

EVALUACIÓN DE LA EFECTIVIDAD DEL MOTOR

Grupo F

Jorge Aleman Gonzalez
Roger Ventura Castelló

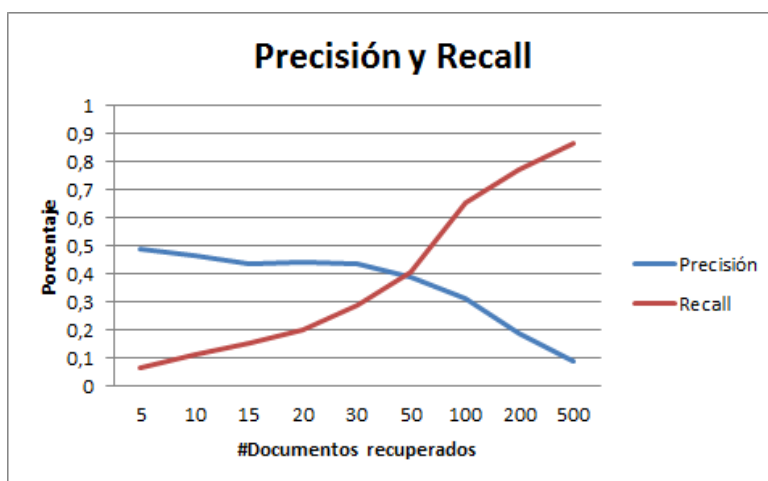
Ferran Cantariño I Iglesias
Gabriel Arroyo Hinojosa

Hemos testado nuestro motor realizando un total de 23 queries. Analizando un poco los resultados obtenidos vemos una recuperación total de 10487 documentos. De los 1152 documentos relevantes se han recuperados 970.

En general obtenemos un motor de búsqueda que empieza con una precisión ligeramente inferior a la media siendo esta del 48,7 % y manteniéndose en una bajada poco pronunciada hasta los 50 documentos cuando empieza a decrecer rápidamente.

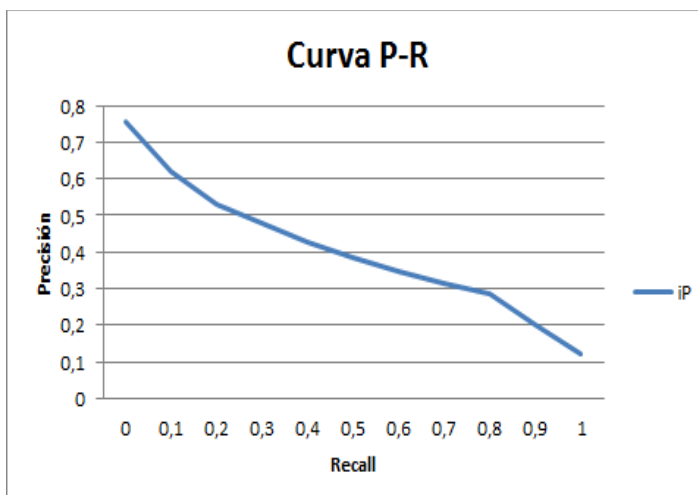
En el caso del recall nos encontramos con que este aumenta a medida que recuperamos más documentos, superando el 50 % a partir de los 100 documentos.

Fijándonos en la progresión de nuestro motor, tenemos una precisión del 48,7% en el cut-off de 5 documentos, una precisión inicial baja que va disminuyendo progresivamente al ir recuperando más información obteniendo precisiones del 18,46% en los 200 primeros documentos y del 8,43% en los 500 primeros.



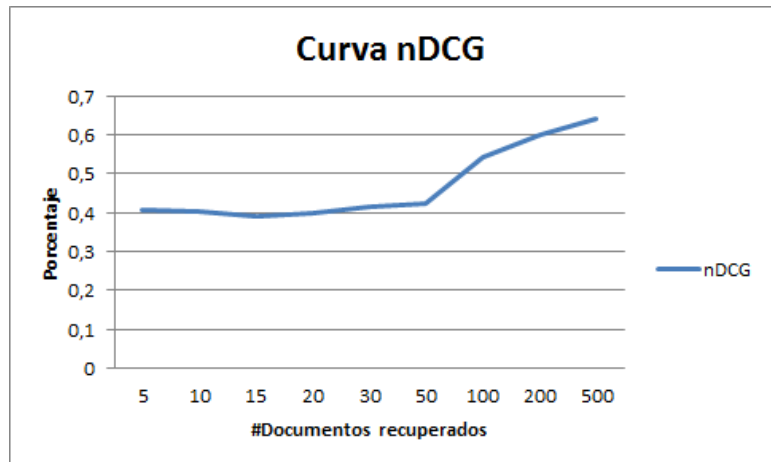
Obtenemos precisiones bastante mejorables en la ejecución de nuestra primera implementación del motor de búsqueda. Como hemos visto en teoría precisión y recall son inversos el uno del otro, por esta razón, inicialmente de 5 documentos recuperamos algo menos de la mitad que son relevantes.

