



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO
CENTRO UNIVERSITARIO UAEM TEXCOCO



**“IMPLEMENTACIÓN DE UNA RED NEURONAL ARTIFICIAL PARA
PRE-DIAGNOSTICAR EL SÍNDROME DE CUSHING”**

TESIS

**QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS DE LA COMPUTACIÓN**

PRESENTA:

MÁRQUEZ SÁNCHEZ LUIS

TUTOR ACADÉMICO

DR. ADRIAN TRUEBA ESPINOSA

TUTORES ADJUNTOS

DRA. CRISTINA JUAREZ LANDIN

M. en C. MACARIO HERNANDEZ CRUZ

TEXCOCO ESTADO DE MÉXICO

JULIO 2011



Universidad Autónoma del Estado de México
UAEM



DICTÁMEN PARA AUTORIZACIÓN DE GRADO DE MAESTRÍA

Texcoco, Méx. , a 30 de junio de 2011.



TÍTULO DEL PROYECTO:



IMPLEMENTACIÓN DE UNA RED NEURONAL ARTIFICIAL PARA PRE-DIAGNOSTICAR EL SÍNDROME DE CUSHING.

TESISTA: L.I.A. MÁRQUEZ SÁNCHEZ LUIS

COPIA



DICTAMEN:

NO. DE REVISIÓN: 3

- RECHAZADO
- SUJETO A MODIFICACIONES
- ACEPTADO, CONDICIONADO
- ACEPTADO



OBSERVACIONES GENERALES: Sin observaciones
Aceptado para la impresión
Aceptado para la defensa de grado



Tutor Académico DR. ADRIAN TRUEBA ESPINOSA	Tutor Adjunto DRA. CRISTINA JUAREZ LANDIN	Tutor Adjunto M. EN C. MACARIO HERNANDEZ CRUZ
--	---	---



“Solo requiero de un momento de un poco de confianza, un tanto de inspiración, un momento de lucidez y algo de espontaneidad para crear una leyenda” SL21

AGRADECIMIENTOS

Gracias a dios por darme la mamá más grandiosa del mundo ya que sin ellos y sin su apoyo incondicional nada de lo que soy ahora sería posible. Al igual que a mi hermana por la inspiración a ser un ejemplo a seguir.

Gracias al Dr. Valentín Sánchez Pedraza por el apoyo académico en la parte médica y por ser un excelente ejemplo a seguir.

Gracias al Dr. Adrian Trueba Espinosa por su apoyo incondicional y comprensión de mi forma de trabajo.

Gracias a mis tutores adjuntos por su tiempo, dedicación, conocimiento y apoyo brindado.

Gracias al COMECYT por el apoyo y confianza en mi proyecto de tesis.

Gracias a todos mis amigos que me brindaron su apoyo y animaron a concluir con el presente.

Gracias a todas y cada una de las mujeres que en su momento me brindaron uno o más momentos de inspiración para continuar escribiendo y no desfallecer en el proceso.

DEDICATORIA

A mi hija Yolotzin Márquez Hernandez

RESUMEN

El Síndrome de Cushing es una de las enfermedades endocrinológicas más complicadas de diagnosticar debido a que la sintomatología que presenta no es exclusiva de la enfermedad, además de no poseer un orden específico de aparición de la sintomatología asociada, el 60% de los casos pre-diagnosticados inicialmente no padecen dicha enfermedad. Con la finalidad de abordar la problemática del SC, se implementó una Red Neuronal Artificial (RNA) multicapa para realizar el pre-diagnostico del SC con el algoritmo *backpropagation*.

Del 100% de los casos presentados a la RNA tiene un promedio 94.43% de clasificación correcta y un error del 5.57% de error de clasificación con lo que se tendría alrededor de un 95% de ahorro en estudios desechados. Por lo tanto se incrementó la clasificación de los mismos hasta un 45% más que los médicos generales que atienden a los pacientes en primera instancia, con lo que se disminuyen en un 95% los gastos del hospital en estudios químicos.

ABSTRACT

Cushing's syndrome is one of the most complicated endocrinological diseases to diagnose, because the symptoms that is not unique to the disease, and of not having a specific order of appearance of the associated symptoms, 60% of cases pre-initially diagnosed as suffering from the disease. In order to address the problems of SC, we implemented an artificial neural network (ANN) multilayer for the pre-diagnosis of SC with the backpropagation algorithm.

100% of the cases presented to the RNA has an average 94.43% correct classification and an error of 5.57% classification error with what would have about 95% savings in studies discarded. Therefore increased their classification up to 45% more than general practitioners who treat patients in the first instance, which is a 95% decrease in hospital costs in chemical studies.

CONTENIDO

	Pág.
1. Introducción	6
2. Planteamiento del problema	8
3. Justificación	9
4. Objetivo General	10
4.1 Objetivos Particulares	10
5. Delimitación	10
6. Hipótesis	11
7. Marco Teórico	12
7.1 Introducción al Síndrome de Cushing	12
7.2 Definición Síndrome de Cushing	15
7.3 Situaciones Especiales	22
7.4 Algoritmo de diagnóstico	23
7.5 Síntomas	27
7.6 Reconocimiento de Patrones	30
7.7 Redes Neuronales	36
7.8 Funcionamiento de Neuronas Biológicas	37
7.9 Funcionamiento RNA	38
7.10 Estructura	39
7.11 Características de operación de las RNA	42
7.12 Tipos de RNA	45

7.13 <i>Backpropagation</i>	61
8. Trabajos relacionados	65
9. Metodología	67
9.1 Pasos metodológicos	67
9.2 Pasos metodológicos	68
10. Resultados	73
11. Discusión	81
12. Conclusiones	84
13. Bibliografía	86
14. Anexo	91

ÍNDICE DE FIGURAS

Figura	Pág.
7.0 Paciente de Cushing	14
7.1 Glándula pituitaria y adrenal	16
7.2 Algoritmo de diagnóstico del SC	24
7.3 Neurona Biológica	37
7.4 Perceptrón con dos dendritas	38
7.5 Axones y sinapsis	45
7.6 Esquema Neuronal de Hopfield	52
7.7 Esquema Neuronal de Kohonen	56
7.8 Vista de caja negra de una red neuronal	62
7.9 Vista interna de una red neuronal sin capas ocultas	62
7.10 Esquema de entrenamiento de una red neuronal	63
7.11 Esquema Neuronal <i>Backpropagation</i>	64
7.12. Gráfica de porcentaje de clasificación	74
7.13. Valores de los errores en épocas de entrenamiento inicial	75
7.14. Valores de los errores en épocas de entrenamiento final	77
7.15. Diferencia entre la salida esperada y la salida generada	77
7.16. <i>K - Fold Cross Validation</i>	79
7.17 Representación de porcentajes de clasificación para entrenamiento, validación y prueba.	80

ÍNDICE DE CUADROS

Cuadro	Pág.
7.1 Escenario clínico del SC	19
7.2 Índice clínico para Grado de severidad en SC	21
7.3. Manifestación Clínica	29
7.4 Media y desviación estándar sobre error de entrenamiento	76
7.5. Resultados obtenidos en la clasificación.	79

1. INTRODUCCIÓN

Diversas disciplinas, como la medicina, metalurgia, química, mecánica, física entre otras, encuentran zonas de convergencia con la informática para incrementar la efectividad en sus estudios y mejorar sus resultados. Los programas computacionales que tratan de emular la inteligencia humana constituyen una muestra de la integración de diversas ciencias y áreas del conocimiento. Su perfeccionamiento futuro permitirá una asistencia apropiada al médico en la toma de decisiones y en el especial a quienes atienden el diagnóstico de enfermedades de difícil identificación.

El desarrollo de tecnologías asociadas con las técnicas de inteligencia artificial (IA), aplicadas a la medicina, representa una novedosa perspectiva, que puede reducir los costos, el tiempo, los errores médicos; así como potenciar el uso de los recursos humanos en las ramas médicas con mayores requerimientos. Y es precisamente donde la presente investigación encuentra vertiente para poder aplicar las redes neuronales y la identificación de patrones permitiendo realizar un pre- diagnóstico del síndrome de Cushing a partir de los datos obtenidos en el diagnóstico diferencial.

Las redes neuronales artificiales han incursionado en diversos campos como la medicina, la química, la física, estudios nucleares y economía entre otras. Permitiendo mejorar los procesos de clasificación e identificación para cada una de las áreas.

El progreso por el uso de los medios informáticos durante los últimos años constituye un indicador del nivel de integración, la utilidad y el papel excepcional que desempeñan las

computadoras en el mundo contemporáneo y en particular, en la esfera médica, debido al veloz procesamiento de los datos y ayuda en la toma efectiva de decisiones.

El amplio espectro en el uso de redes neuronales artificiales en el campo médico, mejorando el tiempo de entrenamiento y en muchos casos ha incrementado el porcentaje de clasificación para las diversas patologías. La presente investigación se enfoca en el estudio del Síndrome de Cushing para poder implementar una red neuronal artificial que ayude en el diagnóstico del Síndrome.

2. **PLANTEAMIENTO DEL PROBLEMA**

La identificación del Síndrome de Cushing (SC) por parte de los médicos familiares es complicada debido a que la mayoría de los síntomas asociados a la enfermedad como las estrías, obesidad central, cara de luna llena, plétora facial, hirsutismo, joroba de búfalo, acné y la alopecia, no son asociados únicamente a este síndrome, también se presentan en ovarios poliquísticos, lupus, obesidad etc. por lo que en la mayoría de los casos los pacientes son enviados con el especialista para tratar este padecimiento y aproximadamente el 60% de los casos pre-diagnosticados inicialmente no padecen dicha enfermedad.

Un pre-diagnóstico erróneo del Síndrome de Cushing, representa costos económicos para el hospital, debido a los estudios químicos previos a la cita con el endocrinólogo. Cuando no se corrobora el padecimiento del SC, los exámenes son desechados, independientemente del gasto de materiales, energía eléctrica, material de laboratorio, espacio, uso de instalaciones y acumulación de trabajo que representa una inversión económica desaprovechada. Por lo tanto se refleja claramente que cuando los análisis son realizados y el pre-diagnostico no es confirmado, dichos estudios no son aprovechados de ninguna manera, y simplemente representan un gasto innecesario.

Por otra parte el especialista Endocrinólogo invierte tiempo al volver a realizar todo el cuadro clínico por omisiones de los médicos familiares al realizar la valoración clínica, lo que evita que atienda a un mayor número de pacientes por periodo laboral, que padecen diversas enfermedades propias de su especialidad.

3. JUSTIFICACIÓN

Considerando la problemática que genera un diagnóstico errado del SC por parte de los médicos familiares, aunado a las consecuencias económicas que esto conlleva se requiere una herramienta que auxilie en el diagnóstico, y que permita prevenir gastos innecesarios.

La red neuronal artificial permitirá que los médicos cuenten con un apoyo para su trabajo y permitirá diagnosticar y dirigir con el especialista los casos pertinentes de mencionada canalización. A su vez, los especialistas podrán realizar mayor número de pruebas en un periodo de tiempo menor, empleando más tiempo para el estudio y análisis de los resultados.

La red neuronal identificará un 80% de los casos prediagnosticados con las características que permitan pre-diagnosticar el SC, y los médicos podrán realizar un mejor pre-diagnostico del SC, apoyados en los resultados otorgados por la red neuronal artificial.

4. OBJETIVO GENERAL

Implementar una red neuronal artificial para clasificar a pacientes con Síndrome de Cushing.

Objetivos Particulares

- Identificar características del Síndrome de Cushing.
- Definición del patrón de entrenamiento
- Seleccionar el tipo de red neuronal artificial a utilizar.
- Entrenar la red neuronal
- Determinación de la topología de la red neuronal que permita identificar los patrones del síndrome de Cushing.
- Modelar la red neuronal artificial.
- Probar la red neuronal artificial.
- Análisis y conclusión de los resultados

5. DELIMITACIÓN

La presente investigación se realizará con pacientes de varios hospitales del Distrito Federal, obteniendo los datos mediante el Hospital General de México, en el departamento de endocrinología, con la colaboración del jefe del servicio de Endocrinología, Dr. Valentín Sánchez Pedraza. Haciendo uso de los casos clínicos que se han detectado como parte de la histología del Síndrome de Cushing, con la finalidad de definir las características que permitan modelar la red neuronal artificial y la detección del SC.

6. HIPÓTESIS

El Síndrome de Cushing puede ser clasificado a partir las características que presentan los pacientes, utilizando una red neuronal como criterio de valoración.

7. MARCO TEÓRICO

7.1 Introducción al Síndrome de Cushing

El síndrome de Cushing también llamado hypercortisolismo, es un desorden hormonal causado por la exposición prolongada de los tejidos corpóreos a elevados niveles de la hormona de cortisol, no es un síndrome común, aparece normalmente en adultos de 20 a 50 años y afecta anualmente en unas 10 a 15 personas sobre cada millón.

Alrededor del 80% se presenta cuando el cuerpo se expone a niveles excesivos de cortisol por períodos prolongados. Muchas personas sufren los síntomas de Cushing por tomar glucocorticoides tales como prednisona contra el asma, artritis reumatoide, lupus u otras enfermedades inflamatorias o como inmunosupresión después de trasplantes.

El síndrome de Cushing por la producción de cortisol, normalmente sigue una cadena precisa de eventos. Primero el hipotálamo, que es una parte del cerebro del tamaño de un cubito de azúcar, manda una hormona adrenocorticotrofina (CRH) hacia la glándula pituitaria. Cuando estas glándulas, situadas justo encima de los riñones reciben la corticotrofina (ACT) responden soltando cortisol al circuito sanguíneo, en el momento en que la producción es excesiva el SC tiene lugar en el cuerpo humano.

Los síntomas varían, pero en la mayoría de los afectados aparece obesidad, cara redonda, aumento de grasas alrededor del cuello, adelgazamiento de brazos y piernas. Los niños tienden a la obesidad con crecimiento retardado. Otros síntomas son piel muy fina y frágil

que se cura mal en caso de tener heridas, marcas de color púrpura como estriadas aparecen en el abdomen, muslos, nalgas, brazos, y pechos. Los huesos se emblandecen y actividades rutinarias como inclinarse, levantarse de cama o de una silla pueden provocar lumbagos o incluso fracturas de costillas o de espina dorsal, también se presenta severa fatiga, músculos blandos, alta tensión y altos niveles de glucosa en sangre, así como irritabilidad, ansiedad y estados depresivos. En la Figura 7.0 se observa una representación y listado de los síntomas del SC [2007, Dorantes] [2006, Newell] .

En muchas ocasiones las mujeres padecen de exceso de vello en la cara, cuello, pechos, abdomen y muslos. Sus períodos menstruales son muy irregulares o incluso desaparecen. En los hombres existe un claro descenso de fertilidad con deseo sexual disminuido o incluso ausente.

El cortisol realiza tareas vitales en el cuerpo, mantiene la presión sanguínea y la función cardiovascular, reduciendo la respuesta inflamatoria del sistema inmunológico, mantiene en balance los efectos de la insulina en la conversión de azúcar en energía, regula el metabolismo de proteínas, carbohidratos y grasas. Una de las tareas más importantes del cortisol es ayudar al cuerpo a superar el estrés, por eso las mujeres en su tercer mes de gestación y los deportistas altamente entrenados normalmente tienen altos niveles de esta hormona en el cuerpo, gente que sufre de depresión, alcoholismo, malnutrición y desordenes de pánicos, suelen también tener los niveles de cortisol elevados.

Cuando los niveles de cortisol en sangre son los adecuados, el hipotálamo y la glándula pituitaria segregan menos CRH y ACTH. Esto asegura el equilibrio en necesidad diaria de

cortisol que segregan las glándulas suprarrenales, sin embargo si algo va mal con las suprarrenales o con sus reguladores en pituitaria y en hipotálamo la producción de cortisol puede estar mal [2007, Dorantes].

