

# Implicaciones éticas del uso de Big Data en entornos profesionales

Iván González Giménez

**Resumen—** En este proyecto se ha analizado como se utiliza el Big Data en entornos profesionales y las implicaciones éticas que esto acarrea. De esta manera se ha realizado un profundo análisis del uso de esta herramienta en los sectores de la política y de la seguridad, sectores donde el debate sobre donde reside el límite del uso de Big Data aún no está resuelto. Con la información recogida del sector de la política, se han aplicado algunas de las ideas presentes en este sector para desarrollar un aplicativo que, mediante un modelo predictor, es capaz de discernir la ideología política de un usuario de Facebook mediante su información pública. Mediante la información recogida de ambos sectores, se ha desarrollado un ensayo ético donde se recogen ideas de diferentes maneras de pensar como el objetivismo y el utilitarismo con el objetivo de esclarecer cual es el uso moralmente correcto del Big Data.

**Palabras clave—** Big Data, Objetivismo, BigML, Facebook, Machine Learning, Python, Utilitarismo, Kant

**Abstract—** In this project we have analyzed how the Big Data is used in professional environments and the ethical implications that causes. In this way, we have made a profound analysis in the use of this tool in the politician and security sectors, sectors where the debate about where's the limit of the use of Big Data isn't resolve yet. With the information collected about the politician sector, we have applied some of the present ideas in this sector to develop an application that, with a predictor model, is capable to discern the politic ideology of a user from Facebook with his public information. With the information collected in both sectors, we have developed an ethic assay where is collected ideas from different ways of thinking like objectivism and utilitarianism with the purpose of bring some light in what the correct moral use of the Big Data is.

**Keywords—** Big Data, Objectivism, BigML, Facebook, Machine Learning, Python, Utilitarianism, Kant

## 1 INTRODUCCIÓN

**E**n la actualidad, cualquier actividad que se realice en Internet queda registrada. Esto ha llevado que el volumen de datos que son almacenados alcance una cantidad inmensa (llegando a alcanzar, según las estimaciones el tamaño de 44 Zettabytes en 2020 [1] ). Al conjunto de estos datos es a lo que denominamos Big Data. El Big Data es utilizado como una fuente de la cual se pueden extraer información de suma importancia. Esto ha llevado que cada vez más entidades hagan uso de esta tecnología

con la intención de conseguir un beneficio. El Big Data se ha convertido en una de las herramientas más utilizadas en este ámbito y ha permitido a las organizaciones conocer mejor el contexto en el que se mueven y actuar en consecuencia. Algunos ejemplos de este hecho se ven reflejados en sectores diversos como la política [2] (donde el Big Data es utilizado para diseñar estrategias de campaña) o la seguridad nacional [3] (donde el Big Data es utilizado para conocer posibles amenazas para el país). Por lo tanto, Big Data no solo trae consigo beneficios para las organizaciones, sino también para la sociedad. En un mundo donde todo está conectado, es imposible discernir un futuro donde la sociedad no sea codependiente del Big Data. Sin embargo, la irrupción del Big Data en los últimos años ha difuminado la línea de aquello que es privado y lo que no, trayendo consigo un nuevo contexto donde el uso de esta herramienta puede comportar debates sobre la moralidad de los actos de las empresas u organizaciones y donde no existe un código

• E-mail de contacto: ivan.gonzalezgi@e-campus.uab.cat  
• Mención realizada: Tecnologías de la Informació  
• Trabajo tutoritzat por: Ruben Rubio Barrera (Departament d'Enginyeria de la Informació i de les Comunicacions)  
• Curso 2018/19

ético que imponga límites al uso de esta herramienta.

Este proyecto busca realizar una reflexión dentro de este marco, introduciendo teorías éticas provenientes de diferentes pensadores y analizar como encajan en este contexto. Los movimientos que se han escogido como referencia para dicha reflexión son el objetivismo, el utilitarismo y la filosofía kantiana. Sin embargo, para realizar este análisis ético, previamente se debe contextualizar como se está utilizando el Big Data en la actualidad y para ello otro de los objetivos es analizar el uso del Big Data en los sectores de la política y de la seguridad nacional. La elección de estos dos sectores para tal propósito no es arbitraria, ya que son dos sectores donde el Big Data tiene un papel muy importante y su uso está causando un gran impacto en la sociedad. Así, por ejemplo, el Big Data es utilizado en política para minar datos de diferentes personas e identificar posibles futuros votantes siendo así uno de los pilares más importantes en el diseño de estrategias de campaña electoral [2]. Esta metodología se ha mostrado muy eficiente y algunos ejemplos de esto los encontramos en el Brexit [5] o en la elección de Donald Trump como presidente de Estados Unidos. [6] Por otra parte, el uso de Big Data en seguridad nacional es utilizado para identificar posibles amenazas terroristas mediante información extraída de la población, aunque a menor escala también es utilizado para la prevención de crímenes en un área local [7].

Otro objetivo es demostrar lo fácil que es acceder a la información necesaria para cumplir los propósitos de los sectores mencionados anteriormente. Para ello, se ha procedido al desarrollo de un aplicativo capacitado para extraer información de un perfil de Facebook y mediante esta predecir la ideología política de una persona.

El presente documento está estructurado como sigue, en la sección 2 se encuentra el estado del arte y posteriormente en la sección 3 se presentan los objetivos del proyecto. Seguidamente, en la sección 4 se procede a explicar la metodología utilizada para el desarrollo del trabajo, junto con la planificación que se ha seguido. A continuación, se exponen de manera más detallada como se utiliza el Big Data en los sectores de la política y la seguridad nacional en los apartados 5 y 6 respectivamente. En la sección 7, se explica cómo se ha construido el aplicativo y los resultados que se han obtenido del mismo en la sección 8, para posteriormente en la sección 9 dar cabida a la reflexión ética del uso de Big Data en entornos profesionales. Finalmente, el trabajo concluye en la sección 10 donde se recogen las conclusiones finales de la realización del estudio.

## 2 ESTADO DEL ARTE

Previamente al análisis ético que ocupa el trabajo, se ha procedido al estudio del contexto de dos de los sectores donde el Big Data tiene un mayor impacto tanto en el propio sector, como en la sociedad. Estos sectores son la seguridad nacional y la política.

En la actualidad el Big Data en conjunción con otros elementos como el Machine Learning y el análisis predictivo son utilizados cada vez más para identificar patrones de actividad criminal [3]. La posibilidad de predecir futuros crímenes es una creciente área crucial de las fuerzas de la ley.

