

EJEMPLOS DE PREGUNTAS DE EVALUACIÓN

CATEGORÍAS GRAMATICALES, TAGGING

1. Indique la categoría gramatical de cada palabra en el siguiente fragmento de texto. Utilice la clasificación en 8 categorías vista en el curso.

Gradualmente, el hermoso universo fue abandonándolo; una terca neblina le borró las líneas de la mano, la noche se despobló de estrellas, la tierra era insegura bajo sus pies. Todo se alejaba y se confundía.

2. ¿Cuál es la diferencia entre etiquetado morfo-sintáctico y tagging? ¿Qué consecuencias para el tagging tiene una de las versiones de la ley de Zipf?

ESTRUCTURAS SINTÁCTICAS, ÁRBOLES DE CONSTITUYENTES Y DE DEPENDENCIAS

a) Realice el análisis en constituyentes y en dependencias de la siguiente oración :

El perro tuvo una lesión en el omóplato izquierdo a raíz de una garrafa que cayó desde la Tribuna Ámsterdam.

b) ¿El árbol de dependencias de la parte a) es proyectivo ? Fundamente.

c) ¿Qué herramientas para análisis sintáctico conoce. Descríbalas brevemente.

EXTRACCIÓN DE INFORMACIÓN : ENTIDADES CON NOMBRE

1. Considere el siguiente ejemplo

El perro tuvo una lesión en el omóplato izquierdo a raíz de una garrafa que cayó desde la Tribuna Ámsterdam y que también hirió al oficial que lo llevaba. "La lesión (de Duque) fue un traumatismo a nivel de la cruz, eso ocasionó un hematoma y ahora está siendo tratado con antiinflamatorios y analgésicos", dijo a El Observador el capitán John Severo, jefe del K-9, la unidad canina de la Guardia Republicana. Duque es uno de los perros que trabaja controlando disturbios y su función es más "psicológica que represiva", dijo Severo.

a) Indique, en el texto anterior, cuáles son las entidades con nombre y cuál es su tipo semántico. Utilice las clases PERSONA, ORGANIZACIÓN, LUGAR y agregue clases si lo considera necesario.

b) ¿En qué consiste la “wikificación”? ¿Por qué puede ser necesaria? Ilustre con el texto anterior, discuta si le parece pertinente en ese caso. ¿Qué procesos serían de interés para ese texto?

c) ¿Qué representan los esquemas BIO y BILOU? Etiquete según el esquema BIO la última oración del texto ejemplo.

d) Indique qué atributos utilizaría para la extracción de entidades con nombre siguiendo un método de aprendizaje supervisado.

2. Considere el siguiente ejemplo

Leonardo Cipriani, presidente de ASSE, redefinió en las últimas horas la situación del Hospital de Bella Unión, en Artigas. En diálogo con el periodista Leonardo Sarro, de Radio Monte Carlo, explicó que no se trataba en sentido estricto de un brote -término que él mismo empelara ayer- debido a las características de los casos presentados.

El jerarca informó que cerca de la pasada medianoche se diagnosticaron seis nuevos casos que "no están relacionados al hospital sino a la familia" de uno de los diagnosticados anteriormente, y que "parte de esa familia vive en Bella Unión, y otra parte en Barra de Quaraí", ya en territorio brasileño.

a) Indique, en el texto anterior, cuáles son las entidades con nombre y cuál es su tipo semántico. Utilice las clases PERSONA, ORGANIZACIÓN, LUGAR y agregue clases si lo considera necesario.

b) La salida de un reconocedor de entidades con nombre para el texto anterior genera la siguiente salida. Las entidades con nombre reconocidas están marcadas con fondo de color :

XXXX Persona

XXXX Organización

XXXX Otra

XXXX Lugar

Leonardo Cipriani, presidente de ASSE, redefinió en las últimas horas la situación del Hospital de Bella Unión, en Artigas. En diálogo con el periodista Leonardo Sarro, de Radio Monte Carlo, explicó que no se trataba en sentido estricto de un brote -término que él mismo empelara ayer- debido a las características de los casos presentados.

El jerarca informó que cerca de la pasada medianoche se diagnosticaron seis nuevos casos que "no están relacionados al hospital sino a la familia" de uno de los diagnosticados anteriormente, y que "parte de esa familia vive en Bella Unión, y otra parte en Barra de Quaraí", ya en territorio brasileño.

Calcule Precision y Recall globales y por Clase.

EXTRACCIÓN DE INFORMACIÓN : RELACIONES

1. Anotación en textos

La célula es la unidad funcional de los tejidos. Las células forman los tejidos, los tejidos con características parecidas forman los órganos y los órganos con funciones similares forman los sistemas o aparatos.

Las células presentan un citoplasma muy compartimentado, con orgánulos separados o interconectados, limitados por membranas biológicas que son de la misma naturaleza esencial que la membrana plasmática. El núcleo es solamente el más notable y característico de los compartimentos en que se divide el protoplasma, es decir, la parte activa de la célula. En el protoplasma distinguimos tres componentes principales, a saber, la membrana plasmática, el núcleo y el citoplasma, constituido por todo lo demás. Las células están dotadas de un citoesqueleto complejo, muy estructurado y dinámico, formado por microtúbulos y diversos filamentos proteicos. Además puede haber pared celular, que es lo típico de plantas, hongos y protistas pluricelulares, o algún otro tipo de recubrimiento externo al protoplasma.

a) Encuentre al menos 3 instancias de relación parte-todo en el texto anterior. Identifique los argumentos y un patrón sintáctico para la relación.

b) Explique en qué consiste la supervisión distante aplicada a un problema de extracción de relaciones. Desarrolle un ejemplo para el texto anterior y pares según uno de los patrones de la parte

LEY DE ZIPF

Conteste las siguientes preguntas fundamentando brevemente.

1 - ¿Qué relación involucra la ley de Zipf?

2 - ¿Cuál es (aproximadamente) el porcentaje de ocurrencia de términos que ocurren una sola vez en un texto.

3- ¿Qué consecuencias tiene, para el manejo de semántica léxica en PLN, la propiedad del punto anterior?

MEDIDAS DE TEORÍA DE LA INFORMACIÓN

1. Sea X la variable aleatoria que representa el experimento de tirar una moneda 2 veces, el conjunto de valores de X es $\{CC, NN, CN, NC\}$, donde C representa Cara y N número

a) Plantee la entropía de X si cara y número son equiprobables

b) Suponga que la probabilidad de Cara es $\frac{3}{4}$. ¿Cómo se compara la entropía de X con la de la parte anterior ?

2. ¿En qué consiste y para qué se usa la medida PMI (*Pointwise Mutual Information*)?

3. ¿En qué consiste la entropía cruzada (*Cross Entropy*)? Dé un ejemplo de su uso.

GRAFOS DE CONOCIMIENTO

1. ¿Cómo es el esquema de organización de Wikidata ?
2. Mencione y dé ejemplos para 3 tipos de relaciones presentes en WordNet.

SEMÁNTICA VECTORIAL Y REDES NEURONALES

- 1- ¿Cómo se representa la semántica léxica mediante vectores ?
- 2- ¿Qué es un modelo disperso? Indique un ejemplo.
- 3- ¿Qué es un modelo denso? Indique un ejemplo.
- 4- Explique en qué consisten las siguientes representaciones:
 - a- one-hot
 - b- skip-gram con negative sampling