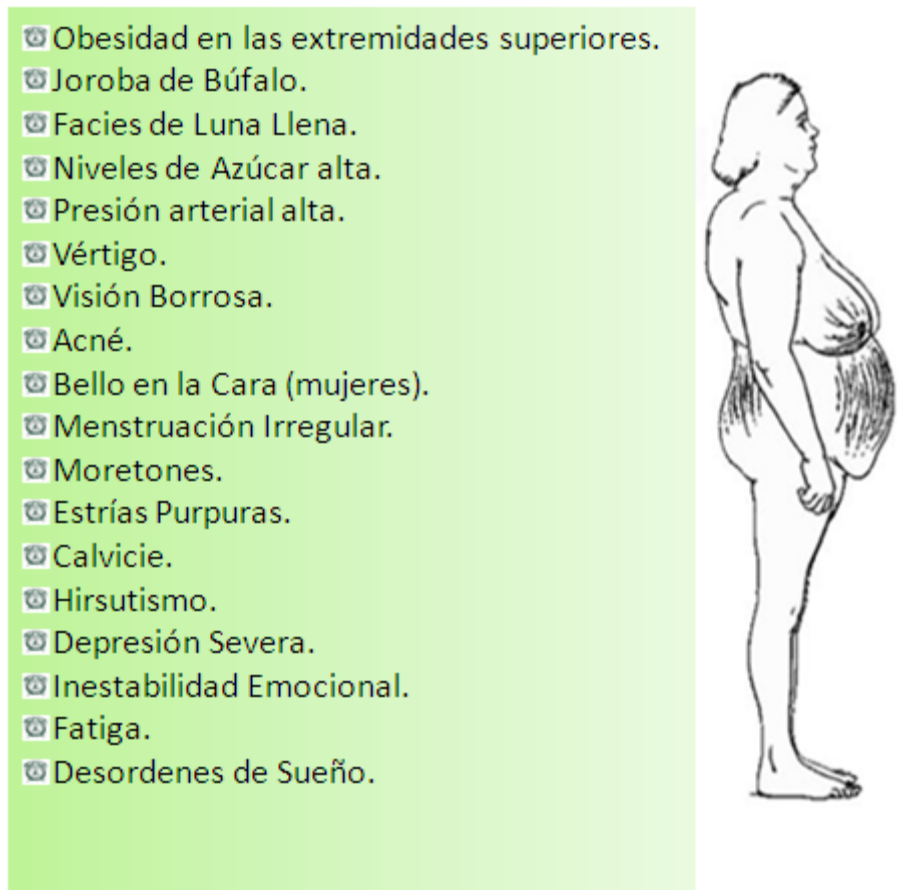


Figura 7.0 Paciente de Cushing. Fuente: [2007, Dorantes]

7.2 Definición del Síndrome de Cushing

El exceso crónico de glucocorticoides (cualquiera que sea su causa), lleva a una constelación de síntomas y características físicas conocida como síndrome de Cushing, que con mayor frecuencia, es atrójeno como resultado del tratamiento crónico con glucocorticoides y en otras ocasiones es espontaneo, provocado por anomalías en la hipófisis o suprarrenal y puede presentarse como consecuencia de la secreción de ACTH por tumores no hipofisarios [2008, Greenspan].

El síndrome de Cushing se clasifica convenientemente como dependiente de ACTH o independiente de ACTH.

Los tipos ACTH dependientes del síndrome de Cushing se caracterizan por hipersecreción crónica de ACTH, que ocasiona hiperplasia de la zona fasciculada y reticular de las suprarrenales y por tanto, aumento en la secreción suprarrenocortical de cortisol, andrógenos y DOC.

El síndrome de Cushing independiente de la ACTH puede deberse a una neoplasia suprarrenal primaria (adenoma, carcinoma ó a hiperplasia nodular suprarrenal). En este caso el exceso al rededor de cortisol suprime la secreción hipofisiaria de ACTH.

Las consecuencias clínicas del hipercortisolismo se deben a las acciones fisiológicas excesivas de los glucocorticoides. Los principales efectos fisiológicos de los glucocorticoides son: aumento en la producción de los glucocorticoides, aumento en la producción hepática de

glucosa (sobre todo, por incremento de la gluconeogénesis), inhibición de la síntesis de glucógeno, elevación de la proteólisis, disminución de la síntesis de glucógeno, elevación de la proteólisis, disminución de la síntesis proteínica y aumento de la síntesis de tejido adiposo en ciertas localizaciones (grasa intra abdominal, huecos supraclaviculares y parte alta de la espalda). Los glucocorticoides disminuyen el proceso de cicatrización(tal vez, al interferir con la síntesis proteínica). También causan resistencia a la insulina y de ese modo, favorecen la síntesis hepática de lipoproteínas de muy baja densidad (VLDL), en la Figura 7.1 se muestra un esquema de la glándula pituitaria y las adrenales.

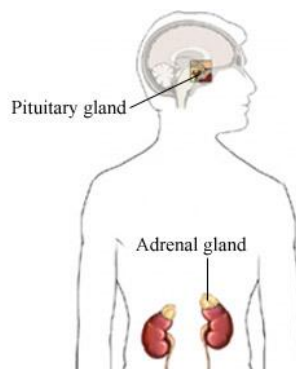


Figura. 7.1 Glándula Pituitaria y adrenal Fuente: [2003, Cortagi]

Variables del Síndrome.

- 1) **Obesidad central:** Se clasifica en base al índice central cadera (ICC) que es un indicador de la distribución de grasa corporal que permite distinguir entre la obesidad de tipo androide con predominio de la grasa en la parte superior del tronco, también conocida como (en forma de manzana o central, de la de tipo ginecoide predominante en las caderas o en forma de pera. Las medidas se realizan mediante el perímetro de la cintura a la altura de la última costilla flotante, o bien a

nivel de las crestas iliacas y el perímetro máximo de la cadera a nivel de los glúteos. Los valores normales para la población mexicana mestiza son de *0.77 m a 0.84 m* para mujeres y *0.78 m a 0.93 m* para hombres.

Las cifras más altas indican distribución de tipo androide y las menores, una distribución ginecoide.

Formula: *Índice central cadera = Cintura cm/Cadera cm*

Se ha visto que una relación entre cintura y cadera superior a 1.0 en varones y 0.8 en mujeres esta asociada a un aumento en la probabilidad de contraer diversas enfermedades (diabetes mellitus, enfermedades coronarias, hipertensión arterial, dislipidemias).

Valores normales son considerados de *menor de 95* en hombres y *menor a 85* en mujeres, obesidad central leve de *0.95 a 1.0* en hombres y de *0.85 a 0.90* en mujeres y obesidad central marcada a valores *mayores a 1.0* en hombres y *mayor a .95* en mujeres.

- 2) **Facies Cushingoide:** Cara de luna llena, pletórica con o sin hipertriosis, y obesidad central.
- 3) **Estrías:** Lesiones cutáneas que aparecen en tronco y extremidades que se observan color rojizo cuando son de reciente aparición o blanquecino cuando son lesiones de mayor tiempo de evolución.

- 4) Fragilidad de la piel: Adelgazamiento y atrofia cutánea con aparición de equimosis aún con contactos leves [2003, Cortagi].

Características del SC

El SC posee diversas características, a continuación se presentan las más destacables.

- SC se presenta como consecuencia de exposición prolongada y crónica de glucocorticoides, la causa más común es el SC Exógeno, con mayor frecuencia en las mujeres, su incidencia es de 0.7 a 2.4 por millón de población por año.

El deterioro en la calidad de vida es frecuente documentado en pacientes con SC posterior a su cura [2001, Boscoro.]

- Los pacientes con SC no controlado tienen un incremento de 5 veces en la mortalidad. En estudios de tamizaje realizados en pacientes diabéticos, especialmente aquellos con obesidad e hipertensión, sugieren que la prevalencia de SC es de 2%-5% [2003, Cortagi].
- El escenario clínico del SC es variable dependiendo de la edad, sexo y duración de la enfermedad. Los hallazgos más comunes son la obesidad central, la cual se debe a exceso de depósito de grasa visceral a nivel abdominal, asociada a adelgazamiento de las extremidades como resultado de atrofia muscular. Los hallazgos típicos son plétora facial, cara de luna llena, joroba de búfalo secundaria al depósito de grasa en la región posterior del cuello [2001, Lindholm].

- Estrías purpuras se presentan principalmente en la región lateral del abdomen, en la región axilar y parte interna de los muslos, atrofia muscular, osteoporosis y litos renales son más frecuentes en hombres. [2006, Newell] [2008, Pironello].

En el Cuadro 7.1 se muestra la prevalencia de los síntomas respecto a los casos que se han presentado, en porcentaje de prevalencia en los pacientes con dicho padecimiento, las manifestaciones clínicas más frecuentes son la obesidad central (90%) que afecta la cara, cuello, tronco y abdomen y generan la facie de luna llena y la giba de búfalo. En algunos casos el acumulo de grasa es retro-orbital produciendo exoftalmos (5 de 10 casos). A nivel de piel pueden presentarse hematomas (35%), acné (35%), estrías (50%) e hirsutismo (75%) [2006, Axelrod].

Cuadro 7.1 Escenario Síndrome de Cushing. Fuente: [2006, Axelrod]

Signos y síntomas	Prevalencia %
Obesidad central	90-100
Cara de luna llena	
Plétora facial	
Disminución lívido	
Estrías púrpuras	70-90
Alteraciones menstruales	
Hirsutismo	
Disfunción eréctil	

Hipertensión	50-70
Debilidad muscular	
Joroba de búfalo	
Fragilidad capilar (brusing)	
Diabetes o intolerancia a la glucosa	
Osteopenia u osteoporosis	
Depresión	
Cefalea	20-50
Lumbalgia	
Edema de extremidades	
Alcalosis hipocalémica	
Nefrolitiasis	
Acné	0-20
Alopecia	

Índice de Gravedad en SC (CSI)

El CSI es una medida médica (clínimétrico) que permite valorar el grado de avance que se presenta en el SC, actualmente no se cuenta con instrumentos clinimétricos para la medición de la gravedad del síndrome de cushing (SC). Solamente se cuenta con un índice de gravedad de Cushing diseñado por Sonino y Cols. Este índice fue validado en pacientes

con SC, consta de 8 variables clínicas que se cifran de 0 a 3(cuadro 7.4) con un intervalo de puntaje total de 0 a 16. La validación se realiza en pacientes con SC evaluados de manera independiente por dos endocrinólogos antes y después del tratamiento, para posteriores análisis.

Cuadro 7.2 Índice clínico para grado de severidad en SC. Fuente: [2006, Axelrod]

Ítems Cs	Ausente	Leve	Grave
Distribución de grasa			
Lesiones cutáneas			
Debilidad muscular			
Trastornos del estado de animo			
Hipertensión			
Diabetes			
Hipocalemia			
Alteraciones relacionadas al genero			

*CSI: Índice de gravedad de Cushing (Cushing syndrome serenty index)

El cuadro anterior se debe de llenar con cada paciente que presente el SC

Factores de Riesgo

Un factor de riesgo es aquello que incrementa la probabilidad de desarrollar una enfermedad o condición, como:

- Uso crónico de medicamentos corticosteroides
- Edad: 20-50 años
- Sexo: femenino (mucho más común en mujeres)

7.3 Situaciones Especiales

En niños con pubertad precoz o tardía y retardo del crecimiento con edad ósea retrasada o adelantada, la clínica del Síndrome de Cushing, es decir, la forma en que se puede presentar, tiene la posibilidad de estar ausente en pacientes con secreción ectópica de ACTH o CRH. La secreción ectópica de ACTH típicamente se presenta con hipotalámia y repercusión general, en el sexo masculino [2006, Newell] ,[2008, Greenspan] .

Clasificación del SC.

- ACTH dependiente

Adenoma hipofisario (enfermedad de Cushing): 70%

Neoplasia no hipofisaria (ACTH ectópica) :15-20%

Predominan en la mujer (8:1) 20-60 años

80-90% < 10 mm, localización hipófisis anterior.

10% macroadenomas, pueden ser invasores. Pocas veces malignos.

- ACTH independiente

Yatrogénico (glucocorticoides, acetato de megestrol)

Neoplasia suprarrenal (adenoma, carcinoma)

Hiperplasia nodular suprarrenal

Enfermedad Suprarrenal nodular pigmentada primaria

Hiperplasia Suprarrenal macronodular masiva

Dependiente de las comidas (GIP)

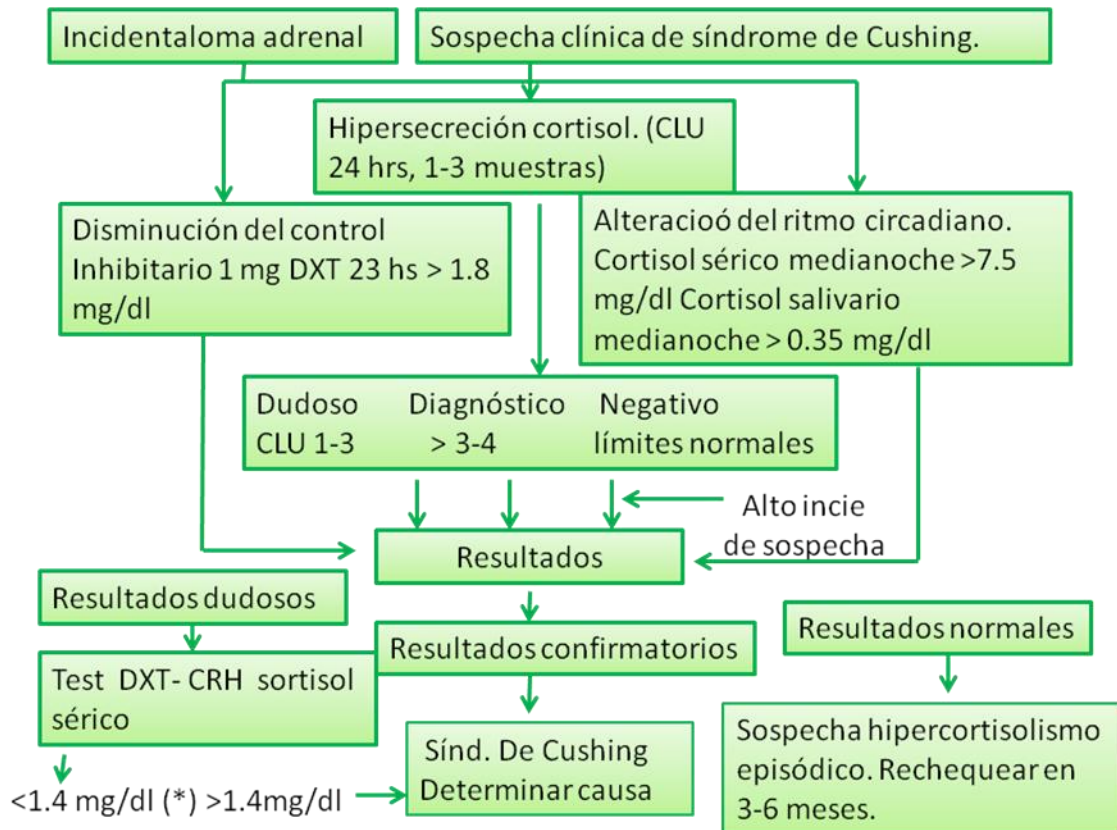
Facticio(Trastornos psicológicos y somáticos)

7.4 Algoritmo de diagnostico

Diagnóstico.

Debe sospecharse diagnóstico de síndrome de cushing ante la presencia de los datos clínicos antes mencionados. Es importante destacar la existencia de cuadros subclínicos, en los cuales las manifestaciones no son perceptibles, aunque existe adenoma suprarrenal y las pruebas dinámicas son con frecuencia anormales. El cuadro ocurre cuando se detecta por accidente un nódulo en una suprarrenal, durante la realización de un estudio de imagen de abdomen (tomografía computarizada o resonancia magnética nuclear) realizado por una

indicación médica independiente. En la Figura 7.2 se presenta el algoritmo de diagnóstico utilizado por los médicos [2003, Arnaldi], [2007, Dorantes].



(*) Probable pseudo SC: Depresión, Obesidad, Alcoholismo

Figura 7.2. Algoritmo de Diagnóstico del SC. Fuente: [2007, Dorantes]

En el estudio diagnóstico, son varias las pruebas útiles, que un médico general debe de conocer, por ejemplo: mediciones basales de cortisol sérico por la mañana (8 a.m.) y por la tarde (4 p.m.); medición de concentraciones séricas de ACTH. Esta prueba se realiza junto con la determinación de creatinina en orina de 24hrs con el propósito de comprobar que la colección de orina haya sido la adecuada. Una colección adecuada contiene entre 10 y 20 mg de creatinina urinaria/kg de peso. En general, la creatinina urinaria depende de la masa

corporal, por lo que tiende a ser mayor en hombres que en mujeres.

El médico endocrinólogo debe de cerciorarse que las mediciones de cortisol sérico y el cortisol urinario sean anormales, y luego determinar la concentración sérica de ACTH.

Las concentraciones séricas de cortisol se ajustan a un ritmo circadiano. Es normal que las concentraciones séricas medidas a las 4 p.m. sean menores (en 2/3) de las medidas a las 8 a.m. Con frecuencia en casos de síndrome de Cushing endógeno se pierde el ritmo circadiano de cortisol, además de que se elevan los valores absolutos en la mañana y la tarde. Es importante saber con qué método se está midiendo el cortisol y cual es la reproducibilidad del ensayo. En casos de síndrome de Cushing iatrogénico, las determinaciones de cortisol se encuentran bajas en forma apropiada, debido a que el eje hipotálamo –hipófisis –suprarrenal detectada una acción glucocorticoide externa(el esteroide exógeno actúa sobre sus receptores en el hipotálamo e hipófisis, inhibiendo en consecuencia, la secreción CRH – ACTH y por lo tanto, la producción endógena de cortisol) Debido a que los esteroides exógenos más utilizados en la clínica no son detectados mediante la prueba que mide cortisol, una persona con síndrome de Cushing exógeno muestra datos clínicos de hipercortisolismo, pero sus determinaciones séricas de cortisol son menores que las normales. Una excepción de esta regla ocurre cuando se utiliza hidrocortisona como esteroide exógeno.

Hidrocortisona es el nombre farmacológico asignado al cortisol cuando se le utiliza como medicamento. Así pues, desde el punto de vista químico, cortisol e hidrocortisona, puede desarrollar datos clínicos de hidropercortisolismo con valores séricos que sugieren un origen endógeno a diferencia de lo que ocurre cuando se utilizan todos los demás glucocorticoides. [2006, Hurd]

Diagnóstico diferencial

Respecto al diagnóstico diferencial del SC, en primera parte se realiza una revisión general de manera visual con el propósito de ubicar la sintomatología periférica, y descartar que se trata de algunas otras enfermedades las cuales tienen a padecer los mismos síntomas, algunas de las enfermedades en las que se presentan algunos de los síntomas son:

La obesidad exógena presenta un patrón de distribución generalizado, y la eliminación de cortisol en orina normal.

