

ANALÍTICA DE DATOS PARA SISTEMAS DE COSTOS BASADOS EN ACTIVIDADES EN LA ERA DE BIG DATA

Juan Bernardo QUINTERO
Universidad de Medellín
(Colombia)

David Manuel VILLANUEVA
Universidad de Medellín
(Colombia)

Fernando Luis GÓMEZ MONTAYA
Universidad de Antioquia
(Colombia)

RESUMEN:

La acelerada evolución de las tecnologías de la información y las comunicaciones, conjuntamente con el aumento del acceso a las mismas, constituyen factores fundamentales para iniciar la era de *Big Data*; una tendencia en la que el análisis de la información debe considerar grandes volúmenes de datos, de diversos orígenes y formas, que deben ser procesados con gran velocidad. Por otro lado, los sistemas de costos y en particular los sistemas de costos basados en actividades, han posibilitado a las compañías conocer su verdadera situación a través de metodologías que permiten obtener el costo de los procesos, subprocesos y actividades, facilitando un consumo efectivo y eficiente de los recursos. La operación de los modelos de costos basados en actividades permite a las compañías realizar analítica descriptiva para conocer el estado actual de la empresa; sin embargo, hacer analítica predictiva para conocer el estado futuro, sugiere considerar información histórica, lo cual puede generar considerables volúmenes de datos tratables fácilmente a través de herramientas *Big Data*. Por su parte, la analítica prescriptiva para sacar mayor provecho a situaciones futuras, hace necesario considerar información de terceros, de redes sociales o de *streaming* de audio y video, para cuyo tratamiento se hacen indispensables las herramientas disponibles en entornos *Big Data*. Este artículo muestra el potencial de integrar disciplinas como el costeo basado en actividades y la analítica de datos en el marco de *Big Data*, ilustrando su implementación en el caso de aplicación del análisis de datos de una universidad.

Palabras claves: Analítica de Datos, *Big Data*, Costeo Basado en Actividades.

INTRODUCCIÓN

El crecimiento exponencial que han experimentado los servicios y posibilidades que brinda Internet, tiene como efecto considerables incrementos en el volumen de datos, en la variedad de formatos de la información y en la velocidad con que es necesario procesar dicha información; este es el contexto en el que surge el fenómeno de los datos masivos o como es denominado en inglés *Big Data*. Al disponer de mayor volumen y variedad de datos, es posible ampliar el espectro de posibilidades en el análisis de información, situación que le puede permitir a las organizaciones conocer más a fondo su verdadera situación, y mejorar el detalle al momento de soportar la toma de decisiones.

Una de las técnicas que pueden usar las organizaciones para analizar la información son los modelos de costeo basados en actividades (llamados ABC por la sigla en inglés de *Activity-Based Costing*), permitiendo conocer los verdaderos costos de los productos o servicios al analizar no solo los recursos y materias primas utilizadas, sino también el esfuerzo realizado para producirlos y demás costos no identificables fácilmente, en los cuales se incurre durante el ciclo de vida de los productos o servicios.

Los modelos de costeo basado en actividades se caracterizan por involucrar diversos elementos de las organizaciones: recursos o cuentas contables, dependencias operativas, dependencias administrativas, macro-procesos, procesos, actividades y los productos o servicios también llamados objetos de costos. Al momento de asociar costos y gastos a cada uno de estos elementos también se deben tener en cuenta las relaciones que existen entre ellos; por ejemplo, cuál es el costo que una dependencia asocia a un proceso, o el monto que una actividad asocia a un producto o servicio. Toda la información que se deriva de estas relaciones se debe almacenar en bases de datos; y si se consideran empresas de tamaño medio o grande que tengan cientos de cuentas, que estén estructuradas en decenas de dependencias, que realicen cientos de actividades y que ofrezcan diversidad de productos; surge entonces la necesidad de almacenar grandes volúmenes de información que se ajustan al marco de *Big Data*, más si se considera que dicha información se debe almacenar periódicamente y es necesario disponer de diversos periodos para hacer análisis predictivos.

El uso de entornos *Big Data* como plataforma de implementación para modelos ABC presenta dos ventajas notorias; la primera tiene que ver con la posibilidad de almacenar grandes volúmenes de información, permitiendo establecer mayor número de elementos a los que se le asocian costos y gastos, e igualmente mayor número de relaciones entre dichos elementos; de esta forma se amplía el espectro de análisis que se pueden realizar. La segunda ventaja tiene que ver con la posibilidad de aplicar técnicas de aprendizaje de máquina para realizar análisis predictivos y prescriptivos basándose en la información de los modelos ABC.

Para definir las estrategias de integración de los modelos ABC y la analítica de datos, este artículo presenta la siguiente estructura: el capítulo dos introduce y

explica la analítica de datos, el capítulo tres analiza el costeo basado en actividades, el capítulo cuatro plantea las consideraciones para integrar los modelos ABC y *Big Data*, el capítulo cinco describe el papel de las herramientas en ABC y en *Big Data*, el capítulo seis presenta un caso de estudio de la integración de estas dos disciplinas, y el capítulo final presenta las conclusiones y propone algunos trabajos futuros al rededor del tema.

Analítica de Datos

El término “analítica” es una palabra de moda (*buzzword*), frecuentemente acompañada de expresiones complementarias para ampliar su significado; analítica de texto, analítica en memoria o analítica predictiva, son algunas de las frases que se encuentran de forma recurrente cuando se habla de tecnologías emergentes en el área de informática (LeHong & Fenn, 2012). Anteriormente se utilizaba el término “análisis” a secas; sin embargo, el concepto de analítica hace uso extensivo de la matemática y la estadística para complementar el análisis, como sucede por ejemplo en los modelos de predicción.

Por su parte, la “analítica de datos” es la disciplina que se encarga de explorar, descubrir e inclusive interpretar patrones en los datos, con el propósito de sacar conclusiones. Cuando estas actividades se realizan en función de soportar la toma de decisiones de negocio o procuran generar valor a una organización, se habla de “analítica de negocio” (Shmueli, Patel, & Bruc, 2016).

