



BIBLIOTECA CENTRAL



UNIVERSIDAD DE GUADALAJARA

CENTRO UNIVERSITARIO DE CIENCIAS BIOLÓGICAS Y AGROPECUARIAS
DIVISION DE CIENCIAS BIOLÓGICAS
DEPARTAMENTO DE CIENCIAS AMBIENTALES
INSTITUTO DE NEUROCIENCIAS

**ELABORACION DE UN SISTEMA EXPERTO DE PREDICCIÓN:
MODELO APLICADO A LA ELECTROFISIOLOGIA DE LA ESQUIZOFRENIA.**

TESIS PROFESIONAL
que para obtener el grado de:

**MAESTRO EN CIENCIAS DEL COMPORTAMIENTO:
OPCION NEUROCIENCIAS**

PRESENTA:

HUMBERTO MADERA CARRILLO

DIRECTOR: DR. MIGUEL ANGEL GUEVARA PEREZ

COMITE: DRA. JULIETA RAMOS LOYO
DRA. MARISELA HERNANDEZ GONZALEZ
DR. EMILIO GUMA DIAZ
DR. ANDRES GONZALEZ GARRIDO

ASESOR EN PROGRAMACION: LIC. DANIEL ZARABOZO E. DE R.

GUADALAJARA, JALISCO.

JULIO DEL 2000

Este trabajo fue realizado en los laboratorios de Psicofisiología y de Análisis Automatizado de Señales Bioeléctricas, del Instituto de Neurociencias de la Universidad de Guadalajara.

Proyecto parcialmente apoyado por CONACYT, clave 4214P-H

A esa maravillosa amiga-esposa, y amante-compañera, **Pilar** de comprensión, ternura, apoyo y consejo. Mi más invaluable cercanía en el tortuoso camino del aprendizaje... y de la vida.

A mi **padre**, el hombre a quien más admiro en este mundo y a quien debo todo lo mejor, de mi mejor formación, las bases de mis mejores ideas y los fundamentos de mi visión del mundo.

A mi **madre y mis hermanos**, quienes dispusieron el rico entorno estimulativo en el que me desarrollé y del que me nutrí. A **Gonzalo** por su compañía.

A **Miguel Angel Guevara Pérez** mi tutor y director de tesis, a quien considero un excelente ingeniero-neurocientífico con una extraña pedagogía que le permite ser al mismo tiempo jefe, maestro, compañero y amigo. Gracias por todo.

A **Julieta Ramos Loyo**, quien me compartió una vasta cantidad de información y me co-dirigió haciendo justas indicaciones, siempre con base en conocimientos de causa, sin dejar un momento sus agradables habilidades diplomáticas.

A **Marisela Hernández González** y su disciplina, quien aportó en cada momento de mi formación el concepto de autoexigencia y perfección.

A **Luis Fco. Cerdán Sánchez** por su tiempo dedicado, siempre atinado en su consejo, formal y respetuoso trato, y amistoso apoyo.

A **Daniel Zarabozo**, quien aportó con sus asesorías en programación, ideas y propuestas de importancia capital para el diseño y funcionalidad del Sistema, y quien además me distinguió con su amistad y su consejo.

A **Emilio Gumá Díaz**, por sus acertadas indicaciones, sus comentarios propositivos y su siempre rápida y solidaria respuesta a mis no siempre oportunos requerimientos académicos.

A **Andrés González Garrido**, por sus señalamientos justos y directos que siempre fueron expresados en compañía de propuestas de solución adecuadas.

A mis entrañables, muy guapas amigas y compañeras de laboratorio, cuya siempre presente belleza contribuyó agradablemente al retraso en la culminación de este proyecto, pero cuyo trato hizo inmensamente grato y enriquecedor mi trabajo:

Claudia Amezcua Gutierrez, por compartirme su alegre visión de la vida y estimular cada día mi capacidad de resolución de problemas informáticos.

Irma Vargas Medina, por aportar a nuestra amistad su sinceridad, sensatez, confiabilidad y cordura.

Araceli Sanz Martín, a quien siempre recordaré, por su incansable reto en el debate de ideas y por lograr cambiar mi concepción del mundo por un color y enfoque diferentes.

A todos mis maestros y compañeros en el Instituto de Neurociencias, principalmente:

Al **Dr. Víctor Manuel Alcaraz**, por su discreto y constante apoyo más allá de lo académico.

A **Esmeralda Matute**, por su preocupación, su intercesión y su constante negociación a mi favor.

A **Sergio Meneses Ortega**, por influir tan favorablemente en mi concepto de las neurociencias con su estilo increíblemente claro de presentar sus clases.

A **Jorge Juárez González**, por inquirir sobre la mejor manera de realizar una aplicación práctica del Sistema y porque sus comentarios me propiciaron profundas reflexiones.

A **Eliana Barrios De Tomasi**, por su especial amistad y cariño, por su tiempo dedicado, y por sus charlas llenas de madurez y respeto.

A **Hilda Gpe. Terríquez Casillas**, por compartir su rivalidad futbolera, y su sincera y *carrillosa* amistad.

A **Ma. Antonia Robles**, por saber ser amiga sin importar las circunstancias.

Al Maestro **José Antonio Curiel Zúñiga**, porque sus conceptos y enseñanzas fueron fundamentales para constituir mis propias bases conceptuales con relación a la psicología como ciencia, porque durante sus clases de filosofía y epistemología (en la licenciatura de psicología) fue mi mayor influencia en la interiorización y autoconstrucción de mis métodos de análisis del mundo.

A mis mejores amigos, quienes me han brindado su cariño y su constante compañía en cada uno de los momentos importantes de mi vida:

A **Laura Cristina Navarro Berumen**, mi mejor amiga de todos los tiempos.

A **Gabriela Ramos Girón**, la más hermosa creación hecha mujer.

A **Miriam E. Jiménez Maldonado**, mi camarada de más altos ideales.

A **Miguel A. Estrada**, el cerebro mejor entrenado en el análisis psicológico.

A **Susana Rodríguez Arellano**, mi ángel.

A **Juan Tomas Coffeen Cabrera**, sus enciclopédicas referencias, su extraño humor tipo Bugy el Aceitoso, y al recuerdo de nuestros tiempos difíciles.

CONTENIDO

RESUMEN

I	INTRODUCCION	1
II	ANTECEDENTES	3
II.1	INFORMATICA	3
	1 Conceptos	3
	2 Areas de aplicación y división	6
	3 Los Sistemas Expertos	9
	3a) Generalidades de los Sistemas Expertos.	9
	3b) El Algoritmo de Ganancia de Informacion	11
II.2	ESQUIZOFRENIA	14
	1 Aspectos generales de la esquizofrenia	14
	2 Clasificación de las esquizofrenias	17
	3 La Neurolepto-Resistencia.	19
	4 El Wisconsin Card Sorting Test y la Esquizofrénia	22
	5 La esquizofrenia y otras pruebas	24
	6 La dimensionalidad	25
II.3	LOS DATOS ELECTROENCEFALOGRAFICOS (EEG)	26
	1 El EEG	26
	2 Características del EEG	27
	2a) 1. Banda delta (δ)	28
	2b) 2. Banda theta (θ)	28
	2c) 3. Banda alfa (α)	28
	2d) 4. Banda beta (β)	29
	2e) 5. Banda Gamma (γ)	29
	3 Análisis del EEG	30
	3a) Transformada Rapida de Fourier (TRF).	30
	3b) Analisis de Coherencia.	31
	3c) Análisis de Correlación:	31
	4 Otros metodos de análisis del EEG	32
	4a) Analisis de Componentes Principales	32
	4b) Regresión Lineal Multiple.	33
	4c) Analisis de Conglomerados	33
	4d) Analisis Discriminante	35

5	Inteligencia Artificial y EEG	35
II.4	EL PROBLEMA DE LA PREDICCIÓN	36
1	El Análisis Discriminante	37
2	La Regresión Lineal	38
III	PLANTEAMIENTO DEL PROBLEMA	39
IV	OBJETIVOS GENERAL Y ESPECÍFICOS	40
IV.1	OBJETIVO GENERAL	40
IV.2	OBJETIVOS ESPECÍFICOS	40
V	HIPÓTESIS	41
VI	METODOLOGÍA	42
VI.1	INSTRUMENTO	42
1	Proceso de construcción del Sistema Experto ESQUIZOR	42
2	Lenguaje elegido: Java	45
3	Elaboración del modelo.	46
4	La Interfaz Gráfica de Usuario (GUI) del sistema	49
4a)	Esquema General	49
4b)	La Barra de Menu	51
4c)	El Panel de Iconos	54
4d)	El Panel Principal	55
VI.2	LOS SUJETOS	58
VI.3	LOS DATOS	60
1	Pruebas Psicopatológicas	60
2	Prueba Wisconsin	61
3	Electroencefalografía (EEG)	61
VI.4	DEFINICIÓN Y NÚMERO DE VARIABLES	62
1	La conformación y la forma de las matrices de datos	63
VI.5	LAS PRUEBAS (GENERALIDADES)	66
1	Conformación de los grupos de Capacitación y de Prueba	66
2	Pruebas generales de aprendizaje: estimado de Error Aparente	67
3	El estimado del Error Verdadero y error por grupo.	68

VI.6	PRUEBA 1: DESEMPEÑO BASICO CON VARIABLES DISCRETAS.	70
1	Objetivo	70
2	Hipótesis	70
3	Metodología	70
3a)	Datos	70
3b)	Procedimiento	71
4	Resultados	73
5	Discusión y Conclusiones	74
VI.7	PRUEBA 2: METODOS DE DISCRETIZACION PARA DATOS DE EEG DIGITALIZADO.	75
1	Objetivo	75
2	Objetivos Específicos	75
3	Hipótesis	75
4	Metodología	75
4a)	Datos.	75
4b)	Procedimiento.	76
5	Resultados	78
6	Discusión y Conclusiones	79
VI.8	PRUEBA 3: PRUEBAS FINALES DE FUNCIONAMIENTO DEL SISTEMA.	81
1	Objetivo	81
2	Hipótesis	81
3	Metodología	81
3a)	Sujetos	81
3b)	Datos	81
3c)	Procedimiento	82
4	Resultados	83
4a)	Del proceso de construcción del Sistema	83
4b)	Del proceso de aprendizaje del Sistema	84
4c)	Del proceso de prueba del Sistema	84
5	Discusión y Conclusiones	98
VI.9	ANALISIS DE REGRESION Y COMPARACION CON EL SISTEMA EXPERTO	100
1	Objetivo	100
2	Objetivos Especificos	100
3	Hipótesis	100

4 Metodología	100
4a) Sujetos	100
4b) Datos	101
4c) Procedimiento	101
5 Resultados	104
5a) De la matriz de Wisconsin	104
5b) Matriz de PA por Frecuencia con ojos cerrados.	105
5c) Matriz de PR por Frecuencia con ojos abiertos.	106
6 Discusión y Conclusiones	106
VII RESUMEN DE RESULTADOS DE LAS PRUEBAS	108
VII.1 PROCESO DE CONSTRUCCION DEL SISTEMA	109
VII.2 PROCESO DE APRENDIZAJE DEL SISTEMA	109
VII.3 PROCESO DE PRUEBA DEL SISTEMA	110
VIII DISCUSION Y CONCLUSIONES	111
VIII.1 EL INSTRUMENTO	111
VIII.2 LOS DATOS	112
VIII.3 LAS PRUEBAS	113
1 Con las matrices de datos psicopatológicos y neuropsicológicos.	113
2 Con la matriz de datos de EEG	114
3 En la comparación del Sistema con el método de regresión lineal.	115
4 Ventajas y Desventajas.	116
IX BIBLIOGRAFIA	119
X ANEXOS	123
X.1 RESULTADOS DEL PROCESO DE PRUEBA DEL SISTEMA	123
X.2 INDICE DE FIGURAS	132
X.3 INDICE DE TABLAS	135

RESUMEN

La informática es considerada la más poderosa herramienta para manejar información del siglo XX, y de ella, la Inteligencia Artificial provee el mayor soporte para las decisiones humanas jamás soñado. Los sistemas expertos, como iniciadores y principales exponentes de los Sistemas de Apoyo en la toma de decisiones han sido diseñados para analizar un gran cúmulo de información y producir la respuesta más cercana al resultado del análisis realizado por un experto humano, y sus aplicaciones abarcan todas las áreas del conocimiento humano.

En este trabajo se probó la eficiencia del Sistema Experto llamado **ESQUIZOR**, elaborado expresamente para abordar un problema existente en el área psiquiátrica: la correcta y temprana clasificación de sujetos esquizofrénicos paranoides en dos grupos; los que no responden a los neurolépticos típicos, también denominados neurolepto-resistentes (RE) y los que sí lo hacen, llamados No Resistentes (NR). Dicho sistema realiza su aprendizaje y clasificación con base en datos electroencefalográficos (EEG).

La esquizofrenia neurolepto-resistente se caracteriza por la falta de remisión de los síntomas a pesar de la medicación con diferentes neurolépticos típicos y puede requerir de varios meses o años para su diagnóstico correcto. En un estudio previo se encontraron diferencias en el electroencefalograma (EEG) de estos pacientes resistentes en relación con los que sí responden (Ramos, y col, en preparación). Con base en ello, y por las ventajas que representaría una herramienta capaz de identificar tempranamente si un paciente es resistente a los neurolépticos.

En el estudio participaron 28 sujetos divididos en tres grupos: 9 esquizofrénicos Resistentes (grupo RE), 10 esquizofrénicos No Resistentes (grupo NR) y 9 sujetos normales (grupo CO), a su vez, los dos grupos de esquizofrénicos conformaron un cuarto grupo denominado EZ.

De cada sujeto se registro su electroencefalograma EEG por medio de 12 electrodos colocados conforme al sistema 10/20 y se obtuvieron sus valores de Potencias (Absoluta y Relativa, por bandas y frecuencias) y Correlaciones (Inter e Intrahemisféricas, por bandas y frecuencias) para cada derivación y en dos condiciones; ojos abiertos y cerrados.

También se obtuvieron los resultados de la aplicación del Wisconsin Card Sorting Test y de un conjunto denominado BPRS-PANSS formado por las pruebas Brief Psychiatric Rating Scale (BPRS) y Positive And Negative Síndrome Scale (PANSS).

Las matrices de datos que se conformaron contenían la anterior información, pero sólo dos de los grupos a la vez. Por lo tanto se conformaron 18 matrices de datos (16 de EEG, 1 de Wisconsin y 1 BPRS-PANSS), par cada una de las siguientes combinaciones; RE y NR, CO y RE, CO y NR, CO y EZ. Con estos datos fue alimentado el Sistema Experto.

Dicho sistema tomó cada matriz por separado para generar con ella un árbol de decisión por aprendizaje de ejemplos, utiliza la U de Mann Whitney para hacer una preselección de las variables con diferencias significativas entre grupos, y la fórmula de la Ganancia de Información para calcular la capacidad de discriminación de cada variable proporcionada, y construye el árbol que proporciona las reglas de clasificación para los datos de nuevos sujetos.

Las pruebas finales del sistema se realizaron bajo la técnica de la intervalación, misma que consiste en extraer de la matriz los datos de un sujeto, alimentar al sistema con el resto de los datos para su aprendizaje, y una vez realizado esto presentar al sujeto extraído para su clasificación, procedimiento que se repite con cada sujeto en la matriz.

La efectividad del sistema se midió con base en el estimado de error en la clasificación de los sujetos en cada grupo y el estimado de error total, también llamado error verdadero, mismo que nos indica la cantidad de fallas en el contexto del total de pruebas realizadas.

Para considerar exitosa la prueba, el sistema debía obtener un número de clasificaciones correctas superior al nivel de incidencia y de azar en cada subtipo de esquizofrenia, cantidades que juntas corresponden al (25 y 75 %, la incidencia y 50%, 50% el azar en cada subgrupo: RE y NR respectivamente), lo que permitió establecer los umbrales máximos de error para el sistema en: 50% de falla en la clasificación de sujetos RE y 25% en la de NR y con un error verdadero inferior al 37.5%.

Las matrices de potencias absoluta con ojos cerrados y relativa con ojos abiertos (ambas por frecuencia) lograron un estimado de error verdadero del 21% e identificaron dos variables que alcanzan la mayor discriminación: 10 Hz en C3 y 16 Hz en F4, respectivamente, la primera presenta diferencias grupales altamente significativas ($p= 0.0006$) y la segunda, una elevada puntuación de Ganancia ($G= 0.8$) mostrando una subdivisión del grupo de RE. En forma general los datos por frecuencia parecen proporcionar mayor información que las bandas y los datos de la Potencia Relativa son más estables para el sistema que los de la Absoluta. El sistema encontró en el EEG información suficiente para la clasificación de los sujetos RE y NR.

Esquizor podría utilizarse para la clasificación de otros tipos de grupos con base en su EEG.

I INTRODUCCION

Actualmente, una de las áreas de mayor interés de la informática es la Inteligencia Artificial, y dentro de esta, la construcción y aplicación de los sistemas expertos es la que ha redituado mayores dividendos en áreas tan diversas como la minería, la química, la medicina y la psiquiatría.

Los sistemas expertos informáticos representan no sólo la oportunidad de almacenar y replicar el conocimiento de los expertos humanos y utilizarlo como si se tuviese al experto presente, sino de que los Sistemas generen su base de conocimientos y reglas sin la intervención del experto humano.

Las técnicas informáticas de inteligencia artificial actuales, permiten programar sistemas capaces de aprender de ejemplos presentados y con ello elaborar sin supervisión, una serie de reglas que les permiten clasificar los datos de un sujeto desconocido.

El modelo escogido para este trabajo fue el de *Sistemas Expertos basados en la generación de árboles de decisión por aprendizaje de ejemplos*, que evalúa a capacidad de cada variable para diferenciar los sujetos de dos grupos distintos.

Este modelo se apoya en un algoritmo que mide la ganancia de información de cada variable y con ello hace una selección de las variables que mejor discriminan a los grupos y su relación entre ellas.

El programa genera los valores de cada variable de manera que se puede conformar un perfil de las relaciones entre las más importantes de ellas y la clasificación final de cada sujeto que conforma el Conjunto de Capacitación (o de ejemplos

El problema enfrentado pertenece al área medico-psiquiátrica: la clasificación de sujetos esquizofrénicos en dos subgrupos, los que responden a los neurolépticos típicos y los que no lo hacen.

La esquizofrenia es una enfermedad psicótica crónica que por sus características específicas se considera altamente inutilizante y en muchas ocasiones atormentadora para el sujeto que la sufre. Le impide su desarrollo laboral y social ya que se presenta típicamente entre la segunda y la cuarta década de la vida.

Generalmente y con una medicación adecuada, los síntomas presentan una remisión importante dentro de las primeras 6-12 semanas.

Sin embargo, se pueden encontrar entre un 10 a 25% de pacientes que no responden a los tratamientos tradicionales y que por ende, su pronóstico es de mala evolución y deteriorante. Los estudios al respecto de esta variante de la enfermedad son escasos y el retraso en el diagnóstico correcto es común.

Actualmente, la única forma de diagnosticar a este tipo de pacientes resistentes al tratamiento (denominados Neurolepto-Resistentes, RE) es a través de un largo proceso de medicación sin observar una disminución mayor ni en la cantidad ni en la calidad de los síntomas padecidos. Esto representa un enorme gasto económico, familiar y social por cada paciente. No existe un sistema o una metodología capaz de identificar por medio del EEG a un paciente esquizofrénico en Resistente y No Resistente (a la medicación típica) de manera que se detecte esta característica antes de hacerlo de forma clínica. Tampoco existe una herramienta que pueda hacerlo de forma semiautomatizada. Sin embargo, se han encontrado diferencias en algunos parámetros electroencefalográficos (EEG) en esquizofrénicos y sujetos normales, y aunque existen resultados contradictorios por la falta de control de algunas variables, entre ellas la edad, sexo, lateralidad, tipo de tratamiento, edad de inicio de la enfermedad, el nivel educativo y los subtipos de esquizofrenia, los resultados de un estudio previo proporcionan información de que existen diferencias en el EEG de pacientes Resistentes, en comparación con los No Resistentes y los sujetos normales. Esto permite suponer que es posible utilizar los parámetros de la actividad electroencefalográfica para clasificar a este tipo de sujetos de manera temprana.

II ANTECEDENTES

II.1 INFORMATICA

1 Conceptos

La inteligencia es un concepto que ha ocupado al ser humano desde tiempos antiguos, la capacidad de aprender y de tomar decisiones correctas se ha interpretado como reflejo de una inteligencia adecuada y funcional, sin embargo, ¿es ésta una capacidad reproducible artificialmente? Actualmente, aún se busca una parte de la respuesta, otra se ha comenzado a contestar, lentamente.

La Inteligencia Artificial (IA) es una disciplina relativamente nueva, se inicia en 1956 al acuñarse el término (propuesto por John McCarthy durante un taller de investigadores interesados en teoría de autómatas, redes neuronales y estudio de inteligencia, organizado en Dartmouth en el verano de ese año), y se facilita su desarrollo con la aparición de las computadoras a principios de la década de los cincuenta con lo que se posibilitó el abordaje experimental, la modelación y la prueba de teorías al respecto.

La IA es tanto un concepto como un área y un campo de aplicación. Dentro del concepto de inteligencia, se pueden identificar dos tendencias: por un lado, la aplicación de la palabra en el ser humano y por otro en las máquinas, en el primer caso, la definición se mueve desde una sencilla acción de comprender... hasta la identificación de las capacidades necesarias para dicha comprensión y aprendizaje sin embargo, en el segundo caso se refiere únicamente, a la capacidad de procesamiento (Freedman, 1994). Esto quiere decir, que no hay mentira en la afirmación de que todas las computadoras son inteligentes, no obstante, la terminología cambia cuando se habla de IA en su conjunto, ya que en términos generales esto implica un tipo de inteligencia similar a la humana... y nuevamente ello nos devuelve al problema de definir lo que es la inteligencia humana y la forma de aplicar dicho concepto a las máquinas.

Algunos conceptos de lo que es la Inteligencia Artificial son los siguientes:

[La automatización de] actividades que vinculamos con procesos de pensamiento humano, actividades como la toma de decisiones, resolución de problemas, aprendizaje.... (Bellman, 1978).

El arte de crear máquinas con capacidad de realizar funciones que realizadas por personas requieren de inteligencia. (Kurzweil, 1990).

El estudio de las facultades mentales mediante el uso de modelos computacionales. (Charniak y McDemott, 1985).

Un campo de estudio que se enfoca a la explicación y emulación de la conducta inteligente en función de procesos computacionales. (Shalkoff, 1990).

Sin embargo, las definiciones de lo que es la IA parecen converger en un punto: el ser humano es el modelo de la IA, debido a que posee la plantilla denominada inteligencia humana. La cuestión se facilita si consideramos al ser humano más como un sistema que como una entidad, de esta forma podremos hablar de él como un sistema inteligente y considerar a un sistema no humano como poseedor de IA si se comporta como un sistema inteligente. Esto divide las definiciones de la IA en cuatro grandes grupos a saber (Russell, y Norvig, 1996):

1. Sistemas que *piensan* como humanos.
2. Sistemas que *actúan* como humanos.
3. Sistemas que *piensan* racionalmente.
4. Sistemas que *actúan* racionalmente.

Considerando para este caso a la racionalidad como un concepto de inteligencia ideal (también bajo un molde conceptual humano). Estas divisiones, a su vez ponen de manifiesto otras subdivisiones: los sistemas que son considerados inteligentes en virtud de su conducta y los que lo son en virtud de su proceso interno o pseudo-pensamiento, además unos son juzgados en su función en comparación con la eficiencia humana mientras que otros lo son contra un ideal.

Lo importante, al mismo tiempo de formular las definiciones de lo que es la inteligencia artificial es definir aquello que requiere hacer la máquina para considerarla

inteligente, las pruebas que debe pasar. Estas pruebas al igual que la definición, se basaron en el modelo humano de percepción, análisis y acción: para el caso de los sistemas que piensan como humanos, no sólo importa la resolución correcta de los problemas, en ocasiones son más importantes los pasos seguidos y su comparación con el proceso de razonamiento humano para encontrar sus similitudes; en el caso de los sistemas que actúan como humanos se ha exigido que la máquina emita respuestas tan parecidas a los humanos que un examinador, en interacción no visual simultánea con ella y con otro humano, se vea imposibilitado para definir correctamente cual de las respuestas recibidas pertenece a la máquina, ésta es la llamada prueba de Turing (Russell y Norvig, 1996); para los sistemas que piensan racionalmente se considera que deben de estar contruidos sobre la base de la lógica formal y que con ella representen aseveraciones relacionadas con los objetos del mundo y sus relaciones mutuas; y por último, actuar racionalmente implica actuar de manera tal, que se logren los objetivos deseados con base en ciertos supuestos. Es en este punto en el que surge la definición de agente, que es algo capaz de percibir y actuar. Desde este enfoque, se considera a la inteligencia artificial como el estudio y la construcción de agentes racionales.

En la mayoría de las definiciones de IA se hace necesario, igual que para el de inteligencia humana, el factor aprendizaje, ya que en los cuatro grupos propuestos es preciso que el sistema adquiriera nuevos conocimientos o información y que en virtud de ello se adapte con éxito a los cambios en el medio en que se encuentre. Ahora bien, el aprendizaje puede definirse como lo hizo Simon en 1983:

... cambios en el sistema que se adaptan en el sentido de que le permiten llevar a cabo la misma tarea a partir de las mismas condiciones de un modo más eficiente y eficaz cada vez.

Sin embargo, en un sentido más amplio el término inteligencia artificial definiría una propiedad de un sistema que realiza los mismos tipos de funciones (siendo estas, procesamiento o acciones) que se pueden juzgar como característicos del pensamiento y/o accionar humano.

2 Areas de aplicación y división

Con el avance de la ciencia moderna la búsqueda de la IA ha tomado dos caminos fundamentales: por un lado, la investigación psicológica y fisiológica de la naturaleza del pensamiento humano, y por el otro, el desarrollo tecnológico de sistemas informáticos cada vez más complejos. Lo anterior debido a que la investigación, el desarrollo y la aplicación de la IA se encuentra en una zona híbrida.

De esta manera, y considerando lo que cada una de estas áreas tiene como objetivo, la IA puede dividirse en campos (Winston, 1994):

El primero, definido por la ingeniería: determinado a resolver problemas reales actuando como un armamento de ideas acerca de cómo representar y utilizar el conocimiento, y de cómo ensamblar sistemas.

El segundo, definido por las ciencias: determinado a conocer ideas acerca de la representación del conocimiento, del uso que se da a este, y del ensamble de sistemas que explican diversas clases de inteligencia.

En la Figura 1, Rich, y Knight, (1994) dividen en tres las principales áreas de aplicación de la IA.

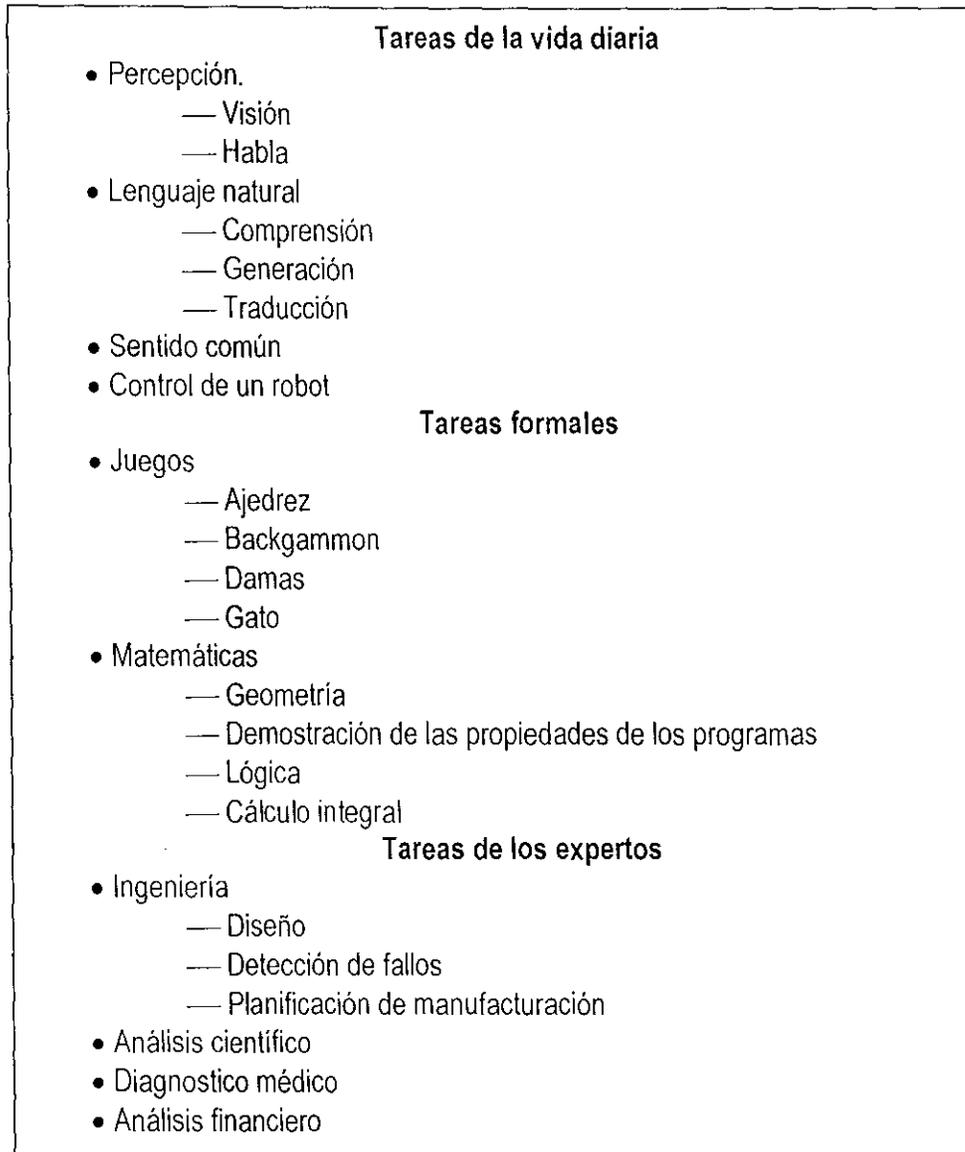


Figura 1. Algunas aplicaciones de la Inteligencia Artificial.

Por otro lado, considerando los tipos de programas que se han desarrollado dentro del universo de la IA, podemos agrupar su tecnología y aplicaciones dentro de tres principales paradigmas (Abud, 1998):

1. Inteligencia artificial basada en comportamiento: es el más nuevo de todos, se inició hace menos de 15 años y su técnica principal la hace adaptarse muy bien a problemas residentes en medios específicos y delimitados (por

ejemplo un agente virtual inteligente realizando tareas específicas dentro de Internet), ya que todas las representaciones del mundo formal se encuentran asociadas por medio de codificaciones y decodificaciones correspondientes al mundo real.

2. Inteligencia artificial basada en conexionismo: Conocida como Redes Neuronales, este paradigma abarca tanto a la tecnología de redes neuronales como a los autómatas celulares y a los algoritmos genéticos, se encuentra constituido por una gran cantidad de elementos sencillos, llamados neuronas, altamente interconectadas, con múltiples entradas y una sola salida. Son sistemas computacionales que imitan la estructura del cerebro para procesar información. Esta tecnología trabaja bien en situaciones en las que se carece o se desconoce el método para llegar a la solución del problema ya que únicamente se entrena a una red con un conjunto de ejemplos hasta que “aprende” a solucionar un problema. En este caso no es necesario conocer el método para llegar a la solución, únicamente la solución misma. Se aplica con éxito en problemas de reconocimiento de patrones, de estimaciones y proyecciones y en problemas de asignación de recursos.
3. Inteligencia artificial basada en símbolos: es llamada IA. Clásica por ser el primer paradigma y se conforma de una serie de patrones y procesos, estipula que la inteligencia reside en el conocimiento (símbolos) y en la forma apropiada de representar y manipular ese conocimiento. De esta tecnología se derivan los sistemas expertos, la propagación de restricciones, el razonamiento matemático, la lógica difusa y las búsquedas heurísticas, entre otros. Se aplica a problemas en los que el espacio de búsqueda de solución es muy grande y complejo y, donde sí se conoce el método para llegar a la solución. Las áreas donde se ha implementado con éxito son: labores de diagnóstico complejos, asignación de recursos en ambientes complicados, programación de televisoras y en entretenimientos como el ajedrez.

3 Los Sistemas Expertos

3a) GENERALIDADES DE LOS SISTEMAS EXPERTOS.

Un sistema experto es una aplicación de inteligencia artificial que usa una base de conocimiento de la experiencia humana para ayudar a la resolución de problemas reales. El grado de resolución de problemas está basado en la calidad de los datos y reglas del experto humano.

También se puede definir a un sistema experto conforme a la propuesta de Weiss y Kulikowski, (1991) quienes consideran que un sistema de este tipo es aquel que maneja los problemas complejos del mundo real requiriendo una interpretación de experto, o que resuelve esos problemas usando un modelo computarizado de razonamiento humano experto, llegando a las mismas conclusiones a las que el experto humano llegaría si se encontrara con un problema comparable.

El sistema experto deriva sus respuestas pasando la base de conocimientos a través de un motor de inferencia, un programa que interactúa con el usuario y procesa los resultados de las reglas y los datos de la base de conocimientos.

Los sistemas expertos deben contener como base reglas, marcos, sistemas de mantenimiento de la verdad y otros muchos mecanismos de razonamiento.

Sin embargo, hay dos cosas que se requiere que un sistema experto haga aparte de lo anterior, explicar sus razonamientos y adquirir nuevos conocimientos, es decir, se requiere que sea capaz de aprender, bajo la definición dada con anterioridad. Esto implica un perfeccionamiento de la habilidad y una adquisición de conocimiento.

La necesidad de que los sistemas expliquen la razón de su afirmación está dada porque los humanos al tomar la decisión final, apoyados por la informática, requieren conocer la validez del proceso de razonamiento que llevó a esa conclusión, para a su vez poder decidir con certeza. Por otro lado, la forma en que los sistemas expertos adquieren nuevos conocimientos puede seguir dos caminos; el primero es a través de la interacción con un experto humano, dada por medio de una interfaz adecuada para la formulación de preguntas, respuestas y explicaciones, la otra manera es que el programa aprenda el comportamiento experto a partir de datos en bruto.

Se pueden definir diferentes métodos para que el sistema experto pueda adquirir conocimiento (Rich y Knight, 1994):

El aprendizaje memorístico es la actividad de aprendizaje más básica, consiste en que la computadora almacene una parte de la información. La forma más sencilla, a la que mucha gente no le llamaría aprendizaje es la que se da en las bases de datos, sin embargo, cuando este método se combina con otros y lo que se almacena son los datos producidos por dicha combinación, y si la interacción y el uso de dichos datos optimizan la ejecución, sí se puede hablar de una metodología de aprendizaje.

