

Segmentando el Big Data con Facebook

CPN Pruvost Andrés Guillermo

Universidad Nacional del Litoral
Facultad de Ciencias Económicas
cpnandres@gmail.com

Abstract. El Big Data es mucho más que grandes datos, constituye la posibilidad de analizar datos de manera masiva y obtener una visión completa de situaciones de negocio y sociales antes medidas solamente observando muestras del objeto de estudio. Los usuarios de Facebook generan diariamente cantidades inmensas de datos que le sirven a la red social para segmentar los avisos de sus anunciantes desde diversas visiones desde los datos básicos de los usuarios hasta su comportamiento de compra, intereses o similitud con su segmento que llevan al marketing de segmentación a un nuevo nivel de detalle.

Keywords: Big Data, segmentación, marketing, Facebook

1 Introducción

El conocimiento del perfil del cliente actual o potencial ha sido siempre un elemento esencial de las actividades de marketing tanto para el diseño del producto o servicio como también su promoción y exposición en el punto de venta.

El comportamiento del cliente y sus gustos pueden además definir tendencias las cuales permiten anticipar acciones a futuro para mantener los niveles de satisfacción y fidelidad. Un ejemplo sencillo de aplicación lo representa la herramienta Google Suggest del motor de búsqueda el cual propone términos de búsqueda al usuario a medida que este coloca palabras en el cuadro de búsqueda, basadas en las coincidencias estadísticas de todas las búsquedas anteriores.

Internet mediante, el registro y trazabilidad de las acciones online de los clientes y prospectos creció de manera exponencial al punto que son necesarios nuevos términos para definir estas medidas, desde el humilde megabyte la evolución fue constante, gigabyte, terabyte, exabyte, zettabyte llegando ahora a los yottabyte, al menos hasta el momento. Ya en 2013 Google procesaba más de 24 petabytes diariamente (Schönberger, 2013).

La cantidad y variedad de los datos disponibles permiten generar una imagen más precisa de los gustos y actividades de los clientes y poder así llegar así con el mensaje adecuado, en el momento preciso, para el cliente específico.

De la mano del Social Media y las aplicaciones del tipo Web 2,0 cualquier usuario con la sola habilidad de navegar por Internet puede ahora publicar y gestionar gran cantidad de datos en diversos formatos, subir videos a YouTube, publicar imágenes en Flickr o Instagram, audios en SoundCloud y un largo etcétera.

El incremento constante de usuarios conectados a Internet mediante Banda Ancha sumados al crecimiento de dispositivos móviles indica que la cantidad de datos generados ira en aumento lo que implica desafíos para las aplicaciones orientadas a su gestión y análisis.

Desarrollar la capacidad de análisis de grandes volúmenes de datos presenta grandes beneficios en distintos ámbitos empresariales, sociales como científicos, básicamente podremos evolucionar del análisis de muestreo a un análisis completo de todos los casos, generando una nueva visión sobre el objeto de estudio reduciendo la incertidumbre.

2 Big Data

El concepto de Big Data o grandes datos en castellano engloba una serie de elementos a los que se suele definir como el conjunto de datos cuyo crecimiento supera la posibilidad de administrarse y analizarse mediante las herramientas tradicionales de procesamiento de datos (Ohlhorst, 2013), compuesto por datos no estructurados y multi estructurados a menudo provenientes de interacciones entre personas y máquinas, aplicaciones Web, redes sociales y sensores (Arthur, 2013).

2.1 Las tres V del Big Data

El Big Data no refleja simplemente una acumulación de datos generados por empresas, sistemas y usuarios, está representado en forma simultánea por el volumen de datos generados, la variedad de tipos de datos y la velocidad con la que cambian los datos.

- Volumen, refiere a la cantidad de datos generados diariamente y su acumulación en distintos dispositivos, desde transacciones comerciales a datos registrados por los satélites orbitando la tierra, solo en Youtube por ejemplo se publican 300 nuevas horas de video cada minuto¹.
- Velocidad con la que se generan y fluyen los datos, duplicando cada año su tamaño. Ya en 2011 de acuerdo a los resultados del McKinsey Global Report se consideraba que el 90% de los datos en el mundo habían sido creados en los últimos dos años y será 44 veces más grande para el año 2020, respecto del 2011.
- Variedad, implica una fuente de datos diversa y en constante crecimiento, desde mensajes de texto, emails, operaciones de comercio electrónico, y un largo etcétera.

