# Using Grip Strength as a Cardiovascular Risk Indicator Based on Hybrid Algorithms

E. F. Bareño-Castellanos[1]*, P. A. Gaona-García[1], J. E. Ortiz-Guzmán[2], C. E. Montenegro-Marin[1]

[1] Faculty of Engineering, Universidad Distrital Francisco José de Caldas, Bogotá (Colombia)
[2] Facultad de Ciencias de la Salud, Universidad de Ciencias Aplicadas y Ambientales, Bogotá (Colombia)

**unir**
LA UNIVERSIDAD
EN INTERNET

## Abstract

This article shows the application and design of a hybrid algorithm capable of classifying people into risk groups using data such as prehensile strength, body mass index and percentage of fat. The implementation was done on Python and proposes a tool to help make medical decisions regarding the cardiovascular health of patients. The data were taken in a systematic way, k-means and c-means algorithms were used for the classification of the data, for the prediction of new data two vectorial support machines were used, one for the k-means and the other for the c-means, obtaining as a result a 100% of precision in the vectorial support machine with c-means and a 92% in the one of k-means.

## I. Introduction

CARDIOVASCULAR diseases are the leading cause of mortality in the world, with ischemic heart disease and cerebrovascular disease being the most frequent [1]. There are factors such as obesity, sedentary lifestyle, hypertriglyceridemia, smoking and family history that are directly related to heart problems and are used by specialists for the early detection of heart disease [2]. For this reason, the importance of finding mechanisms for the assessment of risk factors associated with cardiovascular diseases is growing.

Currently studies are presented where grip strength has received considerable attention not only for the relationships of protein level, muscle mass and strength, but as an indicator of physical capacity and general health [3]. Some studies have suggested that grip strength is a risk factor in multiple diseases such as diabetes, cancer and different heart diseases [4], therefore it is considered that a strong or weak grip of the hand carries more than social cues, as it can also help measure an individual's risk of having a heart attack or stroke, or dying from cardiovascular disease [5].

One of the areas that could contribute in the assessment of these indicators is through processes based on data mining, given its relevance to obtain knowledge of large amounts of data. The data mining-based process enables the integration of multiple disciplines such as statistics, machine learning, neural networks, and pattern recognition. Authors such as [6]-[7] have used techniques such as the nearest neighbor (KNN), decision trees (DT), genetic algorithms (GA) and naive Bayesian classifier (NB) to early detection of heart problems.

Based on the above, a model based on Support Vector Machine (SVM) is presented for classification in risk groups using clustering techniques that allow establishing relationships between grip strength and different physiological variables such as height, weight, body mass index and fat percentage. One of the purposes of this study is to evaluate the behavior and reliability of algorithms based on SVM, so that later a tool can be proposed that can provide support to an expert to determine health status in a quick and complementary way.

For the grouping algorithms two models were used, one of them is a c-means hybrid model and the second one is k-means, to feed these models a dataset was used with data such as age, height, weight, fat percentage and caporal mass index. In addition, the prehensile force was used as a predictive variable in the vectorial support machine.

The following article is structured as follows: section II, related works presents all the background of the research, as well as related works. Section III establishes the work methodology to be followed. Section IV proposes the data mining model for the case study, section V shows the results obtained. Finally, section VI is the discussion of results and conclusions are presented in section VII.

## II. Related Works

Biomarkers are medical indicators that demonstrate bodily functions or pathological processes in addition to providing an objective indication of medical status [8]. The use of grip strength as a biomarker has taken on great importance. In the latest research, it has been determined that low levels of grip strength are associated with various associated morbidities; an increased risk of falls, hospital stay and mortality; and a lower quality of life [9]. For example, the Prospective Urban-Rural Epidemiology (PURE) study determined that the decrease in grip strength is a risk factor for the incidence of

* Corresponding author.

E-mail address: efbarenoc@correo.udistrital.edu.co

cardiovascular diseases and can predict the risk of death in people who develop cardiovascular or non-cardiovascular diseases [10].

Within the context of analysis, data mining is the process of extracting information in order to obtain knowledge from large amounts of data. It is an integration of multiple disciplines such as statistics, machine learning, and neural networks [11]. Data mining allows the exploration of large data sets in search of relationships, knowledge and patterns that are difficult to determine with traditional methods.

In recent years, different machine learning techniques have been used to predict heart disease, using a database with variables such as blood pressure, cholesterol, blood sugar, weight, age and sex. Some of these techniques will be discussed below. Authors such as Rairikar et al. used three data mining techniques to predict heart diseases, which were decision tree, random forests and Naive Bayes, determined that the random forest is the algorithm with the highest precision in addition to the use of a genetic algorithm to select the most important characteristics of the data sets [6]. In a study carried out by Bahrami & Hosseini Shirvani, they verified the different classification techniques in the diagnosis of heart diseases, using techniques such as the decision tree and KNN, the Naive Bayes algorithm was also used for the grouping of relevant characteristics. After ranking and performance evaluation, the decision tree is considered the best algorithm for heart disease diagnosis with the data set that was chosen [7].

