

# SINAI en eHealth-KD Challenge 2020: Combinando Incrustaciones de palabras para el reconocimiento de entidades nombradas en Registros Médicos Españoles

Pilar López-Úbeda , José M. Perea-Ortegab , Manuel C. Díaz-Galianoa, Teresa Martín-Valdivia y L. , METRO.  
Alfonso Ureña-López

a Grupo de investigación SINAI, Universidad de Jaén, España

bUniversidad de Extremadura, España

## Resumen Este

artículo describe el sistema presentado por el grupo de investigación SINAI al desafío eHealth-KD en IberLEF 2020. Se definieron dos sub tareas principales para el descubrimiento de conocimiento en las historias clínicas españolas: reconocimiento de entidades y extracción de relaciones. En el campo del procesamiento del lenguaje natural (NLP), el reconocimiento de entidades nombradas (NER) se puede formular como un problema de etiquetado de secuencia en el que el texto se trata como una secuencia de palabras que se etiquetarán con etiquetas lingüísticas. Dado que los enfoques actuales de última generación para el etiquetado de secuencias suelen utilizar redes neuronales recurrentes (RNN), nuestra propuesta emplea una red neuronal BiLSTM+CRF donde se combinan diferentes incrustaciones de palabras como entrada a la arquitectura. De este modo, pudimos probar el rendimiento de diferentes tipos de incrustaciones de palabras para la tarea NER en historias clínicas españolas: incrustaciones médicas generadas propiamente, incrustaciones no médicas contextualizadas e incrustaciones no médicas previamente entrenadas basadas en transformadores. Los resultados obtenidos para la tarea de reconocimiento de entidades lograron la puntuación F1 más alta entre todos los participantes, mientras que los obtenidos para la tarea de extracción de relaciones muestran que el enfoque propuesto necesita ser explorado más a fondo.

## Palabras clave

Procesamiento de lenguaje natural, reconocimiento de entidades nombradas, extracción de relaciones, incrustaciones de palabras, profundidad Aprendizaje, eSalud, Descubrimiento del Conocimiento, Lengua Española

## 1. Introducción

El volumen de datos digitalizados disponibles en el ámbito sanitario está aumentando exponencialmente. En este escenario, las técnicas de Procesamiento del Lenguaje Natural (PNL) se vuelven imprescindibles para descubrir nuevos conocimientos en textos clínicos, donde se incluye información valiosa sobre síntomas, diagnóstico, tratamientos o uso de medicamentos. Se han llevado a cabo muchas investigaciones sobre el procesamiento clínico de textos escritos en inglés, pero no tantas para idiomas de escasos recursos como el español. La tarea compartida IberLEF eHealth-KD [1] viene a afrontar este reto desde hace dos años [2, 3], proporcionando diferentes tareas de PNL para extraer automáticamente una variedad de conocimientos a partir de documentos electrónicos de salud escritos en español. En 2020, eHealth-KD propone dos sub tareas

Actas del Foro de Evaluación de Lenguas Ibéricas (IberLEF 2020) correo electrónico:

plubeda@ujaen.es (P. López-Úbeda); jmperea@unex.com (JM Perea-Ortega); modiaz@ujaen.com (MC

Díaz-Galiano); maite@ujaen.com (MT Martín-Valdivia); laurena@ujaen.com (LA Ureña-López) orcid: 0000-0003-0478-743X (P.

López-Úbeda); 0000-0002-7929-3 (JM Perea-Ortega); 0000-0001-9298-1 (MC Díaz-Galiano); 0000-0002-2874-0401 (MT Martín-Valdivia);

0000-0001-7540-4 (LA Ureña-López)



© 2020 Copyright de este artículo por parte de sus autores. Uso permitido bajo la Licencia Creative Commons Attribution 4.0 International (CC BY 4.0).

ISSN 1613-0073 Actas del taller CEUR (CEUR-WS.org)

relacionado con la captura del significado semántico de oraciones relacionadas con la salud: reconocimiento de entidades (subtarea A), cuyo objetivo es identificar todas las entidades en un documento y sus tipos; y extracción de relaciones (subtarea B), que busca identificar todas las relaciones semánticas relevantes entre las entidades reconocidas [1].

