Journal of *Personalized Medicine*

MDPI

Article

An Entropy Approach to Multiple Sclerosis Identification

Gerardo Alfonso Perez [1,*] **and Javier Caballero Villarraso** [1,2]

[1] Department of Biochemistry and Molecular Biology, University of Cordoba, 14071 Cordoba, Spain; bc2cavij@uco.es
[2] Biochemical Laboratory, Reina Sofia University Hospital, 14004 Cordoba, Spain
[*] Correspondence: ga284@cantab.net

Abstract: Multiple sclerosis (MS) is a relatively common neurodegenerative illness that frequently causes a large level of disability in patients. While its cause is not fully understood, it is likely due to a combination of genetic and environmental factors. Diagnosis of multiple sclerosis through a simple clinical examination might be challenging as the evolution of the illness varies significantly from patient to patient, with some patients experiencing long periods of remission. In this regard, having a quick and inexpensive tool to help identify the illness, such as DNA CpG (cytosine-phosphate-guanine) methylation, might be useful. In this paper, a technique is presented, based on the concept of Shannon Entropy, to select CpGs as inputs for non-linear classification algorithms. It will be shown that this approach generates accurate classifications that are a statistically significant improvement over using all the data available or randomly selecting the same number of CpGs. The analysis controlled for factors such as age, gender and smoking status of the patient. This approach managed to reduce the number of CpGs used while at the same time significantly increasing the accuracy.

Keywords: multiple sclerosis; DNA methylation; entropy

check for updates

Citation: Alfonso Perez, G.; Caballero Villarraso, J. An Entropy Approach to Multiple Sclerosis Identification. *J. Pers. Med.* **2022**, *12*, 398. https://doi.org/10.3390/jpm12030398

Academic Editors: Youxin Wang and Ming Feng

Received: 21 January 2022
Accepted: 3 March 2022
Published: 4 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Multiple sclerosis (MS) is a chronic autoimmune illness affecting the brain and spinal cord associated with various degrees of disability. In MS, the immune system of the patient attacks the axons, more specifically, the myelin cover; see Figure 1 for a graphical illustration [1]. Inflammation is highlighted by some researchers as one of the drivers of neurodegeneration in MS [2–4]. The evolution of the illness varies greatly from patient to patient, with some individuals experiencing long periods of remissions due to mechanisms that are not yet well understood. The usual manifestation age of the illness is from 20 to 45 years old, but it can occasionally manifest at younger ages, even in children [5]. The causes of MS remain unclear, with a complex underlying combination of genetic and environmental factors the most likely cause [6–10].

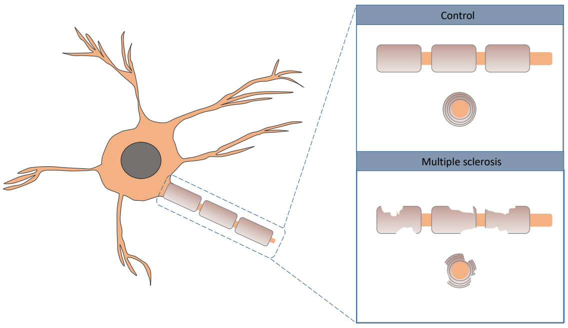


Figure 1. Graphical illustration of neurological damage in MS.

There are some gender considerations to take into account, as the illness is more common in women than men in a 3:1 ratio (and in some countries like Sweden even 5:1).

Some of the common symptoms of the illness include fatigue and numbness, typically in one side of the body [11,12]. Behavioral and cognition abnormalities are also common [13–15]. Currently there are many therapeutic approaches to control or stop the progression of the disease, but no curative treatment is available. However, a large amount of research has been generated regarding this disease. MS has a particularly high prevalence in some areas of Europe and the United States, particularly in northern regions [16].

CpG DNA methylation data has been used to analyze neurodegenerative diseases such as Alzheimer's [17–20] and Parkinson [21–23]. As can be seen in Figure 2, in the context of DNA methylation, CpG dinucleotide (or CpG) refers to cytosine followed by a guanine in the same DNA strand (typically 5′ to 3′), not to be confused with cytosine and guanine pared in two complementary strands.

