**Sistema de Soporte para la Toma de Decisiones en el cobro de Autoavalúos y Pronóstico tributario para la Municipalidad Distrital Mariscal Cáceres - San José en la provincia de Camaná, utilizando Cubos OLAP y Minería De Datos.**

*System Support for Decision Making in collecting Autoavalúos and Forecast tax for the District Municipality Mariscal Cáceres - San Jose in the province of Camana, using OLAP Cubes and Data Mining.*

Milagros Yoshira Cervantes Llanquire[[1]](#footnote-2)

Víctor Manuel Cornejo Aparicio[[2]](#footnote-3)

**RESUMEN**

Este proyecto se basa en las dificultades que presentan la municipalidad, en el tema de Autoavalúos recaudados cada año, donde la institución posee sistemas transaccionales que registran y permiten tramitar los referidos pagos, no permitiendo una explotación adecuada de la data existente, por ello; el presente trabajo muestra la aplicación de minería de datos y cubos OLAP aplicados a los registros existentes de tal manera que permita elaborar elementos de juicio para que las autoridades puedan elaborar planes de desarrollo en función a proyecciones técnicas de la recaudación, y del mismo modo aplicar estrategias de recaudación por sectores.

**Palabras clave:** M*inería de datos, OLAP, Sistema de soporte a las decisiones DSS.*

**ABSTRACT**

This project is based on the difficulties presented by the municipality, on the issue of Autoavalúos raised each year, where the institution pose transactional systems that record and can handle such payments, not allowing adequate exploitation of existing data, therefore, the this paper shows the application of data mining and OLAP cubes applied to existing records so that would develop evidence to allow authorities to prepare development plans according to projections collection techniques, and likewise implement strategies collection by sector.

**Key words:** *Data Mining, OLAP, Decision Support System DSS.*

**INTRODUCCIÓN**

Este proyecto se centra básicamente en lo referente a municipalidades donde hemos tomado a continuación la Municipalidad Mariscal Cáceres De San José ubicado en la provincia de Camaná, debido a la gran cantidad de data histórica existente en el sistema con la cual trabaja el área de rentas debido a sus contribuyentes, al no estar cuantificados de manera detallada, genera un desorden en la información al no estar tratada de una forma óptima.

Por otro lado el presupuesto estimado para el año actual o posterior es sobreestimado o en muchos casos insuficiente debido a que se calcula al azar, por último al no tener clasificados a los contribuyentes no se los puede manejar eficientemente en todo caso sería más sencillo presentarlos en el mapa del distrito para así conocer sus ubicaciones sobre todo de los que no contribuyen presentando un retraso en sus impuestos.

Es por ellos que con el uso de este proyecto se podrá solucionar esas dificultades y manejar a los contribuyentes atrasados que generan problemas de información en la Municipalidad.

**Hipótesis**

Dada la utilización de la Minería de datos, entonces es posible identificar a los tipos de contribuyentes y hacer predicciones exactas para el próximo periodo.

**MATERIAL Y MÉTODOS**

**Material**

1. *Data histórica:* Base de datos de la municipalidad distrital Mariscal Cáceres San José – Camaná.
2. *Programas:* SQL server, Microsoft Visual Studio 2010, Microsoft Office Excel 2007, Full converts Enterprise.

**Métodología**

1. *Integración y recopilación de la data histórica:*

Actualmente la Municipalidad posee una data histórica, la misma que se encuentra en un SGBD Ibexpert, el cual se encuentra estructurado de tal forma que nos permita acceder al detalle de la información que contiene, por tanto; al obtenerla encontramos varias tablas las cuales identificamos las principales para así hacer uso de estas.

El esquema metodológico de trabajo que presentamos a continuación, muestra la forma secuencial como se procede al tratamiento de la información contenida en los repositorios de datos:

1. *Pre-procesamiento:*

Se procede a migrar toda la data del SGBD Ibexpert al SGBD SQL Server, donde efectuaremos las consultas que nos conlleven a la obtención de los elementos para la toma de decisiones planteadas.