En los últimos años, los métodos basados en el aprendizaje profundo se han utilizado ampliamente para algunas tareas relacionadas con la PNL, como el reconocimiento de entidades nombradas (NER) o el etiquetado de parte del discurso (PoS). Estas tareas comúnmente se abordan como un problema de etiquetado de secuencia, donde el texto se trata como una secuencia de palabras que se etiquetan con etiquetas lingüísticas. Los enfoques actuales de última generación para el etiquetado de secuencias proponen el uso de redes neuronales recurrentes (RNN) para aprender representaciones útiles automáticamente, ya que facilitan el modelado de dependencias a larga distancia entre las palabras de una oración. Estas redes generalmente se basan en incrustaciones de palabras, que representan palabras como vectores de números reales. Hay diferentes tipos de incrustaciones de palabras (clásicas [4, 5], a nivel de caracteres [6, 7] o contextualizadas [8, 9]) que comúnmente se entrenan previamente en corpus muy grandes para capturar similitudes sintácticas y semánticas latentes entre palabras. Akbik et al. proponen un nuevo tipo de incrustaciones de palabras llamadas incrustaciones de cadenas contextuales [10], que esencialmente modelan palabras como secuencias de caracteres, contextualizando así una palabra por el texto que la rodea y permitiendo que la misma palabra tenga diferentes incrustaciones dependiendo de su uso contextual.

Este artículo describe el sistema presentado por el equipo del SINAI para ambas sub tareas del eHealth- KD 2020: reconocimiento de entidades y extracción de relaciones. Nuestro grupo tiene experiencia en tareas NER en el dominio biomédico utilizando diferentes metodologías como el aprendizaje automático tradicional [11], RNNs [12] y el aprendizaje automático no supervisado [13, 14], entre otras. Para este desafío, nuestra propuesta se basa en RNN o, más precisamente, la variante bidireccional de memoria a corto plazo junto con una capa de decodificación de campos aleatorios condicionales apilados (BiLSTM+CRF) [15, 6]. Se elige esta arquitectura especializada para abordar NER porque facilita el procesamiento de secuencias de entrada de longitud arbitraria y permite el aprendizaje de dependencias a larga distancia, lo cual es particularmente ventajoso en el caso de la tarea NER enfocada. Además, nuestro enfoque propone la combinación de diferentes tipos de incrustaciones de palabras concatenando cada vector de incrustación para formar los vectores de palabras finales. De esta forma, la probabilidad de reconocer una palabra (entidad) específica en un texto debe incrementarse al combinarse diferentes tipos de representación de esa palabra. Para el caso de las incrustaciones de cadenas contextuales, dado que son robustas frente a palabras mal escritas, suponemos que podrían ser muy adecuadas para NER en textos clínicos.

En la siguiente sección, describimos la arquitectura de nuestro modelo basado en aprendizaje profundo junto con las diferentes incrustaciones de palabras utilizadas como entrada para dicho modelo. Además, explicamos cómo se realizó el aprendizaje utilizando los conjuntos de datos proporcionados por la tarea compartida eHealth-KD.

La Sección 3 informa los resultados obtenidos para los diferentes escenarios de evaluación y, finalmente, concluimos con una breve discusión en la Sección 4 y las conclusiones en la Sección 5.

## 2. Descripción del sistema

El enfoque utilizado para abordar ambas sub tareas del desafío eHealth-KD (NER y extracción de relaciones) se basa en el aprendizaje profundo mediante la implementación del RNN propuesto por Huang et al. [15]. En concreto, hemos utilizado una memoria bidireccional a largo y corto plazo (BiLSTM) con una capa de campo aleatorio condicional (CRF) secuencial. BiLSTM es una extensión del LSTM tradicional

