* + 1. **Práctica 1: *Sentiment Analysis***

*Grupo número:*

1. *Nombre de todos los miembros del grupo*:
2. **Ricardo Champa Bujaico**
3. **Jesús Francisco Lopez San José**
4. **Richard Jordan Cabana Ramirez**

Nota: Siempre deberá entregarse el código fuente de la aplicación a través del CV aunque se haya instalado la aplicación en un servidor web y se indique la URL correspondiente.

**Tratamiento del texto e Implementación**

Para esta aplicación, recuperamos tweets acerca de un usuario, mención o hashtag hasta un límite de entradas descrito por el usuario. Para ello limpiamos el string de cada tweet de caracteres no deseados, y tras aplicar un stemming procedemos a calcular el valor de cada tweet, en base a las palabras con valoración positiva o negativa que contiene y pertenecen al lexicón dado. En caso de haber varias palabras con el mismo stem en el lexicón, se elige un valor aleatorio entre ellas.

Para recuperar los tweets se usa el sistema recomendado por Twitter de recuperar los tweets por ID. Esto se debe a que un nuevo tweet creado por un usuario siempre tendrá un ID superior al último tweet que se creó en el pasado.

El la figura de abajo, los tweets de la izquierda son los tweets que tiene un usuario en un instante y se recuperan 5 de ellos (azules). Puede ocurrir entonces que lleguen nuevos tweets (verdes) y entonces se solicite de nuevo otra tanda de 5 tweets (de la derecha) y se recupere algún tweet que ya habíamos recuperado en el instante anterior (tweet rojo).



Para solucionar esto, simplemente se tiene en cuenta el ID del último tweet recuperado, así en la siguiente llamada sólo se recuperará los tweets con ID menor que el último recuperado.

**Requisitos de ejecución y tecnologías utilizadas**

**Servidor: localhost**

**Lenguajes del lado del cliente: HTML**

**Lenguajes del lado del servidor: PHP**

**Al API de twitter se accede a través de su servicio web restfull (usando GET), con autentificación OAUTH, y las siguientes claves:**

**'oauth\_access\_token' => "321651149-3WDJN7geliDPmQSlO7TWDuOK8cAFgtKYQQFh2Tz1",**

**'oauth\_access\_token\_secret' => "88lttFFu3vs04sJXVTX4tumaNtLjNYFvkfGwwRs6jTODL",**

**'consumer\_key' => "ViOlblew2vsIZ94mgyd9tA",**

**'consumer\_secret' => "702OsJQiF03f4VOVefUG2EwnQkRnHefnHglW6kwi58U"**

Hemos utilizado como tecnología principal PHP, exportando código Javascript para la representación del diagrama de sectores de Google Chart, y HTML para la presentación del formulario.

La librería de stemming español que hemos utilizado está implementada por Rafael Espinosa, y se puede encontrar en <http://www.phpkode.com/scripts/item/php-stemmer/>

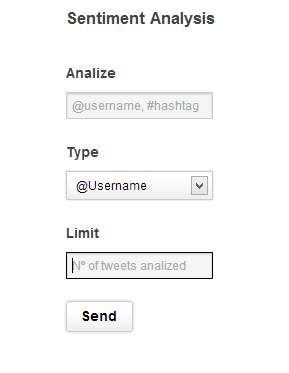
Para nuestro desarrollo hemos utilizado XAMPP , ejecutando Apache para utilizarlo como servidor local.

Para la edición de html y css usamos NetBeans o DreamWeaver Cs4.

El repositorio utilizado para hacer la práctica la hemos creado utilizando GitHub .

**Descripción Interfaz**

La interfaz de la aplicación es la siguiente:



Tiene tres campos en los cuales introducimos el

username , hashtag o un término.

