



Deep Learning. Aplicaciones en Reconocimiento de Lengua de Señas, Generación de Lenguaje Natural e Imágenes Astronómicas.

Franco Ronchetti, Facundo Quiroga, Gastón Ríos, Pedro Dal Bianco, Ivan Mindlin, Laura Lanzarini, Alejandro Rosete, Roberto Gamén, Domènec Puig Valls, Jordina Torrents-Barrena, Yael Aidelman, Carlos Escudero, Anahí Granada.

fronchetti@lidi.info.unlp.edu.ar

Contexto

El Instituto de Investigación en Informática LIDI (III-LIDI) tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos métodos de Deep Learning. Esta presentación corresponde a las tareas de investigación que se llevan a cabo en el III-LIDI en el marco del proyecto F025 "Sistemas inteligentes. Aplicaciones en reconocimiento de patrones, minería de datos y big data" perteneciente al Programa de Incentivos (2018-2021).

Líneas de Investigación y Desarrollo

Reconocimiento de Lengua de Señas con Redes CNN/RNN



Fig. 1 Activaciones de una CNN+RNN para reconocer señas en video.

Las lenguas de señas utilizan un conjunto finito de formas de mano que, en combinación con movimientos de las manos y el cuerpo, se utilizan para señar.

En particular, estamos trabajando con señas de la base de datos LSA64. Esta consiste en un registro de 64 señas de la Lengua de Señas Argentina. Para la tarea se compararon diversas técnicas en el estado del arte del Aprendizaje Automático basadas en Redes Neuronales. Específicamente, se compararon tanto arquitecturas basadas en Redes Recurrentes y Convolucionales, como distintas estrategias de preprocesamiento para optimizar la calidad del reconocimiento. También se están analizando los modelos entrenados para comprender mejor el impacto de estas estrategias de preprocesamiento y de la forma de representación lograda por cada tipo de arquitectura.

Generación de letras de rap con Redes Neuronales Recurrentes

Las redes neuronales tienen un gran potencial también como modelos generativos en el área del lenguaje natural.

Se desarrolló un prototipo de generación automática de letras de rap con estilo freestyle. Se confeccionaron dos bases de datos específicas de este tipo de letras en español, inexistentes hasta la fecha. Se analizaron diferentes sistemas de preprocesamiento del texto incluyendo tokenización a nivel de palabra, de sílaba y el algoritmo BPE. Se realizaron diversos experimentos con redes recurrentes bidireccionales tipo LSTM y se realizó un diccionario específico de rimas en español para controlar cómo el texto las genera.

*De la esquina cómo verás más que ignoraban
 ¿Y dónde está todo eso con lo que señabas
 Hago la diferencia con mi poesía pero nada
 Yo, yeah, entendí que todo importaba*

*La diferencia con el es que me ganó el primer puesto
 La fe mueve montañas mis montañas
 Mi personalidad que me acompañaba
 Mírame a mí ya me voy genio*

Fig. 2 Ejemplo de estrofas generados por el modelo de rap con estilo freestyle

Análisis de Equivarianza en Redes Neuronales

Las redes neuronales son modelos complejos cuya interpretabilidad es difícil. Una forma de añadir interpretabilidad a un modelo es estudiando su equivarianza ante diversas transformaciones. En base a diversas métricas de invarianza y auto-equivarianza (casos especiales de la equivarianza) previamente definidas, caracterizamos diversos modelos de redes neuronales en términos de estas propiedades.

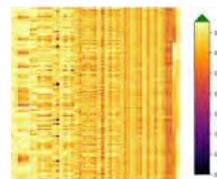


Fig. 3 Invarianza a las rotaciones de las activaciones de una red ResNet entrenada con CIFAR10.

Análisis de Imágenes Astronómicas

En los últimos años, la cantidad de información astronómica disponible se ha multiplicado de forma exponencial. En consecuencia, diversas tareas que anteriormente se realizaban de forma manual o semi-manual deben ahora automatizarse aún más.

En este contexto, se está desarrollando un proyecto con el objetivo de obtener un sistema capaz de encontrar estrellas Be (no-supergigantes cuyo espectro exhibe emisión en línea H α) a partir de información fotométrica de . Se están implementando, a partir de catálogos existentes, modelos clásicos de Aprendizaje Automático que permitan una detección de estrellas Be candidatas, para luego ser evaluadas por un experto.

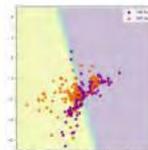


Fig. 4 Red. de dimensionalidad para validar visualmente el comportamiento de un clasificador de estrellas.

Formación de Recursos Humanos

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 1 profesor con dedicación exclusiva, 1 investigador CIC-PBA, 2 becarios de posgrado de la UNLP con dedicación docente, 1 becario CIC, 1 becario CIN, 6 tesistas, 3 profesores extranjeros, y 3 investigadores externos.

Dentro de los temas involucrados en esta línea de investigación, en los últimos dos años se han finalizado 2 tesis de doctorado, 2 tesis de especialización, y 5 tesinas de grado de Licenciatura.

Actualmente se están desarrollando 2 tesis de doctorado, 1 tesis de especialista, 5 tesinas de grado de Licenciatura y 4 trabajos finales de Ingeniería en Computación. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