For their part Jabbar et al. proposed a new hybrid algorithm that combines the KNN algorithm with a genetic algorithm, they used the genetic algorithm to classify the most relevant attributes of the dataset and then used the KNN algorithm, the experimental results showed that the use of a genetic algorithm improved accuracy in diagnosing heart disease [12]. Authors such as Sowmiya used different machine learning techniques in multiple datasets to determine which are the best techniques in the use of prediction of heart disease, diseases such as fibrillation, congenital heart disease, coronary artery disease, heart attack were studied [13]. Along the same lines, Gawande & Barhatte used a 7-layer convolutional neural network for the classification of heart diseases through the analysis of ECG signals (electrocardiogram), achieving a precision of 99.46% in different patients [14]. For their part, Amin et al. carried out a study where they would evaluate multiple algorithms in the prediction of heart problems, the data set used contained the variables: age, sex, Cp, pressure in each person, amount of sugar in the blood, ECG results, maximum heart rate, and heart status. It was determined that the algorithm with the best accuracy was a Support Vector Machine with a result of 86.87% with the use of 9 attributes [15].

The authors López-Martínez et al used a neural network model to predict neonatal sepsis, using a data set of 555 patients divided between 66% positive and 34% negative cases, resulting in an accuracy of 83.1% [16]. In the same way the authors López-Martínez et al. used a neuronal network to estimate the association of variables such as BMI, race, age, among others with hypertension reaching a specificity of 87% with a precision of 57.8% [17].

A work that can be highlighted is the one of Devi et al., which uses the diffuse c-means model for the grouping of skin lesions using non-dermoscopic images resulting in an accuracy of 95.69% and a sensitivity of 90.02% [18]. The problem of this work is that it is only for non-dermoscopic images. The next work by López-Martínez, Núñez-Valdez, García-Díaz, et al., uses big data and artificial intelligence to improve the support to medical decisions in health management of the population, this work resulted in the detection of patterns using different types of data [19].

Taking into account the aforementioned works, the following study aims to use grip strength as a risk indicator for heart problems, data mining algorithms (machine learning) will be used to determine relationships with other biomarkers such as weight, height, sex,

percentage of fat and BMI among others. The objective of this work is to create a tool that can support specialists in decision-making in the medical field and health in general.

## III. Methodology

In this study, machine learning methods were used to create a model based on Support Vector Machine and clustering, this was trained with a database obtained systematically at the university of Rosario. Fig. 1 shows the process of obtaining, cleaning, analysis and results of this investigation.
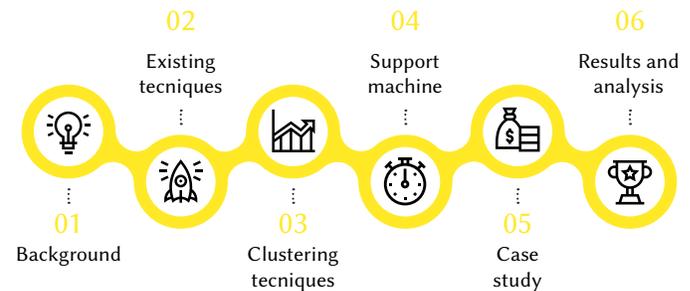


Fig. 1. Methodology and steps for its application.

### A. Sample Population

The sample included in this research corresponded to students from the School of Medicine and Health Sciences of the Universidad del Rosario (Bogotá, Colombia), who were evaluated in the research laboratory of the Physiology Unit. In total, 80 students (50 women and 30 men) were included in the study. The characteristics of the Sample are reported in Table I.

TABLE I. Description of the Sample Composed of 29 Men and 50 Women, the Table Contains the Average (A), the Deviation (D), the Minimum (MI) and the Maximum (MA). The Age Is in Years, the Height Is in Centimeters, the Weight Is in Kilograms, the Body Mass index (BMI) Is in Kilograms Per Square Meter and the Maximum Grip force Is in Newtons

|  | Men n=29 | | | | Women n=50 | | | |
|---|---|---|---|---|---|---|---|---|
|  | A | D | MI | MA | A | D | MI | MA |
| Age | 18,1 | 1,29 | 17 | 22 | 17,83 | 1,04 | 16 | 20 |
| Height | 175,41 | 5,56 | 163 | 186 | 161,14 | 6,83 | 145 | 177 |
| Weight | 70,19 | 11,17 | 53,3 | 95,1 | 58,29 | 8,28 | 48,2 | 76,4 |
| % fat | 16,76 | 6,74 | 6 | 30,4 | 31,94 | 4,87 | 20,6 | 41,7 |
| BMI |  | 3,54 | 16,5 | 31,8 | 22,38 | 2,35 | 17,2 | 28,4 |
| grip force |  | 22,16 | 21,35 | 108,86 | 29,1 | 9,36 | 17,12 | 65,76 |

The dataset can be found through github[1].

The sampling that was followed in this study corresponded to a convenience sampling, since the participants were included after verifying that they met the inclusion criteria and that they did not have any exclusion criteria that did not allow them to be included in the study sample. Regarding the above, the inclusion criteria were: being of legal age, accepting participation in this study by signing the informed consent, not having consumed caffeinated beverages in the last 2 hours or having practiced any intense physical activity moderate or high on the last day and complete all the protocols required to record Heart Rate Variability (HRV) and grip strength. Regarding the exclusion criteria, those participants who were consuming any medication with direct action on the nervous system or on cardiovascular function, who had consumed beverages rich in

---

[1] https://github.com/FrederickUdis/data-Grip-strength.git

caffeine or stimulants of the central nervous system in the last 2 hours were not taken into account. Besides, those who had not signed the informed consent and who did not complete all the protocols required for data collection were excluded.

### B. Methodology

HRV Variability was recorded through Polar V800TM reference monitors, which have been widely validated in the scientific literature for HRV recording. The records were taken in a sitting position, for 15 minutes, in a room at stable temperature and relative humidity. During this period, all participants remained completely silent, without control of the ventilatory rate and trying to be as relaxed as possible. The Polar record was digitized using the Polar ProTrainerTM program and HRV analysis was performed in the KubiosTM program with a 5-minute window of the total record, from which variables were extracted.

