

# Clasificación Automática de la Intención del Usuario en Mensajes de Twitter

**Mauricio Martis**

Escuela de Ingeniería Informática de la  
Pontificia Universidad Católica de Valparaíso  
Valparaíso, Chile  
mauricio.martis.c@gmail.com

**Rodrigo Alfaro**

Escuela de Ingeniería Informática de la  
Pontificia Universidad Católica de Valparaíso  
Valparaíso, Chile.  
rodrigo.alfaro@ucv.cl

**Resumen**— Twitter es una red social de microblogs muy popular hoy en día. Su principal característica es que permite publicar mensajes con un máximo de 140 caracteres. El principal interés en este artículo es comprender qué intención tiene un usuario cuando publica un mensaje y clasificarlo de modo automático. Para ello, se propone una taxonomía que define las principales categorías de intenciones en los mensajes de Twitter y luego se propone un modelo para la clasificación de mensajes de Twitter de forma automática de acuerdo a su intención.

**Keywords**— Taxonomía, Intención, Clasificación Automática de Textos, Twitter.

## I. INTRODUCCIÓN

Twitter es una red social de microblogging que es de gran uso en la actualidad. Su principal característica es que permite escribir mensajes de un máximo de 140 caracteres, es por eso, que el mayor interés por estudiar los mensajes emitidos en esta red se debe principalmente al desafío que involucra la clasificación automática de textos cortos.

La elección de Twitter para el análisis de los mensajes emitidos por sus usuarios son sus características únicas, entre ellas: el enorme volumen generado de información ya que los usuarios publican alrededor de 50 millones de mensaje por día, el largo del mensaje posee un máximo de 140 caracteres lo que lo hace un desafío ya que la intención del mensaje se encuentra concentrada en pocas palabras y además se suele utilizar un lenguaje más informal, lo que en conjunto resulta en un gran desafío.

Nuestra hipótesis es que las técnicas para el análisis automático de la polaridad y minería de opinión se desempeñarán mejor si se aplican a los mensajes donde los usuarios han tenido la intención de expresar una opinión.

En este trabajo se propone una taxonomía que define un conjunto de 8 categorías de las principales intenciones en los mensajes de Twitter definidas a partir de estudios

previos. Esto permite clasificar mensajes automáticamente en categorías como Reporte de Noticia (RN), Opinión de Noticia (ON), Publicidad (PU), Opinión General (OG), Compartir Ubicación/Evento (CU), Chat (CH), Pregunta (PR) y Mensaje Personal (MP). Además se define dos jerarquías que agrupan las categorías definidas en la taxonomía. Finalmente se evalúa y se compara el rendimiento de los clasificadores Naïve Bayes y SVM.

## II. PROBLEMA

Uno de las primeras necesidades que surgen para realizar un análisis de la intención de los mensajes de Twitter, es la definición de una taxonomía que permita poder clasificar los mensajes en una de estas categorías definidas. Esto permite poder separar los mensajes de una manera más fácil y rápida para poder mejorar la toma decisiones.

El uso de máquinas de aprendizaje es fundamental para que la clasificación de los mensajes se realice de manera automática. Es por eso que se debe realizar una comparación de rendimiento de los diferentes algoritmos (SVM, Naive Bayes) que son utilizados actualmente en los problemas de clasificación de textos, y así poder definir cuál de estos presenta mejores resultados y se desenvuelve mejor en un problema de clasificación de mensajes cortos, como es el caso de los mensajes en las redes de microblogging.

## III. TRABAJOS RELACIONADOS

Este trabajo tiene como base varios estudios relacionados con la definición de taxonomías que definen la intención en los mensajes de Twitter.

En [1] Java et al. se define una breve taxonomía que identifica la intención de los usuarios en Twitter entre las cuales se encuentra: a) Charla diaria, b) Conversación, c) Compartir Información (enlaces) y d) Reportar Noticias.

Naaman et al. en [2] examina las actividades sociales

y patrones de comunicación en Twitter analizando mensajes de un total de 350 usuarios de Twitter y aplicando un análisis cuantitativo de los datos pretende identificar la actividad de los individuos en la red social de microblogging Twitter. Se definen 9 categorías entre los que se encuentran: a) Compartir Información, b) Auto promoción, c) Opinión/Quejas, d) Estado actual y pensamientos, e) Yo ahora, f) Preguntas a los seguidores, g) Mantener presencia, h) Anécdota personal y i) Anécdota de Terceros.

