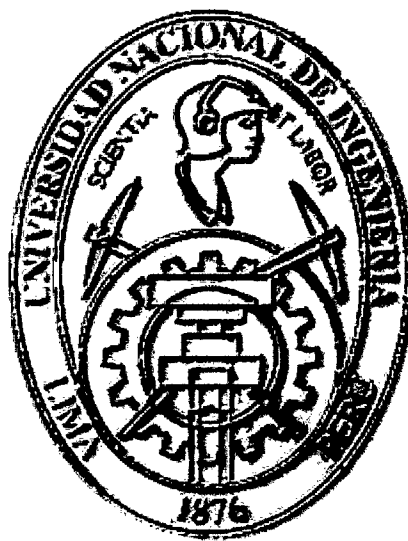


UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE INGENIERIA INDUSTRIAL Y DE SISTEMAS



**“PRONÓSTICO DE DEMANDA DE EFECTIVO EN CAJEROS
AUTOMÁTICOS APLICANDO REDES NEURONALES
ARTIFICIALES”**

TESIS

**PARA OBTENER EL TITULO PROFESIONAL DE:
INGENIERO DE SISTEMAS**

SILVIA GABRIELA GRANADOS FLORES

LIMA – PERÚ

2010

Digitalizado por:

**Consortio Digital del
Conocimiento MebLatam,
Hemisferio y Dalse**

DEDICATORIA

**Dedicado a todas las personas que
con perseverancia y valentía
consiguen sus sueños y hacen de
nuestra realidad un mundo
mucho mejor.**

AGRADECIMIENTO

Gracias a Dios absolutamente.

Gracias a mis padres Haydeé y Andrés por su cariño, ejemplo, apoyo, paciencia y toda representación que su infinito amor me alcanza.

Gracias a mi hermana Haydeé por su profundo amor, ejemplo y apoyo en cada uno de mis emprendimientos.

Gracias a todas las personas valiosas que me rodean día a día cuyo ejemplo y cariño me alientan a seguir alcanzando mis metas.

INDICE

DESCRIPTORES TEMÁTICOS.....	15
RESUMEN	16
INTRODUCCIÓN.....	17
1.1. DESCRIPCIÓN DEL PROBLEMA.....	20
1.2. OBJETIVOS DEL ESTUDIO	23
1.3. ALCANCE Y LIMITACIONES.....	23
1.4. METODOLOGÍA	24
CAPÍTULO I	30
MARCO TEÓRICO.....	30
1.1. ANTECEDENTES.....	30
1.2. PROCEDIMIENTO ACTUAL DE ABASTECIMIENTO DE EFECTIVO	31
1.3. MARCO TEÓRICO CONCEPTUAL	32
1.3.1. CAJERO AUTOMÁTICO (ATM)	32
1.3.1.1. HARDWARE.....	33
1.3.1.2. SOFTWARE	35
1.3.2. SERIES DE TIEMPO.....	36
1.3.3. PRONÓSTICO EN SERIES DE TIEMPO.....	37
1.3.3.1. MODELOS DE PRONÓSTICO	38
1.4. MARCO TEÓRICO INSTRUMENTAL.....	42
1.4.1. REDES NEURONALES ARTIFICIALES.....	42
1.4.1.1. ESTRUCTURA DE UNA RED NEURONAL	44
1.4.1.2. TIPOLOGÍAS DE REDES NEURONALES.....	46

1.4.1.3. MODELO NARX	48
1.4.2. REDES NEURONALES PARA SERIES DE TIEMPO.....	52
CAPÍTULO II	54
ANÁLISIS DE DATOS	54
2.1 DESCRIPCIÓN GENERAL	54
2.2 DESCRIPCIÓN DE DATOS.....	55
2.2.1 CORRESPONDENCIA ENTRE NÓMINAS DE BILLETES Y CASETERAS DEL ATM	55
2.2.2 STOCK DE BILLETES EN CASETERAS	57
2.3 PEDIDOS DE ABASTECIMIENTO.....	69
2.4 DEMANDA DE EFECTIVO	77
2.5 RESUMEN DEL CAPÍTULO	99
CAPÍTULO III	100
LIMPIEZA DE DATOS.....	100
3.1. DESCRIPCIÓN GENERAL.....	100
3.2. MÓDULO LIMPIEZA DE DATOS	101
3.2.1. ANÁLISIS DE VALORES IMPRECISOS	101
3.2.2. ANÁLISIS DE VALORES NULL.....	107
3.2.3. ANÁLISIS DE OUTLIERS.....	119
3.3. RESUMEN DEL CAPÍTULO	124
CAPÍTULO IV.....	126
MODELO DE SOLUCIÓN	126
4.1. DESCRIPCIÓN GENERAL	126
4.2. PLANTEAMIENTO DE LA SOLUCIÓN	126
4.3. RESUMEN DEL CAPÍTULO	129
CAPÍTULO V	131

PREPARACIÓN DE DATOS	131
5.1. DESCRIPCIÓN GENERAL	131
5.2. PREPARACIÓN DE DATOS	132
5.3. NORMALIZACIÓN DE DATOS	136
5.4. RESUMEN DEL CAPÍTULO	138
CAPÍTULO VI	139
DESFASE DE DATOS	139
6.1. DESCRIPCIÓN GENERAL	139
6.2. ARQUITECTURA DEL MODELO DE RNA	140
6.2.1. NÚMERO DE CAPAS	140
6.2.2. NEURONAS DE ENTRADA	142
6.2.3. NEURONAS DE SALIDA	144
6.3. CORRESPONDENCIA ENTRE NEURONAS DE ENTRADA Y SALIDA	147
6.4. RESUMEN DEL CAPÍTULO	150
CAPITULO VII	152
EVALUACIÓN DEL MODELO PROPUESTO	152
7.1. DESCRIPCIÓN GENERAL	152
7.2. MEDICIÓN DEL ERROR	153
7.3. RESUMEN DEL CAPÍTULO	154
CAPÍTULO VIII	156
EXPERIMENTACIÓN	156
8.1. DESCRIPCIÓN GENERAL	156
8.2. DETERMINACIÓN DE CARTERAS	157
8.3. CASOS DE PRUEBA	158
8.4. FUNCIÓN DE TRANSFERENCIA	161
8.5. FUNCIÓN DE ACTIVACIÓN	164

8.6.	FUNCIÓN DE APRENDIZAJE	170
8.7.	RATIO DE APRENDIZAJE.....	177
8.8.	RATIO MOMENTUM.....	184
8.9.	DÍAS HISTÓRICOS	191
8.10.	NEURONAS OCULTAS	197
8.11.	DÍAS PRONOSTICADOS.....	203
8.12.	ERROR FINAL DE VALIDACIÓN	209
8.13.	RESUMEN DEL CAPÍTULO.....	211
	CONCLUSIONES Y RECOMENDACIONES	212
	GLOSARIO DE TÉRMINOS	218
	BIBLIOGRAFÍA.....	221

INDICE DE TABLAS

Tabla 1	: Antecedentes-Trabajos de investigación.....	31
Tabla 2	: Correspondencia entre nóminas de billete y caseteras	57
Tabla 3	: Estructura ejemplo de Stock de Billetes	59
Tabla 4	: Registros recolectados de stock de billetes.....	60
Tabla 5	: Estructura ejemplo de Pedidos de Abastecimiento	70
Tabla 6	: Registros recolectados sobre pedidos de abastecimiento.....	71
Tabla 7	: Estructura ejemplo de Demanda de Efectivo	78
Tabla 8	: Determinación del periodo para cálculo de la demanda	80
Tabla 9	: Procedimiento para cálculo de demanda de efectivo	82
Tabla 10	: Resultado del cálculo de demanda de efectivo	83
Tabla 11	: Número de incidentes reportados por cajero automático	95
Tabla 12	: Indicadores estadísticos e histogramas – Cajeros A y B.....	97

Tabla 13	: Indicadores estadísticos e histogramas – Cajeros C, D y E	98
Tabla 14	: Identificación de registros imprecisos.....	102
Tabla 15	: Número de valores imprecisos de la muestra	102
Tabla 16	: Evaluación del reemplazo de datos imprecisos.....	107
Tabla 17	: Número de valores vacíos de la muestra	108
Tabla 18	: Técnica 1 Promedio diario - Ejemplo.....	109
Tabla 19	: Técnica 1 Promedio diarios - Promedios.....	110
Tabla 20	: Técnica 1 Promedio diarios - Evaluación	110
Tabla 21	: Técnica 2 Media Móvil 2 - Ejemplo.....	111
Tabla 22	: Técnica 2 Media Móvil 2 - Evaluación	112
Tabla 23	: Técnica 3 Media Móvil 3 - Ejemplo.....	113
Tabla 24	: Técnica 3 Media Móvil 3 - Evaluación	113
Tabla 25	: Técnica 4 Interpolación - Ejemplo	115
Tabla 26	: Técnica 4 Interpolación - Evaluación.....	115
Tabla 27	: Reemplazo valores vacíos – Impacto Media	116
Tabla 28	: Reemplazo valores vacíos – Impacto Mediana	116
Tabla 29	: Reemplazo valores vacíos – Impacto Mediana	117
Tabla 30	: Reemplazo valores vacíos – Técnicas empleadas.....	118
Tabla 31	: Identificación de outliers - Ejemplo.....	120
Tabla 32	: Identificación de outliers – Rango de confianza	120
Tabla 34	: Reemplazo de outliers	122
Tabla 33	: Tratamiento de outliers en datos reales	124
Tabla 35	: Equivalencias para día de la semana.....	132
Tabla 36	: Día de la semana - Ejemplo	133
Tabla 37	: Día laborable	133

Tabla 38	: Equivalencias para Día laborable.....	134
Tabla 39	: Día laborable - Ejemplo.....	134
Tabla 40	: Monto de demanda de efectivo - Ejemplo	135
Tabla 41	: Transformación de datos - Ejemplo.....	138
Tabla 42	: Casos de prueba para experimentación.....	159
Tabla 43	: Función de Transferencia – Cajero automático A.....	162
Tabla 44	: Función de Transferencia – Cajero automático B.....	162
Tabla 45	: Función de Transferencia – Cajero automático C	162
Tabla 46	: Función de Transferencia – Cajero automático D	163
Tabla 47	: Función de Transferencia – Cajero automático E.....	163
Tabla 48	: Función de Transferencia – Valores óptimos	163
Tabla 49	: Función de Activación – Error promedio mínimo	170
Tabla 50	: Valores óptimos de parámetros	170
Tabla 51	: Función de Aprendizaje – Error promedio mínimo	177
Tabla 52	: Valores óptimos de parámetros	177
Tabla 53	: Ratio de Aprendizaje – Error promedio mínimo.....	184
Tabla 54	: Valores óptimos de parámetros	184
Tabla 55	: Ratio Momentum – Error promedio mínimo.....	191
Tabla 56	: Valores óptimos de parámetros	191
Tabla 57	: Días históricos – Error promedio mínimo	197
Tabla 58	: Valores óptimos de parámetros	197
Tabla 59	: Neuronas ocultas – Error promedio mínimo	203
Tabla 60	: Valores óptimos de parámetros	203
Tabla 61	: Días pronosticados – Error promedio mínimo	209
Tabla 60	: Valores óptimos de parámetros	209

Tabla 63	: Error promedio final de validación.....	210
----------	---	-----

INDICE DE GRÁFICOS

Gráfico 1	: Instalaciones de ATM's alrededor del Mundo, 2005-2011	17
Gráfico 2	: Instalaciones de ATM's por Región, 2005-2011	18
Gráfico 3	: Problemas alrededor del proceso de abastecimiento	20
Gráfico 4	: Metodología CRISP-DM (Procesos jerárquicos).....	26
Gráfico 5	: Metodología CRISP-DM.....	27
Gráfico 6	: Información del abastecimiento actual de efectivo.....	32
Gráfico 7	: Hardware de un ATM	34
Gráfico 8	: Serie de tiempo	36
Gráfico 9	: Estructura de una Red Neruonal	44
Gráfico 10	: Modelo NARX.....	49
Gráfico 11	: Modelo NARX - Regresión	50
Gráfico 12	: Análisis de Datos.....	54
Gráfico 13	: Distribución de billetes en caseteras del ATM	56
Gráfico 14.1	: Cajero A - Stock de billetes casetera S/.100	61
Gráfico 14.2	: Cajero A - Stock de billetes casetera S/.20	62
Gráfico 14.3	: Cajero A - Stock de billetes casetera US\$20.....	62
Gráfico 15.1	: Cajero B - Stock de billetes casetera S/.100	63
Gráfico 15.2	: Cajero B - Stock de billetes casetera S/.20	63
Gráfico 15.3	: Cajero B - Stock de billetes casetera US\$100.....	64
Gráfico 16.1	: Cajero C - Stock de billetes casetera S/.100.....	64
Gráfico 16.2	: Cajero C - Stock de billetes casetera S/.20.....	65
Gráfico 16.3	: Cajero C - Stock de billetes casetera US\$20	65

Gráfico 17.1	: Cajero D - Stock de billetes casetera S/.100	66
Gráfico 17.2	: Cajero D - Stock de billetes casetera S/.20	66
Gráfico 17.3	: Cajero D - Stock de billetes casetera US\$20	67
Gráfico 18.1	: Cajero E - Stock de billetes casetera S/.100	67
Gráfico 18.2	: Cajero E - Stock de billetes casetera S/.20	68
Gráfico 18.3	: Cajero E - Stock de billetes casetera US\$ 20.....	68
Gráfico 19	: Cajero A - Abastecimiento y remanente de efectivo	72
Gráfico 20	: Cajero B - Abastecimiento y remanente de efectivo	73
Gráfico 21	: Cajero C – Abastecimiento y remanente de efectivo	74
Gráfico 22	: Cajero D – Abastecimiento y remanente de efectivo	75
Gráfico 23	: Cajero E – Abastecimiento y remanente de efectivo.....	76
Gráfico 24	: Extracción de registros sobre stock de efectivo en ATM.....	79
Gráfico 25.1	: Cajero A - Demanda de billetes	85
Gráfico 25.2	: Cajero A - Demanda de billetes	85
Gráfico 25.3	: Cajero A - Demanda de billetes	86
Gráfico 25.4	: Cajero A - Demanda monto total.....	86
Gráfico 26.1	: Cajero B - Demanda de billetes	87
Gráfico 26.2	: Cajero B - Demanda de billetes	87
Gráfico 26.3	: Cajero B - Demanda de billetes	88
Gráfico 26.4	: Cajero B - Demanda de billetes	88
Gráfico 27.1	: Cajero C - Demanda de billetes	89
Gráfico 27.2	: Cajero C - Demanda de billetes	89
Gráfico 27.3	: Cajero C - Demanda de billetes	90
Gráfico 27.4	: Cajero C - Demanda de billetes	90
Gráfico 28.1	: Cajero D - Demanda de billetes	91

Gráfico 28.2	: Cajero D - Demanda de billetes	91
Gráfico 28.3	: Cajero D - Demanda de billetes	92
Gráfico 28.4	: Cajero D - Demanda de billetes	92
Gráfico 29.1	: Cajero E - Demanda de billetes	93
Gráfico 29.2	: Cajero E - Demanda de billetes	93
Gráfico 29.3	: Cajero E - Demanda de billetes	94
Gráfico 29.4	: Cajero E - Demanda de billetes	94
Gráfico 30	: Incidentes en ATM's	96
Gráfico 31	: Análisis de Datos - Resumen	99
Gráfico 32	: Limpieza de Datos.....	101
Gráfico 33	: Cálculo sobre registros imprecisos	106
Gráfico 34	: Reemplazo de outliers.....	123
Gráfico 35	: Limpieza de Datos - Resumen	125
Gráfico 36	: Análisis y limpieza de Datos	126
Gráfico 37	: Proceso de solución	127
Gráfico 38	: Planteamiento de la Solución	130
Gráfico 39	: Preparación de datos	131
Gráfico 40	: Preparación de datos - Resumen	138
Gráfico 41	: Desfase de datos	139
Gráfico 42	: NARX – Red retroalimentación.....	141
Gráfico 43	: Arquitectura de la RNA.....	142
Gráfico 44	: Correspondencia entre neuronas de entrada y salida.....	147
Gráfico 45	: Días históricos.....	147
Gráfico 46	: Días a pronosticar	149
Gráfico 47	: Desfase de datos - Resumen	151

Gráfico 48	: Evaluación del modelo propuesto.....	153
Gráfico 49	: Evaluación del modelo propuesto.....	155
Gráfico 50	: Experimentación del modelo propuesto.....	157
Gráfico 51	: Determinación de carteras.....	158
Gráfico 52	: Función de Activación – Cajero automático A.....	165
Gráfico 53	: Función de Activación – Cajero automático B.....	166
Gráfico 54	: Función de Activación – Cajero automático C	167
Gráfico 55	: Función de Activación – Cajero automático D	168
Gráfico 56	: Función de Activación – Cajero automático E.....	169
Gráfico 57	: Función de Aprendizaje – Cajero automático A.....	172
Gráfico 58	: Función de Aprendizaje – Cajero automático B.....	173
Gráfico 59	: Función de Aprendizaje – Cajero automático C.....	174
Gráfico 60	: Función de Aprendizaje – Cajero automático D.....	175
Gráfico 61	: Función de Aprendizaje – Cajero automático E.....	176
Gráfico 62	: Ratio de Aprendizaje – Cajero automático A	179
Gráfico 63	: Ratio de Aprendizaje – Cajero automático B	180
Gráfico 64	: Ratio de Aprendizaje – Cajero automático C	181
Gráfico 65	: Ratio de Aprendizaje – Cajero automático D.....	182
Gráfico 66	: Ratio de Aprendizaje – Cajero automático E	183
Gráfico 67	: Ratio Momentum – Cajero automático A.....	186
Gráfico 68	: Ratio Momentum – Cajero automático B	187
Gráfico 69	: Ratio Momentum – Cajero automático C.....	188
Gráfico 70	: Ratio Momentum – Cajero automático D.....	189
Gráfico 71	: Ratio Momentum – Cajero automático E	190
Gráfico 72	: Días Históricos – Cajero automático A	192

Gráfico 73	: Días Históricos – Cajero automático B	193
Gráfico 74	: Días Históricos – Cajero automático C	194
Gráfico 75	: Días Históricos – Cajero automático D	195
Gráfico 76	: Días Históricos – Cajero automático E	196
Gráfico 77	: Neuronas Ocultas – Cajero automático A.....	198
Gráfico 78	: Neuronas Ocultas – Cajero automático B.....	199
Gráfico 79	: Neuronas Ocultas – Cajero automático C.....	200
Gráfico 80	: Neuronas Ocultas – Cajero automático D.....	201
Gráfico 81	: Neuronas Ocultas – Cajero automático E.....	202
Gráfico 82	: Días Pronosticados – Cajero automático A	204
Gráfico 83	: Días Pronosticados – Cajero automático B	205
Gráfico 84	: Días Pronosticados – Cajero automático C	206
Gráfico 85	: Días Pronosticados – Cajero automático D	207
Gráfico 86	: Días Pronosticados – Cajero automático E	208

DESCRIPTORES TEMÁTICOS

Redes Neuronales Artificiales

Cajeros Automáticos

Series de Tiempo

Pronóstico

Demanda de Efectivo

Empresa Financiera

RESUMEN

Esta investigación propone un modelo basado en Redes Neuronales Artificiales del tipo Autorregresivo No-Lineal con Entradas Exógenas (modelo NARX) capaz de pronosticar la demanda de efectivo en cajeros automáticos basado en la evolución de variables en serie de tiempo. El desarrollo de esta investigación se efectúa sobre la metodología estándar CRISP-DM y alineado a los procesos jerárquicos que se identifican en su despliegue. Para la construcción de este modelo se obtuvieron registros sobre stock de billetes y pedidos de abastecimiento históricos con una frecuencia diaria, lo que permitió calcular la demanda histórica de efectivo de cinco cajeros automáticos pertenecientes a la red de distribución de una Empresa Financiera Local. Luego de realizar el análisis, limpieza y transformación de los datos, se desarrolló el modelo de solución propuesto con el que se logró construir la red neuronal artificial conformada por una capa de entrada, una oculta y una de salida. Adicionalmente se realizó la experimentación del modelo, ejecutándose 138,240 corridas del sistema por cada cajero automático analizado, identificándose los valores óptimos para los parámetros establecidos en el modelo. En base a las pruebas realizadas se identificó el diseño óptimo del sistema propuesto, obteniendo un error MAPE promedio de validación en el pronóstico entre 7% y 15%, resultado de conjugar los valores óptimos identificados. Asimismo se detallan las conclusiones y recomendaciones a tomar en cuenta para investigaciones futuras de la índole de Redes Neuronales Artificiales en series de tiempo.

INTRODUCCIÓN

Actualmente la competitividad en la industria financiera exige a las empresas ofrecer canales confiables y totalmente disponibles.

Para una empresa de servicios financieros contar con una red de cajeros automáticos (ATM's – siglas en inglés) localizada en puntos estratégicos y que se encuentre operativa las 24 horas del día, significa calidad y servicio que satisface las necesidades del cliente y asimismo, reflejan una gestión efectiva sobre los recursos disponibles para este canal.

En el Gráfico 1 se ilustra la evolución del mercado de cajeros automáticos (ATM's) a nivel mundial y su proyección hasta el 2011 (P. Nicholls [7]).

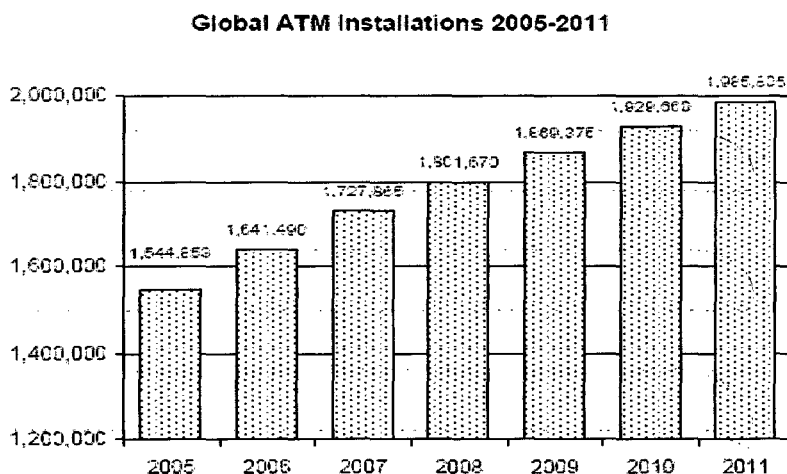


Gráfico 1 : Instalaciones de ATM's alrededor del Mundo, 2005-2011
Fuente : Global ATM Market and Forecasts to 2011
Elaboración : Retail Banking Research Team

En el Gráfico 2 se presenta una comparación entre datos históricos del 2005 y un pronóstico para el 2011 sobre una distribución regional alrededor del mundo (P. Nicholls [7]).

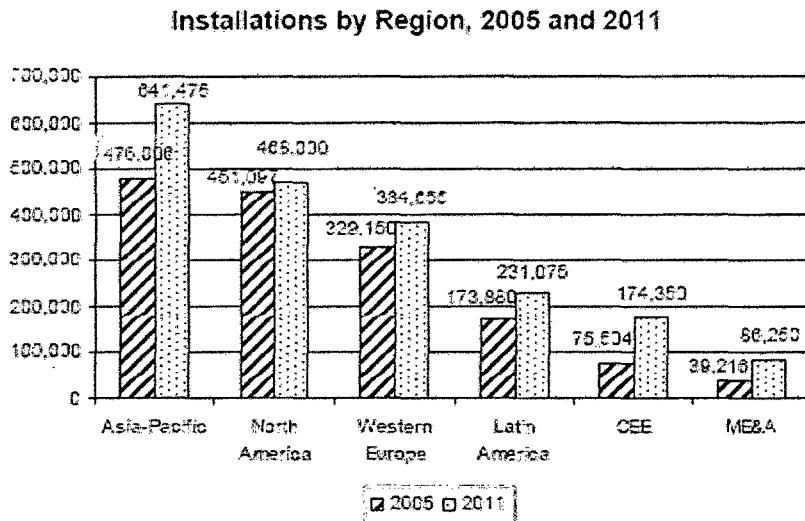


Gráfico 2 : Instalaciones de ATM's por Región, 2005-2011
Fuente : Global ATM Market and Forecasts to 2011
Elaboración : Retail Banking Research Team

Se observa un incremento significativo en el total de instalaciones de cajeros automáticos (ATM's) en todas las regiones mencionadas.

Sin embargo debe considerarse que una ampliación en el número de cajeros automáticos (ATM's) para una organización, significa un incremento de recursos empleados en su operativa. En este sentido el nivel de efectivo inmovilizado en las caseteras de estas máquinas, principal recurso de su operativa, sería el principal impactado (G. Lepecq [4]).

La Inteligencia Artificial, específicamente las Redes Neuronales Artificiales (ANN, siglas en inglés) permiten emplear técnicas y desarrollar modelos capaces de identificar patrones de comportamiento.

Una vez obtenidos estos patrones, es posible pronosticar el posible comportamiento para futuras condiciones (A. Jain et. al. [20]).

Últimamente las empresas de servicios financieros alrededor del mundo emplean diversas soluciones informáticas para pronosticar la demanda de efectivo en su red de cajeros automáticos (ATM's). Entre ellas, se emplean las Redes Neuronales Artificiales, algunos ejemplos de estas soluciones:

- a. **Gescash**© de la empresa Gestran S.A.(Bandera 84 Oficina 309, Santiago de Chile - Chile)
- b. **ATM Network Optimization Suite**© de la empresa Penkiu Kontinentu Bankines Technologijos, UAB (Kalvariju str. 143, Vilnius – Lithuania)
- c. **Diebold ATM Cash Management**© de la empresa Diebold, Incorporated (North Canton, Ohio 44720-8077 – EEUU)
- d. **Cash Optimization & Forecasting Solutions**© de la empresa Fiserv, Inc. (255 Fiserv Drive, Brookfield – EEUU)
- e. **MorphisCM**© de la empresa Morphis, Inc. (6060 North Central Expressway, Suite 442 Dallas, TX 75206 – EEUU)

Además en el mercado se emplean otras técnicas que cuentan con reconocimiento mundial por la performance de sus resultados, tal como el Modelo Autorregresivo Integrado de Media Móvil (ARIMA, siglas en inglés). ARIMA es un modelo estadístico que utiliza variaciones y regresiones de datos estadísticos con el fin de encontrar patrones para su predicción (P. Rossini [2]). Actualmente el Banco de Reserva de Nueva Zelanda emplea una herramienta basada en esta técnica estadística para pronosticar la demanda de moneda nacional y planificar su producción.

Finalmente, esta tesis tiene como objetivo emplear los beneficios de un sistema basado en Redes Neuronales Artificiales para pronosticar la demanda de efectivo en cajeros automáticos (ATM's), tomando como referencia los patrones de comportamiento sobre información histórica (R. Poler et. al. [11]).

1.1. DESCRIPCIÓN DEL PROBLEMA

El proceso de abastecimiento de efectivo en cajeros automáticos origina el análisis de cuatro problemas que la autora considera principales y representa en el Gráfico 3:

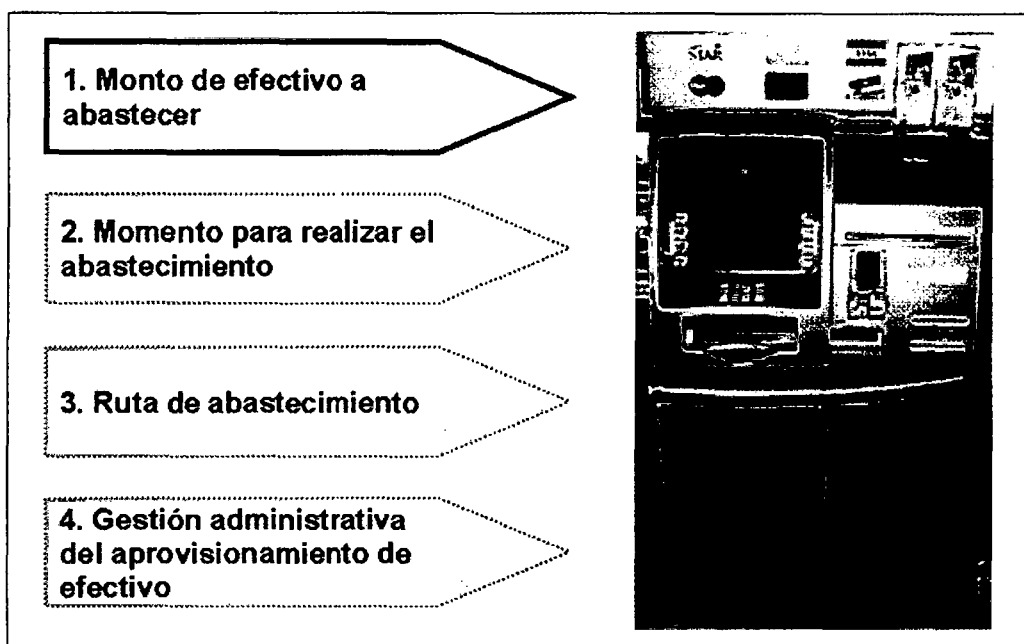


Gráfico 3 : Problemas alrededor del proceso de abastecimiento
Fuente : ATM Networks and Cash Usage
Elaboración : Silvia Granados Flores

1. Monto de efectivo a abastecer: Referido a determinar el monto óptimo de efectivo que permita atender la demanda de efectivo y reducir el remanente.

2. **Momento para realizar el abastecimiento:** Se refiere al punto en el tiempo para efectuar la descarga de remanente y carga de nuevo efectivo en las caseteras del cajero automático (ATM). De esta manera, se busca reducir el tiempo de inoperatividad de la máquina.
3. **Ruta de abastecimiento:** Referido a determinar la secuencia de abastecimiento entre los distintos puntos de la red de cajeros automáticos (ATM's) y así optimizar el costo del servicio de abastecimiento.
4. **Gestión administrativa del aprovisionamiento de efectivo:** Incluye la optimización de las condiciones establecidas para la planificación y control del proceso.

En esta investigación se propondrá una solución al primer problema planteado.

A fin de identificar el monto a abastecer, que optimice requerimientos futuros de efectivo, es necesario identificar tendencias sobre el comportamiento de la demanda para realizar el pronóstico en periodos siguientes.

La demanda de efectivo es una variable que se desarrolla sobre una serie temporal cuyo comportamiento resulta no lineal, lo que eleva el nivel de complejidad para el pronóstico (ver capítulo II).

La predicción de series de tiempo es un problema de investigación que ha sido abarcado con técnicas estadísticas, econométricas, inteligencia artificial, entre otras.

Dentro de las técnicas anteriormente mencionadas, las Redes Neuronales Artificiales (RNA) ofrecen ventajas sobre series no lineales, esto se debe a que los modelos basados en un sistema de RNA pueden aproximar una función continua sobre un dominio y son tolerantes ante el ruido, la información incompleta y los datos contradictorios.

Problema de investigación

¿Cómo determinar un modelo basado en Redes Neuronales Artificiales con el que se pueda pronosticar la demanda de efectivo en cajeros automáticos minimizando el error?

Considerando que el error será medido en base al estándar MAPE (ver capítulo 6).

Hipótesis

La hipótesis formulada para solucionar el problema planteado es con la identificación de variables independientes puede definirse un modelo basado en Redes Neuronales Artificiales que entrenado con data histórica válida, identifique patrones de comportamiento para estimar la demanda futura de efectivo en cajeros automáticos.

Preguntas de investigación

- ¿Cuáles son los factores a considerarse como *inputs* en la arquitectura del sistema de Redes Neuronales Artificiales?
- ¿Cuál es el número de días históricos que se requieren para pronosticar la demanda de efectivo en cajeros automáticos (ATM's), minimizando el error?
- ¿Cuál es el número de días sobre los que se puede pronosticar la demanda de efectivo en (ATM's), minimizando el error?
- Sobre un universo de datos referidos a la demanda histórica, ¿cuál es el error promedio que se obtendrá al experimentar con el modelo de Redes Neuronales Artificiales propuesto?

1.2. OBJETIVOS DEL ESTUDIO

A. Objetivo principal

Modelar un sistema basado en Redes Neuronales Artificiales capaz de pronosticar la demanda de efectivo en cajeros automáticos (ATM's) minimizando el error.

B. Objetivos secundarios

- Identificar las variables independientes (inputs del modelo) que impacten representativamente en el comportamiento de la demanda de efectivo.
- Determinar bajo técnicas de minería de datos, la data histórica óptima que permita identificar patrones de comportamiento en la demanda de efectivo.
- Identificar el error promedio del pronóstico de demanda de efectivo, luego de experimentado el modelo propuesto con la data histórica recolectada.

1.3. ALCANCE Y LIMITACIONES

Este informe de investigación detalla las actividades correspondientes al análisis, construcción, experimentación y resultados del modelo propuesto.

Para el análisis, se detalla el proceso actual de abastecimiento de efectivo en los cajeros automáticos (ATM's) y se identifica la información necesaria para la construcción del modelo. Dentro de esta información se cuenta con 726 registros referentes al stock de billetes en las caseteras de 5 cajeros automáticos (ATM's) durante el periodo del 01/12/2008 al 30/11/2009, generados en turnos de 8:45 a.m., 12:00 p.m. y 5:00 p.m. diariamente. Asimismo se cuenta con el detalle de los pedidos de abastecimiento efectuados durante el mismo periodo. Con la información recopilada se calculará la demanda diaria de efectivo desde

el 01/12/2008 al 30/11/2009. Una vez calculada la demanda, se procederá a reemplazar los valores *null*, imprecisos y *outliers* identificados en la serie de datos, para ello se emplearán técnicas de promedio global, media móvil y/o interpolación para luego transformar los datos a través de una función continua como el logaritmo neperiano. Estos datos finales serán empleados en el modelo a construir.

Para la construcción, se prepararán los datos identificando las variables independientes para la demanda de efectivo, así como los números tentativos de días históricos y de días a pronosticar. De esta manera, se desfazarán los datos siguiendo el orden de las series de tiempo y se determinará la arquitectura inicial del sistema (neuronas de entrada y neuronas de salida).

En la experimentación, se efectuarán las pruebas necesarias para la simulación del modelo, empleando posibles valores para los parámetros como: número de neuronas ocultas, función de transferencia, función umbral, función de aprendizaje, ratio de aprendizaje y momentum.

Una vez obtenidos los resultados de la experimentación, se identificará el conjunto de parámetros que genere el menor error de validación y de esta manera se obtendrá el modelo óptimo.

1.4. METODOLOGÍA

El modelo propuesto sobre un sistema de Redes Neuronales Artificiales generará información que será empleada en la toma de decisiones del proceso de abastecimiento de efectivo en cajeros automáticos (ATM's). Por este motivo, el desarrollo de esta investigación se incluye dentro de un proceso de minería de datos.

Este trabajo de investigación se desarrolla en base a la metodología estándar CRISP-DM, empleada para el diseño y construcción de soluciones enfocadas en la minería de datos.

La metodología CRISP-DM desarrolla el macroproceso de minería de datos desde la identificación de los requerimientos de información hasta la validación y explotación de la misma. Ésta a su vez, está conformada por un conjunto de procesos jerárquicos aplicables, los cuales se presentan en el Gráfico 4.

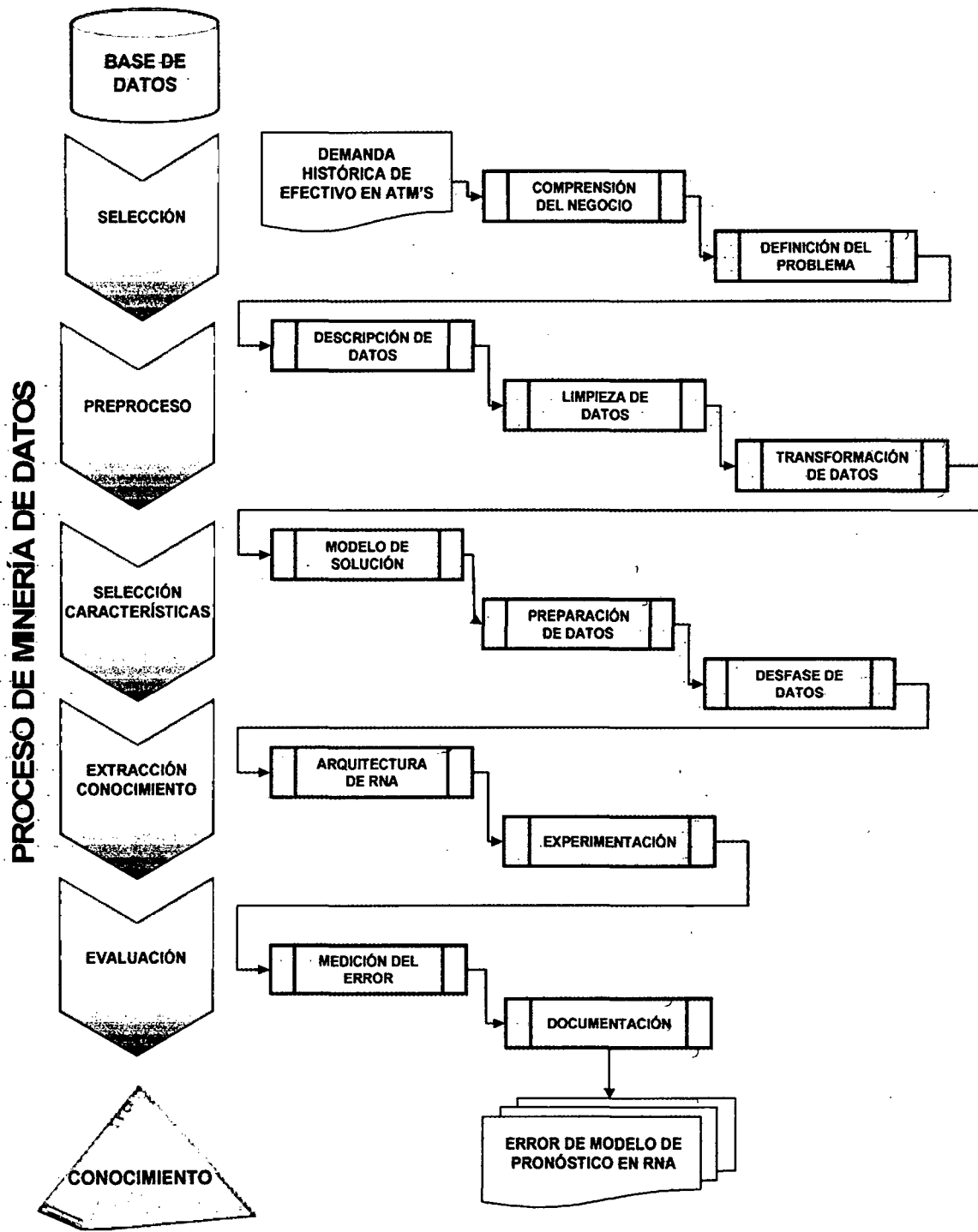


Gráfico 4 : Metodología CRISP-DM (Procesos jerárquicos)
Fuente : Aplicación de la minería de datos al Sistema Cubano de Farmacovigilancia
Elaboración : Propia

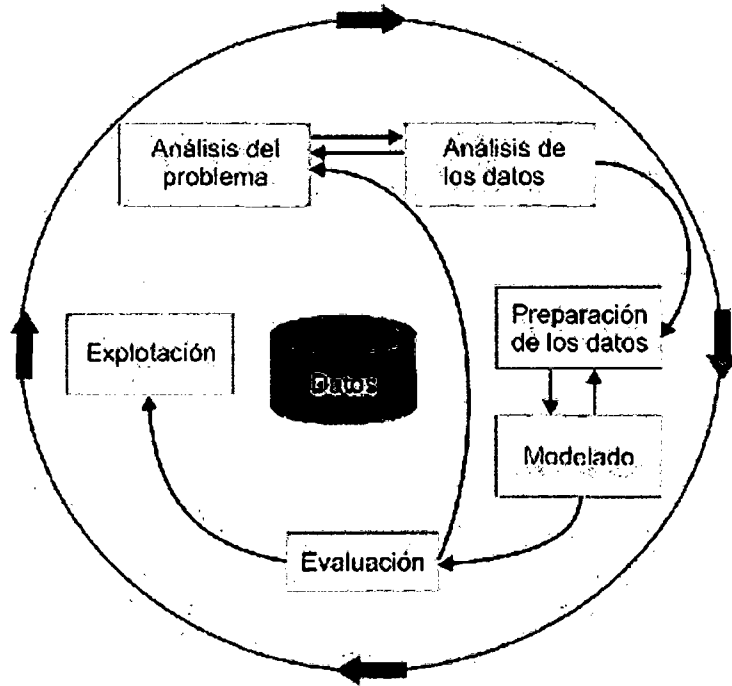


Gráfico 5 : Metodología CRISP-DM
 Fuente : Aplicación de la minería de datos al Sistema Cubano de Farmacovigilancia
 Elaboración : Omar Calzadilla Fernández de Castrol

En el Gráfico 5 representa la metodología CRISP-DM, la cual desarrolla el ciclo de vida de un proyecto de minería de datos en seis fases. Estas fases interactúan entre ellas de forma iterativa durante el desarrollo del proyecto.

En base a los Gráficos 4 y 5 se detallan las fases sobre las cuales se desarrollará la metodología CRIP-DM empleada en esta investigación (I. García Martín [3]):

- A. Fase Análisis del Problema (Procesos: Comprensión del Negocio y Definición del Problema): Comprensión de los objetivos y requerimientos del proyecto desde una perspectiva empresarial para luego ser traducidos en objetivos técnicos.

En el desarrollo del informe, esta fase se efectúa a través de la identificación del proceso actual de abastecimiento de efectivo en cajeros automáticos y el reconocimiento del problema (ver capítulo I).

B. Fase Análisis de Datos (Procesos: Descripción de Datos, Limpieza de Datos y Transformación de Datos): Recolección inicial de datos, identificación de la calidad de los mismos y establecimiento de las relaciones más evidentes que permitan establecer las primeras hipótesis.

En esta fase se emplearán los 726 registros sobre stock de billetes en las caseteras del cajero automático y los pedidos de abastecimiento efectuados dentro del periodo analizado y para el mismo cajero. Se efectuará el cálculo de la demanda histórica de efectivo para luego efectuar la limpieza (reemplazo de valores *null*, imprecisos y *outliers*) y transformación de los datos (ver Capítulo III)

C. Fase Preparación de Datos (Procesos: Modelo de Solución, Preparación de Datos y Desfase de Datos): En base al modelo de solución establecido, se seleccionarán y completarán las variables independientes a la demanda de efectivo (día del mes, día de la semana, sábado/lunes equivalente y día de pago) y que se consideran influyentes en su comportamiento. Finalmente en función a la serie de tiempo, se desfazarán los datos a fin de mantener la continuidad de la variable (ver Capítulo IV y V)

D. Fase Modelado (Procesos: Arquitectura RNA y Experimentación): En esta fase se determinará la arquitectura de la red considerando como *input* los datos analizados y preparados. Se efectuarán las experimentaciones de entrenamiento y validación del sistema, las cuales emplearán los valores identificados como probables para los parámetros del modelo (Ver capítulo VI)

E. Fase Evaluación (Proceso: Medición del Error): Evaluación del modelo desde el punto de vista del cumplimiento de los criterios de

éxito del problema. El criterio seleccionado en esta investigación es la medición del error a través del estándar MAPE (ver capítulo VII).

F. Fase Explotación (Proceso: Documentación): Documentación y presentación de resultados, asegurando el mantenimiento de la solución (ver capítulo VIII).

CAPÍTULO I

MARCO TEÓRICO

1.1. ANTECEDENTES

En la Tabla 1 se listan trabajos de investigación que desarrollan el tema del pronóstico sobre demanda de algún recurso, la misma que se ejecuta sobre una serie de tiempo:

TÍTULO	AUTOR	MODELO PRONÓSTICO	ERROR OBTENIDO
Predicción de la demanda eléctrica de corto plazo mediante Redes Neuronales Artificiales.	D. García, A. Milanes	Redes Neuronales Artificiales	Máximo: 9.55% Mínimo: 0.69%
Pronóstico del precio y demanda de energía eléctrica empleando Redes Neuronales Autorregresivas	D. Yamashita, A. Mohd, R. Yokohama, T. Niimura	Redes Neuronales Artificiales	Máximo: 13.06% Mínimo: 5.94%
Modelando la demanda de materiales de construcción empleando Redes Neuronales Artificiales	Z. Alhosan	Redes Neuronales Artificiales	Máximo: 4.19% Mínimo: 0.06%
Pronóstico de la demanda de agua usando Redes Neuronales Backpropagation en la operación de reservorios de Indonesia	M. Mashudi	Redes Neuronales Artificiales	Máximo: 7.75% Mínimo: 0.4%
Redes Neuronales para pronóstico de demanda de viajes	A. Dantas, K. Yamamoto, M.	Redes Neuronales	Máximo: 15.00% Mínimo: 10.00%

TÍTULO	AUTOR	MODELO PRONÓSTICO	ERROR OBTENIDO
usando GPS y sensores remotos	Lamar, Y. Yamashita	Artificiales	
Pronóstico de demanda turística en Taiwan usando Redes Neuronales Artificiales y Regresiones multivariadas	L. Chang-Jui, C. Hsueh-Fang, L. Tian-Shyug	Redes Neuronales Artificiales	Máximo: 13.32% Mínimo: 8.00%
Pronóstico de la demanda de servicios especiales de telefonía	G. Madden, J. Tan	ARIMA	Máximo: 18.30% Mínimo: 0.04%
Pronóstico de la demanda de electricidad empleando Arima y regresores	B. Ugiliweneza	ARIMA	Máximo: 14.82% Mínimo: 5.96%

Tabla 1 : Antecedentes-Trabajos de investigación
Fuente : Universidad de Guadalajara - CUCEI
Elaboración : Silvia Granados Flores

1.2. PROCEDIMIENTO ACTUAL DE ABASTECIMIENTO DE EFECTIVO

Actualmente la entidad financiera propietaria de la información empleada en esta investigación no posee un sistema automatizado para determinar los montos de abastecimiento de efectivo en su red de cajeros automáticos. El control se lleva manualmente a través de hojas de cálculo, las cuales se actualizan diariamente y almacenan información acerca del stock de billetes en las caseteras de los equipos. Como lo indica el Gráfico 6, para determinar el monto y la oportunidad de abastecimiento de efectivo, se toma como factor base el stock actual de billetes. Asimismo para definir si es necesario realizar el abastecimiento, se consideran las demandas históricas hasta tres meses anteriores (a juicio del analista podría identificarse alguna tendencia) y de ser el caso, el pronóstico enviado por el proveedor de

abastecimiento. Si el resultado fuese positivo, el monto a abastecerse será similar al monto efectuado en las últimas oportunidades.

Hoy en día, el procedimiento no considera casuísticas de series temporales, posibilidades de estacionalidad o tendencia.

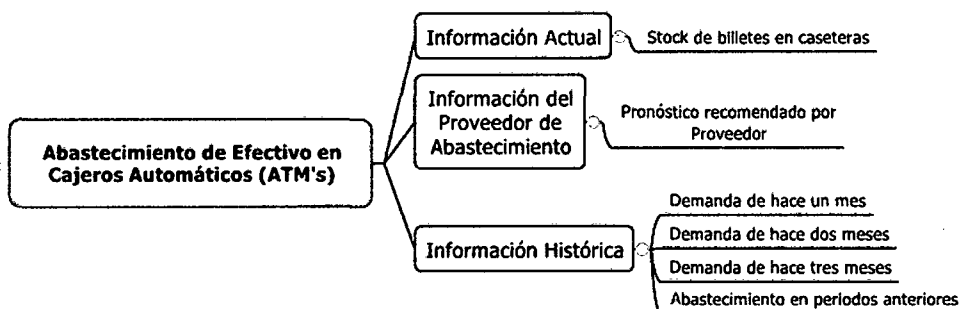


Gráfico 6 : Información del abastecimiento actual de efectivo

Fuente : Entidad Financiera Local

Elaboración : Silvia Granados Flores

1.3. MARCO TEÓRICO CONCEPTUAL

1.3.1. CAJERO AUTOMÁTICO (ATM)

Un cajero automático es una máquina expendedora usada para extraer dinero utilizando una tarjeta magnética.

Algunas transacciones financieras que pueden realizarse a través de un cajero automático:

- Retiro de efectivo de cuentas o tarjeta de crédito
- Abono en cuentas
- Transferencia entre cuentas propias o cuentas de terceros
- Consulta de saldos y movimientos
- Pago de servicios públicos
- Pago de tarjetas de crédito
- Obtención de contraseñas olvidadas de banca online o telefónica
- Recargas de cuentas de servicio

- Envío de remesas de dinero al extranjero

El uso del cajero automático genera ahorro en costos al banco porque reduce el empleo de personal en el registro de las transacciones anteriormente señaladas.

1.3.1.1. HARDWARE

Los dispositivos que constituyen un cajero automático son (H. Snellman [22]):

- CPU, el cual controla la interfase entre el usuario y los dispositivos que procesan la transacción.
- Lector de banda magnética o chip, el cual identifica al cliente.
- PIN Pad, dispositivo de seguridad de pantalla táctil o tipo calculadora para que el usuario introduzca claves de seguridad.
- Criptoprocesador de seguridad.
- Display, dispositivo para que el usuario efectúe sus transacciones.
- Botones funcionales o de pantalla táctil (los cuales están incluidos en el dispositivo de display)
- Impresora, la cual proporciona al usuario el resumen de sus transacciones.
- Bóveda, la cual almacena los billetes a dispensar y/o recibir así como las partes que permitirán estos mecanismos. Mecanismos que incluyen la estructura de la bóveda:
 - o Mecanismo dispensador (provee dinero o elementos de valor)
 - o Mecanismo de depósito, incluyendo modulo de procesamiento de cheques y notas de aceptación.
 - o Sensores de seguridad (magnéticos, térmicos, sísmicos)
 - o Llaves, los cuales aseguran acceso controlado a la bóveda.

- Sistema journal, puede ser sistema electrónico o un dispositivo de impresión, el cual registra el record de operaciones, número de billetes dispensados, etc.
- Cobertura, cuya dimensión y diseño depende de la estética del equipo.

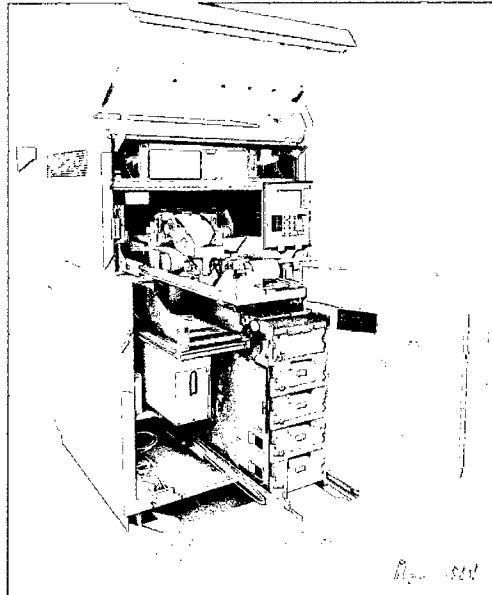


Gráfico 7 : Hardware de un ATM
Fuente : Automated Teller Machine
Elaboración : PrimroseGuy

Actualmente, debido a las nuevas funcionalidades requeridas para estos aparatos así como la reducción de los precios para arquitectura casi computacionales, los cajeros automáticos están adoptando arquitecturas que emplean microcontroladores y/o circuitos integrados con aplicaciones específicas, asimilándose a un computador personal. Existen modelos de cajeros automáticos capaces de emplear sistemas operativos como Microsoft Windows y Linux (D. Bounie et. al. [23]).

1.3.1.2. SOFTWARE

Con la migración del hardware a semi-computadoras personales, el software de un cajero automático puede basarse en sistemas operativos comerciales así como el ambiente de programación. Las plataformas mayormente empleadas en cajeros automáticos son RMX, OS/2 y sistemas operativos de Microsoft como MS-DOS, PC-DOS, Windows NT, Windows 2000, Windows XP Professional o Windows XP. Java, Linux y Unix también pueden ser empleados en este dispositivo.

Linux también está siendo empleado en el Mercado de cajeros automáticos. Por ejemplo, el banco Bansirul de Brasil reemplazó el MS-DOS de sus cajeros por Linux así como el Banco Do Brasil.

Protocolos para las capas de procesamiento de transacciones como Diebold 911 or 912, IBM PBM, and NCR NDC or NDC+ pueden trabajar con generaciones anteriores de hardware así como nuevas plataformas lo que permite construir nuevas funcionalidades, compañías como NCR mantiene innovación continua de sus protocolos generando nuevas versiones como NCR Aprta Advance NDC Version 3.x.y (donde x.y son subversiones).

Con la estandarización de los software base, instituciones financieras han incrementado sus investigaciones a fin de perfeccionar las aplicaciones de programas que les permite manejar sus dispositivos. WOSA/XFS o CEN XFS condiciona un API común para acceder y manipular los diferentes dispositivos del cajero automático.

Los cajeros automáticos que operan bajo la plataforma XFS emplean software como Triton PRISM, Diebold Agilis, CR2 BankWorld, KAL Kalignite, NCR Corporation Aprta Edge, Phoenix Interactive VISTAatm, and Wincor Nixdorf Protopas.

Con la industrialización de los dispositivos computacionales de los cajeros automáticos, las empresas comercializadoras desarrollan planes

a fin de perfeccionar la integridad y potencia del software implementado en la máquina.

1.3.2. SERIES DE TIEMPO

Se denomina Serie de Tiempo al conjunto de mediciones de cierto fenómeno o experimento registradas secuencialmente en el tiempo.

Estas observaciones pueden denotarse por:

$\{x(t_1), x(t_2), x(t_3), \dots, x(t_n)\} = \{x(t) : t \in T \subseteq R\}$ donde $x(t_i)$ es el valor de la variable x en el instante t_i .

El Gráfico 8 representa una serie de tiempo sobre la variable temperatura.

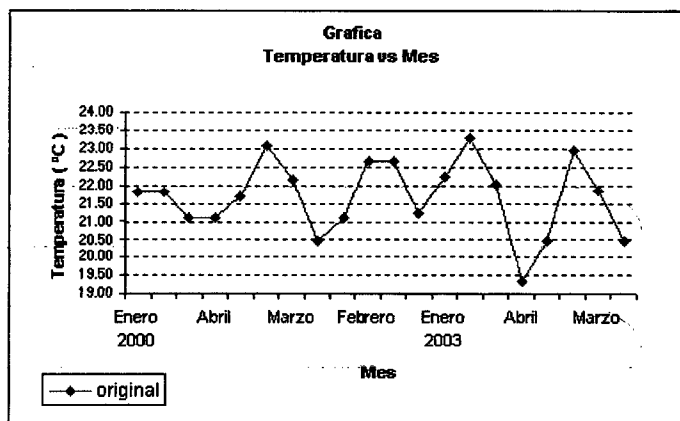


Gráfico 8 : Serie de tiempo
Fuente : Modelos de series de tiempo
Elaboración : Miguel Chong

Componentes de la serie de tiempo

En una serie existen cuatro tipos de variación, los cuales sobrepuestos o actuando de manera sincronizada, contribuyen a los cambios observados en un periodo de tiempo y dan a la serie un aspecto errático. Estas cuatro componentes son (N. Davey et. al. [21]):

- **Tendencia secular:** O tendencia a largo plazo de una serie y es el resultado de factores a largo plazo. Las tendencias a largo plazo se ajustan a diversos esquemas, algunas se mueven continuamente hacia arriba, otras declinan y otras permanecen igual en un cierto periodo o intervalo de tiempo.
- **Variación estacional:** Representa la variabilidad en los datos debida a influencias de las estaciones, se llama también componente estacional. Esta variación corresponde a movimientos de una serie que ocurren año tras año en los mismos meses.
- **Variación cíclica:** Son secuencias alternas de puntos abajo y arriba de la línea de tendencia que duran más de un año, esta variación se mantiene después de que se han eliminado las variaciones o tendencias estacional e irregular.
- **Variación irregular:** Son variaciones ocasionadas por factores a corto plazo, imprevisibles y no recurrentes. Esta variación es impredecible por lo que no se puede descifrar con anterioridad cuál será su impacto sobre la serie.

