

**UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN**



**APLICACIÓN DE LAS REDES NEURONALES ARTIFICIALES
PARA EL PRONÓSTICO DE LA DEMANDA DE AGUA POTABLE
EN LA EMPRESA EPSEL S.A DE LA CIUDAD DE LAMBAYEQUE**

**TESIS PARA OPTAR EL TÍTULO DE
INGENIERO DE SISTEMAS Y COMPUTACIÓN**

YASMÍN ANDREA VIDAURRE SIADÉN

Chiclayo 19 de Abril de 2012

**“APLICACIÓN DE LAS REDES NEURONALES ARTIFICIALES
PARA EL PRONÓSTICO DE LA DEMANDA DE AGUA POTABLE
EN LA EMPRESA EPSEL S.A DE LA CIUDAD DE LAMBAYEQUE”**

POR:

YASMÍN ANDREA VIDAURRE SIADÉN

**Presentada a la Facultad de Ingeniería de la
Universidad Católica Santo Toribio de Mogrovejo
Para optar el título de
INGENIERO DE SISTEMAS Y COMPUTACIÓN**

APROBADA POR EL JURADO INTEGRADO POR

Luis A. Zuñe Bispo

PRESIDENTE

Luis E. Barrueto Chunga

SECRETARIO

Eduardo F. Alonso Pérez

ASESOR

DEDICATORIA

A mis padres y hermanas porque son las personas que me ayudaron a seguir adelante día a día, además porque me apoyaron constantemente a superarme.

AGRADECIMIENTO

Al ver materializada esta meta, logro reafirmar que en compañía de Dios no hay prueba que no pueda ser superada. Quiero expresar mis más profundos agradecimientos a todas las personas que de una manera u otra formaron parte de esta tarea.

A mis padres y hermanas por su apoyo incondicional y por ser fuente de sabiduría.

Especialmente le agradezco a mi asesor Eduardo Alonso, por todos esos conocimientos que me permitió adquirir y por toda su paciencia.

A todos mis amigos y amigas por siempre estar allí y compartir los momentos gratos y menos gratos de esta experiencia

ÍNDICE

I. INTRODUCCION	11
II. MARCO TEÓRICO	17
2.1 ANTECEDENTES DEL PROBLEMA	17
2.2 Bases Teórico-Científicas	20
2.2.1. Sistema de Suministro de Agua	20
2.2.2. Funcionamiento	20
2.2.3. Demanda de Agua Potable	21
2.2.4. Análisis de la Demanda	22
2.2.4.1. Estimación de la Población	22
2.2.4.2. Población Servida de Agua Potable.....	23
2.2.4.3. Demanda Proyectada.....	23
2.2.4. Pronósticos.....	23
2.2.5. Métodos de Pronóstico	24
2.2.6. Medición de Errores	24
2.2.7. Redes Neuronales Artificiales:	25
2.2.7.1. Definición	25
2.2.7.2. Elementos de una Red Neuronal Artificial.....	26
2.2.7.3. Modos de Operación: Aprendizaje y Recuerdo	28
2.2.7.4. Aplicaciones de las RNA.....	32
2.2.7.5. Arquitecturas.....	33
2.2.7.6. Algoritmo de Retropropagación (BP).....	34
2.2.7.7. Aplicación a los Pronósticos.....	35
III. MATERIALES Y MÉTODOS	37
3.1 Diseño Investigación	37
3.1.1. Tipo de Investigación	37
3.2. Variables	37
3.3. Indicadores.....	37
3.4. Población, Muestra y Muestreo.....	38
3.4.1. Población	38
3.4.2. Muestra	38
3.5. Métodos, técnicas e instrumentos de recolección de datos.....	38
3.5.1. Métodos de Pronóstico	38

3.5.2. Técnicas.....	39
3.5.3. Herramientas.....	39
3.5.4. Fuente	39
3.6. Metodología	39
IV. RESULTADOS.....	41
4.1. Metodología para Redes Neuronales de (Kaastra y Boyd 1996).....	41
4.1.2. Diseñar Modelo de RNAs.....	41
4.1.2.1. Selección de la Variable.....	41
4.1.2.2. Recolección de Datos.....	41
4.1.2.3. Pre-procesamiento de Datos	42
4.1.2.4. Definición de Conjunto de Entrenamiento, Validación y Prueba	42
4.1.2.5. Selección de la Arquitectura de Redes Neuronales	43
4.1.2.6. Criterios de Evaluación.....	46
4.1.2.7. Entrenamiento de la Red Neuronal.....	46
4.1.2.8. Implementar Modelo de RNAs.	47
4.1.3. Pesos Calculados por el Algoritmo de Entrenamiento.	55
4.1.4. Cálculo del Error del Modelo de Redes Neuronales Artificiales.	56
4.1.5. Validación de la red con conjunto de prueba.....	57
V. DISCUSIÓN.....	60
5.1. Error de Pronóstico Acumulativo (Indicador).....	60
5.1.1. Error de predicción	61
5.2. Costo (Indicador).....	62
VI. PROPUESTA.....	66
Pronóstico a Futuro	66
VII. CONCLUSIONES.....	69
VIII. REFERENCIAS BIBLIOGRAFICAS	70
IX. ANEXO	72
Anexo I. Código Fuente de la Implementación del Modelo de Redes Neuronales Artificiales en MATLAB.....	72
Anexo II. Código Fuente del Cálculo de Nivel de Error del Modelo de Redes Neuronales Artificiales.....	73
Anexo III. Código Fuente del Cálculo del Pronóstico a Futuro del Modelo de Redes Neuronales Artificiales.....	74

ÍNDICE DE FIGURAS

Figura1. Funcionamiento del Sistema de Suministro de Agua.	21
Figura 2. Neurona Y Conexiones Sinápticas.....	26
Figura 3. Modelo Matemático para una Neurona.	27
Figura4 . Arquitectura Freed-forward.....	33
Figura 5. Arquitectura Recurrente	34
Figura 6. Procedimiento para el Diseño de un Modelo de Predicción basado en RNA (Kaastra y Boyd 1996).	36
Figura7. Función Sigmoidal (Rango de Salida entre -1 y +1).	45
Figura 8 . Modelo de la Red Neuronal.....	45
Figura 9. Editor de Control de MATLAB.	48
Figura 10. Editor de Variables (Representaciones de una Matriz 1 X 252).	49
Figura 11. Resultados del Entrenamiento (Modo Texto).	52
Figura 12. Resultados del Entrenamiento (Modo Gráfico).	53
Figura 13. Performance del Entrenamiento	54
Figura 14. Correlación entre Valores Reales y Pronósticos.....	55
Figura 15. Código para cargar la Data Entrenada	66
Figura16. Los últimos 12 meses.....	67
Figura17. Parámetros de la Red	67
Figura18. Código para la predicción del siguiente año.....	68

ÍNDICE DE TABLAS

Tabla1. Comparación entre la Predicción de Agua Potable y el Consumo Real	14
Tabla 2. Proyecciones de la Población administrativa	22
Tabla3. Proyecciones de la Población Servida.....	23
Tabla4. Proyección del Consumo de Agua Potable 2007-2008.....	23
Tabla5. Método General de Diseño de Pronóstico.....	24
Tabla6. Métricas para Medición del Error del Pronóstico	25
Tabla 7. Indicadores	37
Tabla8. Resumen de Datos de la Red Neuronal	48
Tabla 9. Pesos entre la capa de entrada y la capa oculta de la RNA estimada.....	56
Tabla 10. Pesos entre la capa oculta y la capa de salida de la RNA estimada.	56
Tabla 11. Predicciones de los Últimos 24 meses Aplicando el Modelo de RNAs.....	57
Tabla12. Consumo Real y Predicciones con Redes Neuronales.....	58
Tabla13Consumo Real y Predicciones con Redes Neuronales.....	58
Tabla 14. Consumo Real,Pronóstico de la Empresa y Pronósticode la RN.	61
Tabla 15. Error de Pronóstico de la Empresa y de la Red Neuronal.	62

Tabla 16. Costos Unitarios de Operación y Mantenimiento para el Sistema de Agua Potable.....	63
Tabla 17. Costos Unitarios de Operación y Mantenimiento para el Sistema de Agua Potable.....	64
Tabla 18. Predicción Futura.....	68

ÍNDICE DE GRÁFICOS

Gráfico 1. Consumo Real – Consumo Previsto (M ³).....	59
Gráfico 2. Consumo Real, Consumo Previsto por la Red y Consumo Previsto por la Empresa (M3).	61
Gráfico 3 . Proyección de los Costos de Operación y Mantenimiento para el Sistema de Agua Potable.	65

RESUMEN

La predicción de la cantidad demandada de agua potable es de gran importancia hoy en día para las corporaciones pertenecientes al sector saneamiento. Realiza esta predicción a mediano plazo es una necesidad vital para estas empresas, dado que de ello depende la satisfacción de los usuarios del servicio de saneamiento, los cuales se verían perjudicados en caso este servicio deje de trabajar en horas altamente críticas para ellos.

El presente trabajo pretende demostrar que el empleo de técnicas computacionales basadas en inteligencia artificial, como las Redes Neuronales Artificiales, reducen el nivel de error de las predicciones de la demanda de agua potable.

Palabras Claves: Predicción, demanda de agua potable, Redes Neuronales Artificiales.

ABSTRACT

Nowadays, the drinking water demand forecasting is a major issue to corporations that belong to the sanitation sector. Perform this forecasting at the medium term is a critical necessity for these enterprises, inasmuch as the sanitation service users satisfaction depends on it. Users will turn over damaged in case of this service will turn off in highly critical hours for them.

The present dissertation pretends to prove that by using computational techniques based on artificial intelligence, like Artificial Neural Networks, the drinking water demand forecasting error rate is reduced.

Keywords: Prediction, drinking water demand, Artificial Neural Networks.

I. INTRODUCCION

El agua, como tantas veces se insiste y es evidente, es un recurso limitado e indispensable para satisfacer las necesidades vitales de los seres vivos y un activo básico de primer orden para las actividades económicas, ecológicas y sociales. El agua debería ser un derecho humano y, sin embargo, millones de personas carecen de acceso al agua potable.

Aunque el 70% de la superficie del mundo está cubierta por agua, aproximadamente el 97% del agua del planeta es agua salina, en mares y océanos; apenas el 3% del agua total es agua dulce y de esa cantidad un poco más de dos terceras partes se encuentra congelada en los glaciares y casquetes helados en los polos y altas montañas. La tercera parte de los países en regiones con gran demanda de agua podrían enfrentarse a la escasez severa de agua en éste siglo, y para el 2025, dos tercios de la población mundial probablemente vivan en países con escasez moderada o severa.

La distribución de los recursos de agua dulce es muy desigual. Las zonas áridas y semiáridas del mundo constituyen el 40% de la masa terrestre, y estas disponen solamente del 2% de la precipitación mundial.(Toolbar s.f.)

El agua potable domiciliar es una de las más grandes e irrenunciables conquistas de la sociedad del bienestar, es difícil entender en su justo valor qué significa disponer a cualquier hora del día y de la noche de agua biológicamente potable, libre de gérmenes patógenos y químicamente inocua, sin embargo, la mayor parte de la Humanidad no dispone todavía de agua corriente en sus casas, tienen que buscarla a lugares lejanos que no reúnen las condiciones de potabilidad biológica requeridas y se convierte en transmisora universal de microorganismos patógenos.

La demanda de agua, en un sentido convencional, hace referencia a la necesidad de agua para uno o varios usos, el Perú puede considerarse un país privilegiado al poseer el 5% del agua potable a nivel mundial, sin embargo este recurso es manejado de manera inadecuada, generándose escasez y falta de suministro para aquellas zonas más alejadas y pobres del país. La demanda agrícola y ganadera asciende al 92%. La demanda urbana y rural con 39%.(Naturalle 2007)

En Lambayeque, se cubre el 67.2% de demanda poblacional urbana de agua potable.(Sunass 2009)

La elección e implementación de un método adecuado de pronósticos siempre ha sido un tema de gran importancia para las empresas. Un error significativo en el pronóstico de la demanda de agua podría generarle a una

empresa una pérdida monetaria demasiado grande. En un pronóstico erróneo aumentara los gastos para la Empresa.

La estimación del comportamiento futuro de algunas variables puede realizarse utilizando diversos métodos de pronóstico. Cada una de los métodos de proyección tiene una aplicación de carácter especial que hace de su selección un problema de decisión influido por diversos factores, como por ejemplo, la validez y disponibilidad de los datos históricos, la precisión deseada del pronóstico, el costo del procedimiento, los beneficios del resultado, los periodos futuros que se desee pronosticar y el tiempo disponible para hacer el estudio entre otros. (Sapag Chain y Sapag Chain 2003)

Como se mencionó anteriormente existen diversos métodos para proyectar la demanda de agua. Dentro de los mismos se encuentra el método de las Redes Neuronales Artificiales (RNA). Debido a su constitución y a sus fundamentos, las RNA presentan un gran número de características semejantes a las del cerebro humano. Por ejemplo, son capaces de aprender de la experiencia, de generalizar de casos anteriores a nuevos casos y de abstraer características esenciales a partir de entradas que representan información irrelevante. (Hilera González y Martínez Hernando 2000)

En los últimos años se ha planteado la posibilidad de aplicar las redes neuronales artificiales (RNA), ya que son metodologías novedosas que permiten hacer pronósticos donde hay cierto comportamiento no lineal y prometen un buen desempeño en este campo de investigación.

Desde el punto de vista de las aplicaciones de las redes neuronales, su principal ventaja frente a otras técnicas reside en el procesado paralelo, adaptativo y no lineal. En la actualidad las RNA son un campo muy consolidado en el que continuamente se desarrollan nuevos métodos y aplicaciones.

Dentro del ámbito de predicción de consumo de agua pueden encontrarse investigaciones como: "Aplicación de Redes Neuronales Artificiales para el pronóstico de la demanda de agua en zonas urbanas en Andalucía", este proyecto consiste en la elaboración de una red con mayor capacidad de predicción que otras técnicas, comprobando así que la utilización de las RNA para resolver problemas de pronóstico de series temporales es más eficiente que recurrir a modelos basados en técnicas clásicas. (Daza Sánchez 2008), entre otros.

La demanda de agua, en un sentido convencional, hace referencia a la necesidad de agua para uno o varios usos, y es, por tanto, la que se manifiesta a los niveles de precio actuales.

Esta investigación pone de relieve la necesidad de analizar en profundidad la demanda de agua, lo que motiva el interés en este trabajo ya que se realizará un análisis de la demanda de agua doméstica en la Ciudad de Lambayeque, partiendo de los datos de consumo de agua, correspondientes al periodo 1999-2010. Se propone una nueva técnica que resulta fundamental para predicciones de demanda con mayor precisión

A continuación se describe el actual proceso para la demanda de agua potable que tiene en cuenta la empresa EPSEL S.A

EPSEL S.A es una empresa dedicada a brindar servicios de saneamiento de agua potable, alcantarillado y servicios de laboratorio. En el Agua Potable, el sistema de producción comprende: Capacitación, almacenamiento y conducción de agua cruda, tratamiento y conducción de agua tratada. Y el Sistema de Distribución, que comprende: conexiones domiciliarias, incluye la medición, piletas, unidades sanitarias y otros. Alcantarillado y Sanitario, este servicio de recolección comprende: Conexiones domiciliarias, sumideros, redes y emisores. Y sistema de tratamiento y disposición de aguas servidas. Por último en el Servicio de Laboratorio comprende: Control de calidad de agua potable, control de aguas residuales y razones de un servicio de calidad.

La empresa, actualmente, realiza el proceso de la demanda de agua, a través de un estudio, en el cual se estima la demanda cada 20 años. Este estudio es elaborado por el Área de Planeamiento, que contiene el análisis de la propuesta presentada por EPSEL S.A. sobre la estructura tarifaria para los servicios de saneamiento (demanda de agua). Los parámetros utilizados para el cálculo de la demanda de agua potable son los siguientes: población, conexiones, y volúmenes demandados o consumo de agua, nivel de medición y el porcentaje de agua no contabilizada.

La demanda por el servicio de agua potable está definida por el volumen de agua que los distintos grupos de consumidores están dispuestos a consumir y pagar. Para tal efecto a partir de la estimación de la población administrada se definirán los niveles de cobertura del servicio de agua potable, estimando la población efectivamente servida.

A partir de la determinación de la población servida, se realiza la estimación del número de conexiones por cada categoría de usuario, lo cual dado el volumen requerido por cada grupo de usuarios, determinará la demanda por el servicio de agua potable que enfrentará la empresa en los próximos años.

Las técnicas estadísticas que emplea la empresa EPSEL S.A para la predicción del consumo de agua potable demandada, se basa en la población, teniendo como base la tasa de crecimiento anual de 1.3 % dado por el INEI. La fórmula del cálculo de la predicción de agua lo realizan en una hoja de cálculo de Excel aplicando fórmulas tradicionales, tal como se ha explicado anteriormente.

PRODUCCIÓN Y CONSUMO DE AGUA POTABLE 1990-2001			
PERIODO	PREDICCIÓN DE AGUA POTABLE M3	PERIODO	CONSUMO DE AGUA POTABLE M3
1990	1857485.0	1990	1206182.0
1991	1881648.0	1991	1281465.0
1992	1906103.0	1992	1308783.0
1993	1930886.5	1993	1343684.0
1994	1955962.0	1994	1367290.0
1995	1981402.5	1995	1358100.0
1996	2007171.5	1996	1375473.0
1997	2033269.0	1997	1383540.0
1998	2059695.0	1998	1424768.0
1999	2086449.5	1999	1409690.0
2000	2113569.0	2000	1450390.0
2001	2141053.5	2001	1476802.0

Fuente: EPSEL S.A

Tabla1. Comparación entre la Predicción de Agua Potable y el Consumo Real.

