UNIVERSIDAD MAYOR DE SAN ANDRÉS FACULTAD DE CIENCIAS PURAS Y NATURALES CARRERA DE INFORMÁTICA



TESIS DE GRADO

PREDICCIÓN DEL CONSUMO DE GAS NATURAL APLICANDO REDES NEURONALES

PARA OPTAR AL GRADO DE LICENCIATURA EN INFORMÁTICA MENCIÓN INGENIERÍA DE SISTEMAS INFORMÁTICOS

Postulante: Eddy JahuiraVelasquez

Tutor Metodológico: M. SC. Aldo Ramiro Valdez Alvarado

Asesor: M. SC. Carlos mullisaca Choque

LA PAZ – BOLIVIA

2013



UNIVERSIDAD MAYOR DE SAN ANDRÉS FACULTAD DE CIENCIAS PURAS Y NATURALES CARRERA DE INFORMÁTICA



LA CARRERA DE INFORMÁTICA DE LA FACULTAD DE CIENCIAS PURAS Y NATURALES PERTENECIENTE A LA UNIVERSIDAD MAYOR DE SAN ANDRÉS AUTORIZA EL USO DE LA INFORMACIÓN CONTENIDA EN ESTE DOCUMENTO SI LOS PROPÓSITOS SON ESTRICTAMENTE ACADÉMICOS.

LICENCIA DE USO

El usuario está autorizado a:

- a) visualizar el documento mediante el uso de un ordenador o dispositivo móvil.
- b) copiar, almacenar o imprimir si ha de ser de uso exclusivamente personal y privado.
- c) copiar textualmente parte(s) de su contenido mencionando la fuente y/o haciendo la referencia correspondiente respetando normas de redacción e investigación.

El usuario no puede publicar, distribuir o realizar emisión o exhibición alguna de este material, sin la autorización correspondiente.

TODOS LOS DERECHOS RESERVADOS. EL USO NO AUTORIZADO DE LOS CONTENIDOS PUBLICADOS EN ESTE SITIO DERIVARA EN EL INICIO DE ACCIONES LEGALES CONTEMPLADOS EN LA LEY DE DERECHOS DE AUTOR.

DEDICATORIA

A Dios, por la inspiración y fortaleza en todo momento.

A mis padres queridos ejemplos de voluntad, esfuerzo y amor, quienes son el gran impulso de salir siempre adelante ante cualquier inconveniente.

AGRADECIMIENTOS

Primeramente agradecer a mis padres Felipe y Tereza, a mi hermano(a) que me brindaron todo suapoyo durante este tiempo.

Al M. Sc. Aldo Ramiro Valdez Alvarado, tutor de la presente tesis, ya que gracias a su guía,apoyo, inspiración y consejos durante todo este tiempo de estudios en launiversidad, finalmente me permitieron concluir mi tesis.

Luego agradecer mi revisor al M. Sc.Carlos Mullisaca Choque,por su valiosa colaboración, buena voluntad de guiar y revisar el presente trabajo.

A mi familia porque siempre me apoyaron con sus consejos y me brindaron todo el cariño que es fundamental en la vida.

ÍNDICE

CAPITULO I MARCO INTRODUCTORIO

1.1.Introducción]
1.2.Antecedentes	2
1.3.Planteamiento Del Problema	∠
1.3.1.Problema Central	∠
1.3.2.Problemas Secundarios	5
1.4.Definición De Objetivo	5
1.4.1.Objetivo General	5
1.4.2.Objetivos Específicos	5
1.5.Hipótesis	6
1.5.1.Operacionalización De Variables	6
1.6.Justificación	6
1.6.1.Económica	6
1.6.2.Social	
1.6.3.Científica	
1.7.Alcances y Limites	
1.7.1.Alcances	
1.7.1.1.Alcance Temporal	
1.7.1.2.Alcance Espacial	8
1.7.2.Limites	8
1.8.Aportes	8
1.8.1.Aporte Práctico	8
1.8.2.Aporte Teórico	8
1.9.Metodología	9
CAPITULO II	
MARCO TEÓRICO	
2.1.Inteligencia Artificial	1
2.2.Redes Neuronales	13
2.2.1.Introducción a las Redes Neuronales	13

2.2.2.Historia de las Redes Neuronales	15
2.2.3.El Modelo Biológico	16
2.2.4.Elementos de una Red Neuronal Artificial	18
2.2.5.Backpropagation	20
2.2.5.1.Estructura y Aprendizaje	21
2.2.5.2.Pasos Para Aplicar el Algoritmo de Entrenamiento	22
2.2.5.3.Número de capas ocultas	24
2.2.6.Aplicación de las redes neuronales artificiales	25
2.3.Metodologías para Predicción	26
2.3.1.Metodología para la Predicción con Redes Neuronales Artificiales	26
2.3.1.1.Preparación de los Datos	26
2.3.1.2.Determinación de la Arquitectura de Red	29
2.3.1.3.Diseño de Estrategia de Aprendizaje o Entrenamiento de la Red	29
2.3.1.4.Evaluación Global	30
2.3.2.Método de Medias Móviles como Técnica Estadística de Predicción	30
2.4.Gas Natural	32
2.4.1.Características del Gas Natural	33
2.4.2.Consumo de Gas Natural	33
2.4.3.Producción de Hidrocarburos	35
2.4.3.1.Producción de Gas Natural por Campo	35
2.4.4.Usos Principales e Importancia	37
2.4.4.1.Precios de Gas Natural por Sector y Empresa Demandante	38
2.4.4.2. Ventajas de Seguridad y Ambientales sobre Otros Combustibles	40
CAPITULO III	
MARCO APLICATIVO	
3.1.Introducción	42
3.2.Metodología para Redes Neuronales Predictivas	42
3.3.Preparación de los Datos	42
3.3.1.Adquisición de Datos	42
3.3.2.Estructura de Datos	47
3.3.3.Normalización de Datos	48
3.3.4.División de Datos	50
3.4.Determinación de la Arquitectura de la Red	51

3.4.1.Descripción de la RNA para el Consumo de gas Natural	52
3.5.Diseño de la Estrategia de Aprendizaje o Entrenamiento de la Red	53
3.5.1.Regla de Aprendizaje	54
CAPITULO IV	
PRUEBA DE HIPÓTESIS	
4.1.Introducción	58
4.2.Modelado con medias móviles	58
4.3.Procedimiento de la Prueba de Hipótesis	63
4.3.1.Planteamiento de Hipótesis Nula y Alternativa	64
4.3.2.Selección del nivel de significancia	65
4.3.3.Estadígrafo de Prueba	66
4.3.4.Regla de Decisión	67
4.3.5.Toma de Decisión	68
CAPITULO V	
CONCLUSIONES Y RECOMENDACIONES	
5.1. Introducción	69
5.2. Conclusiones	69
5.2.1. Estado de los Objetivos	69
5.2.2. Estado de la Hipótesis	70
5.3. Recomendaciones	71
5.3.1. Trabajos Futuros	71
5.3.2. Líneas de Investigación	71
Bibliografía	72

ÍNDICE DE FIGURAS

CAPITULO II MARCO TEÓRICO

Figura 2.1. Neurona y Conexiones sinápticas	17
Figura 2.2.Esquema de una Red Neuronal	19
Figura 2.3. Similitudes entre una Neurona Biológica y una Artificial	19
Figura 2.4. Función de activación	20
Figura 2.5. Modelo macro	29
Figura 2.6.Evolución de la producción de hidrocarburos 2012	35
Figura 2.7.Campos de Exploración	36
Figura 2.8.Estructura de la comercialización de hidrocarburos	41
CAPITULO III	
MARCO APLICATIVO	
Figura 3.1. Censo Nacional de población y vivienda	43
Figura 3.2. Consumo de gas natural en el mercado interno	46
Figura3.3. Consumo de gas natural en el mercado interno	46
Figura 3.4. Arquitectura de la red	51
Figura 3.5Interfaz de la presentación	54
Figura 3.6.Interfaz de la construcción de la red	55
Figura 3.7.Interfaz de predicción de datos	55
Figura 3.8.Interfaz de datos para la gestión 2012	56

ÍNDICE DE TABLAS

CAPITULO II MARCO TEÓRICO Tabla 2.1.Instalaciones de redes de gas domiciliario.......37 CAPITULO III MARCO APLICATIVO Tabla 3.3. Regalías por departamento (MM\$us) 201245 **CAPITULO IV** PRUEBA DE HIPÓTESIS Tabla 4.3.Componente estacional 60 Tabla 4.6. Regresión lineal 62

RESUMEN

El gas natural es una de las varias e importantes fuentes de energía no renovables formada por una mezcla de gases y componentes ligeros, y es indispensable para la vida diaria de la población, también el gas natural se usa en la producción de electricidad como combustible de motores de combustión interna, más conocida como "Termoelectricidad", resulta ser una alternativa conveniente por el bajo costo de este combustible pero se hace difícil su implementación en regiones alejadas de los yacimientos de gas, por la necesidad de extender gasoductos de elevado costo. Pero lo extraordinario es el costo bajo de este combustible, también se utiliza en la industria, sectores Residencial, Comercial, Transporte Vehicular y el sector consumidores directos.

Para realizar planes y gestiones es importante tener una herramienta que permita predecir el consumo de gas natural a futuro. Las herramientas que brinda la inteligencia artificial en cuanto a la predicción se ajustan a este requerimiento, las redes neuronales artificiales son herramientas que pueden trabajar con información imprecisa, reduciendo la incertidumbre y devolviendo datos sin ruido.

En el presente trabajo se hace un análisis de variables que influyen en el consumo de gas natural, se desarrolla un modelo de predicción con redes neuronales tomando como variables al recurso energético (electricidad), y las regalías por el IDH, con un 0.05 de error. Posterior al modelado se realizo la predicción del consumo de gas natural para el año 2012 y se obtuvo un error de 0.04, observamos que es menor en un punto respecto a lo planteado lo cual nos da la certeza de que de que el modelo cumple con nuestras expectativas.

ABSTRACT

Natural gas is one of several important non-renewable energy sources formed by a mixture of gas and lighter components, and is indispensable to the daily life of the population, also the natural gas used in the production of electricity and fuel internal combustion engines, better known as "Thermoelectricity", turns out to be a convenient alternative for the low cost of the fuel but is difficult to implement in regions far from the gas fields, the need to extend pipelines expensive. But what is extraordinary is the low cost of the fuel, also used in industry, Residential, Commercial, Vehicular Transport and direct consumer sector.

To make plans and arrangements is important to have a tool to predict natural gas consumption in the future. The tools provided by the artificial intelligence in terms of prediction fit this requirement, artificial neural networks are tools that can work with imprecise information, reducing uncertainty and restoring data without noise.

In this paper an analysis of variables influencing the consumption of natural gas is a model prediction with neural networks using as variables the energy source (electricity), and royalties for the HDI is developed with a 0.05 error. Modeled after predicting natural gas consumption for 2012 was performed and an error of 0.04 was obtained, we note that a point is lower compared to the points which gives us the certainty that the model meets our expectations.

CAPITULO I

MARCO INTRODUCTORIO

1.1. Introducción

La predicción de sucesos futuros ha fascinado al hombre desde los tiempos antiguos. En los primeros relatos escritos las referencias a la predicción del futuro son constantes. La predicción requiere un estudio meticuloso de los sucesos pasados y de las relaciones existentes entre ellos, para partir de ahí, intentar extrapolar un suceso futuro.

En la ciencia la posibilidad de predecir se convierte en una herramienta fundamental, pero particularmente en problemas reales, los métodos matemáticos son diseñados para el tratamiento de este tipo de problemas, el más conocido es el proceso autoregresivo integrado de media móvil, la media móvil. Estos modelos de predicción que se utilizan muy frecuentemente, se fundamentan en el supuesto implícito de la linealidad del sistema que genera la trayectoria de las variables. En este contexto, existe evidencia empírica sobre la eficacia para predecir la evolución y comportamiento de los campos de la medicina, finanzas, economía, y la industria. Basándose en esto hoy en día se pueden encontrar trabajos aplicados a estas disciplinas utilizando técnicas o métodos como algoritmos genéticos, redes neuronales, lógica difusa que son parte de la inteligencia artificial.

El problema que daremos solución con uno de estos métodos en este caso con las redes neuronales, es la predicción del consumo de gas natural, Según el documento de la estatal petrolera, el Sector Eléctrico muestra el mayor consumo de gas natural dentro del mercado interno, le sigue muy de cerca los sectores Residencial, Comercial, Industrial y de Transporte Vehicular, también el sector Consumidores Directos y Otros.

En los últimos años el consumo de gas natural aumentó, y en los posteriores periodos se

incrementara por eso existe la necesidad de calcular cuánto tiempo durara las reservas de acuerdo al consumo de este energético. Las redes neuronales son técnicas paramétricas muy

utilizadas en diversos ámbitos de la ciencia e ingeniería por que permiten resolver problemas complejos que muchas veces no son fáciles de resolver, las redes neuronales permiten obtener un modelo no explicito que relaciona un conjunto de variables de salida con un conjunto de variables de entrada del modelo, más adelante se realizara el desarrollo del modelo de predicción del consumo de gas, para ello se utilizara las redes neuronales para realizar predicciones, obteniendo un resultado confiable que ayudara en la planificación y la toma de decisiones oportunas.

1.2. Antecedentes

En la gestión de recursos energéticos se han implementado técnicas y métodos, particularmente en problemas de predicción, son útiles las redes neuronales.

Actualmente se realizaron diversos estudios del consumo de gas natural con diversos métodos de predicción a nivel mundial, que son de gran importancia para el abastecimiento de gas natural, poder predecir con anticipación el consumo, lo que hace necesario que los responsables de atender el suministro tengan que poner en práctica una serie de operaciones coordinadas para responder a grandes incrementos del consumo de gas natural.

Aplicamos un modelo de red neuronal artificial que basa su funcionamiento en la manera en que el cerebro procesa información y que describe una serie de modelos de redes neuronales que utilizando métodos de aprendizaje supervisado y no supervisado abordan problemas tales como el reconocimiento y la predicción de patrones.

En la actualidad la industria gasífera creció y es uno de los pilares fundamentales en el plano económico del país, la industria gasífera tiene tres actividades la generación,

transmisión y distribución, es necesario una coordinación entre estas tres actividades, porque un problema en cualquiera de las actividades de este sector tendría efectos negativos en la economía, porque está directamente relacionada con la actividad de los demás sectores debido a que es un insumo para el funcionamiento de las mismas, la población no puede dejar de utilizar este producto porque es un elemento indispensable para la vida diaria de la población, también el gas natural se usa en la producción de electricidad como combustible de motores de combustión interna, más conocida como "Termoelectricidad", resulta ser una alternativa conveniente por el bajo costo de este combustible pero se hace difícil su implementación en regiones alejadas de los yacimientos de gas, por la necesidad de extender gasoductos de elevado costo. Pero lo extraordinario es el costo bajo de este combustible, también se utiliza en la industria, sectores Residencial, Comercial, Transporte Vehicular y el sector consumidores directos.

Los recursos de un país como el gas natural son de mucha importancia para el funcionamiento de la economía, por eso el modelo predictivo basado en redes neuronales alimentado con variables que afectan al consumo de gas natural sera capaz de predecir el consumo gas natural a futuro.

Los trabajos realizados en el área de predicción son muchos a continuación mencionaremos algunos trabajos realizados en diversos campos con el fin de predecir comportamientos futuros.

"PREDICCIÓN DEL CONSUMO DEL AGUA POTABLE DE LA CIUDAD DE LA PAZ APLICANDO REDES NEURONALES Y LÓGICA DIFUSA", Lic. Rubén Quispe Condori, [Quispe, 2011]. UNIVERSIDAD MAYOR DE SAN ANDRÉS. Se describen los fundamentos de las Redes Neuronales Artificiales y Lógica Difusa, en cuanto a sus características, modelos y aplicaciones más importantes. A si mismo se realiza estudios sobre las variables que intervienen para el entrenamiento de la red neuronal, para la predicción del consumo de agua.