1.3.3. PRONÓSTICO EN SERIES DE TIEMPO

El análisis en series de tiempo se basa en que se pueden usar los datos relacionados del pasado para realizar pronósticos.

El objetivo de un método de series de tiempo es descubrir en los datos históricos un patrón para después extrapolar ese patrón al futuro; el pronóstico se basa únicamente en los valores de la variable en el pasado o en errores de pronóstico en el pasado.

1.3.3.1. MODELOS DE PRONÓSTICO

Promedios móviles

En este método, para pronosticar el periodo siguiente, se emplea el promedio de valores de los n datos más reciente de la serie de tiempo. El cálculo de un promedio móvil se hace como sigue:

$$y = \frac{\sum_{i=1}^n x_i}{n}, n \neq 0$$

Donde:

y : Promedio móvil

x_i : Valor i de la serie

n : Número de datos de la serie

El término móvil se usa porque cada vez que en la serie de tiempo hay una nueva observación, ésta sustituye a la observación más antigua que se emplee en la ecuación y se calcula un nuevo promedio. De esta manera el promedio se modifica, o se mueve, cada que se tiene una nueva observación.

Promedios móviles ponderados

En el método de los promedios móviles, a todos los datos que se emplean en el cálculo se les da el mismo peso. En este modelo, a cada uno de los valores de los datos se le da un peso diferente y, después, se calcula el promedio ponderado de los valores de los n datos más recientes para obtener el pronóstico. Por ejemplo, a las observaciones más recientes se le da el mayor peso y los pesos disminuyen conforme los datos son más antiguos. Considerando además que la suma de los pesos es igual a 1. El cálculo de un promedio móvil ponderado se hace como sigue:

$$y = \frac{\sum_{i=1}^n d_i x_i}{n}, n \neq 0$$

Donde:

- y : Promedio móvil ponderado
 d_i : Peso ponderado del valor i de la serie
 x_i : Valor i de la serie
 n : Número de datos de la serie

Suavizamiento exponencial

En este modelo se usa un promedio ponderado de los valores pasado de la serie de tiempo; es un caso especial del método de promedio ponderados móviles; en este caso sólo hay que elegir un peso, el peso para la observación más reciente. Los pesos para los demás datos se calculan automáticamente y son más pequeños a medida que los datos son más antiguos. A continuación se presenta el modelo de suavizamiento exponencial básico:

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t$$

Donde:

- F_{t+1} : Pronóstico para el periodo $t+1$ de la serie de tiempo
 Y_t : Valor real en el periodo t de la serie de tiempo
 F_t : Pronóstico para el periodo t de la serie de tiempo
 α : Constante de suavizamiento ($0 \leq \alpha \leq 1$)

Modelo Holt-Winters

El modelo de Holt-Winters es una ampliación perfeccionada del alisamiento exponencial, pues al incorporar la tendencia general de crecimiento o de decrecimiento permite estimaciones a medio y largo

plazo, superando así las restricciones que mostraban tanto las medias móviles como el suavizamiento exponencial. Este método es especialmente aconsejable cuando la serie histórica muestra tendencia a crecer o decrecer. El proceso de predicción de este método está formado por tres componentes: el nivel, la tendencia y el componente estacional y están relacionados por ecuaciones iterativas.

Para usar este método, en un momento i es necesario estimar el valor de la serie suavizada e_i y el valor de la tendencia T_i , de acuerdo con las siguientes fórmulas:

$$e_i = \alpha(e_{i-1} + T_{i-1}) + (1 - \alpha)Y_i$$

$$T_i = \beta T_{i-1} + (1 - \beta)(e_i - e_{i-1})$$

Donde:

- e_i : Nivel de la serie suavizada en el periodo t
- T_i : Valor del componente de tendencia para el periodo t
- Y_i : Valor observado de la serie en el periodo i
- α y β : Constantes de suavizamiento, valores entre 0 y 1

Regresión Múltiple

El modelo de regresión múltiple estudia la relación de una variable dependiente con dos o más variables independientes.

Este método está basado en la determinación de factores que influyen en el comportamiento de la variable dependiente y el cálculo de su correlación. El propósito es determinar la funcionalidad de esta influencia y predecir valores futuros de las variables dependientes.

El proceso se inicia con la identificación de las variables significativas independientes cuyos comportamientos podrían impactar a la variable dependiente. A continuación se procede a validar los factores determinados a través de la evaluación de las relaciones intrínsecas con

la variable dependiente para dar paso a la formulación de las hipótesis que las sustentarán. Asimismo sobre las variables dependientes, se eliminarán las redundantes; es decir, aquellas con bajo grado de relación lineal como datos nulos y outliers. Finalmente se aplican las técnicas seleccionadas para el análisis correlacional y calcular los coeficientes de las variables independientes con el objetivo de estimar el pronóstico para las siguientes series de tiempo. Inicialmente se asume una forma de los datos y se evalúan los resultados, luego se toman otros valores y se vuelve a evaluar su performance, esto continua hasta alcanzar una forma válida de menor error. A continuación el modelo:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_{p1} + \varepsilon$$

Donde:

y : Valor pronosticado

x_1, x_2, x_p : Variables independientes

$\beta_0, \beta_1, \beta_2, \beta_p$: Parámetros del modelo

ε : Error del pronóstico

Métodos Box Jenkins o ARIMA

Este modelo considera que la serie temporal que se trata de predecir es generada por un proceso estocástico cuya naturaleza puede ser caracterizada por un modelo. Para efectuar una estimación basada en este modelo se requiere de un elevado número de observaciones a comparación de los modelos anteriormente descritos. Básicamente, la metodología consiste en encontrar un modelo matemático que represente el comportamiento de una serie temporal de datos y permita efectuar pronósticos únicamente introduciendo el periodo de tiempo correspondiente.

En los modelos del tipo univariante se explica el comportamiento de una serie temporal a partir de las observaciones pasadas de la propia serie y a partir de los errores pasados de previsión. El modelo se representa por la siguiente ecuación:

$$Y_t = \alpha_0 + \alpha_{t-1}y_{t-1} + L + \alpha_{t-p}y_{t-p} + a_t - \theta_1 a_{t-1} - L - \theta_q a_{t-q}$$

Donde:

- Y_t : Valor pronosticado en el momento t
- $\alpha_{t-1}y_{t-1} + L + \alpha_{t-p}y_{t-p}$: Parte autorregresiva (AR) del modelo
- $-\theta_1 a_{t-1} - L - \theta_q a_{t-q}$: Parte medias móviles (MA) del modelo

1.4. MARCO TEÓRICO INSTRUMENTAL

1.4.1. REDES NEURONALES ARTIFICIALES

Las Redes Neuronales Artificiales (RNA) ofrecen un enfoque diferente a las soluciones presentadas en los puntos anteriores, por lo que comúnmente son denominadas como la sexta generación de la computación (P. Rossini [2]). Son modelos predictivos no lineales que aprenden directamente del entrenamiento y reensamblan redes de neuronas biológicas en su estructura. Estos sistemas proveen una herramienta que se reprograma como resultado de su auto-aprendizaje.

Las RNA hacen referencia a una clase de modelos inspirados en el sistema nervioso humano. Están formadas por elementos sencillos de cálculo aritmético, equivalentes a las neuronas, o sea, las células que procesan la información en el cerebro, por tanto la red neuronal equivale a un conjunto de neuronas interconectadas.

El pronóstico a través de las RNA está siendo recientemente considerado debido a los buenos resultados conseguidos tanto en el reconocimiento de patrones como en la predicción de valores, dando a luz una serie de artículos y publicaciones académicas alrededor de este

tema. El concepto está basado en sistemas computacionales que son capaces de aprender a través de la experiencia adquirida por medio del reconocimiento de patrones existentes en el conjunto de datos analizados (Palmer et. al. [12]).

Para la implementación efectiva de las RNA se requiere cumplir con una serie de condiciones, las cuales se detallan a continuación (I. Kaastra et. al. [13]):

- El conjunto de datos disponibles deberá representar información necesaria para caracterizar el problema.
- El tamaño del conjunto de datos disponibles deberá ser el adecuado tanto para el entrenamiento como el testeo.
- Entendimiento de la naturaleza básica del problema a resolver para que de esta manera pueda decidirse iniciar el modelado de la red neuronal.
- Las decisiones también incluirán definir las funciones de activación y transferencias, así como los métodos de aprendizaje.
- Entendimiento de las herramientas de desarrollo.
- Adecuados recursos computacionales (algunas aplicaciones requieren de procesamiento a tiempo real el cual podría exceder los recursos estándares).

Las principales ventajas del uso de redes neuronales son (P. García et. al. [14]):

- Aprendizaje adaptativo: Capacidad de aprender a realizar tareas basadas en un entrenamiento o una experiencia inicial.
- Auto organización: La red crea su propia representación de la información en su interior, descargando al usuario de esto.
- Tolerancia a fallos: Debido a que una red neuronal almacena la información de forma redundante, ésta puede seguir

respondiendo aceptablemente aún si los datos de entrada son ruidosos.

- Flexibilidad: Una red puede manejar cambios no importantes en la información de entrada, como señales con ruido u otros cambios en la entrada (por ejemplo si la información de entrada es la imagen de un objeto, la respuesta correspondiente no sufre cambios si la imagen cambia un poco su brillo o el objeto cambia ligeramente).

1.4.1.1. ESTRUCTURA DE UNA RED NEURONAL

Una red neuronal está constituida por nodos unidos mediante conexiones, y cada una de éstas se le asigna un peso numérico (B. Krose et. al. [1]). Los pesos constituyen el principal recurso de memoria de largo plazo y el aprendizaje se realiza con la actualización de tales pesos. Algunos de los nodos están conectados al ambiente externo y son designadas como unidades de entrada o de salida, como se representa en el Gráfico 8:

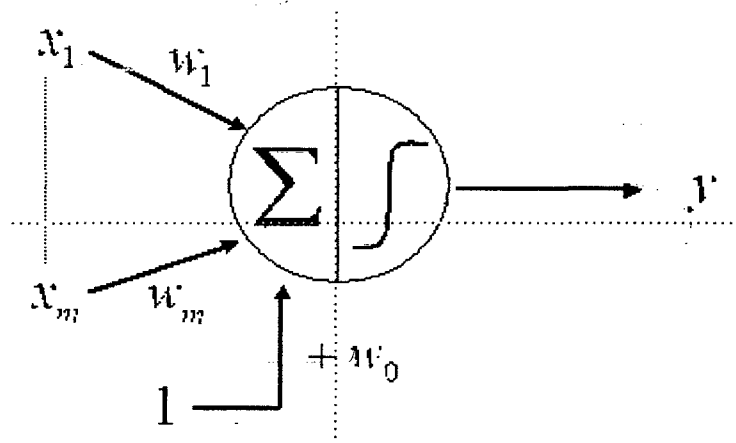


Gráfico 9 : Estructura de una Red Neuronal
Fuente : Introducción a las Redes Neuronales en la Minería de Datos
Elaboración : Faubricio Guillén Varela

Cada nodo realiza el cálculo, recibe señales de sus vínculos de entrada y calcula el correspondiente nuevo nivel de activación que envía a través

de sus vínculos de salida. El cálculo del nivel de activación está basado en los valores de cada una de las señales de entrada que envía un nodo vecino, así como los pesos de cada uno de los vínculos de entrada. El cálculo consta de dos partes, la primera es un componente lineal denominado función de entrada, que calcula la suma ponderada de los valores de la unidad. La segunda es un componente no lineal conocido como función de activación que transforma la suma ponderada en el valor final que sirve como valor de activación de la unidad, todas las unidades utilizan el mismo valor de activación (M. Gómez-Martínez [5]). La entrada ponderada total es la suma de las activaciones de entrada multiplicadas por sus pesos respectivos:

$$y = \gamma\left(\sum_{i=1}^m w_i x_i + w_0\right)$$

Donde:

- y : Salida de la red neuronal
- γ : Función no-lineal
- $\underline{x} = (x_1, \dots, x_m)^T$: Vector de entradas de la red
- $\underline{w} = (w_1, \dots, w_m)^T$: Vector de pesos de la red
- w_0 : Umbral de acción o activación

Función de Propagación

Calcula el valor de base o entrada total a la unidad, como simple suma ponderada de todas las entradas recibidas, es decir, de las entradas multiplicadas por el peso o valor de las conexiones. Equivale a la combinación de las señales excitatorias e inhibitorias de las neuronas biológicas (L. Flores [6]).

Función de Activación

Característica que define el comportamiento de la red neurona, se encarga de calcular el nivel o estado de activación de la neurona en función de la entrada total.

Conexiones ponderadas

Hacen el papel de las conexiones sinápticas, el peso de la conexión equivale a la fuerza o efectividad de la sinapsis. La existencia de conexiones determina si es posible que una unidad influya sobre otra, definen el tipo (excitatorio/inhibitorio) y la intensidad de la influencia.

Salida

Calcula la salida de la neurona en función de la activación de la misma, aunque normalmente no se aplica más que la función identidad, y se toma como salida el valor de activación. El valor de salida cumple la función de la tasa de disparo en las neuronas biológicas.

1.4.1.2. TIPOLOGÍAS DE REDES NEURONALES

A. Topología

En base al patrón de conexiones que presentan, se definen tres tipos básicos de redes:

- **Monocapa:** Red de propagación hacia delante (feed forward) o acíclica, las señales van desde la capa de entrada hacia la salida. Sólo existe una capa de neuronas de entrada y otra de salida. Por ejemplo el perceptrón y Adaline.
- **Multicapa:** Red de propagación hacia delante (feed forward) o acíclica, las señales van desde la capa de entrada hacia la salida. Existe un conjunto de capas intermedias entre la capa de entrada y la de salida (capas ocultas). Por ejemplo, el perceptrón multicapa.

- **Recurrentes:** Red que presenta al menos un ciclo cerrado de activación neuronal. Por ejemplo la red de Elman, la red de Hopfield y la máquina de Boltzmann.

B. Aprendizaje

En función al tipo de aprendizaje (si necesita o no un conjunto de entrenamiento supervisado) (B. Krose et. al. [1]). Para cada tipo de aprendizaje se encuentran los siguientes modelos:

- **Aprendizaje supervisado:** necesitan un conjunto de datos de entrada previamente clasificados; es decir la respuesta objetivo se conoce. Ejemplos de este tipo de redes son el perceptrón (simple y multicapa), la red Adaline y la memoria asociativa bidireccional.
- **Aprendizaje no supervisado o autoorganizado:** no necesitan de tal conjunto previo. Ejemplos de estas redes son las memorias asociativas, redes de Hopfield, máquina de Boltzmann y las redes de Kohonen.
- **Redes híbridas:** son un enfoque mixto en el que se utiliza una función de mejora para facilitar la convergencia. Un ejemplo de estas redes son las funciones de base radial (RBF).
- **Aprendizaje reforzado:** se sitúa a medio camino entre el supervisado y el autoorganizado.

C. Tipo de entrada

Clasificación en base a la capacidad de procesar información de distinto tipo:

- **Redes analógicas:** procesan datos de entrada con valores continuos y, habitualmente acotados. Ejemplos son la red de Hopfield, de Kohonen y redes de aprendizaje competitivo.

- Redes discretas: procesan datos de entrada de naturaleza discreta; habitualmente valores lógicos booleanos. Ejemplos de este tipo de redes son las máquinas de Bolzman y Cauchy.

1.4.1.3. MODELO NARX

El modelo NARX de redes neuronales artificiales está construido sobre una red perceptrón multicapas y sobre la cual se establecen regresores tanto en las entradas como en las salidas.

Estas redes son altamente empleadas cuando se trata de resolver problemas de ecuaciones diferenciales en el mundo continuo, y ecuaciones diferenciales en el tiempo discreto, por ejemplo cuando se trabaja con datos cuya naturaleza se basa en la dinámica de un reloj.

La estructura del modelo NARX es parametrizada con una red neuronal de propagación hacia adelante (llamada también perceptrón multicapa). El modelo consiste de tres capas: una entrada y dos capas de procesamiento.

La capa de entrada está formada por los elementos del vector de regresión. En el modelo NARX, el vector de regresión está constituido por valores pasados de la salida y valores actuales y pasados de la entrada. Las entradas se conectan hacia adelante, propagándose por dos capas procesadoras, hasta obtener la muestra estimada de salida.

Las capas procesadoras están constituidas por una capa oculta y una capa de salida, las cuales consisten en un conjunto de nodos de cálculo basados en neuronas de McCulloch-Pitts. Cada nodo, o neurona, realiza la operación de una suma ponderada de las señales de entrada y un término de sesgo (bias), que alimenta a una función de activación, resultando la salida del nodo. El término de sesgo puede considerarse una entrada unitaria, con su debida ponderación.

La dimensión del vector de regresión determina el número de entradas de la red. En la estructura NARX utilizada, cada una de las entradas se

conectan a cada uno de los nodos de la capa oculta, pero no hay conexión entre nodos dentro de una capa. Cada nodo de la capa oculta, con "n" entradas y una salida, presenta una relación entrada - salida dada por pesos que ponderan las sumas para cada entrada del nodo. La capa de salida contiene un solo nodo, con una función de activación lineal. Este nodo realiza la suma ponderada de cada una de las salidas de los nodos de la capa oculta y un término de sesgo.

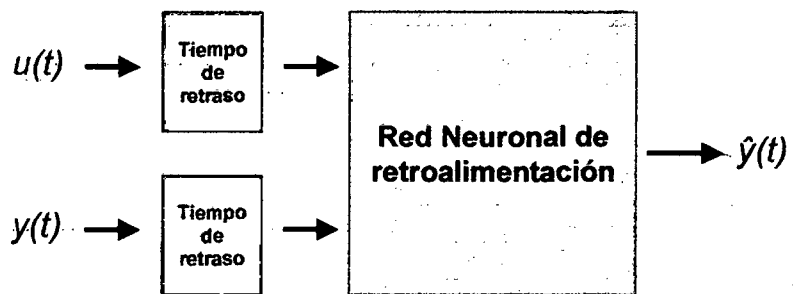


Gráfico 10 : Modelo NARX
Fuente : Neural Network Toolbox
Elaboración : D.C. Sorensen, R.B. Lehoucq, C. Yang y K. Maschhoff

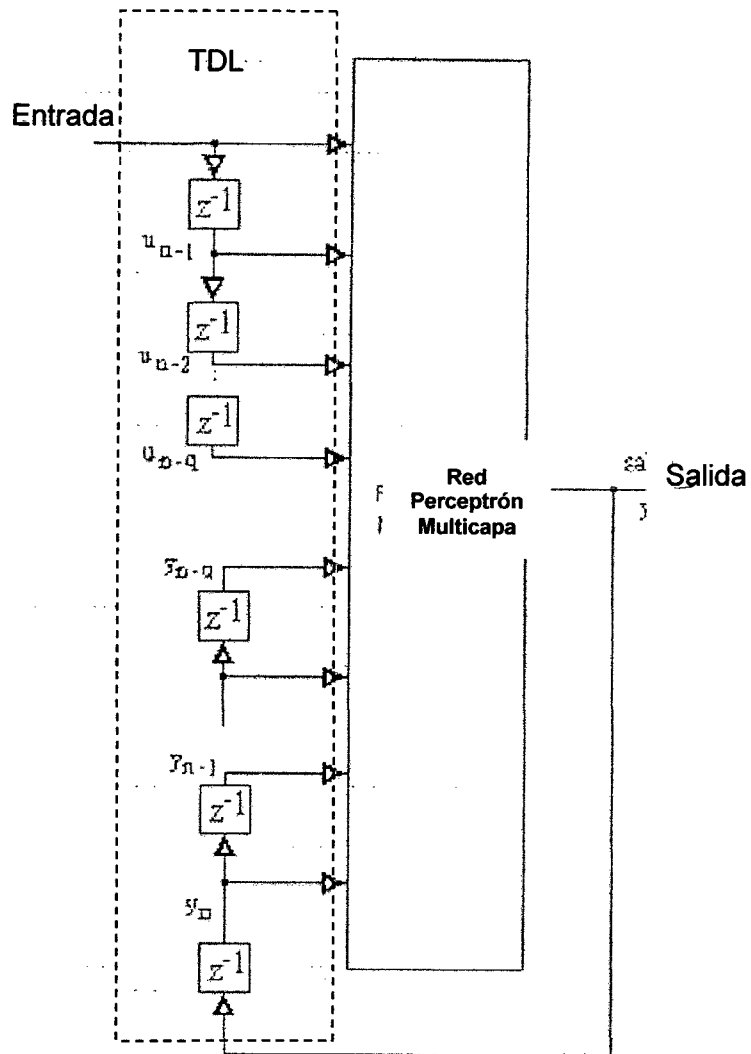


Gráfico 11 : Modelo NARX - Regresión
Fuente : Identification of linear dynamical time-variant systems using feedforward neural network
Elaboración : Ghatol A. A.

Los Gráficos 10 y 11 muestran la arquitectura del modelo NARX que sigue naturalmente del perceptrón multicapa.

El modelo tiene sólo una entrada, la cual se aplica a una red de retardo (memoria) de q unidades. Tiene una sola salida que retroalimenta a la entrada por medio de otra red de retardo también de q unidades. El

contenido de estas dos redes de retardo (memorias) es usado para alimentar la capa de entrada del perceptrón multicapas (P. Chand et. al. [8]).

El valor presente del modelo de entrada es denotado por u_n y el valor correspondiente de la salida del modelo y_{n+1} ; esto es, la salida está adelantada a la entrada por una unidad de tiempo (*TDL: Time delay line*). Por lo tanto, el vector de señales aplicado a la capa de entrada del perceptrón multicapa consiste en un conjunto de datos como sigue:

- Valores presentes y pasados de la entrada, llamados u_n, u_{n-1}, u_{n-q+1} los cuales representan entradas exógenas originadas desde fuera de la red.
- Valores retrasados de la salida y_n, y_{n-1}, y_{n-q+1} en los cuales la salida del modelo y_{n+1} es retroalimentada.

El comportamiento dinámico del modelo NARX es descrito por la siguiente ecuación:

$$y_k = f(y_n, y_{n-1}, \dots, y_{n-q+1}, u_n, u_{n-1}, \dots, u_{n-q+1})$$

Donde:

y_k : Salida de la red neuronal

y_n, y_{n-1}, y_{n-q+1} : Valores de periodos anteriores con tiempo de retardo

u_n, u_{n-1}, u_{n-q+1} : Valores de entrada presentes y pasados

El sistema de entrenamiento consiste en:

- Empezar con unos pesos sinápticos cualquiera (generalmente elegidos al azar).
- Introducir datos de entrada (en la capa de entradas) en base al tiempo de retardo que sea necesario aplicar.

- Dejar que la red genere un vector de datos de salida (propagación hacia delante).
- Comparar la salida generada por la red con la salida deseada.
- La diferencia obtenida entre la salida generada y la deseada (denominada error) se usa para ajustar los pesos sinápticos de las neuronas de la capa de salidas.
- El error se propaga hacia atrás, hacia la capa de neuronas anterior, y se usa para ajustar los pesos sinápticos en esta capa.
- Se continúa propagando el error hacia atrás y ajustando los pesos hasta que se alcance la capa de entradas.

Este proceso se repetirá con los diferentes datos de entrenamiento.

1.4.2. REDES NEURONALES PARA SERIES DE TIEMPO

Para predecir series de tiempo se han diseñado las redes neuronales alimentadas hacia delante (feed-forward) y las redes recurrentes (recurrent). Dentro de las redes alimentadas hacia adelante existen algoritmos como: Linear Associator, Backpropagation, Cascade Correlation, Cascade 2, RAN (Resource Allocating Network) (I. García Martín [3]). Entre las redes recurrentes se tiene: Recurrent Cascade Correlation, Simple Recurrent Networks, Real-Time Recurrent Learning, Sequential Cascaded Network (A. González et. al. [18]).

Las Redes Neuronales Artificiales han sido empleadas principalmente en los casos donde la variable medida en el tiempo presenta características de no ergodicidad, persistencia y no estacionalidad, además de exhibir dependencia a largo plazo, lo cual origina que algunos de los modelos dinámicos de series de tiempo más populares sean inadecuados en este contexto, como son los modelos tipo ARMA y GARCH (J. de Andrés S [9]).

Por un lado, en los modelos de RNA, las redes tipo feedforward con regresores son las más empleadas en series de tiempo, dadas sus

características de estabilidad y convergencia a la solución óptima. Y entre estas, las más populares son las redes perceptrón multicapas con regresores tanto en las entradas como salidas anteriores (O.Martin et al. [10]).

Las redes totalmente recurrentes tipo Jordan también han sido empleadas con muy buenos resultados, debido a que pueden responder según las decisiones tomadas en un pasado inmediato, siendo esto coherente con las propiedades de memoria y dependencia a largo plazo en estos objetos de estudio. No obstante, teniendo en cuenta que para estos casos, siempre se puede disponer de un valor deseado en la respuesta de la red neuronal para el conjunto de entrenamiento, se implementa una instrucción de tipo supervisado.

Asimismo paquetes de software matemáticos de propósito general tales como MATLAB [27], MATHCAD, SIMULINK [30], LABVIEW y aplicaciones estadísticas como SPSS y S-PLUS representan herramientas que permiten explotar las ventajas de las redes neuronales artificiales y aplicarlas a diversos tipos de problemas. La ventaja con respecto a los entornos VENSIM [28] o STELLA [29], los cuales optan por ofrecer un soporte de modelado basado en el símil hidrodinámico de las ecuaciones diferenciales, estos han de ser formulados en términos de depósitos, flujos y reguladores de flujo, de forma similar a los diagramas de Forrester, que son ensamblados gráficamente. La desventaja de este enfoque es que a menudo exige descomponer las ecuaciones del modelo en varios términos, o al menos ser capaz de expresarlas en base al símil utilizado en el software de modelado, como resultado el tiempo de aprendizaje aumenta en función de su familiaridad con ese símil.

CAPÍTULO II

ANÁLISIS DE DATOS

2.1 DESCRIPCIÓN GENERAL

En este capítulo se detallarán las actividades para calcular la demanda de efectivo en cajeros automáticos (ATM's).

Se detallarán los pasos necesarios y consideraciones para calcular la demanda a partir del stock de billetes en caseteras y los pedidos de abastecimiento efectuados en el periodo.

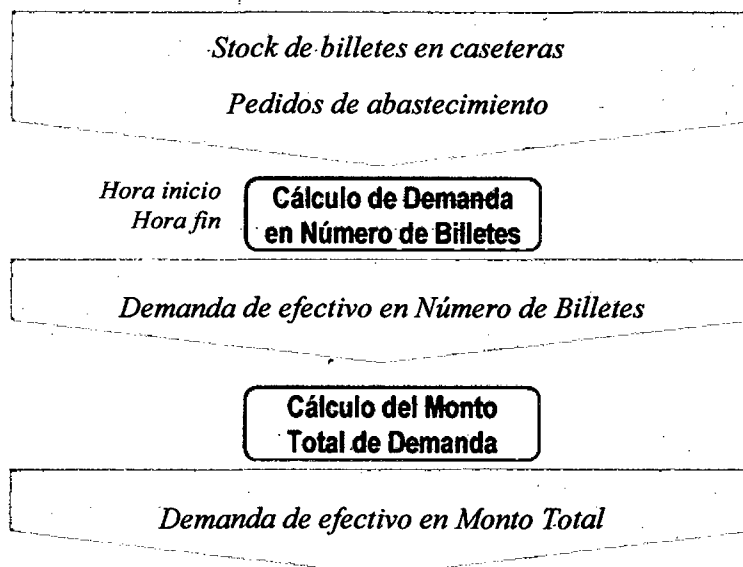


Gráfico 12 : Análisis de Datos
Fuente : Propia
Elaboración : Silvia Granados Flores

2.2 DESCRIPCIÓN DE DATOS

Con la finalidad de predecir la demanda de efectivo se requiere contar con información sobre su demanda histórica, de esta manera el sistema de redes neuronales artificiales podrá aprender a través de los patrones de comportamiento histórico y a través de su generalización, pronosticará el futuro comportamiento de la demanda.

Para la elaboración de este informe, una Empresa Financiera local proporcionó información real sobre cinco cajeros automáticos de su red, a los cuales denominaremos A, B, C, D y E. Para este caso la empresa no contaba con datos directos sobre demanda histórica de efectivo; por lo que fue necesario realizar una serie de cálculos sobre los datos accesibles a fin de obtener la demanda histórica.

Se cuenta con:

- Stock de billetes en caseteras de los cajeros automáticos (ATM's) en ciertos momentos del día.
- Pedidos de abastecimiento de efectivo en un mes.

2.2.1 CORRESPONDENCIA ENTRE NÓMINAS DE BILLETES Y CASETERAS DEL ATM

Antes de efectuar los cálculos para obtener la demanda histórica de efectivo, resulta necesario analizar la estructura de las caseteras de los cajeros automáticos y su correspondencia con las nóminas de billetes.

La distribución de las nóminas de billetes en las caseteras de los cajeros automáticos de esta Entidad Financiera puede graficarse de la siguiente manera:

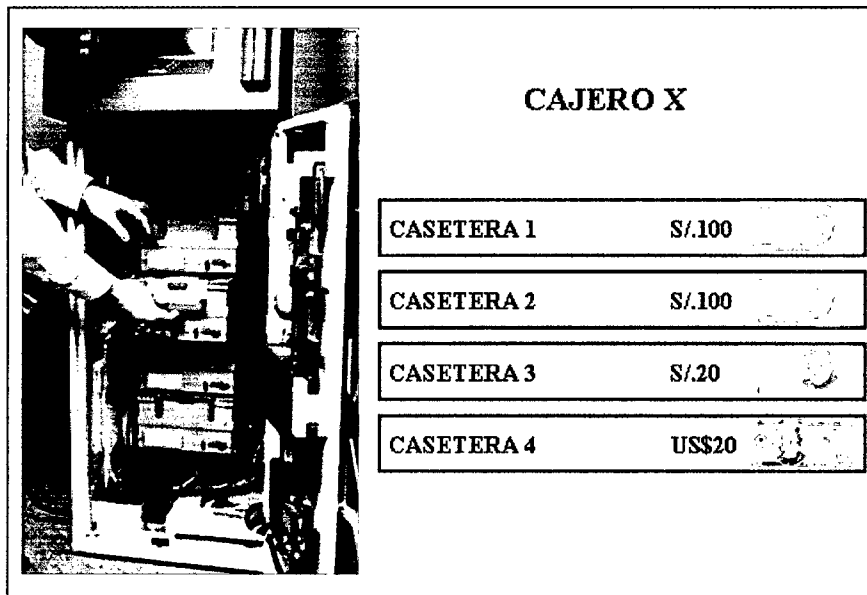


Gráfico 13 : Distribución de billetes en caseteras del ATM

Fuente : Entidad Financiera local

Elaboración : Propia

En la práctica las caseteras 1 y 2 funcionan como una única casetera debido a que al momento de realizar el abastecimiento, el efectivo es repartido por igual en ambas. Por otro lado, la emisión de billetes es realizada por una a la vez, cuando el stock de billetes de la casetera 1 es menor al límite establecido, deja de emitir billetes y la casetera 2 se activa.

Para el desarrollo de este trabajo de investigación se considerarán tres nóminas de billetes y tres caseteras (creándose una casetera virtual que incluya la casetera 1 y 2) generando la siguiente correspondencia:


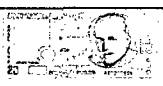

NÓMINA DE BILLETE	CASETERA ORIGINAL	CASETERA FINAL
(1) S/.100 	CASETERA (1)	CASETERA (1)
	CASETERA (2)	
(2) S/.20 	CASETERA (3)	CASETERA (2)
(3) US\$20 	CASETERA (4)	CASETERA (3)

Tabla 2 : Correspondencia entre nóminas de billete y caseteras
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

Por lo indicado en la tabla anterior, todo cálculo a efectuarse sobre la nómina de billete 1, totalizará los montos parciales de las caseteras originales 1 y 2.

2.2.2 STOCK DE BILLETES EN CASETERAS

Estos datos son obtenidos a través de un sistema de información que genera reportes sobre el stock en las caseteras diariamente y en determinadas horas programadas (8:45a.m., 12:00m. y 5:00p.m.).

Se han obtenido datos que corresponden al periodo del 01 de Diciembre del 2008 al 30 de Noviembre del 2009, conformando una base de datos de 1098 registros (3 registros por día).

Estructura de Datos

La estructura del registro sobre stock de billetes en las caseteras de los cajeros automáticos en un determinado momento, tiene la siguiente estructura:

$$\text{Stock de billetes en caseteras}$$
$$S_T^X = \{F_T, H_T, C_X, S_T^{X,S/.100}, S_T^{X,S/.20}, S_T^{X,US\$20}\}$$

Donde:

S_T^X : Stock de billetes en las caseteras del cajero automático "X" en un instante de tiempo "T".

F_T : Fecha del instante de tiempo "T" en el que se mide el stock de billetes en las caseteras del cajero automático "X".

H_T : Hora del instante de tiempo "T" en el que se mide el stock de billetes en las caseteras del cajero automático "X".

C_X : Cajero automático "X".

$S_T^{X,S/.100}$: Stock de billetes con nómina S/.100 en el cajero automático "X" en el instante de tiempo "T". El stock de billetes se representará en número de billetes.

$S_T^{X,S/.20}$: Stock de billetes con nómina S/.20 en el cajero automático "X" en el instante de tiempo "T". El stock de billetes se representará en número de billetes.

$S_T^{X,US\$20}$: Stock de billetes con nómina US\$20 en el cajero automático "X" en el instante de tiempo "T". El stock de billetes se representará en número de billetes.

La siguiente tabla grafica la estructura de datos que correspondería al reporte de stock de billetes en las caseteras de un cajero automático (ATM):

REPORTE DE STOCK DE BILLETES					
FECHA	HORA	CAJERO	CASETERA	CASETERA	CASETERA
			NRO 1 S/.100	NRO 2 S/.20	NRO 3 US\$20
01/12/2008	08:45 a.m.	C123	4755	3172	2145
01/12/2008	12:04 p.m.	C123	4449	3117	2132
01/12/2008	05:08 p.m.	C123	4072	3013	1864
02/12/2008	09:15 a.m.	C123	3336	2673	1833
02/12/2008	12:04 p.m.	C123	3260	2603	1693
02/12/2008	05:00 p.m.	C123	2821	2496	1427
03/12/2008	08:45 a.m.	C123	2421	2257	1365
03/12/2008	12:05 p.m.	C123	2280	2192	1287
03/12/2008	05:01 p.m.	C123	2153	2070	1144
04/12/2008	08:45 a.m.	C123	1838	1856	1092
04/12/2008	12:00 p.m.	C123	1825	1828	1040
04/12/2008	07:20 p.m.	C123	1368	1690	1019
05/12/2008	08:45 a.m.	C123	1087	1537	907
05/12/2008	12:31 p.m.	C123	9168	5179	2020
05/12/2008	05:02 p.m.	C123	8762	5104	1773
06/12/2008	11:45 a.m.	C123	8219	4893	1708
06/12/2008	12:11 p.m.	C123	7909	4758	1635
06/12/2008	05:00 p.m.	C123	7589	4545	1521
07/12/2008	08:45 a.m.	C123	7301	4316	1469
07/12/2008	12:00 p.m.	C123	7090	4220	1446
07/12/2008	05:27 p.m.	C123	6937	4053	1440

Tabla 3 : Estructura ejemplo de Stock de Billetes
Fuente : Propia
Elaboración : Silvia Granados Flores

Datos recolectados

Para esta investigación se han recolectado registros con información sobre stock de billetes en las caseteras de cinco cajeros automáticos (A,

B, C, D y E) correspondiente al periodo del 01 de Diciembre del 2008 al 30 de Noviembre del 2009.

Para cada cajero automático los registros obtenidos se totalizan en la siguiente tabla:

TURNO	HORA	TOTAL REGISTROS
TURNO 1	08:45 a.m.	358
	09:55 a.m.	1
	10:21 a.m.	1
	11:41 a.m.	1
	TOTAL OBSERVACIONES	361
	TOTAL IMPRECISOS	3
TURNO 2	TOTAL VACIOS	4
	12:00 p.m.	332
	12:01 p.m.	11
	12:02 p.m.	3
	12:04 p.m.	1
	12:05 p.m.	2
	12:07 p.m.	1
	12:09 p.m.	2
	12:16 p.m.	2
	12:28 p.m.	1
	12:38 p.m.	1
	12:45 p.m.	1
	12:56 p.m.	1
	01:22 p.m.	1
	01:23 p.m.	1
02:13 p.m.	1	
TOTAL OBSERVACIONES	361	
TOTAL IMPRECISOS	29	
TOTAL VACIOS	4	
TURNO 3	05:00 p.m.	310
	05:01 p.m.	16
	05:02 p.m.	4
	05:03 p.m.	3
	05:05 p.m.	1
	05:06 p.m.	2
	05:07 p.m.	1
	05:08 p.m.	1
	05:09 p.m.	2
	05:11 p.m.	2
	05:12 p.m.	1
	05:14 p.m.	1
	05:16 p.m.	1
	05:19 p.m.	2
	05:21 p.m.	1
	05:22 p.m.	1
	05:23 p.m.	1
	05:27 p.m.	1
	05:36 p.m.	1
	05:41 p.m.	1
	05:50 p.m.	1
	05:56 p.m.	1
	06:02 p.m.	1
	06:19 p.m.	1
	06:27 p.m.	1
	07:20 p.m.	1
	08:29 p.m.	1
08:31 p.m.	1	
TOTAL OBSERVACIONES	361	
TOTAL IMPRECISOS	51	
TOTAL VACIOS	4	
REGISTROS RECOLECTADOS POR ATM		1083

Tabla 4 : Registros recolectados de stock de billetes

Fuente : Entidad Financiera local

Elaboración : Silvia Granados Flores

Comportamiento de los datos

Los datos recolectados pueden ser graficados a través de curvas de evolución para observar su comportamiento. En los Gráficos 14, 15, 16, 17 y 18 se presentan los stock de billetes de las caseteras de S/.100, S/.20 y US\$20 de los cajeros automáticos A, B, C, D y E.

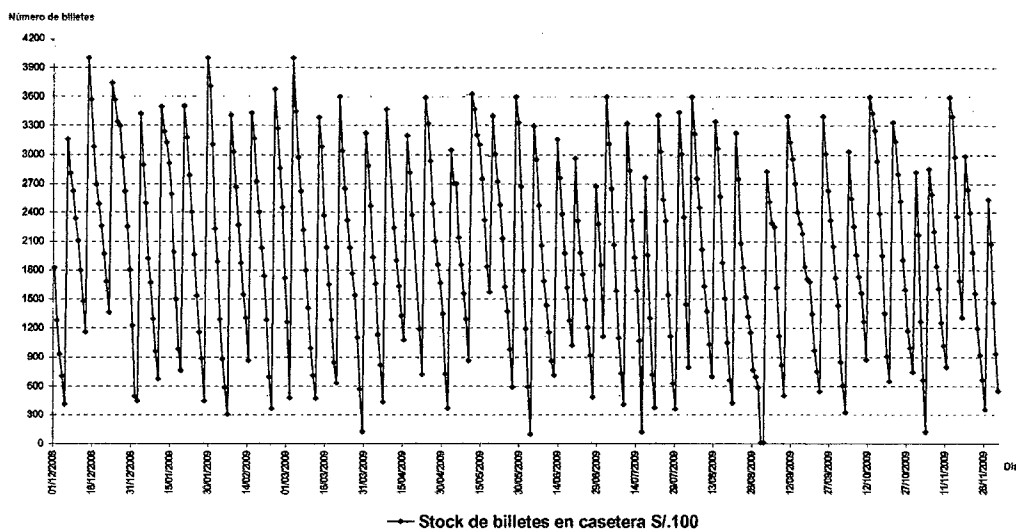


Gráfico 14.1 : Cajero A - Stock de billetes casetera S/.100

Fuente : Entidad Financiera local

Elaboración : Silvia Granados Flores

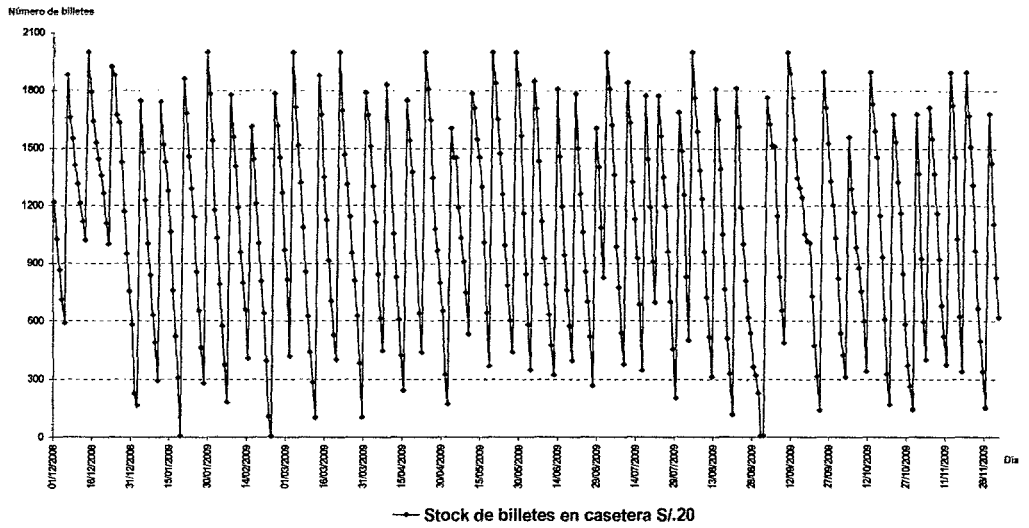


Gráfico 15.2 : Cajero A - Stock de billetes casetera S/.20

Fuente : Entidad Financiera local

Elaboración : Silvia Granados Flores

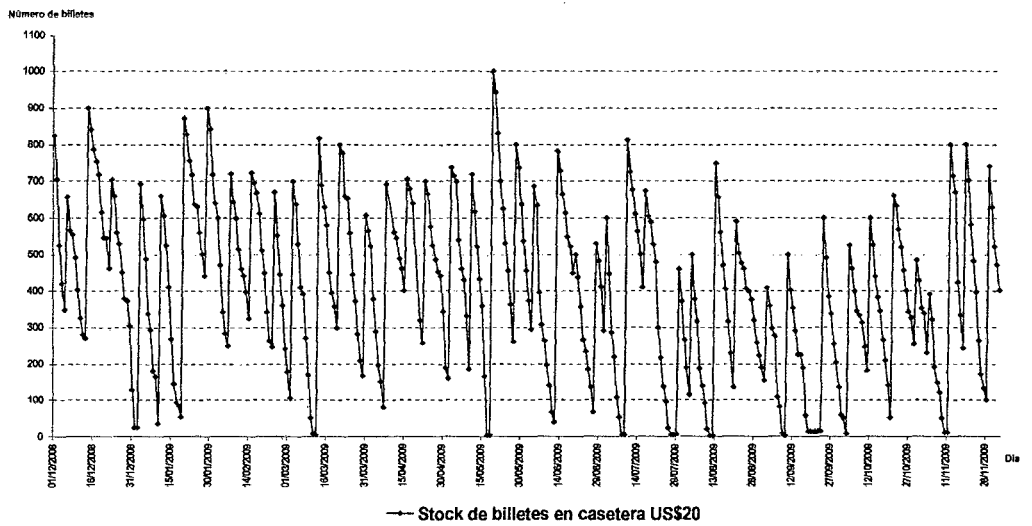


Gráfico 16.3 : Cajero A - Stock de billetes casetera US\$20

Fuente : Entidad Financiera local

Elaboración : Silvia Granados Flores

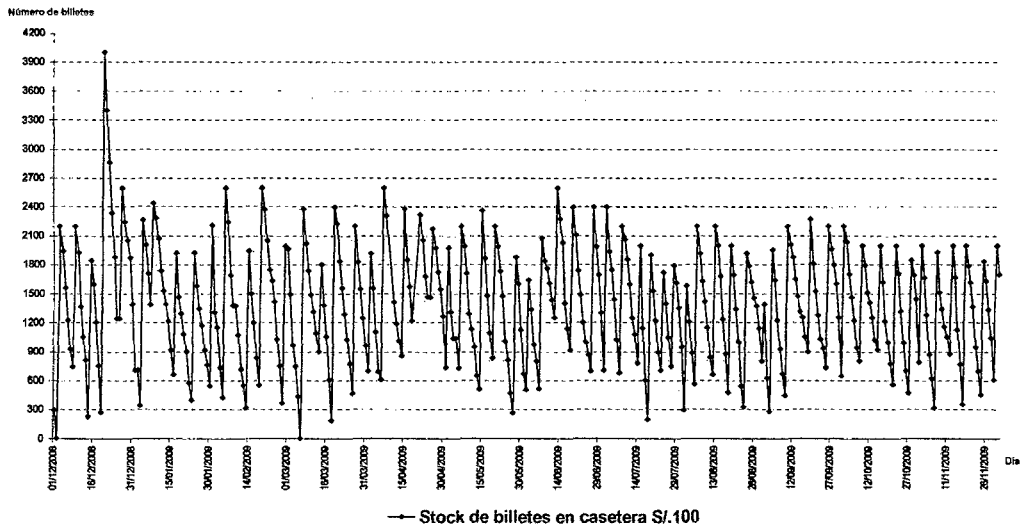


Gráfico 17.1 : Cajero B - Stock de billetes casetera \$/100
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

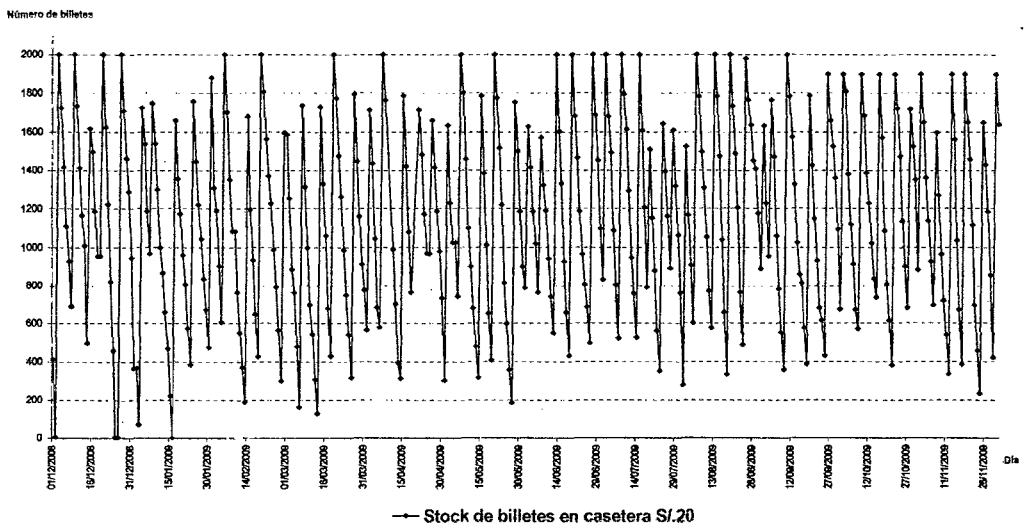


Gráfico 18.2 : Cajero B - Stock de billetes casetera \$/20
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

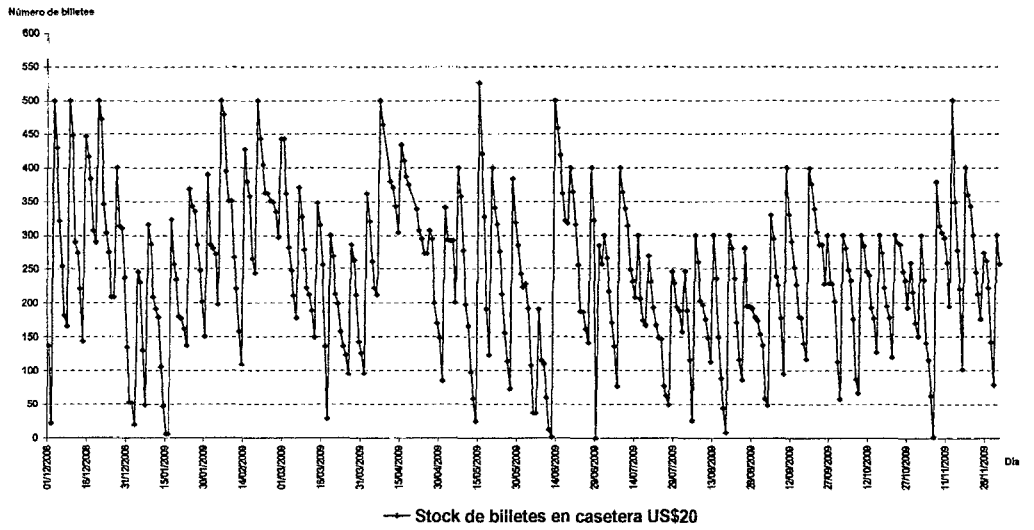


Gráfico 19.3 : Cajero B - Stock de billetes casetera US\$100

Fuente : Entidad Financiera local

Elaboración : Silvia Granados Flores

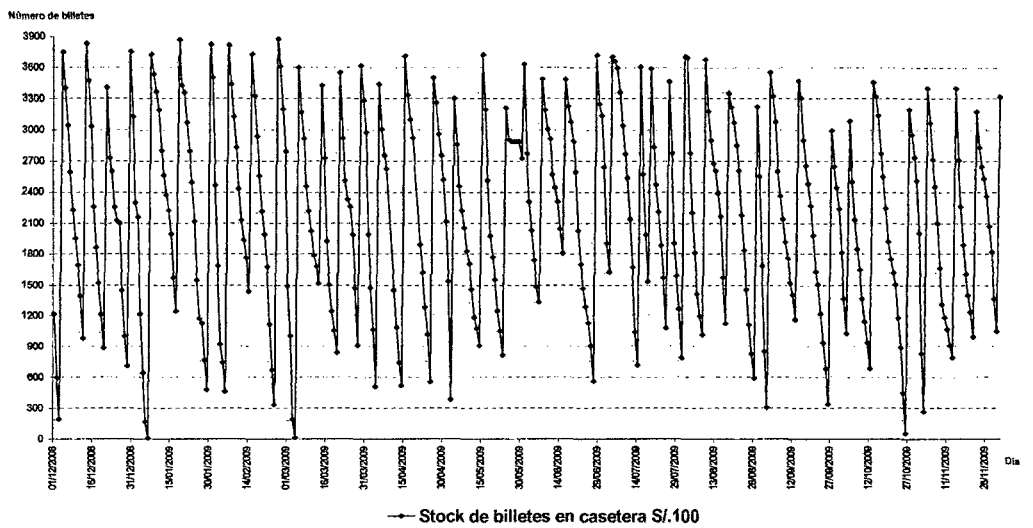


Gráfico 20.1 : Cajero C - Stock de billetes casetera S/.100

Fuente : Entidad Financiera local

Elaboración : Silvia Granados Flores

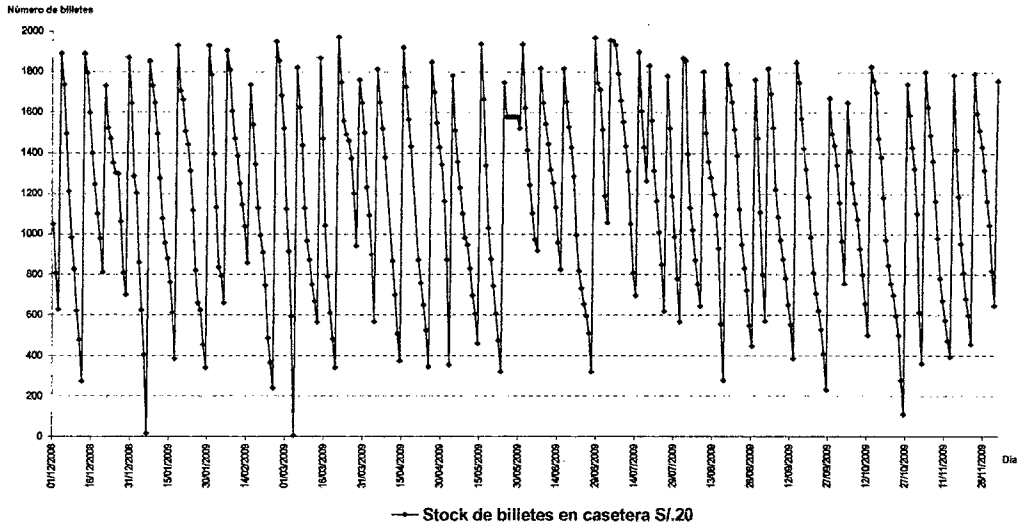


Gráfico 21.2 : Cajero C - Stock de billetes casetera S/.20
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

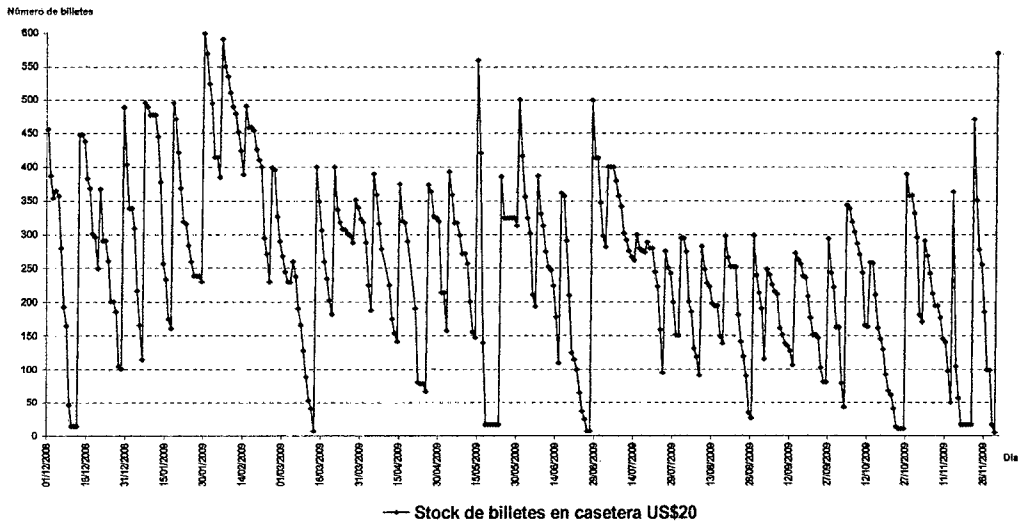


Gráfico 22.3 : Cajero C - Stock de billetes casetera US\$20
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

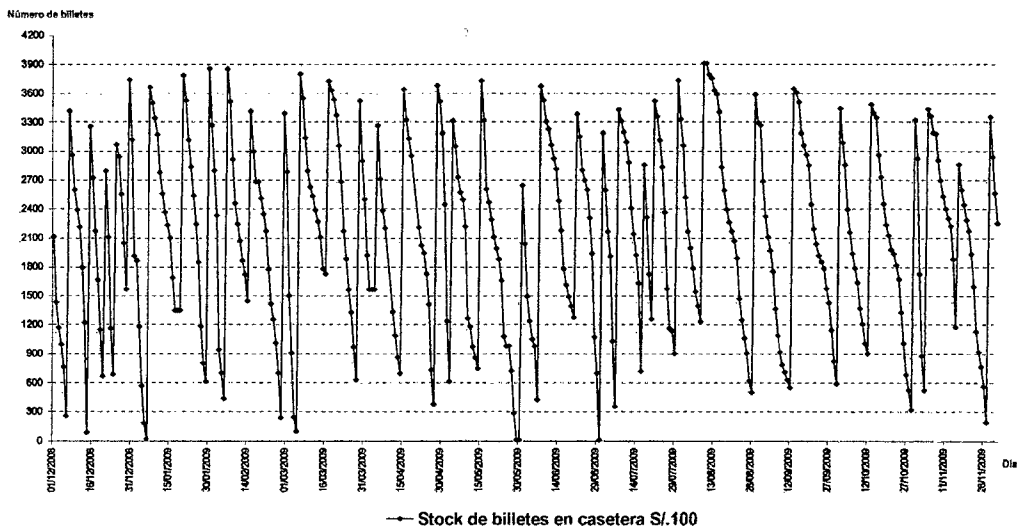


Gráfico 23.1 : Cajero D - Stock de billetes casetera S/.100
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