Otra de las formas de aprendizaje es el aprendizaje a través de consejos, este también puede parecer un método muy rudimentario ya que requiere que el programador escriba una serie de órdenes, sin embargo, se cuenta con un sistema capaz de operacionalizar los conceptos que le son dados en forma de consejos.

Sin embargo, existen otros tipos de aprendizaje que no requieren a un profesor propiamente dicho. Estos programas de aprendizaje se basan en la generalización de la propia experiencia de las máquinas. Denominado como el aprendizaje en la resolución de problemas. Uno de ellos, es el aprendizaje por medio del ajuste de parámetros que consiste en asignarle peso (valor) a los parámetros incluidos en el programa y sobre la base de los cuáles se hará la estimación o predicción, después el programa modificará el valor de dichos parámetros, en virtud de lo cercano o lejano que se encuentren del objetivo, aumentando en el primero de los casos y reduciendo en el segundo, esto es, valdrá más el parámetro más útil para la predicción.

Otro tipo de aprendizaje; El aprendizaje con macro-operadores consiste en un sistema muy parecido al aprendizaje memorístico ya que, el sistema almacena todas las acciones que lo llevaron a conseguir el objetivo, sin embargo, tiene un paso más avanzado, generaliza los pasos para que puedan servir en situaciones que no son completamente iguales a las que acaba de pasar. Esto lo hace cambiando por variables todas las constantes del macro operador almacenado.

Un conjunto diferente de modelos de aprendizaje se refiere al aprendizaje a partir de ejemplos: la inducción, que se utiliza en objetivos de clasificación.

Programas de este tipo son el programa de aprendizaje de Winston (1970) (ver Russell y Norving, 1996), que operaba sobre el sencillo dominio del mundo de los bloques. También el de espacios de versiones de Mitchell (1982), que al igual que el anterior consistía en producir una descripción que fuese consistente con los ejemplos positivos pero no con los negativos.

Un método que se incluye en este tipo de modelo de aprendizaje es el de Árboles de Decisión el cual utiliza una representación de árbol para los conceptos, ya que para clasificar una entrada particular se empieza por la parte superior del árbol y se responden preguntas hasta llegar a una hoja donde se guarda la clasificación. El método se comienza eligiendo un subconjunto aleatorio de ejemplos de entrenamiento, al que se le llama *Ventana* o conjunto de capacitación. El algoritmo clasifica correctamente todos los ejemplos del conjunto de capacitación, el árbol se prueba con los ejemplos que se encuentran fuera de la ventana o conjunto de prueba. Si falla en la clasificación correcta de estos elementos, entonces se escogen nuevos datos para el conjunto de capacitación y se agregan, para probarse nuevamente. Esta metodología también se encuentra basada en la inducción de descripciones de conceptos a partir de ejemplos positivos y negativos. Dentro de este modelo se cuenta con algoritmos que se encargan de ignorar partes del árbol de búsqueda que son irrelevantes para la decisión final, a esto se les denomina 'poda' y por medio de ello se evita que el árbol crezca de manera desproporcionada.

Otros modelos de aprendizaje son: el aprendizaje basado en explicaciones, el descubrimiento y la analogía. Por último, el aprendizaje con redes neuronales y el aprendizaje genético.

3b) EL ALGORITMO DE GANANCIA DE INFORMACION

El esquema utilizado en el aprendizaje bajo el modelo del árbol de decisión para seleccionar atributos está diseñado para reducir a un mínimo la profundidad del árbol final. Lo importante es tomar un atributo que favorezca al máximo la exacta clasificación de los ejemplos. Un atributo perfecto divide los ejemplos en conjuntos que son totalmente positivos o negativos.

Para implantar esta función que escoge el atributo hace falta contar con una medida formal de lo que es "bastante bueno" y lo que es "realmente inútil" El valor

máximo de la tal medida se obtiene cuando el atributo es perfecto, y el valor mínimo cuando el atributo es no sirve en lo absoluto. Una medida adecuada es la cantidad esperada de información que proporciona el atributo.

La teoría de la información el valor de la misma en cada variable se mide en bits. Un bit de información basta para responder a una pregunta de sí o no, sobre la cual no se tiene la más remota idea.

En general si las posibles respuestas v_i tienen probabilidad $P(v_i)$, entonces la cantidad de información I de la respuesta real es obtenida de la siguiente manera:

$$I(P(v_1), \dots, P(v_n)) = \sum_{i=1}^n -P(v_i) \log_2 P(v_i)$$

Lo anterior es precisamente la cantidad de información promedio de los diversos eventos (una cantidad de términos igual a $-\log_2 P$) ponderados mediante las posibilidades de los eventos. Para verificar esta ecuación, veremos que en el lanzamiento de una moneda que no está cargada se obtiene lo siguiente:

$$I\left(\frac{1}{2}, \frac{1}{2}\right) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 1 \text{ bit}$$

En el aprendizaje del árbol de decisiones, la pregunta que necesitamos responder es: en un determinado ejemplo, ¿cuál es la clasificación correcta? Un árbol de decisión correcto responderá la pregunta anterior. La estimación de las probabilidades de las posibles respuestas antes de probar los atributos son obtenidas mediante las proporciones de ejemplos positivos y negativos presentes en el conjunto de capacitación. En el supuesto que en el conjunto de capacitación hay p ejemplos positivos y n negativos. El estimado de la información contenida en la respuesta correcta es:

$$I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) = -\frac{p}{p+n} \log_2 \frac{p}{p+n} - \frac{n}{p+n} \log_2 \frac{n}{p+n}$$

Por lo general, el realizar pruebas en un solo atributo de A no producirá tanta información, pero al menos sí nos dará parte de ésta. Para medir exactamente qué tanta información nos proporcionará basta con determinar qué tanta información necesitamos *después* de la prueba de atributos. Un atributo A divide el conjunto de entrenamiento E en los subconjuntos E_1, \dots, E_v de acuerdo con sus valores respectivos para A , en cuyo caso A tiene valores diferentes. Un subconjunto E_i tiene P_i ejemplos positivos y n_i ejemplos negativos, por lo que si recorremos esa rama necesitaremos $I(P_i/(P_i + n_i), n_i/(P_i + n_i))$ bits más de información para responder la pregunta. Un ejemplo aleatorio tendrá el valor número i para el atributo cuya probabilidad es $(P_i + n_i)/(P + n)$, por lo que en promedio, luego de probar el atributo A , necesitaremos:

$$\text{Restante} = \sum_{i=1}^v \frac{p_i + n_i}{p_i + n_i} I\left(\frac{p_i}{p_i + n_i}, \frac{n_i}{p_i + n_i}\right)$$

bits de información para clasificar el ejemplo. La ganancia de información deducida de la prueba de atributo se define como la necesidad original de información y la nueva necesidad:

$$\text{Ganancia} = I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) - \text{Restante}$$

y la heurística utilizada en la función que escoge el atributo simplemente elegirá el atributo que tenga la ganancia más grande.

II.2 ESQUIZOFRENIA

1 Aspectos generales de la esquizofrenia

La esquizofrenia es una enfermedad psicótica crónica que ocurre típicamente en los últimos años de la segunda década de la vida y la mitad de la cuarta (Cerdán, 1999).

Las características esenciales de la esquizofrenia son una mezcla de signos y síntomas peculiares (tanto positivos como negativos) que han estado presentes una parte significativa de tiempo durante un periodo de 1 mes (o durante un tiempo más breve si ha habido tratamiento con éxito) y con algunos signos de trastornos que han persistido durante al menos 6 meses.

Estos signos y síntomas están asociados a una marcada disfunción social o laboral (DSM-IV).

Los síntomas característicos se dividen principalmente en dos: los positivos y los negativos. Los síntomas positivos reflejan un exceso o distorsión de las funciones normales, incluyendo distorsiones o exageraciones del pensamiento inferencial (ideas delirantes), de la percepción (alucinaciones), el lenguaje y la comunicación (lenguaje desorganizado) y la organización comportamental (comportamiento gravemente desorganizado o catatónico). Estos síntomas pueden dividirse en dos dimensiones: la psicótica que incluye a las dos primeras disfunciones, y la dimensión de desorganización que conlleva las restantes. Los síntomas negativos reflejan una disminución o una pérdida de las funciones normales y comprenden restricciones del ámbito y la intensidad de la expresión emocional (aplanamiento afectivo), de la fluidez y la productividad del pensamiento y el lenguaje (alogia), y del inicio del comportamiento dirigido a un objetivo (abulia).

Las ideas delirantes, son creencias erróneas que habitualmente implican una mala interpretación de las percepciones o las experiencias, se califican como extrañas si son claramente improbables e incomprensibles y si no derivan de las experiencias de la vida

Las alucinaciones, pueden ocurrir en cualquier modalidad sensorial, sin embargo las auditivas son las más comunes y características en la esquizofrenia. Son percibidas como voces, conocidas o no, pero diferentes a los pensamientos de la persona. Estas alucinaciones deben de producirse en un contexto de claridad sensorial y no durante ensoñaciones.

El lenguaje desorganizado, proporciona una inferencia al proceso de pensamiento del sujeto y se evalúan en varios aspectos como son; la continuidad y el flujo en el tema tratado (pérdida del hilo o descarrilamiento), la relación de la respuesta con la pregunta formulada (tangencialidad) o desorganización (incoherencia).

El comportamiento gravemente desorganizado, puede manifestarse desde las tonterías infantiloides hasta la agitación impredecible en un comportamiento con problemas en cualquier forma de dirección a un fin, ocasionando dificultades en la realización de las actividades de la vida cotidiana, sin embargo, debe diferenciarse del comportamiento simplemente carente de sentido.

El aplanamiento afectivo, se caracteriza por la inmovilidad y la falta de respuestas en la expresión facial del sujeto, contacto visual pobre y reducción del lenguaje corporal.

La alogia, se manifiesta por las réplicas breves, lacónicas y vacías. Sus pensamientos aparentemente disminuidos, se reflejan en un descenso de la fluidez y la productividad del habla.

La abulia, tiene como característica la incapacidad para iniciar y persistir en actividades dirigidas a un fin.

Los criterios del DSM-IV para el trastorno de la esquizofrenia son los siguientes:

Síntomas característicos: Dos (o más) de los siguientes, cada uno de ellos presente durante una parte significativa de un período de 1 mes (o menos si ha sido tratado con éxito):

Ideas delirantes

Alucinaciones.

Lenguaje desorganizado

Comportamiento catatónico o gravemente desorganizado.

Síntomas negativos, por ejemplo, aplanamiento afectivo, alergia o abulia.

Disfunción social/laboral: Durante una parte significativa del tiempo desde el inicio de la alteración, una o más áreas importantes de actividad, como son el trabajo, las relaciones interpersonales o el cuidado de uno mismo, están claramente por debajo del nivel previo al inicio del trastorno (o cuando el inicio en la infancia o adolescencia, fracaso en cuando a alcanzar el nivel esperable de rendimiento interpersonal, académico o laboral).

Duración: Persisten signos continuos de la alteración durante al menos 6 meses. Este periodo de 6 meses debe incluir al menos 1 mes de síntomas que cumplan el Criterio A(o menos si se ha tratado con éxito) y puede incluir los periodos de síntomas prodrómicos y residuales. Durante estos periodos prodrómicos o residuales, los signos de la alteración pueden manifestarse sólo por síntomas negativos o por dos o más síntomas de la lista del Criterio A, presentes de forma atenuada (p.ej., creencias raras, experiencias perceptivas no habituales).

Exclusión de los trastornos esquizoafectivo y del estado de ánimo: El trastorno esquizoafectivo y el trastorno del estado de ánimo con síntomas psicóticos se han descartado debido a: 1) no ha habido ningún episodio depresivo mayor, maníaco o mixto concurrente con los síntomas de la fase activa; o 2) si los episodios de alteración anímica han aparecido durante los síntomas de la fase activa, su duración total ha sido breve en relación con la duración de los periodos activo y residual.

Exclusión de consumo de sustancias y de enfermedad médica: El trastorno no es debido a los efectos fisiológicos directos de alguna sustancia (p.ej., una droga de abuso, un medicamento) o de una enfermedad médica.

Relación con un trastorno generalizado del desarrollo: Si hay historia de trastorno autista o de otro trastorno generalizado del desarrollo, el diagnóstico adicional de esquizofrenia sólo se realizará si las ideas delirantes o las alucinaciones también se mantienen durante al menos 1 mes (o menos si se han tratado con éxito).

Los criterios diagnósticos de la CIE-10 son similares a los del DSM-IV en muchos aspectos, pero no idénticos. Los Criterios Diagnósticos de Investigación de la CIE-10 proponen dos caminos distintos para llegar a cumplir los criterios diagnósticos de la esquizofrenia: La existencia de un síntoma de Schneider de primer rango o delirios extraños, o presencia de al menos dos de los otros síntomas característicos (alucinaciones acompañadas de delirios, trastorno del pensamiento, síntomas catatónicos y síntomas negativos). La definición de la esquizofrenia de la CIE-10 tan sólo exige una duración de los síntomas de 1 mes, abarcando de esta forma las categorías diagnósticas del DSM-IV de esquizofrenia y trastorno esquizofreniforme. La definición de la CIE-10 no requiere afectación de la actividad del individuo.

Existen discrepancias en la prevalencia de esquizofrenia en diferentes estudios porque se han empleado metodologías diferentes (p.ej., medio rural vs urbano, población general vs medio clínico u hospitalario) y diferentes definiciones de esquizofrenia (restrictiva vs amplia, basada en criterios vs clínica). Las estimaciones de la prevalencia en muchos estudios extensos han variado desde el 0.2 al 2.0 %. Las tasas de prevalencia son similares en todo el mundo, pero se han descrito la existencia de bolsas de alta prevalencia en algunas áreas específicas. Actualmente se estima entre el 0.5 y el 1 %, 1995, la tasa de incidencia es considerablemente inferior a la de prevalencia debido a que esta enfermedad tiende a ser crónica y se estima en 1/10,000 al año (DSM-IV).

2 Clasificación de las esquizofrenias

En relación con lo psicopatológico, la esquizofrenia se ha clasificado alrededor de dos ejes temporales (ver en Cerdán, 1999): El longitudinal (curso a lo largo de la enfermedad), y el transversal (las características clínicas presentes en el momento de la evaluación). El curso longitudinal ha sido dividido en:

- Episodio con síntomas residuales ínter episódicos.
- Episódico sin síntomas residuales ínter episódicos.
- Continuo.
- Episódico único con remisión parcial.
- Episódico único con remisión total.
- Otro patrón no especificado.

Con relación al curso transversal de la enfermedad, está definido por la sintomatología dominante en el momento de la evaluación. El DSM-IV hace distinción de 5 subtipos:

Tipo paranoide

1. Preocupación por una o más ideas delirantes o alucinaciones auditivas frecuentes.
2. No hay lenguaje desorganizado, ni comportamiento catatónico o desorganizado, ni afectividad aplanada o inapropiada.

Tipo desorganizado, cumple con los siguientes criterios diagnósticos:

1. Predominan; lenguaje desorganizado, comportamiento desorganizado y afectividad aplanada o inapropiada.
2. No se cumplen los criterios para el tipo catatónico.

Tipo catatónico, un tipo de esquizofrenia en el que el cuadro clínico está dominado por al menos dos de los siguientes síntomas:

1. Inmovilidad motora manifestada por catalepsia (incluida la flexibilidad cérea) o estupor.
2. Actividad motora excesiva (que aparentemente carece de propósito y no está influida por estímulos externos).
3. Negativismo extremo (resistencia aparentemente motivada a todas las órdenes o mantenimiento de una postura rígida en contra de los intentos de ser movido) o mutismo.
4. Peculiaridades del movimiento voluntario manifestadas por la adopción de posturas extrañas (adopción voluntaria de posturas raras o inapropiadas), movimientos estereotipados, manierismos marcados o muecas llamativas.
5. Ecolalia o ecopraxia.

Tipo indiferenciado: un tipo de esquizofrenia en que están presentes los síntomas del Criterio A, pero que no cumple los criterios para el tipo paranoide, desorganizado o catatónico.

Tipo residual, cumple con los siguientes criterios diagnósticos:

1. Ausencia de ideas delirantes, alucinaciones, lenguaje desorganizado y comportamiento catatónico o gravemente desorganizado.

2. Hay manifestaciones continuas de la alteración, como lo indica la presencia de síntomas negativos o de dos o más síntomas de los enumerados en el Criterio A para la esquizofrenia, presentes de una forma atenuada (p.ej., creencias raras, experiencias perceptivas no habituales).

Con relación a una de las primeras clasificaciones basada en el EEG llevada a cabo por John (1994) realizó un análisis de clusters de un grupo de esquizofrénicos (agudos, con y sin medicación, de diferente género, y sujetos esquizofrénicos crónicos, de diferente edad) y encontró 5 diferentes subtipos, en los que afirma, es posible clasificar, desde el punto de vista neurofisiológico cualquier población de individuos esquizofrénicos.

Otra clasificación existente se basa en la respuesta del paciente esquizofrénico a los neurolépticos típicos (haloperidol, decanoato de flufenazina, clorpromazina).

3 La Neurolepto-Resistencia.

El concepto de esquizofrenia resistente al tratamiento lleva implícito un pronóstico de mala evolución y deterioro. Esto es conocido desde las primeras descripciones realizadas por Kraepelin y Bleuler en 1911 (Ver. Cerdán L., 1999), ellos afirmaron que de un 20 a 30% de estos pacientes presentan este carácter evolutivo.

En las dos últimas décadas, algunos estudios del seguimiento del grado de recuperación y funcionamiento psicosocial de sujetos esquizofrénicos llaman la atención en la literatura. Manfred Bleuler reportó en 1978 (Harding, 1992) un estudio realizado en la clínica Burghölzi en Zurich a lo largo de 23 años que incluyó a 208 pacientes admitidos durante 1942 y 1943. Los resultados señalan que un 53 a 58% de los sujetos habían mejorado significativamente. Otros estudios señalan mejorías que van desde el 46 hasta el 68% (Ciompi y Müller, 1976, Harding et al 1987, para ambos ver Cerdán L., 1999) con más de 100 pacientes cada uno.

Los estudios anteriores indican que cerca de un 40% de los pacientes esquizofrénicos tienen una mala evolución.

La precisión conceptualmente rigurosa de esquizofrenia resistente al tratamiento con neurolépticos, se debe a Kane (1988), quien señala los criterios fenomenológicos, psicopatológicos, cuantitativos y evolutivos, que son los siguientes:

- Esquizofrenia diagnosticada con el DSM-III.
- Curso permanente de la enfermedad en los últimos 5 años.
- Sin cambio en la sintomatología durante todo este período de tiempo.
- Al menos tres ensayos con neurolépticos diferentes.
 - Con familias químicas de neurolépticos diferentes.
 - Duración mínima del tratamiento de 6 semanas.
 - Con una posología equivalente a 1 gramo de clorpromazina x 24 hs.
- No reducción de los síntomas.
- Puntuación en la escala de apreciación clínica global, igual o mayor de 4.
- Puntuación igual o mayor a 45 en la escala psiquiátrica breve (BPRS).
- Puntuación de 4 o más, al menos en 2 de los 4 reactivos siguientes (desorganización conceptual, desconfianza, comportamiento alucinatorio o pensamiento extraño y poco habitual).

Otros criterios de neurolepto-resistencia son los de Wilson (1989 en Brenner y Merlo, 1995), quien la define señalando que: debe realizarse el diagnóstico con el DSM-III, persistencia de síntomas psicóticos durante 2.5 años, con los siguientes criterios operativos:

- Puntaje en el BPRS de igual o mayor de 45, con 3 o más de los siguientes reactivos de igual o mayor de 3: aislamiento, trastornos del pensamiento y alucinaciones, ideas delirantes y retardo psicomotor.
- Tres clases de neurolépticos usados durante los últimos años de familia química diferente y a dosis equivalente a 1 gramo de clorpromazina, con duración mínima de 8 semanas.

Los criterios de Schüssler et al (1988, en Brenner y Merlo, 1995), quienes la definen como la ausencia de mejoría después de un tratamiento con neurolépticos durante 4 semanas. Con los siguientes criterios operativos: puntaje de impresión clínica global de 3, reducción de la sintomatología no mayor del 5%.

Los criterios de Keefe et al (1991 en Brenner y Merlo, 1995), la define como la ausencia de mejoría después de 6 semanas de tratamiento, a dosis de 40 mgs de haloperidol diario; cambios menores al 20% en el BPRS, y en la escala de impresión clínica global un puntaje no mayor de 2.

Brenner et al (1990) diseñaron una escala que mide la respuesta a neurolépticos, así como el grado de resistencia en los esquizofrénicos, a saber:

Nivel 1: Remisión Clínica

Se obtiene una respuesta rápida y substancial a la medicación neuroléptica sin embargo, el paciente puede manifestar algunas tendencias anhedónicas y otros síntomas negativos. La impresión clínica global es normal, algunos de los reactivos que miden psicoticidad del BPRS, con puntuación igual o menor que 2. El funcionamiento del sujeto es adecuado.

Nivel 2: Remisión Parcial

Se obtiene una rápida reducción de los síntomas esquizofrénicos, con prevalencia de síntomas psicóticos residuales leves. La impresión clínica global es igual a dos, ninguno de los reactivos de psicoticidad es mayor de 3. El paciente funciona con supervisión ocasional en áreas sociales y vocativas.

Nivel 3: Resistencia Leve

Existe una lenta e incompleta reducción de síntomas positivos y necesitan supervisión ocasional. Los puntajes en la impresión clínica global de cuatro. Dos de los reactivos de psicoticidad negativos, con efectos adversos en dos o más áreas del ajuste, individual y social, requiriendo supervisión ocasional. La escala de impresión clínica global es de 3, no más de un reactivo de psicoticidad, igual o mayor de 4.

Nivel 4: Resistencia Moderada

Se presentan reducción en algunos síntomas, pero persisten varios que inciden adversamente en cuatro o más áreas del funcionamiento. Se requieren que el individuo sea supervisado. Obtiene una puntuación en el BPRS de 45 en la versión de 18 reactivos.

Nivel 5: Resistencia Severa

Existe una reducción de síntomas pero la prevalencia de otros síntomas afecta seis o más áreas del funcionamiento social e individual, que hacen que el paciente requiera frecuente supervisión. La impresión clínica global, igual a 5. Puntaje en los reactivos de psicoticidad de 5 en al menos 3 de los reactivos, BPRS con puntaje total de 50.

Nivel 6: Moderadamente Refractario

Leve o ausencia de reducción de síntomas y persistencias de síntomas positivos y negativos que marcadamente alteran el funcionamiento de todas las áreas de ajuste personal y social. En la escala de impresión clínica global calificación de 6. Al menos un reactivo de psicoticidad del BPRS, igual a 6, o dos reactivos iguales a cinco. La puntuación global del BPRS es de un mínimo de 50.

Nivel 7: Severamente Refractario

No existe reducción de los síntomas con alta prevalencia de síntomas positivos y negativos; con trastorno de conducta, disruptivos y peligrosos. Todas las áreas de ajuste social se encuentran severamente alteradas, y hacen que el sujeto requiera una supervisión constante. El CGI es igual a 7, al menos unos de los reactivos del BPRS es igual a 7: el puntaje del BPRS es de un mínimo de 50.

4 El Wisconsin Card Sorting Test y la Esquizofrénia

La prueba de Wisconsin (WCST: Wisconsin Card Sorting Test) fue desarrollada para evaluar la capacidad de razonamiento abstracto y el cambio de estrategias cognoscitivas en respuesta a contingencias y cambios en el medio ambiente (Berg, 1948; Grant y Berg, 1948).

La aplicación de esta prueba consiste en presentar al sujeto cuatro cartas clasificadas de acuerdo a tres características: el color (rojo, verde, amarillo, azul), el número (uno, dos tres y cuatro símbolos), y forma del símbolo (triangulos, circulos, cruces o estrellas), las cuales debe de parear por color forma o número. Esta prueba

requiere de estrategias en la planeación, búsqueda organizada, utilización medioambiental para cambiar sets cognitivos, dirigiendo la conducta hacia un objetivo determinado y modular la respuesta a la impulsividad (Chelune y Baer, 1986; Gnys y Willis, 1991; Welsh y Pennington, 1988).

Ha sido ampliamente reportada una baja ejecución en el Wisconsin en pacientes esquizofrénicos y se ha hipotetizado que este déficit esta relacionado con una disfunción en el lóbulo frontal, específicamente la corteza prefrontal dorsolateral (Weinberger y Cols. 1986; Haut. y cols., 1996; Gold y Cols, 1997; Volz y cols, 1997).

Sullivan y cols (1993), realizaron un estudio en tres grupos de sujetos: 22 pacientes con esquizofrenia, 20 pacientes con alcoholismo crónico y 16 sujetos normales, por medio de la prueba del Wisconsin Card Sorting Test. Sus principales objetivos fueron: 1.- identificar las múltiples dimensiones del funcionamiento del WCST por medio del análisis de componentes principales (ver más adelante); 2.- examinar los criterios de validez de los factores resultantes tanto como su sensibilidad para reflejar deficiencia en pacientes con lesiones que invaden la corteza dorsolateral prefrontal (DLPFC); y 3.- para valorar la validez del constructo de cada factor del WCST contra otras medidas de las funciones del lóbulo frontal y lóbulos no frontales en dos poblaciones siquiátricas.

Se tomaron los valores de rendimiento de otros 7 pacientes de un estudio anterior con lesiones en el DLPFC para el factor de perseveración el de clasificación ineficiente.

A los primeros 3 grupos de pacientes se les administraron cuatro medidas de función ejecutiva (incluyendo el WCST) y cuatro de función de memoria.

Se realizaron comparaciones entre los pacientes del lóbulo frontal con los controles a través de la prueba *t*, y se redujeron las puntuaciones del WCST por medio de un análisis de componentes principales usando rotación varimax. Se tomaron en cuenta 3 factores, a los que se les llamó: Perseveración, Factores de clasificación ineficiente, y Factores de Errores no perseverativos, con el 58%, 19% y 14% de la varianza explicada respectivamente.

Los grupos de esquizofrénicos y alcohólicos mostraron similitud en los patrones de deficiencia que los observados en el de lesión en lóbulos frontales. Al final,

al comparar los patrones de rendimiento en en WCST y comparados los 4 grupos por medio de ANOVAs y la prueba de Sheffé post hoc, las puntuaciones del factor de perseveración de los esquizofrénicos fue significativamente mayor que los del grupo normal, mientras que la del grupo de alcohólicos fue mayor que la de los normales en la puntuación del factor de clasificación ineficiente.

5 La esquizofrenia y otras pruebas

El agrupamiento estadístico de signos y síntomas clínicos es un método utilizado para clasificar la esquizofrenia. Avasthi y Kulhara (1997) a fin de encontrar y clasificar factores de síntomas generales con base a puntajes clínicos, realizaron 3 escalas denominadas la Escala de valoración de síntomas positivos (SAPS), Escala de valoración de síntomas negativos (SANS) y la Escala global de puntaje psicopatológico (CPRS). Estas fueron aplicadas en una muestra de 80 esquizofrénicos estables diagnosticados de acuerdo con el DSM-III-R. Y encontraron que la dicotomía positivo-negativa no pudo ser replicada, por lo que refieren que es muy simplista e inadecuada para la clasificación de la esquizofrenia según sus síntomas o síndromes. Por otro lado la inclusión de la CPRS permitió distinguir los items de depresión y astenia del factor de síntomas negativos, lo cual hasta hoy no se había hecho.

Otros estudios aportan datos en relación con la estabilidad de los síntomas que sirven para su diagnóstico. Salokangas, (1997) señala que los hallazgos en la investigación dependen ampliamente de los métodos utilizados para la recolección de datos así como de la naturaleza de la población. La sintomatología de la esquizofrenia no es muy estable sobre el tiempo, varía de acuerdo a la duración de la enfermedad, nuevas investigaciones sugieren que los síntomas negativos están presentes desde el inicio de la enfermedad, y que son relativamente estables a lo largo del tiempo. Sin embargo, esta dimensión negativa no es independiente de otras dimensiones, como la depresión. En el segundo año, síntomas depresivos se incorporan a la dimensión negativa. Durante los primeros 5 años la distinción entre síntomas negativos y positivos no esta clara, y se va distinguiendo en las fases tardías de la enfermedad. Las dimensiones negativas están ampliamente asociadas con el sexo del paciente, la edad y el estado civil.

Teóricamente, puede ser posible que la depresión durante el periodo de desarrollo de la enfermedad sea consecuencia de los neurolépticos suministrados y no del desorden esquizofrénico por sí mismo.

La estructura del cuadro o síndrome esquizofrénico es muy complejo, y varía considerablemente de acuerdo con la duración de la enfermedad (Salokangas, 1997).

Estos dos estudios señalan que son más estables los síntomas negativos que los positivos, e incluso que en estudios realizados fuera de USA, el Alfa de Cronbach tiene mayor puntuación para los síntomas negativos que para los positivos, y de igual manera parece ser que los síntomas positivos no son patognomónicos de los síndromes esquizofrénicos (Knorrning y Lindström, 1995).

Otro estudio, realizado por Gerez y Tello (1995), realizado con dos grupos de sujetos, uno clasificado con base en sus síntomas positivos y negativos pudo generar un modelo utilizando análisis discriminante de manera tal que obtuvo una clasificación correcta de los ejemplos, y probando dicho modelo con sujetos desconocidos alcanzó un 78% de sensibilidad y un 85% de especificidad. Este resultado fue interpretado como que los sujetos predominantemente negativos y positivos tienen diferentes perfiles, lo que dichos autores encuentran consistente con las hipótesis de hipofrontalidad y disfunción del lóbulo temporal.

6 La dimensionalidad

El mismo estudio de Salokangas (1997) refuerza la idea de que existen más dimensiones derivadas de la sintomatología que las descritas. Las variables proporcionadas por el Comprehensive Psychopathological Rating Scale (CPRS) después de ser factorizadas generaron el descubrimiento de mayores dimensiones, como son:

- Una dimensión negativa bastante estable con lenguaje afectivo y empobrecido e introvertido. Esta dimensión fue más prevalente en pacientes no casados.
- Una dimensión engañosa

- Una dimensión alucinatoria.
- Una dimensión desorganizada
- Una dimensión depresiva con experiencias no reales.

Otro estudio realizado por Knorrning y Lindström (1995) sugiere que manejar un modelo dual como el de síntomas Positivos y Negativos es muy restringido y sugieren uno de 5 factores como son: los síntomas positivos, los negativos, el nivel de excitabilidad, el aspecto cognitivo y la ansiedad/depresión.

Un año después Lenzenweger y Dworkin (1996) apoyan la multidimensionalidad de la esquizofrenia con base en un modelo de cuatro factores (genético, neurobiológico, neurológico y psicosocial). Estos cuatro factores podrían ser utilizados en la clasificación y organización de la esquizofrenia.

Así, se han realizado numerosos estudios con métodos estadísticos de tipo factorial como: Análisis discriminante, Análisis de componentes principales, Análisis de conglomerados y regresión lineal, para tratar de clasificar a pacientes esquizofrénicos en subgrupos, pero estos métodos utilizan muchos sujetos, con un pobre control de variables relacionadas con las características de los sujetos. En ninguno se ha tomado en cuenta la neurolepto-resistencia.

II.3 LOS DATOS ELECTROENCEFALOGRAFICOS (EEG)

1 El EEG

El electroencefalograma o EEG es el registro continuo de las fluctuaciones espontáneas de voltaje generadas por el cerebro (John, 1977). Es llamado un registro grueso, ya que se efectúa sobre la respuesta de una población celular (Thompson, 1986). La actividad eléctrica cerebral, se compone de diferentes ritmos cerebrales con características de voltaje, frecuencia, topografía, reactividad y morfología particulares (Tyner, Knott y Mayer, 1983). La frecuencia es definida como el número de ondas en un segundo (Hertz o ciclos por segundo). El EEG se puede dividir en bandas que se

caracterizan por su frecuencia: Delta (δ) de 0.5 a 3.5 Hz, Theta (θ) de 4 a 7.5 Hz, Alfa (α) de 8 a 13 Hz y Beta (β) de 13.5 a 30 Hz. La amplitud de las ondas del EEG es del orden de los microvolts y su distribución en la corteza es diferente, de tal forma que la banda beta se localiza principalmente en regiones anteriores mientras que la alfa se encuentra preferentemente en las regiones posteriores.

Las bases neurofisiológicas del EEG pueden ser resumidas de la siguiente manera (Kiloh, McComas, Osselton, y Upton, 1981):

1. Las ondas repetitivas que pueden ser registradas de la superficie del cerebro o del cuero cabelludo son sumas de potenciales sinápticos generados por las células piramidales en la corteza cerebral.
2. Los potenciales sinápticos son las respuestas de las células corticales a las descargas rítmicas del núcleo talámico.
3. La frecuencia y la amplitud de las descargas talámicas (y la razón de los potenciales corticales) son determinadas por el especial arreglo de interconexiones excitatorias e inhibitorias entre células talámicas.
4. Durante la 'activación', impulsos de la formación reticular abolen las descargas rítmicas en el núcleo talámico y causan que los potenciales corticales lleguen a desincronizarse.

El registro del EEG, conforme a las técnicas para los grandes grupos celulares, se obtiene utilizando electrodos gruesos o macroelectrodos (generalmente de un tamaño superior a los 0.1 mm de diámetro) (Thompson, 1986). La señal es llevada a un amplificador el cual aumenta el voltaje detectado por los electrodos, después se pasa por un convertidor analógico-digital que convierte la señal continua a cortes pequeños que representan la posición de un punto en una línea y de esta forma es capturada en una computadora.

2 Características del EEG

La actividad EEG es siempre dinámica, y es constantemente influenciada por diversos factores tales como: el estado de conciencia (sueño-vigilia), el nivel de activación (reposo, alerta, atención), las entradas sensoriales (ojos abiertos, cerrados,

la aplicación de estímulos, la madurez del sistema nervioso, la salud-patología cerebral y el estado emocional.

A partir de los cada vez más sofisticados aparatos de registro de la actividad electroencefalográfica, se ha podido determinar la existencia en el EEG del hombre de diversos tipos de ondas que tienen un rango de frecuencia entre los 0.5 y los 50 Hz, con una amplitud entre 1 y 150 μV en vigilia y hasta 300 μV durante el sueño (Corsi-Cabrera, 1983). Los investigadores han dividido estas frecuencias en cinco bandas de acuerdo a su morfología, frecuencia, amplitud, reactividad y área cerebral de localización preferente. Estas cinco bandas son:

2a) BANDA DELTA (δ)

Como actividad fisiológica del EEG se encuentran ondas δ en los lactantes, y a veces en los niños y los adolescentes, mientras que en el adulto, excepto durante el sueño (fases III y IV) y en la hiperventilación, casi siempre tienen significación patológica (Simon, 1983). La actividad δ tiene una frecuencia que va de los 0.5 a los 3.5 Hz y una amplitud entre 100 y 300 μV .

