Cognitive analysis of multiple sclerosis utilizing fuzzy cluster means

Imianvan Anthony Agboizebeta.¹ and Obi Jonathan Chukwuenyi².

¹Department of Computer Science, Faculty of Physical Sciences, University of Benin, Benin City, Edo State, Nigeria.

tonyvanni@yahoo.com

²Department of Computer Science, Faculty of Physical Sciences, University of Benin, Benin City, Edo State, Nigeria.

tripplejo2k2@yahoo.com

Abstract

Multiple sclerosis, often called MS, is a disease that affects the central nervous system (the brain and spinal cord). Myelin provides insulation for nerve cells improves the conduction of impulses along the nerves and is important for maintaining the health of the nerves. In multiple sclerosis, inflammation causes the myelin to disappear. Genetic factors, environmental issues and viral infection may also play a role in developing the disease. MS is characterized by life threatening symptoms such as; loss of balance, hearing problem and depression. The application of Fuzzy Cluster Means (FCM or Fuzzy C-Mean) analysis to the diagnosis of different forms of multiple sclerosis is the focal point of this paper. Application of cluster analysis involves a sequence of methodological and analytical decision steps that enhances the quality and meaning of the clusters produced. Uncertainties associated with analysis of multiple sclerosis test data are eliminated by the system

Keywords:

Fuzzy logic, Fuzzy Set, Fuzzy Cluster Means, Diagnosis, Multiple sclerosis, Membership Function

1. Introduction

Multiple sclerosis (MS) was first described in 1868 by Jean-Martin Charcot after observing the personal diary of Sir Augustus d’Este’, born 1794 grandson of King George III of England, reveals a medical history strongly suggesting that Augustus suffered from multiple sclerosis [3].

Multiple sclerosis (MS) or Demyelinating disease [12] is a disease in which the nerves of the central nervous system (brain and spinal cord) degenerate [13, 17 and 31]. Myelin, which provides a covering or insulation for nerves, improves the conduction of impulses along the nerves and also is important for maintaining the health of the nerves. In multiple sclerosis, inflammation causes the myelin to disappear [17]. Consequently, the electrical impulses that travel along the nerves decelerate, that is, become slower and the nerves themselves are
damaged. As more and more nerves are affected, a person experiences a progressive interference with functions that are controlled by the nervous system such as vision, speech, walking, writing, and memory. Multiple sclerosis is a progressive autoimmune disease and the most common neurological disease diagnosed in young adults [31].

Multiple sclerosis is triggered by a combination of genetic factors and environmental elements, such as exposure to a virus or some infections. Others causes include autoimmune diseases, neuromuscular disease, brain conditions, demyelinating disorder, spinal cord disorders, coordination and balance conditions [13, 17 and 31].

Symptoms of multiple sclerosis can vary greatly between individuals. In the early stages of the disease, symptoms come and go, and people with multiple sclerosis can experience periods of remission, in which symptoms disappear and periods of relapse, in which symptoms reappear. Symptoms may worsen during relapses and complications can include seizures, changes in mutation, muscle spasms, and depression [23 and 31].

Symptoms generally first appear between the ages of 20 and 40 years [17 and 24]. Most often, the first symptom is visual changes. Other symptoms may include dizziness, muscle weakness, in coordination, balance problems, fatigue, trembling, abnormal sensations in the extremities, cognitive impairment, and paralysis. Testing for multiple sclerosis is an involved process because there is no single test that will specifically diagnose the condition. In addition, a definitive diagnosis of multiple sclerosis may often be delayed because early symptoms are often mild, vague, disappear for periods of time, and/or may mimic other conditions. Multiple tests are needed to rule out other conditions that can be confused with multiple sclerosis or that have similar symptoms affecting the nervous system. These include some viral infections, lupus, lime disease, and vitamin deficiencies [13 and 31].

A diagnosis of multiple sclerosis is generally made by a neurologist and is made based on the compilation of a thorough neurological history and physical, blood tests, a lumbar puncture, and an MRI (magnetic resonance imaging), which may show the brain and spinal cord damage and lesions characteristic of multiple sclerosis [23 and 31]. There is no cure for multiple sclerosis, but the disease can be managed and symptoms controlled to various degrees of success with a variety of medications, such as interferons, copolymer, novantrone, natalizumab, corticosteroids, muscle relaxers and Intrathecal Baclofen Pump Therapy [17]. Other therapies include physical therapy, avoiding heat, regular exercise, eating a healthy well-balanced diet and maintaining an optimal weight. Equipment, such as walkers, canes, and braces may also be needed to help maintain independence [17 and 31].

