



Instituto Tecnológico y de Estudios Superiores de Monterrey
Inteligencia artificial avanzada para la ciencia de datos II

Implementación de un modelo de Deep Learning para la clasificación de emociones en texto

Profesor:

Benjamín Valdés Aguirre

Paul Park - A01709885

9 de noviembre de 2025

Introducción

El reconocimiento automático de emociones en texto es un desafío del procesamiento de lenguaje natural (NLP) que busca identificar sentimientos como alegría, tristeza, enojo o miedo a partir del contenido textual.

Este proyecto implementa una red neuronal **LSTM bidireccional** usando

TensorFlow/Keras, capaz de clasificar oraciones en seis emociones: *joy*, *sadness*, *anger*, *fear*, *love*, *surprise*.

Descripción del dataset

El conjunto de datos está compuesto por tres archivos (**train.txt**, **val.txt**, **test.txt**), con miles de oraciones en inglés, cada una asociada a una etiqueta emocional.

Ejemplo de registro:

I am so happy today;*joy*

I'm afraid of what will happen;*fear*

Cada archivo contiene frases preprocesadas para entrenamiento, validación y prueba, respetando una proporción aproximada de 80/10/10.

Metodología

Preprocesamiento

1. Codificación de etiquetas con **LabelEncoder**.
2. Tokenización de texto con **Tokenizer(num_words=10000)**.
3. Padding de secuencias a longitud fija (**max_len = 50**).
4. Separación en tres subconjuntos (train, val, test).

Modelo base

- Embedding de 10,000 palabras → 64 dimensiones.
- LSTM bidireccional (64 unidades).
- Dropout (0.5).
- Capa densa final con activación *softmax*.
- Entrenamiento por 10 épocas, *batch_size* 32.
- Pérdida: *sparse categorical crossentropy*.

Modelo mejorado

- Embedding de 128 dimensiones.
- LSTM bidireccional (128 unidades).
- Dos capas adicionales de Dropout (0.6 y 0.4).
- Capa densa ReLU (64 unidades).
- *EarlyStopping* en **val_loss**.

Resultados

Modelo	Precisión en prueba	Observaciones
Base	0.90	Buen equilibrio entre bias y varianza
Mejorado	0.90	Mayor estabilidad y menor pérdida en validación

Gráficamente, las curvas de *accuracy* y *loss* muestran convergencia rápida sin sobreajuste significativo.

Análisis

El modelo logró identificar correctamente la emoción dominante en la mayoría de las frases. Las categorías **joy** y **sadness** son las más fácilmente diferenciables, mientras que **love** y **surprise** presentan mayor confusión debido a su ambigüedad contextual. La regularización mediante *Dropout* y *EarlyStopping* mejoró la estabilidad sin sacrificar precisión.

Pruebas de predicción

Ejemplos ejecutados en el Colab de demostración:

"I feel amazing today!" → **joy**

"I'm worried something bad will happen." → **fear**

"He yelled at me for no reason!" → **anger**

"I love spending time with you." → **love**

Estas pruebas confirman que el modelo generaliza correctamente en textos nuevos.

Conclusiones

- Se implementó exitosamente una arquitectura **LSTM bidireccional** en TensorFlow/Keras.
- Se usó un **dataset real de emociones**.
- Se compararon versiones del modelo y se justificaron las mejoras.
- Se documentó el proceso completo en dos notebooks: entrenamiento y predicción.
- El modelo alcanza un desempeño del **90 % de precisión**, demostrando la eficacia del enfoque recurrente para NLP.

Futuras mejoras

- Incorporar embeddings preentrenados (GloVe, Word2Vec).
- Experimentar con arquitecturas Transformer (BERT).
- Desplegar una interfaz web o API para interacción con usuarios.

Referencias

- TensorFlow / Keras Documentation — <https://www.tensorflow.org>
- Dataset: Kaggle Emotion Dataset — <https://www.kaggle.com/datasets>
- Chollet, F. (2021). *Deep Learning with Python (2nd Edition)*.