



SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

Clasificación de la demanda sanitaria mediante procesamiento de lenguaje natural (PLN)

Joan Vila Francés

joan.vila@uv.es



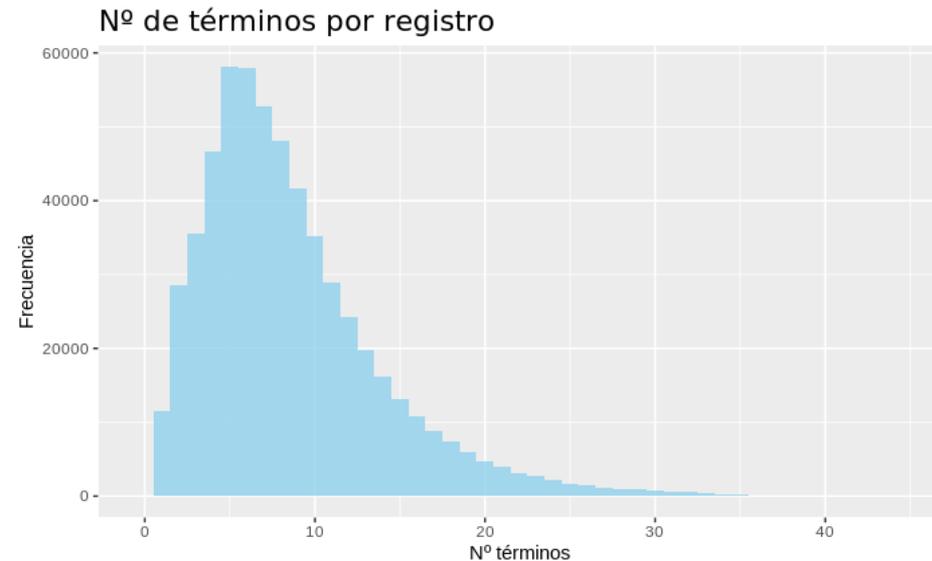
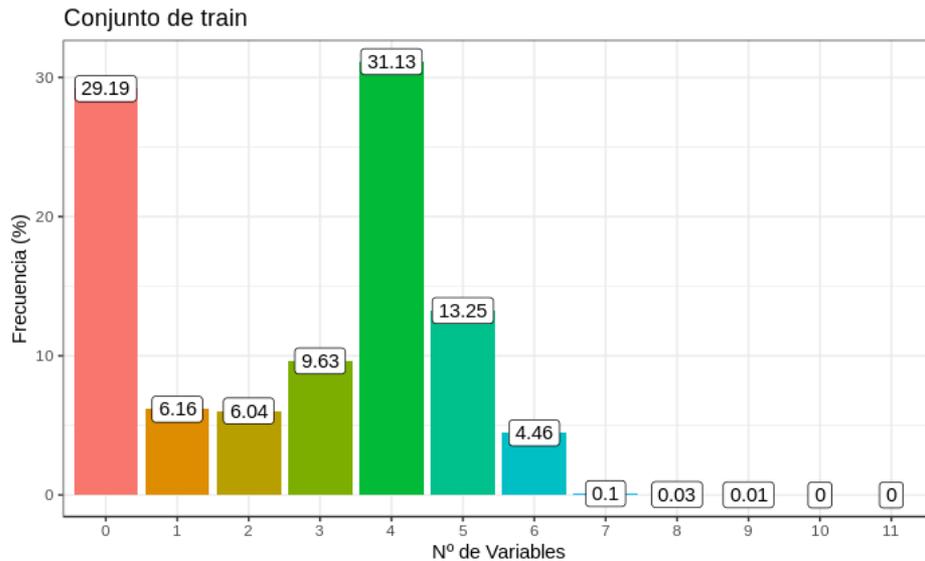
IDAL

Intelligent
Data
Analysis
Laboratory



SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

- *Información disponible*
 - Todos* los incidentes tienen un campo de observaciones relleno