Prognosis of multiple sclerosis is highly variable. Some people have very minor problems, whereas others can end up in a wheelchair, lose vision, or other nasty complications. Depending on which part of the spinal cord or nervous system is attacked, some people have what is called a "single sclerosis" [31] where there is only one incident of a sclerosis (lesion) on the spinal cord, and this may or may not progress to "multiple" sclerosis.

Fuzzy logic provides a means for representing and manipulating data that are not precise, but rather fuzzy [27]. Fuzzy logic presents an inference morphology that enables appropriate
The theory of fuzzy logic encompasses a mathematical strength to capture the uncertainties associated with human cognitive processes. This paper presents a Fuzzy Cluster Mean (Fuzzy C-Mean or FCM) knowledge-based model for the diagnosis of multiple sclerosis focusing on two forms; Relapsing/Remitting Multiple Sclerosis and Primary Progressive Multiple Sclerosis. Fuzzy C-means clustering algorithm is being used as the problem solving and reasoning algorithm in the inference engine of the knowledge base system for the evaluation, classification and matching of patterns to more than one class of multiple sclerosis. Statistics, Neural network Neurology Physiology psychology and are also incorporated in developing the proposed systems. The main objective is to classify and match any individual independently with more than one cluster depending on the degree of membership.

2. Literature Review

The following are type of Multiple Sclerosis [28]

**Benign Multiple Sclerosis** describes a type of relapsing-remitting multiple sclerosis in which few relapses occur. These relapses tend to produce sensory symptoms, which go away and leave very little, or no residual damage or disability [16 and 21]. In Relapsing/remitting multiple sclerosis patients tend to experience an attack or series of attacks (exacerbations) followed by complete or partial remission [6]. Secondary Progressive Multiple Sclerosis (SPMS) begins with relapsing remitting multiple sclerosis. The relapsing remitting (RRMS) stage of the disease may persist for many years before the onset of secondary progressive MS [7].

**Primary progressive multiple sclerosis** (PPMS) is most commonly found in men [8]. It is characterized by gradual clinical decline with no real or distinct periods of remission. There may be temporary periods where the disease appears to plateau, or level out, and this may include some partial, yet minor, relief from some symptoms, however the course of the disease is one of continual decline from the outset.

**Optic-Spinal form of Multiple Sclerosis**

The optic-spinal form of multiple sclerosis (OSMS), characterized by recurrent involvement of optic nerve and spinal cord with rare brain magnetic resonance imaging lesions [18]. Pseudotumoral form of multiple sclerosis is a very rare form of multiple sclerosis.

In Familial multiple sclerosis the MS destroys myelin, the fatty sheath that protects nerve fibers carrying message traffic from various muscles to and from the central nervous system. For reasons currently unknown, in some people the myelin sheath breaks down, resulting in destruction of the nerve fibers and the symptoms of MS [14].

**Acute fulminant multiple sclerosis** Malignant Multiple Sclerosis, is a particularly aggressive form of the disease. It is very rare. This highly aggressive form is defined by its swift and relentless decline to significant disability or even death, often within a few weeks or months after the onset of the initial attack. It is characterized by widespread and progressive cerebral white matter destruction or by severe pathological involvement of clinically strategic regions such as brainstem, resulting in bulbar paralysis [12].
Balo’s Concentric Sclerosis is a rare demyelinating disorder characterized pathologically by concentric rings of alternating demyelinated and relatively myelin preserved white matter. The pathogenesis of the concentric lesion may be explained by periodic suppression of demyelination in a rapidly expanding area of inflammation, allowing remyelination or only transient incomplete demyelination to occur [30].