La tabla 1, muestra la imprecisión que realiza la empresa al predecir el consumo de agua potable, lo que ocasiona gastos innecesarios en la ampliación del sistema de agua potable, como insumos químicos para potabilizar el agua, materiales para el mantenimiento y reparaciones, control de calidad y herramientas para el desarrollo de los trabajos de operación y mantenimiento, etc. Esta falta de precisión, por parte de la Empresa ocasiona problemas que pueden causar graves daños en el desempeño de la empresa a mediano y a largo plazo; como la mala proyección de los costos de operación y mantenimiento del Sistema de agua potable, el cual va cambiando año tras año en un promedio aproximado de S/.127,883.90 nuevos soles. Con una buena predicción estos gastos se podrían reducir.

Teniendo en cuenta la falta de precisión en el pronóstico del consumo de agua potable, y el papel predominante que tiene el agua para la vida y la

importancia que tiene de disponer de pronósticos precisos, se ha planteado el siguiente problema de investigación:

¿De qué manera se puede mejorar el pronóstico de la demanda de agua en la empresa EPSEL S.A?

Para la cual planteamos la siguiente hipótesis: “A través de la aplicación de una Red neuronal Artificial se mejorará el pronóstico de la demanda de agua en la empresa EPSEL S.A”

El objetivo general de la tesis es desarrollar un modelo de RNA para el pronóstico del consumo de agua a mediano plazo, que permita mejorar el proceso de predicción de la demanda de agua en la empresa EPSEL S.A.

Objetivos Específicos:

- Proponer un modelo de predicción de agua basado en RNA.
- Reducir el error de estimación de la demanda de agua.
- Reducir los gastos relacionados con el abastecimiento de agua para la ciudad de Lambayeque.

Esta investigación se justifica en lo científico ya que permitirá establecer una comparación formal entre dos técnicas de predicción. Los resultados obtenidos probarán el predominio de las RNA aplicadas al pronóstico sobre los métodos clásicos que utilizan actualmente. La elección e implementación de un método adecuado de pronósticos siempre ha sido un tema de gran importancia para las empresas. Un error significativo en el pronóstico de la demanda de agua podría generarle a una empresa una pérdida monetaria demasiado grande. En este caso, el pronóstico erróneo disminuye los beneficios de la empresa. Es por ello que mediante las RNA se podrá pronosticar la demanda de agua en la empresa EPSEL S.A.

Se justifica económicamente, ya que los costos de inversión para este proyecto son bajos, gracias a la tecnología de bajo costo utilizada. La empresa reducirá el tiempo para pronosticar la demanda de agua y de esta manera logrará incrementar su rentabilidad, logrando así que la empresa EPSEL S.A. sea líder en el mercado.

Con respecto a lo tecnológico, se justifica ya que hoy en día el agua, como tantas veces se insiste y es evidente, es un recurso limitado e indispensable para satisfacer las necesidades vitales de los seres vivos y un activo básico de primer orden para las actividades económicas, ecológicas y sociales. Donde este proyecto permitirá que la empresa utilice sus recursos tecnológicos de

forma más eficiente. De la misma forma, se beneficiará de las ventajas del uso del enfoque tecnológico de Redes Neuronales y del software Matlab aplicados a los pronósticos de demanda de agua potable.

Y por último con respecto a lo social, mediante los resultados arrojados, afectaran a la empresa positivamente al ofrecer un mejor servicio de calidad, fiabilidad y seguridad beneficiando de esta manera a todos sus clientes y sociedad en general.

II. MARCO TEÓRICO

2.1 ANTECEDENTES DEL PROBLEMA

Antecedente # 01: *Predicción de la Demanda de agua con Redes Neuronales Artificiales en zonas urbanas de Andalucía.*(Daza Sánchez 2008)

La presente tesis doctoral consiste en la elaboración de una red con mayor capacidad de predicción que otras técnicas, se diseñan numerosas redes combinando arquitecturas, funciones de activación, retardos de entrada y nodos en la capa oculta, y de ellas se selecciona una red neuronal de tipo perceptrones multicapa. Estas redes están formadas por varias neuronas o nodos conectados entre sí y organizados por capas; cada capa se compone de una o más neuronas, donde su número es independiente del número de nodos de cualquier otra capa. La arquitectura de la red se define por el número de capas, el número de nodos por capa y las conexiones entre estos.

La RNA seleccionada en esta tesis es de tres capas: una de entrada, una oculta y una de salida. Donde cada neurona de la red procesa la información recibida por sus entradas y produce una respuesta o activación que se propaga, a través de las conexiones correspondientes.

Se logró determinar que la utilización de las redes neuronales para resolver problemas de pronóstico de series temporales es más eficiente que recurrir a modelos basados en técnicas clásicas.

Antecedentes #02: *Pronóstico de Ventas usando Redes Neuronales.*(Toro Ocampo, Mejía Giraldo y Salazar Isaza 2004)

Este artículo muestra algunas técnicas de pronóstico tradicionales que son comparadas con el comportamiento de una Red Neuronal Artificial. Se usa un problema de pronóstico de ventas como referencia. Para obtener un buen comportamiento de la Red Neuronal se aplican diferentes funciones de transferencia así como diferentes criterios de parada. Son obtenidos excelentes resultados que muestran el potencial de esta técnica artificial. Ya que su objetivo es proponer una red neuronal que sea alimentada con los niveles de ventas de períodos de tiempo anteriores al que se desea estimar y con un parámetro que informe el período que se desea estimar dado que el problema que se está abordando es de tipo estacional. Y presentar una aplicación de la teoría de redes neuronales en pronósticos de producción mensuales basada en datos de ventas mensuales a fin de aumentar las alternativas. Porque el uso óptimo de los recursos, la creciente demanda de una mayor variedad de productos, entre otros, obliga a los fabricantes a

realizar programas de producción más estrictos y flexibles para poder maximizar el uso de los costosos equipos de producción, mano de obra, inversiones en materias primas, de forma que se cumplan las fechas de entrega a los clientes finales minimizando los costos. Para realizar eficientes programas de producción que atiendan la demanda sin incurrir en excesivos costos de capital, es deseable un adecuado manejo de la información de ventas con el fin de realizar pronósticos con buen nivel de confianza para todos los departamentos de la empresa.

Es por ello que se propone una red neuronal que sea alimentada con los niveles de ventas de períodos de tiempo anteriores al que se desea estimar y con un parámetro que informe el período que se desea estimar dado que el problema que se está abordando es de tipo estacional. Con toda esta información la red debe pronosticar un nivel de ventas partiendo del comportamiento histórico de dicha variable.

Las redes neuronales se “aprenden” la dinámica del fenómeno y tienen en cuenta el tiempo como una variable que afecta el ciclo de las ventas. De esta forma ofrecen resultados más robustos y cercanos a la realidad.

Antecedente #03: *Inteligencia Artificial en pronóstico de demanda de energía Eléctrica: Una aplicación en optimización de recursos energéticos*(Sarmiento Maldonado y Villa Acevedo 2008)

Este artículo consiste en una aplicación actual de las redes neuronales en la tarea de pronóstico de demanda eléctrica en Colombia. Se utilizan redes del tipo MultiLayerPerceptron (MLP) con algoritmos de entrenamiento Backpropagation, y Radial Basic Function (RBF). La información disponible de tiempo y valor de demanda horaria en megavatios (MW) es organizada de tal forma que se plantea la tarea de pronóstico como un problema de clasificación de información donde las anteriores redes han evidenciado un buen desempeño. La tarea de organización de la información, entrenamiento, validación y pronóstico son desarrolladas con programación en Matlab.

Antecedente #04: *La demanda de telefonía fija y móvil: Una aplicación de redes neuronales artificiales*(Coca Carasila y Villagómez Méndez 2009)

Este artículo nos dice que el catalogar a un determinado bien o servicio como una de las «extensiones de los sentidos y las funciones humanas de la vista, el oído y el tacto», muestra la importancia del papel que desempeñan en nuestras vidas, muestra el desarrollo que alcanzaron impulsadas por las mismas necesidades de los seres humanos, muestra un mercado dinámico e

importante. El servicio de telefonía móvil o celular es el desencadenador de estas expresiones y además tema de comentario, investigación y preocupación de comunidades científicas y organismos internacionales como el World Economic Forum de Davos.

Este artículo se adentra en este mercado, en el que intervienen activamente la demanda y la oferta de servicios y equipos que no terminan de innovar, buscando satisfacer las, cada vez más, exigentes necesidades y deseos de los usuarios. Buscan analizar, específicamente la demanda de telefonía fija y móvil, pretendiendo dilucidar una coyuntura particular y un futuro inmediato e incierto, sobre todo para el participante que sobrelleva las consecuencias, la telefonía fija. Después de una exhaustiva revisión de los datos históricos, se determinó aplicar el Perceptron Multicapa (MLP). El MLP, además de ser un aproximador universal de funciones, ha demostrado suficiencia para la predicción de series de tiempo y regresiones, con excelente resultados en la medición de la demanda de gas y electricidad.

Antecedente #05: *“Redes Neuronales Artificiales y Regresión de Series Temporales para la Predicción de la cantidad demandada de Energía Eléctrica por Empresas del Sector Energético Eléctrico”* (Narváez García 2009)

La presente tesis consiste en demostrar que el empleo de técnicas computacionales basadas en inteligencia artificial, con Redes Neuronales Artificiales, reducen el nivel de error de las predicciones de la demanda de energía eléctrica en comparación al método estadístico de Regresión de Series Temporales, además de las implicancias que esto conlleva.

Los antecedentes 1, 3, 4 y 5 utilizan la misma metodología que es el Perceptron Multicapa (MLP), con algoritmos de entrenamiento Backpropagation. Estas redes están formadas por varias neuronas o nodos conectados entre sí y organizados por capas; cada capa se compone de una o más neuronas, donde su número es independiente del número de nodos de cualquier otra capa. Y además de ser un aproximador universal de funciones, ha demostrado suficiencia para la predicción de series de tiempo y regresiones.

La metodología del Perceptron Multicapa (MLP), es una de las arquitecturas más utilizadas en la resolución del problema, debido a su capacidad como aproximador universal de funciones y a su fácil uso y aplicabilidad. Y será de gran ayuda para realizar un estudio sobre el pronóstico del consumo de agua en la ciudad de Lambayeque.

2.2 Bases Teórico-Científicas

2.2.1. Sistema de Suministro de Agua

El agua potable es uno de los recursos naturales más importantes utilizados por el hombre. El agua juega un papel fundamental en la localización, crecimiento y función de las comunidades; su afirmación sigue estando completamente vigente en la actualidad. Esto ha hecho que la gestión del agua haya sido siempre un aspecto importante en cualquier sociedad.(Marshall 1999)

El sistema de suministro de agua por gravedad con tratamiento es un conjunto de estructuras para llevar el agua a la población mediante conexiones domiciliarias. Consta de diferentes procesos físicos y químicos necesarios para hacer posible que el agua sea apta para el consumo humano, reduciendo y eliminando bacterias, sustancias venenosas, turbidez, olor, sabor, etc. Se dice sistema por gravedad porque el agua cae por su propio peso, desde la captación al reservorio y de allí a las conexiones domiciliarias.

El estado físico del sistema de suministro de agua frecuentemente es pasado por alto en la estimación de la demanda y sin embargo puede tener una influencia muy importante debido a las pérdidas en las tuberías. Las pérdidas en la red constituyen típicamente en torno al 50 % del agua introducida en la red y no facturada. La cantidad de pérdida depende de cada área metropolitana en función de la edad y el estado de las infraestructuras del diseño del sistema, sin embargo las medidas de control de fugas mediante control de presión, sondeo de rutas y otras han demostrado que pueden ser el método más efectivo e inmediato para la reducción de la demanda(White, y otros 2003)

2.2.2. Funcionamiento

El sistema de suministro de agua funciona de la siguiente manera: Inicialmente, el agua potable se genera mediante la captación de agua cruda superficial, del canal San Romualdo, y tratamiento realizado en la Planta N°2. El agua producida es bombeada a un reservorio de carga de 100 m³ ubicado en la misma planta desde el cual mediante una tubería de 14" y 16" conduce el agua hacia las tres cisternas de almacenamiento de 1,674 m³ de capacidad ubicadas en la antigua planta N°1.

Posteriormente, el agua es clorada y bombeada hacia los reservorios elevados de 280 m³ y 3000 m³ de capacidad, ambos reservorio se encuentran ubicados en la ciudad en las instalaciones de la antigua planta de tratamiento. Este proceso es ilustrado en la figura 1.

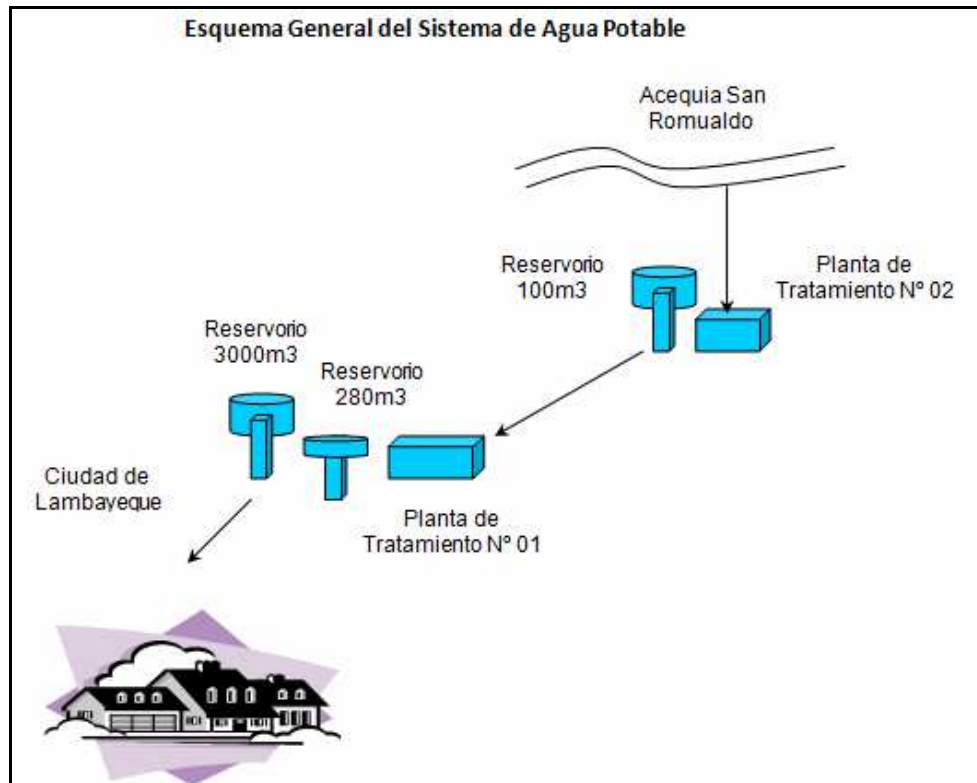


Figura1. Funcionamiento del Sistema de Suministro de Agua.

Los reservorios abastecen de agua a la población de Lambayeque mediante una línea de aducción de 16" de Asbesto Cemento.

Las redes de distribución están conformadas por tuberías que van desde las 3" hasta 14" de diámetro, en su mayoría de Asbesto Cemento.

2.2.3. Demanda de Agua Potable

Para determinar la demanda actual, se estudian los grupos de consumidores de características homogéneas, considerando el tipo de consumo (comercial, residencial e industrial), la cobertura actual del servicio y la facturación promedio mensual para cada grupo definido. Se analiza y estima la demanda futura por grupo de consumidor, en base al crecimiento esperado de la población, a la cobertura proyectada y a la dotación de consumo esperada. Los consumos

comerciales se proyectan proporcionalmente al crecimiento de los residenciales, y los consumos industriales se proyectan de acuerdo al crecimiento del rubro. (Galán Ordax 2007)

2.2.4. Análisis de la Demanda

El estudio que realiza EPSEL S.A para determinar el consumo de agua se proyecta a 20 años. De acuerdo al diagnóstico realizado, se plantea ejecutar el proyecto mediante la estimación de la población.

2.2.4.1. Estimación de la Población

Los resultados de la proyección de la población se utilizan para obtener las proyecciones de la demanda de los servicios de agua potable. La población inicial será la población obtenida por el Censo realizado por el INEI (Instituto Nacional de Estadística e Informática). Y la tasa de crecimiento asumida es de 1.3%.

En la tabla 2 se presenta la proyección de la población urbana para la ciudad de Lambayeque.

AÑO	POBLACIÓN
1993	52,901
1994	53,588
1995	54,285
1996	54,991
1997	55,706
1998	56,430
1999	57,163
2000	57,906
2001	58,659
2002	59,422
2003	60,194
2004	60,977
2005	61,770
2006	62,573
2007	63,386

Fuente: EPSEL S.A.

Tabla 2. Proyecciones de la Población administrativa

2.2.4.2. Población Servida de Agua Potable

Para el cálculo de la población servida de la zona de estudio se ha considerado la población administrada, con una cobertura del 100%. En la Tabla 3, se muestra la proyección de la población servida de agua potable.

