

¿Podemos predecir si Twitter hundirá un banco?

Carlos Perales González



1 ¿Puede Twitter hundir un banco?

2 Análisis de Sentimiento

Usos de este análisis

Análisis en español

Entrenamiento específico para reputación

3 Metodología de clasificación

Extracción de features

Estimación de probabilidades

Clasificadores

4 Resultados

Tasa de acierto

Curva ROC

5 Conclusiones

1 ¿Puede Twitter hundir un banco?

2 Análisis de Sentimiento

Usos de este análisis

Análisis en español

Entrenamiento específico para reputación

3 Metodología de clasificación

Extracción de features

Estimación de probabilidades

Clasificadores

4 Resultados

Tasa de acierto

Curva ROC

5 Conclusiones

¿Puede Twitter hundir un banco?

¿Puede Twitter hundir un banco?

Realmente, no hay datos históricos de que haya ocurrido con anterioridad.

Sin embargo, con la reciente crisis financiera, los reguladores (BCE, FED) están empezando a pedir más medidas de control sobre los bancos.

La regulación está asociada al riesgo:

- Riesgo de mercado
- Riesgo de crédito
- ...

En breve, la Autoridad Bancaria Europea (EBA) exigirá a los bancos una medida del riesgo reputacional.

Riesgo → posibilidad de pérdidas (ganancias, liquidez . . .)

En este caso, a consecuencia del daño de la perspectiva que la sociedad tenga de esta.

Causas y consecuencias

Hay una serie de directrices de la EBA sobre los temas más sensibles a la reputación.

Causas de una mala reputación

- Corrupción
- Mala gestión
- Política de empresa (negocios armamentísticos, deshaucios ...)
- ...

Consecuencias

- Retirada de efectivo
- No apertura de cuentas
- ...

Organización del trabajo

- ① Recogida de tweets (*tweemanager*)
- ② Base de datos con tweets 'positivos' y 'negativos'
- ③ Entrenamiento de un clasificador
- ④ Clasificación a lo largo del tiempo → Indicador reputacional

Clasificador: instrumento matemático que se sirve de un algoritmo estadístico con el fin de asignar una etiqueta clasificatoria.

1 ¿Puede Twitter hundir un banco?

2 Análisis de Sentimiento

Usos de este análisis

Análisis en español

Entrenamiento específico para reputación

3 Metodología de clasificación

Extracción de features

Estimación de probabilidades

Clasificadores

4 Resultados

Tasa de acierto

Curva ROC

5 Conclusiones

Usos de este análisis

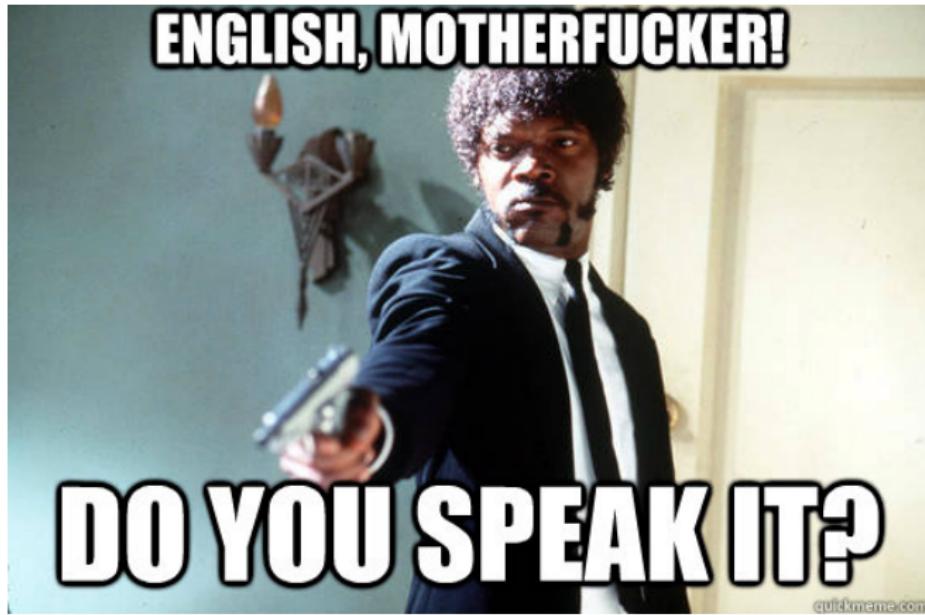
Análisis de sentimiento: determinación automatizada de la subjetividad, polaridad ('positivo', 'negativo') y fuerza que tenga un texto. En pocas palabras, clasifica en 'bueno' y 'malo'. Nos puede servir para ver cómo evolucionan las opiniones.