2b) BANDA THETA (θ)

La actividad θ es un componente normal sobresaliente en el EEG infantil. Se extiende por la banda de frecuencia de los 4 hasta debajo de los 8 Hz y una amplitud entre 50 y 75 μV que en ocasiones puede alcanzar los 100 μV . Se presenta en las primeras fases del sueño lento. Suele obtenerse en las regiones temporal y temporo-parietal de los niños, pero también las hay durante la tensión emocional de ciertos adultos, sobre todo cuando existe frustración, así como en casos de patología cerebral. El ritmo θ se bloquea al abrir los ojos (Simon, 1983).

2c) BANDA ALFA (α)

La banda alfa (α) está compuesta por ondas regulares o irregulares sobre ambos hemisferios cerebrales cuya frecuencia oscila entre los 8 y los 13 Hz. Se presenta principalmente en estado de vigilia relajado, con los ojos cerrados. El ritmo α

es más claro en las regiones posteriores de la cabeza. En la infancia es inicialmente inestable, estabilizándose hacia los 10 años. La amplitud α es variable, siendo la occipital generalmente más alta: por debajo de los 50 μV en los adultos (Simon, 1983).

Comúnmente se presenta una baja amplitud en el ritmo α a causa de cansancio o mal relajamiento. La amplitud tiende a incrementarse mediante distintos estímulos externos e hiperventilación. La actividad α se bloquea totalmente al abrir los ojos y parcialmente en presencia de cansancio. Después de tener por unos cuantos minutos los ojos abiertos, puede presentarse también el ritmo α (Simon, 1983).

2d) BANDA BETA (β)

La banda β se compone de ondas de más de 13 Hz, en las regiones frontales de la cabeza de un sujeto adulto en estado de vigilia con los ojos abiertos. La amplitud de la actividad β es variable, generalmente menor a los 30 μV . Se bloquea mediante movimientos contralaterales o estímulos táctiles.

En las mujeres adultas el ritmo β presenta una amplitud notablemente mayor, siendo ésta una diferencia sexual ausente en la infancia. Así mismo, la actividad β en los niños posee una amplitud más elevada, y una frecuencia de 16 a 24 hz.

En cerca del 8% del promedio general de la población, las ondas β dominan en estado de vigilia relajado con los ojos cerrados, simétricamente en ambos hemisferios. No existe, en este porcentaje de la población, ninguna diferencia de amplitud considerable en las distintas regiones de la cabeza, siendo ocasionalmente superior en los frontales y temporales, pero por lo general no superiores a los 30 μV .

2e) BANDA GAMMA (γ)

Son ondas cuya frecuencia oscila de los 30 hasta los 50 Hz, teniendo una amplitud promedio de 20 μV . Tal ritmicidad se ha asociado tanto a aspectos de integración cognitiva, (Desmedt y Tomberg, 1994; Llinas y Ribary, 1993), como a la activación intensa del sistema nervioso central y la desorganización emocional.

3 Análisis del EEG

El análisis del EEG se ha realizado a través de la inspección visual y de métodos cuantitativos digitales. Los análisis que se usan más comúnmente para estos métodos cuantitativos son la Transformada Rápida de Fourier (TRF) y los análisis de Coherencia y Correlación.

3a) TRANSFORMADA RAPIDA DE FOURIER (TRF).

Mediante la TRF se obtiene el valor de amplitud y de frecuencia del EEG, los cuales se pueden agrupar en bandas de frecuencia particulares (Figura 2). Nos permite obtener la potencia absoluta o energía de la señal, ya sea de cada frecuencia en forma aislada o de las frecuencias agrupadas en bandas: delta, theta, alfa1, alfa2, beta1 y beta2, así como, la potencia relativa de estas bandas (Guevara, 1995).

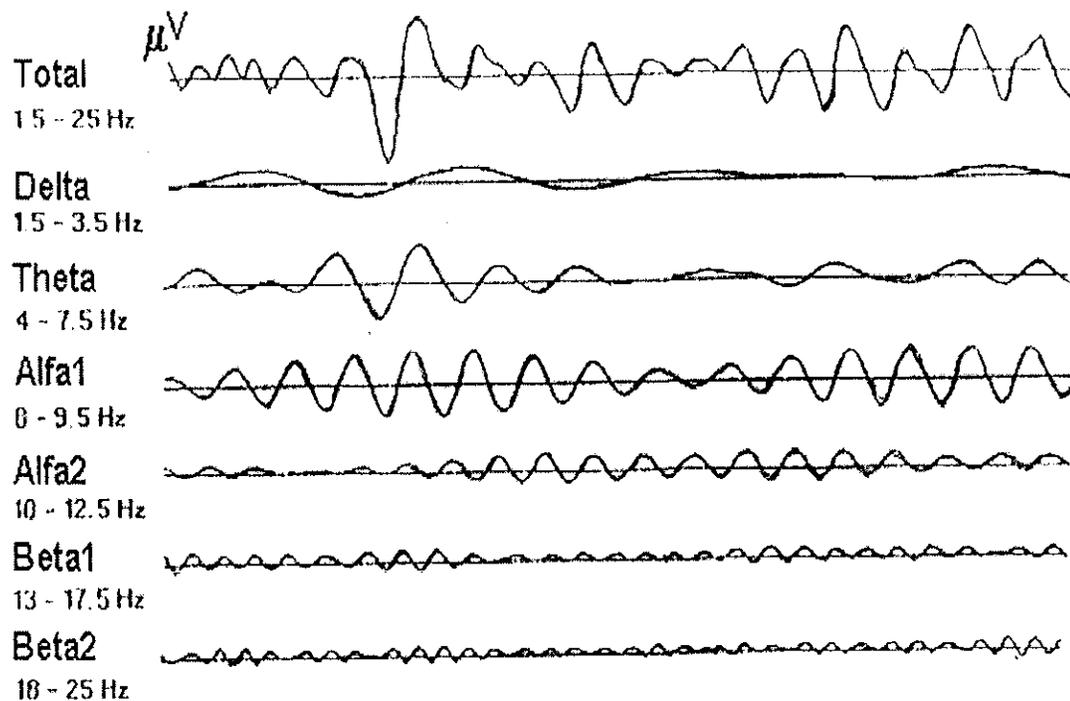


Figura 2. Descomposición de una señal en sus bandas componentes, mediante la Transformada Rápida de Fourier.

3b) ANALISIS DE COHERENCIA.

El cálculo de coherencia se ha empleado para encontrar relaciones de estabilidad entre dos señales estocásticas en el tiempo, que hayan sido digitalizadas, cuya relación sea lineal y de las que se haya obtenido puntuaciones en pares independientes. La coherencia se calcula siempre a través de una transformación al dominio de la frecuencia y se obtiene un coeficiente de coherencia, para más de un par de segmentos de la señal, para cada valor de frecuencia obtenido como componente de la señal que se ha analizado. El valor de cada coeficiente de coherencia oscila entre cero y uno (Guevara, 1995).

Debido a que la coherencia se obtiene de los espectros de la TRF, su cálculo, a partir del avance tecnológico, ha resultado más fácil y rápido en relación al de la correlación, por lo que muchos paquetes comerciales incluyen el análisis de coherencia y no el de correlación. Esto ha provocado que en la mayor parte de las investigaciones se utilice el análisis de coherencia y no el de correlación (Guevara, 1995).

3c) ANÁLISIS DE CORRELACIÓN:

Al igual que el análisis de coherencia, el análisis de correlación se ha empleado para establecer una posible relación funcional entre diferentes regiones del cerebro, solo que en lugar de tomar en consideración la estabilidad de las señales en el tiempo, resalta la relación de fase entre las señales electroencefalográficas de dos áreas. Para el cálculo de correlación entre la actividad EEG de áreas corticales se usa el coeficiente de correlación producto-momento de Pearson, que se ubica entre los índices que tienen mayor potencia estadística para encontrar posibles relaciones entre variables, y se aplica siempre que se tengan dos variables continuas cuya relación sea lineal y de las que se haya obtenido puntuaciones en pares independientes.

En un estudio reciente del Dr. Miguel Ángel Guevara (1995), donde se hizo una comparación entre correlación y coherencia, se llegó a la conclusión de que el análisis de correlación tiene ventajas sobre el de coherencia, ya que, entre otras cosas, permite conocer el sentido de la relación entre las señales a través del signo,

aspecto que se pierde en la coherencia, y no depende del valor de la potencia. Asimismo, Guevara (1995) desarrolló un método para calcular la correlación en un tiempo similar a la de la coherencia mediante el uso de la TRF.

4 Otros metodos de análisis del EEG

4a) ANALISIS DE COMPONENTES PRINCIPALES

El análisis de componentes principales fue desarrollado por Hotelling (1933) (ver en Ato y López, 1994), y su objetivo es describir la variación de un conjunto de sujetos y observaciones en términos de un conjunto de variables no correlacionadas que son combinaciones lineales de las variables originales. La ecuación básica de este método en el primer componente principal es:

$$Y_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1k}X_k$$

Donde Y_1 es el primer componente principal, X_1, \dots, X_k corresponde a las k variables del estudio y a_{1k} corresponde a las cargas factoriales del primer componente principal en las k variables manifiestas.

El análisis de componentes principales (CP) es una técnica de análisis multivariado que se puede aplicar cuando se tienen p variables dependientes (y estocásticas) correlacionadas que explican el comportamiento de una población de datos. Esta población de datos puede ser representada por la variable aleatoria p -dimensional X' donde $X' = [X_1, \dots, X_p]$, con un vector de medias μ y una matriz de covarianzas σ . El análisis lo que pretende es encontrar un nuevo conjunto Y de variables Y_j (con $j = 1, \dots, p$) donde estas Y_j no están correlacionadas entre sí y cuyas varianzas están en orden decreciente en magnitud desde la de Y_1 hasta la de Y_p . El análisis de componentes principales no supone distribución normal en ninguna de las variables con las que va a tratar y esto hace que esté más cerca de la estadística no paramétrica que de la paramétrica con todas las ventajas y desventajas de ésta (Guevara, 1988).

4b) REGRESIÓN LINEAL MULTIPLE.

El análisis de regresión lineal es una herramienta matemática para tratar de explicar la dependencia de una variable, llamada variable dependiente, en función de otras variables, llamadas explicatorias. Al hacer esto, se persigue generar un modelo matemático (una función de regresión) que explique el comportamiento de una población de datos, o bien, predecir el comportamiento de dicha población (Guevara, 1988).

Las suposiciones asociadas son las siguientes (Wayne, 1996):

1. Las X_i son variables no aleatorias (fijas). Esta suposición distingue al modelo de regresión del modelo de correlación múltiple. Esta condición indica que cualquier inferencia que se haga de los datos de la muestra sólo se explica, al conjunto de valores de X observados y no a algún conjunto mayor de X .
2. Para cada conjunto de valores de X_i existe una subpoblación de valores de Y . Para construir ciertos intervalos de confianza y probar la hipótesis, es necesario saber, o bien, el investigador debe inclinarse a suponer, que estas subpoblaciones de valores de Y siguen una distribución normal.
3. Las variaciones de las subpoblaciones de Y son todas iguales.
4. Los valores de Y son independientes. Es decir, los valores de Y seleccionados para un conjunto de valores de X no dependen de los valores de Y seleccionados en otro conjunto de valores de X .

4c) ANALISIS DE CONGLOMERADOS

Es una técnica multivariante cuya función consiste en agrupar los casos (o variables) de una matriz rectangular en subgrupos o agrupamientos. Estos agrupamientos pueden obedecer a leyes naturales, como es el caso de las Ciencias Biológicas, o a cualquier grupo de características comunes a los casos o variables, como ocurre en Ciencias Sociales. Por regla general, esta técnica se utiliza en estudios exploratorios donde no se tiene una idea preconcebida de cómo se agrupan los casos o variables. Además, se utiliza, en ciertos casos, como un primer paso para otras técnicas más complejas como el Análisis factorial o el Escalamiento multidimensional.

Como otras técnicas multivariantes, se aplica sobre una matriz de m medidas en cada uno de n objetos. Esta matriz de $n \times m$ se transforma en una matriz de $n \times n$ (o $m \times m$) de similitudes o distancias, donde n es el número de casos (o variables) que se quiere agrupar. Sobre esta matriz se aplica, entonces, el algoritmo conveniente para determinar los conglomerados. Este algoritmo permitirá establecer las reglas por las que los casos (o variables) se unirán a los conglomerados en base a las similitudes entre los objetos de la matriz (Ato y López, 1994).

Básicamente, los métodos de Análisis de conglomerados se pueden dividir en *métodos jerárquicos* y *métodos de partición* (Jonson, 1967; Hartigan, 1975; Dillon y Goldstein, 1984; en Ato y López, 1994). Los métodos jerárquicos se caracterizan por realizar fusiones o divisiones sucesivas para formar los conglomerados. A su vez se pueden dividir en métodos aglomerativos y métodos divisivos. Los *métodos aglomerativos* parten de la existencia inicial de un conglomerado para cada caso objeto de análisis. Por aproximaciones sucesivas se van uniendo los objetos en conglomerados hasta formar un conglomerado único que los incluye a todos.

Los métodos divisivos parten de un conglomerado único y trabajan realizando divisiones sucesivas hasta dejar todos los objetos en conglomerados individuales. Tanto unos como otros presentan sus resultados en forma de árbol, más comúnmente conocidos como dendogramas.

A diferencia de los métodos jerárquicos, los métodos de partición dividen el grupo de objetos en un número preseleccionado de conglomerados que no tienen una estructura jerárquica. A continuación calculan los centroides de esos conglomerados para después asignar un objeto a un conglomerado en función de su cercanía al centroide. Se vuelven a calcular de nuevo los valores centroides y se vuelve a repetir el proceso hasta que los conglomerados se estabilizan en dos etapas sucesivas. Estadísticamente los métodos divisivos se caracterizan por maximizar la suma de cuadrados (SC) Intersujetos entre conglomerados con respecto a la SC intrasujetos. Así, el proceso que emplean estos métodos es similar a un ANOVA donde se desconocen los grupos.

4d) ANALISIS DISCRIMINANTE

La técnica del análisis discriminante es altamente similar al MANOVA y al análisis de correlación canónica. De hecho, podemos caracterizar el análisis discriminante como un análisis de correlación canónica, donde uno de los grupos de variables está formado por variables ficticias (codificación) que representan la pertenencia de los casos a distintos grupos. Por otro lado, el análisis discriminante descansa matemáticamente sobre el MANOVA; el análisis discriminante efectúa una descomposición en eigenvalores y eigenvectores de la matriz **SCR-1SCM**, donde **SCR** y **SCM** son las matrices de sumas cuadráticas del modelo y sumas cuadráticas de error de un MANOVA (Ato y López, 1994).

5 Inteligencia Artificial y EEG

El abordaje del EEG con base en la Inteligencia Artificial hace énfasis en la identificación de algunas de las características de la señal electroencefalográfica con fines de apoyo en la toma de decisiones o en el análisis de los datos. Este objetivo privilegia la aplicación de la técnica en estudios sobre actividad epiléptica.

En este sentido se han utilizado Sistemas Expertos para la detección de actividad epileptiforme en el EEG, tal como el que elaboraron Davey y cols. en 1989, este sistema consta de dos partes, la primera se encarga de detectar todas las ocurrencias similares a las espigas buscadas, mientras que en un segundo paso un sistema de reglas integrado en el Sistema Experto identifica y elige únicamente aquellas que cumplen con los requisitos para realmente ser clasificadas como actividad epileptiforme. Este sistema basado en reglas debe ser programado con base en los conocimientos de un experto en la materia que integra información tal como: la orientación del canal (con relación a sus más próximos), la sincronía, la amplitud y la duración. En su fase experimental, el sistema ha logrado clasificar correctamente la mayoría de los eventos epileptiformes ocurridos en 320 segmentos con los que se probó.

Sistemas similares, también basados en reglas han incorporado información referente al entorno electroencefalográfico en el que se produce la espiga para identificar a ésta. Tal es el caso del sistema perfeccionado por Dingle y cols. en 1993, probado con un total de 180 minutos de EEG en 16 canales, proveniente de 11

pacientes. Estos sistemas no realizan falsas detecciones (falsos positivos) y el porcentaje de aciertos en la detección de espigas epileptiformes identificadas por un experto humano alcanza hasta el 80%.

II.4 EL PROBLEMA DE LA PREDICCIÓN

El problema de la predicción estadística es definir a qué grupo pertenece un nuevo individuo. En la práctica se sabe que el error que se cometerá con los nuevos individuos es mucho mayor que los errores de clasificación de los individuos que se utilizaron para calcular los límites. (Harmony y Alcaraz, 1987).

La predicción estadística se basa en la consideración de que un sujeto desconocido pertenece a un grupo y no a otro de un par (o más) de opciones dadas. Y estas opciones de pertenencia se respaldan en un número variable de sujetos que le aportan a cada grupo sus valores constantes (de grupo) y minimizándose en mayor medida sus valores individuales.

Por ello, con el fin de disminuir al mínimo la posibilidad de una predicción equivocada en un nuevo sujeto, es recomendable lo siguiente (Harmony y Alcaraz, 1987):

- Haber observado en el análisis de la estructura de los datos, que existen grupos homogéneos de individuos.
- Haber comprobado que existen diferencias muy significativas entre los grupos que se intentará clasificar.
- Obtener clasificaciones correctas de más de 80% de los individuos con los que se calcularon los límites que definen a los grupos.

Una consideración adicional que se debe tener en cuenta es aquella bien conocida que la replicabilidad independiente que cabe esperar de las funciones discriminantes o de los análisis de conglomerados disminuye en función del número de variables para estos cálculos. En general, existe "una regla del pulgar" que establece que deben tenerse por lo menos 5, y preferiblemente 10 veces, más sujetos que variables si se espera que una función discriminante pueda clasificar adecuadamente una muestra independiente.

La predicción basada en EEG ha sido utilizada con varios métodos: Análisis discriminante, regresión lineal, etc, sin embargo, también son comunes las ocasiones en que se conjugan los métodos de análisis del EEG con los métodos predictivos. Tal es el caso de la investigación realizada por Guevara (1988) en la cual se utilizó el análisis de componentes principales, con el objeto de reducir variables en una investigación orientada a determinar la existencia de patrones de respuestas electroencefalográficas diferentes en niños disléxicos con respecto a niños normales. En dicha investigación, se tomaron los 6 primeros componentes (que explicaban el 94.14% de la varianza) y se realizó un análisis de regresión lineal múltiple con el objeto de hacer un modelo que permitiera discriminar entre niños normales y disléxicos, dicho modelo clasificó correctamente a 8 de 9 niños normales y a 7 de 9 disléxicos.

1 El Análisis Discriminante

En 1973, Harmony y cols, utilizaron los valores del coeficiente de correlación de polaridad (CCP) y la relación de energía (RE), para cuantificar las asimetrías cerebrales de un grupo de sujetos normales y otro de pacientes (con accidente vascular encefálico, con tumores cerebrales y con epilepsia). Las dos variables (CCP y RE) se midieron entre distintas parejas de derivaciones homólogas izquierdas y derechas, durante dos condiciones: con el sujeto en reposo con los ojos cerrados y durante la estimulación luminosa intermitente.

Al realizar la comparación entre grupos se encontraron diferencias significativas en ambos valores, principalmente entre los normales y los dos primeros grupos de pacientes

Al calcular una ecuación discriminante entre los sujetos normales y los pacientes con tumores cerebrales, utilizando las 36 variables obtenidas (9 derivaciones X 2 mediciones X 2 condiciones), se obtuvo una eficiencia del 82% en la clasificación de todos los individuos.

Una análisis similar entre el grupo de sujetos normales y el grupo de pacientes con accidente vascular encefálico demostró que la ecuación tenía 80% de eficiencia en la predicción.

No se encontraron diferencias significativas entre los sujetos normales y los pacientes epilépticos en quienes se empleo el mismo tipo de análisis discriminante, por

lo que se concluyó que este procedimiento no era útil en la detección de pacientes cuyas alteraciones electroencefalográficas se caracterizaran por la presencia de actividad paroxística. Sin embargo, el éxito en la clasificación (en la etapa de la definición del modelo) de los sujetos con lesión cerebral en comparación con los normales implica que la unión de las variables con el método es adecuada para discriminar grupos y clasificar sujetos asociados con dichos grupos, lo que tiene valor predictivo.

2 La Regresión Lineal

El equipo de Petersen y sus colaboradores en Gotenburgo ha diseñado un método para la evaluación automática del EEG basado en el análisis de frecuencias. Dicho método se basa en la similitud que tienen los EEGs anormales (por la presencia de ondas lentas) con los EEGs de los individuos en edades más tempranas de la vida.

Primeramente efectuaron el análisis de frecuencias en 650 EEGs, de sujetos normales de 1 a 21 años de edad (Matousek y Petersen, 1973). Se obtuvieron los valores de energía para cada banda electroencefalográfica (alfa, beta, theta y delta) y las interrelaciones entre estos valores para cada derivación. Con estos datos efectuaron un análisis de regresión, para obtener la correlación más alta posible entre el EEG y la edad. De esta forma, encontraron las ecuaciones que, aplicadas a los parámetros electroencefalográficos (valores de energía de las bandas y sus interrelaciones), permiten conocer la edad teórica electroencefalográfica de un individuo. De esta edad teórica se deriva el cociente dependiente de la edad (CDE). Este procedimiento, basado exclusivamente en el análisis de un minuto del EEG, se aplicó a 500 sujetos y tuvo una correspondencia del 80% con la interpretación por inspección visual del trazado completo de 30 minutos por dos expertos (Petersen y Matousek, 1975).

La regresión lineal permite conocer el grado en que un grupo de variables (en muchos casos, las independientes) explican (por eso se definen como explicatorias) a las variables dependientes. En este caso lo importante es que la relación establecida en el modelo creado, pudo acercarse bastante a la clasificación lograda por un experto humano.

III PLANTEAMIENTO DEL PROBLEMA

Dentro del padecimiento de la esquizofrenia se pueden encontrar entre un 10 a 25% de pacientes que no responden a los tratamientos tradicionales y que por ende, su pronóstico es de mala evolución y deteriorante. Los estudios al respecto de esta variante de la enfermedad son escasos y el retraso en el diagnóstico correcto es común.

Actualmente, no existe un sistema o una metodología capaz de identificar por medio del EEG a un paciente esquizofrénico en Resistente y No Resistente (a la medicación típica) de manera que se detecte esta característica antes de hacerlo de forma clínica. Tampoco existe una herramienta que pueda hacerlo de forma semiautomatizada.

La existencia de una herramienta de clasificación temprana facilitaría la identificación de dicha característica y por tanto evitaría el continuo cambio de medicamentos y la prolongación del tiempo que el enfermo sigue padeciendo el problema.

En el campo de la práctica médico-clínica de la psiquiatría, se requieren herramientas que apoyen al médico en el diagnóstico eficiente y temprano de padecimientos complejos y difíciles, como lo es la Esquizofrenia. Esta necesidad se ve incrementada cuando la enfermedad sujeta a diagnóstico se escapa al conocimiento del clínico en virtud del poco conocimiento sobre las características propias del problema. Tal es el caso del grupo de Esquizofrénicos Neurolepto-Resistentes.

En este contexto, la aportación de una herramienta efectiva, capaz de aprender sobre las características del problema, generar diagnósticos y explicaciones de los mismos, es un elemento invaluable para la pronta identificación y atención del padecimiento.

Actualmente, se cuenta con la tecnología y los conocimientos para producir un sistema experto con esas características, que puede ser el primer especialista en este padecimiento que toma el EEG como elemento diagnóstico fiable.

La doble utilidad de esta herramienta es su capacidad para ser reproducida, adecuada y aplicada a otras necesidades de clasificación basadas en EEG.

IV OBJETIVOS GENERAL Y ESPECIFICOS

IV.1 OBJETIVO GENERAL

Elaborar un Sistema Experto basado en un árbol de decisión generado por aprendizaje de ejemplos basado en el algoritmo de la Ganancia de Información, capaz de hacer la clasificación correcta de sujetos esquizofrénicos paranoides en dos grupos: Resistente (a neurolépticos típicos) y No Resistente, tomando como datos de entrada los parámetros de su actividad eléctrica cerebral.

IV.2 OBJETIVOS ESPECIFICOS

1. Determinar cuál(es) método(s) de discretización es(son) más efectivo(s) en combinación con el algoritmo de Ganancia para la identificación de las variables del EEG que poseen el más alto valor de identificación de grupos, al comparar los datos de los sujetos Resistentes con los No Resistentes.
2. Elaborar un árbol de decisión basado en aprendizaje por ejemplos que pueda clasificar satisfactoriamente a la totalidad de los sujetos disponibles que entran a formar parte del Conjunto de Capacitación. Probar el árbol con datos no incluidos en el Conjunto y realizar los ajustes necesarios.
3. Integrar el sistema para que sea capaz de analizar la información de un nuevo EEG y clasificarlo en alguno de los siguientes grupos:
 - Resistente
 - No Resistente

V HIPOTESIS

Primera hipótesis:

Un Sistema Experto informático basado en la metodología de generación de árboles de decisión por aprendizaje de ejemplos, que utilice como elemento principal para valorar la capacidad de discriminación de una variable el algoritmo de la Ganancia de Información, será capaz de hacer la clasificación correcta de sujetos esquizofrénicos paranoides en dos grupos: Resistentes (a neurolepticos típicos) y No Resistentes, tomando como datos de entrada los parámetros de su actividad eléctrica cerebral.

Segunda hipótesis:

El número de clasificaciones correctas logradas por el Sistema Experto será superior al porcentaje de incidencia de cada grupo de esquizofrénicos (Resistente y No Resistente), al logrado por el azar, y al que alcance un Modelo Lineal de Probabilidad (MLP).



VI METODOLOGIA

VI.1 INSTRUMENTO

1 Proceso de construcción del Sistema Experto ESQUIZOR

La construcción del sistema se describe en 4 fases de proceso cíclico continuo de evaluación y desarrollo, modelo tomado de la propuesta de O'Moore, Clarke y Brender de 1990 (Ver, Korpinen, 1993) presentado en la Figura 2.

FASE I

En esta fase del proyecto se definen los procedimientos, maneras de abordaje y consideraciones importantes relativas al tipo de interfaz para interacción con el usuario. El diseño de los arreglos de matrices de datos y las herramientas con que el programa accederá a la información. También se discuten la forma en que se integran las secuencias en el tratamiento de los datos, la forma en que el programa los separa y la importancia que se le asigna a cada dato por separado. Así mismo, se estudia con detalle el método más adecuado para la poda del Arbol.

También se realizan pruebas con cada uno de los prototipos de módulos elaborados por separado verificando que puedan procesar la información de forma adecuada.

FASE II

La validación involucra el examen del sistema con el mundo real. Para esta fase se utilizan ejemplos sencillos o partes de ellos a fin de verificar que el sistema clasifique adecuadamente, se revisa la validez del conocimiento que produce el algoritmo del árbol, esto es, se evalúa la eficiencia del algoritmo de aprendizaje de acuerdo a la teoría de la información que postula un modelo matemático para escoger el mejor de los atributos (características que servirán para diferenciar los grupos), considerando el valor de los mismos en un continuo que va de 0 a 1 y que se

implementa mediante la función DETERMINAR-GANANCIA que mide la cantidad de información que proporciona cada atributo, eligiendo al valor más alto.

FASE III

Esta fase implica la validación del sistema en el ambiente del usuario final para evaluar la interacción del usuario y la interfaz de la máquina, la facilidad de uso, la conveniencia de las secuencias propuestas y la accesibilidad de los menús y las funciones.

Las modificaciones emanadas de esta prueba deben hacer al sistema más amigable, fácil de operar y aprender.

Se examinan detenidamente la facilidad para la transferencia o transporte, instalación e independencia del programa, con la única limitación de correr sobre la plataforma Windows 95, 98 y posteriores.

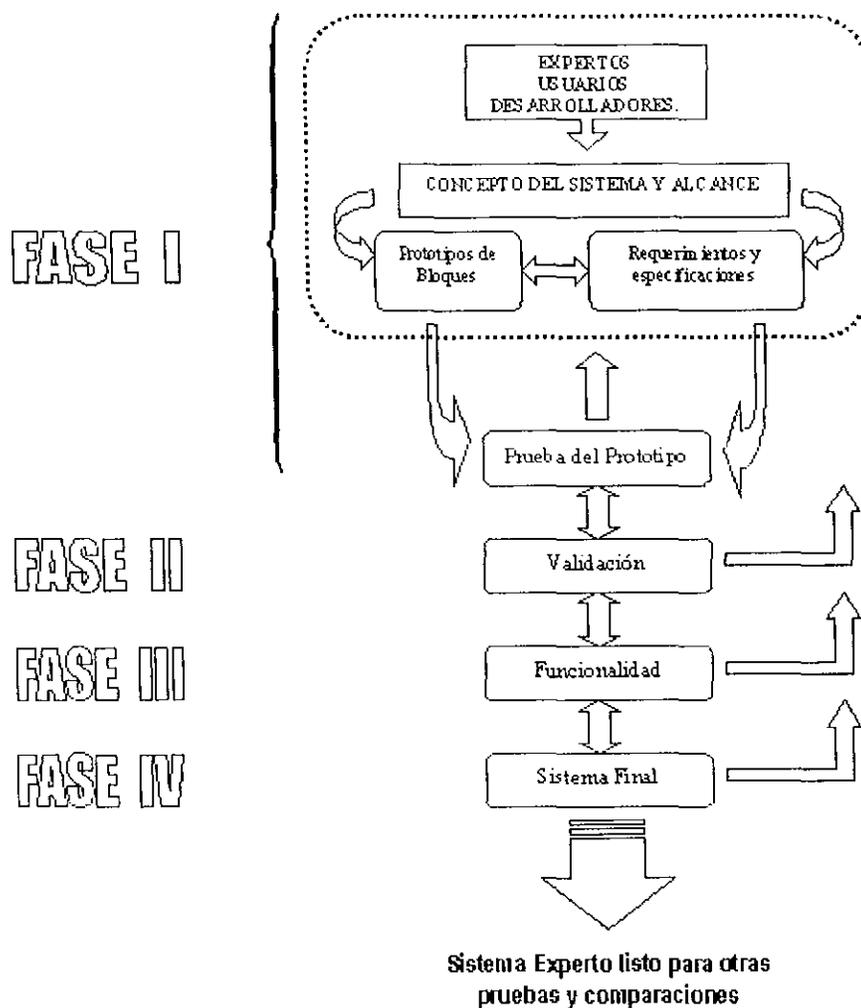


Figura 3. Ciclo continuo de evaluación y desarrollo del proyecto de construcción del Sistema Experto ESQUIZOR.

FASE IV

Esta fase implica pruebas de funcionamiento en cada uno de los módulos, la interfaz de usuario en cada parte, los archivos de ayuda, la accesibilidad, la facilidad de manejo, y los propios resultados de las pruebas de clasificación.

Se alista el sistema para las comparaciones de efectividad con otros sistemas o metodologías de clasificación para la misma área y tipo de pacientes.

Una vez superada con éxito la Fase IV, se realiza una comparación del nivel de eficiencia del sistema con otra herramienta de clasificación.

La valoración de este nivel se llevará a cabo contabilizando el porcentaje de error en la clasificación de sujetos no incluidos en el Conjunto de Capacitación, siendo los valores mínimos los siguientes.

1. Para el grupo de Esquizofrénicos No Resistentes, será considerado como un porcentaje de error inferior al humano, el que sea menor al 25%, en virtud de que el diagnóstico médico se basa en la incidencia de un padecimiento dado, y como se indicó anteriormente un 10 a 25% de los enfermos de esquizofrenia se pueden considerar Neurolepto-resistentes.
2. Para el grupo de Resistentes, será considerado aceptable un porcentaje de error inferior al 50% en virtud de que aún cuando el diagnóstico médico en la práctica se encuentre por debajo del porcentaje de incidencia, el diagnóstico (hipotéticamente) basado en el azar, le concedería un 50% a los aciertos a cada caso de neurolepto-resistencia.

2 Lenguaje elegido: Java

Para la construcción del sistema se eligió el lenguaje Java de Sun microsystems inc., y algunas de las razones que condujeron a esa elección fueron las siguientes:

1. Java fue diseñado como un programa de arquitectura neutral, independiente de la plataforma, con lo que evita el problema que tienen la mayoría de los lenguajes que son diseñados para ser compilados para un destino específico, es decir, se trabajó en él para hacerlo portable y que pudiese ser ejecutado en una serie de CPUs en diferentes entornos.
2. Es orientado a Objetos. Utiliza de manera efectiva las características de herencia, encapsulación y polimorfismo.
3. Es un lenguaje estructurado con alta facilidad de uso.
4. Fue diseñado para cumplir el requisito del mundo real de crear programas en red interactivos, permite escribir programas que hacen varias cosas a la vez, lo que se denomina Multihilos.

5. Java se compila en una representación intermedia llamada código binario (Java bytecode). Este código se puede interpretar en cualquier sistema que tenga un interprete de Java su traducción es directa a código máquina nativo y consigue con ello un alto rendimiento.
6. Fue diseñado para el entorno distribuido de Internet y las características que le han puesto para ello le aportan un nuevo nivel de abstracción en la programación cliente/servidor.
7. Java es dinámico, pudiendo verificar y resolver accesos a los objetos que conforman sus programas durante el tiempo de ejecución.
8. Además tiene características importantes de seguridad.

3 Elaboración del modelo.

El modelo propuesto para este sistema experto de clasificación consta de 11 pasos (Ver Figura 4).

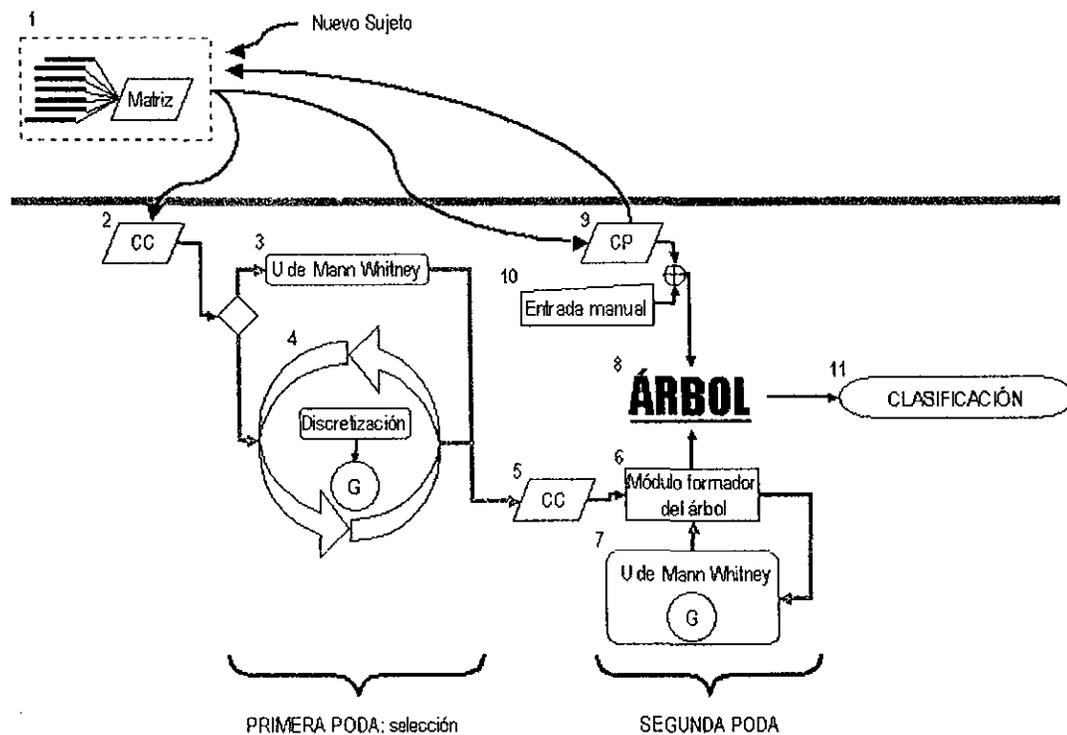


Figura 4. Diagrama de flujo del modelo de funcionamiento modular del sistema Esquizor.