Años	Población	Cobertura	Población servida
		%	(hab)
2007	63,386	100.00	63,386
2008	64,210	100.00	64,210

Fuente: EPSEL S.A

Tabla3. Proyecciones de la Población Servida

2.2.4.3. Demanda Proyectada

La demanda se proyecta a lo largo del horizonte del proyecto que realiza la Empresa. En la Tabla 4, se muestra la proyección de la demanda de agua del año 2007 y 2008.

Año	Consumo de Agua Potable		
	Litro	m3	Litro
	día	año	segundo
2007	6,338,600.0	2,313,589.0	73.36
2008	6,421,000.0	2,343,665.0	74.32

Fuente: EPSEL S.A

Tabla4. Proyección del Consumo de Agua Potable 2007-2008

2.2.4. Pronósticos

Pronosticar es muy importante en muchos tipos de empresas, ya que las predicciones de hechos futuros se pueden incorporar al proceso de toma de decisiones. (Bowerman, O' Connell y Koehler 2007)

Al predecir hechos que ocurrirán en el futuro, quien elabora los pronósticos debe confiar en la información de los hechos que han ocurrido en el pasado y que le interesan. Es decir, con objeto de preparar un pronóstico, quien elabora el pronóstico tiene que analizar

la información anterior y basar la predicción en los resultados del análisis. Las personas que elaboran los pronósticos utilizan los datos pasados de la manera siguiente: Primero, dicha persona analiza los datos para poder identificar un patrón que se pueda utilizar para describirlo. Esta es la forma básica para construir pronósticos basándonos en que el patrón permanecerá igual en el futuro. Si el patrón que se identificó no persiste en el futuro, entonces las predicciones no serán exactas, razón por la cual lo más aconsejable es prever cuando será el cambio de dicho patrón y hacer los cambios necesarios a nuestro sistema de pronósticos mediante la aplicación de otras técnicas para otros tipos de patrones.

2.2.5. Métodos de Pronóstico

Se podría hablar de la existencia de un método general para diseñar modelos de pronósticos el cual se describe en la tabla 5, aplicable a cualquier caso.

N°	Pasos	Descripción
1	Identificar el Horizonte Temporal	Es decir el corto, mediano y largo plazo.
2	Observar los Hechos	Identificar las variables a predecir y sus dependencias
3	Seleccionar una Técnica	Seleccionar las técnicas que más se ajustan al tipo de problemas presentado.
4	Aplicar Técnicas Seleccionadas	Implementar el pronóstico con la técnica seleccionada.

Tabla5. Método General de Diseño de Pronóstico

2.2.6. Medición de Errores

La presencia del componente irregular en una serie de tiempo que representa fluctuaciones impredecibles en los datos, quiere decir que es de esperarse algún error en el pronóstico. Para saber que tan eficaz es un pronóstico, es decir que tan lejos está del valor real de la variable a predecir, existen ciertas métricas estándares que nos permiten medir los errores en los pronósticos. En la tabla 6 se define El Error Cuadrático Medio y el Error del Pronóstico Acumulativo, las cuales

consideran la suma de todos los errores de predicción en el tiempo, arrojando un resultado total.

Métrica	Definición
Error Cuadrático Medio	$ECM = \frac{\sum e_t^2}{n}$ <p>e_t = Error del Pronóstico n = Cantidad total de tiempos observados</p>
Error de Pronóstico Acumulativo	$EPA = \sum_{t=1}^n e_t$ <p>e_t = Error del Pronóstico n = Cantidad total de tiempos observados</p>

Tabla6. Métricas para Medición del Error del Pronóstico

2.2.7. Redes Neuronales Artificiales:

2.2.7.1. Definición

Las redes neuronales son modelos artificiales que buscan reproducir una de las características propias de los seres humanos: la capacidad de memorizar y asociar hechos, por tanto, son un modelo simplificado del cerebro humano.

El cerebro es capaz de procesar a gran velocidad grandes cantidades de información procedentes de los sentidos, combinarla o compararla con la información almacenada y dar respuestas adecuadas incluso en situaciones nuevas. Lo más interesante es su capacidad de aprender a representar la información necesaria para desarrollar tales habilidades sin instrucciones explícitas para ello.

Una de las principales preocupaciones de los investigadores a lo largo de la historia es construir máquinas computacionales que sean capaces de realizar procesos con cierta inteligencia.

El cerebro humano contiene aproximadamente 12 billones de células nerviosas o *neuronas*, que poseen la característica de poder comunicarse. Cada una de las neuronas tiene entre 5.600 y 60.000 *conexiones dendríticas* provenientes de otras neuronas (Figura2), que están conectadas a la membrana de la neurona y transportan los impulsos enviados desde otras neuronas. Cada neurona tiene una salida llamada *axón*, y el contacto de éste con una dendrita se lleva a cabo a través de la sinapsis. El axón y las dendritas transmiten la señal en una única dirección.

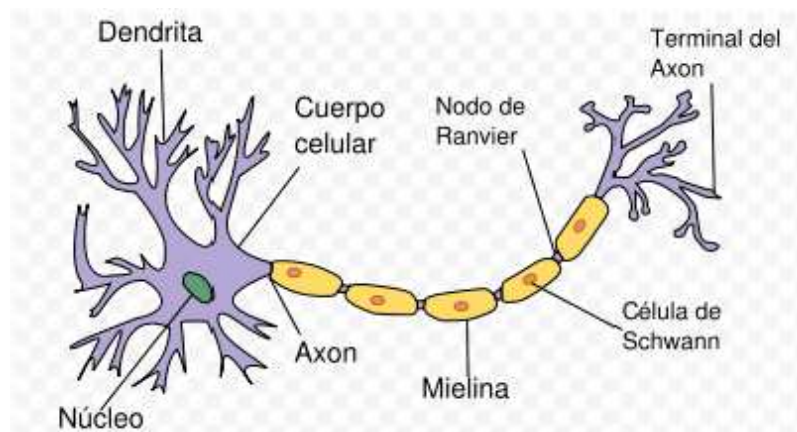


Figura 2. Neurona Y Conexiones Sinápticas

Puede decirse que el cerebro humano está compuesto de decenas de billones de neuronas interconectadas entre sí formando circuitos o redes que desarrollan funciones específicas.

2.2.7.2. Elementos de una Red Neuronal Artificial.

Según la literatura internacional (Dijk, V.; y otros. 2001) la utilización de modelos no lineales de series temporales ha aumentado considerablemente en los últimos años y, en particular, las redes neuronales artificiales (RNA), que han mostrado una gran capacidad predictiva en múltiples campos, aunque históricamente, el dominio lo han tenido las especificaciones lineales, por haber presentado un mayor desarrollo teórico y una mayor difusión³⁰. Son muchos los autores que consideran los modelos de redes neuronales artificiales como una clase de modelos matemáticos no lineales que tratan de simular la organización y el funcionamiento de las neuronas biológicas¹.

¹La estructura cerebral está compuesta de un elevado número de neuronas interconectadas entre sí. Una neurona recibe señales de otras a través de receptores de estructura fina, llamados dendritas, estas señales

Dentro del campo de la inteligencia artificial tratan de modelizar la capacidad de aprendizaje del cerebro humano.

La neurona artificial, denominada así por analogía con el cerebro humano, es el elemento básico de la red, a partir de un vector de entrada procedente del exterior o de otras neuronas proporciona una única respuesta o salida (véase Figura3). Cada neurona se modeliza mediante unidades de proceso y se compone de un núcleo encargado de los cálculos, una red o vector de conexiones de entrada y una salida. Una red neuronal es una estructura que establece la topología de la red, la agrupación de las neuronas en unidades estructurales llamadas capas, pudiéndose hablar de redes de una, de dos o de más de dos capas, de manera que cada una de ellas alimenta a todas las neuronas de la capa siguiente, y la intensidad de las distintas interconexiones. Las neuronas están conectadas entre sí por unos enlaces representados por unos pesos sinápticos o ponderaciones que reflejan la intensidad e importancia del enlace y que se ajustan conforme a un algoritmo de aprendizaje especificado para minimizar la función de coste determinada en base al error en que se incurre comparando los datos obtenidos por la red con los datos reales. La salida de cada neurona es el resultado que suministra una función, denominada función de activación, a la suma de sus entradas multiplicadas por los pesos sinápticos.

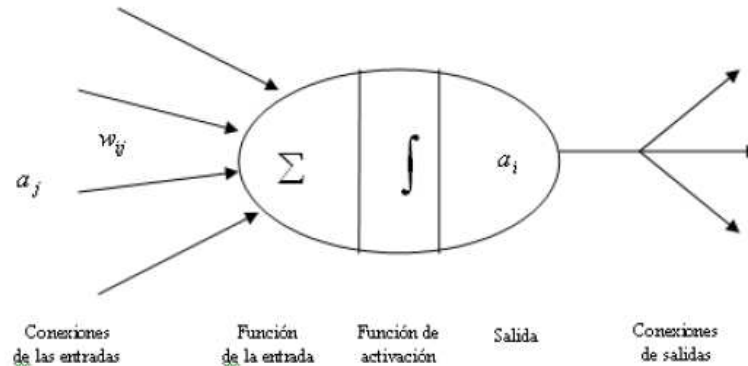


Figura 3. Modelo Matemático para una Neurona.

Se distinguen tres tipos de capas: de entrada, de salida y ocultas. La capa de entrada está formada por las neuronas que reciben datos del entorno, almacenando la información suministrada a la red. En la capa de salida están las neuronas que ofrecen la respuesta de la red

circulan por un canal o medio llamado axón que se divide en cientos de ramas. Una neurona se activa si la señal total recibida de todas las neuronas que conectan con ella excede un cierto nivel, conocido como nivel de activación. Durante el desarrollo de un ser vivo, el cerebro se modela, ya que existen muchas cualidades que se adquieren por la influencia del medio externo, este sería el aprendizaje.

neuronal hacia el exterior y las capas ocultas, que no tienen conexión directa con el entorno, sino que son interna a la red, son las encargadas de extraer, procesar y memorizar la información, el número de capas ocultas puede estar entre cero y un número elevado. Las neuronas de las capas ocultas pueden estar interconectadas de distintas maneras, lo que determina, junto con su número las distintas tipologías de redes neuronales.

2.2.7.3. Modos de Operación: Aprendizaje y Recuerdo

Las RNA están concebidas para presentar características análogas a las del funcionamiento del cerebro humano. La capacidad de aprender es una de las características de la red neuronal que la diferencian de la programación tradicional y que imita, de forma semejante y en pequeña escala, el funcionamiento del cerebro. Son capaces de aprender de la experiencia, de generalizar de casos anteriores a nuevos casos, de abstraer características esenciales a partir de entradas que representan información irrelevante, etc.

Como la red neuronal no posee ningún tipo de conocimiento útil almacenado, es necesario entrenarla, o lo que es lo mismo enseñarla para que sea capaz de realizar una determinada tarea. Se distinguen dos modos de operación en las redes neuronales: el modo aprendizaje o entrenamiento y el modo recuerdo o ejecución.

Una definición de aprendizaje enunciada muchos años antes de que surgieran las redes neuronales y que se puede aplicar al proceso de aprendizaje de estos sistemas, podría ser (Hilera González y Martínez Hernando 2000): La modificación del comportamiento inducido por la interacción con el entorno y como resultado de experiencias conducente al establecimiento de nuevos modelos de respuesta a estímulos externos. Una característica fundamental de las redes neuronales es que son sistemas entrenables, capaces de aprender a partir de un conjunto de patrones de entrenamiento.

En el proceso de aprendizaje de las RNAs se ajustan los parámetros libres de la red a partir de un proceso de interacción con el entorno que rodea la red, es de carácter secuencial (Chung-Ming y White 1994), y adaptativo, de tal manera, que la red adquiere conocimiento continuamente, como resultado de la acumulación de experiencias

ocurridas, similar al proceso de aprendizaje del cerebro humano en lo que se refiere a considerar los errores pasados.

Se puede considerar que el conocimiento se encuentra almacenado en los pesos de las conexiones entre neuronas, al igual que biológicamente el conocimiento parece estar más relacionado con las sinapsis o conexiones entre neuronas que con ellas mismas. Los pesos de las conexiones son las unidades de memoria de la red y sus valores representan el estado actual de conocimiento. El proceso de aprendizaje se entenderá finalizado cuando el valor de los pesos se mantenga estable.

Al construir un modelo neuronal se especifica una determinada arquitectura de red en la que los pesos sinápticos iniciales son nulos o aleatorios. Para que la red resulte operativa es necesario entrenarla, lo que constituye el modo aprendizaje. El entrenamiento o aprendizaje convencional modifica los pesos sinápticos en respuesta a una información de entrada o regla de aprendizaje, conseguida optimizando una función de error o coste, que valora la eficacia de la red. Se suelen considerar dos tipos de reglas: las que responden a un aprendizaje supervisado y las que lo hacen a un aprendizaje no supervisado.

Normalmente es un proceso iterativo, que actualiza los pesos reiteradamente hasta que la red neuronal llega al rendimiento deseado.

Se distinguen cuatro tipos de aprendizaje(Haykin 1999):

- **Aprendizaje supervisado.** La red tiene siempre disponible para cada patrón de entrenamiento de entrada, un patrón de salida, que representa la respuesta deseada para esa entrada y que se indica a la red, por parte de un agente externo o supervisor. La red ajusta sus pesos hasta que su salida tiende a ser la deseada y minimice el error de salida de la red, posteriormente será evaluada por el agente externo o supervisor, que controla el proceso de aprendizaje de la red. El aprendizaje finaliza cuando se ha obtenido una secuencia de modelos que tiene asociado un vector de salida.
- **Aprendizaje no supervisado.** Se presenta a la red un conjunto de patrones para los que no se parte de una respuesta de salida

deseada de la red, sino que es ésta la que reconoce regularidades presentes en los datos de entrada. La red modifica los pesos para poder asociar a la misma unidad de salida los vectores de entrada más similares.

Mediante el aprendizaje (no supervisado) la red tiene que descubrir las regularidades presentes en los datos para poder clasificarlos en categorías no determinadas a priori (Bonilla Musoles y Puertas Medina 1997). La red no recibe ninguna información del entorno que le indique si la salida generada es o no correcta.

- **Aprendizaje híbrido.** Como su nombre indica coexisten en la red los dos tipos de aprendizaje anteriores, el supervisado y el no supervisado, generalmente en distintas capas de neuronas. El modelo de contra-propagación y las RBF son ejemplos de redes que hacen uso de este tipo de aprendizaje.
- **Aprendizaje reforzado.** Es similar a la idea del aprendizaje supervisado en cuanto que se puede definir un criterio de actuación de la red dependiendo de las salidas que origine, aunque en este caso no se suministra explícitamente la salida deseada y, por tanto, la evaluación del resultado de la red es global, se decidirá si está actuando bien o mal pero sin proporcionar detalles. En este caso no se puede determinar la fracción de error correspondiente a cada salida, por lo que son necesarios algoritmos de mayor complejidad y un conjunto de entrenamiento de mayor tamaño. En ocasiones se denomina aprendizaje por premio-castigo.

Una red neuronal tiene una buena capacidad de generalización si responde adecuadamente ante patrones no utilizados en su entrenamiento, es decir, fuera de la muestra en la cual recibió su entrenamiento.

El error de aprendizaje suele calcularse como el error de pronóstico acumulativo de los resultados de la red, además existe un error de generalización, que se puede cuantificar en base a un conjunto de patrones no coincidentes con los utilizados en el entrenamiento.

A partir de que el sistema ha sido entrenado, los pesos y la estructura quedan fijos y la red neuronal ya está dispuesta para procesar datos,

este modo de operación corresponde al modo recuerdo o de ejecución.

Es usual disponer de un conjunto de datos distintos a los utilizados para el entrenamiento, de los que se conoce la respuesta correcta, y usarlos como test, valorando si la red responde adecuadamente frente a datos distintos a los usados en el aprendizaje.

Si la red funciona bien se puede pasar a la fase de aplicación, por lo que se puede usar para obtener un resultado frente a datos totalmente nuevos para los que no se conoce la respuesta correcta. (Moutinho 2000)

Para el caso de las redes unidireccionales, donde no existe realimentación, las neuronas responden proporcionando directamente la salida del sistema, sin problemas de estabilidad, pero en las redes con realimentación, que son sistemas dinámicos no lineales, es más difícil dar una respuesta estable, por ello, algunos teoremas generales recogen las condiciones necesarias para asegurar la estabilidad de la respuesta, bajo determinadas condiciones, para numerosas redes neuronales.

Las redes neuronales son sujeto de sobre ajuste cuando hay demasiados parámetros en el modelo. Si se utilizan redes totalmente conectadas, la preocupación se centra en el número de capas ocultas y su tamaño, el enfoque más usual es probar varias y quedarnos con la mejor, por ello, se necesitan las técnicas de validación cruzada para elegir la arquitectura de la red que proporcione la mayor precisión de predicción en la validación.

Si se consideran redes que no están totalmente conectadas, es necesario utilizar un método de búsqueda efectivo a través del gran espacio de topologías de posibles conexiones, como es el algoritmo del daño cerebral óptimo (optimal brain damage) que comienza con una red totalmente conectada y va eliminando conexiones.

