

Modelo de Oportunidad de Mejora para el cliente en la elección de su plan celular mediante el diseño de una datawarehouse y árboles de decisión

Lorena Patricia Calvas Bayas ¹, Ángela Elvira Vásconez Vasconez ²
Instituto de Ciencias Matemáticas, Escuela Superior Politécnica del Litoral
Campus Gustavo Galindo Velasco, Km. 30.5 Vía Perimetral, Apartado: 09-01-5863
Guayaquil - Ecuador
evascone@espol.edu.ec – lpcb08@hotmail.com

Resumen

Para efectos de este proyecto hemos considerado el proveedor de telefonía móvil con el nombre de Cellcorpo. Cellcorpo es un proveedor de celular que tiene una gran diversidad de planes de llamadas, hemos elegido los dos más populares en nuestro mercado: Controlado y de prepago. Un plan de llamadas controladas se define por tener una cantidad fija de minutos y / o cantidad fija de mensajes de texto, dependiendo de qué tipo de plan que el cliente decida cuando se adquiere la teléfono móvil con el proveedor. Un plan de llamadas pre-pagadas se define por un plan, si el cliente compra minutos pre pagados para activar el teléfono hasta que los minutos estén agotado o se caduquen. En nuestro estudio hemos aplicado árbol de decisión para la base de datos para empresas determinar el cliente se beneficiaría cambiar su plan de llamadas según el últimos tres meses de su plan de llamadas hábitos de consumo. base de datos de Cellcorpo consta de doscientos veintitrés mil doscientos cincuenta y seis clientes. Estos clientes tienen actualmente uno de los dos tipos de planes en nuestra estudio, controlada o prepago. Para llegar a nuestros resultados nos han modelado los árboles de decisión utilizando la tabla "Maestro" con la ayuda de un programa llamado AnswerTree por SPSS. Después de aplicar las variables necesarias para la segmentación de clientes, llegamos a los sesenta y cinco mil seiscientos sesenta y tres clientes que se beneficiarían de un cambio de planes.

Palabras Claves: Autocontrol, prepago.

Abstract

For the purpose of this project we have deemed the cellular provider we have chosen for a case study as Cellcorpo. Cellcorpo as a cellular provider that has a diversity of calling plans, we have chosen the two most popular in our market: Controlled and Pre-paid. A Controlled calling plan is defined by having a fixed amount of minutes and or fixed amount of text messages, depending of what type of plan the client chooses when he acquires the cellular phone with the provider. A pre-paid calling plan is defined by a plan where the client buys prepaid minutes which activate the phone until the minutes run out or they expire. In our study we have applied decision tree's to the companies data base to determine which client would benefit changing their calling plan according the last 3 months of their calling plan consuming habits. Cellcorpo's database consists of two hundred and twenty three thousand two hundred and fifty six clients. These clients currently have one of the two types of plans in our study; controlled or pre-paid. To arrive at our results we have modeled decision trees using the table "Maestro" with the help of a program called AnswerTree by SPSS. After applying the necessary variables to the segmentation of cliental, we arrive at sixty five thousand and six hundred and sixty three clients that would benefit from a change of plans.

Keywords: Controlled, Pre-paid.

1. Introducción

Para realizar este estudio tomamos la base de datos de la empresa CELLCORPO la cual contiene doscientos veinte y tres mil doscientos cincuenta y seis registros, que es el total de nuestra clientela, quienes poseen dos tipos de planes de telefonía celular, plan prepago o plan controlado.

Con la bases de datos de nuestros clientes, realizamos un modelo de minería de datos donde veremos el diagrama estrella con las tablas clientes, consumo cliente plan, consumo cliente tarjeta, contrato cliente, detalle plan, fecha, número celular y la tabla de hecho denominada Maestro, la unión de dichas tablas

2. Marco Teórico

2.1 Diagrama de Influencia

El diagrama de influencia es una herramienta gráfica que muestra las relaciones entre las decisiones, los eventos al azar y las consecuencias para un problema de decisión.

