

# Detección de Depresión con Transformadores

Héctor Gibrán González Leal [A01282778]

Tecnológico de Monterrey

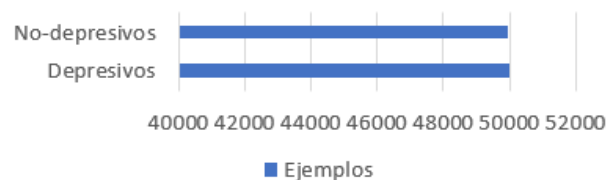
## Introducción

La depresión es un trastorno mental que afecta alrededor de 300 millones de personas alrededor del mundo. Por eso, es un tema que ha sido muy estudiado en el área de Inteligencia Artificial. Este proyecto tuvo como objetivo construir un modelo de NLP para detectar si un texto tiene un tono depresivo o no. Se construyó un set de datos a partir de textos de Reddit y se aplicó transfer-learning en 3 modelos distintos basados en transformadores para hacer el análisis. Los resultados que se muestran se evaluaron usando 3 métricas distintas. Finalmente, se incluyen algunas limitaciones que se identificaron en el desarrollo del proyecto.

## Métodos

El set de datos se construyó a partir de textos recolectados de Reddit. Principalmente, se usaron textos de **r/depression** y **r/SuicideWatch** marcados como depresivos (etiquetados como **1**). Para los no-depresivos (etiquetados como **0**), se tomaron textos de los siguientes 5 subreddits:

- ShowerThoughts
- Wholesome
- MadeMeSmile
- UpliftingNews
- HumansBeingBros

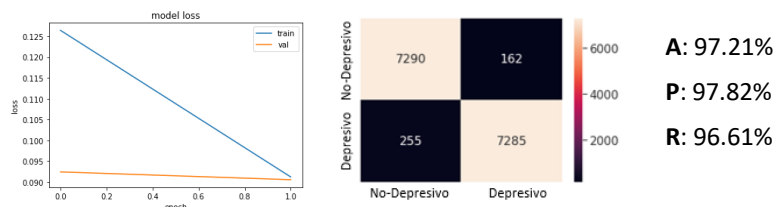


Se aplicó un proceso de preprocesamiento a los datos, tras lo cual se alimentaron a los siguientes 3 modelos: **BERT**, **RoBERTa** y **XLNet**. Para los 3 se usó lo siguiente para realizar el transfer-learning:

Hiperparámetros	Métricas
Batch size de <b>128</b>	Binary Accuracy
Learning rate de <b>0.01</b>	Precision
Epochs de <b>2</b>	Recall
	Confusion Matrix

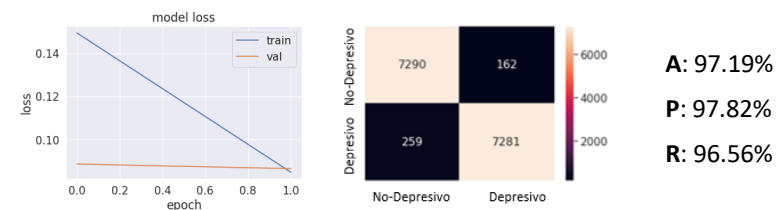
## Resultados

### BERT



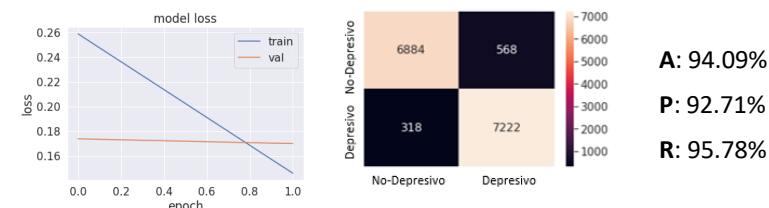
Este modelo desde el principio tuvo métricas muy altas, lo cual se puede validar con la gráfica de costo. Aún con sólo 2 epochs, se puede decir que el transfer-learning fue suficiente para obtener buenas predicciones.

### RoBERTa



Aquí los resultados fueron prácticamente idénticos a BERT. La diferencia principal que hubo entre ambos fue que la gráfica de costo salió incluso más baja.

### XLNet



Las métricas que se obtuvieron con este modelo no fueron bajas, pero tampoco fueron igual de altas a los otros 2 modelos. Esto podría ser porque se debió hacer un transfer-learning más largo.

## Conclusiones

Este proyecto trató de atacar uno de los problemas más presentes en el mundo. Los resultados que se obtuvieron fueron muy buenos, no obstante, se identificaron algunas áreas de oportunidad y limitaciones que tuvo este proyecto:

- Los modelos solamente detectaron textos que muy explícitamente hablen sobre depresión.
- No se está incorporando ningún tipo de metadatos al análisis. Por ejemplo, se pudiera contabilizar la cantidad de pronombres o de palabras con connotación negativa en el texto.
- El transfer learning que se hizo fue bastante pequeño, por lo que con más tiempo se puede lograr un aprendizaje mejor.