



COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

facebook

comillas.edu

Universidad Pontificia Comillas ICAI

Analítica Social y de la Web *Social and Web Analytics*

3.2 Proyecto (fase 1)

Javier Ruiz de Ojeda

Curso 2023-24
Segundo semestre





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Proyecto:

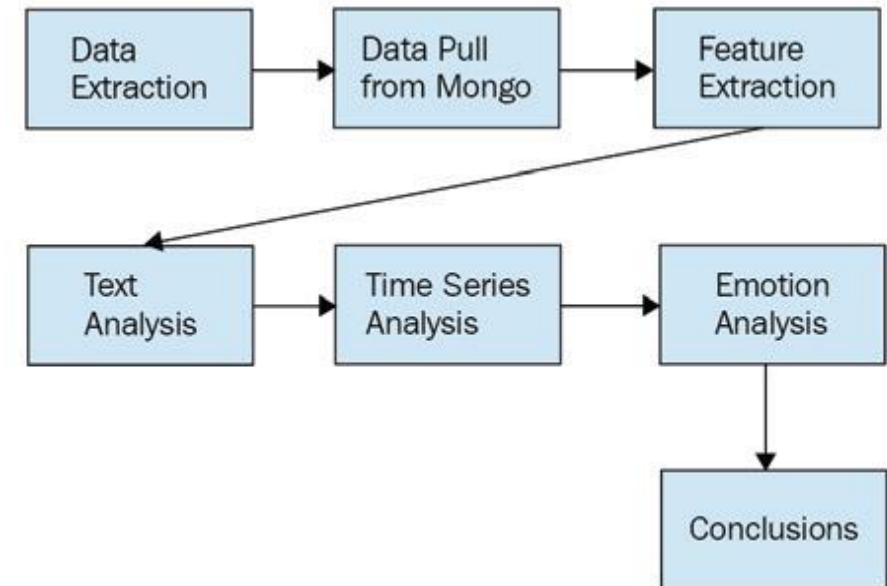
Análisis de datos:



Proyecto - Fase 1

Objetivos

- Extraer posts de vuestra marca permitidos por la API de Facebook
- Extraer los metadatos de cada post (timestamp, número de likes, número de shares, número de comentarios...)
- Extraer los comentarios a cada post y sus metadatos
- Procesar los posts para obtener un número significativo de palabras clave, bigramas y hashtags
- Procesar los comentarios con una API (por ejemplo Alchemy) para obtener las emociones asociadas
- Preparar una presentación de conclusiones



Proyecto - Fase 1

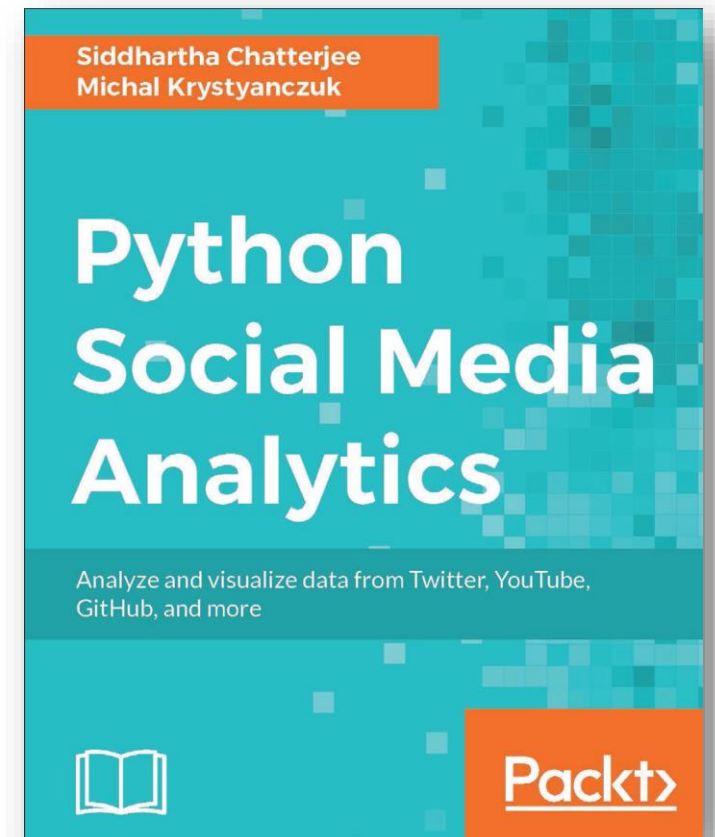
Fuentes

Libro, capítulo 2:

- Crear conexión a la página de marca en FB

Libro, capítulo 3 :

