

Aplicación web para identificar personalidad, género y edad de usuarios en Twitter

Janet V. Hernández-García¹, Gabriela Ramírez-de-la-Rosa¹,
Esaú Villatoro-Tello¹, Héctor Jiménez-Salazar¹, Verónica Reyes-Meza²

¹ Universidad Autónoma Metropolitana Unidad Cuajimalpa,
Departamento de Tecnologías de la Información,
División de Ciencias de la Comunicación y Diseño,
México

² Universidad Autónoma de Tlaxcala,
Centro Tlaxcala de Biología de la Conducta,
México

2113066463@alumnos.cua.uam.mx,
{gramirez,evillatoro,hjimenez}@correo.cua.uam.mx, vrmeza@gmail.com

Resumen. Twitter es una fuente de información para muchas tareas del procesamiento del lenguaje natural. Particularmente para tareas de perfilado de autores, es decir, la tarea de determinar mediante el texto de un autor características demográficas de éste, por ejemplo, género, edad y personalidad. Tradicionalmente este problema se resuelve mediante un enfoque de clasificación supervisado. Uno de los mayores problemas de este enfoque es la necesidad de datos etiquetados (ejemplos) para generar clasificadores confiables; y la complejidad para encontrar datos etiquetados depende del problema. Mientras que es relativamente sencillo obtener información del género, es más complicado obtener información sobre la personalidad de un usuario de Twitter. En este contexto, se presenta una aplicación web que identifica, mediante el análisis del texto de los tuits de un usuario, su perfil que consta de edad, género y rasgos de personalidad. El principal objetivo de la aplicación es que estos usuarios permitan validar la información del perfil generada mediante la respuesta a un cuestionario de personalidad para expandir el conjunto de datos etiquetados. Por otro lado, la aplicación propuesta usa una representación basada en grafos para entrenar modelos para cada problema; esta representación, hasta nuestro conocimiento, no ha sido empleada en clasificación no temática de textos.

Palabras clave: Perfilado de autores, representación basada en grafos, procesamiento del lenguaje natural, aprendizaje supervisado.

Web Application for Analyzing Personality, Genre and Age of Twitter Users

Abstract. Twitter is a very useful information source for diverse natural language processing tasks. Particularly, for the author profiling task, *i.e.*,

identifying demographic aspects of an author based on the text that he/she writes. Normally this problem has been solved by means supervised classification techniques. However, one of the biggest problems with this approach is the lack of labeled data (training examples) to build reliable classification models. The complexity of collecting labeled data depends on the problem; for instance, it is easier to find labelled data for genre than data for personality traits. On this regard, we present a web application to identify the profile from a given user through the analysis of his/her tweets. The main goal of the proposed application is that users of our application can validate the results presented to them by answering a personality test, this aiming to expand our labeled corpus. Additionally, our application uses a graph based representation to build the classifiers which, to the best of our knowledge, it has not been used for non thematic clasification problems.

Keywords: Author profiling, graph based representation, natural language procesing, supervised learning.

1. Introducción

Las redes sociales son una de las mayores fuentes de información en Internet a nivel mundial. De acuerdo a Statisticbrain³, al 1 de diciembre de 2015, el número de usuarios activos de Facebook, la red social con mayor número de usuarios, es de 1374 millones de usuarios; por otra parte Twitter tiene 289 millones de usuarios activos generando un promedio de 58 millones de tuits al día.

Todos estos usuarios de redes sociales generan grandes cantidades de diversos tipos de contenidos como: videos, fotografías, *posts*, revisiones de productos, etc. Utilizando esta información disponible en Internet, se han tratado de resolver problemas de diversos orígenes, por ejemplo: se ha tratado de identificar el estado de ánimo de las personas [10], predecir las fluctuaciones en la bolsa de valores [4], identificar a pedófilos en sitios web de conversaciones [7,21], generar perfiles de usuario [22], entre muchos otros.

En particular, en años recientes, un problema que ha llamado la atención de investigadores en el área de Procesamiento del Lenguaje Natural es el problema denominado *perfilado de autor*, es decir, el problema de identificar, a través del texto que alguien escribe, características demográficas del autor de ese texto; los rasgos más comúnmente identificados son género y rango de edad [11,3,5]. Sin embargo, existen otros aspectos demográficos que son de interés no solo a la comunidad de computación, sino también a áreas de las Ciencias Sociales, particularmente a la Psicología; este aspecto es conocido como la identificación de rasgos de personalidad, mismo que se considera una dimensión más al problema de perfilado de autor.