Fuzzy Logic (FL) is a branch of machine intelligence that helps computers paint pictures of uncertain world. Fuzzy sets were introduced by [32] as a means of representing and manipulating data that are not precise, but rather fuzzy. Fuzzy logic provides an inference morphology that enables appropriate human reasoning capabilities to be applied to knowledge-based systems. The theory of fuzzy logic provides a mathematical strength to capture the uncertainties associated with human cognitive processes, such as thinking and reasoning. A fuzzy set $A$ is called trapezoidal fuzzy number (Figure 1) with tolerance interval $[a, b]$, left width $\alpha$ and right width $\beta$ if its membership function has the following form

$$A(t) = \begin{cases} 
1 - (a - t)/\alpha & \text{if } a - \alpha \leq t \leq a \\
1 & \text{if } a \leq t \leq b \\
1 - (t - b)/\beta & \text{if } a \leq t \leq b + \beta \\
0 & \text{otherwise}
\end{cases}$$

and we use the notation $A = (a, b, \alpha, \beta)$. It can easily be shown that $[A][\gamma] = [a - (1 - \gamma)\alpha, b + (1 - \gamma)\beta], \forall \gamma \in [0, 1]$. The support of $A$ is $(a - \alpha, b + \beta)$.

![Figure 1: Trapezoidal fuzzy number.](image)

A fuzzy number $A$ is a fuzzy set of the real line with a normal, (fuzzy) convex and continuous membership function of bounded support. The family of fuzzy numbers will be denoted by $F$. A quasi fuzzy number $A$ is a fuzzy set of the real line with a normal, fuzzy convex and continuous membership function satisfying the limit conditions.

$$\lim_{t \to -\infty} A(t) = 0, \quad \lim_{t \to \infty} A(t) = 0.$$
2.1 Fuzzy Clustering

Clustering involves the task of dividing data points into homogeneous classes or clusters so that items in the same class are as similar as possible and items in different classes are as dissimilar as possible [2 and 31]. Clustering can also be thought of as a form of data compression, where a large number of samples are converted into a small number of representative prototypes or clusters [10]. Depending on the data and the application, different types of similarity measures may be used to identify classes, where the similarity measure controls how the clusters are formed [2 and 15]. Some examples of values that can be used as similarity measures include distance, connectivity, and intensity [5].

In non-fuzzy or hard clustering, data is divided into crisp clusters, where each data point belongs to exactly one cluster [1]. In fuzzy clustering, the data points can belong to more than one cluster, and associated with each of the points are membership grades which indicate the degree to which the data points belong to the different clusters [18].

2.2 Overview of Fuzzy Cluster Means (FCM) Algorithm

The FCM algorithm is one of the most widely used fuzzy clustering algorithms. The FCM algorithm attempts to partition a finite collection of elements X={X₁, X₂,...,Xₙ} into a collection of c fuzzy clusters with respect to some given criterion [29]. Given a finite set of data, the algorithm returns a list of c cluster centers V, such that V = Vᵢ, i=1, 2,..., c and a partition matrix U such that U = Uᵢⱼ, i=1,..., c   j =1,..., n where Uᵢⱼ is a numerical value in [0, 1] that tells the degree to which the element Xᵢ belongs to the i-th cluster [29].

The following is a linguistic description of the FCM algorithm, which is implemented by fuzzy Logic. The algorithm is as follows;

**Step 1:** Select the number of clusters c (2 ≤ c ≤ n), exponential weight μ (1 < μ < ∞), initial partition matrix U₀, and the termination criterion ε. Also, set the iteration index l to 0.

**Step 2:** Calculate the fuzzy cluster centers {Vᵢᵢ i=1, 2,..., c} by using Uᵢ.

**Step 3:** Calculate the new partition matrix Uᵢ₊₁ by using {Vᵢᵢ i=1, 2,..., c}.

**Step 4:** Calculate the new partition matrix = || Uᵢ₊₁ - Uᵢ || = | Uᵢⱼ i=1, 2,..., c. If > ε, then set l = l + 1 and go to step 2. If ≤ ε, then stop.

The initial cluster centers are computed in two ways; Arithmetic means of all the data points or running FCM several times each starting with different initial cluster centers. In this work the first method is adopted.
3.0 Methodology

This system which employs fuzzy C-Means for the diagnosis of multiple sclerosis is developed in an environment characterized by Microsoft Window, XP professional Operating System, Microsoft Access Database Management system, Visual Basic Application Language and Microsoft Excel. Neuro Solutions and Crystal reports were used for neural network analysis and graphical representation. An approach for analyzing clusters to identify meaningful pattern for determining whether a patient suffers from multiple sclerosis or not is presented. The system provides a guide for diagnosis of a multiple sclerosis within the decision-making framework.