La analítica de datos suele ser responsabilidad de un rol denominado “científico de datos”. Este rol necesita de competencias en diversas áreas: primero debe conocer en profundo el manejo de las tecnologías de información y comunicaciones, segundo tener conocimiento en matemáticas y estadísticas y tercero tener experiencia en un dominio de negocio (Drew Conway Data Consulting, 2013); esto tiene como resultado que el científico de datos ha sido llamado el trabajo más sexi del siglo XXI (Davenport & Patil, 2012). La figura 1 ilustra cómo la combinación de las habilidades del científico de datos define áreas de conocimiento relacionadas como la investigación tradicional y el aprendizaje de máquina.



Figura 1. Competencias del científico de datos

El científico de datos realiza actividades que conducen a diversas modalidades de análisis de información, lo cual deriva en tres tipos de analítica de datos: descriptiva, predictiva y prescriptiva (Davenport, 2013).

Analítica descriptiva (Analítica 1.0 - la era de “Business Intelligence”): se centra en encontrar explicaciones a lo que sucedió o está sucediendo. En esta era se da el apogeo de la inteligencia de negocios, y se fortalecen las prácticas de las llamadas bodegas de datos (*Data Warehouse*). Dichas prácticas incluyen las labores de ETL (*Extract, Transform and Load*), diseño de tableros de control (*Dashboard*) y construcción de reportes (*Reporting*). Los desarrollos en este tipo de analíticas se dirigen por las tecnologías, lo cual, dio paso a la aparición de diversas plataformas de inteligencia de negocio.

Analítica predictiva (Analítica 2.0 - la era del “Big Data”): se centra en determinar con cierto nivel de precisión lo que puede suceder. Inicia con el surgimiento de la minería de datos (*Data Mining*), para posteriormente complementarse con los modelos de predicción apoyados por las matemáticas, la estadística y el aprendizaje de maquina (*Machine Learning*) e incluyen por ejemplo la identificación de patrones. Los desarrollos en este tipo de analíticas se dirigen por los datos, lo cual dio paso a la aparición de las plataformas para tratar grandes volúmenes de datos sacando provecho del Procesamiento en Paralelo Masivo (MPP: *Massively Parallel Processing*) como es el caso de Hadoop, uno de los protagonistas de los entornos *Big Data*.

Analítica prescriptiva (Analítica 3.0 – la era de la “Oferta de Datos Enriquecidos”): adicional a realizar predicciones, se diseñan modelos que apoyan el planteamiento de acciones a seguir para sacar provecho o evitar resultados ad-

versos derivados de dichas predicciones. Este tipo de modelos resultan de gran utilidad en el soporte a la toma de decisiones e incluyen por ejemplo el reconocimiento de sentimientos en audio. Los datos enriquecidos provienen de diversas fuentes como redes sociales, audio, video, dispositivos móviles e incluso sensores en el marco de Internet de las Cosas (IoT: *Internet of Things*). Los desarrollos en este tipo de analíticas se dirigen por la ciencia, lo cual tiene como efecto el fortalecimiento del rol del científico de datos.

Otra estrategia de clasificación de la analítica de datos, se puede basar en el tipo de análisis que se realiza; diferenciando el análisis endógeno para entender las dinámicas internas de la organización, del análisis exógeno para conocer la forma en que la organización se relaciona con su entorno:

- En el frente del **análisis endógeno** se destaca la “analítica contable” (comúnmente llamada contabilidad analítica) la cual se basa en la información financiera para determinar la forma en que la organización consume y recibe recursos.
- Las prácticas de **análisis exógeno** cubren por ejemplo la “analítica web” para conocer la actividad en Internet de una organización o la “analítica social” para entender su posicionamiento en redes sociales.

Este artículo muestra la forma de potenciar el análisis endógeno en una organización usando la “analítica de costos” a través de la integración de una técnica de aprendizaje de máquina, como las redes neuronales artificiales y un modelo de contabilidad analítica como el costeo basado en actividades.

Costeo basado en actividades

La literatura alrededor del “Sistema de Costeo Basado en Actividades – ABC- (por sus siglas en inglés Activity Based Costing), que surgió a finales de la década de los 80’s, como respuesta a las necesidades de las empresas de establecer un costo para sus productos que les permitiera tomar decisiones más adecuadas, y solucionar el problema de los costos indirectos, es así como en sus inicios se mencionó que era una metodología para distribuir estos costos” (Jara, Castañeda, & Gómez, 2003), las empresas encontraron en esta metodología una posibilidad de costear con mayor grado de precisión los productos o servicios teniendo en cuenta las limitaciones de las metodologías tradicionales, tales como las distribuciones de costos indirectos o la asignación de costos a través de todo el ciclo de vida del producto.

Esta metodología en principio logró mostrarse cómo una posibilidad de mejoramiento en la distribución de los costos indirectos de fabricación al punto de que algunos teóricos, sostenían que era exclusiva para la distribución de indirectos, sin embargo tomando su definición, el costeo basada en actividades, en su concepción principal establece que “los recursos no son consumidos por los productos o servicios en forma directa, sino que estos son consumidos por las acti-

vidades desarrolladas en la empresa y que estas a su vez son consumidas por los objetos de costo” (CGN, 2008), es decir bajo esta definición básica, se pueden establecer tres niveles iniciales de información: recursos, actividades y productos o servicios. Sin embargo, dicha definición no puede ser tan aferrada a la teoría al momento de implementar un sistema de gestión de costos, ya que el modelo puede adaptarse a cualquier entidad, teniendo en cuenta sus particularidades, y por ende utilizar niveles de información adicionales para estructurar unas asignaciones de costos más ajustadas a la realidad de cada empresa.