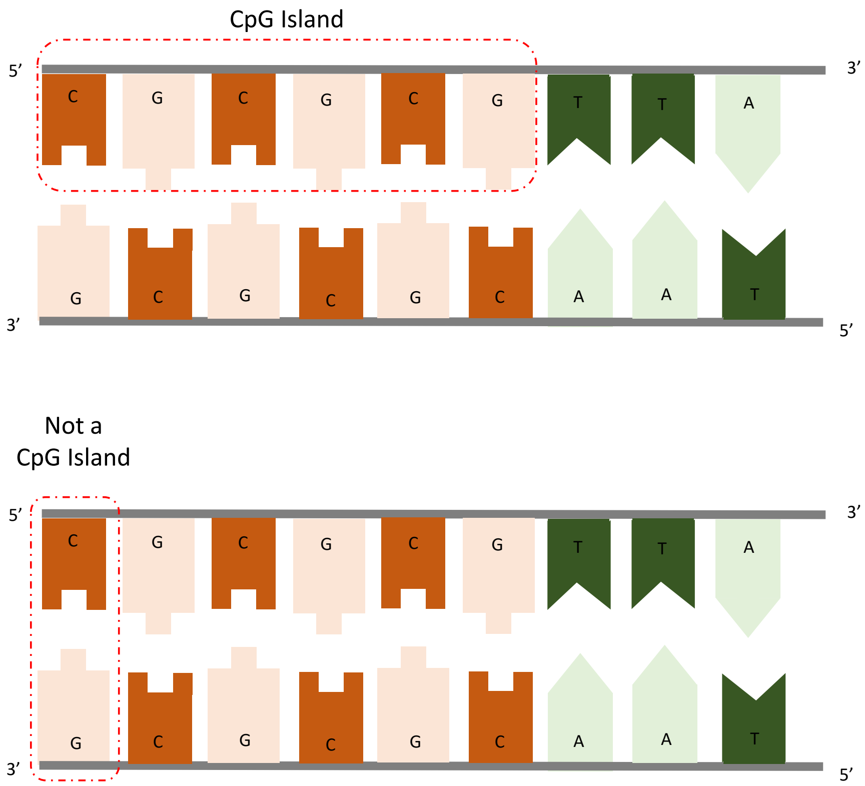


Figure 2. Illustration of CpG islands.

Methylation is simply the addition of a methyl group at the 5-carbon (see Figure 3). DNA methylation has been extensively studied in the context of aging, with several biological clocks built using such types of data. Technological advances in recent years have made possible the analysis of DNA methylation levels on thousands of CpGs in a fast and reliable way. In practice, what is obtained is the percentage level of methylation with a value ranging from 0 to 1 (100% methylated). DNA methylation for cancer diagnostics has made significant progress in the last decades, including many seminal papers [24–27]. There is also a significant body of research covering diabetes [28–32].

DNA methylation has also been used in the context of multiple sclerosis [33,34]. Most of the existing literature on the topic tends to use linear approaches. In this paper, we have followed a non-linear approach, which is in principle more generic and encompassing than a linear approach. Machine learning techniques have been successfully used in multiple applications of different types of diseases [35–38]. More specifically, neural networks have been used as an algorithm for the identification of neurodegenerative illnesses, such as Alzheimer's, using DNA methylation data as the input [39–41].

We applied the concept of Shannon Entropy in the context of DNA methylation applied to multiple sclerosis identification. As far as we are aware, this approach has not been followed before. Shannon Entropy is a concept initially developed in information theory, which attempts to quantify the amount of information contained in a certain set of data [42]. The precise mathematical definition of this concept will be introduced in the materials and methods section. It will be shown that using the concept of Shannon Entropy for CpG selection can generate accurate results.

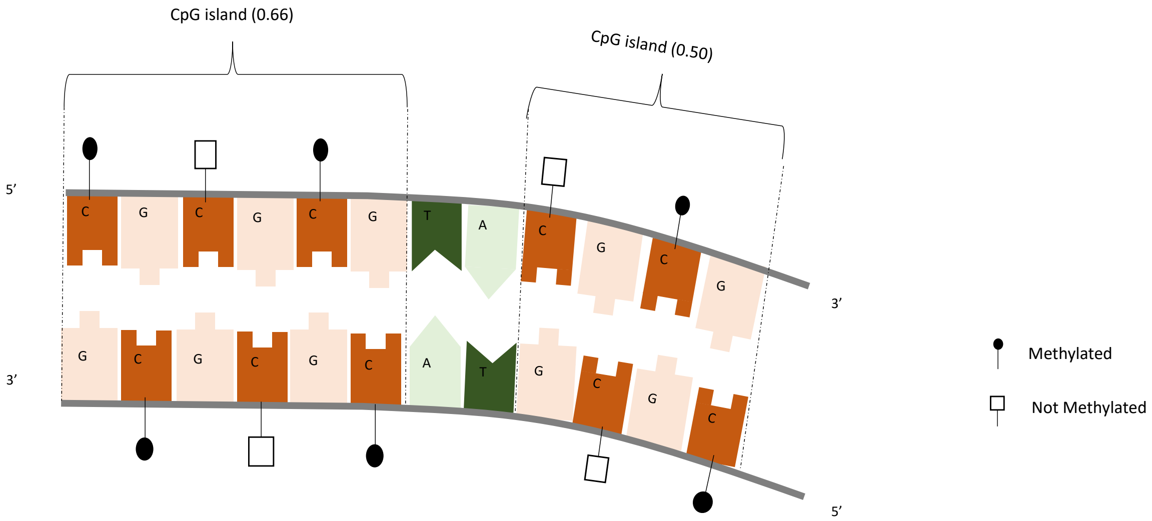


Figure 3. DNA methylation illustration.

Motivation and Aims

Biomarkers are an increasingly important field, particularly when they can be analyzed using non or minimally invasive techniques. In this regard, blood is a particularly interesting tissue as it can be cheaply and quickly obtained from a patient causing only minimal discomfort. Blood has a significant advantage over other tissues such as brain matter, which is much harder to obtain. DNA methylation data can be accurately and rapidly analyzed using technologies such as the Illumina machines. Shannon Entropy is a concept frequently used in machine learning. The motivation to use this approach for data selection is in trying to find techniques that might reduce the dimensionality of the data. Shannon Entropy is one of the few concepts in the existing literature directly related to the amount of information contained in the data, which seems to be a reasonable starting point when trying to reduce the dimensionality of the data while maintaining as much information as possible.

The aim of this article is to develop techniques to identify DNA methylation signatures applicable for the identification of multiple sclerosis patients.