En este contexto cabe destacar el papel del PAN, quien organiza una serie de eventos científicos y competencias internacionales sobre textos digitales

³ <http://www.statisticbrain.com/social-networking-statistics/>

orientados a problemas forenses⁴. En el caso del perfilado de autores, desde el 2013 a la fecha se han enfocado a los problemas de identificación de género y edad en redes sociales. Solamente en el 2015 se contempló la dimensión de rasgos de personalidad. Esto, además de mostrar lo novedoso del problema de identificación de personalidad en redes sociales, habla también de lo difícil que ha resultado generar modelos confiables, pues bajo un esquema supervisado, se necesitan datos etiquetados. Estos datos etiquetados son particularmente difíciles de encontrar para el caso de la identificación de la personalidad de usuarios en redes sociales.

Por lo anterior, en este trabajo se plantea realizar una aplicación web que cumpla dos funciones: *i*) proporcionar un modelo, basado en los datos del PAN-2015, para la identificación de la personalidad, género y edad de usuarios de Twitter, mediante un enfoque de representación basada en grafos; y *ii*) utilizar la plataforma para validar datos mediante un módulo de evaluación, donde los usuarios puedan corroborar o corregir las respuestas de la aplicación web.

Por lo tanto, los objetivos de este proyecto son dos. Primero, evaluar el desempeño de una representación basada en grafos en el problema de perfilado de autor. Hasta nuestro conocimiento, la representación elegida, propuesta por [9] sólo se ha utilizado en tareas de clasificación temática. El segundo objetivo del proyecto es desarrollar una aplicación web gratuita para que usuarios de Twitter puedan conocer su personalidad, y colaborar en la evaluación del modelo generado, mediante un módulo de validación.

El resto del artículo está organizado como sigue: la sección 2 describe algunas plataformas existentes, las cuales se aproximan a los objetivos de nuestro trabajo. En la sección 3 se describe detalladamente la arquitectura del sistema desarrollado. Luego, la sección 4 muestra la funcionalidad de la aplicación web desarrollada. En la sección 5 se evalúa de forma experimental la representación basada en grafos propuesta por [9], utilizada para tareas de perfilado de autor. Finalmente, en la sección 6 se listan las conclusiones del trabajo e ideas de trabajo futuro.

2. Trabajo relacionado

Como se ha mencionado, la identificación de la personalidad mediante el texto es una tarea que tiene retos distintos a los de identificación de género y edad. Por lo tanto, en esta sección se describen de manera muy breve algunos trabajos que han tratado con el problema de la identificación de la personalidad en texto. Posteriormente, se describen las herramientas o aplicaciones más relacionadas con la aplicación desarrollada en este trabajo.

En cuanto a la investigación realizada para la identificación de la personalidad en texto se pueden clasificar trabajos que hacen uso de diccionarios previamente compilados [2,12], como es el caso de LIWC (*Linguistic Inquiry and Word Count*) [20] y la base de datos psicolingüística MRC [6]. Otro grupo de trabajos han

⁴ <http://pan.webis.de/>

dejado a un lado los diccionarios y se han centrado en analizar el contenido de los textos, ya sea el conjunto de palabras independientes o las secuencias de éstas (n -gramas de palabras). Por ejemplo, en [15] los autores evitan el uso de herramientas sociolingüísticas y deciden hacer una representación del texto en forma de n -gramas de palabras de longitud $n = 2$ y $n = 3$.

Otro grupo de investigaciones se han centrado en el análisis de la personalidad en usuarios de redes sociales como Twitter [1,17] y Facebook [16,19]. En general, estos trabajos toman ventaja de la estructura de la red que se forma con otros usuarios así como del comportamiento de cada usuario dentro de esta red para tratar de determinar la personalidad de los usuarios.

Por otro lado, en cuanto a las herramientas y aplicaciones similares a la propuesta en este trabajo, que buscan analizar el texto escrito en redes sociales de usuarios específicos, se pueden listar las siguientes:

AnalyzeWords⁵. Es un sitio experimental que analiza los tuits de una cuenta proporcionada. Se basa en la investigación realizada por Niederhoffer y colaboradores [14], donde la idea principal es que las palabras que se usan para comunicar un tuit revelan no solo pistas de la personalidad, sino del estilo de pensar, estado emocional y la conexión con otras personas. Este sitio permite introducir una cuenta de Twitter, luego realiza un análisis de las 748 palabras más recientes de la cuenta dada. El resultado del análisis es presentado en tres estilos: emocional, social y de pensamiento.

MyPersonality⁶. Fue una aplicación popular de Facebook que permitió a los usuarios realizar pruebas psicométricas reales, y almacenar, con su consentimiento, su perfil psicológico y perfil de Facebook. Actualmente, aunque el sitio ya no está en línea, existe la base de datos recolectada de *myPersonality* que contiene más de 6,000,000 resultados de las pruebas junto con más de 4,000,000 de perfiles individuales de Facebook. Los encuestados provienen de distintos grupos de edad, orígenes y culturas.