- Análisis de mercado
- Inteligencia artificial

Análisis en español

La mayoría de la bibliografía que se puede encontrar sobre el tema está en inglés. ¿Por qué?



Análisis en español

¿Por qué hay más análisis de sentimiento en inglés, pese a ser el español una lengua bastante hablada?

- Construcción gramatical más sencilla
- Menos formas verbales
- Muchos verbos regulares y lexemas sencillos

El español es más difícil que el inglés

Análisis en español

Para un análisis en español hemos tenido que trabajar antes el texto.

- ① Stopwords (artículos, preposiciones . . .)
- ② Stemming (reducir una palabra a su raíz)
- ③ Uso de n-gramas (secuencias de n palabras)

Inicialmente nos apoyamos en el paquete *NLTK* para estos procesos, por la cantidad de reglas que involucraban.

Corpus

Para que el algoritmo clasificador funcione el entrenamiento tiene que ser el adecuado.

Buen entrenamiento → Buen clasificador

El entrenamiento se ha conseguido escogiendo tweets que representen las directrices de la EBA sobre qué afecta a la reputación.

Nuestro corpus se ha obtenido mediante recogida de tweets, usando *tweemanager*, buscando sobre entidades financieras.

<https://github.com/nfqolutions/tweemanager>

Índice

1 ¿Puede Twitter hundir un banco?

2 Análisis de Sentimiento

Usos de este análisis

Análisis en español

Entrenamiento específico para reputación

3 Metodología de clasificación

Extracción de features

Estimación de probabilidades

Clasificadores

4 Resultados

Tasa de acierto

Curva ROC

5 Conclusiones

Estructura

- ① Extracción del tweet en sus features
- ② Reglas de probabilidad para los features
- ③ Composición de features en un clasificador

Features: propiedades de un fenómeno a estudiar. Sirven como variables.

- Nuestras *features* serán n-gramas del texto

Extracción de features

- Se aplica *stemming* y *stopwords*

'CaixaBank, reconocida por su compromiso con la conciliación laboral y familiar'

'caixabank', 'reconoc', 'compromis', 'conciliacion', 'laboral', 'famili',
'caixabank reconoc', 'reconoc por', 'por su', 'su compromis', 'compromis con', 'con la', 'la conciliacion', 'conciliacion laboral', 'laboral y', 'y famili', 'caixabank reconoc por', 'reconoc por su', 'por su compromis', 'su compromis con', 'compromis con la', 'con la conciliacion', 'la conciliacion laboral', 'conciliacion laboral y', 'laboral y famili'

Laplace simple

La probabilidad de Laplace simple (LS) es puramente frecuentista

$$P(ngram|pos) = \frac{d}{N} \quad (1)$$

Donde:

d es el n° de veces que aparece en positivo

N es el n° de veces total que aparece

Laplace suavizado (o adición suavizada)

La probabilidad de Laplace suavizada (AS) es una composición de la probabilidad de Laplace con la prob. uniforme

$$P(ngram|pos) = \frac{d + \alpha}{N + n_{cat}\alpha} \quad (2)$$

Donde:

α es un factor de composición. Se ha escogido 0,5

n_{cat} es el nº de categorías ('positivo' y 'negativo'
 $\rightarrow 2$)

Laplace simple vs suavizado

- Ante variación de N y d, Laplace suavizado funciona mejor
- Con valores d y N altos, tiende a Laplace simple
- Laplace suavizado evita asignar valores extremos

Ejemplos

'CaixaBank impulsará el crédito entre los abogados de Castellón'

'impuls' (LS) → 1,0

'impuls' (AS) → 0,75

'credit' (LS) → 0,9167

'credit' (AS) → 0,8846

Clasificadores

El tweet está formado por *features*

$$\text{tweet} \sim \text{features}$$

Para la clasificación usamos el teorema de Bayes

$$P(\text{pos}|\text{features}) = \frac{P(\text{pos})P(\text{features}|\text{pos})}{P(\text{pos})P(\text{features}|\text{pos}) + P(\text{neg})P(\text{features}|\text{neg})} \quad (3)$$

Se usa junto con la hipótesis de independencia

$$P(f_i \cap f_j) = P(f_i|f_j)P(f_j) = P(f_i)P(f_j) \quad \forall_{i,j} i \neq j \quad (4)$$

Bayesiano ingenuo y nuestro indicador reputacional

¿Relación entre tweet y *features*? El bayesiano ingenuo (*Naive Bayesian* o NB):

$$\text{tweet} = (\text{feature}_1, \text{feature}_2, \dots, \text{feature}_n) = \bigcap_i f_i \equiv \text{features} \quad (5)$$

¡Hipótesis! Cambiamos la concepción del elemento a clasificar

$$\text{tweet} \sim \bigcup_i f_i \equiv \text{features} \quad (6)$$

Comparación. Clasificador de manzanas

Tenemos un objeto. ¿Será una manzana?



