



ANÁLISIS PROMOTER - ESPAÑA

PRACTICA PNL

Izaskun López-Samaniego Palomino | 18/06/2017

OBJETIVO

Se realiza un análisis de los comentarios realizados por los clientes de la empresa de alquiler de parkings para obtener a partir del análisis semántico insight que nos permitan entender como nos valora el cliente y en una segunda fase crear un modelo para segmentar las respuestas con las etiquetas necesarias para tratarlos posteriormente.

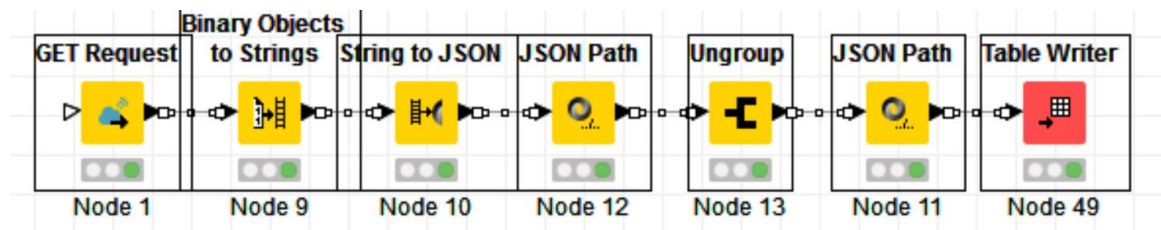
PROCESAMIENTO DE LOS DATOS

La información recogida son las respuestas a la encuesta que se envía a los clientes que han aparcado con la plataforma on-line a través de la herramienta Promoter. En dicha encuesta los clientes valoran su experiencia del 1 al 10 y en ocasiones escriben comentarios que nos ayudan a entender el porqué de dicha puntuación.

DESCARGA DE DATOS

Nos descargamos la información en un fichero JSON conectándonos al API de promoter.io.

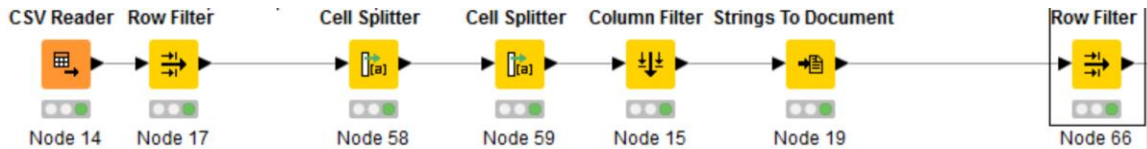
Una vez descargada la información tratamos los datos para convertirlos en una tabla con la que poder trabajar.



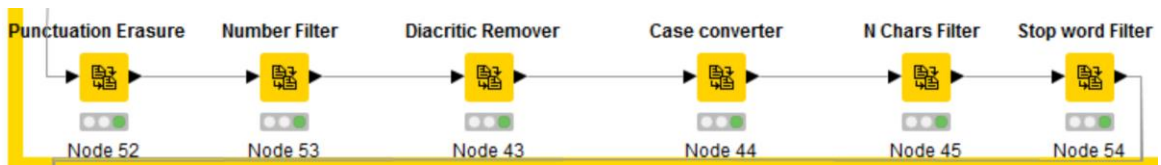
Este API tiene la limitación de descargar un máximo de 100 registros en cada llamada. Para solucionarlo hay que realizar una llamada por página y unificar los registros; para la práctica se ha trabajado con un fichero CSV ya consolidado y con la información anonimizada.

TRATAMIENTO DE LA INFORMACION

Se carga la información desde un fichero *.csv con los datos anonimizados, a partir de la carga del fichero se siguen los siguientes pasos:



1. Selección de los campos y los registros a utilizar:
 - 1.1. Se eliminan los campos que no aportan información
 - 1.2. Seleccionamos únicamente los registros que provienen de clientes españoles.
 - 1.3. Nos quedamos con la primera etiqueta sobre los comentarios de los detractores.
 - 1.4. Se filtra la información para quedarnos únicamente con los registros que tienen comentarios.
2. Depuración del contenido de los comentarios para homogeneizar los datos



- 2.1. Eliminar signos de puntuación, acentos y otro tipo de caracteres.
- 2.2. Convertir los caracteres en mayúsculas.
- 2.3. Eliminar caracteres con una longitud inferior a 4 caracteres.