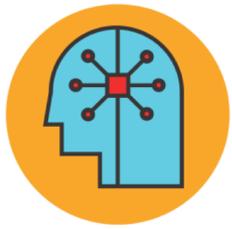




SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

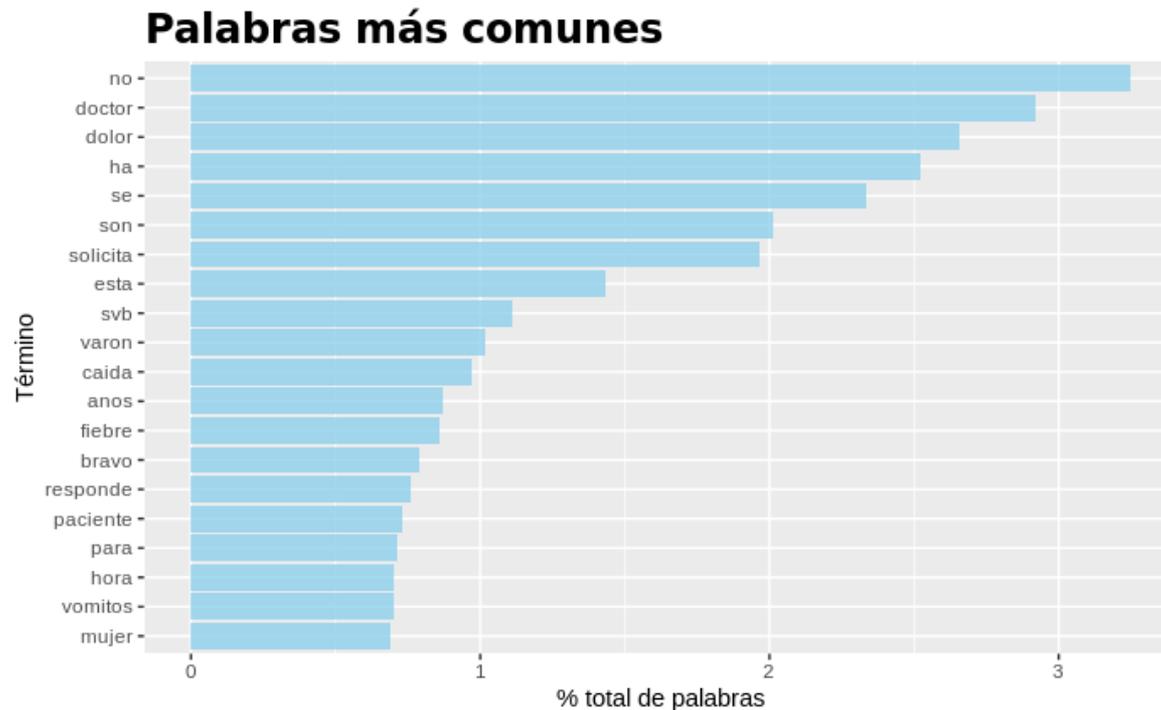
- *¿Cómo es el texto disponible*?*
 - Tenemos un vocabulario de unas 46k palabras diferentes.
 - Hay casi 5 millones de palabras.
 - El 90% de las palabras que aparecen son los 1030 términos más comunes.

**para el conjunto de entrenamiento*



SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

- *¿Cómo es el texto disponible*?*

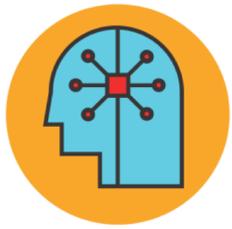


**para el conjunto de entranamiento*



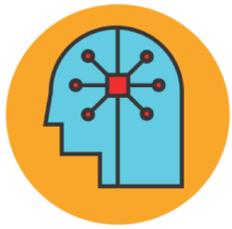
SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

- Análisis del campo de texto
 - Pre-procesado
 - Limpieza/reemplazo de texto
 - Corrección ortográfica
 - Obtención de características
 - Modelo de bolsa de palabras
 - Modelo secuencial
 - Entrenamiento de un clasificador