Las redes de una única capa poseen un algoritmo de aprendizaje más simple y eficiente, pero tienen un poder de expresividad muy limitado. Por otra parte, las redes multicapa son más difíciles de entrenar, pero son más expresivas, es decir pueden representar funciones no lineales generales. Pretendiendo proporcionar lo mejor de ambos modelos han surgido métodos de aprendizaje, relativamente nuevos, denominados máquinas de vectores soporte (SVMs) o máquinas núcleo, que usan un algoritmo de aprendizaje eficiente y pueden representar funciones no lineales complejas, aunque sus resultados son relativos.

2.2.7.4. Aplicaciones de las RNA

- **Clasificación:** puede identificar datos y clasificarlos en clases. A diferencia de las técnicas tradicionales puede generar para un problema concreto estructuras específicas de identificación de características.
- **Asociación:** la red funciona como un proceso de recuperación de un dato a partir de una entrada relacionada con el dato almacenado
- **Agrupamiento:** cuando no se produce una clasificación de los datos de entrada, es la red la que genera esta clasificación a medida que se va produciendo la señal.
- **Generalización:** las redes pueden detectar las relaciones subyacentes entre diferentes elementos, y por tanto, pueden extrapolar las relaciones numéricas presentes en los datos disponibles en situaciones similares.
- **Optimización:** cuando es necesario elegir la opción óptima entre diferentes alternativas, las redes proporcionan un procedimiento relativamente rápido para generar una solución sub-óptima que es suficientemente satisfactoria.
- **Predicción y control:** la flexibilidad de las redes neuronales las hace adecuadas para la predicción ya que el sistema va aprendiendo y puede seguir de manera versátil estos cambios.

Las RNA, en la actualidad se están utilizando principalmente en una doble dirección: predicción y clasificación. (Ineiaragón Torre y Al. 1997) Hay que hacer mención especial a uno de los “motores” en el desarrollo de las redes neuronales: la predicción en series temporales.

En problemas financieros donde existen dificultades de programación y mantenimiento en los sistemas expertos, se ha demostrado la capacidad de las redes neuronales para dar una respuesta satisfactoria a estos problemas, incluso en los que la información está incompleta o presenta algún fallo.

Existen numerosos estudios que analizan series financieras con redes neuronales como son los trabajos de (Figuroa De La Cruz, y otros 2007).

De las múltiples aplicaciones que pueden tener las redes neuronales en el campo económico, como son predicción de precios, segmentación de mercados, etc. Como son los trabajos de (Toro Ocampo, Mejía Giraldo y Salazar Isaza 2004), (Munir Andrés y Martha 2007) y (Lawrence 1997).

Dentro del ámbito de predicción de consumo de agua destacan las aportaciones de (Daza Sánchez 2008).

2.2.7.5. Arquitecturas.

Se pueden clasificar las arquitecturas de una red neuronal de dos maneras; la primera es según el número de capas de neuronas que presenta la red (puede ser unicapa o multicapa) y la otra forma, es de acuerdo a los patrones de conexión entre las neuronas y la propagación de los datos. En base a esta última clasificación podemos distinguir dos tipos e arquitecturas.

A. Redes Freed-forward.

En este tipo de redes, el flujo de datos desde la entrada hasta la salida es estrictamente hacia delante. El procesamiento de datos puede extenderse sobre múltiples capas de neuronas, pero no existen conexiones hacia atrás, ósea conexiones que se extienden desde la salida de una neurona a la entrada de otra en la misma capa o capa anterior.

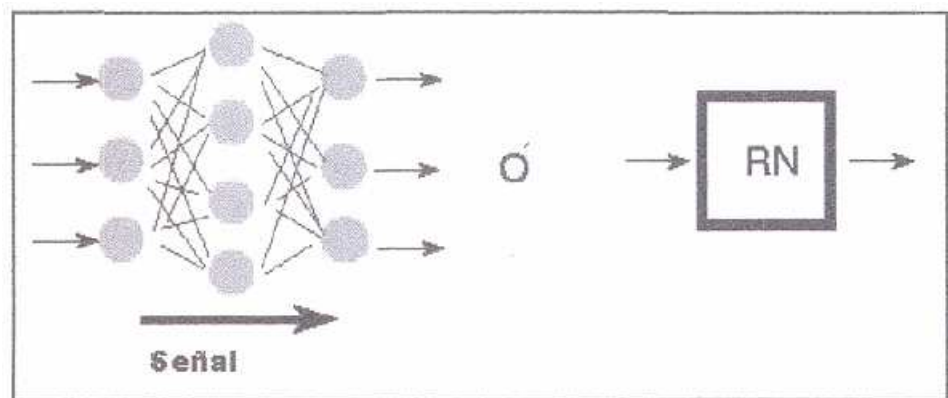


Figura4 . Arquitectura Freed-forward.

En la figura 4 se aprecia la dinámica de esta arquitectura y entre sus principales representantes podríamos nombrar al perceptron monocapa, el perceptron multicapa que es muy útil

para aproximar relaciones no lineales entre datos de entrada y de salida, Adeline, entre otras.

B. Redes Recurrentes

Estas redes a diferencia de las primeras, sí contienen conexiones hacia atrás.

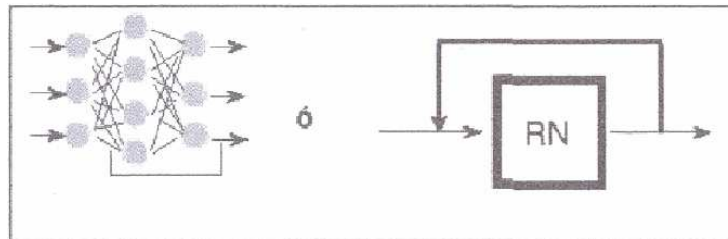


Figura 5. Arquitectura Recurrente

En la figura 5 se puede observar como el flujo de datos se desplaza de la salida una neurona ubicada en capa posterior a la entrada de una neurona ubicada en una capa anterior. Entre los representantes de esta arquitectura tenemos las redes de base radical, redes de Hopfield y los mapas Auto-Organizativos de Kohonen, éstos últimos de mucha utilidad para clasificación de datos.

2.2.7.6. Algoritmo de Retropropagación (BP).

El algoritmo de aprendizaje de retropropagación, consiste básicamente en propagar el error hacia atrás, desde la capa de salida hasta la de entrada, permitiendo así la adaptación de los pesos con el fin de reducir dicho error (Hilera González y Martínez Hernando 2000).

En forma simplificada, el funcionamiento de una red “backpropagation” consiste en el aprendizaje de un conjunto predefinido de pares de entradas y salidas dados como ejemplos, empleando un ciclo de propagación-adaptación de dos fases:

- En la primera fase, o paso hacia delante, se aplica un primer patrón como estímulo para la capa de entrada de la red, éste se va propagando a través de las capas siguientes para generar la salida, la cual proporciona el valor del error al compararse con la que se desea obtener.

- En la siguiente fase o paso hacia atrás, estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa oculta intermedia que contribuyan directamente a la salida, recibiendo el porcentaje del error aproximado a la participación de las mismas en la salida original.

Este proceso se repite hacia atrás, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aporte relativo al error total. Basándose en esta información recibida, se reajustan todos los pesos de conexión, de manera que la siguiente vez que se presente el mismo patrón disminuya la diferencia entre la salida calculada y la deseada.

La importancia de la red backpropagation consiste en su capacidad de auto adaptar los pesos de las neuronas de las capas intermedias para aprender la relación que existe entre un conjunto de patrones dados como ejemplos y sus salidas correspondientes.

2.2.7.7. Aplicación a los Pronósticos

El verdadero poder de las redes neuronales recae en su habilidad para capturar relaciones no lineales inherentes en los datos, mientras que los modelos lineales describen una relación lineal entre observaciones actuales y posteriores, las redes neuronales describen una relación no lineal entre las dos. Debido a este motivo son de gran utilidad en el área de las predicciones.

De acuerdo con los investigadores (Dorffner 2008), el Perceptron Multicapa es el tipo de red neuronal artificial más utilizado y con los mejores resultados para este tipo de aplicación (predicciones). Es por ello que en la literatura (Kaastra y Boyd 1996), se propone un modelo para diseñar una red neuronal artificial con aplicaciones a predicciones. El procedimiento para diseñar el modelo consta de 8 pasos (figura 6).

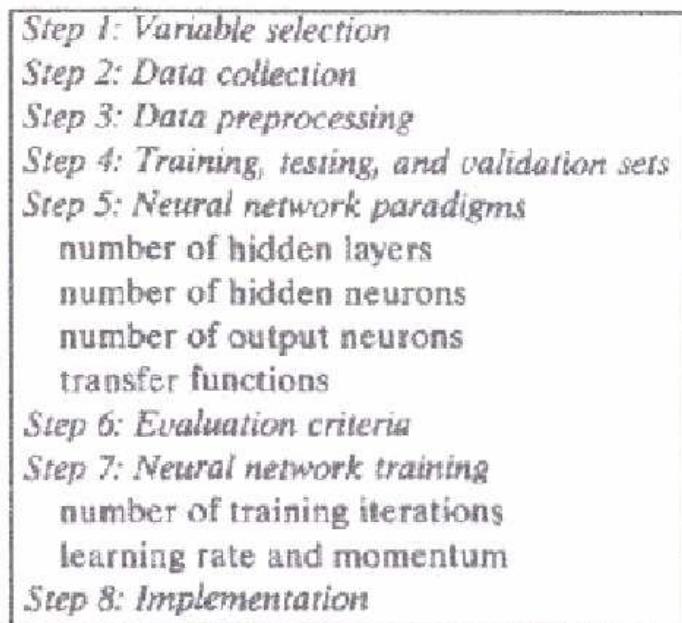


Figura 6. Procedimiento para el Diseño de un Modelo de Predicción basado en RNA (Kaastra y Boyd 1996).

El procedimiento empieza identificando las variables de entrada, seguidamente se procede a la recopilación de datos asegurando la consistencia de los mismos disminuyendo el ruido inherente en ellos (puede ser aplicando estandarización), luego se definen los conjuntos de entrenamiento, prueba y validación en base a los datos obtenidos. A partir de este momento es tiempo de comenzar a tomar en cuenta los paradigmas de diseño de la red neuronal artificial como el número de capas, el número de neuronas en cada capa, el número de salida y la función de transferencia. Una vez diseñada la red deberemos seleccionar los criterios de evaluación de los resultados que arroja la misma y para ello debemos determinar métricas tales como las mostradas en la tabla 3. Inmediatamente después se debe de planificar como va a ser la fase de aprendizaje de la red, teniendo en cuenta factores tales como el número de interacciones y la tasa de aprendizaje. Por último se implementa el modelo haciendo uso de software adecuado.

III. MATERIALES Y MÉTODOS

3.1 Diseño Investigación

3.1.1. Tipo de Investigación

El proyecto propuesto está referido a una investigación experimental y descriptiva. Se considera experimental dada la existencia de una variable independiente y otra dependiente, generando de esta forma una alteración en la variable dependiente. Y es descriptiva dado que se miden los indicadores de la variable dependiente.

3.2. Variables

- **Variable independiente:** Redes Neuronales Artificiales para el pronóstico de la demanda de agua.
- **Variable dependiente:** Predicción de la demanda de agua potable para la Empresa EPSEL S.A.

3.3. Indicadores

VARIABLE	INDICADOR	DESCRIPCIÓN
Error de Pronóstico	Error de Pronóstico Acumulativo $EPA = \sum_{t=1}^n e_t$ e_t = Error del Pronóstico n = Cantidad total de tiempos observados	Indica la suma total de todos los errores de predicción en el tiempo.
Costos	Costo Total	Indica la suma total de todos los gastos en el Sistema de Agua Potable.

Tabla 7. Indicadores

3.4. Población, Muestra y Muestreo

3.4.1. Población

Dado que el propósito de este trabajo es realizar pronóstico de demanda a mediano plazo, se estableció la población conforme al horizonte temporal:

- Mediciones de consumo de agua potable en Metros Cúbicos (M³) agrupados por meses (Enero 1990 – Diciembre 2010), haciendo un total de 252 meses.

3.4.2. Muestra

Debido a que los datos están almacenados y son manipulados electrónicamente, no hay necesidad de establecer una muestra, es más conveniente trabajar con toda la población, por lo tanto hablamos de una población muestral.

3.5. Métodos, técnicas e instrumentos de recolección de datos

3.5.1. Métodos de Pronóstico

- Método General de Diseño de Pronóstico

N°	Pasos	Descripción
1	Identificar el Horizonte Temporal	Es decir el corto, mediano y largo plazo.
2	Observar los Hechos	Identificar las variables a predecir y sus dependencias
3	Seleccionar una Técnica	Seleccionar las técnicas que más se ajustan al tipo de problemas presentado.
4	Aplicar Técnicas Seleccionadas	Implementar el pronóstico con la técnica seleccionada.

Tabla5. Método General de Diseño de Pronóstico

- Método de Diseño de Modelos de Pronósticos con Redes Neuronales Artificiales.

3.5.2. Técnicas

Redes Neuronales Artificiales.

3.5.3. Herramientas

MATLAB 7.8.0.

MATLAB Neural Network Toolbox

3.5.4. Fuente

Los datos a emplear fueron obtenidos de la Empresa EPSEL S.A., la cual proporciono archivos históricos de datos reales y de gran utilidad para nuestro propósito.

3.6. Metodología

La metodología para la implementación de Redes Neuronales en el pronóstico de agua, se basa en el trabajo de(Kaastra y Boyd 1996).

El cual está compuesto por 8 pasos:

- 1. Selección de la Variable.**
- 2. Recolección de Datos.**
- 3. Pre-procesamiento de Datos.**
- 4. Definición de Conjunto de Entrenamiento, Validación y Prueba.**
 - a. Conjunto de Entrenamiento.**
 - b. Conjunto de Validación.**
 - c. Conjunto de Prueba.**
- 5. Selección de la Arquitectura de Redes Neuronales.**
 - a. Número de Neuronas de Entrada.**

IV. RESULTADOS

4.1. Metodología para Redes Neuronales de *(Kaastra y Boyd 1996)*.

4.1.2. Diseñar Modelo de RNAs.

4.1.2.1. Selección de la Variable

Determinar qué variables de entrada son importantes en el pronóstico de series de tiempo es un asunto crítico. En este caso, el interés en la predicción involucra emplear como entradas datos técnicos en lugar de fundamentales por las características mismas de la investigación.

- Variable a Predecir. Cantidad mensual de demanda de Agua Potable de la Empresa EPSEL S.A, la cual utilizaremos durante todo este trabajo para probar nuestra hipótesis.
- Tipo de Variable. Cuantitativa.
- Unidad. Es el Metro Cúbico (M3).

4.1.2.2. Recolección de Datos

Al seleccionar las variables del paso anterior también se debe considerar su disponibilidad.

Los datos utilizados para el presente trabajo fueron obtenidos por parte de la Empresa EPSEL S.A., y consisten en mediciones de la cantidad mensual demandada de agua potable de la ciudad Lambayeque de 20 años detallados mensualmente. Por lo tanto la serie temporal de la demanda de agua, que se presenta a la red, corresponde a los datos del periodo comprendido entre Enero de 1990, hasta Diciembre de 2010, lo que hace un total de 252 meses.

Para poder introducir la serie de tiempo a la red neuronal es necesario tener los datos de tal forma que se puedan manipular con facilidad en el software para el diseño de la implementación de la RN, para lo cual se copiaron todos los datos en una hoja de cálculo de Excel.

4.1.2.3. Pre-procesamiento de Datos

El pre-procesamiento de datos se refiere al hecho de analizar y transformar las variables de entrada y salida para minimizar el ruido, enfatizar relaciones importantes, detectar tendencias, y aplanar la distribución de la variable para ayudar a la red neuronal en el aprendizaje de patrones relevantes. Como las redes neuronales son buscadores de patrones, la representación de los datos es crítica en el diseño de una red exitosa. Las variables de entrada son raramente ingresadas a la red en su forma original.

En nuestro caso la etapa de pre-procesamiento de datos no fue necesaria, dado que los datos ya se encuentran en formato libre de ruido (no contiene valores nulos ni valores con formato inconsistente), sin embargo, el software de implementación ofrece funciones de pre-procesamiento y post-procesamiento implícitas en las funciones de entrenamiento haciendo innecesario el uso individual de éstas.

4.1.2.4. Definición de Conjunto de Entrenamiento, Validación y Prueba

A1. Conjunto de Entrenamiento.

Es el conjunto de datos que le sirve a la red neuronal para aprender los patrones presentes en los datos, es decir la obtención de los pesos de la red. En este caso este conjunto representará un total de 228 registros de los 252 disponibles. Correspondientes a los primeros 19 años.

A2. Conjunto de Validación.

Se utiliza para la comprobación final de la red, donde los datos empleados son los más recientes consecuentemente al último valor de la muestra. El conjunto de validación debe consistir en las más recientes y contiguas observaciones. Se debe procurar no usar el conjunto de validación como conjunto de testeo tras ejecutar repetidamente pasos de series de entrenamiento-testeo-validación y adaptar las variables de entrada basadas en la actuación de la red en el conjunto de validación.

Muchas veces ocurre durante el proceso de entrenamiento de la red, un problema llamado overfitting (se da cuando la red tan solo aprende puntos individuales en vez de generales), para evitarlo; es necesario incluir un conjunto de validación durante el proceso de aprendizaje de la red. En este caso este conjunto se obviará, debido a que estamos utilizando un componente de retardo y por lo tanto el conjunto de entrenamiento se va generando en el tiempo, impidiendo de esta manera que exista un conjunto de validación priori.