Morris et al. en [3] se explora el fenómeno de usar los mensajes en las redes sociales para realizar preguntas. A través de una encuesta se identifican las principales características de las preguntas que se emiten en las redes de microblogging, como por ejemplo, el 80% de las preguntas se realizan de manera explícita, y además incluyen el signo de interrogación.

Sriram propone en [4] un conjunto de 8 características para clasificar mensajes de Twitter con el objetivo de mejorar el filtrado de información además de utilizar BOW. Además define 5 categorías en las que clasifica los mensajes de Twitter entre las que se encuentran: a) Noticias, b) Opinión, c) Ofertas, d) Eventos y e) Mensajes privados. Los resultados obtenidos a través de Naïve Bayes mejoran los resultados obtenidos utilizando solamente la representación BOW gracias a la utilización de las 8 características definidas por el autor.

En conjunto estos estudios demuestran que hay un gran interés en torno a la definición de una taxonomía en las redes de microblogging, y además, conforman la base para definir las 8 categorías de intención en Twitter definidas en este trabajo.

#### IV. PROPUESTA

A continuación se presenta la propuesta que abarca la definición de una taxonomía que detalla las categorías en las cual un mensaje puede ser clasificado dependiendo de su intención. Además se describe el conjunto de datos realizado para este estudio, seguido por la descripción de la etapa de pre-procesamiento que permite considerar las características únicas de Twitter.

##### A. Definición de las características

A continuación se presentan las 8 categorías definidas en este trabajo, que incluyen las principales intenciones en los mensajes de Twitter:

1) **Reporte de Noticia (RN):** Corresponde a una noticia emitida de manera objetiva y formal, por lo general emitida por una cuenta corporativa (CNN, Cooperativa, La Tercera, etc.) acompañada de un hipervínculo a la información completa.

2) **Opinión de Noticia (ON):** Corresponde a una noticia emitida de manera objetiva y formal, similar a un RN, pero se emite comentario positivo o negativo relacionado con la noticia opinada.

3) **Publicidad (PU):** Mensajes generalmente emitidos por empresas que ofrecen ofertas, promociones, etc. Incluyendo incluso un hipervínculo de la promoción.

4) **Opinión General (OG):** Representa la expresión del autor sobre algún tema en particular. Suele ser escrito de manera informal, acortando palabras y destacando sentimientos.

5) **Compartir Ubicación / Evento (CU):** Estos mensajes incluyen información geográfica del autor del mensaje, informando dónde se encuentra este y/o que se encuentra haciendo. Usualmente hacen uso de servicios como Foursquare.

6) **Chat (CH):** Representa un mensaje como parte de una conversación entre uno o más usuarios de la red haciendo uso de las menciones a usuarios en Twitter.

7) **Pregunta (PR):** Una pregunta directa se realiza a uno o más usuarios específicos haciendo uso de la posibilidad de nombrar a los usuarios. También se identifican las preguntas a seguidores, que es una pregunta abierta para que cualquiera pueda intervenir.

8) **Mensaje Personal (MP):** Corresponde a cualquier mensaje en el cual se desee informar de asuntos personales y que no corresponda a ninguna de las otras categorías descritas. Incluye mensajes relacionados a situación sentimental, compartir anécdotas, mantener la presencia, entre otros.

Una de las principales diferencias que se presenta en la Tabla 1 es que se ha buscado agrupar algunas categorías en unas que las engloban, esto se debe a que no se busca tener una granularidad tan fina para ciertos casos. Un claro ejemplo de esto se presenta en la categoría Mensaje Personal, que incluye las categorías que en la propuesta de [2]. Representan las actividades que una persona realiza a diario (por ejemplo, anécdotas personales, anécdotas de otros, mantener presencia). Por

otra parte se ha buscado tener un mayor detalle en el ámbito de la información de noticias, intentando diferenciar cuando una noticia es objetiva, cuando se emiten opiniones sobre estas y además cuando estas se consideran promociones.