Una breve descripción general de los pasos del modelo es la siguiente:

- Primero. El Sistema necesita ser alimentado de información por medio de un archivo de texto, con los datos de los sujetos y las variables acomodados en renglones y columnas.
El formato de la matriz es igual para el Conjunto de Capacitación (CC) o el Conjunto de Prueba (CP).
La inclusión de los datos de un nuevo sujeto o de un sujeto del Conjunto de Prueba se realiza fuera del Sistema.
- Segundo. La eliminación de uno o varios sujetos de una matriz Conjunto de Capacitación puede realizarse con el editor de textos del programa mismo, y de igual manera, la elección de un sujeto del Conjunto de Prueba o la escritura de sus datos para su clasificación pueden realizarse directamente en el editor. Lo anterior termina de conformar al Conjunto de Capacitación (CC) o al Conjunto de Prueba (CP) respectivamente, para su utilización por el sistema.
- Tercero. A cada una de las Variables en una matriz Conjunto de Capacitación se le realiza una prueba U de Mann Whitney para determinar el valor de la p.
- Cuarto. A cada una de las Variables en una matriz Conjunto de Capacitación se le aplica uno a uno los cuatro métodos de discretización y seguido a cada método aplicado, la variable es pasada por el Algoritmo de Ganancia de la Función Determinar-Ganancia para finalmente discretizarla conforme al método que produjo el valor más alto de Ganancia.
Los pasos 3 y 4 conforman lo que se denomina la primera poda o selección de las variables, el fin de este paso es disminuir el número de variables que formarán el árbol. Está considerado como un método de prueba ya que la eliminación de las variables no relevantes favorecerá la formación de un árbol compacto y con el menor número de variables en cada rama. Los criterios de selección y de exclusión de las variables en este paso se describen más adelante.

- Quinto. En este paso el sistema cuenta con una nueva matriz seleccionada con el mínimo de variables posibles y con el mayor grado de discriminación entre grupos (dado por el valor de Ganancia) y el valor más bajo de P. Con esta nueva matriz se llevará acabo el aprendizaje del sistema y la formación del árbol.
- Sexto. La matriz seleccionada pasa por el Modulo Formador del Arbol Método Formador del Arbol), que envía a cada variable a pasar por el séptimo paso y una vez que por ese medio obtiene el valor de P y el de Ganancia;
- Identifica la variable que logre la mayor discriminación de los grupos implicados.
 - Identifica las ramas emanadas de dicha variable que se encuentren ocupadas únicamente por elementos de un grupo o del otro.
 - Identifica las ramas que se encuentran ocupadas por elementos de ambos grupos, y las envía a pasar por el séptimo paso.
- Séptimo. Se encuentra integrado por dos pasos;
- La fórmula de la U de Mann Whitney, por medio de la cual se determina la significancia estadística para cada variable y,
 - La función Determinar-Ganancia. Ambos pasos regresan los respectivos valores de la variable al paso sexto que la envió.
- La unión de estos dos pasos en la elección de cada variable en cada rama del Arbol, se ha definido como el segundo método de poda.
- Octavo. Con todo lo anterior el Arbol queda formado, y en sí mismo cada nodo representa una regla de decisión que implica el nombre de una variable y cada rama el valor de dicha variable con la que deberá ser comparado el nuevo sujeto a clasificar (elemento del Conjunto de Prueba).
- Noveno. El paso noveno representa la utilización de un archivo de texto que contiene los datos de un sujeto que va a ser clasificado.

Décimo. Dichos datos del sujeto pueden ser introducidos de forma manual al sistema por medio del editor de textos.

Undécimo. El sistema reporta en pantalla la clasificación encontrada para el nuevo sujeto.

4 La Interfaz Gráfica de Usuario (GUI) del sistema

4a) ESQUEMA GENERAL

El Sistema fue escrito totalmente en lenguaje de programación Java utilizando el Entorno de Desarrollo Integrado (IDE por sus siglas en inglés) de Visual Café Professional Edition, versión 4 de la compañía Symantec. El compilador utilizado y el tiempo de ejecución son los de Java Development Kit (JDK) versión 1.1.

La interfaz de usuario se basó en el nuevo conjunto de paquetes que fueron presentados casi al mismo tiempo que el JDK 1.2 (e incluso incluidos en él) denominadas las Java Foundation Clases (JFC) versión 1.1 (también llamados Swing 1.0) sobre todo en los denominados Swing componentes (Componentes Swing). Se eligieron estos componentes debido a que son elementos ligeros que no se atan a los objetos asociados del sistema operativo nativo, sino por el contrario, buscan una de tres bibliotecas de aspecto y comportamiento conectable incluidas en el Swing 1.0. Los elementos de Interfaz Gráfica de Usuario (GUI) de Swin basada en la nueva forma de comunicar información sobre eventos del JDK 1.1: el modelo de eventos de delegación (delegation event model), este modelo define a cada elemento de comunicación entre la GUI y el programa como un evento. Para registrar la ocurrencia de un evento particular de un componente dado, las clases de aplicación piden al componente que añada un oyente a una lista (Morgan, 1999).

Otra característica que se decidió utilizar al elegir los componentes Swing en la construcción de la GUI se relaciona con el aspecto y la sensación de la misma, ya que la independencia de la plataforma de Java permite adaptar los aspectos y la sensación (LyF) de los citados componentes. El aspecto y la sensación la conforman las diferencias entre los controles que se usan en las ventanas de diferentes sistemas (Windows vs Macintosh vs UNIX), la forma en que los menús aparecen y la forma en

que los botones se comportan al pulsarse. Esta característica, a la que se denomina Aspecto y Sensación Conectables (PLyF), le permite al programador crear aplicaciones que utilizan un aspecto y sensación que son independientes de la plataforma de ventanas en las que se ejecuten. El PLyF se implementa en términos de la arquitectura del controlador de visualización de modelos (MVC), esto implica que los modelos disponibles de PLyF dependan de los paquetes contenidos en la versión de JDK que se posea o en los paquetes PLyF que se instalen (Jaworski, 1999).

Para este trabajo se eligió el predeterminado como Metal. El Sistema presenta una interfaz que se encuentra dividida en paneles que funcionan interconectados entre sí, estos se encuentran esquematizados en las Figuras 5 y 6.

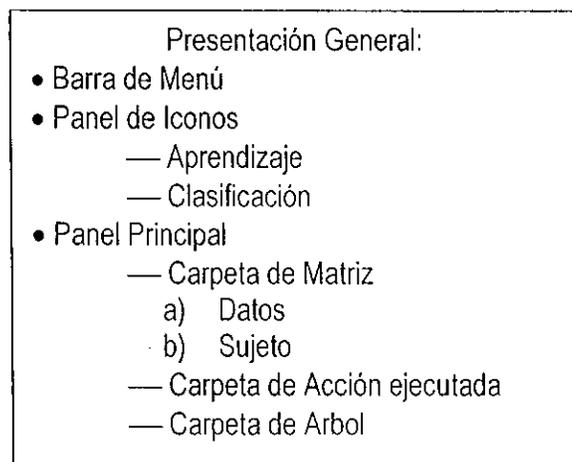


Figura 5. Conformación de la Interfaz gráfica de usuario del programa Esquizor.

Básicamente, el programa fue construido para funcionar en función de dos tareas principales: el aprendizaje/entrenamiento del sistema, o la clasificación de un sujeto. La primera parte incluye los métodos de discretización y de aplicación del algoritmo de la Ganancia de Información, mientras que la segunda se enfoca a la recuperación de los datos contenidos en el archivo de Arbol y a la comparación de los datos de los sujetos para generar un diagnóstico predictivo de su pertenencia a un grupo dado. También se incluye una barra de progreso con una etiqueta que muestra gráficamente el grado de avance de una tarea.

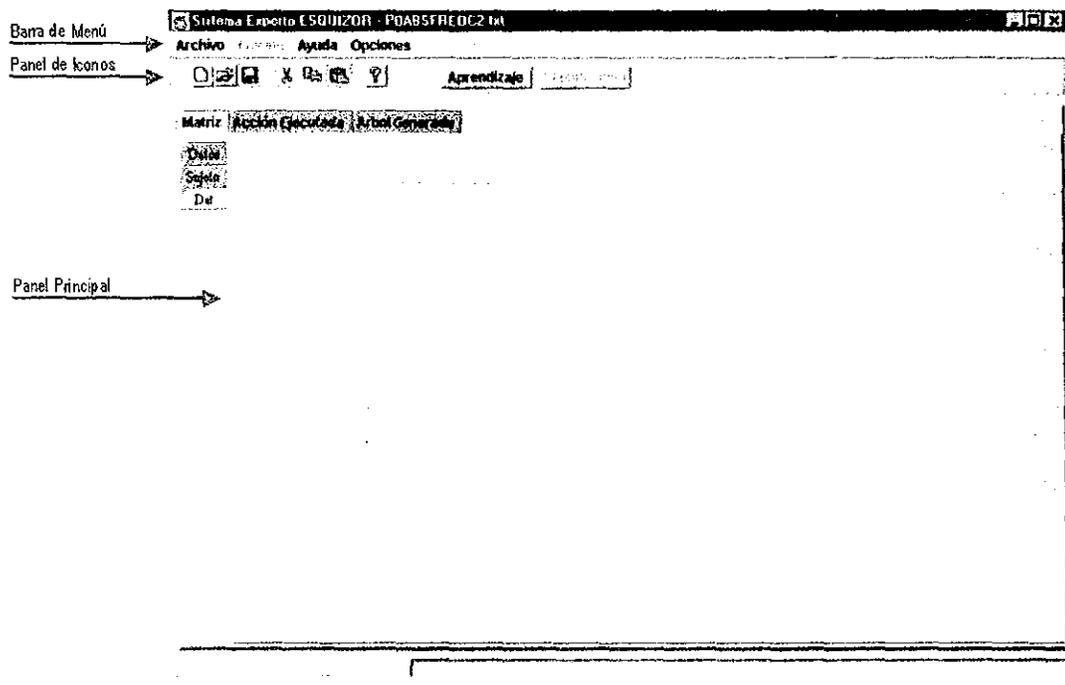


Figura 6. Pantalla principal del Sistema Experto Esquizor.

4b) LA BARRA DE MENU

La Barra de Menú contiene a tres menús principales: *Archivo*, *Ayuda* y *Opciones*.

El Menú **Archivo** contiene a su vez en una cortina vertical a los submenús: *Nuevo*, *Abrir*, *Cerrar*, *Configurar Matriz*, *Clasificar* y *Salir* (ver Figura 6).

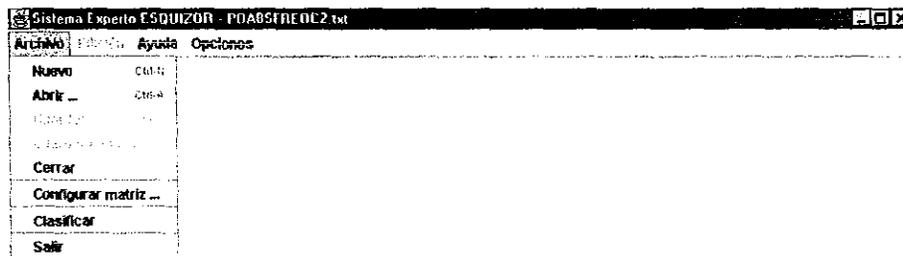


Figura 7. La cortina vertical que se despliega al pulsar el menú *Archivo*.

El submenú *Nuevo* permite la formación de una matriz directamente en la pantalla del Sistema por medio de una tabla en la que se pueden ingresar los datos de los sujetos y sus variables. Se puede decidir el número de las columnas y renglones de la tabla en el momento de su formación o se puede agrandar o recortar durante el proceso de llenado de las celdas. Estas funciones igualmente son llevadas a cabo por medio del primer icono (la hoja en blanco) del Panel de Iconos.

El submenú *Abrir*, utiliza la ventana de dialogo "Open" del sistema Windows para facilitarle al usuario la navegación entre los directorios y archivos de su disco duro, para elegir el archivo de texto que deberá presentar y sobre el que basará su aprendizaje o clasificación, (ver Figura 8).

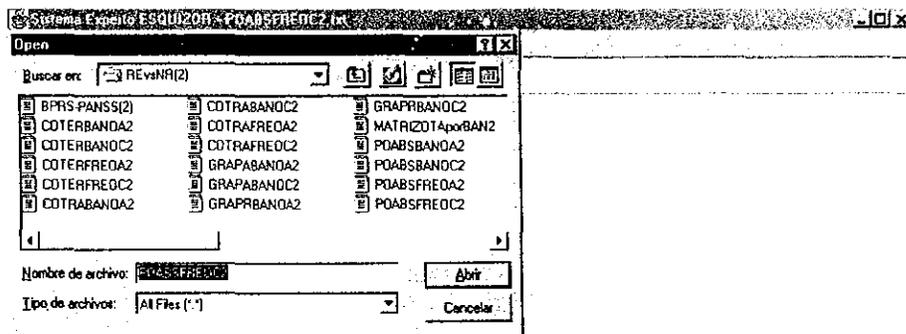


Figura 8. La ventana Open del Sistema de Windows.

De igual manera, esta ventana se puede llamar desde el segundo icono (el fólder abierto) del Panel de iconos, y una vez pulsado el botón de *Abrir* situado en su parte inferior derecha permite ver desplegados los datos del archivo elegido en la zona del Panel Principal, dentro de la Carpeta Matriz, en la Subcarpeta Datos (ver Figura 9).

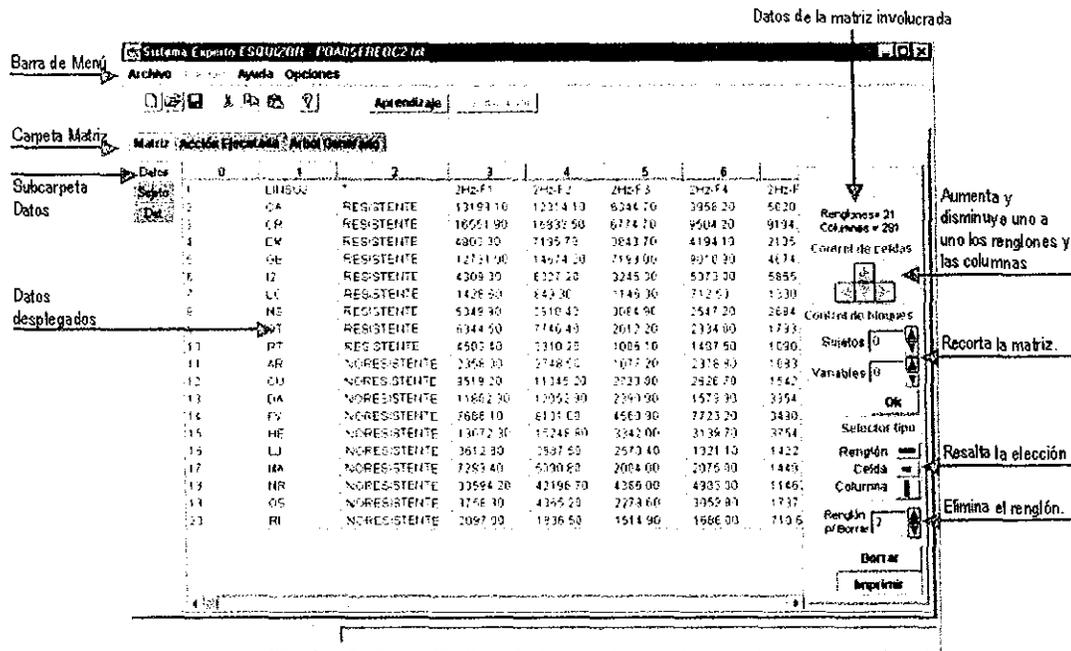


Figura 9. Presentación en pantalla de los datos de una matriz.

Los Submenús *Cerrar* y *Clasificar* se encuentran inhabilitados en esta versión del programa.

El menú *Configurar Matriz* presenta una ventana en la que se definen los parámetros necesarios a fin de que los datos del archivo elegido se presenten en el orden adecuado para su procesamiento por el sistema. En esta ventana se especifica la posición de las variables y los sujetos en la matriz del archivo.

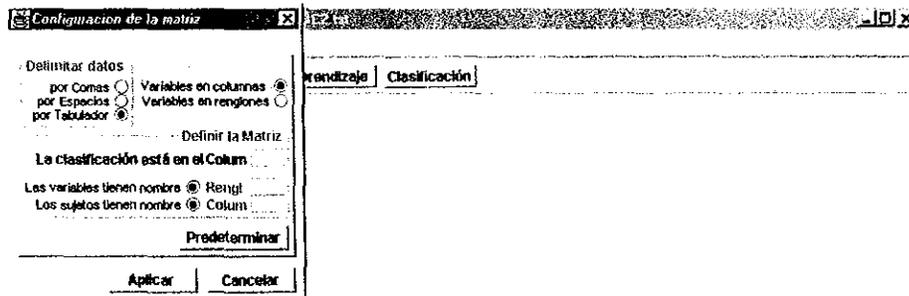


Figura 10. La ventana *Configurar Matriz*.

También, en esta ventana se especifica el tipo de separador de datos utilizado en el archivo, así como la existencia o no de nombres en las variables y/o sujetos.

El Submenú *Salir* permite cerrar todo el programa.

El Menú ***Ayuda*** despliega un submenú llamado *Acerca de..* por medio del cual se presenta una ventana con los datos de la versión del programa, los autores y el lugar en que se desarrolló (ver Figura 11).

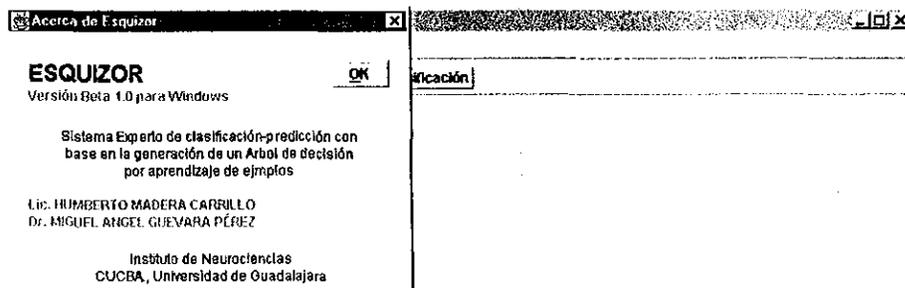


Figura 11. Información del programa.

El Menú ***Opciones*** permite elegir un método de Discretización (ver más adelante; Método de Discretización).

4c) EL PANEL DE ICONOS



Nuevo: sus funciones son iguales que las del submenú *Nuevo*.



Abrir: sus funciones son iguales que las del submenú *Abrir*.



Guardar: sin función en esta versión.



Cortar: sin función en esta versión.



Copiar: sin función en esta versión.



Pegar: sin función en esta versión.



Ayuda: sus funciones son como las descritas para *A cerca de...*

Aprendizaje

Aprendizaje: dispara los mecanismos de aprendizaje integrados en el programa, tomando como datos de entrada a los que se encuentran en la tabla desplegada por la carpeta Matriz en la subcarpeta Datos.

Clasificación

Clasificación: dispara los mecanismos de clasificación integrados en el programa, tomando como datos de entrada a los que se encuentran en la tabla desplegada por la carpeta Matriz en la subcarpeta Sujeto.

4d) EL PANEL PRINCIPAL

La ***Carpeta de Matriz*** contiene en su interior a dos subcarpetas: la de *Datos* y la de *Sujeto*.

La subcarpeta de *Datos* contiene la información de las variables y los sujetos implicados en el proceso de aprendizaje, estos datos se presentan por medio de una tabla en forma de matriz (ver Figura 9).

La subcarpeta de *sujeto* contiene la información del sujeto que se va a clasificar. Cuenta con una tabla que puede contener los datos del sujeto, sean estos ingresados directamente o tomados de una matriz con otros sujetos. También cuenta con dos áreas de texto: una de ellas contiene los resultados de la clasificación y la otra proporciona información relativa a la matriz de datos implicada en el proceso de aprendizaje la(s) variable(s) principal(es) involucradas en dicha clasificación (ver Figura 12).

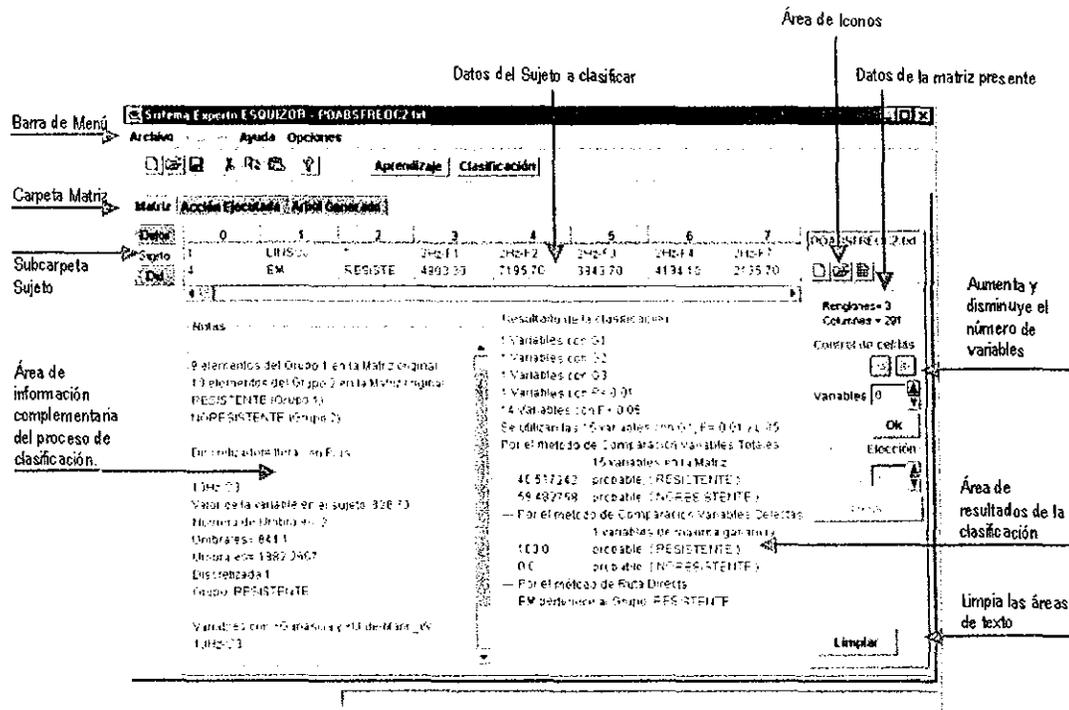


Figura 12. La Subcarpeta de Sujeto en la carpeta de Matriz.

La **Carpeta de Acción Ejecutada** presenta los datos de todo el proceso, tanto el de aprendizaje como el de clasificación. Informa sobre el número de sujetos y variables en la matriz con datos para aprendizaje, como las variables seleccionadas y sus valores. También cita las variables involucradas en la construcción del árbol y describe al árbol mismo, permitiendo la identificación de los sujetos involucrados en el entrenamiento. Define los umbrales especificados para la discretización de las variables y la posición de los datos del sujeto clasificado en las variables que constituyen el árbol (ver Figura 13).

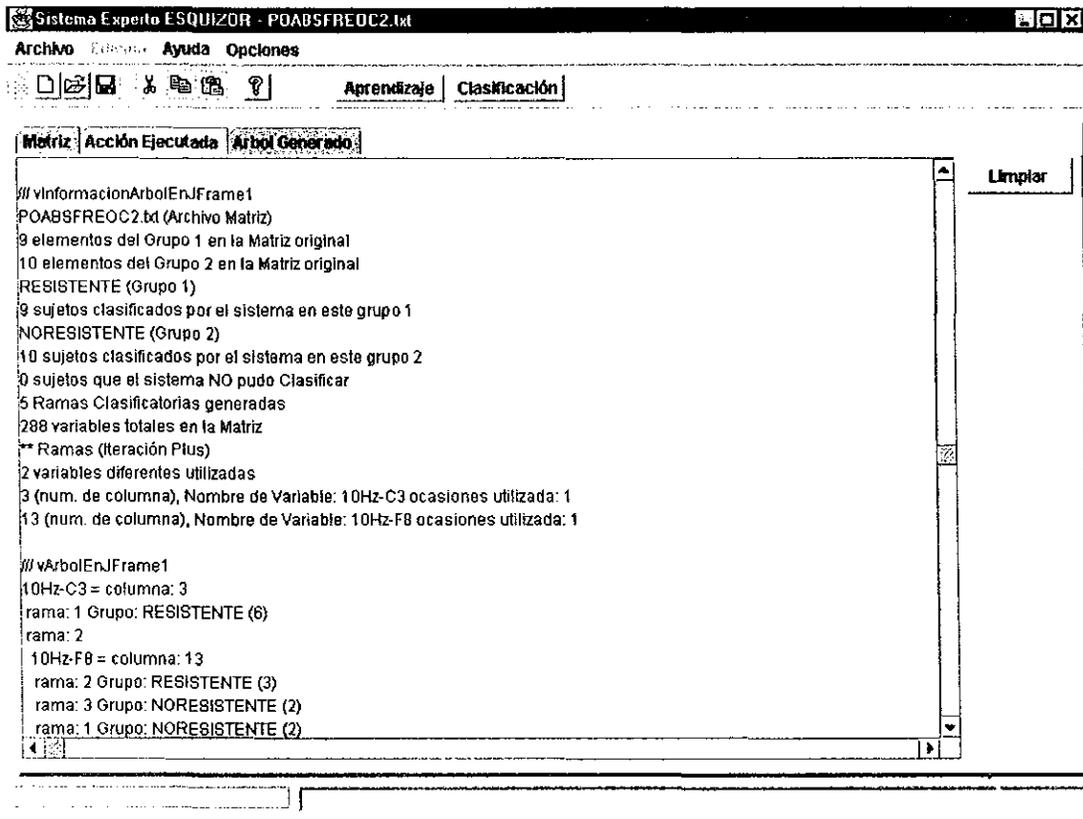


Figura 13. La Carpeta de Acción Ejecutada.

La **Carpeta de Arbol Generado** permite la visualización del árbol constituido por las variables involucradas en el aprendizaje y sus interrelaciones (ver Figura 14). Cita las variables en cada rama y en el tronco, así como la clasificación del grupo que ocupa la posición de la hoja final y el número de los sujetos en esa posición.

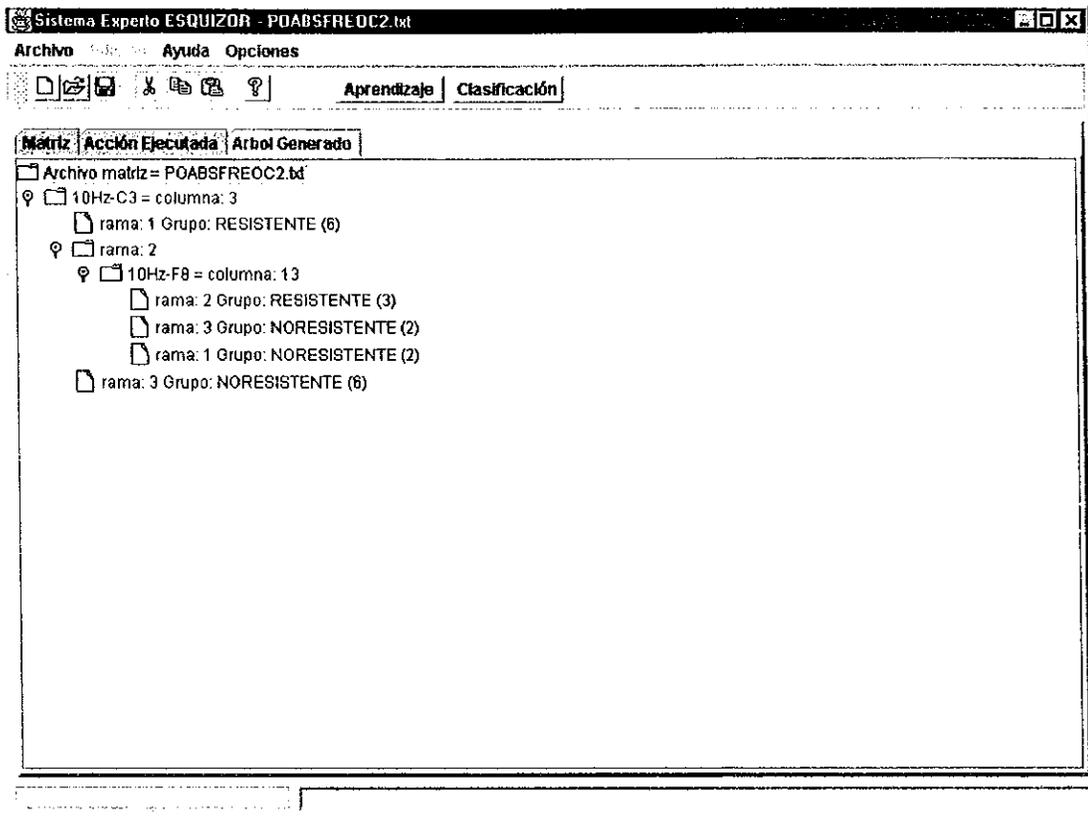


Figura 14. La Carpeta Arbol Generado.

VI.2 LOS SUJETOS

Participaron un total de 28 sujetos divididos en tres grupos independientes: una muestra de 9 pacientes esquizofrénicos Neuroleptoresistentes (RE); 10 pacientes esquizofrénicos No Resistentes (NR); y 9 sujetos Controles sin padecimiento psiquiátrico (CO). A la unión de los sujetos de los grupos RE y NR se les definió como un cuarto grupo: Grupo Esquizofrénico (EZ).

Los dos primeros grupos fueron seleccionados de una población del Centro Comunitario de Salud Mental No. 1 del IMSS.

Todos los grupos están formados por elementos del sexo masculino, apareados por edad, escolaridad y lateralidad (conforme a la prueba de Annett ,1967).

DE LOS SUJETOS QUE CONFORMARON LOS GRUPOS RE Y NR:

Fueron diagnosticados coincidentemente por dos psiquiatras, como esquizofrénicos Paranoides basados en los criterios internacionales vigentes del DSM IV y CIE 10, así como en Resistentes o No Resistentes, de conformidad con los criterios de esquizofrenia resistente de Keefe y cols. (1990) y Brenner y Merlo (1995) que incluyen los siguientes parámetros:

- Ausencia de mejoría después de 6 a 8 semanas de tratamiento.
- Requerimiento de una dosis de 40 mgs. ó más de Haloperidol por 24 horas.
- Cambios menores en 20% en la escala psiquiátrica breve (BPRS).
- Impresión clínica global mínima o mayor de dos.

Los dos grupos de pacientes se encontraban apareados, además, por fecha aproximada de inicio del padecimiento y el tipo de tratamiento.

CRITERIOS DE INCLUSIÓN

De los sujetos esquizofrénicos:

Pacientes esquizofrénicos (tipo paranoide) con tratamiento con neurolépticos típicos (haloperidol y decanoato de flufenazina, y un antiparkinsonico: biperiden).

- Diagnosticados por DSM-IV y CIE 10.
- Esquizofrénicos neurolepto-resistentes conforme a los criterios de Keefe y cols. (1990), y Brenner y Merlo (1995).
- Escolaridad mínima de 3ero. de secundaria.

De los sujetos control:

- Sin antecedentes heredo-familiares de enfermedad mental.
- Sin antecedentes de adicciones o enfermedades neuropsiquiátricas previas.
- Sin encontrarse patología al examen de funciones mentales por psiquiatría.

CRITERIOS DE EXCUSIÓN

- Pacientes esquizofrénicos no adherentes al tratamiento.
- Pacientes esquizofrénicos con familias disfuncionales.
- Pacientes esquizofrénicos con otra patología psiquiátrica o neurológica agregada.
- Antecedentes de terapia electroconvulsiva seis meses antes del tratamiento.

PROCEDIMIENTO

A los dos grupos de esquizofrénicos se les aplicaron por dos psiquiatras diferentes:

- La Escala Psiquiátrica Breve (BPRS).
- La Escala de Síntomas Positivos y Negativos (PANSS).
- La Escala de Neurolepto-Resistencia de Brenner.
- La prueba de lateralidad de Annett.

Durante el estudio los pacientes de los grupos RE y NR estuvieron sometidos a un régimen terapéutico farmacológico con neurolépticos típicos (haloperidol y/o decanoato de flufenazina) y un antiparkinsoniano (Biperiden o triexifenidil).

VI.3 LOS DATOS

1 Pruebas Psicopatológicas

Esta matriz contenía los datos de las pruebas de BPRS y PANSS, y sus datos se conformaron con los valores de los sujetos del par RE vs NR, en virtud de que las pruebas psicopatológicas se aplicaron únicamente a los sujetos esquizofrénicos.

Los resultados de las pruebas BPRS y PANSS, en unión con otros criterios, fueron utilizadas para identificar y clasificar a los sujetos RE y NR para esta prueba, por ello, la primera prueba que debió superar el Sistema fue clasificar correctamente a la totalidad de los sujetos contenidos en dicha matriz.

2 Prueba Wisconsin

En el presente estudio se incluyeron 7 de los resultados que genera el test:

1. Porcentaje de Respuestas Correctas
2. Porcentaje de Errores
3. Porcentaje de Respuestas Perseverativas
4. Porcentaje de Errores Perseverativos
5. Porcentaje de Errores No Perseverativos
6. Porcentaje de Respuestas Conceptuales
7. Categorías Completadas.

3 Electroencefalografía (EEG)

La actividad eléctrica espontánea (EEG) fue registrada en las derivaciones PF1, PF2, F3,F4,F7,F8,T3,T4,C3,C4,O1 y O2, referidas a las orejas cortocircuitadas de acuerdo al sistema 10/20 de colocación de electrodos. Se registraron los movimientos oculares como un control de artefactos.