Apply Magic Sauce - PredictionAPI⁷. Esta aplicación permite conocer cuál es la personalidad de usuarios en redes sociales. La predicción se basa en los Likes en Facebook. El motor de predicción fue montado por los investigadores del Centro de Psicometría de la Universidad de Cambridge y está basado en el conjunto de datos de *myPersonality* de más de 6 millones de usuarios. Como resultado, se pueden visualizar las predicciones realizadas por el modelo interno; estas predicciones son la edad, género, personalidad, y la orientación religiosa y política del usuario.

Test de personalidad TP2010. Ésta era una aplicación de Facebook que obtenía información acerca de la personalidad del usuario a través de un test de personalidad, así como también recolectaba datos de las interacciones de usuario con la red social. La meta de TP2010 era descubrir la relación entre los resultados del usuario en la prueba de personalidad y todos aquellos atributos que describen la interacción en Facebook. Esta aplicación también permitía realizar una

⁵ <http://www.analyzewords.com>

⁶ <http://mypersonality.org/>

⁷ <http://applymagicsauce.com/>

comparación de personalidad entre dos usuarios de Facebook. Adicionalmente, podía hacer recomendaciones de amigos basados en compatibilidad, mostrando a las personas, que también habían realizado el test y cuyos resultados eran más similares a los del usuario específico.

En general, las herramientas mencionadas anteriormente proponen distintos métodos para el análisis e identificación de rasgos de personalidad, género y edad (en algunos casos) de información producida por un usuario en redes sociales. La mayoría de estos sistemas incorpora un esquema para recolectar datos; sin embargo, los datos recolectados son en su mayoría en Inglés. A diferencia de la herramienta que proponemos, pues está orientada a recolectar datos en Español. Adicionalmente, los modelos internos con los que las herramientas descritas realizan el perfilado de los usuarios, se basan en su mayoría en los métodos basados en diccionarios o una representación basada en una bolsa de n -gramas, ya sea de palabras o de caracteres. En contraste, en nuestra propuesta proponemos utilizar un método basado en grafos, originalmente descrito y propuesto por [9]. Este método de clasificación utilizando una representación basada en grafos, hasta nuestro conocimiento, no se ha usado en problemas de clasificación no temática de textos.

3. Sistema propuesto

La Figura 1 muestra el esquema general del sistema desarrollado. Dentro de la línea punteada se muestran todos los procesos que se realizan *en línea*, mientras que fuera de la línea punteada se encuentra la generación de los modelos para cada problema, esto es: clasificación de género, edad y cada uno de los rasgos de personalidad del modelo psicológico denominado BigFive [13].

3.1. Construcción del modelo de clasificación

La etapa fuera de línea consiste en la construcción de los modelos de clasificación requeridos. Cabe resaltar que los problemas de clasificación son distintos para cada aspecto demográfico que se trate. Es decir, para el problema de identificación de género se trata con un problema de clasificación binaria, pues las clases sólo son *femenino* o *masculino*. Para el problema de identificación de edad, dado que el conjunto utilizado para generar el modelo consta de cuatro rangos de edad, el problema trata con 4 clases: *18-24* años, *25-34* años, *35-49* años y *50-XX* años. Mientras que para la identificación de la personalidad, se trabajó con 5 clasificadores binarios, uno para cada rasgo del modelo BigFive. De forma que se tienen cinco clasificadores con dos clases: *polaridad negativa* y *polaridad positiva*.

Para la construcción de cada clasificador se implementó el método desarrollado por George Giannakopoulos *et al.* [9] que consiste en dos etapas, la generación de los grafos modelo para cada clase y la segunda etapa que consiste en la construcción del clasificador empleando un enfoque supervisado, donde cada documento es representado por $3 \times N$ atributos (para N igual al número