A3. Conjunto de Prueba.

Corresponde a los datos restantes, una vez seleccionados los patrones de entrenamiento. Este conjunto de datos se utiliza para evaluar la precisión de la red. Está compuesto de los últimos 24 registros de nuestra data (últimos dos años).

4.1.2.5. Selección de la Arquitectura de Redes Neuronales

Existen diversas formas de determinar la arquitectura a seguir para desarrollar una red neuronal de predicción, cabe destacar que se emplea en la mayoría de los casos redes con algoritmo de entrenamiento Backpropagation.

- **Número de neuronas de entrada.** Corresponde a las mediciones de agua potable de un año. Por tanto la cantidad de neuronas de entrada son 12.
- **Número de Capas Ocultas.**

Las capas ocultas dan a la red la habilidad de generalizar, y en la práctica las redes neuronales con una o dos capas ocultas son las más utilizadas y han tenido un buen desempeño (Kaastra y Boyd 1996). El incremento en el número de capas también incrementa el tiempo de procesamiento y el peligro de sobre ajuste lo que conduce a un pobre desempeño en la predicción fuera de muestra. El sobre ajuste ocurre cuando un modelo de predicción tiene muy pocos grados de libertad. En otras palabras, se tienen relativamente pocas observaciones en relación con sus parámetros y por lo tanto es capaz de memorizar datos

individuales en lugar de aprender patrones generales. Es por eso que se recomienda que todas las redes neuronales comiencen de preferencia con una o a lo mucho con dos capas.

Para nuestro caso hemos optado por una sola capa oculta la cual es suficiente para asegurar la capacidad de generalización de nuestra red dada la cantidad de datos con los que contamos. Esto nos da un total de 3 capas, la capa de entrada, la capa oculta y la capa de salida.

- **Número de Neuronas Ocultas.**

El número de neuronas de nuestra capa oculta será de un 75% del total de entradas (Kaastra y Boyd 1996). Es decir, dado que el número de entradas de nuestra red es igual al número de meses que contiene un año (12 meses), el número de neuronas de nuestra capa oculta sería de 9 neuronas.

- **Número de Neuronas de Salida.**

Decidir el número de neuronas de salida es algo más sencillo porque hay muchas razones para emplear sólo una neurona de salida. Las redes neuronales con múltiples salidas, especialmente si éstas salidas están ampliamente espaciadas, producirán resultados inferiores en comparación con una red con una única salida. Lo recomendable es tener una red especializada para cada una de las salidas deseadas en cada predicción.

El número de neuronas de la capa de salida de nuestra red es una, porque tan sólo intentamos predecir el valor de la variable seleccionada para el siguiente mes y no para un conjunto de meses lo cual disminuiría la precisión de la red.

- **Función de Transferencia.**

El propósito de esta función es prevenir a las salidas de alcanzar valores muy elevados los cuales pueden paralizar la red y detener el entrenamiento de la

misma. Como función de transferencia usaremos la función sigmoideal (cuyo rango de salida se encuentra entre -1 y +1) la cual es preferente para tareas de predicción según (Kaastra y Boyd 1996) y se muestra en la figura 7.

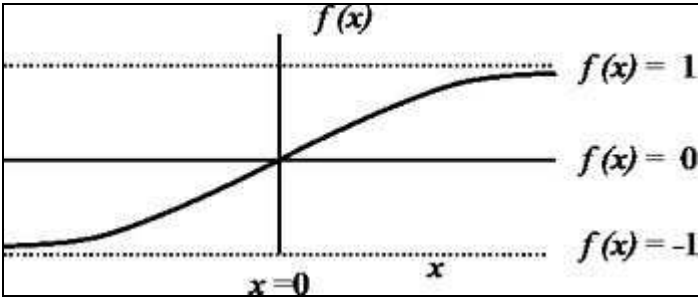


Figura7. Función Sigmoideal (Rango de Salida entre -1 y +1).

La figura 8 muestra la topología y el modelo de neurona.

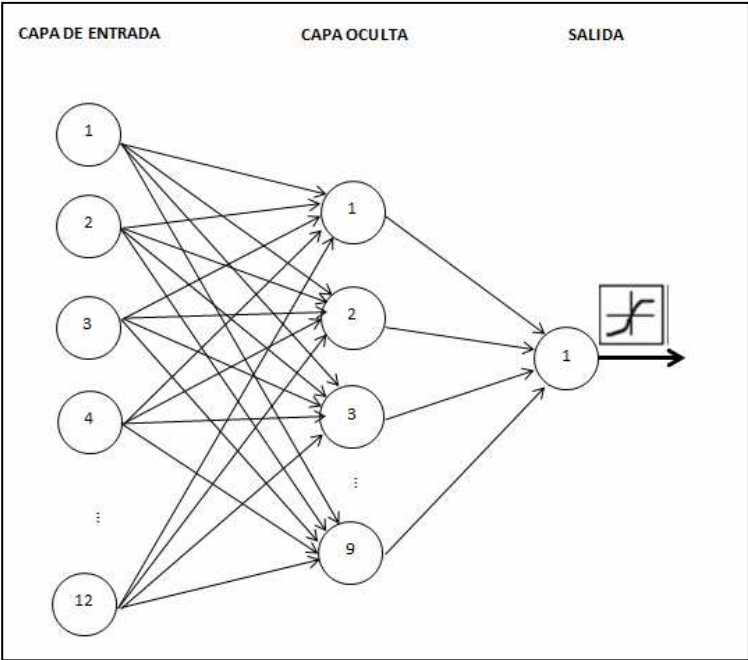


Figura 8 . Modelo de la Red Neuronal.

4.1.2.6. Criterios de Evaluación

En cuanto a los criterios de evaluación para medir la eficiencia de la red, se ha considerado utilizar el Error Cuadrático Medio (ECM) de la tabla 6. El Error Cuadrático Medio calculado como la diferencia entre la salida de la red y la respuesta deseada. El cual se usa como factor de culminación del entrenamiento. Para esta fase se fijó un parámetro en 300 épocas y el factor de terminación empleado del ECM teniendo como umbral un valor de 0.5.

4.1.2.7. Entrenamiento de la Red Neuronal

Entrenar una red neuronal para aprender patrones involucra el presentarle ejemplos de manera iterativa de las respuestas correctas. El objetivo del entrenamiento es encontrar un conjunto de pesos entre las neuronas que determinan el mínimo global de la función de error. A menos que el modelo esté sobre-ajustado, el conjunto de pesos debería proporcionar una buena generalización.

El entrenamiento de la red utiliza la técnica del gradiente descendente inmerso en el algoritmo de retropropagación. El entrenamiento se detendrá básicamente cuando suceda uno de estos 3 motivos: el número de iteraciones excede el número de épocas establecidas, si la función de evaluación (ECM) cae por debajo de la meta establecida, o si el error medio por la función de evaluación se incrementa para un número específico de iteraciones (este último caso requiere de la existencia del conjunto de validación). En cualquier caso los pesos y bias obtenidos son los encontrados en el mínimo error medido por la función de evaluación.

- **Número de Iteraciones (Épocas) Entrenamiento**

Muchas veces el procedimiento de agregar un número mayor de iteraciones en el entrenamiento producirá un menor error en el entrenamiento,

pero éste a su vez no garantiza que se obtendrá el menor error con el conjunto de prueba. En nuestro caso el número máximo fue de 300 épocas que fue suficiente para conseguir un nivel de error aceptable.

- **Tasa de Aprendizaje y Momento**

La tasa de aprendizaje es una constante de proporcionalidad que determina el tamaño del cambio de los pesos. El cambio de los pesos de una neurona es proporcional al impacto del peso de las neuronas sobre el error. Un método de incrementar la tasa de aprendizaje y por lo tanto agilizar el tiempo de entrenamiento sin caer en oscilaciones, es incluir un término de momento en la regla de aprendizaje de *BP*. El término de momento determina cómo los cambios pasados de los pesos afectan a los actuales cambios en los pesos. Este término suprime oscilaciones lado a lado filtrando variaciones de alta frecuencia. Cada nueva dirección de búsqueda es una suma ponderada de los gradientes actuales y previos.

4.1.2.8. Implementar Modelo de RNAs.

Se lleva a cabo la implementación del modelo haciendo uso de la técnica de Redes Neuronales Artificiales.

Para proceder con la implementación de este modelo, antes presentamos un resumen de los datos más relevantes del mismo (tabla 8) que nos ayudará a internarnos en el proceso de su implementación en MATLAB.

Resumen de Datos de la Red Neuronal	
Data Pre-procesada	252 Registros
Tipo de Variable	Cuantitativa (Metros Cúbicos. M3)
Frecuencia de Data	Mensual
Numero de Épocas	300
Tamaño de Conjunto de Entrenamiento	228
Tamaño de Conjunto de Prueba	24 meses
Número de Neuronas de Entrada	12
Número de Capas Ocultas	1
Número de Neuronas Ocultas en cada Capa	9
Número de Neuronas de Salida	1
Función de Transferencia	Función Sigmoideal
Función de Error	Error Cuadrático Medio

Tabla8. Resumen de Datos de la Red Neuronal

El editor de control de MATLAB es el siguiente (figura 9):

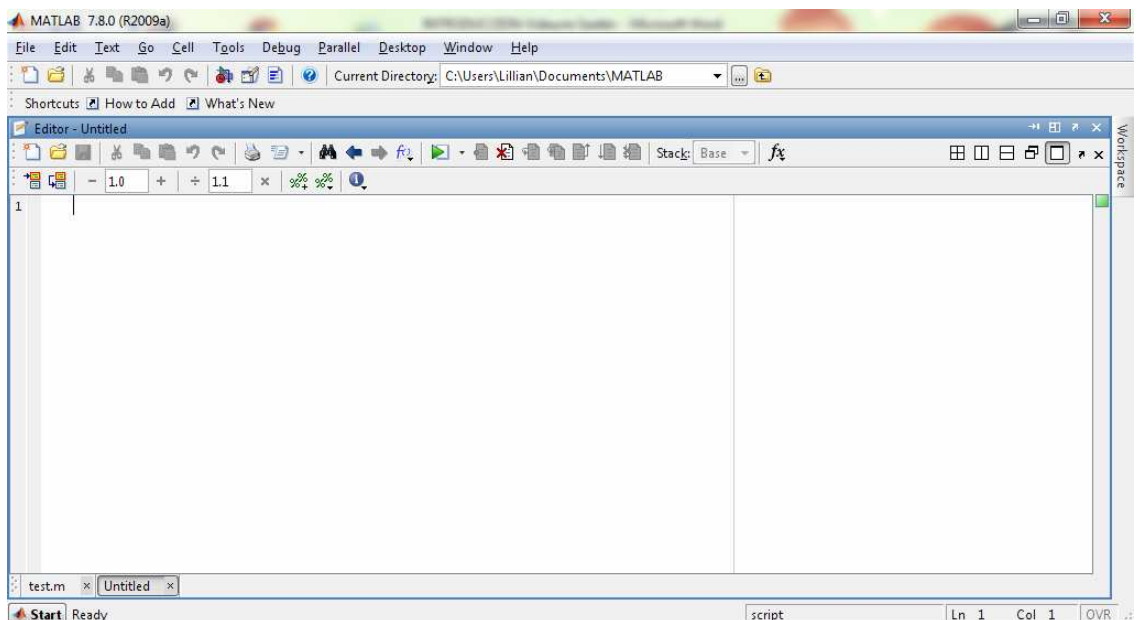


Figura 9. Editor de Control de MATLAB.

MATLAB. Luego filtramos los primeros 228 registros, que corresponden al año 1990 hasta el 2008 que serán necesarios para llevar a cabo el entrenamiento de la red (Los otros 24 registros, son los que corresponden al 2009 y 2010, los cuales se usarán para probar la precisión de la red):

```
dataEntren = con2seq(data)  
dataEntren = dataEntren(1:228);
```

Procedemos a crear nuestra red con la instrucción:

```
red = newfftd(dataEntren,dataEntren,(1:12),9);
```

Dicho código crea un objeto de red neuronal feed-forward con tiempo de retardo de 12 meses al que llamamos “red”. Por defecto crea una capa oculta, que toma como función de evaluación el error cuadrático medio (ECM) y asigna el número de neuronas de salida a la dimensión del conjunto objetivo (en nuestro caso es uno).

Los argumentos de esta orden lo explicamos a continuación:

- El primer argumento es el conjunto de entrenamiento que servirá de entrada. El cual es **dataEntren**.
- El segundo argumento es el conjunto objetivo, (dataEntren). Ambos conjuntos son iguales en nuestro caso dado que la red tiene un componente de retraso que organiza las entradas ingresándolas cada 12 registros. Esto es consecuencia de que nuestra red tendrá 12 neuronas de entrada.
- El tercer argumento representa los retardos que tendrá nuestra red. En nuestro caso es de 12 (1:12), debido a que el número de meses del año son 12.
- El cuarto argumento representa el número de neuronas en la capa oculta (9).

Ahora vamos a implementar los dos motivos básicos en la cual uno de ellos hará que el entrenamiento finalice, por lo cual el entrenamiento debería detenerse; y el otro asigna la meta de la función de evaluación.

Pasamos a asignar el número máximo de épocas que se pueden alcanzar durante el entrenamiento. Si se alcanza este número de épocas, el entrenamiento se detiene:

```
red.trainparam.epochs = 300;
```

Después asignamos la meta de la función de evaluación (ECM) cuyo valor al ser alcanzado, detendrá el entrenamiento.

```
red.trainparam.goal = 0.5;
```

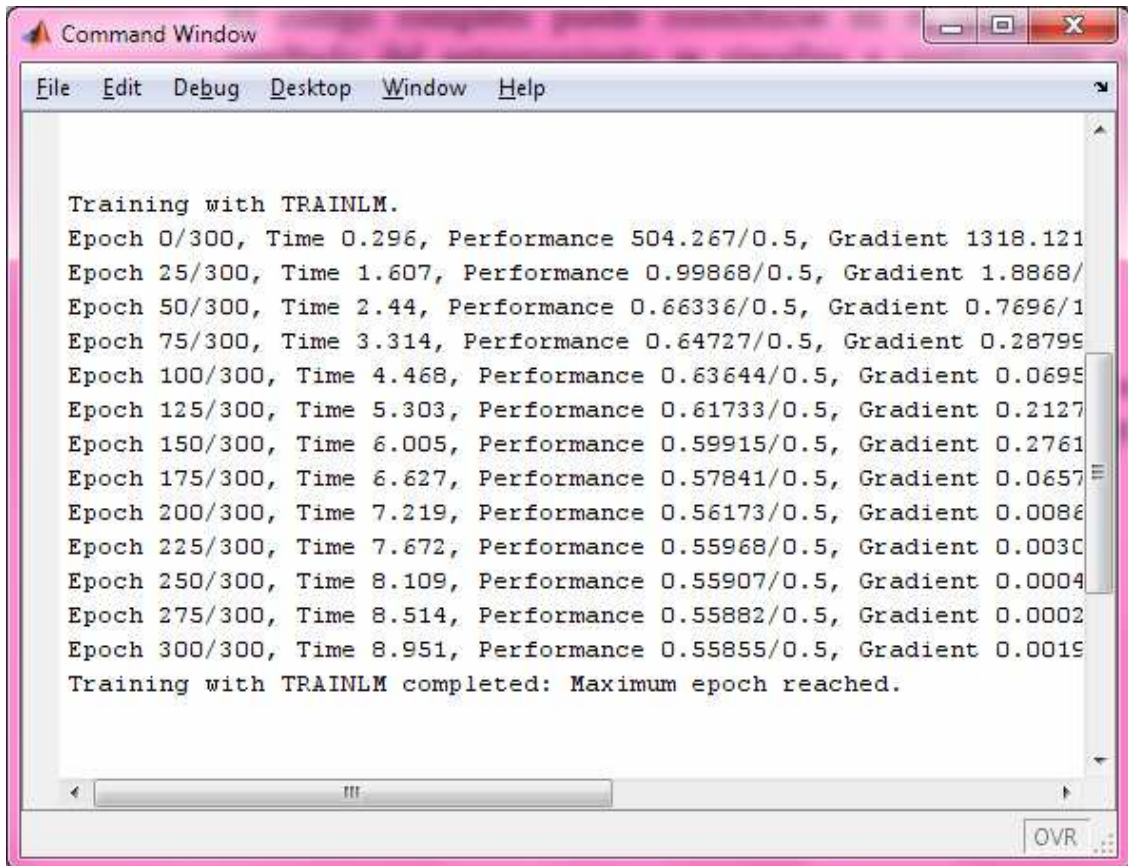
Una vez conformada la red, en seguida procedemos a preparar los argumentos necesarios para el entrenamiento de la misma. Creamos un primer arreglo llamado "P", el cual servirá para almacenar el conjunto de entrada el cual empieza en el registro 13, ya que los anteriores 12 (meses que corresponden al primer año) será usado para completar el componente de retardo de la red. Al segundo arreglo lo llamaremos "T", el cual representara el conjunto de salida y es igual al conjunto de entrada dado que la predicción de la variable se basa en sus valores anteriores. El tercer arreglo llamado "PP", representa los 12 valores iniciales que tendrá el componente de retardo:

```
P = dataEntren(13:end);  
T = dataEntren(13:end);  
PP = dataEntren(1:12);
```

Finalmente procedemos con el entrenamiento de la red, en la cual se utilizara la función "train". Cabe mencionar que el pre-procesamiento y el post-procesamiento de datos que consiste por ejemplo, en anulación de valores nulos, valores repetitivos, etc., ya están inmersos en la función "train".

```
red = train(red,P,T,PP);
```

El código completo puede consultarse en el Anexo I. El resultado del entrenamiento se visualiza a continuación en modo de texto (figura 11) y luego en modo gráfico (figura 12):



```
Command Window
File Edit Debug Desktop Window Help

Training with TRAINLM.
Epoch 0/300, Time 0.296, Performance 504.267/0.5, Gradient 1318.121
Epoch 25/300, Time 1.607, Performance 0.99868/0.5, Gradient 1.8868/
Epoch 50/300, Time 2.44, Performance 0.66336/0.5, Gradient 0.7696/1
Epoch 75/300, Time 3.314, Performance 0.64727/0.5, Gradient 0.28799
Epoch 100/300, Time 4.468, Performance 0.63644/0.5, Gradient 0.0695
Epoch 125/300, Time 5.303, Performance 0.61733/0.5, Gradient 0.2127
Epoch 150/300, Time 6.005, Performance 0.59915/0.5, Gradient 0.2761
Epoch 175/300, Time 6.627, Performance 0.57841/0.5, Gradient 0.0657
Epoch 200/300, Time 7.219, Performance 0.56173/0.5, Gradient 0.0086
Epoch 225/300, Time 7.672, Performance 0.55968/0.5, Gradient 0.0030
Epoch 250/300, Time 8.109, Performance 0.55907/0.5, Gradient 0.0004
Epoch 275/300, Time 8.514, Performance 0.55882/0.5, Gradient 0.0002
Epoch 300/300, Time 8.951, Performance 0.55855/0.5, Gradient 0.0019
Training with TRAINLM completed: Maximum epoch reached.
```

Figura 11. Resultados del Entrenamiento (Modo Texto).