1. Data extraction (posts + comentarios)
2. Data pull
3. Feature extraction
4. Content analysis



Proyecto - Fase 1

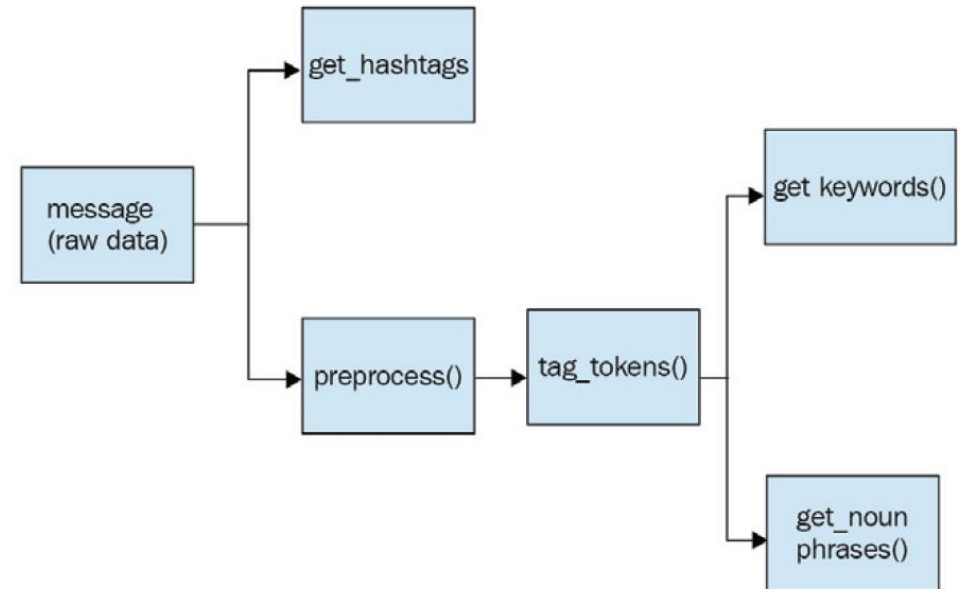
Paso 1: Data Extraction y Paso 2: Data Pull

- ✓ Escoger una marca con actividad reciente y suficiente
- ✓ Obtener al menos 250 posts y hasta 100 comentarios de cada uno
- ✓ Puede usarse MongoDB u otro método de almacenamiento que permita un análisis asíncrono de los datos

Proyecto - Fase 1

Paso 3: Feature Extraction

- ✓ Personalizar una función `preprocess()`
- ✓ Crear *wordclouds* diferenciadas para sustantivos, determinantes, adjetivos y verbos
- ✓ Crear *wordcloud* de los sintagmas nominales (*noun phrases*) más frecuentes
- ✓ Almacenar/imprimir los contenidos de las variables `df_posts` y `df_comments`



Proyecto - Fase 1

Paso 4: Content Analysis (2)

✓ Obtener:

- Tabla con las 10 palabras más frecuentes en posts y en comentarios (y su frecuencia)
- 5 posts asociadas a cada una de las 3 palabras más frecuentes en posts (cada uno con su url)
- 5 comentarios asociados a cada una de las 3 palabras más frecuentes en comentarios (cada uno con su url)

Google:

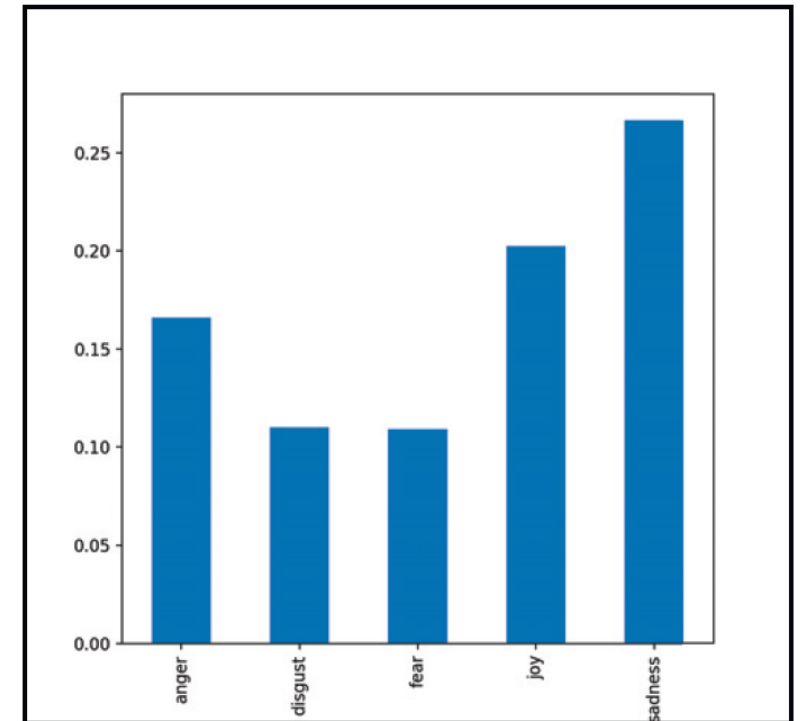
- "Ok Google", how big are polar bears? (<http://goo.gl/RcjL0g>)
- Winter is long and awful. Let's talk tips for planning spring and summer travel with Google Flights (<http://g.co/go/9cj5a>).
- Pick up your wand and experience the magic of the wizarding world with a little help from Google and #FantasticBeasts: <https://landing.google.com/fantasticbeasts/#/>.

User comments	Frequency	Brand posts	Frequency
google	123986	google	516
cancel	62846	new	155
play	62729	search	130
ever	30538	day	96
be	16089	today	95
plz	15447	world	94
change	13431	doodle	92
history	12665	check	91
hey	11562	see	91
anyone	11027	get	87

Proyecto - Fase 1

Paso 4: Content Analysis (5)

- ✓ Obtener:
 - Gráfica de barras que muestre la distribución de emociones de los primeros 1000 comentarios
 - 1 post para cada una de las emociones, donde esa emoción sea superior a 0,5





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

¡Muchas gracias!