Las señales de EEG, fueron registradas en un polígrafo Grass (filtros de 1 a 35 Hz), por medio de electrodos de oro.

La impedancia fue menor a 10 Kohms.

Los sujetos fueron registrados en reposo, con ojos abiertos y cerrados, durante 1.5 minutos en cada condición.

Las señales fueron capturadas en una computadora tipo PC, a través de un convertidor analógico/digital.

Se tomaron 30 muestras de 512 puntos, a una frecuencia de muestreo de 256 Hz y fueron analizadas por medio de la transformada rápida de Fourier, para obtener los valores de potencia absoluta (PA) y posteriormente, de potencia relativa (PR) en cada Hz, desde 2 a 25, y para cada banda:

- Delta, de 2 a 3.5 Hz.
- Theta1, de 4 a 5.5 Hz.
- Theta2, de 6 a 7.5 Hz.
- Alfa1, de 8 a 9.5 Hz.
- Alfa2, de 10 a 12.5 Hz.

- Beta1, de 13 a 17.5 Hz.
- Beta2, de 18 a 25 Hz.

Además, se obtuvieron índices de correlación inter (FP1-FP2; F3-F4; F7-F8; T3-T4; C3-C4; y O1-O2) e intrahemisférica (entre las diferentes áreas de cada uno de los hemisferios) para cada banda.

Los datos fueron constituidos con los resultados del análisis de la transformada rápida de Fourier realizados a los EEG tomados a cada sujeto.

- Potencia absoluta por Hz y banda (delta, theta1, theta2, alfa1, alfa2, beta1 y beta2) y por derivación.
- Potencia relativa para cada Hz y banda, y derivación.
- Índices de correlación inter e intrahemisférica.

VI.4 DEFINICION Y NUMERO DE VARIABLES

VARIABLES CONTROLADAS

Se realizó un pareamiento de cada sujeto de un grupo con sujetos de los restantes dos, tomando en cuenta el tipo de esquizofrenia paranoide, la edad, el sexo, lateralidad y medicación que están recibiendo los pacientes, así como la escolaridad y el tiempo de evolución de la enfermedad desde el primer brote.

VARIABLE INDEPENDIENTE

Esquizofrenia Paranoide neurolepto-Resistente, No Resistente, y sin padecimiento psiquiátrico.

VARIABLES DEPENDIENTES

PRUEBAS PSICOPATOLOGICAS

- Los resultados obtenidos de la aplicación de las pruebas BPRS y PANSS.

PRUEBA WISCONSIN.

- Los resultados obtenidos de la aplicación por computadora del *Wisconsin Card Sorting Test: Computer Versión 3 for Windows (WCST:CV3)*.

EEG

- La Potencia Absoluta, que es igual a la sumatoria del voltaje al cuadrado, de las frecuencias de cada banda del EEG.
- La Potencia Relativa: proporción con la que cada frecuencia contribuye a la potencia total del espectro, expresada en porcentaje.
- La Correlación Interhemisférica: la relación de fase de la actividad eléctrica cerebral entre zonas homólogas de los hemisferios cerebrales.
- La Correlación Intrahemisférica: la relación de fase de la actividad eléctrica cerebral entre diferentes zonas dentro del mismo hemisferio.

1 La conformación y la forma de las matrices de datos

- A. Con estos datos de las pruebas Psicopatológicas, Wisconsin y EEG se conformaron 18 Arreglos de Matrices con las siguientes características (ver Figura 15):
- Una primera columna con numeración progresiva que parte de cero en la casilla [1,1] hasta [n,1].
 - Una segunda columna que contiene el nombre o la clave de cada sujeto a partir de la casilla [3,2] hasta [n,2].
 - Una tercera columna que contiene el nombre del grupo al que pertenece cada sujeto, cuenta con un mínimo dos sujetos de grupos diferentes que pueden estar distribuidos en forma aleatoria y ordenada por cualquier criterio, a partir de la casilla [3,3] hasta [n,3].
 - Un primer renglón con numeración progresiva que parte de cero en la casilla [1,1] hasta [1,m].
 - Un segundo renglón que contiene el nombre de cada variable a partir de la casilla [2,4] hasta [2,m], y cuya disposición puede ser aleatoria u ordenada por cualquier criterio.
 - Los datos de cada uno de los sujetos (renglones) en cada una de las variables (columnas) desde la casilla [3,4] y hasta la [n,m].

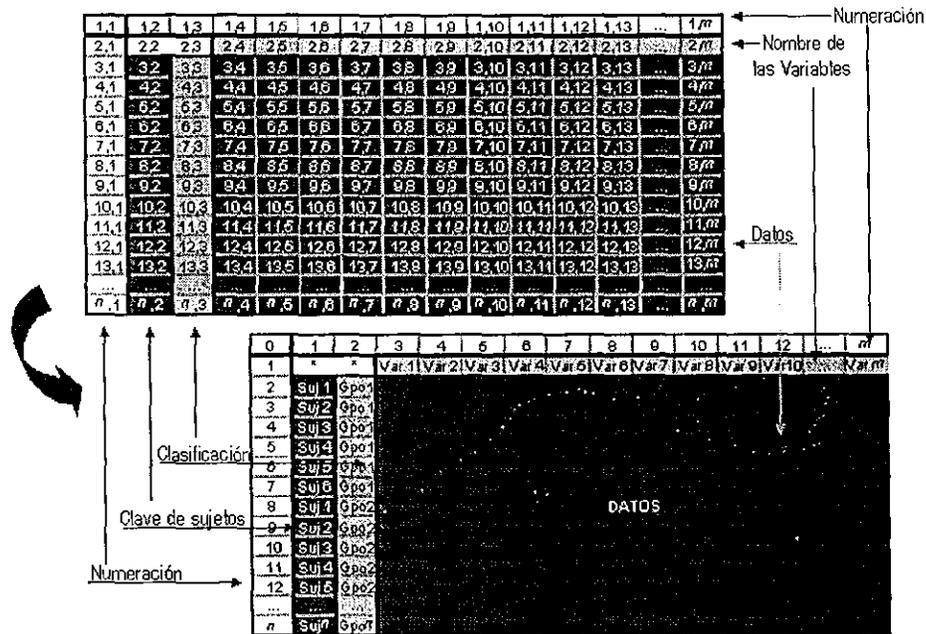


Figura 15. Distribución de los sujetos y las variables en una matriz.

B. Los dieciocho Arreglos de Matrices son los siguientes.

Pruebas Psicopatológicas

1. BPRS y PANSS (sus variables en una sola matriz):

- BPRS
- PANSS Síntomas Positivos
- PANSS Síntomas Negativos
- PANSS Psicopatología General
- PANSS Puntaje Total

Wisconsin

2. Los siete datos utilizados del Wisconsin
 - Porcentaje de RespuestasCorrectas
 - Porcentaje de Errores
 - Porcentaje de RespuestasPerseverativas
 - Porcentaje de ErroresPerseverativos
 - Porcentaje de ErroresNoPerseverativos
 - Porcentaje de RespuestasConceptuales
 - CategoríasCompletadas

EEG

3. De Potencia Absoluta por banda con Ojos Abiertos.
 4. De Potencia Absoluta por banda con Ojos Cerrados.
 5. De Potencia Relativa por banda con Ojos Abiertos.
 6. De Potencia Relativa por banda con Ojos Cerrados.
 7. De Correlación Interhemisférica por banda con Ojos Abiertos.
 8. De Correlación Interhemisférica por banda con Ojos Cerrados.
 9. De Correlación Intrahemisférica por banda con Ojos Abiertos
 10. De Correlación Intrahemisférica por banda con Ojos Cerrados
 11. De Potencia Absoluta por frecuencia con Ojos Abiertos.
 12. De Potencia Absoluta por frecuencia con Ojos Cerrados.
 13. De Potencia Relativa por frecuencia con Ojos Abiertos.
 14. De Potencia Relativa por frecuencia con Ojos Cerrados.
 15. De Correlación Interhemisférica por frecuencia con Ojos Abiertos.
 16. De Correlación Interhemisférica por frecuencia con Ojos Cerrados.
 17. De Correlación Intrahemisférica por frecuencia con Ojos Abiertos
 18. De Correlación Intrahemisférica por frecuencia con Ojos Cerrados
- C. Por cada Arreglo de Matriz se construyeron cuatro matrices ya que cada una puede contener únicamente dos grupos de sujetos diferentes, ello genera cuatro tipos de pares diferentes.
- CO vs EZ: Que contiene los datos de todos los sujetos Control y los de todos los sujetos Esquizofrénicos (Resistentes y No resistentes como parte de un único grupo).

- RE vs NR: Con los datos de los sujetos Resistentes y No Resistentes.
- CO vs RE: Con los datos de los sujetos Controles y los Resistentes.
- CO vs NR: Con los datos de los sujetos Controles y los No Resistentes.

Los cuatro tipos de pares en cada tipo de Arreglo de Matriz: EEG y Pruebas Psicopatológicas, más el de Wisconsin que tiene únicamente un par; RE y NR da como resultado un total de 69 matrices diferentes.

Cada uno de los Arreglos de Matrices anteriormente citadas en B. contenían el siguiente número de variables: BPRS-PANSS, 5 variables; Wisconsin, 7 variables; las de potencias Absoluta y Relativa por banda, 84 cada una y por frecuencia, 288; las de Correlación Interhemisférica por banda, 42 cada una y por frecuencia, 144; las de Correlación Intrahemisférica por banda, 210 cada una y por frecuencia, 720; todas las anteriores tanto en Ojos Abiertos como Cerrados, lo que da una suma un total de 3732 variables.

VI.5 LAS PRUEBAS (GENERALIDADES)

Para probar el sistema se utilizó la técnica de la *intervalidación* que trata de calcular qué tan bien la hipótesis actual podrá predecir datos no vistos (Russell y Norvig, 1996). Para ello, en cada matriz, separamos los datos un sujeto del resto de los datos conocidos, y los utilizamos para probar el desempeño en la predicción de una hipótesis inducida a partir del resto de los datos conocidos, repitiendo el proceso con cada sujeto restante.

1 Conformación de los grupos de Capacitación y de Prueba

Conforme a la técnica de *intervalidación* implementada para nuestras pruebas, definimos como “Conjunto de capacitación” el que está formado por los datos de los cuales el sistema extraerá las reglas (formará el árbol) con las que clasificará a un sujeto desconocido o grupo de ellos (denominados “Conjunto de prueba”).

En el caso de este sistema, y para calcular el estimado de *Error Verdadero*, fue necesario que en cada matriz se conformaran tantos conjuntos de capacitación y de prueba como sujetos contenía. Para ello se siguieron los siguientes pasos.

1. Se seleccionaron los datos de un sujeto en uno de los grupos de pares: CO vs EZ; RE vs NR; CO vs RE; CO vs NR.
2. El sujeto seleccionado fue extraído de dicha matriz.
3. Se integró una nueva matriz con el resto de los sujetos, la que se denominó "Conjunto de Capacitación" (sin selección de variables).
4. El sujeto extraído se consideró "Prueba".
5. Después de dicha prueba se iniciaba la construcción de otro Conjunto de capacitación y otro de prueba, extrayendo un sujeto diferente de la matriz original e iniciando los pasos antes descritos.

2 Pruebas generales de aprendizaje: estimado de Error Aparente

Las pruebas generales de aprendizaje se pueden definir como la evaluación que se realiza del programa con base en los resultados emitidos por este al procesar una matriz de datos, esto es, el árbol generado a partir de un conjunto de sujetos de dos grupos diferentes de los cuales conoce su pertenencia y que conforman el Conjunto de Capacitación.

Los sistemas expertos basados en la generación de árboles de decisión por aprendizaje de ejemplos parten de un estimado error aparente de cero, ya que los ejemplos utilizados en su matriz de aprendizaje o Conjunto de Capacitación son correctamente clasificados para generar el árbol.

Sin embargo, un valor de cero de *error aparente* no garantiza que el sistema continúe manteniendo el mismo valor de error al ser probado en la clasificación de ejemplos no incluidos en la muestra de su matriz de aprendizaje, incluso, el valor de error estimado en el caso de la clasificación de nuevos sujetos (llamado *error verdadero*) tiende a ser mayor que el de *error aparente* (Weiss, y Kulikowski, 1991).

La finalidad de estas pruebas es lograr que el programa alcance los siguientes objetivos:

1. Clasificar correctamente en sus grupos a la totalidad de los sujetos proporcionados y
2. Producir el árbol más compacto posible, esto es con el menor número de ramas y variables utilizadas para evitar la dispersión de los datos.

El estimado de Error Aparente consiste en el número de sujetos pertenecientes al Grupo de Capacitación no clasificados por el sistema divididos entre el número total de sujetos en el citado grupo. Para efectos de manejo en el presente estudio, el resultado de la operación se multiplicará por 100 a fin de considerarlo cual si fuera un porcentaje.

3 El estimado del Error Verdadero y error por grupo.

Una vez conformado un Conjunto de Capacitación y su contraparte Conjunto de prueba, el sistema realizó la clasificación del sujeto Prueba con los datos aportados por el árbol generado con el Conjunto de Capacitación.

Cada resultado de clasificación fue contabilizado como Acierto o Falla, al comparar la clasificación del sistema con la verdadera pertenencia del sujeto a uno de los dos grupos de la matriz.

Se realizaron tantas pruebas a cada grupo como sujetos contuviera, esto es:

- En las matrices del par Resistentes y No Resistentes, se realizaron 9 y 10 clasificaciones respectivamente: 19 en total.
- En las matrices del par Controles y Resistentes, se realizaron 9 y 9 clasificaciones respectivamente: 18 en total.
- En las matrices del par Controles y No Resistentes, se realizaron 9 y 10 clasificaciones respectivamente: 19 en total.
- En las matrices del par Controles y Esquizofrénicos, se realizaron 9 y 19 clasificaciones respectivamente: 28 en total.

El Estimado del Error Verdadero se realizó de la siguiente manera: Se contabilizaron las fallas en la clasificación de todos los sujetos para cada matriz y se convirtió el resultado en un porcentaje.

Sin embargo, también se calculó el porcentaje de error al clasificar a los sujetos de cada grupo (por separado).

VI.6 PRUEBA 1: DESEMPEÑO BASICO CON VARIABLES DISCRETAS.

1 Objetivo

Probar el funcionamiento del Módulo Formador del árbol, del prototipo del Sistema con base en un ejemplo publicado cuyos datos sean discretos.

2 Hipótesis

El prototipo producirá un árbol de decisión idéntico al ejemplo del que se tomaron los datos.

3 Metodología

3a) DATOS

Los datos se tomaron del ejemplo presentado en Russell y Norvig, (1996 pp. 565). Son datos de 12 sujetos (de X1 a X12), discretos, distribuidos en 10 variables (atributos) acomodadas en columnas, más una columna que contiene la clasificación de cada ejemplo (meta esperarán), mismos que se reproducen en la Tabla 1.

Este ejemplo fue diseñado para la construcción de un Sistema Experto que por medio de un árbol de decisión pudiese predecir si un cliente va a decidir esperar una mesa en un restaurante o no. Los datos pertenecen a dos grupos de sujetos 6 de los cuales decidieron esperar. Las variables contienen información referente a las circunstancias siguientes:

1. *Alternativa:* Si algún restaurante cercano ofrece alguna alternativa adecuada.
2. *Bar:* Si el restaurante tiene área de bar cómoda en donde esperar la asignación de la mesa.
3. *Viernes/Sábado:* Es válido en viernes y sábado.
4. *TenerHambre:* Si tenemos hambre.
5. *Clientes:* Cuantas personas hay en el restaurante (los valores son *Ninguno, Algunos, EstáLleno*).

- 6. *Precio:* El rango de los precios del restaurante (\$, \$\$, \$\$\$).
- 7. *Lluvia:* Si está lloviendo afuera.
- 8. *Reservación:* Si se hizo una reservación.
- 9. *Tipo:* El tipo de restaurante (*francés, italiano, tailandés o de hamburguesas*).
- 10. *TiempoEsperaEstimado* Tiempo de espera estimado por el anfitrión (0-10 minutos, 10-30, 30-60, o más de 60).

Tabla 1. Datos del ejemplo presentado por Russell y Norvig, (1996 pp. 565).

Ejemplo	ATRIBUTOS										Meta Esperarán
	Alternativa	Bar	Viernes/Sábado	Tener/Hambre	Clientes	Precio	Lluvia	Reservación	Tipo	Tiempo/Espera	
X1	Si	No	No	Si	Algunos	\$\$\$	No	Si	Francés	0-10	Si
X2	Si	No	No	Si	Lleno	\$	No	No	Tailandés	30-60	No
X3	No	Si	No	No	Algunos	\$	No	No	Hamburguesa	0-10	Si
X4	Si	No	Si	Si	Lleno	\$	No	No	Tailandés	10-30	Si
X5	Si	No	Si	No	Lleno	\$\$\$	No	Si	Francés	> 60	No
X6	No	Si	No	Si	Algunos	\$\$	Si	Si	Italiano	0-10	Si
X7	No	Si	No	No	Ninguno	\$	Si	No	Hamburguesa	0-10	No
X8	No	No	No	Si	Algunos	\$\$	Si	Si	Tailandés	0-10	Si
X9	No	Si	Si	No	Lleno	\$	Si	No	Hamburguesa	> 60	No
X10	Si	Si	Si	Si	Lleno	\$\$\$	No	Si	Italiano	10-30	No
X11	No	No	No	No	Ninguno	\$	No	No	Tailandés	0-10	No
X12	Si	Si	Si	Si	Lleno	\$	No	No	Hamburguesa	30-60	Si

3b) PROCEDIMIENTO

Con los datos de la tabla anterior se elaboró un archivo de texto que los contenía en forma de matriz bajo el nombre *Datos.txt*.

El acomodo de los datos en la matriz fue adecuado al diseño requerido por el sistema, esto es la columna de clasificación (Meta esperarán), fue colocada al lado derecho de la columna de ejemplos.

Una vez realizado lo anterior, el archivo fue leído por el sistema y ejecutadas las funciones de aprendizaje (formación del árbol con el algoritmo de Ganancia de Información para elegir las variables).

El árbol generado fue comparado con el árbol descrito por Russell y Norvig (1996 pp. 565) ver Figuras 16 y 17 para calificar el grado de similitud.

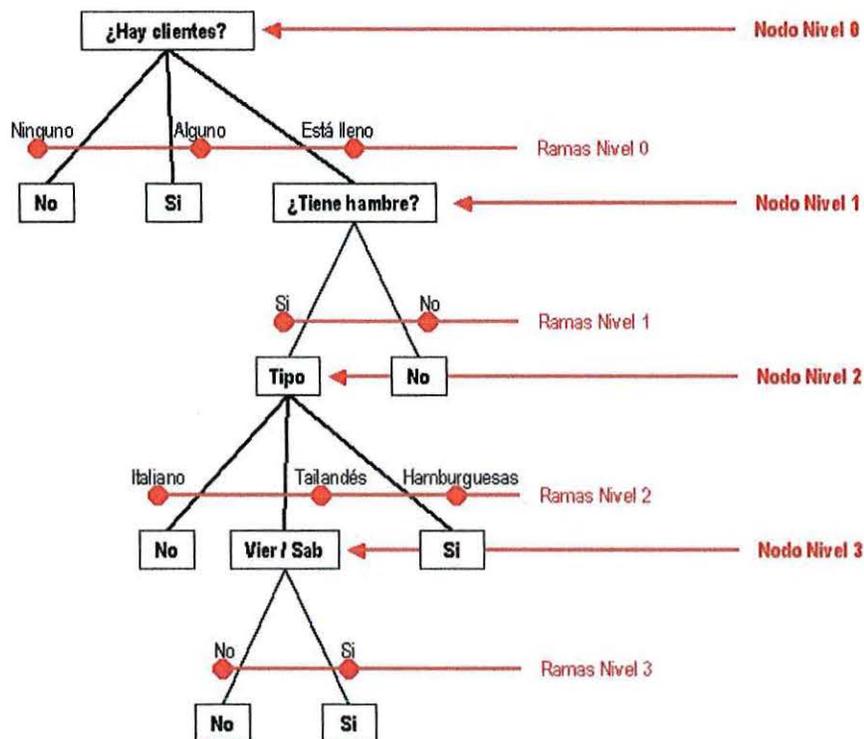


Figura 16. Arbol tomado de Russell y Norvig (1996 pp. 565). Las líneas, círculos y texto rojos han sido agregados para clarificar los Nodos, las Ramas y sus niveles. Cada nodo es una variable y las ramas que se desprenden de la misma representan los diferentes valores que los sujetos de la matriz tienen en dicha variable. El nivel de profundidad de un árbol se determina por el número de niveles de nodos y su amplitud por el promedio de ramas por nodo.

El grado de similitud se evaluó con relación a:

1. Las variables que fueron elegidas en cada nivel: un punto por cada variable correcta (en nombre y nivel). 4 puntos posibles.
2. Las ramas por variable: un punto por cada rama correcta (en nombre y nivel). 10 puntos posibles.
3. La distribución general del árbol: un punto por cada relación Rama-Variable siguiente. 3 puntos posibles.

Lo anterior da un total de 17 puntos posibles para un 100% de similitud.

4 Resultados

El árbol generado por el sistema obtuvo un 100% de similitud al compararse con el descrito por Russell y Norvig (1996 pp. 565), ver la Figura 17. La única diferencia encontrada es su orientación horizontal.

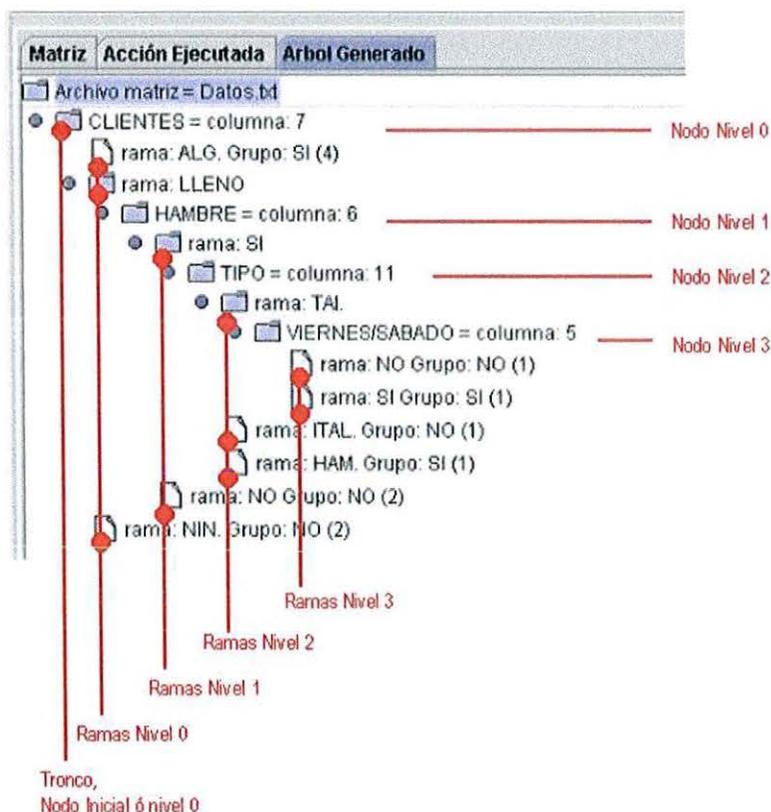


Figura 17. Árbol generado por el Sistema. Las líneas, círculos y texto rojos son nuestros.

5 Discusión y Conclusiones

La presente prueba permitió demostrar que la implementación del Módulo Formador del árbol en el Sistema, funciona adecuadamente y que genera árboles correctos cuando los datos ingresados son discretos.

Sin embargo, los datos del EEG son valores continuos, que aún con la digitalización siguen siendo diferentes de sujeto a sujeto en una misma variable, lo que impide la formación de subgrupos, facilita la sobre-especialización y por ende inhabilita al sistema para extraer las características de cada grupo.

Es importante definir una metodología para discretizar los valores del EEG antes de generar el árbol, para lo cual se realizó la siguiente prueba.

VI.7 PRUEBA 2: METODOS DE DISCRETIZACION PARA DATOS DE EEG DIGITALIZADO.

1 Objetivo

Diseñar un sistema de discretización de los datos del EEG, que permita conservar lo más posible las características que identifican a cada sujeto con su grupo de pertenencia y al mismo tiempo discriminen a los grupos entre sí.

2 Objetivos Específicos

1. Definir el(los) método(s) de discretización para los datos del EEG.
2. Evaluar el valor de Ganancia alcanzado con dicho(s) método(s) y con los datos originales a fin de poder elegir el que tenga mayor capacidad de discriminación grupal.

3 Hipótesis

Si la distribución de los datos en la matriz refleja la diferencia entre los sujetos de los dos grupos que la conforman, esta permitirá encontrar principalmente a los sujetos de dos grupos que son distintos en ambos extremos de un continuo. Ello a su vez permitirá, encontrar uno o dos puntos en los que estos dos grupos se dividan y se separen sus mayorías, lo que nos llevará a convertir los valores de la mayoría de un grupo en un solo valor igual y los valores del otro grupo en otro valor único pero diferente del primero, con lo que se preservarían la mayoría de las cualidades identificadoras de cada grupo.

4 Metodología

4a) DATOS.

Se conformó una matriz con los datos de dos grupos de sujetos: los 9 Resistentes (RE) y los 10 No Resistentes (NR).

Se incluyeron en dicha matriz todas las variables de Potencias (Absoluta y Relativa, por bandas y frecuencias y en condición ojos abiertos y cerrados) y Correlaciones (Inter e Intrahemisféricas, por bandas y frecuencias, y en condiciones ojos abiertos y cerrados)

Debido a la cantidad de variables, se decidió reducir el número con base en su capacidad discriminativa.

Se realizó una prueba U de Mann Whitney en cada variable para identificar aquellas con significancia estadística ($p < 0.01$, eliminando todas las que tuviesen puntajes mayores a este.

La matriz de datos quedó lista para las pruebas con 36 variables únicamente.

4b) PROCEDIMIENTO.

Se definieron 5 métodos de discretización:

1. **Mediana:** Se determina la mediana de todos los datos de la variable y se asigna el valor "1" a los datos por debajo de ella y "2" a los que se encuentran por encima de la misma.
2. **Q1, H, Q4:** Este método determina los valores cuartiles de la variable y utiliza el valor del primer cuartil, y del tercer cuartil para discretizar los datos de los sujetos. Convierte en "1" los datos que son iguales o inferiores al primer cuartil, en "2" los que sean mayores que el primero y menores que el tercer cuartil y en "3" aquellos que sean mayores o iguales al tercer cuartil.
3. **Terciles:** En este trabajo definimos terciles a los tres bloques en los que se divide un conjunto de datos siguiendo algunas de las convenciones para los cuartiles. Esto implica ordenar los datos de la variable de menor a mayor y después determinar el dato que ocupa la posición correspondiente al límite superior del primer tercio y límite inferior del tercer tercio del total de los mismos. Estos dos datos escogidos representan los umbrales de discretización.

En el caso de que el número total de datos sea múltiplo de 3, ese límite corresponda a un solo dato en cada 1/3 como se describió y por ello se toman los valores de los datos que ocupan esas posiciones.

Pero en caso contrario, cuando el número total es igual a un múltiplo de 3+1 o de 3+2, es necesario distribuir los 6/6 de esos dos números entre las 3 partes en las que se dividió el total de los datos.

Para el límite del primer 1/3; se suma el valor de los datos adyacentes y dividiendo entre tres, se suman dos tercios del resultado al valor del dato anterior al límite y uno al primer dato posterior a él.

Para el límite del segundo 1/3, igualmente se suman los datos adyacentes y después de la división se suman dos tercios del resultado al primer valor posterior al él y uno al primer valor anterior.

Para la discretización se asigna un "1" a los datos inferiores o iguales al del límite superior del primer 1/3, un "2" a los datos superiores del límite del primer 1/3 e inferiores al límite superior del segundo 1/3, y un "3" a los datos iguales o superiores al valor de este límite superior del segundo 1/3.

4. **Mediana por grupo:** Se separan los datos de los sujetos que pertenecen a cada uno de los dos grupos en la matriz y se determina la mediana de cada grupo. Se identifica la mediana de mayor y la de menor valor. Se asigna el valor de "1" a los datos que tengan valores menores o iguales a la mediana de menor valor, se asigna "2" a los valores que se encuentren entre los valores de ambas medianas y finalmente se asigna el valor de "3" a los valores iguales o superiores a la mediana mayor.
5. **Iteración por Variable:** En este método se conjugan las características de los otros cuatro, ya que en cada variable a discretizar se corren los cuatro métodos anteriores y después de cada uno se mide el valor de la Ganancia de Información, manteniendo la discretización con el método que haya alcanzado el mayor puntaje. Esto permite que la variable una vez discretizada mantenga la mayor relación con sus datos originales y su capacidad discriminativa entre grupos.

Para valorar el Sistema se utilizó la técnica de *Interválidación* que consiste en la extracción de los datos de un sujeto del total de sujetos disponibles, la elaboración del árbol de decisión con los datos del resto de los sujetos y posteriormente, la clasificación del sujeto extraído con base en dicho árbol. El proceso se repite hasta que todos los sujetos disponibles han sido clasificados de esta manera. Con ello se puede probar el funcionamiento del problema con sujetos desconocidos para él, cuando se tiene una muestra pequeña.

Por tanto, se corrieron 19 pruebas en la matriz utilizada, con los 5 diferentes métodos de discretización lo que resulta en un total de 95 pruebas.

Se obtuvo el *Estimado de Error Verdadero* con cada uno de los métodos de discretización. Dicho *Estimado* representa el porcentaje de fallas del Sistema en la clasificación de sujetos desconocidos para él. Consiste en el conteo de los aciertos y las fallas y la transformación de estas últimas en porcentaje. El *Error Verdadero* permite un adecuado juicio de la efectividad del Sistema para clasificar adecuadamente sujetos desconocidos.

La identificación del método de discretización que permite una adecuada clasificación se requiere que:

1. El porcentaje de error en el grupo de Resistentes se encuentre por debajo del 50%,
2. que el porcentaje de error en el grupo de No Resistente se encuentre por debajo del 25%, y
3. que el porcentaje de *Error Verdadero* se encuentre por debajo del 37.5% y que además los porcentajes de error en los grupos cumplan con los puntos 1 y 2.

5 Resultados

La Tabla 2 presenta los resultados de la valoración del sistema después de la discretización con cada una de los métodos descritos.

Tabla 2. Resultados generales de las pruebas del Sistema. Presenta el *Estimado de Error Verdadero* en porcentaje (primer número en las casillas) y el porcentaje de error en los grupos (siguientes dos renglones en cada casilla RE= Resistentes, NR= No Resistentes) La celda gris indica el método que obtuvo el mejor porcentaje de error.

Matriz conformada por RE y NR con una selección de 36 var.	
% error	Método
31.57 RE = 44.4 NR = 20	Mediana
26.31 RE = 22.2 NR = 30	Q1, H, Q4
26.31 RE = 22.2 NR = 30	Terciles
26.31 RE = 0.0 NR = 50	Mediana por Grupo
26.31 RE = 44.4 NR = 10	Iteración por Variable

Solo dos métodos cumplen con los tres puntos citados para considerar un método de discretización como adecuado: el método de la mediana y el de iteración por variables.

6 Discusión y Conclusiones

El mejor método para la discretización es la Iteración por Variable, primeramente porque al igual que el método de la mediana, cumple con los tres requisitos antes citados y porque obtiene un menor porcentaje de *Error Verdadero*.

Si recordamos como se estructura el método de la Iteración por Variable, estaremos ciertos que cada variable fue discretizada por cada uno de los cuatro primeros métodos numerados en la tabla 2 y se eligió como discretización definitiva a la que alcanzó la mayor ganancia. Esto quiere decir que el método de Iteración significa elegir el mejor de los cuatro anteriores métodos, lo que se refleja en los porcentajes de error, sin embargo, aún es muy alto el número de errores en el grupo de Resistentes.

Estos porcentajes pueden disminuirse si el número de variables que ingresan a formar el árbol son seleccionadas de manera más estricta. Tal parece que el número de variables candidatas a formar el tronco del árbol en muchas de las ocasiones son más de 2.

Es necesario implementar sobre los procesos ya estructurados una metodología de selección más estricta para las variables que son utilizadas para la formación del árbol. Con ello, al reducir las variables candidatas a formar el tronco, y considerando que la variable elegida para ese nodo principal alcanza la mayor discriminación, el número de errores será disminuido.

La siguiente prueba fue diseñada para implementar y probar una selección de variables.

VI.8 PRUEBA 3: PRUEBAS FINALES DE FUNCIONAMIENTO DEL SISTEMA.

1 Objetivo

Integrar al Sistema un módulo que realice una preselección y limite el número de variables en la matriz que ingresan a formar el árbol.

2 Hipótesis

El realizar una preselección de las variables en la matriz utilizada, dejando únicamente las que diferencian en mayor medida a los dos grupos de sujetos, permitirá reducir la sobreadaptación evitando que el Sistema encuentre *regularidades* irrelevantes. De esta forma se tomarán en consideración variables que tienen que ver realmente con la discriminación de los grupos. Igualmente esto evitará que más de una variable pase a formar parte del tronco del árbol y aumentará el nivel de acertividad del Sistema.

3 Metodología

3a) SUJETOS

Participaron los 28 sujetos definidos anteriormente y divididos en los tres grupos independientes ya citados 9 RE; 10 NR; y 9 CO. A la unión de los sujetos de los grupos RE y NR se les definió como un cuarto grupo: Grupo Esquizofrénico (EZ).

3b) DATOS

Se utilizaron los resultados de las pruebas psicopatológicas BPRS y PANSS, los del Wisconsin y los datos del EEG.

3c) PROCEDIMIENTO

Para integrar las matrices se conformaron 4 diferentes pares de grupos: RE y NR; CO y RE; CO y NR; CO y EZ.

El par RE y NR fue el único que contó con matriz de datos de pruebas psicopatológicas (BPRS y PANSS).

Todos los pares contaron con matriz de datos de Wisconsin y de EEG.

En el caso del EEG, se conformaron 16 matrices para cada par, 8 con datos por bandas y 8 por frecuencias (Potencia Absoluta y Relativa, Correlación Intrer e Intrahemisférica; separadas en las condiciones ojos Abiertos y Cerrados).

Al sistema se le incluyó un subprograma que toma cada variable de la matriz y calcula su valor de p con base en la prueba U de Mann Whitney.

Para las pruebas del Sistema y para el cálculo del *Error Verdadero* (incluyendo los Errores por Grupo) se utilizó la técnica de la *Interválidación* y para ello, una vez definido un Conjunto de Capacitación en cada Matriz se llevaron a cabo los siguientes pasos:

1. Se realizó una **selección de variables** en la matriz Conjunto de Capacitación con base en: los resultados de la prueba U de Mann Whitney y el valor de Ganancia, ambos por cada variable. En dicha selección se utilizaron los siguientes criterios de inclusión:
 - Que la variable seleccionada tuviese un valor de $p < 0.01$, y/o
 - Que tuviese el valor de Ganancia más alto, ó
 - Que tuviese un valor de $p < 0.05$ (sin la presencia de otras variables con valores de $p < 0.01$, o que el número de variables con valores menores de p en conjunción con el número de variables con la Ganancia más alta no excediera de 2.)

