

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 2

Crítica

Attention Is All You Need

Este es un paper bien conocido en el área de investigación sobre procesamiento del lenguaje natural. Se presenta un modelo innovador, denominado *Transformer*, que logra resultados superiores al estado del arte en la tarea de traducción de texto entre idiomas. El inicio del artículo está bien construido en mi opinión, se enfoca en explicar el estado del arte, que consiste principalmente en *RNNs* con la arquitectura de *Encoder-Decoder*, y explica las limitaciones de estos modelos (poca paralelización, malos para oraciones largas). Luego, presentan la arquitectura propuesta, que sólo utiliza capas de atención y *fully-connected*. En esta sección del paper encuentro que es muy importante tener conocimiento previo en el área (o revisar los trabajos que propusieron el concepto de atención), ya que se entra rápidamente en detalle y es fácil perder el hilo al no saber de antemano la utilidad y funcionamiento del mecanismo de atención en redes neuronales. Quizás se podría haber explicado un poco sobre las principales cualidades de estos mecanismos antes de entrar de lleno en su aplicación dentro de los *Transformers*. La última parte del artículo se encarga de mostrar los resultados del modelo propuesto, lo que encuentro que es bastante claro y bien presentado (dando a entender que el modelo es mucho más liviano que los del estado del arte, y además logra un mejor desempeño). En general, es claro el valor que presentó esta publicación para el año 2017, y me pareció muy interesante lo ilustrado, además de que el paper en sí era bastante corto (lo que es buena idea cuando no se quiere desgastar al lector con una gran cantidad de texto o explicaciones innecesarias). Como mencioné anteriormente, lo único que cambiaría está relacionado a agregar más información que resuma conceptos clave previos a esta investigación, como lo son el mecanismo de atención y de *encoding-decoding*, de forma que sea más accesible para lectores que quizás estén recién comenzando a meterse en el área (tampoco es necesario entrar en detalle con estas explicaciones, pero al menos dar una idea de cómo funcionan).