

Tecnología y automatización en codificación

Utilizar la tecnología para automatizar el trabajo pesado o peligroso no es ninguna novedad, al menos en el campo de los trabajos manuales. Tampoco lo es en ciertas tareas cognitivas repetitivas o tediosas, en las que una máquina entrenada, es capaz de desenvolverse igual o mejor que un humano. Aparece así el concepto de Inteligencia Artificial (IA) y concretamente en el campo de la codificación, el del Procesamiento del Lenguaje Natural (PLN).

Un ejemplo de la aplicación del PLN en el campo de la codificación son las soluciones de codificación de la actividad de áreas de alto volumen y baja complejidad como la urgencia. Este tipo de soluciones, que tan buenos resultados han dado estos últimos años y que han permitido codificar toda la actividad de un área de actividad tan ingente, con prácticamente los mismos recursos.

Esta tecnología puede estar basadas en el reconocimiento de patrones y generación de modelos matemáticos para la producción de los códigos (aproximación estadística) o en la creación de reglas de codificación aplicadas al texto con contenido clínico (aproximación semántica). Nosotros en Solventum optamos por la semántica para el desarrollo de las soluciones de Solventum™ Coding Expert Solution (CES) y Solventum™ 360 Encompass™ Modulo Urgencias, igual que en el pasado optamos por realizar una categorización clínica para desarrollar el algoritmo de los agrupadores poblacionales (CRGs) y es que la experiencia nos ha enseñado que entender bien las reglas del negocio en el que trabajas, siempre aporta una ventaja diferencial sobre el resto de aproximaciones.

Durante este periodo hemos hecho mucho hincapié en reducir al mínimo los falsos positivos, para que se pueda confiar en los resultados de actividad automática y no sea necesario revisarla; la actualización ágil de contenidos, especialmente ante sucesos como la aparición de nuevas patologías como la COVID o finalmente, conceptos de usabilidad, que permitan trabajar en un entorno claro y con acceso a la mayor cantidad de información relevante, con la menor cantidad de interacciones posibles.

Y es que la capacidad de estas soluciones aún estaba lejos del techo que suponía el trabajo de un codificador humano bien formado y actualizado y la intención era buscar la simbiosis de “tecnología al servicio del codificador” para liberar tiempo del codificador en tareas de baja complejidad para dedicarlas a otras de mayor valor. Sin embargo, la aparición el año pasado de un nuevo tipo de inteligencia artificial, la inteligencia basada en Large Language Models (LLM) o IA generativa, ha abierto un interesante debate en el que parece que, no solo será capaz de sustituir a cualquier otra tecnología, sino también a las personas que estaban a cargo de esas tecnologías.

¿Se trata de una realidad o solo de una moda?

Para contestar a la pregunta de cuál es la capacidad real de los LLM en el área de codificación lo mejor es ponerlos a prueba y eso es lo hicimos con un equipo de investigación y desarrollo de nuevas soluciones en Solventum y los resultados cuantitativos, menos del 40% de precisión (porcentaje de códigos producido sobre códigos esperados) y recuperación (porcentaje de códigos producidos sobre el total de códigos esperados), nos han demostrado claramente que aún estamos lejos de que estos sistemas puedan ser una alternativa real.

Sin embargo, quizá lo más relevante de esta pequeña investigación sean los resultados cualitativos obtenidos y es que los LLM están diseñados para “generar” resultados y esa capacidad creativa, tan útil en otros campos, da lugar a la variabilidad en la respuesta a una misma pregunta y complica mucho cada paso del proceso: a la hora de introducir las instrucciones, que han de ser repetidas cada pocas interacciones, ya que el LLM tiende a modificar la estructura de la respuesta en función de sus “preferencias”; a la hora de reproducir los resultados, ya que el mismo caso analizado con las mismas instrucciones da resultados diferentes en función del entrenamiento previo del modelo y finalmente, en cuanto a la validez de los resultados, ya que es posible encontrar con frecuencia que se añaden patologías no existentes en el informe inicial y que generalmente están relacionadas con otras que sí se mencionan en el informe, clásicas “alucinaciones” de este tipo de modelos.