2. Materials and Methods

The DNA methylation data for each individual was stored in a vector X^i.

$$X^i = \left\{ \begin{array}{c} X_1^i \\ X_2^i \\ \vdots \\ X_m^i \end{array} \right\} \tag{1}$$

where m is the number of CpGs analyzed per patient. A numerical example would be:

$$X^2 = \left\{ \begin{array}{c} 0.211 \\ 0.723 \\ \vdots \\ 0.983 \end{array} \right\} \tag{2}$$

Which represents all the CpG information available for patient number 2. In this example, the methylation level in the first and second CpGs are 21.1% and 72.3%, respectively. As there is a large number of cases analyzed it is more convenient to group the data in a matrix form.

$$X = \begin{pmatrix} X_1^1 & X_1^2 & \dots & X_1^n \\ X_2^1 & X_2^2 & \dots & X_2^n \\ \vdots & \vdots & & \vdots \\ X_m^1 & X_m^2 & \dots & X_m^n \end{pmatrix} \tag{3}$$

In this notation, there are n cases (including both patients and controls) with m CpGs associated with each case. The status of the individual analyzed (multiple sclerosis or control) was defined with a binary variable $\{0, 1\}$ stored in a target vector T, with the value 0 indicating a healthy control case and the value 1 indicating a patient with multiple sclerosis.

$$T = \{0, 1, 0, \dots, 1\} \tag{4}$$

As there are n cases, there will be n entries for this vector. In this example, the first and third cases are control cases, and the second one a patient with MS. As a preliminary step, each CpG was individually linearly modeled against the classification vector T and only those with a p-value below 5% were included. The rest of the CpGs were discarded. The dimension of X was reduced from $(n \cdot m)$ to $(n \cdot l)$, where l is the number of CpGs with a p-value below 5%. p-value prefiltering was carried out in all the data. The Shannon Entropy (H) concept was then used to further filter the number of CpGs used. The Shannon Entropy approach step was carried out only for the training dataset. Shannon Entropy can be intuitively understood as the amount of information contained in some data and it is a concept borrowed from information theory. The mathematical expression for Shannon Entropy is as follows:

$$H = -\sum_i P_i log_2(P_i) \tag{5}$$

This concept is typically applied in discrete mathematics. The probabilities can be estimated empirically. In simple terms, more entropy translates into more information contained. After the initial filtering, the absolute value of the Shannon Entropy was estimated for each CpG.

$$H = \left\{ \begin{array}{c} H_1 \\ H_2 \\ . \\ . \\ . \\ H_l \end{array} \right\} \tag{6}$$

Only CpGs with an entropy value (H_i) bigger than certain predefined value (H_i^f) were considered. All the other CpGs were excluded from the analysis. In this way we obtained H^*.

$$H^* = \left\{ \begin{array}{c} H_1^* \\ H_2^* \\ . \\ . \\ . \\ H_q^* \end{array} \right\} \tag{7}$$

In this notation $q \leq l$. After selecting the CpGs, it is necessary to choose the classification algorithm that is used. A neural network with a hidden layer and an output layer was used. The hidden layer contained 50 artificial neurons, while the output layer contained a single artificial neuron. The 50 neurons in the hidden layer are of the sigmoid symmetric transfer function type. The neuron in the output layer is of the type sigmoid positive transfer function (both of these transfer functions are built-in in Matlab). All the neurons include a bias factor. The neural network was trained with the scaled conjugate backpropagation algorithm. Another four learning algorithms were tested (Levenberg–Marquardt, resilient backpropagation, one-step secant and gradient descent). As in the case of the transfer functions in the artificial neural networks, the learning algorithms are also built-in options in Matlab. Among all the learning algorithms, the best results were obtained using the scaled conjugate backpropagation approach. The data was divided into a training and a testing dataset. The testing dataset accounted for approximately 15% of the data. All the calculations were carried out in Matlab. Neural networks have been extensively used for modeling purposes and can accurately describe many complex underlying dynamics. An important step is to check that the classification error obtained using the above mentioned Shannon Entropy approach for CpG selection is more accurate than the one obtained when using the same number of randomly selected CpGs; in other words, controlling that the improvement in accuracy is not simply due to the reduction in the dimensionality of the data.

All the calculations were done in Matlab, the Shannon Entropy value was calculated using an existing Matlab function. The methylation data was analyzed using two decimals of precision in percentage terms. The analysis did not appear to be very sensitive to an increase to the third decimal place, but it started to have more impact thereafter (four or five decimal places in percentage terms). We believe that using two decimal places is a reasonable precision considering the likely accuracy of the experimental data.

A sensitivity analysis was also carried out. The underlying assumption was that CpGs with very little data variation would be less useful for classification purposes. In an extreme case, if the DNA methylation level for a given CpG was the same for all patients, then this information would not be useful for classification purposes. We did not assume that the CpGs with the most data variation (measured as the standard deviation) were necessarily the best choices, as other factors such as experimental noise (and potentially many others) can increase the variation of the data. However, it seemed reasonable to carry out a sensitivity analysis over reasonable values of the volatility of the DNA methylation data.

Data

DNA methylation data for 279 individuals were obtained from the GEO database (publicly available data) with the accession code GSE 106648 [43]. The database contained both individuals with multiple sclerosis (140) as well as control individuals (139). The age range was from 16 to 66 years old, and there were 77 male individuals. There were more females than male patients. This is consistent with the observation that MS tends to be more common among females than males; 138 of the individuals in the dataset were smokers. Age, gender and smoking status (Table 1) were used as inputs in the model. As in the case of DNA methylation, these factors were allocated to their corresponding training or testing dataset.