Los pacientes con alcoholismo crónico y depresión presentan cifras elevadas de cortisol en orina y resistencia a la supresión con dexametasona nocturna a dosis bajas. Al interrumpir el consumo de alcohol o mejorar el estado de ánimo estas pruebas vuelven a la normalidad. En este caso se realiza la prueba de dexametasona de dos días a dosis bajas para realizar el diagnóstico diferencial.

El estrés en estos individuos interrumpe la regulación normal de la secreción de ACTH.

Síndrome de Cushing: En la medición del cortisol basal en sangre y en orina se obtienen valores bajos por la inhibición del eje hipófisis-suprarrenal.

7.5 Síntomas

En el Síndrome de Cushing, los síntomas varían pero en la mayoría de los afectados aparece obesidad, cara redonda, aumento de grasas alrededor del cuello, adelgazamiento de brazos y piernas. Los niños tienden a la obesidad con crecimiento retardado. Otros síntomas son piel muy fina y frágil que se cura mal en caso de tener heridas. Marcas de color púrpura como estriadas aparecen en el abdomen, muslos, nalgas, brazos, y pechos. Los huesos se emblandecen y actividades rutinarias como inclinarse, levantarse de cama o de una silla pueden provocar lumbagos o incluso fracturas de costillas o de espina dorsal, también se presenta severa fatiga, músculos blandos, alta tensión y altos niveles de glucosa en sangre, así como irritabilidad, ansiedad y estados depresivos. Los síntomas más comunes incluyen cara redonda, obesidad troncal con relativo adelgazamiento de las extremidades, astenia, fatiga de músculos de las cinturas escapular y pélvica (una queja frecuente es dificultad para subir escaleras), equimosis que aparecen con traumatismos relativamente leves, giba dorsal, huesos supraclaviculares llenos, estrías violáceas en abdomen (en ocasiones, en mamas y parte proximal de las extremidades), caída del cabello y piel delgada. Es frecuente que aparezcan o empeoren diabetes mellitus e hipertensión arterial. El hipercortisolismo es también causa de osteoporosis secundaria y por ende, de mayor riesgo de fracturas. La depresión o un franco cuadro psicótico también pueden ser manifestaciones clínicas. En el Cuadro 7.3 se muestran las diferentes manifestaciones clínicas del SC de acuerdo a sus categorías.

Síntomas del SC

- Obesidad de la parte superior del cuerpo.
- Rostro redondo o con forma de luna.
- Aumento de grasa alrededor del cuello.
- Adelgazamiento de los brazos y piernas.
- Piel delgada y frágil.
- Oscurecimiento de la pigmentación de la piel.
- Acné.
- Magulladuras (moretones).
- Marcas de estiramiento de la piel del abdomen, muslos, glúteos, brazos y el pecho.
- Debilidad muscular y de los huesos.
- Fatiga severa.
- Nivel alto de azúcar en la sangre.
- Irritabilidad y ansiedad.

- Crecimiento excesivo del vello en las mujeres.
- Pérdida o irregularidad de los ciclos menstruales en las mujeres.
- Deseo sexual y fertilidad disminuidos en los varones.

Cuadro 7.3. Manifestación Clínica. [1993, S.Y. Kung]. Fuente: [2006, Hurd]

Generales	Neuropsiquiátricas 85%
Obesidad 90% HTA 85%	Labilidad emocional Euforia Depresión Psicosis
Dérmicas	Disfunción gonadal
Plétora 70% Hirsutismo 75% Estrías 50% Acné 35% Equimiosis 35%	Trastornos menstruales 70% Impotencia, disminución de libido 85%
Músculo esqueléticas	Metabólicas
Osteopenia 80% Debilidad 65%	Intolerancia a la glucosa 75% Diabetes 20% Hiperlipemia 70% Poliuria 30% Cálculos renales 15%

7.6 Reconocimiento de patrones

Existen varios trabajos realizados en inteligencia artificial aplicados a la medicina. A continuación se mencionan algunas líneas de investigación y los resultados obtenidos.

Como primera línea de investigación, en lo que respecta a la clasificación y a la detección de anomalías cardíacas en 1971, se ha utilizado algoritmos semiautomáticos capaces de obtener un contorno sobre una región de interés que contiene, por ejemplo, todo el ventrículo izquierdo. Esta metodología se volvió una de las primeras formas de análisis digital de imágenes en ser considerado clínicamente útil y las redes *back propagation* todavía no habían sido desarrolladas [2000, Duncan], [1971, Strauss].

Una segunda línea, también de esa época, seguida por un grupo relativamente pequeño de investigadores se involucró en el tema tratando al análisis de imágenes medicas como un problema de procesamiento de información en una única imagen. Esta línea posee enfoques basados en reconocimiento de patrones, procesamiento de imágenes, procesamiento de señales y visión computarizada. Ejemplos de estos se incluyen el trabajo de [Sklansky 1973] que localiza automáticamente tumores mediante métodos relacionados a los patrones de reconocimiento. Otros esfuerzos, como el trabajo de [Pizer, 1978], enfatizaron el mejoramiento de la imagen y las estrategias de visualización, observando que estos eran problemas críticos para los usuarios finales (radiólogos y otros). Si bien no clasifican ni automatizan, sentó las

bases del pre-procesamiento.

Los primeros intentos de sistemas CAD (Computer Aided Design, Diseño asistido por computadora), completamente automático en mamografías de rayos X fueron propuestos sobre 1987 y fueron basados en las mejoras de la calidad de imagen producidas en la década anterior. Estos esfuerzos obtuvieron a una variedad de operaciones con umbrales de tolerancia y detección de características sobre mamografías digitalizadas y luego funciones discriminantes lineales para intentar clasificar automáticamente tejido normal y calcificaciones [2000, Duncan].

Línea de clasificación y subdivisión orientada a pulmones

Hacia fines de la década de los noventa, se presentó un sistema de diagnóstico asistido por computadora (CAD) para detectar seis patrones de tejidos pulmonares basándose en las características de sus texturas. Este método fue utilizado para determinar el subconjunto óptimo sobre regiones de 31x31 píxeles encontrando así la de mayor interés. Un clasificador Bayesiano fue entrenado para utilizar este subconjunto óptimo y así reconocer los seis tipos de tejidos. Se reportó que este sistema automatizado tuvo una tasa de éxitos comparable a la de un grupo de observadores humanos previamente entrenados.

Sluimer en su trabajo sobre un banco de filtros multi-escala 1995, utilizó este banco para representar la estructura y textura de la imagen. Fueron empleados varios clasificadores para entrenar el sistema y su tasa de error mostró un desempeño muy

similar a la obtenida en un experimento con dos radiólogos.

Casi todos los sistemas de diagnóstico asistidos por computadora dividen la imagen en pequeñas regiones, utilizan técnicas clásicas de procesamiento de imágenes para calcular las características de la imagen pero no toman ventaja del conocimiento anatómico existente. Por eso la idea de [2007, Zrimec & Busayarat] fue primero segmentar y extraer las características anatómicas de las imágenes y luego utilizar ese conocimiento para detectar anomalías causadas por enfermedad. Con esta idea han logrado resultados cercanos al 95% en detección de alguna anomalía en los pulmones [Sluimer, 2003] [1999, Uppaluri] [2007, Zrimec].

Línea de clasificación por imagen media aplicada al torso

Algunos enfoques se basan en la clasificación de píxeles o regiones y el cálculo de probabilidades para obtener un resultado global para la imagen en consideración. Estos enfoques requieren de etiquetado manual de píxeles o regiones para su entrenamiento. Por ejemplo para detectar una enfermedad interna en el torso [2004, Loog & van Ginneken] y [2006, van Ginneken] aplicaron técnicas de clasificación de radiografías de torso basadas respectivamente en la clasificación de píxeles y regiones.

De acuerdo a [2006, Arzhaeva] en la práctica, un buen resultado derivado de un píxel o pequeñas regiones, utilizan técnicas clásicas de procesamiento de imágenes para calcular las características de la imagen pero no toman ventaja del conocimiento

anatómico existente. Por eso la idea de [2007, Zrimec & Busayarat] fue primero segmentar y extraer las características anatómicas de las imágenes y luego utilizar ese conocimiento para detectar anomalías causadas por enfermedad. Con esta idea han logrado resultados cercanos al 95% en detección de alguna anomalía en los pulmones.

En la práctica, un buen resultado derivado de un píxel o una región no se encuentra muy seguido. Por otro lado, los resultados derivados de la imagen están casi siempre disponibles durante la recopilación de un conjunto de datos o son más fáciles de obtener. [2006, Arzhaeva] utilizó un enfoque de clasificación que permitió clasificar una imagen como un todo a partir del entrenamiento sobre etiquetas de imágenes globales. Como la información que indica la presencia o ausencia de patología es local, una representación de una imagen es introducida donde las características de una imagen global son derivadas de características de píxeles locales. El punto de partida de este método es la extracción de las características locales desde píxeles correspondientes espacialmente en todas las imágenes bajo consideración. Una forma de obtener estos píxeles sobre la imagen es mapearla con una imagen media de todas las imágenes. Ésta se segmenta para obtener un número de marcas fijas que pueden ser utilizadas para establecer una función de mapeo. Luego, un nuevo conjunto de características gráficas se deriva de las características locales de la otra imagen. Con muchas imágenes de referencia, es posible construir un banco de imágenes para entrenar a la clasificación de las mismas. Este método puede ser aplicado en puntos donde la anatomía es muy similar, en el caso del cáncer de mama, la clasificación obtendría varios errores ya que los tamaños y tipos de tejidos están sujetos al constante cambio existiendo una infinidad de patrones distintos que representan

mamas sanas y otra infinidad que representa mamas enfermas. [2006, Arzhaeva], [2004, Log]

Resolución espacial y en amplitud.

El termino imagen se refiere a una función de intensidad bidimensiona, que se representan indistintamente del contexto y de la nomenclatura como $f(x,y)$, $I(i,j)$, $f(i,j)$ $E(x,y)$ $q(x,y)$ etc donde las literales son las coordenadas espaciales y el valor de f, I, E o G en cualquier punto (x,y) ó (i,j) es proporcional a la intensidad o nivel de gris de la imagen en ese punto.

Dependiendo del número de píxeles que tenga el dispositivo o en el caso de imágenes analógicas del número de muestras tomadas, la imagen poseerá más o menos resolución espacial. [2004, Pajares]

Problemática que se presenta

Este enfoque presenta varias falencias, la principal es la utilización de píxeles mediante coordenadas cartesianas, las cuales son producto de una arbitrariedad (el ángulo en que fueron adquiridas las imágenes), entrenar una red neuronal para que aprenda con todos los ángulos requiere de un conjunto de datos muy importante y esto decrementa la tasa de éxito. Otra falencia importante es la calidad de la entrada, si son píxeles, se requiere una entrada muy grande generando nuevamente la necesidad de un conjunto de datos acorde, lo cual, al volverse inviable obliga a tomar menos entradas y con ello ignorar más información que puede ser útil; incluso imágenes

con muy alta resolución deberían ser reducidas para poder ser mapeadas a la red neuronal obteniéndose la misma calidad de resultado que una imagen de muy baja resolución. Si bien se toman medias, varianzas y otros operadores estadísticos los lugares donde son aplicados pueden estar lejos de las zonas de interés y por ende no aportan información útil al proceso de clasificación.

Los resultados obtenidos durante los experimentos en la línea aplicada al cáncer pulmonar, por [2003, Uchiyama], tuvieron una tasa de éxito aproximada del 60%, lo cual continúa siendo bajo para la producción e implementación de un software. Cabe mencionar que una clasificación aleatoria logra una tasa media de éxito del 50% que es el mínimo posible desde el punto de vista de la teoría de la información de Shannon 1948; Si un algoritmo logra tasas medias menores que 50%, entonces invirtiendo sus resultados se obtendrían tasas mayores y por lo tanto resultaría útil. Una tasa del 60% solo aporta información del orden de 10% sobre 50%, es decir un 20%; una tasa de éxito media del 75% aporta un 50% de información y ya comienza a ser importante. Lograr un algoritmo con una tasa media de éxito del 0% sería óptimo, ya que es sabido que el algoritmo siempre se equivoca y por lo tanto, al tener dos valores posibles, solo hay que descartar el valor obtenido y se sabrá que el restante siempre es la solución [2006, Arzhaeva] [2004, Pajares] , [2003, Uchiyama].

7.7 Redes neuronales

Una red neuronal artificial es una estructura que permite procesar entradas de forma similar a la que ocurre en nuestro cerebro. Están compuestas, básicamente, por pequeñas unidades llamadas neuronas que pueden estar enlazadas entre sí. Esta estructura recibe una entrada de datos que, luego de ser adaptada para tomar valores admitidos por las neuronas, son ingresados en algunas de ellas (neuronas de entrada) y luego siguiendo las conexiones (sinapsis) entre neuronas se va propagando información hasta llegar a las neuronas de salida, las cuales arrojan un resultado. Las Redes Neuronales Artificiales (RNA) son utilizadas para predicción, minería de datos, reconocimiento de patrones y sistemas de control adaptativo. Constituyen una parte muy importante en el estudio y desarrollo de la inteligencia artificial (AI) y el de la vida artificial (*a-life*).

Las redes neuronales artificiales pueden ser combinadas con otras técnicas como la lógica difusa (lógica fuzzy), algoritmos genéticos, sistemas expertos, estadísticas, transformadas de Fourier, etc. para que sea más potente, su procesamiento, y buscando obtener mejores resultados, en menor tiempo y costo computacional. Son adecuadas para resolver problemas que hasta ahora eran resueltos sólo por el cerebro humano y resultaban difíciles o imposibles para las máquinas lógicas secuenciales. Un procesamiento paralelo realizado por un gran número de elementos altamente interconectados, es la clave de su funcionamiento.

Esta tecnología que puede ser desarrollada tanto en software como en hardware y con ella se pueden diseñar soluciones de software que permiten aprender, adaptarse a condiciones variantes, o inclusive si se dispone de una colección suficiente grande de datos, predecir el estado futuro de algunos modelos.

7.8 Funcionamiento de Neuronas Biológicas

El cerebro humano continuamente recibe señales de entrada de muchas fuentes y las procesa a manera de crear una apropiada respuesta de salida. Nuestros cerebros cuentan con millones de neuronas que se interconectan para elaborar Redes Neuronales. Estas redes ejecutan los millones de instrucciones necesarias para mantener una vida normal.

Las neuronas son las células que forman la corteza cerebral de los seres vivos, cada una está formada por elementos llamados cuerpo, axón y dendritas, como se muestra en la siguiente Figura 7.3 [2008, Martin].

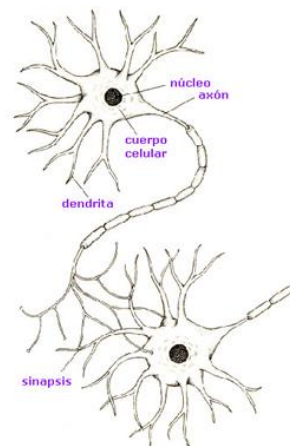


Figura. 7.3 Neurona Biológica. [2008, Martin].

7.9 Funcionamiento de las RNA

Una de las misiones en una red neuronal consiste en simular las propiedades observadas en los sistemas neuronales biológicos a través de modelos matemáticos recreados mediante mecanismos artificiales (como un circuito integrado, un ordenador o un conjunto de válvulas). El objetivo es conseguir que las máquinas den respuestas similares a las del cerebro humano, caracterizadas por su generalización y complejidad. En la Figura 7.4 tenemos la representación de un perceptrón con dos entradas, la suma de sus valores y la salida [2008, Martin] , [1993, S.Y. Kung].

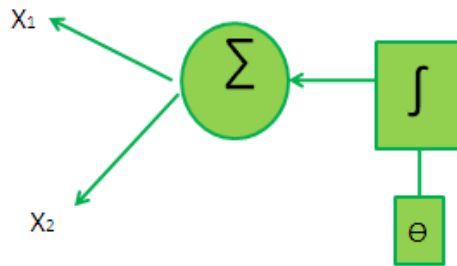


Figura. 7.4 Perceptrón con 2 entradas. [1993, S.Y. Kung].

Una red neuronal se compone de unidades llamadas neuronas. Cada neurona recibe una serie de entradas a través de interconexiones y emite una salida. Esta salida viene dada por tres funciones:

1. La función de activación, consiste en la sumatoria de cada entrada multiplicada por el peso de su interconexión (valor neto). Si el peso es positivo, la conexión se denomina *excitatoria*; si es negativo, se denomina *inhibitoria*.

2. Una función de activación, que modifica a la anterior. Puede no existir, siendo en este caso la salida la misma función de propagación.
3. Una función de transferencia, que se aplica al valor devuelto por la función de activación. Se utiliza para acotar la salida de la neurona y generalmente viene dada por la interpretación que queramos darle a dichas salidas. Una de las más utilizadas es la función sigmoïdal (para obtener valores en el intervalo $[0,1]$) y la tangente hiperbólica (para obtener valores en el intervalo $[-1,1]$).