Algunos conceptos importantes mencionados en la definición del ABC, son los recursos, las actividades y los objetos de costos finales (Cuervo, Osorio, & Duque, 2013). Los recursos hacen referencia a cualquier costo o gasto en el que incurra una entidad para el normal desarrollo de sus funciones, tales como gastos de personal, mantenimiento, reparaciones, seguros, depreciaciones etcétera, aquí se le resta importancia a la separación contable entre costos (lo directamente relacionado con el producto o servicio) y los gastos (lo relacionado con la administración del negocio), ya que todos estos elementos son considerados igualmente importantes al momento de realizar el costeo. Las actividades corresponden a al trabajo realizado por las personas en la empresa, que están orientadas a satisfacer necesidades internas o externas, las cuales están agrupadas en procesos bien sea aquellos relacionados directamente con la misión de la empresa o aquellos que contribuyen a mantener la estructura normal de operación (de soporte o apoyo), y los objetos de costos, hacen referencia a los productos o servicios que se quieren costear en la empresa,

En el caso de las distribuciones en este sistema de costeo, se utiliza el “concepto: inductor, conductor, base de asignación, base de distribución, direccionador o cost driver, que en esencia corresponde a un criterio de aplicación o distribución de costos el cual determina cómo se distribuye la carga de trabajo o los recursos de la compañía a las actividades y objetos de costo, y pueden influenciar directa o indirectamente el costo” (Duque, Gómez, & Cuervo, 2005) dichos criterios de costos que facilitan las asignaciones, es así como se puede establecer el costo, mediante el costeo de las actividades que de manera directa o indirecta tienen incidencia en los productos o servicios.

Es preciso resaltar que en Colombia la metodología de costeo basado en actividades empezó a tener difusión a mediados de los 90, pero tuvo mayor auge a partir del año 1997, cuando por primera vez, una entidad gubernamental (la Superintendencia de Servicios Públicos Domiciliarios) emitió la Resolución 1417, que exigió la implementación de sistemas de costeo basado en actividades para las empresas de servicios públicos domiciliarios (SSP, 1997), sector en el que se vieron obligadas a implementar el modelo más de 3.000 empresas entre públicas y privadas, las cuales estaban bajo el ámbito de aplicación de dicha resolución.

El caso de estudio que se analizará está orientado a Universidades, para las cuales la implementación de sistemas de costos no es exigida de manera explícita, sin embargo, el factor más preponderante a la hora de implementar un sistema de

costos, no es la obligación legal de tenerlo, de hecho esto hace que se pierdan sus bondades o que la entidad se preocupe únicamente por hacer la tarea de entregar información para entes externos, por esto lo más importante a la hora de establecer un sistema de costeo, es la posibilidad que tiene la Universidad de disponer de información adecuada para tomar decisiones en cuanto a sus principales productos o servicios, clasificados estos en Docencia, Investigación y Extensión.

Niveles de un modelo de costos: para toda implementación de un modelo de costeo basado en actividades, siempre se establecen una serie de pasos o fases que comprenden todos los elementos necesarios para disponer finalmente del modelo que suministre la información de costos a múltiples niveles de detalle, pero que además permita realizar los análisis de costos, ya que quizás este es el principal trabajo que debe hacerse al respecto, es decir, para aquellas empresas que se ocupan únicamente de calcular sus costos y que no hacen análisis y propuestas de mejoramiento, el modelo no pasará de ser una simple herramienta que se quede almacenada en el escritorio de los directivos y que en algunos casos, permitirá el reporte de datos a entidades externas, por esto, siempre se ha manifestado la necesidad de disponer de información que permita tomar decisiones sustentadas en costos y por tanto el sistema siempre ha sido concebido en función de su utilidad interna.

La estructura del modelo de costos, corresponde a la representación gráfica del flujo de información a través de todo el sistema y para cada empresa dicha estructura debe establecerse en función de las necesidades de información de la misma, para tal efecto se establecen niveles en el sistema, para que los costos se vayan asignando entre ellos, por ejemplo el modelo más básico derivado de la definición del ABC, se pueden visualizar en tres niveles, Recursos, Actividades y Objetos de Costos, como se ilustra en la figura 2.

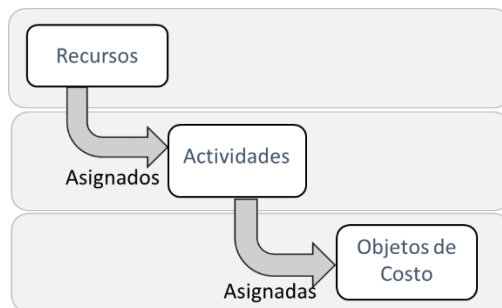


Figura 2. Definición del Costeo Basado en Actividades

Partiendo de la definición de niveles básicos en el sistema, debe identificarse cuáles son los niveles que acorde con las necesidades de la empresa, se requieren utilizar para la estructuración del sistema de costos, tales como Centros de Costos Especiales, Centros de Costos Distribuibles, Centros de Costos Contables, Centros Misionales, Centros Administrativos, Actividades, Procesos, Servicios Intermedios Servicios Finales, Productos etcétera, los posibles niveles mencionados se establecen en la medida en que la entidad lo requiera.

Como puede evidenciarse, los anteriores elementos en su mayoría son comunes a la implementación de un sistema de costeo basado en actividades, sin embargo, cada uno tiene su particularidad en razón de la entidad para la cual se esté diseñando el sistema.

Estructura del Sistema de Costos: como se mencionó, el modelo de costos depende de las características particulares de cada entidad, es así como los niveles de información que se usan, varían frente a la definición original del costeo basado en actividades. En adelante se presenta el detalle de un modelo genérico de costos para el caso de las Universidades, teniendo en cuenta que el caso práctico que se explicará, está relacionado directamente con estas instituciones. Para el caso particular de las Universidades, la estructura de costos, corresponde a una representación gráfica del flujo de costos a través de todo el sistema, dicha representación puede cambiar de acuerdo con las necesidades y particularidades de cada universidad, de manera general se han establecido unos niveles principales de información para el sistema de costos en una Universidad, aunque no quiere decir que esto sea una camisa de fuerza, ya que se pueden hacer variaciones, que permitan mejorar la visualización de los costos y por ende su distribución. Se presenta una estructura genérica del modelo, que permite incorporar todos los elementos a evaluar en el sistema, como se ilustra en la figura 3.