Tabla 1. Tabla comparativa entre categorías

Categorías Propuestas	Categorías de Naaman	Categorías de Java	Categorías de Sriram
Reporte de Noticia	Information Sharing	Reportar Noticia	Noticias
Opinión de Noticia	Information Sharing Opinions/Complaints	-	Opinión
Publicidad	Self Promotion	Compartir Información (Enlaces)	Ofertas
Opinión General	Opinion/Complaints	-	Opinión
Pregunta	Question to Followers	-	-
Mensaje Personal	Presence Maintenance Me Now Anecdote (me) Anecdote (Others)	Charla Diaria	Noticias
Compartir Ubicación / Evento	-	-	Eventos
Chat	-	Conversación	Mensaje privado

## V. EXPERIMENTOS

En esta sección se presentan la descripción del conjunto de datos utilizado, cómo fueron representados los mensajes y las medidas de evaluación.

### A. Conjunto de Datos

El conjunto de datos empleados en este trabajo se compone de 5222 mensajes clasificados manualmente. Los mensajes se encuentran escritos en idioma español y consta de tweets emitidos y donde se mencione a: Banco Santander, Pontificia Universidad Católica de Valparaíso, El Mercurio, La Tercera, Movistar y un conjunto de tweets de usuarios al azar. Los datos sin clasificar han sido provistos por la empresa Analitic y clasificados por los autores.

### B. Pre-procesamiento

En la etapa de pre-procesamiento es necesario no solo considerar las características comunes de los textos sino que además se utilizan características específicas del dominio [5].

En esta etapa se realizan los siguientes pasos: a) Identificación de menciones, b) Identificación de

emoticono, c) Identificación de URL, d) Identificación de puntuaciones, e) Identificación de cifras numéricas y f) identificación de hashtags. Esto con la finalidad de incluir estas características como palabras clave en el mensaje y que así puedan ser consideradas por Bag-Of-Words (BOW).

### C. Representación

Se utiliza Bag-of-words para la representación y además se definen 3 características binarias. a) Mención al inicio del mensaje, b) Mención dentro del mensaje y c) Uso de emoticonos. Se probó utilizando frecuencia de términos (TF) y presencia de términos (TP). Sin embargo, debido a que los mensajes en Twitter son cortos, normalmente hay sólo una ocurrencia para un término, por lo cual no se aprecian diferencias en el desempeño del clasificador utilizando ambas representaciones.

## VI. RESULTADOS

Para obtener resultado se utilizó los clasificadores Naive Bayes y SVM que provee la herramienta WEKA. Debido a que la diferencia de la precisión de los clasificadores no es significativa, los resultados detallados a continuación sólo representan el uso de presencia de términos (TP).

Se puede decir que SVM es superior que Naive Bayes para el tipo de clasificación expuesto en este trabajo, alcanzando hasta un 7% de diferencia en la precisión por tipo de peso de los términos.

Naive Bayes en la categoría que presenta mayor problema en la clasificación es en Opinión de Noticia, en donde alcanza incluso aproximadamente 35%. Por otra parte utilizando SVM la categoría que presenta menor precisión es Opinión General, donde se alcanza aproximadamente 43%.

En cuanto a las categorías que presentan mejores resultados, en Bayes corresponde a Compartir Ubicación/Evento con un 98% y Reporte de Noticia con un 95% de precisión, por otro lado en SVM también corresponden a las mismas categorías, con un 99% y un 93% respectivamente.

A partir de las matrices de confusión, se puede observar que los problemas que ocurren en la clasificación en las categorías Opinión de Noticias y Opinión General se deben a lo siguiente:

a) Muchos mensajes que debiesen ser clasificados como Opinión de noticias son clasificados como Reporte de Noticias, esto se puede deber principalmente a que una Opinión de noticia se compone en el 80% del

mensaje del reporte de noticia que se está opinando, y donde el resto son pocas palabras del autor dando una opinión.

b) Por otra parte los mensajes que debiesen ser clasificados como Opinión General son muchas veces clasificados como Mensaje Personal o Chat, esto se debe a que el lenguaje utilizado en estas 3 categorías es muy similar, además, al existir mensajes de Chat que contienen opinión (y son clasificados como opinión general), estos en su mayoría parecen mensajes de Chat, o sea se encuentran dirigidos a una persona, con la diferencia que se emite un juicio de opinión sobre algún tema particular. Otro punto que afecta la clasificación de estas categorías es el lenguaje informal y muchas veces con faltas de ortografía, falta de acentos, etc. que causa que palabras que debiesen ser consideradas la misma, se consideran diferentes.