- Rojo
- Redondo
- $\sim 7 \text{ cm}$ de diámetro

Bayesiano ingenuo (NB) \rightarrow Rojo y redondo y $\sim 7 \text{ cm}$ de diámetro
Nuestro indicador (IR) \rightarrow Promedio de variables.

NB presenta *overfitting*. Tiende a puntuar con valores extremos.
IR puntúa con valores intermedios. Es más conservador.

Índice

1 ¿Puede Twitter hundir un banco?

2 Análisis de Sentimiento

Usos de este análisis

Análisis en español

Entrenamiento específico para reputación

3 Metodología de clasificación

Extracción de features

Estimación de probabilidades

Clasificadores

4 Resultados

Tasa de acierto

Curva ROC

5 Conclusiones

¿Qué medimos? Acierto y ROC

Escogemos un set de validación para testear nuestros clasificadores.

Set de validación → 10 % del tamaño del entrenamiento

Probamos cuántos es capaz de acertar (**tasa de aciertos**), junto con la curva ROC y el área bajo esta curva.

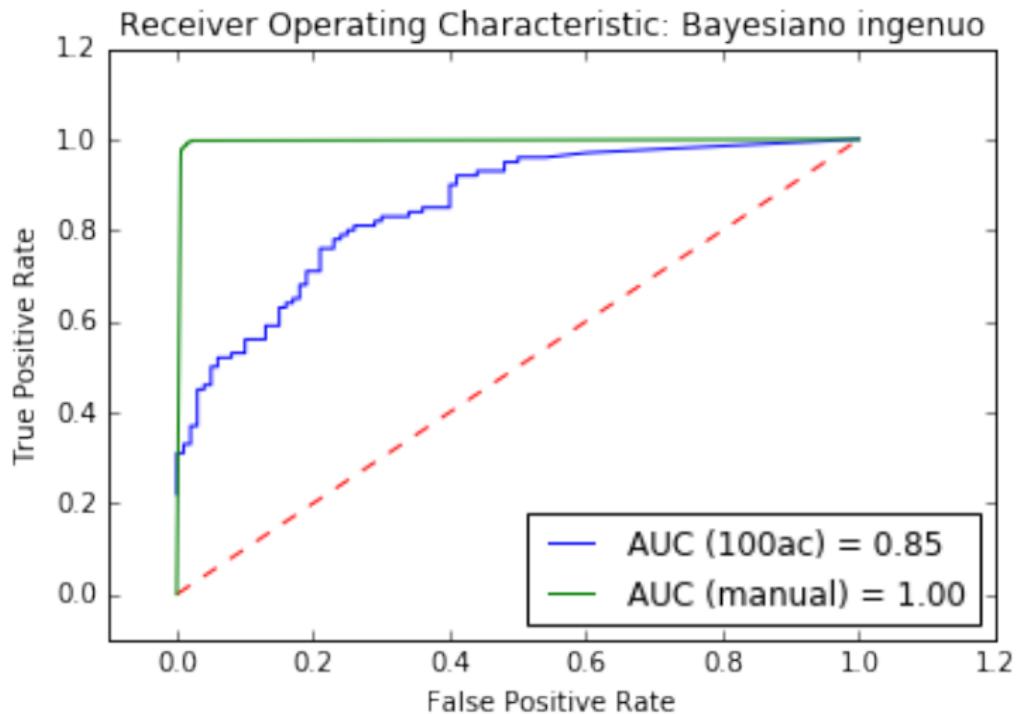
	Acierto (validación)	Acierto (entrenamiento)
NB	0.715	0.995
IR	0.790	0.990