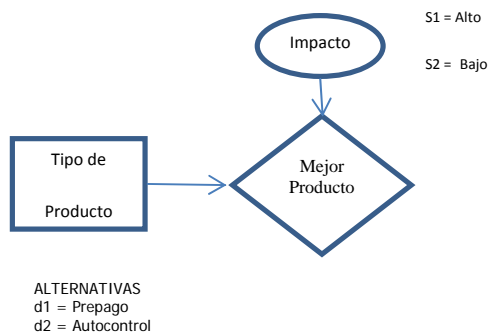


Figura 1. Diagrama de influencia

2.2 Modelo Multidimensional

El Modelo estrella o también conocido como Diagrama Estrella es llamado así debido a que su implementación en un ambiente relacional esta dado por varias tablas que almacenan las jerarquías dimensionales y una tabla de hecho, en este caso la tabla de Hecho se denomina Maestro.

forman el cubo de datos del modelo. Realizamos un estudio con árboles de decisión para obtener como resultado a los clientes que debemos escoger para presentarles una mejor opción con respecto a su plan de telefonía celular actual.

Tenemos como finalidad determinar un mejor beneficio a los clientes de una empresa de telefonía celular en cuanto al plan que necesite y le sea más conveniente, tomando en cuenta sus hábitos de consumo de los tres últimos meses de utilización de su línea telefónica celular.

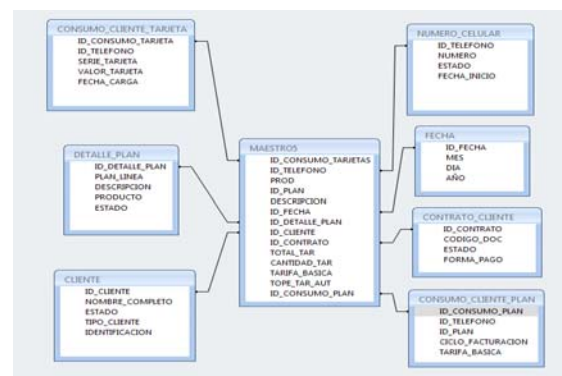


Figura 2. Diagrama Estrella

El cubo de datos es el conjunto formado por todas las tablas Dimensión y tabla de Hecho que al final dan una vista en forma de Cubo cuyas celdas están compuestas por las medidas de la tabla de Hecho.

3. Modelamiento de Árbol de Decisión

3.1 Árbol de decisión

Un árbol de decisión es un modelo de predicción utilizado para seleccionar una de varias alternativas posibles de decisión.

Dada una base de datos se construyen diagramas de construcciones lógicas, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que ocurren de forma sucesiva, para la resolución de un problema.

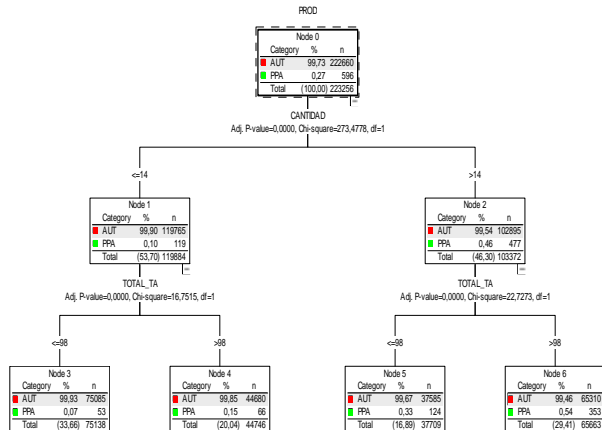
Para el análisis del árbol de decisión nos valemos de un software estadístico, denominado SPSS AnswerTree Versión 3.

SPSS AnswerTree es una herramienta para la creación de sistemas de clasificación en forma de árbol de decisión, lo cual nos ofrece una fácil interpretación de los resultados.