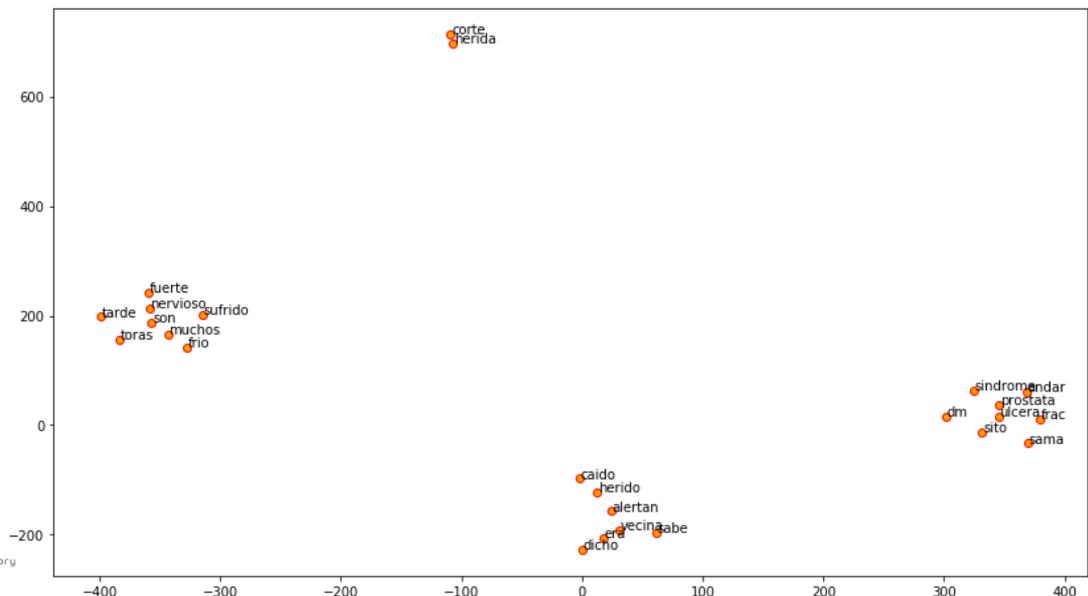
SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

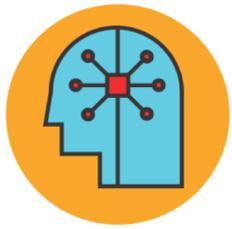
- Modelo de bolsa de palabras (BoW)
 - Contamos las apariciones de cada palabra en el texto.
 - Ponderamos por la "rareza" del término (inversa del nº de documentos que lo contienen) \Rightarrow **TF-IDF**
 - Se extiende a combinaciones de 2 palabras (*bigramas*)



SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

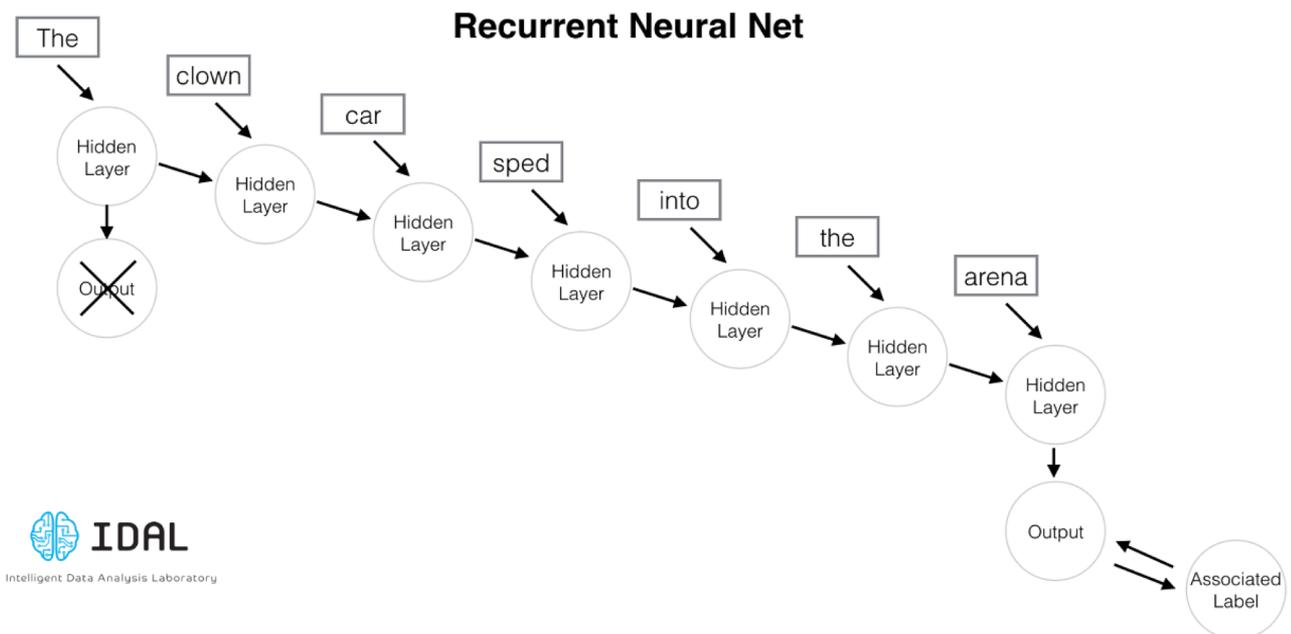
- Modelo secuencial
 - Obtenemos una descripción vectorial para cada palabra del vocabulario (*word embeddings*)