De esta forma el número de variables en la matriz fue disminuído en la medida de lo posible pero sin ser menor de 3, con el fin de mantener en ella únicamente a las variables con mayores probabilidades de diferenciar a los grupos.

2. Una vez conformada la matriz con las variables seleccionadas, cada variable contenida en ella fue discretizada por cada uno de los 4 métodos descritos anteriormente y fue medido su respectivo valor de Ganancia, eligiendo el método que generó la mayor ganancia y el menor número de ramas.
3. La nueva matriz ya discretizada se formó únicamente con las variables seleccionadas.
4. El Conjunto de Capacitación de esta forma definido fue útil para probar la clasificación del único sujeto que conformó su Conjunto de Prueba, por lo que el proceso se repite con un sujeto diferente hasta haber probado al sistema con cada uno de ellos.
5. Sin embargo, si durante la selección de las variables no se encuentra por lo menos una con $p < 0.05$, la matriz en turno se considera eliminada de las pruebas y no se prosigue con su utilización, ésto sin importar cuántos sujetos se tengan probados o falten por hacerlo en la matriz.

La técnica de la *Interválidación* implicó, en forma general, que para cada prueba, se tomasen los datos completos de la matriz en turno, se extrajese de ella los datos completos de un sujeto (borrándolos de la matriz), se eligiesen las variables que mejor discriminaran los grupos, con base en los datos de los sujetos que quedaron, se formase el árbol con base en estas variables selectas y se realizara la clasificación del sujeto extraído con base en este árbol producido.

El número total de pruebas realizadas (sin contar las matrices eliminadas del proceso) fue de 954, distribuidas de la siguiente manera: 285 en el par RE y NR; 216 en el par CO y RE; 285 en el par CO y NR; 168 en el par CO y EZ.

4 Resultados

4a) DEL PROCESO DE CONSTRUCCIÓN DEL SISTEMA

La introducción de una prueba U de Mann Whitney después de que el algoritmo de ganancia evalúa una variable permitió identificar aquellas que presentan

mayor diferencia entre grupos, y esta información relacionada con el valor de la ganancia disminuyó hasta en un 50% el número de las variables candidatas a formar el tronco o las ramas del árbol, sin incidencia negativa importante sobre la efectividad del sistema.

La *Iteración por variable* en conjunción con el nuevo método de selección de variables disminuyó el porcentaje de error en las matrices de mayor utilidad para la clasificación hasta permitir un rango aceptable del 21% en la clasificación de RE y NR con EEG y 5% con Wisconsin. Las matrices que no alcanzaron este porcentaje fueron consideradas sin información suficiente para ser utilizadas con fines de clasificación en este trabajo.

4b) DEL PROCESO DE APRENDIZAJE DEL SISTEMA

Este sistema clasificó correctamente a todos los sujetos contenidos en sus Conjuntos de Capacitación durante la formación de los árboles de decisión, por tanto, el valor del *error aparente* es cero.

4c) DEL PROCESO DE PRUEBA DEL SISTEMA

En forma general, los resultados del proceso de prueba del sistema, expresados como estimado de error verdadero (convertido a porcentaje) se presentan en la Tabla 3.

La información más importante con relación a esta tabla se encuentra en el apartado correspondiente a cada par (RE y NR; CO y RE; CO y NR; CO y EZ) en los resultados de la prueba con datos de EEG más adelante. El desglose particular que cita las variables implicadas y sus frecuencias de aparición se encuentran en el apartado B de los Anexos.

Tabla 3. Resultados generales de las pruebas del Sistema. Resultados del *Estimado de Error Verdadero* en porcentaje (primer número en las casillas) y el porcentaje de error en los grupos (siguientes dos renglones en cada casilla). El “*” indica que por lo menos en un sujeto de esa matriz no se encontró ninguna variable con diferencia intergrupala significativa, y que no se valoró dicha matriz. Las celdas grises indican las matrices que obtuvieron menor porcentaje de error.

RE y NR		CO y RE		CO y NR		CO y EZ		
Banda	Frec.	Banda	Frec.	Banda	Frec.	Banda	Frec.	
32 RE = 44 NR = 20	63 RE = 56 NR = 70	*	44 CO = 56 RE = 33	*	47 CO = 33 NR = 60	*		Potencia Absoluta con ojos Abiertos
37 RE = 33 NR = 40	21 RE = 22 NR = 20	*	*	16 CO = 11 NR = 20	47 CO = 44 NR = 50	*		Potencia Absoluta con ojos Cerrados
32 RE = 33 NR = 30	21 RE = 33 NR = 10	28 CO = 11 RE = 44	44 CO = 33 RE = 55	53 CO = 44 NR = 60	42 CO = 44 NR = 40	46 CO = 55 EZ = 42	Pruebas no efectuadas	Potencia Relativa con ojos Abiertos
32 RE = 44 NR = 20	32 RE = 33 NR = 40	28 CO = 11 RE = 44	33 CO = 22 RE = 44	32 CO = 0 NR = 60	47 CO = 33 NR = 60	46.42 CO = 89 EZ = 26		Potencia Relativa con ojos Cerrados
*	68 RE = 67 NR = 70	78 CO = 67 RE = 89	55 CO = 33 RE = 78	21 CO = 11 NR = 30	32 CO = 44 NR = 20	25 CO = 44 EZ = 16		Correlación INTER. con ojos Abiertos
79 RE = 89 NR = 70	57 RE = 44 NR = 70	*	28 CO = 33 RE = 22	37 CO = 44 NR = 30	37 CO = 33 NR = 40	41 CO = 67 EZ = 28		Correlación INTER. con ojos Cerrados
63 RE = 44 NR = 80	53 RE = 44 NR = 60	61 CO = 55 RE = 67	22 CO = 22 RE = 22	21 CO = 22 NR = 20	26 CO = 22 NR = 30	43 CO = 67 EZ = 32		Correlación INTRA. con ojos Abiertos
58 RE = 78 NR = 40	58 RE = 67 NR = 60	61 CO = 44 RE = 78	72 CO = 78 RE = 67	63 CO = 55 NR = 80	47 CO = 33 NR = 60	58 CO = 78 EZ = 42		Correlación INTRA. con ojos Cerrados

Con Pruebas Psicopatológicas

El sistema Esquizor identificó y clasificó sin errores a todos los sujetos integrantes de la matriz de pruebas psicopatológicas RE y NR. En esta prueba se obtuvo el 0 % de error verdadero en la matriz, identificándose como las variables principales más utilizadas a:

- BPRS (en todas las pruebas: 20/20)
- PANSS Positivo (+) (en 14/20 pruebas)
- PANSS Total (t) (en 2/20 pruebas)

Así mismo, cuando los datos de todos los sujetos en la matriz fueron pasados por el sistema para generar el árbol, las variables seleccionadas como tronco fueron:

- BPRS
- PANSS Positivo (+)

Con Pruebas Neuropsicológicas

Las pruebas realizadas con la matriz de datos del WISCONSIN arrojaron los siguientes resultados:

1. **En el par RE VS NR:** Se obtuvo un 5% de error verdadero, con un 10% de fallas en la clasificación de los sujetos RE y 0% en la de los sujetos NR. La variable identificada como principal tanto a lo largo de las 19 pruebas como con la totalidad de los sujetos fue:
 - Categorías Completadas

Un análisis más minucioso de la distribución de los valores de cada grupo en la variable escogida para el tronco del árbol (*Categorías completadas*) en la matriz de datos de Wisconsin, en el par RE vs NR (ver Figura 18), permitió identificar la presencia de dos bloques con 6 y 7 sujetos pertenecientes a los grupos NR y RE respectivamente, completamente diferenciados, mientras que un tercer bloque con 7 sujetos sin diferenciar. El sistema se vale de las variables *% de Errores Perseverativos* y *% de Respuestas Perseverativas* para clasificar correctamente a los sujetos de este último bloque.

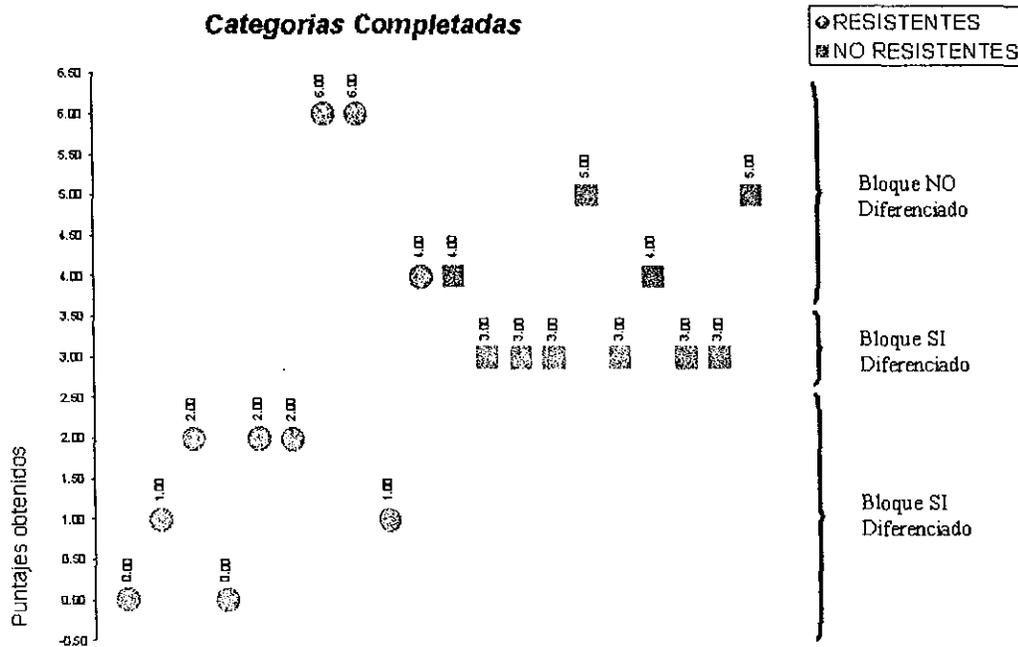


Figura 18. Distribución de los valores de los sujetos de los grupos RE y NR en la variable *Categorías Completadas*. El desplazamiento en el eje X se realizó con el fin de identificar mejor a cada sujeto, no representa valor alguno.

Por otro lado, la distribución de los valores en la variable con menor puntaje de probabilidad ($p = 0.004$), la variable: % de Errores Perseverativos (ver Figura 19), no le permite una mejor discriminación de los grupos involucrados, ya que se obtienen dentro de los dos bloques diferenciados, únicamente 4 y 5 sujetos de los grupos NR y RE respectivamente, y un tercer bloque No diferenciado con 11 sujetos.

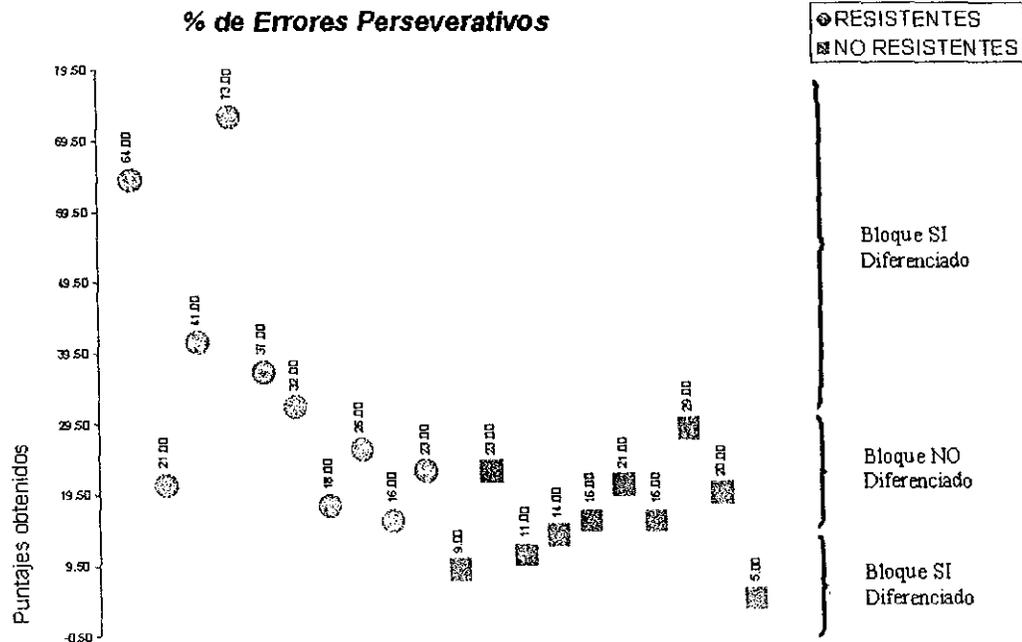


Figura 19. Distribución de los valores de los sujetos de los grupos RE y NR en la variable % de Errores Perseverativos. El desplazamiento en el eje X se realizó con el fin de identificar mejor a cada sujeto, no representa valor alguno.

Lo anterior pone de manifiesto la mayor capacidad discriminativa del sistema al basar la selección de las variables involucradas primeramente en el puntaje de ganancia y después en el puntaje de probabilidad.

De igual forma, cuando los sujetos no se encuentran en una rama (o bloque) completamente diferenciado, la conjunción de sus valores en la variable principal (o tronco) con una segunda e incluso con una tercera, permiten la identificación de su grupo de pertenencia.

- 2. En el par CO VS RE:** Se obtuvo un 15% de error verdadero, con un 20% de fallas en la clasificación de los sujetos CO y 90% en la de los sujetos RE. La variable identificada como principal fue:

— Respuestas Conceptuales

3. En el par **CO VS NR**: Se obtuvo un 40% de error verdadero, tanto en la clasificación de sujetos CO como NR. La variable principal tanto a lo largo de las 20 pruebas como con la totalidad de los sujetos fue:
 - Respuestas Conceptuales
4. En el par **CO VS EZ**: Se obtuvo un 40% de error verdadero, con un 60% de fallas en la clasificación de los sujetos CO y 30% en la de los sujetos RE. La variable identificada como principal tanto a lo largo de las 30 pruebas como con la totalidad de los sujetos fue:
 - Respuestas Conceptuales

Con datos del EEG.

En forma general, los resultados de las pruebas del sistema fueron los siguientes:

- A).- En todos los pares, se observa una tendencia a presentar menores errores de clasificación al utilizar matrices de potencias Absoluta y Relativa por bandas, en contraposición a las de frecuencias. Esta tendencia se acentúa en los pares de CO y RE, y CO y NR.
- B).- De igual manera en la mayoría de los pares se presentan menor número de errores al utilizar las matrices de correlación por frecuencias, en contraposición a la de correlación por bandas.
- C).- Al comparar el rendimiento de las potencias en general (Absoluta y Relativa, tanto por banda como por frecuencia) contra la correlación en general (Inter e Intrahemisférica, por banda y frecuencia), los mejores resultados se encontraron en las correspondientes a las bandas.

Tomando en cuenta el promedio por matriz del número de variables seleccionadas para el Conjunto de Capacitación en cada prueba en los pares CO y RE, CO y NR se obtuvieron valores similares: cercanos al 5 en las matrices por banda y entre los 9 y 10 en las matrices por frecuencia (Ver Tabla 4).

Tabla 4. Promedio de las variables seleccionadas por prueba, en cada matriz. El "***" indica que esa matriz no se valoró

CO y RE		CO y NR		
Banda	Frec.	Banda	Frec.	
*	7	*	6	Potencia Absoluta con ojos Abiertos
*	*	4	5	Potencia Absoluta con ojos Cerrados
6	14	4	8	Potencia Relativa con ojos Abiertos
6	8	9	33	Potencia Relativa con ojos Cerrados
4	7	3	6	Correlación INTER. con ojos Abiertos
*	4	6	9	Correlación INTER. con ojos Cerrados
4	14	4	7	Correlación INTRA. con ojos Abiertos
8	7	7	6	Correlación INTRA. con ojos Cerrados

En forma más específica, las pruebas en las matrices de cada par arrojaron la siguiente información:

1. En el par RE y NR:

Las matrices de Potencias (Absoluta y Relativa, tanto por banda como por frecuencia) obtuvieron menor porcentaje de error que las matrices de Correlación.

La matriz de Correlación Interhemisférica con ojos Abiertos (por banda), fue descartada para terminar las pruebas ya que en una de ellas, ninguna de sus variables obtuvo un valor de < 0.05 .

Las matrices de Correlación Interhemisférica (tanto ojos Abiertos como Cerrados, por banda y por frecuencia) ubicaron su estimado de error verdadero por encima del 50 %.

Se identificaron dos matrices con valores de error verdadero inferiores al 25%, las dos pertenecen a bandas con valores por frecuencia: La Potencia Absoluta con ojos cerrados y la Potencia Relativa con ojos abiertos.

En el caso de la matriz de Potencia Absoluta con ojos Cerrados (también por frecuencia), en ella tanto el valor de Ganancia ($G= 0.6350$) como el de probabilidad ($p= 0.006$) parecen definir como sobresaliente a la variable 10 Hz en C3 (ver Figura 20), en este caso la concordancia está dada porque los dos grupos formadores de la matriz (RE y NR) se congregan mayoritariamente en ambos extremos de un ordenamiento continuo, y se mezclan muy poco (2 sujetos de cada grupo) en una zona nebulosa central. Estadísticamente, esta variable presenta diferencias significativas entre sus grupos, lo que no sucede con la banda a la que pertenece (Alfa2). Esto puede explicarse tomando en consideración que las otras dos variables en la misma derivación, que conforman la banda (11 Hz en C3, $p= 0.01$ y 12 Hz en C3, $P= 0.14$) pueden enmascarar la significancia de la primera.

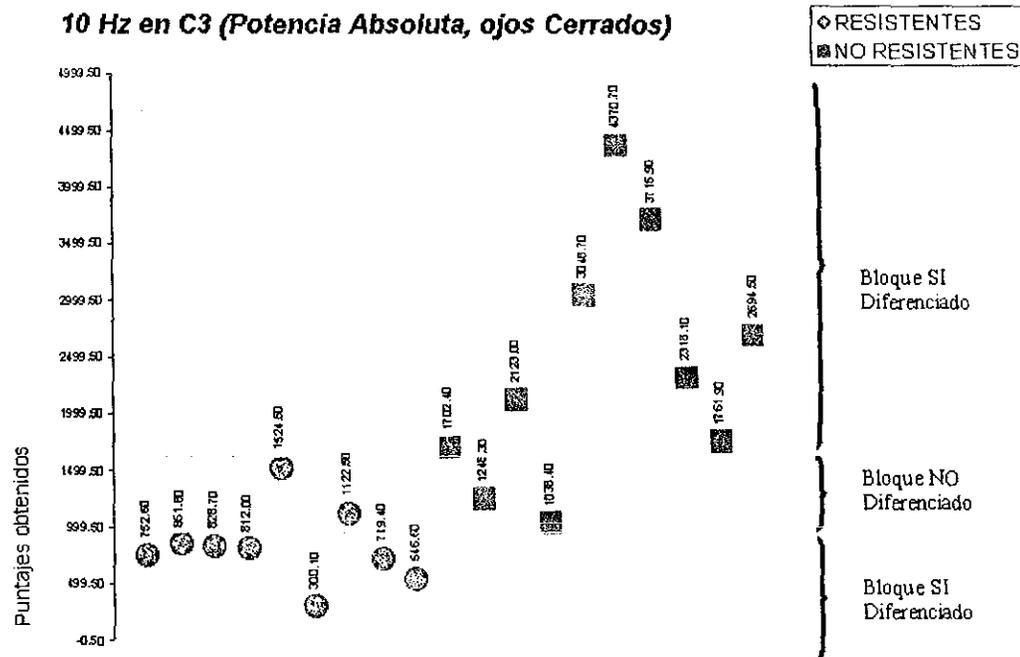


Figura 20. Distribución de los valores de los sujetos de los grupos RE y NR en la variable 10 Hz en C3 (Potencia Absoluta con ojos Cerrados). El desplazamiento en el eje X se realizó con el fin de identificar mejor a cada sujeto, no representa valor alguno.

El árbol formado por el Sistema (ver Figura 21), nos permite identificar a la mayoría de los sujetos de ambos grupos, concentrados en los valores extremos de la variable, esto es en la rama 1 y la rama 3, dejando unos pocos mezclados en la rama central.

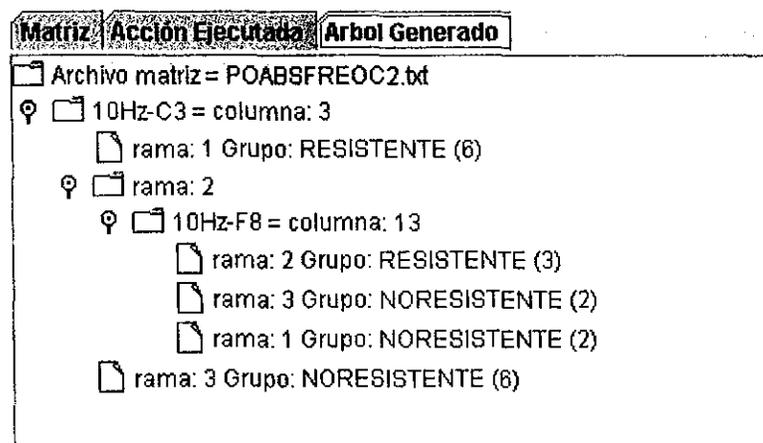


Figura 21. Arbol generado a partir de la matriz de potencia absoluta por frecuencia con ojos cerrados.

Con base en este árbol y los valores umbrales generados por el programa para cada rama, podemos considerar tentativamente, que existen 5 perfiles para la clasificación de sujetos esquizofrénicos paranoides en dos grupos: RE y NR:

- Perfil 1: los sujetos que en la potencia absoluta (con ojos cerrados) de la frecuencia de 10 Hz, en la derivación C3, obtengan un valor **igual o inferior** a $844.10 \text{ d}\mu\text{V}^2$, son Resistentes
- Perfil 2: los sujetos que en la potencia absoluta (con ojos cerrados) de la frecuencia de 10 Hz, en la derivación C3 obtengan un valor **superior** a $1882.27 \text{ d}\mu\text{V}^2$, son No Resistentes
- Perfil 3: este perfil se subdivide en tres partes, e incluye a los sujetos que en la variable y derivación antes citada obtengan puntajes iguales o mayores al perfil 1 y menores al perfil 2:

1. que además, en la potencia absoluta (con ojos cerrados) de la frecuencia de 10 Hz, en la derivación F8, obtenga un puntaje igual inferior a $370.73 \mu V^2$, se clasificará como NR.
2. que obtenga un puntaje superior a $605.46 \mu V^2$, se clasificará como RE, y
3. que obtenga un puntaje igual o superior al primero e inferior al segundo, se clasificará como NR.

Los primeros dos perfiles son más estables y tienen mayor consistencia, ya que también tienen un número mayor de sujetos clasificados.

En el caso de la matriz de Potencia Relativa con ojos Abiertos (datos tomados por frecuencia), la variable que mejor discrimina entre los grupos es 16 Hz en F4, su valor de Ganancia es el más alto de dicha matriz ($G= 0.8080$), sin embargo, su puntaje de probabilidad no es significativo ($p= 0.5676$). Esto se debe a que dentro del grupo de Resistentes, esta variable parece identificar dos subgrupos, o bien, una dispersión muy grande y con dos núcleos en zonas extremas (ver Figura 22). Estadísticamente, esta variable no figuraría como importante, pero los valores generados por los sujetos No Resistentes que se aglutinan mayoritariamente en una zona tan estrecha parecen identificar un perfil de grupo.

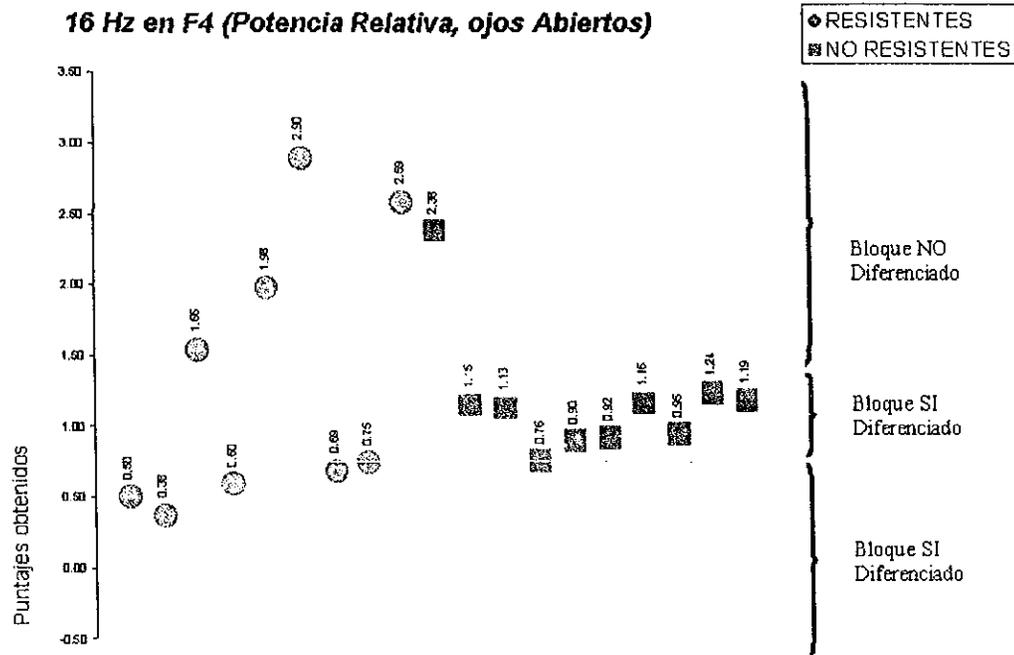


Figura 22. Distribución de los valores de los sujetos de los grupos RE y NR en la variable 16 Hz en F4 (Potencia Relativa con ojos Abiertos). El desplazamiento en el eje X se realizó con el fin de identificar mejor a cada sujeto, no representa valor alguno.

El árbol formado (Figura 23), nos identifica a los sujetos de los grupos

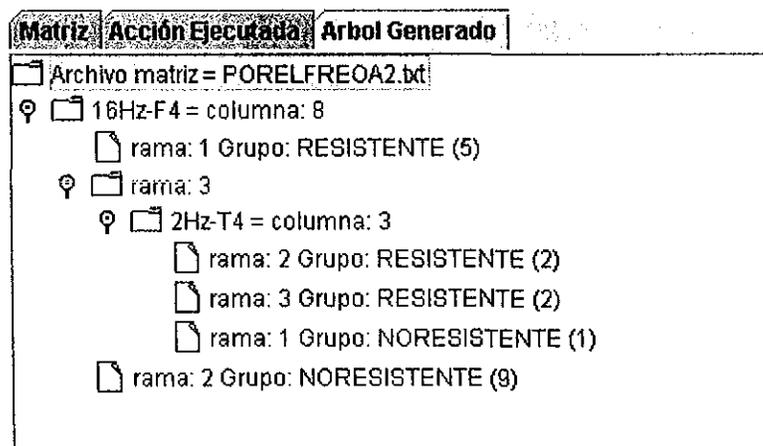


Figura 23. Arbol generado de la matriz de potencia relativa por frecuencia con ojos abiertos.

Con base en este árbol y los valores umbrales generados por el programa para cada rama, podemos considerar tentativamente , que también existen 5 perfiles para la clasificación de sujetos esquizofrénicos paranoides en dos grupos: RE y NR:

Perfil 1: los sujetos que en la potencia relativa (con ojos abiertos) de la frecuencia de 16 Hz, en la derivación F4, obtengan un valor **igual o inferior** a $0.75 \mu V^2$, son Resistentes

Perfil 2: los sujetos que en la potencia relativa (con ojos abiertos) de la frecuencia de 16 Hz, en la derivación C· obtengan un valor **superior** a $0.75 \mu V^2$ pero inferior a $1.55 \mu V^2$, son No Resistentes

Perfil 3: este perfil se subdivide en tres partes, e incluye a los sujetos que en la variable y derivación antes citada obtengan puntajes **iguales o mayores** a $1.55 \mu V^2$, y que además:

1. el que en la potencia relativa (con ojos abiertos) de la frecuencia de 2 Hz, en la derivación T4, obtenga un puntaje inferior a $13.73 \mu V^2$, se clasificará como NR.
2. el que obtenga un puntaje superior $21.74 \mu V^2$, se clasificará como RE, y
3. el que obtenga un puntaje igual o superior al primero e inferior al segundo, se clasificará como NR.

2. En el par CO y RE:

El Sistema identifica mejor a los sujetos del grupo Control que a los del grupo Resistente, ya que el número de errores en la clasificación de este segundo son mayores. Esta diferencia es más evidente en las matrices de potencia relativa (por banda y por frecuencia).

La matriz con más baja estimación de error verdadero (22%) fue la conformada por valores de Correlación Intrahemisférica con ojos abiertos (frecuencias). También sus valores de error por grupo fueron bajos y similares entre sí (22% c/u). En esta matriz, la variable citada como

principal discriminadora de grupos fue: la correlación entre T4 y O2 en 23 Hz con ojos abiertos.

Para esta variable, tanto el valor de Ganancia ($G= 0.666$) como el de probabilidad ($p= 0.0007$) parecen corroborar su alta capacidad de discriminación. Una revisión más minuciosa de la distribución de los valores de los grupos en esa variable permite identificar en los extremos de un continuo a los sujetos de grupos contrarios, y en el centro únicamente 3 sujetos en el bloque no diferenciado (ver la Figura 21).

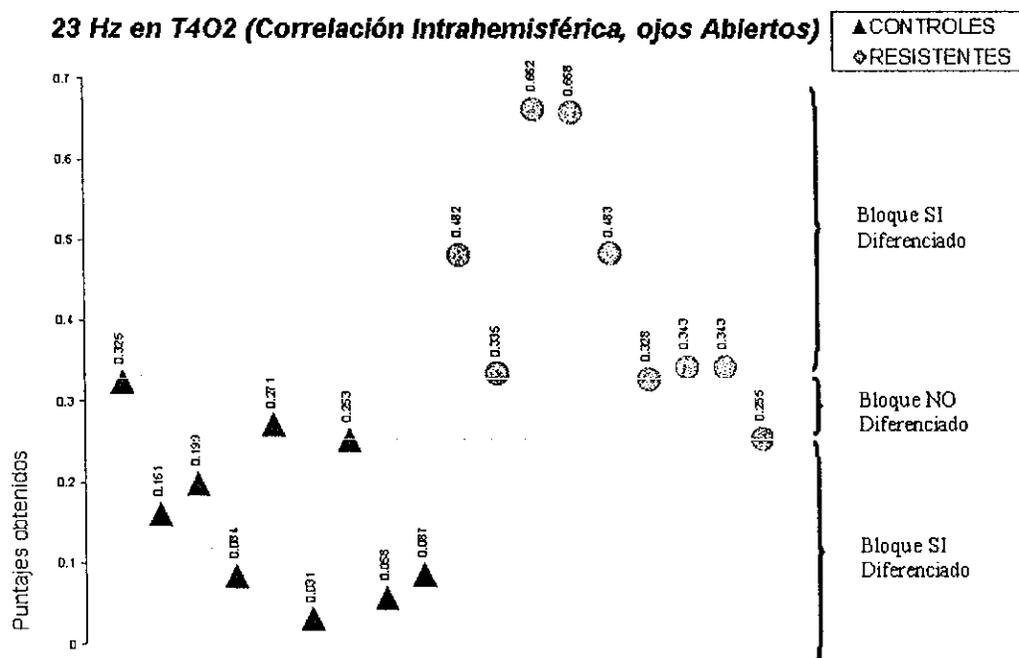


Figura 24. Distribución de los valores de los sujetos de los grupos CO y RE en la variable 23 Hz correlación entre T4 y O2 con ojos abiertos. El desplazamiento en el eje X se realizó con el fin de identificar mejor a cada sujeto, no representa valor alguno.

3. En el par CO y NR:

El Sistema identifica mejor a los sujetos del grupo Control que a los del grupo No Resistente, ya que el número de errores en la clasificación de

este segundo es mayor. Sin embargo, es en este par, en el que se encontró el estimado de error verdadero más bajo.

La matriz con más baja estimación de error verdadero (16%) fue la conformada por valores de la potencia absoluta con ojos cerrados (bandas). También sus valores de error por grupo fueron bajos, 11% en CO y 20% en NR. En esta matriz, la variable citada como principal discriminadora de grupos fue: Alfa1 O2.

Para esta variable, tanto el valor de Ganancia ($G= 0.6350$) como el de probabilidad ($p= 0.003$) parecen corroborar su alta capacidad de discriminación. Una revisión más minuciosa de la distribución de los valores de los grupos en esa variable permite identificar en los extremos de un continuo a los sujetos de grupos contrarios, y en el centro a 7 sujetos en el bloque no diferenciado (ver la Figura 22).

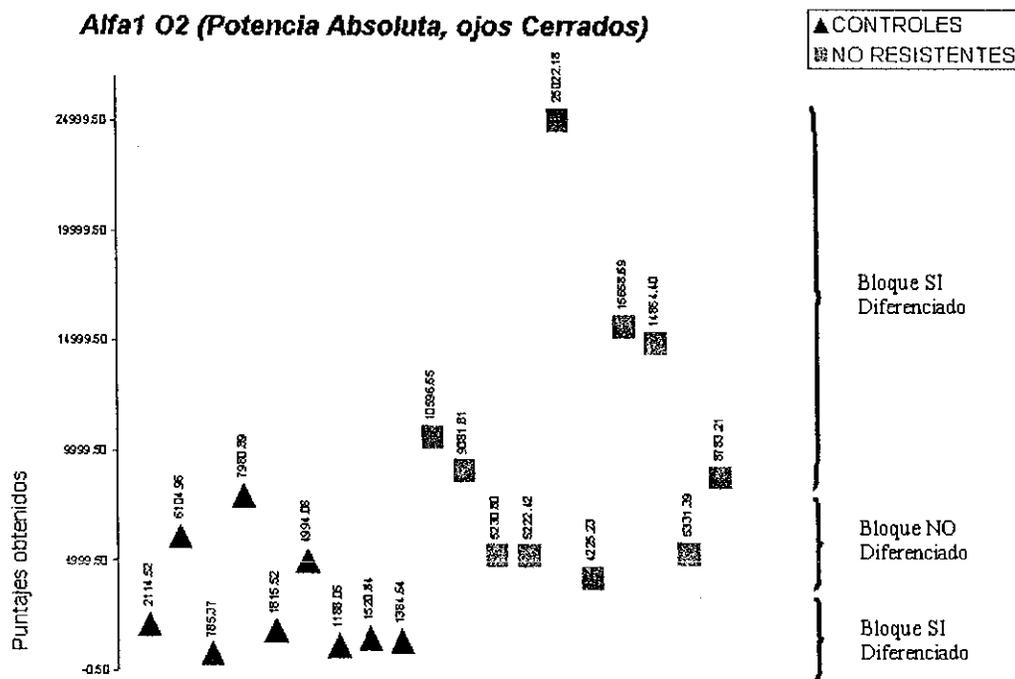


Figura 25. Distribución de los valores de los sujetos de los grupos CO y NR en la variable *Potencia Absoluta de Alfa1 en O2 con ojos cerrados*. El desplazamiento en el eje X no representa valor alguno.