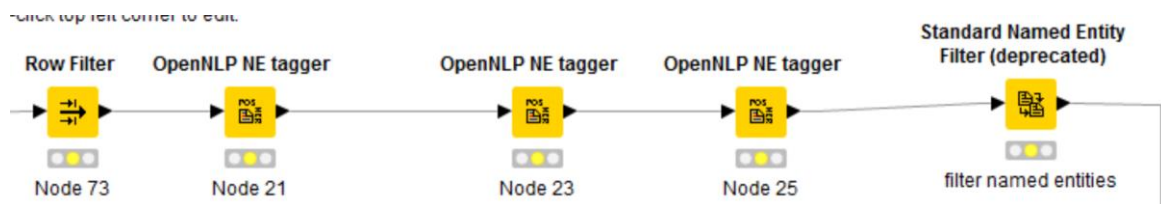
ANALISIS DE LA INFORMACION

INSIGHTS DE CLIENTES

Para entender que están valorando los clientes nos centramos en analizar los términos que utilizan en sus comentarios, viendo aquellos que se usan con más frecuencia a través de una nube de tags.

Para ello ejecutamos el siguiente proceso primero para los promotores y posteriormente para los detractores y obtenemos sus respectivas nubes de tags

1. Filtrar para quedarnos con la información que nos interesa analizar



-
- The diagram illustrates the MapReduce workflow for word frequency analysis. It consists of four main stages connected by arrows:
- Bag of Words Creator**: The first stage, represented by a yellow box with a document icon. It takes input and produces a "Bag of Words" (represented by a list of words).
 - Tags to String**: The second stage, represented by a yellow box with a "TS" icon. It takes the "Bag of Words" and produces a "String" (represented by a list of strings).
 - TF**: The third stage, represented by a green box with a "TF" icon. It takes the "String" and produces a "TF" (Term Frequency) value (represented by a list of numbers).
 - GroupBy**: The final stage, represented by a yellow box with a "GroupBy" icon. It takes the "TF" and produces the final output (represented by a list of groups).

- Color Manager** **Tag CloudImage Port Writer**
- 
- ```
graph LR; CM[Color Manager] --> TCPW[Tag CloudImage Port Writer]; TCPW --> TCPW2[Tag CloudImage Port Writer]
```

## ANALISIS DE PROMOTORES

A partir de la nube de tags podemos obtener unos primeros input que nos indican como los clientes que nos valoran positivamente hablan en sus comentarios en primer lugar del precio, la comodidad/facilidad.

ATENDIDO LLEGAMOS TIEMPO EXTRAORDINARIO MUSTADO PESAR MUCHA FORMA RECOMENDABLE CORRECTAMENTE GRANADA MOVIL UBICADO CUMPLIO REALIZAR EPIS  
ADICIONAL TIPO PASADIZO ASISTIDA PRESTADO TENDRAN JUSTO SALUD TIEMPO PRIMERA GARANTIA GENERAL LUGAR TRAMITE CONTRATACIONES FORMI  
TACION GESTIONES ASIGURAS PROPIO PROXIMIDAD IMPECABLE TODOS DEJAR TRANQUILIDAD SOBRE OFERTA VOLVERE GARANTIA INTUITIVA TRABAJO CONFORME I  
T ATENDIDOR RECOMENDADO PUNTUALIDAD MISMO TUUVIMOS EFICAZ CENTRO CORRECTO CIUDAD DONDE MUCHAS UTILIZADO FUERON INDICACIONES SORPRESA  
3 RESERVADO SINGULA MUCHO CONTRATAR ESTABA EFICACIA EXCELENTE PERSONAL ACCESO BUENOS DISPONIBILIDAD MADRO TODAS FORMALIDAD HEIO  
TO SERVICIOS NUNCA ESTAR EXPERIENCIA ENCONTRAR PODER TRATO COMODO COCHE SALIR FUNCIONO BUENAS LLEGADA TABLETA FALLAS X  
6BUDID INTERNET ENTRAR PROBLEMAS NINGUN FACIL SERVICIO BUENA APARCAMIENTO PARKINGS CERCA AGIL  
ZACIONES LINDAS ATENCION DESDE ECONOMICO SENCILLEZ VECES SITUADO DURANTE 58  
CIUDAD AMABILIDAD ATENCION PARKING RESERVA TE  
RACION RAZONABLE HACER PRACTICO PLAZA AHORRO 1  
SISTEMA POSIBILIDAD SIEMPRE PARQUEO SERIEDAD RAPIDEZ BARATO SEGURO MEJOR RECIBO DESTINO SALUDO 61  
MBEN RAPIDA CENTRO SITUACION SEGURIDAD PERFECTO RAPIDO UBICACION CUANDO FUNCIONAMIENTO  
D RESPECTO TANTO USADO UTILIZAR GRACIAS CALIDAD APARCAR FACILIDAD SENCILLO SITIO FIABLE PROCESO AMPLIAS RESERVA  
SFACCION TELEFONO VEHICULO APLICACION PORQUE FIABILIDAD PRECIOS PROBLEMA RESERVAR PLAZAS FACILIDADES PERFECTAMENTE TRAMITES 1  
LUNDO COMPETITIVO SATISFECHO TRAVES BUSCAR AUNQUE RELAJON GENIAL GESTION CERCANIA ESTANCIA PAGINA SABER OTROS CONTENTO GRANDE SA  
FICA COMPLICACIONES TARIFAS ADEPTOFRONTO CONTRATACION TIENES BUENO AMABLE PAGAR PARTE NUEVO COMUNICACION GESTIONAR SACAR CLAR  
ABORDANDO UENO VIAJE MEJORAR COMPRA TOTALMENTE AHORRAS DESDE HORAS ADEMAS LUGAR MOMENTO MUCHOS CIUDADES FUNCIONADO RESULTA CANCELAR