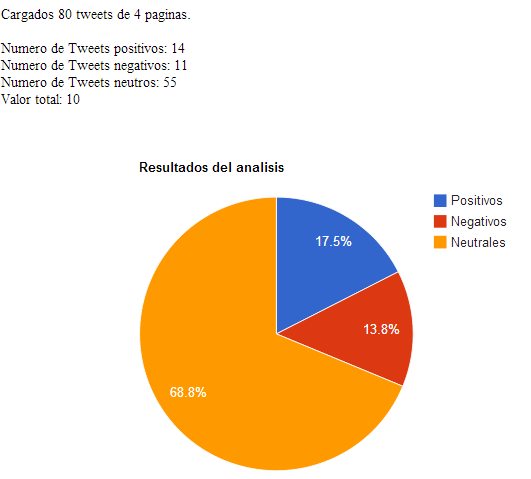
Elegimos el tipo de entrada @Username , @Mention

o #Hashtag.

Número de tweets analizados.

Pulsamos Send para empezar el análisis.

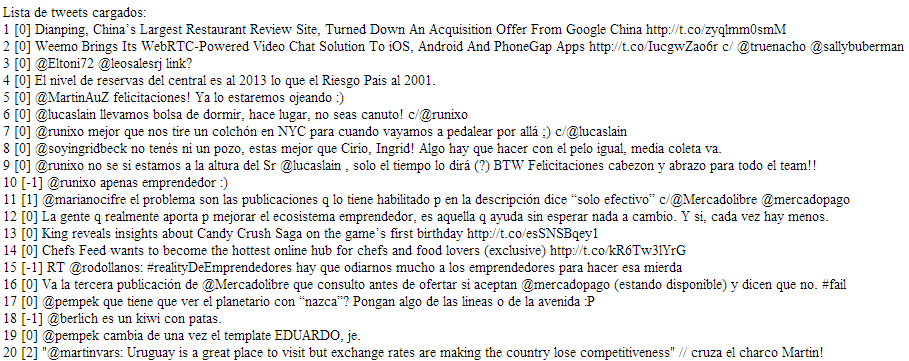
**Se nos mostrará la pantalla siguiente :**



Mostramos el número de tweets

considerados y la clasificación que hacemos de ellos en positivos , negativos o neutrales.

El resultado que obtenemos lo visualizamos con el diagrama de sectores .



Los tweets analizados aparecen en la parte inferior , numerados y con un valor entre llaves que indica la valoración de cada tweet.

**Conclusiones**

Nos parece muy positivo el uso de tecnologías web, ya que en esta carrera es un tema que no se toca casi nada, y en el mundo laboral en general es de lo más exigido y valorado.

En cuanto al tema de la práctica nos parece que es acertado, dentro de lo que cabe es una herramienta curiosa que incluso podría ser utilizada para crear un usuario bot que arroje información sobre los temas candentes del día, etc.

Una extensión que se nos ocurre es poder ver los cambios de humor en el tiempo de una persona concreta, comparando resultados en intervalos de tiempo diferentes de su vida**.**

***Grupo del trabajo***

*Ricardo : Investigó la comunicación con la api de twitter para poder obtener los tweets de un user o hashtag, adicionalmente se encargó de las hojas de estilo para el aspecto de la aplicación.*

*Jesús: Se ha encargado del tratamiento de los tweets que le suministraba Ricardo. En este tratamiento se asigna un valor a cada tweet en base al método ya descrito, con el lexicón. Además se ha encargado en parte de la representación de los resultados obtenidos, y ha realizado la adaptación a PHP del código Javascript realizado por Richard para la representación del diagrama de sectores.*

*Richard: Investigó la api de Google Chart para poder mostrar los resultados visualmente en forma de diagrama de sectores , con los valores que Jesús le suministraba. También se encargó de crear el repositorio del proyecto en GitHub y de realizar la memoria del proyecto.*

***Valoración de la práctica***

**¿Te ha parecido interesante esta práctica?**

Ricardo:

Nivel de dificultad: 2.5

Interés:4

Motivación:1.33

Richard:

Nivel de dificultad: 3

Interés: 3.9

Motivación: 2.5

Jesús:

Nivel de dificultad: 3

Interés: 3

Motivación: 2