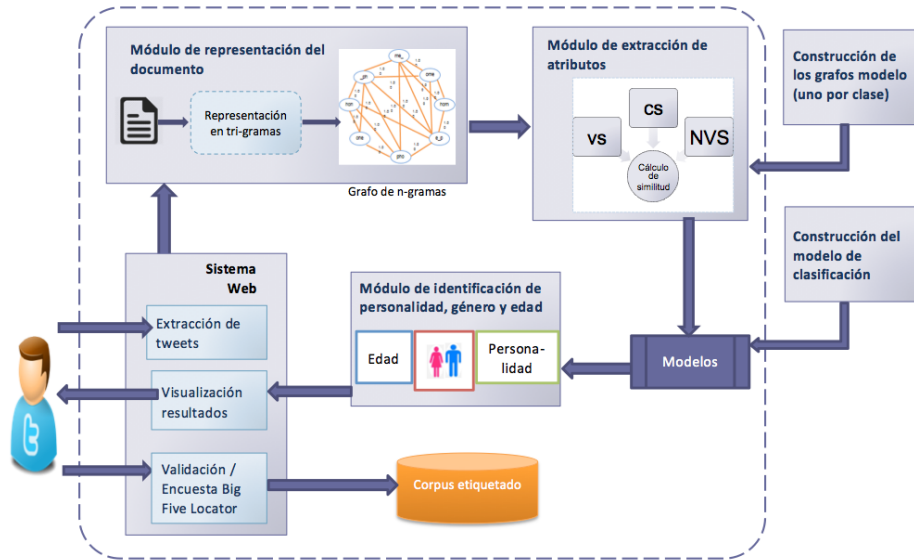


Fig. 1. Esquema general del sistema propuesto para la identificación de personalidad, género y edad de usuarios en Twitter

de clases del problema de clasificación). El cálculo de estos atributos se presenta en la sección 3.3.

Para la primera etapa es necesario crear los grafos modelo de cada clase; para esto es necesario representar cada documento como un grafo, donde cada nodo corresponde a un n -grama dentro del documento, las aristas determinan la co-ocurrencia de dos n -gramas en una misma ventana, y el peso de las aristas representa la frecuencia de co-ocurrencia. Para realizar esta representación se divide el documento en k -ventanas de tamaño t . En este proyecto utilizamos tri-gramas de caracteres con ventanas simétricas de tamaño 4, es decir, cada ventana consiste en el tri-grama analizado, más dos tri-gramas a la izquierda y dos tri-gramas a la derecha.

Una vez que se cuenta con los documentos representados mediante un grafo de n -gramas, todos los grafos de la misma clase se combinan en un solo grafo $G^c = \langle V, E, W \rangle$ al que se le denomina grafo de clase c . Esto es, dada una colección de grafos de los documentos de \mathcal{D}_c , $G_{d_i} = \langle V_i, E_i, W_i \rangle$, éstos se combinan con $G^c = \langle V \cup V_i, E \cup E_i, W \rangle$, donde W es calculado mediante el valor medio de los pesos de ambos grafos. Al final, el grafo modelo G^c tiene las siguientes propiedades: sus nodos incluyen la unión de los nodos de los grafos individuales, y sus pesos se ajustan de manera que converjan con el valor medio de los respectivos pesos. El grafo resultante captura patrones comunes como de co-ocurrencia y secuencias de caracteres vecinos.

3.2. Representación de documentos

Dentro de la aplicación web, una vez que se introduce el nombre de usuario de Twitter, se proceden a descargar mediante la API de Twitter, los tuits más recientes del usuario en cuestión. Para esto sólo se consideran los tuits originales del usuario. El conjunto de los tuits descargados se conjuntan para formar un solo documento. Este documento de texto se preprocesa para eliminar URLs, remplazar símbolos por una etiqueta S y números por un etiqueta N . Posteriormente, estos documentos se representan en tri-gramas de caracteres.

Una vez que el documento se encuentre representado con el conjunto de tri-gramas de caracteres, se construye el grafo no dirigido $G_i = \{V, E, W\}$, donde V son todos los tri-gramas presentes en el documento, E contiene todas las aristas (Vo, Vd) tal que los tri-gramas de Vo y Vd co-ocurren en una misma ventana. Finalmente, W representa la frecuencia de la co-ocurrencia de dos tri-gramas [9].

3.3. Extracción de atributos

La extracción de atributos consiste en obtener medidas de similitud entre el grafo de un documento d_i con cada uno de los grafos de las clases del problema que se intenta resolver. Sean $G^c = \langle V^c, E^c, W^c \rangle$ el grafo de la clase c y $G_i = \langle V, E, W \rangle$ el grafo del documento d_i . Las medidas de similitud calculadas son las siguientes [9]:

- *Containment Similarity* (CS). Esta medida expresa la porción de aristas de un grafo G_i que son compartidos con un segundo grafo (grafo de la clase c) G^c :

$$CS(G_i, G^c) = \frac{\sum_{e \in E} \mu(e, E^c)}{\min(|E|, |E^c|)}, \quad (1)$$

donde $\mu(e, E^c) = 1$ si y sólo si $e \in E^c$, y $|E|$ indica el número de aristas del grafo G_i .

