

Preproceso

Para eliminar el ruido entre una misma palabra → Ej: v.agra, calcular semejanza contra una base de datos de términos spam. Sustituir la palabra por la que maximice la probabilidad a posteriori.

Eliminar espacios en blanco entre una misma palabra, manteniendo la misma estructura del mensaje.

Utilizar stopwords lists para evitar palabras con semántica vacía, realizar la misma comprobación que en la eliminación de ruido. En este caso, eliminar las palabras del mensaje.

- <http://snowball.tartarus.org/algorithms/spanish/stop.txt>
- <http://www.ranks.nl/stopwords/spanish>

Extracción de propiedades

Comparar las frecuencias de los términos en el mensaje, contra la frecuencia de términos en el corpus con las muestras de spam. <http://es.wikipedia.org/wiki/Tf-idf>

Para todo término t con una probabilidad $> \epsilon$ (ϵ mínima probabilidad de acierto), usar la misma probabilidad como una dimensión en las muestras (aprendizaje/test) de Perceptrón:

$$y \in \mathbb{R}^d, s \in y/s = P_{spam}(t)$$

Se tendrán en cuenta aspectos del usuario: número de mensajes, popularidad, tiempo registrado, número de logs,... Constituirán más dimensiones en el vector representación y . Todos los aspectos del usuario en el dominio \mathbb{R} .

2 clases → $\mathbb{C} = \{Spam, NoSpam\}$.

Vectores de peso → W_{Spam}, W_{NoSpam} , inicializados a 0.

Factor de aprendizaje α → Ajustar empíricamente.

Margen b → Ajustar empíricamente.

Firmas AntiSpam

Lista negra (Ip): <http://www.dnsqueries.com>, <http://mxtoolbox.com/blacklists.aspx>

Lista blanca: Configurar lista con usuarios reconocidos o admitidos por la administración.

Resúmenes. Bases de datos con HASHES de mensajes spam conocidos.

Cabecera de mensaje.