"RECONOCIMIENTO DE CROMOSOMAS HUMANOS EN IMÁGENES BASADO EN UNA RED NEURONAL DE RETROPROPAGACION", Lic. Rodrigo SimónQuelcaSirpa, [QUELCA, 2011]. UNIVERSIDAD MAYOR DE SAN ANDRES. Este trabajo propone el análisis de imágenes de cromosomas que se encuentra en bandeo G, de forma automática. Luego se realiza la clasificación tomando en cuenta la etapa previa de la segmentación y la detección de bordes de las imágenes cromosómicas, todo esto para obtener una muestra completamente representativa. La etapa de reconocimiento requiere de una imagen de entrada, para procesarlo en la red neuronal previamente entrenada, y esta devuelve resultados de aproximación con respecto a los tipos de cromosomas.

1.3. Planteamiento Del Problema

1.3.1. Problema Central

El creciente consumo de gas natural es uno de los problemas a futuro con las que cuenta nuestro Bolivia, el sector eléctrico muestra el mayor consumo de gas natural dentro del mercado interno, le sigue muy de cerca los sectores Residencial, Comercial, Industrial y de Transporte Vehicular que en conjunto tuvieron un consumo promedio.

Los Consumidores Directos, también tuvieron un consumo promedio, este sector incluye industrias conectadas directamente a la red troncal de transporte, consumidores propios del sector hidrocarburos como refinerías, estaciones de bombeo y Planta Río Grande, además de las Estaciones de Servicio de gas natural que tienen contrato directo con YPFB. Por lo cual crece la importancia de gestionar eficientemente la información del consumo de gas natural, puesto que si no ponemos recaudos, podría afectarnos de manera negativa.

Considerando el problema, se plantea la siguiente pregunta:

¿Cómo se puede determinar el consumo de gas natural en Bolivia?

1.3.2. Problemas Secundarios

- Tratamiento insuficiente de factores, como el crecimiento de la población, consumo de gas natural por parte de la población, hacen que a la hora de realizar cálculos estadísticos el resultado no sea óptimo.
- Tratamiento incorrecto del uso de los datos estadísticos de consumo de gas natural, permiten una mala administración de los volúmenes de gas para el uso de la población.
- Retraso en el envío de gas natural a la población, genera desconfianza en el uso de este energético imprescindible.
- La información provista por métodos estadísticos y matemáticos son limitados ya que no contemplan con una retroalimentación de variables que fluctúan con el tiempo, así que estos métodos que realizan predicciones no son los mejores.
- Todo este factor lleva a que exista desabastecimiento de gas natural hacia la población, llevando al perjuicio a los ciudadanos, empresas y otros.

1.4. Definición De Objetivo

1.4.1. Objetivo General

Plantear un modelo de simulación para el consumo de gas natural en Bolivia.

1.4.2. Objetivos Específicos

 Analizar y considerar las variables que influyen en el consumo de gas natural por parte de la población. Analizar y encontrar datos precisos de consumo de gas natural.

Elaborar un plan para abastecer a la población de acuerdo con los datos fiables

obtenidos.

Elaborar y validar el modelo neuronal de retroalimentación para realizar mejores

predicciones.

Construir un prototipo en base a la red neuronal.

1.5. **Hipótesis**

El modelo de simulación basado en redes neuronales alimentado con variables que afectan

al consumo de gas natural es capaz de predecir mejor el consumo gas natural a futuro, con

una efectividad del0.05 de error.

1.5.1. Operacionalización De Variables

Variable Dependiente: La predicción del consumo de gas natural.

Variable Independiente: El modelo de simulación.

Variable Interviniente: Las redes neuronales artificiales.

Justificación 1.6.

1.6.1. Económica

En la actualidad se busca la utilización eficiente de los recursos económicos con la

aplicación del modelo. Los precios del gas natural en el mercado interno fluctúan entre un

nivel máximo de 1,9768 \$us/Mpie3 y un nivel mínimo de 0,90 \$us/Mpie3, por eso es

imprescindible ver los aspectos financieros, precisamente porque con el modelo basado en

redes neuronales se podrá ver la fluctuación económica de ahorros o ganancias a futuro,

con la predicción del consumo de gas natural.

6

1.6.2. Social

Se plantea un modelo que pretende ser una ayuda importante a la hora de predecir el consumo de gas natural de la población, en función al sector de destino y la normativa empleada en cada caso, siendo el Sector Eléctrico el mayor consumidor.

El gas natural para el Sector Eléctrico en el punto de entrega, al ingreso de la planta termoeléctrica. El gas natural que es entregado a las empresas distribuidoras, el cual va con destino a los Sectores Residencial, Comercial, Industrial y Transporte Vehicular, vemos que estos sectores son los que consumen gran cantidad de gas natural, por eso es necesario ver el consumo de gas natural y tener una idea de cuánto se gasta al día, al mes al año porque si no tomamos recaudos, este elemento podría disminuir y afectar al desarrollo de nuestro país, porque el consumo de gas natural es un elemento de vital importancia en la sociedad.

1.6.3. Científica

Se justifica científicamente, pues en los últimos años se ha visto una creciente utilización de las redes neuronales en diferentes campos y disciplinas, la interacción que ocurre con estas disciplinas aporta nuevos conocimientos e interrogantes para la ciencia, ya que se ingresa en un campo donde las disciplinas necesitan de otras disciplinas, así también de una aplicación de métodos y conocimientos para una mutua cooperación en este caso mediante el modelo de predicción en base a redes neuronales, se busca aportar al conocimiento científico.

1.7. Alcances y Limites

1.7.1. Alcances

1.7.1.1. Alcance Temporal

Se pretende elaborar un modelo neuronal que se ajuste de forma más exacta en el consumo

de gas natural para su posterior uso en la predicción a futuro como ser: en días, meses, años, pero todo basado en sus valores pasados.

1.7.1.2. Alcance Espacial

Pretendemos construir un modelo de simulación para Bolivia, que nos permite ver la magnitud del incremento del consumo de gas natural para una toma de decisiones acertada.

1.7.2. Limites

Los datos pueden ser variados y visto de distintas formas para su uso, como en industrias vehículos, en el sector energético (termoelectricidad) y otros, pero lo esencial, es tener datos que nos ayuden a ver el comportamiento de la información en Bolivia para un mejor uso de ella.

1.8. Aportes

1.8.1. Aporte Práctico

El modelo basado en redes neuronales artificiales una vez implementado será una herramienta alternativa y será de mucha ayuda en la toma de decisiones en distintos sectores como en industrias, en el sector de transporte y también en el sector energético (termoelectricidad).

1.8.2. Aporte Teórico

Se pretende construir una alternativa real y eficiente, frente a los modelos estadísticos o matemáticos existentes, referidos a la teoría de pronóstico y al crecimiento de la población, también ayudar a la toma de decisiones de las autoridades competentes en el consumo de gas natural.

1.9. Metodología

El método de investigación que se adoptara para realizar el presente trabajo será el método científico, por método o proceso científico se entiende aquellas prácticas utilizadas y ratificadas por la comunidad científica como válidas a la hora de proceder con el fin de exponer y confirmar sus teorías. Las teorías científicas, destinadas a explicar de alguna manera los fenómenos que observamos, pueden apoyarse o no en experimentos que certifiquen su validez, involucra una serie de pasos ya que nos ayudara a poner prueba la hipótesis planteada.

Para el método científico, las etapas que presentan y que son necesarias se presentan a continuación: observación, formulación de hipótesis, experimentación y emisión de conclusiones.

Observación, la observación consiste en examinar los hechos y fenómenos que pueden ser percibidos.

Formulación de Hipótesis, consiste en elaborar una explicación provisional de los hechos observados y de sus posibles causas.

Experimentación, consiste en reproducir y observar varias veces el hecho o fenómeno que se quiere estudiar de esta manera pueden estudiar la relación existe entre una magnitud y otra.

Emisión de conclusiones, permite al científico comprobar si su hipótesis era correcta y consiste en la interpretación de los hechos observados de acuerdo con los datos experimentales y dar una explicación científica al hecho o fenómeno visto.

El tipo de investigación a utilizar será el descriptivo, porque puede ofrecer la posibilidad de

de predicciones aunque sean rudimentarias. Muy frecuentemente el propósito del investigador es describir situaciones y eventos. Esto es, decir cómo es y se manifiesta determinado fenómeno. Los estudios descriptivos buscan especificar las propiedades importantes de personas, grupos, comunidades o cualquier otro fenómeno que sea sometido a análisis (Dankhe, 1986). Miden o evalúan diversos aspectos, dimensiones o componentes del fenómeno o fenómenos a investigar. Desde el punto de vista científico, describir es medir. Esto es, en un estudio descriptivo se selecciona una serie de cuestiones y se mide cada una de ellas independientemente, para así describir lo que se investiga.

Los estudios descriptivos buscan especificar las propiedades importantes de personas, grupos, comunidades o cualquier otro fenómeno que sea sometido a análisis, en nuestro caso el estudio del consumo de gas natural.

CAPITULO II

MARCO TEÓRICO

2.1. Inteligencia Artificial

El propósito de la Inteligencia Artificial es hacer computacional el conocimiento humano no analítico porprocedimientos simbólicos, conexionistas o híbridos. Para el conocimiento analítico existen otras ramas de la computación queestudian los métodos y técnicas adecuadas para su representación formal y posterior desarrollo de los programas de ordenadorcorrespondientes. Para conseguir una visión amplia del contenido de Inteligencia Artificial se usan criterios extensionales, esdecir, proporcionando una relación lo más completa posible de los temas que estudia la Inteligencia Artificial, junto con loscriterios intencionales, que establecen las reglas de clasificación, de forma que se pueda concluir si un problema pertenece ono al campo de la Inteligencia Artificial en función del conocimiento necesario para solucionarlo. Básicamente en la Inteligencia Artificialexisten perspectivas muy bien definidas las cuales son dos básicamente: la Inteligencia Artificial como ciencia de lo natural (análisis), la Inteligencia Artificial como ciencia de lo artificial (Ingeniería de síntesis).

En ambas ramas cooperan dos paradigmas, que constituyen dos formas de analizar un proceso y, esencialmente, dosmetodologías de síntesis de una solución: computación simbólica, de grano grueso y programable, computación conexionista, de grano pequeño y autoprogramable por aprendizaje.

La Inteligencia Artificial como ciencia de lo natural es una ciencia de análisis en su mayoría, su objeto formal es el conocimiento, busca un modelo del conocimiento humano organizado en varios niveles (estático, dinámico y estratégico) y susceptible de serusado en predicción.

La Inteligencia Artificial es el conjunto de hechos asociado a los procesoscognoscitivos y a los principios organizacionales y estructurales que dan lugar al comportamiento humano. Por eso su objetoformal coincide en parte con el de la neurología y la ciencia cognoscitiva y su método pretende aproximarse al de la física,integrando teoría y experimento. Como consecuencia de la evaluación del experimento, se reformula el problema o serediseñan mecanismos de inferencia y se extraen conclusiones para nuevo prototipos. La Inteligencia Artificial como ciencia busca una teoría delconocimiento computable con capacidad predictiva análoga a la de una ley física (impersonal, comprobableexperimentalmente y transferible).

La Inteligencia Artificial como ciencia de lo artificial es una ciencia de síntesis que aspira a convertirse en unaingeniería. Busca un sistema que cumpla una serie de especificaciones funcionales. En Inteligencia Artificial se trabaja coninformación y conocimiento. El resultado es un programa de ordenador sobre una máquina específica desarrollado a partir deun modelo del conocimiento que supuestamente usa el operador humano que realiza esa tarea.

Las tareas de la I.A. de síntesis son de alto nivel y se pueden clasificar en tres grupos que en creciente son:dominios formales, dominios técnicos y formas básicas y genuinas del comportamiento humano

Las tareas que usan conocimiento técnico - científico en dominios estrechos, admiten una representación dentro deuna jerarquía de tareas genéricas de análisis, de modificación o de síntesis que son válidas en muchas aplicaciones con sólomodificar la parte del conocimiento. Lo característico de estas tareas es lo limitado del conocimiento que manejan y laposibilidad de formalizar el conocimiento con las técnicas disponibles. Las tareas en dominios estrechos dan lugar a laingeniería del conocimientoque se desconecta de las raíces de la Inteligencia Artificial buscando procedimientos de síntesis con las siguientesfacetas:

Parte de la descripción de tareas a nivel de conocimiento. Para ello es necesariorealizar un proceso de obtención de ese conocimiento a partir del experto humano que lo posee(Allen Newell).

Se busca una representación de ese conocimiento separándolo de los mecanismos de aplicación delmismo de forma que puede acumularse por procedimientos incrementales.

Seleccionar las técnicas adecuadas para su implementación y desarrollar un primer prototipo con laayuda de ciertas herramientas para la adquisición del conocimiento y de entornos de programación.

Se hace énfasis en el carácter de ingeniería buscando procedimientos sistemáticos de implementación, evaluación y refinamiento de los mismos. Estos procedimientos deben ser explícitos, reproducibles yparcialmente independientes del dominio.

Se usan lenguajes y entornos de programación que faciliten el desarrollo rápido y eficiente de aplicaciones.

2.2. Redes Neuronales

2.2.1. Introducción a las Redes Neuronales

El hombre se ha caracterizado por la búsqueda constante de nuevas vías para mejorar sus condiciones de vida, esto le ha servido para reducir el trabajo en aquellas operaciones en las que la fuerza juega un papel primordial.

Los progresos obtenidos han permitido dirigir sus esfuerzos a otros campos, como por ejemplo, a la construcción de máquinas calculadoras que ayuden a resolver de forma automática y rápida determinadas operaciones que resultan tediosas cuando se realizan a mano.

En 1946 fue construida la primera computadora electrónica, ENIAC. Desde entonces los desarrollos en este campo han tenido un auge espectacular. Las computadoras permiten implementar fácilmente algoritmos para resolver multitud de problemas que antes resultaban engorrosos de resolver. Sin embargo, se observa una limitación importante: ¿qué ocurre cuando el problema que se quiere resolver no admite un tratamiento algorítmico, como es el caso, de la clasificación de objetos por rasgos comunes? Este ejemplo demuestra que la construcción de nuevas máquinas versátiles requiere un enfoque del problema desde otro punto de vista. Los desarrollos actuales de los científicos se dirigen al estudio de las capacidades humanas como una fuente de nuevas ideas para el diseño de nuevas y mejores máquinas.

Las redes neuronales son más que otra forma de emular ciertas características propias de los humanos, como la capacidad de memorizar y de asociar hechos. Si se examinan con atención aquellos problemas que no pueden expresarse a través de un algoritmo, se observará que todos ellos tienen una característica en común: la experiencia.

El hombre es capaz de resolver estas situaciones acudiendo a la experiencia acumulada. Así, parece claro que una forma de aproximarse al problema consista en la construcción de sistemas que sean capaces de reproducir esta característica humana.

Las redes neuronales son un modelo artificial y simplificado del cerebro humano, que es el ejemplo perfecto del que se dispone para un sistema que es capaz de adquirir conocimiento a través de la experiencia.

Una red neuronal es "un nuevo sistema para el tratamiento de la información, cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso humano: la neurona". Todos los procesos del cuerpo humano se relacionan en alguna u otra forma con la (in)actividad de estas neuronas. Las mismas son un componente relativamente

simple del ser humano, pero cuando millares de ellas se conectan en forma conjunta se hacen muy poderosas.