Estas funciones son las que se ocupan en la investigación para modificar los pesos del entrenamiento cuando la salida no es la salida deseada, por lo que se requiere penalizar a las neuronas mediante sus pesos, con la finalidad de tener la salida deseada.

7.10 Estructura

La mayoría de los científicos coinciden en que una RNA es muy diferente en términos de estructura de un cerebro animal [2000, Hilara]. Al igual que el cerebro, una RNA se compone de un conjunto masivamente paralelo de unidades de procesamiento muy simples y es en las conexiones entre estas unidades donde reside la inteligencia de la red. Sin embargo, en términos de escala, un cerebro es muchísimo mayor que cualquier RNA creada hasta la actualidad y las neuronas artificiales también son más simples que su contrapartida animal.

Biológicamente, un cerebro aprende mediante la reorganización de las conexiones sinápticas entre las neuronas que lo componen. De la misma manera, las RNA tienen un gran número de procesadores virtuales interconectados que de forma simplificada simulan

la funcionalidad de las neuronas biológicas. En esta simulación, la reorganización de las conexiones sinápticas biológicas se modela mediante un mecanismo de pesos, que son ajustados durante la fase de aprendizaje. En una RNA entrenada, el conjunto de los pesos determina el conocimiento de esa RNA y tiene la propiedad de resolver el problema para el que la RNA ha sido entrenada. [2000, Hilerá], [1993, S.Y. Kung].

Por otra parte, en una RNA, además de los pesos y las conexiones, cada neurona tiene asociada una función matemática denominada función de transferencia. Dicha función genera la señal de salida de la neurona a partir de las señales de entrada. La entrada de la función es la suma de todas las señales de entrada por el peso asociado a la conexión de entrada de la señal. Algunos ejemplos de entradas son la función escalón o *heaviside*, la lineal o mixta, la sigmodal y la función gaussiana, en la presente investigación se utilizará una función gaussiana para la clasificación y elaboración del hiperplano de separación de las clases, recordemos que la función de transferencia es la relación entre la señal de salida y la entrada.

Modelos

- Perceptrón
- Adaline
- Perceptrón multicapa
- Memorias asociativas
- Máquina de Bolzman

- Máquina de Cauchy
- *Backpropagation*
- Redes de Elman
- Redes de Hopfield
- Red de contrapropagación
- Redes de neuronas de base radial
- Redes de neuronas de aprendizaje competitivo
- Mapas Autoorganizados (RNA)
- Crecimiento dinámico de células
- Gas Neuronal Creciente
- Redes ART (*Adaptative Resonance Theory*)

Las dendritas forman una estructura de filamentos muy fina que rodean el cuerpo de la neurona. El axón es un tubo largo y delgado que se ramifica en su extremo en pequeños bulbos finales que casi tocan las dendritas de las células vecinas. La pequeña separación entre los bulbos finales y las dendritas se le denomina sinapsis [2008, Martin].

Las neuronas, al igual que las demás células del cuerpo, funcionan a través de impulsos eléctricos y reacciones químicas. Los impulsos eléctricos que utiliza una neurona para intercambiar información con las demás, viajan por el axón que hace contacto con las dendritas de la neurona vecina mediante las sinapsis. La intensidad de la corriente transmitida depende de la eficiencia de la transmisión sináptica. Una neurona en especial

transmitirá un impulso eléctrico por su axón si suficientes señales de las neuronas vecinas transmiten a través de sus dendritas en un tiempo corto. La señal que se transmite a la neurona podrá ser además ya sea inhibitoria o excitatoria. La neurona dispara, esto es, manda el impulso por su axón, si la excitación excede su inhibición por un valor crítico, el umbral de la neurona [2004, Viñuela].

7.11 Características de operación de las RNA

Las características de operación de las RNA son las siguientes:

Las RNA puede tener factores de peso fijos o adaptables. Las que tienen pesos adaptables emplean leyes de aprendizaje para ajustar el valor de la fuerza de una interconexión con otras neuronas. Si las neuronas utilizan pesos fijos, entonces su tarea deberá estar previamente definida. Los pesos serán determinados a partir de una descripción completa del problema. Por otra parte, los pesos adaptables son esenciales si no se conoce previamente cual deberá de ser su valor correcto [1993, S.Y. Kung].

Existen dos tipos de aprendizaje: supervisado y no supervisado. El primero ocurre cuando se le proporciona a la red tanto la entrada como la salida correcta, y la red ajusta sus pesos tratando de minimizar el error de su salida calculada. Este tipo de entrenamiento se aplica por ejemplo, en el reconocimiento de patrones. El entrenamiento no supervisado se presenta cuando a la red se le proporcionan únicamente los estímulos, y la red ajusta sus interconexiones basándose únicamente en sus estímulos y la salida de la propia red. Las leyes de aprendizaje determinan como la red ajustará sus pesos utilizando una función de

error o algún otro criterio. La ley de aprendizaje adecuada se determina en base a la naturaleza del problema que se intenta resolver [2000, Hilera].

Las RNA adaptables tienen dos fases en su operación.

1. Entrenamiento de la red. El usuario proporciona a la red un número adecuado de estímulos de entrada, y de salida, la red entonces ajusta sus pesos de interconexión o sinapsis hasta que la salida de la red está lo suficientemente cerca de la salida correcta.

2. Recuperación de lo aprendido. A la red se le presenta un conjunto de estímulos de entrada y esta simplemente calcula su salida. Cuando la red emplea entrenamiento no supervisado, algunas veces será necesario que reajuste su sinapsis durante la fase de recuperación.

La gran diferencia del empleo de las redes neuronales en relación con otras aplicaciones de la computación radica en que no son algorítmicas, esto es no se programan haciéndoles seguir una secuencia predefinida de instrucciones. Las RNA generan ellas mismas sus propias reglas, para asociar la respuesta a su entrada; es decir, aprende por ejemplos y de sus propios errores.

El conocimiento de una RNA se encuentra en la función de activación utilizada y en los valores de sus pesos.

Las redes neuronales formadas por los perceptrones se interconectan en forma muy similar a como las neuronas humanas se disponen en la corteza cerebral humana, y lo más

importante, son capaces de asociar y generalizar sin reglas. Han sido utilizadas con gran éxito para reconocer retornos de sonar bajo el agua, escritura a mano, voz, topografía de terrenos, controlar brazos de robots, evaluar datos personales, modelar fenómenos cognoscitivos, y, predecir tendencias financieras.

La clase de problemas que mejor se resuelven con las redes neuronales son los mismos que el ser humano resuelve mejor: asociación, evaluación, y reconocimiento de patrones. Las redes neuronales son perfectas para problemas que son muy difíciles de calcular pero que no requieren de respuestas perfectas, sólo respuestas rápidas y buenas. Tal y como acontece con el escenario bursátil en el que se quiere saber ¿compro?, ¿vendo?, ¿mantengo?, o en el reconocimiento cuando se desea saber ¿se parece? ¿es el mismo pero tienen una ligera modificación?.

Por otra parte, las redes neuronales son muy malas para cálculos precisos, procesamiento serie, y no son capaces de reconocer nada que no tenga inherentemente algún tipo de patrón. Es por esto, que no pueden predecir la lotería, ya por definición es un proceso al azar.

Existen varias formas de hacer las conexiones en una RNA, así como existen varias formas de conectar neuronas biológicas en el cerebro. Cada tipo sirve para diferentes procesos, el elegir la correcta topología y sus características, es imprescindible para lograr fácilmente la solución del problema [2000, Hilera] [2008, Martin] [2004, Viñuela].

7.12 Tipos de Redes Neuronales Artificiales

Perceptron

En 1943, Warren McCulloch y Walter Pitts originaron el primer modelo de operación neuronal, el cual fué mejorado en sus aspectos biológicos por Donald Hebb en 1948.

En 1962 Bernard Widrow propuso la regla de aprendizaje Widrow-Hoff, y Frank Rosenblatt desarrolló una prueba de convergencia, y definió el rango de problemas para los que su algoritmo aseguraba una solución. Él propuso los 'Perceptrons' como herramienta computacional.

En la Figura. 7.5 se representa una neurona artificial, que intenta modelar el comportamiento de la neurona biológica. Aquí el cuerpo de la neurona se representa como un sumador lineal de los estímulos externos z_j , seguida de una función no lineal $y_j = f(z_j)$.

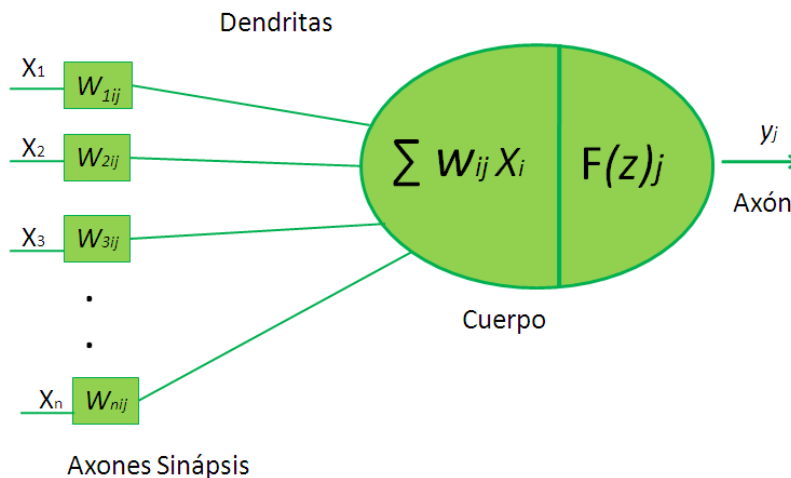


Figura. 7.5 Axones y Sinapsis. Fuente: [2008, Martin]

La función $f(z)_j$ es llamada la función de activación, y es la función que utiliza la suma de estímulos para determinar la actividad de salida de la neurona.

Este modelo se conoce como perceptrón de McCulloch-Pitts, y es la base de la mayor parte de las arquitecturas de las RNA que se interconectan entre sí. Las neuronas emplean funciones de activación diferentes según la aplicación, algunas veces son funciones lineales, otras funciones sigmoidales (p.ej. la tanh), y otras funciones de umbral de disparo. La eficiencia sináptica se representa por factores de peso de interconexión w_{ij} , desde la neurona i , hasta la neurona j .

Los pesos pueden ser positivos (excitación) o negativos (inhibición). Los pesos junto con las funciones $f(z)$ dictan la operación de la red neuronal. Normalmente las funciones no se modifican de tal forma que el estado de la red neuronal depende del valor de los factores de peso (sinapsis) que se aplica a los estímulos de la neurona.

En un perceptrón, cada entrada es multiplicada por el peso W correspondiente, y los resultados son sumados, siendo evaluados contra el valor de umbral, si el resultado es mayor al mismo, el perceptrón se activa.

El perceptrón es capaz tan sólo de resolver funciones definidas por un hiperplano (objeto de dimensión $N-1$ contenida en un espacio de dimensión N) que corte un espacio de dimensión N . Un ejemplo de una función que no puede ser resuelta es el operador lógico XOR.

Una explicación más sencilla de un hiperplano sería, hablando en un plano de dos dimensiones, una línea que separa a los elementos existentes en dos grupos. El perceptrón

sólo puede resolver una función, si todos los posibles resultados del problema pueden separarse de ésta forma (en dos secciones) es decir, que no se combinen entre sí [2004, Viñuela].

El entrenamiento de un perceptrón es por medio de la regla de aprendizaje delta: Para cada peso W se realiza un ajuste dW según la regla de la Ec. 1:

$$d_w = L_R(T - Y)X \text{ Ec. 1}$$

Donde LR es el parámetro de aprendizaje, T el valor deseado, Y el valor obtenido, y X la entrada aplicada al perceptrón.

El Perceptrón básico de dos capas (entrada con neuronas lineales, analógicas, y la de salida con función de activación de tipo escalón, digital) solo puede establecer dos regiones separadas por una frontera lineal en el espacio de patrones de entrada, donde se tendría un hiperplano.

Un Perceptrón con tres niveles de neuronas puede formar cualquier región convexa en este espacio. Las regiones convexas se forman mediante la intersección entre las regiones formadas por cada neurona de la segunda capa, cada uno de estos elementos se comporta como un Perceptrón simple, activándose su salida para los patrones de un lado del hiperplano.

Un Perceptrón con cuatro capas puede generar regiones de decisión arbitrariamente complejas. El proceso de separación en clases que se lleva a cabo consiste en la partición de

la región deseada en pequeños hipercubos. Cada hipercubo requiere neuronas en la segunda capa (siendo n el número de entradas a la red), una por cada lado del hipercubo, y otra en la tercera capa, que lleva a cabo el and lógico de la salida de los nodos del nivel anterior. La salida de los nodos de este tercer nivel se activaran solo para las entradas de cada hipercubo. Los hipercubos se asignan a la región de decisión adecuada mediante la conexión de la salida de cada nodo del tercer nivel solo con la neurona de salida (cuarta capa) correspondiente a la región de decisión en la que este comprendido el hipercubo llevándose a cabo una operación lógica Or en cada nodo de salida. Este procedimiento se puede generalizar de manera que la forma de las regiones convexas sea arbitraria, en lugar de hipercubos [2007, Andina] , [2000, Hilerá] , [2004, Viñuela].

En teoría, el Perceptrón de cuatro capas puede resolver una gran variedad de problemas cuyas entradas sean analógicas, la salida sea digital y sea linealmente separable. El problema práctico radica en el número de neuronas, en el número idóneo de capas ocultas, la extensión de la función de activación, el tiempo de entrenamiento de la red, las implicaciones en la generación de ruido (al tener un número excesivo de neuronas) en contraparte con la ventaja de tener un sistema tolerante a fallas al tener un número de neuronas redundante.

El rango de tareas que el Perceptrón puede manejar es mucho mayor que simples decisiones y reconocimiento de patrones. Por ejemplo, se puede entrenar una red para formar el tiempo pasado de los verbos en ingles, leer texto en ingles y manuscrito.

El Perceptrón multicapa (MLP) puede ser usado para la predicción de una serie de datos en el tiempo, como la medición de la demanda de gas y electricidad para identificar tendencias de consumo, además de la predicción de cambios en el valor de los instrumentos financieros.

Predicción de mercados financieros, diagnósticos médicos, el Perceptrón como una red codificadora, el Perceptrón aprende a sumar enteros [2008, Martin],[1993, S.Y. Kung]

Boltzmann

En la Máquina de Boltzmann, generalización de la red de Hopfield que incluye unidades ocultas, la operación de actualización se basa en un concepto de termodinámica estadística conocido como simulated annealing. La red de Hopfield, la máquina de Boltzmann y un derivado conocido como la máquina del teorema de campo medio se han utilizado en aplicaciones de segmentación y restauración de imágenes y optimización combinatorial.

Características

La red de Boltzmann consiste en un conjunto de N elementos de procesamiento interconectados que actualizan sus valores de activación de forma asíncrona e independiente del resto de los elementos de procesamiento. Todos los elementos son a la vez de entrada y salida. Los valores de activación son binarios.

El estado del sistema está dado por los valores de activación Y_k . La entrada de la neurona k en el ciclo temporal $t+1$ viene dada por la Ec. 2

$$s_k(t + 1) = \sum_{j \neq k} y_j(t) w_{jk} + \theta_k \text{ Ec. 2}$$

Para obtener el nuevo valor de activación se aplica una función umbral.

Cuando un elemento de procesado mantiene su valor de activación se dice que es estable.

Se llama estado estable a aquel en el cual todos los elementos de procesado son estables.

Con la restricción extra de simetría en los pesos de conexión, $W_{jk} = W_{kj}$, el sistema puede ser descrito mediante una función energía de la forma, como se muestra en la Ec. 3 [2007, Andina].

$$E = -\frac{1}{2} \sum_{j \neq k} \sum_k y_j y_k w_j - \sum_k \theta_k y_k \text{ Ec. 3}$$

Funcionamiento

A cada estado de la red se le puede atribuir una cierta cantidad de energía, el sistema evoluciona tratando de disminuir la energía mediante un proceso de relajación, hasta alcanzar un mínimo (valle) donde se estabiliza. Los mínimos de energía se corresponden con los recuerdos almacenados durante el aprendizaje de la red.

Ante la presentación de un estímulo nuevo se obtendrá una configuración inicial más o menos parecida a alguno de los estímulos almacenados, el sistema evolucionará hasta caer en una configuración estable que representa el recuerdo asociado a ese estímulo. Si la configuración inicial discrepa mucho de los recuerdos almacenados podemos alcanzar algún mínimo que no se corresponde a ningún recuerdo almacenado, recuperando en ese

caso una información espuria, o podríamos no alcanzar ningún mínimo, quedando inestable: en ese caso diríamos que la red está "confundida", no es capaz de reconocer el estímulo, no recuerda.

Una tercera posibilidad es que al cabo de unos pasos de evolución empiece a repetir periódicamente una secuencia definida de estados; con esta dinámica se han modelado ciertas excitaciones nerviosas que regulan acciones rítmicas y repetitivas; y se ha tratado de reproducir la memoria de secuencias temporales, pe. el recuerdo de melodías [2000, Hilera].