En este contexto existe una corriente de pensamiento que dictamina que los gobiernos y las instituciones tienen que saber que está pasando en todo momento con el objetivo de descubrir cómo se debe actuar. Los métodos que permiten realizar a las instituciones esta tarea normalmente involucran entidades centralizadas que brinden y faciliten servicios públicos. Sin embargo, en un mundo cada vez más conectado y globalizado estas estructuras se están desmoronando debido a que la información se encuentra cada vez más dispersa. Alex Pentland, una figura importante en la ciencia computacional americana sostiene que la variedad de problemas con los que se enfrenta la sociedad hoy en día (terrorismo, plagas, crisis económicas...) demuestra que se necesita una actualización en como utilizamos los datos [8]. La compartición de información entre distintas agencias de seguridad de manera más eficiente puede suponer un paso gigantesco en la lucha contra el terrorismo y permitiría encontrar a criminales y personas desaparecidas de manera más rápida. Si un país extranjero puede acceder a la información necesaria para encontrar un criminal buscado (foto, ADN, huellas dactilares) de otro país puede poner en marcha mecanismos que le permitan localizar a esta persona dentro de su frontera. Con las estructuras de datos correctas se podrían automatizar mecanismos de búsqueda de personas a nivel global.

Por otra parte, las fuerzas políticas, mediante el uso de esta misma herramienta, son capaces de diseñar estrategias que permitan llamar la atención de un nuevo público y así recoger nuevos votos. Uno de los ejemplos más importantes de estas estrategias lo encontramos en el Brexit. Este movimiento debe su éxito en gran parte al Big Data, ya que este le permitió acercarse a un electorado que no se veía representado por ninguna de las fuerzas políticas del momento y que representó un total de más de tres millones de votos [5].

## 3 OBJETIVOS

Tras un análisis global del tema tratado en el trabajo, se ha procedido a establecer los objetivos que se pretenden conseguir a lo largo de la realización de este. A continuación, se presentan dichos objetivos ordenados de forma ascendente en función de su prioridad dentro del marco del proyecto.

- En primer lugar, construir un marco teórico del uso de Big Data en entornos profesionales en forma de documentación. Para la construcción de dicho marco teórico, se ha procedido a la recolección de información del uso de Big Data en política y en seguridad nacional para posteriormente redactar un documento donde se recojan diferentes maneras de como se utiliza el Big Data en estos sectores y el impacto que tiene en los mismos.
- Seguidamente con la información recolectada, redacción de un ensayo ético donde se analice el uso de Big Data en entornos profesionales desde diferentes perspectivas, las cuales son el objetivismo, el utilitarismo y la filosofía kantiana. El objetivo final de la redacción de dicho ensayo es reflexionar sobre donde reside el límite que define el uso moralmente correcto o incorrecto del uso de Big Data en los sectores mencionados anteriormente.

- Finalmente, se pretende demostrar lo fácil que es utilizar la información de cualquier persona en estos sectores desarrollando un aplicativo que mediante información extraída de un perfil público de Facebook pueda predecir la ideología política del usuario haciendo uso de un modelo predictor desarrollado mediante Machine Learning.

## 4 METODOLOGÍA Y PLANIFICACIÓN

Tras la definición de los objetivos del proyecto se ha procedido a la elección de la metodología y definición de la planificación a seguir. Para la realización del trabajo se ha escogido como metodología a utilizar la metodología de Desarrollo Iterativo. Las diferentes etapas de las que se compone el trabajo han pasado por esta metodología. Para el desarrollo de la documentación (que incluye la documentación del uso de Big Data en seguridad nacional y en política en conjunción con el ensayo ético), se han agrupado en pequeñas etapas repetitivas (iteraciones) diferentes versiones de estos documentos con el objetivo en última instancia de que en cada iteración los documentos vayan mejorando hasta conseguir la versión final. Dicha metodología también se ha aplicado a la hora de desarrollar el aplicativo que se ha mencionado en el apartado de objetivos. Por lo tanto, en cada iteración se ha ido mejorando la aplicación y se le han añadido nuevas funcionalidades hasta conseguir la versión definitiva.

En cuanto a la planificación, esta sufrió algunas modificaciones a lo largo del transcurso de la primera fase del trabajo debido a la ligera modificación que sufrieron los objetivos en esta etapa. Si bien se ha construido el marco teórico que se compone de la documentación referente al uso del Big Data para posteriormente utilizarlo como base para la redacción del ensayo filosófico centrado en el punto de vista ético sobre el uso de Big Data en estos casos de uso, a priori se pretendía redactar un ensayo ético por tema tratado. El motivo de este cambio es que se consideró que un ensayo ético que abarque un espectro más amplio del uso de Big Data en entornos profesionales es más interesante que dos ensayos que traten los temas por separado. Por otra parte, el objetivo del aplicativo también sufrió cambios, pasando de ser un crawler que extraía información de los perfiles públicos de Facebook a un programa que utiliza la propia API de Facebook de tal manera que pueda ser procesada para determinar la ideología política de un usuario. Finalmente, por falta de tiempo el objetivo inicial de buscar información sobre el uso del Big Data en entidades financieras se desechó. La planificación original se muestra en la Tabla 1. Tras la modificación inicial de los objetivos la planificación se ha seguido correctamente tal y como se resume en la Tabla 2 donde se resumen los puntos de control definidos para el proyecto y las salidas que se han obtenido al final de cada uno de estos en conjunción con las fechas de entrega.

## 5 BIG DATA EN POLÍTICA

Uno de los usos más comunes del Big Data en política es como herramienta electoral. Para saber cómo las campañas electorales utilizan el Big Data como herramienta electoral se debe contestar en primer lugar por qué se necesita. El

TABLA 1: PLANIFICACIÓN ORIGINAL

Planificación		
Nombre	Salidas	Entrega
Entrega 1	Definición de planificación Definición de tareas	10/03/2019
Entrega 2	Documentación Seguridad Nacional Documentación uso de Big Data en bancos Concretar funcionalidad aplicativo Ensayo Ético para cada uno de los temas	14/04/2019
Entrega 3	Aplicativo funcional Documentación Big Data en política Ensayo ético restante	26/05/2019

TABLA 2: PUNTOS DE CONTROL

Planificación		
Nombre	Salidas	Entrega
Entrega 1	Definición de planificación Definición de tareas	10/03/2019
Entrega 2	Documentación Seguridad Nacional Documentación Big Data en política Concretar funcionalidad aplicativo	14/04/2019
Entrega 3	Aplicativo funcional Ensayo Ético	26/05/2019

principal objetivo del uso de Big Data en este contexto es proporcionar a los partidos políticos de una lista de ciudadanos que pueden representar potenciales votantes y a los cuales contactar para llevar a efecto tal objetivo. La pregunta que surge ahora es como se consigue esa lista de potenciales votantes. Los primeros integrantes de la lista son relativamente sencillos de identificar, ya que se tratan de personas que se han mostrado activas políticamente, es decir, han participado activamente en otras campañas y han donado a causas afines al partido. De esta manera, un partido puede establecer un primer perfil de persona que es afín a sus ideas y enviarle publicidad para que estas personas le voten en las siguientes elecciones. Esta estrategia es eficaz para mantener una lista de personas que probablemente ya te han votado, pero el uso potencial del Big Data en este paradigma es la recaptación de nuevos votantes. Cuando se habla de recaptación de nuevos votantes se habla de modelos predictivos que establecen una puntuación a cada persona registrada en las bases de datos de votantes. Esta puntuación indica que probabilidad hay de que un individuo determinado vote a un partido. Esto es algo muy potente, ya que mediante un análisis de las predicciones se puede establecer un perfil de votante que no se había tenido en cuenta y encontrar un conjunto demográfico objetivo al cual enviar publicidad para recolectar nuevos votantes [6]. De una manera similar, un partido político puede utilizar el análisis previamente mencionado para encontrar zonas donde el índice de futuros votantes se predice que será muy bajo o muy elevado. Con estos datos se puede optar por dos opciones diferentes, o bien invertir en las zonas donde te van a votar con más probabilidad o bien invertir en las zonas