2.2.2. Historia de las Redes Neuronales

Un neurofisiólogo y un matemático, respectivamente, fueron los primeros teóricos que concibieron los fundamentos de las redes neuronales, al publicar una teoría acerca de la forma de trabajar de las neuronas (Un Cálculo Lógico de la Inminente Idea de la Actividad Nerviosa - Boletín de Matemática Biofísica 5: 115-133). Ellos realizaron el primer modelo de neurona artificial(McCulloch/Pitts, 1943)

Fue el primero en explicar los procesos del aprendizaje (que es el elemento básico de la inteligencia humana) desde un punto de vista psicológico, desarrollando una regla de como el aprendizaje ocurría. Aun hoy, este es el fundamento de la mayoría de las funciones de aprendizaje que pueden hallarse en una red neuronal. Su idea fue que el aprendizaje ocurría cuando ciertos cambios en una neurona eran activados (Hebb, 1949).

Comenzó el desarrollo del Perceptrón. Esta es la red neuronal más antigua; utilizándose hoy en día para aplicación como reconocedor de patrones. Este modelo era capaz de generalizar, es decir, después de haber aprendido una serie de patrones podía reconocer otros similares, aunque no se le hubiesen presentado anteriormente. Sin embargo, tenía una serie de limitaciones, por ejemplo, su incapacidad para resolver el problema de la función OR-exclusiva y, en general, era incapaz de clasificar clases no separables linealmente (Rosenblatt, 1957).

Desarrollaron el modelo Adaline (ADAptativeLINearElements). Esta fue la primera red neuronal aplicada a un problema real (filtros adaptativos para eliminar ecos en las líneas telefónicas) que se ha utilizado comercialmente durante varias décadas (Widrow/ Hoff, 1960).

Desarrolla la Lernmatrix (memoria asociativa que es capaz de recuperar una imagen o conocimiento, aún cuando el medio este viciado) (Steinbuch, 1961).

En este año surgieron críticas que frenaron, hasta 1982, el crecimiento que estaban experimentando las investigaciones sobre redes neuronales. Minsky y Papert, del Instituto Tecnológico de Massachussets (MIT), publicaron un libro Perceptrons, en el cual probaron (matemáticamente) que el Perceptrón no era capaz de resolver problemas relativamente fáciles, tales como el aprendizaje de una función no-lineal. A pesar del libro, algunos investigadores continuaron su trabajo. Tal fue el caso de James Anderson, que desarrolló un modelo lineal, llamado Lineal Associator (Minsky/Papert, 1969).

Desarrolló la idea básica del algoritmo de aprendizaje de propagación hacia atrás (backpropagation); cuyo significado quedó definitivamente aclarado en 1985 (Werbos, 1974).

Ingeniero electrónico de la Universidad de Helsinki, desarrolló un modelo similar al de Anderson, pero independientemente. (Kohonen, 1977).

Provocó el renacimiento de las redes neuronales con su libro: "Computación neuronal de decisiones en problemas de optimización" (Hopfield, 1982).

Redescubrieron el algoritmo de aprendizaje de propagación hacia atrás (backpropagation). A partir de 1986, el panorama fue alentador con respecto a las investigaciones y el desarrollo de las redes neuronales (Rumelhart/Hinton, 1986).

2.2.3. El Modelo Biológico

Las neuronas y las conexiones entre ellas (sinapsis) constituyen la clave para el procesamiento de la información. De alguna manera la neurona elabora una señal de salida a partir de las dendritas, que son la vía de entrada de las señales que se combinan en el

cuerpo de la neurona, las envía al axón, que es el camino de salida de la señal generada por la neurona.

Las sinapsis, son las unidades funcionales y estructurales elementales que median entre las interacciones de las neuronas. En las terminaciones de las sinapsis se encuentran unas vesículas que contienen unas sustancias químicas llamadas neurotransmisores, que ayudan a la propagación de las señales electroquímicas de una neurona a otra ver figura 2.1.

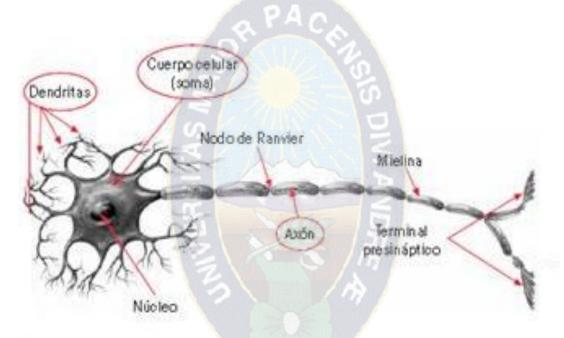


Figura 2.1. Neurona y Conexiones sinápticas **Fuente:**Sanz, 2001

Lo que básicamente ocurre en una neurona biológica es lo siguiente: la neurona es estimulada o excitada a través de sus entradas y cuando se alcanza un cierto umbral, la neurona se dispara o activa, pasando una señal hacia el axón.

Como ya se sabe, el pensamiento tiene lugar en el cerebro, que consta de billones de neuronas interconectadas. Así, el secreto de la "inteligencia" -sin importar como se definase sitúa dentro de estas neuronas interconectadas y de su interacción. La forma que dos

neuronas interactúan no está totalmente conocida. En general, una neurona envía su salida a otras por su axón, el axón lleva la información por medio de diferencias de potencial, u ondas de corriente, que dependen del potencial de la neurona. La neurona recoge las señales por su sinapsis sumando todas las influencias excitadoras e inhibidoras. Si las influencias excitadoras positivas dominan, entonces la neurona da una señal positiva y manda este mensaje a otras neuronas por sus sinapsis de salida.

2.2.4. Elementos de una Red Neuronal Artificial

Las redes neuronales son modelos que intentan reproducir el comportamiento del cerebro humano. Los mismos constan de dispositivos elementales de proceso: las neuronas. A partir de ellas, se pueden generar representaciones específicas, de tal forma que un estado conjunto de ellas puede significar una letra, un número u otro objeto. Generalmente se pueden encontrar tres tipos de neuronas:

Aquellas que reciben estímulos ext<mark>ernos relacionados con el aparato sensorial, que tomarán la información de entrada. Dicha información se transmite a ciertos elementos internos que se ocupan de su procesamiento.</mark>

Es en las sinapsis y neuronas correspondientes a este segundo nivel donde se genera cualquier tipo de representación interna de información. Como no tienen relación directa con la información de entrada ni con la salida, estos elementos se denominan unidades ocultas.

Una vez finalizado el período de procesamiento, la información llega a las unidades de salida, cuya misión es dar la respuesta al sistema.

Como las redes neuronales se asemejan al cerebro tienden a sufrir bajas pero son capaces de reconocer los ruidos o fallas.

A continuación se puede ver el esquema de una red neuronal:

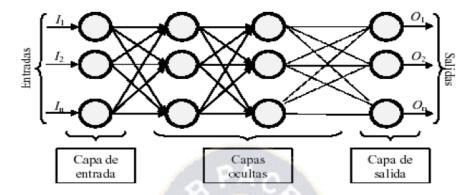


Figura 2.2 Esquema de una Red Neuronal Fuente: Soria, 2010

El esquema está constituido por neuronas interconectadas y arregladas en tres capas (esto último puede variar). Los datos ingresan por medio de la "capa de entrada", pasan a través de la "capa oculta" y salen por la "capa de salida". Cabe mencionar que la capa oculta puede estar constituida por varias capas a su vez.

En la siguiente figura se compara una neurona biológica con una neurona artificial. En la misma se pueden observar las similitudes entre ambas (tienen entradas, utilizan pesos y generan salidas).



Figura 2.3. Similitudes entre una Neurona Biológica y una Artificial **Fuente:** Valencia, 2006

La neurona artificial pretende imitar las características más importantes de las neuronas biológicas. Cada neurona i-ésima está caracterizada en cualquier instante por un valor numérico denominado valor o estado de activación a i (t); asociado a cada unidad, existe

una función de salida, f_i , que transforma el estado actual de activación en una señal de salida. Dicha señal es enviada a través de los canales de comunicación unidireccionales a otras unidades de la red; en estos canales la señal se modifica de acuerdo con la sinapsis (el peso, w_{ji}) asociada a cada uno de ellos según determinada regla. Las señales moduladas que han llegado a la unidad j-ésima se combinan entre ellas, generando así la entrada total.

$$Net_j = \sum_i y_i w_{ji}$$
 (2.1)

Una función de activación, F, determina el nuevo estado de activación a_j (t+1) de la neurona, teniendo en cuenta la entrada total calculada y el anterior estado de activacióna_i(t).

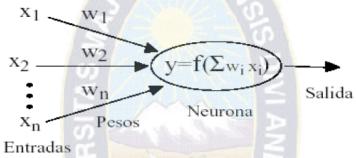


Figura 2.4. Función de activación Fuente: Sánchez, 2006

2.2.5. Backpropagation

En 1986, Rumelhart, Hinton y Williams, formalizaron un método para que una red neuronal aprendiera la asociación que existe entre los patrones de entrada y las clases correspondientes, utilizando varios niveles de neuronas.

El método backpropagation (propagación del error hacia atrás), basado en la generalización de la regla delta, a pesar de sus limitaciones, ha ampliado de forma considerable el rango de aplicaciones de las redes neuronales. El funcionamiento de la red backpropagation (BPN) consiste en el aprendizaje de un conjunto predefinido de pares de entradas-salidas dados como ejemplo: primero se aplica un patrón de entrada como estímulo para la primera capa de las neuronas de la red, se va propagando a través de todas

las capas superiores hasta generar una salida, se compara el resultado en las neuronas de salida con la salida que se desea obtener y se calcula un valor de error para cada neurona de salida. A continuación, estos errores se transmiten hacia atrás, partiendo de la capa de salida hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total. Basándose en el valor del error recibido, se reajustan los pesos de conexión de cada neurona, de manera que en la siguiente vez que se presente el mismo patrón, por lo cual la salida esté más cercana a la deseada.

La importancia de la red backpropagation consiste en su capacidad de autoadaptar los pesos de las neuronas de las capas intermedias para aprender la relación que existe ente un conjunto de patrones de entrada y sus salidas correspondientes. Es importante la capacidad de generalización, facilidad de dar salidas satisfactorias a entradas que el sistema no ha visto nunca en su fase de entrenamiento. La red debe encontrar una representación interna que le permita generar las salidas deseadas cuando se le dan entradas de entrenamiento, y que pueda aplicar, además, a entradas no presentadas durante la etapa de aprendizaje para clasificarlas.

2.2.5.1. Estructura y Aprendizaje

En una red Backpropagation existe una capa de entrada con n neuronas y una capa de salida con m neuronas y al menos una capa oculta de neuronas internas. Cada neurona de una capa (excepto las de entrada) recibe entradas de todas las neuronas de la capa anterior y envía su salida a todas las neuronas de la capa posterior (excepto las de salida). No hay conexiones hacia atrás feedbackni laterales entre las neuronas de la misma capa.

La aplicación del algoritmo tiene dos fases, una hacia delante y otra hacia atrás. Durante la primera fase el patrón de entrada es presentado a la red y propagado a través de las capas hasta llegar a la capa de salida. Obtenidos los valores de salida de la red, se inicia la

segunda fase, comparándose éstos valores con la salida esperada para así obtener el error. Se ajustan los pesos de la última capa proporcionalmente al error. Se pasa a la capa anterior con una retropopagación del error, ajustando los pesos y continuando con este proceso hasta llegar a la primera capa. De esta manera se han modificado los pesos de las conexiones de la red para cada patrón de aprendizaje del problema, del que conocíamos su valor de entrada y la salida deseada que debería generar la red ante dicho patrón.

La técnica Backpropagationrequiere el uso de neuronas cuya función de activación sea continua, y por lo tanto, diferenciable. Generalmente, la función utilizada será del tipo sigmoidal.

2.2.5.2. Pasos Para Aplicar el Algoritmo de Entrenamiento

Inicializar los pesos de la red con valores pequeños aleatorios.

Presentar un patrón de entrada y especificar la salida deseada que debe generar la red.

Calcular la salida actual de la red. Para ello se presentan las entradas a la red y se calcula la salida de cada capa hasta llegar a la capa de salida, ésta será la salida de la red. Los pasos son los siguientes:

Se calculan las entradas netas para las neuronas ocultas procedentes de las neuronas de entrada. Para una neurona j oculta:

$$net_{pj}^{h} = \sum_{i=1}^{N} w_{ji}^{h} x_{pi} + \Theta_{j}^{h}$$
 (2.2)

En donde el índice h se refiere a magnitudes de la capa oculta; el subíndice p, al p-ésimovector de entrenamiento, y j a la j-ésima neurona oculta. El término θpuede ser opcional, pues actúa como una entrada más.

Se calculan las salidas de las neuronas ocultas:

$$y_{pj} = f_j^h = \left(net_{pj}^h\right) \qquad (2.3)$$

Se realizan los mismos cálculos para obtener las salidas de las neuronas de salida:

$$net_{pk}^{o} = \sum_{j=1}^{L} w_{kj}^{o} y_{pj} + \Theta_{k}^{o}$$
 (2.4)

$$y_{pk} = f_k^o = \left(net_{pk}^o\right) \tag{2.5}$$

Calculamos los términos de error para todas las neuronas.

Si la neurona k es una neurona de la capa de salida, el valor es:

$$\varphi_{pk}^o = (d_{pk} - y_{pk}) f_k^o (net_{pk}^o)$$
 (2.6)

La función f debe ser derivable. En general disponemos de dos formas de función de salida:

La función lineal:

$$f_k(\operatorname{net}_{jk}) = \operatorname{net}_{jk}$$
 (2.7)

La función sigmoidal:

$$f_k(net_{jk}) = \frac{1}{1 + e^{-net_{jk}}}$$
 (2.8)

La selección de la función depende de la forma que se decida representar la salida: si se desea que las neuronas de salida sean binarias, se utiliza la función sigmoidal, en otros casos, la lineal.

Para una función lineal, tenemos: $f_k^o = 1$, mientras que la derivada de una función sigmoidal es:

$$f_k^{o} = f_k^o (1 - f_k^o) = y_{pk} (1 - y_{pk})$$
 (2.9)

Por lo que los términos de error para las neuronas de salida quedan:

$$\varphi_{pk}^o = (d_{pk} - y_{pk})$$
, para la salida lineal.

$$\varphi_{pk}^o = (d_{pk} - y_{pk})y_{pk}(1 - y_{pk})$$
, para la salida sigmoidal.

Si la neurona j no es de salida, entonces la derivada parcial del error no puede ser evaluada directamente, por tanto se obtiene el desarrollo a partir de valores que son conocidos y otros que pueden ser evaluados.

La expresión obtenida en este caso es:

$$\varphi_{pj}^{h} = f_{j}^{h} \left(net_{pj}^{h} \right) \sum_{k} \varphi_{pk}^{o} w_{kj}^{o}$$
(2.10)

Donde observamos que el error en las capas ocultas depende de todos los términos de error de la capa de salida. De aquí surge el término propagación hacia atrás.

Actualización de los pesos: para ello utilizamos un algoritmo recursivo, comenzando por las neuronas de salida y trabajando hacia atrás hasta llegar a la capa de entrada, ajustando los pesos de la siguiente forma:

Para los pesos de las neuronas de la capa de salida:

$$w_{kj}^{o}(t+1) = w_{kj}^{o}(t) + \Delta w_{kj}^{o}(t+1)$$

$$\Delta w_{kj}^{o}(t+1) = \alpha \varphi_{pk}^{o} y_{pj}$$
(2.11)

Para los pesos de las neuronas de la capa oculta:

$$w_{ji}^{h}(t+1) = w_{ji}^{h}(t) + \Delta w_{ji}^{h}(t+1) \qquad (2.12)$$
$$\Delta w_{ji}^{h}(t+1) = \alpha \varphi_{pj}^{o} x_{pi}$$

En ambos casos, para acelerar el proceso de aprendizaje se puede añadir un término momento.

El proceso se repite hasta que el término de error:

$$E_p = \frac{1}{2} \sum_{K=1}^{M} \varphi_{pk}^2$$
 (2.13)

Resulta aceptablemente pequeño para cada uno de los patrones aprendidos.

