

5° CONVENCION NACIONAL DE PROFESIONALES EN COBRANZA
5th NATIONAL CONVENTION OF COLLECTING PROFESSIONALS



ASOCIACION DE
PROFESIONALES EN
COBRANZA
Y SERVICIOS
JURIDICOS AC



ConveCob2009

La Profesionalización en la Cobranza
Professionalization of Collections

Redes Neuronales y su Aplicación en Modelos de Cobranza

Dr. Viterbo H. Berberena González

Profesor Ciencias Analíticas

Consultor Senior en Inteligencia Analítica



NOVIEMBRE 12 y 13 | CENTRO BANAMEX | MÉXICO D.F.

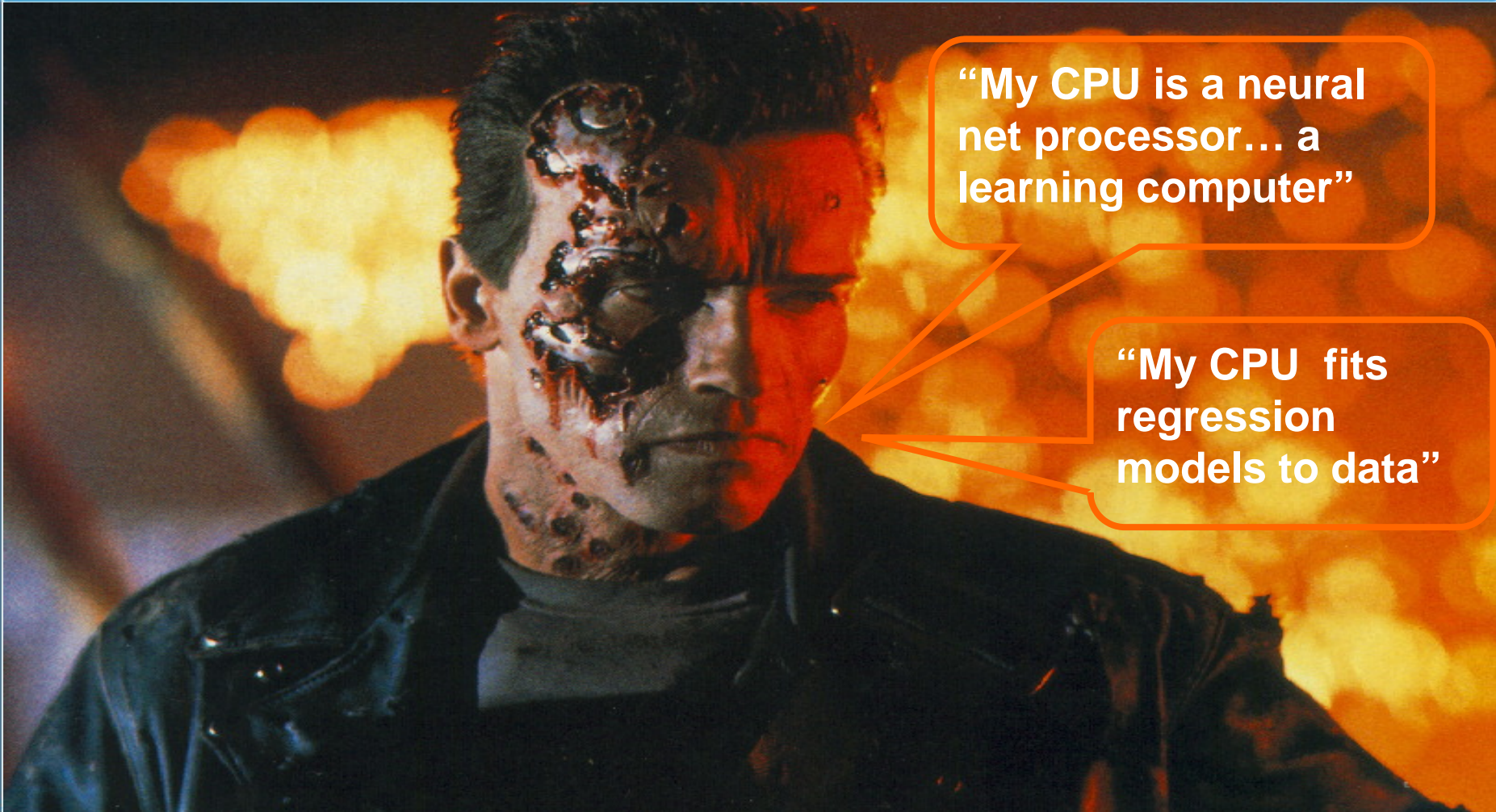


Temas a tratar:

- El desafío analítico de estimar una función multivariada.
- Una aproximación a las redes neuronales artificiales.
- Hipótesis fundamental en la investigación de patrones en bases de datos con redes neuronales.
- Tipos de aplicación de las redes neuronales en los modelos de riesgo.
- Impacto estratégico de los modelos de riesgo en las empresas.
- Características generales de los modelos de riesgo.
- Tipos de modelos de riesgo de crédito.
- Construcción y utilización de los modelos de riesgo con redes neuronales.
- Benchmarking de los modelos de redes neuronales.
- Un modelo versus varios modelos del portafolio de clientes.
- Calibración de los modelos de riesgo de crédito con redes neuronales.
- Modelación holística del riesgo en las organizaciones.



¿Realidad o Ficción?



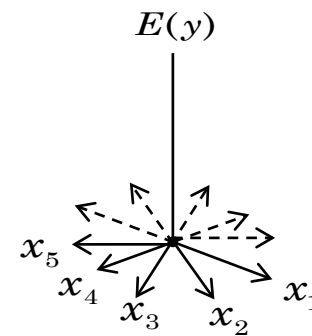
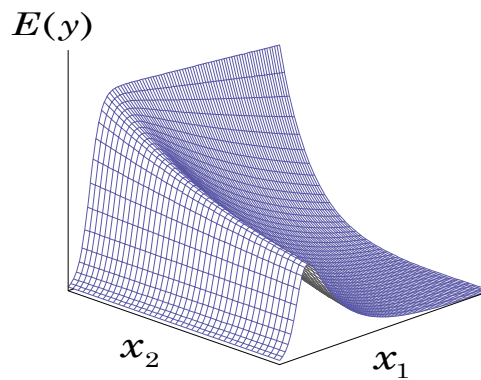
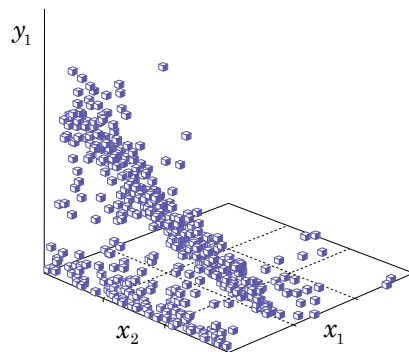
“My CPU is a neural net processor... a learning computer”

“My CPU fits regression models to data”



El Desafío Analítico de Estimar una Función Multivariada

- Métodos de Estimación de Funciones Multivariadas
 - Artificial Neural Networks (Rumelhart, 1986)
 - Projection pursuit regression (Freidman and Stuetzle, 1981)
 - Thin-plate splines (Wahba, 1990)
 - Multivariate adaptive regression splines (Freidman, 1991)





Las Redes Neuronales Artificiales

- Desarrolladas originalmente para intentar imitar la neurofisiología del cerebro humano y reproducir fenómenos complejos tales como la inteligencia.
- Existe una controversia sobre si las redes neuronales son realmente inteligentes, sin embargo de lo que no hay duda es de que han evolucionado en modelos estadísticos muy útiles.
- Se pueden considerar estimadores flexibles universales de funciones multivariadas.
- En general se puede clasificar como modelos de regresión paramétricos no lineales.