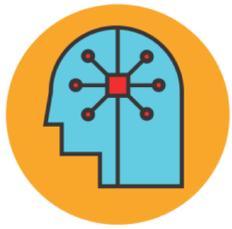




SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

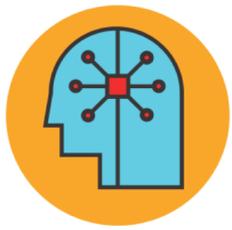
- Modelo secuencial
 - Alimentamos una red neuronal recurrente (LSTM) con los vectores de cada palabra del texto *secuencialmente*.





SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

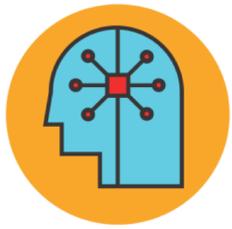
- Modelo TF-IDF vs. LSTM
 - El modelo TF-IDF es fácilmente interpretable \Rightarrow *formación, manual de estilo*
 - El modelo LSTM es más potente \Rightarrow *mayor tasa de acierto*



SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

- Resultados

	Riesgo	Competencia SES	Demorabilidad	Acierto total	Beneficio
TF-IDF	+11,4%	+0,2%	+15%	+6,3%	+3%
LSTM	+11,6%	+0,9%	+15,7%	+7,4%	+3,3%
TF-IDF global				+13,6%	+4,4%
LSTM global				+14,3%	+4,2%



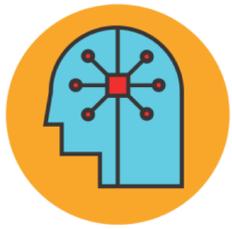
SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

- Resultados
 - Caso particular de incidentes sin variables clínicas registradas

Para RIESGO VITAL

Incremento del **24,64%** en tasa de acierto

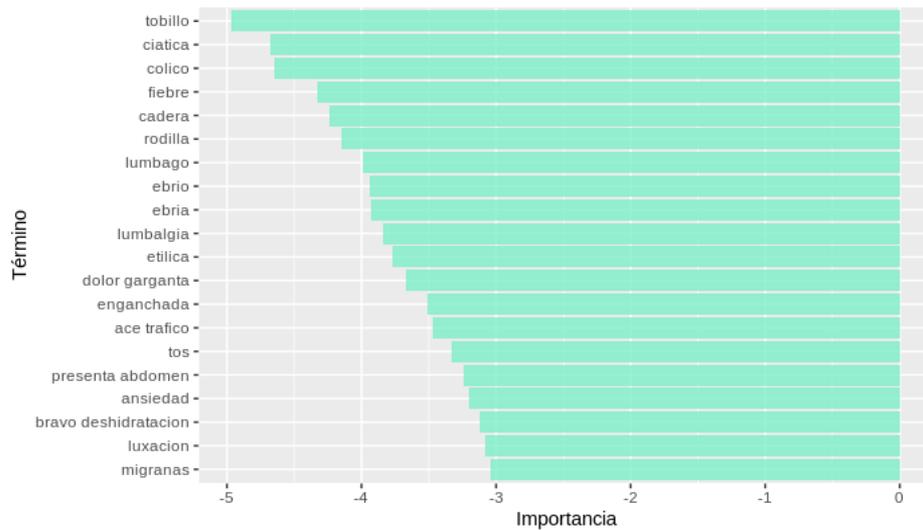
Incremento del **78,48%** en sensibilidad



