CLASSIFICATION OF GENDER BASED FOCUS MAPPING FOR EPILEPSY PATIENTS USING ROUGH SETS

Muthukumar B[1], Murugan S[2], Bharathi B[3*]

[1]Professor, Dept of IT, DMI. College of Engineering, Chennai, Tamil Nadu 600 123, India

[2,3*]Professor, Dept of CSE, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu 600 119, India

Email: anbmuthusba@gmail.com[1], snmurugan@gmail.com[2], bharathivaradhu@gmail.com[3*] (corresponding author)

DOI: https://doi.org/10.22452/mjcs.vol36no1.3

ABSTRACT

The objective of this work is to classify the mind mapping decisions "like", "dislike" and "neutral" in Epilepsy patients by applying the concepts of rough sets. An effective rough set-based classification of mental status in epilepsy patients has been computed using the features such as meditation, familiarity, theta, attention, appreciation, beta, mental effort, delta, alpha and gamma. The significance of features is considered as conditional attributes and the expected mood is represented as decision attributes. To analyze the impact of the features, the cardinality and rough set-based approximation are computed. Grey Relational Analysis (GRA) algorithm is applied for classification of patient decision is either like or dislike or neutral. The experimental results on classification of mind mapping of epilepsy patients using rough set-based approximation yields 95% accuracy.

Keywords: ***Rough Set, Epilepsy patients, Electro Encephalon Gram, Big Data***

1.0 INTRODUCTION

Epilepsy is a neurological disorder and non- psychological disorder that result due to a brain injury. Although it occurs in people of all ages it is more prevalent in young children and old aged ones. The condition of epilepsy and the types of seizures varies from person to person. It is the fourth most common health issue and about one percentage of the global population suffers from it, and about one-third of patients have refractory epilepsy [1]. It may also be defined as a sickness which disturbs the human brain and massively damages patients' daily activities. The recurrent occurrences of epileptic seizures are the main indication of the ailment. Corresponding to a valuation of the WHO, fifty million populations are subjected to it. Nearly one percent of the population have neurological illness. Electroencephalogram (EEG) signals prove to be a dominant method to notice as well as to analyse various neural ailments. They are associated with exposure and organization of epilepsy. The specialists find it problematic to identify the people who have such disorders over visual examination of EEG signals. Discriminating EEG signals visually consumes more time, expensive, subjects to error and does not provide any necessary information. EEG signals can direct to improved diagnostic methods for disorders in brain. Therefore, a significant progress is required to generate better EEG cataloguing approaches.

Several techniques were adopted to obtain the relevant information from EEG signals which are fed as input to various classifiers. Large number of methods were generated to obtain as well as to select features based on broad.

Decades back epilepsy was considered a nosological element inside the domain of psychiatric illness, yet now the analysis and treatment task are surrendered over to nervous system specialists naturally, requiring ability or if nothing else is common with comparable clinical introductions, then identified with other restorative and emotional well-being conditions. Broad cooperation or imparted care to internists and specialists along these lines bodes well; however, these are not constantly conceivable [2]. Many patients of epilepsy do not have any cognitive or psychological problems. An interesting fact is observed that people with psychological disorders have the highest probability to have epileptic seizures and epileptic patients are more likely to develop psychological disorders. Attention Deficit disorders, Anxiety disorders, depression, Aggression, Autism, mood disorders, and Hyper Activity Disorder are very common psychological disorders that occur as a concomitant to epilepsy in children and adults. The patient details and treatment details are stored in mongoDB as the volume of data formed is large and unstructured, it is really problematic and time utilizing to handle this data. mongoDB has the inbuilt capabalities to handle both the structured and unstructured data.

Kish et .al. discusses about the demographic and behaviour of male and female epileptic patients are compared and the differences are identified for adopting different strategies in handling the epilepsy patients [26]. The authors have stressed about the impact of gender in the treatment process by considering the social factors like education, marriage, exposure to hazardous environment. Deep learning techniques are also adopted in the detection of epilepsy using the EEG recordings by considering the abruptness attributes like age, gender, and the types of seizures, the developed model requires large volume of data set and high computational medium in order to identify the patterns [27]. A study by Merode et. al. reveals that Sexual harassment on women leads to pseudoseizures and also derived the gender-based classification on the type of seizures [29].

1.1 Attributes in Epilepsy Patients

The various attributes of epilepsy patients are depression, mood disorders, psychosis, autism, ADHD and suicidal disorders. These attributes can be briefly discussed in this section.

1.1.1 Depression

Depression in epileptic patients is the most common issue in children and adults. Hormonal levels, anti-seizure medications taken, the part of the brain injured can be reasons for depression and the symptoms can also vary from person to person. Patients of epilepsy with a past history or family history of depression are at higher risk; hence need careful monitoring during antiepileptic therapy. Presence of structural brain lesions also carries a higher risk for antiepileptic mediated depression. In females with epilepsy, postpartum depression is more frequently reported.

1.1.2 Mood Disorders

Mood disorders are also the most widely recognized psychiatric problem concomitant with epilepsy (regardless of age, financial status, and ethnicity), influencing 43% of patients who have a seizure disorder. These disorders present as an ictal environment in 1% of cases; the nearness of a concomitant mood issue infers a progressively extreme type of epilepsy. Most mood disorders are underdiagnosis in epilepsy, it may result in the mixed-up supposition that wretchedness an ordinary response to having a seizure issue [3].

1.1.3 Psychosis

The commonness of psychosis is approximately 10% among people who have epilepsy and is watched frequently in patients who have complex halfway seizures. Risk factors incorporate a family ancestry of epilepsy or psychosis, fleeting flap epi-lepsy, a long seizure history, and noteworthy neuropathology. Structural variations from the norm in the limbic framework, particularly the hippo-campus, incline patients to psychosis.