1. *Selección Limpieza y Transformación:*

Una vez migrada la data se realizara el proceso de ETL (Extrayendo las columnas principales, Transformándolas a formatos consistentes y Limpiando datos nulos y/o repetidos), generando nuestro datamart.

1. *Minería de Datos y Cubos OLAP:*

Utilizando minería de datos tendremos un datamart con el modelo tipo estrella donde se mostrarán nuevas tablas dimensionadas, así como también tablas donde estarán las coordenadas, patrones y la neurona en sí; el objetivo de esta fase es producir nuevo conocimiento que se pueda utilizar, y de esta forma; se podrá visualizar mediante un cubo con diferentes vistas datos a detalle y concisos.

1. *Evaluación e Interpretación*

En esta fase los patrones descubiertos tendrán tres cualidades: Precisos, comprensibles e interesantes. Para ello se evaluará lo obtenido e interpretará para un fin donde se pueda explotar mejor el conocimiento.

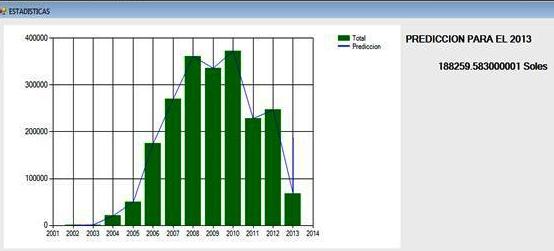
1. *Difusión y uso:*

Con los resultados ya adquiridos se procederá a su difusión es decir que se distribuya y se comunique, donde el nuevo conocimiento extraído deberá integrar el know-how del municipio de esta manera estará apta para tomar las decisiones adecuadas para así tener un mejor manejo de sus contribuyentes, para ello se presentaran los resultados en forma totalizada y disgregada por tipos, en un estadio del tiempo o su progresión a través del tiempo.

**RESULTADOS**

Los resultados obtenidos son los siguientes:

1. *Visualización de los diferentes resultados de la aplicación al utilizar Minería de datos y cubos OLAP.*



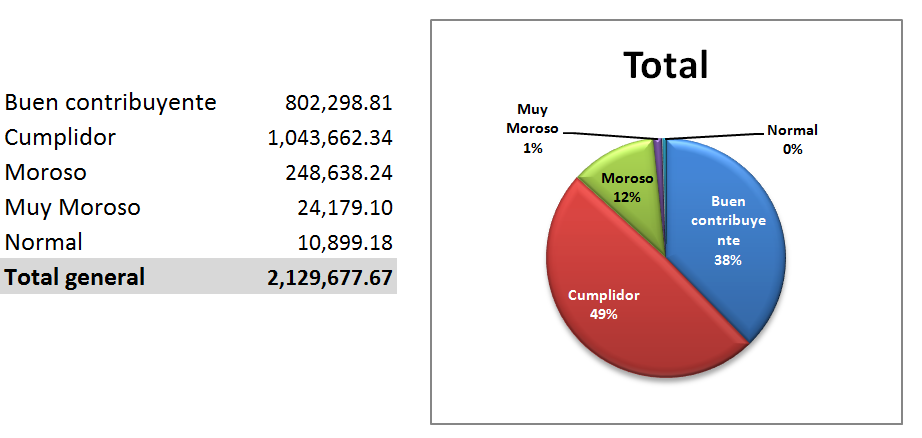
**Figura 1. Cuadro Estadístico del pronóstico para los siguientes años**

Con la data histórica presente se podrán realizar pronósticos anuales de recaudación para los años siguientes, de tal manera llevar a cabo un mejor control y manejo del dinero para así distribuirlo de manera óptima.



