



UNIVERSIDAD
**SAN IGNACIO
DE LOYOLA**

FACULTAD DE CIENCIAS EMPRESARIALES

Carrera de International Business

**REPUTACIÓN CORPORATIVA EN UN MERCADO
EXTRANJERO Y SU RELACIÓN CON VARIABLES
DEL SOCIAL MEDIA - CASO DEL BCP BOLIVIA**

**Tesis para optar por el Título Profesional de Licenciado en
International Business**

CLAUDIA LORENA LAO SAAVEDRA

Asesor:

Mag. Jesús Robles Lara

Lima- Perú

2020

Dedicatoria

A mis padres, Larry y Nelmith, por su amor, apoyo incondicional y por inculcarme el hábito del estudio. A mis hermanos, por su permanente afecto y cariño. Los amo.

RESUMEN

El presente documento buscó evidenciar si existe relación entre la reputación del BCP-Bolivia y ciertos elementos propios del Social Media. Para ello se utilizó el índice de Merco con data anual del 2013 al 2018. Se correlacionó dicho índice con seis potenciales variables explicativas: la cantidad de comentarios en YouTube por año que aparecían en los vídeos subidos por la propia empresa, la cantidad de veces en la que el banco fue objeto de noticias en los portales de periódicos bolivianos, los índices de búsqueda relativo en Google y YouTube acerca del negocio dentro de territorio boliviano, y el puntaje sentimental presentes en los comentarios de YouTube y en el contenido de los diarios. Estas últimas, precisamente, son las que tienen relación significativa y directa con la reputación ($r = 0.945$ para el caso de YouTube y 0.836 en el caso de las noticias). Se ha empleado el análisis sentimental, la cual es una técnica propia del Data Mining, para procesar los textos. Finalmente, fue empleada una red neuronal que buscó predecir la reputación usando todas las variables explicativas. La precisión de esta fue de 43.2%. La regresión lineal, por otro lado, solo incluye al puntaje sentimental de los comentarios en YouTube como variable significativa. El R-cuadrado del modelo de regresión simple fue de 87.5%, pero existe el sesgo de la autocorrelación en los residuos, por lo que no se cumplen con una de las condiciones básicas de validez.

Palabras Clave: Social Media, Minería de Datos, Reputación Corporativa, BCP Bolivia, Redes Neuronales, Análisis Sentimental

ABSTRACT

This document sought to reveal a relationship between the reputation of the BCP-Bolivia and some elements of the Social Media itself. This was used to use the annual index of the Merco index from 2013 to 2018. That index was correlated with six potential explanatory variables: the amount of comments on YouTube per year that appeared in the videos uploaded by the company itself, the number of times the bank was the subject of news in the portals of Bolivian newspapers, the searchable indexes on Google and YouTube about the business within Bolivian territory, and the sentimental score present in the comments on YouTube and the content of the journals. The latter, precisely, are those that have a significant and direct relationship with the reputation ($r = 0.945$ for the case of YouTube and 0.836 in the case of the news). The sentimental, technical analysis of the Data Mining was used to process the texts. Finally, a neural network that sought to predict reputation using all the explanatory variables was employed. The accuracy of this was 43.2%. Linear regression, on the other hand, only includes the sentimental score of the comments on YouTube as a significant variable. The R-Square of the simple regression model was 87.5%, but there is the bias of autocorrelation in residues, so they are not met with one of the basic conditions of validity.

Keywords: Social Media, Data Mining, Corporate Reputation, BCP Bolivia, Neuronal Networks, Sentimental Analysis

ÍNDICE

Introducción..... 1

Capítulo I – Problema de Investigación	2
1.1. Problema de Investigación	2
1.1.1. Planteamiento del Problema	2
1.1.2. Formulación del Problema.....	3
1.1.3. Justificación	3
1.2. Marco Referencial.....	5
1.2.1. Antecedentes.....	5
1.2.2. Marco Teórico	7
1.3. Objetivos e Hipótesis	27
1.3.1. Objetivos.....	27
1.3.2. Hipótesis	28
Capitulo II – Método	29
2.1. Tipo y Diseño de Investigación	30
2.1.1. Tipo de investigación	30
2.1.2. Diseño de investigación	30
2.2. Variables de Estudio	31
2.3. Universo.....	32
2.4. Instrumentos.....	32
2.4.1. NVIVO	32
2.4.2. TextCleanr.....	32
2.4.3. <i>Language Tool</i>	32
2.4.4. Grammarly	33
2.4.5. Google Cloud Natural Language.....	33
2.4.6. Excel.....	33
2.4.7. SPSS.....	34
2.5. Procedimiento para recolección de datos.....	34
2.6. Plan de análisis.....	37

2.6.1. Correlación entre variables.....	37
2.6.2. Red neuronal	37
Capitulo III- Resultados	41
3.1. Presentación de Resultados.....	41
3.2. Discusión	48
3.3. Conclusiones.....	49
3.4. Recomendaciones	50
3.5. Limitaciones.....	51
Bibliografía.....	53

Lista de Tablas

Tabla 1: Ranking de Cartera – Principales Bancos de Bolivia (en miles de bolivianos)	25
Tabla 2: Ranking de Depósitos del público y empresas con participación estatal - Principales Bancos de Bolivia (en miles de bolivianos)	26
Tabla 3: Variables de Estudio.....	31
Tabla 4: Índices de Búsqueda en Google y YouTube acerca del BCP -Bolivia en territorio boliviano.....	35
Tabla 5: Variables consolidadas	36
Tabla 6: Pruebas de normalidad	41
Tabla 7: Matriz de Correlaciones de Variables	42
Tabla 8: Resumen del modelo	44
Tabla 9: Parámetros de la red neuronal y Precisión del modelo	44
Tabla 10: Modelo de Regresión.....	46
Tabla 11: Variables excluidas del modelo de regresión	46
Tabla 12: Resumen del modelo de Regresión Lineal y prueba de Durbin Watson.....	47
Tabla 13: Pruebas de Normalidad para el modelo de regresión	47
Tabla 14: Comparación entre modelos Red Neuronal – Regresión Lineal	48

Lista de Figuras

Figura 1: Interés mensual hacia Huawei (setiembre del 2014 hasta setiembre del 2019) en el mundo de acuerdo al volumen de búsquedas.....	18
Figura 2: Interés mensual hacia United Airlines (setiembre del 2014 hasta setiembre del 2019) en el mundo de acuerdo al volumen de búsquedas	19
Figura 3: Distribución porcentual de las acciones del BCP Bolivia	21
Figura 4: Interés hacia el BCP en territorio boliviano 2014-2018 de acuerdo a Google	22
Figura 5: Bancos relacionados al BCP Bolivia de acuerdo a Google en territorio boliviano	23
Figura 6: Activos y Patrimonio del BCP Bolivia 2014-2018 (en millones de dólares)	23
Figura 7: ROE y ROA obtenidos por el BCP Bolivia durante el periodo 2014-2018...	24
Figura 8: Proceso de extracción de datos de Google y YouTube.....	34
Figura 9: Proceso de Text Mining (o Data Mining) en el caso de YouTube	35
Figura 10: Diagrama de un Perceptrón.....	37
Figura 11: Esquema de una red neuronal con una sola capa oculta	38
Figura 12: Arquitectura de la red neuronal.....	43
Figura 13: Valor predicho vs. Observado en la red neuronal	45
Figura 14: Importancia de las variables predictoras dentro del modelo.....	45

Lista de Anexos

Anexo 1: Calendario y Recursos Disponibles	61
Anexo 2: Interface de Uso de Google Trends	62
Anexo 3: Delimitación de resultados de Google Trends a espacio boliviano y para YouTube	63
Anexo 4: Interface de uso de Google Cloud Natural Language.....	64
Anexo 5: Resultados de Google Cloud Natural Language.....	65
Anexo 6: Uso de NVIVO	66
Anexo 7: Proceso de captura finalizado	66
Anexo 8: Selección y conversión de data formato en Excel	68
Anexo 9: Interface de software de depuración	70
Anexo 10: Interface de softwares traductores	71
Anexo 11: Uso de software correctores.....	72
Anexo 12: Matriz de consistencia	74

Introducción

El BCP es una de las empresas más representativas del Perú. Dicha entidad encontró en Bolivia una oportunidad para expandir sus actividades. Hoy en día la entidad es uno de los principales referentes del sector de servicios financieros en el país altiplánico.

Los elementos del Social Media han ido creciendo y diversificándose a medida que la proporción de personas que accede a internet se incrementa. Actualmente, existen herramientas que permiten medir elementos asociados a dicho entorno digital. Es por ello que métodos como el Data Mining la Minería de Datos han ido ganando popularidad entre quienes desean estudiar variables propias de empresas empleando indicadores que surgen de la interacción digital.

Uno de los indicadores que más debe preocupar a una compañía es la reputación corporativa. Distintas investigaciones han postulado que elementos del Social Media pueden tener correlación con el mencionado indicador. Ello implica, tentativamente, que de alguna manera diferentes elementos presentes en la web pueden influenciar en la reputación de una corporación.

Debido a esto, el presente documento busca determinar si las variables del Social Media tienen relación con la reputación corporativa para el caso del BCP Bolivia. La tesis está dividida en 3 capítulos. En el primero, se abarca la problemática en sí con fundamentos teóricos y artículos de revistas indexadas como antecedentes. Luego, en el segundo capítulo, se hace referencia al método, sección donde se describe de dónde se obtienen una parte de los indicadores Social Media, así como una breve introducción a las redes neuronales. Finalmente, en el capítulo 3 se muestran los resultados de la investigación, así como conclusiones, recomendaciones y limitaciones.

Capítulo I – Problema de Investigación

1.1. Problema de Investigación

1.1.1. Planteamiento del Problema

Horn & Taros (2015) y Abril (2003) indicaron que la reputación corporativa es un concepto que se puso de moda a mediados de la primera década del nuevo milenio. Antes del año 2000, solo un sector de la academia comenzó a elaborar dicho concepto (Abril, 2003). Poco después, los corredores de seguros deciden incluir la pérdida de reputación en las listas de riesgos que pueden afectar a las corporaciones (Grutzmacher, 2011). En la actualidad, el término en sí forma parte del ámbito de la gestión en su totalidad (Abril, 2003).

De acuerdo a Abril (2003), conceptos como marca, relaciones institucionales, imagen y comunicación, constituían la aproximación más esencial de lo que las empresas proyectan. En la actualidad, se menciona con mayor frecuencia la reputación, ética corporativa, buen gobierno y responsabilidad social.

El Banco de Crédito de Bolivia (abreviado como BCP y también conocido como BCP Bolivia) es una empresa financiera con sede principal en La Paz, capital boliviana (BCP-Bolivia, 2019). Es una filial internacional que forma parte del Banco de Crédito del Perú, entidad que pertenece a CREDICORP, una de las entidades empresariales más grande del Perú (BCP-Bolivia, 2019). El Banco de Crédito de Bolivia comenzó sus actividades en el país altiplánico tras la compra del Banco Popular, operación que se concretó en 1994.

Posteriormente el BCP Bolivia continuaría con sus planes de expansión adquiriendo el Banco de la Paz en 1998 y posteriormente parte de la cartera del Banco Boliviano Americano (BCP-Bolivia, 2019). Como consecuencia el BCP Bolivia se ha ido posicionando como uno de los más importantes del país. Luego, la entidad llegó a ser considerada por el Monitor Empresarial de Reputación Corporativa (Merco), empresa afiliada a la multinacional KPMG, como una de las más importantes del sector de servicios financieros. Desde el año 2013 hasta la

actualidad, el BCP Bolivia se ha mantenido dentro del ranking de las empresas financieras más importantes del país altiplánico.

Naturalmente, el BCP Bolivia ha sido objeto de interés en motores de búsqueda como Google. Por otro lado, diversos diarios bolivianos han ido informando sobre distintos sucesos vinculados a la entidad financiera. No solo eso, también en portales como YouTube distintos usuarios han realizado comentarios acerca de la institución. Una proporción de la información en donde se valora al BCP Bolivia se encuentra en formato de texto. Como bien señalaron Ghose, Ipeirotis, & Sundararajan (2007), los textos pueden ser procesados y convertidos a patrones numéricos. Una compañía puede ser objeto de múltiples valoraciones *online*. Actualmente, y debido a la incorporación de recientes tecnologías, es posible obtener patrones numéricos a partir de dicha información textual (Jasso-Hernández, Pinto, Vilariño & Lucero, 2014).

Las búsquedas en Google, las noticias online, los comentarios en YouTube pueden ser considerados como elementos pertenecientes al Social Media (Fuchs, 2014). Investigadores como Peetz, de Rikje & Kaptein (2015) sugieren que los elementos del Social Media pueden tener relación con la reputación. Precisamente, el objetivo de la presente investigación es determinar si la reputación corporativa del BCP Bolivia guarda relación con el entorno Social Media que lo ha rodeado.

1.1.2. Formulación del Problema

¿La reputación corporativa del BCP Bolivia tiene relación con las variables Social Media de su entorno durante el período 2013-2018?

1.1.3. Justificación

Debido a la importancia del Social Media en un entorno digital, es necesario conocer si dicha variable termina impactando sobre la reputación de una empresa. El caso del BCP representa una aplicación importante, ya que es uno de los casos de internacionalización más representativos en el contexto peruano y uno de los muy pocos en el ámbito de servicios financieros.

El hecho de procesar data que proviene de un entorno Social Media es un ejemplo del aprovechamiento de datos secundarios. La exploración de datos estructurados y no estructurados (textos, por ejemplo) provenientes de plataformas electrónicas forma parte de las nuevas tendencias académicas presentes en distintos *journals* internacionales. Tradicionalmente, una investigación de esta naturaleza requeriría la aplicación de instrumentos prediseñados (encuestas principalmente) en un país extranjero (Bolivia). Ello es poco viable dadas las naturales limitaciones de tiempo y recursos disponibles.

La investigación también tiene implicancias sobre el marketing electrónico. Si es que efectivamente existe relación entre la reputación corporativa del BCP Bolivia y las variables del tipo Social Media, sería posible aproximarse a la construcción de un indicador de reputación corporativa, semestral, trimestral e incluso mensual. Esto debido a que los elementos del Social Media (búsquedas en Google, comentarios en YouTube, noticias en los diarios) siempre están presentes; mientras que la medición de la reputación corporativa requiere de un considerable esfuerzo.

El proceso de extracción de datos de plataformas *online* para una posterior identificación de patrones o tendencias forma parte de lo que se conoce como Minería de Datos, este método es un camino alternativo para investigadores que deseen explotar datos de libre acceso (Henriquez, Guzmán, & Salcedo, 2016). Precisamente el *Data Mining* (o Minería de Datos) forma parte de los nuevos rumbos de investigación. En los repositorios académicos de universidades peruanas no es frecuente encontrar trabajos que empleen el *Data Mining* dentro del campo de las ciencias empresariales.

La investigación está circunscrita en las líneas de investigación de la carrera de Negocios Internacionales de la Universidad San Ignacio de Loyola, ya que se trata de un aspecto vinculado a una empresa de origen peruano en un mercado extranjero. Debe destacarse que la producción académica de la universidad se ha centrado más en aspectos de empresas extranjeras o negocios internacionales dentro del ámbito peruano, por lo que estudiar a una empresa peruana en un mercado extranjero resulta un aporte desde otra perspectiva.

1.2. Marco Referencial

1.2.1. Antecedentes

El BBVA (2018) tuvo por objetivo construir un Índice de Percepción de Corrupción en tiempo real y de alta frecuencia a partir de Google Trends. La entidad mostró que la percepción mundial sobre la corrupción ha aumentado desde 2009, con gran heterogeneidad entre países. Sobre el caso de Brasil, se demostró que una mayor percepción de corrupción reduce la aprobación del gobierno así como la confianza. Para demostrar ello se tomó una muestra de 168 series mensuales (desde el 2004 hasta el 2017). Entre otras sugerencias, la entidad sugiere que el estudio puede ser aplicado al ámbito privado.

Barbato & Montesi (2016) investigaron si la reputación corporativa se puede evaluar desde datos de redes sociales. Los autores indicaron que la web 2.0 y la aparición de los medios sociales crearon nuevas formas de comunidad, ya que permitían a la gente compartir su opinión con los demás a escala global, a menudo hablando con extraños sobre el tema que apenas conocen. Incluso si la comunicación es en su mayoría "impersonal" todavía da una opinión compartida subyacente para ser evaluada. Los investigadores se centraron en el escándalo de Volkswagen y en el efecto que creó en la red social Twitter. El escándalo de VW fue elegido debido a su amplia evolución a través del tiempo y sus amplios efectos sobre el rendimiento financiero de VW. Con el fin de cumplir el objetivo de investigación, fueron recopilando tweets sobre VW. El conjunto de datos incluyó tweets de tres semanas antes de la fecha de fuga de escándalos hasta nueve meses después de que saliera. Los investigadores demostraron que existió correlación entre la carga emocional de los comentarios y el valor de la empresa en la Bolsa de Valores.

Vidya, Fanany & Budi (2015) señalaron que la competencia entre los proveedores de telefonía móvil para adquirir nuevos clientes a través de campañas y publicidad representan una guerra, especialmente en medios de comunicación social. Los investigadores señalaron que los problemas surgen sobre cómo medir la reputación de la marca de estos proveedores basados en las respuestas de un "n" personas sobre la calidad de sus servicios. La investigación aborda este problema midiendo la reputación de la marca basada en el análisis de sentimiento de

publicaciones en Twitter. Los datos extraídos fueron 10,000 comentarios de Twitter de enero a marzo de 2015 acerca de tres proveedores de telefonía móvil en Indonesia. Los investigadores emplearon diferentes algoritmos y determinaron con un nivel de precisión de entre 39% y 65% la satisfacción de los consumidores usando como *inputs* los puntajes sentimentales.

Shulze-Horn y otros (2015) sugirieron que los sentimientos evidenciados en los comentarios online pueden incidir sobre la reputación corporativa. Los investigadores llegaron a esta conclusión luego de un análisis de casos en donde se destacó la información de plataformas como YouTube y Vimeo como elementos a tomar en cuenta para el estudio de la reputación de marcas de Inglaterra. Los investigadores sugirieron, entre otras cosas, estudiar los sentimientos reflejados en los comentarios y los niveles de actividad en las plataformas web como variables que inciden en la reputación de las empresas.

Peetz, de Rikje & Kaptein (2015) indicaron que en la gestión de la reputación, saber qué impacto tiene un tweet en la imagen de una marca o la empresa es crucial. La reputación asociada a la polaridad de un *tweet* es una medida de cómo el tweet influye en la reputación de una marca o empresa. Los autores identificaron que el proceso requiere ser automatizado. Para calificar un tweet como positivo, negativo o neutro (polaridad sentimental), los investigadores desarrollaron un modelo fundado en características basadas en tres dimensiones: la fuente del tweet, el contenido y la recepción del mismo. Es decir, cómo se percibe el tweet. Para verificar que existía relación entre la reputación y los datos de Twitter se emplearon datos de la empresa *Replab* entre los años 2012 y 2013 bajo modelos regresivos. Al final los investigadores determinaron que la reputación guardaba correlación con la proporción de valoraciones positivas en los comentarios.

Grutzmacher (2011) tuvo por objetivo examinar el papel de los medios de comunicación en la reputación corporativa. A más detalle, el estudio se centró en la pregunta de que si la participación de las empresas en las redes sociales afecta su reputación. Para ello se eligió a Nokia. El estudio se dividió en tres áreas problemáticas: (1) descubrir qué parte potencial de los usuarios de los medios eran influenciados por las actividades de la empresa en una red social y quién participó

en las actividades de redes sociales de Nokia; (2) examinar si en los medios sociales los usuarios eran potencialmente conscientes de la participación y reputación de Nokia; y (3) observar si la participación de las empresas en las redes sociales influían de manera positiva o negativa en la reputación corporativa en general. El estudio empleó 284 comentarios. Determinó principalmente que la actividad en redes sociales guarda relación con la reputación de la empresa.

Luca (2011) tuvo por objetivo determinar si las opiniones de consumidores en línea afectan la demanda del restaurante. Para ello empleó datos de *Yelp.com* y los datos de un conjunto de restaurantes de New York. El investigador escogió *Yelp.com* como muestra debido a que esta plataforma muestra con prominencia la puntuación media. El investigador presento tres conclusiones sobre el impacto de las opiniones de los consumidores en la industria de restaurantes: (1) un aumento de estrellas en la calificación de *Yelp* conduce a entre un 5 y 9% en el aumento de ingresos, (2) este efecto es impulsado por restaurantes independientes; las calificaciones no afectan a restaurantes franquiciados; (3) las cadenas de restaurantes han disminuido su cuota de mercado a medida que la penetración *Yelp* aumentaba. Esto sugiere que las opiniones en línea son un sustituto del antiguo sistema de recomendación; (4) los consumidores no usan toda la información disponible y responden mejor a la calidad cambios que son más visibles; y (5) los consumidores responden con más fuerza cuando una calificación contiene más información.

1.2.2. Marco Teórico

1.2.2.1. Social Media.

Según Andreas & Michael (2010), el Social Media puede definirse como un conjunto de aplicativos propios del Internet que han sido desarrollados tomando en consideración fundamentos tecnológicos e ideológicos e la Web 2.0, entorno en donde los usuarios son quienes generan el contenido (de ahí que se les denomine creadores de contenido). Luego, Manning (2014) define al Social Media como un término que se usa a menudo para denominar a nuevas formas de medio que dan un papel primordial a la

participación interactiva. Manning (2014) también señaló que el desarrollo del Social Media puede dividirse en dos etapas: la era de transmisión y la era interactiva. En la primera, los medios de comunicación estaban centralizados en una entidad: una estación de TV, de radio, un periódico, etc.