Table 1. Basic descriptive information of the patients.

Description	Amount
Male	77
Female	202
Smokers	138
Non-smokers	141
Age	16, 77

The DNA methylation data [43] was obtained from peripheral blood tissue using the Illumina Human Methylation 450 Beach Chip. There were 485,512 CpG DNA methylation data per patient.

3. Results

As can be seen in Figure 4, the average classification error using all the available data with a *p*-value below 5% was 55.4%, while the error obtained when using only the CpGs with the top 10% Shannon Entropy values (9499 CpGs) was 19.93%, which is a statistically significant improvement. Equivalently, the proposed approach (using Shannon Entropy as a filter) generated a successful classification rate of approximately 80.07%, while the direct approach (using all the data) generated a successful classification rate of approximately 44.6%. The direct approach likely generates poor classifications due to the issue of local minima, which is likely improved by the introduced Shannon Entropy filtering. The model accuracy was substantially improved while at the same time reducing the amount of input data required in the mode. After the two steps (*p*-value filtering and Shannon Entropy filtering), the amount of CpGs was reduced by approximately 98% compared to the total initial data available. These results were obtained by dividing the data into training and testing datasets, with the testing dataset not used during the training phase. The testing dataset contained approximately 15% of the total data. Unless explicitly mentioned, all the results shown below refer to the testing dataset results. All the models controlled for age, gender and smoking status of the patients. As it can be seen in Table 2, the average sensitivity and specificity obtained were 78.3% and 81.8%, respectively. An example showing a confusion matrix and ROC can be seen in Figures 5 and 6. The obtained list of CpGs can be found in the supplementary files.

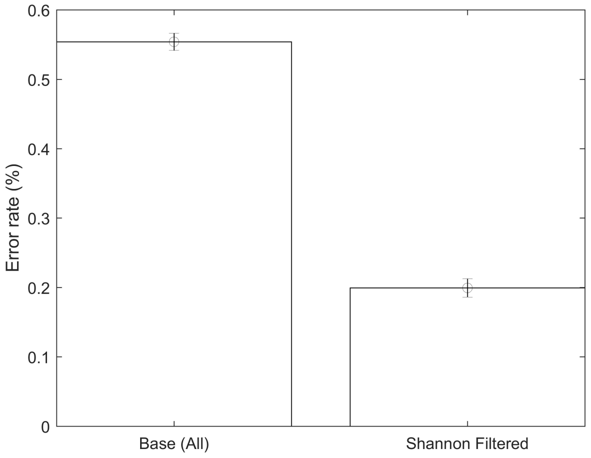
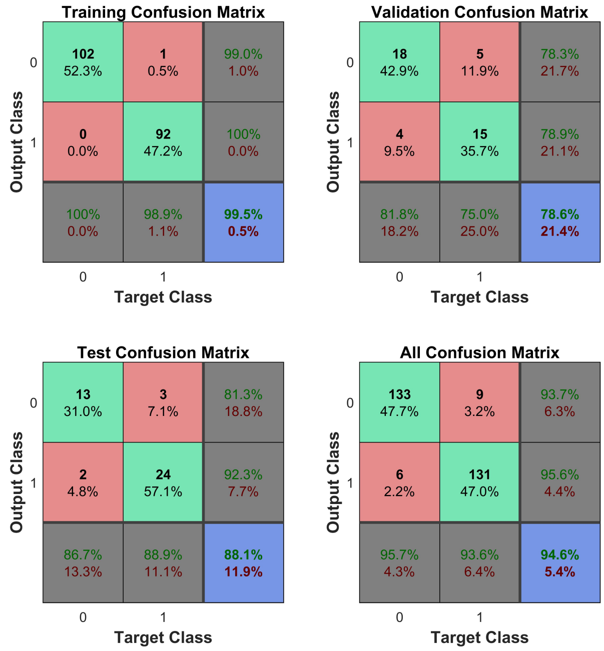


Figure 4. Error rate comparison between direct approach and Shannon Entropy filtered approach.

Table 2. Average classification forecasting accuracy.

Accuracy Measure	Percentage
Average successful classification	80.1%
Sensitivity	78.3%
Specificity	81.8%



Training Confusion Matrix

Output Class \ Target Class	0	1	
0	102 52.3%	1 0.5%	99.0% 1.0%
1	0 0.0%	92 47.2%	100% 0.0%
	100% 0.0%	98.9% 1.1%	99.5% 0.5%

Validation Confusion Matrix

Output Class \ Target Class	0	1	
0	18 42.9%	5 11.9%	78.3% 21.7%
1	4 9.5%	15 35.7%	78.9% 21.1%
	81.8% 18.2%	75.0% 25.0%	78.6% 21.4%

Test Confusion Matrix

Output Class \ Target Class	0	1	
0	13 31.0%	3 7.1%	81.3% 18.8%
1	2 4.8%	24 57.1%	92.3% 7.7%
	86.7% 13.3%	88.9% 11.1%	88.1% 11.9%

All Confusion Matrix

Output Class \ Target Class	0	1	
0	133 47.7%	9 3.2%	93.7% 6.3%
1	6 2.2%	131 47.0%	95.6% 4.4%
	95.7% 4.3%	93.6% 6.4%	94.6% 5.4%

Figure 5. A sample confusion matrix (after *p*-value prefiltering and Shannon Entropy filtering).