En el caso del recall este también empieza también un poco bajo con porcentajes, por ejemplo del 11,32% en los 10 documentos iniciales, pero sube hasta obtener un recall del 86,51% en los primeros 500 documentos acercándose cada vez más al recall total mientras que la precisión se ve reducida lentamente.



Si observamos los resultados del nDCG vemos como para los primeros 50 documentos obtenemos un ganancia de información relevante constante con un valor no superior al 43 %. A partir de los 50 documentos podemos ver como la curva de nDCG aumenta hasta llegar al 64,25 % a los 500 documentos.

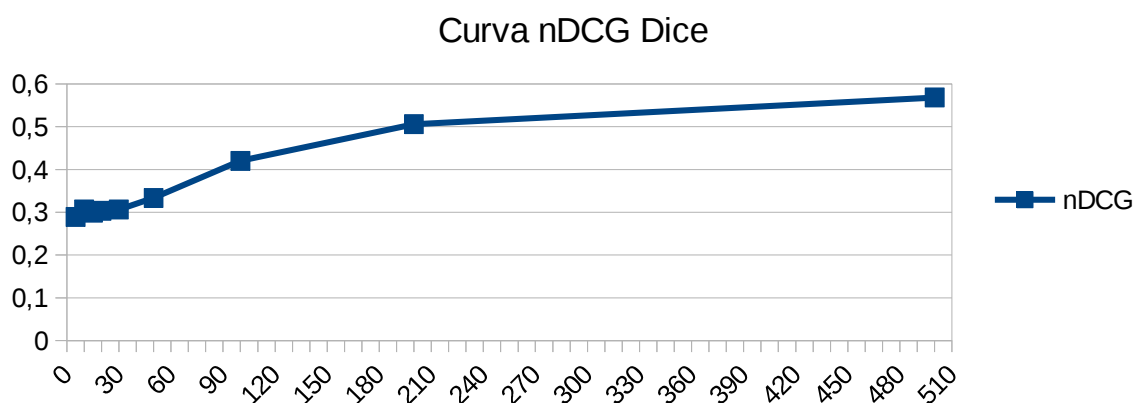
Esta información nos dice pues, que para los primeros 5 – 20 documentos, que suelen ser generalmente a los que se accederá si se realiza una búsqueda poco exhaustiva, la recuperación de la información no es demasiado buena siendo el orden en que se han recuperado los documentos bastante lejano del ideal pues, en nuestro caso, ronda el 40 %.



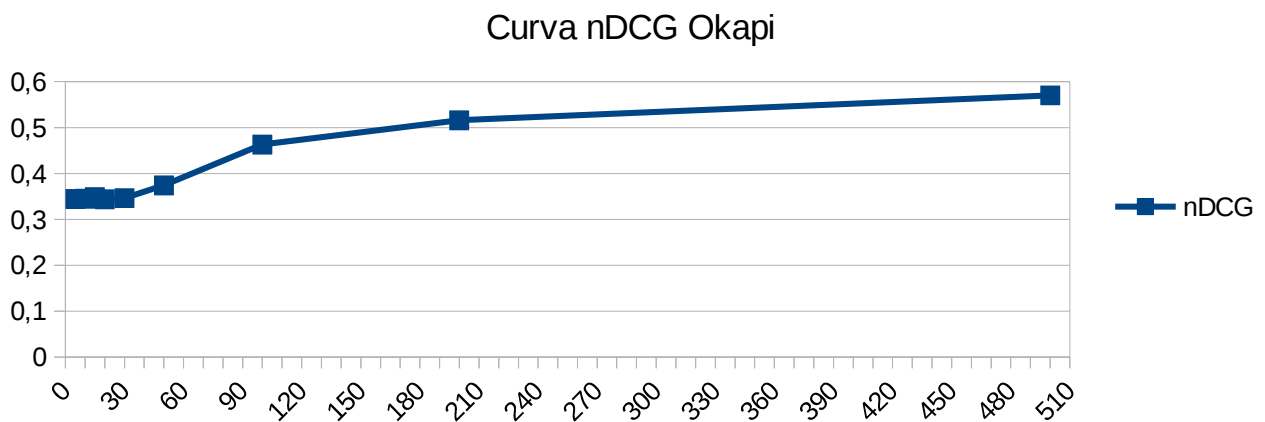
Como conclusión podemos decir que por ahora nuestro motor necesita bastantes mejoras que le permitan obtener mejores resultados para conseguir mejorar esta curva nDCG. Dependiendo de cual sea su función y a quién esté orientado el motor podríamos decir que con los resultados actuales la información encontrada no sería lo suficientemente precisa como para ser de lo suficientemente idónea para un usuario que no tenga intención de realizar una búsqueda muy exhaustiva. También podríamos concluir que el buscador es relativamente exhaustivo siendo capaz de recuperar más del 70% de documentos relevantes a partir de los 200 documentos.