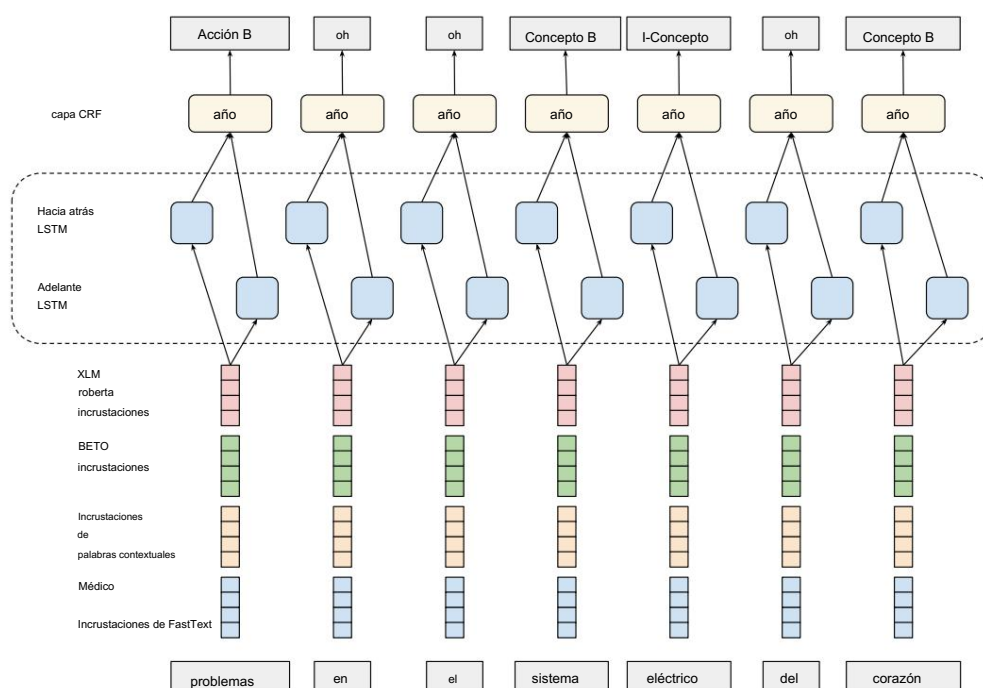


Figura 1: Enfoque propuesto basado en una red neuronal BiLSTM+CRF que utiliza una combinación de diferentes incrustaciones de palabras como capa de entrada.

que mejora el rendimiento del modelo en problemas de clasificación de secuencias [16]. En el uno Por otro lado, el objetivo principal de BiLSTM es dividir las neuronas de un LSTM normal en dos direcciones, una para dirección temporal positiva (estados hacia adelante) y otro para dirección temporal negativa (estados hacia atrás). Utilizando dos direcciones de tiempo, ingrese información del pasado y del futuro del presente. El marco de tiempo se puede utilizar para comprender mejor el contexto del informe médico. En el otro Por otro lado, una capa CRF está representada por líneas que conectan capas de salida consecutivas, de modo que permite utilizar eficientemente etiquetas pasadas y futuras para predecir la etiqueta actual, que es similar al uso de características de entrada pasadas y futuras a través de una red BiLSTM.

Las redes neuronales recurrentes generalmente utilizan una capa de incrustación como entrada, lo que hace que Es posible representar palabras y documentos utilizando una representación vectorial densa. Por lo tanto, la La posición de una palabra dentro del espacio vectorial se aprende del texto, lo que permite que las palabras que se utilizan de manera similar para tener una representación similar. Para encajar la entrada de texto en el sistema neuronal profundo estructura de red, combinamos diferentes tipos de incrustaciones de palabras: biomédica generada propia incrustaciones de palabras para español, incrustaciones de palabras contextuales e incrustaciones de palabras previamente entrenadas Basado en transformadores. La Figura 1 muestra la arquitectura propuesta basada en BiLSTM+CRF.

Red neuronal que utiliza una combinación de diferentes incrustaciones de palabras como capa de entrada.

## 2.1. Incrustaciones de palabras

Se han combinado diferentes incrustaciones de palabras para formar la capa de entrada de la propuesta. Red neuronal BiLSTM+CRF:

### 2.1.1. Incorporaciones médicas de FastText

Aunque hay incrustaciones de palabras biomédicas disponibles para español [17, 18, 19], hemos intentado generar otras nuevas a partir de corpus existentes relacionados con el dominio biomédico en español. Para ello, en primer lugar extrajimos el corpus español de MeSpEN [20]. Luego, se agregó al corpus anterior información adicional en español de diferentes sitios web de información clínica como Mayo Clinic [21], Organización Mundial de la Salud [22] y WebMD [23]. El preprocesamiento realizado para entrenar las incrustaciones de palabras consistió en convertir el texto a minúsculas, eliminar las URL y eliminar las líneas múltiples. Finalmente, se utilizó FastText [24] para realizar el entrenamiento aplicando la siguiente configuración: modelo skip-gram, 0.05 para la tasa de aprendizaje, tamaño de 300 para los vectores de palabras, 10 para el número de épocas y 5 para el número mínimo de ocurrencias de palabras.