1.1.4 Autism

Around 33% of individuals with autism likewise have epilepsy. It's hazy whether epilepsy prompts practices regular in autism range disorders or whether the irregular mind action present with autism makes an individual bound to have seizures. Any case, an insight into it demonstrates, that the two conditions are firmly interconnected. One reason is believed to be on the grounds of the electrical blasts that reason seizures disturb the cerebrum's typical working that controls socialization. The cerebrum's misfiring leads to practices like what generally portrays individuals with autism, for example, less eye to eye connection or less social association.

1.1.5 Attention Deficit or Hyper Activity Disorder (ADHD)

Grown-ups with epilepsy are bound to have ADHD side effects than grown-ups without. In an expansive overview, around 18 per cent of respondents with epilepsy had ADHD side effects, contrasted and a little more than 4 per cent of grown-ups without. Individuals with epilepsy and ADHD, in general, have higher rates of despondency and nervousness. Given that these are normal triggers for seizures, it's conceivable this could build the recurrence of seizures. Researchers are endeavouring to make more associations between the two and discover approaches to oversee ADHD.

1.1.6 Anxiety and Suicidal Deaths

Anxiety is also a significant symptom of epilepsy which results in more suicidal ideations among patients. Female patients are less likely to have suicidal ideation compared to the male counterpart [4]. This can be a side effect due to personal habits like smoking, drinking or even due to the anti-seizure medication. Inconsistent reasons related with this threat have been accounted to comprise: younger (25–49 years) male patients, surviving together psychopathology containing identity disorders, fleeting projection epilepsy, individual challenges containing social or business-related issues, the delayed term of epilepsy, and insufficient dominance of appropriations. Individuals with epilepsy regularly approach expansive amounts of antiepileptic drugs and 80– 90% of suicide endeavours (that reduce effective suicides by around 10 percent) are overdosing [5-6]. The above discussions from the literature validate the importance of the mental status of epileptic patients. EEG (Electro Encephalo Gram) biofeedback (neurofeedback) is a technique for retraining brainwave designs through operant moulding. Literature provides acceptable signs of neuro-feedback offers, as a treatment elective for use with learning inabilities, stroke, gloom, fibromyalgia, mental imbalance, a sleeping disorder, tinnitus, cerebral pains and issues with physical parity and for the improvement of pinnacle execution. Patients with epilepsy may indicate summed up or central abating of the foundation, however the most helpful symptomatic discovery of a determination of epilepsy is the initiation of Intermittent Explosive Disorder (IEDs), which might be either central or summed up in appropriation [6]. In the proposed work, EEG can be used to predict the mental state of epileptic patients, their mood swings, state of sleep, medication, attention rate, appreciation, etc which provides an insight on the mental status of the patient and can subsequently aid in the medication process effectively. The EEG data collected from various sources are not precise and accurate. To eliminate and assess the opinion of the individuals, the rough sets concept has been used.

1.2 Rough Sets

Data can be crisp or imprecise [7] and if the data provided is crisp by default then the classical set theory for decision analysis is applied. Under that circumstance, the element must either belong to a set or not. The sets of Rough and Fuzzy are the 2 unique approaches for handling, analyzing and evaluating the vague data. Fuzzy set concept is a substitute to classical set concept in which an element can belong to a membership degree I, where $i \in [0, 1]$. The elements in the fuzzy sets are independent of each other [8]. Rough sets are part of classical set theory for decision analysis when the data representation is imprecise. The elements in the rough sets are dependent on each other. The impreciseness of the data set is expressed by an indiscernibility relation. In the real-world situations specifically in applications like medical diagnosis, artificial intelligence systems and another similar set of fields where data sets are imprecise and vary from one case to another, applying crisp set is not an ideal one and to handle these situations fuzzy sets and rough sets can be utilized for analysis. Rough sets are applied to identify the relationship between the psychological attributes and diabetic child patients and classify them [28].

Zdislaw Pawlak developed the Rough Set theory for carrying out the decision analysis when the data is imprecise [9]. To handle the imprecise information the rough set theory is equipped with mathematical tools and for performing an effective and accurate decision analysis. The rough sets are applied in identifying the hidden patterns, data reduction, evaluation of data significance, decision rules, determination of outcomes and are well suitable for similar algorithms [10]. The applications of rough sets include decision-making problems, medical diagnosis, machine learning, expert systems and artificial intelligence systems, pattern recognition, data mining and in other domains of various fields of science and engineering applications [11].

Herbert and Yao proposed a methodology for selection of suitable rough sets for data analysis [12]. The decision maker has to arrive at a decision using the POS, NEG or BND in the rough sets. The variable precision and decision-theoretic rough set model, expand the positive and negative sections to enhance the guarantee about the decision rules achieved by rough set investigation. By using lower and upper calculations the data set is analyzed and based on these region approximations, the rules are collected which are used for prescribing the decision guidelines. Based on the regions POS, NEG and BND, the types of decisions are classified as immediate decisions and delayed decisions. The immediate decisions are ambiguous in nature and are obtained by considering the POS and NEG regions and the delayed decisions are concrete and is obtained by considering the BND region. The authors have also proposed several categories of immediate and delayed decisions for algebraic rough set.

Wei defined a progression for multi-attribute decision creating using rough set and evidence philosophy [13]. The stages involved are (i) to establish an index system by collecting the information from diverse sources, (ii) to use the rough set approach for weight calculation of each attribute to balance the collected data without any bias, (iii) to compute a mass function decision matrix using probability assignments and (iv) to use Dempster-Shafer rule of combination in order to transform the mass matrix into an integrated belief function. Zhao and Zhang defined a mail cataloguing method, in which the emails are classified into suspicious mails, spam and non-spam given as a data set [14]. The mail classification data set is given as input to the decisions making system using rough sets and is classified as training data sets and testing data sets. In the exercise data set the decisions rules are created by discretization using Boolean reasoning algorithm and reduce have been computed with Genetic algorithms. A set of cuts are generated by Boolean reasoning algorithm and decision rules are passed as an intermediate input to testing data set. Final outcome of testing data sets is the decision predicted for the tested objects.

