Caso grupal: aplicaciones prácticas de PLN en investigación y en el mercado final

Objetivos

Con esta actividad grupal se pretende que los alumnos obtengan la capacidad de trabajar en equipo y de mantener una interlocución adecuada y enriquecedora con sus colaboradores, además de que adquieran la capacidad de trabajo independiente y autónomo. Es muy importante que los alumnos sepan aplicar los conocimientos adquiridos y desarrollen su capacidad para resolver problemas en entornos nuevos o poco conocidos dentro de contextos más amplios o multidisciplinares relacionados con el procesamiento del lenguaje natural. Por ello, en esta actividad grupal se va a trabajar en una recopilación bibliográfica sistematizada de las diferentes aplicaciones prácticas del Procesamiento del Lenguaje Natural. Se trata por tanto de obtener una visión en profundidad de los diferentes problemas y de las tendencias en el estado del arte, además de extraer sus propias conclusiones sobre este tema tan relevante del procesamiento del lenguaje natural.

Descripción

En esta actividad debes investigar en grupo sobre una de las cinco aplicaciones prácticas de Procesamiento de Lenguaje Natural explicadas en el tema 9 de la asignatura. Los temas disponibles son:

* Traducción automática
* Autocompletado y generación automática de resúmenes
* Análisis de sentimientos
* Question Answering
* Reconocimiento automático del habla y Text-to-speech

Para escoger los temas sobre los que trabajará el grupo, utilizad el script:

* *Random select from list.ipynb*

Se trata de un Notebook de Jupyter que, a partir de los nombres de los miembros del grupo, selecciona **uno** de los cinco temas posibles. Simplemente sigue las instrucciones contenidas en el Notebook.

Cuando hayáis obtenido el tema sobre el que trabajar, deberéis realizar una búsqueda bibliográfica en bases de datos indexadas disponibles de forma pública (Google Scholar), o bien a través de la biblioteca virtual del campus (Web of Science, Scopus).

Obtened, para el tema a trabajar, **tres** referencias bibliográficas diferentes que cumplan con las siguientes características:

* Deben ser posteriores a 1995.
* Sus fechas de publicación deben estar separadas por al menos 4 años.
* Deben describir trabajos experimentales con resultados empíricos.

Para las referencias escogidas, aseguraos de poder descargar el artículo completo, pues en el resto de la actividad deberás trabajar sobre dicho texto.

Para cada una de las referencias deberéis realizar las siguientes tareas:

* Describir la tecnología que se ha utilizado en el experimento descrito.
* Describir la metodología experimental. ¿En qué consistió el experimento?, ¿Qué corpus se utilizó?
* Describir los resultados obtenidos: ¿qué métricas se utilizaron para validar el experimento?, ¿qué resultados se obtuvieron? Valora si dichos resultados son adecuados para el ámbito de aplicación.

Por otra parte, para cada el tema sobre el que estáis trabajando, deberéis realizar una búsqueda en la web de las soluciones comerciales disponibles en la actualidad. Escoged una de entre todas las soluciones encontradas e indicad la respuesta a las siguientes preguntas:

* ¿Qué empresa es el proveedor?, ¿Cuándo fue fundada dicha empresa?, ¿Qué otras soluciones ofrece dicha empresa en el ámbito de la Inteligencia Artificial?
* ¿A qué perfil de cliente va dirigido el producto?
* ¿Cuál es el modelo de negocio de la empresa con respecto al producto?
* ¿Con qué tecnología de Procesamiento de Lenguaje Natural está construido el producto?

**Nota**: es posible que no toda esta información de los productos comerciales esté disponible públicamente. En ese caso, simplemente expresad vuestra opinión al respecto y argumentadla adecuadamente.

Finalmente, responde a las siguientes preguntas, de nuevo para el tema analizado:

* Con relación a los artículos de investigación experimental encontrados, ¿observas una evolución en las tecnologías, metodologías o los resultados obtenidos?
* ¿Se observa una conexión entre la investigación y las soluciones comerciales?
* Según vuestro criterio, ¿son satisfactorios los resultados obtenidos en la actualidad, o existen retos pendientes? ¿Cuáles son dichos retos?

En todas las preguntas del ejercicio, las respuestas deberán estar debidamente argumentadas.

Qué entregar

* El fichero de selección de los temas, ‘*Random select from list.ipynb’*, con los nombres de los integrantes del equipo y el resultado de la ejecución.
* Un informe en Word con los elementos habituales (portada, índices, numeración de páginas, encabezados y pies de página, introducción, conclusiones y **bibliografía**) que contenga los análisis solicitados.
* Un fichero zip con los archivos PDF (o formato correspondiente) de los tres artículos analizados.
* En el informe con las respuestas debes indicar claramente todas las fuentes consultadas y referenciarlas correctamente según la normativa APA.
* Extensión máxima 10 páginas, incluyendo todos los apartados.
* Si se detecta plagio en el informe o actividades que copien parte de las respuestas de otro alumno, todos los alumnos involucrados obtendrán una calificación para la actividad de cero puntos.

**Extensión** máxima: 15 páginas, fuente Calibri 12, interlineado 1,5.

**Rúbrica**

|  |  |  |  |
| --- | --- | --- | --- |
| Etiquetado morfosintáctico (valor real: puntos) | Descripción | Puntuación máxima  (puntos) | Peso  % |
| Referencias adecuadas | Para cada referencia aportada, se trata de un trabajo experimental sobre el tema objeto de estudio en el que se pueden observar resultados empíricos. | 2 | 20% |
| Referencias bien analizadas | El análisis de las referencias en cuanto a tecnologías utilizadas, metodología experimental y resultados obtenidos refleja adecuadamente el contenido de los artículos y está planteado de una forma clara y ordenada. | 4 | 40% |
| Soluciones comerciales adecuadas | La solución comercial identificada es, efectivamente, un producto que utiliza PLN para resolver el tema objeto de estudio. | 1 | 10% |
| Soluciones comerciales bien analizadas | El análisis de las soluciones comerciales en cuanto a las preguntas planteadas en el enunciado de la actividad es coherente, está bien argumentado y no presenta errores de redacción. | 2 | 20% |
| *Preguntas finales* | Las respuestas otorgadas a las preguntas finales son coherentes, están bien argumentadas y no presenta errores de redacción. | 1 | 10% |
|  |  | **10** | **100 %** |