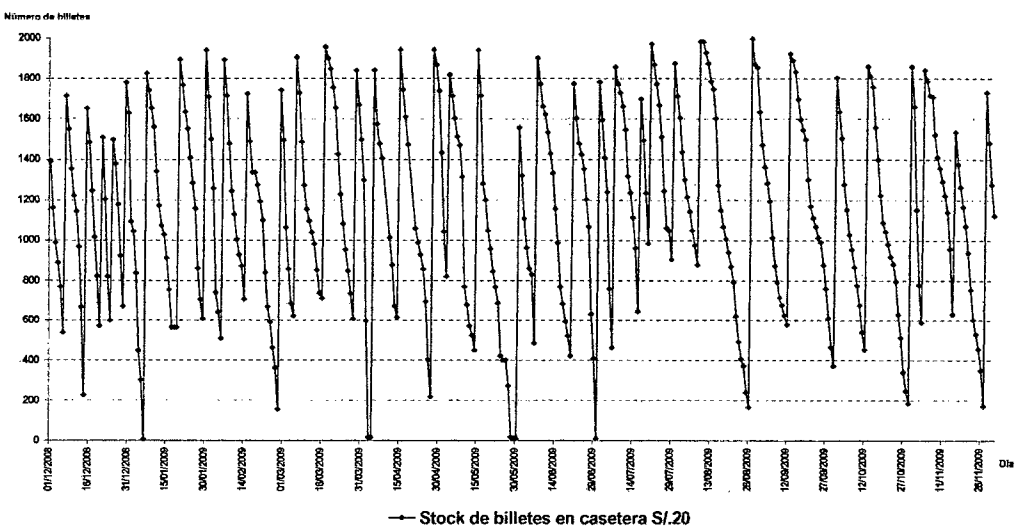


Gráfico 24.2 : Cajero D - Stock de billetes casetera S/.20
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

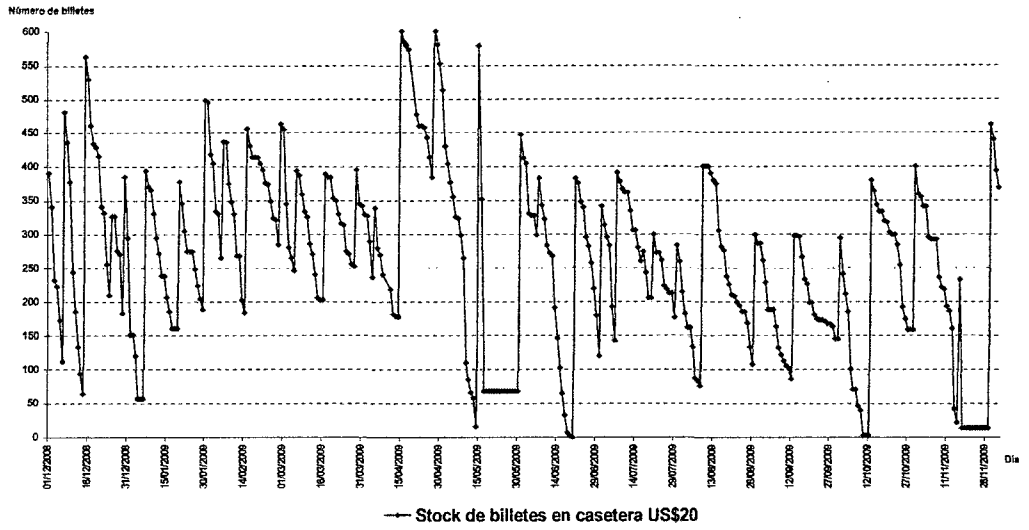


Gráfico 25.3 : Cajero D - Stock de billetes casetera US\$20

Fuente : Entidad Financiera local

Elaboración : Silvia Granados Flores

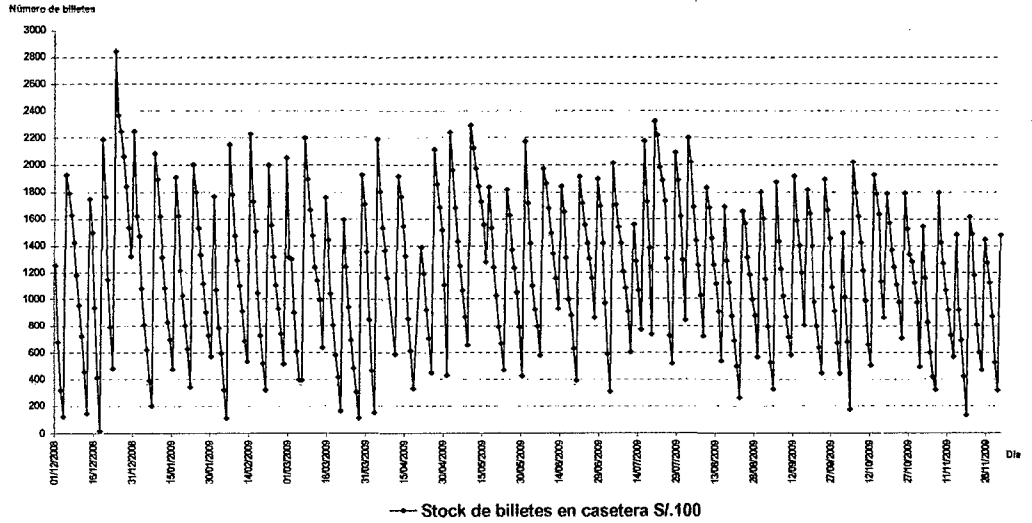


Gráfico 26.1 : Cajero E - Stock de billetes casetera S/.100

Fuente : Entidad Financiera local

Elaboración : Silvia Granados Flores

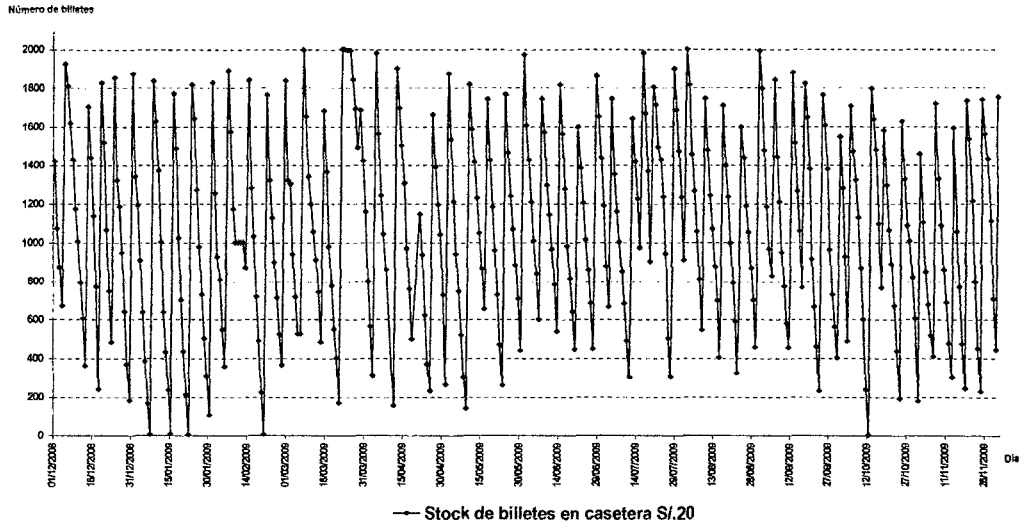


Gráfico 27.2 : Cajero E - Stock de billetes casetera S/.20

Fuente : Entidad Financiera local

Elaboración : Silvia Granados Flores

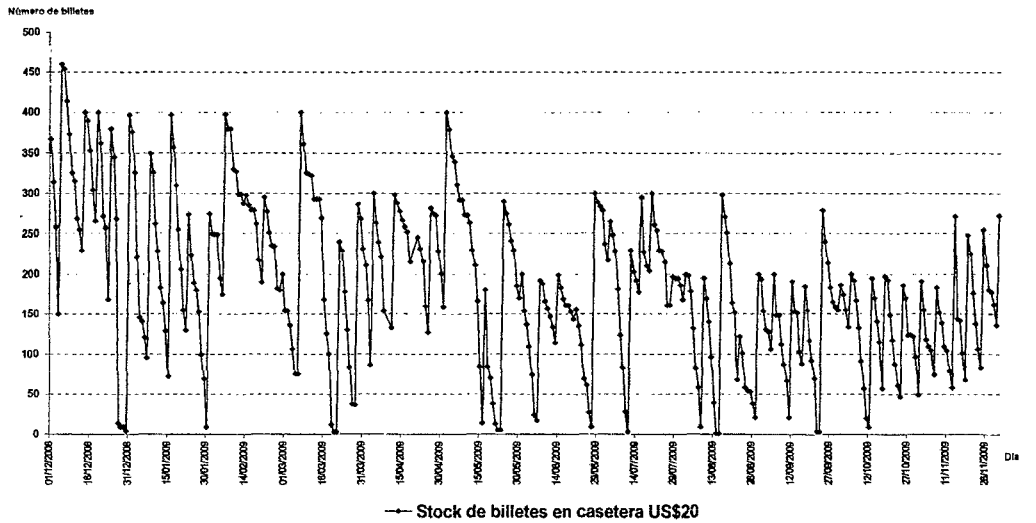


Gráfico 28.3 : Cajero E - Stock de billetes casetera US\$ 20

Fuente : Entidad Financiera local

Elaboración : Silvia Granados Flores

2.3 PEDIDOS DE ABASTECIMIENTO

Estos datos corresponden al consolidado mensual con los pedidos de recarga de efectivo en los cajeros automáticos.

De este modo, se ha obtenido información correspondiente a los pedidos de abastecimiento efectuados del 01 de Diciembre del 2008 al 30 de Julio del 2009.

Estructura de Datos

La estructura del registro correspondiente al pedido de abastecimiento de efectivo en un cajero automático se representará:

Pedidos de abastecimiento

$$A_T^X = \{F_T, H_T, C_X, A_T^{X,S/.100}, A_T^{X,S/.20}, A_T^{X,US\$20}, R_T^{X,S/.100}, R_T^{X,S/.20}, R_T^{X,US\$20}\}$$

Donde:

A_T^X : Abastecimiento de billetes en el cajero automático "X" en un instante de tiempo "T".

F_T : Fecha del abastecimiento en un instante de tiempo "T".

H_T : Hora del abastecimiento en un instante de tiempo "T".

C_X : Cajero automático "X"

$A_T^{X,S/.100}$: Número de billetes abastecidos en la casetera con nómina S/.100 correspondiente al cajero automático "X" en el instante de tiempo "T".

$A_T^{X,S/.20}$: Número de billetes abastecidos en la casetera con nómina S/.20 correspondiente al cajero automático "X" en el instante de tiempo "T".

$A_T^{X,US\$20}$: Número de billetes abastecidos en la casetera con nómina US\$20 correspondiente al cajero automático "X" en el instante de tiempo "T".

$R_T^{X,S/.100}$: Número de billetes remanentes en la casetera con nómina S/.100 correspondiente al cajero automático "X" en el instante de tiempo "T".

$R_T^{X,S/.20}$: Número de billetes remanentes en la casetera con nómina S/.20 correspondiente al cajero automático "X" en el instante de tiempo "T".

$R_T^{X,US\$20}$: Número de billetes remanentes en la casetera con nómina US\$20 correspondiente al cajero automático "X" en el instante de tiempo "T".

La siguiente tabla representa la estructura de datos que correspondería al reporte de pedidos de abastecimiento en un cajero automático (ATM):

REPORT DE PEDIDOS DE ABASTECIMIENTO DE EFECTIVO									
N	FECHA ABASTECIMIENTO	HORA ABASTECIMIENTO	CAJERO	#BILLETES ABASTECIDOS CASETERA NRO 1 S/.100	#BILLETES ABASTECIDOS CASETERA NRO 2 S 20	#BILLETES ABASTECIDOS CASETERA NRO 3 US\$20	#BILLETES REMANENTES CASETERA NRO 1 S/.100	#BILLETES REMANENTES CASETERA NRO 2 S/.20	#BILLETES REMANENTES CASETERA NRO 3 US\$20
1	05/12/08	09:21 a.m.	C123	9000	4800	1920	1003	1418	838
2	17/12/08	09:16 a.m.	C123	9600	4800	2160	1634	2035	521
3	19/12/08	09:05 a.m.	C123	9600	4800	1920	3283	2405	1111
4	05/01/09	11:51 a.m.	C123	9600	4800	1920	1082	403	60
5	07/01/09	09:00 a.m.	C123	9600	4800	1920	1630	706	86
6	22/01/09	03:10 p.m.	C123	9600	4800	1920	1874	29	17
7	26/01/09	08:25 a.m.	C123	9600	4800	1920	470	230	768

Tabla 5 : Estructura ejemplo de Pedidos de Abastecimiento

Fuente : Propia

Elaboración : Silvia Granados Flores

Datos recolectados

Los pedidos de abastecimiento de billetes efectuados desde el 01 de Diciembre del 2008 al 31 de Julio del 2009 se totalizan en la siguiente tabla:

MES	TOTAL PEDIDOS REALIZADOS				
	A	B	C	D	E
Dic - 08	4	5	4	5	5
Ene - 09	3	5	3	3	4
Feb - 09	4	4	3	3	4
Mar - 09	3	5	4	3	4
Abr - 09	4	5	4	4	5
May - 09	3	5	2	1	4
Jun - 09	4	4	3	3	4
Jul - 09	5	6	5	4	6
Ago - 09	4	5	3	2	4
Set - 09	2	4	3	2	5
Oct - 09	4	6	3	2	5
Nov - 09	4	5	4	3	5

Tabla 6 : Registros recolectados sobre pedidos de abastecimiento
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

Comportamiento de los datos

Los datos recolectados sobre pedidos de abastecimiento en los cajeros automáticos A, B, C, D y E se representan en los Gráficos 17, 18, 19, 20 y 21 respectivamente. Se grafican los montos de abastecimiento y remanentes en las caseteras de S/.100, S/.20 y US\$20.

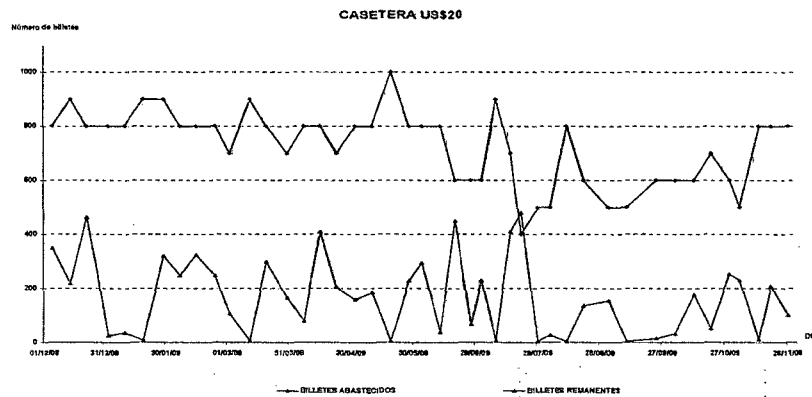
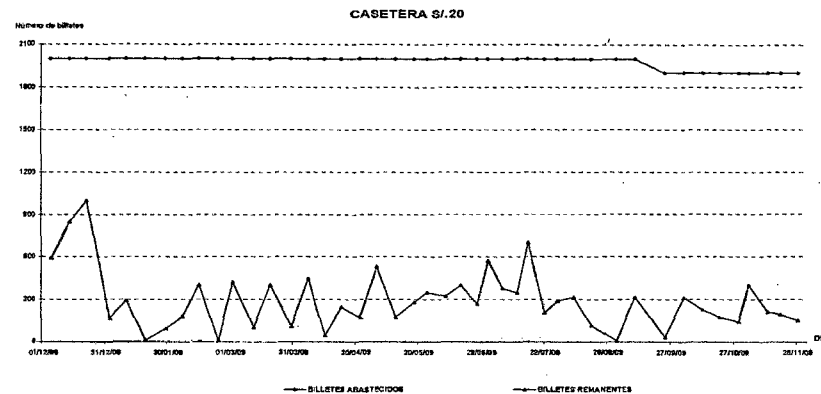
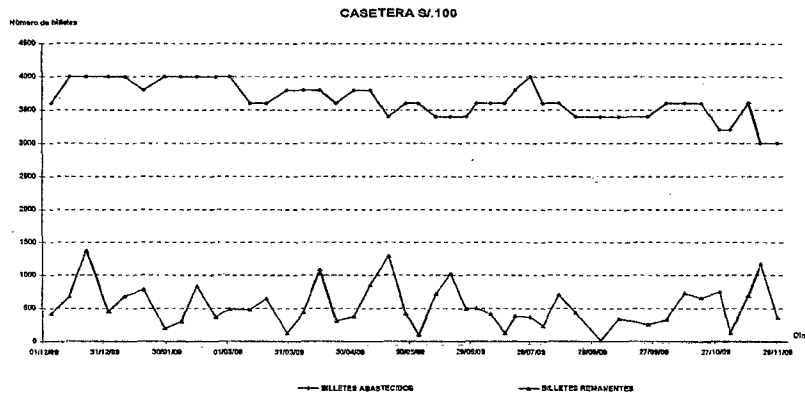


Gráfico 29 : Cajero A - Abastecimiento y remanente de efectivo
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

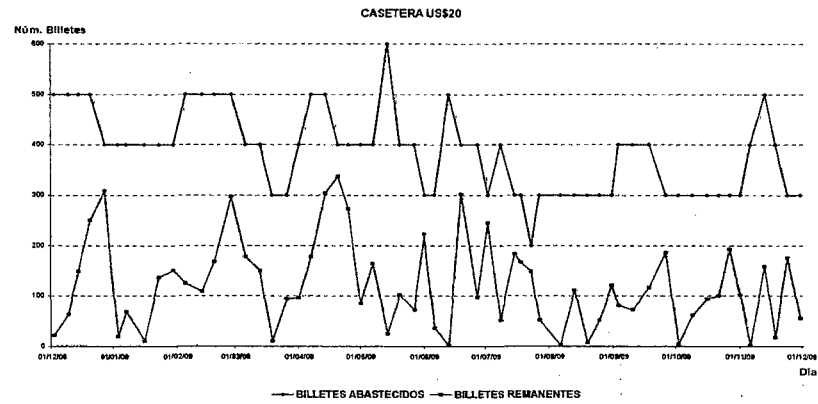
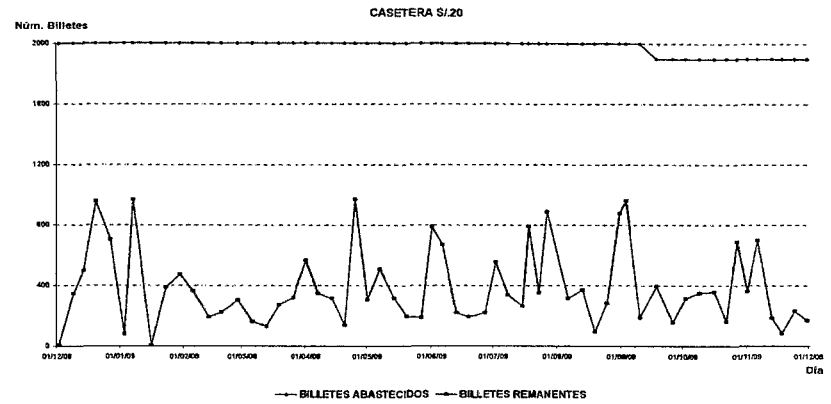
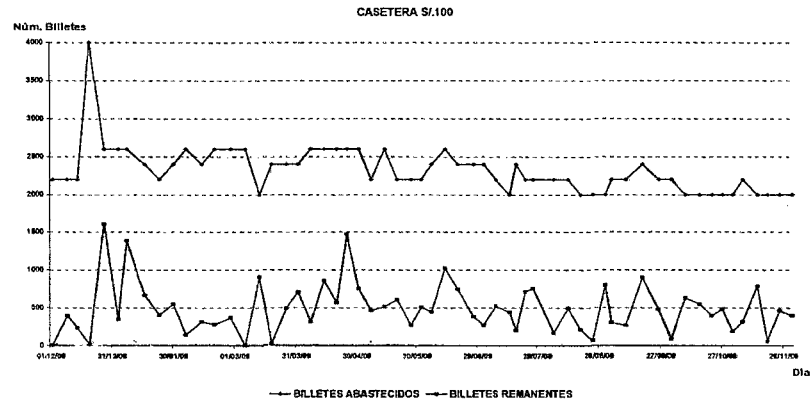


Gráfico 30 : Cajero B - Abastecimiento y remanente de efectivo
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

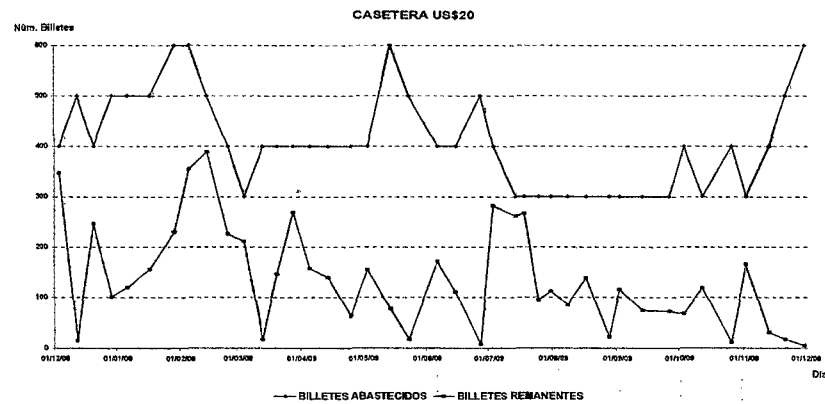
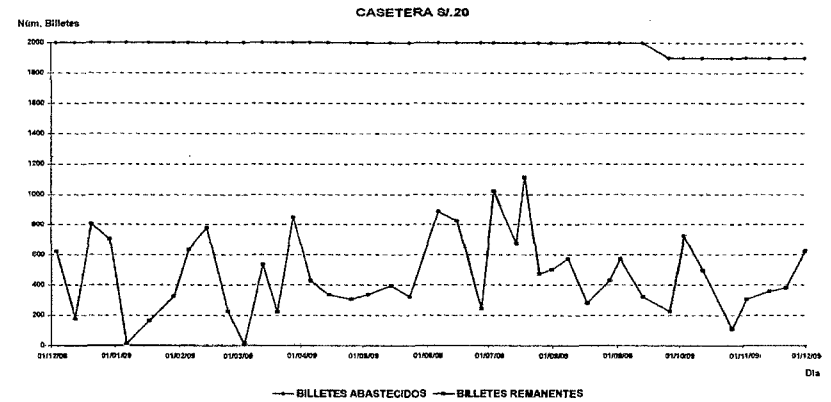
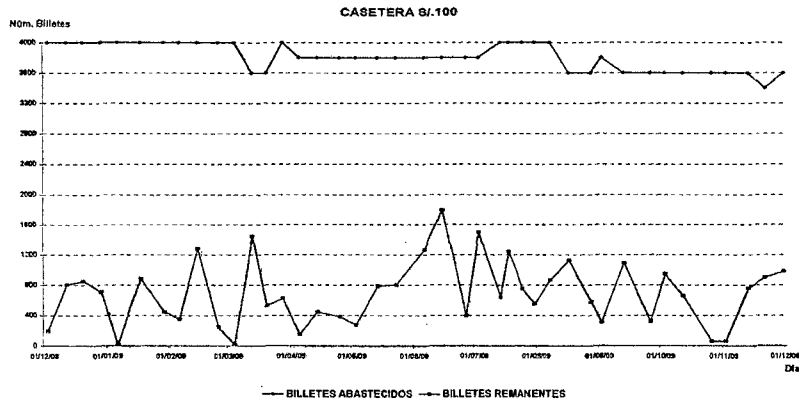


Gráfico 31 : Cajero C – Abastecimiento y remanente de efectivo
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

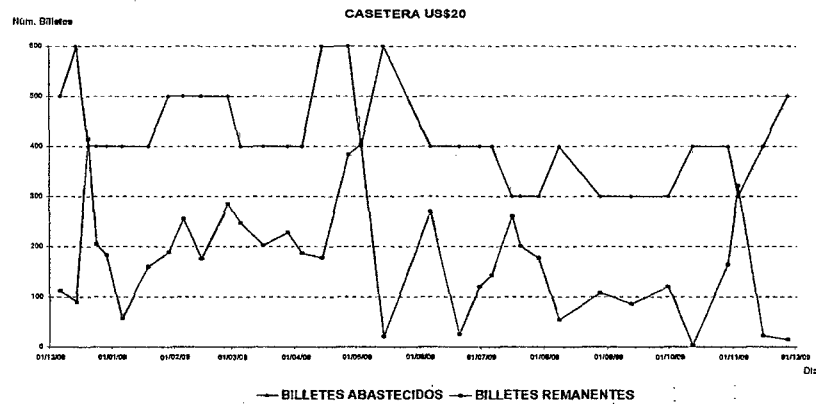
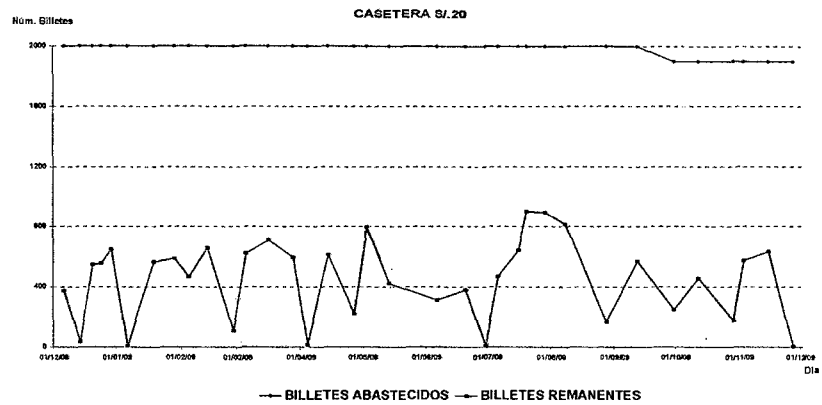
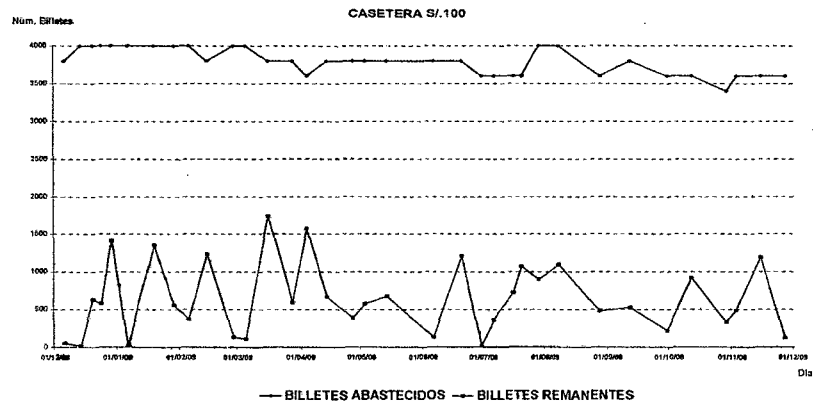


Gráfico 32 : Cajero D – Abastecimiento y remanente de efectivo
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

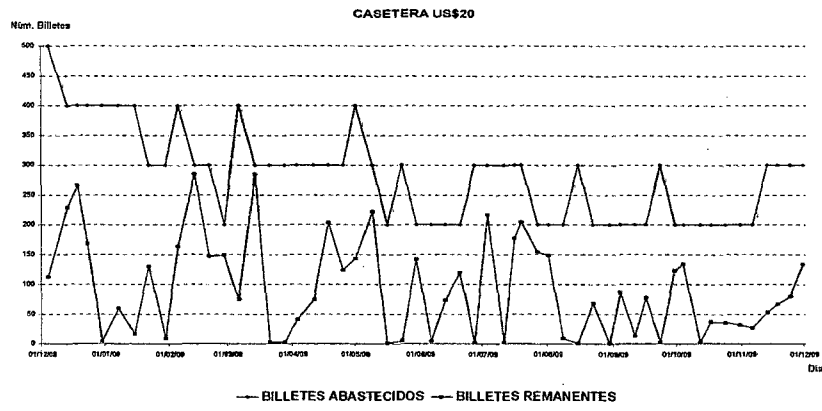
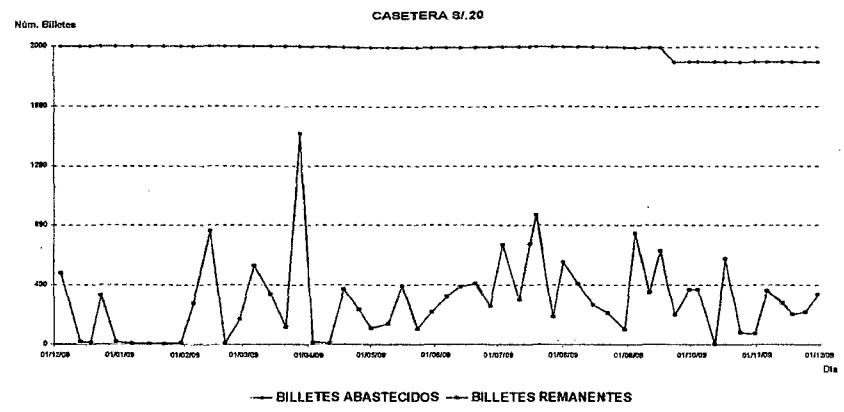
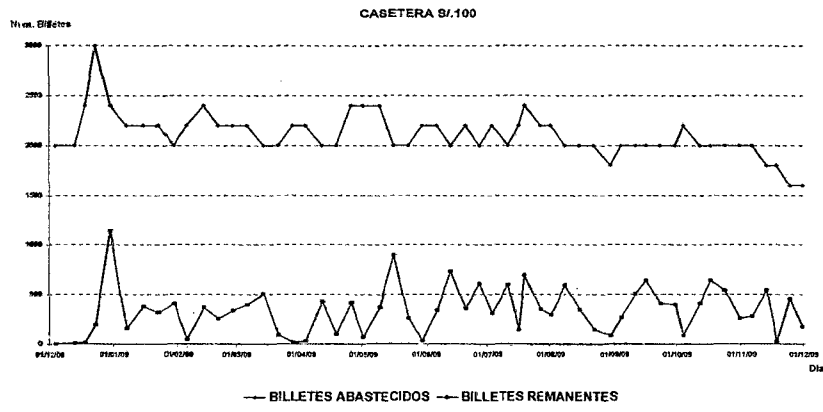


Gráfico 33 : Cajero E – Abastecimiento y remanente de efectivo
Fuente : Entidad Financiera local
Elaboración : Silvia Granados Flores

2.4 DEMANDA DE EFECTIVO

La estructura del dato correspondiente a la demanda de efectivo en número de billetes se representará de la siguiente manera:

$$\text{Demanda en Número de Billetes}$$
$$D_I^X = \{C_X, P_I, Fi_I, Hi_I, Ff_I, Hf_I, D_I^{X,S/.100}, D_I^{X,S/.20}, D_I^{X,US\$20}\}$$

Donde:

D_I^X : Demanda de efectivo en el cajero automático (ATM) "X" en el intervalo de tiempo "I". La demanda de efectivo será representada por número de billetes.

C_X : Cajero automático "X"

P_I : Denominación del periodo "I" en el que se mide la demanda de efectivo. Este intervalo de establece desde el punto de inicio $\{Fi_I, Hi_I\}$ al punto final $\{Ff_I, Hf_I\}$.

Fi_I : Fecha del punto de inicio del intervalo de tiempo "I".

Hi_I : Hora del punto de inicio del intervalo de tiempo "I".

Ff_I : Fecha del punto final del intervalo de tiempo "I".

Hf_I : Hora del punto final del intervalo de tiempo "I".

$D_I^{X,S/.100}$: Demanda de billetes de S/.100 en el cajero automático (ATM) "X" en el intervalo de tiempo "I". La demanda de efectivo será representada en número de billetes.

$D_I^{X,S/.20}$: Demanda de billetes de S/.20 en el cajero automático (ATM) "X" en el intervalo de tiempo "I". La demanda de efectivo será representada en número de billetes.

$D_I^{X,US\$20}$: Demanda de billetes de US\$20 en el cajero automático (ATM) "X" en el intervalo de tiempo "I". La demanda de efectivo será representada en número de billetes.

La siguiente tabla representa la estructura de datos que correspondería a la demanda de efectivo en un cajero automático (ATM):

DEMANDA DE EFECTIVO EN CAJERO AUTOMÁTICO								
PERIODO	FECHA INICIO	HORA INICIO	FECHA FIN	HORA FIN	CAJERO	DEMANDA		
						CASETERA 1 S/100	CASETERA 2 S/20	CASETERA 3 US\$20
01/12/2008	01/12/2008	08:45 a.m.	02/12/2008	09:15 a.m.	C123	1420	499	312
02/12/2008	02/12/2008	09:15 a.m.	03/12/2008	08:45 a.m.	C123	915	416	468
03/12/2008	03/12/2008	08:45 a.m.	04/12/2008	08:45 a.m.	C123	582	400	273
04/12/2008	04/12/2008	08:45 a.m.	05/12/2008	08:45 a.m.	C123	751	320	185
05/12/2008	05/12/2008	08:45 a.m.	06/12/2008	11:45 a.m.	C123	505	25	282
06/12/2008	06/12/2008	11:45 a.m.	07/12/2008	08:45 a.m.	C123	918	577	239

Tabla 7 : Estructura ejemplo de Demanda de Efectivo
Fuente : Propia
Elaboración : Silvia Granados Flores

Cálculo de Demanda en Número de Billetes

A partir del reporte de stock de efectivo en cajeros automáticos se extraerán los registros cuya hora contemple el límite del periodo determinado análisis.

Por ejemplo, podría determinarse (empleando los registros de la Tabla

2) que el periodo P_I será diario e iniciará a las 8:45 horas y finalizará a las 8:44 horas del día siguiente. Por este motivo desde la Tabla 2 se extraerán los registros correspondientes a las 8:45 horas o aquellos de mayor aproximación, tal como se indica en el Gráfico 22

REPORTE DE STOCK DE BILLETES						
FECHA	HORA	CAJERO	CA SETERA	CA SETERA	CA SETERA	
			NRO 1 S/.100	NRO 2 S/.20	NRO 3 US\$20	
01/12/2008	08:45 a.m.	C123	4755	3172	2145	EXTRAER
01/12/2008	12:04 p.m.	C123	4449	3117	2132	
01/12/2008	05:08 p.m.	C123	4072	3013	1864	
02/12/2008	09:15 a.m.	C123	3336	2673	1833	EXTRAER
02/12/2008	12:04 p.m.	C123	3260	2603	1693	
02/12/2008	05:00 p.m.	C123	2821	2496	1427	
03/12/2008	08:45 a.m.	C123	2421	2257	1365	EXTRAER
03/12/2008	12:05 p.m.	C123	2280	2192	1287	
03/12/2008	05:01 p.m.	C123	2153	2070	1144	
04/12/2008	08:45 a.m.	C123	1838	1856	1092	EXTRAER
04/12/2008	12:00 p.m.	C123	1825	1828	1040	
04/12/2008	07:20 p.m.	C123	1368	1690	1019	
05/12/2008	08:45 a.m.	C123	1087	1537	907	EXTRAER
05/12/2008	12:31 p.m.	C123	9168	5179	2020	
05/12/2008	05:02 p.m.	C123	8762	5104	1773	
06/12/2008	11:45 a.m.	C123	8219	4893	1708	EXTRAER
06/12/2008	12:11 p.m.	C123	7909	4758	1635	
06/12/2008	05:00 p.m.	C123	7589	4545	1521	
07/12/2008	08:45 a.m.	C123	7301	4316	1469	EXTRAER
07/12/2008	12:00 p.m.	C123	7090	4220	1446	
07/12/2008	05:27 p.m.	C123	6937	4053	1440	

Gráfico 34 : Extracción de registros sobre stock de efectivo en ATM

Fuente : Propia

Elaboración : Silvia Granados Flores

Inicialmente la matriz correspondiente a la demanda de efectivo se construirá a partir de los registros extraídos. Se considerará como rango del periodo: mayor o igual a la fecha y hora del stock del día y menor que la fecha y hora de stock del día siguiente:

$$P_I = [S_T^X \{ F_T, H_T \}, S_{T+1}^X \{ F_{T+1}, H_{T+1} \} >$$

Entonces:

$$D_I^X \{ F_{i_1}, H_{i_1} \} = S_T^X \{ F_T, H_T \}$$

$$D_I^X \{ F_{f_1}, H_{f_1} \} = S_{T+1}^X \{ F_{T+1}, H_{T+1} \}$$

Por lo que:

$$P_I = [D_I^X \{ F_{i_1}, H_{i_1} \}, D_I^X \{ F_{f_1}, H_{f_1} \} >$$

Donde se espera que las horas de inicio y fin de los periodos sean iguales:

$$P_I \{Hi_I\} = P_{I+1} \{Hi_{I+1}\}$$

$$P_I \{Hf_I\} = P_{I+1} \{Hf_{I+1}\}$$

En la Tabla 7 se presenta la construcción de los registros sobre demanda de efectivo, específicamente en la determinación del periodo.

DEMANDA DE EFECTIVO EN CAJERO AUTOMÁTICO								
PERIODO	FECHA INICIO	HORA INICIO	FECHA FIN	HORA FIN	CAJERO	DEMANDA		
						CAJETERA 1 S/.100	CAJETERA 2 S/.20	CAJETERA 3 US\$20
01/12/2008	01/12/2008	08:45 a.m.	02/12/2008	09:15 a.m.	C123			
02/12/2008	02/12/2008	09:15 a.m.	03/12/2008	08:45 a.m.	C123			
03/12/2008	03/12/2008	08:45 a.m.	04/12/2008	08:45 a.m.	C123			
04/12/2008	04/12/2008	08:45 a.m.	05/12/2008	08:45 a.m.	C123			
05/12/2008	05/12/2008	08:45 a.m.	06/12/2008	11:45 a.m.	C123			
06/12/2008	06/12/2008	11:45 a.m.	07/12/2008	08:45 a.m.	C123			

Tabla 8 : Determinación del periodo para cálculo de la demanda
Fuente : Propia
Elaboración : Silvia Granados Flores

Considerando que los pedidos de abastecimiento se realizan en una fecha y hora conocidas, cada registro será incluido dentro del cálculo de demanda de efectivo del periodo que corresponda.

Para determinar qué registro de demanda de efectivo será incluido en qué periodo debe de considerarse la fecha y hora del pedido de abastecimiento. De esta manera, el registro de abastecimiento será incluido dentro del lapso de tiempo "I" si es mayor o igual a la fecha y hora de inicio y menor a la fecha y hora fin del registro de demanda de efectivo.

Por lo que:

Si $A_T^X \{F_T, H_T\} \in [D_I^X \{Fi_I, Hi_I\}, D_I^X \{Ff_I, Hf\}]$, entonces

A_T^X se incluye en el cálculo de D_I^X

Fin

Esto quiere decir que la demanda de efectivo por nómina de billete se calculará de la siguiente manera:

Si $A_T^X \{F_T, H_T\} \in \left\{ D_I^X \{F_i, H_i\}, D_I^X \{Ff_i, Hf_i\} \right\}$ entonces

$$D_I^{X,S/.100} = S_I^{X,S/.100} + A_T^{X,S/.100} - R_T^{X,S/.100} - S_{I+1}^{X,S/.100};$$

$$D_I^{X,S/.20} = S_I^{X,S/.20} + A_T^{X,S/.20} - R_T^{X,S/.20} - S_{I+1}^{X,S/.20};$$

$$D_I^{X,US\$20} = S_I^{X,US\$20} + A_T^{X,US\$20} - R_T^{X,US\$20} - S_{I+1}^{X,US\$20};$$

Sino

$$D_I^{X,S/.100} = S_T^{X,S/.100} - S_{T+1}^{X,S/.100};$$

$$D_I^{X,S/.20} = S_T^{X,S/.20} - S_{T+1}^{X,S/.20};$$

$$D_I^{X,US\$20} = S_T^{X,US\$20} - S_{T+1}^{X,US\$20};$$

Fin

A continuación en la Tabla 9 se representa la dinámica a seguir para el cálculo de la demanda de efectivo en número de billetes:

Stock inicial 'S(I) 1. SUMA					Remanente 'R(T) 2. RESTA					Abastecimiento 'A(T) 3. SUMA					Stock final 'S(I+1) 4. RESTA					= DEMANDA		
FECHA	HORA	CA SETERA 1 S/100	CA SETERA 2 S/20	CA SETERA 3 US\$20	FECHA	HORA	CA SETERA 1 S/100	CA SETERA 2 S/20	CA SETERA 3 US\$20	FECHA	HORA	CA SETERA 1 S/100	CA SETERA 2 S/20	CA SETERA 3 US\$20	FECHA	HORA	CA SETERA 1 S/100	CA SETERA 2 S/20	CA SETERA 3 US\$20	CA SETERA 1 S/100	CA SETERA 2 S/20	CA SETERA 3 US\$20
01/12/2008	08:45 a.m.	4755	3172	2145											02/12/2008	09:15 a.m.	3336	2673	1833	1420	499	312
02/12/2008	09:15 a.m.	3336	2673	1833											03/12/2008	08:45 a.m.	2421	2257	1365	915	416	468
03/12/2008	08:45 a.m.	2421	2257	1365											04/12/2008	08:45 a.m.	1838	1856	1092	582	400	273
04/12/2008	08:45 a.m.	1838	1856	1092											05/12/2008	08:45 a.m.	1087	1537	907	751	320	185
05/12/2008	08:45 a.m.	1087	1537	907	05/12/2008	11:51 a.m.	1003	1418	838	05/12/2008	11:51 a.m.	9000	4800	1920	06/12/2008	11:45 a.m.	8219	4893	1708	865	25	282
06/12/2008	11:45 a.m.	8219	4893	1708											07/12/2008	08:45 a.m.	7301	4316	1469	918	577	239

Tabla 9 : Procedimiento para cálculo de demanda de efectivo

Fuente : Propia

Elaboración : Silvia Granados Flores

En base al ejemplo anterior y el procedimiento detallado, se obtienen los siguientes registros para la demanda de efectivo:

DEMANDA DE EFECTIVO EN CAJERO AUTOMÁTICO								
PERIODO	FECHA INICIO	HORA INICIO	FECHA FIN	HORA FIN	CAJERO	DEMANDA		
						CASETERA 1 S/.100	CASETERA 2 S/.20	CASETERA 3 US\$20
01/12/08	01/12/08	08:45 a.m.	02/12/08	09:15 a.m.	C123	1420	499	312
02/12/08	02/12/08	09:15 a.m.	03/12/08	08:45 a.m.	C123	915	416	468
03/12/08	03/12/08	08:45 a.m.	04/12/08	08:45 a.m.	C123	582	400	273
04/12/08	04/12/08	08:45 a.m.	05/12/08	08:45 a.m.	C123	751	320	185
05/12/08	05/12/08	08:45 a.m.	06/12/08	11:45 a.m.	C123	865	25	282
06/12/08	06/12/08	11:45 a.m.	07/12/08	08:45 a.m.	C123	918	577	239

Tabla 10 : Resultado del cálculo de demanda de efectivo

Fuente : Propia

Elaboración : Silvia Granados Flores

Cálculo de la Demanda de Efectivo en Monto Total

La variable Monto Total se refiere al valor total de la demanda en moneda de Nuevo Soles. Cabe considerar que para la conversión de dólares a soles se empleó el tipo de cambio promedio para el 2009 de 3.00, según la Superintendencia de Banca y Seguros.

El cálculo de la variable Monto Total se efectúa a través de la siguiente fórmula:

$$\text{Monto Total} = \text{Demanda Billetes S/.100} * 100 + \text{Demanda Billetes S/.20} * 20 + \text{Demanda Billetes US\$20} * 20 * 3.00$$

De esta manera el registro de Monto de demanda de efectivo se obtiene de la siguiente manera:

Monto de demanda de efectivo

$$MD_I^X = \{P_I, Fi_I, Hi_I, Ff_I, Hf_I, C_X, MD_I^{X,S/.100}, MD_I^{X,S/.20}, MD_I^{X,US\$20}\}$$

Donde:

MD_I^X : Valor Total de la demanda de efectivo en el cajero automático (ATM) "X" en el intervalo de tiempo "I".

P_I : Denominación del periodo "I" en el que se mide la demanda de efectivo. Este intervalo se establece desde el punto de inicio $\{F_{i_I}, H_{i_I}\}$ al punto final $\{F_{f_I}, H_{f_I}\}$.

F_{i_I} : Fecha del punto de inicio del intervalo de tiempo "I".

H_{i_I} : Hora del punto de inicio del intervalo de tiempo "I".

F_{f_I} : Fecha del punto final del intervalo de tiempo "I".

H_{f_I} : Hora del punto final del intervalo de tiempo "I".

C_X : Cajero automático "X"

$MD_I^{X,S/.100}$: Valor Total de la demanda de billetes de S/.100 en el cajero automático (ATM) "X" en el intervalo de tiempo "I".

$MD_I^{X,S/.20}$: Valor Total de la demanda de billetes de S/.20 en el cajero automático (ATM) "X" en el intervalo de tiempo "I".

$MD_I^{X,US\$20}$: Valor Total de la demanda de billetes de US\$20 en el cajero automático (ATM) "X" en el intervalo de tiempo "I".

Comportamiento de los datos

En los siguientes gráficos se presentan los comportamientos de la demanda histórica calculada para los cajeros automáticos A, B, C, D y E. Tanto para número de billetes en las tres nóminas así como el Monto Total.

A continuación se presenta el comportamiento de las 4 variables en los Gráficos 25, 26, 27, 28 y 29:

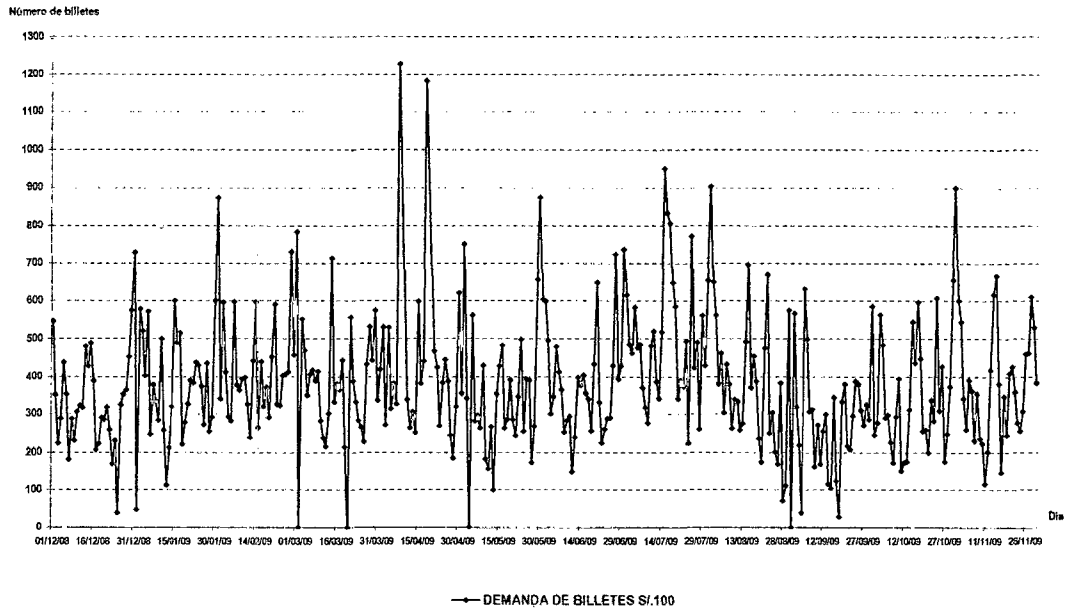


Gráfico 35.1 : Cajero A - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

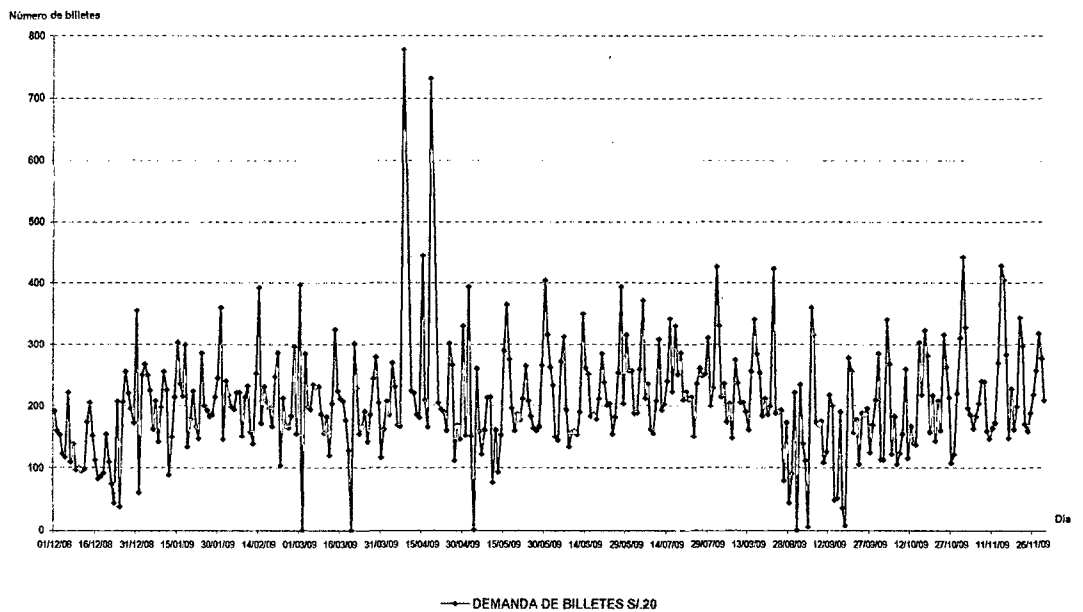


Gráfico 36.2 : Cajero A - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

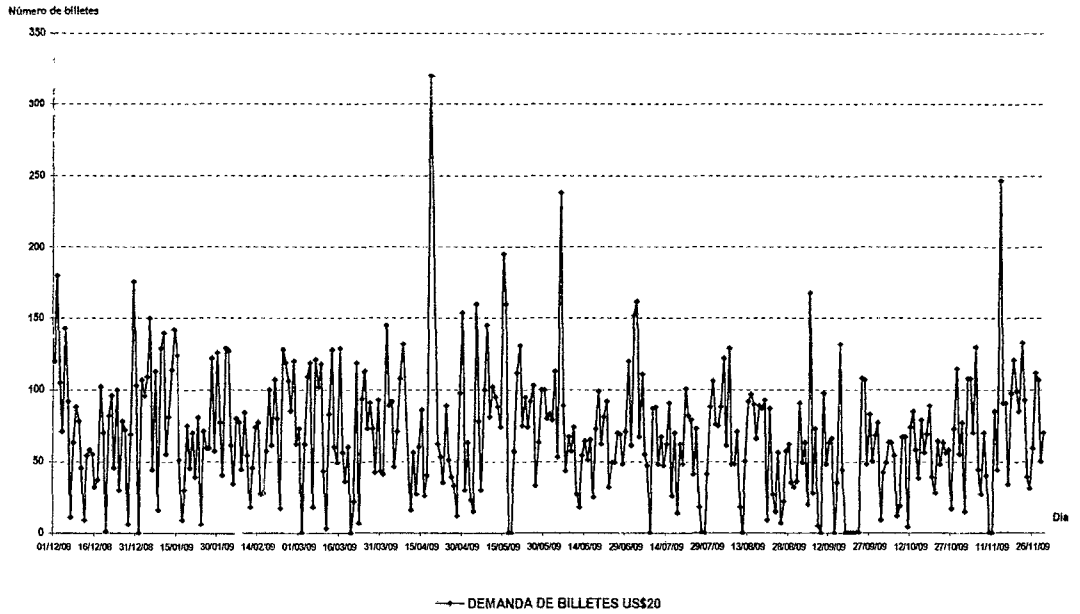


Gráfico 37.3 : Cajero A - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

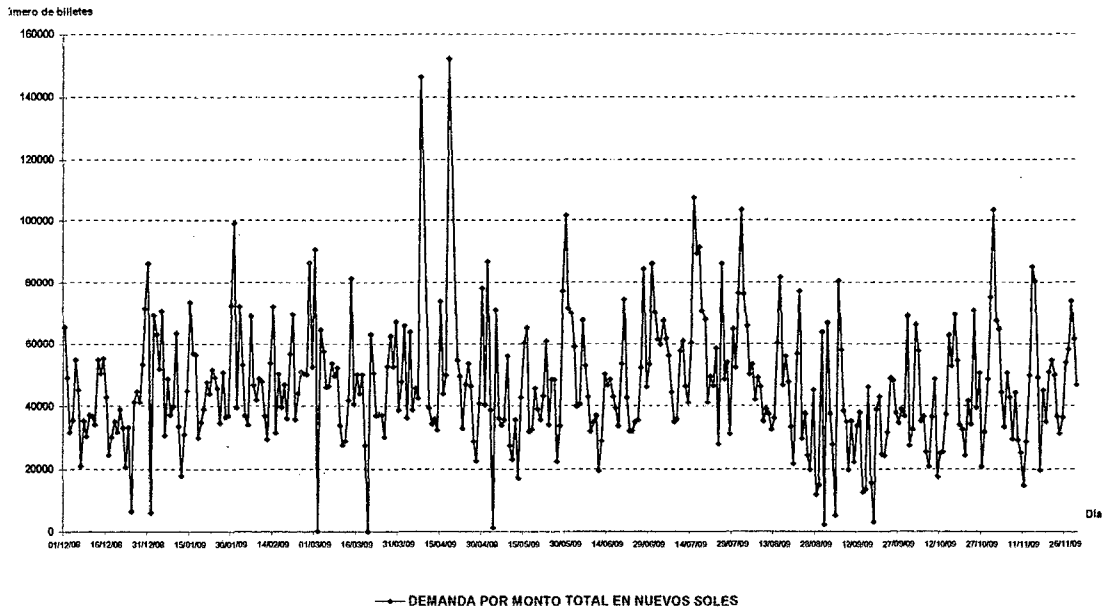


Gráfico 38.4 : Cajero A - Demanda monto total
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