Con respecto a la segunda etapa, Manning (2014) señaló que el auge de las tecnologías digitales (y en particular la tecnología móvil), facilitó la interacción a una escala nunca antes vista. En consecuencia, surge una nueva era de los medios donde la interactividad ocupa un lugar fundamental. Una persona ahora puede comunicarse con muchas, y obtener una respuesta en lapsos muy breves (Andreas & Michael, 2010). Los consumidores a menudo veían limitadas sus voces, ahora comparten sus opiniones con sus pares (Manning, 2014). Ello puede deberse al incremento en el nivel de accesibilidad a las nuevas tecnologías como consecuencia de precios cada vez más bajos. En consecuencia, los individuos son capaces de buscar información de distintas fuentes y dialogar con sus pares a través de foros de mensajes (Andreas & Michael, 2010).

Manning (2014) indicó que una característica inherente a cualquier componente del Social Media es el hecho de requerir algún tipo de plataforma digital (estacionaria o móvil). Sin embargo, de acuerdo a la perspectiva particular del autor, no todo lo que se encuentre en formato digital puede ser denominado Social Media. De acuerdo a Manning (2014) existen dos características esenciales que permiten delimitar con propiedad lo que es un entorno Social Media.

- Alguna forma de participación: en los elementos del Social Media la pasividad no puede ser una característica predominante, aunque plataformas como Tik-Tok, YouTube o Facebook permiten asumir el rol de espectador pasivo ante lo que los generadores de contenido publican. Por lo general, como mínimo, se debe crear un perfil que facilite el inicio de las interacciones (Manning, 2014). Esta característica en sí misma hace que los elementos del Social Media se distingan de otros más tradicionales.

- La naturaleza participativa: el Social Media implica participar, ya sea con el entorno familiar a amical, e incluso con conocidos y con desconocidos con los que se comparten intereses comunes. Según Manning (2014) y Andreas & Michael (2010), citados por Farfan & Zapata (2018) son ejemplos de Social Media: el correo electrónico, foros, blogs, redes sociales, plataformas de *streaming* de videojuegos, etcétera.

Los postulados de Manning (2014) y Andreas & Michael (2010) no son los únicos que buscan definir lo que es un Social Media, existen otros investigadores como Dergiades, Milas & Panagiotidis (2015) que han incorporado el concepto de volumen. Ellos no niegan que el Social Media implica interacción y participación, pero agregan que un Social Media real involucra que los niveles de actividad en la web alcancen una masa crítica.

La masa crítica que mencionan Dergiades, Milas & Panagiotidis (2015) puede medirse con cantidad de *tweets* o búsquedas en Google o YouTube. Otros autores como Fulmer & Shapiro (2016) coinciden con los anteriormente citados usando el caso de la caída de Hosni Mubarak, quien sobrevivió a seis intentos de asesinato, a numerosas ofensas de los partidos de la oposición y a un movimiento islamista insurgente. Él controló el poderoso aparato de seguridad de Egipto, siguió siendo un firme aliado de Estados Unidos, y amplió el poder de una draconiana Ley de Emergencia a través de la cual podría suspender los derechos constitucionales a voluntad. Sin embargo, el 11 de febrero de 2011, una oleada desarmada y no organizada de egipcios derrocó a Mubarak en la mayor serie de protestas en Egipto hasta la fecha (Fulmer & Shapiro, 2016). Los investigadores indican que el Social Media cumplió un rol importante, pero que la caída de Mubarak no hubiera sido posible si es que el sentimiento de rebelión no alcanzaba niveles críticos. El interés hacia tópicos en específico puede ser medido con Google Trends, tal como ocurrió para el caso de la rebelión en Egipto.

1.2.2.2. Los datos textuales como datos no estructurados.

Castillo & Cheang (2019), citando a Thelwall (2017), indicaron que los datos en formato de texto libre, los predominantes en las redes sociales pertenecen a los Datos no Estructurados. Inclusive entre los datos textuales, los datos ingresados en dicho formato poseen una cantidad de estilos demasiado diversos, dicho de otra forma, tienen un alto grado de heterogeneidad. YouTube es un buen ejemplo de ello: los usuarios pueden escribir sus opiniones con libertad (Thelwall, 2017).

Los comentarios en YouTube, además, no evidencian el uso de un lenguaje estandarizado (Kalambe, Pratiba & Shah, 2015). En contraste, las cartas y/o solicitudes poseen estructuras más definidas, de igual manera las hojas de vida publicadas en *LinkedIn* (Thelwall, 2017). Esto también es indicado por Fulmer & Shapiro (2016), quienes además agregaron que los datos textuales estandarizados también pueden encontrarse en la mayoría de la prensa escrita.

1.2.2.3. Minería de Textos.

Una de las áreas de investigación de mayor modernidad y en constante crecimiento en el ámbito del *Computer Science* es el *Text Mining* (también conocida Minería de Datos Textuales) (Silge & Robinson, 2017). Castillo & Cheang (2019), citando a Kalambe, Pratiba & Shah (2015), indican que esta área del conocimiento pretende sacar provecho de la considerable sobrecarga de datos textuales mediante una serie de procesos que evidencia interdisciplinariedad, ya que involucra áreas: Procesamiento del Lenguaje Natural, Matemática, Estadística y *Machine Learning*.

Las principales tareas del *Text Mining* consisten en la extracción, descubrimiento e identificación de vínculos entre una cantidad considerables de datos en formato textual. (Silge & Robinson, 2017). El *Text Mining* posibilita convertir datos textuales a números, números que pueden ser objeto de tratamiento estadístico. (Hofmann & Chisholm, 2015). De hecho, a partir de datos textuales convertidos en índices

numéricos ha podido contruirse modelos de regresión, redes neuronales; así como realizar pruebas de correlación (Silge & Robinson, 2017).

En la investigación se han extraído comentarios de YouTube usando herramientas como NVivo (ver sección 2.4.1) para un posterior depuración y transformación de español a inglés (ver sección 2.5). La extracción y transformación son las dos primeras etapas de la minería de datos empleadas en la tesis.

1.2.2.4. *Análisis Sentimental.*

De acuerdo a Oliva (2014), citados por Castillo & Cheang (2019) , el análisis sentimental, como campo de investigación, tiene un estrecho vínculo con lo que se puede considerar una parte del Procesamiento del Lenguaje Natural, campo orientado al estudio de la interacción entre los ordenadores y el lenguaje empleado por las personas (Nadkarni, Ohno-Machado & Chapman, 2011). Dicho de otra forma, busca mecanismos óptimos para la comunicación hombre-máquina empleando las distintas lenguas existentes en el mundo (Nadkarni, et.al. 2011). Minería de Opinión u *Opinion Mining* son otros de los nomenclaturas que han sido empleadas para referirse al Análisis Sentimental (Castillo & Cheang, 2019).

Oliva (2014) señaló que los sentimientos visibles en textos se ven en dos formas: las primera es explícita, donde la frase expresa una opinión de forma directa (“Fue un día espantoso y asqueroso”); mientras que la segunda es implícita (“El auto se malogró en menos de una semana”). Un sector de la academia se ha concentrado en las expresiones explícitas debido a su facilidad de análisis (Oliva, 2014).

La polaridad sentimental es una rasgo inherente de los textos valorativos (Oliva, 2014). La polaridad se hace presente, regularmente, de forma dicotómica: el texto puede expresar sentimientos positivos o negativos, e incluso pueden producirse escenarios donde exista un equilibrio entre lo positivo y negativo. En estos casos se habla de polaridad neutra (Silge & Robinson, 2017).

El *Sentiment analysis* (análisis del sentimiento) ha concitado la atención de académicos e investigadores en los últimos años. En esta área del conocimiento, se ha procurado medirse los sentimientos presentes en un contenido textual (Blanchard, 2011). El análisis de sentimiento es uno de los ejes transversales de las investigaciones orientadas a la explotación de datos textuales propios de entornos Social Media (Fuchs, 2014).

Por definición, la opinión es una creencia, una decisión en la que no predominan elementos de convicción o prueba, o bien carecen de ellos (Ghose, Ipeirotis, & Sundararajan, 2007). En este contexto, esto es contradictorio con los hechos. Por lo tanto, los sentimientos que expresan opiniones se consideran subjetivos, mientras que las afirmaciones fácticas o reales se consideran objetivas. El sentimiento está casi relacionado con la emoción personal y se usa para evaluar la opinión de la persona (Ghose, et.al. 2007). En este sentido, las declaraciones que transmiten un sentimiento son tanto objetivas como subjetivas (Andreas & Michael, 2010).

1.2.2.5. Análisis de sentimiento como herramienta de investigación.

Analizar la carga sentimental en textos es una tarea que reviste una gran importancia en la actualidad. Existe una multitud de aplicaciones que pueden resultar beneficiadas de procedimientos computacionales que permitan detectar automáticamente si la intención del autor ha sido expresarse de manera “positiva”, “negativa”, “objetiva”, “subjetiva” o “neutra” (Silge & Robinson, 2017).

Por ejemplo, puede darse el caso de una personalidad política requiera saber si la ciudadanía tiene una apreciación positiva o negativa sobre su persona (Whu & Hsieh, 2014). Otro ejemplo, podría ser la determinación de la reputación para una institución pública o privada (Peetz, de Rikje & Kaptein, 2015). En ambos ejemplos se requiere analizar el punto de vista de las personas (usuarios) acerca de las entidades que son objeto de valoración. Si bien, en el pasado era complicado aplicar

cuestionarios hacia los usuarios, en la actualidad esta práctica es raramente usada debido principalmente a los siguientes inconvenientes):

- a. La aplicación de cuestionarios es un proceso que puede ser oneroso, tanto desde un punto de vista financiero como en la cantidad de tiempo invertido (Jasso-Hernández, Pinto, Vilariño, & Lucero, 2014).
- b. Dichos cuestionarios requieren ser tabulados para ser analizados en una etapa posterior (Jasso-Hernández, et.al. 2014).
- c. La aplicación de cuestionarios requiere que el proceso de selección de candidatos sea llevado cuidadosamente a fin de garantizar resultados adecuados (Jasso-Hernández, et.al. 2014).
- d. El analizar los datos es una tarea que debe llevarse a cabo con rapidez a fin de evitar la obsolescencia de las conclusiones (Jasso-Hernández, et.al. 2014).

Según Kumlin (2016), las redes sociales proporcionan elementos que dan paso a la *Big Data*. El aprovechamiento de datos de distintas redes para dar paso a la *Big Data* es una de las innovaciones más importantes de la última década. Esta innovación tecnológica está dando forma a la experiencia del cliente y les ofrece un canal de conexión.

1.2.2.6. Social Media y Negocios.

Las distintas formas del Social Media revelan la capacidad de los individuos por hacer escuchar su voz. Ello se ha dado a través de la democratización de contenidos e ideas que pueden ser expresados en torno a pasiones comunes (Rincón-García, 2017). Las redes sociales pueden también inspirar movimientos y encender el cambio (Kumlin, 2016). Hoy en día la cantidad de datos producidos a nivel mundial es vasta. Por ejemplo, en Internet se producen millones de datos debido a la utilización masiva de las redes sociales, servicios de mensajería, blogs, wikis, correo electrónico, entre otros (Rincón-García, 2017).

Toda esa gama de datos es atractiva para diferentes estamentos comerciales, industriales y académicos, pero la extracción y su respectivo procesamiento, hace que esta tarea sea muy compleja y difícil si se hace

de forma manual (Horn & Taros, 2015). Sumado a esto, las personas comunes participan activamente en Internet dejando sus propios comentarios, opiniones y hasta reseñas, en todo tipo de temas, usando su lenguaje nativo (Blanchard, 2011).

Desde una perspectiva empresarial, monitorear y analizar las opiniones de los usuarios o clientes se convierte en algo casi esencial en un entorno de alta competitividad (Rincón-García, 2017). Sin embargo, en mayor o menor medida, las empresas se enfrentan al problema de analizar un gran volumen de opiniones relativas a productos o servicios propios o de la competencia, y de múltiples fuentes, como pueden ser blogs, artículos de opinión o redes sociales. Lo que hace inviable abordar manualmente el análisis de esas opiniones, siendo necesario establecer procesos automáticos que se encarguen de la tarea (Ghose, Ipeiritis, & Sundararajan, 2007).

1.2.2.7. *Reputación Corporativa y el entorno digital.*

¿Qué es la reputación corporativa? ¿Cómo se puede medir? De acuerdo a Blanchard (2011) estas dos preguntas han sido ampliamente discutidas por los académicos, sin llegar a una definición compartida o a la metodología de evaluación. Cada investigación ha definido la reputación de la empresa en función de su ámbito específico: los economistas consideran a la reputación corporativa como los resultados de las decisiones estratégicas pasadas de la empresa, mientras que los investigadores lo definen como un recurso (Blanchard, 2011).

En marketing, la reputación corporativa es inadecuada debido a que se puede confundir con el concepto de imagen de marca (Horn & Taros, 2015). Los sociólogos se centran en el papel de la interpretación de las acciones de una empresa realizada por intermediarios en un mundo de información incompleta (Luca, 2011). Las teorías organizativas comparten parcialmente la definición de los estudios, pero se centran en el apoyo de las instituciones a la larga duración de la empresa (Horn & Taros, 2015). A pesar de que no existe una definición común, todos los estudios están de

acuerdo en un punto: la reputación corporativa es el resultado de la relación entre una empresa y sus partes interesadas (Henriquez y otros, 2016).

La industria de las redes sociales está experimentando un crecimiento exponencial. Existen millones de dispositivos, sistemas y usuarios conectados a internet que generan estos tipos de datos. Las redes sociales no son solo una herramienta de comunicación crítica con clientes para cualquier marca en todas las industrias, ya que pueden contribuir a estimar la demanda en tiempo real (Blanchard, 2011).

De acuerdo a Barbato & Montesi (2016) la popularidad en el uso de redes sociales ofrece oportunidades que deben ser tomadas en cuenta. Los datos provenientes de redes sociales ayudan a obtener información en tiempo real sobre distintos aspectos del negocio. Ello incluye la experiencia del cliente y reputación de la empresa (Kumlin, 2016). Lo que la gente esté hablando de una empresa o marca en línea, puede que sea escuchado (eso dependerá de la empresa). No obstante, administrar datos de redes sociales y entenderlos es una ventaja (Kumlin, 2016)

Según Blanchard (2011), uno de los errores más comunes cometidos por empresas cuando están planeando tomar las redes sociales en uso, es que se centran demasiado en las herramientas de medios sociales y las plataformas, y no suficientemente en los objetivos del negocio. Lovett (2011) manifestó que medir las redes sociales es una tarea desafiante y gratificante al mismo tiempo. El principal reto es identificar los objetivos y métricas que tengan significado y sean importantes para el negocio porque la medición en sí no es el objetivo (Blanchard, 2011). Otro objetivo consiste en colaborar con las partes interesadas para alinear las métricas que realmente identificarán el progreso hacia los objetivos de negocio (Blanchard, 2011).

Lo siguiente es encontrar las herramientas y los canales apropiados. Es necesario tomarse el momento de identificar los canales donde se concentra el mercado potencial. Porque para la mayoría, el canal más popular puede no ser necesariamente el mismo que para las demás empresas. Los profesionales de la comunicación deben considerar una

combinación de medios para enviar información corporativa a los consumidores (Barbato & Montesi, 2016)

Li (2011) señaló que los gustos de Twitter, el sitio de microblogging; Facebook, la herramienta de redes sociales; y YouTube, la plataforma para compartir videos, es un fenómeno sociocultural. Hecho posible con el advenimiento de Web 2.0, un conjunto de tecnologías que permiten a los usuarios web congregarse, interactuar y compartir información en línea, y conocido por el colectivo sustantivo social media, han transformado el panorama mediático para empresas.

Pero en términos de reputación, Li (2011) se preguntó qué valor ofrecen las redes sociales a las organizaciones. A través de entrevistas con líderes en los campos de las redes sociales y corporativas buscó responder a esa pregunta. Li (2011) concluyó sugiriendo que las redes sociales han anunciado un nuevo amanecer para el concepto de gestión de la reputación, una donde las organizaciones pueden, por primera vez, comenzar a manejar sus reputaciones directamente.

Szwajka (2017) indicó que en los últimos años el Social Media ha contribuido en cambiar el entorno de comunicación en el que las partes interesadas como usuarios pueden publicar y distribuir sus opiniones sobre la empresa y sus productos. Esto genera una serie de potenciales amenazas a la imagen y reputación tanto de personas como de organizaciones. Las empresas no pueden ignorar este hecho y deben utilizar el Social Media no solo como un canal de marketing de comunicación adicional, sino también en un ámbito más amplio (Szwajka, 2017). Es decir, como una herramienta para construir y proteger su reputación.

Las plataformas digitales también han aumentado la importancia estratégica de la gestión de la reputación corporativa, lo que hace más difícil gestionar debido a que no pueden controlar todo lo que se dice acerca de las empresas; las redes sociales son un ejemplo de ello (Henriquez, Guzmán, & Salcedo, 2016). Las empresas están cada vez más preocupadas por la información y las opiniones que se pueden dar a través de las comunidades en línea rápidamente sin ningún control (Blanchard,

2011). Mientras que la actividad en redes sociales aumenta, el poder de las partes interesadas también, ya que representan una gran fuente de información sobre los sentimientos, y opiniones de las personas que es posible monitorear en tiempo real (Grutzmacher, 2011).

1.2.2.8. La reputación corporativa y la medición de Merco.

Zanfrillo, Artola, & Morettini (2016) indicaron que se trata de un instrumento de evaluación reputacional lanzado en el año 2000, basado en una metodología *multistakeholder* compuesta por cinco evaluaciones y doce fuentes de información. Merco, a inicios del 2020, construye ocho rankings (Merco Talento, Merco Consumo, Merco Talento Universitario, Merco Digital MRS, Merco Empresas, Merco Líderes, Merco Responsabilidad y Gobierno Corporativo) (Merco, 2020). Por otro lado, tiene presencia en doce países: España, Argentina, Colombia Chile, Brasil, México, Perú, Ecuador, Bolivia, Costa Rica, Panamá y Portugal (en proceso) (Merco, 2020).

Ferruz (2018) indicó que el índice proporcionado por Merco fue uno de los primeros en ser auditado en el mundo: el seguimiento y la verificación del indicador fue llevada a cabo por KPMG, tomando en consideración la norma Norma Internacional de Encargos de Aseguramiento. Los criterios ponderados son de carácter público, pueden ser consultados en el mismo sitio donde se visualizan los resultados.

Pero ¿de qué manera Merco obtiene los resultados? De acuerdo a Zanfrillo, Artola, & Morettini (2016), para el caso de Reputación Corporativa, las variables analizadas por una muestra de directivo incluye resultados financieros, reputación interna, innovación y ética empresarial, (Carrasco, 2017). Por su parte, un grupo de periodistas especializados en el análisis de información económica valoran la transparencia informativa y la accesibilidad a la información. Un grupo de analistas financieros hace lo propio respecto con los resultados financieros y a la calidad de la información económica, Finalmente, un grupo de ONG's valora, además, el compromiso con la comunidad y la responsabilidad social y medioambiental (Ferruz, 2018).

1.2.2.9. Búsquedas en Google y YouTube, y el contenido noticioso como elementos relacionados a la reputación corporativa.

De acuerdo a Dergiades, Milas & Panagiotidis (2015) y Shulze-Horn y otros (2015), indicadores asociados a las búsquedas en Google y YouTube pueden tener relación con la reputación corporativa. Y es que, por ejemplo, los escándalos y sucesos adversos pueden asociarse no solo a la reputación, sino también a los resultados financieros de una empresa (Dergiades, Milas & Panagiotidis, 2015). Las búsquedas en Google pueden evidenciar hechos importantes que tienen relación con una compañía, como en el caso de Huawei:



Figura 1: Interés mensual hacia Huawei (setiembre del 2014 hasta setiembre del 2019) en el mundo de acuerdo al volumen de búsquedas

Fuente: Google Trends (2019a)

En la figura anterior es posible apreciar que el interés hacia la empresa Huawei alcanzó un pico máximo en la semana del 19 al 25 de mayo del 2019, período en el que se dieron a conocer las sanciones impuestas por la administración Trump a la compañía China. Ello ocasionó un incremento en las búsquedas acerca de la compañía asiática.

Con respecto a YouTube, también es posible identificar externalidades que afectan a la presencia online de una compañía, tal como evidencia la siguiente figura:

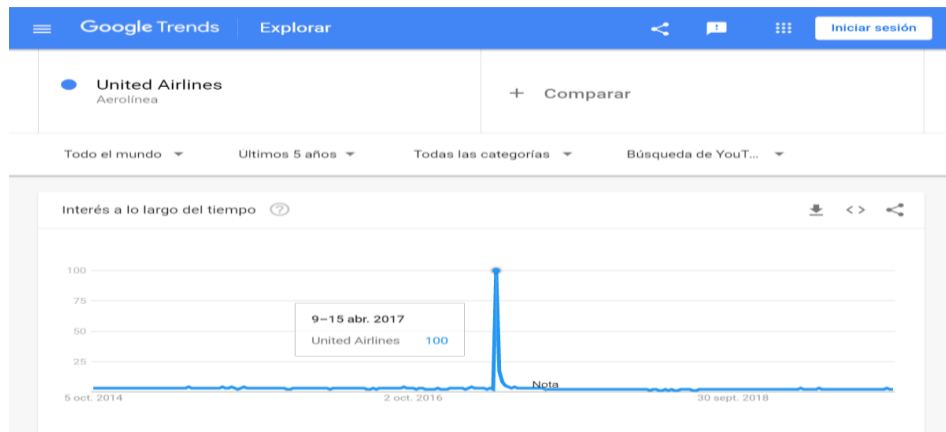


Figura 2: Interés mensual hacia United Airlines (setiembre del 2014 hasta setiembre del 2019) en el mundo de acuerdo al volumen de búsquedas

Fuente: Google Trends (2019b)

En el caso de United Airlines, las búsquedas en YouTube acerca de dicha empresa se incrementaron considerablemente en la semana del 9 al 15 de abril del 2017. Ello ocurrió debido a que en ese lapso se produjo un escándalo que se viralizó principalmente a través del mencionado portal: el desembarco violento de un pasajero de origen chino que se negaba a abandonar su vuelo.