On the other hand, the grip force was recorded using the PowerLabTM system, using the manual dynamometer from the same developer. In total, each participant made 3 attempts, always with the dominant hand and the arm flexed at 90 ° with the elbow as close as possible to the participant's trunk and without moving them neither forward nor backward during the recording. The procedure consisted of squeezing the dynamometer as hard as possible for 3 seconds and this action was repeated 3 times, with a 5-minute rest period between each series. Of the 3 attempts, the best was chosen for the final analysis.

The second part of the first phase consisted of a background study where a rigorous investigation was carried out on the grip force as an indicator of heart problems, as well as the different machine learning techniques used in heart problems. In the second phase, the different techniques sought in phase one were evaluated, determining that the most appropriate algorithms for our problems are two types of algorithms, the first group is determined supervised algorithms and the other is unsupervised.

In the third phase, an unsupervised algorithm was used because [20] it does not depend on specific instructions when performing a classification, instead it is based on the autonomous grouping of data through exploration, this type of learning is very similar to the way humans learn, which offers a very flexible approach when applying it to our problem. It should be noted that in addition to an unsupervised algorithm, a hybrid algorithm was used that uses fuzzy logic rules to increase precision.

In phase four the model based on Support Vector Machine was used to classify the results, this was applied through a case study in phase five.

The result and analysis are carried out in phase six, in addition to obtaining the performance metrics of the algorithm.

## IV. Case Study

The proposed model has two main components that can be observed in Fig. 2, the first is the data processing and the second is the data analysis, in the data processing during the data acquisition we obtain the data in raw and we go to the preparation of the data where we do a cleaning to convert it into a dataset for analysis, the next component has three parts (analysis, results and conclusions).

The main part of the second component is the analysis, which is divided into 4 subcomponents that are the data preprocessing, the clusters, the classification model and finally the output, in section 4.1 to 4.4 we will talk more about these components.

Based on the above, the model is a 2-layer model, the first layer being in charge of acquiring and preparing the data for the second layer, this is done with the Python programming language and the specialized libraries for data treatment.
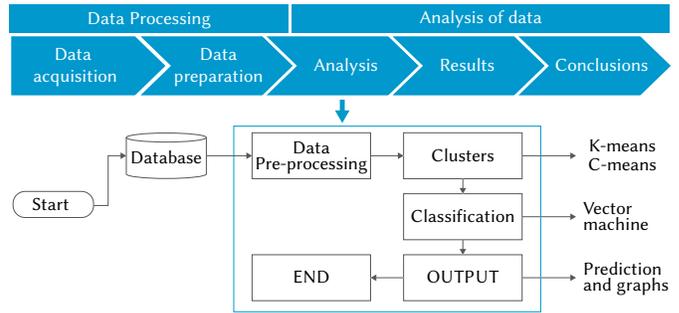


Fig. 2. Model proposed for data processing and data analysis.

The second layer is the constructed algorithm that has two important parts that are the grouping where the two chosen algorithms are used to later pass to the Support Vector Machine which has the function of classifying the result of the grouping.

### A. Data Cleaning

After data collection. All records with missing values were removed from the data set, reducing the number of data from 80 to 78. Then the variable sex (male or female) became a category variable or more commonly called binary (0 or 1). The task of data preprocessing was carried out by cleaning null data, the data was divided between men and women. Subsequently, the characteristics of the names and surnames were eliminated to give a total of 19 characteristics.

### B. Data Processing

After processing the data, two clustering algorithms were used to determine the risk groups, [21] describes the k-means clustering algorithm as a popular algorithm due to its effectiveness in dividing multiple points into k clusters, likewise it has a high adaptability to different problems. For our study, this was the first of the two clustering algorithms that were used, the purpose of this clustering was to find groups of related data establishing relationships between characteristics such as grip strength (grip strength) and body mass index (BMI), fat percentage (% Fat) among others. It works by adding a new column to the dataset which contains the labels of the cluster to which each data belongs to finally be classified by a Support Vector Machine (SVM).

The second algorithm used was the c-means (FMC) which assigns a probability of belonging to each point on a group or cluster to finally choose the best group that belongs to the said point [22].

The implementation of the k-means algorithm was done with the Kmeans library of sklearn, three groups or clusters which are determined by 3 stars were obtained as a result. This can be seen in Fig. 3.
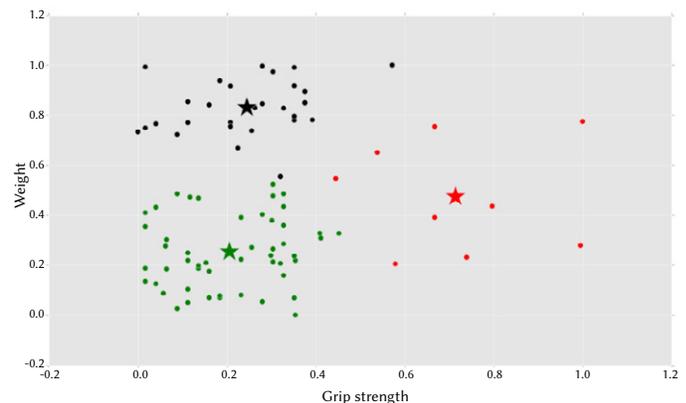


Fig. 3. Result obtained from the k-means algorithm using the python language for analysis and visualization.