- *Value Similarity* (VS). Medida de similitud que indica el número de aristas del grafo G_i contenidas en el grafo de la clase G^c que tengan el mismo peso:

$$VS(G_i, G^c) = \frac{\sum_{e \in E} \frac{\min(w_e, w_e^c)}{\max(w_e, w_e^c)}}{\max(|E|, |E^c|)}. \quad (2)$$

- *Normalized Value Similarity* (NVS). Es una medida derivada de la anterior donde no se considera el tamaño relativo de los grafos comparados:

$$NVS(G_i, G^c) = \frac{\sum_{e \in E} \frac{\min(w_e, w_e^c)}{\max(w_e, w_e^c)}}{\min(|E|, |E^c|)}. \quad (3)$$

3.4. Identificación de personalidad, género y edad

Como se ha mencionado en secciones anteriores, la aplicación desarrollada trata con tres problemas de clasificación diferentes: clasificación de género, clasificación de edad y clasificación de personalidad (en sus cinco rasgos). En este sentido, se construyeron 7 clasificadores, como se explicó en la sección 3.1. Mediante la utilización de la herramienta Weka [8] se realizaron las clasificaciones según cada problema. El resultado de clasificación es presentada al usuario mediante un módulo de visualización de resultados como se discutirá en la sección 4.

3.5. Módulo de evaluación

Dado que uno de los objetivos de este proyecto es la recolección de datos en español etiquetados con los rasgos de personalidad del BigFive, se implementó un módulo de evaluación dentro de la aplicación web. La evaluación consiste en un cuestionario de personalidad: *Big Five Locator* que contiene 25 preguntas con cinco opciones de respuesta. Este módulo consta de tres funciones: *puntaje*, *conversión* e *interpretación*. Finalmente, para la retroalimentación del usuario, el sistema le muestra su personalidad basada, ahora, en el resultado del cuestionario, de modo que le permita contrastar los resultados con el dado por la aplicación web.

Las respuestas al cuestionario se almacenan junto con el conjunto de tuits utilizados por la aplicación web, de forma que con el paso del tiempo y con el uso continuo se puedan almacenar tuits y la correspondiente personalidad, género y edad del usuario. Cabe mencionar que por el momento, no existe retroalimentación entre el *corpus etiquetado* almacenado con la generación de los modelos de clasificación (ver Figura 1); esta etapa es considerada en el trabajo futuro.

4. La aplicación web: Identifying Your Personality-IYP

Con el objetivo de ilustrar el comportamiento de la aplicación web, se realizó una prueba con un usuario de Twitter al que llamaremos *Anónimo* (con cuenta de Twitter *@anonimo*⁸). Una vez que la aplicación descarga los tuits originales del usuario *Anónimo*, y después de realizar todo el procedimiento descrito en la sección anterior, la aplicación propuesta muestra al usuario su perfil, tal como se ilustra en la Figura 2.

En esta Figura 2 puede observarse que el sistema determinó que el género (B en la figura) del usuario *Anónimo* es *Femenino* y su rango de edad (A en la figura) es de 35 a 45 años. Por otro lado, en cuanto a su personalidad, ésta se muestra en función de cada rasgo del BigFive. Para el usuario *@anonimo*, el

⁸ El nombre de la cuenta de usuario es solamente para fines ilustrativos y no corresponde con el usuario de la cuenta <https://twitter.com/anonimo>

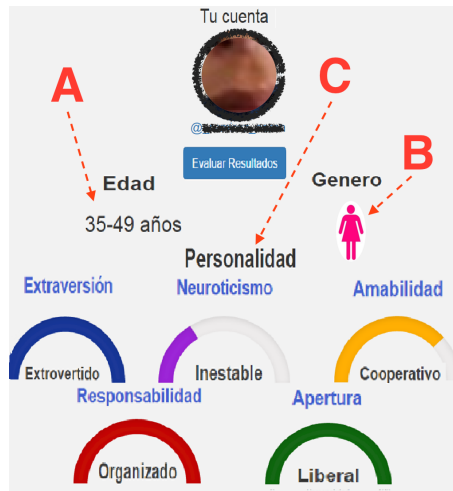


Fig. 2. Ejemplo de la visualización del resultado de edad, género y personalidad para el usuario *@anonimo*

sistema ha identificado que es *Extrovertido*, *Inestable*, *Cooperativo*, *Organizado* y *Liberal*.

La representación de los resultados está basada en la confiabilidad del clasificador para determinar cada uno de los polos (clases) de cada rasgo. Es decir, en el caso del usuario *@anonimo* el sistema arrojó los resultados de clasificación mostrados en la Tabla 1 junto con el grado de confianza del clasificador Naïve Bayes que se utilizó dentro de la aplicación.