De acuerdo a Llorente & Cuenca (2017), esta última situación es un ejemplo de crisis en la reputación. Se está ante el empoderamiento del consumidor: escenario donde las redes sociales hacen que los consumidores y el público en general tomen el control de la conversación durante todo el día utilizando múltiples plataformas. Actualmente, allí es donde se realizan las conexiones y se construyen o destruyen marcas y compañías (Llorente & Cuenca, 2017). El reto para las empresas es participar en la conversación y crear la narrativa adecuada para una compañía basada en un propósito compartido que ponga al ciudadano, consumidor y cliente en el centro de nuestras decisiones (Llorente & Cuenca, 2017).

Los ejemplos antes presentados son ejemplos de cómo existen ciertas noticias que pueden impactar en la reputación corporativa. Pero en esta época la velocidad con la que fluye la información es mucho mayor. Las noticias que vinculan a las corporaciones siempre nos han mostrado hechos que están vinculados a un conjunto de empresas. Los indicadores

vinculados a Google y YouTube proporcionan evidencia para determinar si los distintos sucesos terminan o no llamando la atención por parte de los internautas.

Dicho de otra forma, no basta con que exista un contenido noticioso que pueda afectar la reputación corporativa, sino que además dicho contenido debe generar interés de los potenciales lectores. El interés hacia ciertas noticias podrían explicar mejor la reputación corporativa que la cantidad de noticias acerca de una entidad particular.

BCP – Bolivia.

De acuerdo a Moody's (2017) el Banco de Crédito de Bolivia (abreviado como BCP - Bolivia) es una compañía de servicios bancarios y financieros ubicada en Bolivia. Tiene su sede en La Paz y es una subsidiaria internacional de propiedad absoluta del Banco de Crédito del Perú, el banco más grande de Perú. BCP Bolivia inició sus operaciones en el país luego de la adquisición por parte de BCP del banco boliviano Banco Popular. BCP Bolivia ofrece sus productos y servicios a través de 102 sucursales operativas, 240 cajeros automáticos, centros de llamadas y plataformas bancarias en línea y móviles. El Banco de Crédito de Bolivia es actualmente el cuarto banco más grande de Bolivia por activos totales.

De acuerdo a la Bolsa de Valores de Bolivia (2018), las acciones de la empresa se dividen de la siguiente forma:

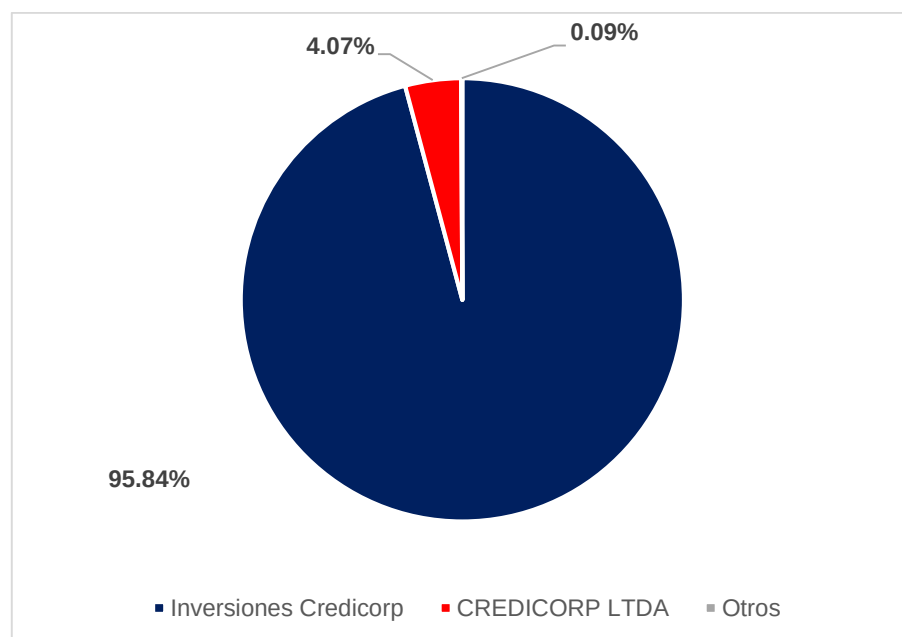


Figura 3: Distribución porcentual de las acciones del BCP Bolivia

Fuente: Bolsa de Valores de Bolivia (2018)

Por otro lado, la misma Bolsa de Valores de Bolivia (2018) indica que la empresa fue constituida en 1992 y que en la actualidad cuenta con 1,708 trabajadores. Por otro lado de acuerdo a la consulta PWC (2017), el BCP – Bolivia registró ingresos financieros por más de 175 millones de bolivianos, lo cual equivale a aproximadamente 24.5 millones de dólares.

En cuanto al interés que suscita el BCP Bolivia en territorio boliviano se cuenta con datos de Google Trends. Esta herramienta registra las búsquedas de un tema determinado en un espacio geográfico específico. Dicha herramienta muestra solo índices relativos y no absolutos. Dicho de otra forma, si el término de búsqueda alcanza un pico de 120,000 búsquedas en un espacio geográfico determinado, y dicho pico es el mayor registrado, entonces Google Trends indicará que el interés hacia dicho término es 100. Si en otra ciudad, las búsquedas son 60,000 entonces, el índice de Google será 50. Esto nace a partir de una regla de tres sencilla.

Los datos acerca de las operaciones del BCP Bolivia por departamento no son públicos. Sin embargo, es tentativo creer que las búsquedas web tienen correlación con las transacciones realizadas. Los resultados para la popularidad del BCP son los siguientes:

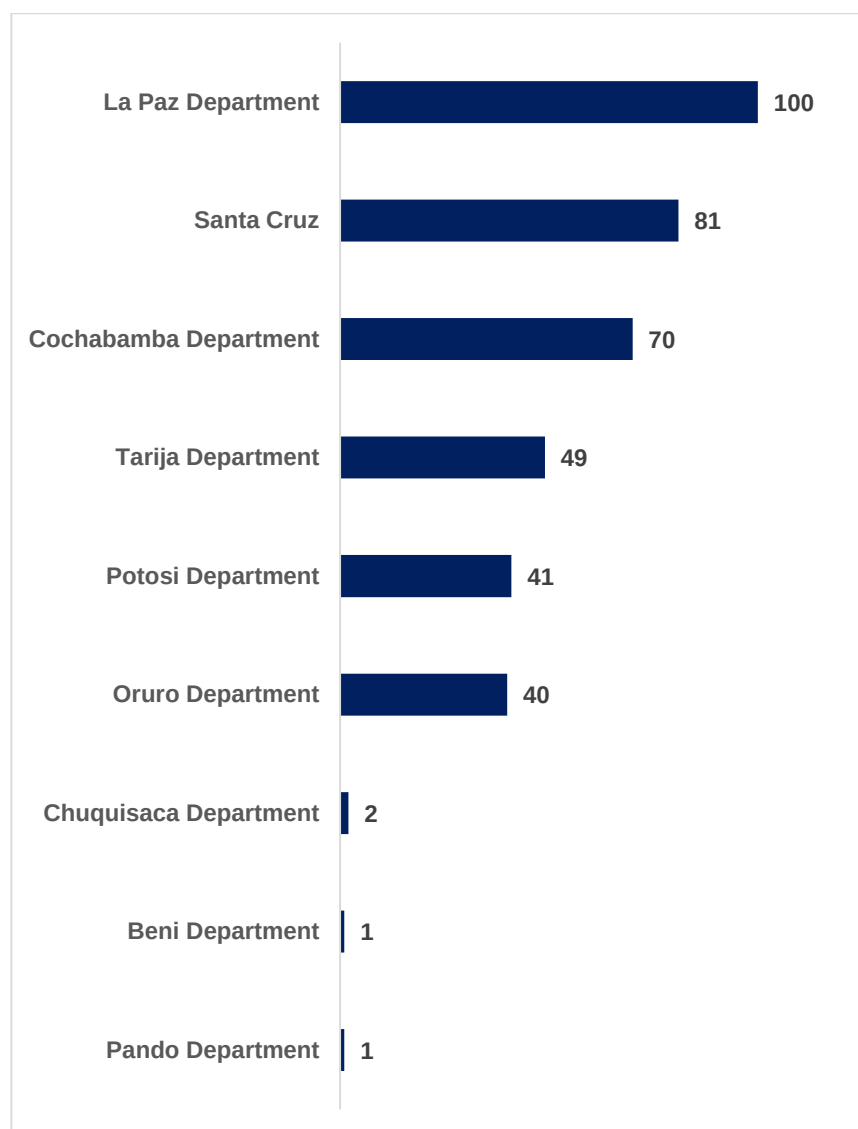


Figura 4: Interés hacia el BCP en territorio boliviano 2014-2018 de acuerdo a Google

Fuente: Google Trends (2019)

Nota: El interés se mide a través de búsquedas. Es un indicador relativo: en aquella zona donde más búsquedas acerca del BCP se han realizado registra un valor de 100, mientras que en aquellas zonas donde las búsquedas han sido mínimas se registra un valor de 1.

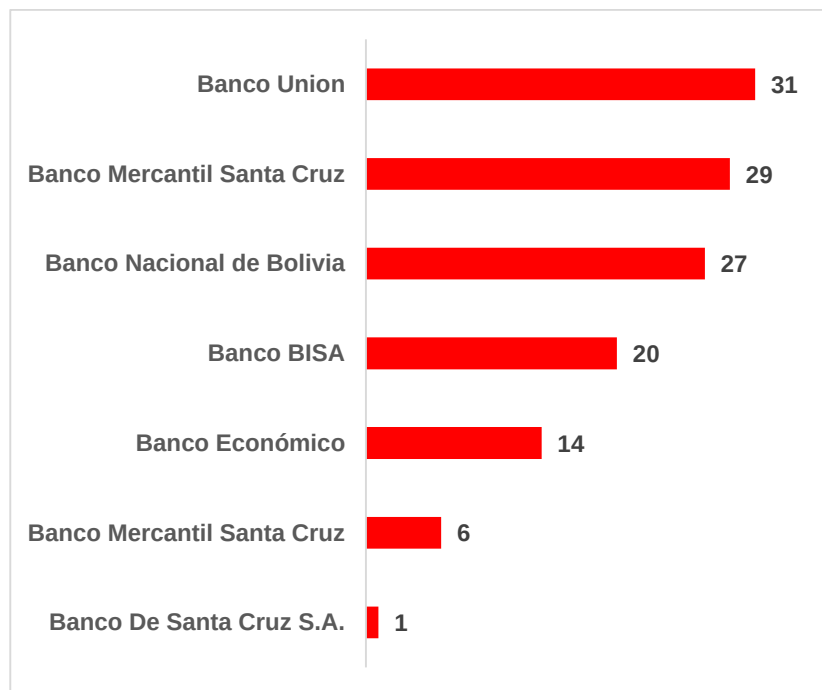


Figura 5: Bancos relacionados al BCP Bolivia de acuerdo a Google en territorio boliviano

Fuente: Google Trends (2019)

Nota1: El indicador es un índice (que va desde 0 a 100) mientras más relacionado este una entidad al BCP Bolivia, más alto será el valor.

Nota2: En la escala original, la entidad financiera que alcanza un valor de 100 es el BCP de Bolivia.

Por otro lado, si se trata de la situación financiera de la empresa, se cuenta con la siguiente información:

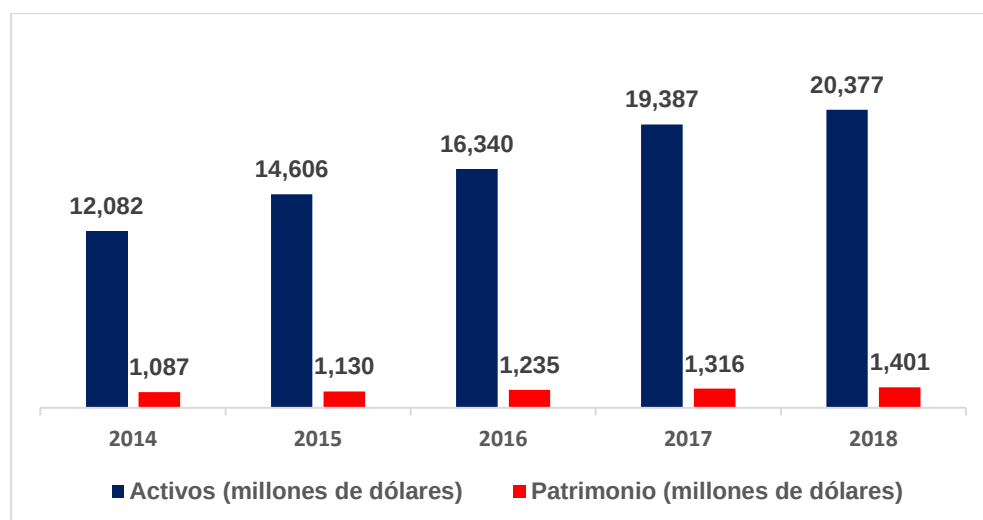


Figura 6: Activos y Patrimonio del BCP Bolivia 2014-2018 (en millones de dólares)

Fuente: Moodys (2019)

Moody's Corporation, agencia calificadoradora de riesgo, investiga el ámbito financiero internacional, de forma más precisa se centra en entidades estatales y comerciales (European Securities and Markets Authority, 2018). La empresa evalúa la solvencia de los deudores empleando una escala propia. La compañía tiene un *market share* de 40 % en el mercado de calificación crediticia global, similar al de su principal rival, Standard & Poor's (European Securities and Markets Authority, 2018). Se ubica dentro del top 3 del mundo en lo relacionado a calificación de crédito, junto a Fitch Rating y Standard & Poor's (European Securities and Markets Authority, 2018).

En cuanto a la rentabilidad del BCP Bolivia en el período 2014-2018 se cuenta con los siguientes datos:

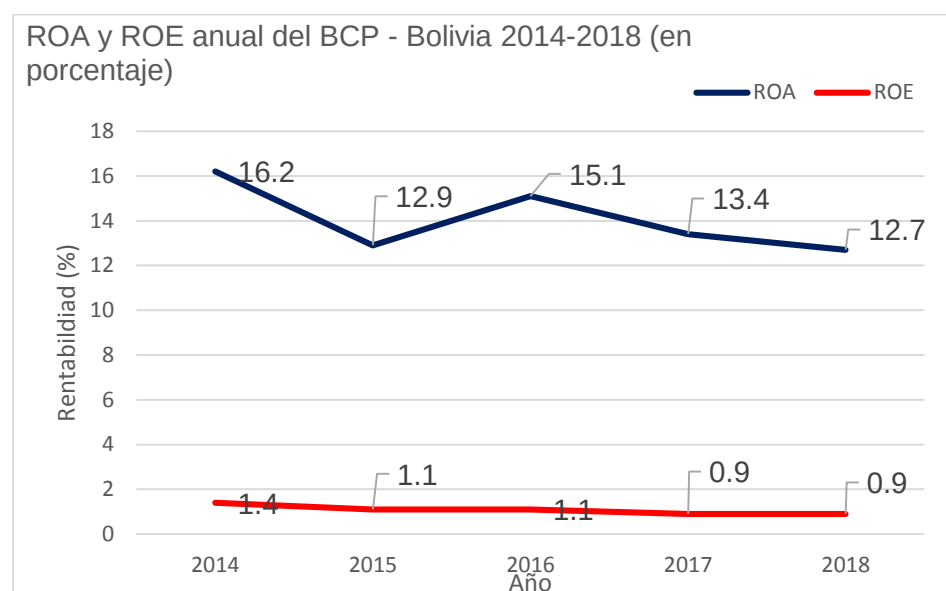


Figura 7: ROE y ROA obtenidos por el BCP Bolivia durante el periodo 2014-2018

Fuente: Moodys (2019)

En cuanto a la participación de mercado del Banco de Crédito, un indicador relevante es la cartera total de préstamos. Con respecto a este indicador se tiene:

INSTITUCIÓN	DIC-16			DIC-17			DIC-18			JUL-19		
	CARTERA	% DE PARTICIP.	RANKING	CARTERA	% DE PARTICIP.	RANKING	CARTERA	% DE PARTICIP.	RANKING	CARTERA	% DE PARTICIP.	RANKING
BANCOS MÚLTIPLES												
BNB - Banco Nacional	11.705.261	10,24%	2	13.283.257	10,14%	2	14.941.829	9,86%	2	16.525.940	10,23%	2
BUN - Banco Unión	10.700.681	9,36%	3	12.731.691	9,72%	3	14.766.444	9,74%	3	15.985.948	9,90%	3
BME - Banco Mercantil	13.865.198	12,13%	1	16.451.470	12,55%	1	22.254.481	14,69%	1	23.299.138	14,43%	1
BIS - Banco Bisa	9.447.518	8,27%	4	11.069.005	8,45%	5	13.113.585	8,65%	5	14.052.101	8,70%	5
BCR - Banco de Crédito	9.438.349	8,26%	5	11.155.537	8,51%	4	13.256.418	8,75%	4	14.395.830	8,92%	4
BGA - Banco Ganadero	6.051.074	5,30%	9	6.886.660	5,26%	9	7.845.483	5,18%	9	8.432.799	5,22%	9
BEC - Banco Económico	5.699.792	4,99%	11	6.653.115	5,08%	10	7.229.273	4,77%	10	7.772.276	4,81%	10
BSO - Banco Solidario	8.043.327	7,04%	7	8.950.797	6,83%	8	9.991.127	6,59%	8	10.553.756	6,54%	7
BNA - Banco de la Nación Argentina	113.560	0,10%	36	111.631	0,09%	43	133.820	0,09%	41	124.104	0,08%	42
BDB - Banco de Brasil	207.636	0,18%	24	288.108	0,22%	25	329.890	0,22%	24	430.274	0,27%	23
BIE - Banco para Iniciativas Empresariales	8.156.438	7,14%	6	9.139.602	6,97%	7	9.998.014	6,60%	7	10.529.126	6,52%	8
BFO - Banco Fortaleza	1.814.510	1,59%	14	2.205.525	1,68%	14	2.586.011	1,71%	13	2.719.868	1,68%	13
BFS - Banco Fassil	7.341.332	6,42%	8	9.515.076	7,26%	6	11.083.750	7,31%	6	11.323.395	7,01%	6
BPR - Banco Prodem	5.789.406	5,07%	10	6.465.186	4,93%	11	7.192.778	4,75%	11	7.564.262	4,68%	11
SUB-TOTAL	98.374.082	86,09%		114.906.661	87,69%		134.722.902	88,90%		143.708.816	89,00%	

Tabla 1: Ranking de Cartera – Principales Bancos de Bolivia (en miles de bolivianos)

Fuente: Autoridad de Supervisión del Sistema Financiero (2019)

Con la información de la tabla anterior, es posible afirmar que el BCP- Bolivia se encuentra entre los Top 5 del mercado boliviano, al menos desde el 2015 y hasta julio del 2018. Otros competidores importantes del sector son el Banco Nacional, Banco Unión, Banco Mercantil y Banco Bisa. Por otro lado, en cuanto al ranking de depósitos se tiene:

Tabla 2: Ranking de Depósitos del público y empresas con participación estatal - Principales Bancos de Bolivia (en miles de bolivianos)

INSTITUCIÓN	DIC-16			DIC-17			DIC-18			JUL-19		
	DEPÓSITOS	% DE PARTICIP.	RANKING	DEPÓSITOS	% DE PARTICIP.	RANKING	DEPÓSITOS	% DE PARTICIP.	RANKING	DEPÓSITOS	% DE PARTICIP.	RANKING
BANCOS MÚLTIPLES												
BNB - Banco Nacional	16.382.422	11,43%	3	17.728.050	12,02%	2	20.058.770	12,34%	2	20.657.087	12,51%	2
BUN - Banco Unión	23.101.601	16,12%	1	16.618.226	11,27%	3	18.946.677	11,65%	3	18.259.738	11,05%	3
BME - Banco Mercantil	21.496.230	15,00%	2	23.312.321	15,81%	1	28.623.183	17,61%	1	28.908.699	17,50%	1
BIS - Banco Bisa	12.665.357	8,84%	4	13.794.722	9,35%	4	15.912.767	9,79%	4	15.359.251	9,30%	5
BCR - Banco de Crédito	12.106.903	8,45%	5	12.324.795	8,36%	5	15.226.029	9,37%	5	15.482.019	9,37%	4
BGA - Banco Ganadero	7.047.770	4,92%	8	7.393.961	5,01%	8	8.429.404	5,18%	7	8.612.467	5,21%	7
BEC - Banco Económico	6.181.830	4,31%	10	7.115.733	4,82%	9	7.742.229	4,76%	9	8.079.067	4,89%	9
BSO - Banco Solidario	6.457.552	4,51%	9	6.985.484	4,74%	10	7.243.213	4,46%	10	7.757.641	4,70%	10
BNA - Banco de la Nación Argentina	75.047	0,05%	44	64.069	0,04%	46	63.412	0,04%	46	66.825	0,04%	46
BDB - Banco de Brasil	219.819	0,15%	23	124.544	0,08%	34	92.680	0,06%	41	91.965	0,06%	42
BIE - Banco para Iniciativas Empresariales	7.255.325	5,06%	7	8.190.860	5,55%	7	8.264.119	5,08%	8	8.565.721	5,19%	8
BFO - Banco Fortaleza	1.944.224	1,36%	14	2.127.895	1,44%	14	2.406.458	1,48%	13	2.570.896	1,56%	13
BFS - Banco Fassil	7.684.756	5,36%	6	10.323.671	7,00%	6	11.019.530	6,78%	6	11.598.692	7,02%	6
BPR - Banco Prodem	5.559.281	3,88%	11	6.124.337	4,15%	11	6.370.720	3,92%	11	6.748.649	4,09%	11
SUB-TOTAL	128.178.116	89,45%		132.228.668	89,65%		150.399.190	92,51%		152.758.717	92,48%	

Fuente: Autoridad de Supervisión del Sistema Financiero (2019)

Con la información de la tabla anterior, es posible afirmar que el BCP- Bolivia se encuentra entre los Top 5 del mercado boliviano en cuanto a depósitos. En este caso la participación porcentual en cuanto a depósitos es superior a las carteras de préstamos (9.37% vs. 8.92%). Los competidores principales son el Banco Nacional, Banco Unión, Banco Mercantil y Banco Bisa.

