

Variaciones de la función de similitud para TextRank

TextRank

Autorizaron a un camión sin chofer a circular por las rutas de Estados Unidos Daimler anunció que su prototipo Freightliner Inspiration se convirtió er autónomo en recibir la autorización para recorrer las rutas y autopistas de junto al tránsito actual. La automotriz alemana dijo que el vehículo, que l prueba de forma previa en Alemania en un circuito urbano cerrado, permitirá de transporte ahorren combustible y puedan tener una flota más segura dura Para obtener la licencia que le permite recorrer los caminos de Nevada, forma autónoma en condiciones reales, Daimler dijo que el Freightliner cumplir más de 16 mil kilómetros de prueba. Como antecedente, en este est autorización para su vehículo autónomo. La autonomía del camión de Daimler es parcial, ya que sólo se habilita esta

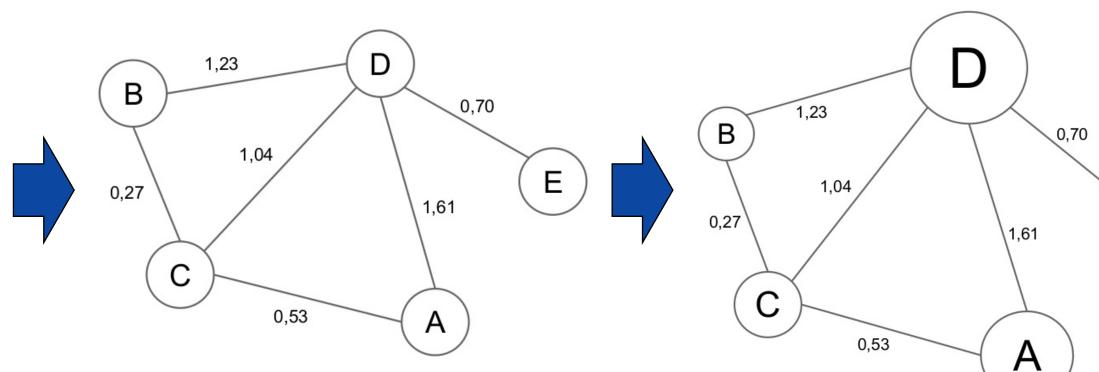
está en una ruta o autopista, manteniendo una distancia prudencial respecto

Ante un obstáculo o dificultad que el sistema no pueda resolver, el Freightline

- 2. daiml anunc prototip freightlin inspiration convirt automotriz aleman vehicul hab puest prueb form
 - ahorr combust pued flot mas segur trayect obten licenci permit recorr camin nev unid form : mas mil kilometr prueb

antecedent googl logr autoriz vehicul autonom

autoriz camion chof circul rut unid



Autorizaron a un camión sin chofer a circular por las rutas de Estados Unidos Daimler anunció que su prototipo Freightliner Inspiration se convirtió er autónomo en recibir la autorización para recorrer las rutas y autopistas del junto al tránsito actual. La automotriz alemana dijo que el vehículo, que l' prueba de forma previa en Alemania en un circuito urbano cerrado, permitirá de transporte ahorren combustible y puedan tener una flota más segura dura Para obtener la licencia que le permite recorrer los caminos de Nevada, forma autónoma en condiciones reales, Daimler dijo que el Freightliner I cumplir más de 16 mil kilómetros de prueba. Como antecedente, en este est autorización para su vehículo autónomo. La autonomía del camión de Daimler es parcial, ya que sólo se habilita esta está en una ruta o autopista, manteniendo una distancia prudencial respecto

sin adelantarse a otros conductores más lentos. Ante un obstáculo o dificultad que el sistema no pueda resolver, el Freightline una alerta al chofer y detendrá lentamente la marcha si no recibe respue

1. Texto original

sin adelantarse a otros conductores más lentos.

2. Preprocesamiento

- Separación en oraciones
- Filtro de palabras
- Lematización

3. Armado del grafo

- Uso de la función de similitud
- Se le asigna un puntaje a cada par de oraciones

4. PageRank

- Se aplica una variación del algoritmo para obtener las oraciones más representativas del texto.

5. Texto resumido

Propuestas

Función de similitud original:

Cuenta las palabras en común entre cada par de oraciones.

$$Similitud(S_i, S_j) = \frac{|\{w_k | w_k \in S_i \& w_k \in S_j\}|}{log(|S_i|) + log(|S_j|)}$$

Subcadena común más larga

Devuelve la longitud de la subcadena más larga presente en las dos oraciones.

Similitud coseno

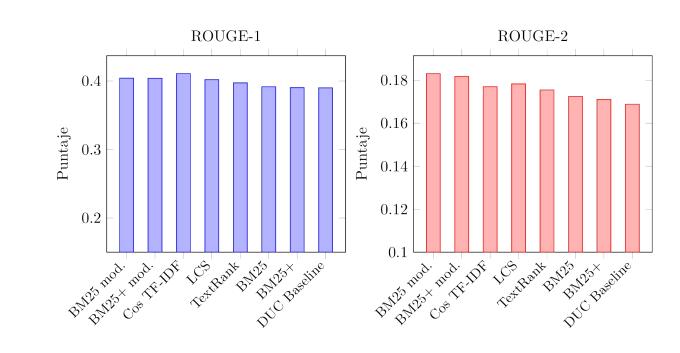
Usando el modelo TF-IDF representa cada oración como un vector y computa el resultado como el coseno del ángulo que comprenden

BM25

Implementa una variación de la famosa función BM25 para recuperación de información; que no permite que haya resultados que tomen valores negativos.

Evaluación y resultados

Aca iria todo el contenido del texto o lo que sea



Método	ROUGE-1	ROUGE-2	ROUGE-SU4	Mejora
BM25 ($\varepsilon = 0.25$)	0,4042	0,1831	0,2018	2,92%
$BM25 + (\varepsilon = 0.25)$	0,404	0,1818	0,2008	2,60%
Similitud coseno	0,4108	0,177	0,1984	$2{,}54\%$
$BM25 + (IDF = \log(N/N_i))$	0,4022	0,1805	0,1997	$2,\!05\%$
$BM25 (IDF = \log(N/N_i))$	0,4012	0,1808	0,1998	$1{,}97\%$
Subcadena común	0,402	$0,\!1783$	0,1971	1,40%
$BM25 + (\varepsilon = 0)$	0,3992	0,1803	0,1976	$1{,}36\%$
$BM25 (\varepsilon = 0)$	0,3991	0,1778	0,1966	0,89%
TextRank	$0,\!3983$	$0,\!1762$	$0,\!1948$	_
BM25	0,3916	0,1725	0,1906	-1,57%
BM25+	0,3903	0,1711	0,1894	-2,07%
Referencia DUC	0,39	0,1689	0,186	-2,84 %

Contribución

Aca iria todo el contenido del texto o lo que sea



Conclusión

Aca iria todo el contenido del texto o lo que sea

Referencias

Aca iria todo el contenido del texto o lo que sea



Autores

Aca iria todo el contenido del texto o lo que sea