Curva ROC

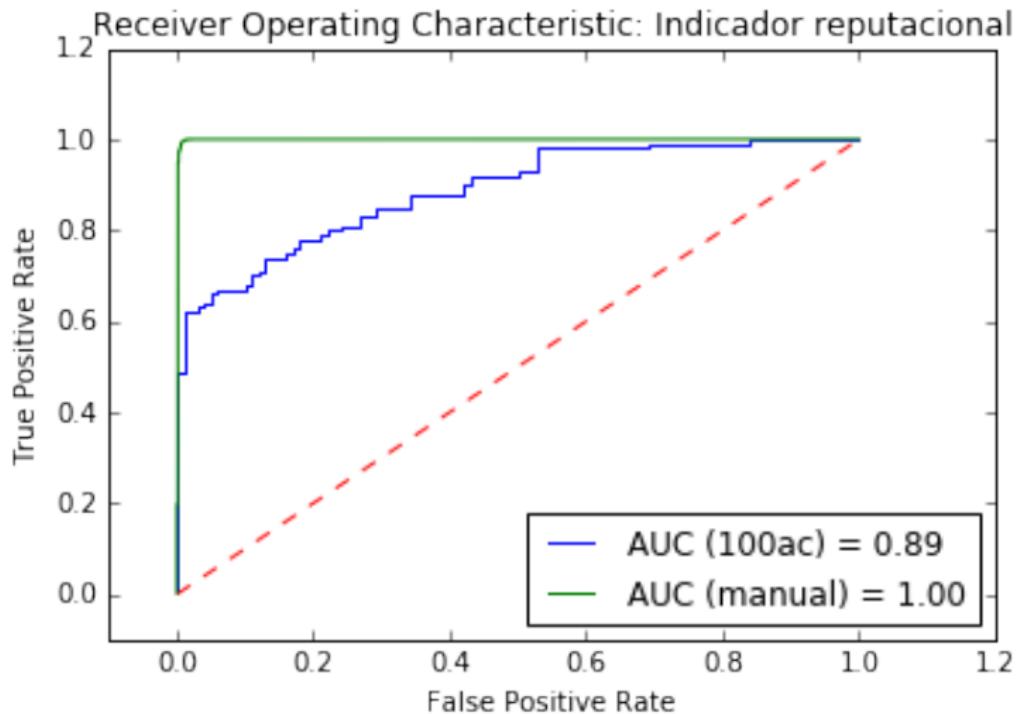
La **curva ROC** es una técnica para ver cómo varía la tasa de verdaderos positivos (TPR) y de falsos positivos (FPR) conforme el umbral discriminante cambia.

El **área bajo esta curva** representa la probabilidad de que un caso que el clasificador haya categorizado como positivo, efectivamente, sea realmente positivo

Gráficas de la curva ROC. NB



Gráficas de la curva ROC. IR



Índice

1 ¿Puede Twitter hundir un banco?

2 Análisis de Sentimiento

Usos de este análisis

Análisis en español

Entrenamiento específico para reputación

3 Metodología de clasificación

Extracción de features

Estimación de probabilidades

Clasificadores

4 Resultados

Tasa de acierto

Curva ROC

5 Conclusiones

Conclusiones

- El clasificador NB tiene *overfitting* que el clasificador IR no tiene.
- Esto se debe, entre otras cosas, al uso de Laplace suavizado y a la hipótesis $tweet \sim \bigcup_i f_i$
- La hipótesis de independencia es ingenua pero funciona
- Los bigramas y trigramas ayudan al análisis en español

Resultados parecidos + ventajas de nuestro clasificador →
 $IR \geq NB$

Conclusiones

- Prueba de una nueva concepción de objeto a clasificar
- Obtención de corpus específico para reputación
- Solución técnica de un problema real
- Desarrollo de la aplicación *Qdos*

The screenshot shows the Qdos application's user interface. At the top, there is a navigation bar with links for 'Reputación Corporativa', 'Riesgo Reputacional', 'Cloud Knowledge', and 'Contacto'. Below the navigation bar, there is a large image of a city skyline with modern skyscrapers. Overlaid on this image is the text: 'Aplicamos valor a la entidad mediante la mejora de su marca reputacional'. To the right of the main image, there is a sidebar with two sections: one featuring a landscape photo and another featuring a person using a tablet. The sidebar contains descriptive text about corporate reputation and risk analysis.

La reputación corporativa puede convertirse en un obstáculo para el crecimiento de la empresa. La razón que una Entidad proyecta/frente al mundo exterior es un riesgo constituye un factor de riesgo adicional para los beneficios, finanzas, reputación o liquidez, derivados de un daño en la reputación de la misma.

Los comentarios sobre la actividad realizada en línea tienen un gran valor. Es fundamental obtener opiniones sobre factores relevantes que permitan detectar evertos que puedan dar lugar a problemas de reputación.

Analizamos de forma automática y en tiempo real la reputación corporativa mediante un conjunto de indicadores previamente definidos que nos permite obtener una medida de riesgo.

Gracias por su atención!!

21



Agradecimientos: Hugo Marrão y Rogelio Rodríguez