4. En el par CO y EZ:

Al realizar las pruebas con las matrices de potencias Absoluta y Relativa y de correlaciones, tanto Inter. como Intrahemisférica, todo ello por bandas, se observó que el estimado de error en el grupo de CO es mucho mayor que el error en el grupo de EZ. Estos resultados fueron consistentes en todas las pruebas y parece que la influencia de la n en los grupos tiene mucho que ver. El grupo de CO tiene 9 elementos mientras que el de EZ está conformado por 19 sujetos. Esta diferencia puede influir en la conformación del árbol de decisión de tal forma que cuando un nuevo sujeto es evaluado para clasificación, la mayor parte de sus datos coincidan con alguno de los datos del grupo de mayoría (EZ), lo que explicaría un mayor número de aciertos en dicho grupo y un mayor número de fallas en el otro.

Visto lo anterior, se decidió no correr los experimentos de las matrices con datos por frecuencias y considerar que es muy alta la probabilidad de que los resultados obtenidos también estén sesgados por la n de los grupos.

Con todos estos inconvenientes, se identificó, a la matriz de correlación Interhemisférica con ojos abiertos como la que obtuvo el más bajo puntaje de error.

5 Discusión y Conclusiones

Los resultados obtenidos por el sistema tanto en la matriz de Wisconsin como en las de Potencia Absoluta por Frecuencia con ojos cerrados y Potencia Relativa por Frecuencia con ojos abiertos, ambas en el par RE y NR, indican que el sistema puede ser considerado como un adecuado instrumento para la clasificación.

La información que el Sistema nos proporciona con relación a la clasificación de los sujetos considerando los datos de los árboles de decisión son los siguientes:

Una de las matrices que obtuvo el 21% de error verdadero, la de Potencia Absoluta por frecuencia con ojos cerrados, nos dice que:

1. La mayoría de los sujetos esquizofrénicos paranoides que presentan valores pequeños en la potencia absoluta (ojos cerrados) de 10 Hz en la derivación C3, pueden ser Resistentes.
2. pero si presentan valores altos en dicha variable, entonces pueden ser No Resistentes.
3. para los que presentan un valor intermedio, se puede recurrir a una segunda variable para tomar la decisión.

De igual forma, los datos de la segunda matriz, la de Potencia Relativa por frecuencias con ojos abiertos, nos dice:

1. La mayoría de los sujetos esquizofrénicos paranoides que presentan valores pequeños en la potencia relativa (ojos abiertos) de 16 Hz en la derivación F4, pueden ser Resistentes.
2. mientras que los que presentan valores altos se debe de recurrir a una segunda variable para clasificarlos.
3. para los que presentan valores intermedios, la clasificación sería No Resistentes.

De igual forma, el hecho de haber obtenido un porcentaje de error verdadero del 22% en el par CO y RE en la matriz de correlación intrahemisférica con ojos abiertos (por frecuencias), y del 16% en el par CO y NR en la matriz de potencia absoluta con ojos cerrados (por bandas) es indicativo de la capacidad del sistema para identificar sujetos de grupos diferentes.

Sin embargo, es importante comparar su eficacia contra una herramienta estadística tradicional como lo es el modelo de Regresión Lineal Múltiple.

VI.9 ANALISIS DE REGRESION Y COMPARACION CON EL SISTEMA EXPERTO

1 Objetivo

Comparar la efectividad de la clasificación del Sistema Experto con la alcanzada por el método de Regresión Lineal Múltiple.

2 Objetivos Específicos

1. Obtener con un modelo lineal de probabilidad (Regresión Lineal Múltiple, con una variable dependiente dicotómica,) el estimado de Error Verdadero de las matrices que obtuvieron el más bajo puntaje de Error Verdadero en el par RE y NR.
2. Realizar una comparación entre los Estimados de Error Verdadero del Sistema Experto y la Regresión Lineal.

3 Hipótesis

El estimado de Error Verdadero logrado por el Sistema Experto será menor al de la Regresión Lineal.

4 Metodología

4a) SUJETOS

Participaron los 19 sujetos definidos anteriormente y divididos en los grupos independientes ya citados 9 RE y 10 NR.

En el caso de la matriz de Wisconsin los grupos tienen igual número de sujetos: 10 RE y 10 NR.

4b) DATOS

Se utilizaron los datos contenidos en las matrices:

- a) Wisconsin, en el par RE y NR
- b) Potencia Absoluta por Frecuencia con ojos cerrados, en el par RE y NR.
- c) Potencia Relativa por Frecuencia con ojos abiertos, en el par RE y NR.

4c) PROCEDIMIENTO

1. Para calcular el Error Verdadero (incluyendo los puntajes de Error por Grupo) se utilizó la técnica de la *Intervalidación*.
2. Al valorar el estimado de Error con la matriz de datos del Wisconsin se incluyeron las 7 variables utilizadas para el Sistema Experto.
3. En el caso de las matrices de EEG, se utilizó el procedimiento de selección de variables descrito en la Prueba 3 con el fin de reducir el número de las mismas, y cumplir con el requisito (Regresión Lineal Múltiple) de $n-1$ variables en cada grupo, de manera que las variables seleccionadas para cada prueba fuesen las mismas que se utilizaron con el Sistema Experto. Las variables utilizadas en cada prueba para la matriz con datos de Potencia Absoluta por Frecuencia con ojos cerrados, se encuentran listadas en la Tabla 5.
4. Se elaboró un modelo lineal de probabilidad del tipo: Regresión Lineal Múltiple, con una variable dependiente dicotómica (Gujarati, 1981) y se valoró su eficacia en la predicción de sujetos con base en el Error Verdadero y la técnica de Intervalidación.
5. En el continuo de los valores de probabilidad que van de 0 a 1, el umbral que divide los valores del grupo RE y NR es el 0.5. Para todas las pruebas, cuando el pronóstico alcanza valores superiores a 0.5 se considera clasificado como RE, y cuando los valores son inferiores a dicho umbral se considera clasificado como NR, ya que el primer grupo se asoció con el extremo que representa el 1 y el segundo grupo con el extremo que representa el 0.

- Se calculó el porcentaje de error con base en el número de clasificaciones correctas sobre el número de pruebas realizadas, tanto para cada grupo como para el total de los sujetos.

Las variables seleccionadas para cada matriz: de Potencia Absoluta por frecuencia con ojos cerrados (ver Tabla 5) y de Potencia Relativa por frecuencia con ojos abiertos (ver Tabla 6) utilizadas para esta prueba, se presentan a continuación.

Tabla 5. Variables seleccionadas para cada prueba de Regresión. **= < 0.01, *= < 0.05, G= Mayor Ganancia, Celda Gris= No Probado.

Variables	RESISTENTES									NO RESISTENTES									
	Sujeto 1	Sujeto 2	Sujeto 3	Sujeto 4	Sujeto 5	Sujeto 6	Sujeto 7	Sujeto 8	Sujeto 9	Sujeto 10	Sujeto 11	Sujeto 12	Sujeto 13	Sujeto 14	Sujeto 15	Sujeto 16	Sujeto 17	Sujeto 18	Sujeto 19
1. 2 Hz. T3						*	*	*	*		*	*		*			*		
2. 2 Hz. T4	*	*		*		*	*	*	*	*	*	*		*	*	*	*	*	*
3. 2 Hz. C3						*			*										
4. 3 Hz. T3						*		*	*		*				*	*	*		
5. 9 Hz. F8												*							
6. 9 Hz. C3	*	**		*		*	**	*	*	*	*	*		*	**	*	*	*	*
7. 9 Hz. C4															*				
8. 9 Hz. O1															*				
9. 9 Hz. O2	*	*		*		*	*	*	*	*	*	*		*	**	*	*	*	*
10. 10 Hz. F1		*	**							*	*								
11. 10 Hz. F2		*	**						*	*	*	*		*	*	*	*	*	*
12. 10 Hz. F3	*	*	*G	*		*	*	*	*	*	*	*		*	*	*	*	*	*
13. 10 Hz. F4	*	*	**	*		*	*	*	*	*	*	*		*	*	*	*	*	*
14. 10 Hz. F7		*																	
15. 10 Hz. F8			**								*								
16. 10 Hz. T3																		*	
17. 10 Hz. T4											*								
18. 10 Hz. C3	*G	*G	**	*G		*G	*G	*G	*G	*G	*G	*G		*G	*G	*G	*G	*G	*G
19. 10 Hz. C4		*	**	*		*				*	*	*		*	*	*	*	*	*
20. 10 Hz. O2	*	*	**	*		*	*	*	*	*	**	*		*	*	*	*	*	*
21. 11 Hz. F1		*								*									
22. 11 Hz. F4										*									

23.	11 Hz. F8			**	*			*	*		*	*								
24.	11 Hz. C3	*	*		*		*	*	*	**	*	*		*	*	*	*	*	*	*
25.	11 Hz. C4				*															
26.	11 Hz. O2						*													
27.	16 Hz. F8											*								
28.	17 Hz. T4																			G
29.	23 Hz. F7																*			
30.	24 Hz. F7																*			
31.	25 Hz. C3																*			
Total var.		8	13	9	11	19	11	12	11	12	14	16	13	17	11	15	11	12	13	10

Tabla 6. Variables seleccionadas para cada prueba de Regresión. ** = < 0.01, * = < 0.05, G = Mayor Ganancia, Celda Gris = No Probado.

Variables	RESISTENTES									NO RESISTENTES									
	Sujeto 1	Sujeto 2	Sujeto 3	Sujeto 4	Sujeto 5	Sujeto 6	Sujeto 7	Sujeto 8	Sujeto 9	Sujeto 10	Sujeto 11	Sujeto 12	Sujeto 13	Sujeto 14	Sujeto 15	Sujeto 16	Sujeto 17	Sujeto 18	Sujeto 19
1. 2 Hz T4	**	**	**	**	**	**	**	G	**	**	**	**	**	**	**	**	**	**	**
2. 9 Hz F2	**	**	**	**	**	**	**	**	**	**			**		**				
3. 10 Hz F8	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**
4. 16 Hz F4	G	G	G	G	G	G		G	G	G	G	G	G	G	G	G	G	G	G
5. 10 Hz F2			**						**	**									
6. 10 Hz F4			**					**											
7. 10 Hz O2			**		**	**						**							
8. 10 Hz T3						**													
9. 19 Hz T4						**			**										
10. 20 Hz T4						**			**										
11. 24 Hz T4						**			**										
12. 25 Hz T4						**			**										
13. 9 Hz F1							**												
14. 17 Hz F3								G											
15. 20 Hz F3								G											
Total var.	4	4	7	4	5	10	4	7	9	5	3	3	4	3	4	3	3	3	3

5 Resultados

5a) DE LA MATRIZ DE WISCONSIN EN EL PAR RE Y NR

El modelo generado por la Regresión Lineal Múltiple (RLM) no encuentra elementos suficientes para identificar separadamente a los dos grupos en la matriz de aprendizaje. Los coeficientes de determinación (porcentaje de varianza explicada) en la mayoría de las pruebas son cercanos a 0.5.

La variable que permitió al Sistema Experto discriminar entre grupos, no tienen un punto de corte que permita una diferenciación lineal, por ello la Regresión tuvo un bajo rendimiento en la matriz de Wisconsin.

El valor del porcentaje de error verdadero se encuentra en el 50% (Tabla 7).

Tabla 7. Resultados de las pruebas de clasificación con RLM sobre la matriz de Wisconsin.

Grupo	Sujetos	Acierto	Falla	Pronóstico	% Error de Gpo.	% Error Verdadero
RESISTENTES	Sujeto 1		X	0.0000	100	50
	Sujeto 2		X	0.1171		
	Sujeto 3		X	0.2129		
	Sujeto 4		X	0.0000		
	Sujeto 5		X	0.1537		
	Sujeto 6		X	0.1177		
	Sujeto 7		X	0.4861		
	Sujeto 8		X	0.0000		
	Sujeto 9		X	0.1102		
	Sujeto 10		X	0.1745		
NO RESISTENTES	Sujeto 11	/		0.4105	0	
	Sujeto 12	/		0.3526		
	Sujeto 13	/		0.2166		
	Sujeto 14	/		0.0737		
	Sujeto 15	/		0.0000		
	Sujeto 16	/		0.3083		
	Sujeto 17	/		0.2975		
	Sujeto 18	/		0.2544		
	Sujeto 19	/		0.2247		
	Sujeto 20	/		0.3605		

5b) MATRIZ DE PA POR FRECUENCIA CON OJOS CERRADOS.

Los valores de Error, tanto en cada grupo como en el estimado de Error Verdadero de las pruebas realizadas con Regresión Lineal Múltiple (Ver Tabla 8), son superiores a los puntajes de Error obtenidos con el Sistema.

Tabla 8. Resultados de las pruebas de clasificación realizadas con Regresión Lineal Múltiple sobre la matriz de Potencia Absoluta por Frecuencia con ojos cerrados en el par RE y NR.

Grupo	Sujetos	Acierto	Falla	Pronóstico	% Error de Gpo.	% Error Verdadero
RESISTENTES	Sujeto 1	/		0.6722	38	29
	Sujeto 2		X	0.4116		
	Sujeto 3	/		1.0000		
	Sujeto 4	/		0.5471		
	Sujeto 5					
	Sujeto 6	/		0.7927		
	Sujeto 7		X	0.4215		
	Sujeto 8		X	0.4867		
	Sujeto 9	/		0.5521		
NO RESISTENTES	Sujeto 10	/		0.3526	22	
	Sujeto 11	/		0.3488		
	Sujeto 12		X	0.8321		
	Sujeto 13					
	Sujeto 14	/		0.0000		
	Sujeto 15		X	0.8615		
	Sujeto 16	/		0.4444		
	Sujeto 17	/		0.2465		
	Sujeto 18	/		0.3664		
	Sujeto 19	/		0.0000		

5c) MATRIZ DE PR POR FRECUENCIA CON OJOS ABIERTOS.

Los valores de Error, tanto en cada grupo como en el estimado de Error Verdadero de las pruebas realizadas con Regresión Lineal Múltiple (Ver Tabla 9), son superiores a los puntajes de Error obtenidos con el Sistema.

Tabla 9. Resultados de las pruebas de clasificación realizadas con Regresión Lineal Múltiple sobre la matriz de Potencia Absoluta por Frecuencia con ojos cerrados en el par RE y NR.

Grupo	Sujetos	Acierto	Falla	Pronóstico	% Error de Gpo.	% Error Verdadero
RESISTENTES	Sujeto 1	/		0.8299	55.5	36.84
	Sujeto 2	/		0.8232		
	Sujeto 3		X	0.0820		
	Sujeto 4	/		0.7110		
	Sujeto 5	/		0.6763		
	Sujeto 6		X	0.0000		
	Sujeto 7		X	0.1939		
	Sujeto 8		X	0.3892		
	Sujeto 9		X	0.2015		
NO RESISTENTES	Sujeto 10		X	0.6466	20	
	Sujeto 11	/		0.2304		
	Sujeto 12	/		0.1514		
	Sujeto 13	/		0.2582		
	Sujeto 14	/		0.0201		
	Sujeto 15	/		0.4249		
	Sujeto 16	/		0.3082		
	Sujeto 17		X	0.6559		
	Sujeto 18	/		0.1454		
	Sujeto 19	/		0.0461		

6 Discusión y Conclusiones

Se pueden notar diferencias importantes entre las dos metodologías comparadas.

- Primero.** La capacidad para realizar la integración de la información de los ejemplos en un modelo o perfil clasificatorio. El Sistema Experto puede incluir a todas las variables presentes en la matriz, el módulo de selección le permite ser más eficiente, sin embargo, aún en el caso de prescindir de él, el sistema funciona. Por otro lado, la Regresión Lineal no puede generar un modelo cuando se incluyen igual o mayor número de variables que de sujetos en su matriz de aprendizaje.
- Segundo.** La efectividad de la Regresión se disminuye dramáticamente cuando trata de establecer la clasificación de un sujeto desconocido con base en un modelo predictivo extraído de variables que no presentan diferencias significativas entre grupos.
- Tercero.** Tanto el puntaje de Error en cada grupo como el Error Verdadero son mayores bajo la metodología de la Regresión Lineal que con el Sistema Experto.
- Cuarto.** Solamente el Sistema experto es capaz de presentar una ruta que muestra la interrelación entre las variables implicadas en la clasificación y los umbrales de estas mismas con los valores del sujeto clasificado.

Lo anterior nos lleva a considerar que el Sistema es superior en efectividad tanto cuantitativa como cualitativamente a la metodología de la Regresión Lineal Múltiple.

VII RESUMEN DE RESULTADOS DE LAS PRUEBAS

Las pruebas se realizaron sobre un total de 69 matrices divididas en cuatro pares de grupos. Esto es un primer par definido por los datos de los sujetos RE y NR, un segundo par con datos de CO y RE, un tercero con CO y NR, y finalmente un cuarto par con los datos de CO y EZ.

Para cada par, se conformaron 16 matrices de EEG, con los datos de las potencias Absoluta y Relativa por bandas y frecuencias ojos abiertos y cerrados, con los datos de las correlaciones Inter e Intrahemisféricas igualmente por bandas y frecuencias ojos abiertos y cerrados.

Para cada par se conformó una matriz con los datos del Wisconsin.

Para la primera matriz únicamente (RE y NR), se conformó una matriz con los datos del BPRS-PANSS.

Una lista de las matrices (para cada par) formadas es la siguiente:

De EEG.

1. De Potencia Absoluta por banda con Ojos Abiertos.
2. De Potencia Absoluta por banda con Ojos Cerrados.
3. De Potencia Relativa por banda con Ojos Abiertos.
4. De Potencia Relativa por banda con Ojos Cerrados.
5. De Correlación Interhemisférica por banda con Ojos Abiertos.
6. De Correlación Interhemisférica por banda con Ojos Cerrados.
7. De Correlación Intrahemisférica por banda con Ojos Abiertos
8. De Correlación Intrahemisférica por banda con Ojos Cerrados
9. De Potencia Absoluta por frecuencia con Ojos Abiertos.
10. De Potencia Absoluta por frecuencia con Ojos Cerrados.
11. De Potencia Relativa por frecuencia con Ojos Abiertos.
12. De Potencia Relativa por frecuencia con Ojos Cerrados.
13. De Correlación Interhemisférica por frecuencia con Ojos Abiertos.

14. De Correlación Interhemisférica por frecuencia con Ojos Cerrados.
15. De Correlación Intrahemisférica por frecuencia con Ojos Abiertos
16. De Correlación Intrahemisférica por frecuencia con Ojos Cerrados

De Wisconsin

17. Matriz con los datos del Wisconsin.

De BPRS-PANSS

18. Matriz con datos de BPRS y de PANSS (sólo para el par RE y NR).

VII.1 PROCESO DE CONSTRUCCION DEL SISTEMA

En la primera prueba, el sistema pudo reproducir el árbol generado por los datos discretos proporcionados por Russell y Norvig, (1996 pp. 565).

En la segunda prueba la utilización de los métodos: Rango cuartil Q1, H, Q4; Terciles; Mediana por Grupo; y Mediana, en forma separada, discretizando la totalidad de las variables en la matriz con cada uno de ellos, mostraron poca efectividad. El rango del valor estimado de *error verdadero* en una misma matriz discretizada con los 4 métodos se encontró entre el 31 y el 26%. Sin embargo, se implementó un sistema denominado *Iteración por Variable* el cual consiste en discretizar cada variable con todos los métodos secuencialmente y obtener el valor de la ganancia después de cada discretización, lo que permitió finalmente, dejar discretizada la variable con el método que generó la mayor ganancia.

En la tercera prueba, el sistema pudo construir un árbol a partir de datos de EEG y alcanzar valores aceptables de predicción de sujetos desconocidos. Además, la ruta presentada por el árbol de decisión proporciona información de los umbrales tomados en cuenta para la clasificación de un sujeto desconocido.

VII.2 PROCESO DE APRENDIZAJE DEL SISTEMA

Los sistemas expertos basados en la generación de árboles de decisión por aprendizaje de ejemplos parten de un estimado error aparente de cero, ya que los

ejemplos utilizados en su matriz de aprendizaje o Conjunto de Capacitación son correctamente clasificados para generar el árbol.

El Sistema fue construido conforme a los cánones de los sistemas expertos citados arriba, utilizando el algoritmo de la Ganancia de Información y durante las pruebas no reportó ningún sujeto del Conjunto de Capacitación sin clasificación correcta, por lo que tiene un estimado de *error aparente* de cero.

VII.3 PROCESO DE PRUEBA DEL SISTEMA

El proceso de prueba del sistema consistió en el cálculo del *valor de error verdadero*. Para ello en cada matriz (de cada comparación de grupos) fueron extraídos los datos de un sujeto a la vez, conformando éste a la prueba y el resto al Conjunto de Capacitación, y realizándose tantas pruebas como sujetos contenía la matriz. En cada prueba, el Conjunto de Capacitación entraba al sistema para formar el árbol, mientras que los resultados de la clasificación del sujeto extraído conformaban el valor de error.

Se identificaron las siguientes variables como contenedoras de información suficiente para realizar la clasificación de sujetos desconocidos:

Para el par RE y NR: dos matrices a saber:

La de potencia absoluta con los ojos cerrados y

La de potencia relativa con los ojos abiertos, ambas por frecuencias.

Para el par CO y RE: la matriz con mayor información discriminativa fue la correlación intrahemisférica con los ojos abiertos (por frecuencias).

Para el par CO y RE: la matriz con mayor información discriminativa fue la potencia absoluta con ojos abiertos (por bandas).

En el par CO y EZ: se identificó una fuerte tendencia a formar el perfil de aprendizaje a favor de los EZ, influida por el mayor número de sujetos que este grupo contenía.

VIII DISCUSION Y CONCLUSIONES

VIII.1 EL INSTRUMENTO

A partir de los resultados del Sistema en las pruebas con datos discretos se consideró óptimo el funcionamiento del Algoritmo de la Ganancia de Información y la estructura básica de los módulos de construcción del árbol de decisión.

El manejo de datos cuasi-continuos como es el caso del EEG digitalizado se pudo superar con los métodos de discretización, y estos en conjunción con la selección de las variables tomando en consideración los valores de ganancia y el resultado de la prueba U de Mann Whitney producen los mejores resultados

Este sistema, igual que muchos otros integra un módulo de poda de árbol, esto es, un módulo que se encarga de evitar que el árbol incorpore variables sin importancia para la clasificación, pero a diferencia de otros Sistemas Expertos, el presente toma en consideración la selección de variables con base en pruebas estadísticas y en combinación con los valores obtenidos con el algoritmo de ganancia de información. Esto quiere decir que la poda se realiza antes de integrar al sistema los datos.

El Sistema pudo manejar adecuadamente los datos discretos de la primera prueba y los resultados con las pruebas efectuadas con matrices de EEG.

La conformación final del Sistema parece ser la más indicada. Los tres pasos principales se pueden resumir en lo siguiente:

Primero, la selección de las variables con mayores diferencias significativas entre grupos permite excluir aquellas que no contienen información discriminatoria, disminuyen la carga de análisis y el procesamiento en el Sistema y evitan que se incluya en los perfiles de grupo, información irrelevante. El reducir en la medida de lo posible la información que no se encuentra directamente relacionada con la discriminación de los grupos es importante, y nuestro trabajo al igual que otros anteriormente realizados como el de McKenzie y cols. (1993), permiten afirmar que una preselección de las variables mejora las reglas derivadas de ellas, y de esta manera, la efectividad en la clasificación.

Segundo, la metodología de discretización incluida permite convertir los numerosos valores digitalizados del EEG en un número menor de valores que engloban grupos de sujetos, permitiendo mantener en la mayor medida posible la información contenida en los datos.

Tercero, el algoritmo de Ganancia de Información integrado al subprograma de formación de árbol genera las reglas de clasificación en forma de rutas desde el tronco y las ramas hasta la hoja final. Cada rama simboliza un conjunto de valores en la variable, y esto implica que en cada variable se determinan varios umbrales (mínimo uno, máximo dos) contra los que se comparan los valores del sujeto a clasificar.

El sistema requiere en promedio entre 3 y 4 variables diferentes para formar un árbol cuando se tiene un aproximado de 17-19 sujetos en el Conjunto de Capacitación

El Sistema funciona eficientemente y puede clasificar adecuadamente los datos de EEG de sujetos desconocidos.

VIII.2 LOS DATOS

El manejo de los datos, tanto del EEG como de resultados psicopatológicos y neuropsicológicos propuso, con la discretización, que en las variables con mayor capacidad discriminativa, sus valores (acomodados en un continuo) aglutinarían en cada extremo a un grupo, y que la mayoría de los valores mezclados se encontrarían en el centro, sin embargo, algunos casos encontraron ambos extremos ocupados principalmente por un grupo, mientras que su contraparte se encontraba en el centro. Esto fue interpretado como la posible subdivisión del primer grupo en dos tipos de sujetos.

En una comparación entre las matrices de pruebas neuropsicológicas y psicopatológicas con las de EEG, las primeras obtuvieron menores valores de error que las segundas.

Con relación a la influencia de la n de los grupos de cada matriz en el sistema, los resultados en relación con los porcentajes de acierto en cada grupo de las matrices COvsEZ, misma que está conformada por 9 y 19 sujetos respectivamente, parecen indicar que el segundo grupo tiene mayor influencia en la construcción del

árbol, lo que demerita el rendimiento en la clasificación de los sujetos del primero. Esto parece indicar que la efectividad del programa puede verse sesgada desfavorablemente a un grupo cuando el número de elementos de su contraparte es notoriamente mayor en la matriz que ambos conforman. Lo anterior se observa con menor frecuencia en el resto de las matrices cuyo número en ambos grupos es similar.

Los resultados apoyan la hipótesis implícita en el presente trabajo, de que la información contenida en el EEG permite la discriminación de dos grupos de esquizofrénicos paranoides (RE y NR), pero también de que se pueden diferenciar por separado los grupos de Resistentes y No Resistentes con relación a los Controles.

VIII.3 LAS PRUEBAS

1 Con las matrices de datos psicopatológicos y neuropsicológicos.

Para el caso de la matriz de datos psicopatológicos (BPRS-PANSS), el sistema fue capaz de identificar con el más alto valor de ganancia, las dos variables con el menor valor de p , esto es, BPRS y PANSS Positivo (+) y con ello clasificar correctamente a todos los sujetos de dicha matriz. La coincidencia se debió a que los datos de cada uno de los grupos se presentan cargados a cada uno de los extremos de un continuo y ello permite que el valor de corte se presente en la parte media. Esto a su vez facilita la discretización en dos valores que engloban cada uno a la totalidad de los sujetos de un grupo.

Por otro lado, no es una coincidencia que el Sistema haya elegido como una de las variables más importantes para la discriminación grupal, a la variable que contiene el puntaje de síntomas positivos, ya que son estos los que se disminuyen o remiten principalmente con la medicación, permaneciendo por más tiempo los negativos. Por tanto, de dos grupos existentes en los que uno de ellos ha respondido (y por tanto han disminuido sus síntomas positivos) y el otro no, las diferencias para con el que no ha respondido se refleja en los puntajes que en esa variable los hacen diferentes entre sí. El Sistema identifica a dicha variable fácilmente, porque su diferencia es lineal y se expresa en un bajo valor de p y un alto puntaje en G .

Sin embargo, la selección de las variables basada en los más bajos valores de p , no permite por sí misma la producción de árboles con adecuada capacidad predictiva. Con la matriz de datos del Wisconsin en el par RE vs NR realizando la primera selección de variables (primera poda) con base en aquellas que obtuviesen valores de $p < 0.01$ ó 0.05 el estimado de error verdadero fue de 40%, mientras que al incluir en la selección a las variables con mayor ganancia, dicho error disminuyó hasta el 5%. Por tanto, los valores de ganancia, unidos a los de p , de cada variable pueden ayudar a identificar las variables con mayor información discriminativa entre los grupos.

2 Con la matriz de datos de EEG

EN LA MATRIZ DEL PAR RE vs NR.

Los menores porcentajes en el estimado del error verdadero en la mayoría de las matrices de datos de frecuencias en comparación con las de datos de bandas, parece indicar que la información contenida en las frecuencias permite identificar y distinguir mejor los grupos contenidos en una matriz, que la información de las bandas.

La diferencia en el número de variables contenidas en la matriz original (antes de la primera poda) entre las matrices de frecuencias y las de bandas, no parece ser un factor decisivo en la estimación del error verdadero, ya que una vez que se pasa a la matriz en cuestión por la primera poda, el número de variables elegidas resultante es muy similar entre las dos. Esto es, en los dos casos de matrices, el número mínimo de variables es 3 y el máximo elegido pocas veces supera los 15, esto permite que la comparación entre las matrices de bandas y frecuencias sea más válida ya que el número de variables que entra a formar el árbol es muy similar. Las diferencias encontradas en relación a la efectividad no se pueden atribuir al número de variables seleccionadas.

Por otro lado, el hecho de que una matriz carezca de la información suficiente en sus datos para clasificar correctamente a cada uno de los sujetos prueba parece correlacionarse con dos factores: el primero tiene que ver directamente con el porcentaje estimado de error verdadero, ya que los valores más grandes de este indicarían una distribución más aleatoria de los sujetos de ambos grupos (ordenados

en un continuo); el segundo, tiene que ver con el número de variables diferentes citadas durante las pruebas de clasificación de todos los sujetos de la matriz, un alto número de variables diferentes quiere decir que cada vez que un sujeto es extraído de la matriz los valores de los sujetos que se quedan, modifican el valor de Ganancia (G) o de significancia (p) de las variables, de tal modo que otra variable diferente es elegida en cada ocasión. Así mismo, ello estaría indicando que los valores de G o de p son similares en un gran número de variables, por ello, una matriz que a lo largo del total de sus pruebas obtenga un alto número de variables diferentes citadas en el tronco del árbol, indicaría que carece de un patrón (proporcionado por los datos de los sujetos que conforman cada grupo) capaz de identificar a los grupos participantes.

En caso contrario, cuando una matriz cuenta con valores pequeños de error verdadero y número reducido de variables diferentes durante sus pruebas, podemos pensar que estamos ante un patrón congruente de clasificación de los grupos implicados en la matriz.

Tal es el caso de las matrices de Potencia Absoluta con ojos Cerrados y Potencia Relativa con ojos Abiertos, ambos por frecuencia. El valor del estimado de error verdadero es de 21, con un número de 2 y 3 variables diferentes durante las pruebas, de las cuáles, la variable principal cuenta con 18 y 17 citas de 19 clasificaciones, respectivamente. Esto permite inferir que dichas variables siguieron obteniendo altas puntuaciones de G y bajos valores de p la mayoría de las veces, sin importar qué sujeto fuera excluido del Conjunto de Capacitación.

3 En la comparación del Sistema con el método de regresión lineal.

La comparación de la efectividad del Sistema Experto con la metodología de la Regresión Lineal se llevó a cabo con base en el estimado de Error, tanto del Error Verdadero como del Error por Grupo. En virtud de lo anterior, el Sistema resultó más efectivo para clasificar correctamente a un sujeto desconocido que el método de la Regresión. Sin embargo, también la posibilidad de incluir en la matriz de capacitación números grandes de variables para su manejo con el Sistema, marca una diferencia con relación a la Regresión lineal. Esta última no permite manejar un número mayor a $n-1$ variables por grupo en la matriz.

Además, el Sistema puede proporcionar la ruta completa de la clasificación y los umbrales que separan a los grupos en una variable, esto es, el Sistema puede explicarnos más fácilmente el porqué de la clasificación llevada a cabo.

La forma no lineal que utiliza el Sistema para extraer información relevante para la discriminación grupal, le permite identificar variables no elegidas como importantes por los sistemas estadísticos tradicionales y por tanto puede integrar modelos predictivos más efectivos.

El Sistema Experto Esquizor se perfila como una herramienta importante en la clasificación de grupos con base en datos Electroencefalográficos, Psicopatológicos y Neuropsicológicos.

Estudios futuros tendientes a aumentar el tamaño de la muestra en cada grupo podrían establecer perfiles clasificatorios más estables y con ello aumentar el grado de efectividad del sistema.

4 Ventajas y Desventajas.

Los resultados del sistema nos permiten enumerar las siguientes:

VENTAJAS:

1. El sistema permite evaluar un número elevado de variables, aún cuando en el resultado de dicha evaluación, permita elegir un número menor para elaborar el árbol.
2. El Sistema realiza la clasificación de sujetos desconocidos con base en un árbol de decisión generado con los ejemplos del Conjunto de Capacitación, y con base en un número realmente pequeño de variables que conforman el citado árbol.
3. El Sistema se adapta cada vez que un nuevo sujeto le es agregado o extraído.
4. El sistema permite visualizar en el árbol generado, las relaciones existentes entre las variables implicadas en la ruta de clasificación y proporciona los umbrales de cada variable implicada en

5. En el caso de tomarlo como selector únicamente, el sistema puede servir para indicar cuando existen variables con diferencia significativa.

También se pueden enumerar las siguientes:

DESVENTAJAS:

1. Precisamente, debido a que el Sistema utiliza pocas variables para formar el árbol, desaprovecha la fortaleza que le significaría considerar más variables en cada rama del árbol.
2. El sistema precisa que las variables sean discretizadas antes de procesar la los datos, lo que puede disminuir la fuerza de la información proporcionada.
3. Solo puede manejar dos grupos diferentes a la vez.
4. Falla cuando falta algún dato en la matriz proporcionada.
5. El tamaño de la n en cada grupo debe de ser relativamente parecido, ya que las grandes diferencias en el número de sujetos que contiene uno y otro grupos se reflejan en perfiles clasificatorios favorecedores del grupo más grande.

ESTUDIO POSTERIOR

Existen nuevas propuestas dentro de la IA, por medio de las cuales se puede mejorar el sistema creado en este proyecto.

Dentro de los sistemas inteligentes, las redes neuronales pueden manejar el aprendizaje de manera similar a los Sistemas Expertos, ya que pueden aprender en base a ejemplos y generalizar dicho aprendizaje, además, este aprendizaje es adaptativo, porque se actualiza con cada nuevo ejemplo. Sin embargo, también es capaz de soportar la falta de datos y no requiere de la discretización.

Un Sistema bajo esa tecnología podría generar los perfiles de varios grupos a la vez sin las restricciones impuestas por este Sistema experto.

