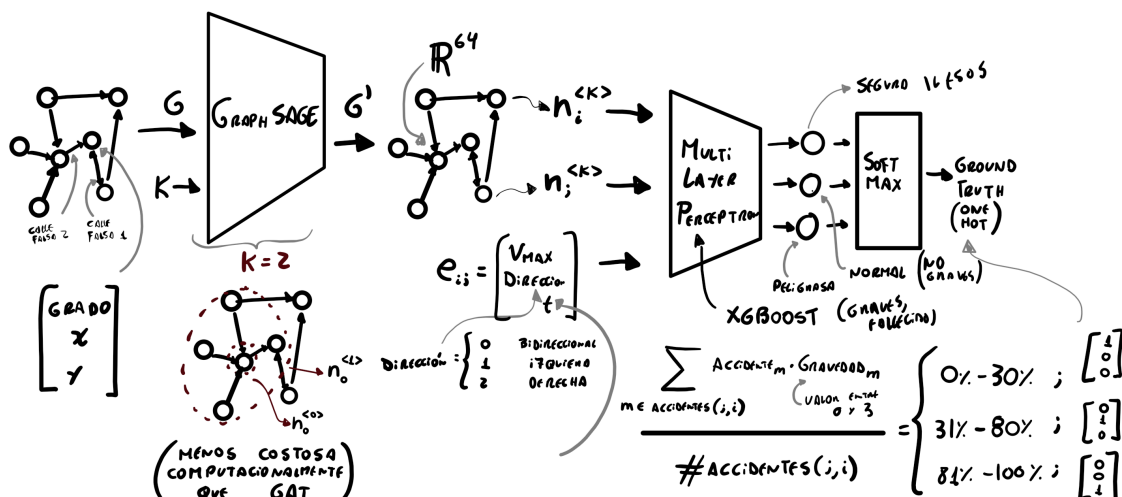


II C3641_avance_proyecto

Presentación

1. Comentar el cambio de proyecto (dataset era parcial y no publico la parte importante, pesado)
2. Contexto: Importancia por la seguridad vial, ayuda a planificación urbana, ayuda a tomar medidas en caso de que sea una calle insegura.
3. Graphical abstract: Hacer refernecia a los dos modelos y al cálculo de la gravedad de una calle.
4. Comentamos los datos recopilados: dataset sinieestros, dataset de los hospitales (especializada, urgencias, atención primaria, etc). Porqué se escogió el graphsage y no el gat y porqué el xgboost
5. Setting experimental: Código del colab
6. Próximos pasos: Implementar el modelo ya final, inclusión de la temporalidad en el modelo. Interpretar las variables determinantes en la gravedad: inclusión de algoritmo de explicabilidad.

GRAPHICAL ABSTRACT



<https://arxiv.org/pdf/1706.02216>

GRAPHSAGE

<https://arxiv.org/pdf/1706.02216>

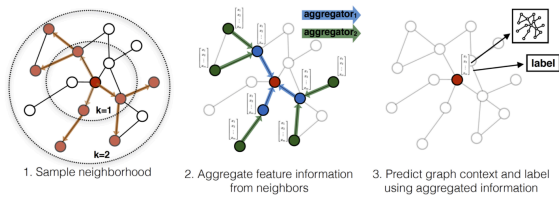


Figure 1: Visual illustration of the GraphSAGE sample and aggregate approach.

Algorithm 1: GraphSAGE embedding generation (i.e., forward propagation) algorithm

Input : Graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$; input features $\{x_v, \forall v \in \mathcal{V}\}$; depth K ; weight matrices $\mathbf{W}^k, \forall k \in \{1, \dots, K\}$; non-linearity σ ; differentiable aggregator functions $\text{AGGREGATE}_k, \forall k \in \{1, \dots, K\}$; neighborhood function $\mathcal{N}: v \rightarrow 2^{\mathcal{V}}$

Output : Vector representations z_v for all $v \in \mathcal{V}$

```

1  $h_v^0 \leftarrow x_v, \forall v \in \mathcal{V}$ ;
2 for  $k = 1 \dots K$  do
3   for  $v \in \mathcal{V}$  do
4      $h_{\mathcal{N}(v)}^k \leftarrow \text{AGGREGATE}_k(\{h_u^{k-1}, \forall u \in \mathcal{N}(v)\})$ ;
5      $h_v^k \leftarrow \sigma(\mathbf{W}^k \cdot \text{CONCAT}(h_v^{k-1}, h_{\mathcal{N}(v)}^k))$ 
6   end
7    $h_v^k \leftarrow h_v^k / \|h_v^k\|_2, \forall v \in \mathcal{V}$ 
8 end
9  $z_v \leftarrow h_v^K, \forall v \in \mathcal{V}$ 

```

- El k será de 2 a 3 porque tener mientras más vecindad más se parecerán los nodos (capturaran la misma información).
- Comentar: Usaremos GraphSAGE primero porque es computacionalmente viable, y luego trabajaremos con GAT para lograr tener mayor riqueza de la info de los vecinos
- Utilizar las categorías de fallecidos, graves, menos graves, leves e ileso para obtener el índice de la gravedad
- Los accidentes en las intersecciones se aproximan según su ubicación espacial a la arista más cercana.

DIFF CON GAT

Característica	GraphSAGE	GAT (Graph Attention)
Cómo agrega vecinos	Agregación simple (promedio, max-pooling).	Promedio ponderado "inteligente" (Atención).
Complejidad	Más simple y rápido de entrenar.	Más complejo y costoso computacionalmente.
"Inteligencia"	Trata a todos los vecinos (muestreados) por igual.	Aprende qué vecinos son más importantes.
Cuándo usarlo	Excelente como primer modelo (baseline). Rápido para grafos enormes.	Cuando la importancia de los vecinos varía mucho.