## ANALISIS DE DETRACTORES

Los detractores son aquellos clientes que han valorado su experiencia con una puntuación por debajo o igual a 7

A partir de la nube de tag vemos como los principales conceptos nombrados por estos clientes es la reserva, el tiempo y el precio. En segundo lugar vemos conceptos como la entrada, la plaza/aparcamiento, el poder entrar/llegar o las llaves.

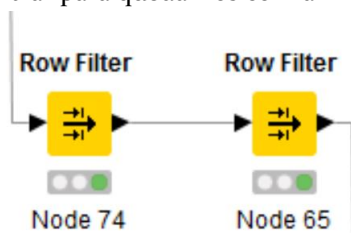


## ETIQUETAR LOS COMENTARIOS

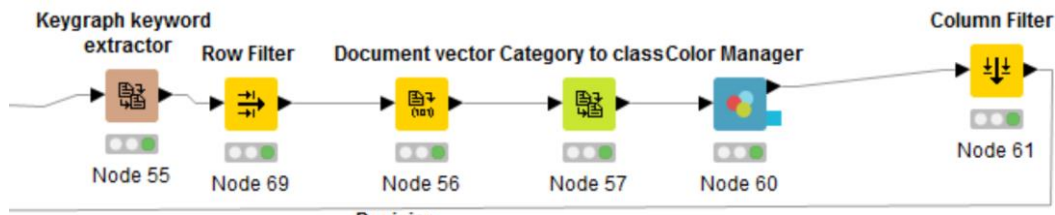
Actualmente los comentarios se etiquetan manualmente para posteriormente realizar acciones sobre los mismos, por lo que vamos a realizar un ejercicio de clasificación que nos permita clasificarlos según un modelo predefinido.

En esta ocasión vamos a clasificar únicamente los comentarios de los detractores siguiendo los siguientes pasos:

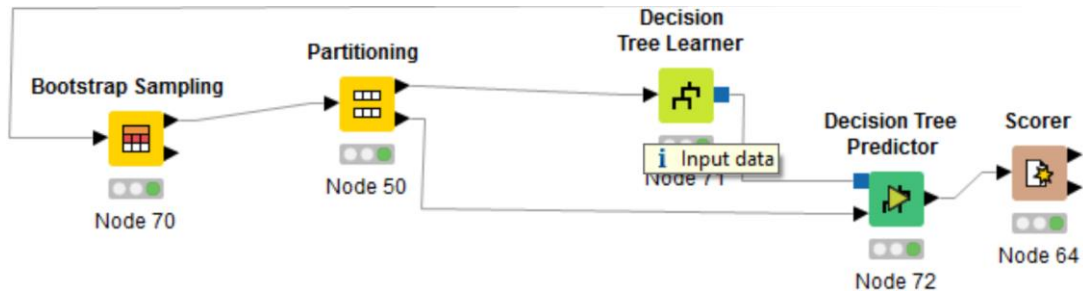
1. Filtrar para quedarnos con la información que nos interesa analizar