1.2.2.10. Variables Social Media a estudiar

Dados los puntos tratados en el marco teórico, serán consideradas las siguientes variables.

- Puntaje Sentimental de los comentarios en el canal del BCP-Bolivia en YouTube: Es una variable que refleja la actitud hacia el contenido publicado en el canal.
- Número de comentarios en el canal del BCP-Bolivia en YouTube: Es una variable asociado al nivel de interacción de los usuarios con las publicaciones en formato de video realizadas por la misma empresa.
- Número de menciones en diarios de Bolivia acerca del BCP.
Cantidad de noticias que involucran al BCP en los principales diarios de Bolivia (El diario, La Razón, Jornada, Cambio, Extra, Pagina 7, El Alto, La Palabra de Beni, Correo del Sur, Libertador, Los Tiempos, Opinión, La Voz, La Patria, El Potosí, El Deber, El Mundo, El día, La Estrella del Oriente, El País y El Norte.).
- Puntaje Sentimental de las noticias acerca del BCP Bolivia: Es una variable que refleja la actitud (desde negativa a positiva) reflejada en las noticias que involucran al BCP.
- Índice de Búsquedas en Google acerca del BCP Bolivia: Es el índice que refleja la cantidad de búsquedas que se han producido acerca del BCP Bolivia en territorio boliviano a través de Google.
- Índice de Búsqueda en YouTube acerca del BCP Bolivia: Es el índice que refleja la cantidad de búsquedas que se han producido acerca del BCP Bolivia en territorio boliviano a través de YouTube.

1.3.Objetivos e Hipótesis

1.3.1. Objetivos

- Objetivo General: Analizar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con las variables del tipo Social Media en el período 2013-2018.
- Objetivo Específico 1: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el puntaje

sentimental de los comentarios en YouTube realizados en el canal de la empresa durante el período 2013-2018.

- Objetivo Específico 2: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el número comentarios en YouTube realizados en el canal de la empresa durante el período 2013-2018.
- Objetivo Específico 3: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el número menciones de la entidad en los principales diarios online de Bolivia en el período 2013-2018.
- Objetivo Específico 4: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el puntaje sentimental acerca de la entidad en los principales diarios online de Bolivia en el período 2013-2018.
- Objetivo Específico 5: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el índice de búsquedas en Google acerca de la entidad en territorio boliviano en el período 2013-2018.
- Objetivo Específico 6: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el índice de búsquedas en YouTube acerca de la entidad en territorio boliviano en el período 2013-2018.

1.3.2. Hipótesis

- Hipótesis General: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con las variables del tipo social media en el período 2013-2018.
- H1: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el puntaje sentimental de los comentarios en YouTube realizados en el canal de la empresa durante el período 2013-2018.
- H2: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el número comentarios en YouTube realizados en el canal de la empresa durante el período 2013-2018.

- H3: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el número menciones de la entidad en los principales diarios online de Bolivia en el período 2013-2018.
- H4: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el puntaje sentimental acerca de la entidad en los principales diarios online de Bolivia en el período 2013-2018.
- H5: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el índice de búsquedas en Google acerca de la entidad en territorio boliviano en el período 2013-2018.
- H6: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el índice de búsquedas en YouTube acerca de la entidad en territorio boliviano en el período 2013-2018.

Capítulo II – Método

2.1. Tipo y Diseño de Investigación

2.1.1. Tipo de investigación

Los objetivos de la investigación pueden alcanzarse recolectando datos y recurriendo al análisis estadístico matemático (Hernández, Fernández, & Baptista, 2014). Este último paso permitirá conocer el valor de verdad de las hipótesis. De acuerdo a Hernandez & Mendoza (2018, estos son rasgos de una investigación cuantitativa. Es necesario destacar que parte de las variables del tipo Social Media son obtenidas a partir de datos en formato de texto libre; no obstante, resulta relativamente sencillo convertir dicho formato a índices numéricos que pueden ser aprovechados para un posterior tratamiento estadístico (Hofmann & Chisholm, 2015).

2.1.2. Diseño de investigación

La investigación que pretende determinar si existe relación entre la reputación del BCP – Bolivia y un conjunto de variables del tipo Social Meda tiene tres características: no experimental, correlacional y transversal.

Es no experimental, debido a que el investigador no pretende ejercer control sobre las variables (Hernandez & Mendoza, 2018). Ello aplica tanto para la variable de reputación corporativa como para los elementos del Social Media. Es correcto afirmar que en el caso de algunos elementos del Social Media son procesados con el análisis de sentimiento, pero se trata únicamente de una conversión de datos textuales a numéricos.

Es correlacional, debido a que el investigador busca determinar si la variación entre los componentes vinculados a la Social Media guarda relación con la reputación corporativa. Esta es una característica de las investigaciones correlacionales (Hernandez & Mendoza, 2018).

Debe mencionarse que esta relación es transversal debido a que el tiempo no es una variable de análisis (Hernandez & Mendoza, 2018). El tiempo únicamente fue empleado para organizar las variables en forma anual. Por otro lado, todos los datos son recolectados en un momento específico. Esta es una característica de las investigaciones transversales (Hernandez & Mendoza, 2018).

2.2. Variables de Estudio

Todas las variables de estudio provienen de fuente secundaria Merco y Social Media. Estas dos últimas no se obtienen directamente y requieren de un proceso de extracción y transformación de datos.

Tabla 3: Variables de Estudio

Variables	Descripción	Operacionalización de Variables
Reputación Corporativa	Es una variable de origen secundario proporcionada por Merco. Es construida en base a una combinación de metodologías cuantitativas y cualitativas. El rango de esta variable va desde 0 hasta 10,000. Es publicada anualmente y es la única variable dependiente.	Operacionalmente se obtiene del índice Merco publicado en http://www.merco.info/bo/ . Es decir, es obtenido mediante información directa.
Puntaje Sentimental de los comentarios en el canal del BCP-Bolivia en YouTube	Es una variable de origen secundario que surge a partir del procesamiento de elementos textuales en la social en los comentarios a los videos del canal del BCP-Bolivia en YouTube). Este indicador fue propuesto por Grutzmacher (2011) y correlacionado con la reputación por Luca (2011) En la investigación será estimada de forma anual. El puntaje sentimental va desde -1 hasta 1. Es una variable independiente que mide el sentimiento presente en los comentarios de YouTube para un año determinado.	Los comentarios son procesados con el instrumento Google Cloud Natural Language (previamente extraídos y depurados con NVIVO). Dichos comentarios han sido anualizados, es decir, organizados en forma anual.
Número de comentarios en el canal del BCP-Bolivia en YouTube	Tal como su nombre lo indica, es el número de comentarios que reciben los videos del BCP Bolivia (en el canal de YouTube de la entidad). Es una variable independiente cuyo valor pertenece a los números naturales. Es una variable anualizada.	Basta con organizar los comentarios de respuestas realizados por usuarios a los videos del BCP Bolivia para luego organizarlos en forma anual.
Número de menciones en diarios de Bolivia acerca del BCP.	Cantidad de noticias que involucran al BCP en los principales diarios de Bolivia (El diario, La Razón, Jornada, Cambio, Extra, Pagina 7, El Alto, La Palabra de Beni, Correo del Sur, Libertador, Los Tiempos, Opinión, La Voz, La Patria, El Potosí, El Deber, El Mundo, El día, La Estrella del Oriente, El País y El Norte.) por periodo anual.	El conteo se realiza desde la página http://www.prensaescrita.com/america/bolivia.php con apoyo del buscador del sitio web.
Puntaje Sentimental de las noticias acerca del BCP Bolivia.	Es una variable de origen secundario que surge a partir del procesamiento de elementos textuales que involucran al BCP de Bolivia. Resume el sentimiento hacia el banco, presente en los comentarios de YouTube Admite valores desde -1 a 1.	Los comentarios son procesados con el instrumento Google Cloud Natural Language (previamente extraídos y depurados con NVIVO). Dichos comentarios han sido anualizados, es decir, organizados en forma anual.
Índice de Búsquedas en Google acerca del BCP Bolivia	Es el índice total anualizado que indica la cantidad de búsquedas que se han producido acerca del BCP Bolivia en territorio boliviano a través de Google.	Obtenido con Google Trends explorando el término Banco de Crédito de Bolivia.
Índice de Búsqueda en YouTube acerca del BCP Bolivia	Es el índice total anualizado que indica la cantidad de búsquedas que se han producido acerca del BCP Bolivia en territorio boliviano a través de YouTube.	Obtenido con Google Trends explorando el término Banco de Crédito de Bolivia.

Elaboración: Propia

2.3. Universo o Población

Los indicadores de la reputación corporativa son publicados por Merco. En el caso de los índices de búsqueda en Google y YouTube acerca del BCP Bolivia, ambos son proporcionados por Google Trends. El número de comentarios en YouTube y la cantidad de menciones en noticias acerca del BCP surgen a partir de un conteo simple en aplicaciones electrónicas. Finalmente, los puntajes sentimentales se obtienen analizando todos los comentarios online del BCP-Bolivia en YouTube y el contenido de las noticias que involucraron a la entidad financiera. Cabe mencionar que la pertinencia de considerar estos datos como poblacionales es un aporte de los investigadores (Mantyla, Graziotin & Kuuttila, 2016). Son un total de seis series de datos anaulizados. La data corresponde a datos anuales desde el 2013 hasta el 2018.

2.4. Instrumentos

2.4.1. NVIVO

Es el instrumento que sirve para extraer los datos de YouTube. Basta con ingresar a la cuenta del BCP de Bolivia en YouTube y copiar los enlaces de los videos. A continuación, utilizando la opción Ncapture de NVIVO exporta los comentarios a una base de Excel. El NVivo es un software que fue diseñado originalmente para el análisis cualitativo, conforme han ido apareciendo nuevas versiones, los desarrollares han ido incorporando funcionalidades de *Data Mining* (Hamed & Said, 2013).

2.4.2. TextCleanr

Es un software especializado en la depuración de contenido textual (Hofmann & Chisholm, 2015). Su uso es adecuado para datos textuales, ya que estos no son del todo homogéneos. El instrumento es de libre disponibilidad y puede ser empleado por cualquier persona que cuenta con una computadora con acceso a internet en la siguiente dirección <https://www.textcleanr.com/> . En la investigación ha sido empleado para depurar los comentarios de YouTube realizados en el canal del BCP Bolivia.

2.4.3. Language Tool

Es un software especializado en correcciones gramaticales, ortográficas y de sintaxis (Hofmann & Chisholm, 2015). Precisamente, dichas funciones han sido aplicadas sobre el contenido textual en español con el fin de poderlo procesar con

una mayor precisión (Hofmann & Chisholm, 2015). La herramienta es de libre disponibilidad y se encuentra disponible en <https://languagetool.org/es/>.

2.4.3. Watson Translate

Es un software de traducción diseñado por IBM. Pertenece a la línea de productos *Watson*, basados en inteligencia artificial (IBM, 2019). Cuenta con una versión gratuita disponible en <https://www.ibm.com/watson/services/language-translator/> (IBM, 2019). En la investigación ha sido empleado para traducir el texto de español a inglés.

2.4.4. Grammarly

Es un software orientado a la corrección ortográfica, gramatical y de sintaxis para textos en idioma inglés (Hofmann & Chisholm, 2015). Funciona como un complemento descargable para *Google Chrome* y puede emplearse en hojas HTML (el espacio para redactar un correo de Gmail, por ejemplo). Funciona como el corrector de un procesador de texto con la ventaja de que cuenta con un diccionario de mayor amplitud.

2.4.5. Google Cloud Natural Language

Es una herramienta de libre diseñada por Google a la que se puede acceder gratuitamente en la versión demo (Hofmann & Chisholm, 2015). El instrumento permite ingresar datos textuales mediante el copiado y pegado (como en un procesador de textos). Luego es necesario pulsar el botón analizar para obtener *outputs*, lo cuales son inmediatos. El instrumento es empleado para el análisis de sentimiento. Este proceso es el que permite convertir los comentarios en YouTube y el contenido noticioso en números (Hofmann & Chisholm, 2015).

2.4.6. Excel

Este instrumento es una hoja de cálculo diseñado por Microsoft (s.f.). Es utilizada como almacén para los datos y para trasladar la data textual y numérica entre los distintos instrumentos. Los formatos más recientes de Excel (con la terminación *xls*.) tienen la ventaja de ser fácilmente importados por la mayoría de software de *Data Mining* (Hofmann & Chisholm, 2015).

2.4.7. SPSS

Este instrumento es un software estadístico cuya licencia pertenece a IBM (IBM, 2019) . Es utilizado para el procesamiento estadístico de los datos. Este proceso permitirá alcanzar los objetivos de la investigación y comprobar el valor de verdad de las hipótesis.

2.4.8. Motor de metabúsqueda acerca de noticias acerca de Bolivia

El sitio web Prensa Escrita Bolivia cuenta con una opción de redireccionamiento y metabúsqueda de noticias en donde pueden rastrearse el contenido noticioso acerca de un tópico específico en los principales diarios online de Bolivia. De acuerdo a Gatos & Nouraros (1998) resulta académicamente valido emplear sitios web de metabúsqueda para el estudio de contenido noticioso. Estos son: El diario, La Razón, Jornada, Cambio, Extra, Pagina 7, El Alteño, La Palabra de Beni, Correo del Sur, Libertador, Los Tiempos, Opinión, La Voz, La Patria, El Potosí, El Deber, El Mundo, El día, La Estrella del Oriente, El País y El Norte.

2.5. Procedimiento para recolección de datos

Los datos de recuperación corporativa son extraídos de Merco mediante observación directa. En el caso de los comentarios en YouTube, basta con el conteo con el instrumento NVIVO. La cantidad de noticias se obtiene mediante conteo simple y anotación.

En el caso de los indicadores vinculados a las búsquedas en Google y YouTube solo se requiere el uso de Google Trends. Los resultados son mensuales y pueden anualizarse mediante operaciones básicas. Los resultados son registrados en una hoja de cálculo.

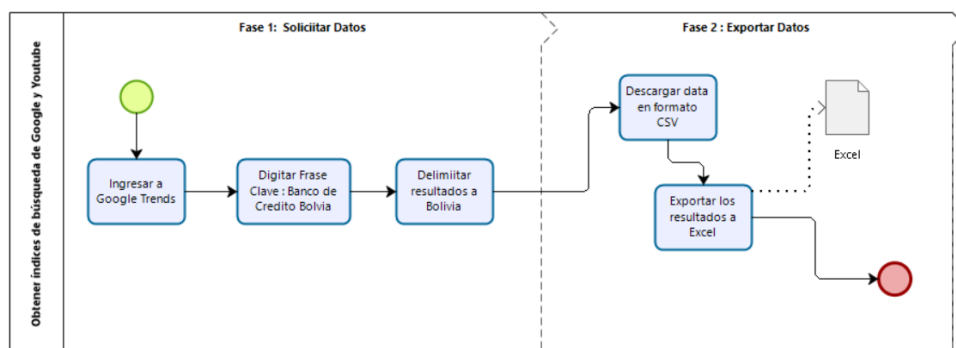


Figura 8: Proceso de extracción de datos de Google y YouTube

Elaboración Propia

La conversión de textos a patrones numéricos sí resulta un proceso más complejo que requiere diferentes etapas. Es necesario extraer los datos desde YouTube con el NVivo, depurarlos (TextCleanr), traducirlos (Watson) y corregirlos (Grammarly y Language Tool). Todo ello para procesarlos con el Google Cloud Natural Language para su posterior registro en Excel. Los datos vinculados a las búsquedas en Google y YouTube se encuentran en la siguiente tabla:

Tabla 4: Índices de Búsqueda en Google y YouTube acerca del BCP -Bolivia en territorio boliviano

Año	2013	2014	2015	2016	2017	2018
Índices Google	952	435	399	499	640	578
Índices YouTube	185	0	0	0	183	71

Fuente: Google Trends (2019)

Por otro lado, el tratamiento de datos textuales cuenta con su propio flujograma, tal como se muestra a continuación:

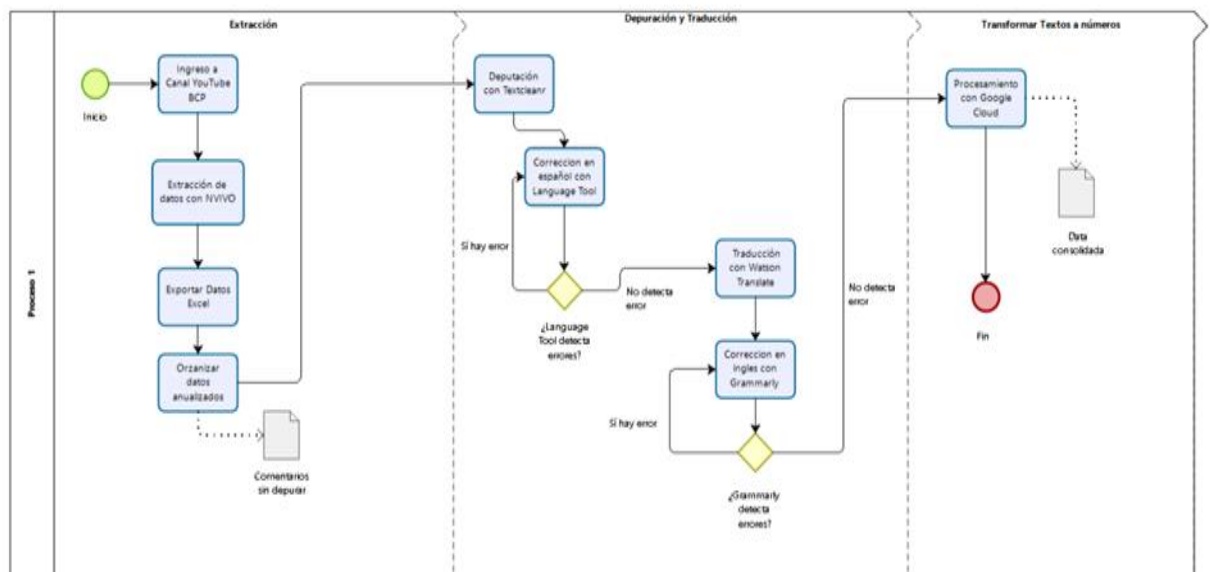


Figura 9: Proceso de Text Mining (o Data Mining) en el caso de YouTube

Elaboración Propia

En el caso del procesamiento de noticias, el proceso es similar al anterior, solo que en lugar de extraer datos de YouTube solo se necesita emplear el sitio de búsqueda de diarios bolivianos. Todos los textos son organizados por año y procesados con Google Cloud Natural Language.

Tabla 5: Variables consolidadas

Año	Reputación Corporativa	Sentimiento Noticias Bolivia	Sentimiento Youtube	Comentarios YouTube	Noticias Medios Bolivia	Indice Google	Indice YouTube
2013	6608	0.4	0.5	120	19	952	185
2014	6682	0.5	0.6	145	23	435	0
2015	6980	0.6	0.6	208	28	399	0
2016	7308	0.7	0.7	306	34	499	0
2017	6444	0.6	0.6	425	42	640	183
2018	8101	0.8	0.9	289	38	578	71

Fuentes: Merco (2019), Google Trends (2019), Prensa Escrita Boliviana (2019)

Elaboración: Propia

2.6. Plan de análisis

2.6.1. Correlación entre variables

En primer lugar, es posible recurrir a pruebas de correlación entre la variable dependiente (reputación) y cada una de las variables independientes (elementos del Social Media). Previamente debe analizarse si las variables siguen o no una distribución normal. Para ello debe aplicarse las pruebas de normalidad. Dependiendo de ello, debe utilizarse una prueba de correlación específica: Pearson o Spearman (Restrepo & González, 2007). Lo adecuado sería emplear una regresión múltiple, pero esta requiere al menos 20 series de datos para que se genere un modelo aceptable (Restrepo & González, 2007).

2.6.2. Red neuronal

La Red Neuronal es un modelo computacional que se componen por un conjunto de elementos simples e interconectados que procesan información (Pérez & De Paz, 2007). De acuerdo a Pérez & Martín (2003), las redes neuronales están inspiradas en el sistema de neuronas biológicas propio del cerebro humano. Las redes neuronales pueden predecir una variable cuantitativa (*output*) en función a una serie de variables de entrada (*inputs*) (Pérez & De Paz, 2007). El perceptrón simple (neurona) de McCulloch-Pitts (1943), desarrolladores de este concepto matemático, está estructurada como una suma ponderada de *inputs*, seguido por una función no lineal conocida como función de activación (Pérez & De Paz, 2007). Entre las funciones de activación que aparecen con mayor frecuencia están la lineal, función logística y la sigmoide (Muñoz, 2015).

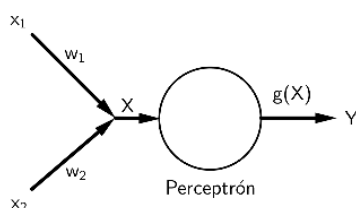


Figura 10: Diagrama de un Perceptrón

Fuente: Díaz (2015)

Nota: Las variables de entrada son X_1 y X_2 (*inputs*), W_1 y W_2 representan los pesos. Y es el resultado del proceso (*outputs*)

Según Pérez & Martín (2003), a medida que se ha ido avanzando en los métodos estadísticos, han sido desarrolladas las herramientas necesarias para el modelado las relaciones que se caracterizan por la no linealidad y, por tanto, resultan complejas. Las redes neuronales artificiales (ANNs por sus siglas en inglés, *Artificial Neural Networks*) son programas de computadora que simulan algunas de las funciones del cerebro humano que pueden ser calificadas como superiores (Pérez & De Paz, 2007). Al igual que en el cerebro, hay neuronas y sinapsis, con varias conexiones sinápticas denominadas 'pesos' (Pérez & Martín, 2003). No obstante, a diferencia del cerebro humano y de forma similar a otros programas de computadoras, existe una neurona que representa la salida dado un conjunto de neuronas de entrada para cada problema específico y de acuerdo a una arquitectura en particular. (Dean, 2014).