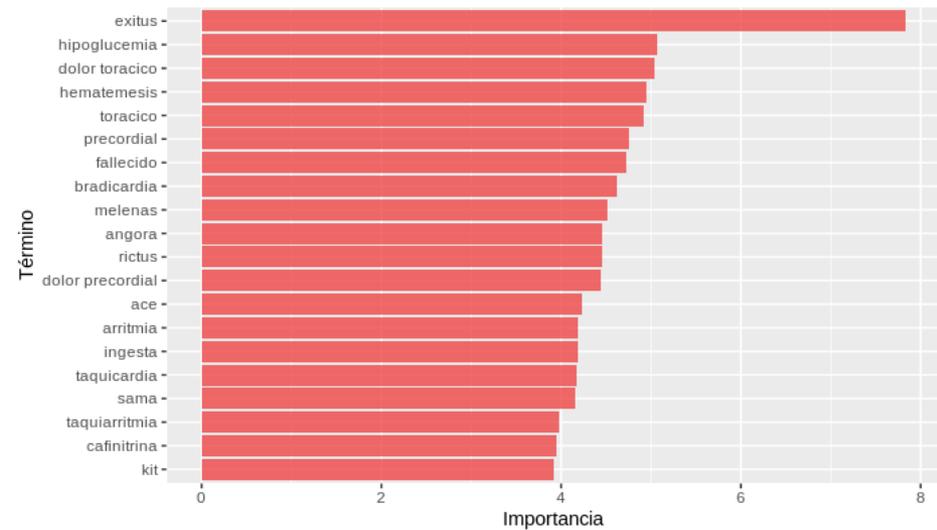
SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