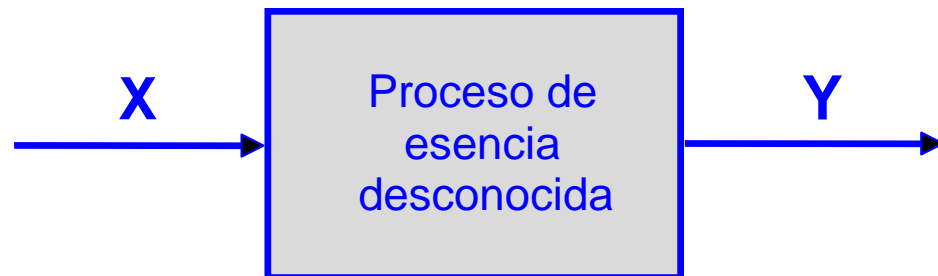
Condiciones para la Aplicación de las Redes Neuronales

- Las redes neuronales artificiales son especialmente útiles para la modelación predictiva cuando:
 - Se desconoce la formulación matemática de la función de respuesta (patrón de comportamiento) que relaciona las variables de entrada o predictores con la variable de salida o variable objetivo (principio de caja negra).
 - La predicción es más importante que la explicación de fenómeno en estudio (la estructura de los modelos reales de redes neuronales son muy complejos y difíciles de entender)
 - Hay suficientes datos para entrenar la red y encontrar una regularidad matemática general en la base.



Las Redes como Modelos Estocásticos

Principio de Caja Negra



donde,

X - es el vector de variables de entrada (que incluye las variables controlables),

Y - es el vector de variables de salida, y el rectángulo o caja negra es el proceso cuya esencia interna se desconoce exactamente.

Modelo Estocástico

$$y = f(p, x)$$

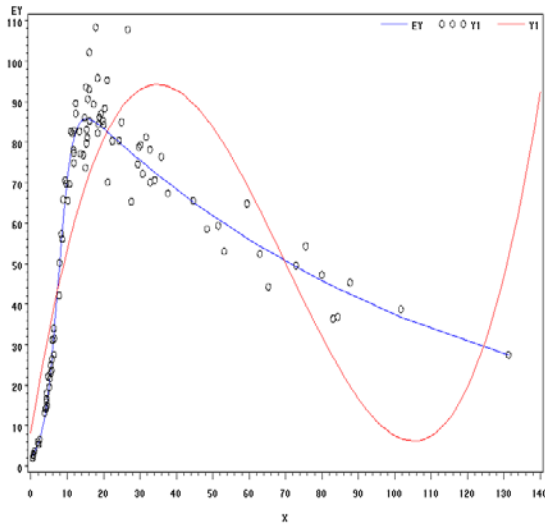
donde,

f - Contiene la información del método que se utiliza para su obtención.