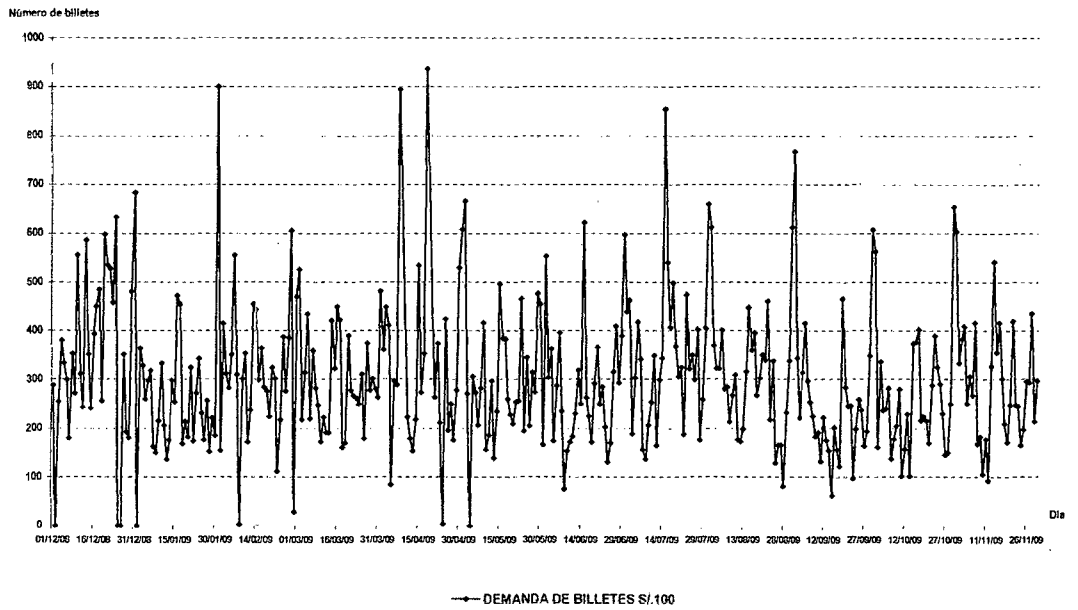


Gráfico 39.1 : Cajero B - Demanda de billetes

Fuente : Empresa Financiera local

Elaboración : Silvia Granados Flores

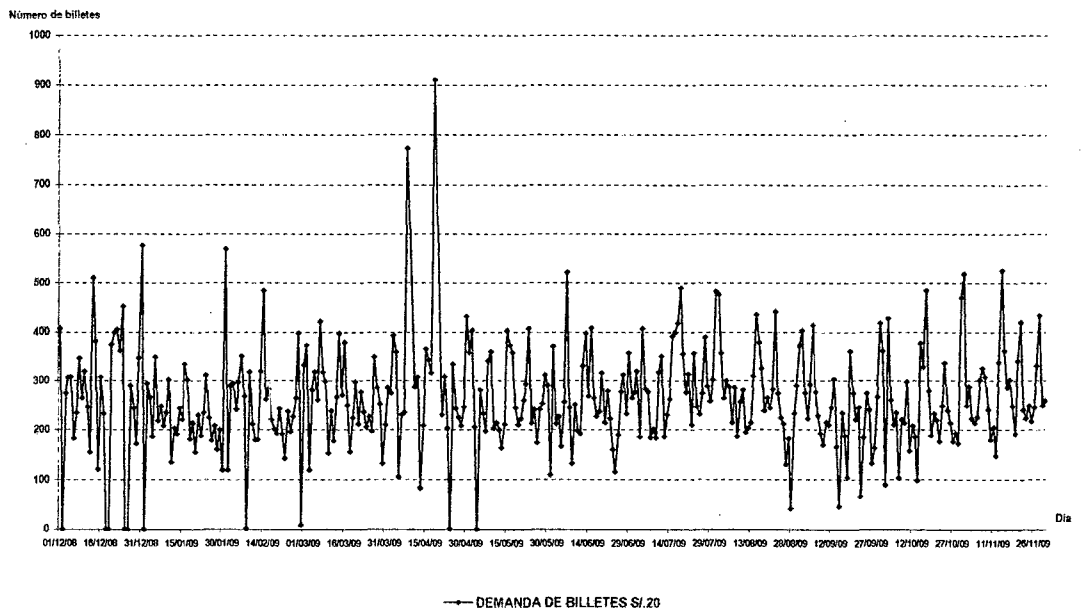


Gráfico 40.2 : Cajero B - Demanda de billetes

Fuente : Empresa Financiera local

Elaboración : Silvia Granados Flores

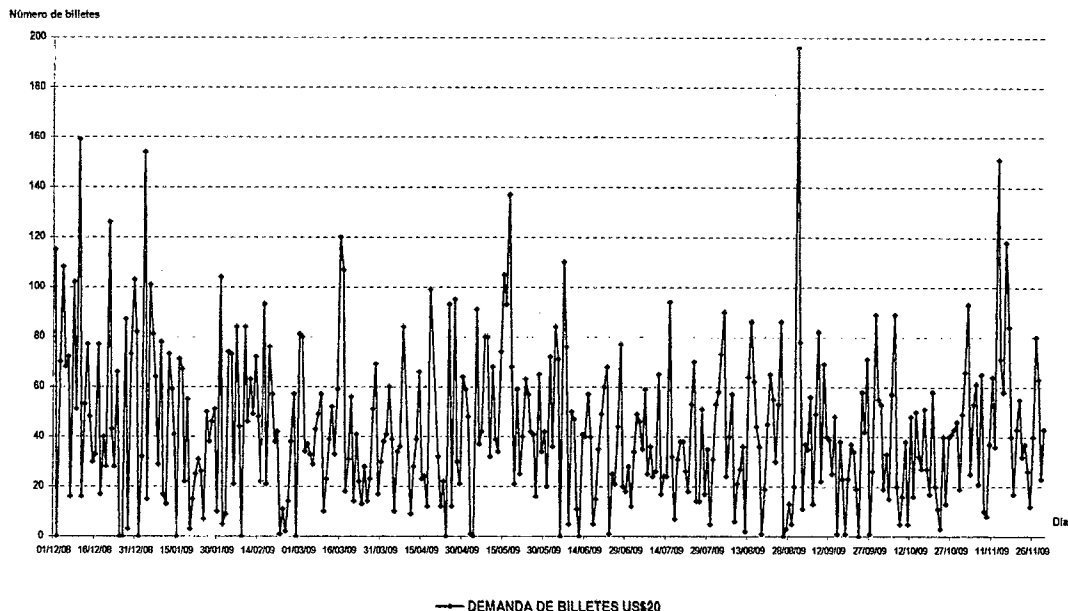


Gráfico 41.3 : Cajero B - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

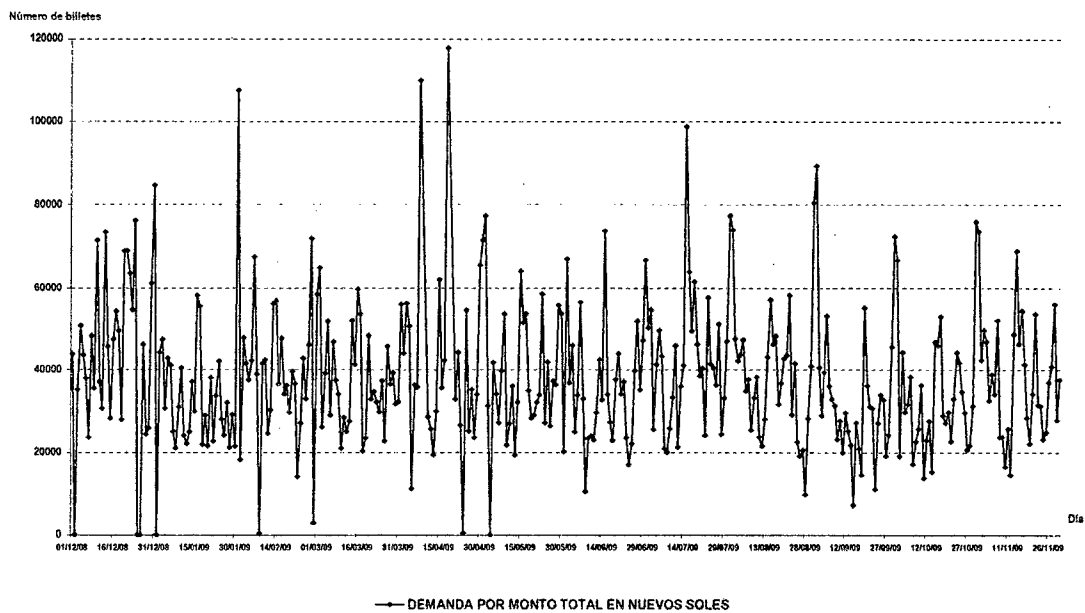


Gráfico 42.4 : Cajero B - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

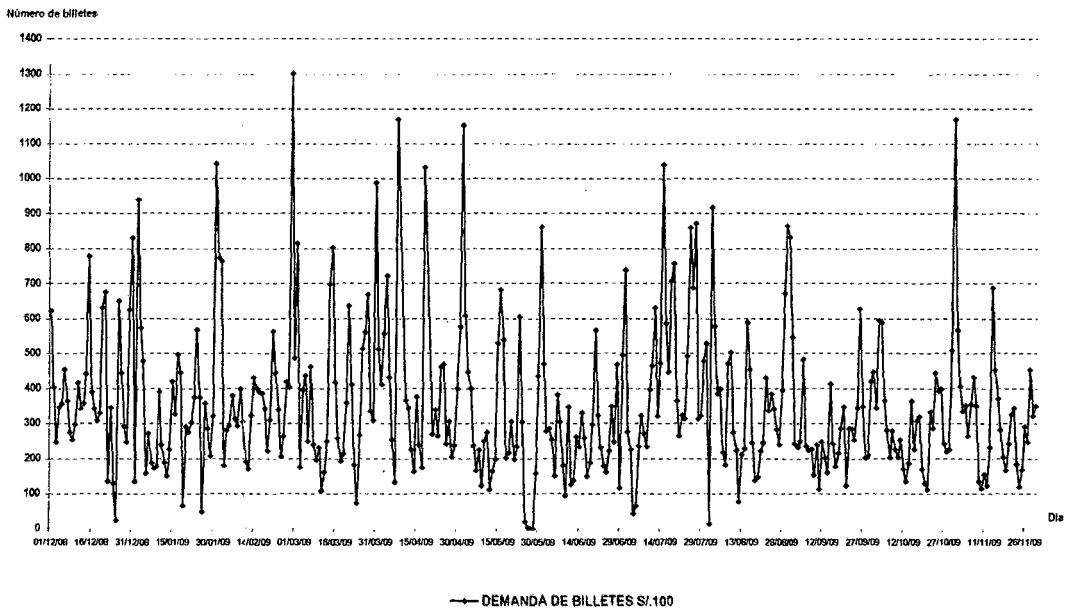


Gráfico 43.1 : Cajero C - Demanda de billetes

Fuente : Empresa Financiera local

Elaboración : Silvia Granados Flores

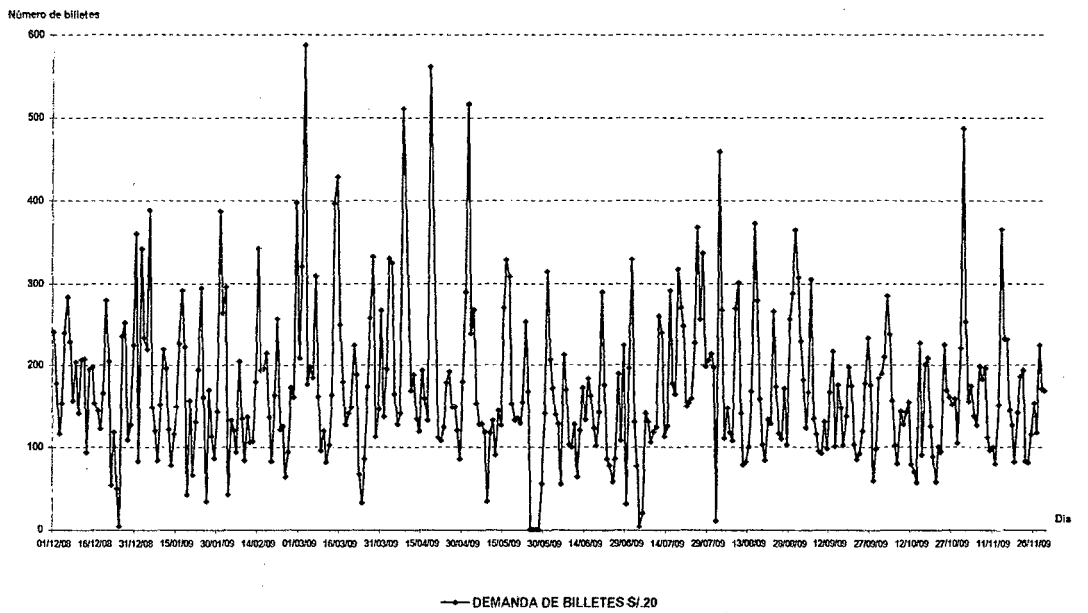


Gráfico 44.2 : Cajero C - Demanda de billetes

Fuente : Empresa Financiera local

Elaboración : Silvia Granados Flores

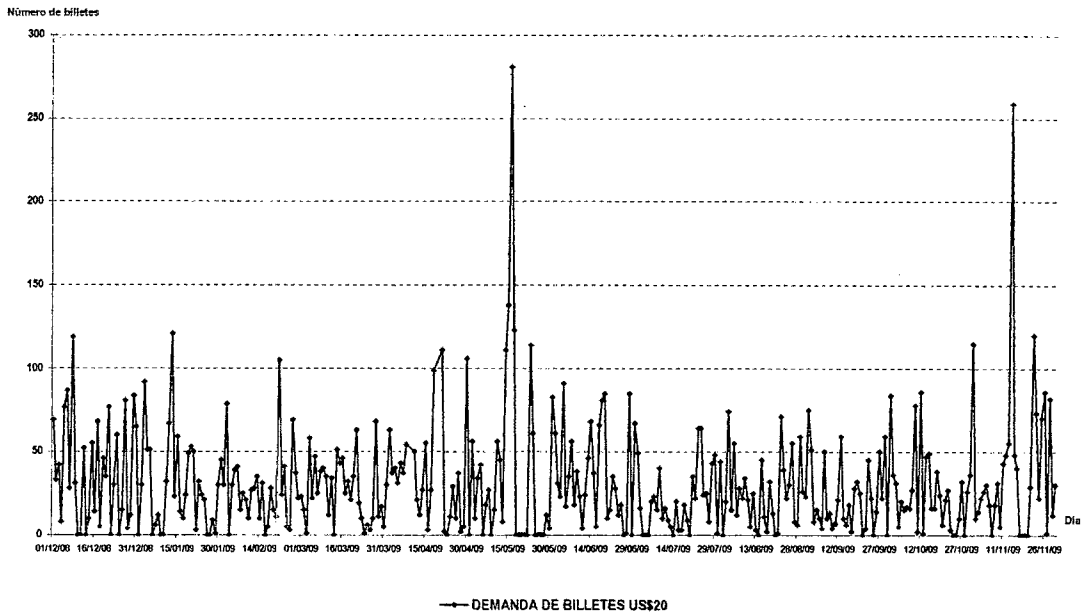


Gráfico 45.3 : Cajero C - Demanda de billetes

Fuente : Empresa Financiera local

Elaboración : Silvia Granados Flores

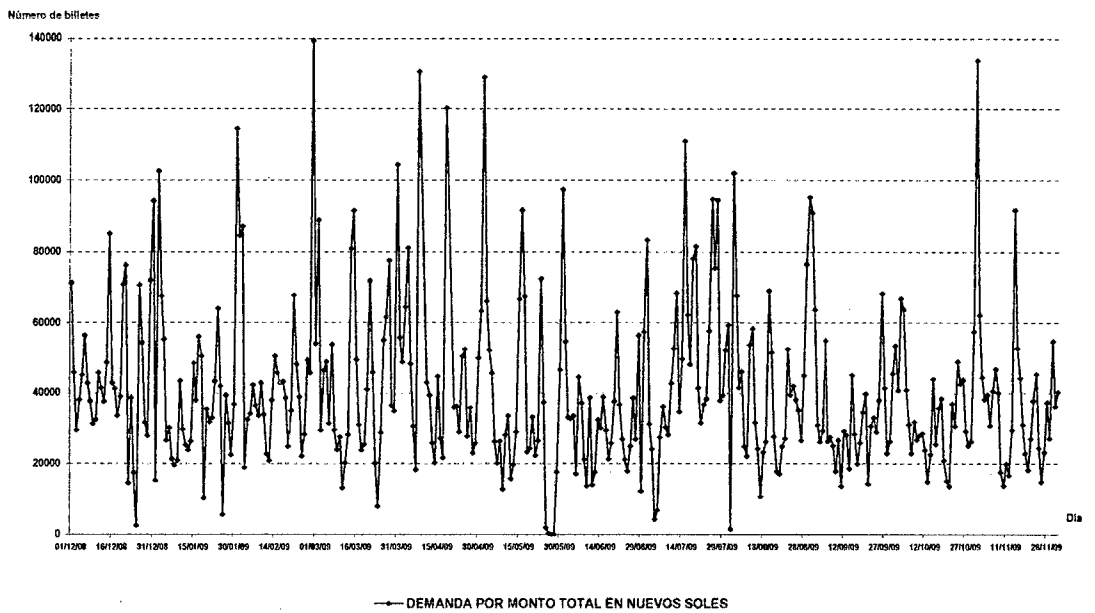


Gráfico 46.4 : Cajero C - Demanda de billetes

Fuente : Empresa Financiera local

Elaboración : Silvia Granados Flores

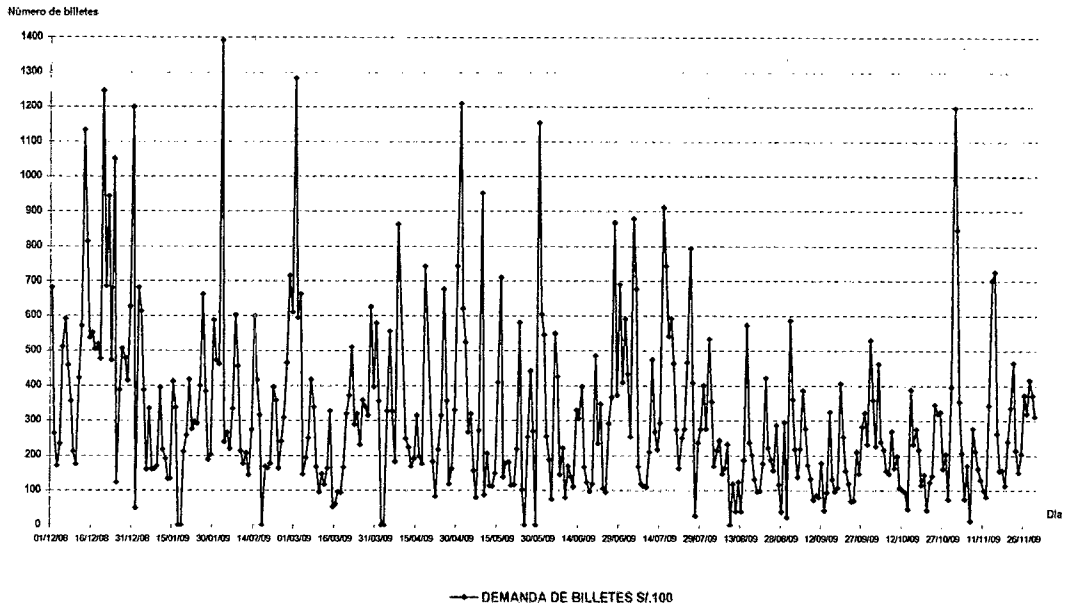


Gráfico 47.1 : Cajero D - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

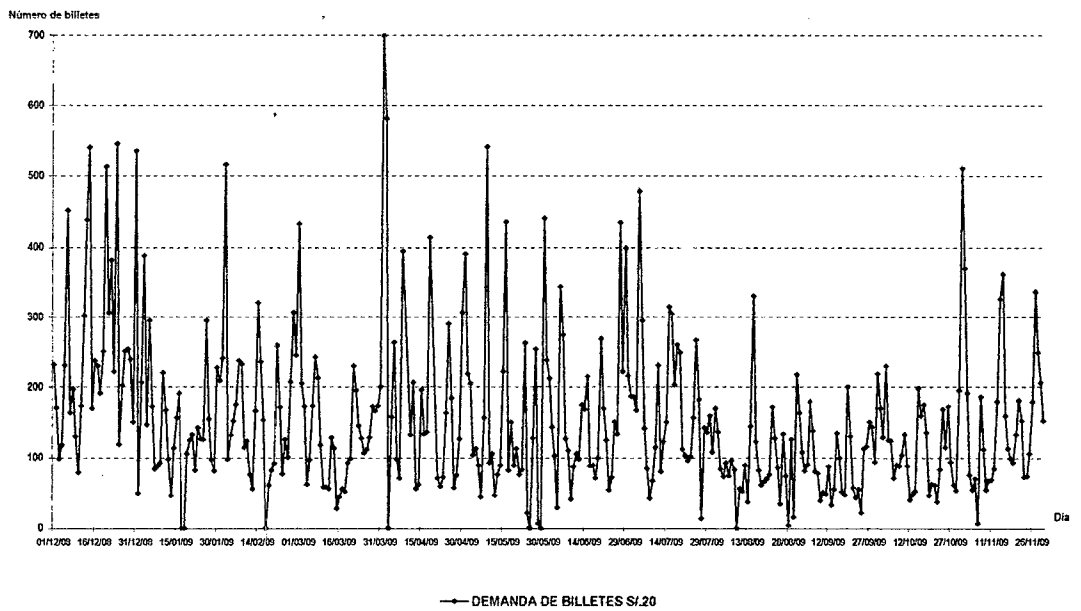


Gráfico 48.2 : Cajero D - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

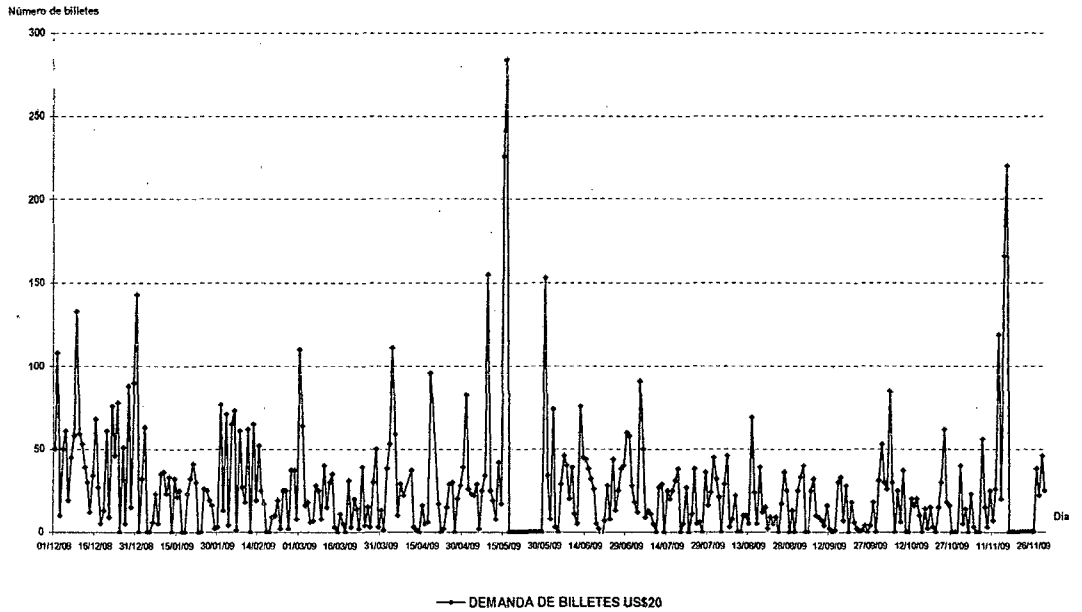


Gráfico 49.3 : Cajero D - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

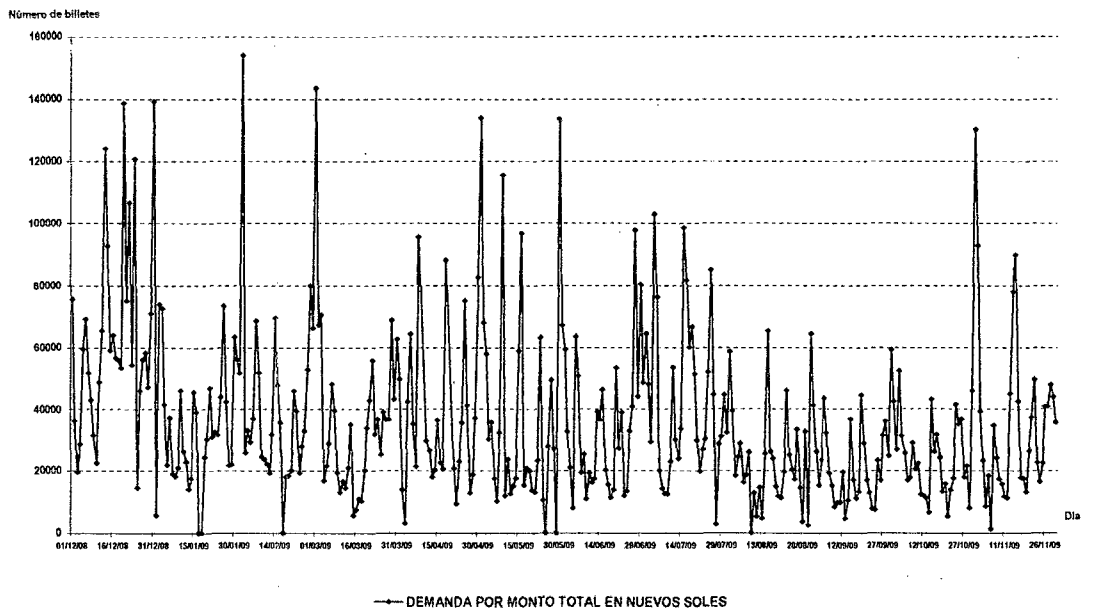


Gráfico 50.4 : Cajero D - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

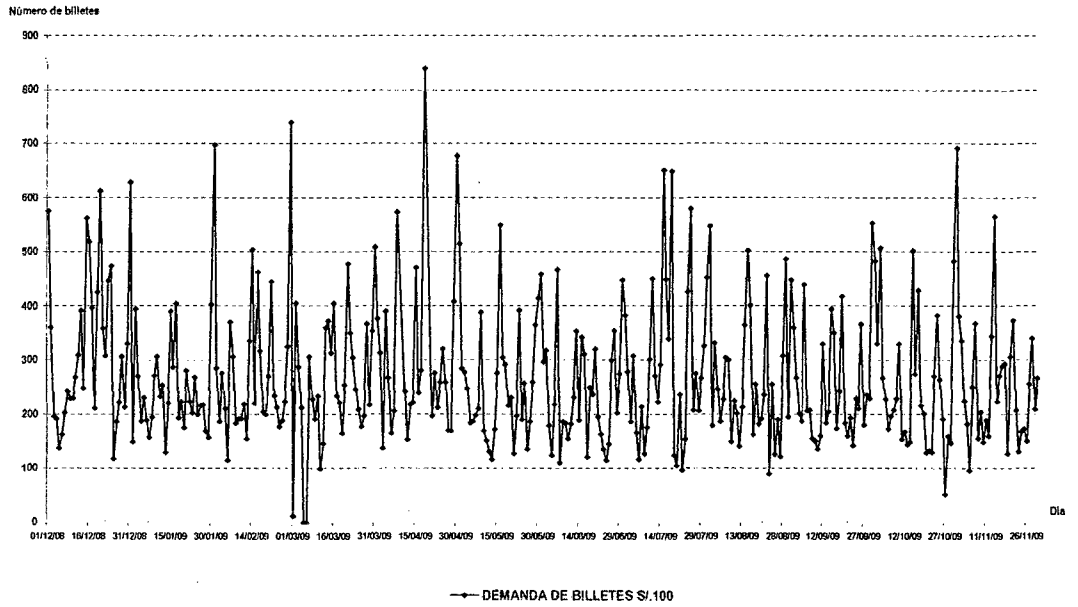


Gráfico 51.1 : Cajero E - Demanda de billetes

Fuente : Empresa Financiera local

Elaboración : Silvia Granados Flores

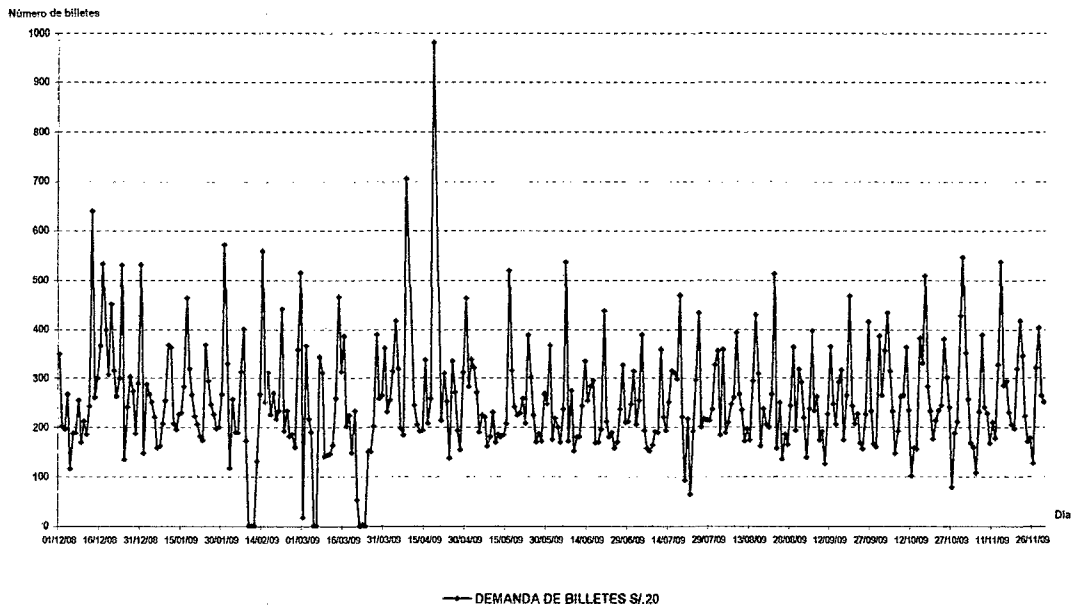


Gráfico 52.2 : Cajero E - Demanda de billetes

Fuente : Empresa Financiera local

Elaboración : Silvia Granados Flores

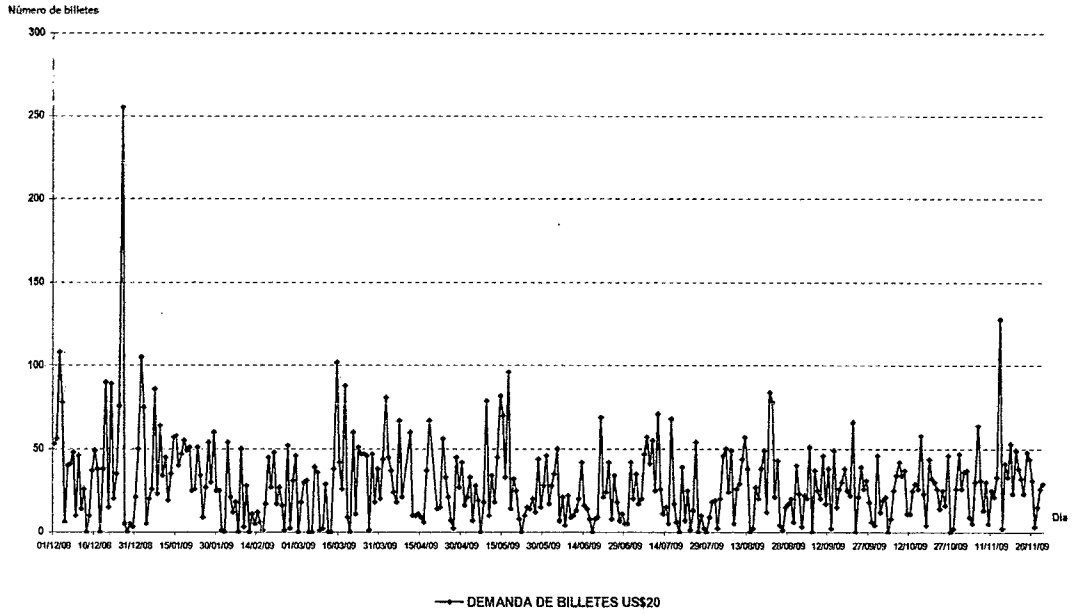


Gráfico 53.3 : Cajero E - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

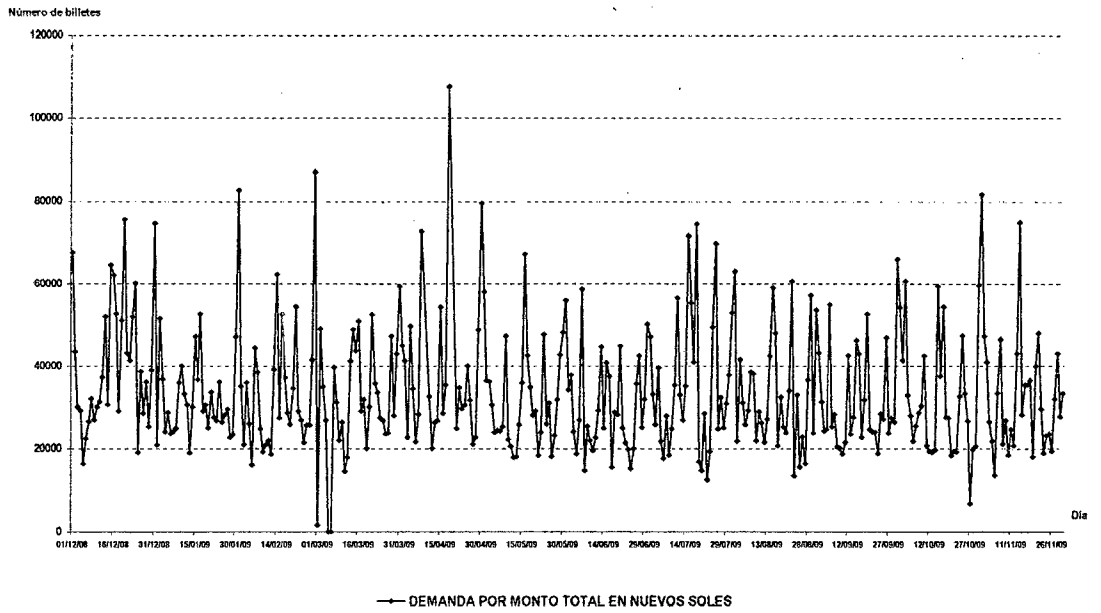


Gráfico 54.4 : Cajero E - Demanda de billetes
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

Impacto en la Demanda por indisponibilidad del servicio

Tal como se indica en el Capítulo I, los cajeros automáticos (ATM's) son dispositivos que ejecutan transacciones a solicitud del usuario y cuya disponibilidad de servicio depende principalmente de los recursos e insumos requeridos para la ejecución de las operaciones. Bajo esta premisa, un cajero automático podría permanecer fuera de servicio por falta de recursos (p.e. falta de papel para log auditor) o fallas técnicas (p.e. fallas de comunicación con el autorizador) impidiendo la atención al público por un lapso de tiempo e impactando principalmente en las estadísticas que permiten calcular el comportamiento de la demanda de efectivo.

A continuación, en la Tabla 11 se presenta un reporte con el número incidentes por mes presentado durante el periodo analizado y sobre los ATM's incluidos en la investigación:

REPORTE DE INCIDENTES Desde 01/12/08 al 30/11/09					
MES	CAJERO AUTOMÁTICO				
	A	B	C	D	E
Dic-08	6	6	9	2	5
Ene-09	2	3	2	4	4
Feb-09	4	4	8	4	7
Mar-09	3	9	9	7	7
Abr-09	5	8	11	2	7
May-09	11	8	9	7	5
Jun-09	3	3	5	4	6
Jul-09	5	6	8	4	6
Ago-09	3	8	4	7	2
Sep-09	13	8	5	4	2
Oct-09	3	5	2	0	2
Nov-09	3	2	4	3	3
PROMEDIO	5	6	6	4	5

Tabla 11 : Número de incidentes reportados por cajero automático

Fuente : Empresa Financiera local

Elaboración: Silvia Granados Flores

En la Tabla 11 puede observarse que el número máximo de incidentes por mes es 13 para el cajero automático A; 8, en B; 11, en C; 7, en D y 7, en E.

En el Gráfico 30 se distinguen cinco principales temas (Hardware, software, comunicación, recursos y un grupo genérico de otros), sobre los cuales se identificaron 17 puntos cuya falla o manifestación podría generar interrupciones en el servicio de los ATM's.

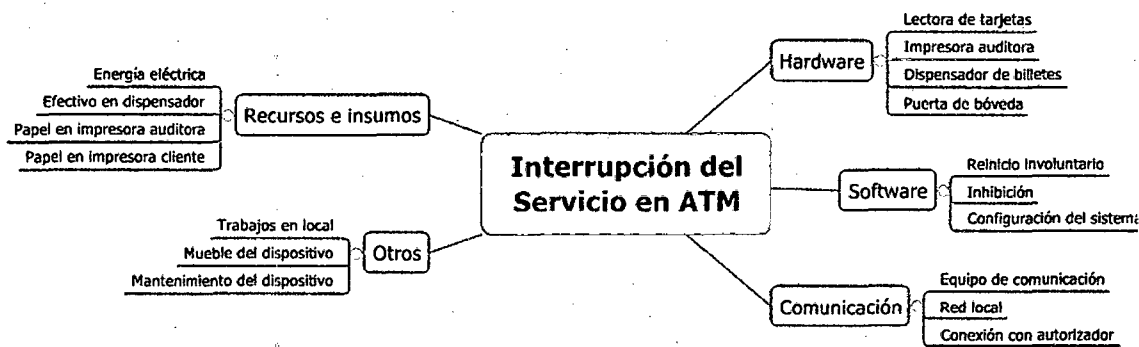


Gráfico 55 : Incidentes en ATM's
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

Actualmente no es posible identificar el tiempo promedio en el cual los cajeros automáticos no estuvieron disponibles, ya que los sistemas de información sobre monitoreo de la red de cajeros automáticos de la Empresa Financiera no almacenan esta información. Sin embargo, para la investigación será necesario considerar que la interrupción del servicio es una de las principales causas de imprecisión en el pronóstico.

Distribución de frecuencia para la Demanda calculada

A fin de conocer la distribución de la frecuencia de los valores calculados, en la Tabla 9 se identifican las características de esta distribución:

CAJERO AUTOMÁTICO	DEMANDA DE BILLETES DE:	MEDIA	MEDIANA	MÁXIMO	MÍNIMO	DESVIACIÓN ESTÁNDAR	OBLICUIDAD	SIMETRÍA	DISTRIBUCIÓN DE FRECUENCIA	DIAGRAMA DE BLOQUE
A	S/. 100	382.31	356.00	1,228.00	0.00	174.16	1.05	Asimétrica a la derecha		
	S/. 20	206.99	199.00	778.00	0.00	87.78	1.52	Asimétrica a la derecha		
	US\$20	68.84	66.00	320.00	0.00	43.05	1.06	Asimétrica a la derecha		
	MONTO TOTAL	46,501.33	44,240.00	152,140.00	0.00	20,029.43	1.07	Asimétrica a la derecha		
B	S/. 100	301.68	282.00	937.00	0.00	146.46	1.11	Asimétrica a la derecha		
	S/. 20	263.36	249.00	911.00	0.00	107.89	0.92	Asimétrica a la derecha		
	US\$20	43.05	38.00	196.00	0.00	30.70	1.11	Asimétrica a la derecha		
	MONTO TOTAL	38,018.50	35,600.00	117,860.00	0.00	17,367.73	1.02	Asimétrica a la derecha		

Tabla 12 : Indicadores estadísticos e histogramas – Cajeros A y B

Fuente : Empresa Financiera local

Elaboración: Silvia Granados Flores

CAJERO AUTOMÁTICO	DEMANDA DE BILLETES DE:	MEDIA	MEDIANA	MÁXIMO	MÍNIMO	DESMIACIÓN ESTANDAR	OBLICUIDAD	SIMETRÍA	DISTRIBUCIÓN DE FRECUENCIA	DIAGRAMA DE BLOQUE
C	S/. 100	353.58	307.00	1,301.00	0.00	210.75	1.50	Asimétrica a la derecha		
	S/. 20	171.02	154.00	588.00	0.00	91.35	1.29	Asimétrica a la derecha		
	US\$20	31.33	24.00	281.00	0.00	32.88	2.74	Asimétrica a la derecha		
	MONTO TOTAL	40,625.26	36,140.00	139,380.00	0.00	23,268.77	1.44	Asimétrica a la derecha		
D	S/. 100	353.04	289.00	1,392.00	0.00	255.02	1.64	Asimétrica a la derecha		
	S/. 20	171.32	143.00	700.00	0.00	118.20	1.66	Asimétrica a la derecha		
	US\$20	28.64	21.00	284.00	0.00	34.95	3.31	Asimétrica a la derecha		
	MONTO TOTAL	36,214.07	29,000.00	153,800.00	0.00	27,086.54	1.63	Asimétrica a la derecha		
E	S/. 100	271.00	234.00	740.00	0.00	126.41	1.24	Asimétrica a la derecha		
	S/. 20	248.29	227.00	640.00	0.00	106.84	1.38	Asimétrica a la derecha		
	US\$20	29.15	22.00	255.00	0.00	27.66	2.66	Asimétrica a la derecha		
	MONTO TOTAL	33,738.61	29,180.00	107,640.00	0.00	14,964.29	1.25	Asimétrica a la derecha		

Tabla 13 : Indicadores estadísticos e histogramas – Cajeros C, D y E

Fuente : Empresa Financiera local

Elaboración: Silvia Granados Flores

En todos los casos se observa que la asimetría por Monto Total es menor al promedio del Número de Billetes demandados en las tres denominaciones con, por lo que en adelante se optará por efectuar el análisis con la variable Monto Total.

2.5 RESUMEN DEL CAPÍTULO

En este capítulo se detallaron las variables a considerar y actividades necesarias para el cálculo de la demanda de efectivo.

Las variables que se emplearon se refieren al stock de billetes en las caseteras del cajero automático y los pedidos de abastecimiento efectuados en el periodo, además se consideran como parámetros la hora inicio y fin del periodo a determinarse para la demanda.

De esta manera el cálculo de la demanda mantiene la siguiente dinámica:

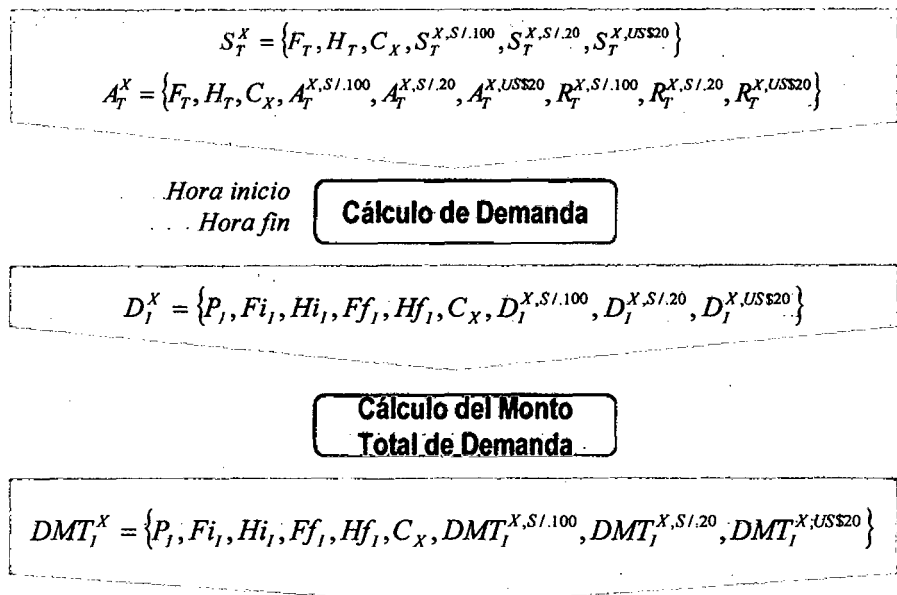


Gráfico 56 : Análisis de Datos - Resumen
Fuente : Propia
Elaboración : Silvia Granados Flores

CAPÍTULO III

LIMPIEZA DE DATOS

3.1. DESCRIPCIÓN GENERAL

En este capítulo se describirán las técnicas a desarrollar a fin de obtener información completa sobre la demanda de efectivo en el lapso de tiempo analizado. De tal manera que permita entregar al modelo de red neuronal artificial un conjunto de patrones base para su pronóstico.

Inicialmente se tratarán los valores null (valores vacíos) de demanda de efectivo y su tratamiento para mantener la continuidad en la serie de tiempo.

En segundo lugar se trabajarán los valores imprecisos, aquellos que se originan porque algún registro de stock de billetes en las caseteras (S_T^x) no es extraído en el límite de inicio o fin del periodo de tiempo establecido como medida.

Asimismo se analizarán los valores outliers (valores anómalos) de demanda de efectivo y se detallarán las posibles técnicas a emplear a fin de suavizar estos extremos de la curva.

Por último se revisarán técnicas de transformación de datos a fin de mejorar la interpretación de los mismos así como elevar la performance de la red neuronal.

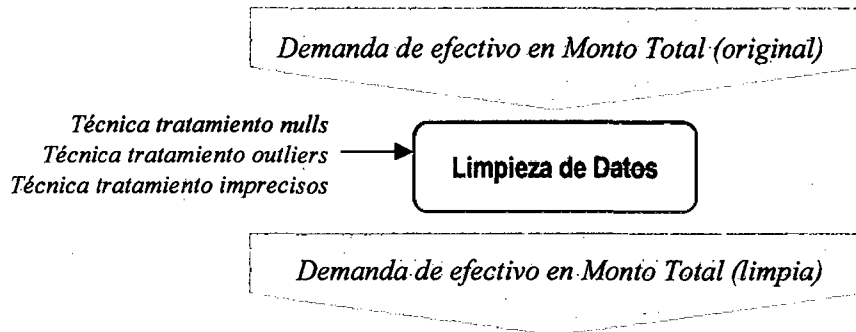


Gráfico 57 : Limpieza de Datos
Fuente : Propia
Elaboración : Silvia Granados Flores

3.2. MÓDULO LIMPIEZA DE DATOS

Luego de efectuarse el proceso de análisis de datos y calculada la demanda histórica de efectivo, se desarrollará el proceso de limpieza de datos.

La información original así como la calculada puede contener valores atípicos, valores faltantes o imprecisos debido a factores externos. En este módulo se suavizará el impacto de los valores atípicos, se imputarán los valores faltantes y se aproximarán los datos imprecisos.

3.2.1. ANÁLISIS DE VALORES IMPRECISOS

En base a la información manejada para este análisis, se identifican como valores imprecisos a los registros sobre el stock de billetes en las caseteras de los cajeros automáticos cuya hora de emisión no corresponde a la hora fin del periodo.

En la Tabla 14 se presentan 7 registros sobre stock de billetes, de los cuales 2 registros representan datos imprecisos debido a que la hora indicada no corresponde a la hora fin del periodo:

REPORTE DE STOCK DE BILLETES TURNO 8:45 a.m.						
--	--	--	--	--	--	--

N°	FECHA	HORA	CAJERO	CASETERA	CASETERA	CASETERA
				NRO 1 S/.100	NRO 2 S/.20	NRO 3 US\$20
1	01/12/2008	08:45 a.m.	C123	4755	3172	2145
2	02/12/2008	09:15 a.m.	C123	3336	2673	1833
3	03/12/2008	08:45 a.m.	C123	0	2257	1365
4	04/12/2008	08:45 a.m.	C123	1838	1856	1092
5	05/12/2008	08:45 a.m.	C123	1087	1537	907
6	06/12/2008	11:45 a.m.	C123	8219	4893	1708
7	07/12/2008	08:45 a.m.	C123	7301	4316	1469

Tabla 14 : Identificación de registros imprecisos

Fuente : Empresa Financiera local

Elaboración: Silvia Granados Flores

En el ejemplo anterior los registros 2 y 5 son considerados registros imprecisos para el cálculo de la demanda de efectivo debido a que el atributo Hora (09:15 a.m. y 11:45 a.m. respectivamente) excede la Hora Inicio/Hora Fin del periodo, por este motivo será necesario efectuar una transformación a fin de reducir el impacto de la imprecisión.

En la Tabla 15 se presenta el número de valores imprecisos dentro de los datos recolectados para cada cajero automático:

VALORES IMPRECISOS Desde 01/12/08 al 30/11/09				
CAJERO AUTOMÁTICO				
A	B	C	D	E
3	3	3	3	3

TOTAL OBSERVACIONES: 361

% VALORES IMPRECISOS: 0.83%

Tabla 15 : Número de valores imprecisos de la muestra

Fuente : Empresa Financiera local

Elaboración: Silvia Granados Flores

Técnica empleada

Para reducir el impacto de los registros identificados como imprecisos se aplicará una aproximación a través de una proporción lineal a fin de aproximar el valor de stock de billetes para las 8:45 a.m.

A. **CASO 1:** La proporción se establece sobre el punto anterior al analizado. Por ejemplo para el Registro 2 se considerará el Registro 1 (01/12/08 8:45 a.m.) con un stock de 4755 billetes. A través de una resta se calculan los minutos transcurridos desde el Registro 1 hasta el Registro 2 (Total: 1470 min.), luego con un cálculo de proporción se calcula el valor que correspondería al 02/12/08 8:45 a.m.

El cálculo se representa así:

$$S_T^X = S_{T-1}^X \frac{[S_{T-1}^X - S_T^X] \times [S_T^X \{F_T, H_T\} - S_{T-1}^X \{F_{T-1}, H_{T-1}\}]}{[S_T^X \{F_T, H_T\} - S_{T-1}^X \{F_{T-1}, H_{T-1}\}]}$$

Donde:

S_T^X : Stock de billetes en las caseteras del cajero automático "X" en un instante de tiempo " T " .

S_{T-1}^X : Stock de billetes en las caseteras del cajero automático "X" en un instante de tiempo " T' " .

$S_T^X \{F_T, H_T\}$: Fecha y hora en la que se registra el stock de billetes en las caseteras del cajero automático "X" en un instante de tiempo " T " .

$S_T^X \{F_T, H_T\}$: Fecha y hora en la que se registra el stock de billetes en las caseteras del cajero automático "X" en un instante de tiempo " T' " .

$S_{T-1}^X \{F_{T-1}, H_{T-1}\}$: Fecha y hora en la que se registra el stock de billetes en las caseteras del cajero automático "X" en un instante de tiempo "T-1" .

B. **CASO 2:** El registro impreciso corresponde a un punto en el tiempo inmediatamente posterior al punto de recarga de efectivo. Por ejemplo para el Registro 6, se identifica que se ha efectuado un abastecimiento entre el Registro 5 y Registro 6 (incremento del stock). Por este motivo se considerará el Registro 1 de la Tabla 4 (Registro de Pedidos de Abastecimiento) como punto de partida para efectuar la proporción. Se calculan los minutos transcurridos desde el 05/12/08 09:21 a.m. hasta el 06/12/08 11:45 a.m. para luego con un cálculo de proporción calcular el stock correspondiente al 06/12/08 08:45 a.m.

El cálculo se representaría con:

$$S_T^X = A_t^X - \frac{[A_t^X - S_T^X] \times [S_T^X \{F_T, H_T\} - A_t^X \{F_t, H_t\}]}{[S_T^X \{F_T, H_T\} - A_t^X \{F_t, H_t\}]}$$

$$S_T^X \{F_T, H_T\} \neq A_t^X \{F_t, H_t\}$$

Donde:

S_T^X : Stock de billetes en las caseteras del cajero automático "X" en un instante de tiempo " T ".

$S_{T'}^X$: Stock de billetes en las caseteras del cajero automático "X" en un instante de tiempo " T' ".

$S_T^X \{F_T, H_T\}$: Fecha y hora en la que se registra el stock de billetes en las caseteras del cajero automático "X" en un instante de tiempo " T ".

$S_{T'}^X \{F_{T'}, H_{T'}\}$: Fecha y hora en la que se registra el stock de billetes en las caseteras del cajero automático "X" en un instante de tiempo " T' ".

$A_t^X \{F_t, H_t\}$: Fecha y hora en la que se realiza el abastecimiento de billetes en el cajero automático "X" en un instante de tiempo "t".

En el Gráfico 33 se representa gráficamente los métodos de cálculo para el Caso 1 y Caso 2:

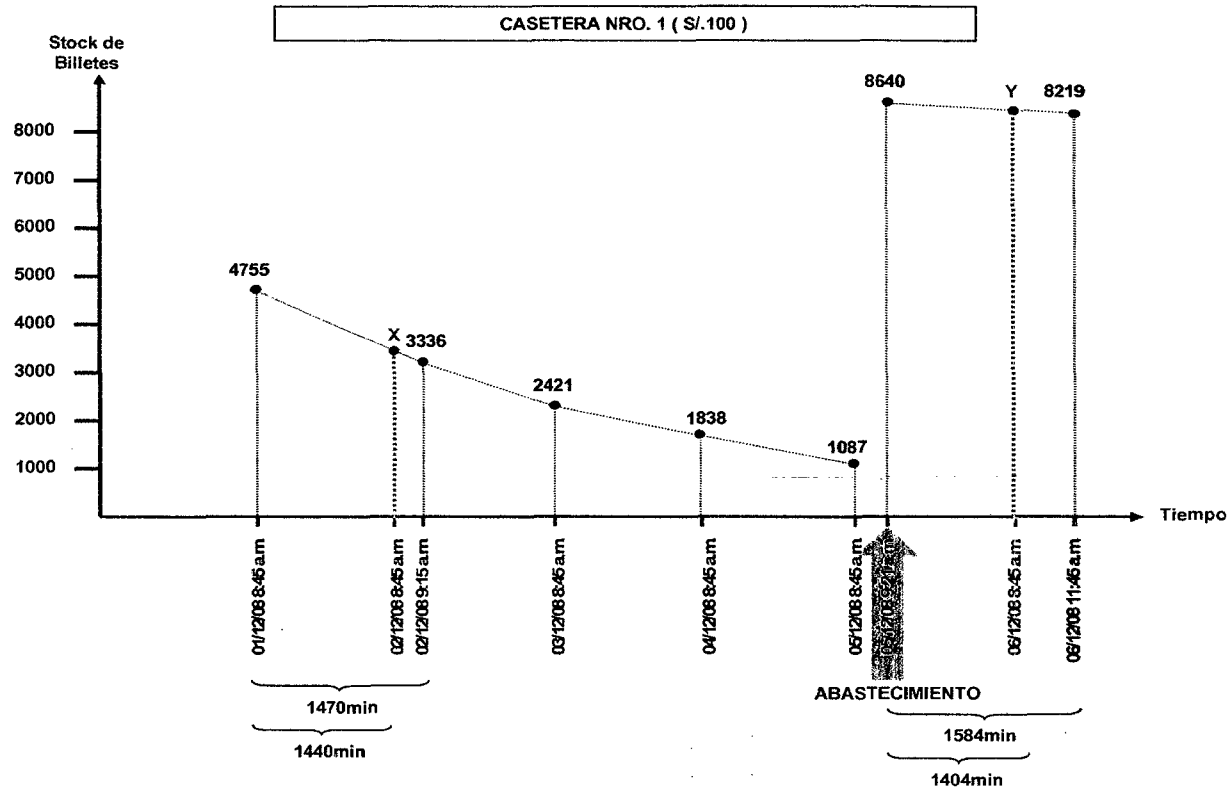


Gráfico 58 : Cálculo sobre registros imprecisos
Fuente : Propia
Elaboración : Silvia Granados Flores

En base a la técnica de reemplazo propuesta para valores imprecisos, en la Tabla 16 se calcula el %desviación de las medidas de tendencia central originales (Media y Mediana) con respecto a las resultantes luego de efectuar el reemplazo a través de la técnica establecida.

		DEMANDA ORIGINAL			DEMANDA SIN VALORES IMPRECISOS			%DESVIACION		
		CASETERA	CASETERA	CASETERA	CASETERA	CASETERA	CASETERA	CASETERA	CASETERA	CASETERA
		NRO 1-2 \$/100	NRO 3 \$/20	NRO 4 US\$20	NRO 1-2 \$/100	NRO 3 \$/20	NRO 4 US\$20	NRO 1-2 \$/100	NRO 3 \$/20	NRO 4 US\$20
A	MEDIA	382.31	206.99	68.84	378.12	204.72	68.09	1.10%	1.10%	1.10%
	MEDIANA	356.00	199.00	66.00	362.17	201.00	66.00	1.73%	1.01%	0.00%
B	MEDIA	301.68	263.37	43.06	298.37	260.48	42.58	1.10%	1.10%	1.10%
	MEDIANA	282.00	249.00	38.00	282.00	249.00	38.00	0.00%	0.00%	0.00%
C	MEDIA	353.58	171.02	30.77	349.71	169.15	30.44	1.10%	1.10%	1.10%
	MEDIANA	307.00	154.00	23.00	310.00	154.00	23.00	0.98%	0.00%	0.00%
D	MEDIA	315.91	154.79	25.45	312.46	153.09	25.25	1.09%	1.10%	0.77%
	MEDIANA	250.00	127.00	17.00	250.00	127.00	17.00	0.00%	0.00%	0.00%
E	MEDIA	269.08	254.73	28.94	266.13	251.94	28.62	1.10%	1.10%	1.10%
	MEDIANA	230.00	234.00	25.00	229.00	233.00	25.00	0.43%	0.43%	0.00%

Tabla 16 : Evaluación del reemplazo de datos imprecisos

Fuente : Propia

Elaboración : Silvia Granados Flores

Como se observa en la tabla anterior, la desviación originada con el reemplazo oscila entre 0% y 1.73% de las medidas originales.