donde se prevé una menor cantidad de votos para mitigar el daño. Si bien el Big Data puede determinar la estrategia electoral de un partido, su uso también puede influenciar en el programa político de un partido ya que de igual manera que un modelo predictivo es capaz de determinar si una persona es un potencial votante, otro modelo predictivo es capaz de determinar si una idea va a ser aceptada por el público general [2].

Los modelos predictivos utilizados en campañas electorales se dividen en dos grupos diferenciados entonces. El primero predice el comportamiento o actitud de posibles votantes de tal manera que se puede determinar si un determinado individuo tiene una alta probabilidad de votar a un partido y el segundo modelo predictivo predice como los votantes responderán ante ciertas campañas de divulgación. Para ambos casos las técnicas que se utilizan para lograr las predicciones son técnicas de regresión simples. Por lo tanto, la mayoría de los problemas de regresión que se tratan en este contexto son problemas de regresión lineal. La regresión lineal se define como la capacidad de un software de predecir una variable continua determinada mediante la definición de una recta cuya construcción se habrá visto influenciada por una serie de parámetros que tienen una estrecha relación con el valor que se quiere predecir. También hace falta contemplar problemas de regresión logística, que también ocupa un gran espacio de los problemas que se pretenden analizar en este contexto. A diferencia de su contraparte lineal, estos modelos no intentan predecir un valor continuo, sino que intentan resolver problemas de clasificación. Por lo tanto, la regresión logística consiste en predecir el resultado de una variable categórica en función de una serie de variables independientes o predictoras. El hecho de que estos modelos utilicen dichas técnicas los hace muy fácil de construir y adaptarlos a diferentes escenarios políticos. El uso de estos sistemas sin embargo trae consigo una serie de desventajas que hace falta considerar. En primer lugar, la utilidad de las técnicas que descubren correlaciones depende en gran medida del talento del analista de datos de la campaña en particular que las emplea. Un analista de datos de campaña capacitado con un gran número de personas que están familiarizadas con las propiedades disponibles en las bases de datos de votantes puede generar puntuaciones altamente exactas. Sin embargo, un analista de datos de campaña un poco menos capacitado podría generar predicciones que son solo ligeramente mejores que los métodos no predictivos que son solo ligeramente mejores que los métodos no sofisticados empleados en campañas anteriores. De esta manera los modelos predictivos en esta cuestión aún sufren de una alta dependencia de un factor humano que actúa como una suerte de cuello de botella para que estos modelos logren alcanzar cotas más altas. Otro fallo significativo de estos sistemas es que se construyen mediante un gran volumen de datos que, si bien son útiles para construir modelos predictivos eficaces a gran escala, no los prepara para construir modelos predictivos eficaces a pequeña escala. En este contexto algoritmos que hacen uso de Machine Learning, en concreto de aprendizaje supervisado han mostrado ser más eficaces. Estos algoritmos se basan en la experiencia y su aprendizaje se alcanza mediante la busca de relaciones entre ciertas acciones y las respuestas que generan. El aprendizaje automático supervisado incluye métodos tales como la clasificación y los árboles de

regresión. En un enfoque de regresión, el algoritmo crea un 'bosque' al dibujar una serie de árboles a partir de muestras a partir de datos existentes. Las muestras se dividen en una base de datos en función de dónde los parámetros discriminan mejor el resultado del interés. Es decir, los valores que muestran una mayor correlación con la variable que queremos predecir ayudan a establecer caminos de interés para su análisis en estos árboles. En entornos reales, la construcción de modelos predictivos mediante esta metodología requiere la construcción de varios modelos con árboles diferentes con el objetivo de contrastar las predicciones del modelo con datos reales para validarlos y al mismo tiempo descartar aquellos modelos que han mostrado una menor precisión en sus predicciones [2]. Un ejemplo de árbol predictor lo encontramos en la Figura 1, donde se muestra un árbol que predice la probabilidad de éxito de una idea.

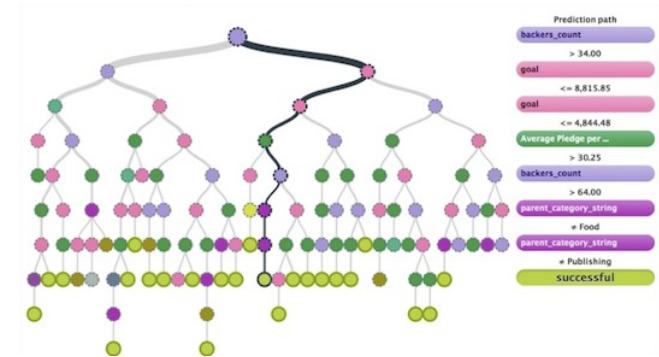


Fig. 1: Ejemplo grafo predictor

## 6 BIG DATA EN SEGURIDAD NACIONAL

En la actualidad el Big Data en conjunción con otros elementos como el Machine Learning y el análisis predictivo son utilizados cada vez más para identificar patrones de actividad criminal. La posibilidad de predecir futuros crímenes es un área crucial para las fuerzas de la ley. En el contexto de utilizar el Big Data como herramienta para luchar contra el crimen existe, tal y como se ha mencionado anteriormente una corriente de pensamiento que dictamina que los gobiernos y las instituciones tienen que saber qué está pasando en todo momento con el objetivo de descubrir qué se tiene que hacer. Es en este contexto donde se propone construir el denominado "Sistema Nervioso Digital", es decir un sistema que monitorice toda la actividad del planeta con el objetivo de recolectar información que prediga dónde se van a acometer acciones delictivas o encontrar fugitivos. Esto en el ámbito de la seguridad nacional es muy relevante ya que puede permitir que los gobiernos prevengan ataques terroristas o encontrar fugitivos con mayor facilidad. Sin embargo, la idea de contar algún día con un sistema nervioso digital parece difícil ya que las agencias gubernamentales son reticentes a compartir información con otras agencias de otros países [8]. Igualmente, se pueden ver ejemplos de usos similares del Big Data en áreas locales utilizados por la policía con resultados prometedores en el ámbito de la predicción del crimen como se puede apreciar en la Figura 2. Hay tres categorías de técnicas que los departamentos de policía utilizan para la predicción de crímenes: análisis de espacio, análisis de tiempo y análisis de redes sociales que

se proceden a explicar a continuación [9].