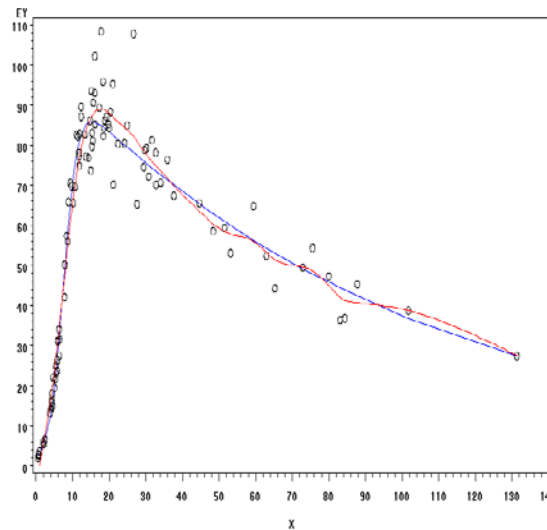


Criterio de Optimalidad de los Modelos de Redes

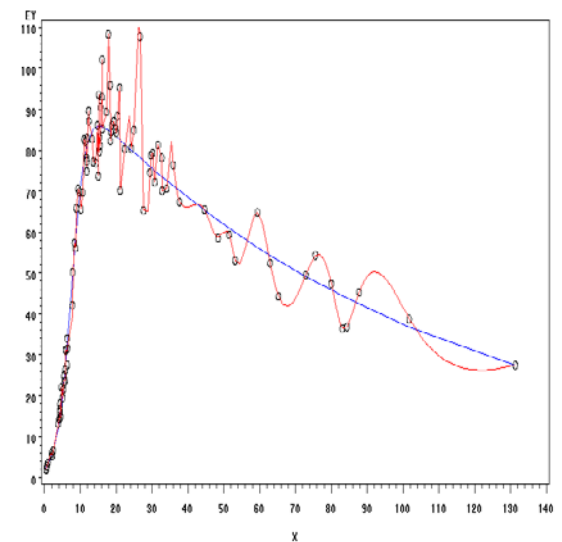
Falta Ajuste



Buen Ajuste



Sobre Ajuste



Criterio de Optimalidad → Minimizar la Falta de Ajuste



La Varianza Explicada por el Modelo Red Neuronal

Datos de Entrenamiento



Datos de Validación

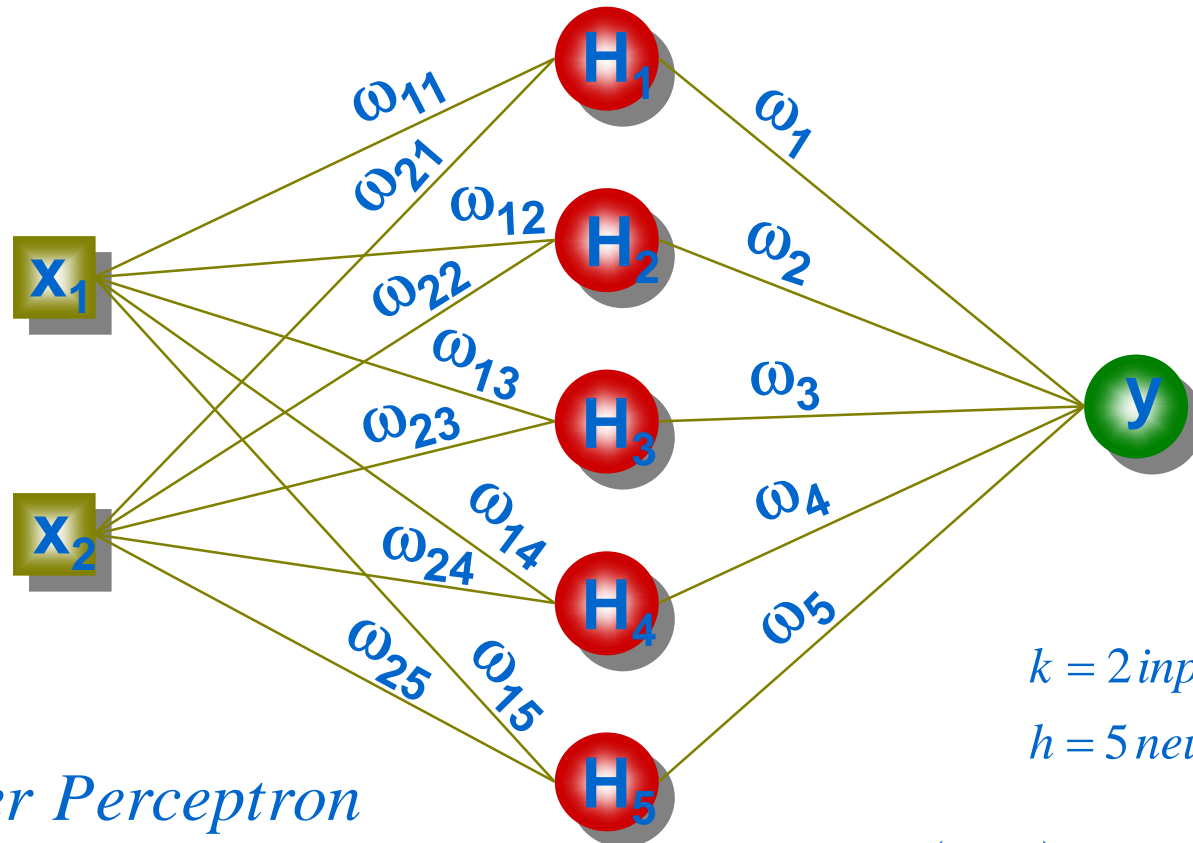


$$VarGen_{Entren} \cong VarGen_{Valid}$$

$$VarEsp_{Entren} \neq VarEsp_{Valid}$$



Arquitectura de una Red Neuronal Artificial Simple



Multilayer Perceptron

$k = 2$ inputs

$h = 5$ neurons

$$\#Coef = h(k + 2) + 1 = 5(2 + 2) + 1 = 21$$



Modelo de la Red Neuronal Artificial

$$g_0^{-1}[E(y)] = \omega_0 + \omega_1 H_1 + \omega_2 H_2 + \omega_3 H_3 + \omega_4 H_4 + \omega_5 H_5$$

$$H_1 = g_1(\omega_{01} + \omega_{11}x_1 + \omega_{21}x_2)$$

$$H_2 = g_2(\omega_{02} + \omega_{12}x_1 + \omega_{22}x_2)$$

$$H_3 = g_3(\omega_{03} + \omega_{13}x_1 + \omega_{23}x_2)$$

$$H_4 = g_4(\omega_{04} + \omega_{14}x_1 + \omega_{24}x_2)$$

$$H_5 = g_5(\omega_{05} + \omega_{15}x_1 + \omega_{25}x_2)$$



Muestra de Datos de Entrenamiento de la Red

	X1	X2	Y1
1	3.9366972053	0.0223795096	14.829487108
2	49.257916406	0.0365619645	16.273161089
3	16.192109679	0.0535658423	53.517753978
4	15.821026412	0.0633866744	43.401479783
5	22.4260905	0.0347430281	50.685108335
6	48.601270373	0.0190147518	42.499518666
7	19.652403337	0.0491321012	39.974782096
8	83.157692749	0.0585614325	0.9003420199
9	29.670866061	0.0103105038	71.150817172
10	52.234389622	0.0001575249	95.362840131
11	16.641904172	0.0244302409	63.269899387
12	16.691757812	0.04218461	56.297404505
13	2.920178612	0.0142847545	8.0317370351
14	11.23756655	0.0307505692	59.372377689
15	12.427484541	0.0318784088	81.493351133
16	24.301471839	0.0470338853	36.356215234
17	38.503333444	0.0199558235	55.306964355
18	8.5225218295	0.0238904535	46.737768819
19	2.8523041337	0.0097380895	7.9699568035
20	0.3053411869	0.0501312855	2.1728792599
21	0.9879643265	0.0348989277	3.9336410113
22	30.312692101	0.0755354159	12.2047084



Valores de los Coeficientes Calculados para la Red

	From	To	Weight	
1	X1	H11	0.0058062274	ω_{11}
2	X2	H11	-2.954682111	ω_{21}
3	X1	H12	-0.615838602	ω_{12}
4	X2	H12	-0.679401372	ω_{22}
5	X1	H13	1.2229899207	ω_{13}
6	X2	H13	-0.182951319	ω_{23}
7	X1	H14	-0.344952944	ω_{14}
8	X2	H14	1.4831246207	ω_{24}
9	X1	H15	-4.779230778	ω_{15}
10	X2	H15	-0.131563911	ω_{25}
11	BIAS	H11	-2.947032253	ω_{01}
12	BIAS	H12	0.0766158325	ω_{02}
13	BIAS	H13	1.6511786161	ω_{03}
14	BIAS	H14	4.1175374386	ω_{04}
15	BIAS	H15	-3.506551483	ω_{05}
16	H11	Y1	5.5892090066	ω_1
17	H12	Y1	35.38679607	ω_2
18	H13	Y1	-106.1462009	ω_3
19	H14	Y1	-126.1787386	ω_4
20	H15	Y1	-62.22003288	ω_5
21	BIAS	Y1	211.46381386	ω_0

Combination function weights (argument of activation functions)

Combination function biases (argument of activation functions)

Inverse of the output activation function weights and bias



Geometría de la Red Neuronal Artificial

$$g_0^{-1} \left(\text{Surface 1} \right) = w_0 + w_1 \cdot \text{Surface 2} + w_2 \cdot \text{Surface 3} + w_3 \cdot \text{Surface 4} + w_4 \cdot \text{Surface 5} + w_5 \cdot \text{Surface 6}$$

The diagram illustrates the geometric representation of an artificial neural network. It shows a large 3D surface on the left, enclosed in large parentheses and labeled g_0^{-1} . This surface is equated to a sum of six smaller 3D surfaces, each multiplied by a weight w_i . The surfaces are: w_0 (a flat plane), w_1 (a surface with a vertical ridge), w_2 (a surface with a smooth curve), w_3 (a surface with a gentle slope), w_4 (a surface with a sharp dip), and w_5 (a surface with a vertical step).



Hipótesis Fundamental de la Investigación de Patrones

1. Tres momentos de la minería de datos.



¿Qué tan reciente?

¿Qué tan futuro?

