ITESM Campus Monterrey

**Proof of Concept: Leveraging Transformers for Text Classification**

Luis Alberto Portilla López

Research Stay - Going beyond Artificial Intelligence: Artificial Emotions

TC3073 | Group 573

## Introduction

The transformer architecture has revolutionized natural language processing (NLP) by enabling efficient and scalable sequence processing through self-attention mechanisms. Transformers form the backbone of state-of-the-art models like BERT, GPT, and T5, which excel in tasks ranging from text classification to question answering. This Proof of Concept (PoC) demonstrates how transformers can be applied to text classification, offering significant improvements in accuracy and scalability.

## Business Problem

Text classification is a critical NLP task with applications in spam detection, topic categorization, and sentiment analysis. However, traditional approaches face limitations such as:

* Context Handling: Difficulty in capturing long-range dependencies in text.
* Efficiency: Inefficient processing of large and diverse datasets.
* Generalization: Challenges in adapting to diverse domains and vocabularies.

## Proposed Solution

This PoC proposes leveraging pre-trained transformer models for text classification tasks. The approach includes:

* Model Selection: Use a pre-trained transformer model such as BERT or DistilBERT for fine-tuning.
* Dataset: Train and evaluate the model on labeled datasets like AG News for topic classification or IMDb for sentiment analysis.
* Implementation: Use frameworks like Hugging Face's Transformers library to fine-tune and evaluate the model.
* Comparison: Benchmark transformer performance against traditional models (e.g., Logistic Regression, LSTMs) on the same datasets.

## Expected Outcomes

The implementation is expected to achieve:

* Higher Classification Accuracy: Improved performance due to transformers' ability to capture complex patterns.
* Scalability: Efficient handling of large datasets with diverse text inputs.
* Generalization: Robust performance across various domains and text types.

## Conclusion

This PoC demonstrates the potential of transformers in advancing text classification tasks. By addressing limitations in traditional approaches, transformers provide a scalable and efficient framework for NLP applications. The findings will serve as a foundation for broader adoption of transformer-based solutions in text classification and other NLP tasks.