Se ve que el performance (la función de ECM) va disminuyendo conforme transcurren las épocas, sin embargo no alcanza su tope (0.5).

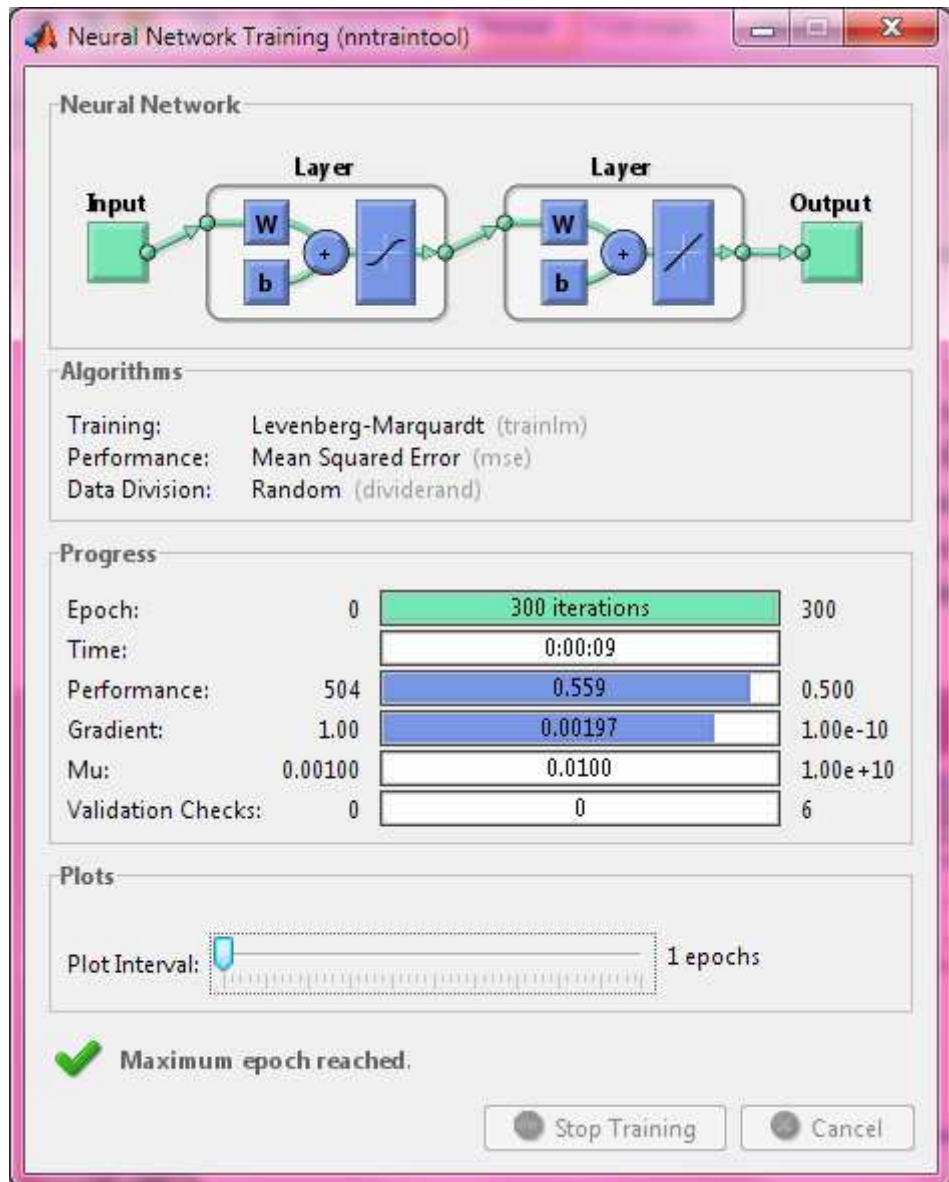


Figura 12. Resultados del Entrenamiento (Modo Gráfico).

Se puede apreciar que el entrenamiento finalizó porque se alcanzó la época 300 que era nuestro máximo número de épocas permitido y una de nuestras condiciones de parada. Seguidamente guardamos la red ya entrenada para utilizarla en el Cálculo del Nivel de Error, y finalmente, ya para concluir con la implementación del modelo, presentamos las gráficas que miden la eficiencia del mismo:

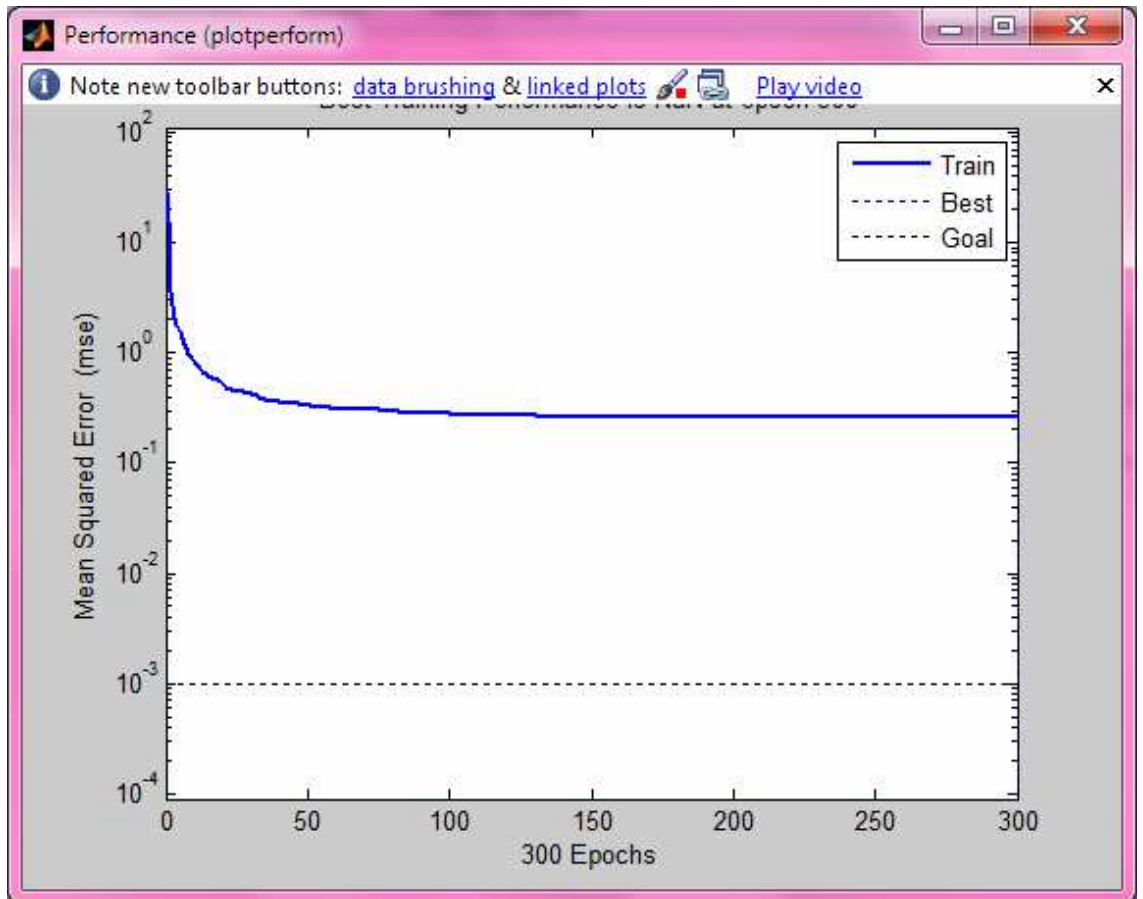


Figura 13. Performance del Entrenamiento

La figura 13 nos muestra cómo va disminuyendo el error a través del tiempo (épocas) hasta alcanzar su mínimo durante el entrenamiento.

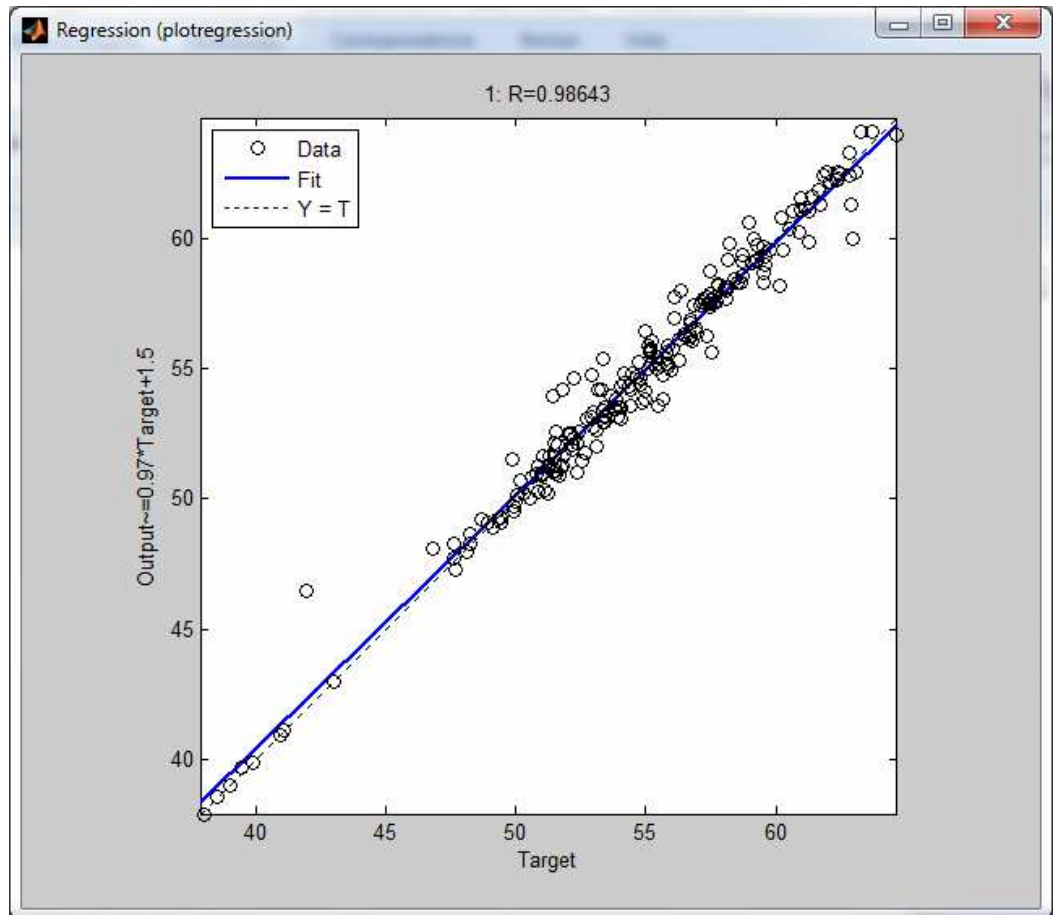


Figura 14. Correlación entre Valores Reales y Pronósticos

En la figura 14 podemos apreciar, que el coeficiente de correlación es bastante elevado ($R=0,98643$) y muy cercano a 1, lo cual quiere decir que existe una muy alta correlación entre los valores reales (abscisas) y los pronósticos (ordenadas). La línea azul indica los puntos en los que los valores reales son iguales a los pronosticados (situación ideal).

4.1.3. Pesos Calculados por el Algoritmo de Entrenamiento.

- Pesos de la Capa de Entrada con la Capa Oculta.

En los pesos se suele guardar la mayor parte del conocimiento que la red neuronal tiene sobre la tarea en cuestión. Los pesos resultantes entre la capa de entrada y la capa oculta del modelo se muestran en la tabla 9.

9 X 12	BÍAS	1	2	3	4	5	6	7	8	9	10	11	12
1	1.5239	0.2977	-1.6808	0.5265	0.5372	0.1738	0.3295	-1.2026	1.2473	1.4086	-1.0949	-0.7322	0.0573
2	3.1285	3.3274	-2.2791	0.8369	-1.1479	0.4965	0.6109	-0.4971	-1.9031	-3.7763	3.4692	3.9556	-1.2105
3	-1.1158	-0.4940	-0.7588	0.9816	2.2254	0.4982	0.6322	-1.7839	0.8741	1.3849	-0.1231	0.1425	0.5711
4	-1.2126	-0.1454	0.2549	0.2455	-0.4767	0.1888	-0.4853	0.0689	1.1177	0.9907	-1.4708	-1.0261	0.1558
5	-2.1662	0.4915	0.4907	1.4419	-0.9389	0.588	0.2344	9.7462	-0.4783	0.1541	-8.1890	11.202	0.9715
6	-0.8685	-0.6249	-0.6889	0.8705	-1.9268	2.3082	-1.0324	-1.2334	3.1703	0.4753	-1.7558	-0.1501	-0.2352
7	0.9311	4.2972	-1.4902	-2.5747	-4.5885	4.1122	-2.0575	-1.5361	2.9947	-5.5039	0.2775	3.3578	-0.3449
8	-0.4891	-1.7637	-1.9817	1.4179	-4.9438	2.8112	-0.9577	0.7112	2.1397	2.0655	1.1096	0.4771	-1.0389
9	3.6291	-5.3370	2.2610	1.0514	6.2774	1.9370	2.8948	-9.1052	2.7831	-3.5226	5.8241	1.1172	-0.2258

Tabla 9. Pesos entre la capa de entrada y la capa oculta de la RNA estimada

- Pesos de la Capa Oculta con la Capa de Salida.

A continuación se presentan los pesos de la capa oculta con la capa de salida en la tabla 10.

1 X 9	BÍAS	1	2	3	4	5	6	7	8	9
1	2.1908	-1.4392	1.5537	0.6066	3.9332	-0.2759	-0.9928	0.3610	0.4068	0.4325

Tabla 10. Pesos entre la capa oculta y la capa de salida de la RNA estimada.

4.1.4. Cálculo del Error del Modelo de Redes Neuronales Artificiales.

Para llevar adelante esta etapa emplearemos los 24 registros restantes de nuestra data que representan los últimos 24 meses medidos (Enero 2009 – Diciembre 2010). La cual procederemos a utilizar para determinar el nivel de error del pronóstico de la predicción de la cantidad demandada de agua potable a mediano plazo (meses).

La predicción para los próximos 24 meses del consumo de agua mediante la Red Neuronal seleccionada, es comparada con los datos del consumo real para comprobar la validez del modelo seleccionado. El código completo que muestra el cálculo del error de este modelo, se puede apreciar en el Anexo II.

En la tabla 11, se muestra para los 2 últimos años de la data los valores reales del consumo de agua y los valores obtenidos por la predicción de la red neuronal.

MES	VALOR REAL	VALOR DE PREDICCIÓN	NIVEL DE ERROR
Enero	166.840	169.628	-2.788
Febrero	167.834	167.782	0.052
Marzo	156.403	169.813	-13.410
Abril	157.339	156.571	0.768
Mayo	160.802	154.584	6.218
Junio	161.506	159.369	2.137
Julio	156.495	161.391	-4.896
Agosto	160.335	162.340	-2.005
Septiembre	162.827	159.397	3.430
Octubre	163.394	162.304	1.090
Noviembre	164.325	163.855	0.470
Diciembre	172.103	166.029	6.074
Enero	168.336	172.349	-4.013
Febrero	169.814	169.769	0.045
Marzo	161.725	172.656	-10.931
Abril	162.489	161.664	0.825
Mayo	164.417	160.478	3.939
Junio	165.742	164.378	1.364
Julio	161.345	165.980	-4.635
Agosto	163.477	166.568	-3.091
Septiembre	165.708	163.258	2.450
Octubre	166.938	165.799	1.139
Noviembre	167.456	167.684	-0.228
Diciembre	174.302	170.041	4.261

Tabla 11. Predicciones de los Últimos 24 meses Aplicando el Modelo de RNAs

4.1.5. Validación de la red con conjunto de prueba

Se evalúa la red para comparar con los datos arrojando el posible resultado.

