



## Becas colaboración curso 2015/2016

Fecha: 18 Junio 2015

### Vicerrectorado de Investigación, Innovación y Transferencia

Subcomisión de I+D+i

Propuesta del departamento *ESTADISTICA E INVESTIGACION OPERATIVA APLICADAS Y CALIDAD*

**Núm Proyecto: 2015/09/00007**

#### **Responsable**

Calduch Losa, Maria De Los Angeles

#### **E-mail**

mcalduch@eio.upv.es

#### **Ext.**

74964

#### **Título proyecto**

Minería de datos para la identificación de nichos de mercado web

#### **Valoración proyecto**

4

#### **Descripción proyecto**

El objetivo del presente proyecto es la aplicación de técnicas de minería de datos para desarrollar y testear una metodología basada en la aplicación del análisis de menciones web, que se realiza para enlazar recursos de información en la web.

Durante el proceso de creación de contenidos, tanto los webmasters (en aquellas webs que controlan) como los usuarios en general (generalmente en sitios de redes sociales) tienden a mencionar conceptos, marcas o entidades, que denotan interés en un determinado nicho de mercado. A partir de la cuantificación de estas menciones (por ejemplo las marcas más mencionadas en un entorno u otro), se puede generar un análisis de menciones (Text Mining o Minería de Textos) para observar qué conceptos analizados se mencionan más, desde dónde (sitios de referencia, intranets, etc.) y en qué forma (uso de jergas, ironías, sentimiento sinónimos, códigos, etc.), así como poder establecer posibles modelos predictivos que permitan identificar tendencias futuras.

El proyecto Trademetrics (CSO2013-46138-P) ha recopilado información masiva sobre personas, marcas y empresas de los sectores españoles del vino y moda, generando con ello una gran base de datos brutos.

Estos datos serán analizados mediante las técnicas pertinentes de minería de datos con el propósito de conocer sus propiedades y relaciones. Con todo ello se construirá y se validará un modelo final predictivo optimizado y aplicado al análisis competitivo de Ecommerce y Emarketing para identificar nichos de mercado internacionales.

#### **Actividades a realizar por el alumno**

• Integrar los datos en herramientas de análisis, text mining y data mining tales como RapidMiner, NodeXL y WEKA, y de visualización de datos.

• Emplear técnicas de clasificación, clustering y asociación ver las relaciones de los datos.

• Obtener y validar un modelo predictivo.

#### **Horario**

Hasta 12 horas semanales. Mañana o tarde, a convenir. Parte del trabajo puede hacerse de manera no presencial.