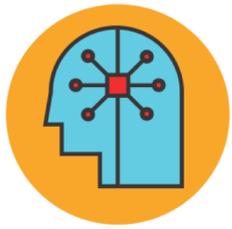
- Interpretación: riesgo vital

Términos más influyentes para NO Riesgo Vital



Términos más influyentes para Sí Riesgo Vital

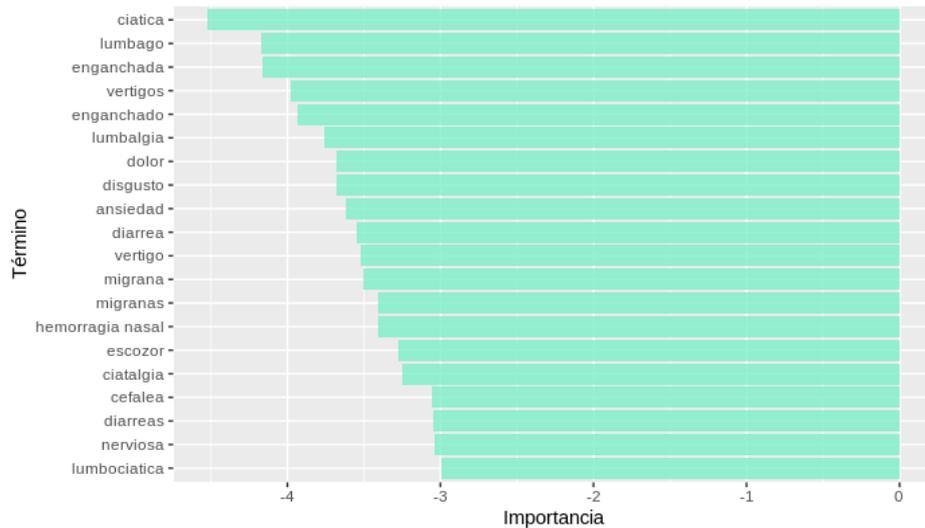




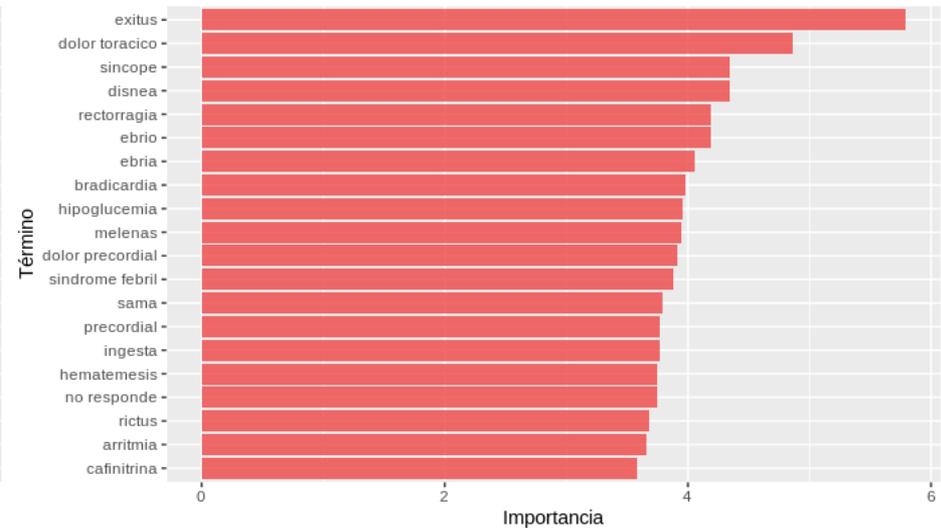
SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

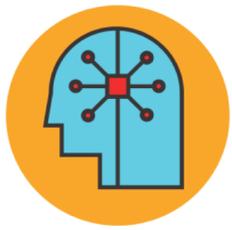
- Interpretación: competencia SES

Términos más influyentes para NO Competencia SES



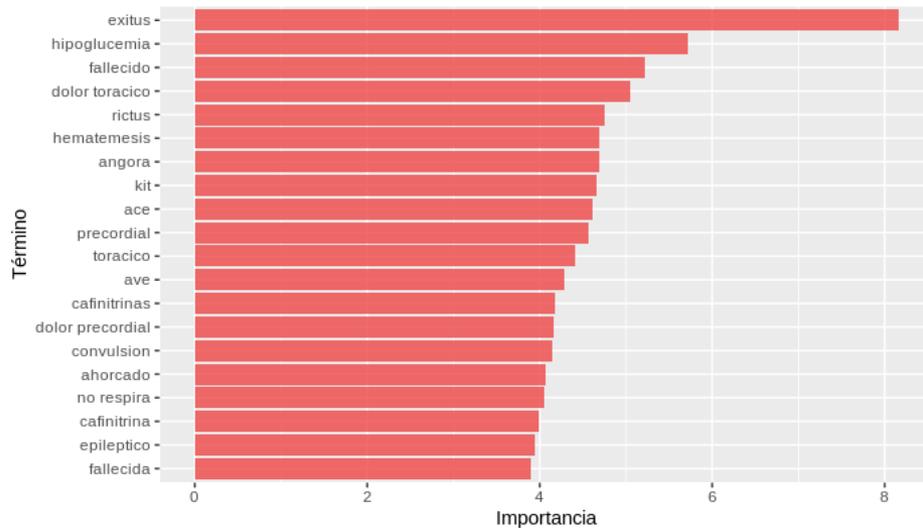
Términos más influyentes para Sí Competencia SES