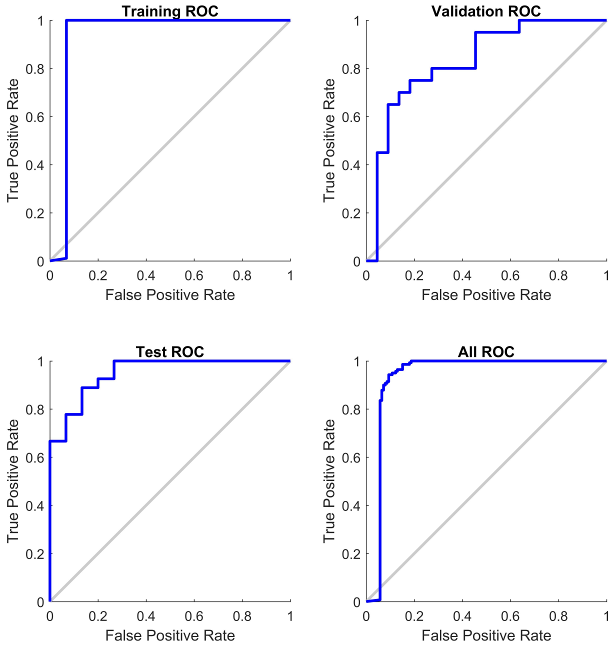


Figure 6. ROC (after *p*-value prefiltering and Shannon Entropy filtering).

In order to compare the results, two baseline values were obtained using the volatility (standard deviation) as an indicator. In the first baseline case, the top 2% most volatile CpGs were selected without any prefiltering (such as *p*-value). This was done in order to have a dimensionality comparable to the results obtained using the proposed approach (*p*-value prefiltering plus Shannon Entropy filtering). The classification success ratio using this technique was approximately 51.6%. A second base line level was obtained. In this case, *p*-value prefiltering was carried out followed by a selection of the most volatile CpGs. The threshold value for the volatility was selected in order to make the final dimension of the data, i.e., number of CpGs selected, approximately the same as the one obtained in the proposed approach (*p*-value plus Shannon filtering). The successful classification rate was 56.1%.

An important test to carry out is comparing the performance of the obtained CpGs by the Shannon Entropy approach (as inputs for the classification algorithm) to the results using a matrix of randomly selected CpGs. In this way, we account for the reduction in dimensionality of the data. Ten randomly selected sets of CpGs of the same size as the one obtained using the Shannon Entropy approach (9499) were selected. All the included CpGs in this random approach had *p*-values of less than 5%, i.e., this analysis was carried out after the initial linear filtering. Ten simulations were carried out for each of the ten different randomly selected sets of CpGs. The average value and the confidence interval can be seen in Figure 7. The Shannon Entropy approach generates classifications that are statistically significantly more accurate than a random selection of the same size.

As mentioned in the methods and materials section, a sensitivity analysis using the standard deviation of the DNA methylation data for each CpG was also carried out. In Figure 8, the results of selecting the CpGs with the highest volatility are shown. The range selected encompassed the top 5% to the top 50%, in 5% increments. For example, the first column shows the error rate (misclassifications) when using the top 5% of CpGs according to their standard deviation from the initial pool containing 9499 CpGs (after the initial filtering using Shannon Entropy filtering).

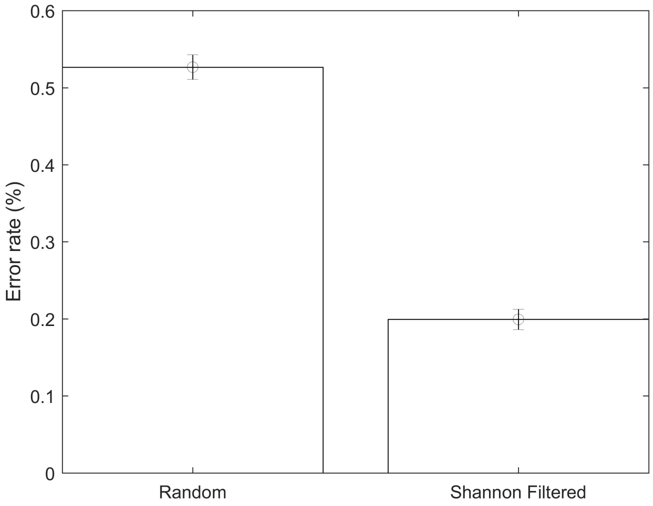


Figure 7. Error rate comparison between the Shannon Entropy filtered approach and random selection of the same size.

The intuition behind this approach is selecting CpGs with variation in the methylation values. As an extreme example, completely flat data (with standard deviation equal to zero) will arguably contain no value from a classification point of view. It is also acknowledged that some of that volatility might be caused by experimental and other sources of noise. The best results were obtained when using the top 15% most volatile CpGs with an average correct classification rate of 81.42%. However, the results were not statistically different (at a 5% significance) when compared with the results obtained by filtering for Shannon Entropy only (no filtering according to the standard deviation of the CpGs).