2.2.5.3. Número de capas ocultas

No se pueden dar reglas concretas para determinar el número de neuronas o número de capas de una red para resolver un problema concreto.

Respecto al número de capas de la red, en general tres capas son suficientes (entrada - oculta-salida). Sin embargo, hay veces que un problema es fácil de resolver con más de una capa oculta. El tamaño de las capas, tanto de entrada como de salida, suele estar determinado por la naturaleza de la aplicación. En cambio, decidir cuántas neuronas debe tener una capa oculta no suele ser tan evidente. El número de neuronas ocultas interviene en la eficiencia de aprendizaje y de generalización de la red. No hay ninguna regla que indique el número óptimo, en cada problema se debe ensayar.

2.2.6. Aplicación de las redes neuronales artificiales

Las Redes neuronales artificiales son interesantes para una gran cantidad de personas de diferentes áreas:

Científicos de la computación que quieren averiguar sobre las propiedades del procesamiento de información no-simbólica con una red neuronal y sobre sistemas de aprendizaje en general.

Estadísticos utilizan redes neuronales como modelos de clasificación y regresión no lineal flexible, ingenieros de varias especialidades se aprovechan de las capacidades de redes neuronales en muchas áreas, como procesamiento de señales y control automático

Los científicos cognoscitivos ven redes nerviosas como un posible aparato para describir modelos de pensamiento y conciencia (función cerebral de alto-nivel). Neuro – fisiólogos utilizan redes neuronales para describir y explorar función cerebral de medio-nivel (memoria del eje, sistema sensorial y motriz).

Físicos usan redes neuronales para modelar fenómenos en mecánica estadística y para muchas otras tareas.

Biólogos utilizan Redes Neuronales para interpretar sucesiones del nucleotide.

Filósofos y otras personas también pueden ser interesados en Redes Nerviosas por las varias razones.

Las Redes neuronales Artificiales con su capacidad para deducir un significado desde datos complicados o imprecisos, pueden ser utilizadas para detectar patrones o detectar una tendencia que es muy compleja de hallar por una persona con modelos determínisticos o por otras técnicas computacionales tradicionales. Una red entrenada puede ser vista como un experto en el manejo de la información que se le ha dado para analizar. Este experto puede ser utilizado para proporcionar proyecciones ante nuevas situaciones de interés.

2.3. Metodologías para Predicción

2.3.1. Metodología para la Predicción con Redes Neuronales Artificiales

Hace referencia al conjunto de procedimientos racionales utilizados para alcanzar una gama de objetivos que rigen en una investigación científica, una exposición doctrinal o tareas que requieran habilidades, conocimientos o cuidados específicos.

Alternativamente puede definirse la metodología como el estudio o elección de un método pertinente para un determinado objetivo.

Según los autores Palit y Popovic, en un estudio realizado el año 2005, la metodología para la predicción es generalmente entendida como un conjunto de enfoques, métodos y herramientas para la recopilación de datos a ser utilizados para la previsión o predicción de los futuros valores de estas series de datos, incluyendo las siguientes etapas como vemos a continuación:

2.3.1.1. Preparación de los Datos

En la preparación de datos se presentan cuatro etapas las cuales son:

a) Adquisición de Datos

La adquisición de datos consiste en la toma de muestras que nos servirán como entradas y salidas para el modelo de red neuronal.

b) Estructuración de Datos

Se debe obtener el par (entrada, salida deseada), es decir se obtendrá la siguiente ecuación:

$$(x(t-(D-1)\Delta), \dots x(t-\Delta), x(t), x(t+R))$$

Donde x(t+R) son los valores predichos basados en los valores pasados de una serie de tiempo hasta un punto t siendo R el instante para el cual se quiere predecir, D será numero de patrones de la muestra, Δ será el intervalo entre los valores de la serie de tiempo y x(t) será el valor de la serie de tiempo en un instante t.

c) Normalización de Datos

En los problemas de integración de información heterogénea en los que además subyace el objetivo de la toma de decisiones, como es el caso del problema de la localización, el de evaluación de impacto ambiental, la selección de contratistas o proveedores, etc., es indispensable disponer de una herramienta que permita expresar los valores de los diferentes criterios en unidades homogéneas con el objetivo de convertirlos en valores comparables y operables aritméticamente entre sí. Los principales motivos que confirman esta necesidad son los siguientes:

La técnica de normalización es quitar las inconsistencias y el ruido de los datos y dar a conocer las relaciones importantes de las variables de entrada y salida por que los datos generalmente son heterogéneos, es necesaria porque en muchas oportunidades los resultados de la predicción dependen de la normalización de datos debido a la discrepancia que existen en los datos originales como ser valores muy grandes o demasiado pequeños.

En general se entiende que la normalización es la operación mediante la cual un conjunto de valores de una determinada magnitud son transformados en otros, de tal suerte que estos últimos pertenezcan a una escala predeterminada, la normalización puede realizarse de los siguientes modos: sin cambio de magnitud, como sería el caso de transformar la mediciones de temperatura realizadas por diversos termómetros en grados Celsius y otros, con cambio de magnitud a escala libre, como sería el caso de transformar los datos de producción de trigo, leche y gas natural, expresados en toneladas, hectólitros y hectómetros cúbicos respectivamente, a una escala de valores de mercado expresados en euros, con cambio de magnitud a escala fija, como sería el caso de transformar los datos de consumo de energía eléctrica, consumo de combustibles fósiles y consumo telefónico, a una escala adimensional acotada entre los valores [0,10].

El procedimiento es el siguiente:

$$v = \frac{b - bmin}{bmax - bmin}$$

Donde:

v: dato normalizado, b: dato a normalizar, bmin: dato más pequeño de la serie,

bmax: dato más grande de la serie

Se adapta a la concentración media de los valores en general, el procedimiento consiste simplemente en modificar la escala en este caso en un rango de [0, 1].

d) División de Datos

En la práctica es común dividir la serie de tiempo en tres distintos conjuntos los cuales son: el conjunto de entrenamiento, de prueba y validación. El conjunto de entrenamiento es la más grande y se usa para que la red neuronal pueda aprender el modelo que presentan los datos. El conjunto de prueba oscila entre un 10% al 30% del conjunto de entrenamiento se usa para evaluar la habilidad de la generalización de una red entrenada. El conjunto de

validación debe consistir en las más recientes observaciones contiguas, se debe tener el cuidado para no usar el conjunto de validación como conjunto de prueba.

2.3.1.2. Determinación de la Arquitectura de Red

Es decir, la selección del tipo de red que se utilizará para la predicción, la determinación del número de nodos de entrada y salida de la red, número de capasocualtas, el número de neuronas dentro de las capas, determinación de las interconexiones entre las neuronas, la selección de la función de activación de las neuronas, entre otros.

En un modelo neuronal estándar para realizar la predicción de una serie de tiempo se debe usar un juego de d-tuplas como entradas y un solo valor designado a la salida de la red. La estructura a nivel macro del modelo predictivo es ilustrativo en la figura 2.6. En el cual se observa un conjunto de d-tuplas como variables predictivas [x(t-d),...x(t-1),x(t)], y la única salida [x(t+1)], como variable predicha [Larrañaga et al, 2002].

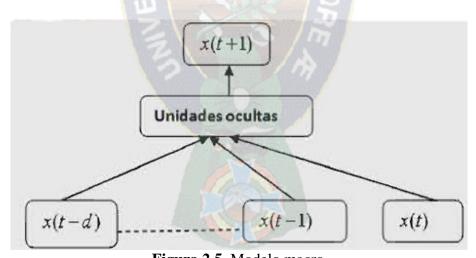


Figura 2.5. Modelo macro **Fuente:** Larrañaga et al, 2002

2.3.1.3. Diseño de Estrategia de Aprendizaje o Entrenamiento de la Red

Es decir, la selección del algoritmo de aprendizaje, el índice de rendimiento, la capacitación y la supervisión. El entrenamiento de la red neuronal es un proceso en el que la red aprende

a reconocer el modelo inherente a los datos de entrenamiento. El objetivo del entrenamiento es encontrar el conjunto conveniente para los valores de los pesos en la conexión de los nodos de la red neuronal a través de la afinación de los mismos durante el entrenamiento de la red, haciendo que la red pueda lograr el atributo de la generalización más alto.

2.3.1.4. Evaluación Global

De los resultados de la predicción mediante la utilización de nuevos conjuntos de datos de observación.

El entrenamiento de la red puede requerir de repetidos ensayos de entrenamiento con varios valores de pesos iníciales. Después de cada corrida de entrenamiento los resultados de la misma tienen que ser evaluados y comparados con los resultados logrados en las corridas anteriores, en orden para seleccionar la mejor corrida.

2.3.2. Método de Medias Móviles como Técnica Estadística de Predicción

El método de medias móviles es una técnica estadística de análisis utilizada para estimar el valor futuro de una serie de datos, denominada Medias Móviles.

Su metodología es simple, consiste en utilizar promedios de las observaciones pasadas más recientes de la serie analizada. Para cada período, el cálculo del promedio se renueva añadiendo el más reciente y eliminando el más antiguo de los considerados. La "amplitud" de la media móvil (el número de observaciones incluidas en el cálculo promedio), se denomina "orden" de la media móvil: media móvil de orden "n":

$$\frac{y_1+y_2+\dots+y_n}{n}, \frac{y_2+y_3+\dots+y_{n+1}}{n}, \frac{y_3+y_4+\dots+y_{n+2}}{n}$$

La media móvil simple es la media aritmética de los n datos anteriores mientras más grande sea n, mayor será la influencia de los datos antiguos.

La media móvil central en lugar de utilizar solo datos pasados, se utilizan también datos posteriores a aquel se quiere obtener la media.

La media móvil ponderada es una media multiplicada por ciertos factores que le dan determinado peso a determinados datos. Luego vamos al siguiente paso que es el componente estacional:

$$Componente estacional = \frac{Consumo mensual}{Mediamovil central}$$

Luego realizamos el cálculo del índice estacional, para esto agregamos una columna para poner el periodo, luego realizamos el cálculo de la serie desestacionalizada:

Luego realizamos una regresión lineal de la forma:

$$y = a + bx$$

Para obtener los datos pronosticados, evaluamos los datos de la siguiente forma:

$$y = (y *) * indiceestacional$$

Comparamos los datos de las medias móviles ponderadas frente a los datos reales y finalmente calculamos el error absoluto.

Ventajas de esta técnica de predicción:

Las medias móviles resultan más apropiadas cuando la aleatoriedad de los datos es elevado y la autocorrelación baja. La media móvil solo tiene de memoria de un periodo, solo debe ser utilizada con fines predictivas a coroto plazo.

Presenta una progresión tanto más suavizada cuanto mayor sea el número de términos incluidos en el promedio.

Desventajas de las medias móviles:

La presencia de tendencia marcada o estacionalidad hace muy arriesgado utilizar las medias móviles para la estimación.

Para evitar la estacionalidad, pueden elaborarse medias móviles de orden igual al orden estacional (técnica de alisado), pero esto, válido a efectos de ajustes promedio general, implicará varios sesgos en materia de predicción. Una media móvil de este tipo "filtra" la estacionalidad y, por tanto, después habrá de nuevo de añadirse a la estimación realizada sobre la serie filtrada.

2.4. Gas Natural

El gas natural se define la mezcla de hidrocarburos livianos en estado gaseoso, donde la mayor proporción corresponde al metano (CH₄) en un valor que oscila entre el 80% al 90%. El porcentaje restante está constituido por etano (C₂H₆), propano, butano y superiores, pudiendo contener asimismo en proporciones mínimas, vapor de agua, anhídrido carbónico, nitrógeno, hidrogeno sulfato, etc.

El gas natural proviene de yacimientos subterráneos que pueden ser de gas o de petróleo y gas, según que en su origen se encuentre asociado o no al petróleo. El gas natural procede generalmente de las perforaciones que se realizan en los yacimientos petrolíferos, de la descomposición de la materia orgánica con el tiempo, el petróleo es más liviano que el agua, puede flotar sobre los lagos subterráneos de agua salada.

En la parte superior se encuentra el gas, que ejerce enormes presiones, con lo cual hace un fluir el petróleo hacia la superficie, se formo hace millones de años, cuando plantas y pequeños animales marinos fueron enterrados por arena y piedra, Capas de lodo, arena, piedra, plantas y materia animal acumulada hasta que la presión y el calor de la tierra, los convirtió en petróleo y gas natural.

2.4.1. Características del Gas Natural

Una de las principales características del gas natural es la economía, ya que al no requerir de procesos industriales en su elaboración, su costo es más bajo que el de otros combustibles. Otro elemento que incide en su bajo costo es el hecho de ser un recurso abundante, distribuido a lo largo de los cinco continentes, lo que permite una transacción libre en los cinco continentes, lo que permite una transacción libre en los mercados. La ventaja competitiva del precio ha llevado a las industrias a incorporar el gas natural, con el fin de hacer más rentables sus procesos de fabricación.

De igual forma las generadoras de electricidad son parte importante de los consumidores de gas natural con sus plantas de generación directa y de ciclo combinado.

El gas natural también tiene una gran aceptación en los hogares, ya que adicionalmente al ahorro que obtienen por el uso de este combustible, la seguridad y el suministro continuo del mismo, permite a los usuarios acceder a una calidad de vida superior. Por otra parte el gas natural es el único combustible más liviano que el aire, por lo que se esparce fácilmente y al no ser toxico, por no tener monóxido de carbono se constituye como un combustible más seguro.

2.4.2. Consumo de Gas Natural

El gas natural es una mezcla de gases compuesta principalmente por metano. Se trata de un gas combustible que proviene de formaciones geológicas, por lo que constituye una fuente de energía no renovable. También de metano, el gas natural puede contener dióxido de carbono, etano, propano, butano y nitrógeno, entre otros gases. Estos componentes hacen que el uso del gas natural sea contaminante. Además de su presencia en yacimientos fósiles, el gas natural puede obtenerse a partir de la descomposición de los restos orgánicos. Este

proceso es promovido en plantas de tratamiento especializadas que producen el denominado biogás.

Cuando las reservas de gas se encuentran en lugares apartados donde no resulta rentable la construcción de gasoductos para llevar el gas a los hogares e industrias, es posible procesar el gas natural para convertirlo en gas natural licuado (GNL). Así, en forma líquida, se facilita su transporte. El GNL suele trasladarse a -161°C, ya que la licuefacción puede reducir el volumen de gas hasta en 600 veces.

El gas natural almacenado a altas presiones (entre 200 y 250 bar), se transforma en gas natural comprimido (GNC), un combustible que se utiliza en vehículos ya que resulta económico en comparación a la gasolina. Otro uso del gas natural aparece con la producción de hidrógeno, que también supone un combustible alternativo para los vehículos. En estos casos, el hidrógeno puede utilizarse a través de la combustión (con un motor de explosión) o mediante una pila de combustible (el hidrógeno se convierte en electricidad y alimenta un motor eléctrico).

El uso de gas natural, como el que conlleva la utilización de toda sustancia tóxica, tiene sus riesgos; es un peligro que, a diferencia de otros, no se ve y apenas puede percibirse a través de los sentidos, por tanto es sumamente importante que se tengan ciertas precauciones para evitar catástrofes que, podrían traer como consecuencia la muerte.

El gas natural es transportado hasta los centros de consumo. Al entrar a la ciudad el gas natural es odorizado para posteriormente ser distribuido mediante redes a los domicilios, estaciones de servicio (GNV), comercios e industrias.

La exportación de gas natural se realiza por gasoductos de mayor capacidad para poder cumplir con los volúmenes de los contratos.

2.4.3. Producción de Hidrocarburos

Durante la gestión 2012, la producción bruta de gas natural, alcanzó un promedio de 51,11 MMm3/día. Asimismo, la producción bruta de petróleo, condensado y gasolina natural obtuvo un promedio de 51,49 MBbl/día.