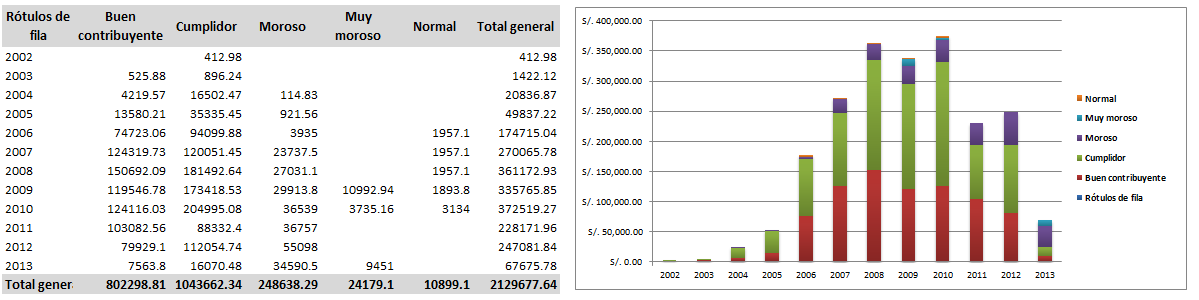
**Figura 2. Mapa donde se muestran ubicados los tipos de contribuyentes**

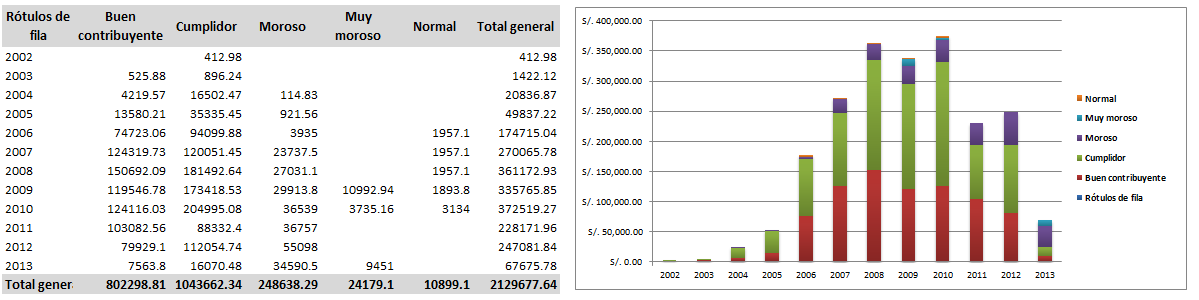
De acuerdo con nuestra red neuronal y la catalogación por tipos para con los contribuyentes se podrá visualizar en el mapa del distrito perteneciente a la Municipalidad la relación de los mismos pero en colores, de tal manera poder distinguirlos en 5 tipos: Cumplidor, moroso, buen contribuyente, normal y muy moroso, para así tener una mejor ubicación de cada tipo.