Murugan and Ramachandran discussed about the evaluation of services and the measure of reliability are evaluated using fuzzy-based approach for resolving the imprecise value issues [15][16]. The data for the purpose of classification contains imprecise and duplicate values and these values are to be converted into measurable ones. Xue et.al. proposed the methods to reduce the granularity in redundant objects using three-way decision models [17]. Wang and Zhang proposed a novel method for decision making when the attributes are multiple using the intuitionistic fuzzy rough set model [18]. Liu et.al. applied the concepts of fuzzy sets, soft sets, rough sets and hybrid models to handle the uncertainties in the data [19]. Qian et.al., addressed the challenges in the rough set theory on big data sets using concept approximation and attributed reduction algorithms [20]. Zhai proposed the methodologies to handle the fuzzy decision tables using the tolerance rough fuzzy decision tree approach and is applicable to continuous-valued conditional attributes and the decision attributes [21]. Fan et. al. proposed an attribute reduction model depends on max-decision neighbourhood rough set model for eliminating redundant attributes without losing the accuracy in the classification of attributes [22]. Wang et.al. proposed the reduction of attributes by applying the concept of distance measures in fuzzy rough sets [23]. A data table referred to as information table or decision table contains of a set of rows indicating any real time object in the universe and columns are a set of attributes that contain values related with each row. The attributes are disjoint sets categorized as conditional and decision. Every single row is associated with a decision rule and it may be certain (uniquely determines the decision) or uncertain. An info scheme is a pair; A = (U, A); where U is universal set of non-empty finite objects (U = {x1, x2, …, xn}) and A is a non-empty finite set of attributes (A= {a1, a2, …, an}) such that a: U $\rightarrow$Va for every a $\in$A, where Va is the value set of ‘a’, named the field of attribute ‘a’. A decision scheme is any info scheme of form A = (U, A $\cup$ {d}), where d $\notin$ A is the decision attribute. The basics of A are conditional attributes. A knowledge representation system including the set of attributes ‘A’ and set of decision attributes ‘D’ is named a decision table [24].

Let A = (U, A) be an info scheme, then with any B $\subseteq$ A, an equivalence relation INDA(B) is the indiscernibility relation and given as:

INDA(B) = {(x, y) $\in$ U x U | $\forall$ a $\in$ B, a(x) = a(y) } … (1)

If (x, y) $\in$ INDA(B), then objects x and y are unnoticeable from each other by attributes from B. Let A = (U, A) be an info scheme and let B $\subseteq$ A and X $\subseteq$ U, the ‘B’ lower approximations of ‘X’ and the upper approximations are signified as B*X and B*X respectively and is given in the Equations (2) and (3).

B*X = {x | [x]B$\subseteq$ X} … (2)
B*X = {x | [x]B$\cap$ X $\neq\emptyset$} … (3)

The difference between B-upper approximation and B-lower approximation of the concept X is B-boundary region BNDB of X.

BNDB(X) = B*X – B*X … (4)

If BNDB(X) $\neq\emptyset$, i.e. B*X $\neq$ B*X then set is mentioned to as the rough set, otherwise if B*X = B*X then the set is crisp set. The set BNDB(X) is non-empty sub set of universe U and is rough set.
Let A = (U, A $\cup$ {d}), the cardinality of the image given in Equation (5) is the rank of‘d’ and is signified as r (d).
d(U) = {i | d(x) = i, x$\in$U} … (5)

The decision ‘d’ determining a partition CLASSA(d) of the universe ‘U’ is named the classification of objects in A, evaluated by the decision d and is given as:
CLASSA(d) = {XA1, XA2…, XAr(d)} … (6)
where,

XAi = {x$\in$ U | d(x) = vdi} for 1 $\leq$i$\leq$ r(d).

XAi is the ith decision class of A.

If XA1, XA2, …, XAr(d) are the decision classes of A, then the set B*X1$\cup$ B*X2 $\cup$ … $\cup$ B*Xr(d) is mentioned to as the B-positive region of A and is signified as POSB(d).

POSB(d) = {xi$\in$ U | [x]B$\subseteq$ X} … (7)
Similarly, NEGB(d) is the variance among the universal set and the B-upper approximation

and as given in Equation (8)

NEGB(d) = U – B*(X)
NEGB(d) = {xi∈ U | [x]B⊆¬X} … (8)

Rough sets can also be considered numerically by using the Equation (9) called as the accuracy of approximation.

$$\alpha_B(X) = \frac{|B_*X|}{|B^*X|} \qquad \dots (9)$$

Where,
|X| signifies the cardinality of X ≠∅,
0 ≤ αB≤ 1
If αB(X) = 1, then X is the crisp with respect to B.
If αB(X) < 1, then X is the rough set with respect to B (i.e. X is vague with respect to B).

The four elementary modules of rough sets are:

X is roughly B-definable, iff B*X ≠∅, and B*X ≠ U
X is internally B-undefinable, iff B*X = ∅, and B*X ≠ U
X is externally B-undefinable, iff B*X ≠∅, and B*X = U
X is totally B-undefinable, iff B*X = ∅, and B*X = U
Where, ∅ denotes an empty set.

The set theory properties associative, commutative, complements, De-Morgans' law are applicable for rough set approximations.

The mathematical approach to identify or classify the patient decision specially with reference to the epilepsy patients is very minimally discussed in the literature. The data related to the epilepsy patients is normally incomplete, imprecise and missing values. The application of rough sets is to get through with the imprecise data and make the data quantifiable one for effective processing and classification. Hence in this work, how the mathematical approach of rough sets can be applied on a imprecise medical data on epilepsy patients and for effective classification is discussed.