Hopfield

Las redes de Hopfield son redes de adaptación probabilística, recurrentes, funcionalmente entrarían en la categoría de las memorias autoasociativas, es decir, que aprenden a reconstruir los patrones de entrada que memorizaron durante el entrenamiento. Son arquitecturas de una capa con interconexión total, funciones de activación booleana de umbral (cada unidad puede tomar dos estados, 0 o 1, dependiendo de que la estimulación total recibida supere el umbral determinado), adaptación probabilística de la activación de las unidades, conexiones recurrentes y simétricas, y regla de aprendizaje no supervisado. Mientras que las redes en cascada (no recurrentes) dan soluciones estables, los modelos recurrentes dan soluciones inestables (dinámicas), lo que no siempre es aconsejable. En la Figura 7.6 podemos ver el esquema representativo de una red hopfield, notando sus aristas dirigidas y con retroalimentación.

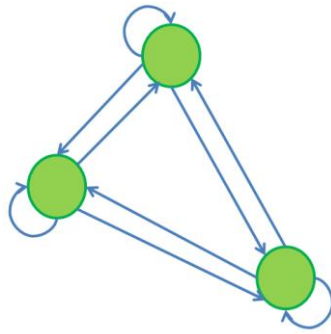


Figura 7.6 Esquema Hopfield. [2000, Hilera]

La principal aportación de Hopfield consistió precisamente en conseguir que tales redes recurrentes fueran así mismo estables. Imaginó un sistema físico capaz de operar como una memoria autoasociativa, que almacenara información y fuera capaz de recuperarla aunque la misma se hubiera deteriorado [1993, S.Y. Kung],[2004, Pajares].

La Red de Hopfield es recurrente y completamente conectada. Funciona como una memoria asociativa no lineal que puede almacenar internamente patrones presentados de forma incompleta o con ruido. De esta forma puede ser usada como una herramienta de optimización. El estado de cada neurona puede ser actualizado un número indefinido de veces, independientemente del resto de las neuronas de la red pero en paralelo.

Las neuronas se conectan todas entre sí, y consigo mismas. Para empezar se le asigna a cada unidad el valor o estado correspondiente del patrón de entrada. En cada ciclo se elige una neurona al azar y se calcula su activación según la función de umbral:

Sea n el número de neuronas en la red, la estimulación total se calcula como el sumatorio de todas las entradas ponderadas, incluida la procedente de la misma unidad, Ec. 4.

$$X_j = \sum_{i=1}^x w_{ij} y_i \quad \text{Ec. 4}$$

Se puede trabajar con cualquier valor de umbral para la función de activación, pero típicamente se usa el 0 como umbral; tiene la ventaja de simplificar las ecuaciones, como en las Ec. 5.

$$y_i = 1 \text{ si } x_i > 0 \quad y_i = 0 \text{ si } x_i < 0 \quad \text{Ec. 5}$$

Por definición los recuerdos (ítems o patrones almacenados) son puntos fijos en la dinámica de la red, es decir, configuraciones que no cambian en el tiempo aunque se siga aplicando la regla de evolución. Para almacenar un recuerdo habrá que lograr que la presentación del patrón de entrada lleve a la red a alcanzar un punto fijo, esto se logrará mediante alguna regla de aprendizaje que modifique los pesos de las conexiones [2004, Pajares].

Aprendizaje de las redes de Hopfield: Regla de Cooper-Hebb

La elección de la regla de aprendizaje no es trivial, depende de la interrelación de los patrones que se desea memorizar. Si estos patrones están poco correlacionados (pseudoortogonales) podemos aplicar la regla de Cooper-Hebb, basada en la regla de Hebb o regla del producto:

Supongamos que tenemos p ítems o patrones que almacenar, entonces calculamos Ec. 6

$$s_i i < > j \text{ (matriz simétrica)} w_{ij} = 0 \text{ si } i = j \text{ (la diagonal} = 0) \quad \text{Ec. 6}$$

Esta regla fortalece las conexiones cuando las unidades i-ésima y j-ésima tienen la misma

activación o valor de estado, y las debilita en caso contrario.

El entrenamiento seguirá los siguientes pasos.

1. Elegir un número de neuronas que cumpla el criterio del 15%
2. Codificar los ítems que queremos memorizar de forma que los patrones para representarlos se parezcan lo menos posible entre sí, para aproximarnos a la condición de pseudo-ortogonalidad.
3. Calcular los pesos de las conexiones según la regla de Cooper-Hebb.

Ya tenemos la red lista para usarla, asignando a las unidades el estado inicial y dejando que la red evolucione hasta alcanzar un mínimo (esto lo conseguiremos comprobando que el decremento de energía es nulo) [2000, Hilerá], [2008, Martín], [1993, S.Y. Kung].

Valoración:

Las redes de Hopfield se han aplicado a campos como la percepción el reconocimiento de imágenes y optimización de problemas, mostrando gran inmunidad al ruido y robustez. Incluso se han llegado a desarrollar chips específicos para este tipo de redes. El estudio de las representaciones de secuencias temporales es un área de gran interés, con aplicaciones en reconocimiento automático de voces y movimientos.

Hopfield ha mostrado como aplicar los mismos principios con funciones de activación continuas como la función sigmodal, con muy pocas modificaciones [1993, S.Y. Kung].

Pero pese a sus evidentes ventajas no están exentas de problemas:

- El número máximo de patrones no correlacionados que puede almacenar es igual al 15% del número de neuronas de la red.
- Requieren mucho tiempo de procesamiento hasta converger a una solución estable, lo que limita su aplicabilidad.
- Otro de los problemas achacados a las redes de Hopfield es su tendencia a caer en mínimos locales, como en las redes de retropropagación. La solución pasa por aplicar los métodos estadísticos que ya comentamos en el apartado dedicado a las redes de retropropagación, el equilibrio termodinámico simulado.

Kohonen

Existen evidencias que demuestran que en el cerebro existen neuronas que se organizan en muchas zonas, de forma que las informaciones captadas del entorno a través de los órganos sensoriales se representan internamente en forma de capas bidimensionales. Por ejemplo, en el sistema visual se han detectado mapas del espacio visual en zonas de córtex (capa externa del cerebro). También en el sistema auditivo se detecta organización según la frecuencia a la que cada neurona alcanza la mayor respuesta (organización tonotópica). En la Figura 7.7, podemos ver el esquema representativo de una topología Kohonen.

Aunque en gran medida esta organización neuronal está predeterminada genéticamente, es probable que de ella se origine mediante el aprendizaje. Esto sugiere, por tanto, que el cerebro podría poseer la capacidad inherente de formar mapas topológicos de las informaciones recibidas del exterior. De hecho, esta teoría podría explicar su poder de operar con elementos semánticos: algunas áreas del cerebro simplemente podrían crear y

ordenar neuronas especializadas o grupos con características de alto nivel y sus combinaciones. Se trataría, en definitiva, de construir mapas espaciales para atributos y características [2007, Andina].

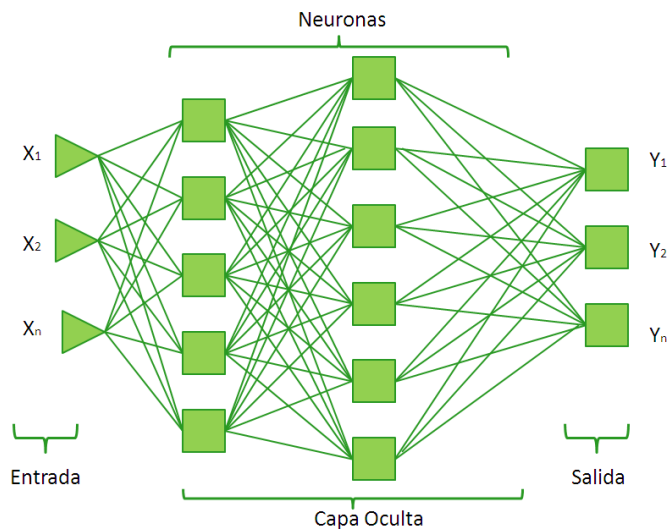


Figura 7.7 Esquema Kohonen. Fuente: [2007, Andina]

Historia

A partir de estas ideas, T. Kohonen presentó en 1982 un sistema con un comportamiento semejante. Se trataba de un modelo de red neuronal con capacidad para formar mapas de características de manera similar a como ocurre en el cerebro.

El objetivo de Kohonen era demostrar que en un estímulo externo (información de entrada) por sí solo, suponiendo una estructura propia y una descripción funcional del comportamiento de la red, era suficiente para forzar la formación de mapas.

Este modelo tiene dos variantes, denominadas LVQ (*Learning Vector Quantization*) y TPM

(*Topology-Preserving Map*) o SOM (*Self-Organizing Map*). Ambas se basan en el principio de formación de mapas topológicos para establecer características comunes entre las informaciones (vectores) de entrada a la red, aunque difieren en las dimensiones de éstos, siendo de una soladimensión en el caso de LVQ, y bidimensional, e incluso tridimensional, en la red TPM.

Características

Pertenece a la categoría de las redes competitivas o mapas de auto-organización, es decir, aprendizaje no supervisado. Poseen una arquitectura de dos capas (entrada-salida) (una sola capa de conexiones), funciones de activación lineales y flujo de información unidireccional (son redes en cascada).

Las unidades de entrada reciben datos continuos normalizados, se normalizan así mismo los pesos de las conexiones con la capa de salida. Tras el aprendizaje de la red, cada patrón de entrada activará una única unidad de salida.

El objetivo de este tipo de redes es clasificar los patrones de entrada en grupos de características similares, de manera que cada grupo activará siempre la(s) misma(s) salida(s). Cada grupo de entradas queda representado en los pesos de las conexiones de la unidad de salida triunfante. La unidad de salida ganadora para cada grupo de entradas no se conoce previamente, es necesario averiguarlo después de entrenar a la red.

Arquitectura

En la arquitectura de la versión original (LVQ) del modelo Kohonen no existen conexiones hacia atrás. Se trata de una de las N neuronas entrada y M de salida. Cada una de las N neuronas de entrada se conecta a las M de salida a través de conexiones hacia adelante (*feedforward*).

Entre las neuronas de la capa de salida, puede decirse que existen conexiones laterales de inhibición (peso negativo) implícitas, pues aunque no estén conectadas, cada una de las neuronas va a tener cierta influencia sobre sus vecinas. El valor que se asigne a los pesos de las conexiones hacia adelante entre las capas de entrada y salida (W_{ji}) durante el proceso de aprendizaje de la red va a depender precisamente de esta interacción lateral.

La influencia que una neurona ejerce sobre las demás es función de la distancia entre ellas, siendo muy pequeñas cuando están muy alejadas. Es frecuente que dicha influencia tenga la forma de un sombrero mexicano.

Por otra parte, la versión del modelo denominada TPM (*Topology Preserving Map*) trata de establecer una correspondencia entre los datos de entrada y un espacio bidimensional de salida, crenado mapas topológicos de dos dimensiones, de tal forma que ante datos de entrada con características comunes se deben activar neuronas situadas en próximas zonas de la capa de salida [2007, Andina], [2008, Martín], [2004, Viñuela].

Aprendizaje

Supongamos que tenemos patrones de entrada n -dimensionales.

El procedimiento a seguir sería el siguiente, aleatorizar los pesos de las conexiones. Normalizar los pesos de las conexiones incidentes de cada unidad de salida sobre la unidad: dividir cada conexión por la raíz cuadrada de la suma de los cuadrados de las conexiones de cada unidad. Normalizar igualmente los datos de entrada

1. Aplicar un patrón de entrada.
2. Calcular alguna medida de similitud/disimilitud (producto interno, distancia euclídea o de Mahalanobis, etc.) entre las entradas y los pesos de las conexiones.
3. Declaración de la unidad ganadora de salida con los pesos más parecidos al patrón de entrada. El vector de pesos de la unidad ganadora, se convierte en el centro de un grupo de vectores cercanos a él.
4. Modificar los pesos de los vectores de pesos W_j "cercaños" a W_c (distancia menor a D), según la fórmula: De esta manera conseguimos que los vectores de pesos de la unidad ganadora y de su "vecindario" se parezcan cada vez más al patrón de entrada que hace ganar a esa unidad.
5. Repetir los pasos 1 a 4 con todos los patrones de entrada.

A medida que avanza el aprendizaje hay que ir reduciendo D y A , Kohonen recomienda empezar con un valor de A cercano a 1 y reducirlo gradualmente hasta 0.1. D puede empezar valiéndolo la máxima distancia existente entre los pesos de las conexiones al

principio y acabar siendo tan pequeño que no quede ninguna unidad en el vecindario de la unidad ganadora. En ese momento solo se entrenará una unidad, que al final tendrá su vector de pesos igual al vector de entrada.

La precisión de la clasificación de los patrones de entrada aumenta con el número de ciclos de aprendizaje. Kohonen recomienda una cantidad de ciclos no inferior a 500 veces el número de neuronas de salida para obtener buenos resultados.

Aplicación

Una vez entrenada, podemos usar a la red para clasificar patrones de entrada similares en el espacio n -dimensional. Una clase o grupo de patrones similares tiende a controlar una neurona específica, que representará el centro de una esfera n -dimensional (de radio unitario, pues normalizamos los datos sobre la unidad). Esa neurona resultará la más activada frente a los patrones más parecidos a su vector de pesos.

Después del aprendizaje, la clasificación consiste en presentar una entrada y seleccionar la unidad más activada. Además, el vector de pesos nos servirá para reconstruir el patrón de entrada [2008, Martin].

Los modelos de redes presentados anteriormente fueron escritos con la finalidad de dar un panorama general sobre los tipos de RNA, el algoritmo que utilizaremos para la investigación es el *Backpropagation* por lo que le dedicamos y describimos en el siguiente apartado.

7.13. Backpropagation

Con este tipo de red neuronal artificial se pretende implementar el reconocimiento de patrones y por ende la ayuda para el predigagnóstico del síndrome de cushing, a continuación se destacan las principales características por las cuales se ha elegido.

Introducción

El perceptrón, es el ejemplo más elemental de una red neuronal, de hecho, no puede siquiera ser considerado una red, puesto que no intervienen otros elementos. Si se combinan varios perceptrones en una capa, y los estímulos de entrada después se suman, tendremos ya una red neuronal. Una red neuronal muy eficaz para resolver fundamentalmente problemas de reconocimiento de patrones es la red neuronal de propagación hacia atrás, en inglés *backpropagation network*.

Una red neuronal de tipo *backpropagation* permite aprender mediante un conjunto de ejemplos (entrada-salida) comúnmente denominado *training set*. Al haber aprendido mediante este conjunto, se puede obtener una salida coherente para una entrada dada.

En la Figura 7.10 es posible observar cómo se obtiene una salida a partir de la entrada, en este caso vemos a la red neuronal como una caja negra, con dos capas, una de entrada y otra de salida [2000, Hilera].

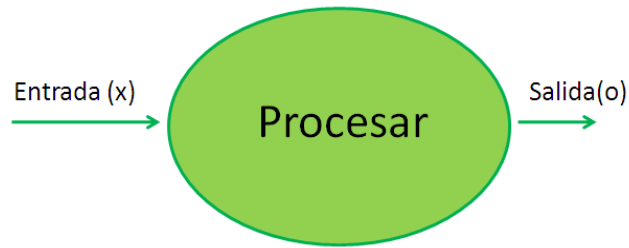


Figura.7.8 Vista de caja negra de una red neuronal. Fuente: [2000, Hilera]

Como podemos apreciar en la Figura. 7.8, cada neurona de entrada, que posee un valor en el rango $[0;1]$, y se pasa ese valor a todas las neuronas de salida. Ese valor es multiplicado por el peso $W_{i,j}$ representado en las aristas. El valor de ij es igual al de una función (denominada de transferencia) aplicada a la sumatoria de todos los productos definidos como i para j de esa neurona de salida.

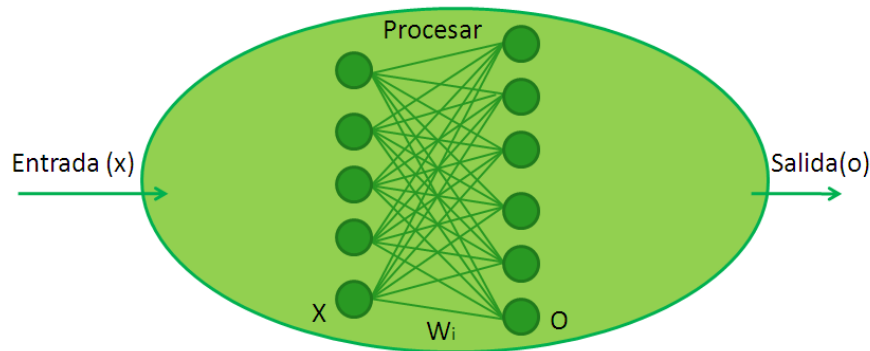


Figura. 7.9 Vista interna de una red neuronal sin capas ocultas. Fuente: [2000, Hilera]

La información que almacena el aprendizaje de una red se encuentra en los pesos $W_{i,j}$ y en ningún lado más. Es muy importante comprender esto, ya que son los pesos los que hay que ajustar en el proceso de entrenamiento, en el estudio que realizamos dicha información de aprendizaje permitirá identificar a los pacientes con SC de los que no lo tienen. En la Figura 7.9 podemos observar el proceso de entrenamiento tomando la red en forma de caja negra. Vemos que existen dos salidas: la que obtenemos mediante la

red y la deseada. El objetivo del proceso de entrenamiento es minimizar el error de la predicción y para ello, como se mencionó anteriormente, solo es posible modificar los pesos de la red [2007, Andina],[2000,Hilera] , [1993, S.Y. Kung].