Con los datos de la tabla de Hecho, Maestro, la cual contiene doscientos veinte y tres mil doscientos cincuenta y seis registros, y con ayuda del software antes mencionado ingresamos dichos registros para hacer el árbol de decisión, podemos por medio del árbol tipo CHAID para realizar la predicción.

El árbol de segmentación general que se muestra en el Gráfico 3.1 se divide en dos niveles a partir del nodo parental (nodo 0). Cada nivel está representado por una variable predictora colocada en orden de prioridad. De hecho, la cantidad resulta ser el mejor predictor, por tener el X^2 más alto (273,48). De igual forma, es la variable en la que se observa mayor diferencia estadísticamente significativa entre los grupos de categorías que la integran, como es el caso de las menores y mayores de 14 tarjetas ingresadas durante los tres últimos meses.

ÁRBOL DE DECISION. PREDICTOR: Cantidad



Fuente: AnswerTree SPSS 3.0

Elaborado: Lorena Calvas y Ángela Vásconez

El nodo raíz muestra el desglose de los casos de toda la muestra (223.256 clientes).

En este conjunto de datos se observa que los clientes poseen plan autocontrol en su mayoría (99,73%) y el resto plan prepago (0,27%).

Cabe destacar que la mayoría de los clientes que ingresan la cantidad menor o igual a catorce tarjetas durante los tres meses poseen plan autocontrol (99,90%), así mismo los clientes que ingresan mayores de catorce tarjetas durante los tres meses poseen plan autocontrol (99,54%).

El estudio se realiza durante tres meses, es por ello que cabe recalcar que pasado este período de tiempo los clientes que poseen planes pre-pagos si no han ingresado minutos de tiempo aire en su línea celular, dicha línea será anulada totalmente. La caducidad de las tarjetas se muestra a continuación:

Tiempos de vigencia:

\$3: 7 días - **\$6:** 30 días - **\$10:** 30 días - **\$20:** 45 días - **\$30:** 45 días

Es decir que para mantener activa su línea celular debe haber realizado durante los últimos tres meses al menos una recarga a su plan pre-pago, de lo contrario perderá su número o línea celular.

4. Resultados del Análisis

Con la segmentación ya realizada y guiándonos por las variables predictoras que el árbol de decisión nos dio, llegamos a sesenta y cinco mil seiscientos sesenta y tres a los cuales podemos ayudar con la mejora del plan que posee.

5. Conclusiones y Recomendaciones

Realizando la segmentación respectiva llegamos a la conclusión de que de los doscientos mil registros aproximados sesenta y cinco mil seiscientos sesenta y tres son los clientes a los cuales debemos “atacar”, proponiendo un nuevo plan que le darán mayores beneficios que el plan que tiene actualmente.

De los sesenta y cinco mil seiscientos sesenta y tres clientes, más del 90% de los clientes poseen un plan autocontrol según análisis de estudio.

De los sesenta y cinco mil seiscientos sesenta y tres clientes, menos del 10% de los clientes poseen un plan prepago según análisis de estudio.

El 30% de los clientes son aquellas personas que deben cambiar su plan por una mejor opción, debido a su hábito de consumo y así ahorrarían dinero.

La variable Cantidad de tarjetas es la que nos ayuda a predecir por medio de árboles de decisión los clientes que deben cambiar de plan.

Los clientes que son resultado de nuestro estudio ingresan más de catorce tarjetas en el trimestre.

Los clientes que son resultado de nuestro estudio consumen más de noventa y ocho dólares en el trimestre.

Los clientes que son resultado del estudio deben cambiar a un nuevo plan de telefonía celular.

Se debería tomar en cuenta la base de datos la tarifa promedio que utilizan todos los clientes en ambos planes, así sería más rápido obtener a los clientes que mas ingresan tarjetas durante los tres meses tanto en plan autocontrol como en planes prepagos.

