

# Uso de conocimiento experto para la mejora de modelo de codificación automática de morfología y topografía

Sergio Peñafiel, Nicolás Bravo, Marcela Aguirre, Felipe Buscaglia, Veronica Sanhueza, Manuel Meneses, Inti Paredes



## RESULTADOS

Luego de implementadas las mejoras discutidas con los patólogos expertos, se realizó nuevamente el análisis con los mismos documentos procesados por la plataforma web y se comparó la precisión para las top N sugerencias, con una mejora de +8% para N=10 y +13% para N=20.

### PROBLEMA

Fundación Arturo López Pérez (FALP) junto con la Universidad de Chile desarrollaron la plataforma web Topomorfo [1], basada en un modelo de Deep Learning para el reconocimiento automático en texto libre en español de menciones de topografía y morfología de un tumor, y el mapeo a sus respectivos códigos CIE-O [2]. Se muestra el diagrama del modelo en la Figura 1. Su rendimiento en las primeras 5 sugerencias es 72% y 65% de precisión en morfología y topografía, respectivamente. Se busca aumentar este rendimiento, basado en las sugerencias de anatomopatólogos de FALP.

### SOLUCIÓN

Luego de 6 meses de utilización de la plataforma web de Topomorfo con usuarios internos FALP, se analizaron los resultados de las predicciones de morfología y topografía de cada uno de los informes procesados y se comparó con la información codificada de biopsias del sistema de información patológico de FALP. Se realizaron sesiones de trabajo con los expertos en las cuales se identificaron los principales problemas del modelo actual y cuáles son las posibles causas metodológicas de estos.

### HIPÓTESIS

El conocimiento experto de los patólogos permite mejorar el rendimiento de Topomorfo al dar cuenta de información adicional del contexto de un informe de biopsia, que el aprendizaje automático no es capaz de inferir.

### OBJETIVOS

Evaluar el impacto de las mejoras al modelo Topomorfo derivadas de las discusiones con el equipo de Anatomía Patológica de FALP.

### MÉTODO

Se decidió implementar las siguientes 5 mejoras que aprovechan el contexto de un término:

1. Filtro de morfología por topografía según libros de la OMS.
2. Reconocer procedimientos para detectar topografía.
3. Actualización de ranking (priors) de los códigos más frecuentes según el historial del sistema de información patológico de FALP.
4. Incluir nuevas formas de escritura.
5. Re-entrenamiento en base a los casos etiquetados validados.

### VARIABLES Y MÉTRICAS

Se compararán los resultados de predicción usando Topomorfo con los informes de biopsias realizadas por patólogos antes y después de aplicar las mejoras. Se analizaron un total de 8400 informes de anatomía patológica que fueron codificados por patólogos. Se mide la precisión promedio de las primeras N sugerencias del modelo, con N entre 1 y 20. También se reporta el recíproco del ranking promedio [3].

### RESULTADOS

Luego de implementadas las mejoras discutidas con los patólogos expertos, se realizó nuevamente el análisis con los mismos documentos procesados por la plataforma web y se comparó la precisión para las top N sugerencias, con una mejora de +8% para N=10 y +13% para N=20 (Figura 2).

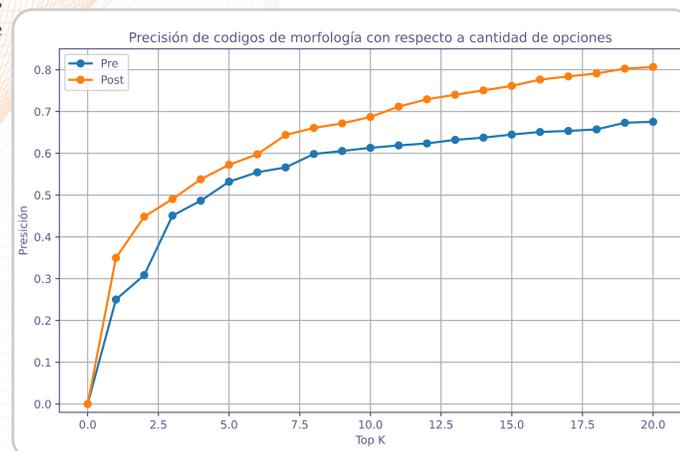


Figura 2. Resultados de precisión del modelo antes y después de mejoras sugeridas por patólogos.

