

Anticipándose a las decisiones del consumidor

Analizando correctamente la información no es necesario tener una bola de cristal para detectar patrones de comportamiento para predecir el comportamiento de los clientes.



En entornos enormemente cambiantes como los que enfrentamos en la actualidad, donde un cliente está siendo abordado por múltiples canales, con gran variedad de productos y facilidades para el pago, es necesario moverse rápido. El que no llega a tiempo con la propuesta correcta para cada cliente, lo perderá y le será muy difícil recuperarlo.

Aunque parezca extraño, mediante el análisis de la información es posible detectar patrones que nos permitirán predecir comportamientos de los clientes. Sin tener una bola de cristal o ser un vidente, usted podrá identificar qué clientes serán, por ejemplo, los más propensos a fugarse, a comprar un producto, a realizar un reclamo y varias otras aplicaciones que están teniendo las empresas actualmente y que le son de gran ayuda.

La predicción ayuda a anticiparse a la decisión del consumidor permitiendo entregarle una propuesta que

él tomaría con una alta probabilidad.

Resultados concretos

La predicción, en esencia, funciona generando grupos de clientes que poseen una alta probabilidad de generar algún evento determinado. Esta probabilidad es asignada en base al comportamiento anterior del grupo de clientes.

Las aplicaciones de la predicción en los negocios son múltiples, siendo las más utilizadas la **fuga** y la **propensión de compra**. Sin embargo, es posible aplicarlo en cualquier análisis en que sea necesario responder a la siguiente pregunta: ¿quiénes son los más propensos a ...? Lo importante es tener el registro de la generación del evento que se requiere estudiar, es decir, la variable objetivo.

Una predicción exitosa puede generar importantes beneficios para la empresa. Por ejemplo, si logra predecir correctamente a los clientes con propensión de fuga, puede impactar de manera relevante en los resultados finales de la empresa. Cada cliente retenido supondrá ingresos que hubiera perdido y, por lo tanto, mejora en la utilidad. Usted podría dimensionar el beneficio de su modelo haciendo un simple cálculo como el siguiente:

	Antes	Después
Total Clientes	1,000	1,000
Clientes Fugados	100	85
Tasa Efectividad	10%	8.5%
Ingreso Promedio por Cliente	\$100	\$100
Ingreso Adicional por menor Fuga	\$0	\$15,000

Por lo tanto, el haber logrado reducir la tasa de fuga de 10% a 8,5% supone un incremento en ingresos para la empresa de \$15.000.

De la misma forma es posible medir el beneficio marginal para otras aplicaciones. Si, por ejemplo, su objetivo es medir propensión **de compra**, el beneficio será por el menor costo que supone contactar a un grupo de clientes más reducidos. Luego, el análisis sería de la siguiente forma:

	Antes	Después
Cientes Contactados	100	50
Total Clientes que contratan	5	4
Tasa Efectividad	5%	8%
Ingreso por Cliente	\$1,500	\$1,500
Ingreso Total	\$7,500	\$6,000
Costo por Contacto	\$50	\$50
Costo Total	\$5,000	\$2,500
Utilidad Total	\$2,500	\$3,500

En este ejemplo el haber incrementado la efectividad supone un ingreso adicional de \$1.000. Es importante notar que es posible que incluso la venta absoluta sobre el nuevo grupo sea menor - tal como se ve en el ejemplo la venta se redujo en un 20%- , sin embargo, la disminución de costos compensa con creces esa menor venta.

Como siempre: de lo simple a lo sofisticado

Tal como se ha visto en los módulos anteriores, el BI se mueve en un espectro muy amplio en términos de complejidad y sofisticación. La predicción no es una excepción. Los modelos analíticos existentes permiten generar tasas de éxito impresionantes. Por ejemplo, los usuarios de Gmail se quedan admirados con la capacidad de clasificación de los Spam de la empresa; de hecho, posee tasas muy cercanas al 100% de efectividad. Este sistema de clasificación se realiza a partir de la aplicación de modelos de predicción no triviales, que detectan las características de un email Spam, logrando acertar en la mayoría de los casos.

Ahora, si usted aún no ha abordado complejos modelos de predicción, es recomendable comenzar con la tradicional estadística descriptiva, para ir detectando ciertos patrones que lo puedan llevar a encontrar ese grupo. Además, al pasar por este proceso se dará cuenta de cuáles son las variables más relevantes que antes usted no conocía o sólo intuía.

Supongamos, por ahora, que no conocemos ninguna herramienta y modelo de predicción, por lo que tendremos que resolver el problema de la siguiente empresa mediante **estadística descriptiva**:

Una compañía de prensa escrita y online está estudiando incorporar la totalidad (hoy es parcial) de la lectura de su diario por internet y, para ello, requiere conocer cuántos clientes estarían dispuestos a pagar por este nuevo servicio. Para ello le solicita a usted que realice el análisis para poder conocer si es una decisión rentable o no.