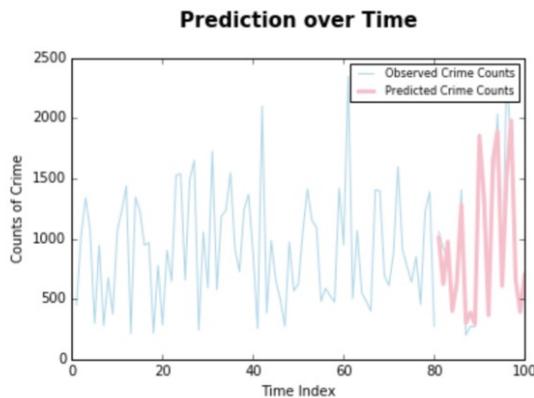


Fig. 2: Predicciones de delitos en Atlanta con IBM Watson Studio

## 6.1. Análisis de Espacio

Uno de los usos originales del mapeo de crimen es la identificación de puntos calientes delictivos, que son áreas en las que hay una mayor probabilidad de que un crimen ocurra. La definición de estos puntos calientes se realiza mediante el análisis de diversos factores como el histórico de criminalidad en dichas áreas (ya que, por lo que muestran diversos análisis estadísticos, es muy probable que, en un lugar donde ha ocurrido un crimen ocurra otro, que en algún lugar donde nunca han ocurrido) o los antecedentes penales de la gente que habita el área en cuestión. En un contexto retrospectivo, la detección de puntos calientes ha aumentado la comprensión de las características asociadas a los altos índices de criminalidad en áreas, tales como rutas de transporte, lugares de ocio, y espacios urbanos con una alta densidad de población. En términos de vigilancia predictiva, la detección de estos puntos calientes supone un factor clave en la asignación de recursos y en la toma de decisiones a corto plazo. Esto es debido a que los puntos calientes van a requerir una mayor inversión por parte de la policía y dicha inversión se verá reflejada en mujeres y hombres y coches patrulla para poder enfriar dichos puntos. Es importante tener en cuenta que un punto caliente no es un espacio definido, sino que su dimensionalidad es indefinida y puede variar a lo largo del tiempo expandiéndose o estrechándose según el caso, por lo tanto, los puntos calientes no tienen límites claramente definidos ya que la colocación de dichos límites se realiza de manera arbitraria (en función claro está de un marco de datos que ayuda a definir el punto caliente). La ubicación final, el tamaño y la forma de los puntos calientes son influenciados por juicios hechos por los analistas. Dichos juicios vienen determinados por los incidentes delictivos que se incluyen en el análisis, por las características utilizadas para definir los puntos calientes (incidentes criminales pasados, ambientes característicos asociados con el crimen, o ambas) mediante agrupación o concentración y el esquema de ponderación aplicado sobre incidentes penales pasados.

Estos análisis pueden llegar a ser automatizados por software con la capacidad de interpretar los datos. Esto permite conocer a la policía que sectores de la ciudad requieren un

cuidado más exhaustivo y por lo tanto instalar más unidades. Hay que destacar que, si bien esta información sirve para prevenir crímenes, el hecho de conocer los puntos calientes también aumenta la probabilidad de arrestos cuando se comete un crimen en esas zonas.

## 6.2. Análisis de Tiempo

Existen varios métodos estadísticos para analizar el clustering que están dirigidos en la identificación de áreas con altos niveles de criminalidad. En un contexto donde se pretende predecir donde ocurrirá un crimen, los métodos de agrupación detectan ubicaciones o áreas donde es probable que ocurra un crimen basado en donde ha ocurrido el crimen con anterioridad. Estos métodos, sin embargo, no aprovechan los patrones temporales en el crimen. A pesar de que algunos algoritmos de agrupamiento ponderan los eventos recientes más fuertemente, generando pronósticos, estos no ilustran cómo la incidencia del crimen cambia con el tiempo. La agrupación no aporta gran información sobre los movimientos en la actividad delictiva. En este mismo contexto encontramos algunas tareas que exigen atención a los patrones temporales. Si un departamento de policía ha observado una erupción de robos y está tratando de predecir el siguiente incidente en la cadena, es fundamental identificar los patrones en la trayectoria de robos tomada por un presunto delincuente para averiguar dónde ocurrirá el siguiente. Estas tareas pueden ser realizadas con sistemas de software capaces de interpretar los datos o bien por analistas.

## 6.3. Análisis de Redes Sociales

Esta categoría se utiliza principalmente para detectar personas de interés, es decir, personas que por su perfil en las redes sociales se ha detectado que es probable que cometan un crimen. Mediante bases de datos, la policía puede identificar a individuos que son centrales para organizaciones criminales como bandas, organizaciones terroristas y redes de tráfico de drogas entre muchas otras. La relevancia de las redes sociales para el análisis criminológico es bien establecida. La delincuencia organizada, como el narcotráfico, la violencia de pandillas, y el robo en serie, requiere la creación y mantenimiento de diversas relaciones. Una red de tráfico de drogas, por ejemplo, puede incluir proveedores, distribuidores, contrabandistas, compradores, y blanqueadores de dinero. Además, las redes criminales son incrustadas en el contexto social en el que operan; en ellas se nutren y buscan captar nuevos clientes. Las agencias policiales pueden utilizar herramientas para mapear estas personas. En las aplicaciones de lucha contra el crimen, el análisis de redes sociales es frecuentemente utilizado para identificar nodos centrales, individuos que tienen un alto nivel de conectividad dentro de la red. Usando medidas de centralidad, un analista puede identificar individuos de interés en el contexto de un problema dado. Si se busca adquirir información sobre una red sin desmontarla, esto se puede lograr, contactando a un actor con un alto nivel de cercanía. Alternativamente, un objetivo podría insertar información en una red, algo que podría lograrse mejor utilizando un actor con una alta influencia en el círculo de la propia red. En conclusión, encontrando personas de interés y estudiando sus

perfíles en las redes sociales se puede lograr a discernir con mayor o menor precisión la estructura de una red criminal y las personas que la componen. Incluso se pueden incautar dichos perfíles con el objetivo de buscar pruebas incriminatorias o de descubrir redes criminales que hasta ahora se consideraban desconocidas.

## 7 DESARROLLO DEL APlicativo

En las anteriores secciones se han expuesto algunos ejemplos de uso de Big Data en entornos profesionales. A continuación, se va a proceder a la explicación del aplicativo realizado que funciona como un caso práctico de como se utiliza el Big Data en política para la recaptación de nuevos votantes. En términos generales el funcionamiento del aplicativo se resume en la extracción de datos de usuarios de Facebook para posteriormente crear un modelo predictor (mediante la API de BigML) que sea capaz de predecir a qué ideología política pertenece una persona en base a sus datos. A continuación, se recogen una serie de apartados que explican el diseño y la funcionalidad del aplicativo en cuestión.

