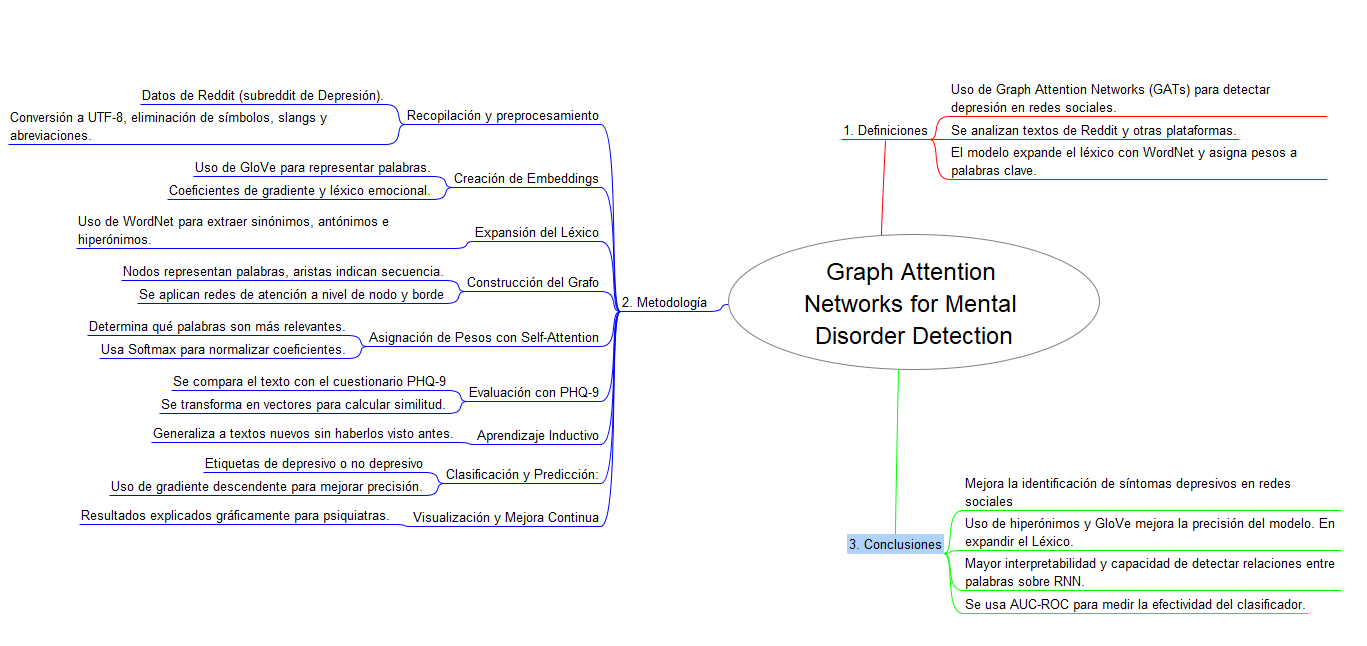
Daniel Alberto Urrego León

Juan David García Acevedo

Guion sustentación **Graph Attention Networks for Mental Disorder Detection**



**TABLA DE CONTENIDO**

[1. Introducción 1](#_Toc190906819)

[2. Definiciones 2](#_Toc190906820)

[3. Metodología 2](#_Toc190906821)

[3.1. Recopilación y preprocesamiento de datos 2](#_Toc190906822)

[3.2. Creación de embeddings 2](#_Toc190906823)

[3.3. Expansión del léxico emocional 3](#_Toc190906824)

[3.4. Construcción del grafo 3](#_Toc190906825)

[3.5. Asignación de pesos con Self Attention 3](#_Toc190906826)

[3.6. Puntuación con cuestionarios psicométricos 4](#_Toc190906827)

[3.7. Aprendizaje inductivo GATs 4](#_Toc190906828)

[3.8. Clasificación y predicción 4](#_Toc190906829)

[3.9. Visualización de resultados 4](#_Toc190906830)

[3.10. Mejora Continua del modelo 4](#_Toc190906831)

[4. Conclusiones 5](#_Toc190906832)

1. **Introducción**

El artículo habla de un método nuevo para detectar la depresión en las redes sociales mediante Grafos de atención (GATs). El modelo que se propone analiza el lenguaje utilizado en como las personas se expresan en sus redes para identificar síntomas depresivos. Lo que hace especial el grafo de atención en comparación a sistemas RNN es que asigna pesos a cada palabra “Vecindario Lingüístico” con las emociones transmitidas. Luego expande el léxico usando hiperónimos para mejorar la precisión. Los autores hacen una experimentación fuerte y los resultados que obtienen demuestran que GATs supera a otros modelos para la clasificación de textos depresivos, y resultados interpretables. Y por ultimo se hace una reflexión sobre el potencial de la inteligencia artificial para facilitar tempranamente casos de suicidio y su aporte a la salud mental.

1. **Definiciones**

Este método se centra en utilizar el lenguaje de las personas en sus redes sociales, y comenzar a realizar una asignación de pesos a cada palabra para identificar síntomas depresivos.