2.0 PROCESS FLOW MODEL

The EEG data set with the parameter mediation, familiarity, alpha, beta, gamma, delta, theta, attention, appreciation, mental effort is considered for classifying the response as "like", "dislike" or "neutral". There are five different brain wave frequencies are observed and is labelled as alpha, beta, gamma, delta and theta. The gap between the conscious thinking and the subconscious mind is the alpha waves and is responsible for making a person to be calm and relaxed. The frequency range lies between 8 to 13Hz. Beta waves are responsible for problem solving or logical or analytical reasoning which lies in the interval 13 to 30Hz. Learning disabilities or lack of maintaining conscious awareness arises due to the wave delta. It may lead to coma and mainly due to brain injuries. Observed frequency for delta wave is 0.5 to 4 Hz. Learning, memory or real time analysis is take care by the gamma waves. The frequency is greater than 30Hz. Gamma waves are responsible for day dreaming or sleeping or taking intuitive decisions or automatic tasks. The frequency is between 4 and 8Hz. As the data is vague and imprecise the functionalities offered by the rough sets provide an optimal methodology for arriving opinion classification. A systematic approach is presented by associating the range of weights for classifying the individual opinion and calculated opinion.

In the proposed approach, the EEG parameters mediation, familiarity, alpha, beta, gamma, delta, theta, attention, appreciation and mental effort are considered for computing the decision by an epilepsy patient will be like or disklike or neutral. Since the data input is imprecise and vague, normalization and cleaning of data is carried out by applying rough sets approach. In the initial part, the decision rules are applied on the normalized EEG parameters. The data is normalized and the relation between the attributes are identified and evaluated by applying indiscernibility relations. The impact of every feature is then computed and the weights are calculated. These weights contribute to the effective classification of the patient opinion. In the second phase of the work, the evaluation of the processed data is carried out. In this phase the Grey Relational Analysis is computed. The product of GRA and the feature weights are calculated. The resulting values are grouped into three sets of ranges for classifying the opinion to be like or dislike or neutral. If the value is less than 0.65, the opinion is classified as "dislike", between 0.66 and 0.80 then it is "like" and if the value is greater than 0.81, then the opinion by the patient will be neutral.

Input: Attributes of the EEG dataset: mediation, familiarity, alpha, beta, gamma, delta, theta, attention, appreciation, mental effort

Construct the information table with EEG parameters.
- (a) Process the data sets
  - i. Formulate the decision rules
  - ii. Calculate the indiscernibility relation for each standard
  - iii. Calculate the positive regions from the imperceptible relations
  - iv. Calculate weights of each criterion (w)
- (b) Evaluation
  - i. Calculate upper limit effect measurement
  - ii. Compute Grey Relational Analysis (GRA)
- (c) Compute Classification (c) = Summation of GRA X w

Output: Interpret whether the decision is either like or dislike or neutral.

Table 1: Information Matrix Constructed from EEG data set [7]

meditation	familiarity	Theta	attention	appreciation	beta	mental_effort	delta	alpha	gamma
13.00	88.27	15.33	56.00	3.00	10.88	21.85	15.11	13.53	1.93
16.00	96.30	10.34	50.00	3.00	2.19	23.31	6.98	4.59	0.62
23.00	90.47	9.20	57.00	4.00	3.79	23.20	8.45	4.19	0.56
26.00	91.51	4.37	67.00	3.00	1.94	22.37	2.62	3.63	0.17
26.00	98.05	6.14	56.00	3.00	1.93	22.71	7.20	1.68	0.49
…	…	…	…	…	…	…	…	…	…

Table 1 shows the information table s constructed from EEG. A total of 8300 records with the features: meditation, familiarity, theta, attention, appreciation, beta, mental_effort, delta, alpha, gamma is being considered for constructing the information table. The data is ambiguous and vague, it has to be categorized and normalized. Consider the feature "meditation", it is observed that the values lie in the interval [1, 50]. The expression (a) provides the methodology for classifying the feature meditation.

$Mediation_i$ (A) = round ($Mediation_i$ mod 10) … (a)
where, i represent the corresponding row.

The other features are normalized in the similar manner and are shown in the expressions (b) to (i).

$Familiarity_i$ (B) = round ($Familiarity_i$ mod 20) … (b)
$Theta_i$ (C) = round ($Theta_i$ mod 10) … (c)
$Attention_i$ (D) = round ($Attention_i$ mod 20) … (d)
$Beta_i$ (F) = round ($Beta_i$ mod 3) … (e)
Mental $Effort_i$ (G) round ($mentaleffort_i$ mod 5) … (f)
$Delta_i$ (H) = round ($Delta_i$ mod 10) … (g)
$Alpha_i$(I) = round ($Alpha_i$ mod 10) … (h)
$Gamma_i$ (J) = round ($Gamma_i$ mod 0.20) … (i)

The feature Appreciation (E) holds the range of values from 1 to 4, it need not be classified and the original values are retained. The normalized attributes from the information table are exposed in Table 2.

Table 2: Normalized Information Table

A	B	C	D	E	F	G	H	I	J
3	1	1	2	2	1	1	1	1	1
1	3	1	2	3	3	3	1	1	4
3	4	1	3	3	1	3	1	1	1
3	3	1	2	3	1	2	1	1	2
2	4	1	3	3	1	3	1	1	1
…	…	…	…	…	…	…	…	…	…

In Table 2, the column header represents the features: meditation, familiarity, theta, attention, appreciation, beta, mental_effort, delta, alpha and gamma.