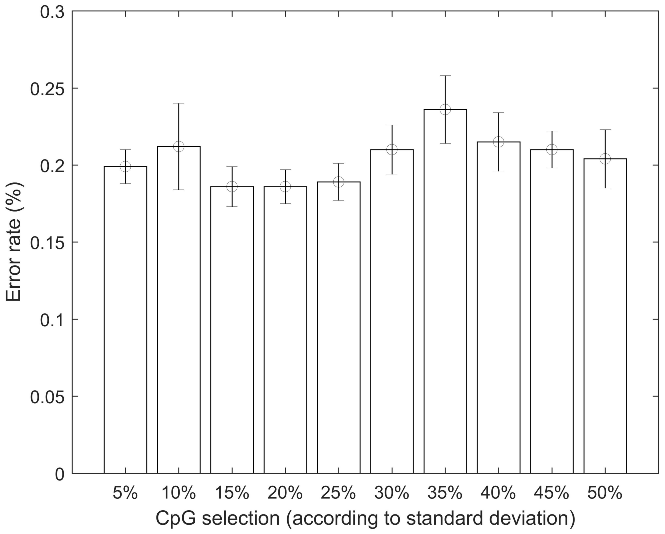


Figure 8. Sensitivity analysis according to the standard deviation of the value of the CpGs. Error rate as a function of the amount of CpGs selected according to their standard deviation.

4. Discussion

An innovative approach is shown for the selection of DNA methylation CpGs to be used in non-linear classification models. This approach is based on the concept of Shannon Entropy, which it is an idea borrowed from the information theory field. Shannon Entropy, in simple terms, can be understood as a measure of the amount of information contained in a set of data. The overall data was first filtered, discarding the CpG with *p*-values above 5%. A quality pre-check of the data was also carried out, excluding CpGs with missing data. The analyzed dataset appeared to be of good quality with no major data issues. Using the two steps approach of *p*-value prefiltering followed by the proposed Shannon Entropy filtering, the dataset was reduced from an original size of approximately 485,512 to a final size of 9499 CpGs, which represents a 98% reduction. The classification analysis, distinguishing between control and multiple sclerosis patients, using the entire dataset, did not generate accurate results. The error rate when using the Shannon Entropy approach was 19.93% (80.07% correct classification), which is a statistically significant improvement over the base case. These error rates were obtained using artificial neural networks as the classification algorithm. All the analyses were carried out controlling for age, gender and smoking status of the patients. It was also tested if the increase in accuracy was due simply to the reduction in the dimensionality of the data. In order to do this, several random CpG configurations of the same size (9499 CpGs) as the one obtained using the Shannon Entropy approach were tested. Their average error rate was 52.66%, which is statistically significantly higher than the results obtained using the Shannon Entropy. This suggests that the Shannon Entropy approach might be a reasonable approach to select potential CpGs relevant for the classification analysis. This type of tool might become rather useful in the future, as the amount of CpGs analyzed per person increases and the computational costs increase accordingly. Another interesting analysis is controlling for the volatility, i.e., the standard deviation, of the CpGs. A sensitivity analysis was carried out in this regard by selecting CpGs according to their standard deviation (in buckets of 5%), i.e., top 5%, top 10%, and so on. When carrying out this type of analysis, there were some improvements in the average accuracy, but these improvements were not statistically significant.

These results were consistent with other articles that found a relationship between DNA methylation in other tissues such as the hippocampus [44]. Using blood as the selected tissue [43] is better suited for clinical purposes. Having a simple test, such as one based on DNA methylation data, which can be applied to many different diseases in a rapid and inexpensive way, can be useful. Multiple sclerosis is a relatively difficult illness to diagnose. Using only clinical symptoms and imaging, such as MRI, is frequently requested when the presence of illness is suspected. From a clinical point of view, it might be practical to have techniques, such as DNA methylation levels in the blood, which can be identified, with a reasonable level of accuracy, the presence of MS with a simple blood test. The physician can use the results from the blood-based biomarker combined with the clinical assessment to decide if it is necessary to carry out further tests, such as imaging.

A very interesting area of future research is the temporal evolution of the DNA methylation in multiple sclerosis, given the diverse evolution of the illness, particularly the long periods of remission experienced by some patients. Further research is necessary to determine feasibility, but it might be possible to use this type of approach for early detection. As more data becomes available, it might be possible to distinguish between different types of illness progression using DNA methylation data. It is possible that differentiating between the different types of evolution might help in targeting therapies in a more precise way.

5. Conclusions

Technical improvements are making possible the generation of large amounts of epigenetic data, such as DNA CpG methylation data, that can be used for the detection of several different types of illnesses, such as multiple sclerosis (MS). Multiple sclerosis is a complex illness with genetic and environmental factors, and importantly, an uncertain evolution with some patients experiencing long periods of remission. In this paper, we present a technique based on the Shannon Entropy concept for the selection of CpGs as inputs for MS identification using non-linear techniques such as artificial neural networks. It was shown that using the proposed approach, the number of CpGs used decreased while the accuracy of the classifications significantly improved. As more DNA methylation data becomes available, it is important to have techniques to efficiently filter these large amounts of information. In this regard, borrowing concepts like Shannon Entropy from other disciplines, such as information theory, might be an interesting approach. Having more data is likely beneficial but not all the new data will be helpful for analysis with a large percentage potentially adding noise. Therefore, it is important to develop techniques to further facilitate quantitative data analysis.

In the future, as more DNA CpG methylation data becomes available, it might be possible to extend this type of analysis in order to identify patients with different types of MS evolution. Currently, MS has no cure, but it is a field of intense research. It is possible that differentiating between the different types of evolution might help in targeting therapies in a more precise way, and this is a very appealing area of future research.