### 7.1. Entorno Virtual

El aplicativo por si solo no es funcional, si no que se requiere previamente la instalación de diversas librerías que permitirán que este se ejecute sin ningún error. Dado que a priori, solamente será nuestro aplicativo quien haga uso de estas librerías, es altamente recomendable instalar estas en un entorno virtual que será utilizado exclusivamente por el programa en cuestión. A continuación, se proceden a exponer cuáles son esas librerías y porqué son necesarias.

- **Facebook-SDK:** Librería que encapsula los métodos de la API Facebook Graph API. Estos métodos serán utilizados a lo largo del flujo de ejecución del programa para extraer información de distintos usuarios de Facebook. Para extraer información de diferentes usuarios, el uso de esta librería requiere de un token de acceso por cada usuario. La obtención de dichos tokens no es fácil de conseguir y requiere que estos usuarios den su permiso de acceso y que tengan una aplicación del desarrollador instalada. Por lo tanto, en un escenario real, cuando un usuario da su permiso para que una aplicación acceda a sus datos y se la instala, los desarrolladores de dicha aplicación pueden acceder a sus datos de perfil, pero no antes de que el usuario de permisos. Si se quisiera entonces que el aplicativo trajera datos de una cantidad notable de usuarios, se requeriría de una subida de la aplicación a la app store de Facebook y que estos se instalarán la aplicación para obtener sus tokens de acceso respectivos. Dado que realmente no se quiere subir una aplicación sin funcionalidad a la app store, se ha procedido a extraer los datos de perfil del desarrollador y usar estos como guía para simular la extracción de datos de personas diferentes de tal manera que se puede replicar el escenario deseado para la creación de un modelo predictor sin tener que recurrir a la burocracia de subir una aplicación a la app store y solicitar permiso a usuarios. En este aspecto es necesario destacar que la página de desa-

rrollo de aplicaciones de Facebook permite la creación de usuarios de prueba que simulan usuarios reales pero la información que se puede extraer de estos es muy pobre (nombre, apellidos y dirección de correo) y no se considera útil para el objetivo del aplicativo.

- **JSON:** La información que se extrae de los usuarios viene en forma de diccionario, algo que resulta extremadamente útil para tratar los datos en Python. Sin embargo, dado a los problemas que se han explicado en el apartado anterior, la extracción de datos de usuarios será simulada. Los datos de estos usuarios ‘inventados’ se encuentran almacenados en distintos ficheros json que serán procesados para la obtención de la información y posteriormente para la creación del modelo predictor. Es en este contexto donde entra en juego la librería json, ya que será esta la que permitirá trabajar con este tipo de ficheros.
- **CSV:** La API de BigML es capaz de crear un modelo predictor, pero para poder realizarlo, se necesita un fichero de tipo csv que contenga datos de los distintos usuarios. Mediante dicho fichero, el modelo predictor será capaz de establecer relaciones entre diferentes variables y la ideología política del sujeto. El uso de esta librería nos va a permitir pasar los datos procesados de los json que se mencionan en el apartado anterior a un csv que pueda ser procesado por BigML.
- **BigML:** Esta librería encapsula los métodos de la API de BigML que serán utilizados a lo largo del flujo de ejecución del aplicativo para generar un modelo predictor en base a unos datos almacenados en un csv y posteriormente generar una predicción en base a unos datos de un usuario nuevo.

### 7.2. Clases

El programa no se compone de una gran complejidad, y su funcionalidad se ve distribuida en 3 clases diferente cuyas funciones y propiedades se proceden a explicar en cada uno de los apartados que componen esta sección y las diferentes realaciones entre ellas se pueden apreciar en la Figura 3.

- **FacebookExtractor:** Tal y como indica su nombre esta clase será la encargada de extraer datos de los usuarios y almacenar la información extraída en una estructura de datos que pueda ser utilizada por otra clase para lograr el objetivo final del programa.
- **generateCSV:** Esta clase se encarga de la construcción del csv que a posteriori será utilizado para la construcción del modelo predictor.
- **Machine:** Esta clase es la encargada de conectarse a la API de BigML y mediante el csv generado por la clase del apartado anterior, crear un modelo predictor que es capaz de predecir la ideología política de un nuevo usuario.

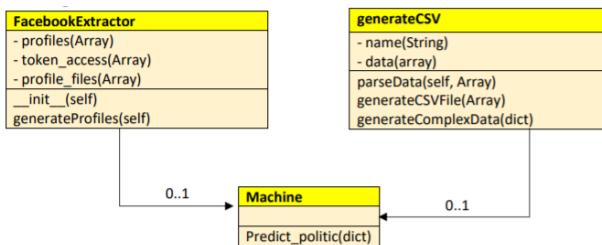


Fig. 3: Diagrama de clases

## 8 RESULTADOS DEL APLICATIVO

Tras la ejecución del aplicativo se obtiene por consola los resultados de dicha ejecución de forma que se muestra por consola las propiedades del nuevo usuario y la ideología política que se ha predicho. De esta manera para el usuario descrito en la Figura 4, obtenemos que el programa predice que es de izquierdas. El motivo de que el aplicativo

```

{
  "age": "27",
  "team": "Barça",
  "gender": "Male",
  "movies": [
    "1833108910253271", "105687816178274", "273894216403422", "57166529496",
    "298908497110378", "1111111111111"
  ],
  "television": [
    "141676752595826", "1180269352032163", "615555128584174",
    "1555156204781510", "1653393938209126", "339799489548766"
  ],
  "books": [
    "12548837258889", "41752661163558678"
  ],
  "music": [
    "39393223796", "7619396355", "2383886528778", "9770929278", "19691681472"
  ],
  "hometown": "Barcelona, Spain"
}
  
```

Fig. 4: Propiedades usuario de prueba

haya determinado este resultado está estrechamente relacionado a que el 80 % de los usuarios mediante los cuales se realiza el aprendizaje son partidarios de esta ideología, de esta manera el aplicativo ha encontrado más fácilmente características similares en usuarios de izquierdas que no de su contraparte. Es interesante probar el aplicativo con ligeras modificaciones en la base de datos de los usuarios, así por ejemplo si modificamos la base de datos de tal manera que el reparto sea más equitativo el resultado puede variar ya que encontrará más similitudes con usuarios de derechas que con el anterior caso algo que efectivamente ocurre ya que el nuevo resultado es que el usuario es de derechas. En este caso el resultado ha variado y el aplicativo ha encontrado más similitudes con usuarios de derechas que no de izquierdas. Finalmente, si adulteramos más la base de datos para que la mayoría de los usuarios sean de derechas los resultados obtenidos no varían respecto al anterior caso ya que el nuevo usuario sigue teniendo más propiedades en común con usuarios de derechas. De esta manera se puede apreciar que el proceso de aprendizaje del aplicativo funciona

correctamente y si se dispusiera de un conjunto de usuarios más grande los resultados obtenidos serían aún más significativos.