Según Pérez & Martín (2003) y Dean (2014) esta técnica fue desarrolladas en los años 50. Si ha recobrado presencia en los últimos años se debe a que ahora los computadores actuales cuentan con *hardware* de mayor potencia (Dean, 2014) Las redes neuronales tienen por objetivo el obtener combinaciones de los *inputs* con características derivadas para luego modelar el destino (*output*) como una función no lineal; así se obtiene un método de aprendizaje aplicable en una diversidad de ámbitos (Pérez & De Paz, 2007). Esta metodología puede considerarse como un modelo de regresión de dos etapas (Muñoz, 2015), tal como evidencia la siguiente figura:

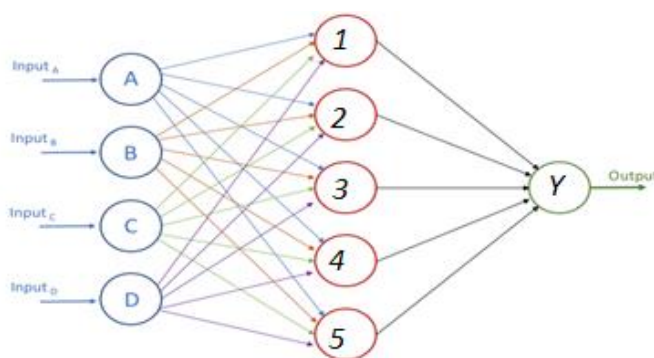


Figura 11: Esquema de una red neuronal con una sola capa oculta

Fuente: Adaptado de Muñoz (2015) citando de Dean (2014)

Nota: En la red neuronal anterior existen 4 elementos de entrada (A,B,C,D) que forman 5 unidades intermedias (perceptrones, círculos rojos del 1 al 5). La combinación de estos últimos origina el elemento de salida (output, círculo verde, Y)

Muñoz (2015), citando a MacKay (2003), afirma que cuando algoritmo de red neuronal es descrito deben ser analizados tres componentes: arquitectura, regla de actividad y regla de aprendizaje. La red neuronal comienza con un grupo aleatorio de pesos y los va modificando en base a una función del error que se calcula con el diferencial entre las observaciones y los valores esperados (Pérez & De Paz, 2007). Este tipo de método se establece como aprendizaje supervisado, porque ajusta los pesos de la red inspeccionando las respuestas deseadas que son conocidas de antemano, de esta forma una red neuronal es entrenada (Pérez & Martín, 2003). La ventaja de este tipo de métodos es su alta flexibilidad en comparación a otros algoritmos como la regresión lineal múltiple, donde es necesario que se cumplan ciertas condiciones (Pérez & Martín, 2003).

Debe tomarse en consideración, que, si bien es cierto que la data es pequeña, puede emplearse la red neuronal, debido a que el algoritmo lo soporta, aunque desde una perspectiva teórica no es lo idóneo (Pérez & Martín, 2003). Sin embargo, puede realizarse a modo de ejercicio que compruebe si la combinación de elementos vinculados al *Social Media* puede aproximarse a la Reputación Corporativa. Este es un ejercicio aplicable en muestras pequeñas (Pérez & Martín, 2003).

2.6.3. Regresión Lineal

A fin de alcanzar una mejor comprensión de los resultados se ha considerado el uso de la regresión lineal. En este caso se pretenden evidenciar variable que la variable dependiente (Reputación Corporativa) puede ser expresada como una combinación lineal de las variables independientes.

De acuerdo a investigadores como Montgomery Peck & Vining (2012), se requiere que al menos 20 series de datos para que la regresión lineal para evitar resultados con sesgos. No obstante, en los libros introductorios de Estadística se suelen realizar ejercicios con pocas series a fin de facilitar una comprensión inicial de los distintos modelos (Besley, 1991).

La regresión lineal, a diferencia de la red neuronal, debe cumplir con un conjunto de condiciones. Entre ellas cabe destacar que las variables Social Media que incluya el modelo no deben tener una relación de dependencia. Ello se verifica

con VIF, Valor Inflacionario de Varianza, el cual no debe exceder a 2 (Besley, 1991). Por otro lado, los residuos (diferencia entre los valores observados y predichos de la Reputación corporativa) deben seguir una distribución normal. Ello se verifica con la prueba de Shapiro-Wilk (Besley, 1991). Finalmente, no debe darse la autocorrelación entre los residuos generados por el modelo. Esta condición se verifica con el estadístico de Durbin-Watson (Besley, 1991). Cabe mencionar que se ha empleado la regresión lineal en su versión por pasos (*step forward*), ello implica la generación de un modelo que únicamente incluye a las variables que aportan significativamente a conocer la variable Reputación Corporativa.

2.6.4. Comparación entre modelos

Independientemente de las condiciones de validez antes mencionadas para la regresión, se comparan ambos modelos en cuanto a su precisión y variables involucradas.

Capítulo III- Resultados

3.1. Presentación de Resultados

El primer paso consiste en aplicar las pruebas de normalidad para los datos

Tabla 6: Pruebas de normalidad

	Kolmogórov-Smirnov			Shapiro-Wilk		
	Estadístico	Gl.	Sig.	Estadístico	Gl	Sig.
Reputación Corporativa	,210	6	,200*	,888	6	,308
Sentimiento Noticias Bolivia	,167	6	,200*	,982	6	,960
Sentimiento YouTube	,308	6	,077	,857	6	,178
Comentarios YouTube	,152	6	,200*	,950	6	,742
Noticias Medios Bolivia	,146	6	,200*	,965	6	,855
Índice Google	,223	6	,200*	,873	6	,237
Índice YouTube	,291	6	,121	,767	6	,059

Dado que el número de elementos de la serie de datos es menor a 20, entonces debe tomarse como referencia la prueba el indicador de Shapiro Wilk (Razali & Wah, 2011). En este caso los niveles de significancia superan al valor de 0.05. Por tanto, los valores siguen una distribución normal (Razali & Wah, 2011).

En caso los niveles de significancia no superaran el valor de 0.05, entonces era posible afirmar que los datos no seguían una distribución normal (Razali & Wah, 2011).

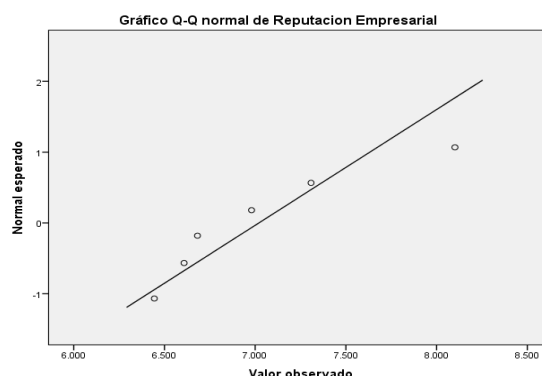


Figura 12: Gráfico de la distribución normal para la variable Reputación Corporativa

Nota: El hecho de que los valores estén más cerca de la recta favorece a que la variable tenga una distribución normal

El test de normalidad se aplica para determinar si las variables siguen una distribución normal (Razali & Wah, 2011). A partir de ellos se puede determinar si es mejor aplicar la correlación de Pearson o Spearman (Razali & Wah, 2011)

Debido a que las variables siguen una distribución normal, lo más adecuado, entonces, sería emplear la correlación de Pearson para determinar si existe una relación lineal: Una vez correlacionadas las variables se obtuvo:

Tabla 7: Matriz de Correlaciones de Variables

		Reputación Corporativa	Sentimiento Noticias Bolivia	Sentimiento YouTube	Comentarios YouTube	Noticias Medios Bolivia	Índice Google	Índice YouTube
Reputación Corporativa	Correlación de Pearson	1	,836*	,935**	,138	,363	-,255	-,381
	Sig.		,038	,006	,794	,480	,625	,456
	N	6	6	6	6	6	6	6
Sentimiento Noticias Bolivia	Correlación de Pearson	,836*	1	,923**	,619	,779	-,481	-,358
	Sig.	,038		,009	,190	,068	,334	,486
	N	6	6	6	6	6	6	6
Sentimiento YouTube	Correlación de Pearson	,935**	,923**	1	,390	,604	-,339	-,308
	Sig.	,006	,009		,444	,204	,510	,552
	N	6	6	6	6	6	6	6
Comentarios YouTube	Correlación de Pearson	,138	,619	,390	1	,965**	-,171	,219
	Sig.	,794	,190	,444		,002	,746	,677
	N	6	6	6	6	6	6	6
Noticias Medios Bolivia	Correlación de Pearson	,363	,779	,604	,965**	1	-,263	,109
	Sig.	,480	,068	,204	,002		,615	,837
	N	6	6	6	6	6	6	6
Índice Google	Correlación de Pearson	-,255	-,481	-,339	-,171	-,263	1	,860*
	Sig.	,625	,334	,510	,746	,615		,028
	N	6	6	6	6	6	6	6
Índice YouTube	Correlación de Pearson	-,381	-,358	-,308	,219	,109	,860*	1
	Sig.	,456	,486	,552	,677	,837	,028	
	N	6	6	6	6	6	6	6

De la tabla anterior es posible inferir que la reputación corporativa tiene correlación con el sentimiento de los comentarios en YouTube ($r = .935$) y el sentimiento de las Noticias ($r = .836$). En ambos casos las correlaciones son significativas. La reputación corporativa por otro lado, no tiene relación significativa con la cantidad de comentarios en el canal de YouTube del BCP ($p\text{-valor} = .794$), ni las noticias acerca del BCP en la prensa boliviana ($p\text{-valor} = .480$). Por otro lado, la reputación corporativa tampoco tiene correlación con los índices de búsqueda en Google ($p\text{-valor} = .625$) y YouTube ($p\text{-valor} = .456$) acerca de la entidad boliviana. Esto permite cumplir los objetivos específicos de la investigación. También permite verificar cada una de las hipótesis específicas.

Por otro lado, cabe destacar el caso de YouTube. No hay correlación entre las variables. Pero eso sucede porque miden cosas distintas. El sentimiento presente en los comentarios es una variable que no depende de la cantidad de comentarios, pero si refleja la actitud de quienes comentan. Por otro lado, los comentarios en el canal de YouTube del BCP Bolivia no son los únicos comentarios que involucran a la entidad: los hay en muchos otros canales, pero no hay forma eficiente de extraerlos (como las búsquedas en Twitter). En este sentido los índices de búsqueda en YouTube son variables de nivel macro.

Con respecto a la red neuronal, los resultados son los siguientes:

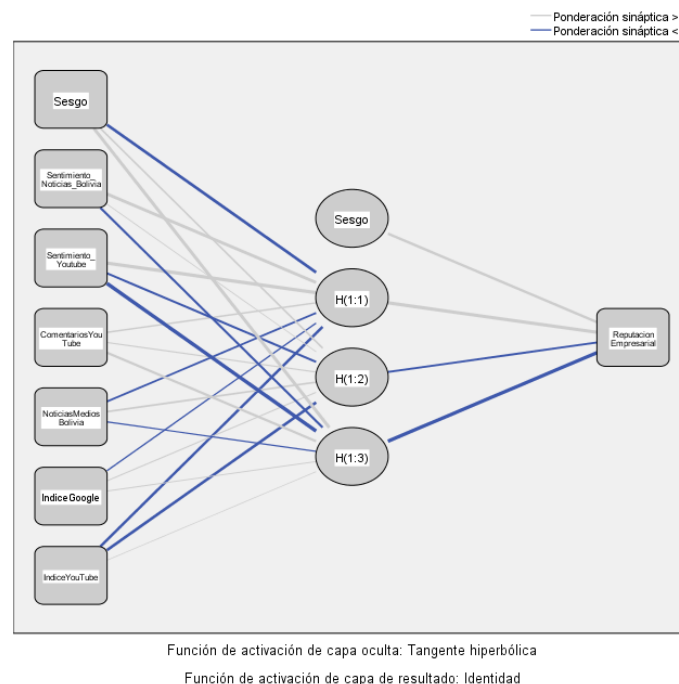


Figura 13: Arquitectura de la red neuronal

De la figura anterior es posible apreciar que se han formado 3 unidades intermedias: H (1:1), H (1:2), H (1:3), más el sesgo. Las mismas surgen a partir de diversas combinaciones de las variables de entrada ya mencionadas. La función de activación es Identidad (análoga a una regresión lineal).

En cuanto al resumen del modelo que pretende relacionar a la reputación corporativa del BCP con elementos del Social Media se obtuvo el siguiente resultado:

Tabla 8: Resumen del modelo

Partición	Descripción	
Entrenamiento	Error de suma de cuadrados	0,067
	Error relativo	0,043
	Regla de parada utilizada	Pasos consecutivos sin disminución del error
	Tiempo de preparación	00:00:00
Pruebas	Error de la suma de cuadrados	1,234
	Error relativo	0,331

Lo anterior se complementa con las estimaciones de cada parámetro para los componentes de la capa oculta:

Tabla 9: Parámetros de la red neuronal y Precisión del modelo

Predictor		Pronosticado			Reputación Corporativa
		Capa oculta			
		H(1:1)	H(1:2)	H(1:3)	
Capa de entrada	(Sesgo)	-,528	,247	,562	
	Sentimiento_Noticias_Bolivia	,562	,008	-,324	
	Sentimiento_Youtube	,690	-,307	-,725	
	Comentarios YouTube	,227	,111	,338	
	Noticias Medios Bolivia	-,269	,262	-,172	
	Índice Google	-,222	,092	,031	
	Índice YouTube	-,436	-,496	,001	
Capa oculta	(Sesgo)				,393
	H(1:1)				,595
	H(1:2)				-,277
	H(1:3)				-,694
Precisión del Modelo					43.2%

La tabla anterior representa una forma simplificada de cómo las distintas variables de entrada (más el sesgo) forman las unidades de la capa oculta (perceptrones), y cómo estos a su vez aportan a la variable respuesta final: la reputación del BCP-Bolivia. Por otro lado, en cuanto a la precisión del modelo se alcanzó un valor de 43.2%, un índice que puede ser considerado bajo tomando en consideración que bajo metodologías como CRISP-DM un modelado resultad adecuado cuando alcanza una precisión de 60% (IBM Knowledge Center, s.f.).

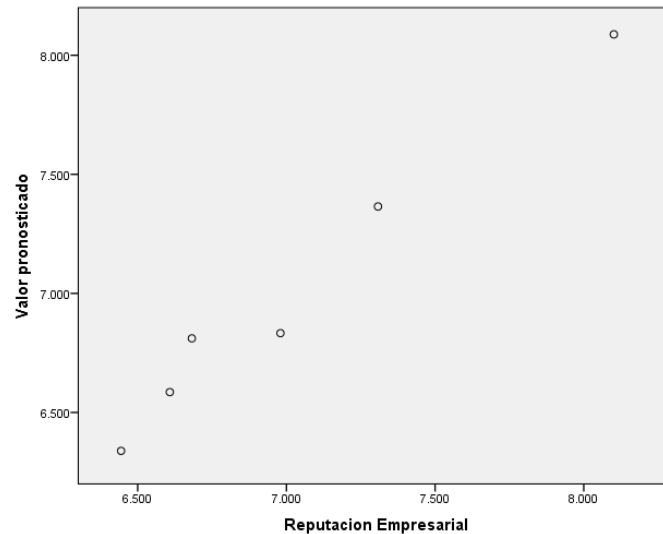


Figura 14: Valor predicho vs. Observado en la red neuronal

Naturalmente no todas las variables tienen la misma importancia. Son seis las variables predictoras (*inputs*)

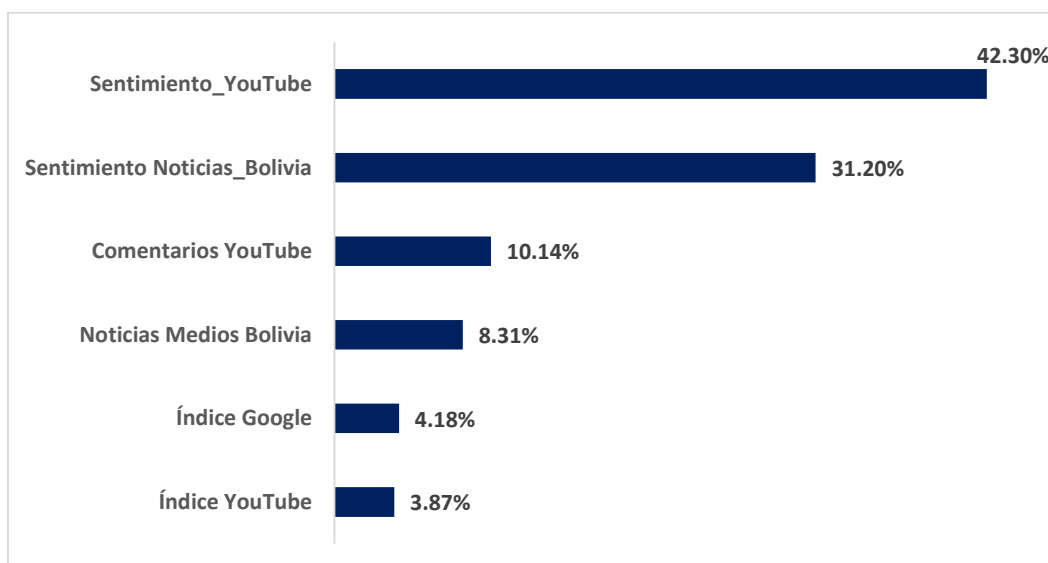


Figura 15: Importancia de las variables predictoras dentro del modelo

Nota: La importancia de las variables predictoras siempre suma 1 (Wendler & Gröttrup, 2016).

De la figura anterior queda evidencia, nuevamente, que las dos variables de mayor importancia son aquellas vinculadas a los sentimientos y emociones. Pero en este caso no se excluye los demás elementos predictores. Sucede que, bajo el modelo basado en redes neuronales, las otras cuatro variables independientes aportan a la estimación de la reputación corporativa. Ello la diferencia de las correlaciones, donde solo hay relación entre los puntajes sentimentales y la variable dependiente.

Con respecto a los resultados de la regresión lineal se obtuvo:

Tabla 10: Modelo de Regresión

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
	B	Error estándar	Beta		
1 (Constante)	4323.684	519.050		8.330	.001
Sentimiento_YouTube	4148.947	783.984	.935	5.292	.006

De la tabla anterior cabe mencionar que no se obtiene una regresión múltiple, sino una simple en donde la única variable dependiente es el sentimiento reflejado en los comentarios de YouTube. La ecuación tendría la siguiente forma:

$$\text{Reputación Corporativa} = 4324.684 + 4148.947 * \text{Sentimiento_YouTube}$$

El detalle de las variables retiradas del modelo (debido a que sus niveles de significancia superaron el .05 de significancia establecida) se muestra a continuación.

Tabla 11: Variables excluidas del modelo de regresión

Modelo	En beta	t	Sig.	Correlación parcial	Estadísticas de colinealidad
					Tolerancia
1 Sentimiento_Noticias_Bolivia	-.192 ^b	-.368	.737	-.208	.147
Comentarios YouTube	-.267 ^b	-1.679	.192	-.696	.848
Noticias Medios Bolivia	-.318 ^b	-1.778	.173	-.716	.635
Índice Google	.070 ^b	.329	.764	.187	.885
Índice YouTube	-.103 ^b	-.499	.652	-.277	.905

Las condiciones de colinealidad no han de ser consideradas para el modelo de regresión lineal simple debido a que solo hay una variable predictora.

Con respecto al resumen del modelo de regresión lineal simple se tiene:

Tabla 12: Resumen del modelo de Regresión Lineal y prueba de Durbin Watson

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Durbin-Watson
1	.935 ^a	.875	.844	241.6402	2.502

La precisión del modelo es 87.5%. Pero existe un sesgo que se evidencia con el estadístico de Durbin Watson. Idealmente debe ser próximo a 2 (en una escala que va de 1 hasta 4). Ello evidencia que existe autocorrelación entre los residuos.

Finalmente, ha sido necesario verificar otra condición vinculada a los residuos del modelo. Estos deben seguir una distribución normal para poder afirmar que la regresión lineal simple es un modelo adecuado (Besley, 1991).

Tabla 13: Pruebas de Normalidad para el modelo de regresión

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Unstandardized Residual	.246	6	.200*	.902	6	.389

En este caso es adecuado usar la prueba de Shapiro-Wilk (son menos de 20 series). Dado que el nivel de significancia supera el .05, entonces es posible afirmar que los residuos del modelo de regresión simple se distribuyen normalmente.

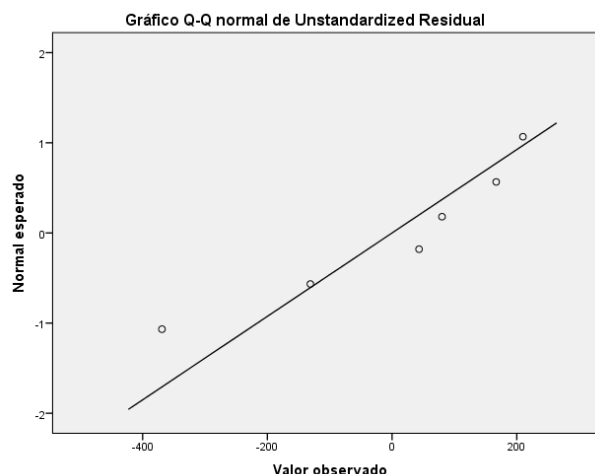


Figura 16: Gráfico de la distribución normal para los residuos de la regresión lineal

Finalmente ha sido necesario comparar los modelos (Red neuronal vs. Regresión Lineal):

Tabla 14: Comparación entre modelos Red Neuronal – Regresión Lineal

Criterio	Red Neuronal	Regresión Lineal
Cantidad de Variables	6 (Todas las variables	
Predictoras	independientes)	1 (Sentimiento_YouTube)
Precision (Accuracy)	43.20%	87.50%
Sesgos	Pocas series para analizar	Pocas series para analizar y autocorrelación de residuos

Ciertamente, la regresión lineal alcanza una mayor precisión, pero el hecho de que no se cumpla una de las condiciones básicas de la regresión puede llegar a desestimar dicho algoritmo. Las redes neuronales no tienen ese inconveniente. Nuevamente hay que mencionar que el tamaño de la muestra no favorece a ninguno de los dos modelos presentados.