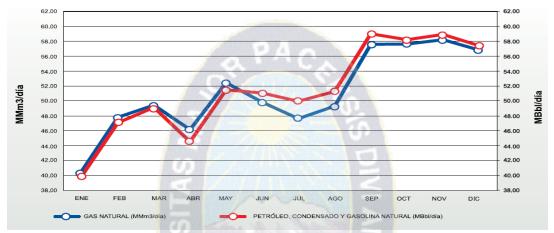


Figura 2.6. Evolución de la producción de hidrocarburos 2012

Fuente: YPFB, 2012

La producción bruta de gas natural en la gestión 2012 en promedio, se incremento en 13,41% en relación a las gestión 2011. A partir del mes de septiembre la producción de gas Natural alcanzo los volúmenes más altos, registrándose un volumen promedio máximo de 58,19 MMm3/día en el mes de noviembre.

2.4.3.1. Producción de Gas Natural por Campo

La producción sujeta al pago de IDH, regalías y participaciones es aquella producción medida en punto de fiscalización la salida de la planta más los volúmenes de producción sujetos a penalidades por quemas. A su vez, es menor a la producción bruta contemplada en el **Cuadro Nº 1** debido a que esta última es medida en boca de pozo antes del proceso de separación de los componentes licuables y se realicen las actividades de uso del gas como combustible en los campos, quema y venteo.

La producción de gas natural sujeta al pago de IDH, regalías y participaciones, durante la gestión 2012, alcanzó un promedio de 49,46 MMm3/día, mostrando un incremento de 13,77% en relación al 2011. Esta producción fue entregada en su totalidad a YPFB por las diferentes empresas que operan los campos bajo contratos de operación.



Figura 2.7. Campos de Exploración

Fuente: YPFB, 2012

Los campos con mayor producción fueron Sábalo y San Alberto que durante la gestión 2012 representaron el 31,9% y 21,2% respectivamente, de la producción total. Así mismo, el campo Margarita - Huacaya represento el 11,5% de la producción en la gestión 2012. Otros campos que tuvieron una producción significativa son Bulo Bulo y Yapacaní, que representó el 4,2% y 3,5% respectivamente, Vuelta Grande, Itaú, Río Grande, Tacobo y Margarita representaron el 3,1%, 3,1%, 3,0%, 2,7% y 1,6% del total de la producción de gas natural del 2012. El resto de los campos representaron el 14,2% del total producido, que

incluye los campos con una producción menor a 0,94 MMm3/día para el 2011 y menor a 1,44 MMm3/día para el año 2012.

La producción de hidrocarburos líquidos sigue la misma tendencia del gas natural alcanzando su máxima producción en el mes de septiembre con 59 MBbl/día. El valor promedio de esta se incrementó en 15,46% en relación a la gestión 2011 también podemos observar los datos de las instalaciones de gas natural por departamento en la tabla 2.1.

Tabla: 2.1. Instalaciones de redes de gas domiciliario

REDES DE GAS DOMICILIARIO								
DEPARTAMENTOS	2006	2007	2008	2009	2010	2011	2012	TOTAL
LA PAZ	2.205	10.983	19.881	37.816	47.541	31.057	21.804	171.287
ORURO	1.099	917	1.727	5.981	6.188	5.049	3.535	24.496
COCHABAMBA		1	7.1	750	4.335	14.020	9.256	28.361
CHUQUISACA	79 <mark>7</mark>	1.718	11.393	2.957	3.372	4.227	2.259	26.723
SANTA CRUZ	396	968	9.581	2.314	5.715	9.810	6.129	34.913
POTOSÍ	1.520	964	1.427	2.922	3.031	3.253	2.473	15.590
		II B						301.370

Fuente: Mendizábal, 2012

2.4.4. Usos Principales e Importancia

Una de principales aplicaciones del gas natural ha sido en la producción de vapor, sustituyendo o complementando, en instalaciones mixtas, la acción de los combustibles sólidos y líquidos. En cuando a las aplicaciones de la industria.

Cerámica: debido al menor contenido de contaminantes y el poder calorífico de los combustibles gaseosos, es posible efectuar el calentamiento directo del producto, lo que permite obtener un grado de combustión elevado y construir hornos más pequeños.

Industria de vidrio: las operaciones térmicas de la industria del vidrio se clasifican en dos grupos de naturaleza esencialmente distinta, uno de ellos es la utilización del gas.

Industria textil: se utiliza para el acabado de las fibras, este proceso requiere mantener una presión constante del gas natural.

Industria química: se considera como una de las materias primas básicas para las síntesis químicas industriales más importantes.

Industria del cemento: consume una considerable cantidad de energía térmica, representando el combustible entre un 25 y un 40% del costo total del producto.

2.4.4.1. Precios de Gas Natural por Sector y Empresa Demandante

Los precios de gas natural en el mercado interno fluctúan entre un nivel máximo de 1,9768 \$us/Mpie3 y un nivel mínimo de 0,57 \$us/Mpie3, en función al sector de destino y la normativa empleada en cada caso, siendo el Sector Eléctrico en el que se aplican los mayores precios y el Sector Residencial, Comercial, Industrial y Transporte Vehicular en el que se aplican los menores precios.

El precio de gas natural para el Sector Eléctrico en el punto de entrega, al ingreso de la planta termoeléctrica, es igual a 1,30 \$us/Mpie3. Para generadoras pertenecientes al Sistema Interconectado Nacional y para generadoras pertenecientes al Sistema Aislado el precio varía entre 1,02 \$us/Mpie3 y 1,9768 \$us/Mpie3.

El precio de gas natural que es entregado a las empresas distribuidoras en Puerta de Ciudad (citygate), cuyo destino son los Sectores Residencial, Comercial, Industrial y Transporte Vehicular, es de 0,98 \$us/Mpie3. Sólo en el caso de la Empresa Tarijeña del Gas S.A.M. – Carapari, es de 0,57 \$us/Mpie3. Asimismo, los precios de gas natural para consumo propio del sector hidrocarburos como combustible para el transporte varían entre 1,0108 \$us/Mpie3 y 1,4208 \$us/Mpie3 y para el uso como combustible en refinación es de 1,31 \$us/Mpie3.

A continuación observamos los precios finales de gas naturalpor sectory empresa demandante, el sector eléctrico es un importante consumidor de gas natural, también las residenciales comerciales industrias y transporte vehicular.

Tabla 2.2.PRECIOS FINALES DE GAS NATURAL POR SECTOR Y EMPRESA DEMANDANTE

	SECTOR					
		Empresa Eléctrica Guaracachi S.A.	1,3000			
	SISTEMA INTERCONECTADO	Empresa Eléctrica Valle Hermoso S.A.	1,3000			
	NACIONAL - SIN R.A. SSDH	Compañía Boliviana de Energía Eléctrica S.A.	1,3000			
	N° 0440/2008	Compañía Eléctrica Central Bulo Bulo S.A.	1,3000			
		ENDE Andina S.A.M	1,3000			
	2	Emp <mark>resa Servici</mark> os Eléctricos Tarija S.A.	1,0500			
ELÉCTRICO	S	Empresa Servicios Eléctricos Tarija S.A Bermejo	1,1000			
	₹ //	Cooperativa Rural de Eléctrificación Ltda.	1,1100			
	SISTEMAS AISLADOS R.A. SSDH N° 0450/2008	Cooperativa Rural de Eléctrificación Lt <mark>d</mark> a Ipias	1,9768			
		Cooperativa Rural de Eléctrificación Ltda Puerto Suárez	1,5485			
		Gas y Electricidad S.A.	1,0200			
		Cooperativa Monteagudo	1,1000			
	5	Cooperativa Muyupampa	1,1000			
		YPFB Redes de Gas Cochabamba	0,9800			
		YPFB Redes de Gas Santa Cruz	0,9800			
		YPFB Redes de Gas Chuquisaca	0,9800			
	ERCIAL, INDUSTRIAL Y TRANS- ITE VEHICULAR	YPFB Redes de Gas La Paz	0,9800			
	DE GAS NATURAL)	YPFB Redes de Gas Oruro	0,9800			
		YPFB Redes de Gas Potosi	0,9800			
		Empresa Tarijeña del Gas S.A.M.	0,9800			
		Empresa Tarijeña del Gas S.A.M Carapari	0,5700			

Fuente: YPFB, 2012

Finalmente tenemos los precios de los consumidores directos y otros, en donde tenemos el

uso de combustible para refinación, también la industria minera, la industria alimenticia y el uso de combustible para el transporte.

Tabla 2.3. PRECIOS FINALES DE GAS NATURAL POR SECTOR Y EMPRESA DEMANDANTE

		Refinería Oro Negro S.A.	1,3100
	USO COMBUSTIBLE PARA REFINACION	Refinería Santa Cruz S.R.L.	1,3100
	R.A. SSDH N° 0452/2008	Refinería Parapetí S.R.L.	1,3100
	N 04527 2008	YPFB Refinación S.A.	1,3100
	INDUSTRIA MINERA – CALERA	Compañía Minera Paitití S.A.*	
CONSUMIDORES	R.A. SSDH N° 0452/2008	Empresa Minera y Calera Sayari S.A.	1,7000
DIRECTOS Y OTROS	INDUSTRIA ALIMENTICIA R.A. SSDH N° 0452/2008	Gravetal Bolivia S.A.*	
***		Empresa COMASA	1,7000
		Laguna Volcán S.R.L.	1,7000
	USO COMBUSTIBLE PARA TRANSPORTE R.A. SSDH	Planta de Compresión Río Grande	1,4208
		YPFB Logística	1,4208
	N° 0695/2008	YPFB Transporte	1,0108
	GNV D.S. N° 29510	Caiguami**	1,5000

Fuente: YPFB, 2012

2.4.4.2. Ventajas de Seguridad y Ambientales sobre Otros Combustibles

El gas natural tiene rango inflamabilidad muy ilimitado, en concentraciones en el aire, por debajo del 4% y por arriba de aproximadamente el 14% no se encenderá el gas natural. Además la temperatura de ignición alta y el rango de inflamabilidad limitado reducen la posibilidad de un incendio o expulsión accidental.

Los vehículos a combustible líquido emiten compuestos orgánicos volátiles en sus escapes y también emisiones de vapores tanto en el momento de la carga del tanque del vehículo como durante la marcha.

Estas emisiones contribuyen a un elevado porcentaje de la contaminación del aire y reaccionan en la atmósfera en presencia de la luz solar para producir ozono a nivel de suelo

en cambio el gas natural comprimido no contribuye a la contaminación.

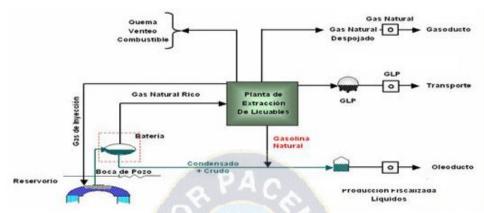


Figura 2.8. Estructura de la comercialización de hidrocarburos **Fuente:** YPFB, 2012

Las ventajas son que: es un combustible relativamente barato, presenta una combustión completa y limpia, la cual prácticamente no emite bióxido de azufre, seguridad en la operación, debido a que en caso de figuras, al ser más ligero que el aire, se disipa rápidamente de la atmosfera. Únicamente se requiere buena ventilación, promueve una mayor eficiencia térmica en plantas de ciclo combinación para generación de electricidad.

CAPITULO III

MARCO APLICATIVO

3.1. Introducción

En este capítulo seguiremos la metodología propuesta para la Predicción del consumo de gas natural aplicando redes neuronales, entendida como un conjunto de enfoques, métodos y herramientas a ser utilizados para la predicción, desde la selección de variables que intervendrán en el aprendizaje de la red, el diseño del modelo de red neuronal predictiva, hasta la preparación de los datos de entrenamiento.

3.2. Metodología para Redes Neuronales Predictivas

Empezaremos a desarrollar el conjunto de enfoques vistos en el capitulo anterior para el consumo de gas natural.

3.3. Preparación de los Datos

3.3.1. Adquisición de Datos

En la adquisición de datos se tuvo que identificar a las variables de entrada y salida, como un detalle adicional tenemos a la población porque ellos son los primeros consumidores del gas natural.

En la figura 3.1 observamos el censo de población y vivienda 2012, son los datos oficiales del total de habitantes en nuestro país, además observamos el porcentaje de crecimiento de cada departamento de la gestión 2001 hasta el último censo del 2012 el porcentaje de crecimiento es alto.

Por otro lado vemos que el incremento poblacional femenino es más, con un 50.1%

que el masculino que es de 49.9%, a continuación observamos los datos Nacionales de la Población del CENSO 2012

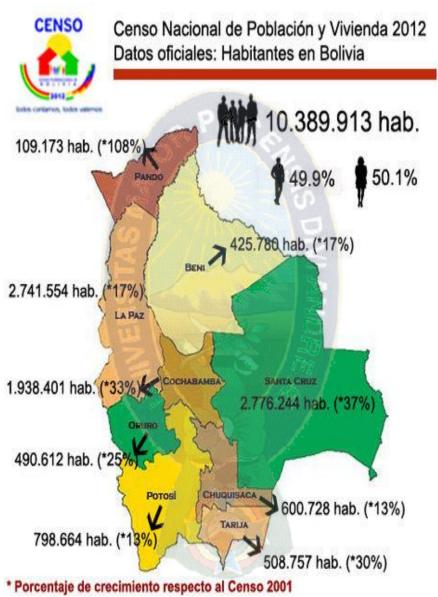


Figura: 3.1. Censo Nacional de población y vivienda **Fuente:** INE, 2012

Las variables entrada son dos el primero es: la producción de energía eléctrica (termoelectricidad) los datos fueron recopilados del Ministerio de Hidrocarburos y Energía,

ver tabla 3.1 por que la producción de energía eléctrica utiliza casi el 50% del consumo de gas natural ver tabla 3.2.

Tabla 3.1. Centrales de Generación existentes Termoeléctricas

AGENTE	CENTRAL	CAPACIDAD EFECTIVA (MW)
GUARACACHI	GUARACACHI	356,0
	KARACHIPAMPA	13,4
	ARANJUEZ	34,7
	TOTAL	404,1
BULO-BULO	BULO-BULO	82,6
ENDE	ENTRE RIOS	97,1
V. HERMOSO	CARRASCO	99,6
	VALLE HERMOSO	67,1
	TOTAL	166,7
COBEE	KENKO	17,5
GUABIRA	GUABIRA	20,7
ENDE	MOXOS	5,0
TOTAL TERMOELÉCTRICAS		793,4

Fuente: Gomes, 2010.

Las unidades en que se medirá la energía eléctrica (termoelectricidad) será de acuerdo al requerimiento de gas que necesiten en este caso será en MMm3/día.

Tabla: 3.2. El sector eléctrico en Bolivia (Termoelectricidad)

SECTOR	ELECTRICO			
150	2011	2012		
ENE	4,05	4,32		
FEB	3,31	3,76		
MAR	2,89	3,49		
ABR	3,53	3,28		
MAY	4,32	4,13		
JUN	4,59	4,17		
JUL	4,67	4,18		
AGO	4,86	4,59		
SEP	4,85	4,87		
OCT	4,55	4,88		
NOV	4,63	4,6		
DIC	4,41	3,93		
PROM	4,22	4,18		

Fuente: Aliaga, 2012

Y finalmente las regalías por IDH ver tabla 3.3, y tabla 3.4 los datos fueron recopilados del Ministerio de Hidrocarburos y Energía en el año 2012.