Tabla 1. Resultados de la aplicación web para el usuario *@anonimo*. La tercera columna muestra la confianza del clasificador Naïve Bayes de la implementación de Weka que es utilizada para la visualización

| Problema | Clase asignada | Confianza (%) |
|------------------------|----------------|----------------|
| Género | Femenino | 97.3 |
| Edad | 34-45 | 100 |
| Rasgos de personalidad | | |
| Extraversión | Positivo | 100.0 |
| Neuroticismo | Negativo | 65.7 |
| Amabilidad | Positivo | 80.4 |
| Responsabilidad | Positivo | 100.0 |
| Apertura | Positivo | 100.0 |

Algo a notar de la Tabla 1 es que el porcentaje de confianza del clasificador se utiliza para mostrar los resultados de cada rasgo. Observe por ejemplo el

rasgo Neuroticismo en el cual se ha clasificado al usuario Anónimo con la clase negativa en un 65.7 % de confianza, esta información se observa en la Figura 2 en la segunda barra.

5. Validación del modelo de clasificación basado en una representación que usa grafos

En esta sección se presenta el conjunto de experimentos que se realizaron a modo de validar el desempeño de la representación basada en grafos que se describió en secciones anteriores. Note que el modelo que se implementó en este proyecto es el desarrollado por Giannakopoulos y colaboradores [9], mismo que hasta nuestro conocimiento no se ha utilizado en problemas de clasificación no temática.

5.1. Conjunto de datos

Para la evaluación experimental del esquema de clasificación se utilizó un subconjunto del corpus de entrenamiento proporcionado por el PAN 2015 [18] para la tarea de perfilado de autor. El subconjunto que se utilizó contiene tuits en español y cuenta con 100 usuarios diferentes anotados con su género, rango de edad y un valor numérico entre $[-0.5, 0.5]$ para cada rasgo de personalidad. Para nuestras pruebas, cada rasgo de personalidad se consideró como un problema de clasificación binaria, por lo tanto un usuario con valor del rasgo menor o igual a cero se asigna a la clase negativa, mientras que usuarios con valor del rasgo mayores a cero, se asigna a la clase positiva.

5.2. Algoritmos de aprendizaje

Para esta evaluación experimental se utilizaron cuatro algoritmos de aprendizaje de los más representativos y utilizados en el campo del aprendizaje automático. A continuación éstos se describen brevemente:

- *Naïve Bayes (NB)*: Método probabilístico, de los más utilizados por su simplicidad y rapidez, que asume la independencia de los atributos entre las diferentes clases del conjunto de entrenamiento.
- *Árboles de decisión (J48)*: Un algoritmo que permite generar un árbol de decisión, el cual selecciona los atributos más discriminativos basándose en su medida de entropía.
- *Máquina de soporte vectorial (SVM)*: Un discriminante lineal que busca un hiperplano óptimo para separar dos clases.
- *Redes neuronales (RBF)*: Un algoritmo basado en una red neuronal artificial que puede predecir las salidas mediante una combinación lineal de funciones de base radial de los parámetros de entrada.

5.3. Medidas de evaluación

Para evaluar la clasificación se utilizó la medida-F la cual está definida en función de la *precisión* y el *recuerdo* y se define como:

$$medida - F = \frac{(1 + \beta^2)Precision * Recuerdo}{\beta^2 Precision + Recuerdo}, \quad (4)$$

donde $\beta = 1$ representa la media armónica entre la precisión y el recuerdo. La función de β es la de controlar la importancia relativa entre las medidas de precisión y recuerdo. Es común asignar un valor de 1 indicando igual importancia a ambas medidas.

5.4. Resultados

Se realizaron tres experimentos: 1) clasificación de género con dos clases (Femenino y Masculino) , 2) clasificación de rangos de edad con cuatro clases (18-24, 25-34, 35-49, 50-XX)y 3) Clasificación de la personalidad para 5 rasgos definido como un problema de clasificación binaria.

Los resultados de las Tablas 2 y 3 son el promedio de la medida-F de realizar los experimentos dos veces sobre dos particiones distintas de 80 %-20 % sobre el conjunto de 100 usuarios.

Tabla 2. Resultados de clasificación (en medida-F) para el problema de Edad y Género

| | <i>Género</i> | | <i>Edad</i> | | | |
|-----|---------------|-----------|-------------|-------------|-------------|-------------|
| | Femenino | Masculino | 18-24 | 25-34 | 35-49 | 50-XX |
| NB | 0.81 | 0.47 | 0.74 | 0.57 | 0.55 | 0.00 |
| J48 | 0.81 | 0.47 | 0.70 | 0.47 | 0.50 | 0.60 |
| SVM | 0.81 | 0.47 | 0.83 | 0.57 | 0.55 | 0.33 |
| RBF | 0.78 | 0.47 | 0.64 | 0.85 | 0.70 | 0.00 |