Pero no todo han sido malas noticias para los LLM en el campo de la codificación. Su capacidad de síntesis, utilizando instrucciones que minimicen su creatividad es muy buena y utilizado conjuntamente con modelos de codificación con Inteligencia artificial tradicional han demostrado ser útiles para abordar, con buenos resultados preliminares, tareas que quedaban fuera del alcance hasta ahora de los modelos tradicionales de procesamiento del lenguaje natural como: la codificación de procedimientos en urgencias o la pre codificación básica de episodios de hospitalización complejos a partir de resúmenes que eliminan buena parte de la ambigüedad y el ruido contenido en los informes originales.

En ese sentido, decidimos continuar nuestra investigación para validar esta hipótesis y los resultados, de nuevo, son concluyentes con una mejora de casi el doble tanto en la precisión como en la recuperación de códigos mediante las soluciones híbridas (IA generativa para hacer resúmenes de la información disponible + IA tradicional para codificar los resúmenes). El siguiente paso ahora, es trabajar en segmentar los errores que tienen que ver con la calidad del resumen (LLM) y los que tienen que ver con la codificación (IA tradicional) para definir las líneas de trabajo necesarias para seguir mejorando en cada parte, ya que aún seguimos por debajo del 80%.

Pero, de nuevo, es el análisis cualitativo de los datos el que nos proporciona una información valiosa respecto al uso de IA tradicionales para codificar y es que no es lo mismo usar una aproximación estadística que una semántica.

Pero, ¿por qué no da lo mismo?

Para explicar la diferencia de los resultados, más allá de lo cuantitativo (precisión), entre las tecnologías disponibles actualmente, es el momento de volver al presente y entender la forma en la que funciona cada una de ellas. Por un lado, tenemos la tecnología semántica, que analiza un texto buscando palabras con sentido clínico capaces de disparar el proceso de codificación y por el otro, el estadístico que establece relaciones de correspondencia entre una determinada secuencia de letras con los códigos que ha generado en el set de informes con que ha sido entrenado.

Así las cosas, más allá de que los resultados de precisión puedan ser parecidos, existe una diferencia muy importante derivada de los fundamentos mismos del proceso de codificación manual, que se replica con la tecnología semántica y que consiste en la selección de una palabra relevante, a la que se va asociando luego el resto de información relacionada, para poder así realizar su indexación para obtener el código o códigos más apropiados, siguiendo siempre la misma secuencia marcada en la norma.

Esta manera de hacer las cosas garantiza, con las técnicas semánticas, la relación del código propuesto con la parte de la documentación que lo ha generado, así como la secuencia de reglas del manual aplicadas, de manera que siempre se va a obtener el mismo resultado y va a ser posible explicar por qué se ha generado ese código, las asunciones que hemos tenido que hacer para lograrlo y qué otras opciones ofrece el manual en esos casos e incluso, qué otra información relevante sería interesante disponer para poder llegar a otros códigos aún más precisos.

La tecnología estadística suple esta carencia con la capacidad de replicar este proceso si dispone del suficiente material para su entrenamiento, pero incluso entonces, se ve siempre sujeta a trabajar con unos márgenes de confianza en su respuesta. Es verdad que esta tecnología queda liberada del trabajo que implica el desarrollo manual de reglas de codificación y que es mucho más flexible a la hora de tratar las erratas más frecuentes, que se solventan sin esfuerzo si disponemos de suficientes casos para el entrenamiento, pero a cambio, tiene la desventaja de que, una vez identificado un error, es más difícil saber por qué se ha producido y cómo corregirlo, perdiéndose así la relación entre documentación, norma y código producido, que es la base de la actividad de codificación.

Finalmente, estamos convencidos de que ese diálogo entre máquina-codificador para explicar el código propuesto, así como el resto de posibles alternativas, constituye la base neurofisiológica para que sea posible establecer una relación de confianza en las soluciones de ayuda automáticas y es la única manera de convertirlos en aliados de los verdaderos protagonistas, los codificadores, para que puedan mejorar la eficiencia de sus departamentos, haciendo su trabajo de una manera más provechosa, enriquecedora y agradable.

Si te ha gustado este artículo y aún no estás suscrito, [**suscríbete a nuestra newsletter**](#) para recibir más actualizaciones y consejos sobre cómo optimizar tu proceso de codificación.

Contáctanos para conocer cómo podemos apoyarte en tu proceso de codificación con nuestras soluciones, y si ya eres cliente, podemos dar formación a tu equipo para sacar el máximo provecho de nuestras soluciones.