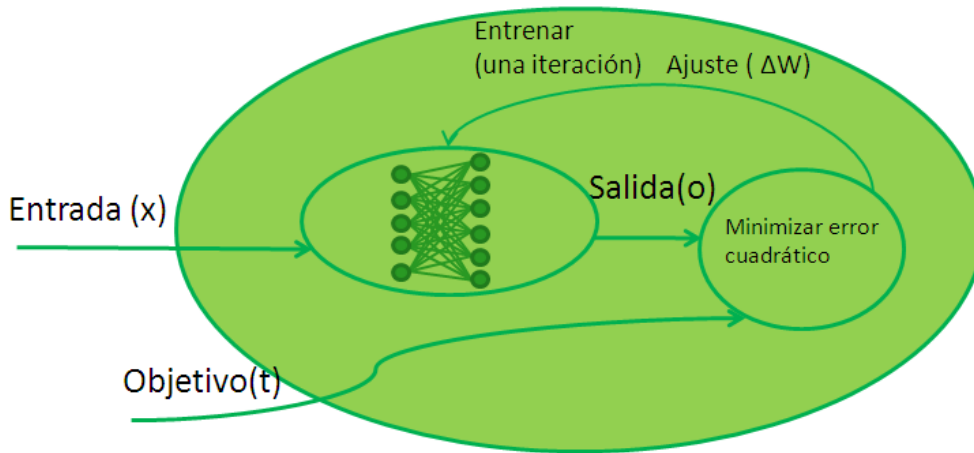


Figura 7.10 Esquema de entrenamiento de una red neuronal. Fuente [2007, Andina]

El proceso de entrenamiento es iterativo, se inicializa la red con pesos cargados de manera arbitraria como configuración inicial y luego se tiende a modificarlos de la mejor forma posible. Para ellos se utiliza la propagación del error hacia atrás mediante sus derivadas y es por ello que la red toma el nombre de back propagation. Para el error obtenido se encuentra un vector $\Delta \rightarrow w$ que sumado al vector de pesos $\rightarrow W$ se obtiene una red que arroja un error más pequeño para esa entrada.

Si se corre para la misma entrada este proceso varias veces, el resultado final sería una red con error nulo para ese valor (siempre y cuando la configuración de la red lo permita). Ese no es el objetivo, lo que se desea es entrenar a la red con varias entradas y luego ver qué sucede cuando ingresamos alguna que no estaba en el conjunto de

datos de entrenamiento. Es por esta razón que no se itera sobre un mismo elemento del conjunto de datos hasta eliminar el error sino que se realiza un acercamiento con un elemento, luego con otro y así hasta recorrer todo el conjunto de datos. A esta recorrida sobre el conjunto de datos se la suele denominar época de entrenamiento (*epoch*). El error no será bajo, pero la red se habrá acercado hacia una zona donde convergen todos los elementos. Al repetir el proceso varias veces, la red comenzará a entrenarse.

Figura 7.11

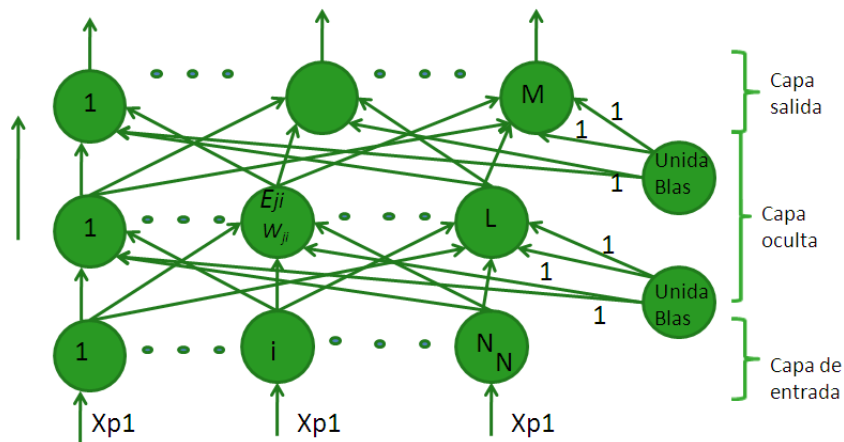


Figura. 7.11 Esquema Neuronal *Backpropagation*. Fuente [1993, S.Y. Kung].

En esta red, se interconectan varias unidades de procesamiento en capas, las neuronas de cada capa no se interconectan entre sí. Sin embargo, cada neurona de una capa proporciona una entrada a cada una de las neuronas de la siguiente capa, esto es, cada neurona transmitirá su señal de salida a cada neurona de la capa siguiente.

La Figura 7.11 muestra un ejemplo esquemático de la arquitectura de este tipo de redes neuronales [2008, Martin] [1993, S.Y. Kung].

8. TRABAJOS RELACIONADOS CON REDES NEURONALES APLICADAS A LA MEDICINA

Las tecnologías de la información en los últimos años se han ayudado en diversas disciplinas para analizar datos y proveer información para la toma de decisiones. En el campo médico se han desarrollado herramientas para la elaboración de diagnósticos alternos o como soporte a diagnósticos realizados por los especialistas [1997, Chan] , [2000, Hiler], estas herramientas se han empleado para procesar datos, en menor tiempo de lo que le lleva a un ser humano hacerlo [2006, Arzhaeva], [2007, Andina] permitiendo estudiar de mejor manera diversos comportamientos de virus, bacterias, cáncer [2003, Arnaldi], [2007, Dorantes] e incluso para hacer proyecciones de efectos secundarios que pudieran provocar ciertos medicamentos o medir el grado de implicaciones patológicas en seres humanos, así como, diagnósticos de acuerdo a un conjunto de manifestaciones clínicas que presentan los pacientes [2006, Axelrod].

Un campo de desarrollo en el cómputo que se ha estado trabajando favorablemente, es la inteligencia artificial, para el diagnóstico de diversas enfermedades en seres humanos. Se ha trabajado [2000, Duncan] para diagnosticar el cáncer de mama, utilizando redes neuronales artificiales (RNA) particularmente *backpropagation* y algoritmos genéticos, destacando que los algoritmos genéticos han dado mejores resultados que las redes neuronales. También se han empleado las RNA en la urología para la detección de cáncer de próstata utilizando el índice *ProstAsure*. Con resultados del 98% de aserciones positivas en el diagnóstico [2003, Cortagi]. En otro trabajo [2001, Boscoro] se implemento una RNA para diagnosticar el cáncer de mama, mediante historiales de mamografías, con memorias

asociativas. La red propuesta tiene una arquitectura flexible, lo que permite actuar no solo como una herramienta de predicción, sino como una herramienta de análisis. Los resultados del estudio mostraron que el índice ROC (*Receiver Operating Characteristic*, ó Característica Operativa del Receptor) área fue 0.8460.02. Estos son competitivos respecto a los obtenidos por *backpropagation*, permitiendo mejorar los resultados obtenidos por estas redes. Así mismo, [2004, Log] propuso un método híbrido basado en redes neuronales y algoritmos genéticos, para la clasificación de Síndrome de Cushing con lesiones concomitantes adrenocorticales. Se utilizaron marcadores para caracterizar el tejido como canceroso o no canceroso, incorporando técnicas de maquinas vectoriales, redes neuronales, arboles de decisión, mapas de características auto-organización y arboles recursivos de contraste máximo. Como conclusión se comentó que la herramienta de diagnóstico sirve para tener una mejor comprensión de los mecanismos implicados en la transformación maligna e información útil para planificar el tratamiento y prevención del cáncer.

9. METODOLOGIA

9.1 Pasos metodológicos

El proyecto se realizó en las siguientes fases:

1.- Identificación de las características del Síndrome de Cushing.

Estudiar la sintomatología del SC determinando los rasgos que permitan entrenar a la red neuronal artificial con la finalidad de que clasifique a los pacientes con SC a partir de los datos de comportamiento patológico que los médicos generan del diagnóstico diferencial.

2.- Determinación de los patrones que permitan identificar el Síndrome de Cushing.

A partir del estudio de las características del SC y los datos generados por los médicos generales se establecerá el conjunto de características del SC que permite a los especialistas determinar si un paciente es positivo para ser diagnosticado con SC. Este conjunto de datos son los que se emplearán en el presente trabajo.

3.- Selección del tipo de Red Neuronal a utilizar.

En base al patrón de SC obtenido se elegirá una arquitectura de red neuronal artificial que permita trabajar satisfactoriamente con los datos y con la finalidad de obtener un porcentaje de clasificación cercano al cien por ciento.

4.- Modelo la Red Neuronal Artificial.

En esta etapa se construyó la RNA propicia para clasificar datos que aportan evidencias de los síntomas del SC.

5.- Entrenamiento y Pruebas de la Red Neuronal a partir de las características del SC obtenidas.

El entrenamiento de la RNA se llevará a cabo mediante un subconjunto de datos exclusivos para el entrenamiento y otro subconjunto de datos exclusivos para las pruebas de clasificación, llevando un control estadístico de las corridas, representadas en gráficas de entrenamiento, prueba y porcentaje de clasificación para cada corrida.

9.2 Desarrollo Metodológico

Identificación de las características del Síndrome de Cushing.

Se proporcionaron Historiales Clínicos de 138 pacientes ingresados por sospecha clínica del Síndrome de Cushing (SC). Las características que se eligieron de acuerdo a los especialistas endocrinólogos, son: Estrías mórbidas, Hirsutismo, Obesidad central, Facies de luna llena, Plétora facial, Joroba de búfalo, Alopecia y Acné.

Determinar los patrones que permitan identificar el Síndrome de Cushing.

Para la representación de datos, se utilizó la siguiente nomenclatura: 0 para ausencia del síntoma, 1 para presencia del síntoma.

Etiqueta de Clasificación: -1 paciente con diagnóstico de SC y 1 paciente descartado del diagnóstico de SC.

Los datos se dividieron aleatoriamente, después se seleccionaron 110 casos (patrones) para el Entrenamiento y 28 de Prueba, equivalente al 80% y 20% respectivamente.

El siguiente ejemplo muestra la presencia y ausencia de los síntomas en un paciente candidato a ser diagnosticado con SC.

Estrías mórbidas	Hirsutismo	Obesidad central	Facies de luna llena	Plétora facial	Joroba de búfalo	Alopecia	Acné
1	0	1	0	1	1	1	0

Selección el tipo de Red Neuronal a utilizar:

Se definió la estructura de la Red *Back Propagation*, $net_1 = \sum_{j \in A} O_j w_{ji} \forall i: i \in \beta$, con 8 neuronas de entrada que corresponden a cada uno de los síntomas, 50 neuronas ocultas en la capa intermedia y una neurona de salida.

Se aplicó una distribución normal $\mu=0$ y $\sigma=1$, para la estandarización de valores de entrada, con el objetivo de acotar los valores entre -1 y 1 , para una mejor identificación de patrones por la red.

Modelar la Red Neuronal Artificial

En el presente trabajo se ha implementado una RNA con el algoritmo de retro-propagación para clasificar a los pacientes con SC, que permite aprender mediante un conjunto de

ejemplos (entrada-salida) comúnmente denominado conjunto de entrenamiento. Después de haber entrenado a la red mediante este conjunto, se puede obtener una salida coherente para una entrada dada. El algoritmo de aprendizaje proporciona una forma de entrenar una red multicapa con alimentación hacia adelante. Comienza alimentando los valores de la entrada de acuerdo a las ecuaciones que se muestran a continuación, utiliza el error de la salida de una neurona j en la iteración t calculándolo mediante: la diferencia entre la entrada dada y la salida deseada, como se muestra en la ecuación 6.

$$\varepsilon(t) = \frac{1}{2} \sum_{j \in C} e_j^2(t) \quad \text{Ec. 7}$$

La salida de la i -ésima neurona en la k ésima capa puede ser calculada mediante:

$$y_i^k = f \left(\sum_{j=1}^{n^{k-1}} w_{jl}^k \cdot y_j^{k-1} \right) \quad \text{Ec. 8}$$

donde $1 \leq k \leq K$, n_k se refiere al número de neuronas en la capa k . El error medio cuadrático de la salida es calculado mediante:

$$E = \frac{1}{2} \sum_{j=1}^{n^k} (d_j - y_j)^2 = \frac{1}{2} \sum_{j=1}^{n^k} \left[d_j - f \left(\sum_{j=1}^{n^{k-1}} w_{ij}^k \cdot y_i^{k-1} \right) \right]^2 \quad \text{Ec. 9}$$

donde i actualiza los pesos mediante paso descendiente

$$w_{ij}^k(t+1) = w_{ij}^k(t) - \eta \frac{\partial E}{\partial w_{ij}^k} \quad \text{Ec. 10}$$

donde η es el parámetro de aprendizaje y

$$\frac{\partial E}{\partial w_{ij}^K} = \frac{\partial E}{\partial y_j^K} \cdot \frac{\partial y_i^k}{\partial w_{ij}^K} = -(d_j - y_j) \cdot f'(net_j^K) \cdot y_i^{K-1} \quad Ec. 11$$

El error de la salida es retropropagado a la capa K-1 como:

$$\delta_j^K = \frac{\partial E}{\partial net_j^K} = -\frac{\partial E}{\partial y_j^K} \cdot \frac{\partial y_j^K}{\partial net_j^K} = (d_j - y_j) \cdot f'(net_j^K) \quad Ec. 12$$

mientras que el aprendizaje en la capa oculta esta dado por:

$$\sum_{j=l}^{n^{k+1}} (\delta_i^{k+1} \cdot w_{ji}^{k+1}) \quad Ec. 13$$

Donde δ_i^{k+1} es el error retropropagado desde la próxima capa, para el error en la k_{th} capa $\delta_i^k = e_i^k \cdot f'(net_i^K)$ será entonces propagado a la capa $k - 1$.

Entrenar y Probar a la Red Neuronal a partir de las características del SC obtenidas

Se obtuvieron los valores óptimos de η , neuronas ocultas y periodos de entrenamiento, con la finalidad de incrementar la precisión de clasificación, con los recursos óptimos.

Se aplicó la función *shuffle*: $[Y,I] = SHUFFLE(X)$ *regresa el índice de la matriz I, si X es un vector; entonces Y = X(I). Si X es un elemento de la matriz entonces, for j = 1:n, Y(:,j) = X(I(:,j),j);*. La función *shuffle* consiste en mezclar el orden de los patrones de entrada para el entrenamiento y no siempre se ingresen en el mismo orden, con la finalidad de incrementar las probabilidades de generalizar correctamente.

Se corrieron las funciones de *K-Fold* y *Cross Validation*, dividiendo los patrones de entrenamiento en k-1 conjuntos de datos, siendo k-1 el set de entrenamiento y k-n el conjunto de validación, para verificar las precisiones de clasificación del patrón de

Entrenamiento, Validación y Prueba.

Las precisiones de clasificación para los datos de entrenamiento, prueba y validación, se obtuvieron a partir del promedio de diez sesiones, entrenando, validando y probando cada uno de los conjuntos de datos.

10. RESULTADOS

En una primera aproximación de la propuesta planteada se trabajó con una Estructura de la red neuronal *BackPropagation*, conformada por 50 Neuronas en la capa oculta, 300 épocas de entrenamiento, y un parámetro de aprendizaje representado por $\eta=0.5$.

Se obtuvieron porcentajes de clasificación con 10 iteraciones (Corridas) con la estructura anterior, obteniendo un porcentaje de clasificación del conjunto de entrenamiento del 95.0909 %, con media de error de 0.47598, y una desviación estándar de 0.6976, el porcentaje obtenido permitió continuar con las pruebas de clasificación ya que es un rango aceptable, apoyándonos en la media del error y la desviación estándar para observar que la muestra utilizada es representativa, dado que la media y la desviación estándar son mínimas.

Respecto al conjunto de Prueba el porcentaje de clasificación fue de 89.0909 %, con una media del error de 0.5759 y una desviación estándar de 0.9976 permitiendo observar que el conjunto de prueba es representativo para el problema implementado, el incremento de la desviación estándar es debido a que las redes neuronales clasifican mejor en grandes conjuntos de datos por lo que el porcentaje de clasificación disminuye al ser el conjunto de datos de prueba, sin embargo, el porcentaje es válido para la clasificación. Con estos resultados se continuó la investigación para buscar parámetros óptimos e incrementar el porcentaje de clasificación al igual que precisión al generalizar.