**Figura 3. Diagrama de los tipos de contribuyentes con el monto total de todos los años de recaudación**

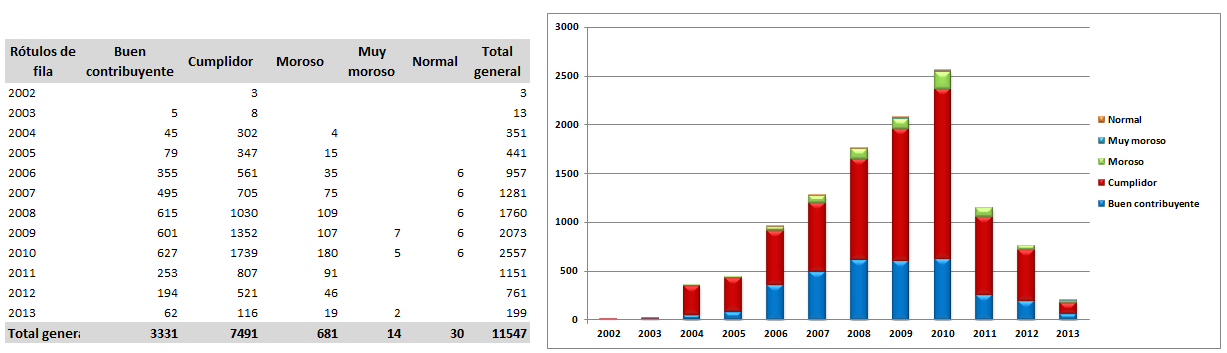
Con este diagrama se puede observar a los tipos de contribuyentes con los montos totales desde el 2002 hasta el 2013, se puede distinguir que el tipo de contribuyente “Cumplidor” (Aquel que cada año paga sus Autoavalúos) genera un mayor ingreso (49%), seguido del “Buen contribuyente” (Principales contribuyentes “Prico” poseedores de mas predios o extensiones de terrenos) (38%), el tipo “Moroso” (demora 3 a 4 años en sus pagos) es representado en menor proporción (12%) lo cual es bueno y con el tiempo tenderá a disminuir si se toman mejores decisiones en el área de rentas contribuido por el alcalde, los del tipo “Muy moroso”(los que se han registrado y solo algunos han pagado) está representado en mínima proporción (1%) lo cual se deberá manejar junto a los morosos, por último los del tipo "Normal” (los que pagan con retraso de 1 año) no genera porcentaje puesto que es escaso el ingreso generado al no presentar gran recaudación en los 11 años.