1 <https://www.youtube.com/yt/press/>

Algunos autores incluyen también Veracidad y Valor como las V restantes de este objeto de estudio refiriendo a la necesidad de acercarse a la mayor exactitud posible de los datos para generar el valor como recurso.

2.2 Origen y clasificación de los datos

Los datos pueden provenir de diversas fuentes y formatos, datos generados por usuarios como en el caso de los videos publicados en Youtube, la creación de contenidos para Blogs o la inmensa cantidad de datos publicados diariamente en redes sociales como Facebook o Instagram.

A nivel transaccional las operaciones de comercio electrónico, el comportamiento registrado de los usuarios en sus recorridos por los sitios Web mediante herramientas de tracking online como Google Analytics, es información que hoy puede ser relevada fácilmente para establecer patrones de comportamiento.

Datos generados por sistemas administradores de clientes como CRM (Customer Relationship Management), o administradores de recursos como ERP (Enterprise Resource Planning), transacciones financieras mediante aplicaciones online de Banking y transferencias electrónicas. Datos biométricos vinculados a seguridad como registros de entrada y salida de operarios, lectores de huellas digitales, registros de movimientos por GPS, son algunos de los datos que máquinas y sistemas día a día generan.

De acuerdo a su formato los datos pueden ser clasificados como:

- Estructurados. Son datos de longitud y formato definido, por lo general almacenados en tablas provenientes de sistemas como CRM, ERP o incluso planillas de cálculo.
- No estructurados. Lo opuesto a los estructurados, carecen de un formato específico y constituyen una de las categorías de mayor crecimiento, no se encuentran organizados y no son fáciles de interpretar por bases de datos tradicionales, mensajes vía Twitter u otras redes sociales, catálogos fotográficos, opiniones y comentarios en redes sociales son ejemplos de esta categoría. Diversas técnicas son utilizadas para darle estructura a este tipo de datos como la categorización semántica, taxonomías, reconocimiento de lenguaje natural, entre otras.
- Semi-estructurados. Estos datos presentan características compartidas, son datos no estructurados en bases de datos con campos, pero con marcadores para separar los distintos elementos, el lenguaje de marcador de hipertexto es un ejemplo.

3 Marketing de segmentation

El Big Data representa una oportunidad por el volumen de datos disponibles, pero un desafío también por los requerimientos para generar valor agregado de los mismos. En los tiempos de Internet el desafío no consiste en obtener los datos, ya que los

mismos están disponibles en enormes cantidades, el verdadero valor agregado proviene del análisis de los mismos para la toma de decisiones.

“Estoy seguro que he malgastado la mitad del dinero invertido en publicidad, el problema es que no sé qué mitad”, famosa frase del mundo de la publicidad atribuida a Jhon Wanamaker que muestra algunos de las fallas de la publicidad tradicional para medir el alcance real de la inversión publicitaria.

A partir del 2006 con el auge de las llamadas aplicaciones Web 2.0 la participación del usuario se hace más activa en todo sentido, en la generación de contenidos, mayor influencia en las decisiones de compra mediante sus comentarios y una comunicación más fluida y bidireccional entre el usuario y la empresa.

Estos cambios obligan a las empresas a dirigir su comunicación publicitaria del monólogo unidireccional hacia la conversación.

Toda la participación y conversación con el usuario es realizada casi en su totalidad de manera online, situación que permite el registro de todo el proceso, desde su visita en el sitio de Ecommerce en donde es posible analizar las páginas visitadas, el tiempo transcurrido dentro del sitio e incluso las palabras clave que utilizaron para llegar al sitio Web, luego de decidida la compra es viable registrar el proceso de compra, medios de pago utilizados, historial de compras anteriores, todos datos que le permiten a la empresa generar promociones y descuentos a medida para cada usuario registrado.