Usted toma el proyecto y pide toda la información existente que le pueda ser de utilidad. Durante los días que pasan mientras obtiene la información se le ocurre una buena idea: estudiar el comportamiento de la navegación por internet de los clientes actuales y así conocer el nivel de importancia que le dan a la lectura de este medio por internet.

Una de las bases de datos que le han entregado posee todas las transacciones realizadas por internet dentro del portal personal del cliente. Luego, usted se plantea conocer lo siguiente:

- Frecuencia de visita
- Tiempo de la visita
- Hace cuántos días fue la última visita
- Páginas más visitadas

- Edad del cliente
- Antigüedad del cliente
- Si el cliente posee suscripción física

Resulta que tiene buena suerte y ha logrado encontrar todas las respuestas en las bases entregadas por la empresa. Y no sólo eso: además, ha detectado la presencia de un grupo que poseería, según su criterio, una alta probabilidad de contratar el servicio por internet. Este grupo está determinado por clientes con alta frecuencia de visita, que se mantienen mucho tiempo cada vez que visitan, que no dejan pasar muchos días sin entrar en el portal, que visitan al menos diez páginas distintas, entre 20 y 60 años y que poseen actualmente una suscripción. Resulta que esto representa un 10% de la totalidad de la base de clientes, pero la empresa que le solicitó el análisis está muy contenta, porque ahora se enfocará en abordar con una buena oferta a este tipo de clientes. Además, le cuenta que aplicarán las mismas reglas para aquellos nuevos clientes que aún no han suscrito la lectura de la totalidad del diario.

Como puede observar, usted ha creado un modelo de predicción realizando un cruce simple entre variables. No fue necesario realizar modelos muy complejos, sino que primó el conocimiento que usted tenía del perfil de clientes que navega por internet y del negocio de la prensa online.

La siguiente etapa en su modelo sería detectar el perfil de los clientes que finalmente contrataron. De esta manera su variable objetivo sería la real y no una aproximación, tal como lo hizo con aquellos que tendrían propensión a comprar la suscripción. De hecho, la empresa podrá medir la efectividad de su modelo inicial haciendo la comparación mediante un grupo de control con clientes que no poseen propensión, según su modelo. Si la tasa de contratación sobre los clientes abordados mediante una oferta es mucho mayor que la del grupo de control, el modelo aplicado fue correcto. De lo contrario, será necesario rehacer o afinar la puntería con modelos de predicción más complejos que logran altas tasas de efectividad para ciertos problemas.

Modelos Predictivos

No es el objeto de este curso ahondar en cada uno de los modelos, pero sí es importante conocer cuáles son los más utilizados al día de hoy:

Árboles de Decisión: los árboles de decisión buscan separar dos o más grupos en cada nodo, de manera homogénea entre sí y heterogénea entre grupos. De esta forma es posible encontrar nodos donde se concentran grupos “puros” con respecto a la variable de predicción y es posible explicar el grupo a partir de las variables discriminantes.

Naive-Bayes: es un clasificador que utiliza la teoría de probabilidades condicionales (teorema de Bayes) para modelar la predicción.

Redes Neuronales: este complejo modelo matemático-computacional logra resolver una predicción a partir de un sistema de nodos interconectados que se asemejan al funcionamiento de nuestro sistema neuronal.

Support Vector Machines: este modelo, que posee similitudes con las redes neuronales, busca separar los grupos de predicción maximizando la separación entre ellos. Este tipo de modelos ha mostrado muy buenos resultados en la actualidad y, para algunos problemas, es el que mejor se comporta.

Regresión Logística: Utilizada principalmente para la predicción de variables continuas, como por ejemplo, proyectar la demanda en un cierto período de tiempo.

Midiendo resultados

Independiente del modelo que utilice, es de gran importancia medir los resultados obtenidos de la predicción. Esta medición puede hacerse de la siguiente forma:

Grupos de Control: al elegir un grupo que se abordará para alguna acción específica, es importante separar otro grupo de clientes que, según el resultado de su modelo, no posee alta propensión. De esta forma, al comparar las tasas de fuga, compra u otras, de un grupo con otro, sería esperable que el grupo de control posea tasas considerablemente menores.

Curvas de Ganancia: las curvas de ganancia muestran “probabilidad acumulada” de los clientes ordenados de mayor a menor propensión. Es decir, representa la ganancia de efectividad generada al abordar a los clientes con mayor propensión.

El siguiente gráfico muestra la curva de ganancia de efectividad de un modelo de predicción de respuesta de emails enviados. La curva verde muestra que, por ejemplo, enviando al primer 30% más propenso para responder, es posible obtener un 60% de las respuestas. Luego, existe una ganancia del 30% con respecto a la situación base. Si esto mismo, por ejemplo, se aplica a clientes que compran un producto por telemarketing, es posible reducir los costos de contacto al cliente de manera importante ya que es necesario contactar sólo a un grupo reducido de clientes para obtener altas tasas de efectividad.

