Sentiment Analysis Project Using CNN

Introduction:

In this report, we present a comparison between two neural network architectures for text classification tasks. Both architectures utilize deep learning techniques, specifically designed to analyze textual data and make predictions.

Steps:

1. **Data Collection and Labeling:**
   - The Sentiment140 dataset, consisting of **1.6** million tweets, was obtained for analysis. All tweets were considered for the subsequent steps.
   - We assigned binary labels: **'0'** for Negative sentiment and **'1'** for Positive sentiment.
2. **Data Preprocessing:**
   - Prior to analysis, we performed data preprocessing steps. This involved the elimination of **stop words**, **digits**, **special characters**, and **punctuation** from the tweets. As a result, we obtained a structured format where each sentence comprised a list of unique words.
3. **Word Embedding:**
   - We employed word embedding techniques to represent the textual data numerically. This step aimed to convert words into dense vectors, facilitating computational analysis.
4. **Model Selection and Comparison:**
   - We utilized two Convolutional Neural Network (CNN) models to analyze the embedded words. The objective was to compare the performance and efficacy of each model in sentiment analysis tasks.

Architecture 1: Convolutional Neural Network (CNN)

The first architecture employs a Convolutional Neural Network (CNN) for text classification. The architecture consists of the following layers:

- Embedding Layer: Transforms input text into dense vectors of fixed size.
- Dropout Layer: Regularizes the model to prevent overfitting by randomly dropping a fraction of input units.
- Convolutional Layers: These layers use convolutional filters to capture local patterns in the input data.
- Global Max Pooling Layer: Extracts the maximum value across each feature map, reducing the dimensionality of the data.
- Flatten Layer: Flattens the input data into a one-dimensional vector.
- Dense Layers: Fully connected layers with ReLU activation function.
- Output Layer: A single neuron with sigmoid activation function for binary classification.

Architecture 2: Inception Model

The second architecture employs an Inception model inspired by the Inception architecture used in computer vision tasks. The architecture comprises the following components:

- Inception Blocks: Each block contains multiple parallel convolutional pathways to capture features at different scales.
- Global Average Pooling Layer: Computes the average value of each feature map across the spatial dimensions.
- Dense Output Layer: A single neuron with sigmoid activation for binary classification.

Comparison

Accuracy:

- CNN Architecture: Achieved a peak validation accuracy of approximately **78.96%.**
- Inception Model: Achieved a peak validation accuracy of approximately **79.17%.**

Both models demonstrate competitive performance, with the Inception model slightly outperforming the CNN architecture.

Parameter Comparison:

- CNN Architecture: Total parameters: **138,574,397**; Trainable parameters: **46,191,465**.
- Inception Model: Total parameters: **46,170,249**; Trainable parameters: **46,168,009**.

The Inception model has significantly fewer parameters compared to the CNN architecture, which can lead to faster training and inference times, and potentially better generalization on smaller datasets.

Test Inception Model:

We use the trained model to predict **20** random sample from **the Test data**, here's an example:

```
Sample 8:
Tweet: I want tuna &amp;amp; salmon sashimi, B.C. rolls and dragon eye.
Predicted Sentiment: Negative
True Sentiment: Negative

Sample 9:
Tweet: cant find a new show to watch
Predicted Sentiment: Negative
True Sentiment: Negative

Sample 10:
Tweet: @marfita again based on how i feel i'll show up
Predicted Sentiment: Negative
True Sentiment: Negative

Sample 11:
Tweet: @Raquel_Soto Good for you. Things like that are always an adventure. And I'm sure you can turn lemons into lemonade, sweets
Predicted Sentiment: Positive
True Sentiment: Positive

Sample 12:
Tweet: And now for the reason we came here. Frosted flakes   NemoNemesis
Predicted Sentiment: Positive
True Sentiment: Positive
```

Conclusion:

Both architectures offer effective solutions for text classification tasks. While the Inception model demonstrates slightly better accuracy and has fewer parameters, the choice between the two architectures depends on various factors such as computational resources, dataset size, and specific requirements of the task. Further experimentation and fine-tuning may be necessary to determine the optimal architecture for a particular application.