The Liberia scikit-fuzzy, which contains the skfuzzy package, was used for the application of the c-means algorithm. An evaluation of up to a maximum of 10 clusters determined that the optimal number of clusters is 2, this can be seen in Fig. 4.
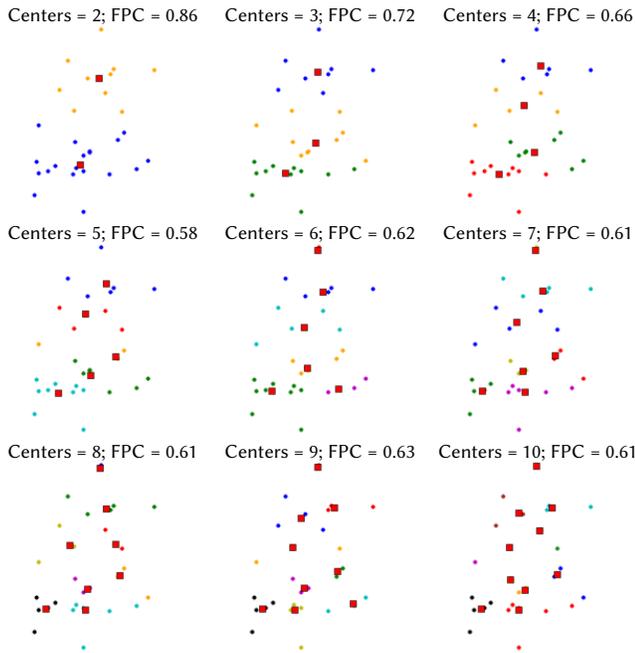


Fig. 4. Result of the c-means algorithm from 2 to 10 clusters.

### C. Classification

Pitale et al. [23] establishes two steps for the application of the classification algorithms: the first phase consists of the definition of the model, in the second phase the method is selected and finally a method is applied to classify it. For this study, a Support Vector Machine (SVM) was used, which aims to classify new data in three risk regions (low-medium-high) according to grip strength and fat percentage.

The vectorial support machine was made in Python with the sklearn. SVM library which offers everything necessary to apply the vectorial support machine as shown in Fig. 5.

```
from sklearn import svm
model = svm.SVC()
model.fit(X_train, y_train)

SVC(C=1.0, break_ties=False, cache_size=200, class_weight=None, coef0=0.0,
    decision_function_shape='ovr', degree=3, gamma='scale', kernel='rbf',
    max_iter=-1, probability=False, random_state=None, shrinking=True,
    tol=0.001, verbose=False)
```

Fig. 5. Code for the use of the sklearn library in addition to the application of the Support Vector Machine (SVM).

## V. Results and Analysis

Of the 79 participants in the analysis, 29 were men and 50 were women. 48.27% of the men had percentage fat greater than 20% and a BMI greater than 20 kg / m2 in addition to grip strength of 40 and 70 Newtons.

As for the women, 78% had a high percentage of fat above 30% but only 14% had a BMI above 25 kg/m 2 in addition to a grip strength of 24 to 40 as shown in Tables II, III, IV. The description in Table II provides more data on the results obtained. There was no significant difference in terms of age, height and weight.

TABLE II. Body Mass index (BMI) Measurements, the Following Parameters Were Used to Obtain the Measurements Corresponding to Table I: Insufficient Weight <18,5 Kg/m2, Normal Weight 18,5 - 24,9 Kg/m2, Overweight Degree I 25,0-26,9 Kg/m2, Overweight Degree II 27,0-29,9 Kg/m2, Type I obesity 30,0-34,9 Kg/m2, Type II Obesity 35,0-39,9 Kg/m2, Type III Obesity 40,0-49,9 Kg/m2 and Type IV Obesity 40,0-49,9 Kg/m2, these Parameters Were Obtained from [25]

| | Men n = 29 | | Women n = 50 | |
|---|---|---|---|---|
| | F | % | F | % |
| Insufficient | 1 | 3,45 | 4 | 8 |
| Normal | 21 | 72,41 | 39 | 78 |
| Overweight | 5 | 17,24 | 4 | 8 |
| Overweight | 0 | 0 | 2 | 4 |
| Type I obesity | 2 | 6,9 | 1 | 2 |
| Type II obesity | 0 | 0 | 0 | 0 |
| Type III obesity | 0 | 0 | 0 | 0 |
| Type IV obesity | 0 | 0 | 0 | 0 |

TABLE III. Guidelines for Determining which Subjects Had a Large Percentage of Body Fat

| | Fat% limit values | |
|---|---|---|
| characteristics | MEN | WOMEN |
| thin | <8% | <15% |
| Optimum | 8,1 a 15,9% | 15,1 a 20,9% |
| Slightly overweight | 16,0 a 20,9% | 21,0 a 25,9% |
| Overweight | 21,0 a 24,9% | 26,0 a 31,9% |
| Obesity | ≥25% | ≥32% |

TABLE IV. Measurements of Fat Percentage, these Values Were Obtained Through the Grouping Analysis Carried Out Before the Prediction, the Results Obtained Determined that Women Possess a higher Percentage of Fat than Men, as Determined in [25].

| Men n = 29 | | Women n = 50 | |
|---|---|---|---|
| F | % | F | % |
| 4 | 13,79 | 0 | 0 |
| 11 | 37,93 | 4 | 8 |
| 5 | 17,24 | 4 | 8 |
| 6 | 20,69 | 19 | 38 |
| 3 | 10,34 | 23 | 46 |