3.2. Discusión

El BBVA (2018) evidenció que las búsquedas acerca de la corrupción en Brasil tenían relación con la aprobación del gobierno de dicho país. En el caso del BCP- Bolivia sucede que el interés hacia la entidad financiera, tanto en Google como en YouTube no tiene relación significativa con su reputación corporativa. Ello es verificable usando la prueba de correlación de Pearson. En cambio, si se usan las redes neuronales, de alguna manera ambos indicadores contribuyen a pronosticar la reputación del Banco.

Barbato & Montesi (2016) investigaron si la reputación corporativa se puede evaluar desde datos de redes sociales, específicamente delimitaron los estudios a Twitter y el rendimiento financiero de Volkswagen. Existía correlación entre ambas variables. En el caso del BCP de Bolivia, existe coincidencia entre la reputación corporativa y el sentimiento presente en los comentarios de YouTube y en las noticias de los diarios bolivianos. Algo similar ocurrió con los hallazgos de Vidya, et.al. (2015) para el caso de la satisfacción de los consumidores en base a comentarios de Twitter.

Lo indicado por Shulze-Horn y otros (2015) resulta válido. Dichos investigadores indicaron que los comentarios en YouTube podrían reflejar la reputación corporativa de

una empresa. En este caso del BCP- Bolivia ello se ha podido comprobar con evidencia estadística. Cabe destacar que, para el caso de la entidad financiera, el puntaje sentimental de las noticias también tiene relación con la reputación.

Peet y otros (2015) indicaron que la reputación online de la empresa Replab podía determinarse con el análisis sentimental. La presente investigación coincide en que dicha técnica de alguna manera guarda relación con la reputación. La diferencia radica en que los investigadores citados emplearon Twitter, mientras que para el caso del BCP Bolivia fue empleado YouTube y las noticias como *inputs* para el análisis sentimental.

Grutzmacher (2011) tuvo por objetivo examinar el papel de los medios de comunicación en la reputación corporativa en el caso de Nokia. En ese caso si existe coincidencia, ya que el investigador citado uso la data de las redes sociales de la empresa, mientras que para el caso del BCP fue extraída la data de la cuenta oficial de YouTube de la empresa. El análisis sentimental nuevamente es una técnica que permite estudiar lo que piensa el consumidor acerca de un producto o servicio.

Luca (2011) tuvo por objetivo determinar si la reputación de un restaurante que ha recibido comentarios en Yelp.com termina impactando sobre la demanda del mismo. Si bien es cierto que no se cuentan con datos de la demanda, ciertamente la reputación corporativa puede tener relación con variables de un entorno Social Media.

3.3. Conclusiones

Todos los objetivos de la investigación han sido cumplidos. La hipótesis general resulta verdadera, es decir, La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con las variables del tipo social media en el período 2013-2018.

Sin embargo, no todas las variables del Social Media guardan relación con la reputación corporativa. La hipótesis específica 1 es verdadera: la reputación del Banco BCP Bolivia de acuerdo a Merco guarda relación con el puntaje sentimental de los comentarios en YouTube en el período 2013-2018 ($r = 0.935$). En contraste, la hipótesis específica 2 resulto falsa, no hay correlación significativa entre la cantidad de comentarios y la reputación corporativa de la entidad que es objetivo de estudio.

Con respecto a la hipótesis específica 3 la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco no tiene relación significativa con el número de menciones

de la entidad en los principales diarios online de Bolivia para el periodo señalado. En contraste si existe correlación entre la reputación de la entidad bancaria con el puntaje sentimental acerca de la entidad en los principales diarios ($r = 0.836$) y, por tanto, la hipótesis específica 4 es verdadera.

Con respecto a las hipótesis específicas 5 y 6, en ambas resulta que no existe correlación significativa entre las búsquedas en Google y YouTube realizadas en territorio boliviano y la reputación corporativa del BCP Bolivia para el periodo 2013-2018.

En resumen, las variables sentimentales son las que guardan relación con la reputación corporativa del BCP y no las de tráfico. Las búsquedas son sinónimo de interés, pero este puede ser positivo (querer afiliarse al banco) o negativo (realizar una queja acerca del banco). Esa es su limitación. En cambio, las variables sentimentales sí recogen bien el efecto antes mencionado.

Los comentarios en YouTube surgen en su mayoría a partir de personas comunes, mientras que las noticias en los diarios bolivianos son producidas por las áreas encargadas de cubrir hechos de importancia corporativos. Entonces la reputación corporativa del BCP Bolivia tiene relación tanto con las opiniones de las personas como con las noticias provenientes de los medios escritos.

Empleando el método de redes neuronales es posible aproximarse a la reputación corporativa usando todas las variables independientes del modelo. En este caso, tanto las variables vinculadas al tráfico web aportan al modelo, pero nuevamente las variables sentimentales son las de mayor importancia.

En el caso de la regresión lineal, solo la variable asociada al sentimiento de los comentarios en el canal de YouTube del BCP. No obstante, existen sesgos a considerar en el dicho modelo: el tamaño de la muestra y el hecho de que exista autocorrelación entre los residuos.

3.4. Recomendaciones

El Banco debe monitorear constantemente los indicadores del Social Media vinculado no solo a YouTube, sino a otras redes. De alguna manera estos guardan relación con la reputación corporativa. Debe tomarse en cuenta, además, que el BCP-Bolivia tiene acceso a todas sus plataformas: Facebook, Twitter, etc. Para ellos es más sencillo recolectar data.

Si bien es cierto que el tráfico en Google y YouTube no tiene relación con la reputación corporativa, es adecuado tener monitoreados ambos indicadores debido a que permitirán identificar en qué momento el banco tendrá la atención de los internautas que se encuentran en territorio boliviano, ya sea por un hecho positivo o negativo.

El estudio puede incluir otras variables del tipo visual. Técnicas como el análisis sentimental es aplicable a imágenes y videos. Con software especializado es posible medir las emociones presentes en ambos formatos. Esto posibilitaría el análisis de datos de redes como Instagram, plataforma también usada por el BCP – Bolivia.

El estudio puede ser replicable a otras empresas de origen peruano que operen en mercados extranjeros, siempre y cuando sean considerados por Merco a la hora de medir la reputación corporativa. Del mismo modo, el método empleado es aplicable a empresas extranjeras que operen en el Perú.

3.5. Limitaciones

Debe tomarse en consideración que los comentarios que se encuentran en YouTube pueden admitir múltiples significados. Existe el lenguaje figurado, diferente al lenguaje literal. Así, por ejemplo, si en un comentario se menciona: “Mataría por poder conseguir esa tarjeta de crédito”; ello no implica que esa persona tenga intenciones negativas (matar). Las computadoras todavía no han llegado a comprender plenamente el lenguaje humano. Una situación en la que un computador pueda comprender perfectamente las emociones del ser humano equivaldría a una inteligencia artificial que ha alcanzado la perfección, escenario muy poco probable actualmente.

Las personas no siempre comprendemos a totalidad las emociones que otros desean transmitirnos. Los mensajes que se intercambian por internet ejemplifican ello. El uso continuo de servicios de mensajería instantánea como WhatsApp, Telegram, Facebook Messenger y otros han terminado por evidenciar la existencia del lenguaje SMS. Si bien es cierto que lenguas como el español, procuran la economía del lenguaje, la escritura en entornos electrónicos parece llevar esta característica al extremo.

En la investigación se procuró construir un esquema que permita mantener el íntegro de los textos. Extraer, depurar y transformar una serie de datos en formato de texto libre a patrones numéricos es un proceso que siempre puede perfeccionarse (Hofmann & Chisholm, 2015). A medida que aparecen softwares con algoritmos de mayor potencia que permitan procesar datos textuales, los resultados pueden mejorarse (Hofmann &

Chisholm, 2015). Esta es una limitación para todas las investigaciones que pretendan trabajar con datos que no poseen una estructura definida.

Bibliografía

- Abril, L. (2003). Tribuna: La Gestión Empresarial - La reputación corporativa como concepto. *El País*. Recuperado el 1 de Octubre de 2019, de https://elpais.com/diario/2003/11/29/economia/1070060414_850215.html
- Andreas, K., & Michael, H. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, 53(1), 59-68.
- Autoridad de Supervisión del Sistema Financiero. (2019). *Estadísticas del sector*. Recuperado el 30 de Agosto de 2019, de <https://www.asfi.gob.bo/index.php/bancos-multiples-boletines.html>
- Barbato, M., & Montesi, D. (2016). *Measuring corporate reputation*. Tesis de maestría, University of Bologna, School of Science. Obtenido de http://amslaurea.unibo.it/12434/1/Tesi_Barbato.pdf
- BBVA. (2018). *Analizando la corrupción con Big Data*. Documento de discusión. Recuperado el 30 de Agosto de 2019, de <https://www.bbvaresearch.com/publicaciones/analizando-la-corrupcion-con-big-data/>
- BCP-Bolivia. (2019). *Nuestro Banco BCP*. Recuperado el 14 de Agosto de 2019, de <https://www.bcp.com.bo/NuestroBanco>
- Besley, D. (1991). *Conditioning Diagnostics: Collinearity and Weak Data in Regression*. John Wiley & Sons.
- Blanchard, O. (2011). Social media ROI: managing and measuring social media efforts in your organization. Indianapolis: Que.
- Bolsa de Valores de Bolivia. (2018). Recuperado el 21 de Octubre de 2018, de https://www.bbv.com.bo/Media/Default/Archivos/Fichas/BTB_CAR.pdf
- Carrasco, J. (2017). Percepción y medida de la reputación empresarial. *Revista Económica Industrial*, 12(1), 117-131.
- Castillo, O., & Cheang, V. (2019). *Elementos de la satisfacción en base a comentarios de Tripadvisor durante el 2017 bajo el enfoque Data Mining: caso de peruanos*

- que se desplazaron en ruta nacional y clase económica en Latam. Tesis de licenciatura, Universidad San Ignacio de Loyola, Facultad de Ciencias Empresariales, Lima. Recuperado el 31 de Julio de 2019, de <http://repositorio.usil.edu.pe/handle/USIL/8628>
- Dean, J. (2014). *Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners*. Hoboken, New Jersey: John Wiley & Sons.
- Dergiades, T., Milas, C., & Panagiotidis, T. (2015). Tweets, Google trends, and sovereign spreads in the GIIPS. *Oxford Economic Papers*, 67(2), 406-432. doi:10.1093/oep/gpu046
- Díaz, M. (2015). Características de un perceptrón. Recuperado el 2 de Octubre de 2019, de <https://www.geogebra.org/m/YPPTGAey>
- European Securities and Markets Authority. (2018). *Report on CRA Market Share Calculation*. Recuperado el 2 de Mayo de 2019, de https://www.esma.europa.eu/sites/default/files/library/cra_market_share_calculation_2018.pdf
- Farfan, S., & Zapata, V. (2018). *Consumo influenciado por el Social Media y su relación con variables psicomográficas, caso de las seguidoras de fashion bloggers peruanas*. Tesis de licenciatura, Universidad San Ignacio de Loyola, Facultad de Ciencias Empresariales, Lima. Recuperado el 3 de Agosto de 2019, de http://repositorio.usil.edu.pe/bitstream/USIL/8535/3/2018_Farfan-Miranda.pdf
- Ferruz, Z. (2018). *Reputación corporativa. Estudio del concepto y las metodologías para su medición: propuesta de un con.* Tesis doctoral, Universidad Complutense de Madrid, Facultad de Ciencias de la Información.
- Fuchs, C. (2014). *Social Media, a critical introduction*. London: Sage.
- Fulmer, R., & Shapiro, J. (2016). *Searching For Revolution Google Trends and Egypt's Arab Spring*. Princeton University, Woodrow Wilson School. Recuperado el 16 de Agosto de 2019, de <http://arks.princeton.edu/ark:/88435/dsp011c18dj23h>

- Gatos, B., & Nouraros, M. (1998). A New Method for Segmenting Newspaper Articles. *Research and Advanced Technology for Digital Libraries*, 695-696.
doi:10.1007/3-540-49653-X_66
- Ghose, A., Ipeirotis, P., & Sundararajan, A. (2007). Opinion Mining Using Econometrics: A Case Study on Reputation Systems. *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, (págs. 416-423).
- Google Cloud. (2019). *Demo de Google Cloud*. Obtenido de <https://cloud.google.com/natural-language/?hl=es>
- Google Trends. (2019a). *Interés hacia la empresa Huawei*. Recuperado el 1 de Octubre de 2019, de <https://trends.google.es/trends/?geo=ES>
- Google Trends. (2019b). *Interés hacia United Airlines en YouTube*. Recuperado el 1 de Octubre de 2019
- Google Trends. (2019c). *Interes hacia el BCP en territorio boliviano*.
- Grammarly. (2019). *Demo*. Recuperado el 30 de Octubre de 2019, de <https://www.grammarly.com/>
- Grutzmacher, A. (2011). *Reputation 2.0: The role of social media in corporate reputation - Case Nokia*. Tesis doctoral, Aalto Univerity, School of Economics .
Recuperado el 24 de Agosto de 2018, de http://epub.lib.aalto.fi/fi/ethesis/pdf/12502/hse_ethesis_12502.pdf
- Hamed, A., & Said, S. (2013). Using NVivo for data analysis in qualitative research. *International Interdisciplinary Journal of Education*, 2(2), 181-186.
- Henriquez, C., Guzmán, J., & Salcedo, D. (2016). Minería de Opiniones basado en la adaptación al español de ANEW sobre opiniones acerca de hoteles. *Procesamiento del Lenguaje Natural*, 56, 25-32.
- Hernandez, R., & Mendoza, C. (2018). *Metodología de la investigación*. Mc Graw Hill.
Obtenido de <https://www.ebooks7-24.com:443/>
- Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la Investigación* (Quinta ed.). México: McGraw Hill.

- Hofmann, M., & Chisholm, A. (2015). *Text Mining and Visualization: Case Studies Using Open-Source Tools*. Chapman and Hall/CRC .
- Horn, I., & Taros, T. (2015). Business Reputation and Social Media: A Primer on Threats and Responses. *Journal of Direct, Data and Digital Marketing Practice* , 16(3), 41-50.
- IBM. (2019). *Language Translator*. Obtenido de <https://www.ibm.com/watson/services/language-translator/>
- IBM. (2019). Live Event Streaming Services & Solutions. Recuperado el 1 de Noviembre de 2019, de <https://www.ibm.com/watson/media/live-event-streaming-services>
- IBM. (2019). SPSS Statistics. Obtenido de <https://www.ibm.com/pe-es/products/spss-statistics>
- IBM Knowledge Center. (s.f.). CRISP-DM Help Overview. Recuperado el 7 de Enero de 2020, de https://www.ibm.com/support/knowledgecenter/SS3RA7_15.0.0/com.ibm.spss.crispdm.help/crisp_overview.htm
- Jasso-Hernández, M., Pinto, D., Vilariño, D., & Lucero, C. (2014). Analisis de sentimientos en Twitter: impacto de las características morfológicas. *Research in Computing Science*, 72(2014), 37-45.
- Kalambe, Y., Pratiba, D., & Shah, P. (2015). Big Data Mining Tools for Unstructured Data: A Review. *International Journal of Innovative technology and Research*, 3(2).
- Kothari, C. (2004). *Research Methodology, Methods and Techniques* (Segunda ed.). New Age International.
- Kumlin, M. (2016). *Social media as a tool for corporate reputation management*. Tesis de maestría, LAHTI UNIVERSITY OF APPLIED SCIENCES. Recuperado el 21 de Junio de 2018, de https://www.theseus.fi/bitstream/handle/10024/117074/Kumlin_Marina.pdf?sequence=2

- Kurtner, M., Nachtsheim, C., & Neter, J. (2004). *Applied Linear Regression Models* (Cuarta ed.). McGraw-Hill Irwin.
- Language Tool. (2019). *Demo*. Recuperado el 30 de Octubre de 2019, de <https://languagetool.org/>
- Li, C. (2011). The impact of social media on corporate reputation. *Vista: Education, Economy and Community*, 1(2), 62-71.
- Llorente & Cuenca. (2017). *El caso United: cómo han cambiado las crisis de reputación en el mundo de hoy*. Recuperado el 28 de Agosto de 2019, de https://ideas.llorenteycuenca.com/wp-content/uploads/sites/5/2017/06/170622_DI_Articulo_Aerolineas-2.pdf
- Lovett, J. (2011). *Secrets : Social Media Metrics Secrets*. Wiley.
- Luca, M. (2011). *Reviews, Reputation, and Revenue: The Case of Yelp.com*. Working Paper, Harvard Business School. Recuperado el 14 de Setiembre de 2018, de https://www.hbs.edu/faculty/Publication%20Files/12-016_a7e4a5a2-03f9-490d-b093-8f951238dba2.pdf
- Manning, J. (2014). Definition and classes of Social Media. En K. Harvey, *Encyclopedia of social media and politics* (págs. 1158-1162). Ca: Sage.
- Manterola, C., & Pineda, V. (2008). El valor de “p” y la “significación estadística”. *Revista Chilena de Cirugía*, 60(1), 86-89.
- Mantyla, M., Graziotin, D., & Kuuttila, M. (2016). The Evolution of Sentiment Analysis - A Review of Research Topics, Venues, and Top Cited Papers. *Computer Science Review*, 27. doi:10.1016/j.cosrev.2017.10.002
- Merco. (2020). ¿Qué es Merco? Recuperado el 4 de Marzo de 2020, de <http://www.merco.info/pe/que-es-merco>
- Microsoft. (s.f.). Microsoft Excel. Obtenido de <https://products.office.com/es/excel>
- Moniz, A., & De Jong, F. (2014). *Reputational DAMAGE: Classifying the impact of allegations of irresponsible corporate behavior expressed in the f*. University of Twente, Netherlands. Recuperado el 20 de Julio de 2018, de https://www.eur.nl/sites/corporate/files/Reputational_DAMAGE_-

_Classifying_the_impact_of_allegations_of_irresponsible_corporate_behavior_expressed_in_the_financial_media.pdf

- Montgomery, D., Peck, E., & Vining, G. (2012). *Introduction to Linear Regression* (Quinta ed.). Wiley.
- Muñoz, E. (2015). *financiera, Aplicación de redes neuronales y regresión logística para predecir el éxito de compra de deuda de una entidad*. Tesis para obtener el grado de magíster en estadística, Lima. Recuperado el 2 de Marzo de 2018, de <http://repositorio.lamolina.edu.pe/bitstream/handle/UNALM/1748/E13-M8-T.pdf?sequence=5&isAllowed=y>
- Nadkarni, P., Ohno-Machado, L., & Chapman, W. (2011). Natural language processing: an introduction . *Journal of the American Medical Informatics Association*, 18(5), 544–551. doi:10.1136/amiajnl-2011-000464
- Oliva, F. (2014). *Minería de Opinión y Análisis de Sentimiento*. Tesis para optar por grado de ingeniero, Pontificia Universidad Católica de Valparaíso, Facultad de Ingeniería. Recuperado el 22 de Diciembre de 2018, de http://opac.pucv.cl/pucv_txt/txt-5500/UCE5729_01.pdf
- Osborne, J., & Waters, E. (2002). Four Assumptions Of Multiple Regression That Researchers Should Always Tests. *Practical Assement, Research & Evaluation*, 8(2).
- Peetz, M., de Rikje, M., & Kaptein, R. (2015). Estimating Reputation Polarity on Microblog Posts. *Information Processing and Management*.
- Pérez, M., & De Paz, Y. (2007). *Cuadernos de Estadística - Aplicación de las Redes Neuronales artificiales a la regresión* (Vol. 35). Madrid: La Muralla.
- Pérez, M., & Martín, Q. (2003). *Cuadernos de Estadística - Aplicaciones de las redes neuronales a la estadística* (Vol. 25). Madrid: La Muralla.
- Petasis, G., Tsirakis, N., & Spiliotopoulos, D. (2013). Large-scale Sentiment Analysis for Reputation Management. *Pathos Workshop SCGL*, 21, págs. 1-14. Darmstadt.

- Razali, N., & Wah, Y. (2011). Power Comparison of Shapiro Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson Darling Test. *Journal of Statistical Modeling and Analytics*, 2(1), 21-33.
- Restrepo, B., & González, J. (2007). Revista Colombiana de Ciencias Pecuarias. 20(2), 183-192.
- Rincón-García, S. (2017). *Minería de textos y análisis de sentimientos en sanidadysalud.com*. Tesis de maestría, Universidad Complutense de Madrid. Recuperado el 21 de Octubre de 2018, de https://eprints.ucm.es/39524/1/memoriaTFM_sergio_rincon_garcia.pdf
- Shulze-Horn, I., Taros, T., Dirkes, S., Hüer, Rose, M., Tietmeyer, R., & Constantinides, E. (2015). Business Reputation and Social Media: A Primer on Threats and Responses. *Journal of Direct Data and Digital Marketing Practice* , 16(3), 194-208. doi:10.1057/dddmp.2015.1
- Silge, J., & Robinson, D. (2017). *Text Mining with R: A Tidy Approach*. O'Reilly Media.
- Szwajka, D. (2017). The Role of Social Media in Corporate Reputation Management - The Result of Polish Enterprise. 9(1), 161-174. doi:10.1515/fman-2017-0013
- Textcleanr. (2019). *Textcleanr - Demo*. Recuperado el 20 de Octubre de 2019, de <https://www.textcleanr.com/>
- Thelwall, M. (2017). Social media analytics for YouTube comments: potential and limitations. *International Journal of Social Research Methodology*, 21(3), 303-316. doi:10.1080/13645579.2017.1381821
- Vidya, N., Fanany, M., & Budi, I. (2015). Twitter Sentiment to Analyze Net Brand Reputation of Mobile Phone Providers. *Procedia - Computer Science*, 72, 519-526.
- Wendler, T., & Gröttrup, S. (2016). *Data Mining with SPSS Modeler, Theory, Exercises and Solutions*. Berlin: Springer.
- Whu, Y., & Hsieh, S. (2014). Public Opinion Toward CSSTA:A Text Mining Approach. *Computational Linguistics and Chinese Language Processing*, 19(4), 19-28.