### 2.1.2. Incrustaciones de palabras contextuales

Las incrustaciones de palabras contextuales se consideran incrustaciones poderosas porque capturan información sintáctico-semántica latente que va más allá de las incrustaciones de palabras estándar [10]. Estas incrustaciones se basan en el modelado del lenguaje a nivel de caracteres y su uso es particularmente ventajoso cuando la tarea NER se aborda como un problema de etiquetado secuencial. Para nuestros experimentos, hemos utilizado las incrustaciones contextualizadas agrupadas propuestas por Akbik et al. [25] para NER. En ese trabajo, los autores proponen un enfoque en el que se aplica una operación de agrupación para destilar una representación global de palabras de todas las instancias contextualizadas de esa palabra, produciendo así representaciones de palabras en evolución que cambian con el tiempo a medida que se observan más instancias de la misma palabra en la palabra . datos.

### 2.1.3. Incrustaciones de palabras previamente entrenadas basadas en Transformers

Finalmente, hemos combinado incrustaciones de palabras previamente entrenadas basadas en transformadores [26], un mecanismo de atención popular utilizado en enfoques de aprendizaje por transferencia. El aprendizaje por transferencia [27] se basa en la aplicación del conocimiento aprendido en tareas anteriores a un nuevo dominio de tarea. Este tipo de incrustaciones comúnmente se entrenan previamente en corpus muy grandes para capturar similitudes sintácticas y semánticas latentes entre palabras. Para nuestros experimentos, hemos utilizado dos incrustaciones de palabras específicas previamente entrenadas: BETO [28], que sigue un modelo BERT entrenado en un gran corpus de español, y XLM-RoBERTa [29], que se generaron utilizando un gran modelo de lenguaje multilingüe. entrenado en 2,5 TB de datos filtrados de CommonCrawl.

## 2.2. Fase de entrenamiento

Los organizadores de la tarea proporcionaron varios conjuntos de datos (entrenamiento, validación y prueba) para permitir a los participantes entrenar sus sistemas adecuadamente. Además, un conjunto de datos no revisado llamado "conjunto" que contenía los envíos de ediciones anteriores, y un conjunto de datos llamado "transferencia" del español

Los	oh	Los	oh
tumores	Concepto B	doctores	oh
malignos de	Concepto B	diagnostican B-implica	
	oh	talasemias O	
glándulas B-Concept		mediante	oh
suprarrenales I-Concept		pruebas de	yo-implica
hijo	oh		yo-implica
poco	oh	sangre	yo-implica
comunes	Concepto B	.	oh
.	oh		

Figura 2: Ejemplo de esquema de etiquetado BIO para reconocimiento de entidades y relaciones.

También se proporcionó una versión de Wikinews. Para los escenarios 1, 2 y 3, utilizamos la capacitación, ensamblar y transferir (800 + 3000 + 100 oraciones respectivamente) conjuntos de datos para entrenamiento, mientras El conjunto de desarrollo (200 frases) se utilizó para validar nuestro sistema. Sin embargo, para el escenario 4, los conjuntos de datos de entrenamiento, ensamblaje y desarrollo (800 + 3000 + 200 oraciones, respectivamente) se utilizaron para el entrenamiento, mientras que el conjunto de transferencia (100 frases) se utilizó para validar el sistema.

Para todos los escenarios, primero se tokenizó cada oración. Para la tarea de reconocimiento de entidades, el Las anotaciones proporcionadas se codificaron utilizando el esquema de etiquetado BIO. Así, cada ficha en un La oración estaba etiquetada con O (no entidad), B (símbolo inicial de una entidad) o I (símbolo interior de una entidad). Este esquema es el más popular en la tarea NER aunque presenta problemas cuando la entidad contiene tokens discontinuos. La Figura 2 muestra un ejemplo de la anotación codificada. realizado para ambas tareas.