### PERSPECTIVA

Luego de las mejoras, el modelo cuenta con mejor rendimiento y está validado por expertos de la institución. Se espera en los próximos meses analizar la performance con informes de anatomía patológica de otras instituciones.

Esta herramienta puede ser utilizada para la codificación automática de grandes volúmenes de informes de anatomía patológica ayudando a la estructuración de la información para reportar, por ejemplo, al Registro Nacional del Cáncer.

El modelo de predicción se puede consultar en línea en el siguiente link: <https://topomorfo.oncodata.org/>

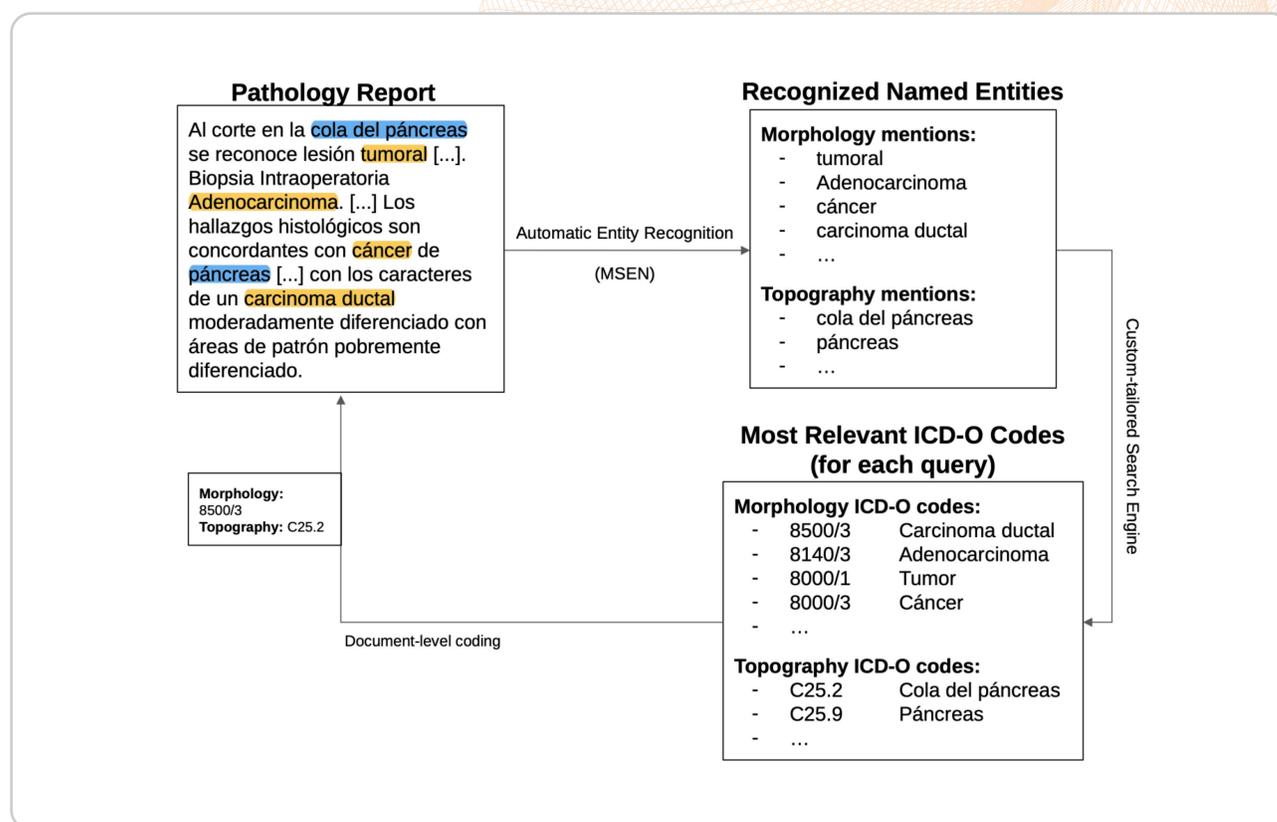


Figura 1. Modelo Deep Learning Topomorfo

### Referencias

1. Villena, F., Báez, P., Peñafiel, S., Rojas, M., Paredes, I., & Dunstan, J. Automatic support system for tumor coding in pathology reports in Spanish.
2. Organización Mundial de la Salud. Classification of tumours. <https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours>
3. E.M. Voorhees (1999). "Proceedings of the 8th Text Retrieval Conference" (PDF). TREC-8 Question Answering Track Report. pp. 77–82.