YouTube - Canal de BCP - Bolivia. (2019). Recuperado el 30 de Octubre de 2019, de https://www.youtube.com/channel/UCy9hNZ6PNM3RRC_Fq4v8vfg

Zanfrillo, A., Artola, M., & Morettini, M. (2016). Comparación entre el Índice MERCO de reputación corporativa e indicadores de popularidad web. *FACES*, 22(47), 7-22.

Anexo 1: Calendario y Recursos Disponibles

Tabla A1: Calendario de Actividades

Actividad	nov-18	dic-18	ene-19	feb-19	mar-19	abr-19	may-19	jun-19	jul-19	ago-19	sep-19	oct-19
Revisión del Plan de Tesis												
Recolección de Data - Social Media												
Segunda Revisión del Plan de Tesis												
Procesamiento de Data- Social Media												
Tercera Revisión del Plan de Tesis												
Procesamiento Estadístico de la Data												
Redacción Versión Final de Tesis												
Sustentación												

Elaboración: Propia

Recursos Disponibles

Todos los datos necesarios para la investigación son de libre acceso, por lo tanto, no requieren desembolso de dinero alguno. Los instrumentos si tienen costo. No obstante, los instrumentos si cuentan con un costo. Entre los recursos disponibles está el Excel. En el caso del Google Cloud Natural Language, NVIVO y SPSS es posible acceder a versiones Beta u *Open Source* (recurso abierto). Es posible acceder a los mismos por el tiempo necesario para la investigación.

Anexo 2: Interface de Uso de Google Trends

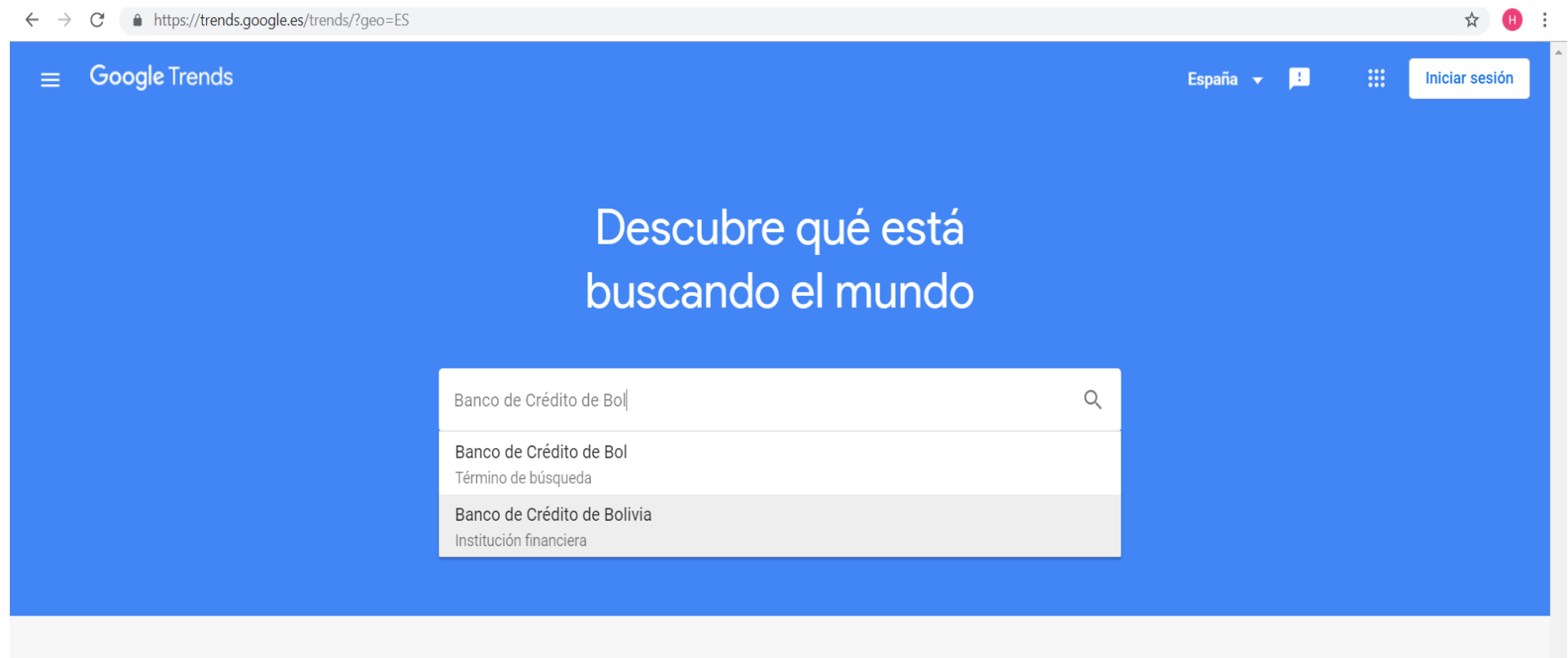


Figura A1: Rastreo de dato de BCP Bolivia en Google Trends

Fuente: Google Trends (2019)

Nota: Debe tomarse en consideración que la data ha sido depurada por el navegador de Google al usar la función autocompletado.

Anexo 3: Delimitación de resultados de Google Trends a espacio boliviano y para YouTube

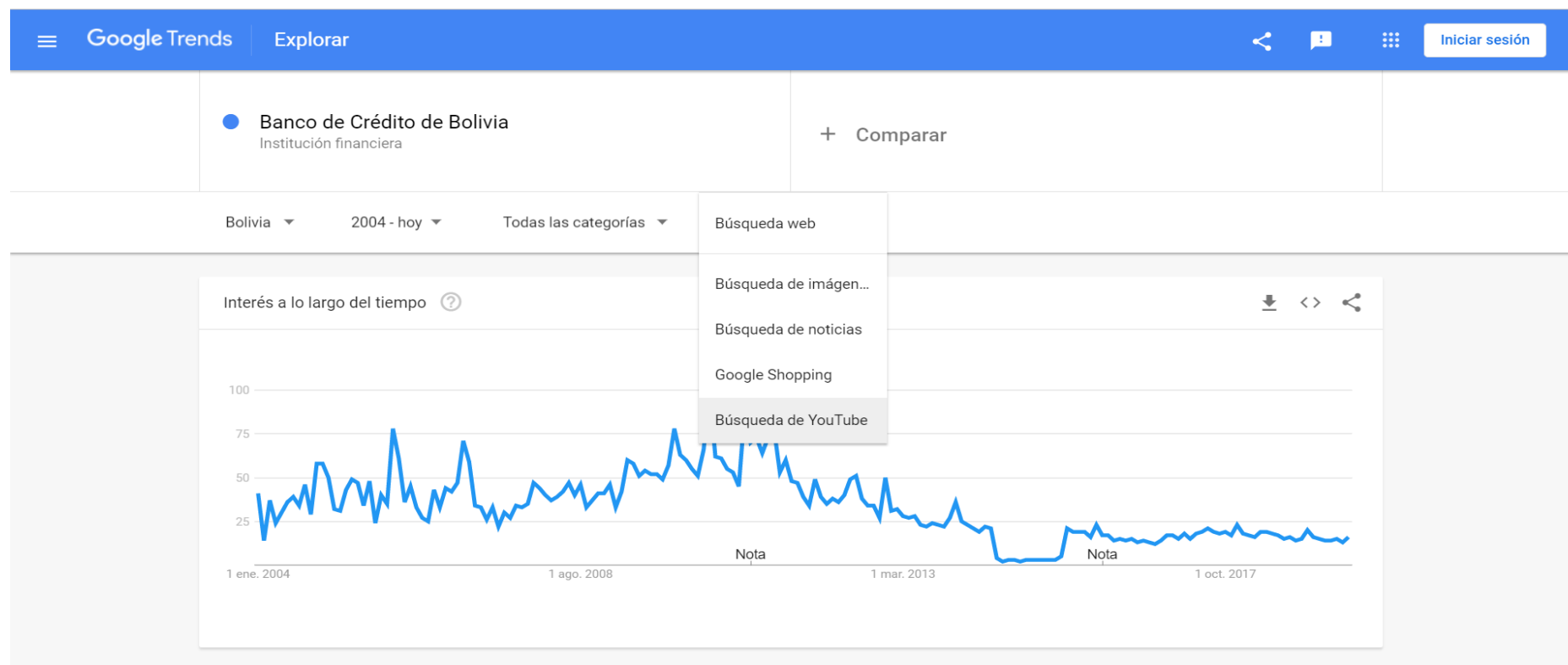


Figura A2: Filtrado de búsqueda

Fuente: Google Trends (2019c)

Nota: Mediante los comandos en forma de cajas con contenido textual es posible seleccionar país, lapso y plataforma de donde desea obtenerse la información,

Anexo 4: Interface de uso de Google Cloud Natural Language

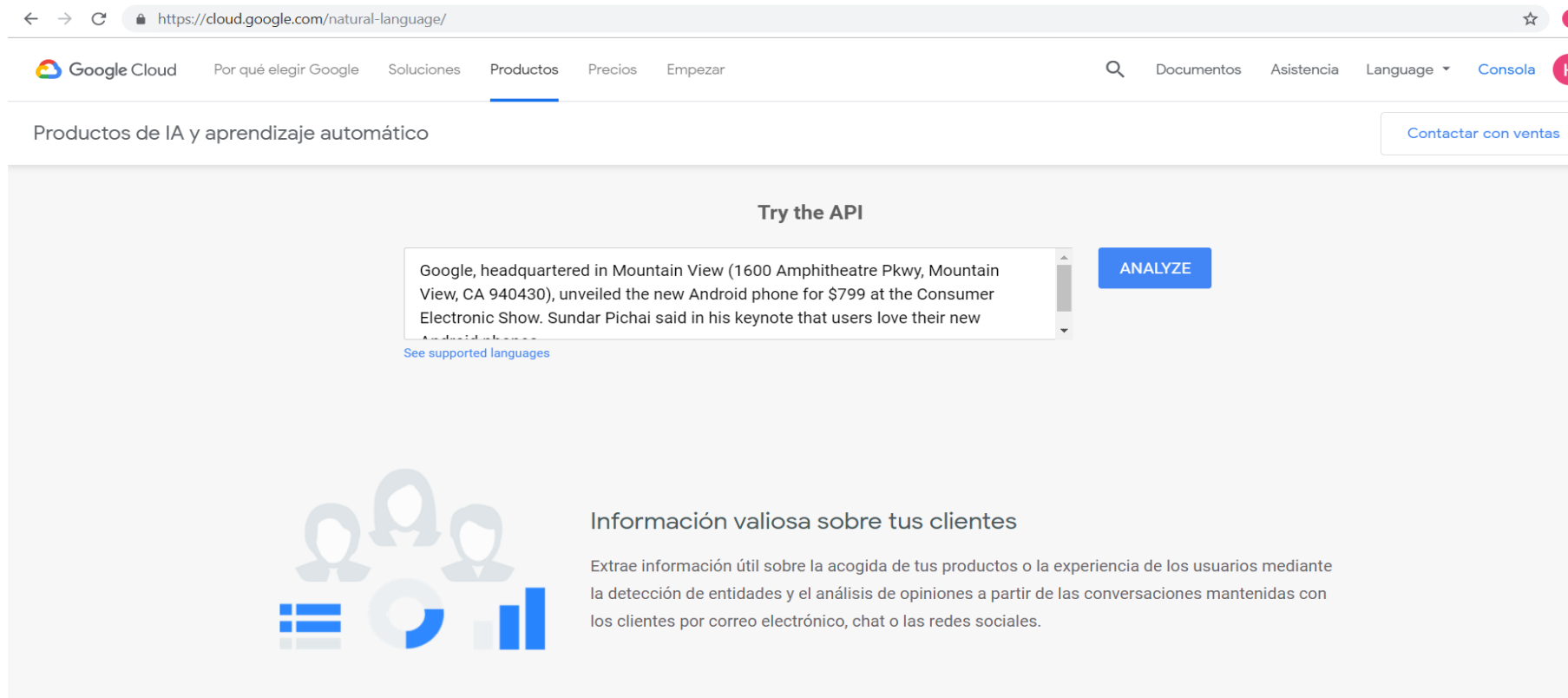


Figura A3: Ventana de inicio de Google Cloud Natural Language

Fuente: Google Cloud (2019)

Nota: Para obtener resultados basta con presionar el botón *Analizar*. El texto puede ser ingresado mediante una operación de copiado y pegado

Anexo 5: Resultados de Google Cloud Natural Language

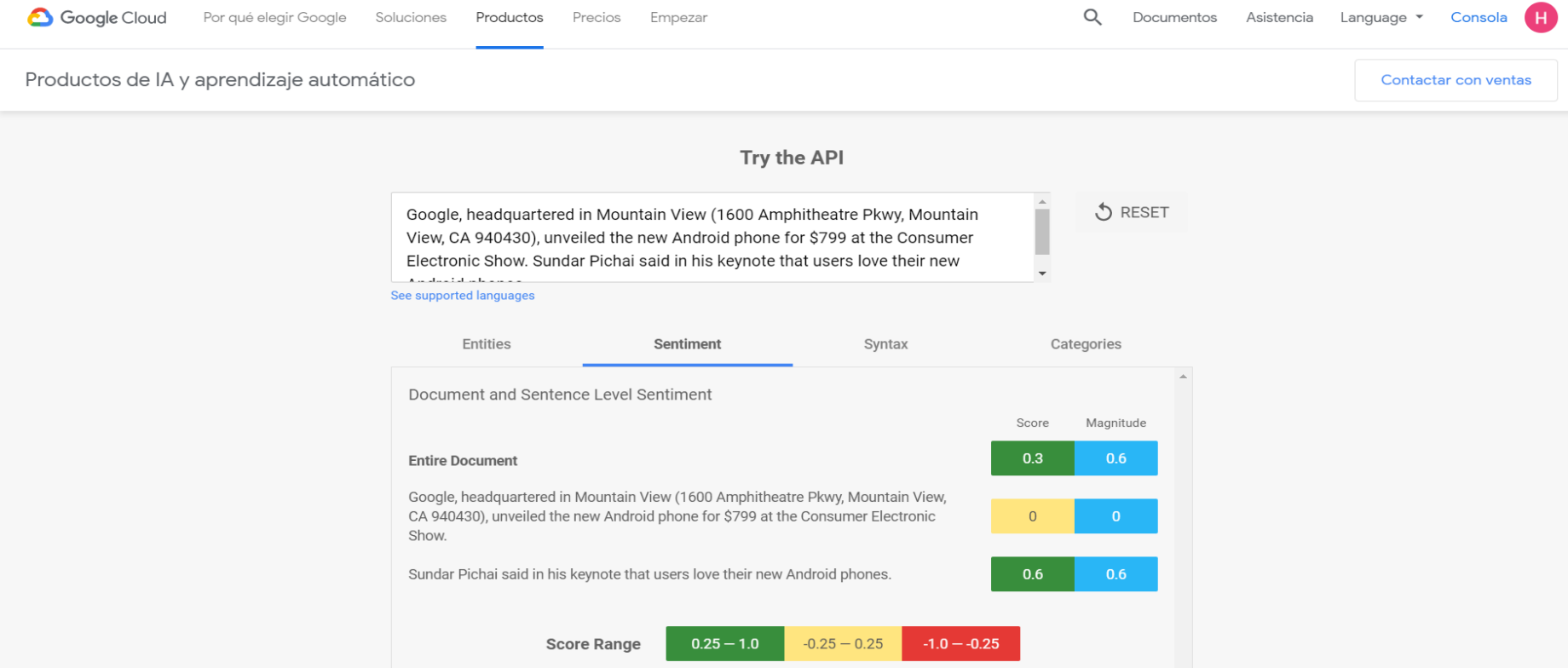


Figura A4: Resultados de Google Cloud para la data de ejemplo

Fuente: Google Cloud (2019)

Nota: La data a emplear será el sentimiento para el documento entero, el cual es un consolidado del contenido textual ingresado.

Anexo 6: Uso de NVIVO

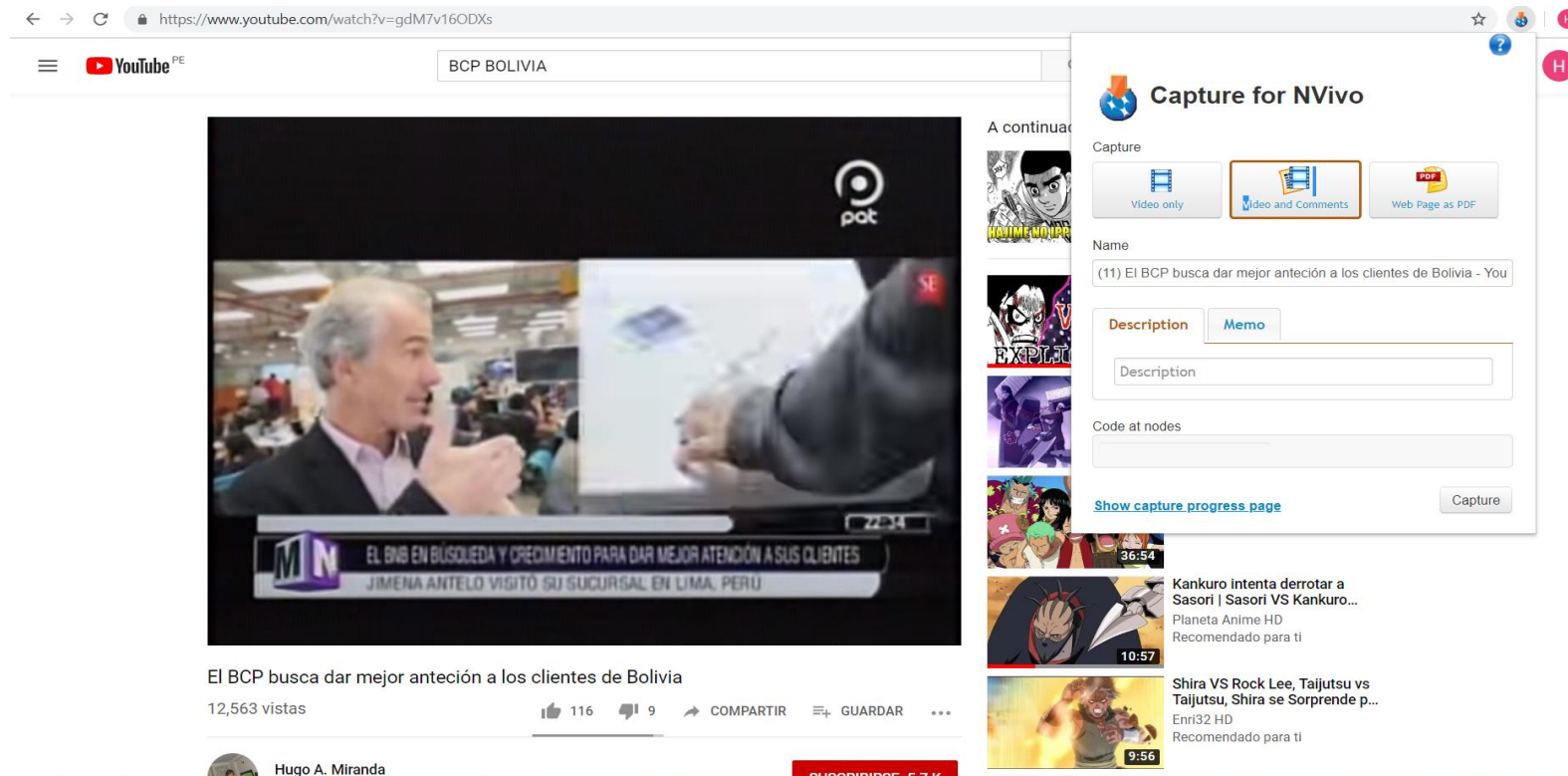


Figura A5: Interface inicial del capturado de NVIVO

Fuente: YouTube - Canal de BCP – Bolivia (2019)

Nota: Basta con ingresar al link del video y activar la pestaña circular azul para activar el comando de extracción. Con ello se extraen tanto el video como los comentarios

