

Detección Automática de Patologías Combinando Modelos de Lenguaje y Ontologías Médicas

Adam Lei Yi Chen Abolacio

Clasificación de patologías

El objetivo es clasificar informes médicos para detectar la patología del paciente.

Testimonio del paciente

"I've had these light patches on my neck and face for years, but they seem to be spreading lately. I haven't noticed any pain or itching, but I'm starting to worry about sun exposure and potential complications."

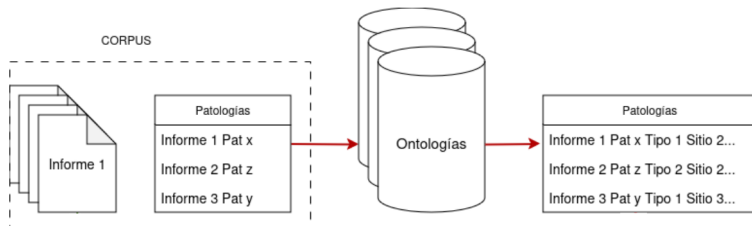
Patología

Vitiligo

Retos

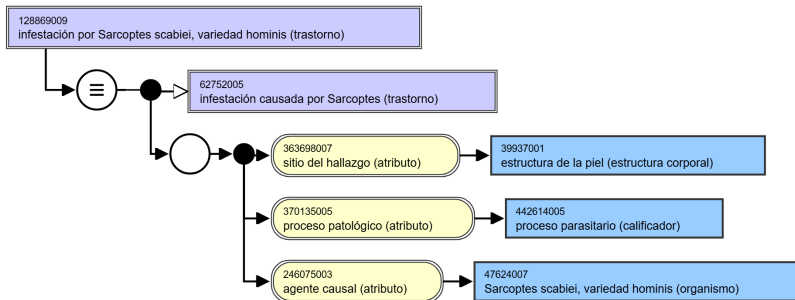
- ▶ Pocos conjuntos de datos disponibles de manera pública.
- ▶ Pocos conjuntos de datos médicos con texto, la mayoría son imágenes y NER.
- ▶ Gran aportación en español: DermatES

Enriquecimiento mediante ontologías

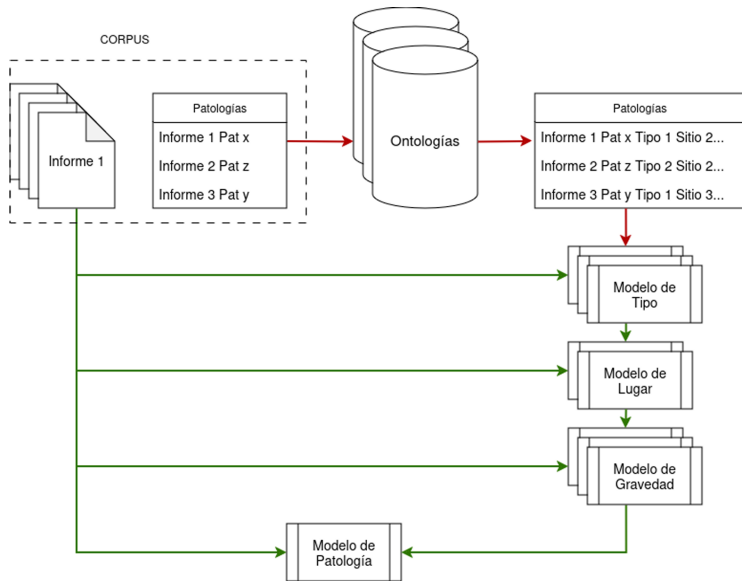


Ontologías utilizadas

- ▶ SNOMED CT
- ▶ ICD-10
- ▶ Tipo, localización y gravedad de la patología.



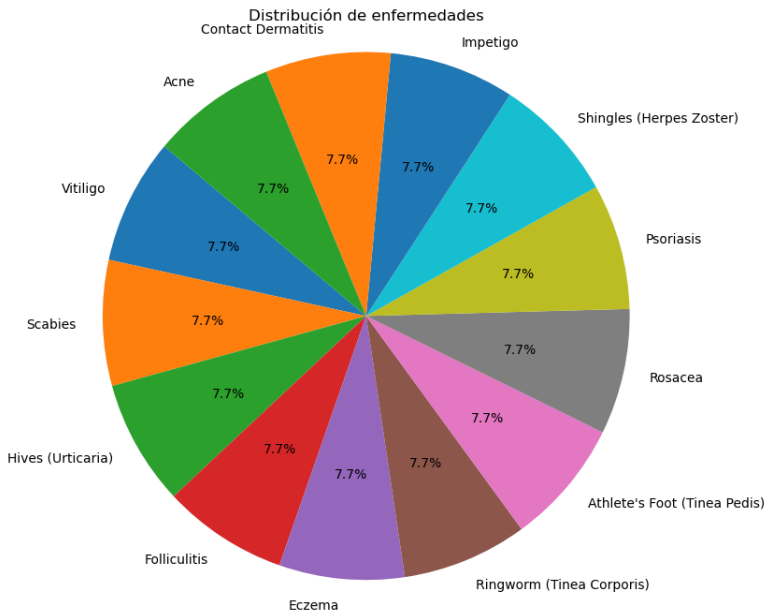
Arquitectura en cascada



Dataset

- ▶ Skin Disease Text Classification
- ▶ 144 casos
- ▶ Enriquecimiento mediante ontologías de forma semi-automática

Dataset



Modelo Transformer

- ▶ Pocos datos, modelo preentrenado.
- ▶ Modelo BERT Small entrenado en corpus inglés
- ▶ 28.76M
- ▶ Entrenamiento de 4 modelos por separado.

Resultados solo texto

Model	Acc	Prec	F1
Logistic Regression	68.2	72.0	66.7
Random Forest	63.6	71.6	62.7
SVC	40.9	42.4	36.2
BERT Small	77.3	83.3	78.5

Resultados cascada

Modo entrenamiento

Model	Acc	Prec	F1
Type	68.2	65.3	64.4
Location	81.8	75.9	77.4
Severity	90.9	92.3	90.8
Disease	95.5	97.7	95.5

Modo oráculo

Model	Acc	Prec	F1
Type	68.2	79.5	72.0
Location	59.1	67.7	61.3
Severity	68.2	68.8	68.3
Disease	54.5	81.8	56.1

Conclusiones

- ▶ Uso de modelos más avanzados.
- ▶ Mejora del rendimiento con ontologías.
- ▶ Arquitectura compleja.
- ▶ Mayor tiempo de computación.
- ▶ Distribuciones desbalanceadas.