2. Se mantiene aproximadamente el patrón.

3. Carácter temporal del modelo (perecedero).

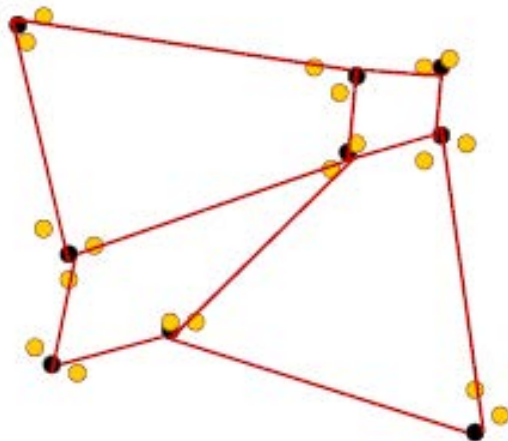


Tipos de Aplicaciones de las RN en el Análisis de Riesgo

Modelos No Supervisados Segmentación

Self-Organizing Maps

$$\varphi_i(x) = \exp\left[-\frac{\|x - \mu_i\|^2}{2\sigma_i^2}\right]$$



Modelos Supervisados Predicción

Multilayer Perceptron

$$g_0^{-1}(E(y)) = w_0 + w_1H_1 + w_2H_2$$

$$H_1 = \tanh(w_{01} + w_{11}x_1 + w_{21}x_2 + w_{31}x_3)$$

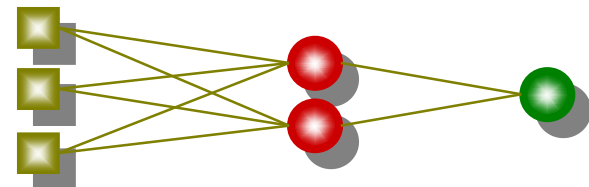
$$H_2 = \tanh(w_{02} + w_{12}x_1 + w_{22}x_2 + w_{32}x_3)$$

Radial Basis Functions

$$g_0^{-1}(E(y)) = w_0 + w_1H_1 + w_2H_2$$

$$H_1 = \exp\left(-w_{01}^2 \left((x_1 - w_{11})^2 + (x_2 - w_{21})^2 + (x_3 - w_{31})^2\right)\right)$$

$$H_2 = \exp\left(-w_{02}^2 \left((x_1 - w_{12})^2 + (x_2 - w_{22})^2 + (x_3 - w_{32})^2\right)\right)$$





Impacto Estratégico de los Modelos de Riesgo

- Más que nunca ante los modelos de riesgo soportan las decisiones estratégicas de las instituciones financieras, a niveles de país, región y globales.
- Las reservas y provisiones de las instituciones financieras están determinadas teniendo en cuenta, o por:
 - Los modelos de riesgo de crédito.
 - Los modelos de riesgo de mercado.
 - Los modelos de riesgo operacional.
 - Los modelos de riesgos de fraude.
 - Los modelos de riesgos de seguros.
 - Otros modelos de riesgo.



Impacto Estratégico de los Modelos de Riesgo

- La regulaciones como Basel II, Solvency II, etc., hacen necesario el uso de modelos más precisos.
- El error de los modelos impacta directamente:
 - La rentabilidad de las instituciones.
 - La solvencia de las empresas.
 - El valor de las acciones comerciales de las organizaciones.
 - La macroeconomía del país.
 - ⋮
 - A toda la sociedad en su conjunto.



Características Generales de los Modelos de Riesgo

- “Essentially, all models are wrong, but some are useful” (George E P E. P. Box, 1987).
- Los modelos nunca son perfectos, de hecho algunos son muy malos pero es lo mejor que se tiene cuando no hay otra alternativa (máxima de la minería de datos).
- En la construcción de los modelos se suponen y se asumen cuestiones subjetivas, inclusive desde la propia construcción de la variable objetivo.
- No se tiene en cuenta el efecto de las redes sociales y tampoco el de la macroeconomía.



Tipos de Modelos de Riesgo de Crédito

1. El modelo para la determinación del riesgo de los clientes en la fase de solicitud del crédito (**Application scoring model**).
2. El modelo para la calificación de riesgo de las cuentas de crédito existentes (**Behavioral scoring model**).
3. El modelo para estimar el monto probable de deuda que el prestador puede esperar recuperar del portafolio de clientes morosos (**Collection scoring model**).
4. El modelo para predecir las moras de los clientes con el objetivo de determinar el segmento óptimo para una revisión o auditoría (**Audit Scoring**).



Construcción y Utilización de Modelos de Riesgo con RN

- De forma muy general, la metodología para la construcción y utilización de modelos predictivos de riesgo con redes neuronales consta de las siguientes etapas:
 1. Definición de la variable objetivo.
 2. Preparación de los datos.
 3. Construcción y evaluación del modelo
 4. Calificación de los clientes con el modelo.
 5. Actuar según las calificaciones de riesgo de los clientes.
 6. Monitorear la calidad de las predicciones.



1. Definición de la Variable Objetivo

Dicotomía Subjetiva

Evento

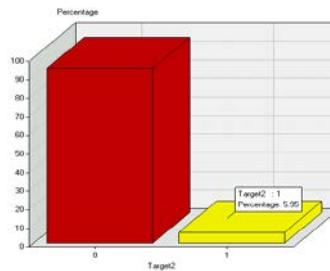
Y=1

¡No pagó!

¡Pagó menos del mínimo!

¡No pagó nada!

**¡Tres o más pagos
vencidos en el período!**



No Evento

Y=0

¡Pagó!

¡Pagó al menos el mínimo!

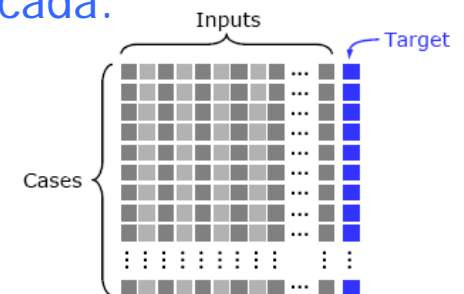
¡Pagó todo!

**¡Al menos dos pagos
vencidos en el período!**



2. Preparación de los datos

- Consiste en un análisis estadístico exploratorio de las variables e incluye las siguientes operaciones generales:
 - 2.1 Descripción de variables, clasificación y revisión de escalas.
 - 2.2 Revisión de los valores perdidos y las variables de negocio redundantes.
 - 2.3 Identificación estadística de las variables de negocio.
 - 2.4 Determinación de la varianza de la matriz de covarianzas.
 - 2.5 Tamizado de variables con poca varianza explicada.
 - 2.6 Análisis de colinealidad.
 - 2.7 Tamizado de variables colineales.
 - 2.8 Revisión final de la base.





3. Construcción y Evaluación de Modelo

- Consiste en la obtención del modelo de red neuronal que mejor describa el patrón de riesgo.
Incluye las siguientes actividades generales:
 - 3.1 Carga de la base en el minero.
 - 3.2 Exploración de los datos.
 - 3.3 Discretización de las variables con criterio de riesgo (creación de las WOE).
 - 3.4 Particionamiento de los datos.
 - 3.5 Entrenamiento y evaluación del modelo con diferentes algoritmos de aprendizaje supervisado.
 - 3.6 Comparación de los algoritmos y selección del mejor modelo.



