

NOMBRE: Benjamín Farías Valdés

N.ALUMNO: 22102671



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
DEPARTAMENTO DE CIENCIA DE LA COMPUTACIÓN

IIC3692 — Tópicos Avanzados en Inteligencia Artificial — 2' 2022

Lectura 9

Crítica

End-To-End Memory Networks

En este artículo se presenta una arquitectura de *RNN* basada en un modelo de atención que hace uso de una memoria externa, denominada *MemN2N*. En la introducción se indica que la propuesta comparte aspectos similares a otras redes de memoria ya existentes, pero innova en el hecho de que la representación interna de la información es continua y por tanto más flexible. En la sección de trabajo relacionado se mencionan otros modelos del estado del arte que inspiran las características del propuesto (*LSTM*, *Memory Networks*, *Bidirectional RNNs*), se destaca el objetivo de lograr almacenar información secuencial de largo plazo (que suele ser problemático para las *RNNs* clásicas). En cuanto al *approach* del paper, se opta por explicar primero una versión base del modelo *MemN2N* con 1 sola capa, lo que encuentro una buena decisión para aclarar los conceptos centrales de forma clara y simplificada. La idea central se basa en entrenar 3 capas de *embedding*, una para la *query* (utilizada en aplicaciones basadas en realizar consultas sobre un texto, estilo *Question Answering*), otra para el *output* y la última para realizar atención sobre el *input* (basada tanto en la información de entrada como en la *query*). Esta idea se extiende posteriormente a varias capas, combinándolas de forma secuencial para formar la arquitectura recurrente. Me parece muy interesante el hecho de que cada capa de la red actúa sobre los mismos datos de entrada (que corresponden a todo el texto que está guardado en la memoria), a diferencia de las *RNNs*, que acceden a la información de forma secuencial y local. Dándole unas vueltas, creo que lo que se logra de cierta forma es darle al modelo la capacidad de leer el texto para encontrar relaciones entre sus palabras, y posteriormente asumir esas relaciones y buscar otras relaciones nuevas basadas en ese contexto ya aprendido. Esto se asemeja mucho a lo que un ser humano hace al leer un texto y tratar de entenderlo, acumulando conocimiento en cada pasada realizada (recurrencia) y siempre pudiendo leer todo desde el inicio sin perder contexto ni olvidar información (memoria). Para la parte de experimentos se muestra como la propuesta es competitiva con redes recurrentes clásicas, y tiene un rendimiento cercano a otros modelos de *Memory Networks* que tienen altos niveles de supervisión (sacrificando flexibilidad por rendimiento). En general me gustó bastante el artículo, si bien tuve que leerlo un par de veces para entender los detalles de la propuesta, realmente es un cambio de paradigma bien ingenioso para modelar más de cerca la capacidad del ser humano al momento de leer y comprender información textual. A pesar de esto, creo que una limitación importante de abarcar en futuros trabajos será la capacidad máxima de la memoria externa, puesto que para grandes cantidades de texto eventualmente podría no ser realista el supuesto de que se puede guardar todo en memoria.