3.2.2. ANÁLISIS DE VALORES NULL

Se consideran como valores null aquellos valores vacíos que podrían originarse por fallas en el detector de stock, fallas en el sistema, fallas en la transmisión de la información desde los cajeros automáticos al sistema central, entre otros.

Como paso previo, debe analizarse la implicancia de los valores null en la base de datos integral, independientemente de la causa que los origina, analizándose cuánto representa la cantidad de valores null dentro de los datos totales (J. Navarro et. al. [15]).

En el caso analizado, se identifica que existen registros vacíos en los reportes de stock de billetes de las caseteras de los cajeros automáticos debido a que la rutina de software programada para recolectar la información no se ejecutó en determinados días. Sin embargo los

registros null con respecto a los registros totales significan aproximadamente el 2% del total de la base, una proporción no significativa para el total de la base.

Debido a la continuidad de una serie de tiempo, los valores null no serán eliminados de la base de datos sino serán reemplazados por valores no vacíos a través de técnicas que se explicarán a continuación y cuyo efecto de reemplazo no impacten considerablemente las medidas de tendencia central del conjunto de datos (L. Useche [19]).

En la Tabla 13 se presenta el consolidado de valores vacíos dentro de los datos recolectados para cada cajero automático:

VALORES NULL				
Desde 01/12/08 al 30/11/09				
CAJERO AUTOMÁTICO				
A	B	C	D	E
4	4	4	4	4

TOTAL OBSERVACIONES: 361
 % VALORES VACÍOS: 1.11%

Tabla 17 : Número de valores vacíos de la muestra

Fuente : Empresa Financiera local

Elaboración: Silvia Granados Flores

Técnicas propuestas de reemplazo

Se analizarán los resultados de aplicar cuatro técnicas a fin de reemplazar los valores perdidos.

Para seleccionar la metodología que mejor se adecue a la naturaleza de los datos recolectados, se considerará como la técnica de mejor performance aquella que cumple las siguientes restricciones:

- Impacte en menor proporción a las medidas originales de tendencia central: media y mediana.

- Cuyo valor de reemplazo sea menor o igual al stock de billetes del día inmediato anterior y mayor o igual al stock del día inmediato siguiente.
- En caso de producirse una recarga de billetes entre el reporte anterior y el analizado, el valor de reemplazo deberá ser menor o igual al pedido de abastecimiento. Por el contrario, si la recarga se efectúa entre el periodo analizado y el siguiente reporte de stock, el valor de reemplazo deberá ser mayor o igual al monto recargado.

A. Técnica 1: Promedio diario

Esta técnica busca clasificar los registros por el día de la semana al que corresponden y reemplazar los valores vacíos por el promedio del día de la semana al que hacen referencia.

En la Tabla 18 se presenta un ejemplo, el día lunes 08/12/2008 en el Reporte de Stock de Billetes correspondiente al turno de las 8:45a.m. se identifica un registro con valores vacíos.

REPORTE DE STOCK DE BILLETES TURNO 8:45 a.m.							
N°	FECHA	DÍA	HORA	CAJERO	CASETERA	CASETERA	CASETERA
					NRO 1 S/.100	NRO 2 S/.20	NRO 3 US\$20
1	01/12/2008	Lunes	08:45 a.m.	C123	4755	3172	2145
2	02/12/2008	Martes	09:15 a.m.	C123	3336	2673	1833
3	03/12/2008	Miércoles	08:45 a.m.	C123	0	2257	1365
4	04/12/2008	Jueves	08:45 a.m.	C123	1838	1856	1092
5	05/12/2008	Viernes	08:45 a.m.	C123	1087	1537	907
6	06/12/2008	Sábado	11:45 a.m.	C123	8219	4893	1708
7	07/12/2008	Domingo	08:45 a.m.	C123	7301	4316	1469
8	08/12/2008	Lunes	08:45 a.m.	C123	NULL	NULL	NULL
9	09/12/2008	Martes	08:45 a.m.	C123	4470	3582	1345
10	10/12/2008	Miércoles	08:45 a.m.	C123	3244	3024	1301
11	11/12/2008	Jueves	08:45 a.m.	C123	2463	2488	1295
12	12/12/2008	Viernes	08:45 a.m.	C123	1456	2059	1216
13	13/12/2008	Sábado	08:45 a.m.	C123	17259	10276	3587
14	14/12/2008	Domingo	08:45 a.m.	C123	9783	5783	1968

Tabla 18 : Técnica 1 Promedio diario - Ejemplo

Fuente : Propia

Elaboración : Silvia Granados Flores

A continuación se calcula el stock promedio por día de semana en base al total de registros recolectados, se presenta en la Tabla 19:

PROMEDIOS DIARIOS			
Día	Casetera 1	Casetera 2	Casetera 3
LUNES	2378	1586	1073
MARTES	3903	3127	1589
MIÉRCOLES	1622	2640	1333
JUEVES	2151	2172	1194
VIERNES	1272	1798	1062
SÁBADO	12739	7584	2648
DOMINGO	8542	5050	1719

Tabla 19 : Técnica 1 Promedio diarios - Promedios
Fuente : Propia
Elaboración : Silvia Granados Flores

Finalmente se compara el valor promedio de stock correspondiente al día de semana de la referencia contra los valores reales de stock para el día anterior y posterior al analizado. En la Tabla 20 se presenta el resumen de la comparación, obteniendo como resultado que el promedio calculado satisface la primera restricción; sin embargo el stock promedio no es mayor o igual al stock del día siguiente, lo que podría originar incoherencia en la dinámica de empleo de stock.

Por lo indicado en el párrafo anterior, esta técnica podría generar inconsistencias en el manejo de la información.

EVALUACION			
Promedio diario Lunes (Caso 18/12/2008)	2378	1586	1073
<i>¿Menor o igual al stock del 07.12.08?</i>	<i>SI</i>	<i>SI</i>	<i>SI</i>
<i>¿Mayor o igual al stock del 09.12.08?</i>	<i>NO</i>	<i>NO</i>	<i>NO</i>

Tabla 20 : Técnica 1 Promedio diarios - Evaluación
Fuente : Propia
Elaboración : Silvia Granados Flores

B. Técnica 2: Media Móvil a 2

Se reemplazarán los valores nulos por la media de los dos registros anteriores a éste.

En la Tabla 21 se presenta un ejemplo, el día lunes 08/12/2008 en el Reporte de Stock de Billetes correspondiente al turno de las 8:45a.m. se identifica un registro con valores vacíos.

REPORTE DE STOCK DE BILLETES TURNO 8:45 a.m.							
N°	FECHA	DIA	HORA	CAJERO	CASETERA	CASETERA	CASETERA
					NRO 1 S/.100	NRO 2 S/.20	NRO 3 US\$20
1	01/12/2008	Lunes	08:45 a.m.	C123	4755	3172	2145
2	02/12/2008	Martes	09:15 a.m.	C123	3336	2673	1833
3	03/12/2008	Miércoles	08:45 a.m.	C123	0	2257	1365
4	04/12/2008	Jueves	08:45 a.m.	C123	1838	1856	1092
5	05/12/2008	Viernes	08:45 a.m.	C123	1087	1537	907
6	06/12/2008	Sábado	11:45 a.m.	C123	8219	4893	1708
7	07/12/2008	Domingo	08:45 a.m.	C123	7301	4316	1469
8	08/12/2008	Lunes	08:45 a.m.	C123	NULL	NULL	NULL
9	09/12/2008	Martes	08:45 a.m.	C123	4470	3582	1345
10	10/12/2008	Miércoles	08:45 a.m.	C123	3244	3024	1301
11	11/12/2008	Jueves	08:45 a.m.	C123	2463	2488	1295
12	12/12/2008	Viernes	08:45 a.m.	C123	1456	2059	1216
13	13/12/2008	Sábado	08:45 a.m.	C123	17259	10276	3587
14	14/12/2008	Domingo	08:45 a.m.	C123	9783	5783	1968

Tabla 21 : Técnica 2 Media Móvil 2 - Ejemplo

Fuente : Propia

Elaboración : Silvia Granados Flores

Para el caso anterior, los valores que reemplazarán los registros vacíos del 08/12/2008, se calcularán a través del promedio entre el stock del 06/12/2008 y 07/12/2008 en cada casetera respectivamente.

Se compara el valor promedio de stock correspondiente a los dos días anteriores contra los valores reales para el día anterior y posterior al analizado. En la Tabla 22 se evalúa el resultado de aplicar esta técnica, obteniendo como resultado que el promedio calculado no satisface la primera restricción; sin embargo es positivo con que el stock promedio es mayor o igual al stock del día siguiente.

EVALUACIÓN			
Media móvil 2	7760	4605	1589
¿Menor o igual al stock del 07.12.08?	NO	NO	NO
¿Mayor o igual al stock del 09.12.08?	SI	SI	SI

Tabla 22 : Técnica 2 Media Móvil 2 - Evaluación
Fuente : Propia
Elaboración : Silvia Granados Flores

En conclusión esta técnica también podría originar incoherencia en el registro de stock tal como la primera técnica.

Cabe recordar que en caso entre el primer corte y la fecha analizada se haya efectuada una recarga de billetes, se procederá a reemplazar el corte más antiguo por el monto recargado dentro del cálculo de la media móvil a 2.

C. Técnica 3: Media Móvil a 3

Se reemplazarán los valores vacíos por la media de los tres registros anteriores a éste.

En la Tabla 23 se presenta un ejemplo, el día lunes 08/12/2008 en el Reporte de Stock de Billetes correspondiente al turno de las 8:45a.m. se identifica un registro con valores vacíos.

REPORTE DE STOCK DE BILLETES
TURNO 8:45 a.m.

N°	FECHA	DIA	HORA	CAJERO	CASETERA	CASETERA	CASETERA
					NRO 1 S/.100	NRO 2 S/.20	NRO 3 US\$20
1	01/12/2008	Lunes	08:45 a.m.	C123	4755	3172	2145.
2	02/12/2008	Martes	09:15 a.m.	C123	3336	2673	1833
3	03/12/2008	Miércoles	08:45 a.m.	C123	0	2257	1365
4	04/12/2008	Jueves	08:45 a.m.	C123	1838	1856	1092
5	05/12/2008	Viernes	08:45 a.m.	C123	1087	1537	907
6	06/12/2008	Sábado	11:45 a.m.	C123	8219	4893	1708
7	07/12/2008	Domingo	08:45 a.m.	C123	7301	4316	1469
8	08/12/2008	Lunes	08:45 a.m.	C123	NULL	NULL	NULL
9	09/12/2008	Martes	08:45 a.m.	C123	4470	3582	1345
10	10/12/2008	Miércoles	08:45 a.m.	C123	3244	3024	1301
11	11/12/2008	Jueves	08:45 a.m.	C123	2463	2488	1295
12	12/12/2008	Viernes	08:45 a.m.	C123	1456	2059	1216
13	13/12/2008	Sábado	08:45 a.m.	C123	17259	10276	3587
14	14/12/2008	Domingo	08:45 a.m.	C123	9783	5783	1968

Tabla 23 : Técnica 3 Media Móvil 3 - Ejemplo
Fuente : Propia
Elaboración : Silvia Granados Flores

Para el caso anterior, los valores que reemplazarán los valores vacíos del 08/12/2008 se calcularán a través del promedio entre el stock del 05/12/2008, 06/12/2008 y 07/12/2008 en cada casetera respectivamente.

Se compara el valor promedio de stock correspondiente a los tres días anteriores contra los valores reales para el día anterior y posterior al analizado. En la Tabla 24 se evalúa el resultado de aplicar esta técnica, obteniendo como resultado que el promedio calculado satisface ambas restricciones.

EVALUACIÓN			
Media móvil 3	5535	3582	1362
¿Menor o igual al stock del 07.12.08?	SI	SI	SI
¿Mayor o igual al stock del 09.12.08?	SI	SI	SI

Tabla 24 : Técnica 3 Media Móvil 3 - Evaluación
Fuente : Propia
Elaboración : Silvia Granados Flores

En conclusión, esta técnica podría emplearse para reemplazar los valores vacíos de los datos recopilados. Sin embargo, cabe recordar que es posible identificarse casos en los cuales no se cumpla con alguna de las restricciones.

En caso que entre el primer corte y la fecha analizada se haya efectuada la recarga de billetes, se procederá a reemplazar el corte más antiguo por el monto recargado dentro del cálculo de la media móvil a 3.

D. Técnica 4: Interpolación Lineal

Los valores null serán reemplazados por los promedios calculados a través del último registro anterior y primer registro siguiente válidos.

En la Tabla 25 se presenta un ejemplo, el día lunes 08/12/2008 en el Reporte de Stock de Billetes correspondiente al turno de las 8:45a.m. se identifica un registro con valores vacíos.

En caso que entre el anterior y siguiente corte válidos se haya efectuada la recarga de billetes, se procederá a reemplazar el corte que corresponde por el monto recargado dentro del cálculo.

REPORTE DE STOCK DE BILLETES
TURNO 8:45 a.m.

N°	FECHA	DIA	HORA	CAJERO	CASETERA NRO 1 S/.100	CASETERA NRO 2 S/.20	CASETERA NRO 3 US\$20
1	01/12/2008	Lunes	08:45 a.m.	C123	4755	3172	2145
2	02/12/2008	Martes	09:15 a.m.	C123	3336	2673	1833
3	03/12/2008	Miércoles	08:45 a.m.	C123	0	2257	1365
4	04/12/2008	Jueves	08:45 a.m.	C123	1838	1856	1092
5	05/12/2008	Viernes	08:45 a.m.	C123	1087	1537	907
6	06/12/2008	Sábado	11:45 a.m.	C123	8219	4893	1708
7	07/12/2008	Domingo	08:45 a.m.	C123	7301	4316	1469
8	08/12/2008	Lunes	08:45 a.m.	C123	NULL	NULL	NULL
9	09/12/2008	Martes	08:45 a.m.	C123	4470	3582	1345
10	10/12/2008	Miércoles	08:45 a.m.	C123	3244	3024	1301
11	11/12/2008	Jueves	08:45 a.m.	C123	2463	2488	1295
12	12/12/2008	Viernes	08:45 a.m.	C123	1456	2059	1216
13	13/12/2008	Sábado	08:45 a.m.	C123	17259	10276	3587
14	14/12/2008	Domingo	08:45 a.m.	C123	9783	5783	1968

Tabla 25 : Técnica 4 Interpolación - Ejemplo
 Fuente : Propia
 Elaboración : Silvia Granados Flores

Se compara el valor promedio de stock correspondiente al día anterior y al día posterior contra los valores reales para el día anterior y posterior al analizado. En la Tabla 26 se evalúa el resultado de aplicar esta técnica; obteniendo como resultado que el promedio calculado satisface ambas restricciones.

EVALUACIÓN			
Interpolación	5885	3949	1407
¿Menor o igual al stock del 07.12.08?	SI	SI	SI
¿Mayor o igual al stock del 09.12.08?	SI	SI	SI

Tabla 26 : Técnica 4 Interpolación - Evaluación
 Fuente : Propia
 Elaboración : Silvia Granados Flores

Selección de técnica a emplear

El criterio para seleccionar la técnica de reemplazo de los valores vacíos o null será aquella que impacte en menor proporción las medidas de tendencia central (Media y Mediana) de los datos originales.

El cálculo del impacto en la media se realizará de la siguiente manera:

$$\% \text{ Impacto } _ \text{ Media} = \left| \frac{\text{Media } _ \text{ final} - \text{Media } _ \text{ original}}{\text{Media } _ \text{ original}} \right| \times 100\%$$

De la misma forma, el impacto de la mediana se calculará:

$$\% \text{ Impacto } _ \text{ Mediana} = \left| \frac{\text{Mediana } _ \text{ final} - \text{Mediana } _ \text{ original}}{\text{Mediana } _ \text{ original}} \right| \times 100\%$$

En la Tabla 27 se identifica el impacto en la media para cada una de las cuatro técnicas mencionadas.

El Impacto Total Media será la suma de los % de impacto calculados:

REEMPLAZO DE VALORES VACÍOS EN STOCK DE BILLETES
08/12/2008 8:45 a.m.

MEDIA ORIGINAL:	5016.22	3685.79	1633.25
MEDIANA ORIGINAL:	3335.80	3024.11	1365.00

N°	TÉCNICA	VALORES REEMPLAZADOS			IMPACTO EN LA MEDIA						IMPACTO TOTAL MEDIA
		CASETERA NRO 1 SI.100	CASETERA NRO 2 SI.20	CASETERA NRO 3 US\$20	CASETERA NRO 1 SI.100		CASETERA NRO 2 SI.20		CASETERA NRO 3 US\$20		
					MEDIA FINAL	%	MEDIA FINAL	%	MEDIA FINAL	%	
1	Promedio diario	2378	1586	1073	4828	3.76%	3536	4.07%	1593	2.45%	10.28%
2	Media Móvil a 2	7760	4605	1589	5212	3.90%	3751	1.77%	1630	0.20%	-5.87%
3	Media Móvil a 3	5535	3582	1362	5053	0.73%	3678	0.21%	1613	1.24%	2.18%
4	Interpolación	5885	3949	1407	5078	1.23%	3705	0.52%	1617	0.99%	2.75%

Tabla 27 : Reemplazo valores vacíos – Impacto Media

Fuente : Propia

Elaboración : Silvia Granados Flores

De la misma manera, en la Tabla 28 se calcula el impacto en la Mediana para cada una de las técnicas empleadas.

El Impacto Total Mediana será la suma de los % de impacto calculados:

REEMPLAZO DE VALORES VACÍOS EN STOCK DE BILLETES
08/12/2008 8:45 a.m.

MEDIA ORIGINAL:	5016.22	3685.79	1633.25
MEDIANA ORIGINAL:	3335.80	3024.11	1365.00

N°	TÉCNICA	VALORES REEMPLAZADOS			IMPACTO EN LA MEDIANA						IMPACTO TOTAL MEDIANA
		CASETERA NRO 1 SI.100	CASETERA NRO 2 SI.20	CASETERA NRO 3 US\$20	CASETERA NRO 1 SI.100		CASETERA NRO 2 SI.20		CASETERA NRO 3 US\$20		
					MEDIANA FINAL	%	MEDIANA FINAL	%	MEDIANA FINAL	%	
1	Promedio diario	2378	1586	1073	3290	1.37%	2848	5.82%	1355	0.73%	7.93%
2	Media Móvil a 2	7760	4605	1589	3903	17.00%	3098	2.44%	1417	3.81%	23.26%
3	Media Móvil a 3	5535	3582	1362	3903	17.00%	3098	2.44%	1363	0.15%	19.59%
4	Interpolación	5885	3949	1407	3903	17.00%	3098	2.44%	1386	1.54%	20.99%

Tabla 28 : Reemplazo valores vacíos – Impacto Mediana

Fuente : Propia

Elaboración : Silvia Granados Flores

Finalmente, en la Tabla 29 se determina el Impacto Total (suma del Impacto Total Media e Impacto Total Mediana) y además si los valores de reemplazo cumplen con las restricciones anteriormente señaladas.

N°	TÉCNICA	IMPACTO TOTAL	¿Menor o igual al stock del 07.12.08?			¿Mayor o igual al stock del 09.12.08?		
			CASETERA NRO 1 S/.100	CASETERA NRO 2 S/.20	CASETERA NRO 3 US\$20	CASETERA NRO 1 S/.100	CASETERA NRO 2 S/.20	CASETERA NRO 3 US\$20
			1	Promedio diario	18.21%	SI	SI	SI
2	Media Móvil a 2	29.13%	NO	NO	NO	SI	SI	SI
3	Media Móvil a 3	21.78%	SI	SI	SI	SI	SI	SI
4	Interpolación	23.73%	SI	SI	SI	SI	SI	SI

Tabla 29 : Reemplazo valores vacíos – Impacto Mediana

Fuente : Propia

Elaboración : Silvia Granados Flores

La técnica de Media Móvil 3 es la que genera el menor impacto y cumple con las restricciones de stock, por lo que es la mejor técnica a emplearse para el ejemplo.

Técnica empleada para los datos recolectados

Sobre los valores recolectados, en la Tabla 30 se presentan los resultados de aplicar las técnicas anteriormente detalladas para reemplazar los valores vacíos.

REEMPLAZO DE VALORES VACÍOS EN STOCK DE BILLETES

N	TECNICA	IMPACTO EN LA MEDIA				IMPACTO EN LA MEDIANA				IMPACTO TOTAL	¿MENOR O IGUAL AL STOCK DEL DIA ANTERIOR Y MAYOR O IGUAL AL STOCK DEL DIA SIGUIENTE?		
		CASETERA NRO 1 S/.100	CASETERA NRO 2 S/. 20	CASETERA NRO 3 US\$20	IMPACTO TOTAL MEDIA	CASETERA NRO 1 S/. 100	CASETERA NRO 2 S/. 20	CASETERA NRO 3 US\$20	IMPACTO TOTAL MEDIANA		CASETERA NRO 1 S/. 100	CASETERA NRO 2 S/. 20	CASETERA NRO 3 US\$20
CAJERO A													
1	Promedio diario	0.01%	0.01%	0.00%	0.02%	0.15%	0.09%	0.00%	0.24%	0.26%	NO	SI	NO
2	Media Móvil a 2	0.15%	0.19%	0.15%	0.49%	0.15%	0.18%	0.00%	0.33%	0.82%	NO	NO	NO
3	Media Móvil a 3	0.06%	0.02%	0.06%	0.14%	0.15%	0.09%	0.00%	0.24%	0.39%	NO	NO	NO
4	Interpolación	0.24%	0.15%	0.30%	0.69%	0.15%	0.09%	0.00%	0.24%	0.94%	SI	SI	SI
CAJERO B													
1	Promedio diario	0.00%	0.00%	0.01%	0.01%	0.00%	0.14%	0.38%	0.52%	0.53%	NO	NO	NO
2	Media Móvil a 2	0.66%	0.59%	0.89%	2.14%	0.58%	0.23%	0.95%	1.76%	3.90%	NO	NO	NO
3	Media Móvil a 3	0.40%	0.34%	0.72%	1.46%	0.56%	0.05%	0.95%	1.57%	3.04%	NO	NO	NO
4	Interpolación	0.57%	0.49%	0.71%	1.77%	0.56%	0.23%	0.95%	1.76%	3.53%	SI	SI	SI
CAJERO C													
1	Promedio diario	0.07%	0.06%	0.05%	0.17%	0.05%	0.03%	0.51%	0.59%	0.76%	NO	NO	NO
2	Media Móvil a 2	0.27%	0.21%	0.01%	0.48%	0.16%	1.01%	0.17%	1.34%	1.83%	NO	NO	NO
3	Media Móvil a 3	0.33%	0.28%	0.07%	0.67%	0.16%	1.01%	0.51%	1.69%	2.36%	NO	NO	NO
4	Interpolación	0.05%	0.10%	0.19%	0.34%	0.09%	0.67%	0.34%	1.10%	1.44%	SI	SI	SI
CAJERO D													
1	Promedio diario	0.02%	0.01%	0.01%	0.04%	0.02%	0.04%	0.00%	0.07%	0.10%	NO	NO	NO
2	Media Móvil a 2	0.20%	0.30%	0.27%	0.77%	0.18%	0.22%	0.00%	0.40%	1.17%	NO	NO	NO
3	Media Móvil a 3	0.27%	0.36%	0.30%	0.93%	0.18%	0.22%	0.00%	0.40%	1.33%	NO	NO	NO
4	Interpolación	0.07%	0.07%	0.14%	0.27%	0.02%	0.04%	0.00%	0.07%	0.34%	SI	SI	SI
CAJERO E													
1	Promedio diario	0.14%	0.14%	0.08%	0.36%	0.91%	0.24%	0.90%	2.05%	2.41%	NO	NO	NO
2	Media Móvil a 2	0.07%	0.12%	0.09%	0.28%	0.24%	0.10%	0.45%	0.78%	1.07%	NO	NO	NO
3	Media Móvil a 3	0.06%	0.09%	0.00%	0.15%	0.91%	0.10%	0.45%	1.46%	1.61%	NO	NO	NO
4	Interpolación	0.13%	0.31%	0.20%	0.63%	0.24%	0.10%	0.45%	0.78%	1.42%	SI	SI	SI

Tabla 30 : Reemplazo valores vacíos – Técnicas empleadas
Fuente : Empresa Financiera local
Elaboración : Silvia Granados Flores

Para los cinco cajeros automáticos analizados, la técnica de Interpolación permite reemplazar los valores vacíos acorde a las restricciones sobre límite de stock; sin embargo sólo para el Cajero C genera el menor impacto sobre la Media y Mediana comparado con las otras técnicas.

En conclusión, debido a que las cuatro técnicas detalladas impactan sobre la Media y Mediana en menos de 1%, se seleccionará la técnica que cumple con las restricciones sobre límite de stock. Debido a ello se empleará la técnica de Interpolación.

3.2.3. ANÁLISIS DE OUTLIERS

Los outliers o anomalías son observaciones atípicas en la muestra, las cuales impactan sobre el comportamiento promedio de los datos.

Al trabajar con datos atípicos en la demanda histórica de efectivo, el sistema de redes neuronales tratará de identificar algún patrón sobre este comportamiento anormal, lo que originaría empleo de recursos innecesariamente y principalmente, afectaría la efectividad de la generalización (D. Iranzo [24]).

Identificación de Outliers

En las Tablas 12 y 13 se identificaron las características en la distribución de frecuencias de la variable Demanda de Efectivo para los cinco cajeros automáticos, tanto a nivel de caseteras como monto total valorado.

Para identificar los outliers de los registros recopilados, inicialmente se identificó el intervalo de confianza formado por:

[Media – 3 x Desviación Estándar; Media + 3 x Desviación Estándar]

El rango de valores típicos para la demanda de efectivo se establece:

- Límite inferior = Media – 3 x Desviación Estándar

- Límite superior = Media + 3 x Desviación Estándar

A fin de ilustrar los conceptos anteriores, en la Tabla 31 se representa una muestra ejemplo de registros sobre demanda de efectivo:

PERIODO	FECHA INICIO	HORA INICIO	FECHA FIN	HORA FIN	CAJERO	CASETERA NRO 1-2 \$/ 100	CASETERA NRO 3 \$/ 20	DEMANDA CASETERA NRO 4 US\$20	MONTO TOTAL
01/01/2009	01/01/2009	08:45:05 a.m.	02/01/2009	08:45:02 a.m.	A	710		21	199
02/01/2009	02/01/2009	08:45:02 a.m.	03/01/2009	08:45:03 a.m.	A	682		338	185
03/01/2009	03/01/2009	08:45:03 a.m.	04/01/2009	08:45:03 a.m.	A	473		362	187
04/01/2009	04/01/2009	08:45:03 a.m.	05/01/2009	08:45:02 a.m.	A	558		304	150
05/01/2009	05/01/2009	08:45:02 a.m.	06/01/2009	08:45:05 a.m.	A	722		356	158
06/01/2009	06/01/2009	08:45:05 a.m.	07/01/2009	08:45:04 a.m.	A	684		370	194
07/01/2009	07/01/2009	08:45:04 a.m.	08/01/2009	08:45:06 a.m.	A	450		384	183
08/01/2009	08/01/2009	08:45:06 a.m.	09/01/2009	08:45:05 a.m.	A	608		781	191
09/01/2009	09/01/2009	08:45:05 a.m.	10/01/2009	08:45:05 a.m.	A	488		214	182
10/01/2009	10/01/2009	08:45:05 a.m.	11/01/2009	08:45:05 a.m.	A	646		315	185
11/01/2009	11/01/2009	08:45:05 a.m.	12/01/2009	08:45:05 a.m.	A	695		372	165
12/01/2009	12/01/2009	08:45:05 a.m.	13/01/2009	08:45:05 a.m.	A	670		299	181
13/01/2009	13/01/2009	08:45:05 a.m.	14/01/2009	08:45:03 a.m.	A	754		367	169
14/01/2009	14/01/2009	08:45:03 a.m.	15/01/2009	08:45:03 a.m.	A	906		375	178
15/01/2009	15/01/2009	08:45:03 a.m.	16/01/2009	08:45:04 a.m.	A	801		321	185
16/01/2009	16/01/2009	08:45:04 a.m.	17/01/2009	08:45:06 a.m.	A	670		239	178
17/01/2009	17/01/2009	08:45:06 a.m.	18/01/2009	08:45:04 a.m.	A	435		321	180
18/01/2009	18/01/2009	08:45:04 a.m.	19/01/2009	08:45:04 a.m.	A	471		358	180
19/01/2009	19/01/2009	08:45:04 a.m.	20/01/2009	08:45:05 a.m.	A	615		364	400
20/01/2009	20/01/2009	08:45:05 a.m.	21/01/2009	08:45:02 a.m.	A	604		451	187
21/01/2009	21/01/2009	08:45:02 a.m.	22/01/2009	08:45:06 a.m.	A	1658		338	173
22/01/2009	22/01/2009	08:45:06 a.m.	23/01/2009	08:45:04 a.m.	A	547		367	181
23/01/2009	23/01/2009	08:45:04 a.m.	24/01/2009	08:45:03 a.m.	A	401		369	178
24/01/2009	24/01/2009	08:45:03 a.m.	25/01/2009	08:45:04 a.m.	A	488		439	159
25/01/2009	25/01/2009	08:45:04 a.m.	26/01/2009	08:45:04 a.m.	A	560		384	169
26/01/2009	26/01/2009	08:45:04 a.m.	27/01/2009	08:45:04 a.m.	A	699		437	173
27/01/2009	27/01/2009	08:45:04 a.m.	28/01/2009	08:45:03 a.m.	A	1490		374	187
28/01/2009	28/01/2009	08:45:03 a.m.	29/01/2009	08:45:03 a.m.	A	771		385	13
29/01/2009	29/01/2009	08:45:03 a.m.	30/01/2009	08:45:03 a.m.	A	694		412	195
30/01/2009	30/01/2009	08:45:03 a.m.	31/01/2009	08:45:06 a.m.	A	685		368	174
31/01/2009	31/01/2009	08:45:06 a.m.	01/02/2009	08:45:03 a.m.	A	680		380	188

Tabla 31 : Identificación de outliers - Ejemplo

Fuente : Propia

Elaboración : Silvia Granados Flores

En la Tabla 32 se identifican las principales tendencias estadísticas (Media, desviación estándar, primer cuartil, tercer cuartil y desviación intercuartil) asimismo la relación simétrica de la distribución de frecuencias para las muestra de registros del ejemplo. Finalmente se calcula el límite inferior y superior:

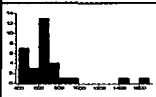
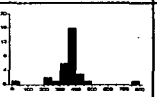
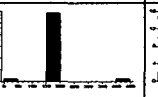
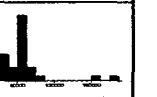
RANGO DE CONFIANZA				
MEDIA	687.58	359.45	180.81	86795.48
DESVIACIÓN ESTÁNDAR	285.10	110.58	51.61	28498.93
Q1	552.50	328.50	173.00	71890.00
Q3	704.50	377.50	187.00	86910.00
IQ	78.00	24.50	7.00	8510.00
FORMA DE LA CURVA				
	Asimétrico a la derecha	Relativamente simétrico	Relativamente simétrico	Asimétrico a la derecha
LÍMITE INFERIOR	-107.71	27.72	25.96	7298.70
LÍMITE SUPERIOR	1482.87	691.18	335.65	166292.27

Tabla 32 : Identificación de outliers – Rango de confianza

Fuente : Propia

Elaboración : Silvia Granados Flores

Sobre el ejemplo de la Tabla 31 se identifican los valores atípicos (outliers); es decir los registros que se encuentran fuera del rango de confianza establecido entre el límite inferior y superior. Se identificaron:

- Casetera S/.100: 2 outliers que exceden el límite superior.
- Casetera S/.20: 1 outlier que excede el límite inferior y 1 outlier que excede el límite superior.
- Casetera US\$20: 1 outlier que excede el límite inferior y 1 outlier que excede el límite superior.
- Monto Total: 2 outliers que exceden el límite superior.

Reemplazo de Outliers

El reemplazo de los outliers se realizará con el límite al cual exceden; es decir, si el valor atípico excede el límite inferior será reemplazado por éste, caso contrario con el límite superior correspondiente.

En base al ejemplo desarrollado anteriormente, en la Tabla 34 se presentan los valores resultantes con la técnica de reemplazo mencionada:

LIMITE INFERIOR	-107.71	27.72	25.96	7298.70
LIMITE SUPERIOR	1482.87	691.18	335.65	166292.27

PERIODO	FECHA INICIO	HORA INICIO	FECHA FIN	HORA FIN	CAJERO	DEMANDA			MONTO TOTAL
						CAJETERA NRO 1-2 SI 100	CAJETERA NRO 3 SI 20	CAJETERA NRO 4 USS 20	
01/01/2009	01/01/2009	08:45:05 a.m.	02/01/2009	08:45:02 a.m.	A	710	27.72	.599	83360
02/01/2009	02/01/2009	08:45:02 a.m.	03/01/2009	08:45:03 a.m.	A	682	338	185	86060
03/01/2009	03/01/2009	08:45:03 a.m.	04/01/2009	08:45:03 a.m.	A	473	362	187	65780
04/01/2009	04/01/2009	08:45:03 a.m.	05/01/2009	08:45:02 a.m.	A	558	304	160	70580
05/01/2009	05/01/2009	08:45:02 a.m.	06/01/2009	08:45:05 a.m.	A	722	356	158	88900
06/01/2009	06/01/2009	08:45:05 a.m.	07/01/2009	08:45:04 a.m.	A	684	370	194	87440
07/01/2009	07/01/2009	08:45:04 a.m.	08/01/2009	08:45:06 a.m.	A	450	364	183	63260
08/01/2009	08/01/2009	08:45:06 a.m.	09/01/2009	08:45:05 a.m.	A	608	691.18	191	87880
09/01/2009	09/01/2009	08:45:05 a.m.	10/01/2009	08:45:05 a.m.	A	488	214	182	64000
10/01/2009	10/01/2009	08:45:05 a.m.	11/01/2009	08:45:05 a.m.	A	646	315	185	82000
11/01/2009	11/01/2009	08:45:05 a.m.	12/01/2009	08:45:05 a.m.	A	695	372	165	86840
12/01/2009	12/01/2009	08:45:05 a.m.	13/01/2009	08:45:05 a.m.	A	670	299	181	83840
13/01/2009	13/01/2009	08:45:05 a.m.	14/01/2009	08:45:03 a.m.	A	754	367	169	92880
14/01/2009	14/01/2009	08:45:03 a.m.	15/01/2009	08:45:03 a.m.	A	906	375	178	108780
15/01/2009	15/01/2009	08:45:03 a.m.	16/01/2009	08:45:04 a.m.	A	801	321	185	97620
16/01/2009	16/01/2009	08:45:04 a.m.	17/01/2009	08:45:06 a.m.	A	670	238	176	82340
17/01/2009	17/01/2009	08:45:06 a.m.	18/01/2009	08:45:04 a.m.	A	435	321	180	60720
18/01/2009	18/01/2009	08:45:04 a.m.	19/01/2009	08:45:04 a.m.	A	471	358	180	65060
19/01/2009	19/01/2009	08:45:04 a.m.	20/01/2009	08:45:05 a.m.	A	615	364	335.66	92780
20/01/2009	20/01/2009	08:45:05 a.m.	21/01/2009	08:45:02 a.m.	A	604	451	187	80640
21/01/2009	21/01/2009	08:45:02 a.m.	22/01/2009	08:45:06 a.m.	A	1482.87	336	173	196292.27
22/01/2009	22/01/2009	08:45:06 a.m.	23/01/2009	08:45:04 a.m.	A	547	367	181	72900
23/01/2009	23/01/2009	08:45:04 a.m.	24/01/2009	08:45:03 a.m.	A	401	369	178	58160
24/01/2009	24/01/2009	08:45:03 a.m.	25/01/2009	08:45:04 a.m.	A	468	438	159	67120
25/01/2009	25/01/2009	08:45:04 a.m.	26/01/2009	08:45:04 a.m.	A	560	384	169	73820
26/01/2009	26/01/2009	08:45:04 a.m.	27/01/2009	08:45:04 a.m.	A	699	437	173	89020
27/01/2009	27/01/2009	08:45:04 a.m.	28/01/2009	08:45:03 a.m.	A	1482.87	374	187	196292.27
28/01/2009	28/01/2009	08:45:03 a.m.	29/01/2009	08:45:03 a.m.	A	771	365	26.94	85580
29/01/2009	29/01/2009	08:45:03 a.m.	30/01/2009	08:45:03 a.m.	A	694	412	195	89340
30/01/2009	30/01/2009	08:45:03 a.m.	31/01/2009	08:45:06 a.m.	A	685	368	174	85300
31/01/2009	31/01/2009	08:45:06 a.m.	01/02/2009	08:45:03 a.m.	A	680	380	188	86880

Tabla 33 : Reemplazo de outliers

Fuente : Propia

Elaboración : Silvia Granados Flores

El objetivo de reemplazar los valores outliers es aminorar el impacto de anomalías en la información, ya que estos fuerzan a la red neuronal generalizar sobre comportamientos atípicos. Sin embargo, es importante que el reemplazo efectuado modifique en menor proporción posible el comportamiento original de los datos; es decir, que las tendencias de crecimiento y decrecimiento se mantengan.

En el Gráfico 34 se presenta el comportamiento de los datos originales versus el comportamiento de los datos resultantes producto del reemplazo en el ejemplo. De esta manera, se puede observar que las tendencias se mantienen en todas las variables analizadas para el ejemplo anterior.

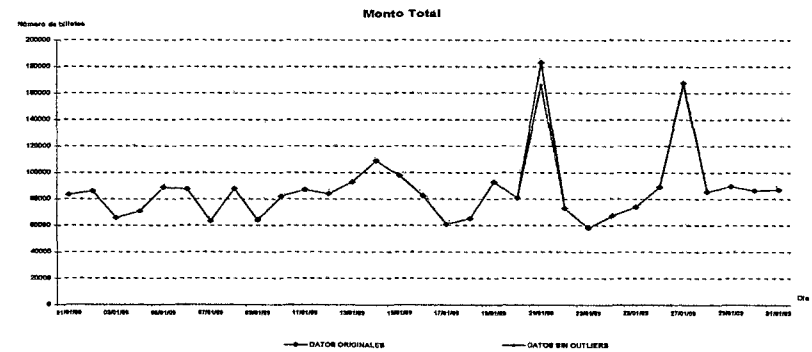
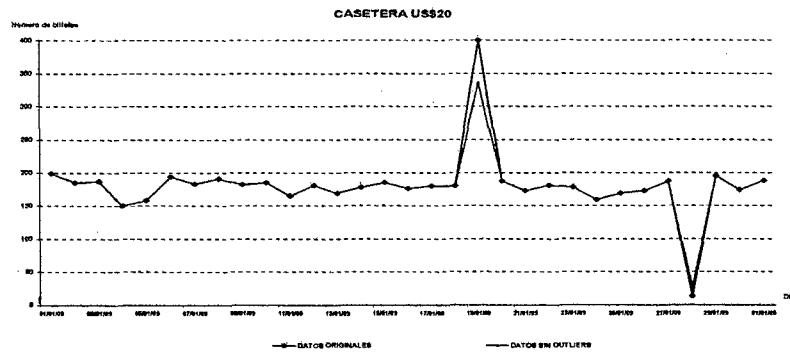
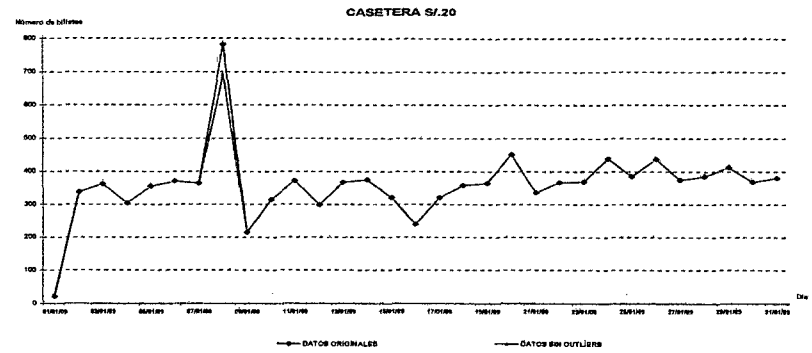
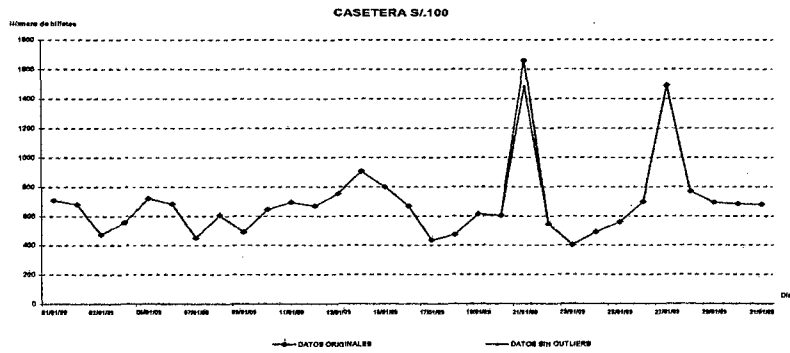


Gráfico 59 : Reemplazo de outliers
 Fuente : Propia
 Elaboración : Silvia Granados Flores

En base a la técnica detallada se identificaron los valores outliers de los registros correspondientes a la demanda de efectivo real calculada para los cinco cajeros automáticos. Estos valores así como su participación del total de registros y el impacto en las medidas de tendencia central por empleo de la técnica mencionada, se detallan en la Tabla 33:

VALORES OUTLIERS Desde 01/12/08 al 30/11/09					
	CAJERO AUTOMATICO				
	A	B	C	D	E
TOTAL OBSERVACIONES:	365.00	365.00	365.00	365.00	365.00
TOTAL OUTLIERS:	4	4	5	9	4
% VALORES OUTLIERS:	1.10%	1.10%	1.37%	2.47%	1.10%
MEDIA ORIGINAL	45,991.73	37,601.86	40,180.05	35,823.23	33,368.88
MEDIA SIN VALORES OUTLIERS:	46,069.00	38,436.16	41,095.36	36,100.64	33,628.79
% DESVIACIÓN MEDIA	0.17%	2.22%	2.28%	0.77%	0.78%
MEDIANA ORIGINAL	44,540.00	35,780.00	36,200.00	29,380.00	29,137.64
MEDIANA SIN VALORES OUTLIERS	45,280.00	36,140.00	36,420.00	30,040.00	29,140.00
% DESVIACIÓN MEDIANA	1.66%	1.01%	0.61%	2.25%	0.01%
% DESVIACIÓN TOTAL	1.83%	3.22%	2.89%	3.02%	0.79%

Tabla 34 : Tratamiento de outliers en datos reales

Fuente : Propia

Elaboración : Silvia Granados Flores

Como se puede observar en la tabla anterior, el empleo de esta técnica genera una desviación entre el 0.01% y 5.65% sobre las medidas originales de tendencia central de la información.

3.3. RESUMEN DEL CAPÍTULO

En este capítulo se detallaron las metodologías a seguir a fin de limpiar los datos de la muestra, específicamente con el tratamiento de los valores imprecisos, valores vacíos y outliers o valores anómalos. Asimismo se detallan los criterios a emplear a fin de identificar la técnica que mejor se adecua a los datos recolectados.

El resumen del capítulo se representa en el Gráfico 35:

$$D_I^X = \{P_I, Fi_I, Hi_I, Ff_I, Hf_I, C_X, D_I^{X,S/100}, D_I^{X,S/20}, D_I^{X,US\$20}\}$$

*Técnicas tratamiento de
nulls, outliers e imprecisos*

Limpieza de Datos

$$D_I^X = \{T_I, C_X, D_I^{X,S/100}, D_I^{X,S/20}, D_I^{X,US\$20}\}$$

Gráfico 60 : Limpieza de Datos - Resumen
Fuente : Propia
Elaboración : Silvia Granados Flores

CAPÍTULO IV

MODELO DE SOLUCIÓN

4.1. DESCRIPCIÓN GENERAL

En este capítulo se detallará el planteamiento de solución para el problema sobre pronóstico de demanda de efectivo. De esta manera se detallarán los pasos a realizar con la finalidad de construir la Red Neuronal Artificial que pronostique la demanda de efectivo.

4.2. PLANTEAMIENTO DE LA SOLUCIÓN

Hasta el momento, se han desarrollado actividades a fin de conseguir datos precisos y limpios para ser procesados por la Red Neuronal Artificial, los cuales son:

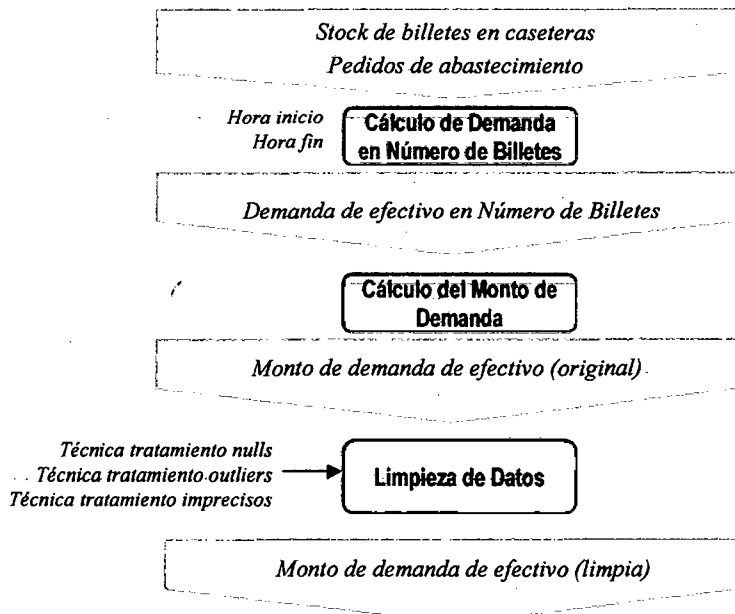


Gráfico 61 : Análisis y limpieza de Datos

Fuente : Propia
Elaboración : Silvia Granados Flores

El proceso de solución para el problema sobre pronóstico de demanda de efectivo a partir de datos históricos, tiene los siguientes elementos de entrada, parámetros y elementos de salida:

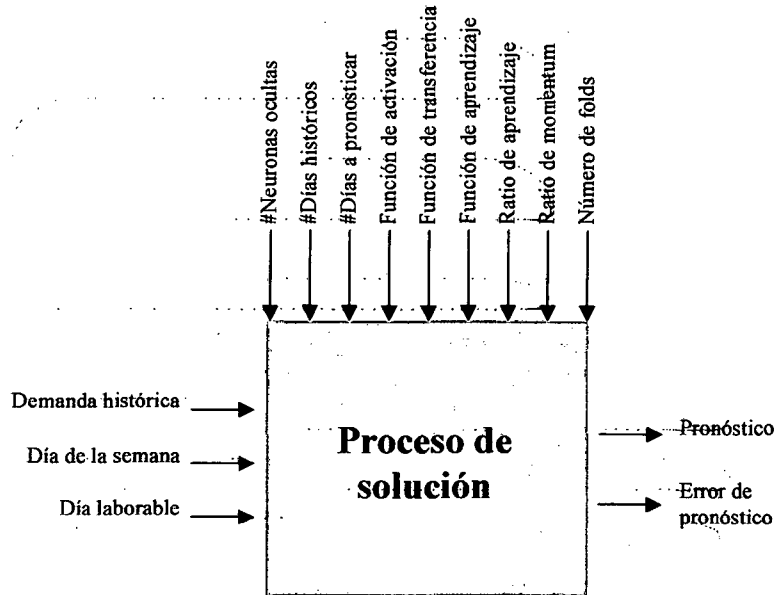


Gráfico 62 : Proceso de solución
Fuente : Propia
Elaboración : Silvia Granados Flores

El proceso de solución a describirse en los siguientes capítulos tiene como elementos de entrada:

- Demanda histórica de efectivo, registros sobre demanda de efectivo en el periodo analizado como histórico, sobre los cuales se ha aplicado técnicas para reemplazar valores imprecisos, vacíos y outliers.
- Día de la semana, número de día dentro de la semana que representa la fecha observada.
- Día laborable, números para identificar si la fecha observada corresponde a un día laborable o no.

En el Módulo Preparación de Datos se construirá la matriz base para diseñar las neuronas de entrada y salida del sistema.

Por otro lado, en el desarrollo del proceso de solución existen parámetros que condicionan la performance del modelo, los cuales se identifican:

- **Número de neuronas ocultas:** Número de neuronas que se establecerán en la capa oculta, aquellas que no tienen contacto con el exterior y sus elementos pueden tener diferentes conexiones.
- **Número de días históricos:** Número de observaciones referidas a fechas anteriores al momento analizado, sobre los que la red neuronal generalizará el comportamiento de la demanda.
- **Número de días pronóstico:** Número de observaciones futuras al momento analizado y sobre los que se efectuará el pronóstico.
- **Función de activación:** Función con resultado binario y sobre cuyo resultado la neurona se activará o no.
- **Función de transferencia:** Función que toma el resultado de la función de activación y lo acota según la interpretación que se quiera dar a la salida.
- **Función de aprendizaje:** Función mediante la cual se modifican los pesos de las conexiones (pesos sinápticos), los cuales son los responsables de almacenar el conocimiento.
- **Ratio de aprendizaje:** Velocidad con que se produce la convergencia de la red hacia el error mínimo.
- **Ratio de momentum:** Acelera el descenso de la velocidad de variación en los pesos sinápticos.
- **Número de folds:** Indica en cuántas partes se dividirá el conjunto de datos a fin de emplear un grupo para la validación y el resto para entrenamiento de la Red.

Finalmente se realizará la experimentación, simulación del sistema diseñado en Redes Neuronales Artificiales, con lo que se establecerán los valores óptimos para los parámetros y que generen el menor margen de error comparado con la demanda real.

Como salida del proceso se obtendrá:

- Error de pronóstico: Es el margen de desviación entre la demanda real y la pronosticada por el sistema de Redes Neuronales Artificiales. Este error se obtendrá a través de la técnica seleccionada para la evaluación de modelo.

4.3. RESUMEN DEL CAPÍTULO

En este capítulo se detalló el proceso de solución diseñado, describiendo las entradas, parámetros y salidas, a fin de construir el sistema de Redes Neuronales Artificiales el cual resolverá el problema de pronóstico de demanda de efectivo.

Este detalle puede visualizarse a través del Gráfico 38:

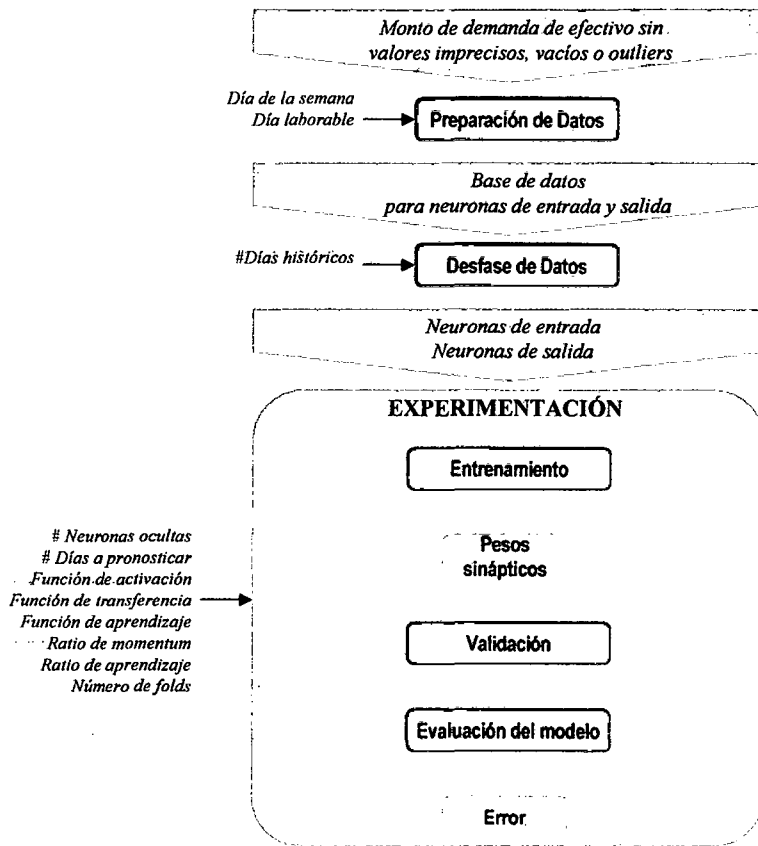


Gráfico 63 : Planteamiento de la Solución

Fuente : Propia

Elaboración : Silvia Granados Flores

CAPÍTULO V

PREPARACIÓN DE DATOS

5.1. DESCRIPCIÓN GENERAL

En este capítulo se detallarán las actividades y consideraciones a tomarse en cuenta para armar la base de datos que servirá de fuente en la construcción de las neuronas de entrada y salida.

A este proceso ingresará la demanda histórica de efectivo luego de haberse aplicado las técnicas para el reemplazo de valores imprecisos, vacíos y outliers. Asimismo los parámetros de entrada (día de la semana y día laborable) ingresarán a la base de datos para determinar las neuronas de entrada de la red y la demanda histórica, las neuronas de salida. El Gráfico 39 ilustra el procedimiento a seguir:

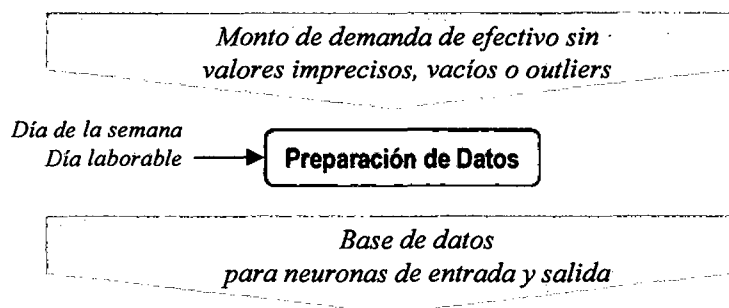


Gráfico 64 : Preparación de datos
Fuente : Propia
Elaboración : Silvia Granados Flores

5.2. PREPARACIÓN DE DATOS

Como entrada de este proceso se tienen los datos de monto de demanda histórica de efectivo, bajo la siguiente estructura:

Monto de demanda de efectivo sin valores imprecisos, vacíos o outliers

$$MD_I^X = \{P_I, Fi_I, Hi_I, Ff_I, Hf_I, C_X, MD_I^{X,S/100}, MD_I^{X,S/20}, MD_I^{X,US\$20}\}$$

Para observar el detalle de esta estructura de datos, revisar el punto 2.2.4 Estructura de Datos (ver Capítulo II).