3. Construcción y Evaluación de Modelo

3.1 Carga de la base de datos en el minero.

Es la conversión del formato de datos original a formato de datos del minero.

- .xls (MS Excel).
- .mdb (MS Acces).
- .dbf (dBASE).
- .wk? (Lotus 1-2-3).
- .db (Paradox).
- .txt (Texto).
- .html (HTML).
- ⋮



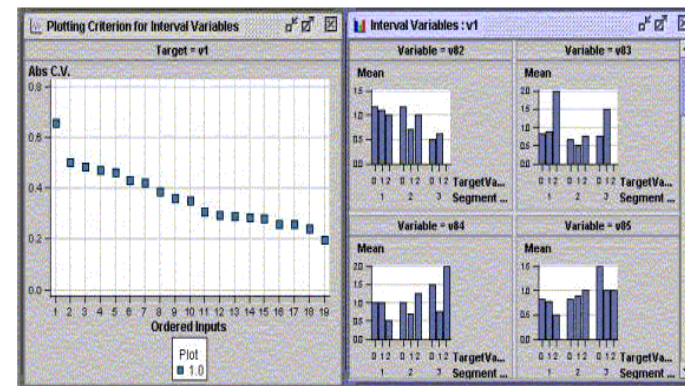
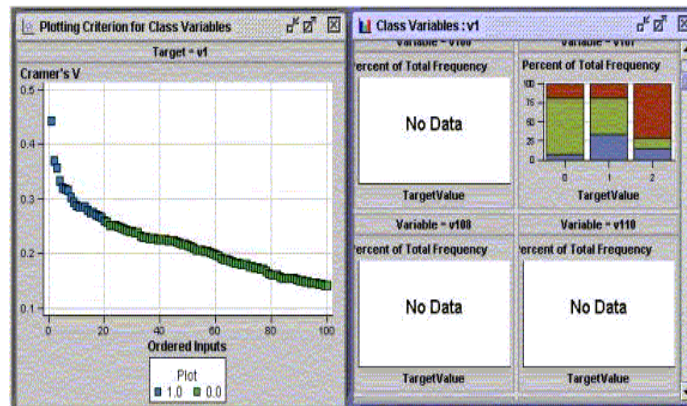
- .sas7bdat



3. Construcción y Evaluación de Modelo

3.2 Exploración de los datos.

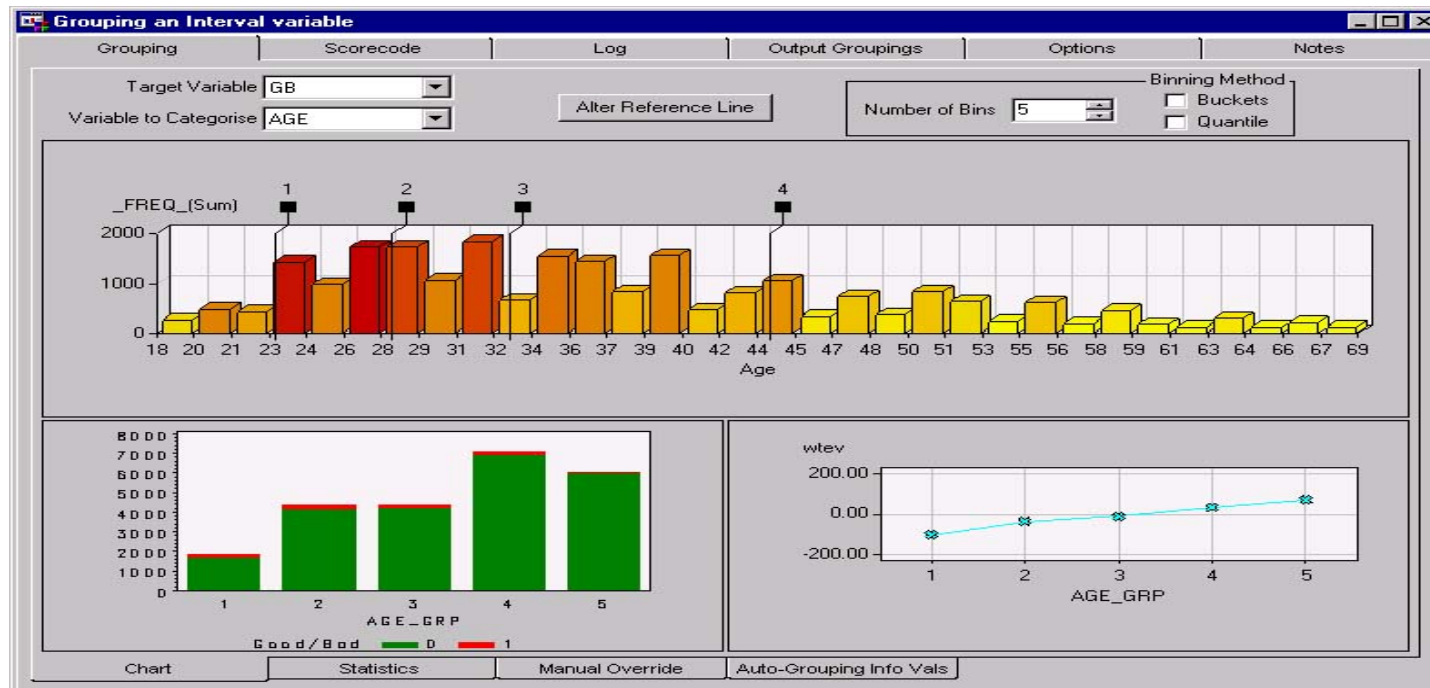
Consiste en calcular algunos estadígrafos descriptivos a la base para verificar que no hubo problemas con la importación.





3. Construcción y Evaluación de Modelo

3.4 Discretización de las variables con criterios de riesgo. Creación de las WOE's, (Weights of Evidence).





3. Construcción y Evaluación de Modelo

3.4 Discretización de las variables con criterios de riesgo. Creación de las WOE, (Weights of Evidence).

El WOE mide el riesgo relativo de un atributo o el nivel de grupo. Si se tiene una variable objetivo binaria de forma que Bad, (Evento=1) y Good (NoEvento=0), el WOE se define como:

