**Model Card: Emotion Detection in Text**

**1. Objective**

This project aims to develop a deep learning model capable of detecting human emotions in short text (tweets). By analyzing language used on social media, the model can classify text into predefined emotion categories such as *joy*, *sadness*, *anger*, *fear*, *surprise*, and *love*. This tool can be valuable in applications like mental health monitoring, crisis detection, marketing analysis, and social listening platforms.

**2. Data**

**Dataset:**

* Source: Kaggle Emotion Dataset (https://www.kaggle.com/datasets/praveengovi/emotions-dataset-for-nlp)
* Format: CSV file with text and emotion columns
* Number of samples: ~20,000
* Emotion Labels: joy, sadness, anger, fear, love, surprise

**Class Distribution Example:**

|  |  |
| --- | --- |
| **Emotion** | **Count** |
| joy | 5362 |
| sadness | 4965 |
| anger | 2159 |
| fear | 1937 |
| love | 1304 |
| surprise | 827 |

**Preprocessing Steps:**

* Lowercased text
* Removed punctuation and stopwords
* Tokenized using Keras Tokenizer
* Padded sequences to fixed length
* Split into train (80%), validation (10%), and test (10%) sets

**3. Model**

**Architecture:**

|  |  |
| --- | --- |
| Layer | Configuration |
| Embedding | input\_dim=10000, output\_dim=128 |
| LSTM | 128 units with dropout |
| Dense | Fully connected layer with softmax |

**Hyperparameters:**

* Optimizer: Adam
* Loss Function: Categorical Crossentropy
* Epochs: 10
* Batch Size: 32

**4. Training and Evaluation**

* Trained for 10 epochs
* Monitored accuracy and loss on validation set
* Final test accuracy: **89.40%**

**Sample Results:**

* Text: "I am so sad and disappointed." → Predicted: sadness (confidence: 0.99)
* Text: "This is amazing, I love it!" → Predicted: surprise (confidence: 0.82)
* Text: "I'm feeling anxious about tomorrow." → Predicted: fear (confidence: 1.00)

**Metrics:**

|  |  |  |  |
| --- | --- | --- | --- |
| Emotion | Precision | Recall | F1-score |
| joy | 0.93 | 0.92 | 0.92 |
| sadness | 0.90 | 0.91 | 0.90 |
| anger | 0.85 | 0.84 | 0.84 |
| fear | 0.91 | 0.89 | 0.90 |
| love | 0.88 | 0.85 | 0.86 |
| surprise | 0.81 | 0.79 | 0.80 |

**5. Limitations**

* Only handles short English tweets
* Ignores multi-label emotions (e.g., sadness and anger together)
* Model performance depends on vocabulary coverage
* Does not account for sarcasm or indirect expressions
* Bias in data may lead to misclassification in real-world texts

**6. Group Members and Responsibilities**

**1. Kevin Sangani**

* Data preprocessing
* Model architecture and training
* Final model testing and prediction examples

**2. Parth Navadiya**

* Dataset analysis and class distribution
* Evaluation metrics and accuracy analysis
* Preparing classification reports and visualization support

**3. Faizan Ali**

* Project documentation and model card creation
* Contribution to training, limitation analysis, and formatting
* Support for presentation content and slide organization