El gigante Google registra Terabytes de información segundo a segundo sobre el comportamiento de los usuarios en el motor de búsqueda, las palabras clave más utilizadas, los sitios Web visitados y videos visualizados entre otros recursos, todos estos datos son luego utilizados en su plataforma de publicidad Adwords la cual permite a los anunciantes segmentar el aviso publicitario de acuerdo a distintos criterios como la localidad donde se mostrará el aviso, el dispositivo PC, Tablet o teléfono celular donde se desea mostrar el anuncio pautado, seleccionar o excluir sitios Web donde mostrar el aviso, entre otros ítems.

No obstante el eje central de la plataforma lo constituyen el uso de las palabras clave como dato para obtener presencia en búsquedas relacionadas, sitios Web y canales de video con temas vinculados a las palabras clave seleccionadas, de esta manera la plataforma se acerca al ideal de mostrar el mensaje adecuado, en el momento oportuno y al cliente correcto.

El Big data puede transformarse en una ventaja competitiva para el mercado minorista, Amazon.com que comenzara vendiendo libros es hoy uno de los principales jugadores en el mercado del E-commerce, los datos relevados en la interacción con sus clientes actuales y potenciales le permiten predecir tendencias y prepararse para incrementos de demanda futuras, brindar promociones cruzadas a clientes existentes así como estrategias dinámicas de precios y ofertas para acciones de fidelización.

Amazon captura todo dato posible de cada usuario para compararlo luego con el resto de los datos y así poder generar recomendaciones personalizadas de acuerdo al historial de compras y productos visualizados del catálogo.

Obviamente Facebook, la red social dominante en el mercado, no queda afuera de la lista en lo relativo a la explotación del Big Data, cuenta con una comunidad de 936

millones de usuarios activos promedio para Marzo de 2015², 800 millones de usuarios de la red de mensajes Whatsapp, 300 millones de usuarios de su red de imágenes Instagram y 600 millones de usuarios de su plataforma de mensajería Messenger, comunidad que genera 45 billones de mensajes diariamente y 4 billones de visualizaciones de videos, base suficiente para hacer de la analítica en tiempo real una ventaja competitiva importante y convertir a la red social en un canal publicitario a escala global.

4 Facebook Ads

Constituye la plataforma publicitaria de la red social Facebook y permite la promoción de páginas de empresa, publicaciones, derivación de tráfico a sitios Web, entre otros objetivos.

De la misma manera que Google Adwords, la plataforma permite lograr una mayor exposición en este caso dentro de la red social, pudiendo de esta manera acceder a otros usuarios segmentados de acuerdo a determinados parámetros que el anunciante puede configurar.

El valor diferencial de Facebook Ads consiste en las posibilidades de filtrado y segmentación de los anuncios, segmentación que utiliza como materia prima los datos personales del perfil de los usuarios de la red social.

Facebook ofrece distintas opciones publicitarias o también denominados Objetivos tales como:

1. Clics en el sitio web, para derivar visitas al sitio Web
2. Conversiones en el sitio web, promover una conversión o acción específica como una venta, descarga de un formulario, para lo cual se requiere un pixel de conversión instalado
3. Interacción con una publicación, promover los posteos de la Fanpage
4. Me gusta de la página para la generación de Fans
5. Instalaciones de tu aplicación, para promover la descarga e instalación de aplicaciones móviles o de escritorio
6. Interacción con tu aplicación, para aumentar el uso de aplicaciones móviles de la empresa
7. Difusión local, permite crear anuncios dirigidos a usuarios próximos al negocio, opción disponible solamente en algunos países
8. Respuestas a eventos, para promover la participación
9. Solicitudes de ofertas, para la promoción de ofertas o descuentos puntuales
10. Reproducciones de video, promover visualizaciones de los videos de la empresa

² <http://newsroom.fb.com/company-info/>

Para cada uno de los objetivos publicitarios habrá un tipo de anuncio de mayor conveniencia que el anunciante puede contratar.