Para la tarea de extracción de relaciones, decidimos generar la misma red neuronal para entrenar cada una de las 13 relaciones semánticas definidas para la tarea.

Por lo tanto, las únicas características lingüísticas de entrada utilizadas para la fase de entrenamiento fueron los tokens. (palabras originales) de cada oración y las etiquetas BIO generadas.

### 3. Configuración experimental y resultados

Para generar los diferentes modelos de aprendizaje utilizamos Flair como marco de PNL [30]. Nosotros usó la configuración de hiperparámetro predeterminada en Flair con la siguiente configuración: tasa de aprendizaje como 0,1, tamaño de lote como 32, probabilidad de abandono como 0,5 y época máxima como 150. Todos los experimentos se realizaron en una única GPU Tesla-V100 de 32 GB con 192 GB de RAM.

El equipo del SINAI presentó 3 ejecuciones para cada escenario propuesto, donde cada ejecución representa un combinación diferente de incrustaciones de palabras, como se describe en la Sección 2:

- Ejecución 1: incrustaciones de Medical FastText + incrustaciones de palabras contextuales.
- Ejecución 2: incrustaciones de Medical FastText + incrustaciones de palabras contextuales + incrustaciones de BETO.
- Ejecución 3: incrustaciones de Medical FastText + incrustaciones de palabras contextuales + incrustaciones de BETO + Incrustaciones XLMRoBERTa.

tabla 1

Resultados de la evaluación obtenidos por el equipo del SINAI en el eHealth-KD de IberLEF2020.

Guión	Correr	Precisión	Recordar	Puntuación F1
1	1	0.6515	0.3106	0.4207
	2	0.6451	0.3005	0,41
	3	0.6446	0.3014	0.4107
2	1	0.8303	0.8094	0.8197
	2	0.8432	0.7932	0.8174
	3	0.8446	0.8067	0.8252
3	1	0,6145	0.325	0,4252
	2	0.6271	0.3654	0.4617
	3	0.6124	0.3615	0.4547
4	1	0.6366	0.1685	0.2664
	2	0.6263	0.1813	0.2812
	3	0.611	0,1795	0.2775

Las métricas definidas por el desafío eHealth-KD para evaluar los experimentos enviados son aquellos comúnmente utilizados para algunas tareas de PNL como NER o clasificación de texto, es decir, precisión, recuerdo y puntuación F1. La Tabla 1 muestra los resultados obtenidos por el equipo del SINAI para cada escenario y ejecutar enviado.

Como se muestra en la Tabla 1, los resultados obtenidos para el escenario 4 son bajos (28,12% de la puntuación F1), aunque Este comportamiento puede deberse a la cantidad limitada de documentos disponibles para entrenar una alternativa. dominio. Para el escenario 3 (subtarea B), logramos nuestra mejor puntuación de F1 con la ejecución 2 (46,17%), que significa que las incorporaciones XLMRoBERTa no proporcionaron información valiosa durante la fase de aprendizaje. En cuanto al escenario 1 (evaluación principal), logramos una puntuación de F1 relativamente pobre. de 42.07%, lo que de alguna manera está en línea con el comportamiento obtenido para el escenario 3, donde el uso de Las incorporaciones previamente entrenadas basadas en transformadores no han mejorado el rendimiento. Finalmente, Cabe señalar que la puntuación de F1 obtenida en nuestra ejecución 3 presentada para el escenario 2 (82,52 %) superó al resto de carreras de la competencia para ese escenario entre todos los participantes.

## 4. Discusión

Según los resultados obtenidos, encontramos que la combinación de incrustaciones de palabras generalmente Proporciona un valor extra para la fase de aprendizaje de la red neuronal. En el caso de la entidad reconocimiento (subtarea A, escenario 2), la combinación de todas las incrustaciones (ejecución 3) mejora todas las métricas obtenidas cuando se combinan menos tipos de incrustaciones. Cabe señalar que nuestro El punto de partida (ejecución 1) combina incorporaciones que se generaron a partir del dominio biomédico. y por el cual ya obtuvimos una alta puntuación en F1. Por lo tanto se puede considerar positivo han mejorado ligeramente esos resultados cuando se agregaron otras incorporaciones no biomédicas.

