer-Rau, D.; Zhang, L.; Traue, H.C. Performance Evaluation of Various Emotion Classification Approaches from Physiological Signals. *Int. J. Artif. Intell. Appl. (IJAIA)* **2018**, *9*, 31–41.
15. Zhong, J.; Hu, X.; Gu, M.; Zhang, J. Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms. In Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation, Vienna, Austria, 28–30 November 2005, doi:10.1109/CIMCA.2005.1631619.