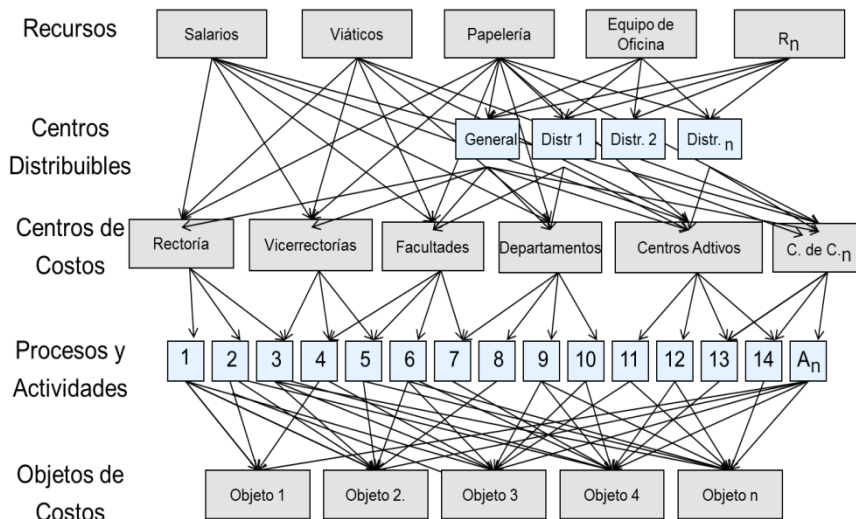


Figura 3. Estructura del Sistema de Costos

La figura anterior se puede analizar por niveles de información, ya que cada uno de los 5 niveles presentados, permite reportar datos de costos y a su vez muestra las relaciones que se establecen entre ellos, en el nivel de **Recursos**, se tienen todos los costos y gastos en concordancia con el plan de cuentas o utilizando recursos que agrupan cuentas que presenten el mismo comportamiento en cuanto a su posibilidad de asignación, los **Centros de Costos Distribuibles**, se utilizan

para agrupar todos los recursos que en una primera distribución deben asignarse mediante criterios a otros niveles, el de **Centros de costos** se utiliza para calcular los costos de todas las áreas, o centros específicos para los cuales se requiera conocer su costo, en este normalmente se encuentran las áreas de la estructura organizacional y vehículos, espacios físicos, proyectos, etcétera, a los cuales se les quiera realizar un seguimiento; en el nivel de **Procesos y actividades**, se incorpora todo el diccionario de actividades que recibe los costos distribuidos de los niveles superiores y finalmente todas estas actividades o procesos son asignados al nivel de **Objetos de costo**, en el cual se establecen todos los productos o servicios que en última instancia son costeados en el modelo.

El modelo planteado para Universidades, es genérico, debido a que las particularidades de cada una de estas, pueden requerir diferentes niveles de detalle para garantizar una mayor precisión al momento de establecer el costo.

Consideraciones de integración de ABC y Big Data

Para plantear un escenario de integración del modelo de costos ABC y *Big Data* se hace necesario definir el modelo desde las características que lo hacen un caso aplicable a los métodos de un entorno como el de *Big Data*.

Variedad, velocidad y volumen (Las tres V)

El modelo ABC requiere de datos que pueden provenir de diferentes fuentes, la razón es que las empresas, especialmente las grandes y medianas, están conformadas por diferentes partes que trabajan para unos objetivos en común, esta es su definición de origen sistémico. En consecuencia, para un análisis cualquiera producto de la distribución de costos de la empresa, se puede necesitar información de sistemas, por ejemplo, gestión de talento humano, contable, proveedores, transaccionales, expertos, entre otros. Lo anterior sugiere que los datos tienen una naturaleza variada desde su fuente y amerita un tratamiento adicional para el procesamiento en el modelo.

Adicionalmente, esta diversidad de fuentes también sugiere preguntarse por la velocidad con la que se produce información aprovechable por el modelo ABC por parte de las empresas. La respuesta puede variar de un tipo de compañía a otra, pero en general las compañías tienen actividades que consumen recursos diariamente o en tiempo real, lo que implica que de una forma u otra, una distribución de costos requiere de información constantemente para actualizarse y eso solo pensando en las actividades efectuadas por el personal, habría que pensar en las transacciones diarias por objeto de compras a proveedores o los servicios de mantenimiento frecuente para soportar el negocio.

Por su parte el volumen está relacionado con la forma computacional como se representaría estructuralmente el modelo ABC, es decir que, para implementar un modelo ABC a través del tiempo (luego de uno o muchos periodos históricos) implica procesar grandes volúmenes de información.

Suponiendo que el modelo se representa mediante un Sistema Gestor de Base de Datos –SGBD- tradicional, la estructura resultante debería permitir guardar un registro en una tabla por cada traza que evidenciaría la relación de costos desde un recurso, pasando por los niveles intermedios que utilizó hasta el objeto de costo que genera. Todos los elementos que se utilizan en el modelo y que se utilizan para recibir o distribuir costos en los diferentes niveles se denominaran en adelante colectores y serán clasificados acorde con el nivel respectivo La figura 4 muestra un ejemplo de los registros que generaría representar dos niveles únicamente.

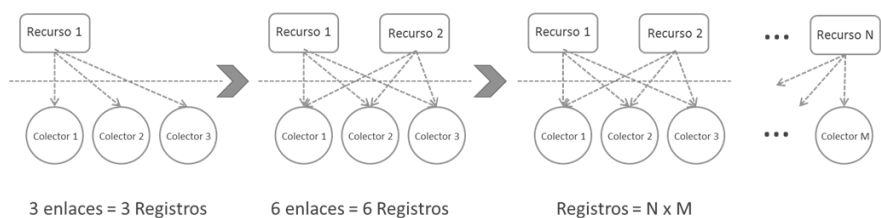


Figura 4. Crecimiento de registros de acuerdo con el número de colectores de los niveles

En general, el número de registros potenciales necesarios estaría dado por la multiplicación del número de elementos de un nivel, por el número de elementos del nivel siguiente y así sucesivamente hasta el nivel final, por lo tanto, habría tantos registros como combinaciones de elementos con subniveles existan en la representación:

$$R * CD * CC * A * OC = \# \text{ de registros potenciales}$$

R: #Recursos (nivel 1), CD: # de centros de distribución (nivel 2), CC: # de centros de costos (nivel 3), A: # número de actividades (nivel 4) y OC: # de objetos de costos (Nivel final)

El cálculo anterior sería un escenario de necesidad extrema, en donde cada elemento de un nivel necesitara relacionarse con todos los elementos del nivel siguiente, lo cual puede o no ocurrir, sin embargo y a pesar de que la estructura debe tener una profundidad (niveles) fijos también se debe proveer almacenamiento para los enlaces y tener en cuenta guardar todo esto en periodos que sean aprovechables para la empresa, usualmente mensual, trimestral o semestralmente.

