Proyecto Integrador Final  
Grupo 2

# Integrantes

Fernando Roa

Kevin Tasat

Martha Alvarez

# Pre Entregable

## Tema de Investigación

Interpretación de emociones a partir de procesamiento de lenguaje natural.

## Antecedentes

A partir de lo investigado a lo largo del curso, nos llama la atención la aplicación de modelos de machine learning sobre el área de procesamiento de texto. De este modo una aplicación que consideramos viable es la identificación de emociones a partir de textos.

## Aporte Esperado

Nuestro objetivo es poder abordar el área del procesamiento de lenguaje natural para interpretar, a través de algoritmos, la semántica de una frase. Si bien sabemos que es la primera vez que abordaremos este tema y existen técnicas que no veremos en el curso, como Redes Neuronales, intentaremos generar un pipeline que nos permita conocer la descomposición de frases en palabras relevantes e irrelevantes, la unión de palabras relevantes para tratar de construir el sentido, y obtener, mediante aprendizaje supervisado, un modelo con la librería NLTK que permita interpretar la emoción subyacente de la frase.

## Disponibilidad de Datos e Infraestructura

Utilizaremos el siguiente [dataset](https://www.kaggle.com/datasets/praveengovi/emotions-dataset-for-nlp) de Kaggle para este trabajo.

## Plan de Trabajo y Cronograma Tentativo

1. **Abordar el tema**: Comenzar a investigar sobre el procesamiento del lenguaje natural, basado en las futuras clases de DH respecto a este tema.
2. **Interacción con el Dataset**: EDA.
3. **Modelo**: Adaptación del algoritmo y revisión de resultados.
4. **Tuning**: Lograr la mejor capacidad predictiva mediante la revisión de las técnicas utilizadas e investigación sobre nuevas técnicas.
5. **Revisión Final**: Documentar el proceso, los resultados obtenidos y nuestro aprendizaje.