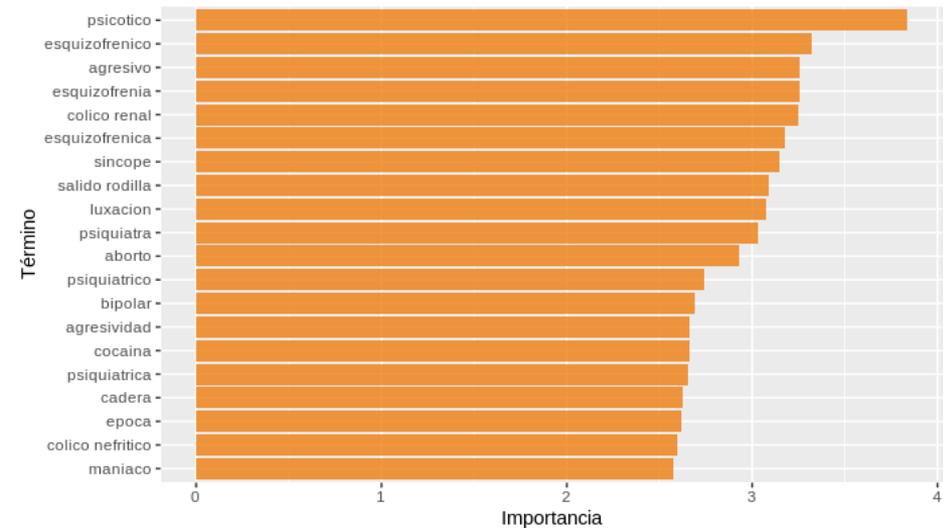


SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

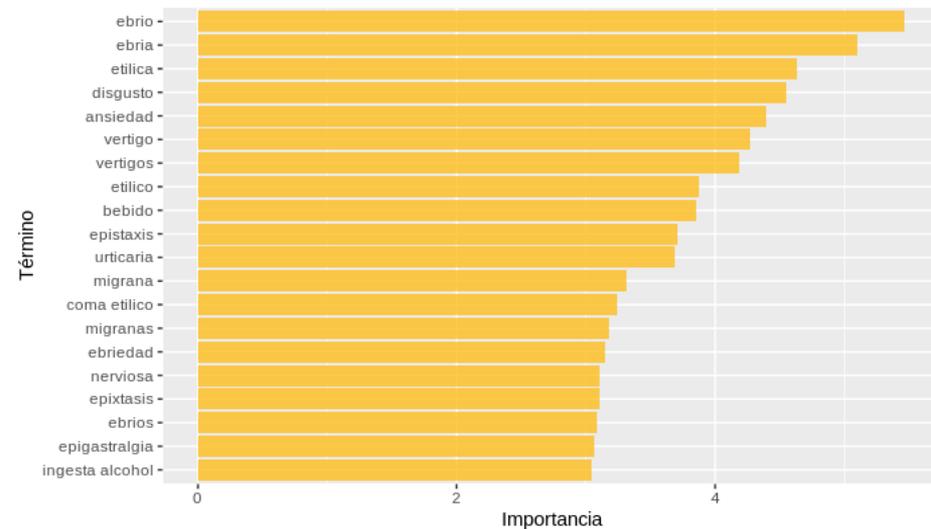
Términos más influyentes para Demorabilidad 0 (inmediata)



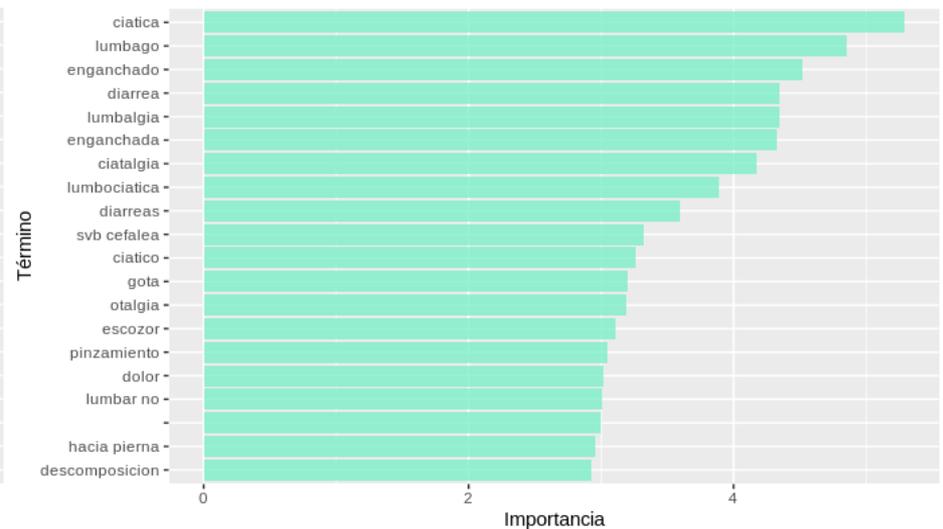
Términos más influyentes para Demorabilidad 2 (MINUTOS)

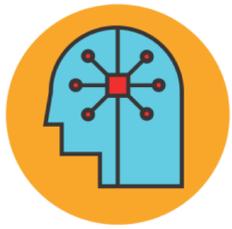


Términos más influyentes para Demorabilidad 3 (HORAS)



Términos más influyentes para Demorabilidad 4 (DÍAS)





SISTEMA EXPERTO DE CLASIFICACIÓN DE LA DEMANDA SANITARIA DE URGENCIAS Y EMERGENCIAS EXTRAHOSPITALARIAS DEL 112

Clasificación de la demanda sanitaria mediante procesamiento de lenguaje natural (PLN)

Joan Vila Francés



IDAL

Intelligent
Data
Analysis
Laboratory