Pontificia Universidad Católica de Chile

Facultad de Matemáticas

Departamento de Matemáticas

# Actividad 5

Programa de Habilidades Comunicativas Escritas para Ciencias Naturales y Matemáticas - LET172E

Fecha de Entrega: 2020-01-17

## Nicholas Mc-Donnell

El “Machine Learning” se ha vuelto una vital herramienta para múltiples tipos de problemas, desde clasificaci´on de im´agenes (Sultana, Sufian, y Dutta, 2019) hasta ayudar a traducir documentos. En este último, las Redes Neuronales (RNs)[[1]](#footnote-1) han sido un cambio impresionante al área, cada persona que ha usado Google TranslateTM ha notado la impresionante mejora en los últimos años, en los cuales el servicio transicionó de usar “Statistical Machine Learning” a RNs (Turovsky, 2016).

El gran cambio que se ha visto viene de la reciente investigación en distintas arquitecturas de RNs, por ejemplo Rashid, Do-Omri, Haidar, Liu, y Rezagholizadeh (2019) propusieron un modelo el cual utiliza avances en otro problema[[2]](#footnote-2) para además avanzar en la calidad de traducción, otros como Vaswani et al. (2017) proponen un modelo simple usando el “Transformador” como base. Esto claramente muestra una variedad importante en las arquitecturas de RNs que intentan solucionar el problema de traducción.

Por lo cual, este trabajo tiene el objetivo de comparar distintas arquitecturas de Redes Neuronales para el problema de mapeo multilingüe[[3]](#footnote-3). Para lograr esto, se realizó una revisión de literatura de varios art´ıculos sobre el tema, de los cuales tres tienen como enfoque presentar una nueva arquitectura de RNs y el resto da mejoras y apoyo a los primeros. Para un facilitar la comparación se introducirán las distintas arquitecturas separadamente, primero se ver´a la arquitectura de Vaswani et al. (2017), despúes la de Rashid et al. (2019) y por último la de Artetxe et al. (2018), despúes de introducirlas se har´a una comparación de los resultados de cada arquitectura.

### El Transformador, arquitectura de Vaswani et al. (2017)

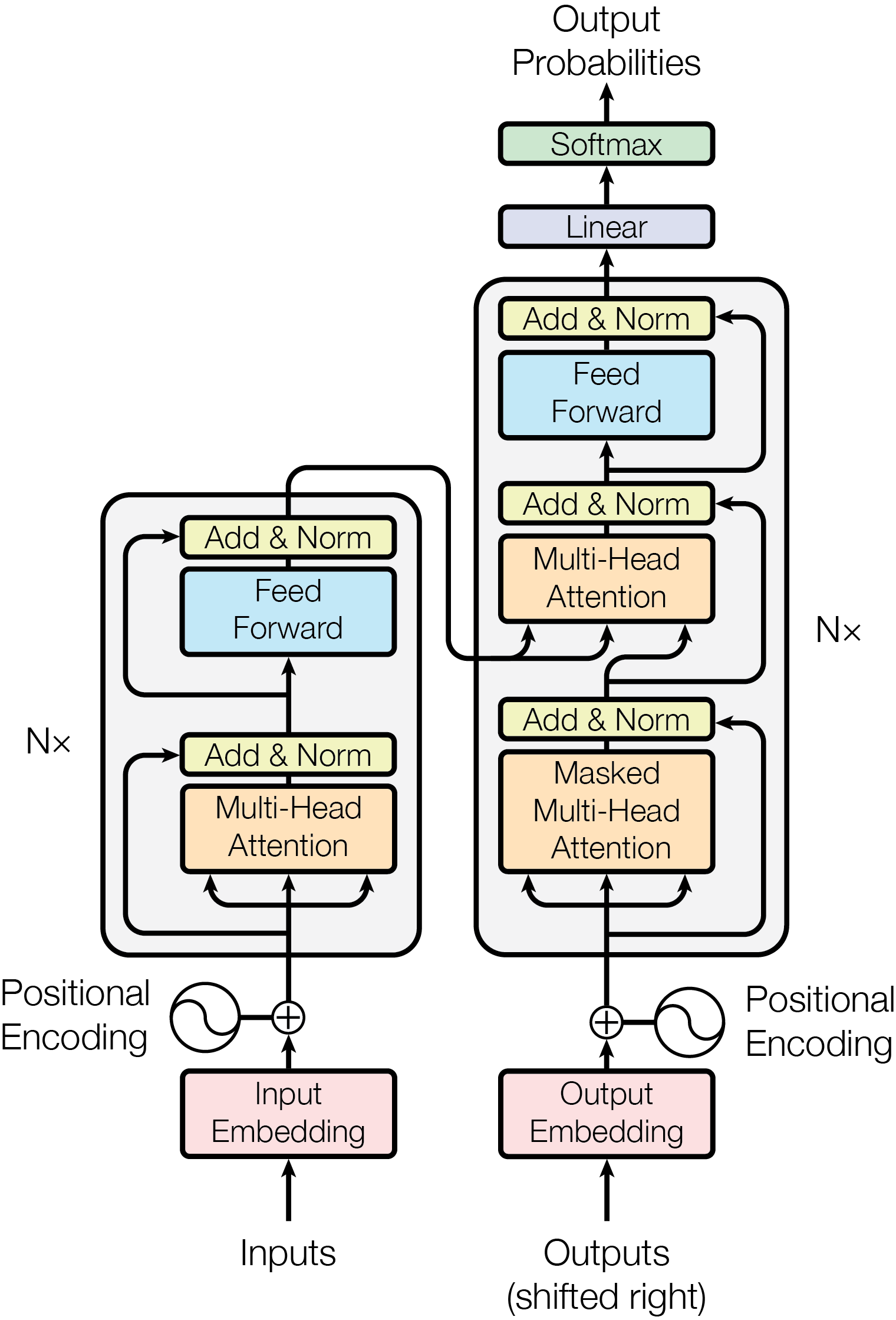
La arquitectura presentada por Vaswani et al. (2017) es una que solamente usa un mecanismo de atención entre el decodificador y el codificador, a este la llaman el Transformador. La idea solo usar el mecanismo de atención proviene de art´ıculos anteriores donde se usaban arquitecturas de RNs convolucionales que inclúıan un decodificador y un codificador, pero además agregaban un mecanismo de atenci´on entre el decodificador y el codificador, logrando as´ı mejores resultados que las RNs convolucionales que no usan este mecanismo.