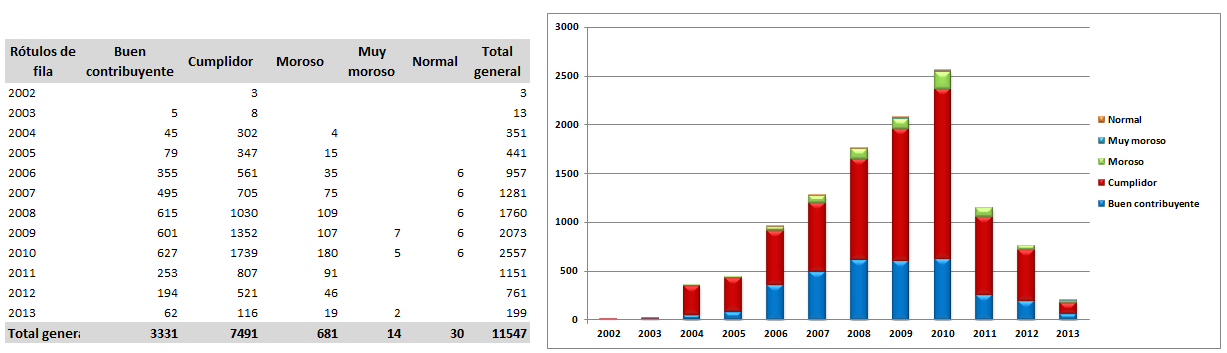




**Figura 4. Diagrama de los tipos de contribuyentes con el monto total acumulado por cada año**

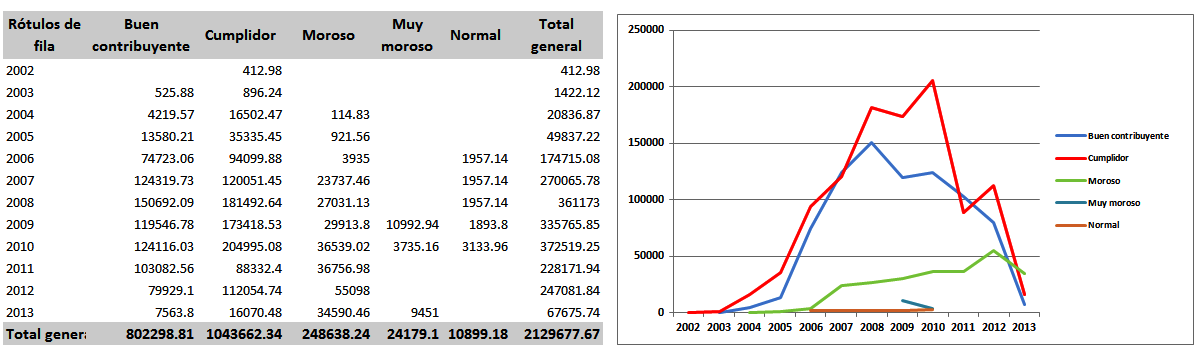
En esta sección se pueden observar los tipos de contribuyentes y sus pagos realizados por cada año donde se muestra que el tipo “Cumplidor” abarca un mayor porcentaje es decir que las personas optan por realizar sus pagos anuales y no dejar pasar 2 o 3 años, etc.; por otro lado los del tipo “Buen contribuyente” presenta un crecimiento y va casi a la par con los del tipo “cumplidor”, lo que significa que los “Prico” generan un elevado ingreso en las recaudaciones puesto que son los que poseen más predios o extensiones; los del tipo “Moroso” a tomado un tercer lugar y a tenido un incremento considerable a partir del año 2007 puesto que ellos prefieren dejar pasar 3 o 4 años para así luego cancelar todo, esto se debe porque aquellas personas no radican en la zona o por motivos económicos optan reunir el monto generalizado para así luego cancelar todo junto; los del tipo “Muy moroso” empiezan a formar parte a partir del año 2009, 2010 y 2013 con montos no tan elevados como los anteriores tipos puesto que los de este sector son las personas que solamente se han registrado y que solo algunas empezaron a pagar sus Autoavalúos, con el tiempo este tipo de contribuyentes tenderá a desaparecer para formar parte de los demás o simplemente disminuirá; por ultimo tenemos los del tipo “Normal” ellos presentan un ingreso mínimo puesto que dejan pasar 1 año al ser conformado como minoría igual contribuyen.