## 9 ENSAYO ÉTICO

En el mundo actual todo está conectado. Cada segundo que pasa se generan toneladas de información que son almacenadas en diferentes lugares del mundo. Este conjunto de información es lo que se denomina como Big Data. El Big Data ha supuesto una revolución en la vida de las personas, posibilitando la era de las redes sociales que al mismo tiempo ha posibilitado que cambie la manera de comunicarse y la interacción entre personas. Gracias a esto han nacido en las últimas décadas gigantes tecnológicos que han encontrado en este ecosistema la manera de ganarse la vida y ofrecer nuevos productos a sus clientes. Evidentemente el Big Data ha supuesto un gran impacto en el mundo de los negocios, pero tampoco debemos olvidarnos lo que ha supuesto a nivel de inteligencia (es decir, para organizaciones tales como la CIA o la NSA entre muchos otros). El Big Data ofrece un gran poder que hasta entonces era desconocido y en algunos de los ejemplos que se han expuesto en este documento (ver secciones 5 y 6) se han podido apreciar algunos casos donde este poder es utilizado de manera correcta o incorrecta. Sin embargo, saber si el uso es correcto o incorrecto desde un punto de vista ético no es algo fácil de discernir ya que según la manera de pensar de las personas te puedes encontrar con diversas respuestas diferentes. Este dilema ocurre a raíz de que no siempre es fácil diferenciar entre un acto moralmente correcto y otro incorrecto. Es en este marco donde entran las teorías éticas. Estas teorías éticas son teorías que nos ayudan a justificar nuestros actos y nos sirven de guía moral en la vida. En el caso del uso del Big Data en entornos profesionales se pueden encontrar diversas teorías éticas que justifican el uso del Big Data de una manera u otra. En este documento en concreto se exponen tres teorías éticas notablemente aceptadas por la sociedad, aunque debido a que el Big Data es un elemento que ha parecido recientemente, ninguna de estas teorías éticas hace referencia explícitamente al Big Data y las conclusiones que se extraerán en este documento son solo interpretaciones personales de estos movimientos en este contexto.

### 9.1. Implicaciones Éticas

En el mundo empresarial, las empresas que trabajan con Big Data se deben enfrentar a una serie de debates éticos que se exponen a continuación [10]:

- **Contexto:** Este debate hace referencia a la finalidad para la cual se utilizan los datos. Un ejemplo de este tipo de debates lo podemos encontrar en entidades financieras donde la información recogida puede ser utilizada para decidir las inversiones a realizar. Estas inversiones pueden hacer referencia a grandes movimientos en el sector bancario al mismo tiempo que pueden hacer referencia a la aprobación de préstamos. En este último, se utilizan datos tales como la liquidez y el historial financiero para determinar las probabilidades de que un cliente devuelva un préstamo.

■ **Consentimiento y elección:** En este marco la empresa se debe preguntar si ha adquirido los datos de manera legítima. Se entiende manera legítima en este caso cuando el usuario que ha aceptado ceder sus datos tenía la opción de declinar y ha entendido lo que ha aceptado haciendo que la cesión de datos sea una decisión suya y no una obligación. Ejemplos de este contexto los encontramos al registrarte en una página donde antes de confirmar el registro se debe leer y firmar un contrato. Por otra parte, se podría considerar como obtención de datos de manera ilegítima en los casos en los cuales se roban o bien mediante web scrapping se recuperan datos personales sin previo consentimiento de las partes afectadas.

■ **Razonable** Este apartado hace referencia a si la cantidad de datos y la información que contienen es razonable en relación al uso que se le va a aplicar. Por ejemplo, no es razonable conocer el estado civil de los usuarios cuando se pretende conocer la cantidad de personas que están utilizando una plataforma.

■ **Substancia:** En este marco se plantea la pregunta de si las fuentes de información usadas son apropiadas, autoritarias, completas y adecuadas en el tiempo para la aplicación. Se puede juzgar como moralmente incorrecto que una organización haya adquirido datos de una fuente inapropiada como por ejemplo una organización de crackers que han robado los datos en cuestión.

■ **Justicia:** Reflexiona sobre como de equitativo son los resultados de la aplicación a todas las partes implicadas.

■ **Consideración:** Este debate hace referencia al uso final que se le va a dar al aplicativo que utiliza estos datos. En este contexto la organización se debe preguntar como encaja su aplicación desde un punto de vista ético.

■ **Responsabilidad:** Hace referencia a como se detectan y se reparan los errores y las consecuencias que pueden tener dichos errores. Uno de los puntos clave del uso del Big Data es que puede provocar grandes errores y eso en un mundo globalizado como el actual puede costar una crisis que afecte a mucha gente de manera negativa.

Estás son algunas de las implicaciones éticas con las que debe convivir una empresa que trabaje con información y tecnología. Toda organización que encaje con esta descripción debe saber moverse por estas implicaciones no simplemente por un tema de moralidad sino también por un tema de legalidad. Esto es representado de manera gráfica en la Figura 5.

## 9.2. Teorías Éticas

Tal y como se ha mencionado en apartados anteriores, las teorías éticas son normas de conducta que intentan resolver las cuestiones que se plantean en los dilemas éticos. Sin embargo, existe una gran diversidad de teoría éticas que dan respuestas diferentes a un mismo dilema ético. En el caso que nos ocupa, es decir, el uso del Big Data en entornos

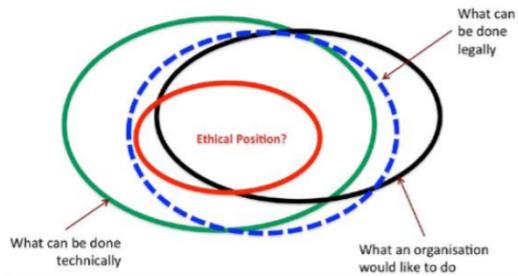


Fig. 5: Diagrama de Venn y moral en el uso de Big Data

profesionales es uno de esos escenarios precisamente, donde depende de la teoría ética que se decida poner en práctica se tendrá una respuesta u otra. Normalmente si preguntas a cualquier persona que si robar para dar de comer a un niño que está hambriento es correcto te contestará que sí. Sin embargo, si le preguntas si está bien matar a miles de personas para salvar a millones la respuesta no es tan sencilla. Existen diversas teorías éticas que dan respuesta a este dilema, pero destacan dos extremos en concreto, el utilitarismo y la filosofía kantiana.

El utilitarismo es una filosofía que define el valor moral de una acción en base a sus consecuencias y el valor que aportan a la sociedad [11]. Poniendo en este paradigma el ejemplo anterior se puede afirmar que un utilitarista clásico estaría de acuerdo con sacrificar miles de vidas para salvar a millones. Si ponemos en práctica dicha ideología en el contexto del Big Data, el utilitarismo está a favor de no poner restricciones al uso del Big Data ya que el quitar estas restricciones, el Big Data ofrece un mayor beneficio a la sociedad. Por ejemplo, se podrían eliminar las restricciones sobre la privacidad para permitir a las agencias gubernamentales identificar amenazas terroristas de una manera más eficiente e incluso capturar a criminales de forma más rápida. Entonces la acción de que un gobierno espíe a sus ciudadanos es una acción moralmente correcta ya que es una acción que trae consigo un gran beneficio para la sociedad.