Se puede definir una evaluación detallada de estos sistemas en términos de comparar la respuesta

Tabla 2

Resultados de evaluación obtenidos por el equipo del SINAI en términos de comparar la respuesta del sistema contra la anotación dorada.

		Tarea A					Tarea B				
		Ejecución del escenario		correcto	incorrecto	parcial	faltante	espurio	correcto	faltante	espurio
1	1	250	15	21	288	46	75	431	108		
	2	250	18	19	287	44	---	441	107		
	3	250	18	19	287	44	66	440	108		
2	1	436	30	28	62	48	x	x	x		
	2	424	27	34	71	38	x	x	x		
	3	432	24	33	67	42	x	x	x		
3	1	x	x	x	x	x	169	351	106		
	2	x	x	x	x	x	109	330	113		
	3	x	x	x	x	x	188	332	119		
4	1	348	33	21	808	48	47	1150	140		
	2	350	20	25	815	46	74	1123	182		
	3	345	27	26	812	50	74	1123	185		

del sistema contra la anotación dorada [31]. Entonces, la evaluación de nuestro sistema considerando

Estas diferentes categorías de errores se muestran en la Tabla 2.

Como se muestra en la Tabla 2, nuestro sistema funciona bien en la primera tarea (reconocimiento de entidades) y escenario 2, pero tiene algunas deficiencias de rendimiento para la tarea de extracción de relaciones. Para En el trabajo futuro, deberíamos centrarnos en descubrir por qué el sistema no captura correctamente el relaciones entre conceptos. Por último, también nos gustaría destacar el comportamiento de nuestros sistema en el escenario alternativo (escenario 4), en el que no podría ser capaz de reconocer entidades que no estaban relacionados con el ámbito médico. Creemos que podría deberse al uso de la palabra.

incorporaciones específicamente entrenadas previamente en ese dominio.

## 5. Conclusiones

Este artículo presenta la participación del grupo de investigación SINAI en el desafío eHealth-KD en IberLEF 2020 (Foro de Evaluación de Lenguas Ibéricas), donde dos tareas principales de PNL involucran Se definieron el idioma español: reconocimiento de entidades y extracción de relaciones. Varios

También se propusieron escenarios de desafío donde se evaluaron diferentes sub tareas de PNL.

Nuestra propuesta sigue un enfoque basado en el aprendizaje profundo para el reconocimiento de entidades nombradas (NER) en documentos sanitarios españoles. Está enfocado en el uso de una red neuronal BiLSTM+CRF donde Se combinan diferentes incrustaciones de palabras como entrada a la arquitectura. Entonces este nervio La red se entrena utilizando los conjuntos de datos anotados proporcionados por la organización, que fueron previamente tokenizado y etiquetado mediante el uso del esquema BIO. Nuestro principal objetivo era demostrar la realización de diferentes tipos de incrustaciones de palabras para la tarea NER en el ámbito médico:

incorporaciones médicas generadas propiamente, incorporaciones no médicas contextualizadas e incorporaciones no médicas previamente entrenadas basadas en transformadores. Los resultados obtenidos para la tarea de reconocimiento de entidades lograron el puntaje de evaluación más alto entre todos los participantes, logrando 82,52% de puntaje F1, 84,46% de precisión y 80,67% de recuerdo. Sin embargo, nuestra propuesta reveló ciertas debilidades para la tarea de extracción de relaciones, cualquiera de las relaciones existentes en la prueba de oro no ha sido capturada por nuestro sistema, lo que significa una pérdida en la recuperación, necesitamos analizar en detalle si el problema radica en la notación BIO de la relación porque al usar esta notación, cuando el sistema encuentra un comienzo de relación, se asocia con el final de la relación más cercana anotada, por lo que será necesario explorar más a fondo el enfoque utilizado.

Para trabajos futuros, estudiaremos el rendimiento del uso de más características lingüísticas como etiquetas Part-Of-Speech como entrada en la red neuronal, así como el uso de ontologías relacionadas con el dominio biomédico y otros tipos de incrustaciones de palabras. Además, intentaremos mejorar la extracción de relaciones implementando otra red neuronal que capture de mejor manera las relaciones entre conceptos.

## Expresiones de gratitud

Este trabajo ha sido parcialmente apoyado por el proyecto LIVING-LANG (RTI2018-094653-B-C21) del Gobierno de España, la Junta de Extremadura (GR18135) y el Fondo Europeo de Desarrollo Regional (FEDER).