Tabla 3. Resultados de clasificación (en medida-F) para el problema de Personalidad

| | <i>Extroversión</i> | | <i>Neuroticismo</i> | | <i>Amabilidad</i> | | <i>Responsabilidad</i> | | <i>Apertura</i> | |
|-----|---------------------|------|---------------------|------|-------------------|------|------------------------|------|-----------------|------|
| | Neg | Pos | Neg | Pos | Neg | Pos | Neg | Pos | Neg | Pos |
| NB | 0.91 | 0.97 | 0.93 | 0.93 | 0.92 | 0.98 | 0.90 | 0.98 | 0.88 | 0.99 |
| J48 | 0.87 | 0.92 | 0.97 | 0.96 | 0.96 | 0.99 | 0.99 | 0.99 | 0.88 | 0.99 |
| SVM | 0.86 | 0.91 | 0.91 | 0.92 | 0.91 | 0.98 | 0.89 | 0.97 | 0.87 | 0.99 |
| RBF | 0.87 | 0.92 | 0.94 | 0.94 | 0.94 | 0.98 | 0.97 | 0.99 | 0.87 | 0.99 |

Para el experimento 1) Clasificación por género, puede observarse, en la Tabla 2, que el desempeño de los cuatro clasificadores es similar; sin embargo, siempre es mejor la clasificación para la clase *femenino*.

En la clasificación por edad (experimento 2) se puede observar que no existe un algoritmo que obtenga el mejor desempeño para todas las clases. Por otro lado, es notorio que la clase que es más fácil de clasificar es la del rango de edad de 18-24 años, mientras que la más difícil es la de 50-xx.

Finalmente, para el experimento 3) Clasificación de rasgos de personalidad (Tabla 3) se observa que en general todos los clasificadores pueden identificar muy bien todos los rasgos. Es importante notar que pese a los buenos resultados obtenidos en estos experimentos, el corpus con el que se trabajó es muy pequeño por lo tanto no se podría generalizar el uso de esta representación para identificar en más del 90 % el rasgo de personalidad de usuarios de Twitter.

Es necesario realizar un estudio más profundo con una cantidad de datos etiquetados mayor para determinar con mayor confianza la pertinencia de esta representación en la tarea de perfilado de autor tanto para género, edad como para personalidad.

6. Conclusiones y trabajo futuro

En este artículo se presentó una aplicación web *Identifying Your Personality - IYP* para el análisis de la personalidad, género y edad de usuarios de Twitter. El objetivo de esta aplicación es doble; por un lado, se presenta la aplicación como una herramienta para generar corpus etiquetado de usuarios de Twitter con sus rasgos de personalidad de acuerdo al modelo psicológico del Big Five, además de su rango de edad y género. Este objetivo se logró mediante la validación de los usuarios de la aplicación web IYP a través de un cuestionario de 25 preguntas que posteriormente se almacena junto con los tuits descargados al momento de realizar el análisis.

Por otro lado, el segundo objetivo fue evaluar el uso de una representación de documentos basado en grafos (de acuerdo al modelo propuesto en [9]). Este objetivo se cumplió con la construcción de grafos modelos para los tres problemas de clasificación atendidos en la aplicación web. Para la construcción de estos modelos se usaron 100 usuarios del corpus proporcionado por PAN-2015. Los resultados de esta evaluación son prometedores, pero es necesario realizar más experimentos con corpus más grandes que gracias a la herramienta se puede construir.

En este sentido, como trabajo futuro se pretende, por un lado, utilizar la información de usuarios recopilada por la aplicación web para generar mejores modelos que a su vez puedan incorporarse a la aplicación. Por otro lado se realizará una evaluación exhaustiva de esta representación variando los parámetros de construcción de grafos, como el tamaño de las ventanas y el tamaño del n-grama.

Agradecimientos. Agradecemos a la Coordinación de la licenciatura Tecnologías y Sistemas de Información de la Universidad Autónoma Metropolitana Unidad Cuajimalpa, al CONACyT a través del proyecto No. 258588 y al SNI.