Existen principalmente 3 formas de desarrollar una campaña de publicidad en Facebook:

- Herramienta básica de creación de avisos. La forma más sencilla de promocionar un objetivo, procedimiento rápido y útil para anunciantes que se inician con presupuestos reducidos. Se accede mediante los botones azules de Promocionar página, Promocionar publicación o Promocionar sitio Web que figuran al ingresar a la Fanpage como administrador.
- Programación a través de la API. Debido a su coste y complejidad esta opción es utilizada por grandes agencias y puede estar fuera del alcance de la mayoría.
- Power Editor. Una herramienta completa y profesional para la generación de avisos con opciones de configuración avanzadas.

4.1 Power Editor

Consiste en un agregado o Plugin que puede ser accedido por el momento solo con el navegador Chrome ingresando a la URL www.facebook.com/powereditor, una vez descargado permite editar avisos en la PC con amplias opciones de configuración, los cuales luego del proceso deben ser subidos a Facebook para quedar integrados a la cuenta publicitaria del anunciante.

La mayor complejidad que demanda la creación de avisos en Power Editor es compensada con las opciones de segmentación y gestión masiva de anuncios disponible, a continuación algunas características distintivas:

- Acceder a novedades, Facebook pone a disposición en Power Editor los últimos avances y novedades publicitarias las que luego de cierto tiempo están disponibles en el administrador básico de anuncios, lo que implica que el anunciante puede utilizarlas meses antes que el resto de los usuarios.
- Acceder a audiencias similares (Lookalike Audiences), permite llegar con el anuncio a usuarios con intereses similares a la actual comunidad de la Fanpage, público similar a usuarios de una determinada Fanpage, similar a clientes que convirtieron en el sitio Web como una compra u otro tipo de acción específica entre otras coincidencias lo cual permite llegar con los avisos a usuarios con el mismo potencial que la actual lista de clientes.
- Publicaciones no visibles (Dark Posts), permite saltar el proceso tradicional de publicar un post para luego crear un anuncio para promocionarlo, con esta opción se puede crear un post que no será publicado en el Timeline pero se puede promocionar como anuncio, esta opción permite crear avisos dirigido a audiencias específicas con el fin de realizar testeos del tipo A/B y evaluar el rendimiento de cada aviso de acuerdo al contenido publicado.
- Segmentación por dispositivo, los avisos podrán ser configurados para visualizarse en PC o teléfonos celulares, en este caso de acuerdo al tipo de

sistema operativo, teléfonos inteligentes o básicos e incluso para los casos de conexión a Facebook utilizando Wifi.

- Edición masiva, Power Editor permite editar avisos en bloque o incluso importar avisos creados mediante planillas de cálculo.
- Seguimiento de conversiones, mientras que el administrador básico de avisos permite monitorear las conversiones de un objetivo, Power Editor permite medir la conversión de varios objetivos de manera simultánea mediante el uso de los pixeles de seguimiento instalados en el sitio Web, carrito de compras, formulario de email marketing, o el destino en el que se desea medir el resultado.

Power Editor es una herramienta con muchas opciones y detalles a configurar, la mayoría de las cuales son agregadas progresivamente a la herramienta básica de avisos, su uso no siempre será la opción por defecto a utilizar, situación que será necesario analizar de acuerdo al volumen de avisos y necesidades de segmentación.

4.2 Segmentación social

Además de la segmentación geográfica y etaria provista también por Google Adwords, Facebook permite segmentar el aviso de acuerdo al perfil social del mercado objetivo, como la información sobre su relación sentimental, estudios cursados, actividad laboral, afinidad étnica, entre otros datos que el mismo usuario carga en su perfil.

Permite segmentar el anuncio hacia personas vinculadas o no con la Fanpage o publicación a promocionar, podremos incluso excluir de la visualización de los avisos a personas que estén asociadas con determinada página de Facebook, opción que permite llegar con el mensaje publicitario a usuarios con un tipo de interés en particular.

Segmentar por comportamiento, permite llegar con el aviso a usuarios con determinados comportamientos, como usuarios de juegos online, quienes utilicen medios de pago en Facebook o quienes se conectan a la red social mediante Tablets, todas audiencias específicas disponibles para anunciantes del sector.