A continuación se detallan los parámetros que se incluirán dentro de la base de datos que será la fuente de las neuronas de entrada y salida:

A. Día de la semana

Número ordinal que hace referencia al día de la semana correspondiente al periodo sobre el cual se mide la demanda de efectivo. Las equivalencias por día de la semana se determinan en la Tabla 35:

Día	Equivalente
Lunes	1
Martes	2
Miércoles	3
Jueves	4
Viernes	5
Sábado	6
Domingo	7

Tabla 35 : Equivalencias para día de la semana

Fuente : Propia

Elaboración : Silvia Granados Flores

En la Tabla 36 se refleja la mecánica a seguir para incluir el parámetro de Día de la semana dentro de la base de datos:

CAJERO AUTOMÁTICO	PERIODO	Día de la semana
A	01/01/2009	4
A	02/01/2009	5
A	03/01/2009	6
A	04/01/2009	7
A	05/01/2009	1
A	06/01/2009	2
A	07/01/2009	3
A	08/01/2009	4
A	09/01/2009	5
A	10/01/2009	6
A	11/01/2009	7
A	12/01/2009	1
A	13/01/2009	2
A	14/01/2009	3
A	15/01/2009	4
A	16/01/2009	5
A	17/01/2009	6
A	18/01/2009	7

Tabla 36 : Día de la semana - Ejemplo

Fuente : Propia

Elaboración : Silvia Granados Flores

B. Día laborable

Corresponde a categorizar el periodo del registro de demanda de efectivo sobre las características de día laborable y no laborable. En la Tabla 37 se definen los criterios para identificar la categoría a asignar:

CATEGORÍA	VALOR
Día no laborable	1
Día laborable	2

Tabla 37 : Día laborable

Fuente : Propia

Elaboración : Silvia Granados Flores

En la Tabla 38 se describen las equivalencias por día de semana:

CATEGORÍA	VALOR
Día no laborable	Domingo Feriados
Día laborable	Lunes Martes Miércoles Jueves Viernes Sábado

Tabla 38 : Equivalencias para Día laborable

Fuente : Propia

Elaboración : Silvia Granados Flores

Este parámetro se asigna con el fin de identificar patrones de comportamiento sobre la demanda de efectivo originados por la característica (laborable/no laborable) del periodo analizado.

Para el ejemplo de la Tabla 36 se incluirá la característica Día laborable, obteniendo la Tabla 39:

CAJERO AUTOMÁTICO	PERIODO	Día de la semana	Día laborable
A	01/01/2009	4	1
A	02/01/2009	5	2
A	03/01/2009	6	2
A	04/01/2009	7	1
A	05/01/2009	1	2
A	06/01/2009	2	2
A	07/01/2009	3	2
A	08/01/2009	4	2
A	09/01/2009	5	2
A	10/01/2009	6	2
A	11/01/2009	7	1
A	12/01/2009	1	2
A	13/01/2009	2	2
A	14/01/2009	3	2
A	15/01/2009	4	2
A	16/01/2009	5	2
A	17/01/2009	6	2
A	18/01/2009	7	1

Tabla 39 : Día laborable - Ejemplo

Fuente : Propia

Elaboración : Silvia Granados Flores

C. Monto de demanda de efectivo

Luego de incluirse los parámetros Día de la semana y Día laborable, se incluirá el Monto de demanda de efectivo del periodo al cual corresponde.

Para el ejemplo de la Tabla 39, en la Tabla 40 se presenta la base de datos donde se incluye el monto de demanda de efectivo.

CAJERO AUTOMÁTICO	PERIODO	Día de la semana	Día laborable	Monto de demanda
A	01/01/2009	4	1	138600
A	02/01/2009	5	2	94260
A	03/01/2009	6	2	67120
A	04/01/2009	7	1	75200
A	05/01/2009	1	2	107160
A	06/01/2009	2	2	95620
A	07/01/2009	3	2	44220
A	08/01/2009	4	2	74600
A	09/01/2009	5	2	64060
A	10/01/2009	6	2	78720
A	11/01/2009	7	1	77920
A	12/01/2009	1	2	72280
A	13/01/2009	2	2	115780
A	14/01/2009	3	2	106620
A	15/01/2009	4	2	116680
A	16/01/2009	5	2	91060
A	17/01/2009	6	2	51680
A	18/01/2009	7	1	63680

Tabla 40 : Monto de demanda de efectivo - Ejemplo
 Fuente : Propia
 Elaboración : Silvia Granados Flores

Luego de incluir los factores anteriormente detallados, los registros que compondrán la base de datos tienen la siguiente estructura:

Base de datos para neuronas de entrada y salida

$$R_I^X = \{C_x, P_I, ds'_I, dl''_I, MD_I^X\}$$

Donde:

R_I^X : Registro de base de datos correspondiente al cajero automático "X" y a un periodo de tiempo "I".

C_X : Cajero automático "X"

P_I : Periodo "I" en el que se mide el monto de demanda de efectivo.

ds'_I : Día de la semana del periodo de tiempo "I".

dl'_I : Día laborable del periodo "I".

MD_I^X : Monto de demanda de efectivo en el cajero automático "X" del periodo "I".

5.3. NORMALIZACIÓN DE DATOS

Este procedimiento realizará la transformación lineal de datos con el fin de mejorar su interpretación.

Esta transformación no cambia la distribución y por tanto no genera cambios en las variables. El objetivo es acortar la distancia entre valores de un mismo factor, a fin de reducir el uso de recursos por la red neuronal para generalizar el comportamiento de los datos.

Para realizar esta transformación, se aplicará la función logaritmo neperiano a cada uno de los valores incluidos en la base de datos.

La estructura de datos resultante de este procedimiento se describe así:

Base de datos transformada para neuronas de entrada y salida.

$$R_I^X = \{C_X, P_I, ds'_I, dl'_I, MD_I^X\}$$

Donde:

R_I^X : Registro de base de datos transformado y correspondiente al cajero automático "X" y a un periodo de tiempo "I".

C_X : Cajero automático "X".

P_I : Periodo de tiempo "I".

ds_I : Día de la semana del periodo "I" transformado bajo la siguiente función:

$$ds_I = \ln(ds'_I)$$

dl_I : Día laborable del periodo "I" transformado bajo la siguiente función:

$$dl_I = \ln(dl'_I)$$

MD_I^X : Monto de demanda de efectivo en el cajero automático "X" del periodo "I" transformado bajo la siguiente función:

$$MD_I^X = \ln(MD_I'^X)$$

Para el ejemplo de la Tabla 40, los datos transformados se observan en la Tabla 41:

CAJERO AUTOMÁTICO	PERIODO	Día de la semana	Día laborable	Monto de demanda
A	01/01/2009	1.386	0.000	11.839
A	02/01/2009	1.609	0.693	11.454
A	03/01/2009	1.792	0.693	11.114
A	04/01/2009	1.946	0.000	11.228
A	05/01/2009	0.000	0.693	11.582
A	06/01/2009	0.693	0.693	11.468
A	07/01/2009	1.099	0.693	10.697
A	08/01/2009	1.386	0.693	11.220
A	09/01/2009	1.609	0.693	11.068
A	10/01/2009	1.792	0.693	11.274
A	11/01/2009	1.946	0.000	11.263
A	12/01/2009	0.000	0.693	11.188
A	13/01/2009	0.693	0.693	11.659
A	14/01/2009	1.099	0.693	11.577
A	15/01/2009	1.386	0.693	11.667
A	16/01/2009	1.609	0.693	11.419
A	17/01/2009	1.792	0.693	10.853
A	18/01/2009	1.946	0.000	11.062

Tabla 41 : Transformación de datos - Ejemplo
Fuente : Propia
Elaboración : Silvia Granados Flores

5.4. RESUMEN DEL CAPÍTULO

En este capítulo se detallaron los procedimientos para construir la base de datos sobre la cual se extraerán las neuronas de entrada y salida. Además se describió la transformación de los datos a fin de mejorar su interpretación para el módulo de Experimentación (ver Capítulo VIII).

El resumen de este capítulo se representa en el Gráfico 40:

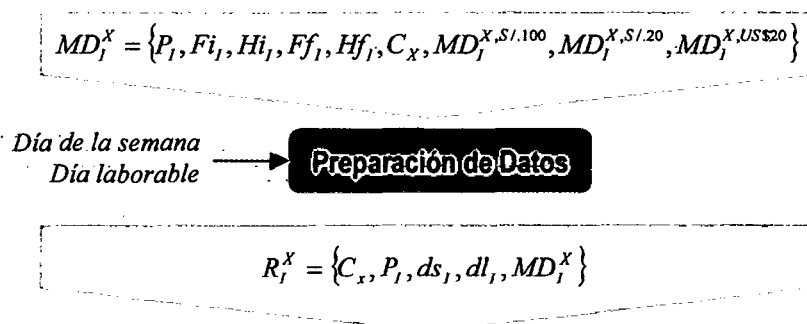


Gráfico 65 : Preparación de datos - Resumen
Fuente : Propia
Elaboración : Silvia Granados Flores

CAPÍTULO VI

DEFASE DE DATOS

6.1. DESCRIPCIÓN GENERAL

En este capítulo se detallará el procedimiento para construir las neuronas de entrada y salida a partir de la base de datos generada a lo largo de los capítulos anteriores.

Inicialmente se describirá la arquitectura del sistema de Redes Neuronales Artificiales para luego detallar el procedimiento con el que se obtendrán las neuronas de entrada y salida del sistema.

En el Gráfico 41 se representa el diagrama de proceso correspondiente a este capítulo:

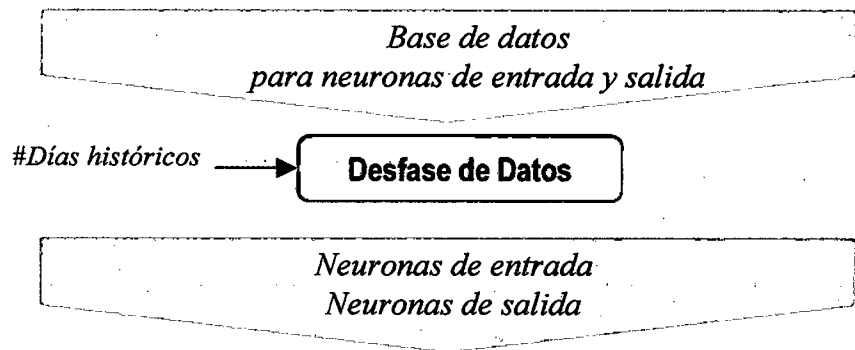


Gráfico 66 : Desfase de datos
Fuente : Propia
Elaboración : Silvia Granados Flores

6.2. ARQUITECTURA DEL MODELO DE RNA

Para identificar la arquitectura del sistema de Redes Neuronales Artificiales se debe tener en cuenta inicialmente las siguientes actividades:

- Seleccionar el número de capas del sistema
- Seleccionar el número de neuronas de entrada
- Seleccionar el número de neuronas de salida

Cabe recordar que en caso se seleccionara un número insuficiente de neuronas en cada capa, el sistema tendrá menor capacidad para generalizar el comportamiento de todos los puntos observados en la serie de tiempo. Por el contrario, si este número fuera determinado en exceso ocasionaría un sobreajuste en el tratamiento y generaría una menor performance en el pronóstico (I. Leal et. al. [25]).

Asimismo la experimentación del modelo (ver Capítulo VIII) permitirá identificar el número óptimo de neuronas de entrada, ocultas y salida.

6.2.1. NÚMERO DE CAPAS

El sistema de Redes Neuronales Artificiales propuesto se basa en el modelo autorregresivo no lineal de series paralelas con entradas exógenas (conocido como NARX, ver Capítulo I).

Este modelo de redes neuronales se caracteriza por ser una red dinámica y recurrente, estableciendo conexiones de retroalimentación dentro de las diferentes capas que constituyen su arquitectura. La ecuación definida para los valores obtenidos a través de la neurona de salida de un modelo NARX se determina así (V. Cassino et. al. [16]):

$$y(t) = f(y(t-1), y(t-2), \dots, y(t-n_y), u(t-1), u(t-2), \dots, u(t-n_u))$$

Donde:

$y(t)$: Valor pronosticado en momento "t"

$y(t-1)$: Valor real de salida en momento "t-1"

$y(t-n_y)$: Valor real de salida en momento "t-ny". Donde ny hace referencia al tiempo de retraso a emplear para el pronóstico.

$u(t-1)$: Valor de entrada en momento "t-1"

$u(t-n_u)$: Valor de entrada en momento "t-nu". Donde nu hace referencia al tiempo de retraso a emplear para el pronóstico.

Adicionalmente, durante el entrenamiento de una red NARX de series paralelas es posible usar el valor real en un momento "t" para la retroalimentación del valor pronosticado de ese mismo momento. La representación del modelo se muestra en el Gráfico 42 (F. Segovia [17]):

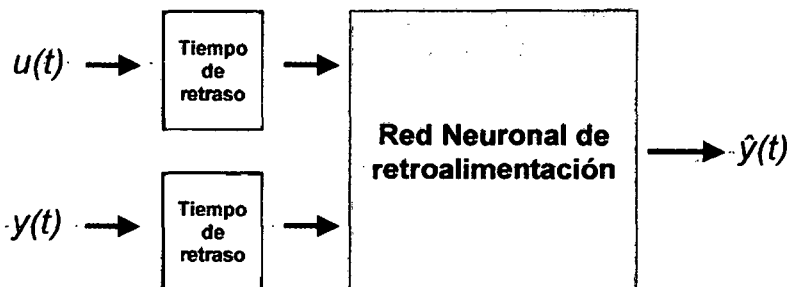


Gráfico 67 : NARX – Red retroalimentación

Fuente : Neural Network Toolbox – Matlab 7.8.0

Elaboración : C. Sorensen, R.B. Lehoucq

Donde:

$y(t)$: Valor real de salida en momento "t"

$u(t)$: Valor de entrada en momento "t"

$\hat{y}(t)$: Valor pronosticado para el momento "t"

De esta manera se propone un modelo de tres capas en el Gráfico 43:

- Capa de neuronas de entrada
- Capa de neuronas ocultas
- Capa de neuronas de salida

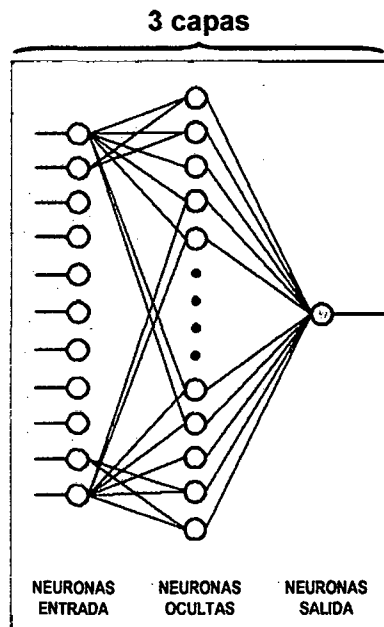


Gráfico 68 : Arquitectura de la RNA

Fuente : Propia

Elaboración : Silvia Granados Flores

6.2.2. NEURONAS DE ENTRADA

La capa de neuronas de entrada considerará los siguientes parámetros:

- Día de la semana
- Día laborable
- Monto de demanda de efectivo histórico (considerando tiempo de retraso)

Estructura de Datos

La estructura de datos que ingresa a este procedimiento está conformado por:

Base de datos transformada para neuronas de entrada y salida

$$R_I^X = \{C_x, P_I, ds_I, dl_I, MD_I^X\}$$

Para construir las neuronas de entrada se extraerá de la estructura anterior las variables correspondientes a:

- Día de la semana del periodo "I" (ds_I)
- Día laborable del periodo "I" (dl_I)
- Monto de demanda de efectivo en el cajero automático "X" del periodo "I" (MD_I^X)

De esta manera la estructura de una neurona de entrada está compuesta por:

Neurona de entrada

$$E_I^X = \{ds_I, dl_I, MD_I^X\}$$

Donde:

E_I^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "I".

Las neuronas de entrada construidas conformarán la Matriz de Neuronas de Entrada:

Matriz de Neuronas de Entrada

$$ME^X = \{E_I^X, E_{I+1}^X, \dots, E_{I+n-1}^X, E_{I+n}^X\}$$

Donde:

ME^X : Matriz de Neuronas de Entrada correspondiente al cajero automático "X".

E_I^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "I".

E_{I+1}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "I+1".

E_{I+n-1}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "I+n-1". Siendo "n" el número total de observaciones para la serie de tiempo.

E_{I+n}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "I+n". Siendo "n" el número total de observaciones para la serie de tiempo.

6.2.3. NEURONAS DE SALIDA

La capa de neuronas de salida considerará el siguiente parámetro:

- Monto de demanda de efectivo pronosticado

Estructura de Datos

Al igual que en la capa de neuronas de entrada, los registros que ingresan a este procedimiento siguen la siguiente estructura:

Base de datos transformada para neuronas de entrada y salida

$$R_I^X = \{C_x, P_I, ds_I, dl_I, MD_I^X\}$$

Para construir las neuronas de salida se extraerá de la estructura anterior la variable:

- Monto de demanda de efectivo en el cajero automático "X" del periodo "I" (MD_I^X)

De esta manera la estructura de una neurona de salida está compuesta por:

Neurona de salida

$$S_I^X = \{MD_I^X\}$$

Donde:

S_I^X : Neurona de salida correspondiente al cajero automático "X" en el periodo "I".

El conjunto de neuronas de salida conformará la Matriz de Neuronas de Salida y se representará de la siguiente forma:

Matriz de Neuronas de Salida

$$MS^X = \{S_I^X, S_{I+1}^X, \dots, S_{I+n-1}^X, S_{I+n}^X\}$$

Donde:

MS^X : Matriz de Neuronas de Salida correspondiente al cajero automático "X".

S_I^X : Neurona de salida correspondiente al cajero automático "X" en el periodo "I".

S_{I+1}^X : Neurona de salida correspondiente al cajero automático "X" en el periodo "I+1".

S_{I+n-1}^X : Neurona de salida correspondiente al cajero automático "X" en el periodo "I+n-1". Siendo "n" el número total de observaciones para la serie de tiempo.

S_{I+n}^X : Neurona de salida correspondiente al cajero automático "X" en el periodo "I+n". Siendo "n" el número total de observaciones para la serie de tiempo.

6.3. CORRESPONDENCIA ENTRE NEURONAS DE ENTRADA Y SALIDA

En base al procedimiento indicado en el punto anterior, se determinan las estructuras a emplear dentro del sistema de redes neuronales.

En el Gráfico 44 se representa la relación entre las neuronas de entrada y salida en base al periodo "I" y los siguientes periodos:

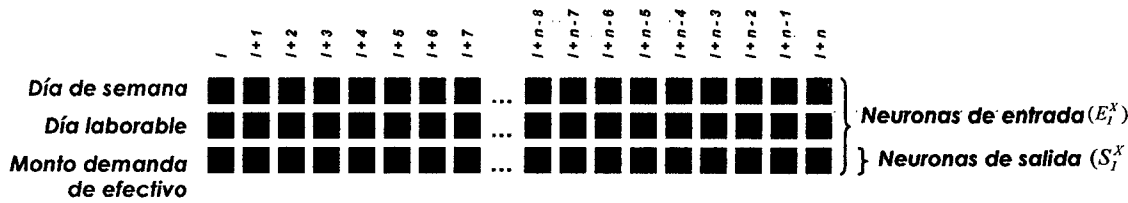


Gráfico 69 : Correspondencia entre neuronas de entrada y salida

Fuente : Propia

Elaboración : Silvia Granados Flores

Las redes neuronales NARX requieren del parámetro "Tiempo de retraso" para determinar la dependencia de las neuronas de salida sobre las neuronas de entrada.

Para el problema desarrollado se considera como "Tiempo de retraso" a los días históricos (número de periodos anteriores a partir del cual se está efectuando el pronóstico).

De esta manera, en el Gráfico 45 se representa el parámetro días históricos como "dh":

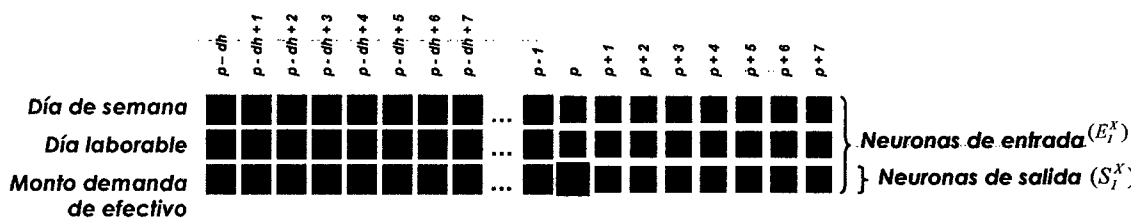


Gráfico 70 : Días históricos

Fuente : Propia

Elaboración : Silvia Granados Flores

Funcionalmente el Gráfico 45 puede ser representado:

$$S_p^X = f(E_{p-dh}^X, E_{p-dh+1}^X, E_{p-dh+2}^X, E_{p-dh+3}^X, \dots, E_{p-1}^X)$$

Donde:

S_p^X : Neurona de salida correspondiente al cajero automático "X" en el periodo "p"

E_{p-dh}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p-dh".

E_{p-dh+1}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p-dh+1".

E_{p-dh+2}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p-dh+2".

E_{p-dh+3}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p-dh+3".

E_{p-1}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p-1".

De la misma manera, el número de días a pronosticar determinan el alcance sobre el cual se evaluará la efectividad del modelo.

El Gráfico 46 ilustra el sistema de pronóstico siendo el número de días pronosticar igual a 3:

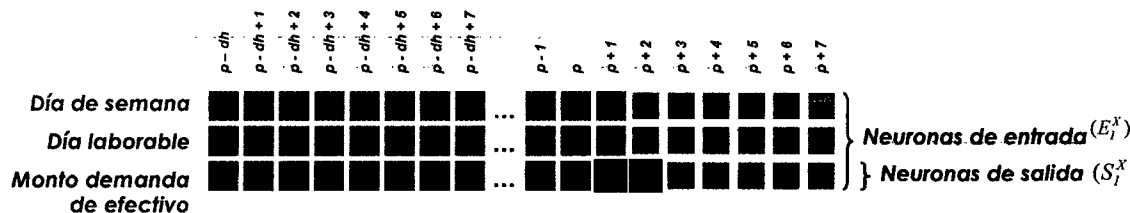


Gráfico 71 : Días a pronosticar
Fuente : Propia
Elaboración : Silvia Granados Flores

El Gráfico 46 puede ser representado:

$$S_{p+1}^X = f(E_{p-dh+1}^X, E_{p-dh+2}^X, E_{p-dh+3}^X, E_{p-dh+4}^X, \dots, E_p^X)$$

$$S_{p+2}^X = f(E_{p-dh+2}^X, E_{p-dh+3}^X, E_{p-dh+4}^X, E_{p-dh+5}^X, \dots, E_{p+1}^X)$$

Donde:

S_{p+1}^X : Neurona de salida correspondiente al cajero automático "X" en el periodo "p+1"

S_{p+2}^X : Neurona de salida correspondiente al cajero automático "X" en el periodo "p+2"

E_{p-dh+1}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p-dh+2".

E_{p-dh+2}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p-dh+2".

E_{p-dh+3}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p-dh+3".

E_{p-dh+4}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p-dh+4".

E_{p-dh+5}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p-dh+5".

E_{p+1}^X : Neurona de entrada correspondiente al cajero automático "X" en el periodo "p+1".

En el módulo Experimentación se identificará el valor óptimo para los días históricos y días a pronosticar (ver Capítulo VIII).

6.4. RESUMEN DEL CAPÍTULO

En este capítulo se han detallado los procedimientos a seguir a fin de construir la capa de neuronas de entrada y de salida del sistema.

Para la construcción de estas dos capas debe definirse el número de días históricos y a pronosticar.

La estructura de datos generada a través del desfase de datos, se representa en el Gráfico 47, donde se observa que los registros de la base de datos final ingresan al procedimiento y generan los registros del sistema de redes neuronales:

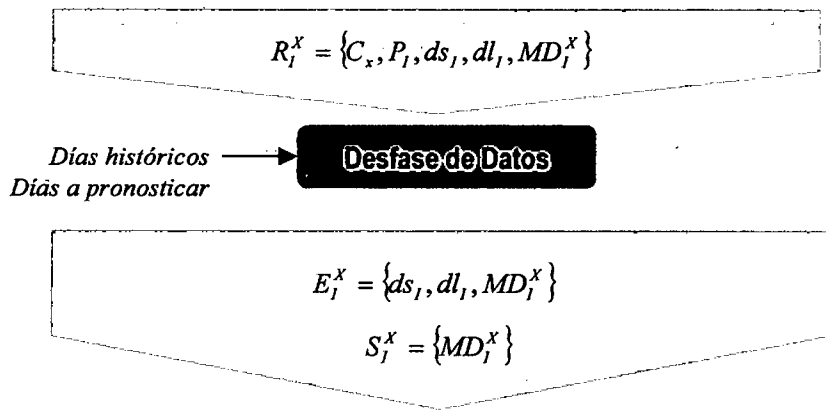


Gráfico 72 : Desfase de datos - Resumen

Fuente : Propia

Elaboración : Silvia Granados Flores

CAPITULO VII

EVALUACIÓN DEL MODELO PROPUESTO

7.1. DESCRIPCIÓN GENERAL

En este capítulo se detallará el método a emplearse a fin de medir la performance del modelo.

La evaluación del modelo propuesto, señalado por la performance, será representada numéricamente a través de la medición del error entre la demanda real y la demanda pronosticada. El estándar a emplearse será la medición MAPE (media absoluta del porcentaje de error), el cual mide la exactitud de una curva predicha comparada con la real (D. Mandic et. al. [26]).

El error se medirá en las corridas de experimentación del modelo. A través del resultado de esta medición, se identificarán los parámetros que mejor se ajustan al modelo (ver Capítulo VIII) y el error que se obtiene con respecto a los valores esperados.

El Gráfico 48 representa el proceso de evaluación del modelo propuesto en Redes Neuronales Artificiales:

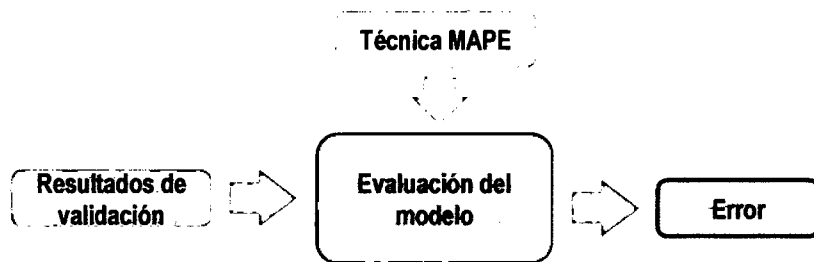


Gráfico 73 : Evaluación del modelo propuesto
 Fuente : Propia
 Elaboración : Silvia Granados Flores

7.2. MEDICIÓN DEL ERROR

La desviación entre la demanda pronosticada por el modelo de Redes Neuronales Artificiales y la demanda real, será medida a través de un procedimiento estándar denominado MAPE.

El cálculo del error MAPE se efectúa de la siguiente manera:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Vr_t - Vp_t}{Vr_t} \right|$$

Donde:

Vr_t : Valor real en tiempo "t"

Vp_t : Valor pronosticado en tiempo "t"

n : Número de puntos pronosticados en la serie de tiempo

La diferencia entre el valor real y pronosticado es dividido por el valor real. El valor absoluto de este cociente se suma con todas las diferencias en los puntos de la serie de tiempo. El promedio de esta suma en términos de porcentaje corresponde al error MAPE.

Cabe recordar que los valores reales fueron transformados a través de la función logaritmo neperiano; por lo tanto dentro del cálculo del error

MAPE deberá considerarse una actividad previa que consiste en efectuar la reversa de la transformación a fin de sincerar el error. Finalmente el cálculo del error MAPE se efectuará así:

$$Error = MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\varepsilon^{Vr'_t} - \varepsilon^{Yp'_t}}{\varepsilon^{Vr'_t}} \right|$$

Donde:

Vr'_t : Valor real en periodo "t" (valor empleado en el modelo; es decir, valor con transformación por logaritmo neperiano)

Yp'_t : Valor pronosticado en periodo "t" (valor resultado del modelo; es decir, valor con transformación por logaritmo neperiano)

n : Número de puntos pronosticados en la serie de tiempo

Para el cálculo de este error debe tomarse en cuenta las siguientes consideraciones:

- La muestra de valores reales no puede contener valores 0 debido a que se produciría una división entre 0.
- Cocientes significativamente grandes o pequeños entre valores pronosticados y reales impactan considerablemente el promedio.

7.3. RESUMEN DEL CAPÍTULO

En este capítulo se detalló el procedimiento para medir el error de pronóstico como resultado del modelo propuesto. Estos resultados se obtendrán en la etapa de experimentación para seleccionar los valores de los parámetros que mejor se ajustan al modelo y originan la mayor performance.

En el Gráfico 49 se representa el resumen del capítulo:

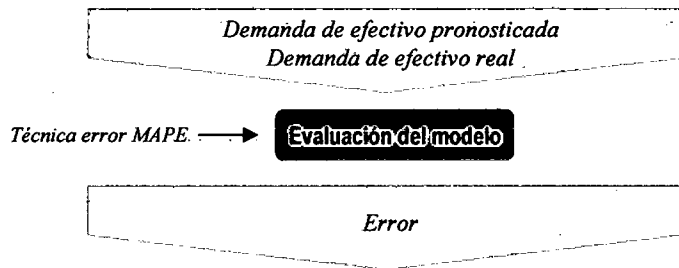


Gráfico 74 : Evaluación del modelo propuesto

Fuente : Propia

Elaboración : Silvia Granados Flores

CAPÍTULO VIII

EXPERIMENTACIÓN

8.1. DESCRIPCIÓN GENERAL

En este capítulo se detallará la experimentación desarrollada en base al modelo propuesto con la finalidad de:

- Determinar los valores óptimos para los parámetros identificados en el modelo.
- Identificar el resultado sobre la evaluación del modelo, el cual se ve reflejado en la medición del error según el método MAPE (ver Capítulo VII).

Los parámetros identificados en el modelo propuesto de Redes Neuronales Artificiales son los siguientes (ver Capítulo IV):

- Función de transferencia
- Función de activación
- Función de aprendizaje
- Ratio de aprendizaje
- Ratio de momentum
- Días históricos
- Días pronosticar
- Neuronas ocultas

En el Gráfico 50 se representa la dinámica a seguir a lo largo del capítulo:

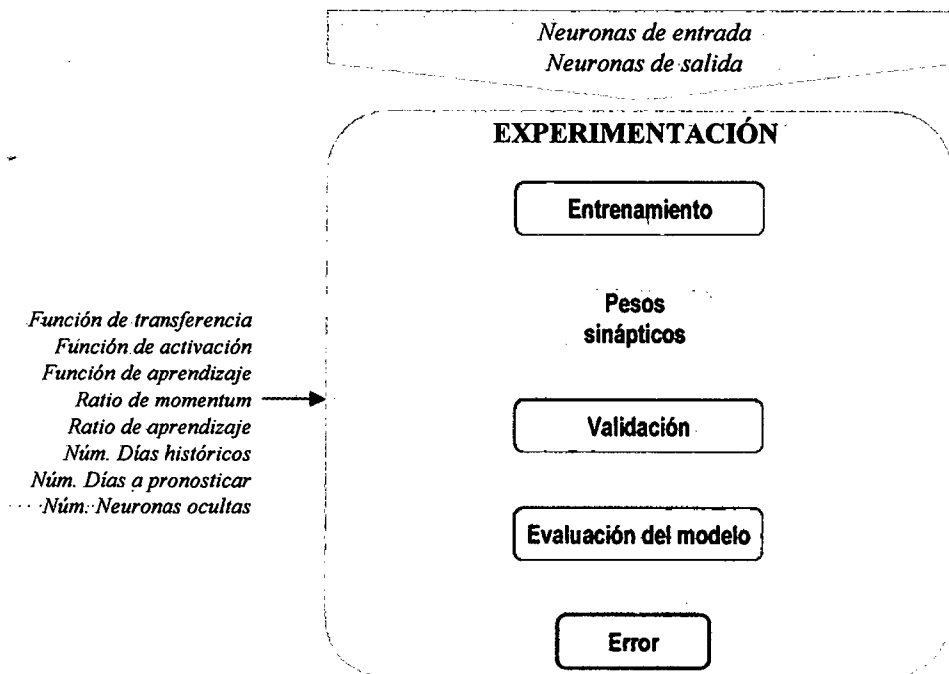


Gráfico 75 : Experimentación del modelo propuesto
 Fuente : Propia
 Elaboración : Silvia Granados Flores

Como puede observarse en el Gráfico 50, el modelo será entrenado y validado en cuantas corridas como combinación de posibles valores puedan tener los parámetros a identificar.

8.2. DETERMINACIÓN DE CARTERAS

De la misma manera, los registros serán separados en carteras (folds) a fin de emplearse un grupo para el proceso de entrenamiento y el resto para la validación. La dinámica que se seguirá en el empleo de carteras se representa a través de un ejemplo en el Gráfico 51:

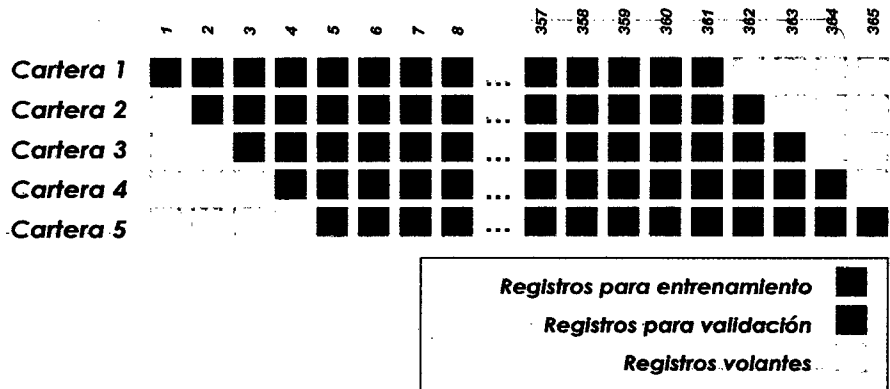


Gráfico 76 : Determinación de carteras
 Fuente : Propia
 Elaboración : Silvia Granados Flores

En el ejemplo del Gráfico 51 el número de carteras es 5; por este motivo, se reservan 4 registros del total para emplearse como volantes (registros que no entrarán en esa corrida del modelo) en la cartera que corresponda.

Cabe indicar que se seleccionó como 3 (tres) el número de carteras para la distribución de registros entre el entrenamiento y la validación.

8.3. CASOS DE PRUEBA

Los casos de prueba a desarrollarse en esta etapa de experimentación se detallan en la Tabla 42.

CASOS DE PRUEBAS

	PARÁMETROS A IDENTIFICAR							
	Función de Transferencia	Función de Activación	Función de Aprendizaje	Ratio de aprendizaje	Ratio de momentum	Número Días Históricos	Número Neuronas Ocultas	Número Días Pronóstico
	tansig - tansig	purelin	traingdm	0.05	0.05	7	(Núm. Días Históricos)/4	1
	tansig - logsig	tansig	traingda	0.50	0.50	15	(Núm. Días Históricos)/2	3
	logsig - tansig	logsig	traingdx	0.75	0.75	30	(Núm. Días Históricos)	5
	logsig - logsig		trainrp	0.90	0.90	45		7
						60		
Número posibles valores	4	3	4	4	4	5	3	4
Número de carteras	3 folds							
Total pruebas del modelo	138240 pruebas							

Tabla 42 : Casos de prueba para experimentación

Fuente : Propia

Elaboración : Silvia Granados Flores

En la Tabla 42 se detallan los posibles valores que tomarán los parámetros identificados en el modelo propuesto; de esta manera, se calcula un total de 138,240 corridas experimentales que se efectuarán.

El procedimiento para identificar los posibles valores a emplearse en la experimentación se basa en aquellos recomendados para redes de retroalimentación del tipo NARX. Dentro de la Tabla 42 se indican en lenguaje y rangos empleados en el paquete de software Matlab 7.8.0 (R2009a) y a continuación de detallan:

1. Función de transferencia: Se emplearán las combinaciones entre las funciones logaritmo sigmoidal (logsig) y tangente hiperbólica sigmoidal (tansig), de la siguiente manera:
 - Tansig – tansig
 - Tansig – logsig
 - Logsig – tansig
 - Logsig - logsig
2. Función de activación: Se emplearán la función lineal (purelin), logaritmo sigmoidal (logsig) y tangente hiperbólica sigmoidal (tansig).
3. Función de aprendizaje: Se emplearán las funciones Aprendizaje adaptable con retroalimentación y pendiente de gradiente (traingda), Aprendizaje retroalimentado con pendiente de gradiente y momentum (traingdm) y Aprendizaje adaptable con retroalimentación, momentum y pendiente de gradiente (traingdx).
4. Ratio de aprendizaje: Debido a que este factor se maneja en un rango [0,1] se emplearán los valores 0.05, 0.5, 0.75 y 0.9 para identificar el comportamiento del modelo con valores extremos y centrales.

5. Ratio de momentum: Debido a que este factor se maneja en un rango $[0,1]$ se emplearán los valores 0.05, 0.5, 0.75 y 0.9 para identificar el comportamiento del modelo con valores extremos y centrales.
6. Número Días Históricos: Los valores a emplearse se diferenciarán por semanas, siendo los valores a emplearse 7, 15, 30, 45 y 60 días.
7. Número Neuronas Ocultas: Los valores a emplearse dependerán de una relación proporcional con los días históricos (neuronas de entrada), en base a ello se realizarán pruebas con neuronas ocultas igual, mitad y cuarta parte de los días pronósticos.
8. Número Días Pronóstico: Los valores a emplearse se diferenciarán por días, siendo los valores a emplearse 1, 3, 5 y 7 días.

Estos valores serán combinados y empleados en las 138,240 pruebas que se efectuarán sobre el modelo propuesto.

Asimismo se empleará el paquete de software Matlab 7.8.0 (R2009a) a fin de ejecutar las pruebas y obtener los resultados de la experimentación.

8.4. FUNCIÓN DE TRANSFERENCIA

En las Tablas 43, 44, 45, 46 y 47 se detallan los resultados (error MAPE promedio) de la experimentación realizada sobre el modelo propuesto para el parámetro Función de Transferencia.

En cada una de las tablas se ha resaltado el valor óptimo que resultó de este parámetro en cada uno de los cinco cajeros automáticos analizados. Cabe recordar que los valores propuestos (ver Capítulo VII) para este parámetro son:

- Tansig – tansig
- Tansig – logsig
- Logsig – tansig
- Logsig - logsig

CAJERO AUTOMÁTICO A				
Función de Transferencia	Número de Experimentos	Error mínimo	Error máximo	Error Promedio
tansig - tansig	34,560	1.78%	118.97%	25.75%
tansig - logsig	34,560	1.67%	74.86%	23.43%
logsig - tansig	34,560	3.56%	205.36%	32.35%
logsig - logsig	34,560	4.65%	103.35%	31.08%

Tabla 43 : Función de Transferencia – Cajero automático A
Fuente : Propia
Elaboración : Silvia Granados Flores

CAJERO AUTOMÁTICO B				
Función de Transferencia	Número de Experimentos	Error mínimo	Error máximo	Error Promedio
tansig - tansig	34,560	2.83%	164.47%	29.10%
tansig - logsig	34,560	2.84%	118.59%	25.68%
logsig - tansig	34,560	2.89%	178.41%	30.65%
logsig - logsig	34,560	2.79%	119.54%	29.15%

Tabla 44 : Función de Transferencia – Cajero automático B
Fuente : Propia
Elaboración : Silvia Granados Flores

CAJERO AUTOMÁTICO C				
Función de Transferencia	Número de Experimentos	Error mínimo	Error máximo	Error Promedio
tansig - tansig	34,560	2.26%	285.92%	35.29%
tansig - logsig	34,560	2.58%	279.52%	35.61%
logsig - tansig	34,560	3.27%	285.92%	35.82%
logsig - logsig	34,560	3.02%	291.47%	36.47%

Tabla 45 : Función de Transferencia – Cajero automático C
Fuente : Propia

Elaboración : Silvia Granados Flores

CAJERO AUTOMÁTICO D				
Función de Transferencia	Número de Experimentos	Error mínimo	Error máximo	Error Promedio
tansig - tansig	34,560	2.14%	228.70%	36.26%
tansig - logsig	34,560	2.21%	273.49%	37.04%
logsig - tansig	34,560	1.14%	272.37%	35.83%
logsig - logsig	34,560	1.98%	259.15%	36.24%

Tabla 46 : Función de Transferencia – Cajero automático D

Fuente : Propia

Elaboración : Silvia Granados Flores

CAJERO AUTOMÁTICO E				
Función de Transferencia	Número de Experimentos	Error mínimo	Error máximo	Error Promedio
tansig - tansig	34,560	2.64%	148.76%	22.49%
tansig - logsig	34,560	2.48%	164.29%	23.41%
logsig - tansig	34,560	3.41%	159.18%	22.44%
logsig - logsig	34,560	2.95%	160.21%	22.95%

Tabla 47 : Función de Transferencia – Cajero automático E

Fuente : Propia

Elaboración : Silvia Granados Flores

Finalmente los valores óptimos para el parámetro Función de Transferencia en cada uno de los cajeros automáticos, se detallan en la Tabla 48:

Cajero automático	Función de Transferencia
A	tansig - logsig
B	tansig - logsig
C	tansig - tansig
D	logsig - tansig
E	logsig - tansig

Tabla 48 : Función de Transferencia – Valores óptimos

Fuente : Propia

Elaboración : Silvia Granados Flores

8.5. FUNCIÓN DE ACTIVACIÓN

Los Gráficos 52, 53, 54, 55 y 56 representan el comportamiento del error promedio resultado de la experimentación realizada en base a los valores propuestos para el parámetro Función de Activación, los cuales son (ver Capítulo VII):

- Purelin
- Logsig
- Tansig

En cada uno de los gráficos se ha resaltado el valor óptimo, el cual se relaciona con el identificado para el punto anterior sobre Función de Transferencia (ver Tabla 48).

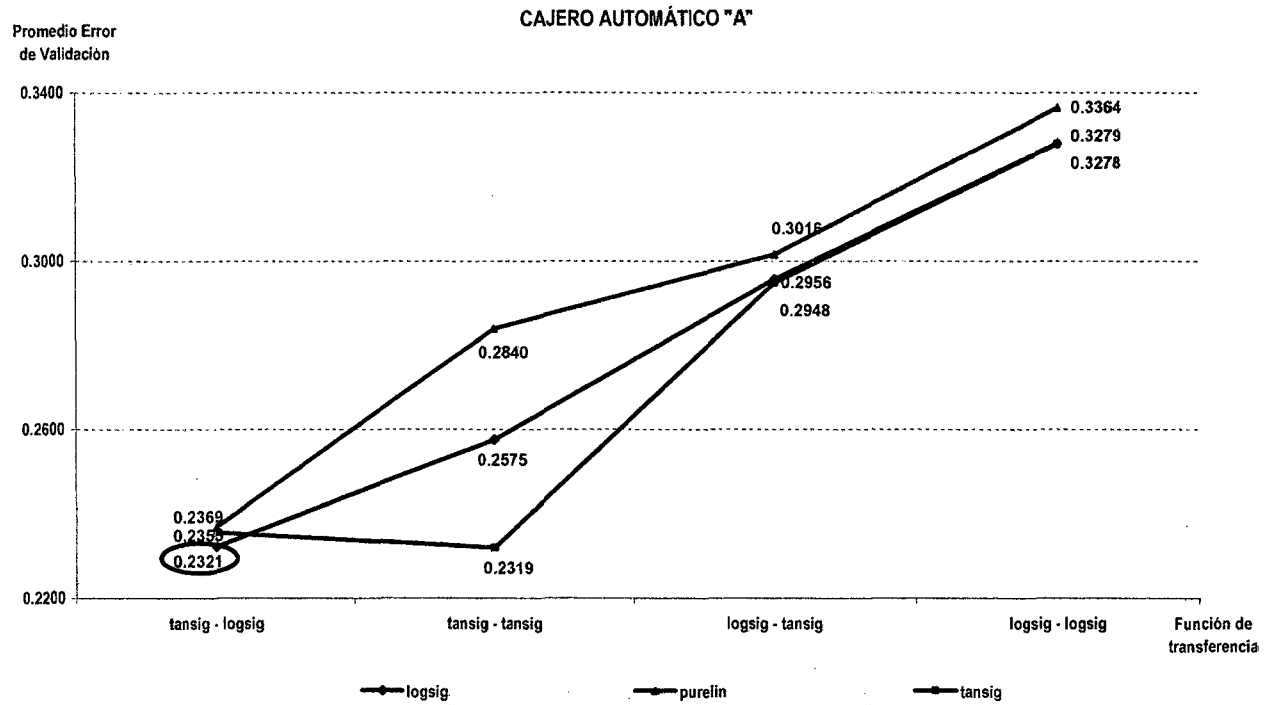


Gráfico 77 : Función de Activación – Cajero automático A
Fuente : Propia
Elaboración : Silvia Granados Flores

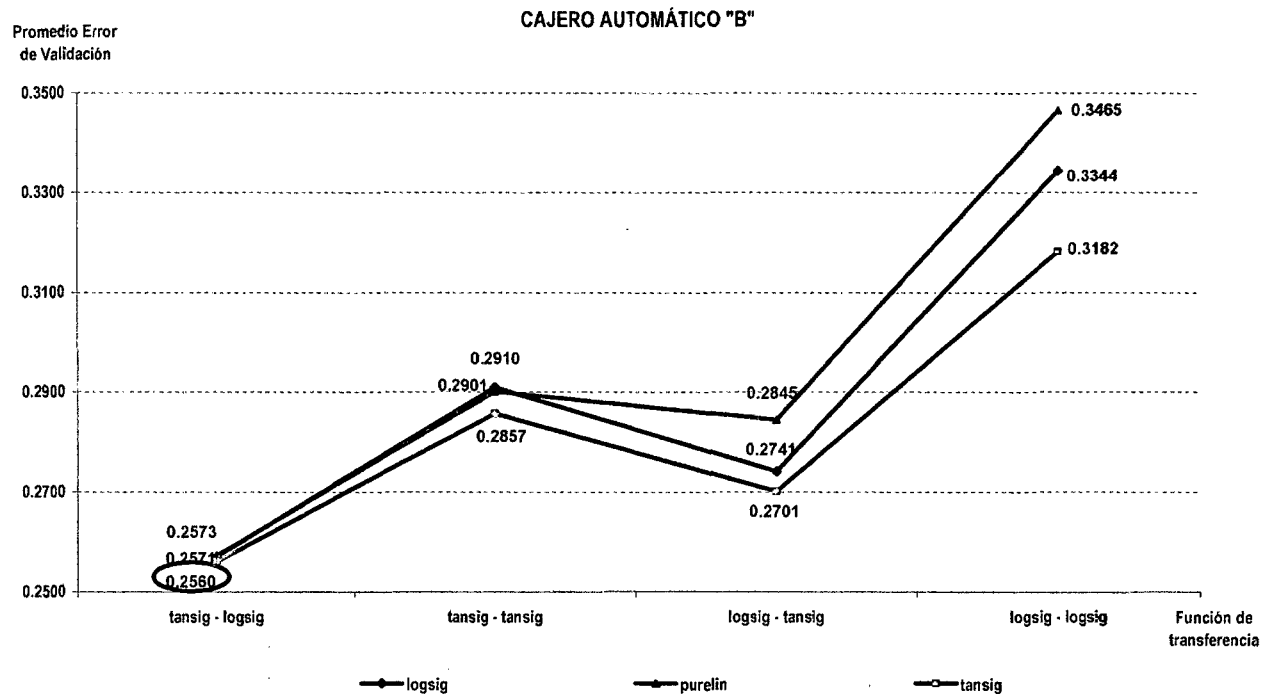


Gráfico 78 : Función de Activación – Cajero automático B
Fuente : Propia
Elaboración : Silvia Granados Flores

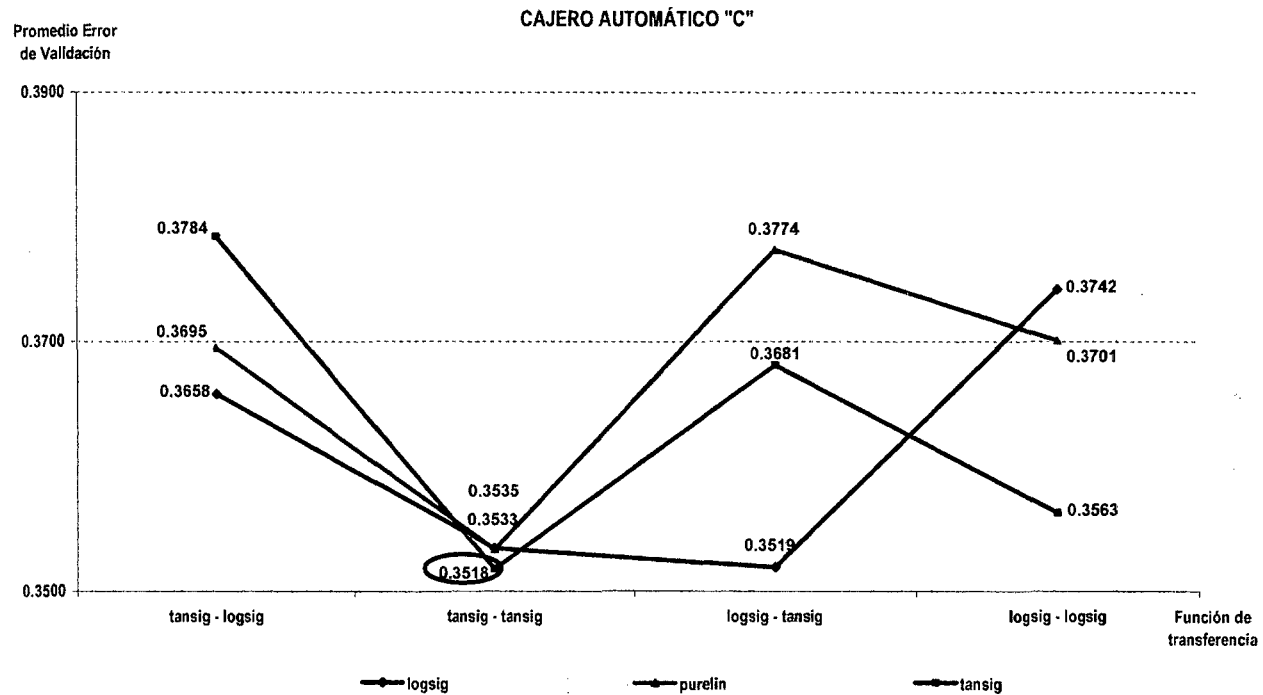


Gráfico 79 : Función de Activación – Cajero automático C
Fuente : Propia
Elaboración : Silvia Granados Flores

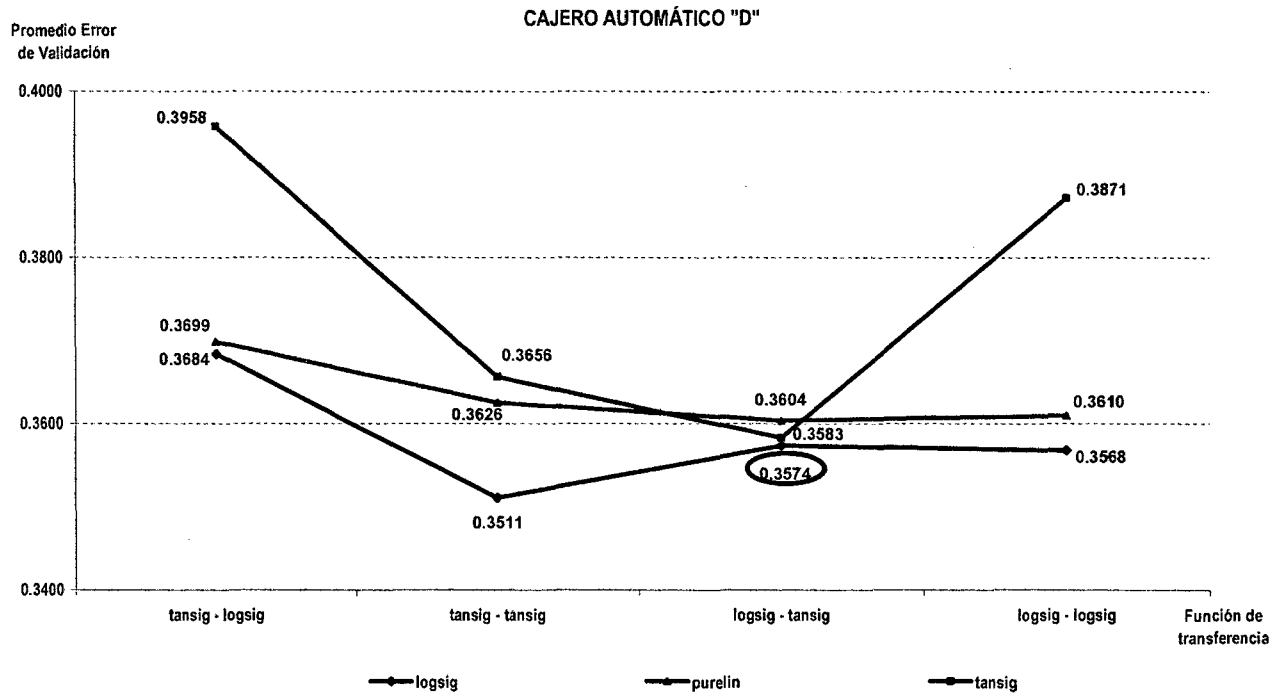


Gráfico 80 : Función de Activación – Cajero automático D
 Fuente : Propia
 Elaboración : Silvia Granados Flores

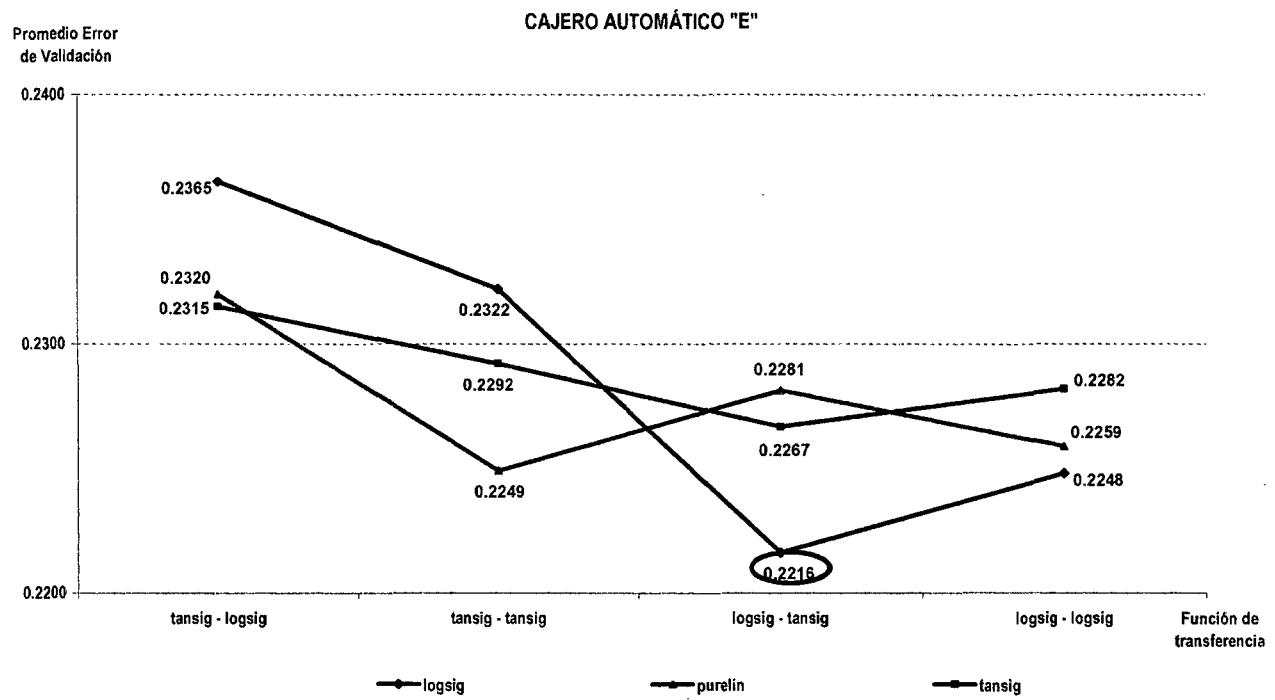


Gráfico 81 : Función de Activación – Cajero automático E
Fuente : Propia
Elaboración : Silvia Granados Flores

Los valores del parámetro Función de Activación que presentan el menor error promedio se detallan en la Tabla 49:

Cajero automático	Función de Activación	Error Promedio Mínimo
A	logsig	23.21%
B	tansig	25.60%
C	tansig	35.18%
D	logsig	35.74%
E	logsig	22.16%

Tabla 49 : Función de Activación – Error promedio mínimo

Fuente : Propia

Elaboración : Silvia Granados Flores

En la Tabla 50 se consolidan los valores óptimos de los parámetros hasta el momento identificados:

VALORES ÓPTIMOS		
Cajero automático	Función de Transferencia	Función de Activación
A	tansig - logsig	logsig
B	tansig - logsig	tansig
C	tansig - tansig	tansig
D	logsig - tansig	logsig
E	logsig - tansig	logsig

Tabla 50 : Valores óptimos de parámetros

Fuente : Propia

Elaboración : Silvia Granados Flores

8.6. FUNCIÓN DE APRENDIZAJE

Los Gráficos 57, 58, 59, 60 y 61 representan el comportamiento del error promedio de validación para los valores propuestos (ver Capítulo VII):

- Traingda
- Traingdm
- Traingdx
- Trainrp

En cada uno de los gráficos se ha resaltado el valor óptimo, el cual se relaciona con el identificado para la Función de Activación (ver Tabla 49).