Redes neuronales artificiales

"Una red neuronal artificial –RNA- es un conjunto de algoritmos matemáticos que procesan información y encuentran relaciones no lineales entre el conjunto de datos, y cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso humano: la neurona" (Caicedo & Lopez, 2009).

Dentro del aprendizaje de máquina las RNA son una técnica que permite recrear el aprendizaje humano propio de las neuronas biológicas mediante símiles artificiales que colaboran para aprender dado un conjunto de estímulos. Las RNA en el ámbito del aprendizaje de máquina, incorporan el símil del aprendizaje a través de un proceso que, dado el cálculo de una salida y el error que representa comparado con datos de entrenamiento, reconfigura cíclicamente el peso de sus enlaces internos hasta producir una configuración que genere una salida con una probabilidad adecuada. Este proceso usa el algoritmo llamado *Backpropagation* y constituye la base del uso de las RNA en el ámbito de la predicción y prescripción analítica.

Semejanza entre elementos del modelo de costos ABC con una red neuronal artificial

Según (Caicedo & Lopez, 2009) la neurona artificial es una abstracción de la neurona biológica en la medida que toma la funcionalidad de algunos elementos como: nodo, entradas, pesos y salidas con su homólogo biológico: neurona, dendritas, sinapsis y axón respectivamente.

Ahora bien, en el modelo de costos ABC, se plantean comportamientos y atributos que permiten plantear que hay elementos que guardan igual similitud con la neurona artificial. Es así como en el modelo ABC, llamamos a estos elementos colector, recursos, costo transferido y transferencia de costo en relación con nodo, entradas, pesos y salidas en la neurona artificial. La figura 5 muestra la relación descrita en el párrafo anterior.

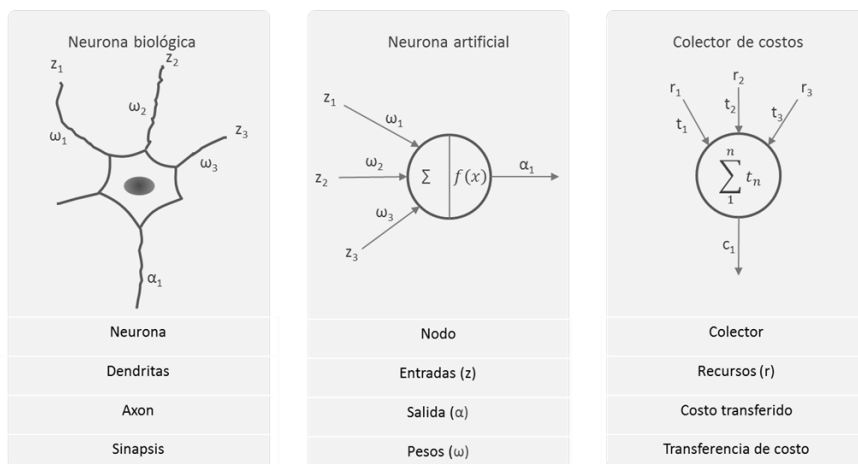


Figura 5. Analogía entre elementos de las RNA y elementos de los modelos de costos

Es necesario agregar que, así como la neurona artificial solo obtiene de la biológica algunas características de la neurona biológica, la definición del colector propuesto puede interpretarse como otra versión simplificada de la neurona biológica y artificial.

Semejanza entre la estructura bajo el modelo de costos ABC con una red neuronal artificial

De igual manera la unidad básica de la neurona biológica produjo una neurona artificial y el colector propuesto, se propone una relación entre colectores agrupados por capas que se alimentan por entradas de recursos y de cuyo procesamiento de información, se producen unas salidas conocidas como el costo de los objetos de costos. En ese orden de ideas, una propuesta para la estructura del modelo ABC, tendría los mismos componentes de las RNA y sus elementos antes descritos interactuarían de manera similar. En la figura 6 se muestran estas semejanzas y como el modelo ABC comparten algunas similitudes estructurales, en esencia se pueden enumerar las siguientes: 1. Entradas, capas intermedias y salidas. 2. En ambas (RNA y ABC) se puede tener horizontalmente hasta n entradas, m elementos para las capas intermedias. Y 3. Salidas producto del proceso interno llevado a cabo al procesar la información de entrada en las capas intermedias. Una representación de esta semejanza se muestra en la siguiente figura.

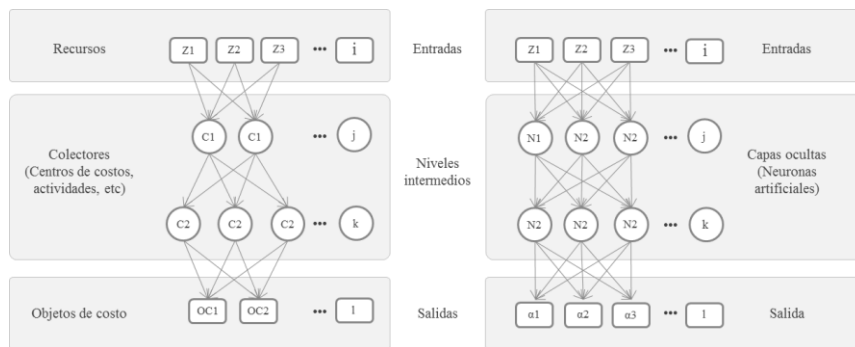


Figura 6. Analogía entre las RNA y los modelos de costos ABC

Una vez establecida la similitud estructural se hace necesario tener en cuenta las siguientes observaciones:

- Un elemento de cualquiera de las capas de un modelo ABC transfiere un valor hacia uno o varios elementos de la siguiente capa a través de un criterio llamado “inductor de costo” (*Cost Driver* en inglés) y en el colector (Neurona en RNA) se realiza una sumatoria de los valores transferidos. A su vez, esta sumatoria se convierte en el valor disponible para transferir hacia otra capa teniendo en cuenta inductores iguales o diferentes. Esto implica que en el modelo el colector a diferencia de la neurona en RNA, no posee una función de transferencia, ni necesita una función de activación para definir la activación del proceso y en consecuencia la salida.
- Debido a lo anterior, subyace otra diferencia funcional, la salida de una neurona artificial es un número entre 0 y 1 que está disponible como entrada para

cualquiera de las neuronas de la siguiente capa. En el modelo ABC la salida del colector en un valor recolectado y de allí su nombre.