## Referencias

- [1] Piad-Morffis, Y. Gutierrez, H. Cañizares-Diaz, S. Estevez-Velarde, Y. Almeida-Cruz, R. Muñoz y A. Montoyo, Overview of the eHealth Knowledge Discovery Challenge en IberLEF 2020, en : Actas del Foro Ibérico de Evaluación de Lenguas compartido con la 36ª Conferencia de la Sociedad Española para el Procesamiento del Lenguaje Natural, IberLEF@SEPLN 2020, España, septiembre de 2020.
- [2] Análisis de sistemas de descubrimiento de conocimiento en eSalud en el Taller TASS 2018, Language Processing Natural (2019). doi:10.26342/2019-62-1.
- [3] Piad-Morffis A, Gutiérrez Y, Consuegra-Ayala JP, Estevez-Velarde S, Almeida-Cruz Y, Muñoz R, Montoyo A, Descripción general del desafío de descubrimiento de conocimientos sobre eSalud en IberLEF 2019, en: Actas del taller CEUR,
- [4] J. Pennington, R. Socher, CD Manning, GloVe: Vectores globales para representación de palabras, en: EMNLP 2014 - Conferencia 2014 sobre métodos empíricos en el procesamiento del lenguaje natural, Actas de la conferencia, 2014. doi:10.3115/v1/d14-1162.
- [5] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean, Representaciones distribuidas de palabras y frases y su composicionalidad, en: Advances in Neural Information Processing Systems, 2013. arXiv:1310.4546.
- [6] X. Ma, E. Hovy, Etiquetado de secuencias de un extremo a otro mediante LSTM-CNNs-CRF bidireccional, en: 54ª Reunión Anual de la Asociación de Lingüística Computacional, ACL 2016 - Long Papers, 2016. doi: 10.18653/v1/p16-1101. arXiv:1603.01354.



- [7] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, C. Dyer, Arquitecturas neuronales para el reconocimiento de entidades nombradas, en: Conferencia de 2016 del Capítulo Norteamericano de la Asociación de Lingüística Computacional: Human Language Technologies, NAACL HLT 2016 - Actas de la conferencia, 2016. doi:10.18653/v1/n16-1030. arXiv:1603.01360.
- [8] ME Peters, W. Ammar, C. Bhagavatula, R. Power, Etiquetado de secuencia semisupervisado con modelos de lenguaje bidireccional, en: ACL 2017 - 55a Reunión Anual de la Asociación de Lingüística Computacional, Actas de la Conferencia (artículos extensos ), 2017. doi:10.18653/v1/P17-1161. arXiv:1705.00108.
- [9] ME Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, L. Zettlemoyer, Representaciones de palabras contextualizadas profundas, en: NAACL HLT 2018 - 2018 Conference of the North American Chapter of the Asociación de Lingüística Computacional: Tecnologías del lenguaje humano - Actas de la conferencia, 2018. doi:10.18653/v1/n18-1202. arXiv:1802.05365.
- [10] A. Akbik, D. Blythe, R. Vollgraf, Incrustaciones de cadenas contextuales para etiquetado de secuencias, en: Actas de la 27ª Conferencia Internacional sobre Lingüística Computacional, 2018, págs. 1638-1649.
- [11] PL Úbeda, MCD Galiano, MT Martín-Valdivia, LAU López, Usando métodos de aprendizaje automático y aprendizaje profundo para encontrar menciones de reacciones adversas a medicamentos en redes sociales, en: Actas de la Cuarta Minería de Redes Sociales para Aplicaciones en Salud (#SMM4H ) Taller y tarea compartida, 2019, págs. 102-106.
- [12] PL Úbeda, MCD Galiano, LAU López, MT Martín-Valdivia, Uso de Snomed para reconocer e indexar menciones de sustancias químicas y medicamentos, en: Actas del 5º Taller sobre Tareas Compartidas Abiertas de BioNLP, 2019, págs.
- [13] P. López-Úbeda, MC Díaz-Galiano, MT Martín-Valdivia, LA Urena-Lopez, Sinai en tass 2018 tarea 3. clasificación de acciones y conceptos con umls en medline, Actas de TASS 2172 (2018).
- [14] P. López-Úbeda, M. C. Díaz-Galiano, A. Montejo-Ráez, M.-T. Martín-Valdivia, L. A. Ureña-López, An Integrated Approach to Biomedical Term Identification Systems, Applied Sciences 10 (2020) 1726.
- [15] Z. Huang, W. Xu, K. Yu, Modelos LSTM-CRF bidireccionales para etiquetado de secuencias, arXiv preimpresión arXiv:1508.01991 (2015).
- [16] M. Schuster, KK Paliwal, Redes neuronales recurrentes bidireccionales, IEEE Transactions on Signal Processing 45 (1997) 2673-2681.
- [17] F. Soares, M. Villegas, A. González-Agirre, M. Krallinger, J. Armengol-Estapé, Incorporaciones de palabras médicas para español: desarrollo y evaluación, en: Actas del 2º Taller clínico de procesamiento del lenguaje natural, Asociación para Lingüística Computacional , Minneapolis, Minnesota, EE. UU., 2019, págs. 124-133. URL: <https://www.aclweb.org/anthology/W19-1916>. doi:10.18653/v1/W19-1916.
- [18] S. Santiso, A. Casillas, A. Pérez, M. Oronoz, Incrustaciones de palabras para la detección de negaciones en historias clínicas escritas en español, Soft Computing (2019). doi:10.1007/s00500-018-3650-7.
- [19] I. Segura-Bedmar, P. Martínez, Simplificación de los prospectos de medicamentos escritos en español mediante la incrustación de palabras, Journal of Biomedical Semantics (2017). doi:10.1186/