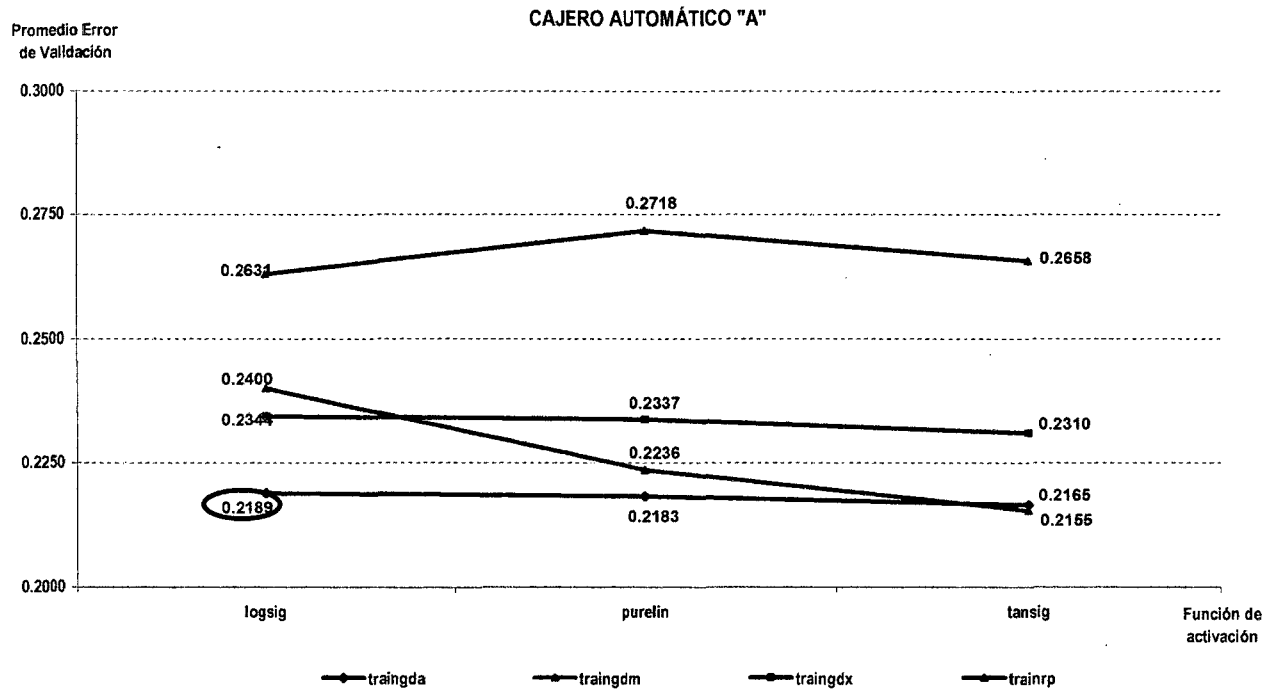


Gráfico 82 : Función de Aprendizaje – Cajero automático A
Fuente : Propia
Elaboración : Silvia Granados Flores

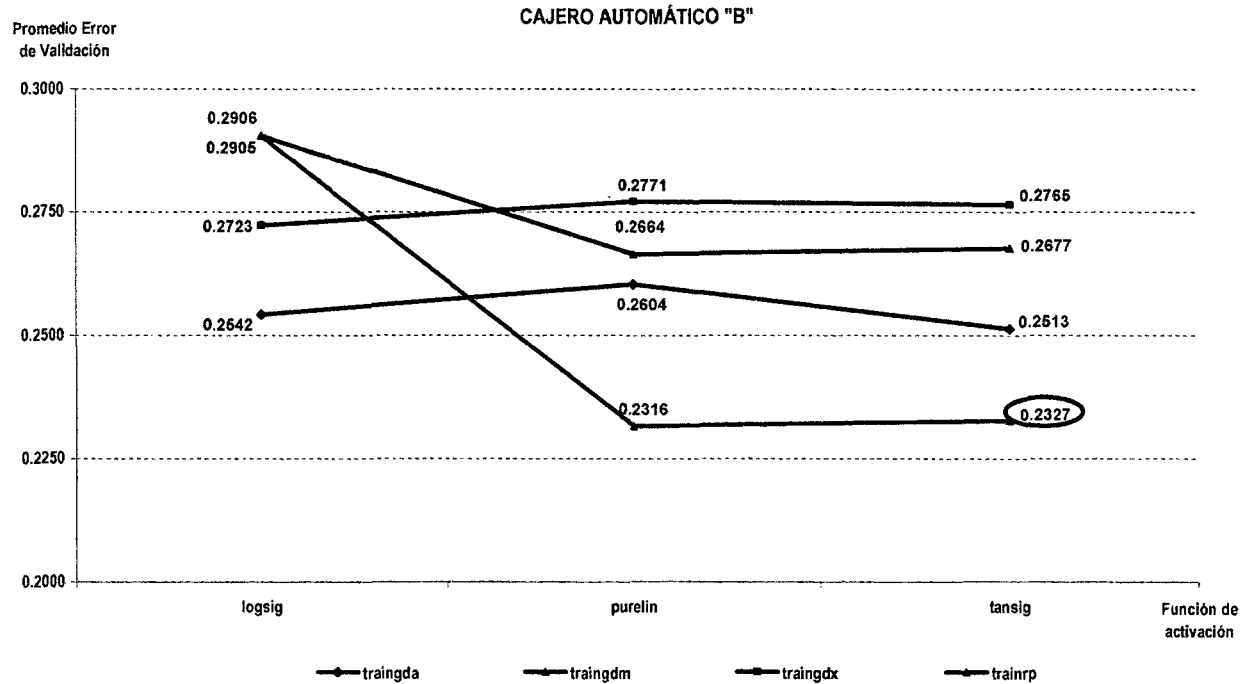


Gráfico 83 : Función de Aprendizaje – Cajero automático B
Fuente : Propia
Elaboración : Silvia Granados Flores

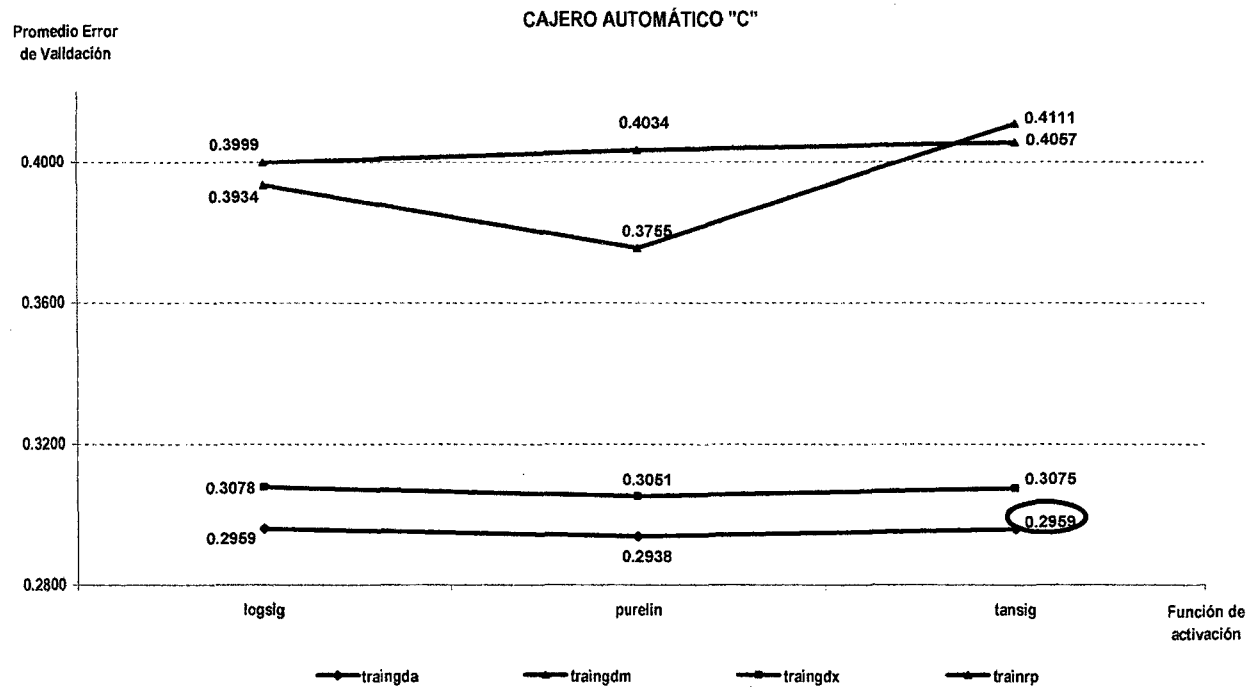


Gráfico 84 : Función de Aprendizaje – Cajero automático C
Fuente : Propia
Elaboración : Silvia Granados Flores

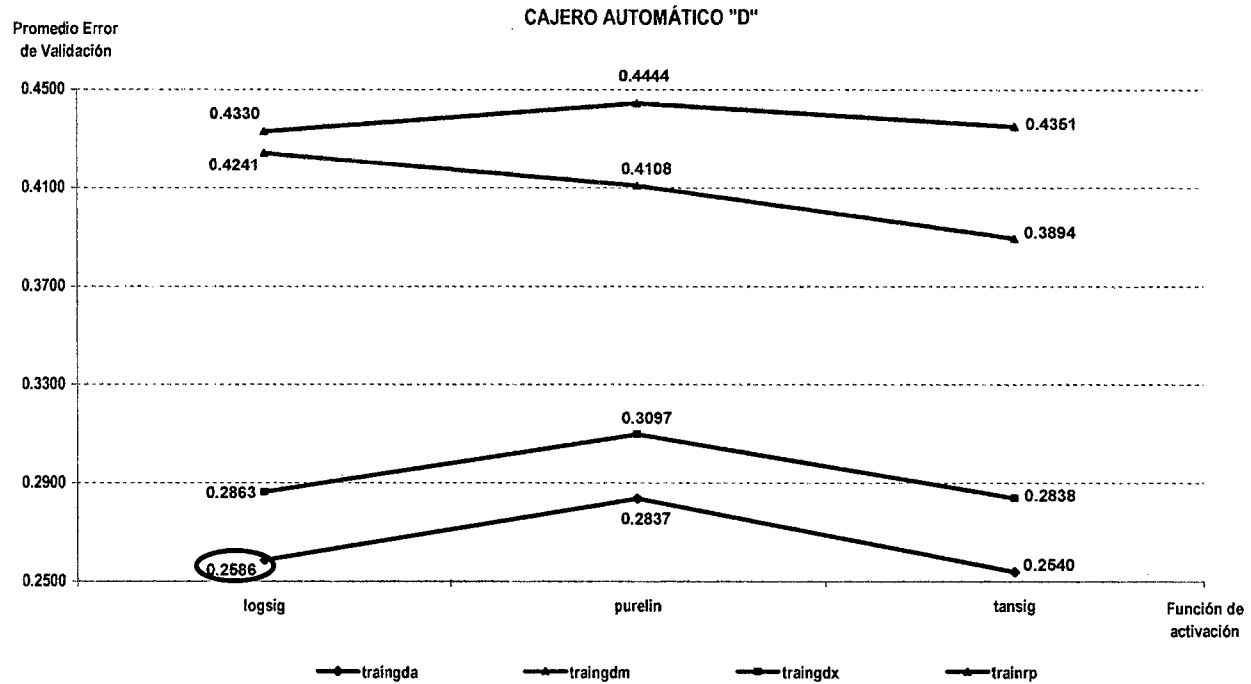


Gráfico 85 : Función de Aprendizaje – Cajero automático D
Fuente : Propia
Elaboración : Silvia Granados Flores

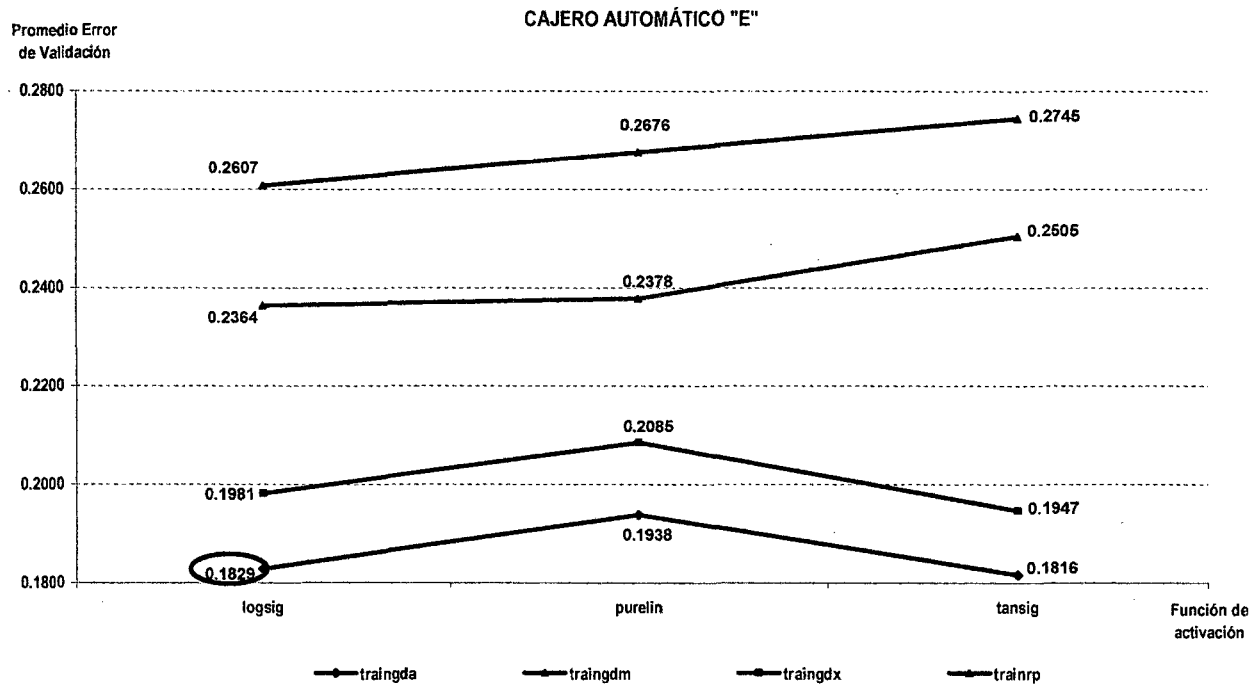


Gráfico 86 : Función de Aprendizaje – Cajero automático E
Fuente : Propia
Elaboración : Silvia Granados Flores

Los valores del parámetro Función de Aprendizaje que presentan el menor error en cada uno de los cajeros automáticos, se detallan en la

Tabla 51:

Cajero automático	Función de Aprendizaje	Error Promedio Mínimo
A	traingda	21.89%
B	traingdm	23.27%
C	traingda	29.59%
D	traingda	25.86%
E	traingda	18.29%

Tabla 51 : Función de Aprendizaje – Error promedio mínimo

Fuente : Propia

Elaboración : Silvia Granados Flores

En la Tabla 52 se consolidan los valores óptimos de los parámetros hasta el momento identificados:

VALORES ÓPTIMOS			
Cajero automático	Función de Transferencia	Función de Activación	Función de Aprendizaje
A	tansig - logsig	logsig	traingda
B	tansig - logsig	tansig	traingdm
C	tansig - tansig	tansig	traingda
D	logsig - tansig	logsig	traingda
E	logsig - tansig	logsig	traingda

Tabla 52 : Valores óptimos de parámetros

Fuente : Propia

Elaboración : Silvia Granados Flores

8.7. RATIO DE APRENDIZAJE

Los Gráficos 62, 63, 64, 65 y 66 presentan el comportamiento del error promedio de validación para los valores propuestos:

- 0.05
- 0.50
- 0.75
- 0.90

En cada uno de los gráficos se ha resaltado el valor óptimo con respecto al identificado para la Función de Aprendizaje (ver Tabla 51).

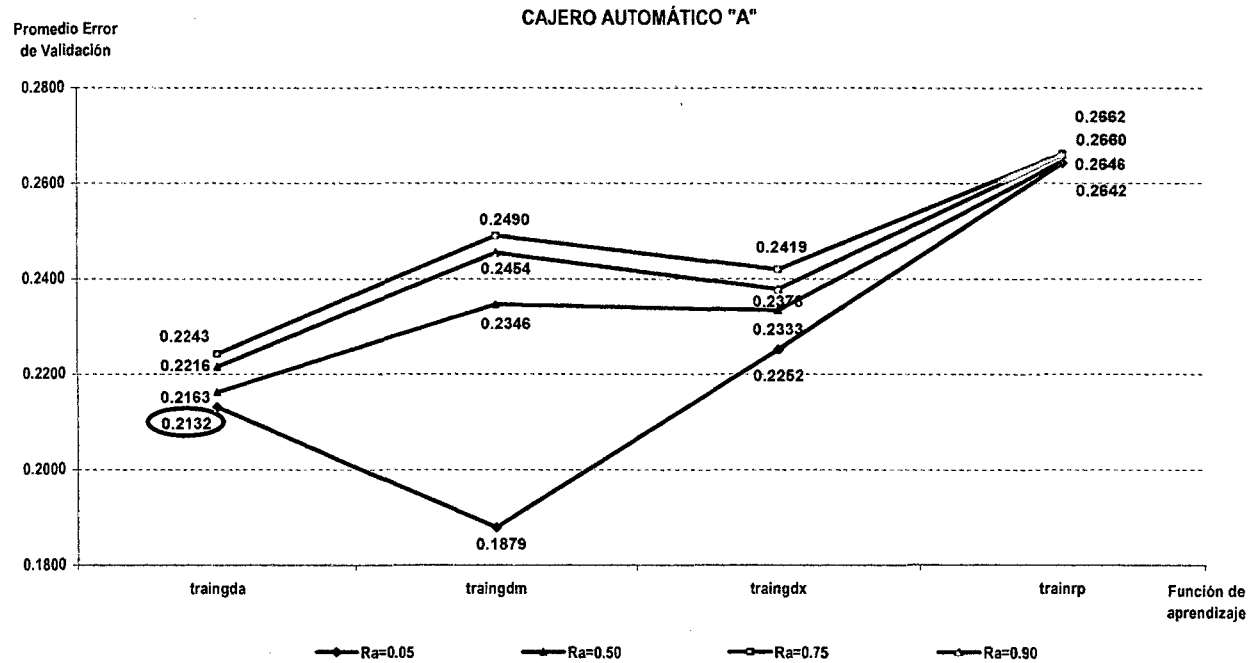


Gráfico 87 : Ratio de Aprendizaje – Cajero automático A
Fuente : Propia
Elaboración : Silvia Granados Flores

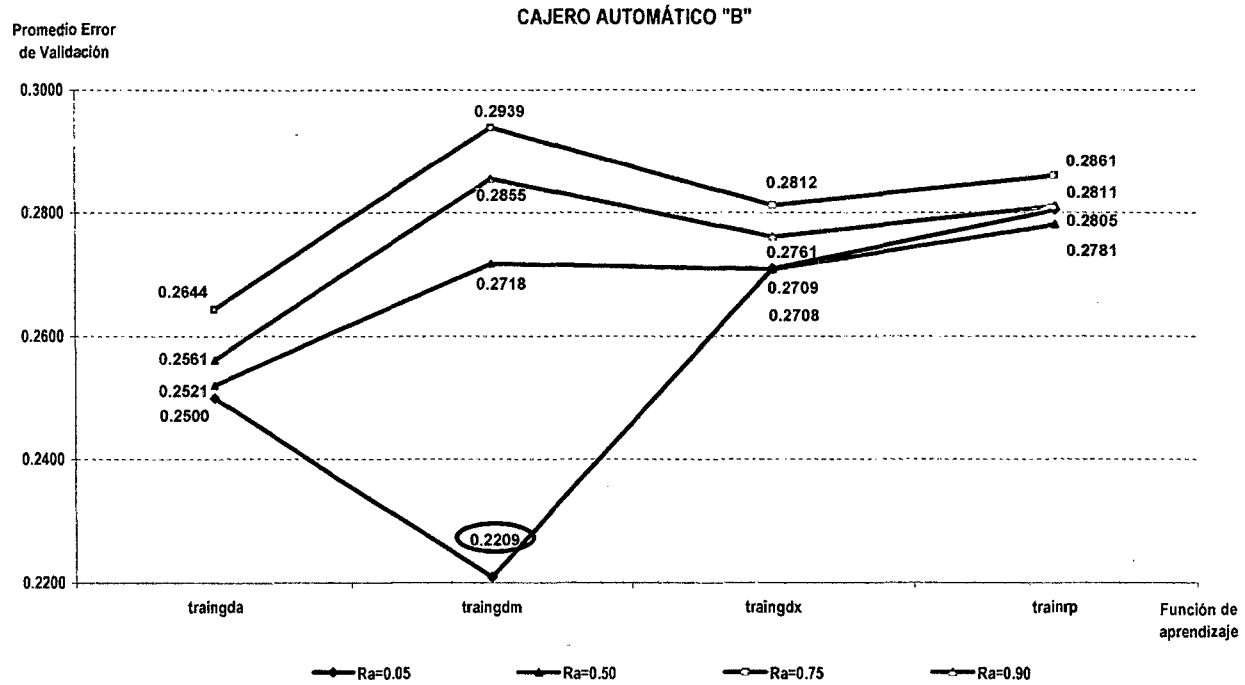


Gráfico 88 : Ratio de Aprendizaje - Cajero automático B
Fuente : Propia
Elaboración : Silvia Granados Flores

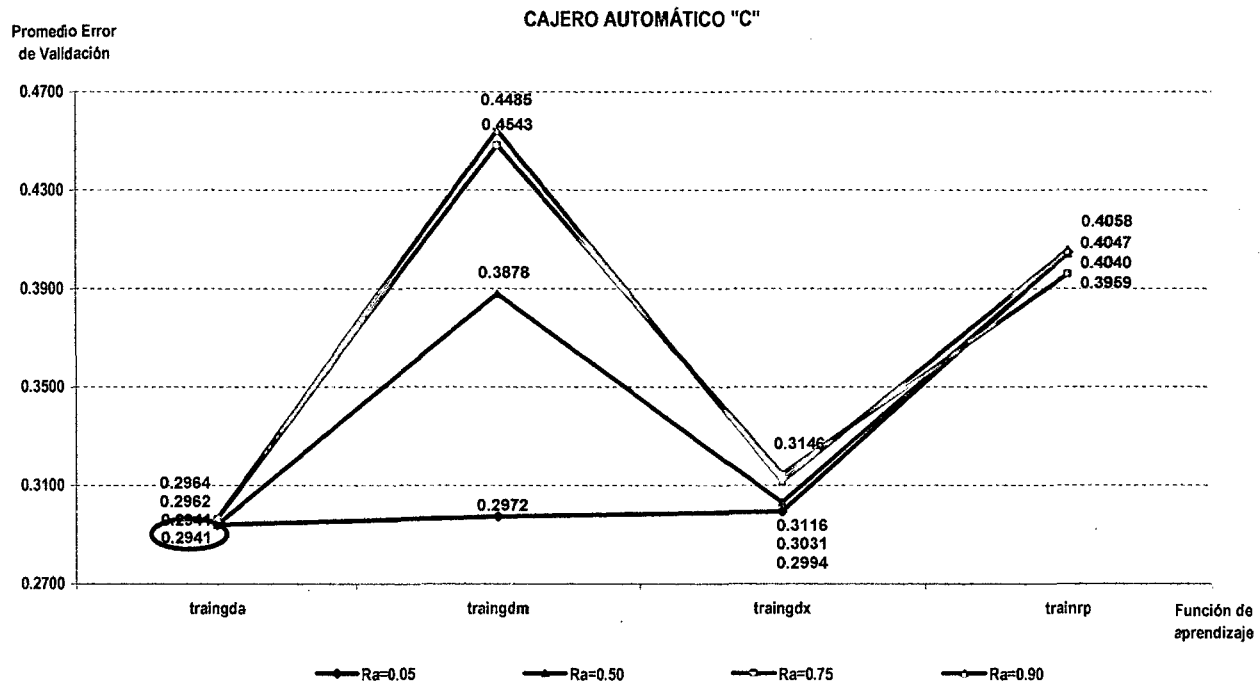


Gráfico 89 : Ratio de Aprendizaje – Cajero automático C
Fuente : Propia
Elaboración : Silvia Granados Flores

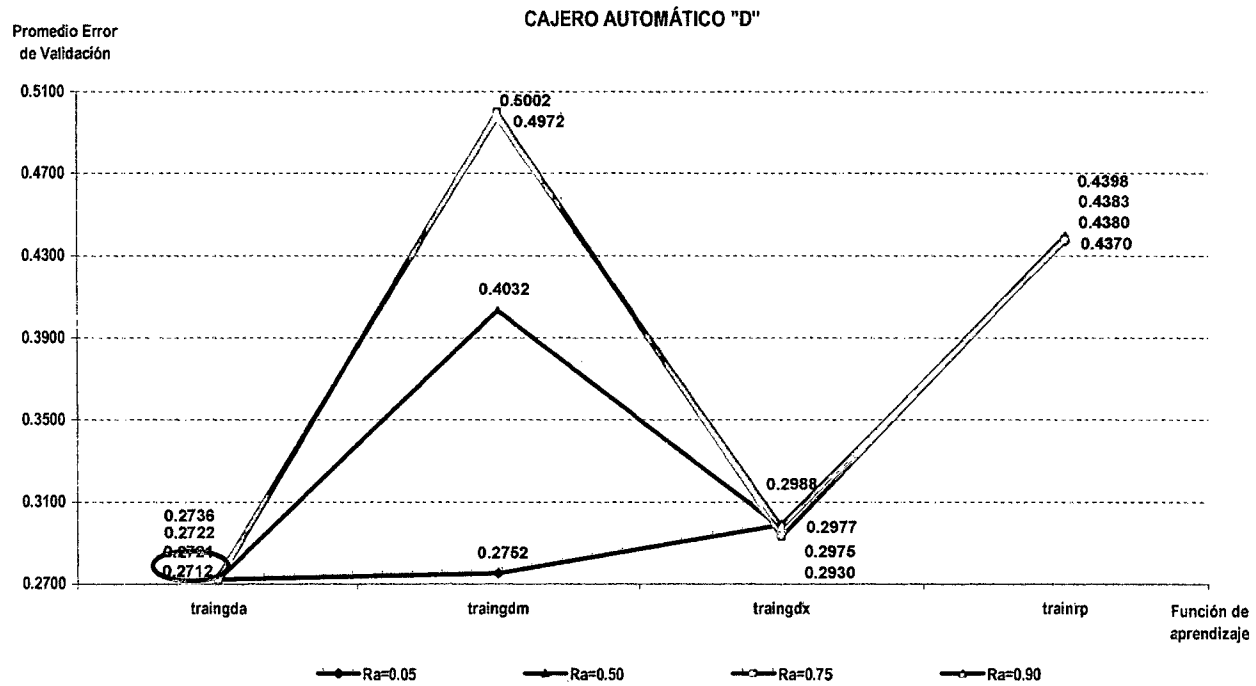


Gráfico 90 : Ratio de Aprendizaje – Cajero automático D
Fuente : Propia
Elaboración : Silvia Granados Flores

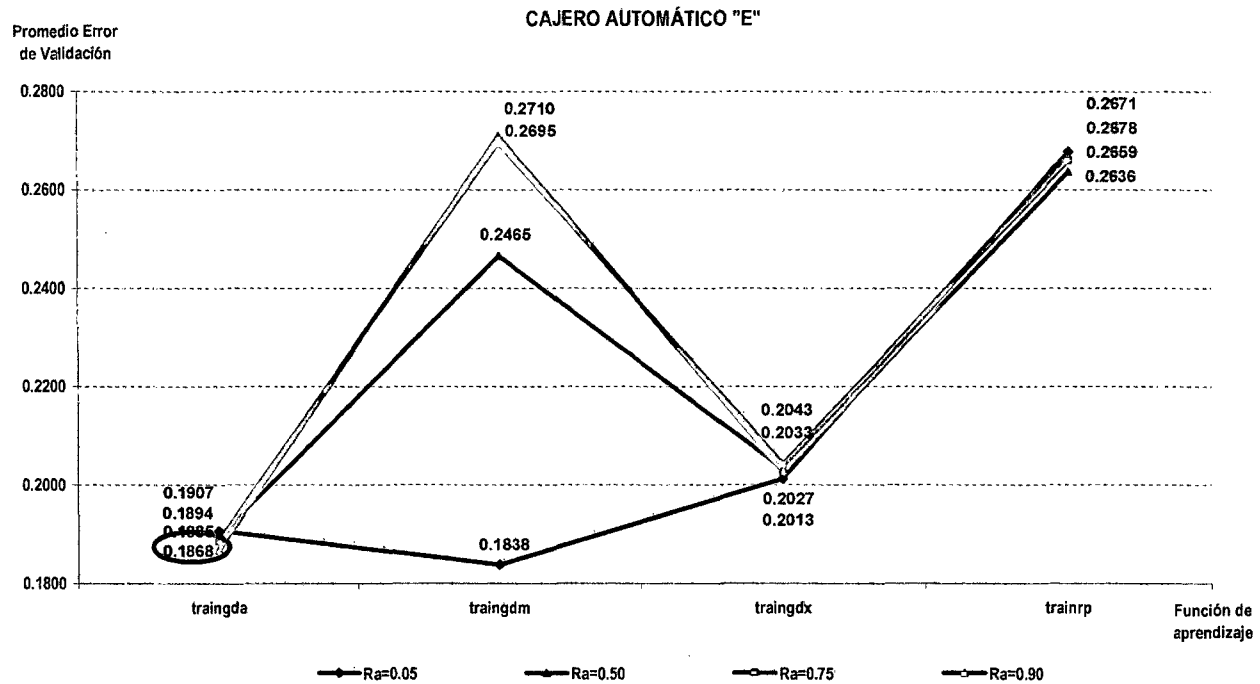


Gráfico 91 : Ratio de Aprendizaje – Cajero automático E
Fuente : Propia
Elaboración : Silvia Granados Flores

Los valores del parámetro Ratio de Aprendizaje que presentan el menor error en cada uno de los cajeros automáticos, se detallan en la Tabla 53:

Cajero automático	Ratio de Aprendizaje	Error Promedio Mínimo
A	0.05	21.32%
B	0.05	22.09%
C	0.05	29.41%
D	0.75	27.12%
E	0.90	18.68%

Tabla 53 : Ratio de Aprendizaje – Error promedio mínimo

Fuente : Propia

Elaboración : Silvia Granados Flores

En la Tabla 54 se consolidan los valores óptimos hasta el momento identificados:

VALORES ÓPTIMOS				
Cajero automático	Función de Transferencia	Función de Activación	Función de Aprendizaje	Ratio de Aprendizaje
A	tansig - logsig	logsig	traingda	0.05
B	tansig - logsig	tansig	traingdm	0.05
C	tansig - tansig	tansig	traingda	0.05
D	logsig - tansig	logsig	traingda	0.75
E	logsig - tansig	logsig	traingda	0.90

Tabla 54 : Valores óptimos de parámetros

Fuente : Propia

Elaboración : Silvia Granados Flores

8.8. RATIO MOMENTUM

Los Gráficos 67, 68, 69, 70 y 71 presentan el error promedio identificado para los valores propuestos:

- 0.05
- 0.50
- 0.75
- 0.90

En cada uno de los gráficos se ha resaltado el valor óptimo con respecto al identificado para el Ratio de Aprendizaje (ver Tabla 53).

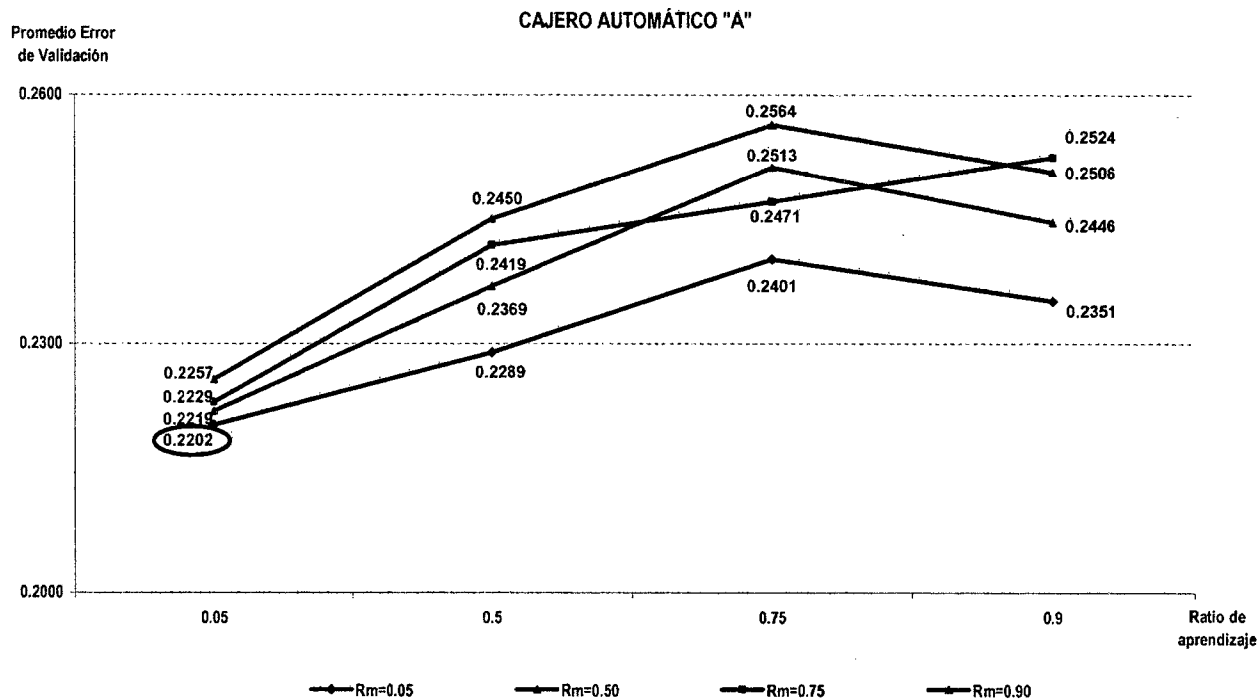


Gráfico 92 : Ratio Momentum – Cajero automático A
Fuente : Propia
Elaboración : Silvia Granados Flores

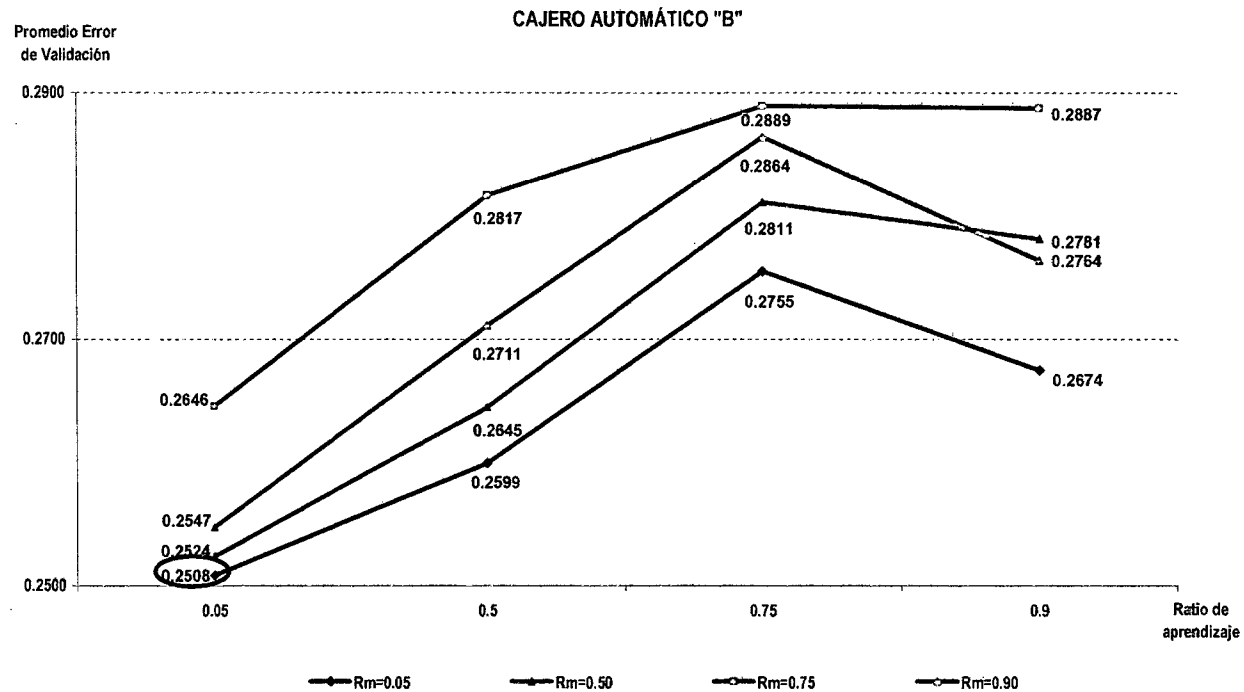


Gráfico 93 : Ratio Momentum - Cajero automático B
Fuente : Propla
Elaboración : Silvia Granados Flores

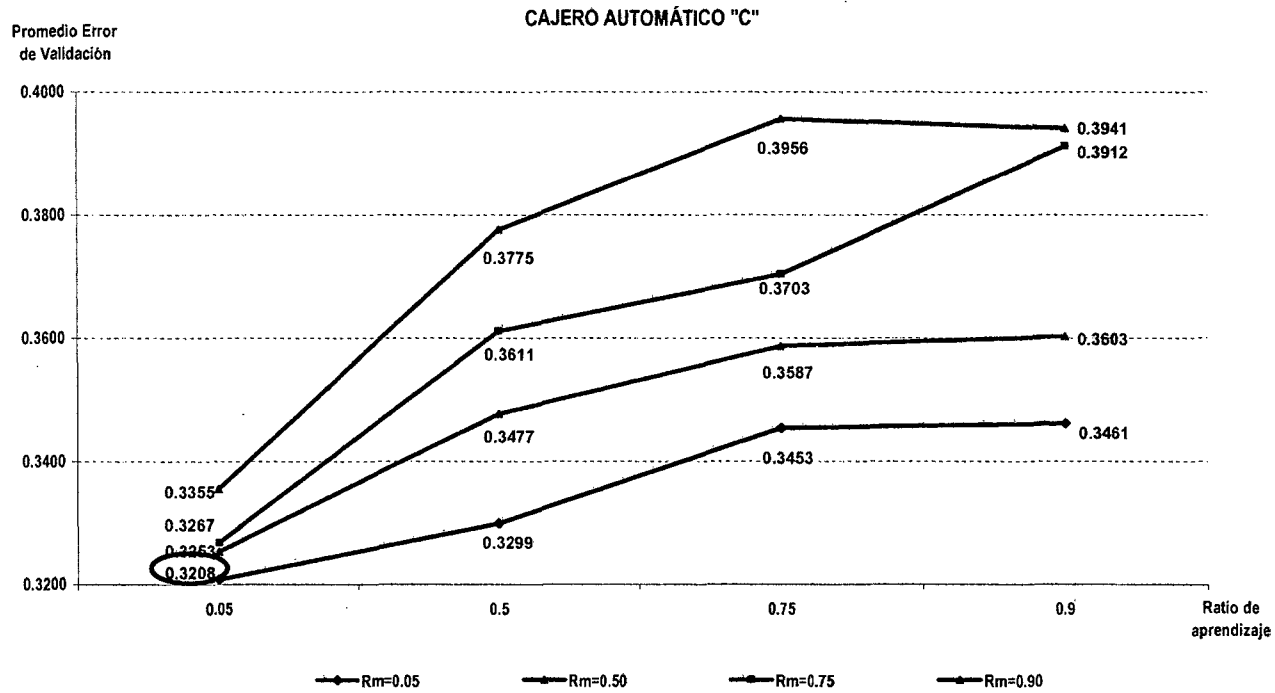


Gráfico 94 : Ratio Momentum – Cajero automático C
Fuente : Propia
Elaboración : Silvia Granados Flores

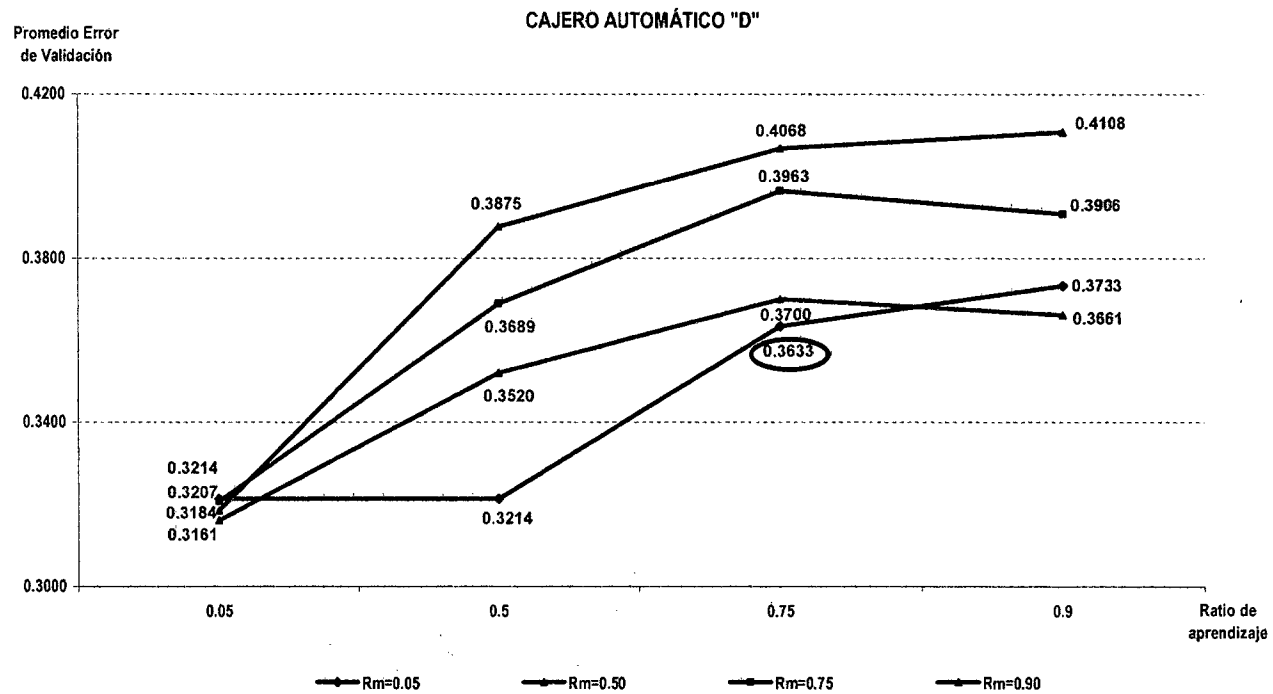


Gráfico 95 : Ratio Momentum – Cajero automático D
Fuente : Propia
Elaboración : Silvia Granados Flores

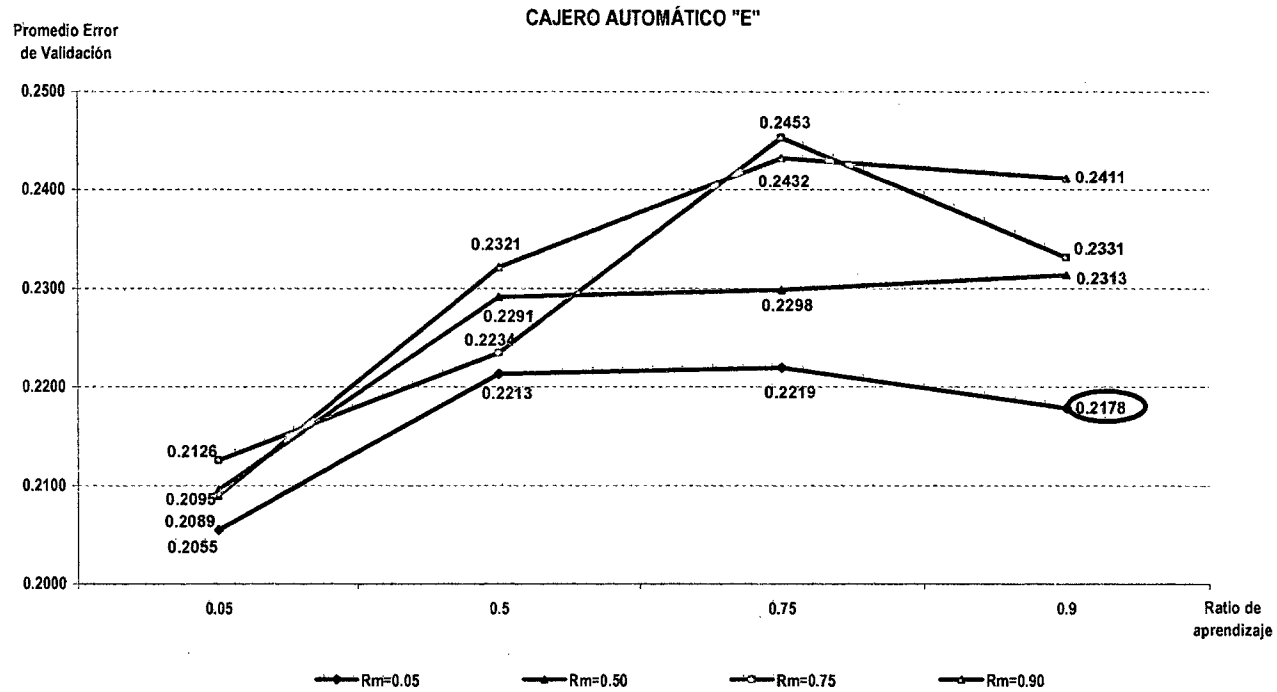


Gráfico 96 : Ratio Momentum – Cajero automático E
Fuente : Propia
Elaboración : Silvia Granados Flores

Los valores del parámetro Ratio Momentum que presentan el menor error se detallan en la Tabla 55:

Cajero automático	Ratio Momentum	Error Promedio Mínimo
A	0.05	22.02%
B	0.05	25.08%
C	0.05	32.08%
D	0.75	36.33%
E	0.90	21.78%

Tabla 55 : Ratio Momentum – Error promedio mínimo

Fuente : Propia

Elaboración : Silvia Granados Flores

En la Tabla 56 se representan los valores óptimos hasta el momento identificados:

VALORES ÓPTIMOS					
Cajero automático	Función de Transferencia	Función de Activación	Función de Aprendizaje	Ratio de Aprendizaje	Ratio Momentum
A	tansig - logsig	logsig	traingda	0.05	0.05
B	tansig - logsig	tansig	traingdm	0.05	0.05
C	tansig - tansig	tansig	traingda	0.05	0.05
D	logsig - tansig	logsig	traingda	0.75	0.75
E	logsig - tansig	logsig	traingda	0.90	0.90

Tabla 56 : Valores óptimos de parámetros

Fuente : Propia

Elaboración : Silvia Granados Flores

8.9. DÍAS HISTÓRICOS

Los Gráficos 72, 73, 74, 75 y 76 presentan el error promedio resultante de la experimentación realizada a fin de identificar el valor óptimo de días históricos que deben emplearse para efectuar el pronóstico de la demanda. Los valores empleados en la experimentación son 7, 15, 30, 45 y 60 días históricos.

En cada uno de los gráficos se ha resaltado el valor óptimo con respecto al identificado para el Ratio Momentum (ver Tabla 55).