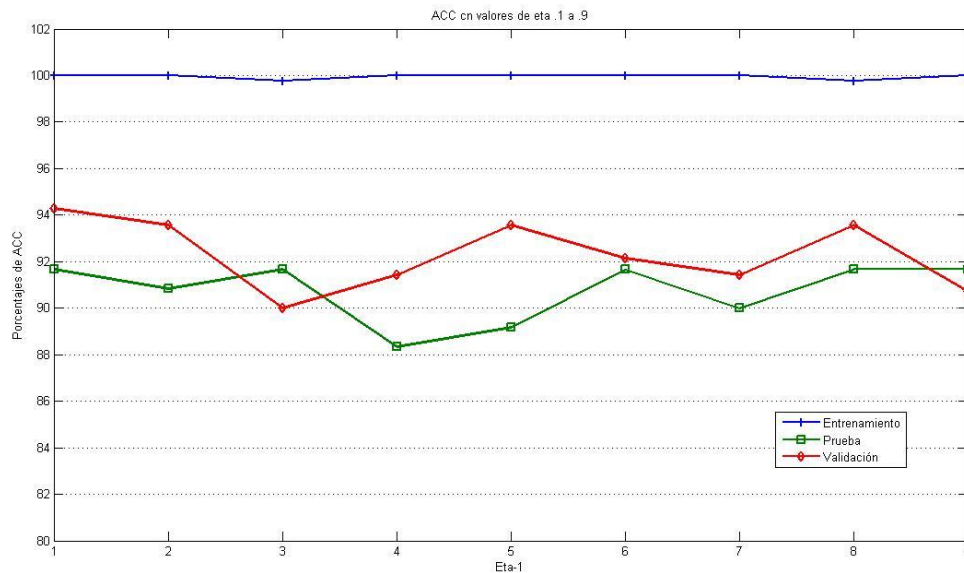


Figura 7.12. Gráfica de porcentaje de clasificación. Elaboración Propia

En la Figura 7.12 se grafica el porcentaje de clasificación, considerando el valor promedio de diez sesiones de clasificación de la RNA. La primera línea con marcas verticales en cada punto de *Eta* muestra el porcentaje de clasificación de las clases que promedian un 99.93% lo que significa que el entrenamiento de la RNA es óptimo con una clasificación adecuada. En la segunda línea (con marcas cuadradas) se muestran los porcentajes de clasificación para el conjunto de Prueba con un promedio de 91.03%, el porcentaje decremento respecto al conjunto de datos de entrenamiento debido a que se le presentan datos nuevos a la RNA, este porcentaje es óptimo y garantiza buenos resultados al generalizar con nuevos patrones. En la tercera línea (con marcas en forma de rombos) con un promedio de 93.89% se observa que los resultados de la clasificación para el conjunto de datos de validación se encuentra muy cercana a los arrojados por el conjunto de prueba. Con esto se asevera que los resultados arrojados por la red para las pruebas son resultados fehacientes y garantizan

la veracidad de los resultados de la red en el momento de generalizar.

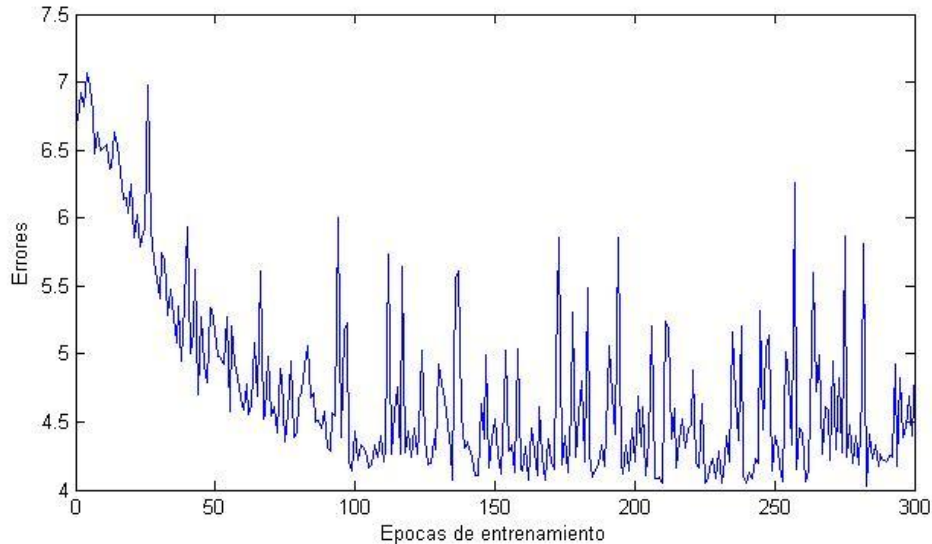


Figura 7.13. Valores de los errores en épocas de entrenamiento inicial Elaboración Propia

En la Figura 7.13 se muestra el comportamiento de la red neuronal artificial en 300 épocas de entrenamiento graficadas en el eje horizontal y en el eje vertical los errores de clasificación, con la oscilación que presenta la gráfica nos permite mostrar que la red no ha convergido y que se requiere hacer una modificación de sus parámetros de entrenamiento para que la red converga.

Para la búsqueda de los valores óptimos se realizaron 50 ejecuciones de la red con el entrenamiento, prueba y validación, para cada parámetro, esto con la finalidad de obtener la media y la desviación estándar del error cuadrático medio y poder elegir la mejor. Se seleccionaron cinco neuronas de la capa intermedia en base al análisis estadístico de la media y la desviación estándar en error, se probaron desde 4 hasta 10 neuronas en la capa

intermedia, los datos obtenidos se muestran en el Cuadro 7.4 de la media y la desviación estándar.

Cuadro7.4 Media y desviación estándar sobre error de entrenamiento. Elaboración Propia

Neuronas	Mean Error	Desviación Error
4	2.978221	0.040058
5	2.976321	0.037985
6	2.97848	0.06101
7	2.979598	0.043243
8	2.976627	0.039273
9	2.977013	0.038926
10	2.976195	0.036884

Los datos muestra un valor de *eta* fijado en 0.65 debido a que se probaron valores desde 0.01 hasta 0.1 utilizando búsqueda mediante malla fina, permitiendo fijar el valor óptimo de *eta* en 0.65. Con la modificación de *eta*, las neuronas en capa oculta y épocas de entrenamiento, se logro que la red convergiera, que significa que la red a convergido y ha encontrado una solución en el hiperplano de trabajo y se corroboraran los valores óptimos dado que la red converge, como se muestra en la Figura 7.14.

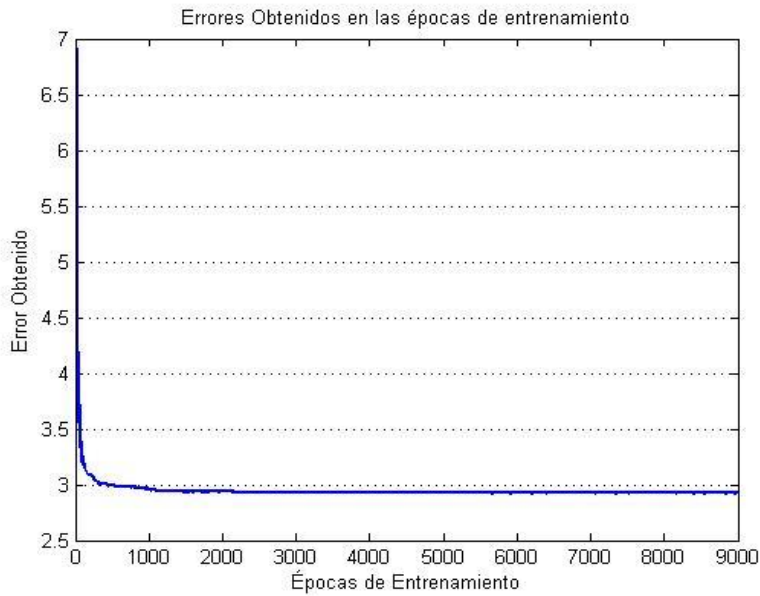


Figura 7.14 Valores de los errores en épocas de entrenamiento final. Elaboración Propia

En la Figura 7.15 se observa la diferencia entre la salida esperada y la obtenida, permitiendo observar que la oscilación de la gráfica describe de 5 a 8 errores en la clasificación, con ello se deduce que la variación de la grafica superan el 0.5 de error obtenido.

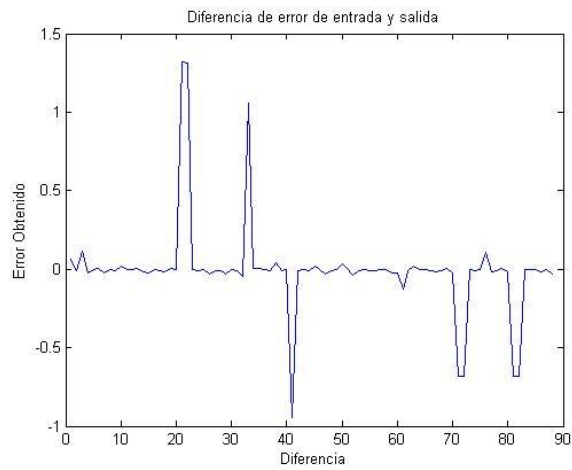


Figura 7.15. Diferencia entre la salida esperada y la salida generada. Elaboración Propia

Para obtener los valores óptimos, empleamos la búsqueda de malla, es decir, se va realizando la acotación de los valores específicos, con la finalidad de obtener el valor exacto que permita incrementar el porcentaje de clasificación, así como las neuronas en la capa oculta. Se utilizó validación cruzada para validar los modelos generados a partir de un conjunto de datos o muestra. La validación consiste en dividir en dos conjuntos complementarios los datos de la muestra, usar uno de ellos para construir el modelo y usar el otro para medir el rango de error del modelo construido como conjunto de prueba. En cada corrida, todos los pesos fueron aleatoriamente inicializados entre 0 y 1. Se implementó validación cruzada con $k=5$ (*k-fold – cross validation*) para el conjunto de datos de entrenamiento con la finalidad de validar los resultados obtenidos y garantizar la similitud del porcentaje de clasificación en el momento de generalizar.

La validación cruzada es una práctica estadística, que consiste en partir una muestra de datos en subconjuntos de tal modo que el análisis es inicialmente realizado en uno de ellos, mientras los otros subconjuntos son retenidos para su uso posterior en la confirmación y validación del análisis inicial. Es una técnica muy utilizada en Inteligencia Artificial para validar los modelos generados a partir de un conjunto de datos o muestra.

La validación consiste en dividir en dos conjuntos complementarios los datos de la muestra, usar uno de ellos para construir el modelo (*training set*) y usar el otro para medir el ratio de error del modelo construido (*test set*) como se muestra en la Figura 7.16.

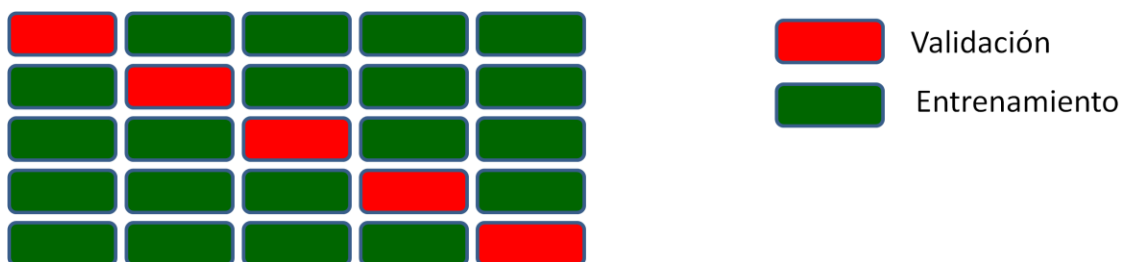


Figura 7.16. *K Fold Cross Validation.* Elaboración Propia

Los resultados fueron obtenidos al entrenar porcentajes de entrenamiento/prueba de 70/30 80/20 y 90/10. Las precisiones de clasificación para los datos de Entrenamiento (**ACC_E**), Validación (**ACC_V**) y Prueba (**ACC_P**), se obtuvieron a partir del promedio de diez sesiones, entrenando, validando y probando cada uno de los conjuntos de datos. El Cuadro 7.5 expone los resultados del estudio por efecto de utilizar diferente numero de capas ocultas y epocas de entrenamiento en la precisión de clasificación. Esto permite asegurar que la red clasificará adecuadamente al presentarle patrones nuevos para su clasificación y que el trabajo realizado es aceptable.

Cuadro 7.5. Resultados obtenidos en la clasificación. Elaboración Propia

ACC_E ENTRENAMIENTO	ACC_V VALIDACIÓN	ACC_P PRUEBA
100	91.666667	91.428571
100	91.666667	90
99.77273	89.666667	91.428571
99.77273	91.666667	91.428571
100	91.666667	93.571429
100	91.666667	93.571429
100	91.666667	91.428571
99.77273	91.666667	93.571429
100	91.666667	92.857143
100	91.666667	90.714286

En la Figura 7.17 se observan los datos graficados en el Cuadro 7.5, donde se muestra con claridad como los porcentajes de validación y prueba están dentro de los rangos de los considerados en el entrenamiento de la red. Así mismo, se muestran el porcentaje de clasificación en cada una de las sesiones, una vez que se tienen implementados los valores óptimos de *eta*, neuronas y épocas de entrenamiento. La diferencia que se hay entre el conjunto de entrenamiento y de prueba es menor a 10 puntos porcentuales. Esto indica que la red generaliza adecuadamente cuando se le presentan los nuevos casos que no correspondieron al entrenamiento. Se incrementó el porcentaje de clasificación para todos los conjuntos, por lo tanto la red clasificara adecuadamente al generalizar.

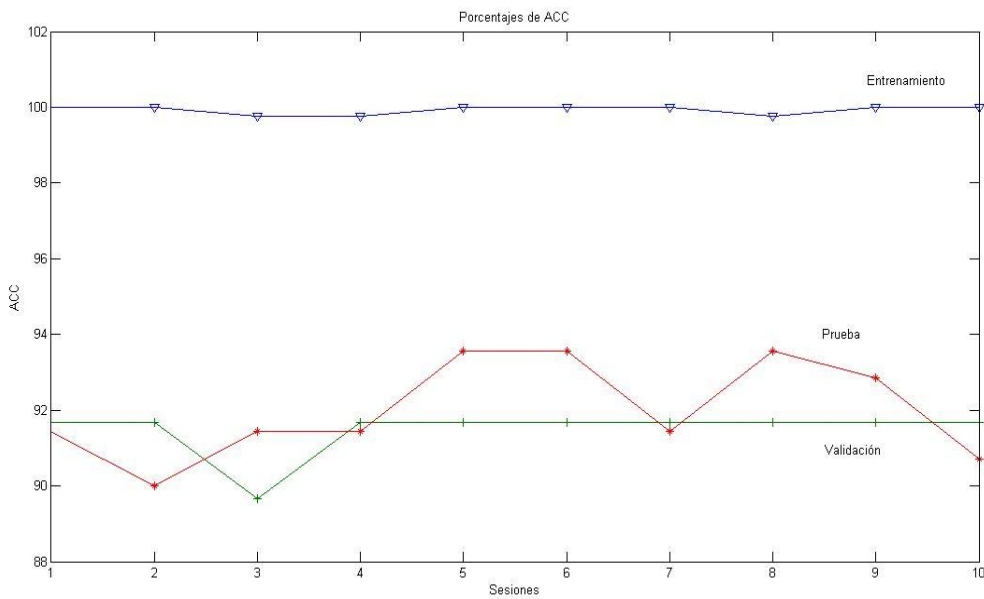


Figura 7.17 Representación de porcentajes de clasificación para entrenamiento, validación y prueba.

Elaboración Propia

11. DISCUSIÓN

Con los 8 síntomas que aportan los médicos para diagnosticar el Síndrome de Cushing (SC), la topología de la red neuronal *Backpropagation* propuesta y el promedio de los resultados de clasificación en el entrenamiento de 99.93 %, en la validación de 91.46 % y 92 % de prueba, permiten observar que la RNA propuesta ha clasificado adecuadamente. Con lo que se garantiza un funcionamiento similar en los casos nuevos donde se requiere generalizar puesto que son datos que la RNA nunca ha visto. De igual manera se puede considerar que del 100% de los casos presentados a la RNA tiene un promedio 94.43% de clasificación correcta y un error del 5.57% de error de clasificación con lo que se tendría alrededor de un 95% de ahorro en estudios desechados. Por lo tanto se incremento la clasificación de los mismos hasta un 45% más que los médicos generales que atienden a los pacientes en primera instancia, con lo que se disminuyen en un 95% los gastos del hospital en estudios químicos.

También se considera que el pre-diagnóstico del SC puede ser certero en un promedio de 95%. Esto significa que si el diagnóstico del síndrome de Cushing por los médicos generales de acuerdo a los datos proporcionados por el Hospital tienen una certeza del 40%, el 60% restante genera gastos económicos innecesarios. Con los resultados obtenidos con la red neuronal propuesta se incrementa la certeza de clasificación en un 54%, representando un ahorro aproximado de 8 mil pesos por paciente.

La RNA *backpropagation* ha permitido clasificar a pacientes con SC, de acuerdo a las ocho características que presentaban en el historial médico, es importante mencionar que las red

neuronal se modificó hasta obtener los parámetros óptimos, 300 épocas de entrenamiento. Al realizar mayor número de épocas o mayor número de neuronas en la capa oculta el error cuadrático medio no varía significativamente, sin embargo si se desean incrementar estos valores la RNA caerá en un sobre entrenamiento, decrementando los el porcentaje de clasificación.

La implementación considera una desventaja propia de las RNA's debido a que siempre presenta una variación en el resultado arrojado puesto que el método de aprendizaje estocástico repercute directamente en los resultados, variaciones que no se presentan con las máquinas de soporte vectorial, en las que se podría trabajar con la finalidad incrementar la clasificación de los patrones y de mantener el porcentaje de clasificación sin variaciones.