Author Contributions: Conceptualization, G.A.P.; methodology, G.A.P. and J.C.V.; software, G.A.P.; validation, G.A.P. and J.C.V.; formal analysis, G.A.P. and J.C.V.; investigation, G.A.P. and J.C.V.; resources, G.A.P. and J.C.V.; data curation, G.A.P. and J.C.V.; writing—original draft preparation, G.A.P.; writing—review and editing, G.A.P. and J.C.V.; visualization, G.A.P. and J.C.V.; supervision, G.A.P. and J.C.V.; project administration, G.A.P. and J.C.V.; funding acquisition, G.A.P. and J.C.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is available in the GEO Database with accession code GSE 106648.

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

References

1. Sospedra, M.; Martin, R. Immunology of multiple sclerosis. *Annu. Rev. Immunol.* **2005**, *23*, 683–747. [CrossRef]
2. Dendrou, C.; Fugger, L.; Friese, M. Immunophatology of multiple sclerosis. *Nat. Rev. Immunol.* **2015**, *15*, 545–558. [CrossRef] [PubMed]
3. Lassmann, H. Multiple sclerosis phatology. *Cold Spring Harb. Perespect. Med.* **2018**, *l8*. [CrossRef]
4. Frohman, E.; Racke, M.; Raine, C. Multiple sclerosis—The plaque abd its pathogenesis. *N. Engl. J. Med.* **2006**, *9*, 942–955. [CrossRef] [PubMed]
5. Goldenberg, M. Multiple sclerosis review. *Pharm. Ther.* **2012**, *37*, 175–184.
6. Dobson, R.; Giovannoni, G. Multiple sclerosis a review. *Eur. J. Neurol.* **2019**, *26*, 27–40. [CrossRef] [PubMed]
7. Ebers, G. Environmental factors and multiple sclerosis. *Lancet Neurol.* **2008**, *7*, 268–277. [CrossRef]
8. Dyment, D.; Ebers, G.; Sadovnick, D. Genetics of multiple sclerosis. *Lancet Neurol.* **2004**, *3*, 104–110. [CrossRef]
9. Rudick, R.; Cohen, J.; Weinstock-Guttman, B.; Kinkel, R.; Ransohoff, R. Management of multiple sclerosis. *N. Engl. J. Med.* **1997**, *22*, 1604–1611. [CrossRef]
10. Wu, G.; Alvarez, E. The immunopathophysiology of multiple sclerosis. *Neurol. Clin.* **2011**, *29*, 257–278. [CrossRef]
11. Krupp, L. Fatigue in multiple sclerosis. *Arch. Neurol.* **1988**, *45*, 435–437. [CrossRef]
12. Rudick, R.; Schiffer, R.; Schwetz, K.; Herdon, R. Multiple sclerosis: The problem of incorrect diagnosis. *Arch. Neurol.* **1986**, *43*, 578–583. [CrossRef]
13. Feinstein, A. The neuropsychiatry of multiple sclerosis. *Can. J. Psychiatry* **2004**, *49*, 157–163. [CrossRef]
14. Chiaravalloti, N.; DeLuca, J. Cognitive impairment in multiple sclerosis. *Lancent Neurol.* **2008**, *7*, 1139–1151. [CrossRef]
15. Heldner, M.; Kaufmann-Ezra, S.; Gutbrod, K.; Bernasconi, C.; Bigi, S.; Blatter, V.; Kamm, C. Behavioral changes in patients with multiple sclerosis. *Front. Neurol.* **2017**, *8*, 437. [CrossRef]
16. McFarlin, D.; McFarland, H. Multiple sclerosis. *N. Engl. J. Med.* **1982**, *307*, 1246–1251. [CrossRef]
17. Liu, D.; Wang, Y.; Jing, H.; Meng, Q.; Yang, J. Mendelian randomization integrating GWAS and mQTL data identified novel pleiotropic DNA methylation loci for neuropathology of Alzheimer's disease. *Neurobiol. Aging* **2021**, *97*, 18–27. [CrossRef]
18. Mastroeni, D.; Grover, A.; Whiteside, C.; Coleman, P. Epigenetic changes in alzheimer's disease: decremens in DNA methylation. *Neurobiol. Aging* **2010**, *31*, 2025–2037. [CrossRef]
19. Bollati, V.; Galimberti, D.; Pergoli, L.; Dalla Valle, E.; Barretta, F.; Cortini, F. DNA methylation in repetitive elementsand Alzheimer disease. *Brain Behav. Inmunity* **2011**, *25*, 1078–1083. [CrossRef]
20. Blanch, M.; Mosquera, J.; Ansoleaga, B.; Ferrer, I.; Barrachina, M. Altered mitochondrial DNA methylation pattern in Alzheimer disease-related pathology and in Parkinson disease. *Am. J. Pathol.* **2016**, *186*, 385–397. [CrossRef]
21. Masliah, E.; Duamop, W.; Galasko, D.; Desplants, P. Distinctive patterns of DNA methylation associated with Parkinson disease: Identification pf concordant epigenetic changes in brain and peripherak blood leukocytes. *Epigenetics* **2013**, *8*, 1030–1038. [CrossRef]
22. Miranda-Morales, E.; Meier, K.; Sandoval-Carrillo, A.; Salas-Pacheco,J.; Vazquez-Cardenas, P.; Arias-Carrion, O. Implications of DNA methylation in Parkinson's disease. *Front. Mol. Neurosci.* **2017**, *10*, 225. [CrossRef]
23. Wulner, U.; Kaut, O.; Piston, D.; Schmitt, I. DNA methylation in Parkinson's disease. *J. Neurochem.* **2016**, *139*, 108–120. [CrossRef]
24. Chan, K.A.; Jiang, P.; Chan, C.W.; Sun, K.; Wong, J.; Hui, E.P.; Ng, S.S.; Chan, H.L. Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18761–18768. [CrossRef]
25. Lehmann-Werman, R.; Neiman, D.; Zemmour, H.; Moss, J.; Magenheim, J.; Vaknin-Dembinsky A.; Rubertsson, S.; Nellgard, B.; Blennow, K.; Zetterberg, H.; et al. Identification of tissue specific cell death using methylation patterns of circulating DNA. *Proc. Natl. Acad. Sci. USA* **2016**, *29*, 1826–1834. [CrossRef]
26. Guo, S.; Diep, D.; Plongthongkum, N.; Fung, H.L.; Zhang, K.; Zhang, K. Identification of methylation haplotype blocks aids in deconvolution pf heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. *Nat. Genet.* **2017**, *49*, 635–642. [CrossRef]
27. Chen, L.H.; Pan, C.; Diplas, B.H.; Xu, C.; Hansen, L.J.; Wu, Y.; Chen X.; Geng, Y.; Sun, T.; Sun, Y.; et al. The integrated genomic and epigenomic landscape of brainsteam glioma. *Nat. Genet.* **2020**, *11*, 3077.
28. Bell, C.; Christopher, G. Genome-wide DNA methylation analysis for diabetic nephropathy in type 1 diabetes melitus. *BMC Med. Genet.* **2010**, *3*, 33.
29. Bansal, A.; Pinney, S. DNA methylation and its role in the pathogenesis of diabetes. *Pediatr. Diabetes* **2017**, *18*, 167–177. [CrossRef]
30. Davegardh, C.; Garcia-Calzon, S.; Bacos, K.; Ling, C. DNA methylation in the pathogenesis of type 2 diabetes in humans. *Mol. Metab.* **2018**, *14*, 12–25. [CrossRef]
31. Rakyan, V.; Beyan, H.; Down, T.; Hawa, M.; Maslau, S.; Anden, D.; Leslie, R. Identification of type 1 diabetes-associated DNA methylation variable positions that precede disease diagnosis. *Epigenomics* **2015**, *7*, 451–460. [CrossRef] [PubMed]
32. Ronn, T.; Ling, C. DNA methylation as a diagnostic and therapeutic target in the battle agaisnt Type 2 diabetes. *PLoS Genet.* **2011**, *7*, 451–460.