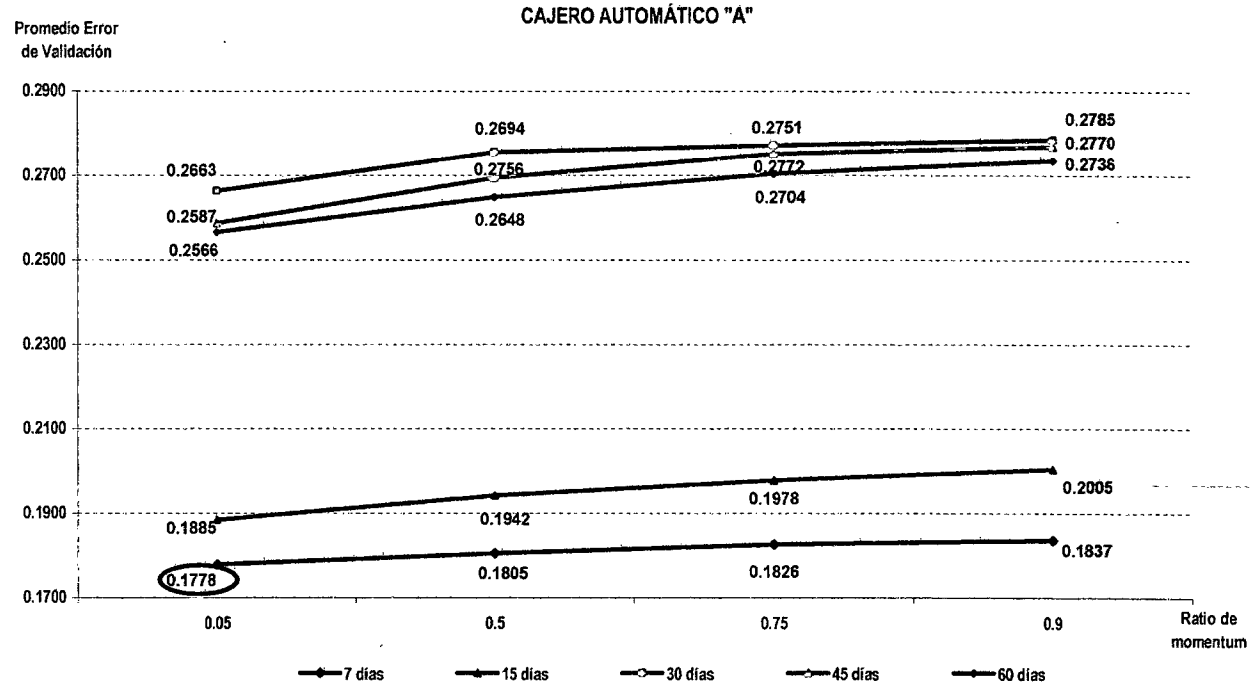


Gráfico 97 : Días Históricos - Cajero automático A
Fuente : Propia
Elaboración : Silvia Granados Flores

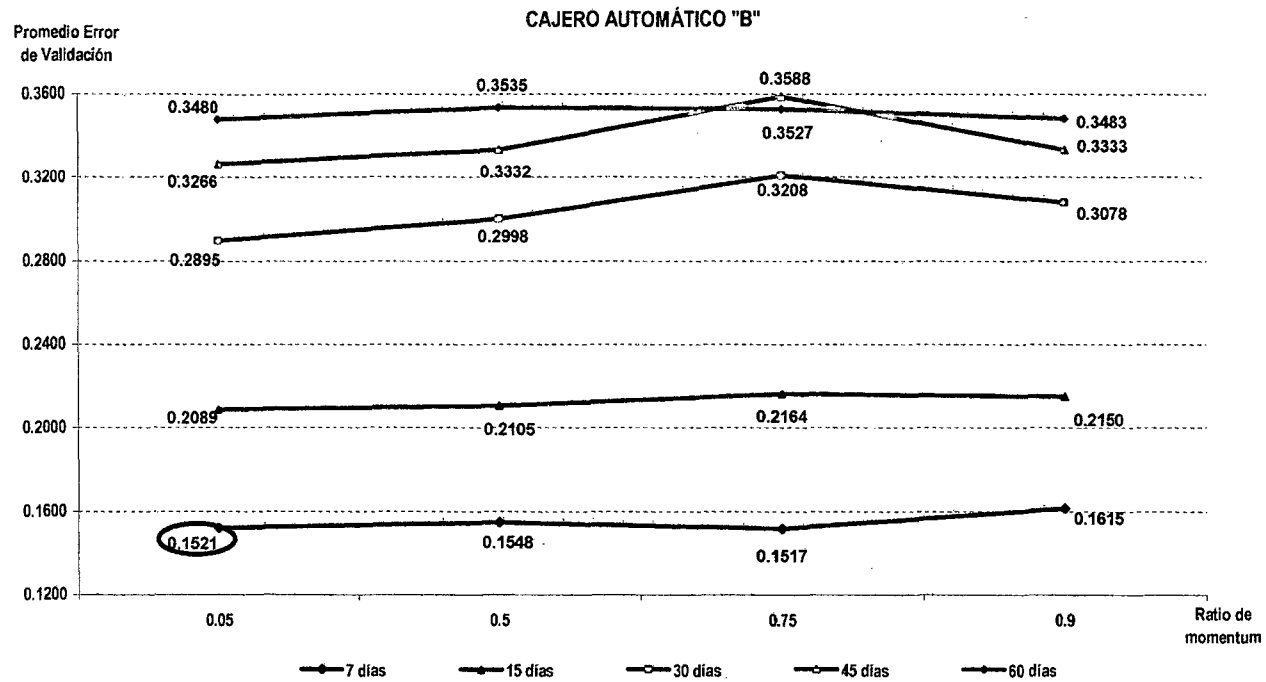


Gráfico 98 : Días Históricos – Cajero automático B
Fuente : Propia
Elaboración : Silvia Granados Flores

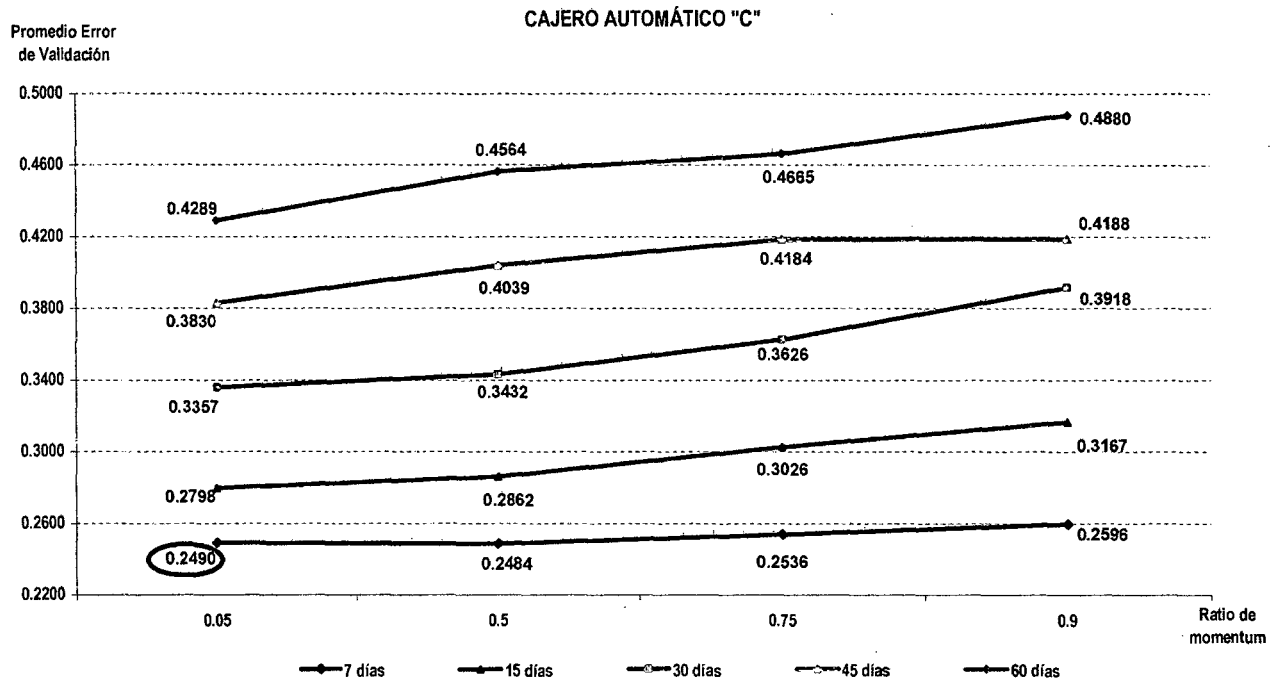


Gráfico 99 : Días Históricos - Cajero automático C
Fuente : Propia
Elaboración : Silvia Granados Flores

CAJERO AUTOMÁTICO "D"

Promedio Error de Validación

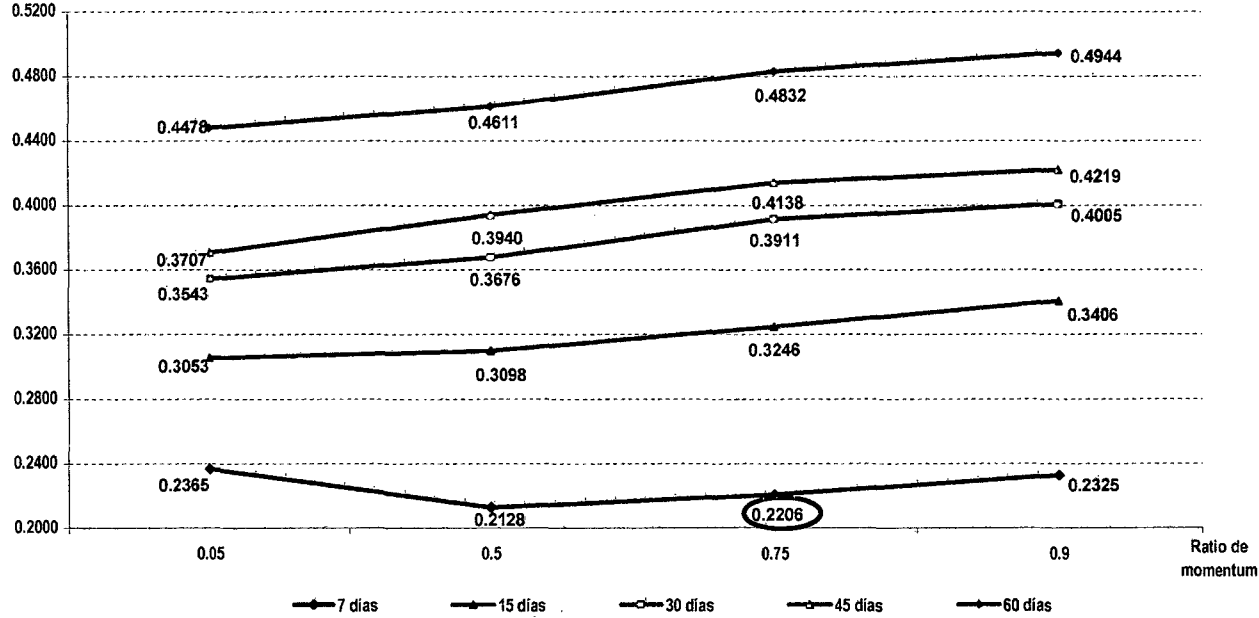


Gráfico 100 : Días Históricos - Cajero automático D
 Fuente : Propia
 Elaboración : Silvia Granados Flores

Promedio Error
de Validación

CAJERO AUTOMÁTICO "E"

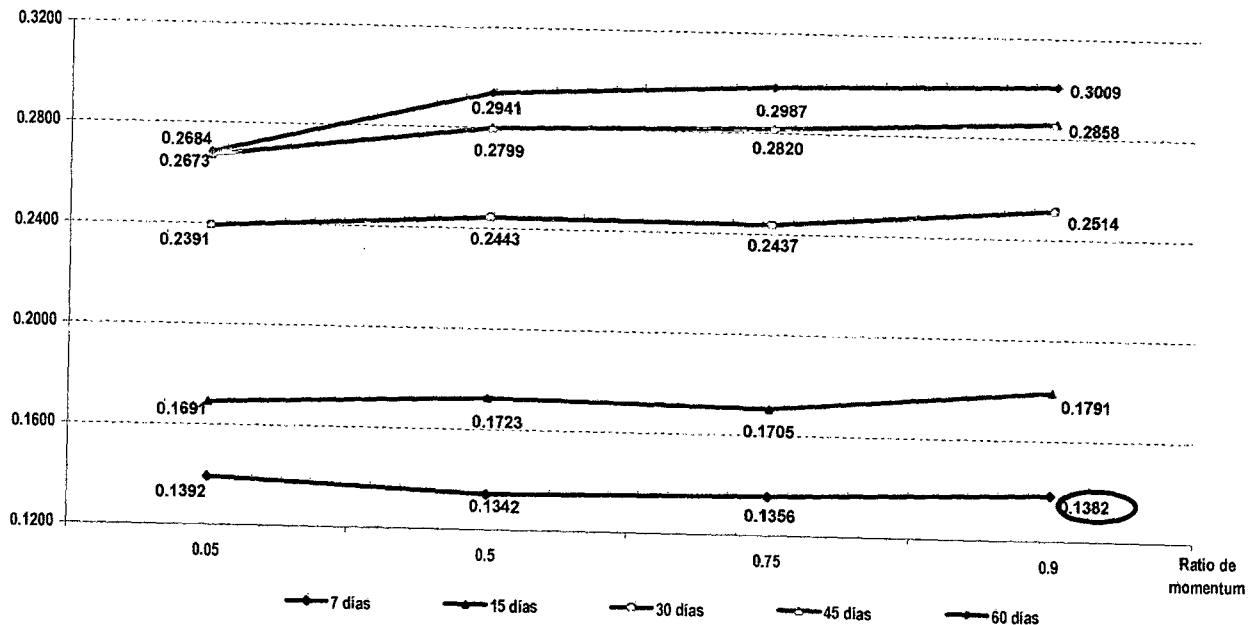


Gráfico 101 : Días Históricos - Cajero automático E
Fuente : Propia
Elaboración : Silvia Granados Flores

Los días históricos empleados para el pronóstico y que representan el menor error se detallan en la Tabla 57:

Cajero automático	Días Históricos	Error Promedio Mínimo
A	7	17.78%
B	7	15.21%
C	7	24.90%
D	7	22.06%
E	7	13.82%

Tabla 57 : Días históricos – Error promedio mínimo

Fuente : Propia

Elaboración : Silvia Granados Flores

En la Tabla 58 se representan los valores óptimos que se han identificado hasta el momento:

VALORES ÓPTIMOS						
Cajero automático	Función de Transferencia	Función de Activación	Función de Aprendizaje	Ratio de Aprendizaje	Ratio Momentum	Días Históricos
A	tansig - logsig	logsig	traingda	0.05	0.05	7
B	tansig - logsig	tansig	traingdm	0.05	0.05	7
C	tansig - tansig	tansig	traingda	0.05	0.05	7
D	logsig - tansig	logsig	traingda	0.75	0.75	7
E	logsig - tansig	logsig	traingda	0.90	0.90	7

Tabla 58 : Valores óptimos de parámetros

Fuente : Propia

Elaboración : Silvia Granados Flores

8.10. NEURONAS OCULTAS

Los Gráficos 77, 78, 79, 80 y 81 muestran el comportamiento del error promedio resultado de la experimentación sobre los valores propuestos, los cuales hacen referencia proporcionalmente a los días históricos:

- Neuronas ocultas = Días históricos
- Neuronas ocultas = Días históricos/2
- Neuronas ocultas = Días históricos/4

En cada uno de los gráficos se ha resaltado el valor óptimo con respecto al identificado para el parámetro Días Históricos (ver Tabla 57).

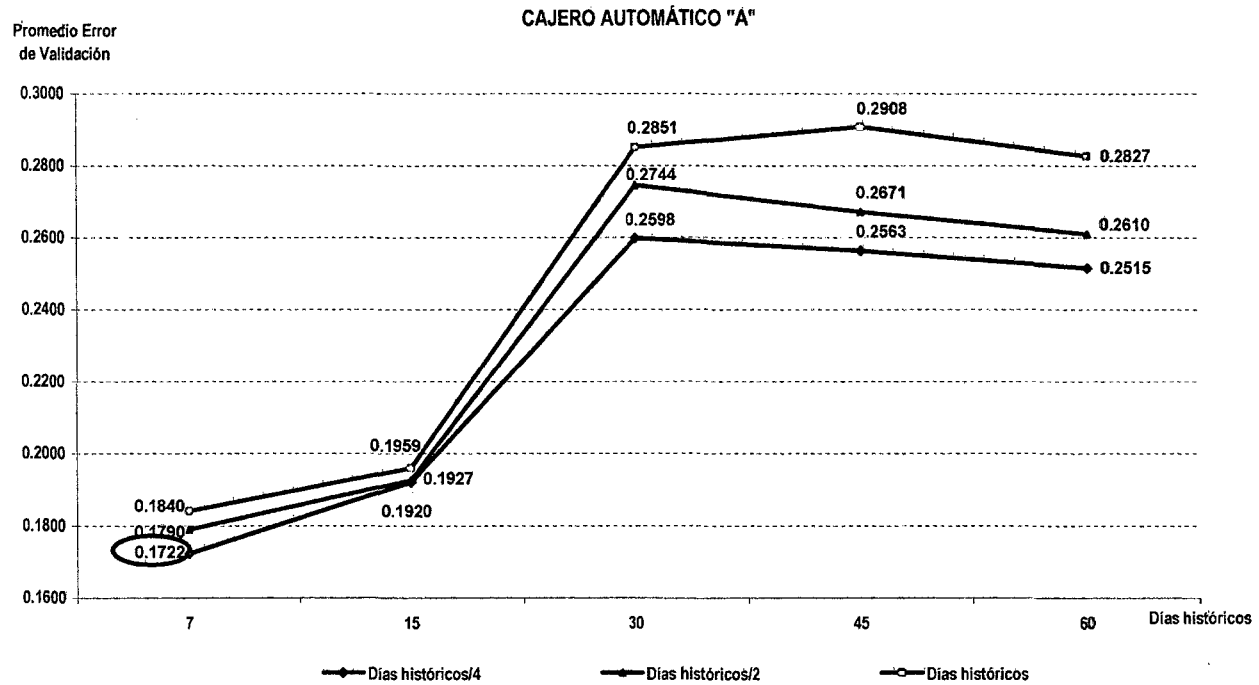


Gráfico 102 : Neuronas Ocultas - Cajero automático A
Fuente : Propia
Elaboración : Silvia Granados Flores

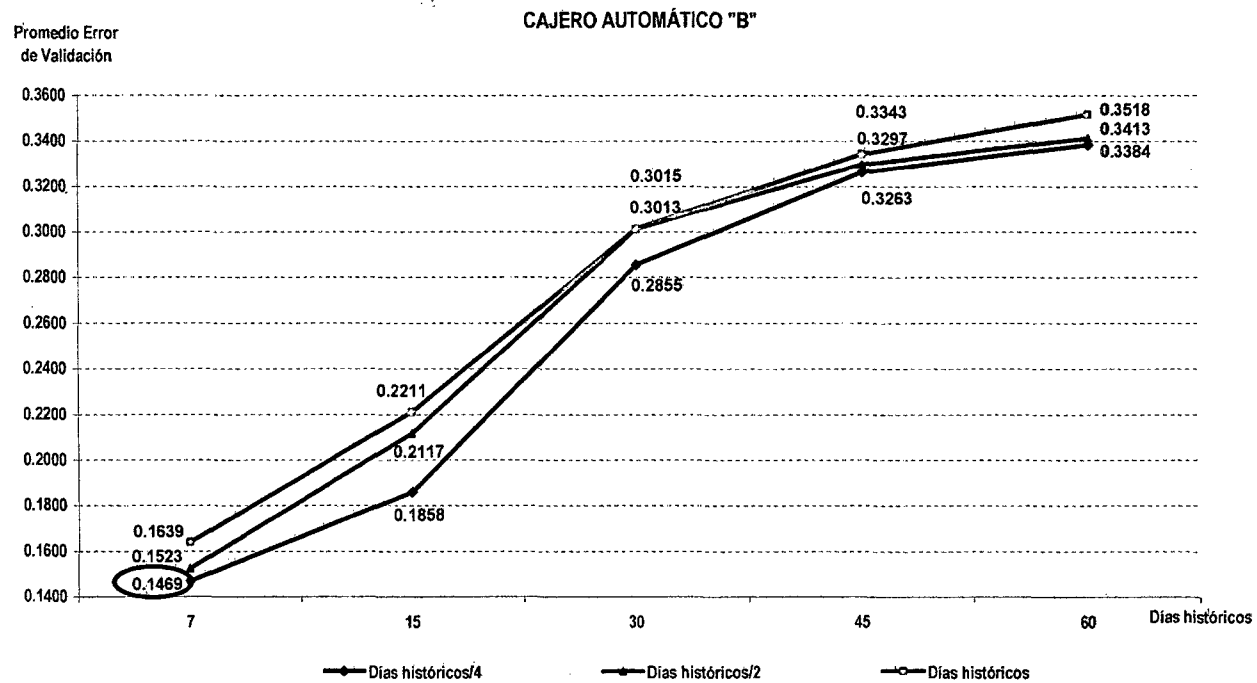


Gráfico 103 : Neuronas Ocultas – Cajero automático B
Fuente : Propia
Elaboración : Silvia Granados Flores

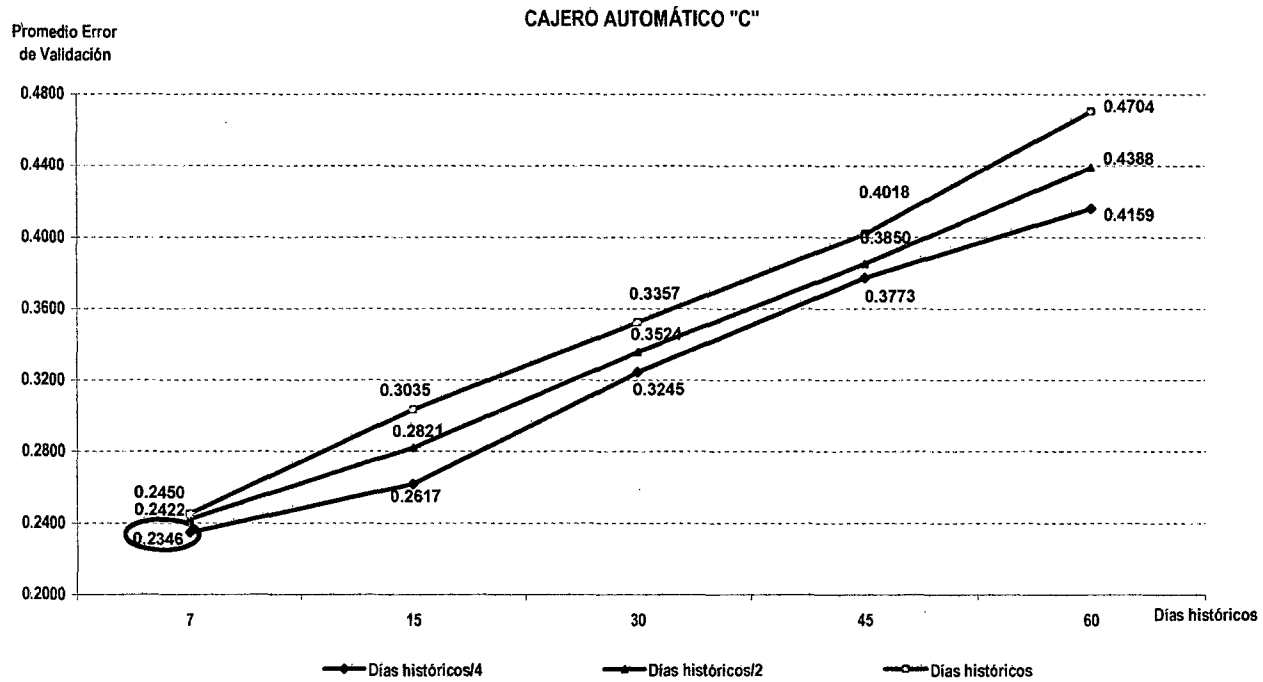


Gráfico 104 : Neuronas Ocultas – Cajero automático C
Fuente : Propia
Elaboración : Silvia Granados Flores

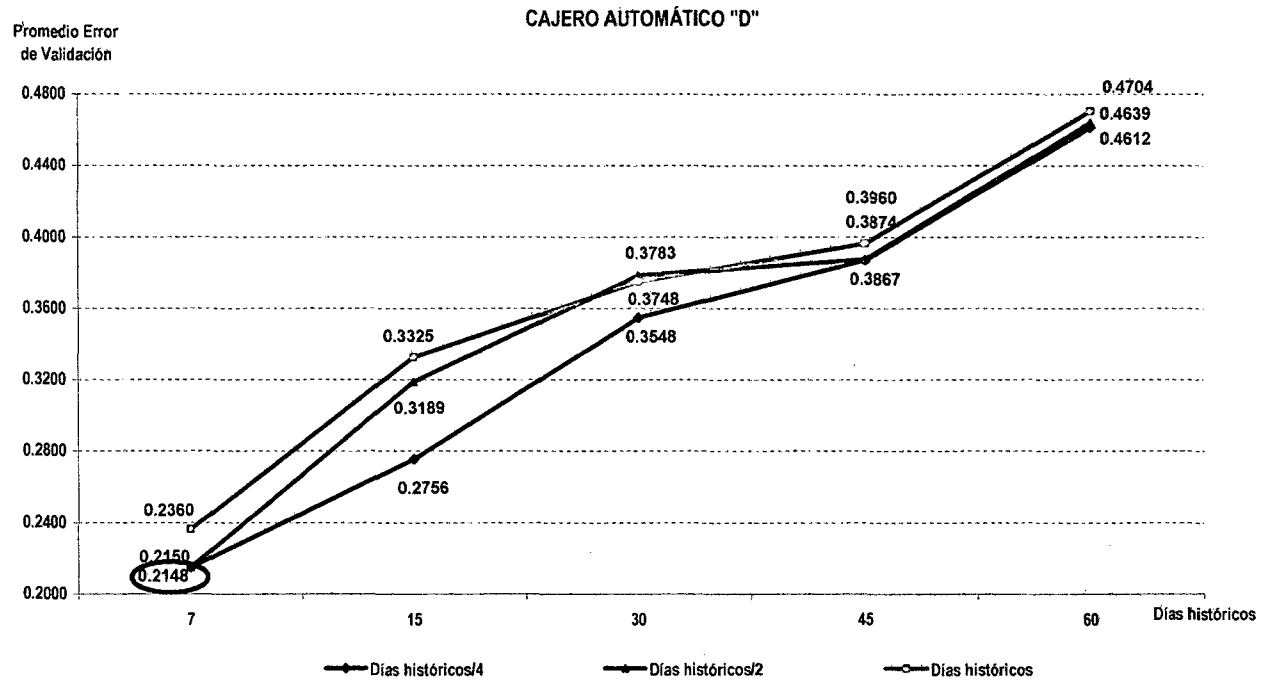


Gráfico 105 : Neuronas Ocultas - Cajero automático D
Fuente : Propia
Elaboración : Silvia Granados Flores

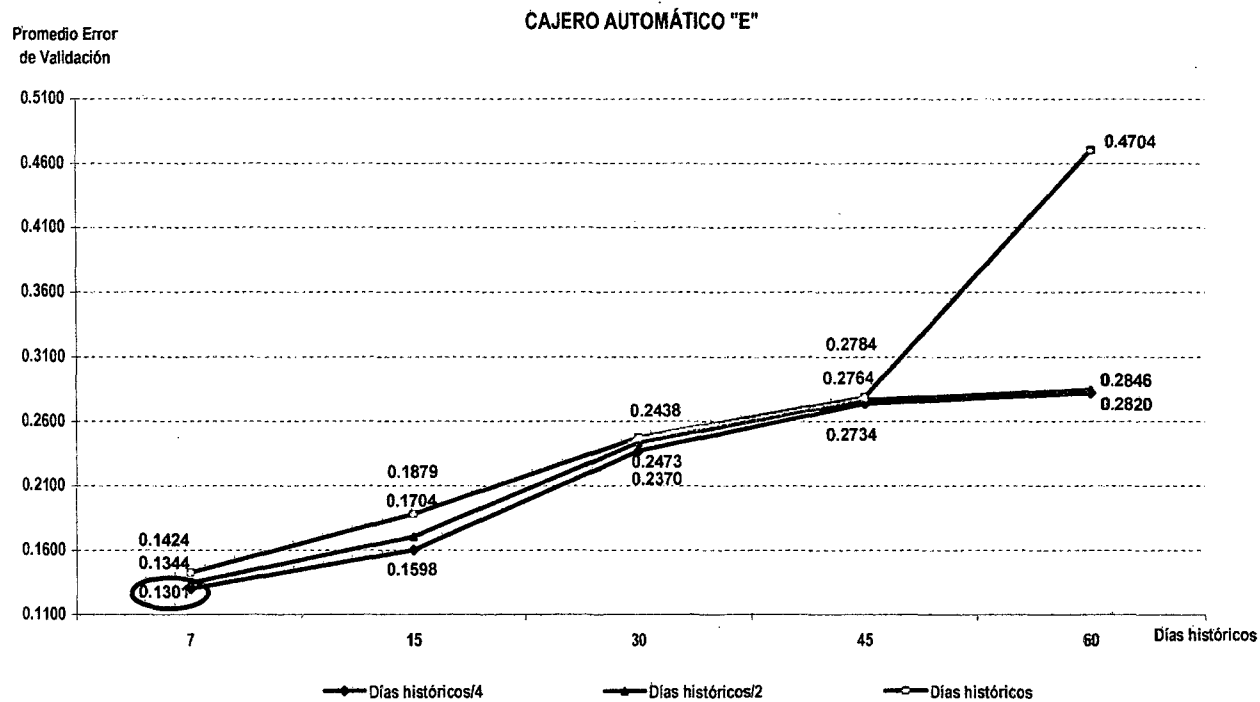


Gráfico 106 : Neuronas Ocultas - Cajero automático E
Fuente : Propia
Elaboración : Silvia Granados Flores

El error promedio mínimo que se obtuvo a través de la experimentación del parámetro en cada uno de los cajeros automáticos se detalla en la Tabla 59:

Cajero automático	Neuronas Ocultas	Error Promedio Mínimo
A	2	17.22%
B	2	14.69%
C	2	23.46%
D	2	21.48%
E	2	13.01%

Tabla 59 : Neuronas ocultas – Error promedio mínimo
Fuente : Propia
Elaboración : Silvia Granados Flores

En la Tabla 60 se representan los valores óptimos de los parámetros que se han identificado hasta el momento:

VALORES ÓPTIMOS							
Cajero automático	Función de Transferencia	Función de Activación	Función de Aprendizaje	Ratio de Aprendizaje	Ratio Momentum	Dias Historicos	Neuronas Ocultas
A	tansig - logsig	logsig	traingda	0.05	0.05	7	2
B	tansig - logsig	tansig	traingdm	0.05	0.05	7	2
C	tansig - tansig	tansig	traingda	0.05	0.05	7	2
D	logsig - tansig	logsig	traingda	0.75	0.75	7	2
E	logsig - tansig	logsig	traingda	0.90	0.90	7	2

Tabla 60 : Valores óptimos de parámetros
Fuente : Propia
Elaboración : Silvia Granados Flores

8.11. DÍAS PRONOSTICADOS

Los Gráficos 82, 83, 84, 85 y 86 muestran el comportamiento del error promedio resultado de la experimentación realizada a fin de definir el número de neuronas de salida. Los valores propuestos son 1, 3, 5 y 7 días a pronosticar.

En cada uno de los gráficos se ha resaltado el valor óptimo con respecto al identificado para el parámetro Neuronas Ocultas (ver Tabla 59).

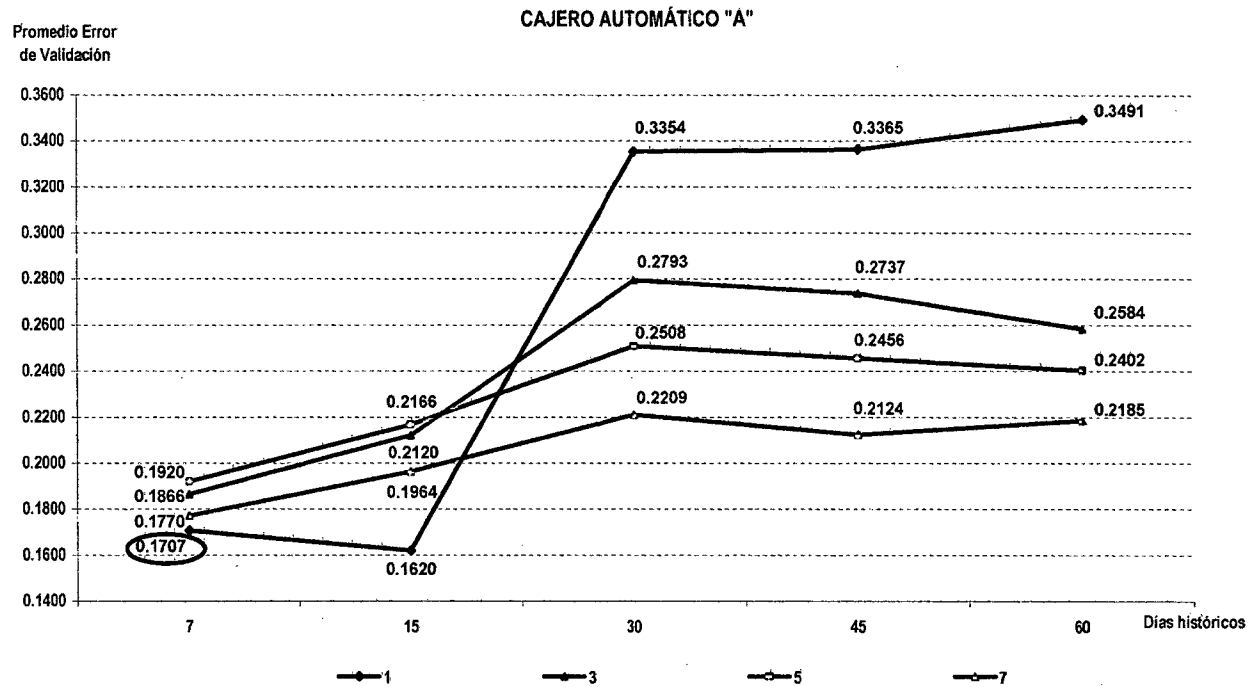


Gráfico 107 : Días Pronosticados – Cajero automático A
Fuente : Propia
Elaboración : Silvia Granados Flores

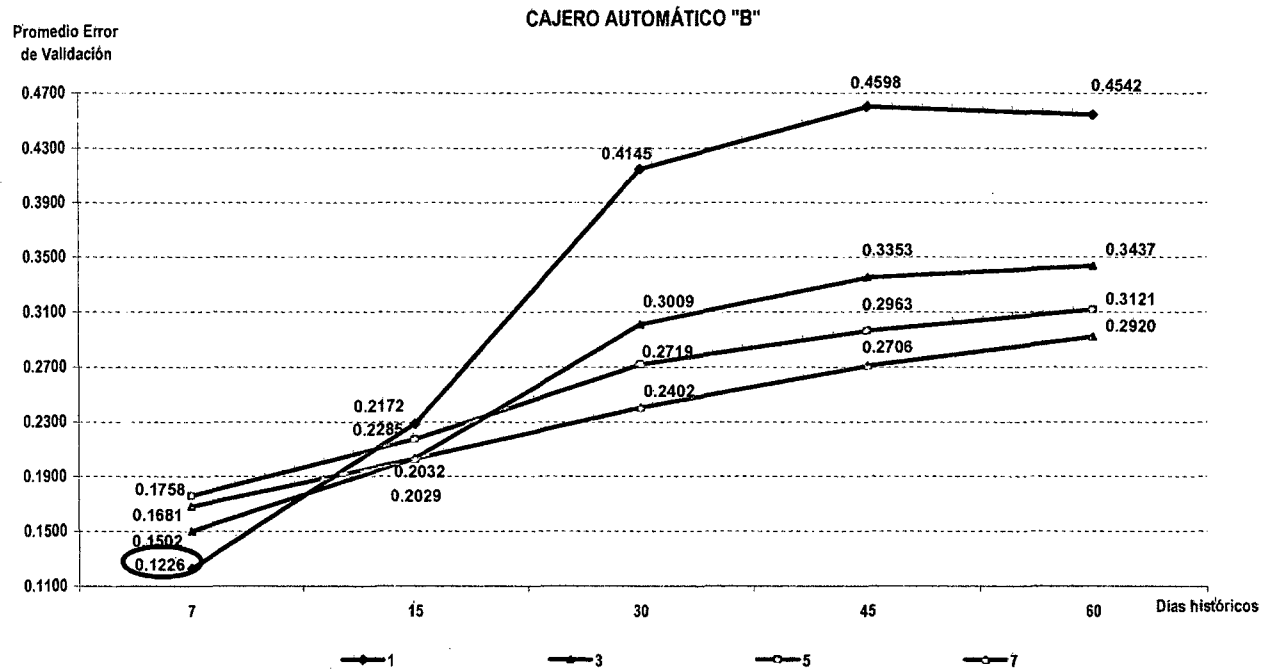


Gráfico 108 : Días Pronosticados – Cajero automático B
Fuente : Propia
Elaboración : Silvia Granados Flores

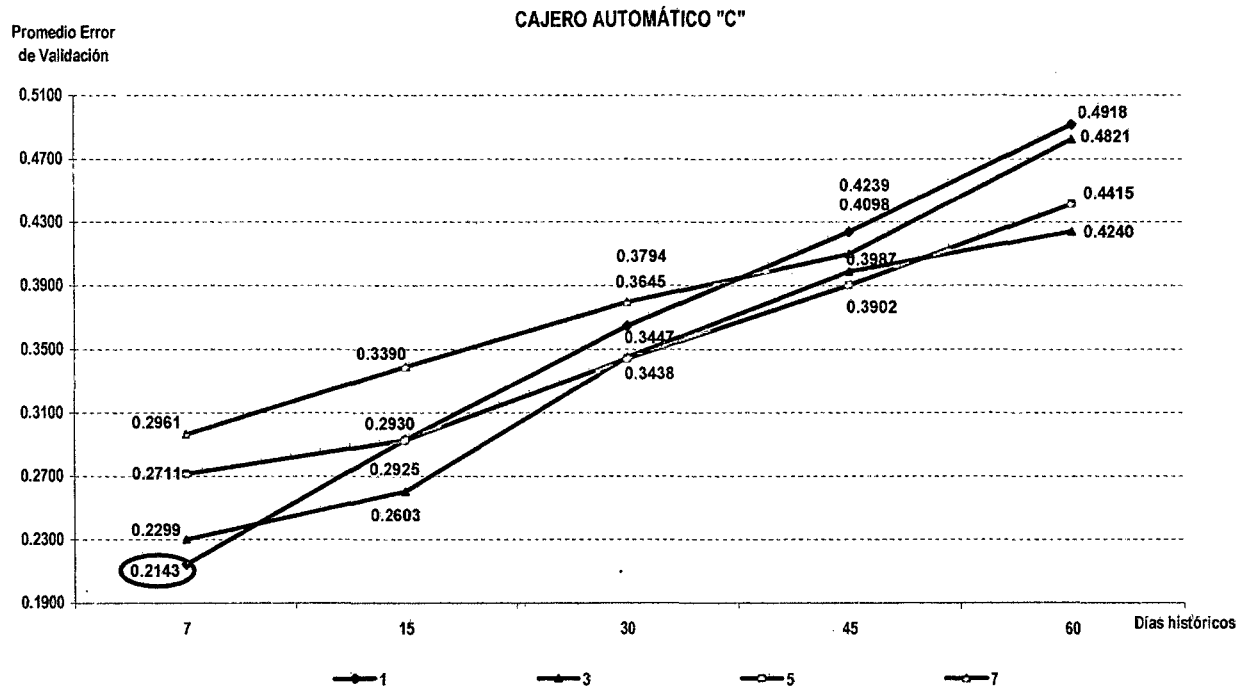


Gráfico 109 : Días Pronosticados – Cajero automático C
Fuente : Propia
Elaboración : Silvia Granados Flores

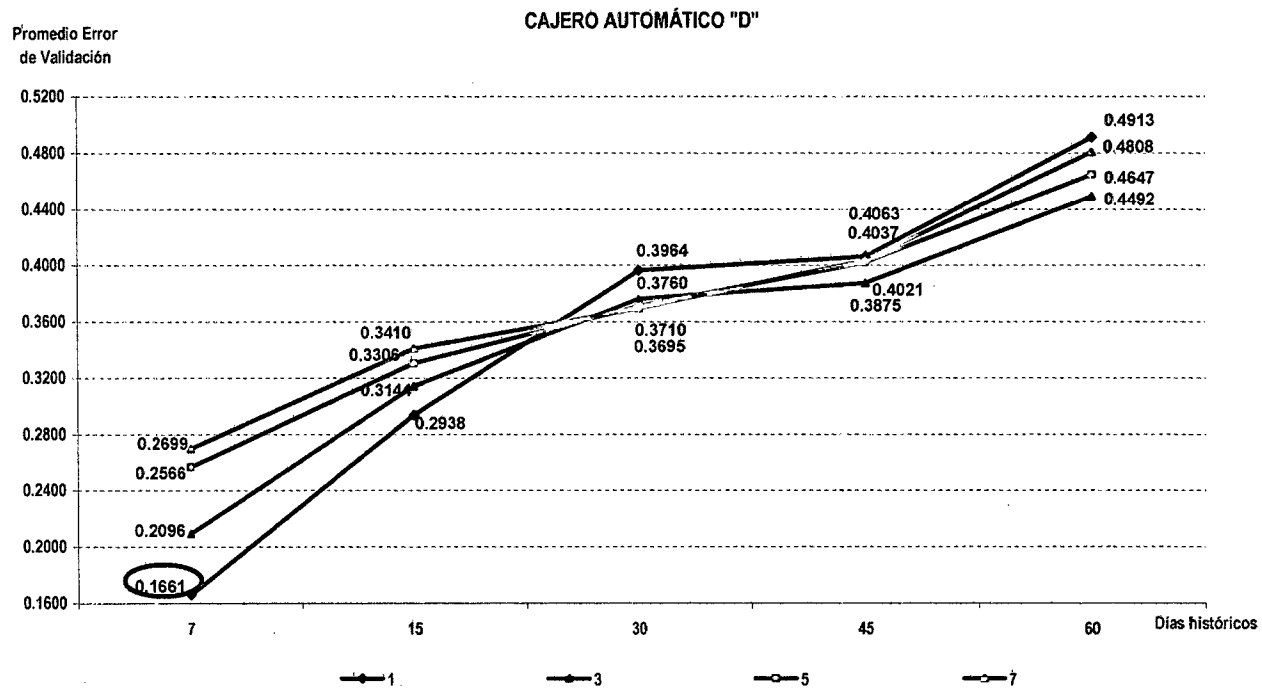


Gráfico 110 : Días Pronosticados – Cajero automático D
Fuente : Propia
Elaboración : Silvia Granados Flores

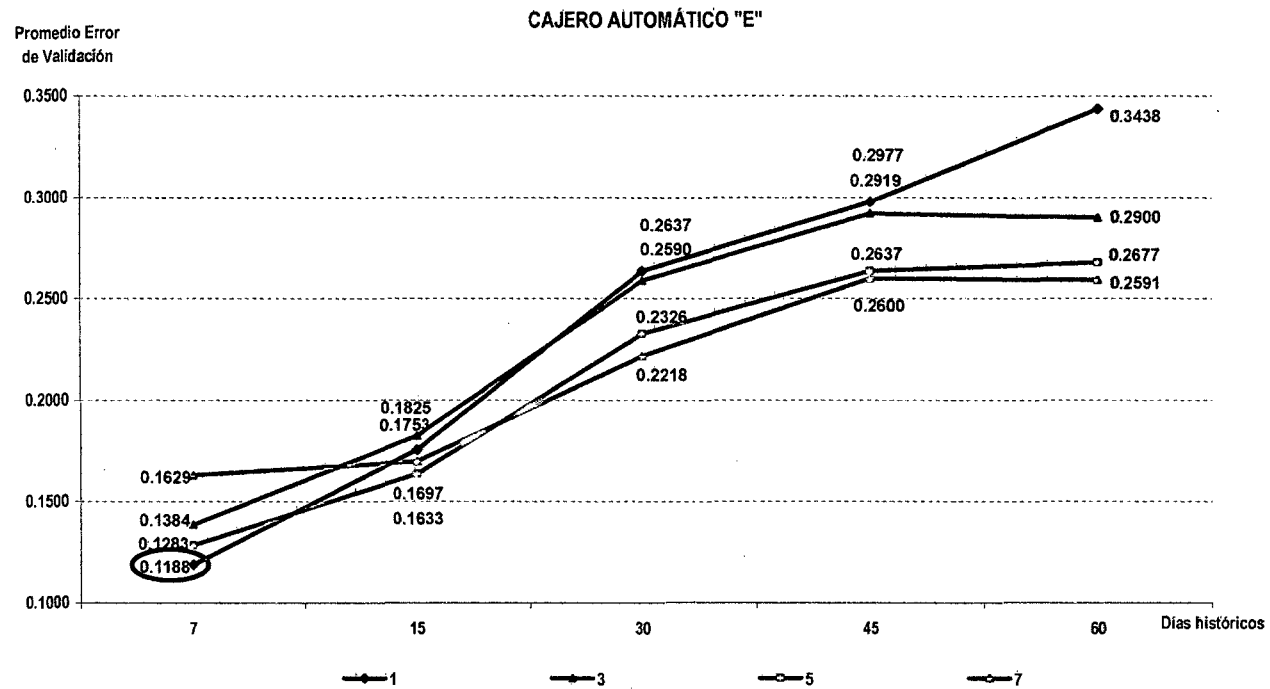


Gráfico 111 : Días Pronosticados – Cajero automático E
Fuente : Propia
Elaboración : Silvia Granados Flores

El valor propuesto que generó un error promedio menor a comparación del total para cada uno de los cajeros automáticos se detalla en la Tabla 61:

Cajero automático	Días Pronosticados	Error Promedio Mínimo
A	1	17.07%
B	1	12.26%
C	1	21.43%
D	1	16.61%
E	1	11.88%

Tabla 61 : Días pronosticados – Error promedio mínimo

Fuente : Propia

Elaboración : Silvia Granados Flores

En la Tabla 62 se presentan los valores óptimos de los parámetros que se han identificado hasta el momento:

VALORES ÓPTIMOS								
Cajero automático	Función de Transferencia	Función de Activación	Función de Aprendizaje	Ratio de Aprendizaje	Ratio Momentum	Días Historicos	Neuronas Ocultas	Días Pronosticados
A	tansig - logsig	logsig	traingda	0.05	0.05	7	2	1
B	tansig - logsig	tansig	traingdm	0.05	0.05	7	2	1
C	tansig - tansig	tansig	traingda	0.05	0.05	7	2	1
D	logsig - tansig	logsig	traingda	0.75	0.75	7	2	1
E	logsig - tansig	logsig	traingda	0.90	0.90	7	2	1

Tabla 62 : Valores óptimos de parámetros

Fuente : Propia

Elaboración : Silvia Granados Flores

8.12. ERROR FINAL DE VALIDACIÓN

En base a las experimentaciones realizadas, se identificaron los valores óptimos para los parámetros del modelo propuesto.

En base a los valores identificados para el modelo, en la Tabla 63 se presenta el error de validación obtenido al conjugarlos en el modelo propuesto. Cabe precisar que se trata de un error promedio considerando que se está trabajando con tres carteras de datos (ver punto 8.2)

ERROR FINAL DE VALIDACION

Cajero automático	Función de Transferencia	Función de Activación	Función de Aprendizaje	Ratio de Aprendizaje	Ratio Momentum	Días Históricos	Neuronas Ocultas	Días Pronosticados	Error de Validación
A	tansig - logsig	logsig	traingda	0.05	0.05	7	2	1	10.47%
B	tansig - logsig	tansig	traingdm	0.05	0.05	7	2	1	9.49%
C	tansig - tansig	tansig	traingda	0.05	0.05	7	2	1	15.84%
D	logsig - tansig	logsig	traingda	0.75	0.75	7	2	1	10.74%
E	logsig - tansig	logsig	traingda	0.90	0.90	7	2	1	7.32%

Tabla 63 : Error promedio final de validación

Fuente : Propia

Elaboración : Silvia Granados Flores

8.13. RESUMEN DEL CAPÍTULO

En este capítulo se determinaron los valores óptimos para los parámetros establecidos en el modelo de redes neuronales artificiales construido.

El modelo fue implementado en MATLAB ® Release 2009a realizándose 138,240 corridas al sistema para cada uno de los cajeros automáticos con la combinación de los valores propuestos para los parámetros analizados.

Estas corridas permitieron que a través de la evaluación del modelo, se puedan identificar los valores óptimos para los 8 parámetros establecidos en el modelo.

El método para identificar el valor óptimo de cada parámetro fue establecer una trazabilidad entre cada ítem analizado. De esta manera el valor que optimiza un primer parámetro, sirve como referencia para la identificación del valor que genere el menor error promedio en el siguiente.

Finalmente con la identificación de los valores óptimos, se calculó el error de validación que se genera con la implementación de estos valores en el modelo. De esta manera, se obtuvieron valores que oscilan entre 7% y 15% (ver Tabla 63).

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

En base a los resultados obtenidos a través de este análisis, se formulan las siguientes conclusiones:

1. El cálculo de la demanda de efectivo a nivel de número de billetes, establecido a partir de la información recolectada sobre stock de billetes y pedidos de abastecimiento durante el periodo analizado, se acopla a la metodología de Gestión de Inventarios donde: $\text{Stock inicial (stock en momento de corte inicial)} + \text{Compras (pedidos de abastecimiento)} - \text{Stock final (stock en momento de corte final)} = \text{Consumo (demanda de efectivo)}$.
2. El momento y la frecuencia de corte para establecer la unidad de tiempo en la que se medirá la demanda de efectivo, se basa en la calidad de la información recolectada. De esta manera, al establecer 8:45a.m. como hora de corte diario tanto para el reporte de stock de billetes como el referido a los pedidos de abastecimiento, se obtuvieron 1.92% registros anómalos (imprecisos o vacíos) del total recolectados, frente a 9.04% y 15.07% que se hubieran obtenido al establecerse las 12:00p.m. y 5:00p.m. respectivamente como momento de corte.

3. El pronóstico de la demanda de efectivo para cada uno de los tres tipos de billetes dentro del cajero automático, establece tres variables para el modelo sobre las cuales la Red Neuronal Artificial tendría que analizar cada uno de sus comportamientos y poder llegar así a una generalización. Al valorizar monetariamente la demanda de efectivo a nivel cajero automático se acota el análisis a una única variable.
4. Para los cinco cajeros automáticos incluidos en la investigación se detectaron un promedio de 4 a 6 incidentes en el mes, los cuales interrumpieron por un lapso de tiempo la disponibilidad del servicio ofrecido por el ATM. Por este motivo, la confiabilidad sobre la exactitud de la demanda de efectivo calculada se ve impactada.
5. El establecimiento de un patrón de medición del impacto sobre las medidas de tendencia central (media y mediana) permitió identificar las mejores técnicas para el reemplazo de las anomalías identificadas en la información (valores imprecisos, vacíos y outliers), las cuales de no ser así se hubiesen introducido en la red neuronal y hubiesen impedido una generalización óptima.
6. El 0.83% de los registros recolectados sobre stock de billetes se refieren a valores imprecisos. Al efectuar el reemplazo de estos valores a través del Método de Proporción Lineal, las medidas de tendencia central (media y mediana) se vieron impactadas entre 0% y 0.73%. Lo que permite concluir que la desviación generada a través del reemplazo no es trascendente.
7. El 1.11% de los registros obtenidos sobre stock de billetes se refieren a valores vacíos. Estos valores fueron reemplazados por el Método de Interpolación Lineal, generando una desviación del

rango entre 0.34% y 1.44% sobre las medidas de tendencia central (media y mediana) de los datos recolectados. Lo que permite concluir que la desviación generada a través del reemplazo no es trascendente.

8. Para identificar los valores outliers dentro del comportamiento de la variable demanda de efectivo, fue necesario reconocer si la disposición de estos valores correspondían a una distribución simétrica o sesgada hacia alguno de los lados a fin de establecer el límite inferior y superior de las variables.
9. Entre el 1.10% y 2.47% de los registros calculados sobre demanda de efectivo se refieren a valores outliers. Estos valores fueron reemplazados por los límites superiores o inferiores según corresponda la ubicación del valor atípico, originando una desviación entre 0.79% y 3.22% de las medidas de tendencia central (media y mediana). Lo que permite concluir que la desviación generada a través del reemplazo no es trascendente.
10. La arquitectura de la red neuronal NARX establece regresores tanto en las entradas como en las salidas. Por este motivo, el error se propaga hacia las capas anteriores hasta llegar a la capa de entrada para de esta manera ajustar los pesos de las conexiones entre neuronas. Asimismo dentro de la estructura de la red NARX se establece un tiempo de retraso entre las entradas con respecto a la salida de la red. Para el modelo propuesto, el tiempo de retraso se estableció en base al número de días históricos que justifican el comportamiento de los días pronosticados.
11. En base a la arquitectura del modelo propuesto sobre red neuronal NARX, el número de días históricos multiplicado por tres (incluyendo esta variable, Día Semana y Día Laborable)

determinan la dimensión de la capa de entrada. Asimismo el número de días a pronosticar define el número de neuronas de la capa de salida del modelo.

12. A fin de reducir la distancia entre los valores empleados para el aprendizaje de la red neuronal pero sin impactar la distribución de las variables, se efectuó la normalización de los datos a través de la función logaritmo neperiano.
13. Se identificaron ocho parámetros dentro del modelo propuesto. Los valores propuestos para estos parámetros fueron determinados a partir de trabajos de investigación previos y en los se consideran como aplicables para la naturaleza del problema. De esta manera cada uno de los ocho parámetros tenía por lo menos 3 valores propuestos y se identificaron valores óptimos que permitieron hallar errores de validación que oscilan en 7% y 15%.
14. En base a la experimentación realizada, se identifica que el error MAPE del modelo incrementa a razón del número de días pronosticados.

RECOMENDACIONES

Como resultado al desarrollo de esta investigación, se generan las siguientes recomendaciones:

1. Sobre los datos recolectados se recomienda obtener información acerca de los factores que generan anomalías en el comportamiento de las variables así como sus ocurrencias; es decir, obtener información sobre número de incidencias, frecuencia, tiempo de ocurrencia, causa del incidente y solución efectuada. Esto con la finalidad de estimar el impacto.

2. Las Redes Neuronales Artificiales se han convertido en una herramienta potente para la solución de problemas complejos y los enfocados a problemas empresariales están evolucionando; sin embargo se recomienda a las organizaciones implementar Sistemas de Información que proporcionen información íntegra y de calidad, para de esta manera potenciar los beneficios que proporciona esta herramienta.
3. Sobre la arquitectura de la Red Neuronal Artificial cabe recordar que en caso se seleccionara un número insuficiente de neuronas en cada capa, el sistema tendrá menor capacidad para generalizar el comportamiento de todos los puntos. Por el contrario, si este número fuera determinado en exceso ocasionaría un sobreajuste en el tratamiento y generaría una menor performance en el pronóstico. Se recomienda que dentro de una investigación de este tipo, en la etapa de experimentación se invierta un mayor tiempo para el análisis y los recursos a emplearse.
4. Se recomienda seleccionar el tipo de modelo de Red Neuronal Artificial en base a la naturaleza del problema y la información disponible sobre las variables que se consideren dentro de la solución.
5. Para identificar los valores atípicos dentro de los datos considerados en el modelo de Redes Neuronales Artificiales, se recomienda reconocer las características estadísticas de las variables como medidas de tendencia central, distribución e intervalos de confianza.
6. Se recomienda determinar los posibles valores de los parámetros identificados en el modelo de Redes Neuronales Artificiales a

través de estudios preliminares realizados y en base a los que se indiquen como aplicables para el tipo de modelo empleado.

GLOSARIO DE TÉRMINOS

ARIMA: Autoregressive integrated moving average (Modelo autorregresivo integrado de media móvil).

ABASTECIMIENTO: Conjunto de actividades que permite identificar y adquirir los bienes y servicios que se requieren para la operación, ya sea de fuentes internas o externas.

ARQUITECTURA: Conjunto de elementos adecuadamente estructurados dentro de un sistema.

ATM: Automatic Teller Machine

CALIDAD: Habilidad que posee un producto, componentes de producto o un proceso para cumplir con los requerimientos de los clientes.

CAPA: Nivel dentro de una estructura de red que proporciona atributos y conectividad con los siguientes niveles.

CRISP-DM: Cross Industry Standard Process for Data Mining

DEMANDA: Manifestación de voluntad de adquisición de un número de unidades de un recurso a un precio dado en el mercado.

ERROR: Medida del error cometido al aproximar una magnitud numérica por una expresión aproximada a la expresión original exacta.

FUNCIÓN: Relación entre un conjunto dominio y otro conjunto rango de elementos, de forma que a cada elemento del dominio le corresponde un único elemento del rango.

IMPRECISO: Calificativo referido a un ente que no es exacto o es poco definido.

INCIDENTE: Suceso inesperado, no planeado y que impacte la actividad normal o esperada de un proceso.

INTERPOLACIÓN: Construcción de nuevos puntos en una función partiendo de un conjunto conocido de puntos.

MAPE: Mean absolute percentage error.

MATLAB: MATrix LABoratory. Software matemático que ofrece un entorno de desarrollo integrado con un lenguaje de programación propio.

METODOLOGÍA: Conjunto particular de procedimientos que se emplean para tratar un caso o situación.

NARX: Nonlinear AutoRegressive Exogenous.

NULL: Registro de información con valor vacío.

OUTLIER: Registro con valor atípico o anómalo.

PROCEDIMIENTO: Es el modo de ejecutar determinadas acciones que suelen realizarse de la misma forma a través de una serie de pasos.

PROCESO: Conjunto de actividades o eventos que se realizan o suceden con un determinado fin.

PRONÓSTICO: Proceso de estimación en situaciones de incertidumbre y que comúnmente se efectúa en series temporales.

RATIO: Relación entre dos elementos o conjuntos cuantitativos.

RELEASE: Nueva versión de una aplicación en informática.

RNA: Red Neuronal Artificial

SET DE PRUEBAS: Conjunto de casos que deben ser desarrollados para probar parte de la funcionalidad de un sistema de software.

STOCK: Cantidad de insumos existentes en una entidad.

TÉCNICA: Procedimiento o conjunto de reglas, normas o protocolos, que tienen como objetivo obtener un resultado determinado.

BIBLIOGRAFÍA

- [1] B. Krose, P. Van der Smagt. An Introduction to Neural Networks. Universidad de Amsterdam. Páginas 30-45, 1996.
- [2] P. Rossini. Using Expert Systems and Artificial Intelligence for Real Estate Forecasting. Sixth Annual Pacific-Rim Real Estate Society Conference Sydney, Australia. Páginas 2-3, 2000.
- [3] I. García Martín. Análisis y predicción de la serie de tiempo del precio externo del café colombiano utilizando redes neuronales artificiales. Facultad de Ciencias, Pontificia Universidad Javeriana. Páginas 46-47, 2003.
- [4] G. Lepecq. The Future of Cash: Survey Predicts Changes in Consumer Habits. Páginas 3-7, 2006
- [5] M. Gómez Martínez. Redes Neuronales Artificiales. IIIA-Instituto de Investigación en Inteligencia Artificial, Barcelona, España. Páginas 56-61, 2004.
- [6] L. Flores. Red Neuronal Artificial. Instituto Tecnológico de Nogales, área de Ingeniería en Sistemas Computacionales. Páginas 11-12, 2008.
- [7] P. Nicholls. 2010 ATM Software, Trends and Analysis. Páginas 16-18, 2009.
- [8] P. Chand K., E. Walia. Cash Forecasting: An Application of Artificial Neural Networks in Finance. Páginas 61-63, 2006.

- [9] J. de Andrés S. Dos Aplicaciones Empíricas de las Redes Neuronales Artificiales a la Clasificación y la Predicción Financiera en el Mercado Español. Páginas 63-71, 2003.
- [10] O. Martín, M. López, F. Martín. Redes neuronales artificiales para la predicción de la calidad en soldadura por resistencia por puntos. Páginas 346-347, 2006.
- [11] R. Poler, J. Mula, D. Peidro. Un Método de Selección de Modelos de Previsión de Demanda. Páginas 2-5, 2007.
- [12] Palmer, M. Beltrán, J. Montaña, R. Jimenez, A. Sesé, J. Franconetti. Minería de datos en Economía. Aplicaciones de redes neuronales artificiales a datos de la Encuesta de gasto turístico de las Islas Baleares. Páginas 4-7, 2007.
- [13] I. Kaastra, M. Boyd. Designing a neural network for forecasting financial and economic time series. Páginas 215-236, 1996.
- [14] P. García E. Aplicaciones de las Redes Neuronales en las Finanzas. Páginas 5-6, 2002.
- [15] J. Navarro P., J. Losilla V. Análisis de Datos Faltantes mediante Redes Neuronales Artificiales. Página 505, 2000.
- [16] V. Cassino, P. Misich, J. Barry. Forecasting the demand for currency. Páginas 29-30, 1997.
- [17] F. Segovia. Redes de Neuronas Recurrentes Para el Reconocimiento de Patrones Temporales. Páginas 96-122, 1992.
- [18] A. González, M. Ávila, E. Gómez, X. Vilasis, O. Mulet, F. Mazzanti. Redes Neuronales para Identificación y Predicción de Series de Tiempo. Páginas 48-49, 2000.
- [19] L. Useche, D. Mesa. Una Introducción a la Imputación de Valores Perdidos. Páginas 11-14, 2006.

- [20] A. Jain, J. Mao. Artificial Neural Networks: A tutorial. Páginas 31-43, 1996.
- [21] N. Davey, S. P. Hunt, R. J. Frank. Time Series Prediction and Neural Networks. Páginas 1-2, 2004.
- [22] H. Snellman. Automated Teller Machine network market structure and cash usage. Páginas 53-58, 2006.
- [23] D. Bounie, N. Houy. A Model of Demand for Cash and Deposits. Páginas 5-8, 2007.
- [24] D. Iranzo. Análisis de outliers: Un caso de estudio. Páginas 7-11, 2007.
- [25] I. Leal, P. Melín. Predicción de las Series de Tiempo del Precio del Tomate en México utilizando Redes Neuronales Modulares y Procesamiento en Paralelo. Páginas 7-14, 2003.
- [26] D. Mandic, J. Chambers. Recurrent Neural Networks for Prediction. 41-44, 2001.
- [27] M. Hudson, M. Hagan, H. Demuth. Neural Network Toolbox™ 7 User's Guide. Páginas 195-200, 2010.
- [28] Vensim (2001). Ventana Systems, Inc. <http://www.vensim.com/>
- [29] Simulink (2001). The MathWorks, Inc. <http://www.mathworks.com/>
- [30] Stella & Ithink (2001). High Performance Systems, Inc. <http://www.hps-inc.com>