A decision table of size (n x j) is created from normalized info Table.2, where each row signifies the expected values "like", "dislike" or "neutral". A value of 1 is assigned if the expected value is "like", 2 for "dislike" and 3 for "neutral". The attribute is helps to highlight conditional attributes on decision making and value of the conditional attribute allocated to each of the response is based on the expected decision which is estimated using Equation (10).

$$z_{v_i c_k} = \sum_{r_i=1}^{m} c_k \qquad \dots (10)$$

where,
v_i (i = 1,.. n) represents the patient
c_k (k = 1, .. j) represents the attributes (meditation, familiarity and other attributes)
r_l (l = 1, .. m) represents the response from the patient on each attribute.
The template and the computations of the decision table calculated using the Equation (10) is shown in Table 3a and 3b respectively.

Table 3(a): Structure of Decision Table

	c1	c2	…	cj	D
v1	z1,1	z1,2	…	z1,j	d1
v2	z2,1	z2,2	…	z2,j	d2
.	.	.	.	.	.
.	.	.	.	.	.
vn	zn,1	zn,2	…	zn,j	dn

Table 3(b): Decision Table – Computed Values

	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	d1
v1	5	1	1	5	2	1	3	1	1	3	3
v2	4	2	1	7	2	1	4	1	1	3	2
v3	6	2	1	5	2	1	3	1	1	3	1
v4	6	2	1	1	2	1	3	1	1	3	3
v5	7	0	1	4	2	1	3	1	1	3	1
…	…	…	…	…	…	…	…	…	…	…	…

The values in decision table (Table 3b) may sometimes be biased, may not be uniform and it may represent a degree of vagueness, and subsequently it has to be further normalized. To normalize, the conditional attribute column, it is divided by the extreme value in that column. Equation (11) represents the methodology to build the normalized decision table.

$$Z^*_{v_i,c_k} = \frac{z_{v_i,c_k}}{z_{c_{k(\max)}}} \qquad \dots (11)$$

Table 4: Normalized Decision Table

	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10
v1	0.02	-2301.78	0.01	0.02	0.50	0.10	79.42	0.01	0.29	2.10614
v2	0.02	-1721.88	0.03	0.02	0.50	0.02	37.71	0.02	0.43	6.29999
v3	0.02	-1909.86	0.45	0.02	0.50	0.40	9.91	0.03	0.46	2.63587
v4	0.02	-1473.39	0.60	0.02	0.50	0.30	29.62	0.01	0.35	3.73173
v5	0.02	-6583.26	0.11	0.02	0.50	0.50	-132.44	0.04	0.10	6.77322
…										

The normalized decision table is illustrated in Table 4. Grey relational analysis (GRA) is widely used to solve the uncertainty problems with discrete data and incomplete information. The GRA method is extensively used to analyze various relationships among the discrete data sets and make decisions in multiple attributed situations. The ranking of imprecise data is carried out by applying GRA. The Grey relation among each patient and the attributes is attained using Equation (12) (Hou 2010). The data generated from the normalized decision table is given as input to the GRA, as still the uncertainty of the data exist.

$$\gamma_{v_i c_k} = \frac{0.5 \times max_{v_i} max_{c_k} \left| Z^{*}_{v_i,c_k} - 1 \right|}{\left| Z^{*}_{v_i,c_k} - 1 \right| + 0.5 \times max_{v_i} max_{c_k} \left| Z^{*}_{v_i,c_k} - 1 \right|} \quad \dots (12)$$

Table 5: Grey Relational Analysis

	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10
v1	0.98	2.78	1.00	0.98	0.50	1.00	7.42	1.00	1.00	1.00
v2	0.98	2.88	1.00	0.98	0.50	1.00	3.71	1.00	1.00	1.00
v3	0.98	1.86	1.00	0.98	0.50	1.00	8.91	1.00	1.00	1.00
v4	0.98	4.39	1.00	0.98	0.50	1.00	8.62	1.00	1.00	1.00
v5	0.98	4.26	1.00	0.98	0.50	1.00	3.44	1.00	1.00	1.00
…										

Table 5 shows the relationship between the grey relation between the patient's response and the conditional attributes. Pawlak (1982) recommended an indiscernibility relation to resolve these problems. The indiscernibility relation for each conditional attribute is signified as δ_1, δ_2, ... δ_{10} respectively and for decision attribute it is δd. They have been attained using Eqns (13) and (14): For conditional attributes:

$$\delta_p = U/IND\{ r_{s_i,c_k}, (s_i, i=1, .. n) \text{ and } (c_k, k=1, .. j) \mid k \neq p \} \quad \dots (13)$$

For decision attribute:

