

Modelos de Scoring para Riesgo de Crédito

Los modelos de scoring de riesgo, dentro del proceso de otorgamiento de crédito, están orientados a anticipar comportamiento futuro.

Podemos dividirlos en tres grandes grupos; Admisión de créditos (Aplicant Scoring y Behavior Scoring), Cobranza de créditos (Collection Scoring) y Proyección de incobrables (Modelos de Provisiones)

La sanidad de las colocaciones de una institución financiera estará dada por la calidad de la cartera, la cual la podemos medir por la PE (Pérdida esperada) Que en el desagregado estará representada por el comportamiento de sus vintage, en particular por los que hayan alcanzado más de 90 días, presentando madurez en su desarrollo, excluyen mora técnica (Entre 1 y 30 días) y libres de efectos de castigos y recuperaciones.

Anterior a la constitución de la cartera, está el proceso de evaluación, el cual tiene distintos componentes: políticas de exclusión, políticas propias de cada productos, posibles componentes de tasa para el caso de modelos Risk to Price y finalmente para el caso de carterizados, los modelos de scoring (Aplicant para clientes nuevos y Behavior para clientes con comportamiento interno)

Para el caso de las provisiones, el fin último es si la cobertura de la cartera que se proyecta a pérdida neta es la adecuada o no (Modelo de Bis II o IFRS)

Otros factores a considerar al evaluar el desempeño de riesgo (Admisión y Cartera) son, además de los antes señalados, aspectos como: rentabilidad objetivo de la empresa, determinar Z de corte para modelos de Risk to Price, distribución de la población dentro de los cluster de riesgo y la tasa de crecimiento de las colocaciones. *Ceteris paribus* los inherentes a la gestión de riesgo, pueden distorsionar indicadores de gestión (Prima de riesgo, Índice de riesgo, Índice de pérdida, Cobertura, Otros) Por el solo hecho de modificarse la distribución de clientes en la admisión o simple aritmética.

Una vez referidos de forma muy somera, acerca de la administración de la gestión de riesgo de crédito, profundizaremos en el desarrollo de los modelo de Scoring y en particular, para estos desarrollos presentar la técnica de maquinas de vectores de soporte o "Support Vector Machine" (SVM). La elección de esta técnica (SVM). La hemos basado en que posee varias ventajas sobre otras técnica y por sus excelentes resultados en clasificar. La clasificación consiste en realizar la separación entre los clientes buenos, que responderán afirmativamente a sus compromisos de mora, de aquellos que no lo harán.

El primer objetivo a presentar es lograr determinar la mejor separación de las clases de los clientes "Pagarán" de los "No Pagarán". El segundo objetivo es demostrar que la técnica elegida pueda trabajar con muestras muy desequilibradas (4.312 Respuestas de no pago y 91.100 respuestas de pago) y como objetivo final que pueda procesar grandes volúmenes de datos.

La técnica SVM utiliza cada registro de la tabla como un vector, logrando soportar el modelo de clasificación resultante en un número pequeño de vectores.

La técnica "Support Vector Machine" (SVM) La cual a pesar de ser una técnica nueva frente a otras como redes neuronales y árboles de decisión presenta, en este caso, mejores resultados.

El método de los vectores soporte es un método general para la resolución de problemas de clasificación, regresión y estimación. El método fue propuesto por V. Vapnik en 1964 en problemas de reconocimiento de patrones. En los años 90 el método fue generalizado y en la actualidad es objeto de un gran interés, por su mejor desempeño que las Redes Neuronales.

SVM ha sido utilizado en aplicaciones en reconocimiento de patrones, reconocimiento de rostros, reconocimiento de caracteres, genética, diagnóstico del tejido humano, clasificación de documentos, modelación de sólidos e interfaces humano-computador.

Algunas de las razones por las que SVM ha tenido éxito es que no padece de mínimos locales y el modelo sólo depende de los datos con mayor información llamados vectores de soporte.

En este sentido las ventajas que tiene SVM son:

Una excelente capacidad de generalización, debido a la minimización del riesgo estructural. Existen pocos parámetros a ajustar, el modelo sólo depende de los datos con mayor información.