En este tipo de ejemplos, se pone en práctica el eterno debate entre libertad y seguridad, un debate que hasta la fecha de hoy no está cerca de ser resuelto. Siempre que ocurre un ataque terrorista la primera reacción de los países es atribuirse más poder y eliminar restricciones que dificulten la ocurrencia de estos hechos. Ejemplos de esto lo encontramos en la ley patriota que lanzó Estados Unidos tras los sucesos del 11 de septiembre de 2001. Dado que el uso de Big Data sin restricciones ofrece grandes beneficios, un juicio utilitarista suena realmente bien a primera vista, sin embargo, dejemos que entre al campo la filosofía kantiana. La gran teoría ética en la que se basa gran parte de la teoría de lo moralmente correcto se define en la siguiente frase: '*Obra sólo según aquella máxima por la cual puedes querer que al mismo tiempo se convierta en ley universal*' [12]. Lo que viene a decir esta frase es que cuando una persona decide seguir una regla personal es porque cree que el resto de las personas también deben seguirla. Por ejemplo, si esa persona cree que cuidar de los mayores es moralmente correcto es porque piensa que todos deberían hacerlo. En el caso del Big Data y repitiendo el ejemplo del espionaje, si se considera que espiar a personas por parte de un gobierno es una acción moralmente correcta como dicen los utilitaristas entonces se considera como moralmente correcto también

que otras personas o gobiernos espíen a las personas que están de acuerdo con esta afirmación. Ahora bien, si le preguntas a la gente si está de acuerdo con que otras personas desconocidas tengan información confidencial sobre ellas lo más probable es que te digan que no. Sin embargo, cabe la posibilidad de que una persona crea que la cesión de datos sin excepción debería ser una ley universal debido a que considera que dicha acción merece la pena ya que trae consigo un gran beneficio para la sociedad. En este contexto podemos encontrar como confluyen al mismo tiempo ideas de Kant así como del utilitarismo como muestra de la gran complejidad que existe en este escenario. Sin embargo, se reitera la idea de que la mayoría de las personas no estarían dispuestas a ceder sus datos personales (por simple sentido común) y por la regla de Kant sería moralmente incorrecto que esas personas decidan utilizar datos sensibles con cualquier propósito ya que sería un acto hipócrita. Sin embargo, esta afirmación no se puede convertir en una ley universal ya que como se ha visto, las máximas que se quieren convertir en leyes universales dependen de la persona en cuestión.

Otro ejemplo del uso de Big Data en entornos profesionales en el cual se puede entrever un dilema ético es el de uso de este elemento en la política. En la actualidad tal y como se explica detalladamente en el apartado de Big Data en política, el Big Data es utilizado para identificar a potenciales votantes nuevos. Esta técnica ha llevado a que se lleven a efecto resultados inesperados como el Brexit o la elección de Donald Trump. Este tipo de factura plantean la pregunta de si se ha Hackeado la sociedad. En este aspecto las teorías éticas anteriormente mencionadas no pueden ofrecer una respuesta concreta ya que en función de la ideología política de la persona que ponga en práctica dichas teorías se encuentra una respuesta u otra, es decir, un utilitarista considera como una acción moralmente correcta el uso de datos personales para manipular elecciones si el partido o movimiento político que se ve beneficiado por dichas manipulaciones es un movimiento que aportará beneficio a la sociedad. Sin embargo, las personas que se familiarizan más con las ideas de Kant dirán todo lo contrario.

Estos análisis éticos ofrecen una perspectiva interesante, pero conviene antes de concluir este ensayo mirar otra perspectiva. Hasta ahora se ha analizado éticamente el comportamiento de entidades abstractas, es decir, partidos políticos o agencias gubernamentales. A continuación, se va a proceder a analizar el uso de Big Data en entornos profesionales desde un punto de vista del individuo, por lo tanto, aquella persona cuyos datos van a ser utilizados. Para este análisis, se va a poner en práctica un movimiento filosófico que ha sido y es en la actualidad uno de los más influyentes, el objetivismo. Una muestra de esta influencia la podemos ver en el número de copias vendidas de la obra que sirve como mayor expositor de este movimiento, La Rebelión de Atlas, escrita por la propia Ayn Rand autora del propio movimiento. En una encuesta que la Biblioteca del Congreso de Estados Unidos realizó en 1991 fue considerado el más influyente en la vida de los estadounidenses, sólo superado por la Biblia. El objetivismo es un movimiento filosófico que se centra en un individualismo radical y que basa todas sus teorías éticas en la razón [13]. Para entender la postura que tiene este movimiento en los debates anteriormente expuestos se debe entender en primera instancia la definición de ética que ofrece este movimiento. En concreto, la ética del objetivis-

mo se fundamenta en un egoísmo racional, es decir, que, por su condición de personas, cada individuo tiene que definir y procurar sus propios valores, y cada uno ha de ser beneficiario de sus propias acciones. En esta definición podemos encontrar ramificaciones tanto del utilitarismo como de la filosofía de Kant ya que una acción moralmente correcta es aquella que no compromete los valores del propio individuo y a la vez aporte un beneficio (es decir, una utilidad) al mismo. Entonces un objetivista verá como moralmente correcto que utilicen sus datos para capturar terroristas o prevenir ataques de la misma índole siempre y cuando esta acción no entre en conflicto con su propio código moral. En cuanto al segundo ejemplo, se encuentra el mismo razonamiento. Por lo tanto, se puede justificar como moralmente correcto que un partido político utilice datos personales (con previa autorización de los afectados) si esta acción no entra en conflicto con el código moral y supone un beneficio para los dueños de dicha información.