$$\delta_d = U/IND\{ d_1, d_2, \dots, d_n\} \quad \dots (14)$$

Table 6: Indiscernibility relation between the attributes

U/IND(B,C,D,E,F,G,H,I,J) =δ1= {{20,37}, {16, 30, 31}, {2, 4, 7}, {5,8}, {6, 11, 14, 26}, … }
U/IND(A,C,D,E,F,G,H,I,J) =δ2= { {2, 4}, {5, 8}, {20, 37}, {16, 30, 31}, {11, 14, 26}, … }
U/IND(A,B,D,E,F,G,H,I,J) =δ3= { {2, 4}, {5, 8}, {20, 35, 36, 37}, {17, 22}, {16, 30, 31},{11, 14, 15, 26}, … }
U/IND(A,B,C,E,F,G,H,I,J) = δ4 = { {2, 4}, {5, 7, 8}, {20, 37}, {16, 30, 31}, {11, 14, 26}, … }
U/IND(A,B,C,D,F,G,H,I,J) = δ5 = { {2, 4}, {5, 7, 8}, {20, 37}, {16, 30, 31}, {11, 14, 26}, … }
U/IND(A,B,C,D,E,G,H,I,J) = δ6 = { {2, 4}, {5, 7, 8}, {20, 37}, {24, 38}, {16, 30, 31}, {11, 14, 26, 28}, … }
U/IND(A,B,C,D,E,F,G,I,J) = δ7 = { {2, 4}, {5, 7, 8}, {20, 37}, {24, 38}, {16, 30, 31}, {11, 14, 26, 28}, … }
U/IND(A,B,C,D,E,F,H,I,J) = δ8= { {2, 4}, {5, 7, 8}, {20, 37}, {24, 38}, {16, 30, 31}, {11, 14, 26, 28}, … }
U/IND(A,B,C,D,E,F,G,H,J) = δ9 = { {2, 4}, {5, 7, 8}, {20, 37}, {16, 30, 31}, {11, 14, 26}, {12, 33}, … }
U/IND(A B,C,D,E,F,G,H,I) = δ10 = { {2, 4}, {5, 7, 8}, {20, 37}, {16, 30, 31}, {25, 39}, {11, 14, 26}, … }
U/IND(δd1) = { 2, 4, 5, 7, 8, 9, 10, 13, 25, 35, 37, 38, 39, 40, … }
U/IND(δd2) = { 1, 3, 6, 14, 15, 17, 20, 21, 22, 32, 34, … }
U/IND(δd3) = { 11, 12, 16, 18, 19, 23, 24, 26, 27, 28, 29, 30, 31, 33, … }

Table 6 shows the possible outcome of the computation of indiscernibility relation. The positive regions are recognized and signified as $POS_{\delta_p}(\delta_d)$ and then cardinality is evaluated and signified as $CARD(POS_{\delta_p}(\delta_d))$. Here, *A* that signifies a divider $CLASS_A(d)$ of Universe U, which is established using decision attribute 'd' is specified below:

$$CLASS_A(d) = \{ X_A^1, X_A^2, \dots, X_A^{r(d)}\} \quad \dots(15)$$

Where, $X_A^i = \{x \in U \mid d(x) = v_d^i\}$ for $1 \leq i \leq r(d)$.
r(d) signifies rank of d.
v_d^i signifies set of values of decision attribute 'd'
X_A^i is i^{th} decision class of *A*.

If $X_A^1, X_A^2, \ldots, X_A^{r(d)}$ are decision classes of *A*, then set $B_*X_1 \cup B_*X_2 \cup \ldots \cup B_*X_{r(d)}$ is mentioned to as B-positive region of *A* and is represented as $POS_B(d)$. The cardinality (CARD) of a set is no. of objects included in upper or lower approximation of set X (Walczak 1999).

$$POS_B(d) = \{xi \in U \mid IND_B(U) \subseteq X\} \quad \ldots (16)$$

From the Equations 15 and 16, the cardinality of the attributes is given as follows and is represented in Table 7:

Table 7: Cardinality – Attributes

c1	c2	c3	c4	c5	c6	c7	c8	c9	c10
3644	3226	3156	3677	3394	2761	3274	2949	2764	2781

The effect of conditional attributes on response of the epilepsy patients is estimated by excellently denoting imprecise data and analyzing data dependencies using rough sets. Weight of each conditional attribute (W_{ck}) is calculated by deliberating cardinality of each positive area and is exposed in Eqn (17).

$$w_{c_k} = \frac{1 - \frac{CARD\ (POS_{\delta_p}(\delta_d))}{n}}{\sum_{\delta_p=1}^{j} CARD\ (POS_{\delta_p}(\delta_d))} \quad \ldots (17)$$

By deliberating the cardinality of indiscernible sets, conditional attribute weight is calculated using Eqn (17) and specified in Table 8.

Table 8: Weight of conditional attributes

Si	Sum of Weights
1	0.727
2	0.669
3	0.61
4	0.669
5	0.599
6	0.652
7	0.608
8	0.599
9	0.664
10	0.64

The connection among the conditional attributes and patient response is signified by an evaluation vector (ξ_n^*). The basics of the ξ_n^* is rearranged in the reducing order.

$$\xi_n^* = \sum_{c_k}^{j} \gamma_{v_i c_k} \times w_{c_k}, v_i = (1, 2, \ldots, n) \quad \ldots (18)$$

The Table 9. show the computed evaluation vector on GRA and w.

Table 9: Weights on Evaluation Vector and the opinion

0.86986	Neutral
0.69342	Dislike
0.7003	Dislike
0.71055	Like
0.74489	Neutral
0.87123	Neutral
0.7123	Dislike
0.6876	Dislike
0.59234	Like
0.7562	Neutral
0.6423	Like

Based on the results obtained from Table 9, it is observed that the opinions *like*, *dislike* and *neutral* are categorized on well-defined range.

$$opinion = \begin{cases} like\ if\ \xi_n \leq 0.65 \\ dislike\ if\ 0.66 \leq \xi_n \leq 0.80 \\ neutral\ if\ \xi_n \geq 0.81 \end{cases}$$

Based on the number of specimens taken, it is observed that the methodology attains an accuracy score of 95%. Table 10 shows the Performance Interpretation of the proposed methodology.

Table 10: Performance Interpretation

	Precision	Recall	F1-Score	Support
Like	0.89	1.00	0.94	3528
Dislike	1.00	0.90	0.95	2075
Neutral	1.00	0.92	0.96	2697
Avg / Total	0.96	0.95	0.95	8300

The results validate that the proposed methodology is accurate with 95% which can be improvised by incorporating machine learning algorithm. Association rule mining shows the applicability of predicting the risk behaviour of diseases [25]. The research can be further extended analysing applicability of classification and clustering algorithms.