La propuesta actual, a la luz de los resultados obtenidos por el programa, y una vez que lo encontrado refuerza la idea de que el EEG puede contener información suficiente para realizar clasificaciones o predicciones, es que se aborte este problema con la tecnología más robusta, confiable y avanzada posible, misma que se encuentra idealmente representada por las Redes Neuronales.

IX BIBLIOGRAFIA

- Abud J. (1998). Inteligencia Artificial. *PC Magazine en español*. 9:5, 113-14.
- Anett, M. (1967) The binomial distribution of right, mixed and left handedness. *Quarterly Journal of Experimental Psychology*, 61, 303-321.
- Ato, M. y López, J.J. (1994). *Fundamentos Estadísticos con SYSTAT*. Addison-Wesley Iberoamericana, E.U.A.
- Avasthi, A., Kulhara, P. (1977) Subsyndromes of chronic schizophrenia: a phenomenological study. *Acta Psychiatr Scand*: 96: 225-229.
- Bellman, R.E. (1978). *An Introduction to Artificial Intelligence: Can Computers Think?*. Boyd y Fraser Publishing Company, San Francisco.
- Berg, E.A. (1948). A simple Objective test for measuring flexibility in thinking. *Journal of General Psychology*, 39, 15-22.
- Brenner, H.D. y Merlo, M.C.G. (1995). Definition of therapy resistant schizophrenia, and its assesment. *European Psychiatry*, 10(1): 11-18.
- Brenner, H.D., Dencker, S.J., Goldstein, Mj., et al. (1990). Defining Treatment refractoriness in schizophrenia. *Schizophr Bull*. 16: 551-61.
- Cerdán, L. (1999). *Características morfológicas y funcionales cerebrales de pacientes esquizofrénicos resistentes a neurolépticos*. Tesis de maestría en ciencia del comportamiento, opción neurociencias, Instituto de Neurociencias, Universidad de Guadalajara, México.
- Corsi-Cabrera., M. (1983) *Psicofisiología del Sueño*. México:Trillas.
- Charniak E., and McDermott D. (1985). *Introduction to Artificial Intelligence*. Addison-Wesley, Reading, Massachusetts.
- Chelune, G.J., y Baer, R.L. (1986). Developmental norms for the Wisconsin Card Sorting Test. *Journal of Clinical and Experimental Neuropsychology*, 8, 219-228.
- Davey, B., Fright, W., Carroll, G. and Jones, R. (1989). Expert System approach to detection of Epileptiform activity in the EEG. *Medical y Biological Engineering y Computing*. 27, 365-370.
- Desmedt, J. E. y Tomberg, C. (1994). Transient phase-locking of 40 Hz electrical oscillations in prefrontal and parietal human cortex reflects the process of conscious somatic perception. *Neuroscience Letters*. 168, 129.
- Dingle, A., Jones, D., Carroll, G., Fright, W. (1993). A Multistage System to Detect Epileptiform Activity in the EEG. *IEEE Transactions on Biomedical Engineering*. Vol. 40, No. 12, 1260-1268.

- DSM-IV. (1995). *Manual diagnóstico y estadístico de los trastornos mentales*. Masson, S.A., Barcelona, España.
- Freedman Alan. (1994). *Diccionario de Computación*. Quinta edición, McGraw-Hill, México, p. 13.
- Gnys, J.A., y Willis, W.G. (1991). Validation of executive function tasks with young children. *Developmental Neuropsychology*, 7, 487-501.
- Gold, J.M., Carpenter, C., Randolph, C., Goldberg, T.E. Weinberger, D.R. (1997). Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Archives of General Psychiatry*, 54, 2159-65.
- Grant, D.A., y Berg, E.A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology*, 34, 404-411.
- Guajarati, D. (1981). *Econometría Básica*. McGraw Hill. México.
- Guevara, M .A. (1995). *Aplicación y Análisis de Correlación a la Actividad Electroencefalográfica*. Tesis Doctoral. Universidad Nacional Autónoma de México. México
- Guevara, M.A. (1988). *Análisis de Potenciales Evocados Mediante los Métodos de Componentes Principales y Regresión Lienal Múltiple*. Tesis de Maestría en Ingeniería Electrica, Universidad Nacional Autónoma de México, México, D.F.
- Harmony T. y Alcaraz V.M. (Coordinadores). (1987). *Daño Cerebral, diagnóstico y tratamiento*. Editorial Trillas, S.A. México.
- Harmony, T., Otero, T., Ricardo, J. Y Fernández, G. (1973). Polarity coincidence correlation coefficient and signal energy ratio of the ongoing EEG activity. I. Normative data. *Brain Res*. 61: 133-140.
- Haut, M.W., Cahill, J., Cutlip, W.D., Stevenson, J.M., Makela, E.H., Bloomfield, S.M. (1996). On the nature of Wisconsin Card Sorting Test performance in schizophrenia. *Psychiatry Research*, 65, 115-22.
- Heaton, R., Chelune, G., Talley, J., Kay, G., Curtiss, G. (1993). *Wisconsin Card Sorting Test Manual*. Psychological Assessment Resources, Inc. (PAR). USA.
- Jaworski, J. (1999). *Java 1.2 al descubierto*. Prentice Hall, España.
- Gerez M. y Tello A. (1995). Selected quantitative EEG (QEEG) and Event-Related Potential (ERP) variables as discriminators for positive and negative schizophrenia. *Biological Psychiatry*, 38: 34-49.
- John, R. R. (1977). Clinical applications of quantitative electrophysiology. *Functional Neuroscience*. Neurometric, 2.
- John, E.R., Pricheps, L.S., Kenneth R., Francis, G., Cancro R., Easton P. y Esverdlof, I. (1994). Quantitative electrophysiological characteristics and subtyping of schizophrenia. *Biological Psychiatry*, 36: 801-826.

- Kane J., Honigfeld G., Singer J. y Meltzer H.Y. (1988). Clozapine for the treatment-resistant schizophrenic: A double blind comparison with chlorpromazine. *Archives of General Psychiatry*, 45, 789-796.
- Keefe, R., Mohs R. y Silverbam, J.M. (1990) Characteristics of Kraepelinian Schizophrenia and their relation to premorbid social functioning. In *The Neuroleptic nonresponsive Patient: Characterization and Treatment* (Angrist B.O. Schulz Ch. Eds.) Washington: American Psychiatric Press.
- Kiloh, L., McComas, A., Osselton, J., y Upton, A. (1981). *Clinical electroencephalography*. Fourth edition, Butterwoeths y Co, Inglaterra.
- Knorrning, L. y Lindström, E. (1995). Principal components and further possibilities with the PANSS. *Acta Psychiatrica Scandinavica*. 91 (Suppl 388): 5-10.
- Korpinen, L. (1993). Computer-aided decision-making for epilepsy and sleep diagnostics. *Acta Neurologica Scandinavica (Supplementum)*. 144: 1-101.
- Kurzweil, R. (1990). *The Age of Intelligent Machines*. MIT Pres, Cambridge, Massachusetts.
- Lenzenweger, M., and Dworkin 1996. The Dimensions of Schizophrenia Phenomenology Not one or two, at least three, perhaps four. *British Journal of Psychiatry*, 168: 432-440.
- Llinás, R. y Ribary, U. (1993). Coherent 40-Hz. oscillation characterizes dream state in humans. *Neurobiology*. 90, 2078-2081.
- Matousek, M. Y Petersen, I. (1973). Frequency analysis of the EEG in normal children and adolescents, en P. Kellaway e I. Petersen (dirs.), *Automation of Clinical Electroencephalography*, Raven Press, Nueva York.
- McKenzie, D.P., McGorry, P.D., Wallace, C.S., Low, L.H., Copolov, D.L. y Singh B.S. (1993). Constructin a Minimal Diagnostic Decisión Tree. *Methods of Information in Medicine*, 32, 161-6.
- Mitchell T.M. (1982). Generalization as search. *Artificial Intelligence*. 18(2), 203-226.
- Morgan, M. (1999). *Descubre Java 1.2*. Prentice Hall. España.
- Rich, E. y Knight, K. (1994). *Inteligencia Artificial*. Segunda edición, McGraw-Hill, España.
- Russell, S, y Norvig, P. (1996). *Inteligencia Artificial: un enfoque moderno*. Prentice Hall Hispanoamericana, S.A., México, pp. 3-5.
- Salokangas RK.(1997) Structure of schizophrenic symptomatology and its changes over time: prospective factor – analytical study. *Acta Psychiatr Scand*: 95: 32-39.
- Shalkoff R.J. (1990). *Artificial Intelligence: An Engineering Approach*. McGraw-Hill, New York.

- Simon, D. (1983). Interview in Wittenburg Door, June-July, P. 20. Citado por Myers, D. G. *Psicología Social*. México: McGraw-Hill, Inc., 1987
- Simon, H.A. (1983). Why should machines learn?. En *Machine Learning, An Artificial Intelligence Approach*. ed. Michalski, R.S., Carbonell, J.G., y Mitchell, T.M. Tioga Press, Palo Alto, CA.
- Sullivan, E.V., Mathalon, D.H., Zipursky, R.B., Kersteen-Tuker, Z., Knight, R.T. and Pfefferbaum, A. (1993). Factors of the Wisconsin Card Sorting Test as Measures of Frontal-Lobe Function in Schizophrenia and in Chronic Alcoholism. *Psychiatry Research*. 46: 175-199.
- Thompson, R. (1986). *Fundamentos de Psicología Fisiológica*. Editorial Trillas, México.
- Tyner, R., Knott, J., Mayer, W. (1983). *Fundamental of EEG Technology. Basic Concepts and Methods*, New York: Raven PrePress.
- Volz, H.P., Gaser, C., Häger, F., Rzanny, R., Mentzel, H. J., Kreitschmann-Andermahr, I., Kaiser, W. A., Sauer, H. (1997). Brain activation during cognitive stimulation with the Wisconsin Card Sorting Test: a functional MRI study on healthy volunteers and schizophrenics. *Psychiatry Research*, 75, 3145-57.
- Wayne W.D. (1996). *Bioestadística. Base para el análisis de las ciencias de la salud*. (3era. edición). Editorial Limusa, S.A. de C.V., México.
- Weinberg, D.R., Berman, K.F., y Zec, R.F. (1986). Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia: I. Regional cerebral blood flow (rCBF) evidence. *Archives of General Psychiatry*, 43, 114-124.
- Weiss, S.M., y Kulikowski, C.A. (1991). *Computer Systems that learn*. Morgan Kaufmann Publishers, Inc., San Francisco, California.
- Welsh, M.C., y Pennington, B.F. (1988). Assessing frontal lobe functioning in children: Views from developmental psychology. *Developmental Psychology*, 4, 199-230.
- Winston, P.H (1994). *Inteligencia Artificial*. Tercera edición, Addison-Wesley Iberoamerica, Delaware, U.S.A.

CUCBA**BIBLIOTECA CENTRAL**

X ANEXOS

X.1 RESULTADOS DEL PROCESO DE PRUEBA DEL SISTEMA

Tabla 10. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separado por bandas, en el par de los grupos RE y NR.

RESISTENTES y NO RESISTENTES (POR BANDAS)				
MATRIZ	% Estimado de Error Verdadero	Error en los grupos	Variables identificadas como principales (tronco de árbol) a lo largo de las 19 pruebas	Variable(s) en el tronco al formar el árbol con todos los sujetos.
Potencia Absoluta con ojos Abiertos	32	RE = 44 NR = 20	Alfa1 F1 (en 11/19) Beta1 T4 (en 6/19) Alfa1 C3 (en 1/19) Alfa 2 O1 (en 1/19) Beta2 F7 (en 1/19)	Beta1 T4.
Potencia Absoluta con ojos Cerrados	37	RE = 33 NR = 40	Alfa1 C3 (en 18/19) Alfa1 C4 (en 1/19)	Alfa1 C3
Potencia Relativa con ojos Abiertos	32	RE = 33 NR = 30	Alfa1 F8 (en 6/19) Delta F1 (en 5/19) Delta T4 (en 5/19) Delta F8 (en 2/19) Alfa1 O2 (en 1/19)	Alfa1 F8
Potencia Relativa con ojos Cerrados	32	RE = 44 NR = 20	Delta T4 (en 17/19) Alfa1 C3 y T4 (en 1/19) c/u	Delta T4
Correlación Inter. con ojos Abiertos	Descartada por falta de variables con diferencia significativa.			
Correlación Inter. con ojos Cerrados	79	RE = 88 NR = 70	Delta F3F8 (en 8/19) Theta1 F1F2 (en 7/19) Beta1 C3C4 (en 2/19) Alfa1 F1F2 (en 1/19) Alfa2 F3F4 (en 1/19)	Delta F3F4
Correlación Intra con ojos Abiertos	63	RE = 44 NR = 80	Theta1 C4O2 (en 5/19) Alfa1 F1F7 (en 3/19) Alfa1 C4O2 (en 3/19) Alfa2 F7O1 (en 3/19) Delta F3C3 (en 1/19) Delta C4O2 (en 1/19) Theta1 F8T4 (en 1/19) Alfa1 F2F8 (en 1/19) Beta2 F4C4 (en 1/19)	Theta1 C4O2
Correlación Intra con ojos Cerrados	58	RE = 77 NR = 40	Theta2 C3O1 (en 15/19) Theta1 C3T3 (en 2/19) Alfa2 C4O2 (en 1/19) Beta2 F7T3 (en 1/19)	Theta2 C3O1

Tabla 11. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separada por frecuencias, en el par que contiene a los grupos RE y NR.

RESISTENTES y NO RESISTENTES (POR FRECUENCIAS)				
MATRIZ	% Estimado de Error Verdadero	% Error en los grupos	Variables identificadas como principales (tronco de árbol) a lo largo de las 19 pruebas	Variable(s) en el tronco al formar el árbol con todos los sujetos.
Potencia Absoluta con ojos Abiertos	63	RE = 55 NR = 70	15 Hz T4 (en 4/19) 19 Hz C3 (en 3/19) 24 Hz T4 (en 2/19) 19 Hz F8 (en 2/19) 19 Hz F7 (en 1/19) 19 Hz T3 (en 1/19) 18 Hz F4 (en 1/19) 18 Hz T4 (en 1/19) 12 Hz O1 (en 1/19) 11 Hz C3 (en 1/19) 22 Hz C4 (en 1/19) 24 Hz C4 (en 1/19) 9 Hz F2 (en 1/19) 9 Hz F4 (en 1/19)	15 Hz T4
Potencia Absoluta con ojos Cerrados	21	RE = 22 NR = 20	10 Hz C3 (en 18/19) 17 Hz T4 (en 1/19)	10 Hz C3
Potencia Relativa con ojos Abiertos	21	RE = 33 NR = 10	16 Hz F4 (en 17/19) 17 Hz F3 (en 1/19) 2 Hz T4 (en 1/19)	16 Hz F4
Potencia Relativa con ojos Cerrados	32	RE = 33 NR = 40	20 Hz O1 (en 8/19) 9 Hz C3 (en 4/19) 3 Hz F7 (en 2/19) 20 Hz O2 (en 1/19) 19 Hz F3 (en 1/19) 14 Hz F3 (en 1/19) 7 Hz F2 (en 1/19) 2 Hz T4 (en 1/19)	20 Hz O1
Correlación Inter. con ojos Abiertos	68	RE = 66 NR = 70	16 Hz F3F4 (en 7/19) 24 Hz F1F2 (en 2/19) 20 Hz T3T4 (en 2/19) 9 Hz O1O2 (en 2/19) 5 Hz F3F4 (en 2/19) 22 Hz O1O2 (en 1/19) 17 Hz F7F8 (en 1/19) 4 Hz F1F2 (en 1/19) 2 Hz T3T4 (en 1/19)	16 Hz F3F4
Correlación Inter. con ojos Cerrados	58	RE = 44 NR = 70	5 Hz F1F2 (en 11/19) 12 Hz T3T4 (en 2/19) 10 Hz F7F8 (en 2/19) 8 Hz F1F2 (en 2/19) 11 Hz F3F4 (en 1/19) 11 Hz F7F8 (en 1/19) 4 Hz F3F4 (en 1/19)	5 Hz F1F2

Correlación Intra con ojos Abiertos	53	RE = 44 NR = 60	18 Hz F4O2 (en 14/19) 22 Hz F4C4 (en 2/19) 15 Hz F4F8 (en 1/19) 5 Hz C3O1 (en 1/19) 4 Hz F2F4 (en 1/19) 3 Hz F8F4 (en 1/19)	18 Hz F4O2
Correlación Intra con ojos Cerrados	58	RE = 66 NR = 60	21 Hz C4O2 (en 4/19) 7 Hz C4O2 (en 3/19) 6 Hz F1O1 (en 3/19) 21 Hz C4O2 (en 2/19) 6 Hz C4O2 (en 2/19) 24 Hz F2F4 (en 1/19) 18 Hz T3O1 (en 1/19) 11 Hz C4O2 (en 1/19) 8 Hz C3O1 (en 1/19) 6 Hz F7O1 (en 1/19)	21 Hz C4O2

Tabla 12. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separada por bandas, en el par que contiene a los grupos CO y RE.

CONTROLES y RESISTENTES (POR BANDAS)				
MATRIZ	% Estimado de Error Verdadero	% Error en los grupos	Variables identificadas como principales (tronco de árbol) a lo largo de las 19 pruebas	Variable(s) en el tronco al formar el árbol con todos los sujetos.
Potencia Absoluta con ojos Abiertos	Descartada por falta de variables con diferencia significativa.			
Potencia Absoluta con ojos Cerrados	Descartada por falta de variables con diferencia significativa.			
Potencia Relativa con ojos Abiertos	28	RE = 11 NR = 44	Theta1 F1 (en 14/18) Beta2 T4 (en 1/18) Beta1 C3 (en 1/18) Beta1 F1 (en 1/18) Beta2 T3 F7 (en 1/18) c/u	Theta1 F1
Potencia Relativa con ojos Cerrados	28	RE = 11 NR = 44	Beta1 F3 (en 11/18) Beta2 C4 (en 6/18) Delta T3 (en 1/18)	Beta1 F3
Correlación Inter. con ojos Abiertos	78	RE = 67 NR = 89	Beta2 F1F2 (en 7/18) Theta2 F1F2 (en 2/18) Delta T3T4 (en 2/18) Theta1 T3T4 (en 2/18) Delta C3C4 (en 2/18) Beta1 F7F8 (en 1/18) Alfa2 F1F2 (en 1/18) Theta1 F1F2 (en 1/18)	Delta C3C4
Correlación Inter. con ojos Cerrados	Descartada por falta de variables con diferencia significativa.			
Correlación Intra con ojos Abiertos	61	RE = 55 NR = 67	Alfa1 F3F7 (en 6/18) Delta F2F4 (en 2/18) Beta2 F3O1 (en 1/18) Beta2 C4T4 (en 1/18) Beta2 C3T3 (en 1/18) Beta2 T4O2 (en 1/18) Alfa2 C3O1 (en 1/18) Alfa1 C4T4 (en 1/18) Delta F1T3 (en 1/18) Theta1 F2F4 (en 1/18) Theta1 F4O2 (en 1/18) Theta1 T4O2 (en 1/18)	Beta2 F2F4
Correlación Intra con ojos Cerrados	61	RE = 44 NR = 78	Beta2 T4O2 (en 6/18) Alfa2 F1T3 (en 4/18) Alfa2 F1F3 (en 2/18) Beta1 F4O2 (en 1/18) Alfa2 F1T3 (en 1/18) Alfa1 F2O2 (en 1/18) Delta F1T3 (en 1/18) Theta2 F1O1 (en 1/18) Theta1 F2F3 (en 1/18)	Beta1 F4O2

Tabla 13. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separada por frecuencias, en el par que contiene a los grupos CO y RE.

CONTROLES y RESISTENTES (POR FRECUENCIAS)				
MATRIZ	% Estimado de Error Verdadero	% Error en los grupos	Variables identificadas como principales (tronco de árbol) a lo largo de las 19 pruebas	Variable(s) en el tronco al formar el árbol con todos los sujetos.
Potencia Absoluta con ojos Abiertos	44	RE = 5 NR = 33	20 Hz C4 (en 5/18) 24 Hz C4 (en 4/18) 19 Hz T3 (en 3/18) 2 Hz F2 (en 2/18) 17 Hz C4 (en 1/18) 19 Hz C3,4 y F7(en 1/18) c/u 24 Hz T4 (en 1/18)	19 Hz T3
Potencia Absoluta con ojos Cerrados	Descartada por falta de variables con diferencia significativa.			
Potencia Relativa con ojos Abiertos	44	RE = 33 NR = 55	22 Hz C3 (en 12/18) 8 Hz F1 (en 2/18) 6 Hz O2 (en 1/18) 7 Hz O2 (en 1/18) 15 Hz F3 (en 1/18) 18 Hz T3 (en 1/18) 20 Hz F7 y C3 (en 1/18) c/u 23 Hz T4 (en 1/18)	22 Hz C3
Potencia Relativa con ojos Cerrados	33	RE = 22 NR = 44	20 Hz F2 (en 14/18) 7 Hz C3 (en 1/18) 16 Hz F3 (en 1/18) 17 Hz F4 (en 1/18) 18 Hz F2 (en 1/18) 19 Hz F2 (en 1/18)	18 Hz F2
Correlación Inter. con ojos Abiertos	55	RE = 33 NR = 78	24 Hz F1F2 (en 11/18) 19 Hz T3T4 (en 4/18) 2 Hz C3C4 (en 1/18) 7 Hz F1F2 (en 1/18) 21 Hz F1F2 (en 1/18)	24 Hz F1F2
Correlación Inter. con ojos Cerrados	28	RE = 33 NR = 22	3 Hz C3C4 (en 16/18) 4 Hz F3F4 (en 1/18) 13 Hz T3T4 (en 1/18)	3 Hz C3C4
Correlación Intra con ojos Abiertos	22	RE = 22 NR = 22	23 Hz C4T4 (en 9/18) 4 Hz F1C3 (en 6/18) 23 Hz T4O2 (en 3/18) 25 Hz C3O1 (en 1/18)	23 Hz T4O2

Correlación Intra con ojos Cerrados	72	RE = 78 NR = 67	6 Hz C4O2 (en 3/18) 11 Hz F1T3 (en 2/18) 15 Hz F8O2 (en 2/18) 20 Hz C4O2 (en 2/18) 3 Hz C3T3 (en 1/18) 6 Hz C3O1 (en 1/18) 8 Hz F2F4 (en 1/18) 11 Hz C3O1 (en 1/18) 12 Hz F3F7 (en 1/18) 19 Hz F4O2 (en 1/18) 21 Hz C4O2 (en 1/18) 23 Hz C4O2 (en 1/18) 25 Hz F2O2 (en 1/18)	11 Hz F1T3
--	----	--------------------	--	------------

Tabla 14. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separada por bandas, en el par que contiene a los grupos CO y NR.

CONTROLES y NO RESISTENTES (POR BANDAS)				
MATRIZ	% Estimado de Error Verdadero	% Error en los grupos	Variables identificadas como principales (tronco de árbol) a lo largo de las 19 pruebas	Variable(s) en el tronco al formar el árbol con todos los sujetos.
Potencia Absoluta con ojos Abiertos	Descartada por falta de variables con diferencia significativa.			
Potencia Absoluta con ojos Cerrados	16	CO= 11 NR= 20	Alfa1 O2 (en 12/19) Alfa1 O1 (en 6/19) Delta T4 (en 1/19)	Alfa1 O2
Potencia Relativa con ojos Abiertos	53	CO= 44 NR= 60	Alfa1 F8 (en 8/19) Theta1 F2 (en 3/19) Theta2 T3 (en 2/19) Beta1 C3 (en 1/19) Theta1 F4 (en 1/19) Theta2 F1 (en 1/19) Theta2 F3 (en 1/19) Delta F1 (en 1/19) Delta F3 (en 1/19)	Theta1 F4 Theta2 F1
Potencia Relativa con ojos Cerrados	32	CO= 0 NR= 60	Beta1 F3 (en 9/19) Beta2 F3 (en 5/19) Beta1 O2 (en 1/19) Beta2 F1 (en 1/19) Delta F3 (en 1/19) Delta T4 (en 1/19)	Beta1 F3
Correlación Inter. con ojos Abiertos	21	CO= 11 NR= 30	Theta2 F1F2 (en 10/9) Alfa2 F1F2 (en 8/19) Theta1 C3C4 (en 1/19)	Theta2 F1F2
Correlación Inter. con ojos Cerrados	37	CO= 44 NR= 30	Delta O1O2 (en 10/19) Beta1 F1F2 (en 4/19) Theta2 F1F2 (en 3/19) Theta1 F1F2 (en 2/19)	Delta O1O2
Correlación Intra con ojos Abiertos	21	CO= 22 NR= 20	Beta2 F2T4 (en 16/19) Delta F3O1 (en 3/19) Beta1 C3T3 (en 1/19)	Beta2 F2T4
Correlación Intra con ojos Cerrados	63	CO= 55 NR= 80	Alfa1 F2O2 (en 8/19) Beta2 T4O2 (en 2/19) Theta2 F1O1 (en 2/19) Beta2 F7C3 (en 1/19) Alfa1 F4O2 (en 1/19) Theta2 F1T3 (en 1/19) Theta2 F3O1 (en 1/19) Theta2 F7O1 (en 1/19) Theta1 F2C4 (en 1/19) Delta F2F8 (en 1/19) Delta F7C3 (en 1/19)	Theta2 F1T3

Tabla 15. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separada por frecuencias, en el par que contiene a los grupos CO y NR.

CONTROLES y NO RESISTENTES (POR FRECUENCIAS)				
MATRIZ	% Estimado de Error Verdadero	% Error en los grupos	Variables identificadas como principales (tronco de árbol) a lo largo de las 19 pruebas	Variable(s) en el tronco al formar el árbol con todos los sujetos.
Potencia Absoluta con ojos Abiertos	47	CO= 33 NR= 60	4 Hz T3 (en 13/19) 2 Hz O2 (en 1/19) 3 Hz F2 (en 1/19) 10 Hz O2 (en 1/19) 13 Hz F1 (en 1/19) 13 Hz F7 (en 1/19) 21 Hz F2 (en 1/19)	4 Hz T3
Potencia Absoluta con ojos Cerrados	47	CO= 44 NR= 50	10 Hz O2 (en 9/19) 4 Hz T3 (en 5/19) 2 Hz T4 (en 3/19) 3 Hz T3 (en 1/19) 9 Hz C3 (en 1/19)	4 Hz T3
Potencia Relativa con ojos Abiertos	42	CO= 44 NR= 40	4 Hz C3 (en 8/19) 9 Hz F8 (en 6/19) 19 Hz F3 (en 3/19) 2 Hz T4 (en 1/19) 5 Hz F4 (en 1/19) 16 Hz F4 (en 1/19)	4 Hz C3
Potencia Relativa con ojos Cerrados	47	CO= 33 NR= 60	16 Hz F3 (en 10/19) 6 Hz F7 (en 5/19) 2 Hz T4 (en 1/19) 15 Hz F3 (en 1/19) 16 Hz O1 (en 1/19) 17 Hz O2 (en 1/19) 20 Hz F1 (en 1/19)	16 Hz F3
Correlación Inter. con ojos Abiertos	32	CO= 44 NR= 20	12 Hz F1F2 (en 12/19) 24 Hz F1F2 (en 5/19) 11 Hz F1F2 (en 1/19) 20 Hz T3T4 (en 1/19) 25 Hz F7F2 (en 1/19)	12 Hz F1F2
Correlación Inter. con ojos Cerrados	37	CO= 33 NR= 40	25 Hz F3F4 (en 13/19) 12 Hz T3T4 (en 3/19) 6 Hz F1F2 (en 1/19) 8 Hz F7F8 (en 1/19) 17 Hz F1F2 (en 1/19)	25 Hz F3F4
Correlación Intra con ojos Abiertos	26	RE = 22 NR = 30	24 Hz F8T4 (en 7/19) 25 Hz F2T4 (en 6/19) 4 Hz F7C3 (en 2/19) 3 Hz F1T3 (en 1/19) 10 Hz F4F8 (en 1/19) 12 Hz F1F7 (en 1/19) 16 Hz F2F8 (en 1/19) 25 Hz T4O2 (en 1/19)	25 Hz F2T4

Correlación Intra con ojos Cerrados	47	RE = 33 NR = 60	16 Hz F8T4 (en 11/19) 13 Hz F4F8 (en 3/19) 22 Hz F8O2 (en 1/19) 6 Hz F2O2 (en 1/19) 7 Hz C4O2 (en 1/19) 8 Hz F1O1 (en 1/19) 8 Hz F3O1 (en 1/19) 13 Hz F1T3 (en 1/19) 15 Hz F4T4 (en 1/19) 18 Hz T3O1 (en 1/19)	16 Hz F8T4
--	----	--------------------	---	------------

X.2 INDICE DE FIGURAS

Figura 1. Algunas aplicaciones de la Inteligencia Artificial. _____	7
Figura 2. Descomposición de una señal en sus bandas componentes, mediante la Transformada Rápida de Fourier. _____	30
Figura 3. Ciclo continuo de evaluación y desarrollo del proyecto de construcción del Sistema Experto ESQUIZOR. _____	44
Figura 4. Diagrama de flujo del modelo de funcionamiento modular del sistema Esquizor. _____	46
Figura 5. Conformación de la Interfaz gráfica de usuario del programa Esquizor. _____	50
Figura 6. Pantalla principal del Sistema Experto Esquizor. _____	51
Figura 7. La cortina vertical que se despliega al pulsar el menú <i>Archivo</i> . _____	51
Figura 8. La ventana Open del Sistema de Windows. _____	52
Figura 9. Presentación en pantalla de los datos de una matriz. _____	53
Figura 10. La ventana <i>Configurar Matriz</i> . _____	53
Figura 11. Información del programa. _____	54
Figura 12. La Subcarpeta de Sujeto en la carpeta de Matriz. _____	56
Figura 13. La Carpeta de Acción Ejecutada. _____	57
Figura 14. La Carpeta Arbol Generado. _____	58
Figura 15. Distribución de los sujetos y las variables en una matriz. _____	64

- Figura 16.**Arbol tomado de Russell y Norvig (1996 pp. 565). Las líneas, círculos y texto rojos han sido agregados para clarificar los Nodos, las Ramas y sus niveles. Cada nodo es una variable y las ramas que se desprenden de la misma representan los diferentes valores que los sujetos de la matriz tienen en dicha variable. El nivel de profundidad de un árbol se determina por el número de niveles de nodos y su amplitud por el promedio de ramas por nodo. _____ 72
- Figura 17.**Arbol generado por el Sistema. Las líneas, círculos y texto rojos son nuestros. _____ 73
- Figura 18.**Distribución de los valores de los sujetos de los grupos RE y NR en la variable *Categorías Completadas*. El desplazamiento en el eje X se realizó con el fin de identificar mejor a cada sujeto, no representa valor alguno. _____ 87
- Figura 19.**Distribución de los valores de los sujetos de los grupos RE y NR en la variable *% de Errores Perseverativos*. El desplazamiento en el eje X se realizó con el fin de identificar mejor a cada sujeto, no representa valor alguno. _____ 88
- Figura 20.**Distribución de los valores de los sujetos de los grupos RE y NR en la variable *10 Hz en C3 (Potencia Absoluta con ojos Cerrados)*. El desplazamiento en el eje X se realizó con el fin de identificar mejor a cada sujeto, no representa valor alguno. _____ 91
- Figura 21.**Arbol generado a partir de la matriz de potencia absoluta por frecuencia con ojos cerrados. _____ 92
- Figura 22.**Distribución de los valores de los sujetos de los grupos RE y NR en la variable *16 Hz en F4 (Potencia Relativa con ojos Abiertos)*. El desplazamiento en el eje X se realizó con el fin de identificar mejor a cada sujeto, no representa valor alguno. _____ 94
- Figura 23.**Arbol generado de la matriz de potencia relativa por frecuencia con ojos abiertos. _____ 94

Figura 24. Distribución de los valores de los sujetos de los grupos CO y RE en la variable *23 Hz correlación entre T4 y O2 con ojos abiertos*. El desplazamiento en el eje X se realizó con el fin de identificar mejor a cada sujeto, no representa valor alguno. _____ 96

Figura 25. Distribución de los valores de los sujetos de los grupos CO y NR en la variable *Potencia Absoluta de Alfa1 en O2 con ojos cerrados*. El desplazamiento en el eje X no representa valor alguno. _____ 97

X.3 INDICE DE TABLAS

Tabla 1.	Datos del ejemplo presentado por Russell y Norvig, (1996 pp. 565).	71
Tabla 2.	Resultados generales de las pruebas del Sistema. Presenta el <i>Estimado de Error Verdadero</i> en porcentaje (primer número en las casillas) y el porcentaje de error en los grupos (siguientes dos renglones en cada casilla RE= Resistentes, NR= No Resistentes) La celda gris indica el método que obtuvo el mejor porcentaje de error.	79
Tabla 3.	Resultados generales de las pruebas del Sistema. Resultados del <i>Estimado de Error Verdadero</i> en porcentaje (primer número en las casillas) y el porcentaje de error en los grupos (siguientes dos renglones en cada casilla). El “*” indica que por lo menos en un sujeto de esa matriz no se encontró ninguna variable con diferencia intergrupala significativa, y que no se valoró dicha matriz. Las celdas grises indican las matrices que obtuvieron menor porcentaje de error.	85
Tabla 4.	Promedio de las variables seleccionadas por prueba, en cada matriz. El “*” indica que esa matriz no se valoró	90
Tabla 5.	Variables seleccionadas para cada prueba de Regresión. **= < 0.01, *= < 0.05, G= Mayor Ganancia, Celda Gris= No Probado.	102
Tabla 6.	Variables seleccionadas para cada prueba de Regresión. **= < 0.01, *= < 0.05, G= Mayor Ganancia, Celda Gris= No Probado.	103
Tabla 7.	Resultados de las pruebas de clasificación con RLM sobre la matriz de Wisconsin.	104

Tabla 8. Resultados de las pruebas de clasificación realizadas con Regresión Lineal Múltiple sobre la matriz de Potencia Absoluta por Frecuencia con ojos cerrados en el par RE y NR. _____	105
Tabla 9. Resultados de las pruebas de clasificación realizadas con Regresión Lineal Múltiple sobre la matriz de Potencia Absoluta por Frecuencia con ojos cerrados en el par RE y NR. _____	106
Tabla 10. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separado por bandas, en el par de los grupos RE y NR. _____	123
Tabla 11. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separada por frecuencias, en el par que contiene a los grupos RE y NR. _____	124
Tabla 12. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separada por bandas, en el par que contiene a los grupos CO y RE. _____	126
Tabla 13. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separada por frecuencias, en el par que contiene a los grupos CO y RE. _____	127
Tabla 14. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separada por bandas, en el par que contiene a los grupos CO y NR. _____	129
Tabla 15. Resultados de las pruebas del Sistema sobre las matrices de datos del EEG separada por frecuencias, en el par que contiene a los grupos CO y NR. _____	130