### 9.3. Dataísmo

El Big Data es un elemento controvertido que ha conformado una multitud de opiniones al respecto dentro de la cúpula de pensadores que existe en el presente. Dentro de esta cúpula existe una corriente que considera que la existencia del Big Data es perjudicial para la humanidad y se contraponen al Dataísmo. El Dataísmo se describe como una filosofía (incluso algunos la describen como religión) creada a partir del significado emergente del propio big data, la inteligencia artificial y el internet de las cosas. Dicha filosofía defiende que el flujo de información es el valor supremo y la libertad de la información es el mayor bien de todos, factor que coincide con una mentalidad utilitarista. Uno de los mayores antagonistas a este movimiento es el filósofo Byung-Chul Han, que argumenta que los datos están llevando a la sociedad a una decadencia cada vez más extrema ya que según él, el dataísmo renuncia totalmente al sentido que radica en la narración, haciendo así un paralelismo con el nihilismo[14]. La conclusión a la que llega este autor en su obra Psicopolítica es la siguiente: '*Los dataístas copulan con datos. Así, se habla de «datasexuales». Son «inexorablemente digitales» y encuentran los datos «sexys». El dígito se aproxima al falo.*'

Una de las figuras que destacan en el marco de opiniones del dataísmo es Yuval Noah Harari que argumenta que las estructuras políticas o sociales competidoras pueden ser vistas como sistemas de procesamiento de datos. De esta manera, Horari plantea que se puede interpretar que toda la especie humana es un solo sistema de procesamiento de datos siendo cada uno de los seres humanos un chip. El conjunto total de la historia del ser humano puede leerse como un proceso de mejora de la eficiencia de este sistema incrementando el número y variedad de procesadores/chips del sistema, incrementando el número de conexiones entre procesadores e incrementando la libertad de movimiento junto con las conexiones existentes. Según Horari, llegará un momento en que la humanidad cederá poderes de decisión a los algoritmos, ya que estos en un futuro mediante el Big Data serán capaces de conocer a una persona mejor que la propia persona en sí misma. El autor concluye que el dataísmo tiene mandamientos prácticos como cualquier otra religión.

Cabe destacar que el dataísmo es una postura filosófica

válida como cualquier otra y que al igual que tiene detractores también tiene a defensores. Una muestra de estos defensores se puede apreciar en los movimientos hacktivistas personificados en figuras como Julian Assange o Aaron Swarth (quien se suicidó y es considerado como el primer mártir del dataísmo según el propio Horari).

## 10 CONCLUSIONES

El Big Data es un elemento que ha venido para quedarse en la vida de las personas, pero su uso evidentemente puede suponer conflictos morales por la información sensible que se maneja, así como las decisiones que se toman a través de ella. En los ejemplos que se han mencionado anteriormente se exponen escenarios donde la moralidad del uso de este tipo de información se pone en duda precisamente por la excepcionalidad de los escenarios. En el caso del uso del Big Data en seguridad nacional se puede apreciar que el uso de esta herramienta realmente ayuda a la disminución del índice de criminalidad ya que las técnicas vistas ayudan a los cuerpos de seguridad en sus tareas para atrapar criminales y evitar que ocurran crímenes. Sin embargo, su uso indiscriminado puede suponer una ruptura con la confidencialidad de la información. En cuanto al escenario relacionado con la política se puede apreciar como el Big Data puede ser utilizado para manipular a un sector importante de la población planteando la cuestión de si realmente se pueden hackear unas elecciones. Aquí el uso del Big Data para tal finalidad colisiona con lo que se podría considerar un uso moral de la misma, sobre todo teniendo en cuenta el éxito que ha provocado en campañas muy polémicas como el Brexit o la elección de Donald Trump como presidente de Estados Unidos. Tal y como se expone en el apartado denominado Ensayo Ético, no existe una ley universal que defina que es ético y que no lo es, sobre todo en un escenario tan complejo como este. En este mismo ensayo se exponen diversos movimientos que tienen sus propias teorías éticas y por lo tanto tienen enfoques distintos. Por ejemplo, un enfoque utilitarista considera como moralmente correcto el uso indiscriminado de esta tecnología sin importar que se rompa la privacidad de las personas ya que esta acción es la que aporta, a priori, un mayor beneficio para la sociedad. En cambio, un enfoque Kantiano pone en duda esta afirmación ya que según las teorías éticas de este movimiento no es moralmente justificable el uso indiscriminado de Big Data, ya que una persona se debe regir por aquella máxima por la cual pueda querer que al mismo tiempo se convierta en ley universal y la mayoría de las personas no estarían de acuerdo en que espiar sea una acción ética. La conclusión del alumno en esta cuestión es que el Big Data se debe utilizar en beneficio de la humanidad, pero con ciertas restricciones, pero el saber donde reside ese límite solo se encuentra en el corazón de la humanidad.

## AGRADECIMIENTOS

La realización de este proyecto no hubiera sido posible sin la colaboración y la atención del tutor Rubén Rubio así que en primer lugar le quiero dar las gracias a él. En segundo lugar, pero no por ello menos importante le quiero dar las gracias a mis compañeros que me han dado conse-

jos (como utilizar la API de Facebook en lugar de realizar un crawler desde cero) que han permitido que el acabado del trabajo sea de mayor calidad. Entre ellos quiero destacar a Jose María Malaguilla que al haber realizado el TFG en el primer semestre ha sido la persona que me ha aportado consejos más útiles.

También quiero hacer una especial mención a Frank Miller, escritor que con sus obras me inspiró a investigar sobre el objetivismo y Ayn Rand que con su libro la Rebelión de Atlas me motivó a expandir mis horizontes en el ámbito de la filosofía y a coger este proyecto.

## REFERENCIAS

- [1] AGUILAR, Luis Joyanes. Big Data, Análisis de grandes volúmenes de datos en organizaciones. Alfaomega Grupo Editor, 2016.
- [2] NICKERSON, David W.; ROGERS, Todd. Political campaigns and big data. *Journal of Economic Perspectives*, 2014, vol. 28, no 2, p. 51-74.
- [3] VLAHOS, James. The department of pre-crime. *Scientific American*, 2012, vol. 306, no 1, p. 62-67.
- [4] CADWALLADR, Carole. The great British Brexit robbery: how our democracy was hijacked. *The Guardian*, 2017, vol. 7.
- [5] GONZALEZ, Roberto J. Hacking the citizenry?: Personality profiling, 'big data' and the election of Donald Trump. *Anthropology Today*, 2017, vol. 33, no 3, p. 9-12.
- [6] ANSOLABEHÉRE, Stephen; HERSH, Eitan. Validation: What big data reveal about survey misreporting and the real electorate. *Political Analysis*, 2012, vol. 20, no 4, p. 437-459.
- [7] PAUL, Anand, et al. Smarter cities series: Understanding the IBM approach to public safety. IBM Corporation, 2011.
- [8] PENTLAND, Alex. Social physics. Penguin, 2014
- [9] BACHNER, Jennifer. Predictive policing: preventing crime with data and analytics. IBM Center for the Business of Government, 2013.
- [10] HERSCHEL, Richard; MIORI, Virginia M. Ethics & big data. *Technology in Society*, 2017, vol. 49, p. 31-36.
- [11] MILL, John Stuart. Utilitarianism. En *Seven masterpieces of philosophy*. Routledge, 2016. p. 337-383.
- [12] ROSS, William David. Kant's ethical theory. Oxford: Oxford University Press, 1954.
- [13] A. Z. Rosenbaum, La Rebelión De Atlas, Ed.1.1. Estados Unidos: Titivillus 1957
- [14] HAN, Byung-Chul. Psicopolítica: neoliberalismo y nuevas técnicas de poder. Herder Editorial, 2014.
- [15] HARARI, Yuval Noah. Homo Deus: breve historia del mañana. Debate, 2016.