Description of train_goemotions_distilbert_colab_v2.py

This script trains and evaluates a multi-label emotion classifier on the GoEmotions dataset using Hugging Face's Transformers library. It is written to be Colab-safe and handles different dataset schemas and missing validation splits. Key functionalities: 1. Dataset Loading and Validation Split - Loads the GoEmotions dataset with either the 'raw' (27 emotions + neutral) or 'simplified' (6 emotions + neutral) configuration. - If the dataset does not provide a validation split, it automatically creates one by taking a fraction of the training set (default 10%). 2. Label Schema Handling - Detects whether the dataset uses a 'labels' column (list of label IDs per example) or a wide schema (one binary column per emotion). - Converts both formats into multi-hot vectors suitable for multi-label classification. 3. Tokenization - Uses AutoTokenizer to convert input text into token IDs, padding/truncating to a maximum length (default 128). 4. Model Setup - Loads a transformer model (default: distilbert-base-uncased) as a multi-label classification head with sigmoid outputs. - The number of labels is automatically set based on the dataset. 5. Training - Uses Hugging Face Trainer with simplified TrainingArguments to avoid compatibility issues across transformer versions. - Default parameters: 3 epochs, batch size 32, learning rate 5e-5. 6. Evaluation - Evaluates on validation and test sets using micro, macro, and weighted F1 scores. - Also generates a full classification report (precision, recall, F1 for each label). 7. Efficiency Snapshot - Counts the number of trainable parameters in the model. - Measures inference latency (ms per batch of 32 examples). 8. Outputs - metrics.json: training, validation, and test metrics, along with run arguments. - classification_report.json: detailed per-class evaluation results. - efficiency_snapshot.json: trainable parameters and average inference latency. - Model checkpoints saved in the specified output directory. This script provides a robust baseline for lightweight emotion detection experiments on GoEmotions. It can be adapted to use alternative transformer models such as ALBERT or ModernBERT by changing the --model_name argument.