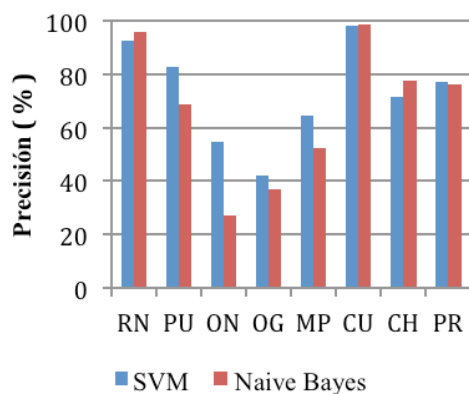


Fig. 1. Precisión de los clasificadores por categorías (BOW y Presencia de Términos)

### A. Clasificación Jerárquica

Luego de los resultados obtenidos se presentan 2 jerarquías que agrupan las 8 categorías definidas en la taxonomía definida. El criterio para definir estas dos jerarquías se explica a continuación.

La primera jerarquía agrupa las categorías según un criterio de similitud entre las definiciones de las categorías, o sea, las opiniones generales y las opiniones de noticias se agrupan en una la categoría superior opinión.

La segunda jerarquía agrupa las categorías según los resultados obtenidos previamente. En donde a partir de la información que entrega la matriz de confusión se agrupan las categorías que son mas similares según las instancias pertenecientes a una clase y que son mal clasificadas en otra clase. Como ejemplo se tiene que las

Opinión de Noticias cuando son mal clasificadas se clasifican como Reporte de Noticias, por lo que se crea una macro categoría que junto estas dos.

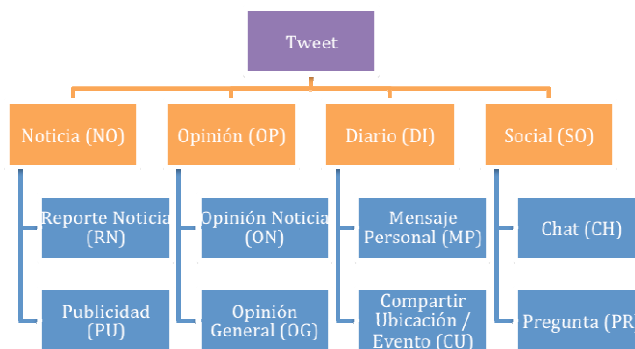


Fig. 2. Diagrama Jerarquía I

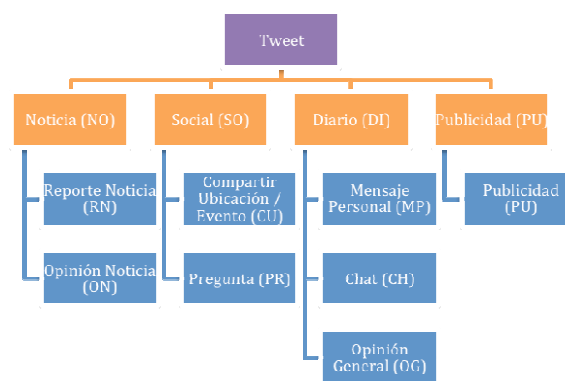


Fig. 3. Diagrama Jerarquía II

De los resultados presentados se puede determinar que una vez definida las dos jerarquías y comparar el rendimiento de los clasificadores SVM y Naive Bayes, la jerarquía II presenta un mejor desempeño en comparación a la jerarquía I.

Este resultado era esperable ya que en la jerarquía II se agrupó las categorías mediante los resultados obtenidos en la sección 6, en donde se toma en consideración la similitud de las categorías agrupadas.

Con respecto a la clasificación de las sub categorías, SVM sigue presentando mejores resultados que Naive Bayes en todos los casos. Las subcategorías de mensajes Social es clasificada con un 99% de exactitud, en cambio las subcategorías de mensajes Diarios es la que presenta mayores problemas. Observando los resultados de la matrices de confusión, los mensajes pertenecientes a la categoría Opinión General y Mensaje Personal siguen siendo las 2 categorías que son mal clasificadas, aunque los resultados se consideran bastante aceptables.

Realizando una comparación a partir de las métricas precisión, cobertura y F-Measure se puede decir las jerarquías I y II tienen un comportamiento aceptable comparado con los resultados obtenidos sin jerarquía, lo que indica que ambas jerarquías entregan buenos resultados y son completamente válidas para poder agrupar los mensajes en categorías superiores.