Referencias

1. Adali, S., Golbeck, J.: Predicting personality with social behavior. In: Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference on. pp. 302–309 (Aug 2012)
2. Argamon, S., Dhawle, S., Koppel, M., Pennebaker, J.W.: Lexical predictors of personality type. In: In Proceedings of the Joint Annual Meeting of the Interface and the Classification Society of North America (2005)
3. Argamon, S., Koppel, M., Fine, J., Shimoni, A.R.: Gender, genre, and writing style in formal written texts. *TEXT* 23, 321–346 (2003)
4. Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. *Journal of Computational Science* 2(1), 1 – 8 (2011), <http://www.sciencedirect.com/science/article/pii/S187775031100007X>
5. Burger, J.D., Henderson, J., Kim, G., Zarrella, G.: Discriminating gender on twitter. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. pp. 1301–1309. EMNLP '11, Association for Computational Linguistics, Stroudsburg, PA, USA (2011), <http://dl.acm.org/citation.cfm?id=2145432.2145568>
6. Coltheart, M.: The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology* 33A, 497–505 (1981)
7. Escalante, H.J., Villatoro-Tello, E., Juarez, A., Montes-y-Gomez, M., Villaseñor, L.: Sexual predator detection in chats with chained classifiers. In: Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. pp. 46–54. Association for Computational Linguistics, Atlanta, Georgia (2013)
8. Garner, S.R.: Weka: The waikato environment for knowledge analysis. In: In Proc. of the New Zealand Computer Science Research Students Conference. pp. 57–64 (1995)
9. Giannakopoulos, G., Mavridi, P., Paliouras, G., Papadakis, G., Tserpes, K.: Representation models for text classification: A comparative analysis over three web document types. In: Proceedings of the 2Nd International Conference on Web Intelligence, Mining and Semantics. pp. 13:1–13:12. WIMS '12, ACM, New York, NY, USA (2012), <http://doi.acm.org/10.1145/2254129.2254148>
10. Golder, S.A., Macy, M.W.: Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* 333(6051), 1878–1881 (2011), <http://www.sciencemag.org/content/333/6051/1878.abstract>
11. Koppel, M., Argamon, S., Shimoni, A.R.: Automatically categorizing written texts by author gender. *Literary and Linguistic Computing* 17(4), 401–412 (2002), <http://llc.oxfordjournals.org/content/17/4/401.abstract>
12. Mairesse, F., Walker, M.A., Mehl, M.R., Moore, R.K.: Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research JAIR* pp. 457–500 (2007)
13. McCrae, R.R.: Cross-cultural research on the five-factor model of personality. *Online Readings in Psychology and Culture* 4(4) (2002), <http://dx.doi.org/10.9707/2307-0919.1038>

14. Niederhoffer, K.G., Pennebaker, J.W.: Linguistic style matching in social interaction. *Journal of Language and Social Psychology* 21(4), 337–360 (2002), <http://jls.sagepub.com/content/21/4/337.abstract>
15. Oberlander, J., Nowson, S.: Whose thumb is it anyway? classifying author personality from weblog text. In: *Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions*. pp. 627–634. Association for Computational Linguistics, Sydney, Australia (July 2006), <http://www.aclweb.org/anthology/P06-2081>
16. Ortigosa, A., Carro, R.M., Quiroga, J.I.: Predicting user personality by mining social interactions in facebook. *Journal of Computer System Sciences* 80(1), 57–71 (Feb 2014), <http://dx.doi.org/10.1016/j.jcss.2013.03.008>
17. Quercia, D., Kosinski, M., Stillwell, D., Crowcroft, J.: Our twitter profiles, our selves: Predicting personality with twitter. In: *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, 2011 IEEE Third International Conference on. pp. 180–185 (Oct 2011)
18. Rangel, F., Celli, F., Rosso, P., Potthast, M., Stein, B., Daelemans, W.: Overview of the 3rd Author Profiling Task at PAN 2015. In: *Cappellato, L., Ferro, N., Jones, G., San Juan, E. (eds.) CLEF 2015 Evaluation Labs and Workshop – Working Notes Papers*, 8-11 September, Toulouse, France. CEUR-WS.org (Sep 2015)
19. Schwartz, H., Eichstaedt, J., Dziurzynski, L., Kern, M., Blanco, E., Kosinski, M., Stillwell, D., Seligman, M., Ungar, L.: Toward personality insights from language exploration in social media. In: *Proceedings of the AAAI Spring Symposium Series* (2013), <https://www.aaai.org/ocs/index.php/SSS/SSS13/paper/view/5764>
20. Tausczik, Y.R., Pennebaker, J.W.: The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of Language and Social Psychology* 29, 24–54 (2010), <http://homepage.psy.utexas.edu/homepage/students/Tausczik/Yla/index.html>
21. Villatoro-Tello, E., Juárez-González, A., Escalante, H.J., Montes-y-Gómez, M., Villaseñor-Pineda, L.: A two-step approach for effective detection of misbehaving users in chats. In: *Forner, P., Karlgren, J., Womser-Hacker, C. (eds.) CLEF (Online Working Notes/Labs/Workshop)* (2012)
22. Villatoro-Tello, E., Ramírez-de-la-Rosa, G., Sánchez-Sánchez, C., Jiménez-Salazar, H., Luna-Ramírez, W.A., Rodríguez-Lucatero, C.: UAMCLyR at RepLab 2014: Author profiling task. In: *Working Notes for CLEF 2014 Conference*, Sheffield, UK, September 15-18, 2014. pp. 1547–1558 (2014)