En las siguientes tablas 12 y 13 se recogen los consumos previstos obtenidos al aplicar redes neuronales, los cuales con comparados con los datos reales para el año 2009 y 2010.

AÑO 2009	RED NEURONAL	CONSUMO REAL
Enero	169,628	166,840
Febrero	167,782	167,834
Marzo	169,813	156,403
Abril	156,571	157,339
Mayo	154,584	160,802
Junio	159,369	161,506
Julio	161,391	156,495
Agosto	162,340	160,335
Septiembre	159,397	162,827
Octubre	162,304	163,394
Noviembre	163,855	164,325
Diciembre	166,029	172,103
Total	1,953,063.00	1,950,203.00

Tabla 12. Consumo Real y Predicciones con Redes Neuronales

AÑO 2010	RED NEURONAL	CONSUMO REAL
Enero	172,349	168,336
Febrero	169,769	169,814
Marzo	172,656	161,725
Abril	161,664	162,489
Mayo	160,478	164,417
Junio	164,378	165,742
Julio	165,980	161,345
Agosto	166,568	163,477
Septiembre	163,258	165,708
Octubre	165,799	166,938
Noviembre	167,684	167,456
Diciembre	170,041	174,302
Total	2,000,624.00	1,991,749.00

Tabla 13 Consumo Real y Predicciones con Redes Neuronales.

Por lo tanto, el consumo global del agua previsto para el año 2009, se estima en 1, 953,063.0 m³ y para el año 2010 se estima en 2, 000,624.0 m³.

La representación gráfica de los valores estimados durante el periodo 2009-2010 con la red neuronal seleccionada y los valores reales se encuentran en la gráfico 1. Ambas líneas representadas se encuentran muy próximas, lo que significa que la red ha aprendido y que tiene capacidad de generalización.

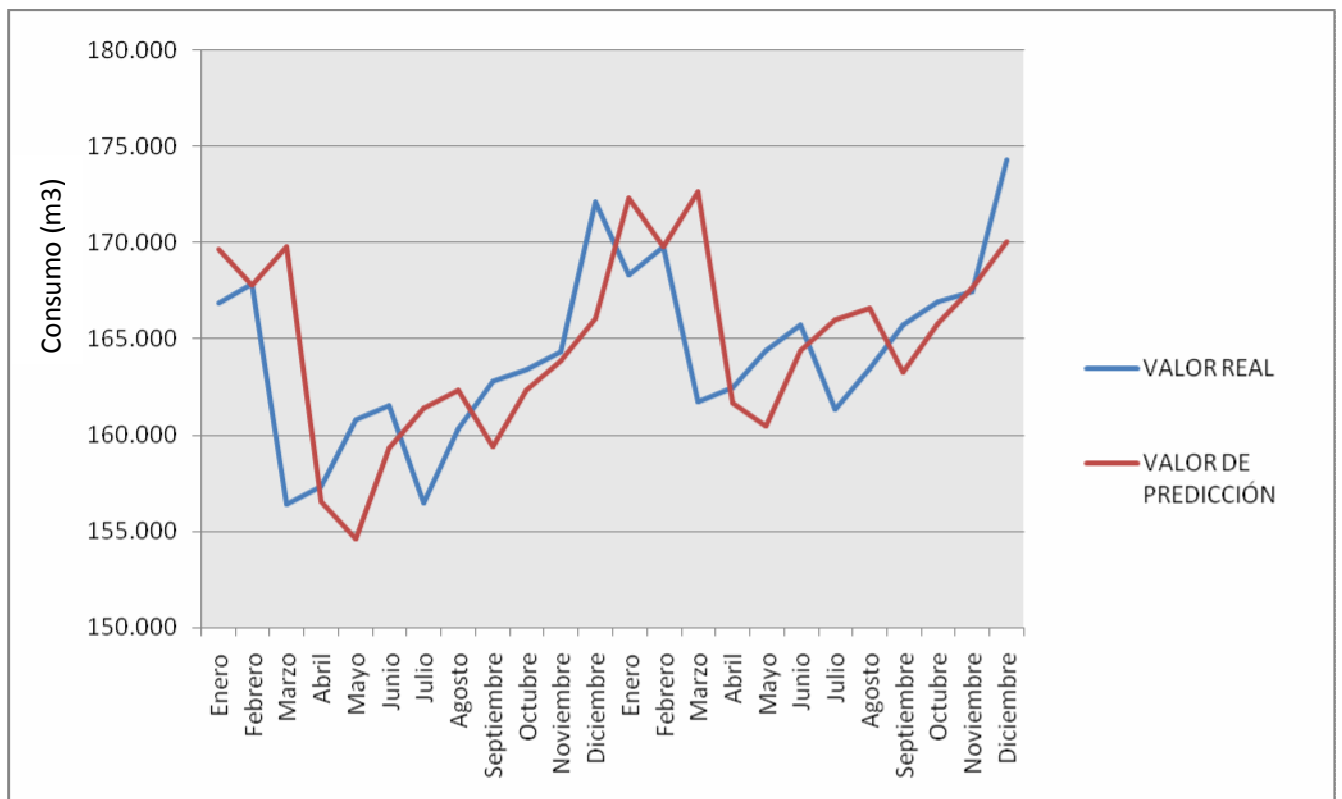


Gráfico 1. Consumo Real – Consumo Previsto (M³).

V. DISCUSIÓN

En este capítulo se analizan y discuten los resultados de la aplicación del sistema desarrollado, en la Empresa EPSEL S.A. El análisis estará enfocado a evaluar los indicadores planteados en el capítulo III. Estos indicadores son:

- Error de Pronóstico Acumulativo.
- Costo.

Para la evaluación de nuestro trabajo, se analizarán los datos de estos indicadores, el Consumo de Agua Real, con el pronóstico del Consumo de Agua, utilizando Redes Neuronales Artificiales.

5.1. Error de Pronóstico Acumulativo (Indicador).

Los valores predichos del consumo de agua mediante la Red Neuronal seleccionada, son comparados con los datos del consumo real y con los datos pronosticados por la empresa, que se muestran en la tabla 14.

AÑO	CONSUMO REAL	PRÓNOSTICO DE LA EMPRESA	PRONÓSTICO DE LA RED
1991	1281465.00	1881648.00	1270089.00
1992	1308783.00	1906103.00	1305906.00
1993	1343684.00	1930886.50	1331162.00
1994	1367290.00	1955962.00	1363103.00
1995	1358100.00	1981402.50	1379907.00
1996	1375473.00	2007171.50	1375644.00
1997	1383540.00	2033269.00	1392820.00
1998	1424768.00	2059695.00	1409648.00
1999	1409690.00	2086449.50	1442193.00
2000	1450390.00	2113569.00	1448981.00
2001	1476802.00	2141053.50	1481591.00
2002	1559860.00	2168903.00	1536522.00
2003	1637647.00	2197081.00	1628409.00
2004	1663320.00	2225660.50	1676403.00
2005	1693177.00	2254605.00	1710208.00
2006	1701288.00	2283914.50	1707034.00
2007	1945255.00	2313589.00	1915147.00
2008	1926613.00	2343665.00	1932021.00
2009	1950203.00	2374142.50	1953063.00

2010	1991749.00	2404985.00	2000624.00
-------------	------------	------------	------------

Tabla 14. Consumo Real, Pronóstico de la Empresa y Pronóstico de la RN.

Como se puede observar los resultados obtenidos en la predicción aplicando la RNA mejoran a los resultados que se obtienen con el Modelo Tradicional utilizado por la Empresa.

En el gráfico 2, se puede apreciar de forma más visual y clara, que las predicciones durante el periodo 1991 – 2010 aplicando redes neuronales artificiales se aproximan mucho más a los valores reales que las predicciones realizadas por la empresa.

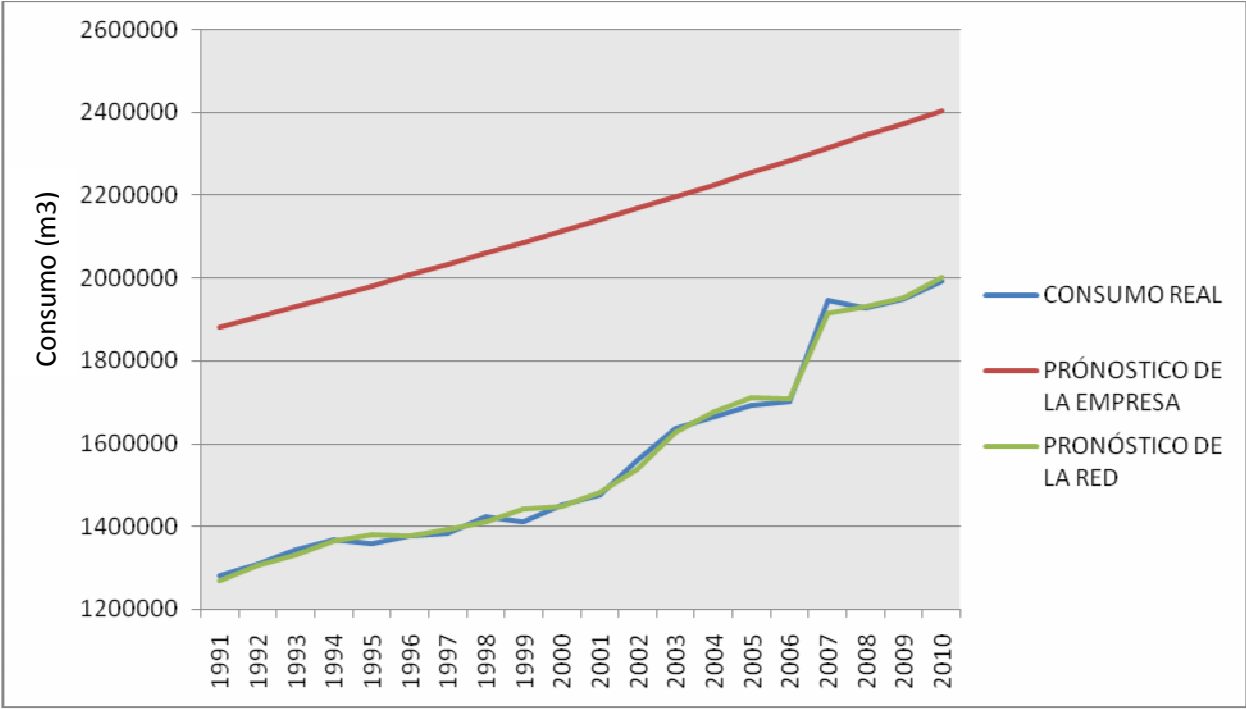


Gráfico 2. Consumo Real, Consumo Previsto por la Red y Consumo Previsto por la Empresa (M3).

5.1.1. Error de predicción

A continuación, en la tabla 15, se representan los errores cometidos por la empresa y por la red neuronal propuesta. Se puede observar cuán grande es la magnitud del error de la predicción hecha por la empresa, en cambio la predicción de la red neuronal tiene un nivel de error muy bajo.

AÑO	ERROR DE LA EMPRESA	ERROR DE LA RED
1991	600183	-11376
1992	597320	-2877
1993	587203	-12522
1994	588672	-4187
1995	623303	21807
1996	631699	171
1997	649729	9280
1998	634927	-15120
1999	676760	32503
2000	663179	-1409
2001	664252	4789
2002	609043	-23338
2003	559434	-9238
2004	562341	13083
2005	561428	17031
2006	582627	5746
2007	368334	-30108
2008	417052	5408
2009	423940	2860
2010	413236	8875
Total	11414658	11378

Tabla 15. Error de Pronóstico de la Empresa y de la Red Neuronal.

5.2. Costo (Indicador).

Costo de Operación y Mantenimiento

Los costos de operación y mantenimiento son aquellos que se refieren a costos incurridos en mantener el sistema actual de agua potable de la ciudad, estos costos han sido estimados basándose en información recopilada en la filial de EPSEL. Estos costos ascienden a S/127,883.90 soles anuales.

Se consideran los rubros de mano de obra, insumos químicos para potabilizar el agua, materiales para el mantenimiento y reparaciones, control de calidad y herramientas para el desarrollo de los trabajos de operación y mantenimiento. Los detalles de los costos de operación y mantenimiento se presentan en la Tabla 16.

DESCRIPCION	UND	CANTIDAD	COSTO UNITARIO	COSTO PRECIOS DE MERCADO
1.0 Personal				
Administrador (1)	mes	12.00	750.00	9000.00
Operarios (3)	mes	24.00	650.00	15600.00
2.0 Herramientas	Glb	1.00	3744.00	3744.00
3.0 Rep. Acceso. Lubric	Glb	1.00	11450.00	11450.00
4.0 Insumos				
HIPOCLORITO DE SODIO	Kg.	60.00	2.10	126.00
Cloro gas	Kg.	1451.00	4.50	6529.50
5.0 Limpieza y mantenimiento de equipos (2 equipos)	mes	12.00	886.20	10634.40
6.0 Servicios				
Energía eléctrica	mes	12.00	4780.00	57360.00
Teléfono	mes	12.00	120.00	1440.00
7.0 Gastos de comercialización	mes	12.00	1000.00	12000.00
COSTO TOTAL ANUAL				127,883.90
VOLUMEN DE AGUA DEMANDADA (AÑO 1990)				2321856
COSTO UNITARIO (S/.m3)				0.06

Tabla 16. Costos Unitarios de Operación y Mantenimiento para el Sistema de Agua Potable

Proyección de Costos de Operación y Mantenimiento

La empresa realiza una proyección de los costos de operación y mantenimiento, para saber cuánto va a gastar anualmente en mantener el Sistema de Agua Potable de la ciudad. Teniendo en consideración los costos unitarios que se obtiene de los egresos totales, entre la cantidad de m³ del volumen de agua, la cual se aprecia en la tabla 16, que corresponde al año 1990 de la operación del proyecto para la ciudad de Lambayeque. Una vez obtenido el volumen de agua producida para los demás años, se proyectan los costos de operación y mantenimiento, los cuales en todos los años se considera el costo unitario multiplicado por el volumen de agua, tal como se aprecia en la Tabla 17, la proyección de los costos de operación y mantenimiento por parte de la empresa y por parte de la red neuronal.

EMPRESA			
AÑO	COSTO DE OPERACIÓN Y MANTENIMIENTO		
	VOLUMEN DE AGUA POTABLE (m3/año)	COSTO UNITARIO	COSTO ANUAL-EMPRESA
		(S/. M3)	(S/. M3)
1991	2676317	0.06	147,406.99
1992	2711129	0.06	149,324.37
1993	2746351	0.06	151,264.36
1994	2782076	0.06	153,232.00
1995	2818256	0.06	155,224.77
1996	2854893	0.06	157,242.67
1997	2891986	0.06	159,285.69
1998	2929581	0.06	161,356.36
1999	2967678	0.06	163,454.67
2000	3006231	0.06	165,578.11
2001	3644421	0.06	200,728.51
2002	3835530	0.06	211,254.49
2003	3842685	0.06	211,648.54
2004	3855432	0.06	212,350.65
2005	3876814	0.06	213,528.34
2006	3936187	0.06	216,798.52
2007	3968563	0.06	218,581.73
2008	4018440	0.06	221,328.84
2009	4020821	0.06	241,249.25
2010	4031468	0.06	241,888.09
		TOTAL	3,752,726.95

RED			
AÑO	COSTO DE OPERACIÓN Y MANTENIMIENTO		
	VOLUMEN DE AGUA POTABLE (m3/año)	COSTO UNITARIO	COSTO ANUAL-RED
		(S/. M3)	(S/. M3)
1991	1587611	0.06	87,442.93
1992	1632383	0.06	89,908.86
1993	1663953	0.06	91,647.68
1994	1703879	0.06	93,846.75
1995	1724884	0.06	95,003.67
1996	1719555	0.06	94,710.17
1997	1741025	0.06	95,892.70
1998	1762060	0.06	97,051.27
1999	1802741	0.06	99,291.93
2000	1811226	0.06	99,759.27
2001	1851989	0.06	102,004.40
2002	1920653	0.06	105,786.28
2003	2035511	0.06	112,112.50
2004	2095504	0.06	115,416.79
2005	2137760	0.06	117,744.19
2006	2133793	0.06	117,525.67
2007	2393934	0.06	131,853.81
2008	2415026	0.06	133,015.55
2009	2441329	0.06	134,464.24
2010	2500780	0.06	137,738.72
		TOTAL	2,152,217.36

Tabla 17. Proyección de los Costos de Operación y Mantenimiento para el Sistema de Agua Potable

Como se puede apreciar en la tabla 17, la proyección realizada por la Empresa durante los años de 1991 hasta 2010 suma un total de S/. 3, 752,726.95 nuevos soles. Mientras que la proyección realizada por la Red Neuronal suma un total de S/. 2, 152,217.36 nuevos soles. Con lo cual se hubiera ahorrando un total de S/. 1, 600,509.59 nuevos soles.

En el gráfico 3, se puede apreciar de forma más visual y clara, que la proyección de los costos de operación y mantenimiento realizado por la red neuronal durante el periodo 1991 – 2010 es mucho menor que la proyección hecha por la empresa.