Anexo 7: Proceso de captura finalizado

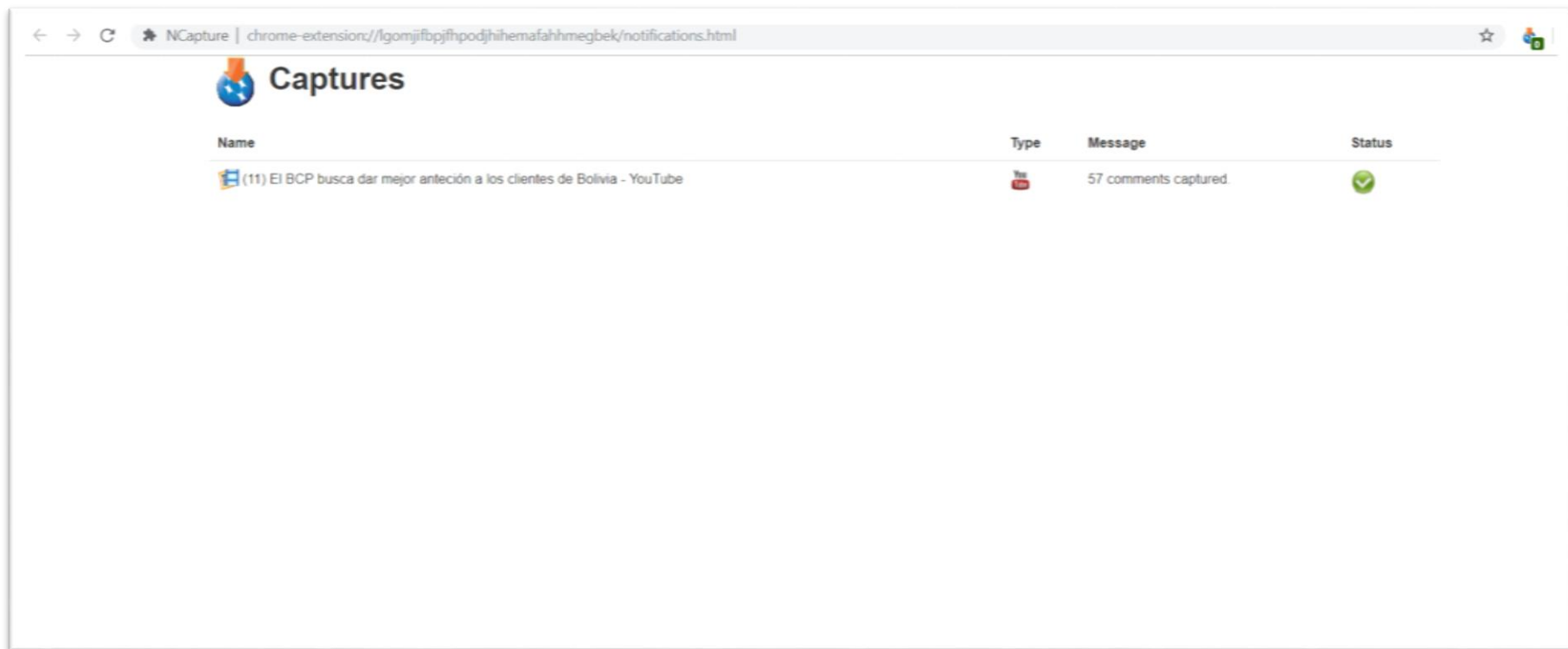


Figura A7: Captura de NVivo finalizada

Fuente: Adaptado de Hamed & Said 2013) a YouTube - Canal de BCP – Bolivia (2019)

Nota: En caso de que la descarga de comentarios no se haya completado aparece una cruz en vez de un *check* en Status

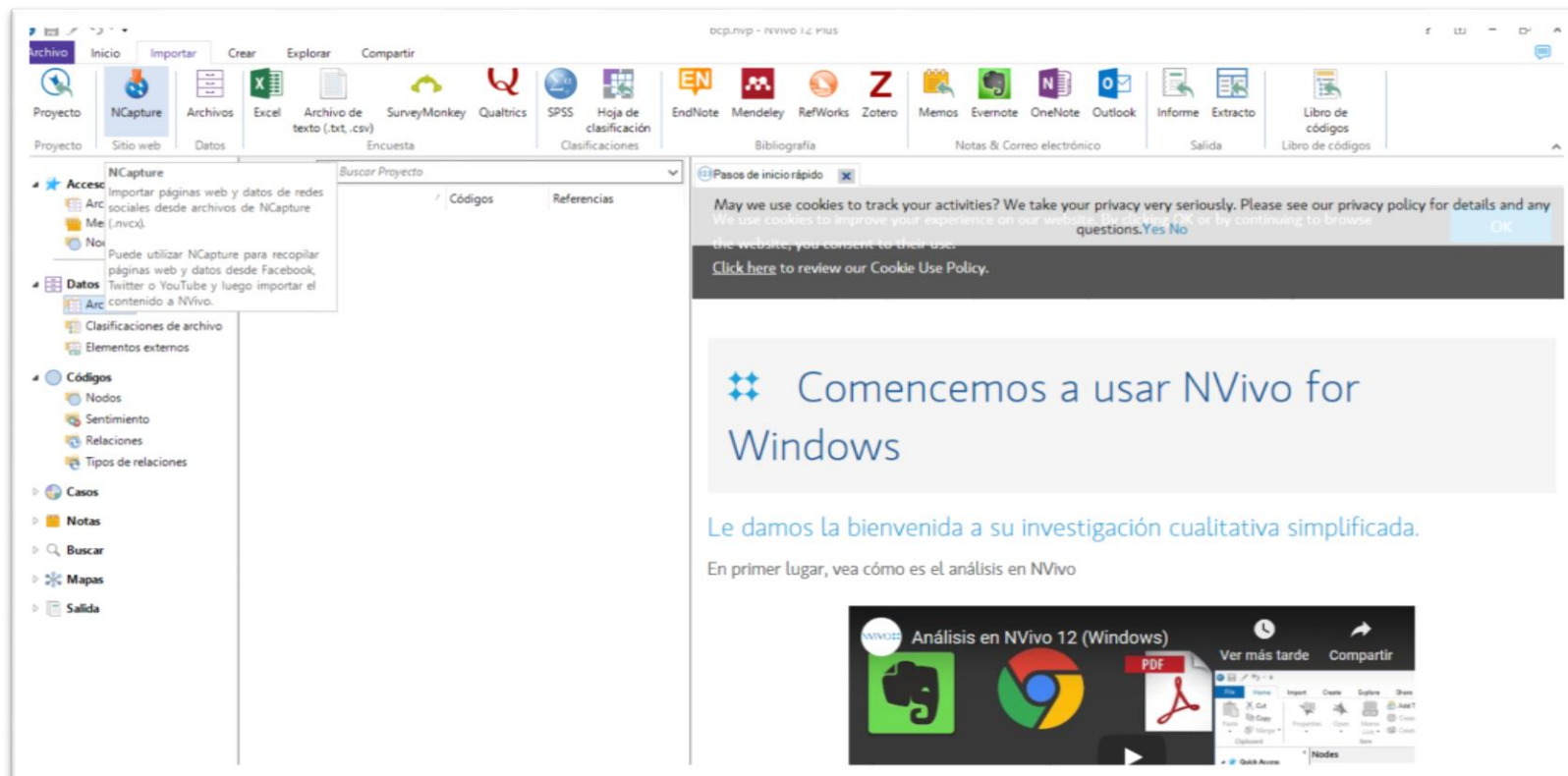


Figura A8: Primer paso para el procesamiento de datos

Fuente: Adaptado de Hamed & Said (2013) a YouTube - Canal de BCP – Bolivia (2019)

Nota: Activar en la pestaña Archivo la opción NCapture

Anexo 8: Selección y conversión de data formato en Excel

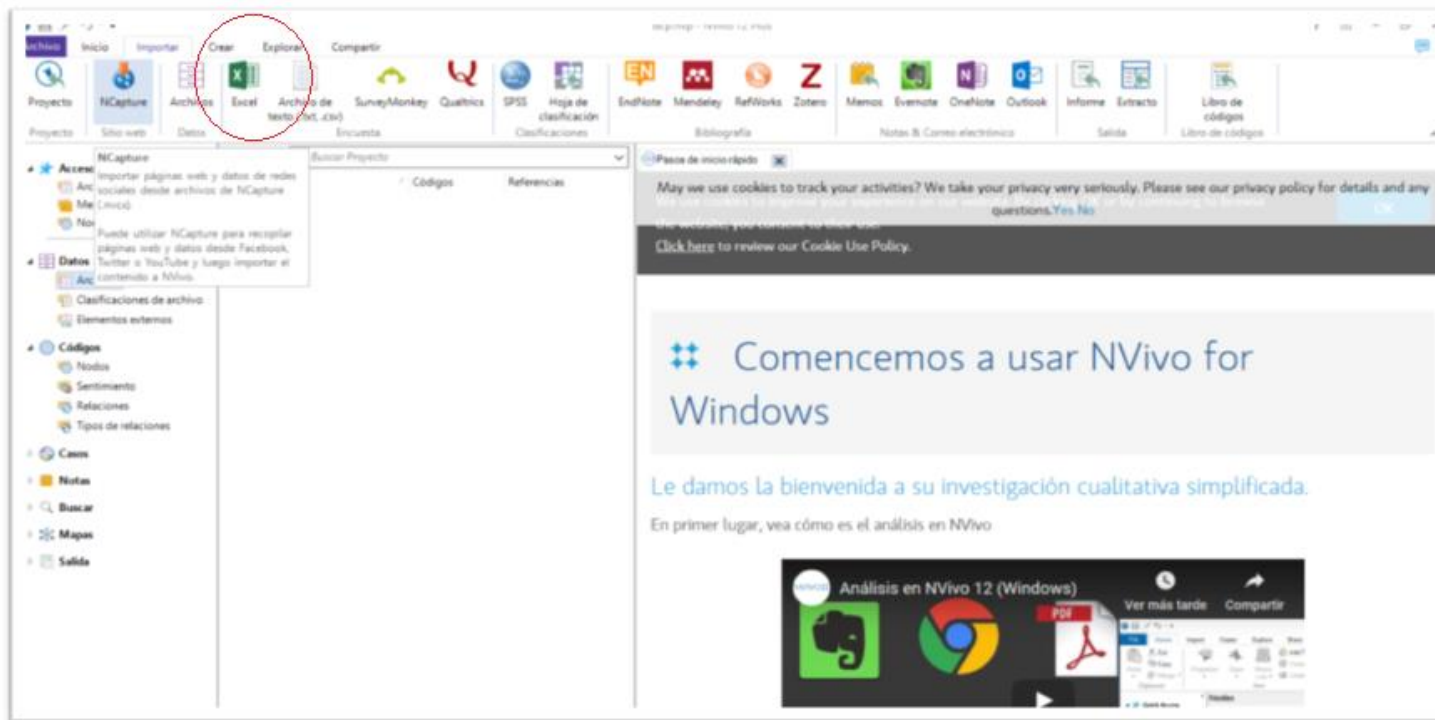


Figura A9: Selección de data con Nvivo

Fuente: Adaptado de Hamed & Said (2013) a YouTube - Canal de BCP – Bolivia (2019)

Nota: Una vez seleccionada la data ya puede manipularse exportarse a Excel con el comando Excel (elipse en rojo)

Anexo 9: Interface de software de depuración

TextCleanr

Home

TextCleanr

The quick, easy, web based way to fix and clean up text when copying and pasting between applications. Remove email indents, find and replace, clean up spacing, line breaks, word characters and more. Perfect for tablets or mobile devices. If you like this tool, check out my [URL & Text Shortener](#)

Paste Text to clean in the box below:

Paste text here, Select Options, Click Clean.

Email Indents

Do Nothing

Text Case

Do Nothing

HTML

Do Nothing

Line Breaks

Do Nothing

Spaces

Do Nothing

Related Tools of Mine:

- [TextSendr - Send Text Messages](#)
- [Tiny.tw URI & Text Shortener](#)
- [NoSlang Internet Slang Dictionary](#)
- [Server Header Status Tool](#)

Figura A10: Uso de Textcleanr

Fuente: Textcleanr (2019)

Anexo 10: Interface de *softwares* traductores

The screenshot shows the IBM Watson Language Translator web interface. At the top, the browser address bar displays 'ibm.com/watson/services/language-translator/'. The navigation bar includes links for 'Watson', 'About', 'Products', 'Use Cases', 'Stories', 'Learn', and 'With Watson', along with a 'Get Started Free' button. The main content area is titled 'Translate Text' in orange. It is divided into two columns: 'Input' and 'Output'. The 'Input' column has a sub-header 'Enter or paste text from a passage.' and a dropdown menu labeled 'Choose Language'. Below this is a tabbed interface with 'Text' and 'Rest API' tabs, where 'Text' is selected. The 'Output' column has a sub-header 'Copy output from this field to clipboard.' and a dropdown menu labeled 'Choose Language'. Below this is a tabbed interface with 'Text' and 'JSON' tabs, where 'Text' is selected. A 'Let's talk' button is located in the bottom right corner.

Figura A11: Uso de Watson Translate

Fuente: IBM (2019)

Anexo 11: Uso de software correctores

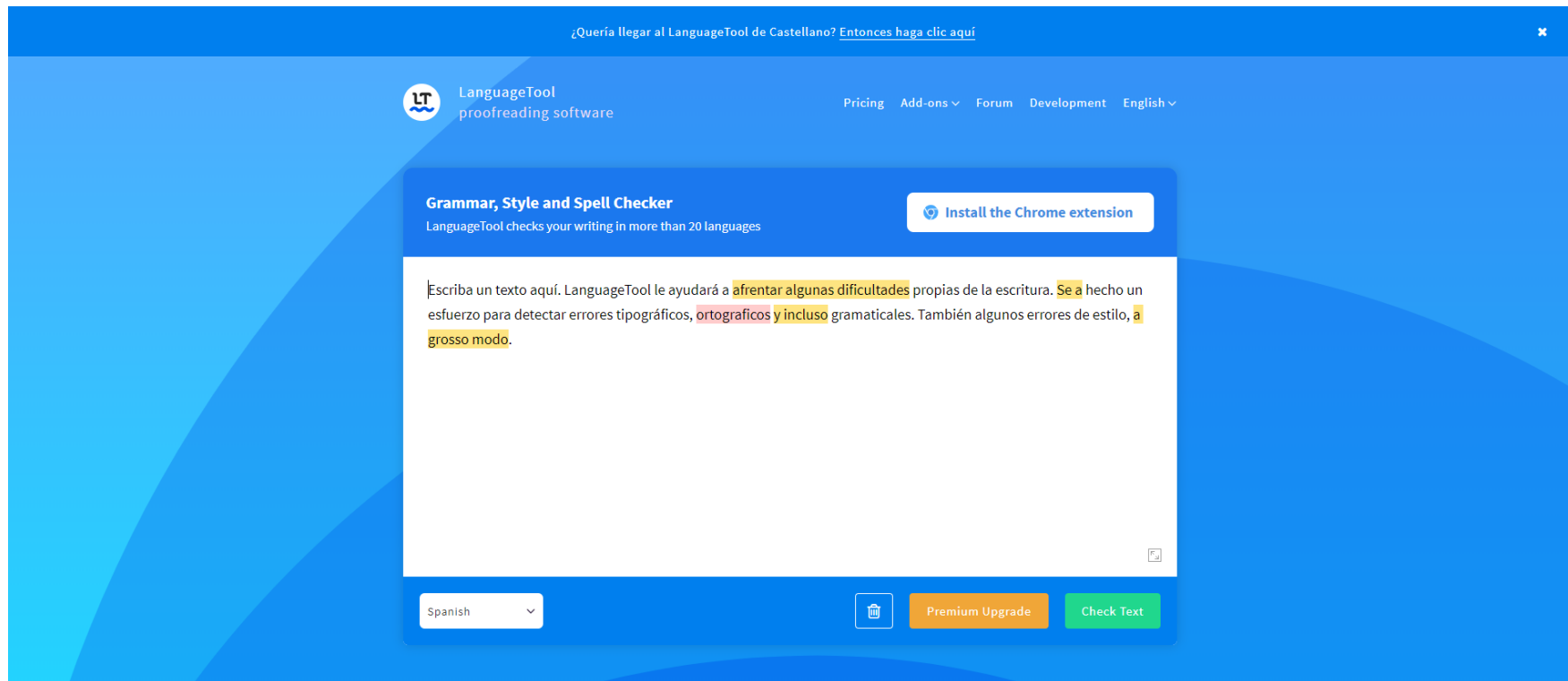


Figura A12: interface de Language Tool

Fuente: Language Tool (2019)

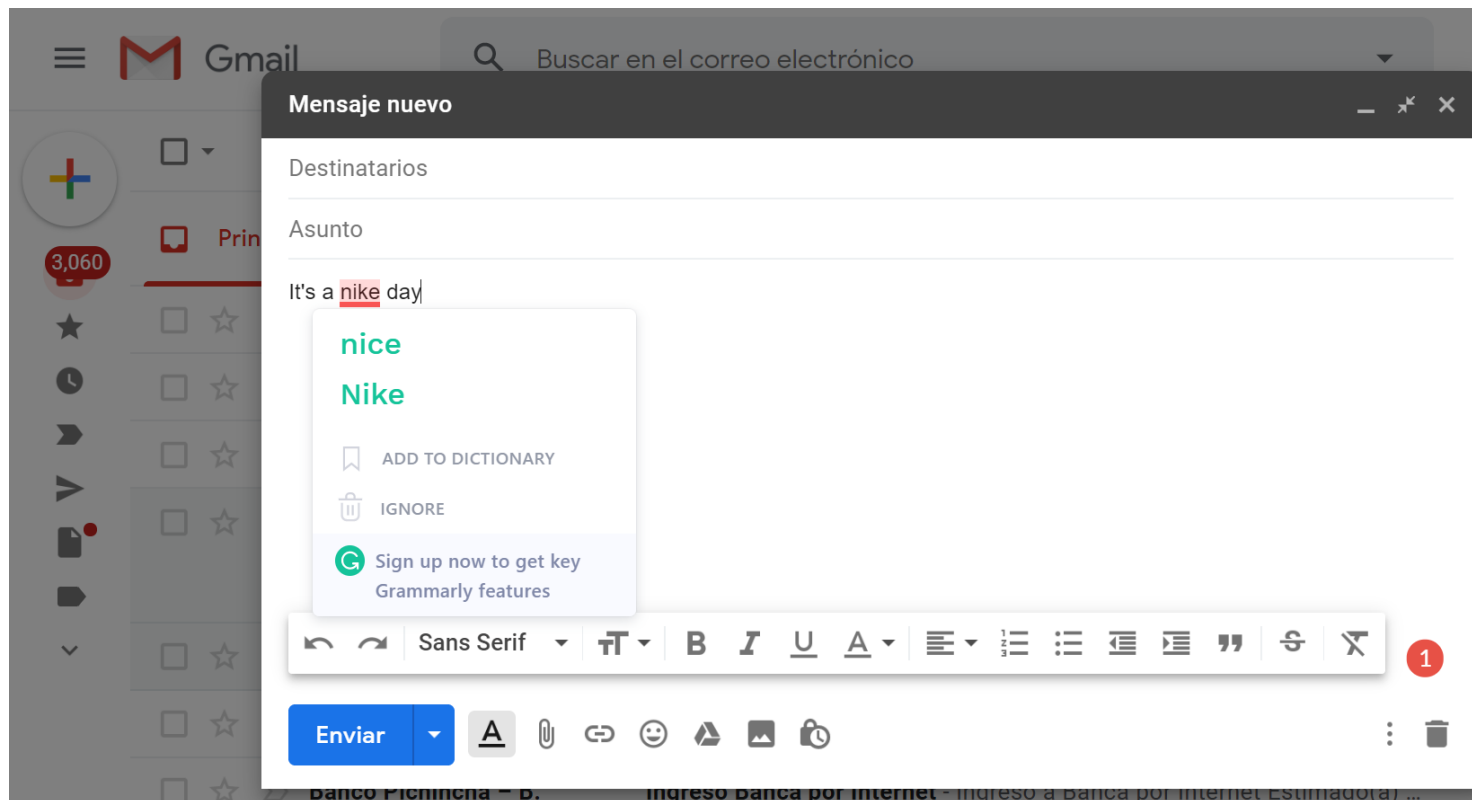


Figura A13: interface de Grammarly incorporado a hoja HTML (correo)

Fuente: Grammarly (2019) integrado a Gmail.

Anexo 12: Matriz de consistencia

Problema de Investigación	Objetivo General	Hipótesis General	Variable Dependiente	Tipo de Investigación	Población
¿La reputación corporativa del BCP Bolivia tiene relación con las variables Social Media de su entorno durante el período 2013-2018?	·Objetivo General: Analizar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con las variables del tipo Social Media en el período 2013-2018.	Hipótesis General: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con las variables del tipo social media en el período 2013-2018.	Reputación Corporativa	Cuantitativa	Datos de Merco -BCP
	Objetivos Específicos	Hipótesis Específicas	Variables Independientes	Diseño de Investigación	Plan de análisis
	Objetivo Específico 1: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el puntaje sentimental de los comentarios en YouTube realizados en el canal de la empresa durante el período 2013-2018.	H1: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el puntaje sentimental de los comentarios en YouTube realizados en el canal de la empresa durante el período 2013-2018.	Puntaje Sentimental de los comentarios en el canal del BCP-Bolivia en YouTube	Correlacional	Correlación de Pearson
	Objetivo Específico 2: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el número comentarios en YouTube realizados en el canal de la empresa durante el período 2013-2018.	H2: a reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el número comentarios en YouTube realizados en el canal de la empresa durante el período 2013-2018.	Número de comentarios en el canal del BCP-Bolivia en YouTube		
	Objetivo Específico 3: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el número menciones de la entidad en los principales diarios online de Bolivia en el período 2013-2018.	H3: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el número menciones de la entidad en los principales diarios online de Bolivia en el período 2013-2018.	Número de menciones en diarios de Bolivia acerca del BCP.	No experimental	Redes Neuronales
	· Objetivo Específico 4: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el puntaje sentimental acerca de la entidad en los principales diarios online de Bolivia en el período 2013-2018.	H4: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el puntaje sentimental acerca de la entidad en los principales diarios online de Bolivia en el período 2013-2018.	Puntaje Sentimental de las noticias acerca del BCP Bolivia.	Transversal	
	Objetivo Específico 5: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con EL índice de búsquedas en Google acerca de la entidad en territorio boliviano en el período 2013-2018.	H5: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el índice búsquedas en Google acerca de la entidad en territorio boliviano en el período 2013-2018.	Índice de Búsquedas en Google acerca del BCP Bolivia		
	Objetivo Específico 6: Determinar si la reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el índice de búsquedas en YouTube acerca de la entidad en territorio boliviano en el período 2013-2018.	H6: La reputación corporativa del Banco BCP Bolivia de acuerdo a Merco guarda relación con el índice de búsquedas en YouTube acerca de la entidad en territorio boliviano en el período 2013-2018.	Índice de Búsqueda en YouTube acerca del BCP Bolivia		

Elaboración Propia