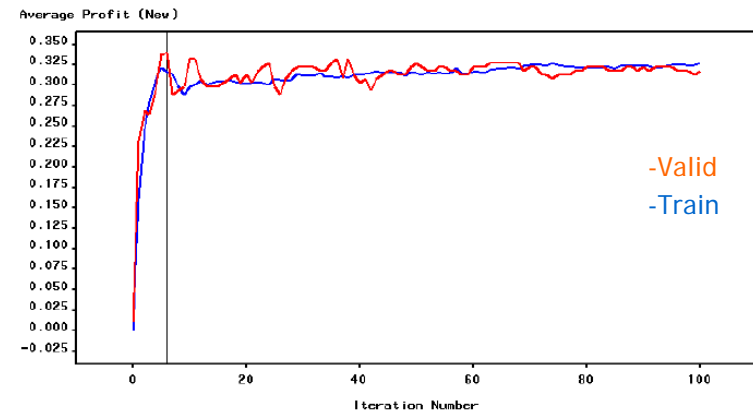
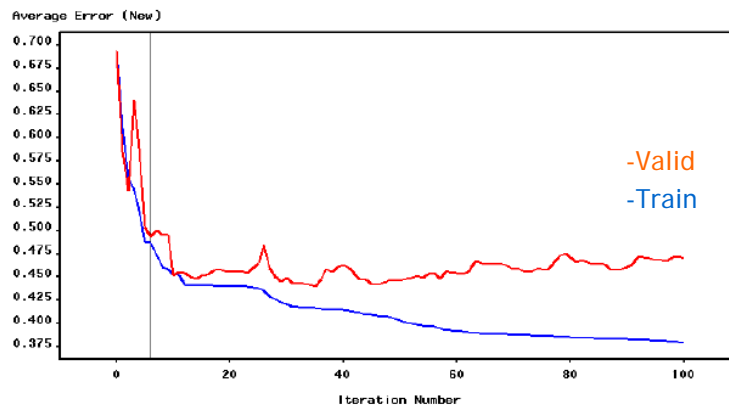
$$WOE_{Attribute} = \ln \frac{P_{Attribute}^{NoEvent}}{P_{Attribute}^{Event}}$$
$$P_{Attribute}^{NoEvent} = \frac{n_{Attribute}^{NoEvent}}{N_{Data}^{NoEvent}}$$
$$P_{Attribute}^{Event} = \frac{n_{Attribute}^{Event}}{N_{Data}^{Event}}$$



3. Construcción y Evaluación de Modelo

3.4 Particionamiento de los datos.

Se realiza con el objetivo de evitar el sobre ajuste del modelo, cuestión especialmente importante en las redes neuronales, las cuales utilizan todas las variables de entrada sin ningún tipo de discriminación.





3. Construcción y Evaluación de Modelo

3.5 Entrenamiento y evaluación del modelo con diferentes algoritmos de aprendizaje supervisado.

En este paso se entrena la red neuronal a través de un mecanismo iterativo que se detiene cuando se alcance el criterio de optimalidad: minimización de la falta de ajuste.

Se evita el sobre ajuste mediante el entrenamiento y validación de los datos al mismo tiempo.

En la práctica se construyen árboles de decisión con algoritmos de particionamiento recursivo para tratar de entender la interrelación entre las variables.



3. Construcción y Evaluación de Modelo

3.5 Entrenamiento y evaluación del modelo con diferentes algoritmos de aprendizaje supervisado.

Para la evaluación del ajuste del modelo se usan estadígrafos tales como:

- Akaike's information criterion.
- Average squared error.
- Degrees of freedom for error.
- Mean squared error.
- Schwarz's bayesian criterion.
- Misclassification rate.
- Total profit for bad.
- Average profit for bad.
- Otros.

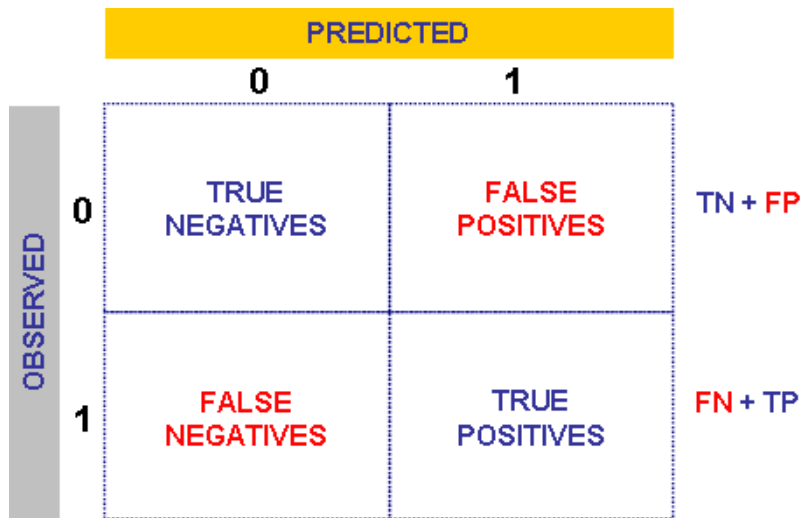
Fit Statistic	Training	Validation
Akaike's Information Criterion	1702.3402228	.
Average Squared Error	0.1884140763	0.20629268
Average Error Function	0.5513835052	0.5912328542
Degrees of Freedom for Error	1358	.
Model Degrees of Freedom	66	.
Total Degrees of Freedom	1424	.
Divisor for ASE	2848	2150
Error Function	1570.3402228	1271.1506365
Final Prediction Error	0.2067282575	.
Maximum Absolute Error	0.9993101248	0.9758885141
Mean Square Error	0.1975711669	0.20629268
Sum of Frequencies	1424	1075
Number of Estimate Weights	66	.
Root Average Sum of Squares	0.4340669031	0.4541945398
Root Final Prediction Error	0.4546737925	.
Root Mean Squared Error	0.4444897827	0.4541945398
Schwarz's Bayesian Criterion	2049.5810789	.
Sum of Squared Errors	536.60328941	443.52926194
Sum of Case Weights Times Freq	2848	2150
Misclassification Rate	0.2914325843	0.343255814
Total Profit for BAD	452	342
Average Profit for BAD	0.3174157303	0.3181395349



3. Construcción y Evaluación de Modelo

3.5 Entrenamiento y evaluación del modelo con diferentes algoritmos de aprendizaje supervisado.

La matriz de confusión permite determinar la precisión con la que la red neuronal captura y predice el patrón.



		Predicted		
		0	1	
Real	Count			
	Total %			
	Col %			
	Row %			
0	157810	65755	223565	
	36.25	15.10	51.35	
	72.51	30.20		
	70.59	29.41		
1	59816	151983	211799	
	13.74	34.91	48.65	
	27.49	69.80		
	28.24	71.76		
	217626	217738	435364	
	49.99	50.01		



3. Construcción y Evaluación de Modelo

3.5 Entrenamiento y evaluación del modelo con diferentes algoritmos de aprendizaje supervisado.

Se define la sensibilidad como:

- Probabilidad de clasificar correctamente un caso cuyo estado original es 'positive' (1).
- Proporción de 'True Positives', 1s observados predichos como 1s (HITS).
- A mayor sensibilidad, mayor potencia para clasificar o identificar correctamente a los 1s.

$$\text{Sensibilidad} = \frac{TP}{FN + TP}$$



3. Construcción y Evaluación de Modelo

3.5 Entrenamiento y evaluación del modelo con diferentes algoritmos de aprendizaje supervisado.

Se define la especificidad como:

- Es la probabilidad de clasificar correctamente un caso cuyo estado original es 'negative' (0).
- Proporción de 'True Negatives', 0s observados predichos como 0s (Accurate rejection).
- A mayor especificidad, mayor potencia para clasificar o identificar correctamente a los 0s.