33. Bos, S.; Page, C.; Andreassen, B.; Elboudwarej, E.; Gustavsen, M.; Briggs, F.; Barcellos, L. Genome-wide DNA methylation profiles indicate CD8+ T cell hypermethylation in multiple sclerosis. *PLoS ONE* **2015**, *10*, e0117403. [CrossRef] [PubMed]
34. Kukalova, O.; Kabilov, M.; Danilova, L.; Popova, E.; Baturina, O.; Tsareva, E. Whole-Genome DNA methylation analysis of peripheral blood mononuclear cells in multiple sclerosisi patients with different disease courses. *Acta Nat.* **2016**, *8*, 103–110.
35. Cruz, J.; Wishart, D. Applications of machine learning in cancer prediction and prognosis. *Cancer Inform.* **2006**, *2*, 59–77. [CrossRef]
36. Fan, Y.; Li, Y.; Bao, X.; Zhu, H.; Lu, L.; Yao, Y.; Li, Y.; Su, M.; Feng, F.; Feng, S.; et al. Development of Machine Learning Models for Predicting Postoperative Delayed Remission in Patients With Cushing's Disease. *J. Clin. Endocrinol. Metab.* **2021**, *106*, 217–231. [CrossRef]
37. Kourou, K.; Exarchos, T.; Exarchos, K.; Karamouzis, M.; Fotiadis, D. Machine learning applications in cancer prognosis and prediction. *Comput. Struct. Biotechnol.* **2015**, *13*, 8–17. [CrossRef]
38. Li, Y.; Chen, Z. Performance evaluation of machine learning methods for breast cancer prediction. *Appl. Comput. Math.* **2018**, *7*, 212–216. [CrossRef]
39. Park, C.; Ha, J.; Park, S. Prediction of Alzheimer's disease based on deep neural network by integrating gene expression and DNA methylation dataset. *Expert Syst. Appl.* **2020**, *140*, 112873. [CrossRef]
40. Alfonso Perez, G.; Caballero Villarraso, J. Alzheimer Identification through DNA Methylation and Artificial Intelligence Techniques. *Mathematics* **2021**, *9*, 2482. [CrossRef]
41. Spolnicka, M.; Pospiech, E.; Peplonska, B. DNA methylation in EVOVL2 amd C!orf132 correctly predicted chronologicalage of individuals from three disease groups. *Int. J. Leg. Med.* **2018**, *132*, 1–11. [CrossRef]
42. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
43. Kular, L.; Liu, Y.; Ruhrmann, S.; Zheleznyakova, G.; Marabita, F.; Gomez-Cabrero, D.; James, T.; Ewing, E.; Linden, M.; Gornikiewicz, B.; et al. DNA methylation as a mediator of HLA-DRB1*15:01 and a protective variant in multiple sclerosis. *Nat. Commun.* **2018**, *9*, 2397. [CrossRef]
44. Chomyk, A.M.; Volsko, C.; Tripathi, A.; Deckard, S.A.; Trapp, B.D.; Fox, R.J.; Dutta, R. DNA methylation in demyelinated multiple sclerosis hippocampus. *Bell Syst. Tech. J. Sci. Rep.* **2017**, *18*. [CrossRef]