- 1.1. Seleccionar promotores/detractores:
    - 1.1.1. Promoters: 8-9
    - 1.1.2. Detractores: 1-6
  - 1.2. Registros etiquetados.
2. Analizamos los términos y su frecuencia para incluir esta información en el modelo de clasificación:



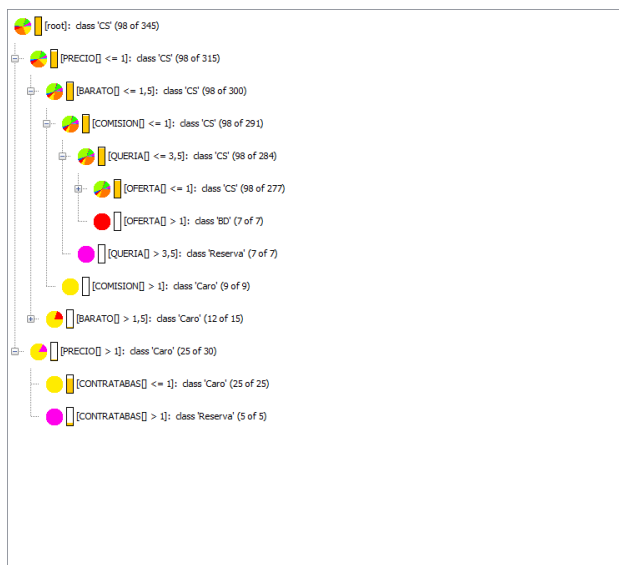
3. Ejecutamos el modelo:



- 3.1. Bootstrap: El volumen de información con la que trabajamos es muy pequeño por lo que con la técnica del bootstrapping incrementamos el volumen de información a tratar manteniendo la “esencia” de la muestra (variabilidad)
- 3.2. Dividimos la muestra en test (60%) y testing (40%)
- 3.3. Ejecutamos el modelo
- 3.4. Ejecutamos la predicción
- 3.5. Obtenemos la información para verificar la viabilidad del modelo.

## ARBOL DE DECISION

Extraemos el árbol de decisión donde podemos observar que los términos que más discriminan son “precio” y “comisión”, seguidos de “oferta” y “quería”



## TEST DE SIGNIFICACION

El ajuste al contrastarlo con la muestra de testing es alto, aunque hay que comprobar en próximos pasos que el modelo no este sobreajustado debido al remuestreo realizado.

| Row ID         | TruePo... | FalsePo... | TrueNe... | FalseN... | D Recall | D Precision | D Sensitivity | D Specifity | D F-meas... | D Accuracy | D Cohen'... |
|----------------|-----------|------------|-----------|-----------|----------|-------------|---------------|-------------|-------------|------------|-------------|
| BD             | 14        | 0          | 216       | 0         | 1        | 1           | 1             | 1           | 1           | ?          | ?           |
| Parking        | 55        | 15         | 155       | 5         | 0.917    | 0.786       | 0.917         | 0.912       | 0.846       | ?          | ?           |
| Caro           | 37        | 0          | 188       | 5         | 0.881    | 1           | 0.881         | 1           | 0.937       | ?          | ?           |
| CS             | 65        | 5          | 156       | 4         | 0.942    | 0.929       | 0.942         | 0.969       | 0.935       | ?          | ?           |
| Contenido      | 19        | 0          | 206       | 5         | 0.792    | 1           | 0.792         | 1           | 0.884       | ?          | ?           |
| Web            | 7         | 0          | 222       | 1         | 0.875    | 1           | 0.875         | 1           | 0.933       | ?          | ?           |
| Económico      | 3         | 0          | 227       | 0         | 1        | 1           | 1             | 1           | 1           | ?          | ?           |
| Fácil/Sencillo | 1         | 0          | 229       | 0         | 1        | 1           | 1             | 1           | 1           | ?          | ?           |
| Producto       | 3         | 0          | 227       | 0         | 1        | 1           | 1             | 1           | 1           | ?          | ?           |
| Reserva        | 6         | 0          | 224       | 0         | 1        | 1           | 1             | 1           | 1           | ?          | ?           |
| Falta          | 0         | 0          | 230       | 0         | ?        | ?           | ?             | 1           | ?           | ?          | ?           |
| Multipass      | 0         | 0          | 230       | 0         | ?        | ?           | ?             | 1           | ?           | ?          | ?           |
| Overall        | ?         | ?          | ?         | ?         | ?        | ?           | ?             | ?           | ?           | 0.913      | 0.889       |