La información médica empleada en el cuadro clínico utilizado para los diagnósticos diferenciales proporcionada para el presente estudio fue ambigua en algunos de los casos, por lo que la incorporación de datos más específicos podría mejorar el funcionamiento y efectividad de la RNA

Se alcanzaron los objetivos del presente trabajo, mediante la identificación de las características del SC, definiendo el patrón de entrenamiento, seleccionando el tipo de red neuronal artificial a utilizar, modelándola y entrenando a la red neuronal para la clasificación satisfactoria de los pacientes con SC.

Las características del SC utilizadas para formar los conjuntos de datos que permitieron entrenar, probar y validar la RNA, impactan de manera significativa en el diagnóstico

diferencial del SC, a pesar de que los especialistas no han encontrado una combinación específica y única que permita realizar el diagnóstico irrefutable de padecer el SC sin requerir pruebas químicas, en la tesis de especialidad endocrinológica titulada: Diagnóstico diferencial del síndrome de Cushing de acuerdo al impacto de cada síntoma, del residente Zarate Gutiérrez Francisco, en el Hospital General de México, Especialidades de la Facultad de Medicina de la UNAM, demostró que las diversas combinaciones de los síntomas que se manejaron para los casos presentados en esta tesis como patrón de funcionamiento de la RNA, determinan la presencia del SC, sin embargo un paciente con SC no puede ser diagnosticado como positivo sino hasta después de los estudios de química sanguínea y orina de cortisol de 24 y 48 hrs.

Aproximadamente el 60% de los casos pre-diagnosticados inicialmente no padecen dicha enfermedad y sin embargo son enviados con los especialistas tras haber realizado estudios de química sanguínea y de orina, por lo que el trabajo del médico general produce un 60% de error en la canalización de los pacientes. La RNA incrementó el porcentaje de clasificación en un promedio de 54% respecto a los médicos generales, sin embargo no se puede garantizar que el trabajo de la RNA siempre sea el mismo.

12. CONCLUSIONES

Con el uso de la RNA *Backpropagation*, se pueden clasificar los pacientes con SC, dependiendo de los síntomas presentados, incrementando en un 32 % la certeza de clasificación respecto a la clasificación de un médico general.

Se pueden disminuir en un 80% los estudios a realizarse, permitiendo un ahorro aproximado de 40 mil pesos, aunado al tiempo de los especialistas e instalaciones del servicio médico

El tiempo de entrenamiento de la RNA se estableció en 300 períodos, puesto que si se seguía entrenando la red la precisión de clasificación caía hasta un 82% debido a un sobreentrenamiento provocando un descenso de la precisión de clasificación.

Se realizaron pruebas mediante graficas de malla para el valor de eta, en este caso particular fijándolo en .60, permitiendo obtener medias de clasificación de 92% a diferencia de otros valores.

En la presente tesis, se modelo una RNA implementando el algoritmo de retro-propagación para clasificar pacientes con SC, la RNA auxilia para obtener un diagnostico rápido y en la mayoría de los casos más precisa que los médicos generales, basándonos en la precisión de clasificación la cual fue mejorada en un 80%.

La información médica que se obtuvo para la presente investigación proporcionada con un mayor grado de sintomatología representaría una oportunidad a identificar un patrón de comportamiento sobre las características que incrementen la asertividad en el pre-diagnostico del SC por el conjunto de síntomas que se manejaron, a pesar de que los médicos especialistas no tengan una combinación especifica de síntomas únicos del SC

para realizar su diagnóstico sin prueba química.

Con la implementación de la RNA se ha incrementado el porcentaje de clasificación de los pacientes con SC en un 50 % respecto a la clasificación que realizan los médicos generales al realizar el pre-diagnóstico del SC.

13.-BIBLIOGRAFIA

[2006, **Arzhaeva**] Arzhaeva Y. Van Ginneken, B.; Tax, *Image Classification from Generalized Image Distance Features: Application to Detection of Interstitial Disease in Chest Radiographs Tax Image Sciences Institute*, Delft University of Technology, Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06) 0-7695-2521-0/06 , 2006

[2007, **Andina**] Andina Diego, **Duc Truong Pham Hardcover** , *Computational Intelligence (Hardcover)*, **Feb 20, 2007**

[2003, **Arnaldi**] Arnaldi G, Angeli A, Atkinson AB, et al. *Diagnosis and complications of Cushing's syndrome: A consensus statement*, J Clin Endocrinol Metab. 2003;88-5595-5602.

[2006, **Axelrod**] Axelrod L, *Glucocorticoid therapy*. Medicine 2006; 56:39-65.

[2001, **Boscoro**] Boscoro M, Bazan. Cushing's syndrome. Lancet 2001; 357:783-91

[1997, **Chan**] Chan, H. P.; Doi, K.;Galhotra, S.; Vyborny, C. J.; MacMahon, H.; Jokich,, *Image Feature Analysis and Computer-Aided Diagnosis in Digital Radiography: Part 1 Automated Detection of Microcalcifications in Mammography*, P. M. 1997. Medical Physics, vol. 14, pp. 538-548.

[2003, Cortagi] Cortagi B, et al. Occult Cushing's syndrome in type 2 diabetes. *JClin Endocrinol Metab* 2003; 88:5808-13

[2007, Dorantes] Dorantes Alicia Yolanda, *Endocrinología Clínica, Manual Moderno*. 3° Edición Sociedad Mexicana de Nutrición y Endocrinología, 2007.

[2000, Duncan] Duncan, J. S.; Ayache, N. 2000. Medical Image Analysis: Progress over Two Decades and the Challenges Ahead. *Ieee transactions on pattern analysis and machine intelligence*, vol. 22, no. 1, january 2000

[2007, Espinsa] Espinosa de los Monteros-Sánchez AL, Valdivia-López. *Consejo en el diagnostico y tratamiento del síndrome de Cushing*. *Rev Endocrinología y Nutrición* 2007;15(4) 53-512.

[2008, Greenspan] Greenspan Francis y Davis Garner, *Endocrinología Básica y clínica*, 5° Edición, Manual Moderno, Sociedad Mexicana de Nutrición y Endocrinología, 2008.

[2000, Hilera] Hilera José R y Martínez Víctor J., *Redes Neuronales Artificiales*, 2000. Alfaomega. Madrid. España.

[2006, Hurd] Hurd Robert, MD, *Síndrome de Cushing*,, Department of Biology, College of Arts and Sciences, Xavier University, Cincinnati, OH. Review provided by VeriMed Healthcare Network.

Traducción y localización realizada por: DrTango, Inc.(2006).

[**2001, Lindholm**] Lindholm J, Jull SJ, Jorgensen JO, et al. *Incidence and late prognosis of Cushing's syndrome: a population based study. Clin Endocrinol Metab* 2001; 86:117-23.

[**2004, Log**] Loog, M.; van Ginneken, *Static posterior probability fusion for signal detection: applications in the detection of interstitial diseases in chest radiographs, in International. 2004, Conference on Pattern Recognition, J. Kittler, M. Petrou, and M. Nixon, Eds.*

[**2008, Martin**] Martin del brío Alfredo, *Redes Neuronales y Sistemas Borrosos*, 3° ed. SANZ, Bonifacio, 440 págs. Coedición: Ra-Ma, Alfaomega,2008 .

[**2006, Newell**] Newell Price H, *Cushing's syndrome. Lancet* 2006; 367:1605-17

[**1993, S.Y. Kung**] S. Y. Kung, *Digital Neural Networks*, 1993. PTR Prentice Hall, Inc.

[**1999, Sklansky**] Sklansky, J.; Ballard, D. 1999. *Tumor Detection in Radiographs, Computers and Biomedical Research*, vol. 6, no. 4, pp. 299-321, Aug. 1999.

[**Sluimer, 2003**] Sluimer, I.; van Waes, P.F.; Viergever, M. A.; *Van Ginneken, Computer-aided diagnosis in high resolution CT of the lungs*, 2003., *Medical Physics*, Vol. 30(12): 3081–3090.

[**1999, Stevens**]Stevens, *Ejection Fraction in Man Without Cardiac Catheterization*, *Am. J. Cardiology*, vol. 28, pp. 985- 991, Nov. 1999

[1971, Strauss] Strauss, H.W. *A Scintiphotographic Method for Measuring Left Ventricular*, 1971.

Pizer, S.; Todd-Pokropek, A.E. 1978. Improvement of Scintigrams by Computer Processing, *Seminars in Nuclear Medicine*, vol. 8, no. 2, pp. 125-146, Apr. 1978.

[2004, Pajares] Pajares Gonzalo, *Imágenes Digitales, procesamiento práctico*, 2004, 970-15-0951, pág 106. Alfaomega

[2008, Pironello] Pironello R, *Cushing's Syndrome. Endocrinol Metab Clin*, 2008: 37:135-49

[2003, Uchiyama] Uchiyama, Y.; Katsuragawa, S.; Abe, H.; Shiraishi, J.; Li, F.; Li, Q.; Zhang, C.-T.; Suzuki, K.; Doi, K. 2003. *Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography*, *Medical Physics*, Vol. 30(9): 2440–2454.

[1999, Uppaluri] Uppaluri, R.; Hoffman, E.A.; Sonka, M.; Hartley, P.G.; Hunninghake, G.W. *Computer recognition of regional lung disease patterns, American Journal of Respiratory and Critical Care Medicine*, McLennan, G. 1999. , Vol. 160(2): 648–654. Walpole, R. E; Myers, R. H. 1999. *Probabilidad y estadística para ingenieros*, Pearson Educación, ISBN 9701702646, 9789701702642, p 229.

[2004, **Viñuela**] Viñuela Pedro Isasi, *Redes de Neuronas Artificiales Un enfoque práctico*, 1a edición 2004, Páginas 248

[2007, **Zrimec**] Zrimec, T.; Busayarat, S. *A System for Computer Aided Detection of Diseases Patterns in High Resolution CT Images of the Lungs*. Computer-Based Medical Systems. 2007, CBMS apos;07. Twentieth IEEE International Symposium on Volume, 20-22 June, pp. 41

[1973, **Sklansky**] Sklansky, J.; Ballard, D. 1973. Tumor Detection in Radiographs. Computers and Biomedical Research, vol. 6, no. 4, pp. 299-321, Aug. 1973.

[1978, **Pizer**] Pizer, S.; Todd-Pokropek, A. E. 1978. Improvement of Scintigrams by Computer Processing. Seminars in Nuclear Medicine, vol. 8, no. 2, pp. 125-146, Apr. 1978.

ANEXO



Servicio de Endocrinología



SECRETARÍA
DE SALUD

SALUD

México, D. F. a 6 de Julio del 2011.

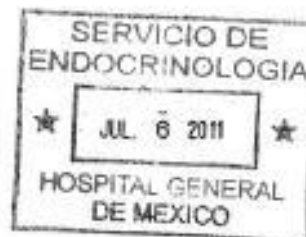
A QUIEN CORRESPONDA

Por medio de la presente hago de su conocimiento que se le proporciono información al C. Luis Márquez Sánchez, sobre las manifestaciones clínicas que presentan los pacientes con Síndrome de Cushing que acudieron a nuestro Hospital, cabe mencionar que la base de datos fue resultado de uno de los múltiples proyectos de investigación que se llevan a cabo en nuestro servicio y cuyos resultados han sido publicados con el título "Comparación del Índice de Gravedad del Síndrome de Cushing entre pacientes con Origen Endocrino y Yatrogeno de la Enfermedad", Publicado en Organó de la Sociedad Española de Endocrinología y Nutrición 2010; 57:426-33.

Sin otro particular quedo de usted para cualquier duda o aclaración con respecto a la información proporcionada.

ATENTAMENTE

Dr. Valentín Sánchez Pedraza
Endocrinólogo
Profesor Adjunto de Curso de Postgrado
en Endocrinología



Anexo cuestionario



MÉXICO

HOSPITAL GENERAL DE



**CARTA DE CONSENTIMIENTO INFORMADO
AUTORIZACIÓN PARA PARTICIPAR EN UN PROYECTO DE INVESTIGACIÓN
TÍTULO DEL PROTOCOLO**

Comparación del Índice de Gravedad en Cushing (CSI).

1. LA NATURALEZA Y PRÓPOSITO DE ESTE ESTUDIO.

Le ha sido solicitado tomar parte en un estudio para conocer el impacto del uso de glucocorticoides (cortisona) en su cuerpo y los cambios y complicaciones que ha generado. Si usted no ha tomado cortisona su participación nos ayudará a comparar las diferencias con los que si lo han hecho. Es importante que usted lea y comprenda información en este consentimiento informado; sienta la libertad de hacer preguntas hasta que usted entienda completamente.

2. EXPLICACION DE LOS PROCEDIMIENTOS A SEGUIR.

Una vez que Usted decida participar en este estudio, durante la primera consulta se le realizará una historia clínica completa, exploración física. Se le solicitará que llene 3 encuestas que le serán proporcionadas por una persona designada para estos fines, la cual se identificara con Usted. El llenado de estas encuestas le tomará alrededor de 30 min. Si usted no toma cortisona solo llenará 2 encuestas. Además se le realizará una venopunción para tomar muestras de sangre para el estudio de sus niveles de glucosa, potasio y hormonas como el cortisol y la ACTH que suelen afectarse por el consumo de cortisona.

Posteriormente y solo para los que toman cortisona, se le solicitará que acuda a revisión a las 6 semanas para repetir la revisión clínica y el llenado de dos cuestionarios y también se tomará una nueva muestra de sangre. Cada cita será confirmada previamente vía telefónica.

A) DURACION ESPERADA DEL ESTUDIO Y NÚMERO DE SIJETOS ESPERADOS EN PARTICIPAR.

Este estudio durará 6 meses, durante los cuales Usted asistirá a una o dos citas (la inicial y a las 6 semanas); aunque si Usted requiere atención médica antes de su cita podrá solicitarlo con su médico tratante.

3. POSIBLES RIESGOS Y MOLESTIAS.

En este estudio, las posibles molestias que le pueden ocasionar, es el tiempo que le lleve contestar los cuestionarios y la venopunción

4. POSIBLES BENEFICIOS DEL ESTUDIO.

Los resultados de este estudio le servirán para tener un mejor conocimiento de su enfermedad y evitar complicaciones y ayudar a otras personas que conozcan el impacto del uso de cortisona. Sin embargo, es posible que usted no reciba beneficios directos de

su participación en este estudio.

5. OFRECIMIENTO PARA CONTESTAR LAS PREGUNTAS SOBRE ESTE ESTUDIO.

Si usted tiene preguntas sobre la conducción de este estudio, usted deberá de contactar a: Dr. Edgar Fernando Duran en las instalaciones del hospital en un horario de 8:00 a 14:00.

6. DERECHO DE CANCELAR SU PARTICIPACION EN ESTE ESTUDIO /PARTICIPACION VOLUNTARIA.

Usted es libre de dejar este estudio en cualquier momento. Su decisión tanto de no participar en este estudio como de dejarlo en cualquier momento después de haberlo iniciado no afectará su atención médica futura.

7. CONFIDENCIALIDAD.

Todos los registros obtenidos mientras usted participa en este estudio así como todos los registros hospitalarios y del consultorio, estarán disponibles solo para el personal médico de la institución involucrado en su atención.

Los resultados de este estudio, incluyendo prueba de laboratorio, podrán publicarse para fines científicos bajo el entendido que su identidad no será revelada.

He leído, o lo han leído para mí, la información anteriormente escrita antes de firmar esta forma de consentimiento. Yo he tenido siempre la oportunidad amplia de preguntar y he recibido las respuestas que han satisfecho dichas preguntas. Si no participo o si descontinúo mi participación en este estudio no seré penalizado y no renunciaré a ninguno de mis derechos legales. Por lo tanto yo acepto voluntariamente tomar parte en este estudio.

He recibido copia de este acuerdo de consentimiento informado.

Por la firma de este consentimiento informado, no renuncio a ninguno de mis derechos legales.



**HOSPITAL GENERAL DE MÉXICO
PARA SER LLENADO POR EL MÉDICO**



Nombre Paciente: _____
Edad: _____ Sexo: _____ Escolaridad: _____
Expediente: _____
Alcoholismo: _____ Tabaquismo: _____
Otras toxicomanías: _____
Enfermedades concomitantes: _____
Diagnósticos de filtro: _____

PARA EL ÍNDICE DE GRAVEDAD DEL SÍNDROME DE CUSHING.

1. Lesiones Cutáneas (Estrías Mórvidas, Violáceas – Purpuras, mayores a 1 cm de ancho).
A. Si
B. No
2. Obesidad central (Estomago abultado, con extremidades delgadas).
A. Si
B. No
3. Distribución facial de la grasa (Facies de luna llena).
A. Si
B. No
4. Plétora facial (mejillas enrojecidas).
A. Si
B. No
5. Hirsutismo (Bigotes en mujeres, bello en muslos, hombros, brazos, espalda, estomago).
A. Si
B. No
6. Joroba de Búfalo (Pólipo de grasa parte inferior del cuello, base de la nuca).
A. Si
B. No
7. Alopecia (Falta de cabello en las mujeres).
A. Si
B. No
8. Acné facial y corporal.
A. Sí
B. No