The process for the medical diagnosis of multiple sclerosis starts when an individual consults a physician (doctor) and presents a set of complaints (symptoms). The physician then requests further information from the patient or from others close to him who knows about the patient’s symptoms in severe cases. Data collected include patient’s previous state of health, living condition and other medical conditions. A physical examination of the patient condition is conducted and in most cases, a medical observation along with medical test(s) is carried out on the patient prior to medical treatment. From the symptoms presented by the patient, the physician narrows down the possibilities of the illness that corresponds to the apparent symptoms and make a list of the conditions that could account for what is wrong with the patient. These are usually ranked in possibility order (Low, Moderate and high). The physician then conducts a physical examination of the patient, studies his or her medical records and ask further questions, as he goes in an effort to rule out as many of the potential conditions as possible. When the list has been narrowed down to a single condition, it is called differential diagnosis and provides the basis for a hypothesis of what is ailing the patient. Until the physician is certain of the condition present; further medical test are performed or schedule such as medical imaging, scan, X-rays in part to conform or disprove the diagnosis or to update the patient medical history. Other Physicians, specialist and expert in the field may be consulted (sought) for further advices.

Despite all these complexities, most patient consultations are relatively brief because the disease is obvious or the physician’s experience may enable him to recognize multiple sclerosis condition quickly. Upon the completion of the diagnosis by the physician, a treatment plan is proposed, which includes therapy and follow-up (further meeting and test to monitor the ailment and progress of the treatment if needed). Review of diagnosis may be conducted again if there is failure of the patient to respond to treatment that would normally work. The procedure of diagnosing a patient suffering from multiple sclerosis is synonymous to the general approach to medical diagnosis. The physician may carry out a precise diagnosis, which requires a complete physical evaluation to determine whether the patient have multiple sclerosis. The examining physician accounts for possibilities of having multiple sclerosis through an interview, physical examination and laboratory test. Some primary health care physicians use screening tools for multiple sclerosis evaluation.

A thorough diagnostic evaluation may include a complete history of the following:

a. When did the symptoms start?
 b. How long have the symptoms lasted?
 c. How severe are the symptoms?
 d. Have the symptoms occurred before, and if so, were they treated and what treatment was received?
4.0 Results and Discussion

To design the FCM Knowledge Base System for diagnosis of multiple sclerosis, we design a system which consists of a set of parameters needed for diagnosis, because nerves in any part of the brain or spinal cord may be damaged, patients with multiple sclerosis can have symptoms in many parts of the body (here, we are using ten basic and major parameters) presented in Table 1.

Figure 4 presents the model of the FCM system for the diagnosis of multiple sclerosis. It comprises of knowledge base system, fuzzy c-means inference engine and decision support system. The knowledge base consists of the database Engine. The knowledge base system holds the symptoms of multiple sclerosis.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>P01</td>
<td>Loss of balance</td>
</tr>
<tr>
<td>P02</td>
<td>Frequent need to urinate</td>
</tr>
<tr>
<td>P03</td>
<td>Double vision</td>
</tr>
<tr>
<td>P04</td>
<td>Facial pain</td>
</tr>
<tr>
<td>P05</td>
<td>Difficulty solving problem</td>
</tr>
<tr>
<td>P06</td>
<td>Depression or feeling of sadness</td>
</tr>
<tr>
<td>P07</td>
<td>Problem with erection</td>
</tr>
<tr>
<td>P08</td>
<td>Trouble chewing and swallowing</td>
</tr>
<tr>
<td>P09</td>
<td>Hearing loss</td>
</tr>
<tr>
<td>P10</td>
<td>Decreased attention span and poor judgment</td>
</tr>
</tbody>
</table>

Table 1: Symptoms of multiple sclerosis
The values for the diagnosis of multiple sclerosis symptoms are not precise hence the adoption of fuzzy logic as a means of analyzing these information. The different forms of multiple sclerosis symptoms, therefore, constitute the fuzzy parameters of the knowledge base system. The fuzzy set of parameters is represented by ‘P’, which is defined as P= {P₁, P₂, …, Pₙ}
where \( P_i \) represents the \( j^{th} \) parameter and \( n \) is the number of parameters (in this case \( n=10 \)). Neural network provides the structure intelligent learning for all forms of multiple sclerosis symptoms, which serves as a platform for the inference engine. The inference engine consists of reasoning algorithms, driven by production rules. These production rules are evaluated by using the forward chaining approach of reasoning. The fuzzy logic and fuzzy C-means algorithm provides the rules for the partitioning of patients into a number of homogenous clusters with respect to a suitable similarity measure.