La estimación de los parámetros se realiza, a través, de la optimización de una función de costo convexa, lo cual evita la existencia de un mínimo local.

La solución es sparse, o sea la mayoría de las variables son cero en la solución de SVM. Esto se traduce en que el modelo final puede ser descrito como una combinación de un número pequeño de vectores de entrada, llamados vectores de soporte.

Las Maquinas de Vectores Soporte o SVM son nuevas estructuras basadas en la teoría estadística del aprendizaje. Se basan en transformar el espacio de entrada en otro de dimensión superior (Espacio de llegada) en el que el problema puede ser resuelto mediante un hiperplano óptimo (De máximo margen).

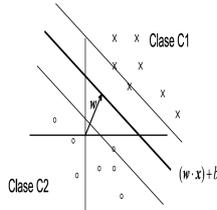
El método de los vectores soporte ("Support Vector" o SV) es un método general para resolver problemas de clasificación cuyo objetivo es encontrar una función $y = f(x)$ dados un conjunto de patrones entrada-salida (posiblemente con ruido), $(x_1, y_1), \dots, (x_N, y_N)$.

La idea detrás del método consiste en transformar los vectores de entrada x (n -dimensionales) en vectores de dimensión más alta z (incluso de dimensión infinita) en los que el problema pueda solucionarse linealmente:

$$y = (w \bullet z) + b$$

De esta forma dado el conjunto de muestras $(x_1, y_1), \dots, (x_N, y_N)$, y siendo $y_i = \{+1, -1\}$, se desea separar las dos clases mediante un hiperplano (suponiendo en este punto que el problema es separable) Existe un único hiperplano óptimo que es el que proporciona un mayor margen de separación entre las clases: $\max \{w, b\} \min\{|x - x_i|, (w \cdot x + b)\}$, con $i = 1, \dots, n$

El problema de encontrar el hiperplano óptimo puede formularse como:
 Sujeto a: minimizar $g(w) =$



El problema de encontrar el hiperplano óptimo puede formularse como:

$$\text{minimizar } g(w) = \frac{1}{2} \|w\|_2^2$$

$$y_i \{(w \cdot x) + b \geq 1\} \quad i = 1, \dots, N$$

La solución a dicho problema toma siempre la forma:

$$w = \sum \alpha_i y_i x_i$$

los α_i son multiplicadores de Lagrange, muchos de los cuales son cero. Los vectores asociados a los multiplicadores no nulos se denominan "vectores de soporte", por lo que:

$$w = \sum_{\text{vectores de soporte}} \alpha_i y_i x_i$$

la constante b se obtiene de alguna de las condiciones (Karush-Kuhn-Tucker)

Clasificadores no lineales

Para extender el concepto de SVM a clasificadores no lineales, se realiza una transformación del espacio de entrada a otro de alta dimensionalidad en el que los datos son separables linealmente

Bajo ciertas condiciones (Condición de Mercer), el producto escalar en el espacio de salida se puede escribir a través de un cierto núcleo o Kernel $\{K(x, y)\}$

Para el desarrollo de nuestro modelo en que tenemos como objetivo, demostrar que SVM tiene mejor performance para scoring de Riesgo de Crédito, inicialmente hemos trabajado con una tabla de 491 variables y 95.412 registros, con data de iniciación y de comportamiento de instituciones financieras, doce vintage y un horizonte de evaluación de dieciocho meses. Estos datos serán analizados, depurados y reparados lo que permitirá determinar la cantidad de variables y registros útiles para el modelo.

En una segunda etapa, de los datos validos se seleccionamos las variables y/o registros para lograr un entrenamiento más rápido, mejorar la capacidad de generalización y evitar sobre ajuste. Todo lo anterior para hacer más eficiente el modelo, hacerlo descifrable para el negocio (Cualquier institución financiera) y de fácil implementación en registros nuevos (Iniciación de clientes - Applicant Scoring y Behavior Scoring)

Realizamos un análisis descriptivo de los datos en estudio. Esta parte estuvo compuesta en primera instancia, del cálculo de las estadísticas básicas para cada una de las variables involucradas. También y siguiendo las etapas involucradas en Data Mining se efectuó una limpieza y reparación de la tabla. En la aplicación de técnicas estadísticas de limpieza de datos se identificaron los datos perdidos y los datos anómalos u outliers.

