Ex. No. 9	**Transformer model for a text classification task**
Date of Exercise	**10/11/2025**

Aim:

To implement a transformer model for text classification, utilizing the self-attention mechanism to improve performance.

Description:

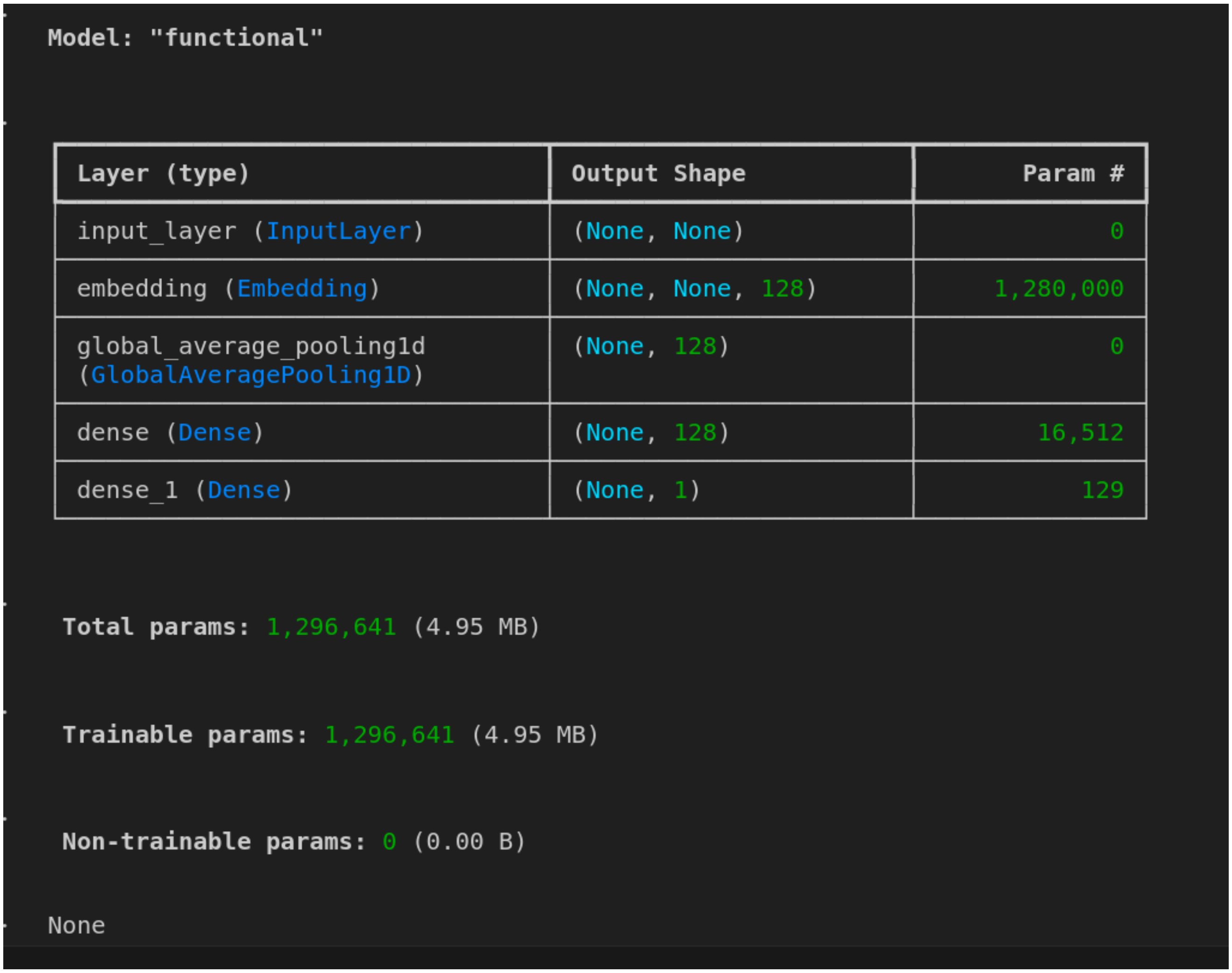
The transformer model is a deep learning architecture that relies entirely on self-attention mechanisms instead of RNNs or CNNs. It is widely used for NLP tasks such as sentiment analysis and text classification.

Code:

```
import tensorflow as tf
from tensorflow.keras.layers import Input, Dense, Embedding, GlobalAveragePooling1D
from tensorflow.keras.models import Model
from tensorflow.keras.optimizers import Adam
# Define input parameters
vocab_size = 10000
embedding_dim = 128
# Define the Transformer-based model
inputs = Input(shape=(None,))
embedding = Embedding(vocab_size, embedding_dim)(inputs)
x = GlobalAveragePooling1D()(embedding)
```

```
x = Dense(128, activation='relu')(x)
x = Dense(1, activation='sigmoid')(x)
model = Model(inputs, x)
model.compile(optimizer=Adam(), loss='binary_crossentropy', metrics=['accuracy'])
print(model.summary())
```

Sample Output:



```
Model: "functional"

┏━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━┳━━━━━━━━━━━━━━━━━━━━━━━━━┳━━━━━━━━━━━━━━━━┓
┃ Layer (type)                     ┃ Output Shape            ┃        Param # ┃
┡━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━╇━━━━━━━━━━━━━━━━━━━━━━━━━╇━━━━━━━━━━━━━━━━┩
│ input_layer (InputLayer)         │ (None, None)            │              0 │
├──────────────────────────────────┼─────────────────────────┼────────────────┤
│ embedding (Embedding)            │ (None, None, 128)       │      1,280,000 │
├──────────────────────────────────┼─────────────────────────┼────────────────┤
│ global_average_pooling1d         │ (None, 128)             │              0 │
│ (GlobalAveragePooling1D)         │                         │                │
├──────────────────────────────────┼─────────────────────────┼────────────────┤
│ dense (Dense)                    │ (None, 128)             │         16,512 │
├──────────────────────────────────┼─────────────────────────┼────────────────┤
│ dense_1 (Dense)                  │ (None, 1)               │            129 │
└──────────────────────────────────┴─────────────────────────┴────────────────┘

 Total params: 1,296,641 (4.95 MB)

 Trainable params: 1,296,641 (4.95 MB)

 Non-trainable params: 0 (0.00 B)

None
```

Result:

The above experiment of Transformer model for a text classification task is done successfully and the output is beed obtained