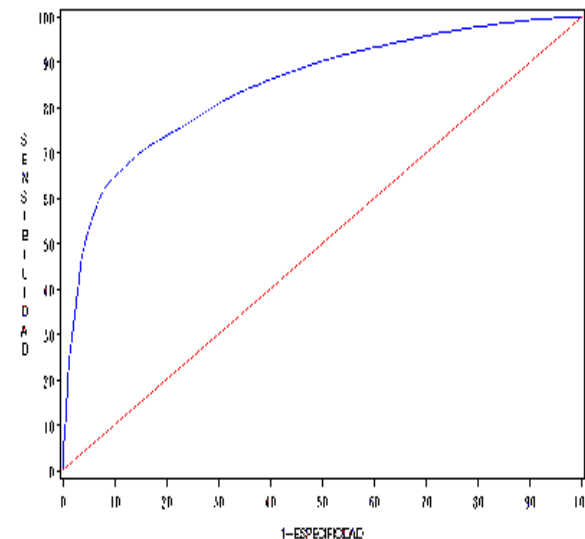
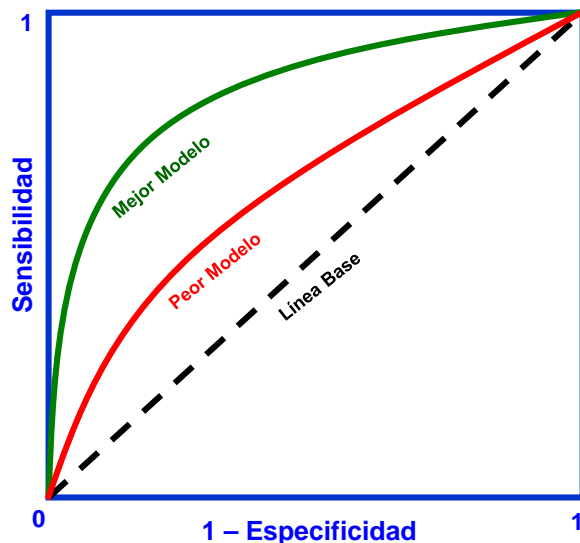
$$\textit{Especificidad} = \frac{TN}{TN + FP}$$



3. Construcción y Evaluación de Modelo

3.5 Entrenamiento y evaluación del modelo con diferentes algoritmos de aprendizaje supervisado.

Se construye una curva ROC que recrea la capacidad predictiva del modelo para diferentes puntos de corte.





3. Construcción y Evaluación de Modelo

3.6 Comparación de los algoritmos y selección del mejor modelo.

En la práctica se evalúan varios algoritmos de modelación y se comparan mediante sus matrices de confusión y curvas ROC.

Se selecciona el que ofrecen mayor rendimiento o el que tiene mayor capacidad para separar los eventos ($y=1$) de los no eventos ($y=0$).

Existen otras medidas las que comparan la distribución de scores de los clientes en clase=0 con la distribución de los clientes en clase=1. Estas no tienen en cuenta el punto de corte (break point).



3. Construcción y Evaluación de Modelo

3.6 Comparación de los algoritmos y selección del mejor modelo.

Coeficiente GINI (CG):

Se define usualmente como el doble del área entre la curva ROC y la diagonal del cuadrado del gráfico de ROC. Es equivalente al área bajo la curva ROC. Algunos valores de este estadígrafo:

CG = 0% El modelo no puede distinguir los buenos de los malos.

CG = 100% El modelo distingue perfectamente los buenos de los malos.

CG = 40% - 60% Modelo de crédito típicos en general.

CG = 70% - 80% Modelos de comportamiento de crédito.

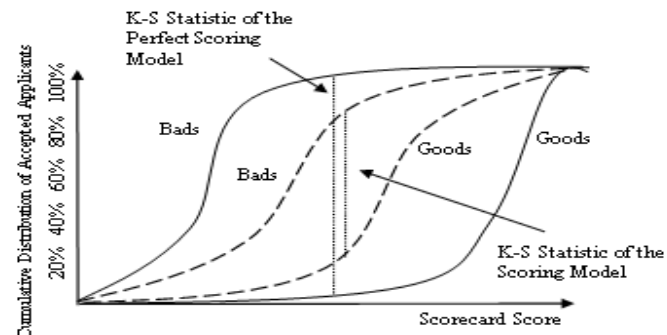
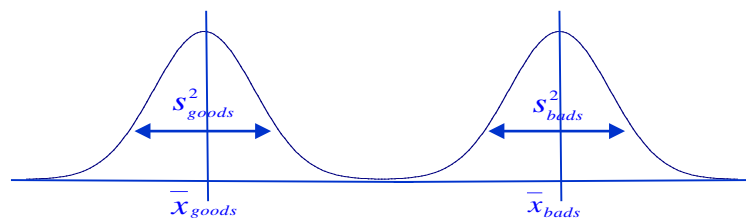


3. Construcción y Evaluación de Modelo

3.6 Comparación de los algoritmos y selección del mejor modelo.

Prueba de Kolmogorov-Smirnov (KS):

El estadígrafo de Kolmogorov-Smirnov es un estimado de la máxima diferencia entre las funciones de distribución acumulativas de los scores de los clientes clasificados como "0" y los scores de los clientes clasificados como "1".





3. Construcción y Evaluación de Modelo

3.6 Comparación de los algoritmos y selección del mejor modelo.

Information Value:

Mide la potencia predictiva de una característica (variable), su habilidad para separar solicitudes de alto riesgo de las de bajo riesgo. Algunos valores de este estadígrafo:

$IV > 0.02$ La característica se incluye en el modelo.

$IV < 0.10$ La característica es débil.

$IV < 0.30$ La característica tiene una fortaleza media.

$IV < 0.50$ La característica es fuerte.

$IV \geq 0.50$ La característica está sobre predicha.

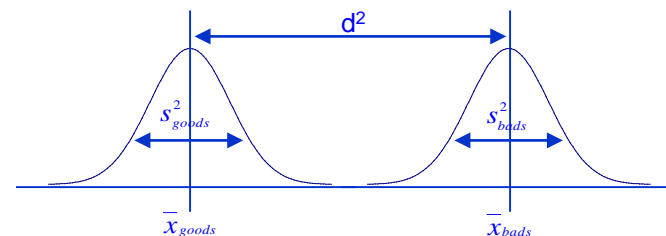


3. Construcción y Evaluación de Modelo

3.6 Comparación de los algoritmos y selección del mejor modelo.

Divergencia (FICO):

Es el cuadrado de la distancia entre las medias muestrales de las distribuciones de ceros y unos, referida a la dispersión promedio de las dos distribuciones. La mayoría de los modelos predictivos de riesgo tiene una divergencia entre 0.5 y 3.0. Puede tomar valores $[0, \infty]$.





4. Calificación de los Clientes con el Modelo

- Comúnmente se usan escalas aceptadas en el sector financiero:

$$\begin{aligned} score &= \log(odds) * factor + offset = \\ &= \left(-\sum_{i=1}^n (woe_i * \beta_i) + \alpha\right) * factor + offset = \\ &= \left(-\sum_{i=1}^n \left(woe_i * \beta_i + \frac{\alpha}{n}\right)\right) * factor + offset = \\ &= \sum_{i=1}^n \left(-\left(woe_i * \beta_i + \frac{\alpha}{n}\right) * factor + \frac{offset}{n}\right) \end{aligned}$$

- Los puntajes de la escala son enteros conforme a normas del sector.
- También se pueden utilizar índices entre 0 y 10, entre 0 y 100.