- El en modelo ABC el inductor de costo es calculado con criterios contables, mientras que una RNA su símil ω_n es calculado como el resultado de la función de transferencia.

Una red neuronal artificial, es en esencia un grupo de entes (neuronas) que de manera similar a las neuronas biológicas, colaboran para producir un resultado esperado de manera automatizada, una de sus ventajas es que, es un modelo utilizado el ámbito del aprendizaje de máquinas, lo que permitiría introducir la posibilidad de realizar analítica predictiva y prescriptiva sobre datos de costos ABC.

Escenarios de uso e implementación de RNA en el modelo ABC

Al establecer una relación de similitud entre la estructura de una RNA y sus elementos con el modelo de costos ABC, permite que se puedan proponer algunos escenarios en los cuales algoritmos y conceptos que aplican en la primera, proporcionen procesos y ventajas a la segunda. Los siguientes pueden ser algunos de los escenarios propuestos:

- **Distribuciones iniciales:** la implementación del modelo ABC para una compañía exige un trabajo inicial cuidadoso en el cual se plantea la estructura de cara a las características específicas de la empresa. Es posible que al inicio de un periodo no existan todos los inductores de costos calculados, sobre todo los que dependen para su cálculo del tiempo de las actividades que aún no se han llevado a cabo. Para este caso en particular, la distribución de costos podría utilizar la fase inicial del algoritmo *Backpropagation*, para que con inductores de costos de prueba, genere la primera salida de objetos de costo así la distribución general no tenga todos los elementos.
- **Análisis predictivo basado en posibles recursos:** por otra parte, y haciendo uso de datos históricos, se puede plantear la posibilidad de interpretar la estructura del modelo ABC, como unas entradas (recursos) con unas capas intermedias ocultas que generan una probabilidad de salida (objetos de costos). Siendo así los periodos históricos de recursos y objetos de costos pasados podrían utilizarse para alimentar/entrenar una RNA y responder preguntas como: ¿dado unos recursos X, Y y Z, cual es la probabilidad de obtener costos T y W.
- **Optimización de drivers de costos:** este es un escenario muy parecido al anterior, la diferencia es que sitúa la salida probable en un nivel de la estructura del modelo que se desea, de ese modo, la pregunta del escenario anterior cambia a: ¿Dados los recursos/colectores (entradas), cual es la probabilidad de obtener los inductores de salida?. Esta aproximación parte de que el pro-

ceso de aprendizaje mediante el algoritmo *Backpropagation*, garantiza que los pesos ω_n , son los mínimos posibles luego del entrenamiento.

El papel de las herramientas en ABC y en Big Data

Actualmente se ofrecen en el mercado soluciones de costos para empresas bajo la metodología de costos ABC con implementaciones que van desde líneas de negocio provistas por compañías como SAP y Oracle con soluciones como SAP BO Profitability and Cost Management (SAP SE, 2017) y Oracle Hyperion (Oracle Corporation, 2017); compañías intermedias como Ignitech (Ignite Technologies, 2017) y SAS (SAS Institute Inc, 2017), hasta compañías con un producto específico como MyABC (Amidia Digital Marketing, 2017) y rapidABC (RapidABC, 2017). También es común encontrar plantillas pagas o gratuitas en Microsoft Excel que permiten realizar implementaciones de la metodología de costos por parte de las empresas.

En el frente de Big Data, el gran protagonista en lo que a herramientas concierne es Hadoop, un entorno que permite el uso de diversos computadores (llamados nodos esclavos) como si fueran uno solo, los cuales, son controlados por otro computador (llamado nodo m) que centraliza las solicitudes y los resultados; de esta forma se potencia considerablemente la capacidad de procesar información voluminosa, a través de técnicas de Procesamiento en Paralelo Masivo (MPP por la sigla en inglés *Massively Parallel Processing*).

Para sacar provecho de esta estrategia de procesamiento, aparece un modelo de programación llamado MapReduce, el cual permite distribuir las tareas complejas, en tareas más sencillas que son procesadas por los nodos esclavos, los cuales entregan sus resultados al nodo maestro para que los compile y entregue una respuesta unificada.

En el caso particular de los modelos de costos basados en actividades, para organizaciones de gran tamaño e incluso para pymes que realicen procesos minuciosos, se pueden alcanzar modelos complejos cuyo procesamiento requiere de considerables recursos computacionales y genera grandes volúmenes de información. En estos escenarios resulta de gran utilidad Hadoop y el modelo programático de MapReduce para aprovechar el MPP y disminuir los tiempos de procesamiento para llevarlos a valores aceptables que no requieran de demasiados recursos computacionales.

Otros componentes fundamentales para la analítica de datos, son las herramientas de visualización, que permiten ver los resultados de los procesamientos que se realizan en los motores de bases de datos y en entornos como Hadoop. En este frente se destacan herramienta como QlikView (QlikTech Internationa, 2017) y Tableau (Tableau Software, 2017), que sirven para la construcción de reportes y tableros de control que usan como base la información procesado.

Una de las estrategias de visualización que está cobrando mayor popularidad, entre analistas de costos y también entre científicos de datos, es el uso de complementos en Excel para la ejecución de reportes, el manejo de tableros de control y la realización de presentaciones que sirven como base para la toma de decisiones. Dicha popularidad se debe al amplio número de funcionalidades matemáticas y estadísticas que brinda Excel para los científicos de datos, y las operaciones de manejo de datos y finanzas para los analistas de costos; características que se integran con las opciones de conexión a bases de datos y la extracción de información de otras fuentes de datos para ampliar las posibilidades de visualización.