Para profundizar, el Transformador tiene dos partes, el decodificador y el codificador. El codificador mapea una secuencia de representaciones con s´ımbolos (*x*1*,...,xn*) a una secuencia de representación continua *z* = (*z*1*,...,zn*)[[4]](#footnote-4). El decodificador a su vez, toma *z* y genera una secuencia de s´ımbolos (*y*1*,...,ym*) un s´ımbolo a la vez. Agregando a lo anterior, el modelo es auto-regresivo en cada paso, en otras palabras el modelo usa de input el s´ımbolo generado anteriormente para generar el siguiente s´ımbolo. Como se ve en la Fig. 1 del art´ıculo de (Vaswani et al., 2017), lo anterior se traduce en en m´odulos de “Atención Multi-Cabezal”, de “Atención Multi-Cabezal Enmáscarada”, de “Retroalimentaci´on” y por último de “Softmax”. De estos, el de mayor inter´es es el de Atención Multi-Cabezal, este puede verse como una función que mapea una consulta y un conjunto de pares llave-valor a una salida, donde la consulta, las llaves y los valores son vectores. En este m´odulo es donde aparecen grandes mejoras, ya que este modulo es altamente paralelizable, lo que se traduce a una reducci´on importante de la cantidad de horas que se necesitan para entrenar el modelo completo.

**GAN Bilingüe, arquitectura de Rashid et al. (2019)**

# Figuras

Figura 1: Fig. 1 del art´ıculo de (Vaswani et al., 2017)



# Referencias

Artetxe, M., Labaka, G., y Agirre, E. (2018). *A robust self-learning method for fully unsupervised cross-lingual mappings of word embeddings.* Descargado de [https://arxiv.org/ abs/1805.06297](https://arxiv.org/abs/1805.06297)

Rashid, A., Do-Omri, A., Haidar, M. A., Liu, Q., y Rezagholizadeh, M. (2019). *Bilingualgan: A step towards parallel text generation.* Descargado de [https://arxiv.org/abs/ 1904.04742](https://arxiv.org/abs/1904.04742)

Sultana, F., Sufian, A., y Dutta, P. (2019). Advancements in image classification using convolutional neural network. *CoRR*, *abs/1905.03288*. Descargado de [http://arxiv](http://arxiv.org/abs/1905.03288)

[.org/abs/1905.03288](http://arxiv.org/abs/1905.03288)

Turovsky, B. (2016). *Found in translation: More accurate, fluent sentences in google translate.* Descargado 2020-01-17, de [https://blog.google/products/translate/ found-translation-more-accurate-fluent-sentences-google-translate/](https://blog.google/products/translate/found-translation-more-accurate-fluent-sentences-google-translate/)

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). *Attention is all you need.* Descargado de [https://arxiv.org/abs/1706](https://arxiv.org/abs/1706.03762)

[.03762](https://arxiv.org/abs/1706.03762)

1. La tecnolog´ıa detr´as del “Machine Learning” [↑](#footnote-ref-1)
2. El problema de procesamiento de lenguaje natural (NLP). [↑](#footnote-ref-2)
3. i.e, uno de los problemas asociados al problema de traducci´on [↑](#footnote-ref-3)
4. i.e., es un punto en R*n* [↑](#footnote-ref-4)