1. **Metodología**
   1. **Recopilación y preprocesamiento de datos**

La metodología inicia planteando una recopilación de datos de la red Reddit en un hilo (Los hilos son como un portal de noticias y solo hay comentarios de temas afines al hilo) r/depression. Al obtener los datos de Reddit se hace una limpieza (usando técnicas de PLN) las cuales son:

* Convertir los textos a minúsculas.
* Eliminar espacios y caracteres no deseados.
* Reemplazar abreviaciones y expresiones informales por su equivalente estándar.

Con el fin de asegurar que el modelo trabaje con datos procesados y que no se presente ruido y haya una homogeneidad.

* 1. **Creación de embeddings**

Después de la recopilación y preprocesamiento, se representa el texto en forma de embeddings (Embeddings con GloVe), que son los vectores numéricos para capturar la semántica de las palabras. Un truco que hacen es la incorporación de léxicos emocionales, y esto le permite al modelo identificar la carga emocional de cada palabra en un determinado contexto. ¿Por qué? Porque en la detección no basta solo con analizar palabras aisladas como veíamos con Alejandro y Miguel reuniones pasadas, donde la palabra Matar pesa mucho en los modelos. Entonces la adición de léxicos emocionales ayuda al modelo a comprender el significado dentro del contexto de la frase.

* 1. **Expansión del léxico emocional**

WordNet es un sistema de referencia léxica, una gran base de datos con sinónimos, hipónimos, merónimos y otras cantidades de relaciones semánticas. Se utiliza esta base de datos para en cierta palabra enriquecer el modelo y su capacidad de léxico emocional. WordNet permite:

* Sinónimos, para capturar variaciones en que las personas expresan emociones.
* Antónimos, para identificar contrastes emocionales.
* Hiperónimos, que ayudan a generalizar términos y mejorar la detección de patrones lingüísticos.

Y con los embeddings GloVe se encuentran términos semánticamente similares, así mismo la base de palabras analizada aumenta.

* 1. **Construcción del grafo**

El GATs o grafo de atención es el corazón de este modelo. Para su construcción cada **nodo** representa una palabra. Los bordes son las relaciones semánticas entre palabras. Y es un grafo dirigido y acíclico, entonces las palabras están ordenadas según su secuencia en la frase. No todas las palabras tienen la misma relevancia entonces los autores presentan un mecanismo Self Attention o autoatención.

* 1. **Asignación de pesos con Self Attention**

Este proceso de autoatención permite que el modelo determine qué palabras son más importantes para tomar acción rápida con las frases y detectar pronta la ideación suicida. Cada palabra tiene su peso en el vecindario lingüístico, es decir, las palabras que la rodea. Se usa la función softmax para normalizar coeficientes de atención. Y esto hace que la importancia entre todas las palabras de la frase se distribuya. Por tanto, el modelo aprende a enfocarse en los términos más relevantes sin necesidad de procesar mucho texto.

* 1. **Puntuación con cuestionarios psicométricos**

Hay una herramienta clínica que se llama Cuestionario PHQ-9 y se usa para diagnosticar síntomas depresivos. Lo que hace el modelo es transformar los textos Reddit del usuario y el resultado del cuestionario PHQ-9 en vectores para calcular su similitud. Si esta es alta, significa que el usuario está expresando síntomas clínicos de depresión.

* 1. **Aprendizaje inductivo GATs**

El aprendizaje inductivo es que el modelo no solo clasifica textos basados en lo que ya ha vistos, puede también generalizar nuevos datos con palabras que no conoció durante el entrenamiento. Entonces el sistema esta preparado para datos que no conoce.

* 1. **Clasificación y predicción**

Hay unos embeddings pre-entrenados, y el modelo clasifica como **Depresivo** o **No depresivo**. Ya el sistema para clasificar se basa en probabilidad con **Gradiente Descendiente**.

* 1. **Visualización de resultados**

Como se dijo en la introducción, el modelo es interpretable por tanto, no solo predice si un usuario tiene depresión, sino que también explica cómo llegó a esa conclusión. Ya se realiza una visualización un poco especializada por que fue para que los psiquiatras y especialistas puedan ver qué palabras fueron las más como importantes en la clasificación que presenta el modelo. Entonces la colaboración entre la IA y los expertos de la salud mental es buena.

* 1. **Mejora Continua del modelo**

Ya para la mejora continua los autores pensaron en seleccionar oraciones con embeddings promedios y revisar su similitud para ajustar la clasificación del modelo. Ajustan la gradiente para suavizar pesos del grafo. Y solo ingresan nuevos términos y relaciones semánticas en el léxico emocional. Todo hace que la detección sea más precisa y efectiva.

1. **Conclusiones**

El artículo demuestra que GATs con expansión de léxico y aprendizaje inductivo es poderoso para la detección de ideación suicida y usuarios depresivos de Reddit. Lo bueno de la estrategia es que explica cómo llego a sus resultados. Y la reflexión sobre como la IA es potencialmente clave en el sector de la salud mental.