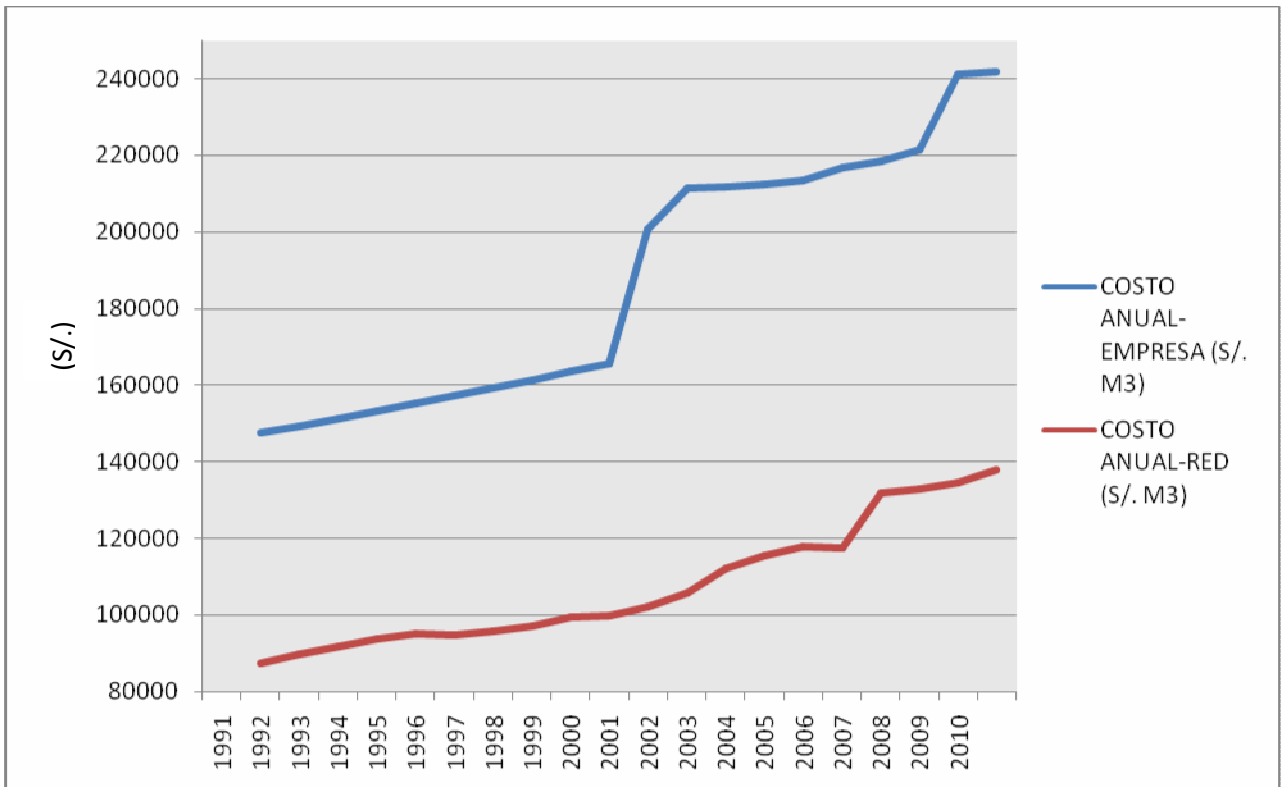


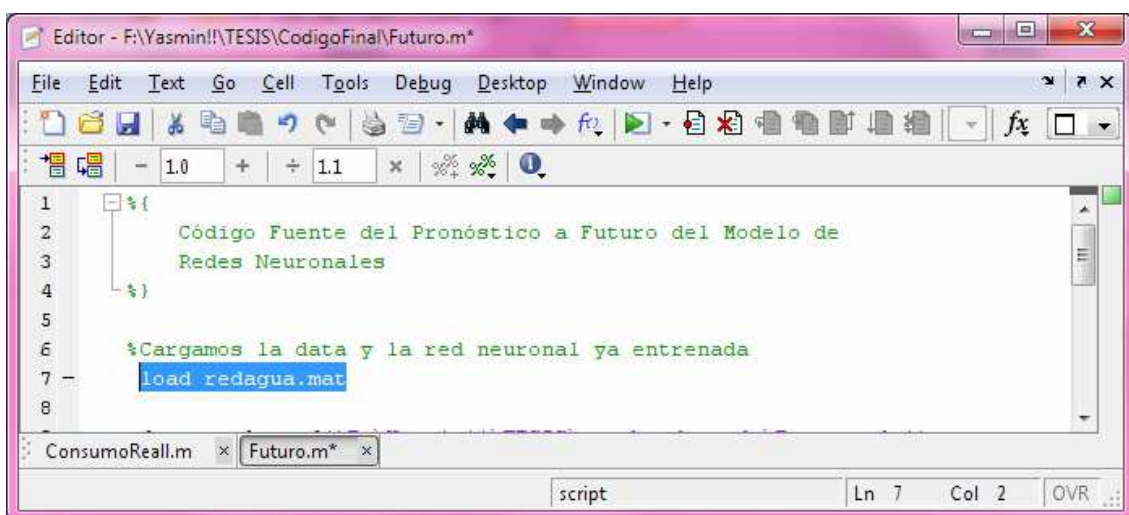
Gráfico 3 . Proyección de los Costos de Operación y Mantenimiento para el Sistema de Agua Potable.

VI. PROPUESTA

Ya se ha destacado el papel predominante que tiene el consumo de agua para la vida y la importancia que tiene el disponer de predicciones precisas para la planificación y gestión sostenible del agua. Por ello se propone que la Empresa, sugiera el Sistema de la predicción del agua utilizando la Red Neuronal. A continuación se describen las pantallas del sistema desarrollado para el pronóstico de agua a partir de la red neuronal que ya fue construida.

Pronóstico a Futuro

Cargamos la data y la red ya entrenada.



```
1 % {  
2     Código Fuente del Pronóstico a Futuro del Modelo de  
3     Redes Neuronales  
4 % }  
5  
6 %Cargamos la data y la red neuronal ya entrenada  
7 load redagua.mat  
8
```

Figura 15. Código para cargar la Data Entrenada

Luego ingresamos un archivo en Excel de los últimos 12 meses anteriores a lo que se quiere predecir.

```
1 %{\n2     Código Fuente del Pronóstico a Futuro del Modelo de\n3     Redes Neuronales\n4 %}\n5\n6 %Cargamos la data y la red neuronal ya entrenada\n7 load redagua.mat\n8\n9 data = xlsread('F:\Yasmin!\TESIS\prueba-demanda\Futuro.xls');\n10
```

Figura16. Los últimos 12 meses.

Una vez cargada la red, en seguida procedemos a preparar los argumentos necesarios para la simulación de la red. Creamos un primer arreglo llamado “P”, el cual servirá para almacenar el conjunto de entrada el cual empieza en el registro 12. Al segundo arreglo lo llamaremos “PP”, representa los 12 valores iniciales que tendrá el componente de retardo:

```
8 data = xlsread('F:\Yasmin!\TESIS\prueba-demanda\Futuro.xls');\n9 dataEntren = con2seq(data);\n10\n11 P = dataEntren(1:12);\n12 PP = dataEntren(1:12);\n13\n14 prediccion = sim(red,P,PP);\n15\n16\n17
```

Figura17. Parámetros de la Red

Finalmente procedemos con la predicción para el siguiente año.

```

8 - data = xlsread('F:\Yasmin!\TESIS\prueba-demanda\Futuro.xls');
9 - dataEntren = con2seq(data);
10
11 - P = dataEntren(1:12);
12 - PP = dataEntren(1:12);
13
14 - prediccion = sim(red,P,PP);
15
16
17

```

Figura18. Código para la predicción del siguiente año

El código completo puede consultarse en el Anexo III.

Los valores del consumo de agua para la predicción futura, se muestran en la tabla 18.

Año 2011	Predicción de la Red Neuronal
Enero	176,329
Febrero	170,120
Marzo	172,444
Abril	161,237
Mayo	160,757
Junio	165,203
Julio	165,929
Agosto	166,678
Septiembre	163,036
Octubre	165,641
Noviembre	167,536
Diciembre	169,867
Total	2,004,777.00

Tabla 18. Predicción Futura.

Por lo tanto, el consumo global del agua previsto para el año 2011, se estima en 2, 004,777.0 m³.

VII. CONCLUSIONES

Mediante la aplicación del software desarrollado en la presente tesis en la empresa EPSEL S.A, durante el tiempo en que se han realizado las pruebas se concluye que:

Mediante la aplicación de técnicas de Inteligencia Artificial como las Redes Neuronales Artificiales con Arquitectura Perceptrón Multicapa es posible predecir la cantidad demandada de agua potable a mediano plazo.

Se logró determinar que el error del modelo de predicción de la cantidad demandada de agua potable a mediano plazo aplicando Redes Neuronales Artificiales es significativamente menor que el obtenido por la empresa.

En lo que respecta a los Costos de Operación y Mantenimiento se ahorraría S/. 1, 600,509.59 nuevos soles al emplear Redes Neuronales Artificiales.

Las Redes Neuronales Artificiales poseen un valor como método eficaz y potencial de predicción aplicable a funciones de planificación en la alta dirección para empresas del sector saneamiento.

La construcción de los modelos de la red para el pronóstico demostró un alto nivel de efectividad y una alta capacidad de adaptación adecuándose al tipo de problema que se modeló. En general, los resultados obtenidos en el entrenamiento y generalización alcanzó un Error Cuadráticos Medio muy bajo. Se demuestra un alto nivel de acierto para el pronóstico.

VIII. REFERENCIAS BIBLIOGRAFICAS

- Bonilla Musoles, M, y R Puertas Medina. «Análisis de las Redes Neuronales: Aplicación a problemas de predicción y clasificación financiera.» Facultad de C. Economía y Empresariales, Universidad de Valencia, Valencia, 1997.
- Bowerman, B, R O' Connell, y A Koehler. *Pronósticos, Series Temporales y Regresión*. Cuarta. Mexico: Internacional Thomson Editores, 2007.
- Chung-Ming, Kuan, y Halbert White. «Artificial Neural Network an Econometric Perspective.» *Econometric Reviews*, 1994: 1-91.
- Coca Carasila, Andrés, y Juan Villagómez Méndez. «La Demanda de Telefonía Fija y Móvil: Una Aplicación de Redes Neuronales Artificiales.» *Cuadernos de Gestión* Vol. 9, nº 2 (2009): 55-72.
- Daza Sánchez, Francisca. «Demanda de Agua en Zonas Urbanas en Andalucía.» Tesis Doctoral, Departamento de Estadística, Econometría, I.O y Organización de Empresas, Universidad de Córdoba, Córdoba, 2008.
- Dorffner, G s.f. «Neural Network for Times Series Processing.» Departament of Medical Cybernetics and Artificial Intelligence, University of Viena and Austrian Research Institute for Artificial Intelligence, 2008.
- Figueroa De La Cruz, Mario, Pablo Rovarini Diaz, Ignacio Del Monte, y Claudia Solórzano. «Aplicación de redes expertas para la predicción de tendencias del mercado financiero.» Proyecto de Investigación, Universidad de Palermo, Buenos Aires, 2007.
- Galán Ordax, José M. «Evaluación Integradora de Políticas de Agua: Modelado y Simulación con Sociedades Artificiales de Agentes.» Tesis Doctoral, Burgos, 2007.
- Haykin, Simon. *Neural Networks: A Comprehensive Foundation*. Segunda. New Jersey: Pretice Hall, 1999.
- Hilera González, José, y Víctor Martínez Hernando. *Redes Neuronales Artificiales: Fundamentos, Modelos y Aplicaciones*. Madrid: Ra-ma, 2000.
- Ineiaragón Torre, A, y Et. Al. «Aplicaciones de redes neuronales en Economía.» *Actas V Jornadas ESEPUMA*, 1997.
- Kaastra, L, y M Boyd. «Designing a Neural Network for forecasting Financial and Economic Times Series.» *Neurocomputing*, 1996: 215-236.

- Lawrence, Ramon. «Using Neural Networks to Forecast Stock Market Prices.»
Department of Computer Science, University of Manitoba, Canada, 1997.
- Marshall, Alfred. «Water as an Element of National Wealth.» New York, 1999.
- Moutinho, L. «Efeitos Interactivos de Planeamiento Estratégico de Planeamento de Marketing Estratégico sobre a Performance: Uma Análise da Rede Neuronal.»
Revista Portuguesa de Marketing, nº 10 (2000): 53-62.
- Munir Andrés, Jail, y Misas Martha. «Evaluación de pronósticos del tipo de cambio utilizando redes neuronales y funciones de pérdidas asimétricas.» *Revista Colombiana de Estadística* Vol. 30, nº 1 (Junio 2007): 143-161.
- Narváez García, Robin. «Redes Neuronales Artificiales y Regresión de Series Temporales para la Predicción de la cantidad demandada de Energía Eléctrica por Empresas del Sector Energético Eléctrico.» Tesis, Trujillo, 2009.
- Naturalle. *Naturalle*. 24 de Marzo de 2007.
<http://naturalle.blogspot.com/2007/03/problemtica-del-agua-en-per.html>
(último acceso: 22 de Agosto de 2011).
- Sapag Chain, Nassir, y Reinaldo Sapag Chain. *Preparación y Evaluación de Proyectos*. Cuarta. McGraw-Hill Interamericana, 2003.
- Sarmiento Maldonado, Henry, y Walter Villa Acevedo. «Inteligencia Artificial en Pronóstico de Demanda de Energía Eléctrica: Una Aplicación en Optimización de Recursos Energéticos.» *Revista Colombiana de Tecnologías de Avanzada* Vol. 2, nº 12 (2008).
- Sunass. «Sunass.» *Estudio Tarifario*. Noviembre de 2009.
http://www.sunass.gob.pe/estudios_tarifarios.php (último acceso: 22 de Agosto de 2011).
- Toolbar, Alexa. *Solo Ciencia*. s.f. <http://www.solociencia.com/ecologia/problemtica-global-agua-estadisticas-clave.htm> (último acceso: 22 de Agosto de 2011).
- Toro Ocampo, Eliana, Diego Mejía Giraldo, y Harold Salazar Isaza. «Pronóstico de Ventas usando Redes Neuronales.» *Redalyc* Vol. X, nº 26 (Diciembre 2004): 25-30.
- White, S, J Robinson, D Cordell, M Jha, y G Milne. *Urban Water Demand Forecasting and Demand Management: Research Needs Review and Recommendations*. Water Services Association of Australia Occasional Paper 9., 2003.-

IX. ANEXO

Anexo I. Código Fuente de la Implementación del Modelo de Redes Neuronales Artificiales en MATLAB.

```
%{
    Código de Implementación del Modelo de Pronóstico de Consumo
de
    Agua usando Redes Neuronales
}%
data = xlsread('D:\Yasmin\TESIS\prueba-demanda\prueba.xls');

dataEntren = con2seq(data);
dataEntren = dataEntren(1:228);
red = newfftd(dataEntren,dataEntren,(1:12),9);

red.trainParam.epochs = 300;
red.trainParam.goal = 0.5;

%Permite ver la salida en consola cada 25 epocas
red.trainParam.show = 25;
red.trainParam.showCommandLine = 1;

P = dataEntren(13:end);
T = dataEntren(13:end);
PP = dataEntren(1:12);

[red,tr] = train(red,P,T,PP);

%Permite graficar la eficiencia en el entrenamiento
plotperform(tr);

pronostico = sim(red,P,PP);
reales = cell2mat(P);
pronostico = cell2mat(pronostico);
error = reales - pronostico;

%Gráfica de Regresión
pronostico = sim(red,P,PP);
plotregression(P(1,:),pronostico(1,:), 'l');
```

Anexo II. Código Fuente del Cálculo de Nivel de Error del Modelo de Redes Neuronales Artificiales.

```
%{
    Código Fuente del Cálculo de Nivel de Error del Modelo de
    Redes Neuronales
}%

%Cargamos la data y la red neuronal ya entrenada
data = xlsread('F:\Yasmin!!\TESIS\prueba-demanda\prueba.xls');
loadredagua.mat

dataEntren = con2seq(data);

%Los ultimos 24 meses + los 12 anteriores que sirviran
%como entrada de retardo
dataEntren = dataEntren(217:252);

%Creamos el conjunto de entrada
P = dataEntren(13:end);
PP = dataEntren(1:12);

prediccion = sim(red,P,PP);

%Valores de las ultimas 24 mediciones
reales = cell2mat(dataEntren(13:end));

prediccion = cell2mat(prediccion);
error = reales - prediccion;
disp('-----');
disp('');
disp('***Resultado del Modelo de Redes Neuronales Artificiales***');
disp(['El Error de Pronóstico Acumulativo es: ',num2str(sum(error))]);
disp('-----');
disp('');
```

Anexo III. Código Fuente del Cálculo del Pronóstico a Futuro del Modelo de Redes Neuronales Artificiales.

```
%{  
    Código Fuente del Pronóstico a Futuro del Modelo de  
    Redes Neuronales  
%}  
  
%Cargamos la data y la red neuronal ya entrenada  
  
loadredagua.mat  
data = xlsread('F:\Yasmin!!\TESIS\prueba-demanda\Futuro.xls');  
dataEntren = con2seq(data);  
  
P = dataEntren(1:12);  
PP = dataEntren(1:12);  
  
prediccion = sim(red,P,PP);
```