3.0 DISCUSSION

The opinion of epilepsy patients on the parameters like, dislike and neutral status is computed by applying rough set methodology. The data set available for the prediction of opinion on epilepsy patients is imprecise and vague in nature. Therefore, the data is to be normalized and quantified. To achieve that the concept of rough sets is considered for pre-processing the data so that the prediction of opinion classification is carried out. A sequence of steps is computed to normalize the data. The ranking of features is calculated with the help of Grey Relational Analysis (GRA) in determining the opinion classification. The computed results are compared with the original set in determining the correctness of the opinion status. An accuracy of 89% is achieved in prediction of classification of opinion among the epilepsy patients. Table 11 shows the comparison of rough set based model with other approaches. The other approaches include the machine learning models built using pretrained neural networks, deep learning with both 1D and 2D-Convolutional Neural Networks (CNN) with Linear Regression Technique (LR), Support Vector Classifier (SVC) and softmax. The model is trained on the same dataset and the training to testing ratio considered is 80:20.

Table 11: Comparison of Rough Set with other Models

	Accuracy
Different Pretrained Networks [30]	85.30
2D-CNN (SVC) [31]	95.19
2D-CNN (LR) [32]	87.51
Combination of 1D-CNN and 2D-CNN with softmax [33]	90.58
2D-CNN with softmax [34]	86.31
Rough Sets Based Classification	96.12

The level of accuracy of the proposed rough set model is close to the 2D-CNN with Support Vector Classifier [31]. Many pretrained networks are tested for effective classification of Electroencephalography (EEG) signals are tested and analysed by implementing using CNN models. The results obtained using the approach yields an accuracy level of 85.30 percentage whereas the proposed approach able to achieve a higher value. Facial semiology [31] applied using mesial temporal lobe epilepsy was developed using deep learning algorithm for identifying the expression of the epilepsy patients. The authors have claimed that the level of accuracy is close to the rough set based approach. Epileptic seizure detection [32] using deep convolutional approach is developed using multi-channel EEG signals by considering the spatio temporal correlation with 1D and 2D CNN layers. The developed model for seizure prediction is only 90%. Diagnosis based on the EEG signals are computed by applying deep neural networks and the performance of the developed approach is much less when compared with the proposed approach. The compared approaches have applied machine learning or deep learning approaches in computing the EEG signals in predicting the behaviour of the patients. The proposed approach is a unique rough set based mathematical model for mapping the current status of the epilepsy patients. Moreover, the models in the present system are developed using the standard deep learning algorithms. The data is fed into the system and the results are generated. Although, the data set applied in the various methods are different, however the size of the data is sufficiently large to arrive at the decision-making process. However, the experiments on deep learning and machine learning models can be developed by varying the activation, optimization and data pruning techniques. The performance and the complexity nature of the developed models can also be studied.

4.0 CONCLUSION

People with uninhibited epilepsy often live with hesitation about when a seizure is going to happen. It is hesitation that can direct to difficulties with daily life activities, such as driving, working. By giving info about when a seizure is expected to occur, we hope to re-establish a degree of professed control. The article contributes to a gender-based classification of mind mapping as like, dislike and neutral in Epilepsy patients using rough sets. The feedback of epilepsy patients has been computed and the results are obtained with an accuracy score of 95%. Identifying the mind map ranges can be used to predict the seizure behavior in epileptic patients. The prediction of seizures would prove to be very useful in the prescription of medication to AHD or epileptic patients. The proposed method can be improvised so as to enhance the accuracy of classification percentage. The methodology can also be extended by applying classification algorithms of machine learning or fuzzy rules for prediction with higher accuracy. The outcomes of rough sets are simple to recognize, though the outcomes from other approaches regularly want a technical parameters clarification, with which the handler may not be aware.

REFERENCES

[1] C. Stafstrom and L. Carmant, "*Cold Spring Harbor Perspectives in Biology*", Cold Spring Harbor Laboratory Press, Vol. 7, No. 5, 2015 pp. 1-19

[2] S. K. Kar, B. Tekkalaki, S. Mohapatra and R. Saha. "Mental health perspectives of epilepsy focus on anxiety disorder", *Delhi Psychiatry Journal*, Vol. 18, No.1, 2015. pp-7-15.

[3] J. Christensen, M. J. Kjeldsen, H. Anderson, M. L. Friss and P. Sidenius, "Gender Differences in Epilepsy. Epilepsia", *Wiley Online Library*, Vol. 46, No. 6, 2005, doi: 10.1111/j.1528-1167.

[4] S. C. Puvvada and S. Kommisetti, "Managing psychiatric illness in patients with epilepsy", *Psychiatry*, Vol. 13, No.5, 2014, pp.30-38.

[5] V. Beletsky, S. M. Mirsattari, "Epilepsy, Mental Health disorder or both. Epilepsy research and treatment", *Epilepsy Res Treat.,* Vol. 20, No. 12, 2012. doi: 10.1155/2012/163731.

[6] V. V. Kalinin and D. A. Polyanskiy. "Gender differences in risk factors of suicidal behaviour in epilepsy", *Epilepsy and Behavior*, Vol. 6, No. 3, 2005, pp.424-429.

[7] Mind Heart Reader, An experiment with EEG signals, http://jannah.github.io/Mind-Heart-Reader/ (last accessed on 03.06.2020)

[8] S. Murugan and V. Ramachandran, "Rough Sets Based Trustworthiness Evaluation of Web Services", *International Review on Computers and Software*, Vol. 7, No. 6, 2012, pp. 2974-2980.

[9] Z. Pawlak, "Rough Sets", International Journal of Computer and Information Sciences, Vol. 11, 1982, pp. 341–356.

[10] Z. Pawlak, "Rough Set. Approach to Knowledge-Based Decision Support". *European Journal of Operational Research*, 1997, pp.48-57.

[11] Z. Pawlak. "Rough Set Theory and its Applications", *Journal of Telecommunications and Information Technology*, Vol 3, 2002, pp.7-10.

[12] J. P. Herbert and J. T. Yao. "Rough Set Model Selection for Practical Decision Making", in *Proceedings of the Fourth International Conference on Fuzzy Systems and Knowledge Discovery, China*, 2007, pp. 203-207. DOI: 10.1109/FSKD.2007.500.