Caso de Estudio

Para el montaje del caso de estudio de un modelo de costos ABC en una universidad, adicional a las metodologías de costos, se usa una metodología para proyectos de analítica de datos llamada CRISP-DM (IBM Corp., 2012), la cual plantea como sus dos primeras fases: la comprensión del negocio y la comprensión de los datos. Para la realización de estas dos fases, se debe tener en cuenta que las universidades necesitan diferentes estrategias de evaluación de la información, de acuerdo con el nivel administrativo de los interesados; para el caso los niveles de evaluación serán: rectorías, facultades y programas. Por otro lado, aparecen las áreas de interés en las cuales los interesados de cada nivel desean analizar la información; para el caso las áreas son: calidad, académica, financiera, infraestructura, investigación y proyección social.

Para ampliar el detalle en la fase de comprensión de los datos, se plantea la realización de una matriz en la que las filas representan los niveles de interés y las columnas las áreas de análisis, y en cada intersección se ubiquen los indicadores claves de desempeño (KPI: *Key Performance Indicator*) y/o los reportes necesarios para el análisis en cuestión. La tabla a continuación presenta un ejemplo de la matriz de áreas de análisis por nivel de interés, para el caso de una universidad.

Tabla 1. Matriz de áreas de análisis por nivel de interés para una universidad

Área ► ▼ Nivel	Calidad	Académica	Financiera	Infraestructura	Investigación	Proyección Social
Rectorías	Deserción total	Docente de planta	Costo x estudiante	Estudiantes x m2	Producción	Educación continua
	Movilidad total	Docente de cátedra	Costo x centro			Percepción en redes
	Rankings					Consultoría y extensión
Facultades	Deserción x programa	Estudiantes x programa	Costo x estudiante	Estudiantes x m2	Producción	Idiomas
	Movilidad x programa	Estudiantes x docente	Margen x programa	Espacios de estudio	Categoría de grupos	Percepción en redes
	Servicio exportado	Servicio exportado		Ocupación de aulas		Consultoría y extensión
Programas	Deserción programa	Estudiantes x docente	Costo x estudiante	Estudiantes x m2	Producción	Idiomas
	Movilidad programa	Estudiantes x curso				Percepción en redes
		Rendimiento académico				Consultoría y extensión

Por cada una de las filas (niveles de interés) se puede construir un tablero de control o *dashboard*, en el que se implementen los reportes e indicadores que permitan el análisis de la información. La figura 7 muestra un tablero de control

para el nivel de las facultades, cabe aclarar que podrían existir otros tableros para los niveles de rectorías y programas.

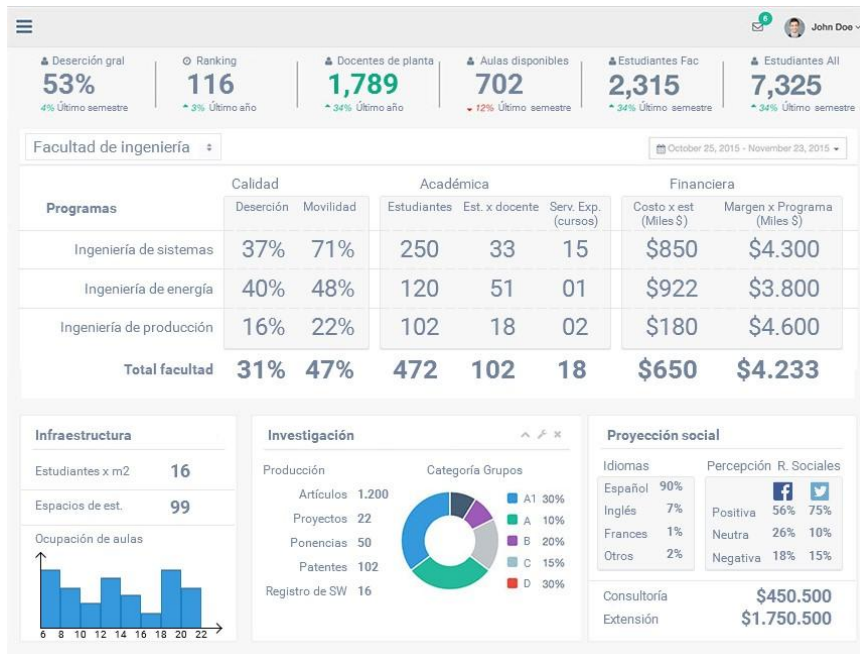


Figura 7. Ejemplo de tablero de control para el nivel de las facultades

La implementación del modelo de costos ABC, resulta de gran utilidad para la generación de los indicadores del área financiera, o algunos del área de proyección social, como consultoría y extensión; pues la mayoría de estos valores podrían resultar de cálculos derivados de la ejecución de los procesos de distribución de costos en la red de colectores. Por su parte, los inductores del modelo de costos, podrían apoyar algunos indicadores del área académica, infraestructura e investigación.

Por otro lado, indicadores del área de proyección social, en particular percepción en redes, se pueden apoyar en una disciplina de mucho auge en *Big data* como lo es la “analítica social” para entender el posicionamiento en redes de la institución, las facultades y los programas.

Otro frente de gran utilidad de la integración de *Big data* y ABC, lo constituye la posibilidad de manipulación de grandes volúmenes de datos, por ejemplo, el reporte de la figura 8 debe procesar decenas de millones de registros que han sido generados por el proceso de distribución de costos en diferentes periodos. Este reporte muestra una serie de tiempo sumalizando los costos de los servicios prestados por la universidad (objetos de costos) en los últimos semestres, y resulta de gran utilidad en los proyectos de analítica.