Sería de mucho beneficio que la empresa realice una prueba piloto con los clientes que han sido resultado del estudio, para que ellos se den cuenta del gran beneficio que ellos lograrían cambiando su plan actual. La prueba piloto se basaría en hablar con dichos clientes y con la aprobación de ellos cambiar al plan que se le recomiende por al menos tres meses y así ellos podrían comparar con los tres meses anteriores.

6. Bibliografía

1. SPSS (2002). Answer Tree 2.0. User's Guide. Chicago: SPSS. Statistics Package for the Social Sciences (2002). Answer Tree 2.0. User's Guide. Chicago: SPSS.
2. Statistics Package for the Social Sciences (2002). Answer Tree 2.0. User's Guide. Chicago: SPSS.
3. 2008, http://www.google.com.ec/imgres?imgurl=http://www.ingelsi.com.ec/spss/marketing-3.gif&imgrefurl=http://www.ingelsi.com.ec/spss/marketing-pack.htm&usq= mK11nN-M6RiHYsXVLRdj-qmEC3U=&h=375&w=500&sz=18&hl=es&start=6&um=1&itbs=1&tbnid=x3XajYQFb_j3BM:&tbnh=98&tbnw=130&prev=/images%3Fq%3Dinterpretacion%2Bde%2Blos%2Bresultados%2Bde%2Bbarboles%2Bde%2Bdecision%2Bpor%2Banswertree%26um%3D1%26hl%3Des%26sa%3DN%26tbs%3Disch:1
4. 2008, http://www.telematica.com.pe/Product/spss_AnsTree.htm
5. 2009, http://es.wikipedia.org/wiki/Minería_de_datos
6. 2009, http://www.answermath.com/mineria_de_datos.htm
7. 2009, http://www.aprendedynamics.com/conceptos_olap.html
8. 2009, <http://www.slideshare.net/dannoblack/datawarehouse-y-datamining-parte-i>
9. 2009, http://www.google.com.ec/imgres?imgurl=http://www.sinnexus.com/images/datamart_cube.jpg&imgrefurl=http://www.sinnexus.com/business_intelligence/datamart.aspx&usq= 2MHYfKY5SO_aA9sCJFus8FuhpQ=&h=167&w=199&sz=6&hl=es&start=3&um=1&itbs=1&tbnid=ua2NqymvrxmzsM:&tbnh=87&tbnw=104&prev=/images%3Fq%3Ddatamart%26um%3D1%26hl%3Des%26sa%3DN%26rlz%3D1G1GGLQ_ESEC377%26tbs%3Disch:1
10. http://www.google.com.ec/imgres?imgurl=http://www.sinnexus.com/images/datamart_cube.jpg&imgrefurl=http://www.sinnexus.com/business_intelligence/datamart.aspx&usq= 2MHYfKY5SO_aA9sCJFus8FuhpQ=&h=167&w=199&sz=6&hl=es&start=3&um=1&itbs=1&tbnid=ua2NqymvrxmzsM:&tbnh=87&tbnw=104&prev=/images%3Fq%3Ddatamart%26um%3D1%26hl%3Des%26sa%3DN%26rlz%3D1G1GGLQ_ESEC377%26tbs%3Disch:1
11. 2009, http://www.google.com.ec/imgres?imgurl=http://www.sinnexus.com/images/datamart_cube.jpg&imgrefurl=http://www.sinnexus.com/business_intelligence/datamart.aspx&usq= 2MHYfKY5SO_aA9sCJFus8FuhpQ=&h=167&w=199&sz=6&hl=es&start=3&um=1&itbs=1&tbnid=ua2NqymvrxmzsM:&tbnh=87&tbnw=104&prev=/images%3Fq%3Ddatamart%26um%3D1%26hl%3Des%26sa%3DN%26rlz%3D1G1GGLQ_ESEC377%26tbs%3Disch:1

Msc. Pedro Fabricio Echaverria Briones
Director de Tesis