Finalmente nuestro modelo de Scoring presentó los siguientes estadígrafos, que nos permiten indicar que presenta excelente estabilidad, predictibilidad y capacidad de separación de los "Buenos Pagadores" de los "Malos Pagadores" demostrando que efectivamente, aplicado a los procesos de riesgo, es superior a las otras técnicas revisadas en nuestro análisis:

Ilustración 1: Optimización de la separación de las clases

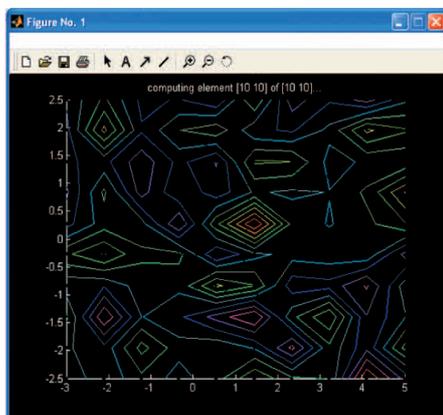


Ilustración 2: Criterios de conclusión

Valor indice	Discriminación	Modelo
K-S		
$KS \leq 0.2$	Nula	No funciona
$0.2 < KS \leq 0.4$	Leve	Satisfactorio
$0.4 < KS \leq 0.5$	Aceptable	Bueno
$0.5 < KS \leq 0.6$	Excelente	Muy bueno
$0.6 < KS$	Extraordinaria	Extraordinario
AUC		
$AUC \leq 0.5$	Nula	No funciona
$0.5 < AUC \leq 0.7$	Leve	Satisfactorio
$0.7 < AUC \leq 0.8$	Aceptable	Bueno
$0.8 < AUC \leq 0.9$	Excelente	Muy bueno
$0.9 < AUC$	Extraordinaria	Extraordinario
ROC		
$ROC \leq 0.5$	Nula	No funciona
$0.5 < ROC \leq 0.65$	Leve	Satisfactorio
$0.65 < ROC \leq 0.8$	Aceptable	Bueno
$0.8 < ROC \leq 0.9$	Excelente	Muy bueno
$0.9 < ROC$	Extraordinaria	Extraordinario

Ilustración 3: Curva de ROC SVM v/s Técnica de Clúster

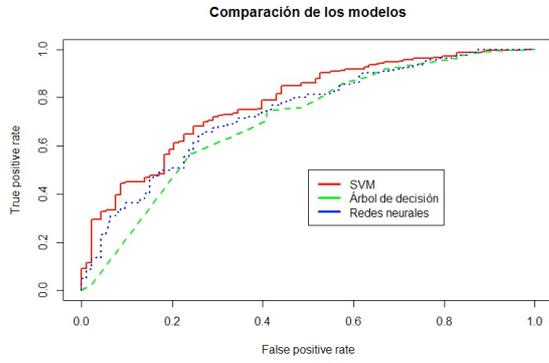


Ilustración 4: Conclusión SVM v/s Clúster v/s Redes Neuronales

Modelo	KS	AUC
SVM	0.53649	0.83148
Redes Neuronales	0.42291	0.74738
Árbol de Decisión	0.41579	0.74611

Ilustración 5: Matrices de confusión para cada modelo

		Valor Real			
		0	1	Total	
Predicción	0	2.827	13.075	15.902	Correctos
	1	1.485	78.025	79510	
Total		4.312	91.100	95.412	
SVM					
		Valor Real			
		0	1	Total	
Predicción	0	1.908	23.039	24.947	Correctos
	1	2.404	68.061	70.465	
Total		4.312	91.100	95.412	
Redes Neuronales					
		Valor Real			
		0	1	Total	
Predicción	0	2998	25719	28718	Correctos
	1	1314	65380	66694	
Total		4.312	91.100	95.412	
Árboles de Decisión					