[13] Y. Liu, L. Martínez and K. Qin, "A Comparative Study of Some Soft Rough Sets", *Symmetry*, Vol. 9, No. 11, 2017, 252.

[14] W. Zhao and Z. Zhang. "An Email Classification Model Based on Rough Set Theory", in *Proceedings of the International Conference on Active Media Technology*, 2005, pp.403-408.

[15] B. Walczak and D. L. Massart, "Tutorial: Rough Sets Theory. Chemometrics and Intelligent Laboratory Systems", *Chemometrics and Intelligent Laboratory Systems*,Vol. 47, No. 1, 1999, pp. 1–16.

[16] S. Murugan and V. Ramachandran, "Byzantine Fault Tolerance in SOAP Communication Services", Malaysian Journal of Computer Science, Vol. 25 No. 2, 2012, pp. 67-75, doi: 10.22452/mjcs.vol25no2.1

[17] Z. -A. Xue, D. -J. Han, M. -J. Lv and M. Zhang, "Novel Three-Way Decisions Models with Multi-Granulation Rough Intuitionistic Fuzzy Sets", *Symmetry*, Vol. 10, No. 11, 2018, 662, doi: 10.3390/sym10110662

[18] J. Wang and X. Zhang. "Two Types of Intuitionistic Fuzzy Covering Rough Sets and an Application to Multiple Criteria Group Decision Making", *Symmetry*, 10, 2018, 462.

[19] Y. Liu, L. Martínez and K. Qin, "A Comparative Study of Some Soft Rough Sets", *Symmetry*, 9, 2017, 252.

[20] Y. Qian, X. Liang, Q. Wang, J. Liang, B. Liu, A. Skowron, Y. Yao, J. Ma and C. Dang. "Local rough set: A solution to rough data analysis in big data", *International Journal of Approximate Reasoning*, Vol. 97, 2018, pp 38-63, doi: 10.1016 /j.ijar.2018.01.008.

[21] J. Zhai, X. Wang, S. Zhang and S. Hou, "Tolerance rough fuzzy decision tree", *Information Sciences*," Vol 465, pp 425-438. 2018, doi: 10.1016/j.ins.2018.07.006.

[22] X. Fan, W. Zhao, C. Wang and Y. Huang. "Attribute reduction based on max-decision neighborhood rough set model", *Knowledge-Based Systems*, Vol. 151, 2018, pp. 16-23, doi: 10.1016/j.knosys.2018.03.015.

[23] C. Wang, Y. Huang, M. Shao and X. Fan, "Fuzzy rough set-based attribute reduction using distance measures". *Knowledge-Based Systems*, Vol 164, 2019, pp 205-212, doi: 10.1016/j.knosys.2018.10.038.

[24] B. Walczak and D. L. Massart, "Tutorial: Rough Sets Theory", *Chemometrics and Intelligent Laboratory Systems*, Vol 47, No. 1, 1999, pp. 1–16.

[25] B. Bharathi and M. D. Kamalesh. "Predicting the risk of diabetes mellitus to subpopulations using association rule mining", in *Proceedings of the International Conference on Soft Computing Systems, Advances in Intelligent Systems and Computing*, Vol.397, 2016, pp.59-665.

[26] N. Kishk, H. Mourad, S. Ibrahim, R. Shamloul, A. Al-Azizi and N. Shalaby, "Sex differences among epileptic patients: a comparison of epilepsy and its impacts on demographic features, clinical characteristics, and management patterns in a tertiary care hospital in Egypt", *Egypt J Neurol Psychiatry Neurosurg*, 55, 39 (2019), doi: 10.1186/s41983-019-0078-7.

[27] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo and J. Xiang (2018), "Epileptic Seizure Detection Based on EEG Signals and CNN", *Frontiers in Neuroinformatics*, 2018 Dec 10; 12; 95, doi: 10.3389/fninf.2018.00095.

[28] A. e. Hassanien, M. E. Abdelhafez and H. S. Own, "Rough Sets Data Analysis in Knowledge Discovery: A Case of Kuwaiti Diabetic Children Patients", *Advances in Fuzzy Systems*, Vol. 2008, pp. 1-13, 2008, doi: 10.1155/2008/528461.

[29] T. v. Merode, M. C. d. Krom and J. A. Knottnerus, "Gender-related differences in non-epileptic attacks: a study of patients' cases in the literature", *Seizure*, Vol. 6, No. 4, 1997, pp. 311-316, doi: 10.1016/s1059-1311(97)80079-9.

[30] P. Bizopoulos, G. I. Lambrou, D. Koutsouris, "Signal2image modules in deep neural networks for eeg classification" in *Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, 23–27 July 2019; pp. 702–705.

[31] D. Ahmedt-Aristizabal, C. Fookes, K. Nguyen, S. Denman, S. Sridharan and S. Dionisio, "Deep facial analysis: A new phase I epilepsy evaluation using computer vision," *Epilepsy Behav.*, 2018, 82, 17–24.

[32] A. Antoniades, L. Spyrou, C. C. Took and S. Sanei, "Deep learning for epileptic intracranial EEG data", in *Proceedings of the 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, Vietri sul Mare, Italy, 13–16 September 2016; pp. 1–6.

[33] C. Park, G. Choi, J. Kim, S. Kim, T. -J. Kim, K. Min, K. -Y. Jung and J. Chong, "Epileptic seizure detection for multi-channel EEG with deep convolutional neural network" in *Proceedings of the 2018 International Conference on Electronics, Information, and Communication (ICEIC), Honolulu, HI, USA*, 24–27 January 2018; pp. 1–5.

[34] S. Roy, I. Kiral-Kornek and S. Harrer, "Deep learning enabled automatic abnormal EEG identification", in Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 2756–2759.