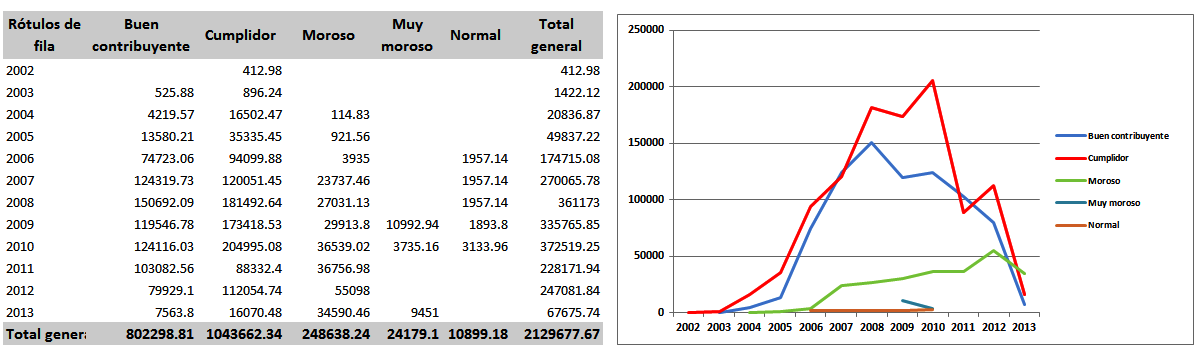




**Figura 5. Diagrama de los tipos de contribuyentes con el número de personas que lo conforman**

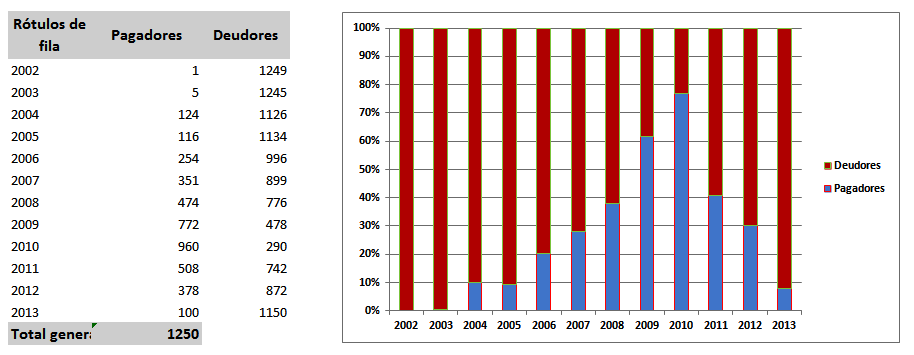
En este diagrama se puede visualizar al total de personas por tipos que han contribuido por cada año, se nota claramente que los del tipo buen contribuyente toman un segundo lugar después de los del tipo cumplidor pero generan casi los mismos ingresos que ellos debido a que lo conforman los “Prico”.

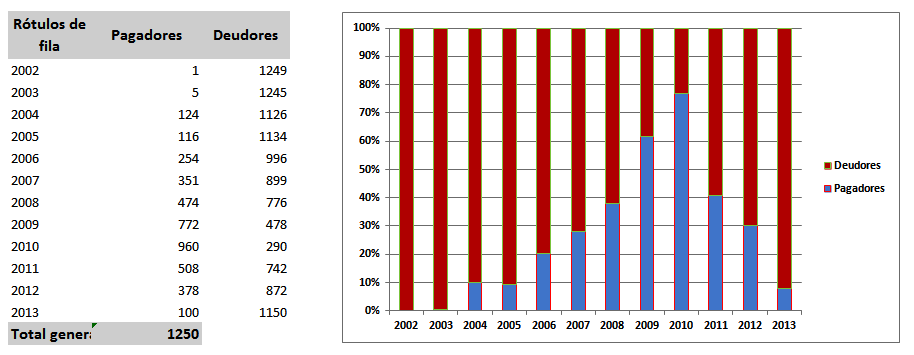




**Figura 6. Diagrama donde se muestra la tendencia de los tipos de contribuyentes del distrito**

En este diagrama se puede visualizar una tendencia por parte de los tipos de contribuyentes a través del tiempo (por cada año), donde los del tipo “Cumplidor” toma una buena alza hasta el año 2010 pero luego muestra una caída en sus Autoavalúos, por otra parte los del tipo “Buen contribuyente” también generan un alza en la recaudación pero solo hasta el año 2008 para así luego descender, los del tipo “Moroso” van tomando un alza a través de los años en sus recaudaciones y probablemente siga así puesto que los de este tipo prefieren dejar pasar 3 o 4 años para luego cancelar, seguido se encuentran los del tipo “Normal” quienes se tornan estables, por ultimo tenemos a los “Muy moroso” quienes presentan una caída pero con el tiempo tendera a desaparecer puesto que son las personas que están registradas y solo algunas han cancelado.





**Figura 7. Diagrama donde se muestra a los que han contribuido y no contribuido por año**

En este gráfico se puede visualizar la cantidad de personas que han contribuido y las que no han contribuido anualmente, notándose claramente que en el año 2010 la mayoría de contribuyentes cancelaron sus Autoavalúos esto se debe que para ese entonces se lanzo campaña tributaria (amnistía) donde las personas haciendo uso de ella decidieron cancelar sus deudas, recaudándose una suma mayor que la de los años anteriores y posteriores a ella.

**ANÁLISIS Y DISCUSIÓN**

Los resultados obtenidos muestran en primer lugar un mejor manejo de los contribuyentes al tenerlos ya clasificados como “Morosos, Muy morosos, Normal, Cumplidor y Buen contribuyente. Segundo el sistema puede predecir la recaudación para el próximo año basándose en las recaudaciones anteriores.