Tabla3.3. Regalías por departamento (MM\$us) 2012

MES	Regalía Departamental del 11%		galía Dptal.		galia Dptal.	Regalia	Nacional toria del 1%	Total Regalia Compensa-	TOTAL
	CBBA	CHUQ	STA. CRUZ	TJA (**)	del 11%	BENI	PANDO	toria del 1%	
ENE (*)	3,38	1,93	12,59	44,91	62,81	3,81	1,90	5,71	68,52
FEB (*)	3,11	1,80	9,35	31,28	45,54	2,76	1,38	4,14	49,68
MAR (*)	3,11	1,80	9,35	31,28	45,54	2,76	1,38	4,14	49,68
ABR	3,79	2,09	11,53	35,42	52,83	3,20	1,60	4,80	57,63
MAY	3,32	1,66	9,75	30,60	45,33	2,75	1,37	4,12	49,45
JUN	3,21	1,73	9,55	30,94	45,44	2,75	1,38	4,13	49,57
JUL	3,04	1,47	10,86	24,40	39,78	2,41	1,21	3,62	43,39
AGO	3,36	4,80	14,10	44,90	67,15	4,07	2,03	6,10	73,26
SEP	3,24	4,33	11,65	39,26	58,48	3,54	1,77	5,32	63,79
ОСТ	3,24	4,33	11,65	39,26	58,48	3,54	1,77	5,32	63,79
NOV	3,24	4,33	11,65	39,26	58,48	3,54	1,77	5,32	63,79
DIC	3,24	4,33	11,65	39,26	58,48	3,54	1,77	5,32	63,79
TOTAL	39,31	34,59	133,66	430,77	638,33	38,69	19,34	58,03	696,36

Fuente: Ministerio de Hidrocarburos y Energía Balance Energético Nacional

En la siguiente tabla tenemos los datos de las regalías por IDH del año 2011.

Tabla 3.4. Regalías por departamento (MM\$us) 2011

MES	R	Regalia Departamental del 17%		Regalía Departamental del 11% Total Regalía Nacional Regalía Compensatoria del 1'			Total Regalía Compensatoria	TOTAL	TOTAL	
MES	СВВА	CHUQ	STA. CRUZ	TJA (**)	Dptal. del 11%	BENI	PANDO	del 1%	2011	2010
ENE (*)	2,48	1,63	6,13	23,39	33,63	2,04	1,02	3,06	36,68	24,63
FEB (*)	2,48	1,63	6,13	23,39	33,63	2,04	1,02	3,06	36,68	24,63
MAR (*)	1,90	0,87	5,99	19,71	28,46	1,76	0,88	2,64	31,10	21,63
ABR	2,19	1,87	6,26	17,76	28,08	1,70	0,85	2,55	30,63	24,47
MAY	2,34	1,75	6,19	20,57	30,85	1,87	0,93	2,80	33,66	27,99
JUN	2,59	1,71	7,06	20,96	32,32	1,96	0,88	2,84	35,16	31,45
JUL	2,34	1,66	6,12	24,17	34,29	2,08	1,14	3,22	37,51	28,01
AGO	3,97	1,81	8,92	27,27	41,96	2,54	1,37	3,92	45,88	35,98
SEP	2,97	1,73	7,39	22,46	34,55	2,09	1,05	3,14	37,69	35,90
ОСТ	3,32	1,86	9,79	32,21	47,18	2,86	1,43	4,29	51,47	37,60
NOV	3,21	1,73	9,56	30,94	45,44	2,75	1,38	4,13	49,57	37,79
DIC	3,26	1,89	9,67	29,01	43,83	2,66	1,33	3,98	47,82	36,40
TOTAL	33,06	20,13	89,21	291,83	434,22	26,35	13,28	39,63	473,85	366,48

Fuente: Ministerio de Hidrocarburos y Energía Balance Energético Nacional

Se toma como variable de salida el consumo de gas natural total en Bolivia, incluyendo, residenciales, comerciales, industria, transporte vehicular, el sector eléctrico, consumidores directos y otros ver figura 3.2.

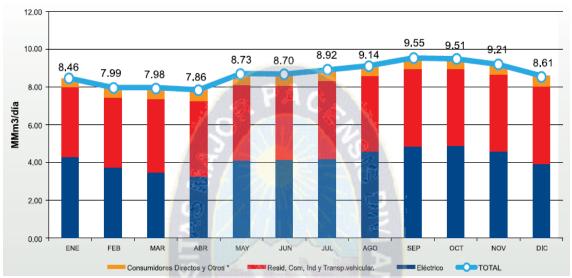


Figura: 3.2. Consumo de gas natural en el mercado interno **Fuente: Y**PFB, 2012

Observamos el consumo de gas natural en el 2011 ver figura 3.3.

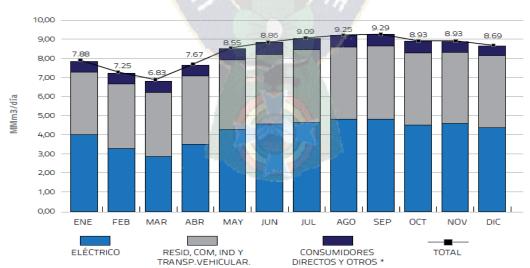


Figura: 3.3. Consumo de gas natural en el mercado interno **Fuente:** YPFB, 2011

Los datos del consumo de gas natural en el mercado interno ver tabla 3.5.

Tabla 3.5. Consumo de gas natural en el mercado interno (MMm3/día)

MES	2011	2012
ENE	7.88	8.46
FEB	7.25	7.99
MAR	6.83	7.98
ABR	7.67	7.86
MAY	8.55	8.73
JUN 🧪	8.86	8.70
JUL /	9.09	8.92
AGO	9.25	9.14
SEP	9.29	9.55
OCT	8.93	9.51
NOV	8.93	9.21
DIC	8.69	8.61

Fuente: YPFB

3.3.2. Estructura de Datos

Ahora trabajamos con las variables que intervienen en el aprendizaje de la red neuronal los cuales son: energía eléctrica (termoelectricidad) que está en base del gas natural y las regalías del IDH, los datos se muestran en la siguiente tabla 3.4.

En estos periodos el uso del gas natural subió de una manera sorprendente especialmente en el sector energético que actualmente consume el 50% del consumo interno.

También aumentaron las regalías gracias a la explotación de este recurso natural que satisface a la población de manera que cada departamento tiene más recursos para seguir progresando.

En este punto ordenaremos las variables de entrada las cuales son el consumo de la energía eléctrica(termoelectricidad) y las regalías por IDH que ofrece la venta de este energético a países aledaños a Bolivia.

Datos de la gestión 2011 y 2012 de las regalías del IDH del consumo del sector eléctrico y el consumo de gas natural en el país.

Tabla 3.6. Estructuración de datos

A = a	Año MES		REGALÍAS	SECTOR	CONSUMO DE
Ano			IDH	ELÉCTRICO	GAS NATURAL
	ENE	1	36.68	4,05	7.88
	FEB	2	36.68	3,31	7.25
	MAR	3	31.10	2,89	6.83
	ABR	4	30,63	3,53	7.67
	MAY	5	33,66	4,32	8.55
2011	JUN	6	35,16	4,59	8.86
2011	JUL	7	37,51	4,67	9.09
	AGO	8	45,88	4,86	9.25
	SEP	9	37,69	4,85	9.29
	OCT	10	51,47	4,55	8.93
	NOV	11	49,57	4,63	8.93
	DIC	12	47,82	4,41	8.69
	ENE	13	68,52	4,32	8.46
	FEB	14	49,68	3,76	7.99
	MAR	15	49.68	3,49	7.98
	ABR	16	57,63	3,28	7.86
	MAY	17	49,45	4,13	8.73
2012	JUN	18	49,57	4,17	8.70
2012	JUL	19	43,39	4,18	8.92
	AGO	20	73,26	4,59	9.14
	SEP	21	63,79	4,87	9.55
	OCT	22	63,79	4,88	9.51
	NOV	23	63,79	4,6	9.21
	DIC	24	63,79	3,93	8.61

Fuente: Elaboración propia

3.3.3. Normalización de Datos

Debido a la discrepancia de datos que existe en la tabla 3.5 es necesario normalizar los datos con la ayuda de la ecuación:

$$v = \frac{b - bmin}{bmax - bmin}$$

Donde: v: dato normalizado, **b**: dato a normalizar, **b**min: dato más pequeño de la serie y **b**max: dato más grande de la serie

Para que tengamos datos homogéneos, por que los resultados muchas veces dependen de la normalización de datos.

Tabla 3.7Datos Normalizados

MES		REGALÍAS	SECTOR	CONSUMO DE GAS
	R	IDH	ELÉCTRICO	NATURAL
ENE	1	0.1419	0.5829	0.3860
FEB	2	0.1419	0.2111	0.1544
MAR	3	0.0110	0	0
ABR	4	0	0.3216	0.3088
MAY	5	0.0711	0.7186	0.6324
JUN	6	0.1063	0.8543	0.7463
JUL	7	0.1614	0.8945	0.8309
AGO	8	0.3577	0.9899	0.8897
SEP	9	0.1656	0.9849	0.9044
OCT	10	0.4889	0.8342	0.7721
NOV	11	0.4443	0.8744	0.7721
DIC	12	0.4032	0.7638	0.6838
ENE	13	0.8888	0.7186	0.5993
FEB	14	0.4469	0.4372	0.4265
MAR	15	0.4469	0.3015	0.4228
ABR	16	0.6334	0.1960	0.3787
MAY	17	0.4415	0.6231	0.6985
JUN	18	0.4443	0.6432	0.6875
JUL	19	0.2993	0.6482	0.7684
AGO	20	1	0.8543	0.8493
SEP	21	0.7779	0.9950	1
OCT	22	0.7779	1	0.9853
NOV	23	0.7779	0.8593	0.8750
DIC	24	0.7779	0.5226	0.6544

Fuente: Elaboración propia

3.3.4. División de Datos

Para realizar un estudio bien detallado, la división de datos es necesaria como vimos en el capitulo anterior tenemos que dividir los datos en tres partes las cuales son: el conjunto de entrenamiento el cual tendrá 16 datos, el de prueba con 4 datos y también el de validación.

Tabla 3.8. Conjunto de entrenamiento

	REGALÍAS	SECTOR SECTOR	CONSUMO DE
MES		_	
	IDH	ELÉCTRICO	GAS NATURAL
1	0.1419	0.5829	0.3860
2	0.1419	0.2111	0.1544
3	0.0110	0	0
4	0	0.3216	0.3088
5	0.0711	0.7186	0.6324
6	0.1063	0.8543	0.7463
7	0.1614	0.8945	0.8309
8	0.3577	0.9899	0.8897
9	0.1656	0.9849	0.9044
10	0.4889	0.8342	0.7721
11	0.4443	0.8744	0.7721
12	0.4032	0.7638	0.6838
13	0.8888	0.7186	0.5993
14	0.4469	0.4372	0.4265
15	0.4469	0.3015	0.4228
16	0.6334	0.1960	0.3787

Fuente: Elaboración propia

Datos de prueba, para nuestro caso los datos son 4 ver tabla 3.9.

Tabla 3.9. Conjunto de prueba

MES	REGALÍAS IDH	SECTOR ELÉCTRICO	CONSUMO DE GAS NATURAL
17	0.4415	0.6231	0.6985
18	0.4443	0.6432	0.6875
19	0.2993	0.6482	0.7684
20	1	0.8543	0.8493

Fuente: Elaboración propia

Datos de validación, para nuestro caso son 4 ver tabla 3.10.

Tabla 3.10. Conjunto de validación

Tubia estas conjunto de variaderon					
MES	REGALÍAS IDH	SECTOR ELÉCTRICO	CONSUMO DE GAS NATURAL		
21	0.7779	0.9950	1		
22	0.7779	1	0.9853		
23	0.7779	0.8593	0.8750		
24	0.7779	0.5226	0.6544		

Fuente: Elaboración propia

3.4. Determinación de la Arquitectura de la Red

La arquitectura de red está basada principalmente en el ingreso de datos por que las variables de entrada que en nuestro caso son dos afectan de gran manera en el aprendizaje de la red, también tenemos capas ocultas que contaran con una cantidad de neuronas y finalmente tenemos una capa de salida que es el consumo de gas natural.

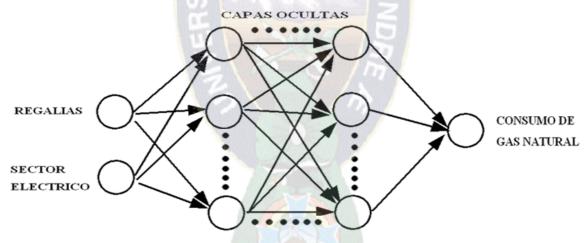


Figura3.4. Arquitectura de la red Fuente: Elaboración propia

Como observamos el número el número de capas ocultas esta en base a un error admisible, el objetivo de la red neuronal es dotar de un número adecuado de neuronas, una vez que se han inicializado los pesos de la red y los del umbral, la red está lista para ser entrenada. La red puede entrenarse para: la aproximación de una función (regresión no lineal), la

asociación del modelo, o la clasificación del modelo. El proceso de entrenamiento requiere de los patrones de conducta apropiados para la red, las entradas de la red p y las salidas en blanco t. Durante el entrenamiento los pesos y el umbral de la red son iterativamente ajustados para minimizar la función de desempeño.

3.4.1. Descripción de la RNA para el Consumo de gas Natural

La neurona artificial que en nuestro caso será RN, para la predicción del consumo de gas natural viene dada por la función.

$$RN=(X, H, Y)$$
 (3.1)

Donde:

X: Es el conjunto de entradas para la red.

H: Es el conjunto de neuronas de la capa oculta, ver ecuación (2.11).

Y: Es la única salida de la red, ver ecuación (2.12).

Podemos representar la salida de la siguiente forma: Y=f(X, H) (3.2)

La entrada a la red neuronal artificial para la predicción del consumo de gas natural de Bolivia, es el vector de la entrada X pertenece a RN con la forma:

$$X=(x_1, x_2)$$
 (3.3)

Donde: x_1 son las regalías y x_2 energía eléctrica.

La salida de la red neuronal está dada por el vector: $Y=(y_i)$, donde y_i es el consumo de gas natural en Bolivia.

Como estamos trabajando en un rango de [0, 1], ver ecuación (2.8), debido a la normalización de datos que hicimos anteriormente la función que se adecua a la red neuronal artificial es la función sigmoidea:

$$y_j = \frac{1}{1 + e^{-Netj}}$$
 (3.4)

Con:
$$Net_j = \sum_{i=1}^{N} w_{ij} y_i$$
.

3.5. Diseño de la Estrategia de Aprendizaje o Entrenamiento de la Red

Para iniciar el entrenamiento se le presenta a la red un patrón de entrenamiento, el cual se describe a continuación, primeramente se realizo experimentos con distintas capas ocultas buscando el error mínimo para así obtener el numero de capas ocultas y obtuvimos que para el entrenamiento necesitemos dos capas ocultas, con 8 neuronas cada una, para el entrenamiento de la red se tiene una matriz tomando valores aleatoriamente como vemos a continuación.

Matriz de 8x2				
	0.53	0.13		
	0.23	0.54		
8/	0.54	0.24		
W =	0.87	0.44		
vv –	0.21	0.21		
	0.54	0.43		
	0.45	0.64		
は一個で色	0.87	0.23		

La salida también viene en forma de matriz el cual presenta el siguiente formato:

onsumo de gas natural					
	dato 1				
	dato 2				
$A_i =$	dato 3				
	dato n				

Donde A_i, representa el consumo de gas natural mensual, para luego tomar los datos y realizar la predicción.

Para unir la capa oculta con la capa de salida es necesario una capa umbral que también presente datos aleatorios que estén en una rango [0, 1] para la inicialización de pesos, porque es imposible conocer por adelantado los puntos mínimos de error. Así es necesaria una regla que combine las entradas de una neurona con los pesos de las conexiones, también se requiere una regla que combine las entradas con el estado actual de la neurona

para producir un nuevo estado de activación, la función sigmoidea también se aplica en la capa de entrada, en las capas ocultas y entre la capa oculta y de salida.