In this paper, the patients were classified into two form of multiple sclerosis according to physician. **Relapsing/remitting multiple sclerosis (RMS)** refers to patient experiencing 5 or more of the listed symptoms during an attack or series of attacks (exacerbations) followed by complete or partial remission while **Primary Progressive Multiple Sclerosis (PPMS)** refer to patients experiencing at least four of the listed symptoms because of their gender (commonly found in men).

Each of the symptoms highlighted in Table 1 is represented with \( P \) (starting from 01 – 10 i.e. \( P_{01} - P_{10} \)).

Table 2, we form two clusters namely Relapsing/remitting multiple sclerosis and Progressive Multiple Sclerosis (PPMS) which represents a typical scenario of the degree of membership, for instance, \( P_{10} \) in cluster 1, we notice it has 0.61. In term of percentage it can be represented as 61% and in cluster 2, 39% .This means that the degree of symptoms of \( P_{10} \) matches 61% of relapsing/remitting multiple sclerosis and 39% Progressive Multiple Sclerosis (PPMS). The FCM clustering distribution shown in Table 2, depicts a total of six symptoms with high degree of membership of Relapsing/remitting multiple sclerosis (RMS) and four symptoms with high degree of membership of Primary Progressive Multiple Sclerosis (PPMS).
Table 2: FCM membership grade of all patients in all clusters of Multiple Sclerosis

<table>
<thead>
<tr>
<th>Codes</th>
<th>DEGREE OF MEMBERSHIP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CLUSTER 1 (RMS)</td>
</tr>
<tr>
<td>P01</td>
<td>0.60</td>
</tr>
<tr>
<td>P02</td>
<td>0.70</td>
</tr>
<tr>
<td>P03</td>
<td>0.25</td>
</tr>
<tr>
<td>P04</td>
<td>0.10</td>
</tr>
<tr>
<td>P05</td>
<td>0.62</td>
</tr>
<tr>
<td>P06</td>
<td>0.20</td>
</tr>
<tr>
<td>P07</td>
<td>0.27</td>
</tr>
<tr>
<td>P08</td>
<td>0.80</td>
</tr>
<tr>
<td>P09</td>
<td>0.89</td>
</tr>
<tr>
<td>P10</td>
<td>0.61</td>
</tr>
</tbody>
</table>

5.0 Conclusion

The need to design a system that would assist doctors in medical diagnosis has become imperative and hence cannot be over emphasized. This paper present a diagnostic fuzzy cluster means system to help in diagnosis of Multiple sclerosis using a set of symptoms. This advanced system which uses clustered data set is more precise than the traditional system. The classification, verification and matching of symptoms to the two groups of clusters (Relapsing/remitting multiple sclerosis and Primary Progressive Multiple Sclerosis) was necessary especially in some complex scenarios. This paper demonstrates the practical application of ICT (Information and communication technology) in the domain of diagnostic pattern appraisal in medicine by determining the extent of membership of individual symptoms. The model proposed allows for the classification of and matching of cluster groups to multiple sclerosis symptoms. The fuzzy-cluster means model proposed in this paper appears to be a more
natural and intelligent way of classification and matching of symptoms to multiple sclerosis groups.

References


Imianvan Anthony Agboizebeta received his B.Sc and M.Sc degree in Computer Science from the University of Benin, Benin City, Nigeria in 1993 and 1998 respectively. He obtained Ph.D degree in Computer Science from Federal University of Technology Akure, Nigeria in 2009. Imianvan teaches at the University of Benin, Benin City, Nigeria. Field of interest: Soft Computing, Intelligent Agent, Tele-medicine and related fields.
Imianvan Anthony Agboizebota

Jonathan Chukwuyeni Obi received his B.Sc degree in 2008 from the University of Benin, Nigeria. He is currently pursuing his Master of Science degree (M.Sc) at the University. He is working as a Graduate Assistant at the Department of Computer Science, University of Benin. His research interests are soft computing, Artificial Intelligence, Telemedicine and Environmental data mining. He has published several papers in international journal and conference.

Jonathan Chukwuyeni Obi