Durante la construcción de nuestro datawarehouse se tomaron en cuenta y crearon algunas tablas esenciales para dimensionarlas y así proceder con la creación de nuestro modelo estrella. El análisis de la aplicación de minería de datos nos permitió obtener respuesta en tiempo real a la consulta de datos lo cual no genera un retraso sobreestimado ni pérdida de la data de manera que al momento de visualizar consultas de manera entendible y de fácil acceso se procedió a elegir la plataforma del Excel puesto que es la que están acostumbrados a manejar en el área de rentas.

Para nuestro cubo OLAP la cantidad de datos obtenidos es eficiente puesto que de él se puede extraer datos relevantes como la cantidad de contribuyentes morosos, así como la clasificación distribución y ubicación de los mismos en el mapa del distrito generando de tal forma un valor agregado a la municipalidad.

**CONCLUSIONES**

1. Se concluye que se pudo desarrollar el sistema de Soporte para la Toma de Decisiones en el cobro de Autovalúos y pronóstico de obras para la Municipalidad Distrital Mariscal Cáceres - San José en la provincia de Camaná, utilizando Cubos OLAP y Minería De Datos de manera eficiente, alcanzando un nivel de eficacia del 98%.
2. Dada la utilización de la Minería de datos, cubos OLAP, redes neuronales y progresión aritmética es posible identificar a los 5 tipos de contribuyentes en el mapa de la provincia, también se pudo desarrollar el hacer predicciones para el próximo año, realizar reportes, hacer uso de gráficas para una mejor visualización y manejo estadístico de los contribuyentes, lo cual le permitió al autor de decisiones, afrontar medidas para la mejora de la recaudación tributaria.

**REFERENCIAS BIBLIOGRÁFICAS**

* Javier Nader, Sistema de apoyo gerencial: En: http://www.itba.edu.ar/archivos/secciones/nader-tesisdemagister.pdf, argentina, 2008, 456pp.
* Javier Nader, Sistema de apoyo gerencial: En: http://www.itba.edu.ar/archivos/secciones/nader-tesisdemagister.pdf, argentina, 2008, 456pp.
* Claudio Cesar, Data warehousing. En: http://personal.lobocom.es/claudio/gen006.htm, Argentina, 2007.
* Méndez Martinez, Andrea, fundamentos de dataWarehouse, En: http://artemisa.unicauca.edu.co/~ecaldon/docs/bd/fundamentosdedatawarehouse.pdf. instituto tecnológico de buenos aires, centro de actualización permanente en ing. de software, 2008. 8pp.
* Luis garrido, Minería de datos, En: Helsinki University of Technology, Finland, Document Server (Finland), Finlandia, 2006, 611-626pp.
* Cubos OLAP, En: http://es.wikipedia.org/wiki/Cubo\_OLAP, 26 abril 2012.
* Pedro Larrañaga, Redes Neuronales, Departamento de ciencias de la computación e inteligencia artificial, universidad país Vasco, España, 2005, 19pp
* Vollmann, T.E.; Berry, W.L.; Whybark, D.C. (1997). Manufacturing Planning and Control Systems. 4th ed. Irwin / McGraw-Hill.

**Agradecimientos:**

El presente trabajo se logró gracias a la coordinación con la municipalidad Mariscal Cáceres San José. Arequipa - Camaná con quienes se logro establecer un nexo de coordinación y mutuo apoyo, a quienes se les agradece por brindar todo el apoyo logístico para la consecución y éxito del presente proyecto, para quienes va nuestro reconocimiento y profundo agradecimiento.

1. Bachiller en Ingeniería de Sistemas e Informática de la UAP. E-mail: [yoshikarin\_2008@hotmail.com](mailto:yoshikarin_2008@hotmail.com) [↑](#footnote-ref-2)
2. Dr(c) en Ciencias de Computación. Docente en la Universidad Alas Peruanas Filial Arequipa y Universidad Nacional de San Agustín. E-mail: vcornejo5@hotmail.com [↑](#footnote-ref-3)