3.5.1. Regla de Aprendizaje

Esta regla se conoce como la modificación del comportamiento inducido por la interacción con el entorno y como resultado de experiencias conducentes al establecimiento de nuevos modelos de respuestas ante estímulos externos.

El aprendizaje esta realizado por el algoritmo de entrenamiento Backpropagation que desarrollamos en el capítulo anterior, en MathLab utilizamos el algoritmo Levenber – Marquardt que está basado en el algoritmo backpropagation debido a su rapidez de convergencia.

Ahora observamos la interfaz grafica de la simulación de predicción del consumo de gas natural aplicando redes neuronales artificiales.



Figura3.5.Interfaz de la presentación **Fuente:** Elaboración propia

Ahora vemos la interfaz de construcción con 2 capas ocultas de 8 neuronas cada una.

PRONOSTICO DEL CONSUMO DE GAS

Construcción de la Red Neuronal						
Datos de Entrada :	Seleccionar archivo entradas.txt					
Datos de Salida :	Seleccionar archivo salidas.txt					
Número de Neuronas Capa 1 :	8 💌					
Número de Neuronas Capa 2 :	8 🔻					
	Entrenar y continuar					

Figura3.6.Interfaz de la construcción de la red **Fuente:** Elaboración propia

Como resultado del entrenamiento obtenemos la matriz de pesos que nos ayudara con la predicción.



Figura3.7.Interfaz de predicción de datos **Fuente:** Elaboración propia

Primeramente se hace una normalización y luego se procede a entrenar, encontrando la matriz de pesos del entrenamiento, finalmente obtenemos, los datos de predicción.

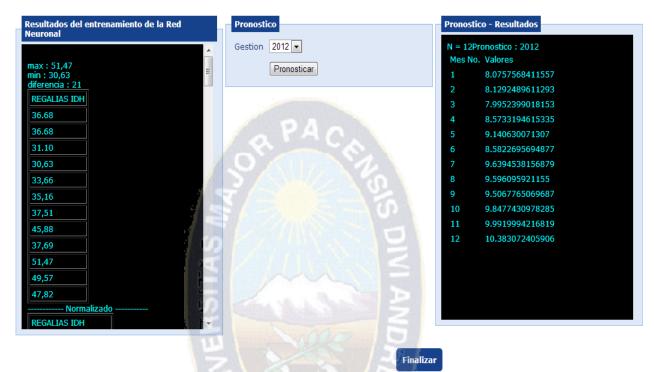


Figura 3.8. Interfaz de datos para la gestión 2012 **Fuente:** Elaboración propia

Para calcular el Error tenemos:

$$Et = Yt - Ft$$

Finalmente calculamos los errores de la red neuronal de la siguiente manera:

$$MAPE = \frac{\sum_{t=1}^{N} \frac{|Et|}{|Yt|}}{N}$$

Donde:

Et: Es el error del pronostico del periodo t.

Yt: Es el valor real para este periodo t.

Ft: Es el valor que se habia pronosticado para el periodo t.

N: Numero de datos.

El error de la red neuronal es:

Tabla 3.11Error de la red neuronal

Tubia 5.112.1101 de la fed ficational						
Datos	Datos de la	Error				
Originales	Red Neuronal	LITOI				
8.46	8.664	0.024				
7.99	8.049	0.007				
7.98	8.631	0.063				
7.86	8.353	0.062				
8.73	8.872	0.016				
8.70	8.505	0.022				
8.92	9.689	0.053				
9.14	9.774	0.069				
9.55	9.072	0.033				
9.51	9.997	0.051				
9.21	9.972	0.082				
8.61	9.592	0.094				
	Error total	0.585				

Fuente: Elaboración propia

El error de la red neuronal es:

$$MAPE = \frac{0.585}{12} = 0.04$$

Para calcular el error porcentual tenemos:

$$MAPE \% = 0.04 * 100\% = 4\%$$

CAPITULO IV

PRUEBA DE HIPÓTESIS

4.1. Introducción

En este capítulo la hipótesis se somete a pruebas o escrutinio para determinar si son apoyadas o refutadas de acuerdo a lo que se observo, en realidad no podemos probar que una hipótesis sea verdadera o falsa, si no argumentar que de acuerdo con ciertos datos obtenidos en la investigación es apoyado o no, desde el punto de vista técnico no se acepta una hipótesis a través de un estudio, si no que aporte evidencia en su favor o en su contra , cuantas más investigaciones apoyen una hipótesis más credibilidad tendrá esta, es un procedimiento basado en la teoría de la probabilidad, se emplea para determinar si la hipótesis es una afirmación razonable.

Este es uno de los aspectos más útiles de la inferencia estadística, puesto que muchos tipos de problemas de toma de decisiones, pruebas o experimentos en el mundo de la ingeniería, pueden formularse como problemas de prueba de hipótesis.

4.2. Modelado con medias móviles

Los datos del consumo de gas natural en Bolivia están enMMm3/día, ver Tabla 3.5, para empezar necesitamos suavizar los datos empleando el método de medias móviles.

$$\frac{y_1+y_2+\cdots+y_n}{N}$$
; $\frac{y_2+y_3+\cdots+y_{n+1}}{N}$; $\frac{y_3+y_4+\cdots+y_{n+2}}{N}$;

Donde: y_1 , es el consumo de gas natural en el mes de enero, y_2 , el consumo de gas natural en el mes de febrero y así sucesivamente hasta y_{12} , en nuestro caso será N=3, por que el estudio de los casos será trimestral, entonces tenemos lo siguiente:

$$\frac{7.88+7.25+6.83}{3}, \frac{7.25+6.83+7.67}{3}, \dots, \frac{8.93+8.93+8.69}{3}$$

La tabla queda de la siguiente manera:

Tabla 4.1. Media móvil

MES		CONSUMO DE	MEDIA	
		GAS NATURAL	MÓVIL	
ENE	1	7.88		
FEB	2	7.25	7.320	
MAR	3	6.83	7.250	
ABR	4	7.67	7.683	
MAY	5	8.55	8.360	
JUN	6	8.86	8.833	
JUL	7	9.09	9.066	
AGO	8	9.25	9.210	
SEP	9	9.29	9.156	
ОСТ	10	8.93	9.050	
NOV	11	8.93	8.850	
DIC	12	8.69		

Fuente: Elaboración propia

Ahora calculamos la media móvil centrada.

Media móvil centrada : $\frac{7,320+7,250}{2}$; $\frac{7,250+7.683}{2}$; ...; $\frac{9.050+8.850}{2}$

:7.285 ;7.466; ... ; 8.9<mark>5</mark>0

Tabla 4.2 Media móvil centrada

	Tabla 4.2. Wedia movii centrada					
MES		CONSUMO DE	MEDIA	MEDIA MÓVIL		
IVIES		GAS NATURAL	MÓVIL	CENTRADA		
ENE	1	7.88				
FEB	2	7.25	7.320			
MAR	3	6.83	7.250	7.285		
ABR	4	7.67	7.683	7.466		
MAY	5	8.55	8.360	8.021		
JUN	6	8.86	8.833	8.596		
JUL	7	9.09	9.066	8.949		
AGO	8	9.25	9.210	9.138		
SEP	9	9.29	9.156	9.183		
ОСТ	10	8.93	9.050	9.103		
NOV	11	8.93	8.850	8.950		
DIC	12	8.69				

Fuente: Elaboración propia

1) Procedemos a calcular el componente estacional que es igual al consumo mensual de gas natural dividido entre la media móvil centrada:

Tabla 4.3. Componente estacional

. 456		CONSUMO DE	MEDIA	MEDIA MÓVIL	COMPONENTE
MES		GAS NATURAL	MÓVIL	CENTRADA	ESTACIONAL
ENE	1	7.88			
FEB	2	7.25	7.320		
MAR	3	6.83	7.250	7.285	0.937
ABR	4	7.67	7.683	7.466	1.027
MAY	5	8.55	8.360	8.021	1.022
JUN	6	8.86	8.833	8.596	1.030
JUL	7	9.09	9.066	8.949	1.015
AGO	8	9.25	9.210	9.138	1.012
SEP	9	9.29	9.156	9.183	1.011
ОСТ	10	8.93	9.050	9.103	0.981
NOV	11	8.93	8.850	8.950	0.997
DIC	12	8.69			

Fuente: Elaboración propia

Calculamos en índice estacional partiendo del componente estacional como vemos a continuación.

Tabla 4.4 Índice estacional

			Tabla	.4 indice esta	acionai	
MES		COMPONENTE	COMPONENTE		ÍNDICE	
		ESTACIONAL			ESTACIONAL	
ENE	1		1.022	1.011		1.016
FEB	2		1.030	0.981		1.005
MAR	3	0.937	0.937	1.015	0.997	0.983
ABR	4	1.027	1.027	1.012		1.019
MAY	1	1.022	ASS	480		1.016
JUN	2	1.030	ANY			1.005
JUL	3	1.015				0.983
AGO	4	1.012				1.019
SEP	1	1.011				1.016
ОСТ	2	0.981				1.005
NOV	3	0.997				0.983
DIC	4					1.019

Fuente: Elaboración propia

Ahora realizamos la serie desestacionalizada (SD), que se realiza como vemos a continuación:

$$SD = \frac{ConsumoMensual}{IndiceEstacional}$$

Los datos ya no contienen movimientos irregulares como vemos a continuación.

Tabla 4.5. Serie desestacionalizada

MES		CONSUMO DE	ÍNDICE	SERIE
		GAS NATURAL	ESTACIONAL	DESESTACIONALIZADA
ENE	1	7.88	1.016	7.755
FEB	2	7.25	1.005	7.213
MAR	3	6.83	0.983	6.948
ABR	4	7.67	1.019	7.526
MAY	5	8.55	1.016	8.415
JUN	6	8.86	1.005	8.815
JUL	7	9.09	0.983	9.247
AGO	8	9.25	1.019	9.077
SEP	9	9.29	1.016	9.143
ОСТ	10	8.93	1.005	8.885
NOV	11	8.93	0.983	9.084
DIC	12	8.69	1.019	8.527

Fuente: Elaboración propia

Ahora realizamos una regresión lineal después de obtener los datos de la serie desestacionalizada.

$$y *= a + bx$$

Donde:

La variable dependiente es la columna de la serie desestacionalizada y la variable independiente es la numeración de los meses para encontrar a y b recurrimos a los mínimos cuadros reemplazando en la formula de mínimos cuadros encontramos a y b, como vemos a continuación.

Haciendo los cálculos tenemos que:

$$a = 7.263$$
 y $b = 0.102$

Ahora con los datos de a y b procedemos a utilizar la regresión lineal para encontrar los datos futuros de predicción del año 2012.

Por regresión lineal encontramos los datos de enero del 2012 hasta diciembre del 2012, pero cabe señalar que estos datos no son los finales.

Tabla 4.6. Regresión lineal

y(13) = 7.263 + 0.102(13) = 8.589
y(14) = 7.263 + 0.102(14) = 8.691
y(15) = 7.263 + 0.102(15) = 8.793
y(16) = 7.263 + 0.102(16) = 8.895
y(17) = 7.263 + 0.102(17) = 8.997
y(18) = 7.263 + 0.102(18) = 9.099
y(19) = 7.263 + 0.102(19) = 9.201
y(20) = 7.263 + 0.102(20) = 9.303
y(21) = 7.263 + 0.102(21) = 9.405
y(22) = 7.263 + 0.102(22) = 9.507
y(23) = 7.263 + 0.102(23) = 9.609
y(24) = 7.263 + 0.102(24) = 9.711
The state of the s

Fuente: Elaboración propia

Para encontrar los datos de la predicción realizamos lo siguiente:

$$y = (y *) * indice estacional$$

Luego realizaremos una comparación del método con los datos reales y finalmente calculamos el error.

Tabla 4.7. Error de las medias móviles

(y *) * indice estacional	Error
8589 * 1.016 = 8.726	0.031
8691 * 1.005 = 8.734	0.093
8793 * 0.983 = 8.643	0.083
8895 * 1.019 = 9.064	0.153
8997 * 1.016 = 9.140	0.047
9099 * 1.005 = 9.144	0.051
9201 * 0.983 = 9.044	0.013
9303 * 1.019 = 9.479	0.037
9405 * 1.016 = 9.555	0.00052
9507 * 1.005 = 9.554	0.0046
9609 * 0.983 = 9.445	0.025
9711 * 1.019 = 9.895	0.149
Error total	0.647

Fuente: Elaboración propia

Calculamos el error absoluto porcentual de la media, el error total esta dado por:

$$MAPE = \frac{\sum_{t=1}^{N} \frac{|Et|}{|Y_t|}}{N}$$

Donde: Et: Es el error del pronóstico del periodo t.

Yt :Es el valor real para este periodo t.

$$MAPE = \frac{0.647}{12} = 0.05$$

Para calcular el error porcentual tenemos:

$$MAPE \% = 0.05 * 100\% = 5\%$$

4.3. Procedimiento de la Prueba de Hipótesis

Se realiza mediante un procedimiento de pasos los cuales son: plantear la hipótesis nula y alternativa, seleccionar el nivel de significancia, identificar el estadígrafo de prueba, formular la regla de decisión y finalmente se toma la muestra y se decide.

Siguiendo este procedimiento al llegar al último paso se puede o no rechazar la hipótesis pero debemos de tener cuidado con esta determinación ya que en la consideración estadística no proporciona evidencia de que algo sea verdadero.

Para la prueba de hipótesis utilizaremos la distribución t de student con una muestra de tamaño 7, para la prueba de hipótesis trabajaremos con los datos predichos en el 2012 comenzando del mes de Mayo hasta el mes de noviembre, con el error relativo tanto de las medias móviles y redes neuronales artificiales precedemos a calcular la media y la desviación típica.

Calculo de errores de las medias móviles y las redes neuronales.

Tabla 4.8.Error de la red neuronal y media móvil

MESES	ERROR RELATIVO		
WIESES	MM	RNA	
5	0.016	0.047	
6	0.022	0.051	
7	0.053	0.013	
8	0.069	0.037	
9	0.033	0.0005	
10	0.051	0.004	
11	0.082	0.025	
Media	0.046	0.025	
Desv. Tip.	0.167	0.017	

Fuente: Elaboración propia

4.3.1. Planteamiento de Hipótesis Nula y Alternativa

La hipótesis nula (H_o), es una afirmación que no se rechaza a menos que los datos muéstrales proporcionen evidencia convincente de que es falsa, también contiene un signo de igualdad con respecto con el valor especificado.

La hipótesis alternativa (H_1) , es la que difiere con la hipótesis nula, es una afirmación que se acepta si los datos proporcionan evidencia de que (H_0) es falsa.

Como estuvimos trabajando con errores tanto en las medias móviles y las redes neuronales artificiales, es necesario establecer que la red neuronal tiene menor error que las medias móviles y es mejor en la predicción.

En nuestro caso la hipótesis nula y alternativa son:

 H_o : $\mu_1 = \mu_2$ Los errores son iguales.

 $H_1 \hbox{:} \qquad \mu_1 < \ \mu_2 \quad \hbox{El error de la red neuronal es menor a la de la media móvil}.$

Donde: μ_1 : Es el error de la red neuronal.

 μ_2 : Es el error de la media móvil.

4.3.2. Selección del nivel de significancia

El nivel de significancia es la probabilidad de rechazar la hipótesis nula cuando sea verdadera, se le denota mediante la letra griega α , también denominada como nivel de riesgo, porque se corre el riesgo de rechazar la hipótesis nula, cuando es verdadera.

Si suponemos que la hipótesis planteada es verdadera, entonces, el nivel de significancia indicara la probabilidad de no aceptarla, es decir, estén fuera del área de aceptación. El nivel de confianza $(1 - \alpha)$, indica la probabilidad de aceptar la hipótesis planteada, cuando es verdadera.