Selección de ubicación. Podremos configurar el aviso para que sea visible en la sección noticias de Facebook, sección noticias de celulares o en la columna derecha de la página. El acceso a la red social por medio de teléfonos celulares crece de manera sostenida alcanzando los 798 millones de usuarios promedio en Marzo de 2015³ lo que representa un crecimiento del 31% respecto al año anterior, esto convierte a la plataforma online en un canal publicitario importante por su movilidad y frecuencia de uso.

Los anuncios pueden ser programados para ser mostrados en días y horarios específicos de acuerdo al uso horario del público objetivo, estos avisos serán mostrados en el momento configurado por el anunciante pudiendo así tener alcance en

3 <http://investor.fb.com/releasedetail.cfm?ReleaseID=908022>

el momento del día y la semana de mayores accesos a la red social de acuerdo a la localización del usuario.

Diversas opciones de tasación de los avisos, CPC (Costo por Clic, el anunciante paga cuando se hace clic en su aviso), CPM (Costo por Mil, el anunciante paga por exposición) y OCPM (Optimizado Costo por Mil), similar al costo por mil pero optimizado para una acción específica lo cual permite ajustar el presupuesto de acuerdo a la estrategia publicitaria o el presupuesto asignado por la empresa a publicidad online.

5 Impacto social del Big Data

Vistas las posibilidades prácticas de segmentar un mensaje publicitario de acuerdo a los datos registrados, cabe analizar su impacto en otras áreas no comerciales.

La transición del análisis de muestreo al análisis del total de datos es un cambio de gran impacto en la detección de tendencias sociales, los cambios de patrones de comportamiento son una valiosa información para la toma de decisiones en caso de pandemias o catástrofes naturales como inundaciones o terremotos tal como lo plantea Nuria Oliver en el proyecto “*Big Data for Social Good*”, y el uso de los datos generados por la telefonía celular como fuente de datos del comportamiento humano a gran escala.

En su proyecto Oliver logró demostrar como el registro de las antenas de telefonía celular permitió conocer los movimientos y consultas en Hospitales y centros de atención médica en la crisis de la gripe A o H1N1, datos que permitieron reducir los índices de contagio de la misma.

El proyecto “*The human face of Big Data*⁴” intenta ser un proyecto colectivo en el que se pone en evidencia el potencial del Big Data para el análisis, triangulación y visualización de datos para la toma de decisiones.

En asociación con la consultora EMC y la empresa Tableau Software se analizaron un total de 1000 millones de mensajes en Twitter para elaborar visualizaciones de datos en diversos campos de interés, indicadores a nivel mundial sobre la percepción de seguridad de los ciudadanos, hábitos y seguridad alimentaria, pasatiempos preferidos o temas más sensibles como las preocupaciones sobre la economía o los avances de la cura de la Malaria.

Incluso el cuidado del medio ambiente, con el impacto que representa en la calidad de vida de los ciudadanos, puede obtener beneficios del Big Data, con el objeto de incrementar el uso de la energía eólica se procesan diariamente 8 millones de gigabytes de información sobre simulación climática, sensores de viento en todo el mundo establecen las regiones más ventajosas para el desarrollo de esta fuente de energía no contaminante.

Junto al crecimiento de la población crecen las necesidades del transporte de personas y bienes, en Londres por ejemplo la agencia gubernamental *Transport For*

4 <http://thehumanfaceofbigdata.com>

London⁵ (TFL) utiliza análisis de Big Data para generar soluciones para los usuarios y prevenir demoras y embotellamientos de autos particulares, autobuses, taxis y bicicletas.

Utilizando datos de provenientes de las tarjetas prepagas de transporte, sensores instalados en vehículos de transporte público, información de las señales de tránsito, mensajes en social media y encuestas, desarrollan vías alternativas para distintos destinos locales, conexiones entre diversos medios de transporte o anticipar posibles congestionamientos debido a obras de reparación, información que es compartida en tiempo real mediante 23 cuentas distintas de Twitter de acuerdo a los recorridos que el usuario del transporte usa regularmente.