According to Table IV, it was possible to establish a group of women who have a percentage Fat greater than 32% with a grip force that ranges between 20 and 40 Newtons, this can be seen in Fig. 6, which implies that women with obesity have a grip strength of less than 30, which is the average seen in Table I.
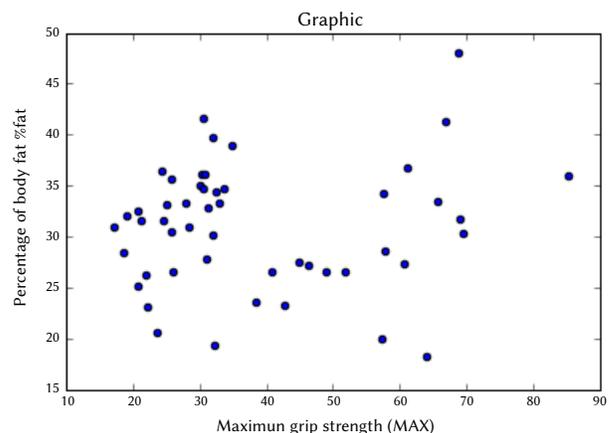


Fig. 6. Dispersion graph of the data referring to the women of the percentage of fat (% fat) Vs maximum grip strength (MAX), performed with Python.

Based on table IV, 90% of this group also has a BMI greater than 22 kg / m2, which indicates a clear relationship between overweight and grip strength in women with these characteristics; this can be clearly seen in Fig. 7.
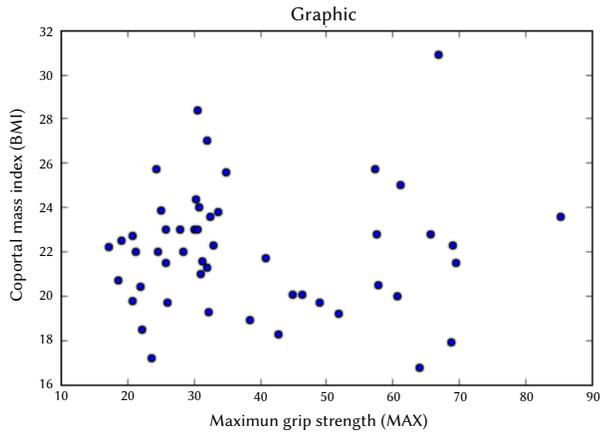


Fig. 7. Dispersion graph of the data referring to the body mass index (BMI) Vs maximum grip strength (MAX), performed with Python.

With respect to men, the group cannot be clearly seen due to the lack of data, but data that can be used as a group or clusters can be seen, this can be seen in Fig. 8 and Fig. 9.
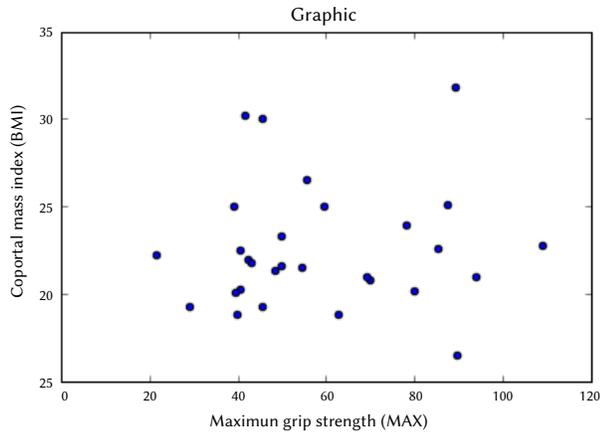


Fig. 8. Dispersion graph of the data referring to the men's body mass index (BMI) Vs maximum grip strength (MAX), performed with Python.
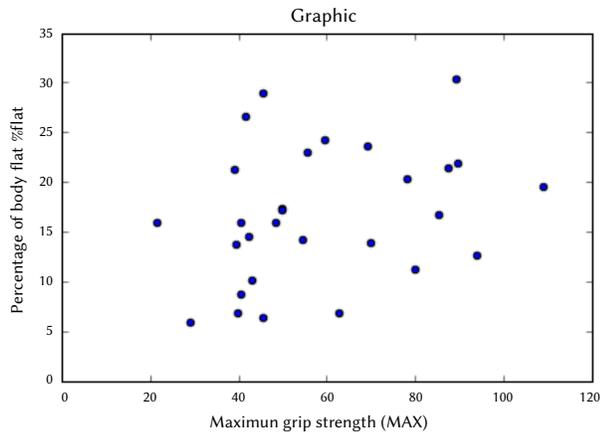


Fig. 9. Dispersion graph of the data referring to men of the percentage of fat (% Fat) (BMI) Vs maximum grip strength (MAX), performed with Python.

According to the results of the use of the k-means technique, 3 risk groups were revealed (low, medium, high). A Support Vector Machine was used to classify these groups and predict new data in each of these groups. Fig. 10 shows the classification of the Support Vector Machine.
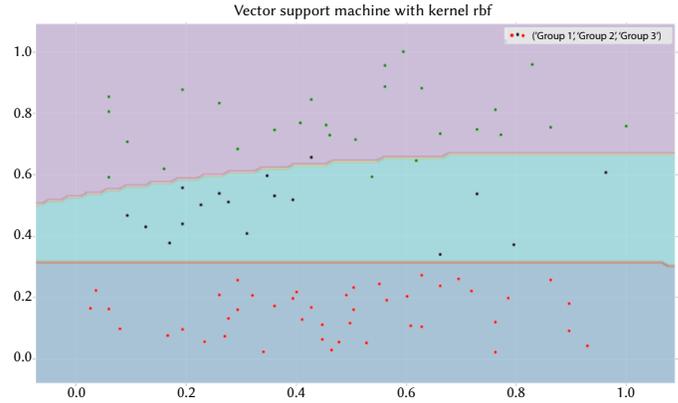


Fig. 10. Result of Support Vector Machine training for data classification using python and sklearn.