Enfrentar el problema desde un punto de vista jerárquico permite clasificar los mensajes en distintos niveles de detalle, si se desea poder separar los mensajes sociales, noticiosos de los mensajes diarios y de opinión, se puede utilizar las 4 categorías definidas en las jerarquías I o II, lo que permite a su vez tener una granularidad más fina al poder bajar al siguiente nivel y obtener la clasificación de estos mensajes en una categoría más detallada.

## VII. CONCLUSIONES Y TRABAJO FUTURO

El trabajo presentado en este artículo describe la problemática entorno a la clasificación de mensajes de Twitter de acuerdo a su intención. Se ha descrito el problema, y se ha definido una propuesta que permite clasificar automáticamente los mensajes de Twitter. Esta propuesta incluyó la definición de una taxonomía en la que se definen las categorías que abarcan los mensajes de Twitter de acuerdo a su intención y un modelo de clasificación.

La etapa de pre-procesamiento, al igual que en el proceso de minería de datos, forma un papel principal para la obtención de buenos resultados. Transformar los mensajes, realizar limpieza a los datos, convertirlos para que el clasificador entienda el texto es un punto fundamental.

Con respecto a los resultados obtenidos en los análisis se puede apreciar claramente la ventaja que presenta SVM por sobre Naive Bayes. Además se puede notar que el tipo de peso utilizado para los términos utilizando SVM se podría considerar poco relevante, ya que los resultados obtenidos son muy similares, esto se debe principalmente a las características propias de los mensajes en Twitter, o sea que el largo de estos es bastante pequeño (máximo 140 caracteres) haciendo que la presencia y frecuencia de los términos sea muy similar.

La taxonomía, respecto a los resultados obtenidos, muestra que existe una gran similitud entre los reportes de noticia y las opiniones de noticias, esto se puede observar principalmente en que una opinión de noticia contiene en gran parte del mensaje el reporte de noticia del cual se está emitiendo juicio, es por eso que los resultados obtenidos por el clasificador son bajos en Opinión de Noticias.

Se definió dos jerarquías que agrupan las 8 categorías

definidas en la taxonomía. Una de estas fue definida a partir de los resultados obtenidos anteriormente y la otra jerarquía reúne las clases a partir de la definición de estas categorías. Los resultados obtenidos con ambas jerarquías son considerados buenos y permiten poder clasificar los mensajes a distinto nivel de detalle, permitiendo distinguir los mensajes emitidos con la intención de informar sobre noticias de los mensajes emitidos con la intención de sociabilizar, y si se requiere poder detallar a que sub categoría pertenecen.

Finalmente de este estudio se desprende que lo mejor es utilizar Máquinas de Soporte Vectorial para la clasificación automática de los tweets. Utilizando para la representación de los tweets Bag of Words y debido a que no existe mucha diferencia entre los pesos de términos, se propone utilizar presencia de términos debido a que es menos costoso que utilizar frecuencia de términos.

Como trabajo futuro se propone realizar mejoras en el proceso de pre-procesamiento, enfocado principalmente en la corrección de palabras mal escritas, detección de palabras acortadas y uso de jergas.

Además se propone realizar un análisis descriptivo de los datos utilizando una representación con n-gramas (bigramas y trigramas) para identificar cuales son los bigramas y trigramas que se repiten con mayor frecuencia, y además evaluar el rendimiento de los clasificadores con esta representación.

## REFERENCIAS

- [1] Akshay Java, Tim Finin, Xiaodan Song, and Belle Tseng, "Why We Twitter: Understanding Microblogging Usage and Communities," in *WebKDD/SNA-KDD '07 Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, 2007..
- [2] Mor Naaman, Jeffrey Boase, and Chih-Hui Lai, "Is it Really About Me?: Message Content in Social Awareness Streams," in *CSCW '10 Proceedings of the 2010 ACM conference on Computer supported cooperative work*, 2010.
- [3] Meredith Ringel Morris, Jaime Teevan, and Katrina Panovich, "What do People Ask Their Social Networks, and Why?: A Survey Study of Status Message Q&A Behavior," in *CHI '10 Proceedings of the 28th international conference on Human factors in computing systems*, 2010.
- [4] Bharath Sriram, "Short Text Classification in Twitter to Improve Information Filtering," in *Master of Science Thesis, The Ohio State University*, 2010.
- [5] Vipul Pandey and C.V.Krishnakumar Iyer, "Sentiment Analysis of Microblogs," in *Stanford University*, 2009.