Cualquier sea la decisión tomada de una prueba de hipótesis, ya sea de aceptación de la (H_0) o de la (H_1) , puede incurrirse en error del cual se tiene dos tipos los cuales son:

El error del tipo uno, se presenta si la hipótesis nula (H_o) es rechazada cuando es verdadera y debía ser aceptada, la probabilidad de cometer un error de tipo uno se denomina con la letra alfa α .

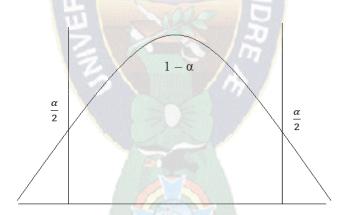
El error del tipo dos, se denota con la letra griega β se presenta si la hipótesis nula es aceptada cuando de hecho es falsa y debía ser rechazada.

	INVESTIGADOR		
HIPÓTESIS NULA	Se acepta (Ho)	Se rechaza (Ho)	
NULA	(110)	(110)	
(H ₀) ES VERDADERA	Decisión	Error tipo I	
	Correcta	1	
(Ho) ES FALSA	Error tipo II	Decisión	
		correcta	

Para cualquier ensayo de hipótesis debe diseñarse de forma que minimice los errores de decisión, en la práctica un tipo de error puede tener más importancia que el otro, y así se tiene que conseguir poner una limitación al error de mayor importancia. La única forma de reducir ambos tipos de errores es incrementar el tamaño de la muestra, lo cual puede ser o no ser posible.

En la práctica se establece el nivel α y para disminuir el error de β se incrementa el número de observaciones en la muestra, pues así se acortan los limites de confianza respecto a la hipótesis planteada. La meta de las pruebas estadísticas es rechazar la hipótesis planteada.

En otras palabras, es deseable aumentar cuando esta es verdadera, o sea, incrementar $(1-\alpha)$, la aceptación de la hipótesis planteada debe interpretarse como interpretarse como que la información aleatoria de la muestra disponible no permite detectar la falsedad de esta hipótesis.



En nuestro caso α será: $\alpha = 0.05$, el nivel de confianza o significancia será de un 95%.

4.3.3. Estadígrafo de Prueba

El valor determinado se consigue a partir de la información muestral, que se utiliza para determinar si se rechaza la hipótesis nula, en nuestro caso utilizaremos es estadístico t.

El tipo de prueba a utilizar en nuestro caso será la prueba bilateral o de dos extremos, la hipótesis planteada se formula con la igualdad, en nuestro caso será de la siguiente forma:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sigma \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$

Donde:

$$\sigma = \sqrt{\frac{N_1 S_1^2 + N_2 S_2^2}{N_1 + N_2 - 2}}$$

Calculando tenemos:

$$\sigma = \sqrt{\frac{7(0.167)^2 + 7(0.017)^2}{7 + 7 - 2}} = 0.444$$

Ahora calculamos t:

$$t = \frac{0.046 - 0.025}{0.444\sqrt{\frac{1}{7} + \frac{1}{7}}} = 0.166$$

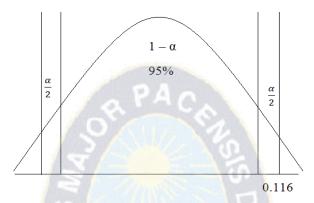
4.3.4. Regla de Decisión

Se establece las condiciones específicas en la que se rechaza la hipótesis nula y las condiciones en que no se rechaza la hipótesis nula.

La región de rechazo define la ubicación de todos los valores que son tan grandes o tan pequeños, que la probabilidad de que se presenten bajo la suposición de que la hipótesis nula es verdadera, es muy remota.

Formulamos la regla de decisión teniendo en cuanta que esta es una prueba de colas, por consiguiente la regla de decisión es rechazar la hipótesis nula y aceptar la hipótesis alternativa, si el valor z calculado no queda en la región comprendida $^+_{-}$ 0.166.

Ahora observamos el intervalo de _ 0.166.



4.3.5. Toma de Decisión

Calculamos el estadístico de prueba, comparamos con el valor crítico y tomamos la decisión de rechazar o no la hipótesis nula. Tenga presente que en una prueba de hipótesis solo se puede tomar una de dos decisiones: aceptar o rechazar la hipótesis nula.

Debe subrayarse que siempre existe la posibilidad de rechazar la hipótesis nula cuando no debería haberse rechazado (error del tipo I), también existe la posibilidad de que la hipótesis nula se acepte cuando debería haberse rechazado (error del tipo II).

Por la regla de decisión el nivel de significación α es la siguiente:

Si $t_{exp} > t_{\alpha}$, entonces se acepta la hipótesis alternativa en nuestro caso tomamos 0.2 de la región crítica y calculamos t_{exp} y nos dio 0.922 y vemos que se cumple:

0.922 > 0.116. Por lo cual se acepta la hipótesis alternativa.

CAPITULO V

CONCLUSIONES Y RECOMENDACIONES

5.1. Introducción

En este capítulo se describirá las conclusiones y recomendaciones a las que se llegaron en el punto final de la investigación para posteriores trabajos en el área de la Inteligencia Artificial.

5.2. Conclusiones

En la presente investigación de tesis podemos concluir que las redes neuronales artificiales tienen varios usos en distintas áreas, en nuestro caso fue de gran utilidad en la predicción del consumo de gas natural, además el pronóstico se realizó a nivel nacional excluyendo algunos factores que todavía no tienen tanta relevancia como el gas domiciliario, que a futuro tendrá una gran importancia, una de las variables de entrada fue el consumo de la termoelectricidad ya que utiliza como fuente a el gas natural y casi el 50% del consumo de gas natural es por parte de las termoeléctricas, las regalías por la venta de gas natural también forma parte de la variable de entrada.

La investigación realizada ofrece información confiable debido a que se recurrió a un método estadístico (medias moviles) para validar la predicción de la red neuronal artificial, y los datos son de una fuente confiable Yacimientos Petrolíferos Fiscales Bolivianos (YPFB), y del instituto nacional de estadística(INE).

5.2.1. Estado de los Objetivos

Los objetivos se fueron cumpliendo a medida que avanzamos en la investigación por ejemplo los objetivos específicos.

Analizar y considerar las variables que influyen en el consumo de gas natural por parte de la población.

Se consideró dos variables de entrada, el consumo de la energía termoeléctrica y las regalías por la venta de gas natural, ya que ambos son factores que afectan en el consumo de gas natural.

Los datos fueron recopilados de fuentes confiables como ser de Yacimientos Petrolíferos Fiscales Bolivianos y el Instituto Nacional de Estadística.

Con los datos de la predicción podemos predecir el consumo de gas natural a futuro teniendo la certeza de que los datos son confiables previendo así el desabastecimiento de este energético.

En la elaboración del modelo de simulación neuronal se tuvo la certeza de que es un modelo confiable ya que se implementó un modelo estadístico (medias móviles), y se comparó con el modelo neuronal de acuerdo al error de cada uno de los modelos en nuestro caso la diferencia del error de ambos modelos es de 0.5 siendo la red neuronal el que tiene un error menor.

Para la construcción de redes neuronales artificiales se aplico como herramienta de desarrollo el software MatLab, lo cual facilito el desarrollo de la herramienta en gran medida posteriormente se codifico en php.

5.2.2. Estado de la Hipótesis

Se realizó un contraste de diferencia de medias con una distribución t student, donde se llegó a la conclusión de que las redes neuronales presentan menos error que el modelo estadístico (medias móviles). Afirmamos que efectivamente realizamos el modelo desimulación para la red neuronal para predecir en consumo de gas natural en un tiempo determinado.

5.3. Recomendaciones

Si queremos que la efectividad de la red neuronal sea lo más mínimo posible es necesario ampliar los datos de tal manera que la red neuronal pueda entrenarse mejor.

5.3.1. Trabajos Futuros

Tomar otro tipo de variables de entrada que a futuro darán una gran eficiencia al consumo de gas natural como ser las conexiones de gas natural en todos los departamentos, el número de habitantes y el precio por el servicio.

5.3.2. Líneas de Investigación

El área de la Inteligencia Artificial es muy grande y las redes neuronales artificiales forman parte de ella, las áreas de trabajo son enormes por que se pueden implementar, en cualquier área que deseemos

Aplicar otro método de la Inteligencia Artificial como ser los algoritmos genéticos y variables de incertidumbre.

Bibliografía

C. ROBERTO HERNÁNDEZ SAMPIERI(1991) "METODOLOGÍA DE LA INVESTIGACIÓN", Escuela Superior de Comercio y AdministraciónInstituto Politécnico Nacional(TESIS) México.

RAMIRO CUENTAS GONZALES(2004) "Sistema Experto de soporte técnico asistido vía telefónica para un ISP", Licenciatura en mención Ing. Sistemas Informática Universidad Mayor de San Andrés (TESIS) La Paz Bolivia.

RODRIGO SIMÓN QUELCA SIRPA(2011) "Reconocimiento de Cromosomas Humanos en imágenes basado en una Red Neuronal de Retropropagación", Licenciatura en mención Ing. Sistemas Informática Universidad Mayor de San Andrés (TESIS) La Paz Bolivia.

RUBÉN QUISPE CONDORI(2011) "Predicción del consumo de agua potable de la ciudad de La Paz aplicando Redes Neuronales y Lógica Difusa", Licenciatura en mención Ing. Sistemas Informática Universidad Mayor de San Andrés (TESIS) La Paz Bolivia.

QUISPE TAMBO FÉLIX CELESTINO(2009)" Mesa de ayuda para requerimientos de hardware y software caso: Instituto Nacional de Estadística La Paz Bolivia", Licenciatura en mención Ing. Sistemas Informática Universidad Mayor de San Andrés (TESIS) La Paz Bolivia.

DAMIÁN JORGE MATICH (2001)"Redes neuronales: Conceptos Básicos y Aplicaciones", Universidad Tecnológica Nacional - Facultad Regional Rosario.

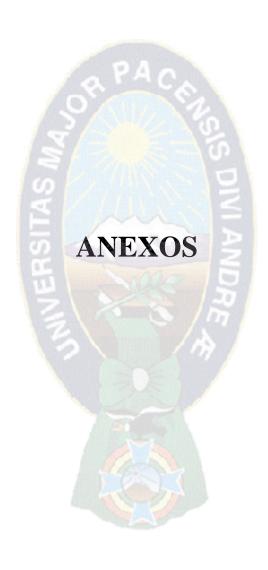
PEDRO LAGRAÑA, IÑAQUI INZA, ABDELMALIKMOUJAHID (2002)"Redes Neuronales, Departamento de Ciencias de la Computación e Inteligencia Artificial", Universidad del País vasco.

GUILLERMO CHOQUE A. (2002)"Inteligencia Artificial- perspectivas y realizaciones", Disponible: http://academic.umsanet.edu.bo/docentes/gchoque

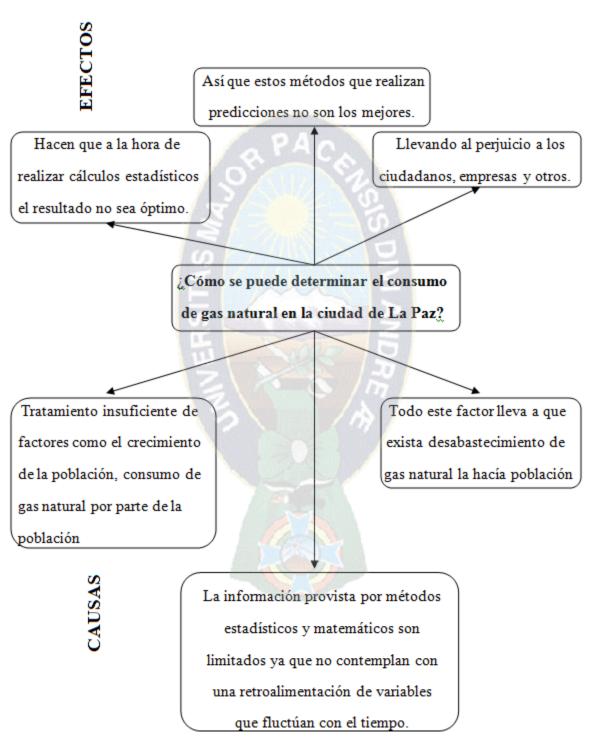
JUAN MANUEL LIZARRAZO MARRIAGA (2006)"Desarrollo de un Modelo de redes neuronales para predecir la resistencia a la comprensión y la resistividad eléctrica del concreto", Universidad Nacional de Colombia – Facultad de Ingeniería.

ROGER PRESSMAN S. (2003)"Ingeniería del Software en un enfoque practico, Quinta Edición", Editorial Concepción Fernández Madrid España.

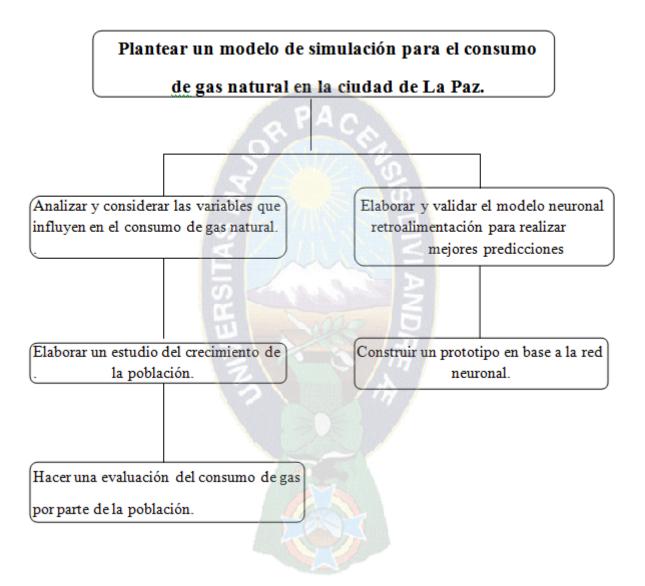
CASIMIRO GONSALES M.P. (2009)"Técnicas de predicción económica", Universidad del País del Vaco.



ANEXO A ÁRBOL DEPROBLEMAS



ANEXO B ÁRBOL DE OBJETIVOS



ANEXO C. MATRIZ DE MARCO LÓGICO

OBJETIVOS	INDICADORES	MEDIOS DE VERIFICACION	SUPUESTOS
FIN Plantear un modelo de simulación para el consumo de gas natural.	Datos del consumo de gas natural mensual o anual para una óptima toma de decisiones.	Datos estadísticos e informes del consumo de gas natural de Yacimientos Petrolíferos Fiscales Bolivianos.	Verificación y aceptación de los métodos de investigación en el ámbito local.
PROPOSITO El modelo de simulación basado en redes neuronales alimentado con variables que afectan al consumo de gas natural es capaz de predecir el consumo gas natural a futuro.	Alto porcentaje de aceptación del Modelo de simulación por parte de los investigadores.	Informes, datos estadísticos incluyendo fechas y resultados.	Obtener la aprobación de investigadores.
PRODUCTOS Método de simulación para la predicción del consumo de gas natural mediante Redes Neuronales.	Investigación del modelo de consumo de gas natural.	Información del consumo de gas natural mensual o anual de Yacimientos Petrolíferos Fiscales Bolivianos.	Se cuenta con los datos necesarios para desarrollar el modelo.
ACTIVIDADES Estudio de la situación actual, identificación de los problemas. Identificación de la metodología a utilizar. Recopilación de información del consumo de gas natural para el modelo de simulación	La metodología que se utilizara será el método descriptivo, y el modelo de simulación estará basado en Redes Neuronales y tendrá una duración de 3 a 4 meses.	Nota de los datos estadísticos de Yacimientos Petrolíferos Fiscales Bolivianos acerca del consumo de gas natural.	Existencia fiable de los datos obtenidos con el modelo.