In other part, the results of the hybrid C-means technique, determined the existence of two risk groups, with the help of the Support Vector Machine it was possible to classify new data into these two groups. Fig. 11 shows the results of the Support Vector Machine applied to the results of the c-means algorithm.
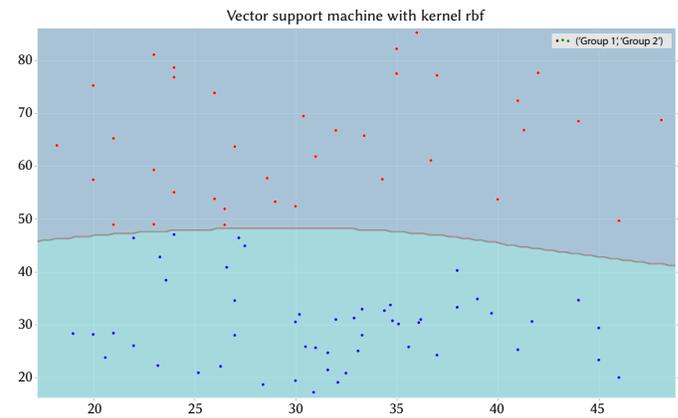


Fig. 11. Result of Support Vector Machine training for data classification using python and sklearn.

To identify the best combination of algorithms, the confusion matrix incorporated in the Python sklearn library was used. This matrix evaluates multiple metrics to determine the best algorithm. The evaluation parameters used are:

- Accuracy
- Recall
- F1-score
- Support

These parameters are obtained through the confusion matrix, which is an evaluation tool used in the machine learning area. The columns of a Confusion Matrix represent the results of the prediction class, and the rows represent the results of the class [24]. Fig. 12 shows the results of the Support Vector Machine with the k-means algorithm.

|   | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.92 | 1.00 | 0.96 | 11 |
| 1 | 1.00 | 1.00 | 1.00 | 5 |
| 2 | 1.00 | 0.83 | 0.91 | 6 |
| accuracy |  |  | 0.95 | 22 |
| macro avg | 0.97 | 0.94 | 0.96 | 22 |
| weighted avg | 0.96 | 0.95 | 0.95 | 22 |

Fig. 12. Metrics of the k-means algorithm and Support Vector Machine obtained through the sklearn python library.

|   | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 1.00 | 1.00 | 1.00 | 9 |
| 1 | 1.00 | 1.00 | 1.00 | 13 |
| accuracy |  | 1.00 | 1.00 |  |
| macro avg | 1.00 | 1.00 | 1.00 |  |
| weighted avg | 1.00 | 1.00 | 1.00 |  |

Fig. 13. Metrics of the c-means algorithm and Support Vector Machine obtained through the sklearn python library.

According to the results obtained, they can be considered acceptable, since the two algorithms have an accuracy higher than 90%, being the best the c-means algorithm with the support vector machine, which presents an accuracy of 100%, the high accuracy rate of the two algorithms is due to the low amounts of data obtained, which are not more than 80. By reducing the number of variables it was possible to obtain great precision with the two algorithms and to determine the influence of the percentage of fat in the prehensile force in women, this was achieved thanks to the use of unsupervised algorithms (c-means and k-means), the use of vectorial support machines gave good results and can be used as a tool to help heart specialists.

## VI. Discussion

In this study it was possible to demonstrate that male volunteers who had a low percentage body fat (<23%) and a lower BMI had a grip strength level of less than 40 Newtons, compared to people who have a body mass index and level of average fat, aspect that corroborates the work carried out by Alberto Cardozo [25]. However, in women it was presented that body weight and BMI were higher than in men, coinciding with what was found by other authors [26], although the latter presented a higher percentage of fat (> 30%); These elements could be induced by various factors such as hormones, eating habits, body composition, etc. Regarding the grip strength, it was found that several women who had percentage of fat greater than 28% had a grip strength less than 40 Newtons, but the BMI was normal.

During the creation and evaluation of the algorithm, multiple problems and drawbacks were found, such as the lack of important characteristics associated with smoking and comorbidities in the dataset. Similarly, at the time of making the correlation matrix, relationship values of less than 30% were found, for this reason it was decided to perform unsupervised algorithms to validate relationships that are not seen in the correlation matrix. In addition, the model presented a problem of a perfect fit to the training set, this situation was caused by the fact that the amount of training data was less compared to the greater number of characteristics, a very common problem presented by supervised classification algorithms [27].

Finally, to highlight, the average value of precision of the applied algorithms were higher in precision than the results obtained by

(Amin et al., 2019), since when they used a Support Vector Machine fed with 9 variables, they obtained a precision of 86% while we obtained 100% accuracy due to the use of a clustering algorithm before using the Support Vector Machine. On the contrary, the work carried out by Gawande & Barhatte used a 7-layer neural network using electrocardiograms to find heart problems and obtained an accuracy of 98.7 [14].

## VII. Conclusion

With the use of clustering techniques (k-means and c-means), groups with low grip strength relationships and high fat percentage were found, as well as the body mass index, which indicates a relationship. Likewise, the use of a Support Vector Machine allowed the classification of new data into risk groups.

