Caso grupal: Análisis de sentimientos

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 investigar sobre el problema del **Análisis de Sentimientos** para que los alumnos puedan obtener una visión en profundidad del problema y de las tendencias, además de extraer sus propias conclusiones sobre este tema tan relevante del procesamiento del lenguaje natural.

Descripción

En esta actividad debes investigar sobre el **Análisis de Sentimientos**, también conocido como Minería de Opiniones. Esta disciplina del procesamiento del lenguaje natural se enfoca en la detección de información subjetiva en el contenido de un texto y su clasificación. Por lo tanto, el análisis de sentimientos permite la identificación del sentir u opinión expresada en un texto, ya sea en un Tweet, en la reseña de una película, en la revisión de un producto…

Para investigar sobre el tema del **Análisis de Sentimientos** se propone que estudies el Emotion Mining Toolkit (EMTk) una herramienta creada recientemente por investigadores de la Università degli Studi di Bari Aldo Moro en Italia.

EMTk se presentó en SEmotion 2019, la conferencia 2019 IEEE/ACM 4th International Workshop on Emotion Awareness in Software Engineering. El artículo donde se describe esta herramienta está publicado en el acta de la conferencia.

Accede al artículo a través del aula virtual o desde la siguiente dirección web:

<https://www.researchgate.net/publication/335644237_EMTk_-_The_Emotion_Mining_Toolkit>

EMTk está publicado bajo licencia de código abierto (MIT License) y está disponible para su descarga en github.

Accede al recurso a través del aula virtual o desde la siguiente dirección web:

<https://collab-uniba.github.io/EMTk/>

Una vez analizada la documentación y el código de EMTk y consultadas otras fuentes bibliográficas que puedas requerir para ampliar el tema del análisis de sentimientos, debes resolver las siguientes cuestiones:

* + Describe como realizarías el análisis de sentimientos utilizando una herramienta como EMTk. Indica el proceso que se sigue para realizar el análisis de sentimientos a partir de un texto. Complementa tu respuesta con un diagrama de bloques (o una representación gráfica similar) donde se muestre este proceso.
  + Define con tus propias palabras los conceptos de “emotion mining” (reconocimiento de emociones) y “polarity mining” (clasificación de la polaridad) en el ámbito del procesamiento del lenguaje natural. Identifica las diferencias entre estos dos conceptos del análisis de sentimientos. Utiliza ejemplos basados en EMTk para argumentar tu respuesta sobre las diferencias de estos dos términos.
  + Identifica las diferentes técnicas de procesamiento del lenguaje natural que implementa EMTk para resolver el reconocimiento de emociones y la clasificación de la polaridad, que has explicado en el apartado anterior. Explica con tus propias palabras el funcionamiento de cada una de estas técnicas. ¿Cuáles consideras que podrían ser las mejores técnicas? Indica el por qué y susténtalo con referencias a artículos o publicaciones científicas del estado del arte que lo demuestren.
  + Describe las librerías de terceros que usa EMTk. Identifica las funcionalidades de estas librerías que usa EMTk y para que procesos se usan.
  + Describe los *datasets* que incluye EMTk y que han sido usados para entrenar los modelos. Explica claramente el proceso que se ha usado para transformar los *datasets* que permiten el reconocimiento de emociones en *datasets* que permitan la clasificación de la polaridad. Busca artículos o publicaciones científicas que muestren cómo han abordado este problema otros investigadores y describe sus soluciones.

En el informe con las respuestas debes indicar claramente todas las fuentes consultadas y referenciarlas correctamente según la normativa APA.

Criterios de evaluación

* Se debe entregar un informe con la descripción detallada de la actividad y las respuestas a todo lo que se solicita en el enunciado.
* Se valorará la explicación clara y argumentada.
* El informe de la actividad debe tener el formato marcado en el apartado **Extensión máxima**. Esta no se puede sobrepasar en ningún caso, por lo tanto, si el informe presentado tiene más páginas que el valor establecido, solo se corregirán y evaluarán las N primeras páginas. Por ejemplo, si se presenta un documento con 10 páginas y la extensión máxima es de 8 páginas, se evaluará solamente el contenido de las páginas 1 a la 8, por lo que todo el contenido presentado en las páginas 9 y 10 no contará para la nota de la actividad.
* 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: 8 páginas, fuente Calibri 12, interlineado 1,5.

**Rúbrica**

|  |  |  |  |
| --- | --- | --- | --- |
| Etiquetado morfosintáctico (valor real: puntos) | Descripción | Puntuación máxima  (puntos) | Peso  % |
| Proceso análisis de sentimientos | Explicar de forma clara el proceso de análisis de sentimientos y aportar el diagrama representativo. | 1.5 | 15% |
| “emotion mining” y “polarity mining” | Describir de forma correcta “emotion mining” y “polarity mining” e identificar razonadamente sus diferencias. | 1 | 10% |
| Módulos de EMTk | Identificar correctamente las técnicas que implementa EMTk y razonar cuales son las mejores | 2.5 | 25% |
| Librerías usadas por EMTk | Nombrar las librerías que usa EMTk e identificar correctamente para que se usan. | 2.5 | 25% |
| *Datasets* en EMTk | Describir de forma clara los datasets usados en EMTk y explicar el proceso de transformación. | 2.5 | 25% |
|  |  | **10** | **100 %** |