Como punto extra hemos realizado un seguido de comprobaciones con diversas formas de calcular la similitud e implementando el método Okapi. En estos casos hemos obtenido resultados en general inferiores a los obtenidos con el método principal. Así pues:

Si usamos el cálculo de la similitud con la fórmula de Dice obtenemos 931 documentos relevantes, con una precisión inicial del 33,04 % que llega hasta el 36,52 % como punto máximo a los 10 documentos y con una exhaustividad que llega al 82,55% al final de los 500 documentos. La precisión media final fue de: 27,66%.



En el caso de Okapi con $k = 2$ y $b = 0.75$, hemos obtenido 877 documentos relevantes, con una precisión inicial y máxima del 40,86 % y un recall del 78,47 %. La precisión media final ha sido del 34,44%.

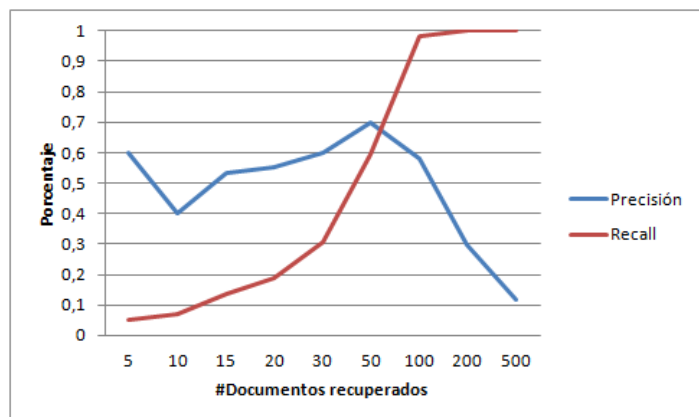


TOPIC #12:

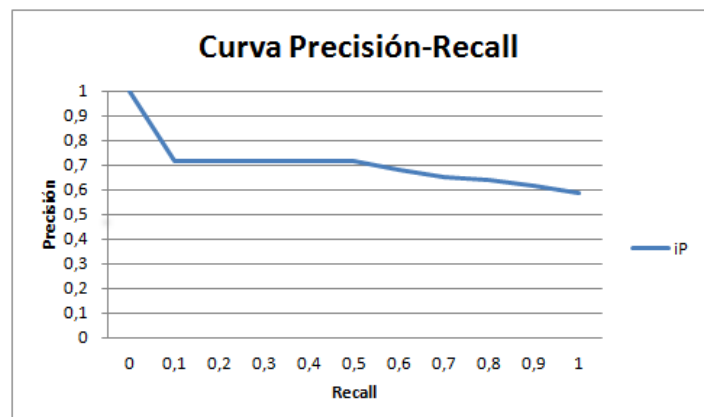
Comentando este tópico en concreto, observamos que el motor nos devuelve los 59 documentos relevantes de un total de 500 recuperados.

Se inicia con una precisión buena del 60% en el cut-off de 5 documentos, lógicamente con un recall bajo al principio.

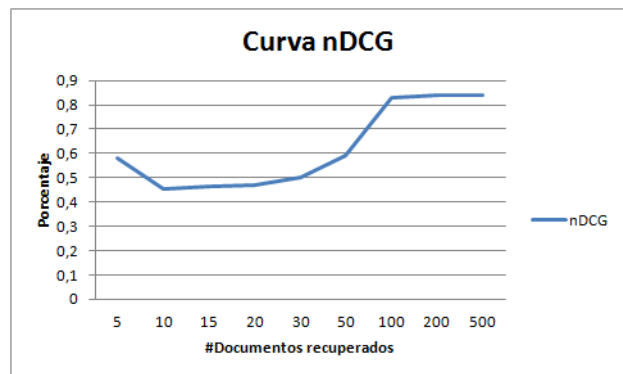
La precisión disminuye en los 10 primeros documentos al 40% pero vuelve a subir al 53,33% cuando recuperamos 15 documentos, experimentando una oscilación en la precisión que se repite, subiendo hasta una precisión mejor que la inicial en los primeros 50 documentos, llegando al 70%. Por último vuelve a bajar hasta el 11,80% al finalizar el recorrido de los 500 documentos.



En cuanto al recall, se mantiene en aumento constante, empezando con un 7% en los 10 primeros documentos, pero experimentando un crecimiento muy notable, sobretodo entre los cut-off 50 y 100, ya que pasa del 59,32% del recall al 98,31%, obteniendo casi el total de documentos relevantes. Al llegar a los 200 documentos vemos que el recall ya ha llegado al 100%.



Analizando los gráficos, sobretudo la curva nDCG, podemos observar como los resultados del topic sufren ciertas oscilaciones. Inicialmente durante los primero 50 documentos los resultados se encuentran cercanos al 50 %, lo cual nos indica que en general el motor presenta la información medianamente de forma correcta. Hacia el final de la gráfica vemos que este valor se dispara llegando prácticamente al 80 al superar los 200 documentos.



En esta consulta en particular podemos decir que los resultados son bastante positivos y obtenemos unos resultados bastante interesantes con una precisión media bastante notable del 63,24 %.