One of the main advantages of using easily obtained variables with grouping and prediction algorithms was the rapid classification of new data, which can be used as a tool for medical decision-making in patients with cardiovascular risk. These algorithms can be applied to any type of patient consulted by a specialist, as long as more variables that imply cardiovascular risk are taken into account.

It is recommended for use in conjunction with tools capable of detecting heart problems and under the supervision of qualified specialists.

The proposed algorithms are suggested as a useful and complementary tool for people who want to know their risk group, since it combines variables that can be easily obtained such as grip strength, weight and body mass index.

As a future study, it is expected to complement the experiment with additional variables associated with smoking, sedentary lifestyle and family history of cardiac comorbidities associated with grip strength. Likewise, other classification techniques such as neural networks, decision trees or genetic algorithms, will be evaluated.

## References

[1] H. Wang et al., "Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015," The Lancet, vol. 388, no. 10053, pp. 1459–1544, 2016, doi: 10.1016/S0140-6736(16)31012-1.

[2] D. Mercy et al., "Riesgo cardiovascular global y edad vascular: herramientas claves en la prevención de enfermedades cardiovasculares Global cardiovascular risk and vascular age: key tools in cardiovascular diseases prevention," Revista Médica Electrónica, Artic. Revisión, pp. 211–226, 2014.

[3] O. Prasitsiriphon and W. Pothisiri, "Associations of Grip Strength and Change in Grip Strength With All-Cause and Cardiovascular Mortality in a European Older Population," Clinical Medicine Insights: Cardiology, vol. 12, 2018, doi: 10.1177/1179546818771894.

[4] C. A. Celis-Morales et al., "Associations of grip strength with cardiovascular, respiratory, and cancer outcomes and all cause mortality: Prospective cohort study of half a million UK Biobank participants," BMJ (Online), vol. 361, pp. 1–10, 2018, doi: 10.1136/bmj.k1651.

[5] M. D. Howard LeWine, "Grip strength may provide clues to heart health - Harvard Health Blog - Harvard Health Publishing," Grip strength may provide clues to heart health, 19-May-2015. [Online]. Available: https://www.health.harvard.edu/blog/grip-strength-may-provide-clues-to-heart-health-201505198022. [Accessed: 03-Jul-2020].

[6] A. Rairikar, V. Kulkarni, V. Sabale, H. Kale, and A. Lamgunde, "Heart disease prediction using data mining techniques," Proceedings of 2017 International Conference on Intelligent Computing and Control, I2C2 2017, vol. 2018-Janua, no. October, pp. 1–8, 2018, doi: 10.1109/I2C2.2017.8321771.

[7] B. Bahrami and M. Hosseini Shirvani, "Prediction and Diagnosis of

Heart Disease by Data Mining Techniques," Journal of Multidisciplinary Engineering Science and Technology, vol. 2, no. 2, pp. 3159–40, 2015.

[8] R. W. Bohannon, "Grip strength: An indispensable biomarker for older adults," Clinical Interventions in Aging, vol. 14, pp. 1681–1691, 2019, doi: 10.2147/CIA.S194543.

[9] sevtap güllüoğlu badıl, "Does Vitamin D Level Affect Grip Strength: a Cross-sectional Descriptive Study," Erciyes Medical Journal, vol. 42, no. 1, pp. 7–11, 2019, doi: 10.14744/etd.2019.15428.

[10] D. P. Leong et al., "Prognostic value of grip strength: Findings from the Prospective Urban Rural Epidemiology (PURE) study," The Lancet, vol. 386, no. 9990, pp. 266–273, 2015, doi: 10.1016/S0140-6736(14)62000-6.

[11] B. Deekshatulu and P. Chandra, "Classification of heart disease using artificial neural network and feature subset selection," Global Journal of Computer Science and Technology, vol. 13, no. 3, 2013.

[12] M. A. jabbar, B. L. Deekshatulu, and P. Chandra, "Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm," Procedia Technology, vol. 10, pp. 85–94, 2013, doi: 10.1016/j.protcy.2013.12.340.

[13] C. Sowmiya, "Comparative Study of Predicting Heart Disease By Means Of Data Mining," Int. J. Eng. Comput. Sci., vol. 5, no. 12, pp. 19580–19582, 2016.

[14] N. Gawande and A. Barhatte, "Heart diseases classification using convolutional neural network," Proceedings of the 2nd International Conference on Communication and Electronics Systems, ICCES 2017, vol. 2018-Janua, no. Icces, pp. 17–20, 2018, doi: 10.1109/CESYS.2017.8321264.

[15] M. S. Amin, Y. K. Chiam, and K. D. Varathan, "Identification of significant features and data mining techniques in predicting heart disease," Telematics and Informatics, vol. 36, pp. 82–93, 2019, doi: 10.1016/j.tele.2018.11.007.

[16] F. López-Martínez, E. R. Núñez-Valdez, J. Lorduy Gomez, and V. García-Díaz, "A neural network approach to predict early neonatal sepsis," Computers and Electrical Engineering, vol. 76, pp. 379–388, 2019, doi: 10.1016/j.compeleceng.2019.04.015.

[17] F. López-Martínez, E. R. Núñez-Valdez, R. G. Crespo, and V. García-Díaz, "An artificial neural network approach for predicting hypertension using NHANES data," Scientific Reports, vol. 10, no. 1, pp. 1–14, 2020, doi: 10.1038/s41598-020-67640-z.

[18] S. S. Devi, N. H. Singh, and R. H. Laskar, "Fuzzy C-Means Clustering with Histogram based Cluster Selection for Skin Lesion Segmentation using Non-Dermoscopic Images," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 6, no. 1, p. 26, 2020, doi: 10.9781/ijimai.2020.01.001.