s13326-017-0156-7.

- [20] M. Villegas, A. Intxaurre, A. González-Agirre, M. Marimon, M. Krallinger, El recurso MeSpEN para traducción automática y terminologías médicas inglés-español: censo de corpus paralelos, glosarios y traducciones de términos, LREC MultilingualBio : Procesamiento de textos biomédicos multilingües (Malero M, Krallinger M, González-Agirre A, eds.) (2018).
- [21] Clínica Mayo, 1998-2020. URL: <https://www.mayoclinic.org/es-es>.
- [22] Organización mundial de la salud, 2020. URL: <https://www.who.int/es>.
- [23] Webmd - mejor información. mejor salud., 2005-2020. URL: <https://www.webmd.com/>.
- [24] P. Bojanowski, E. Grave, A. Joulin, T. Mikolov, Enriquecimiento de vectores de palabras con información de subpalabras, Transacciones de la Asociación de Lingüística Computacional 5 (2017) 135–146.
- [25] A. Akbik, T. Bergmann, R. Vollgraf, Incorporaciones contextualizadas agrupadas para el reconocimiento de entidades nombradas , en: Actas de la Conferencia de 2019 del Capítulo Norteamericano de la Asociación de Lingüística Computacional: Tecnologías del Lenguaje Humano, Volumen 1 (Largo y artículos breves), 2019, págs. 724–728.
- [26] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, AN Gomez, Ł. Kaiser, I. Polosukhin , La atención es todo lo que necesita, en: Avances en los sistemas de procesamiento de información neuronal, 2017, págs. 5998–6008.
- [27] S. Thrun, ¿Es más fácil aprender la enésima cosa que aprender la primera?, en: Avances en Sistemas de procesamiento de información neuronal, 1996, págs. 640–646.
- [28] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Entrenamiento previo de transformadores bidireccionales profundos para la comprensión del lenguaje, preimpresión de arXiv arXiv:1810.04805 (2018).
- [29] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzman, E. Grave, M. Ott, L. Zettlemoyer, V. Stoyanov, Aprendizaje de representación multilingüe no supervisado en escala, preimpresión de arXiv arXiv:1911.02116 (2019).
- [30] A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter, R. Vollgraf, FLAIR: Un marco fácil de usar para la PNL de última generación, en: Actas de Conferencia de 2019 del Capítulo Norteamericano de la Asociación de Lingüística Computacional (Demostraciones), 2019, págs. 54–59.
- [31] N. Chinchor, B. Sundheim, Métricas de evaluación MUC-5, en: Actas de la quinta conferencia sobre comprensión de mensajes, Asociación de Lingüística Computacional, 1993, págs.