Totales_Titulo: Visor
 Fichero Ver Ayuda

30/11/2016 Informe de Valores Totales por Periodo. Pag. 1 de 1

U de Prueba (Administrativo)

Nivel del Colector : Productos o Servicios Misionales

Periodos	Desde	01-01-2014	01-07-2014	01-01-2015	01-07-2015	Total por
Concepto	Hasta	30-06-2014	31-12-2014	30-06-2015	31-12-2015	Concepto
1		11.186.394.956	11.317.903.152	12.763.573.977	16.030.651.186	51.298.523.271
Pregrado						
2		784.990.527	718.944.289	990.126.973	1.318.837.662	3.812.899.450
Posgrado						
3		5.146.338.720	6.454.350.273	7.035.716.022	9.945.680.550	28.582.085.565
Investigación						
4		24.139.696.517	18.119.616.103	19.938.126.129	26.327.431.912	88.524.870.661
Extensión						
6		5.764.999.992	47.504.536.861	4.268.681.630	16.011.969.876	73.550.188.359
Especial						
Total por Periodo		47.022.420.712	84.115.350.678	44.996.224.730	69.634.571.186	245.768.567.306

Figura 8. Ejemplo de reporte de costos por servicios sumariado por semestre

Los procesos de distribución de costos que generan la información para este reporte se pueden beneficiar de los algoritmos de redes neuronales disponibles en librerías de aprendizaje de máquina.

Conclusiones y trabajos futuros

La principal conclusión de este trabajo muestra el beneficio que se puede obtener de la fusión de disciplinas como el costeo basado en actividades y áreas relativas a *Big Data* como analítica y aprendizaje de máquina. Esto amplía en gran medida las posibilidades en cuanto a la implementación de los modelos de costos ABC, pues permiten la integración de nuevas técnicas y adicionalmente, posibilita el aumento en el volumen y complejidad de los modelos de costos a ser implementados sin perjudicar el rendimiento.

El *Big Data*, puede ser utilizado no solo para generar información mediante Dash Board, sino que eventualmente, puede convertirse en una herramienta que contribuya al análisis y depuración de datos necesarios para la implementación de un modelo de costos, mediante el análisis de sistemas de información de nómina, activos fijos y administración académica entre otros.

Como trabajo futuro se puede formular una metodología que permita la integración del montaje de un modelo de costos ABC con metodologías como CRISP-DM, de tal manera que las actividades planteadas en este trabajo se puedan realizar de forma genérica y llevar esta propuesta a su implementación en una universidad que necesite refinar sus esquemas de gestión y analítica.

REFERENCIAS BIBLIOGRÁFICAS

- Amidia Digital Marketing. (15 de 04 de 2017). *Cost and Profit Management*. Obtenido de <https://www.myabcm.com/myabcm-enterprise/>
- Caicedo, E., & Lopez, J. (2009). *Una aproximación práctica a las Redes Neuronales Artificiales*. Cali, Colombia: Universidad del Valle.
- CGN, C. (2008). *Marco de Referencia para la implantación del sistema de costos en entidades del sector público*. Bogota: CGN.
- Cuervo, J., Osorio, J., & Duque, M. (2013). *Costeo basado en actividades ABC: gestión basada en actividades ABM (Segunda ed.)*. Medellín, Colombia: Ecoe Ediciones.
- Davenport, T. H. (12 de 2013). Analytics 3.0: In the new era, big data will power consumer products and services. *Harvard Business Review*. Recuperado el 20 de 07 de 2016, de Analytics 3.0: <https://hbr.org/2013/12/analytics-30>
- Davenport, T. H., & Patil, D. (2012). Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*. Recuperado el 20 de 07 de 2016, de <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/>
- Drew Conway Data Consulting. (26 de 3 de 2013). *The Data Science Venn Diagram*. Recuperado el 15 de 07 de 2016, de <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>
- Duque, M., Gómez, L. F., & Cuervo, J. (2005). Los direccionadores: ¿solución o problema del ABC? *Contaduría Universidad de Antioquia*, 39-58.
- IBM Corp. (2012). *Manual CRISP-DM de IBM SPSS Modeler*. Obtenido de IBM Knowledge Center: <ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/15.0/es/CRISP-DM.pdf>
- Ignite Technologies. (15 de 04 de 2017). *Acorn Performance Analyzer*. Obtenido de <http://www.ignitetech.com/solutions/finance/performance-analyzer/>
- Jara, G. N., Castañeda, J., & Gómez, L. F. (2003). Sistema de Costeo Basado en Actividades como Herramienta del Presupuesto Inteligente para el Distrito Capital, Santa Fé de Bogotá - Colombia-Secretaría de Hacienda Distrital Bogotá D.C. *Contaduría Universidad de Antioquia*, 43, 179-204.
- LeHong, H., & Fenn, J. (18 de 09 de 2012). *Key Trends to Watch in Gartner 2012 Emerging Technologies Hype Cycle*. Gartner, Inc. Recuperado el 13 de 07 de 2016, de Forbes / Tech : <http://www.forbes.com/sites/gartnergroup/2012/09/18/key-trends-to-watch-in-gartner-2012-emerging-technologies-hype-cycle-2/>
- Oracle Corporation. (15 de 04 de 2017). *Oracle Hyperion Profitability and Cost Management*. Obtenido de <https://www.oracle.com/applications/performance-management/products/business-planning/hyperion-profitability-cost-management/index.html>
- QlikTech Internationa. (15 de 04 de 2017). *QlikView® Guided Analytics*. Obtenido de <http://www.qlik.com/us/products/qlikview>

- RapidABC. (15 de 04 de 2017). *Rapid ABC*. Obtenido de <http://www.rapidabc.com/>
- SAP SE. (15 de 04 de 2017). *SAP BusinessObjects Profitability and Cost Management*. Obtenido de <http://go.sap.com/product/financial-mgmt/profitability-cost-mgmt.product-capabilities.html>
- SAS Institute Inc. (15 de 04 de 2017). *SAS Cost and Profitability Management*. Obtenido de http://www.sas.com/en_us/software/performance-management/cost-profitability-management.html
- Shmueli, G., Patel, N., & Bruce, P. (2016). *Data Mining for Business Analytics*. New Jersey: John Wiley & Sons.
- SSP, S. (1997). Resolución 1417, “Por la cual se expide y adopta el Sistema Unificado de Costos y Gastos para Entes Prestadores de Servicios Públicos Domi-ciliarios de Acueducto, Alcantarillado, Aseo, Energía y Gas Combusti-ble.”.
- Tableau Software. (15 de 04 de 2017). *Tableau Desktop*. Obtenido de <https://www.tableau.com/products/desktop>