4. Calificación de los Clientes con el Modelo

- Los puntajes se pueden calcular directamente con los modelos o asignarse con la ayuda de ScoreCards.

Characteristic Name	Attribute	Scorecard Points
EDAD	. -> 27	56
EDAD	27 -> 30	70
EDAD	30 -> 31	61
EDAD	31 -> 38	60
EDAD	38 -> 44	60
EDAD	44 -> .	70
ESTADO	"CHIAPAS", "COAHUILA", "JALISCO", "NAYARIT", "NL", "SINALOA", "SONORA"	49
ESTADO	"CHIHUAHUA", "GUERRERO", "SLP"	55
ESTADO	"AGS", "DF"	56
ESTADO	"GUANAJUATO", "MICHOACAN", "PUEBLA", "QUERETARO", "YUCATAN", "ZACATECAS"	67
ESTADO	"HIDALGO", "MEXICO", "TAMAULIPAS"	70
ESTADO	"BCALIFORNIA", "CAMPECHE", "DURANGO", "MORELOS", "OAXACA", "QUINTANAROO", "TABASCO", "TLAXCALA", "VERACRUZ"	80
TIPO TRABAJO	"0", "8"	56
TIPO TRABAJO	"1", "2"	74
TIPO DEPTO	. -> 4	76
TIPO DEPTO	4 -> .	-102



5. Actuar según las Calificaciones de Riesgo de los Clientes

- A esta escala monótona creciente (risk ranking) se le asigna un punto del corte (break point) basado en consideraciones, teóricas, empíricas y comerciales.

SCORECARD / CORTE	Por encima del punto de corte	Por debajo del punto de corte
Originación	Se acepta	Se rechaza
Comportamiento	Bueno	Malo
Cobranza	Pagará	No Pagará



6. Monitorear la Calidad de las Predicciones

- Es necesario monitorear periódicamente el modelo con matrices de confusión y curvas ROC para determinar si la precisión se asemeja a lo obtenido en la etapa de entrenamiento y validación inicial.
- Pequeñas desviaciones se pueden corregir cambiando el punto de corte.
- Grandes variaciones indican que ha cambiado el patrón de riesgo y es necesario construir un nuevo modelo.
- Los monitoreos se pueden organizar en forma de auditorias a segmentos críticos de la población de clientes (alto riesgo, mayor exposición).
- Las auditorias se realizan con el minero de datos y las mismas herramientas que se usaron en la construcción.



Benchmarking de los Modelos de Redes Neuronales

- En la etapa de desarrollo del modelo:
 - Se compara con otras técnicas de aprendizaje supervisados como modelos estructurados de particionamiento recursivo y regresiones logísticas, entre otras.
- Comparación con otros modelos comerciales externos de empresas de análisis de riesgo:
 - FICO, Experian, ...
- Evaluaciones cualitativas a partir de opiniones de expertos reconocidos en el tema.
- Finalmente, la predicción del riesgo se tiene que enfocar como un proceso de mejora continua, más en tiempo de crisis.



Un Modelo versus Varios Modelos del Portafolio

- En algunos casos, en el portafolio, existen grupos de clientes significativamente diferentes, entonces se logra un mejor rendimiento segmentando el portafolio y construyendo un modelo (de la misma forma que acabamos de ver) para cada segmento.
 - Segmentación cualitativa a partir de opiniones de expertos.
 - Segmentación cuantitativa del portafolio con herramientas de minería de datos.
 - Una mezcla de ambos enfoques: quali y quanti.





Calibración de los Modelos de Riesgo de Crédito con RN

- En la actualidad y debido al fenómeno de sobre endeudamiento de los clientes (en estudio aún), los modelos de riesgo de crédito no son suficiente para una administración eficiente del riesgo, por lo que se hace necesario utilizar otros modelos complementarios para calibrar (segunda discriminación) los puntajes asignados a los clientes por los modelos actuales en uso.
- Entre estos modelos complementarios están:
 - Modelos dinámicos que incluyen la influencia de las variables macroeconómicas (de país, regionales y globales) y de las redes sociales y el efecto combinado de estas dos cosas.
 - Modelos comerciales como CII (Consumer Indebtedness Index) de Experian y el CCI (Credit Capacity Index) de FICO.



Modelación Holística del Riesgo en las Organizaciones

- La división entre los diferentes tipos de riesgo en una organización integrada va perdiendo importancia.
- Hay que calcular el riesgo a nivel de organización de manera integrada y en línea con la estrategia corporativa y las regulaciones nacionales e internacionales.
- La utilización de los modelos, como los de redes neuronales, en una administración amplia e integral del riesgo de las organizaciones.
- Una visión del riesgo integral y coherente en correspondencia con el aprovisionamiento y reservas de la organización.



Conclusiones

- Existen las herramientas y la experiencia para desarrollar modelos de redes neuronales en la predicción del riesgo.
- En el entrenamiento de las redes neuronales es necesario poner especial interés en el fenómeno de sobre ajuste.
- En las condiciones actuales de mercado un sólo modelo de riesgo no es suficiente para discriminar los clientes.
- Hay que incluir en los modelos de riesgo el efecto de las variables macro económicas (segunda discriminación o calibración de la evaluación de riesgo del cliente).
- Es necesario tener en cuenta el efecto del portafolio en su conjunto en lugar de los clientes individuales - redes sociales.
- Desarrollar un enfoque holístico del análisis de riesgo.



Referencias Bibliográficas

- Berberena V. & Híjar G. Impact of the Analytics Program on an Emergent Economy: The Mexican Case. 12th Annual Data Mining Conference. October 2009, Caesar´s Palace, Las Vegas Nevada, USA.
- Berberena V. & Paredes J. Modeling and Optimization of Marketing Campaigns. 11th Annual Data Mining Conference. October 2008, Caesar´s Palace, Las Vegas Nevada, USA.
- Berberena V. Una Aproximación al Análisis Cuantitativo del Riesgo. Forum Annual Scoring 2009 del Buró de Crédito de México. Julio 2008, Hotel Nikko, Ciudad de México.
- Van Gestel T. & Baesens B. Credit Risk Management. Oxford University Press Inc. New York, 2009.
- Anderson R. The Credit Scoring Toolkit. Oxford University Press Inc. New York, 2007.



Datos de Contacto

Dr. Viterbo H. Berberena González

- Profesor de Ciencias Analíticas del Centro de Alta Dirección en Ingeniería y Tecnología de la Universidad Anáhuac México Norte.
- Consultor Senior en Inteligencia Analítica.

✉ vberberena@anahuac.mx

✉ viberberena@pearson-research.com

✉ viterbo@berberena.com

☎ (52 55) 5328 8060 (CADIT, Universidad Anáhuac)

☎ (52 55) 5531 5324 ext.146 (Pearson Research S.A. de C.V.).