

Facultad de Ciencias, UNAM  
Redes Neuronales  
*Proyecto final: Fake News Detection Dataset*

Rosado Cabrera Diego  
Rubí Rojas Tania Michelle

24 de mayo de 2020

## 1. Introducción

La diseminación de noticias falsas, bulos o rumores con el objetivo de manipular la opinión pública es un asunto que cada vez preocupa más en todo el mundo.

Las *fake news* nos llaman la atención por que están protagonizadas por instituciones o personajes públicos que han hecho o dicho algo controvertido, o bien, relatan hechos sorprendentes. Suelen ser noticias polémicas, que provocan la indignación de la sociedad en general.

## 2. Proyecto

Nuestro proyecto consistirá en construir un modelo para la detección de *fake news* usando el algoritmo de clasificación *Passive-Aggressive* (ya que puede clasificar conjuntos masivos de datos, se puede implementar de forma rápida y se ha comprobado que es mejor que el *MLP*).

El conjunto de datos lo obtendremos de la siguiente página

<https://www.kaggle.com/c/fake-news/data>

Este es un archivo *CSV* que contiene 7796 filas con 5 columnas. Donde las columnas refieren el id de la noticia, el título, el autor, el texto de la noticia y una etiqueta de *True (0)* o *False (1)*.

Aunque existe una controversia entre cual es la *verdad*, nosotros nos basaremos en el concepto de *verdad* que fue utilizado para construir este conjunto de datos.