[19] F. López-Martínez, E. R. Núñez-Valdez, V. García-Díaz, and Z. Bursac, "A case study for a big data and machine learning platform to improve medical decision support in population health management," Algorithms, vol. 13, no. 4, pp. 1–19, 2020, doi: 10.3390/A13040102.

[20] A. Roohi, K. Faust, U. Djuric, and P. Diamandis, "Unsupervised Machine Learning in Pathology: The Next Frontier," Surgical Pathology Clinics, vol. 13, no. 2, pp. 349–358, 2020, doi: 10.1016/j.path.2020.01.002.

[21] S. K. Majhi and S. Biswal, "Optimal cluster analysis using hybrid K-Means and Ant Lion Optimizer," Karbala International Journal of Modern Science, vol. 4, no. 4, pp. 347–360, 2018, doi: 10.1016/j.kijoms.2018.09.001.

[22] V. Manikandan, V. Porkodi, A. S. Mohammed, and M. Sivaram, "Ensemble Classification Based Microarray Gene Retrieval System," ICTACT Journal on Soft Computing, vol. 9, no. 1, pp. 1806–1812, 2018, doi: 10.21917/ijsc.2018.0252.

[23] R. Pitale, K. Tajane, and J. Umale, "Heart Rate Variability Classification and Feature Extraction Using Support Vector Machine and PCA: An Overview," Journal of Engineering Research and Applications www.ijera.com, vol. 4, no. 1, pp. 381–384, 2014.

[24] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view," Information Sciences, vol. 507, pp. 772–794, 2020, doi: 10.1016/j.ins.2019.06.064.

[25] L. Alberto Cardozo, L. Alberto, C. Guzman, Y. Andrés, M. Torres, and J. Alejandro, "Artículo Original Porcentaje de grasa corporal y prevalencia de sobrepeso-obesidad en estudiantes universitarios de rendimiento deportivo de Bogotá, Colombia Body fat percentage and prevalence of overweight-obesity in college students of sports performanc," Nutrición clínica y dietética hospitalaria, vol. 36, no. 3, pp. 68–75, 2017, doi: 10.12873/363cardozo.

[26] M. E. Piché, P. Poirier, I. Lemieux, and J. P. Després, "Overview of Epidemiology and Contribution of Obesity and Body Fat Distribution to Cardiovascular Disease: An Update," Progress in Cardiovascular Diseases, vol. 61, no. 2, pp. 103–113, 2018, doi: 10.1016/j.pcad.2018.06.004.

[27] C. M. Hernández-Ruiz, S. A. Villagrán Martínez, J. E. Ortiz Guzmán, and P. A. Gaona Garcia, "Model based on support vector machine for the estimation of the heart rate variability," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11140 LNCS, pp. 186–194, 2018, doi: 10.1007/978-3-030-01421-6_19.

### Edvard Frederick Bareño Castellanos

Is a student in his final year of systems engineering in the faculty of engineering at Universidad Distrital Francisco José de Caldas (UDFJC) in Bogotá Colombia, graduated from the district school Antonio Nariño in 2011 in Bogotá Colombia. Since 2016 I am part of the GIIRA research group at Universidad Distrital Francisco José de Caldas (UDFJC). The areas of interest in research are artificial intelligence, artificial vision, neural networks, applied mathematics, programming and data analysis. Currently he works at Universidad Distrital Francisco Jose de Caldas in systems advisory office performing the functions as a junior developer in multiple systems of the university.

### Paulo Alonso Gaona García

He earned a Ph.D. in Information and Knowledge Engineering from University of Alcalá in 2014 and working as a collaborator researcher in the Information Engineering Research Unit at University of Alcalá since 2012. Is full professor at Engineering Faculty of Universidad Distrital Francisco José de Caldas, Bogotá – Colombia since 2008. He is director of Multimedia Research Group and active member of GIIRA research group since 2008. He has a Master in Information Science and communication from Universidad Distrital Francisco José de Caldas (Bogotá - Colombia - 2006). He is Systems Engineer at Universidad Distrital Francisco José de Caldas (2003). His research interest includes web science, semantic web, network and communications, e-learning, information visualization and visual analytics.

### Johan Enrique Ortiz Guzmán

I am Johan Enrique Ortiz Guzmán, professional in Sports Sciences, graduated from the University of Applied and Environmental Sciences U.D.C.A, in Bogotá Colombia, in 2009 and Master in Physiology, graduated from the University of Valencia, Spain, in 2011. I have been a university professor since 2012, working in faculties of medicine and health sciences such as the Universidad del Rosario, the UDCA university and the Fundación Universitaria del Area Andina, teaching classes in the area of physiology and sports control and integrating the area of basic sciences from the physiology unit. In terms of research, the focus of my work has been oriented towards cardiovascular adaptations to exercise, especially from the area of modulation exerted by the autonomic nervous system on cardiovascular function and this measured through the analysis of the heart rate variability.

### Carlos Enrique Montenegro Marin

PhD in Computer Science from Oviedo University (2012). Master in Web Site Management and Engineering at the International University of the Rioja - UNIR (2013). Master in Information Science at the Universidad Distrital Francisco José de Caldas (2006). Systems Engineer at the District University (2003). He was a Dean of Engineering Faculty (December 2018 - 2019) and full professor, attached to Engineering Faculty of District University "Francisco José de Caldas" since 2006. He was Coordinator of committee of accreditation in the bachelor's in systems engineering (2007 to 2010), He was Coordinator of the bachelor's in systems engineering (2012 to 2014).