Como vemos el análisis del Big Data constituye potencialmente una opción para el diseño de políticas públicas con la posibilidad técnica de tomar en cuenta la totalidad de los datos generados sobre el tema en cuestión y no solamente una muestra, el proyecto del billón de precios *The Billion Price Project*⁶ del MIT dirigido por Rigobon Roberto y Cavallo Alberto propone desarrollar un índice de precios pero no de una canasta básica, sino de la evaluación de millones de precios registrados diariamente utilizando Internet y los sitios de Ecommerce como fuente de los datos, lo cual constituye una innovación en el estudio de la macroeconomía utilizando datos masivos a nivel mundial.

Finalmente el uso del Big Data en el cuidado de la salud es prácticamente un paso inevitable por los beneficios que genera, la digitalización de los estudios de diagnóstico desde exámenes de sangre hasta radiografías y tomografías permite su almacenamiento y procesamiento online, el análisis y cruce de información médica a gran escala favorece el desarrollo del conocimiento médico sobre distintas patologías, evitar los reingresos a centros de salud y analizar tendencias de enfermedades para el desarrollo de medicamentos adecuados por parte de la industria farmacéutica.

El análisis de enfermedades a gran escala permitirá crear perfiles de riesgo y anticiparse así de acuerdo a cada población que comparta similares características y estilos de vida haciendo más eficientes los presupuestos en salud.

Lamentablemente no toda la información procesada en el Big Data es totalmente anónima e inocua, los registros contenidos en historias clínicas digitales por ejemplo pueden potencialmente ser utilizados con criterio discriminatorio en casos de VIH u otras patologías de tipo hereditario, situaciones donde leyes de protección de datos personales como la N° 25326 a nivel nacional y su modificatoria N° 26343, como la ley N° 26388 referente a delitos informáticos en el Código Penal son instrumentos disponibles para reglamentar la actividad en el manejo de datos y evitar abusos relacionados.

5 <https://www.tfl.gov.uk>

6 <http://bpp.mit.edu/>

6 Conclusiones

El Big data constituye una oportunidad para el análisis de datos a gran escala en tiempo real, estos datos incluyen a los estructurados en bases de datos como los no estructurados de los cuales forman parte, entre otros, los generados por los usuarios en las redes sociales.

La inmensa cantidad de datos generados diariamente por los diversos dispositivos existentes requiere nuevos enfoques para su correcta explotación y capitalización para la toma de decisiones.

El comportamiento en Internet del usuario puede ser hoy registrado y almacenado, esta información junto con los datos personales provistos al crear su cuenta facilitan la tarea de la segmentación a escala global, de esta manera la publicidad puede acercarse al ideal de llegar con su mensaje en el momento adecuado, al público adecuado, haciendo así más eficiente el presupuesto publicitario.

Google y Amazon son dos grandes jugadores en el mercado del Big Data evaluando tendencias y realizando predicciones de acciones futuras a partir de los hábitos de sus usuarios, por su parte Facebook suma a estas acciones los datos generados por la inmensa comunidad de la red social y demás empresas adquiridas como Instagram y Whatsapp convirtiéndola hasta el momento en la plataforma publicitaria con mayores posibilidades de segmentación.

Los beneficios del Big Data abarcan diversas áreas como transporte, seguridad, salud y medio ambiente, no son los datos los que generan el valor sino el análisis de los mismos para la toma de decisiones.

Estos beneficios no están exentos de desafíos, la heterogeneidad de los datos de origen provenientes de distintas fuentes y formatos es el principal y genera la necesidad de nuevos enfoques en el tratamiento de los datos, la defensa de la privacidad de los usuarios generadores de los datos y el uso de información personal sensible requerirá la aplicación de la legislación adecuada para evitar el uso indebido de los mismos.

7 Bibliografía

Lisa, A. Big data marketing, engage your customers more effectively and drive value, Jhon Wiley and Sons, 2013, Canada

Ohlhorst F. Big data analytics, turning big data into Big Money, Jhon Wiley and Sons, 2013, Canada

Schönberger, V. Cukier, K. Big data: la revolución de los datos masivos, Turner, 2013