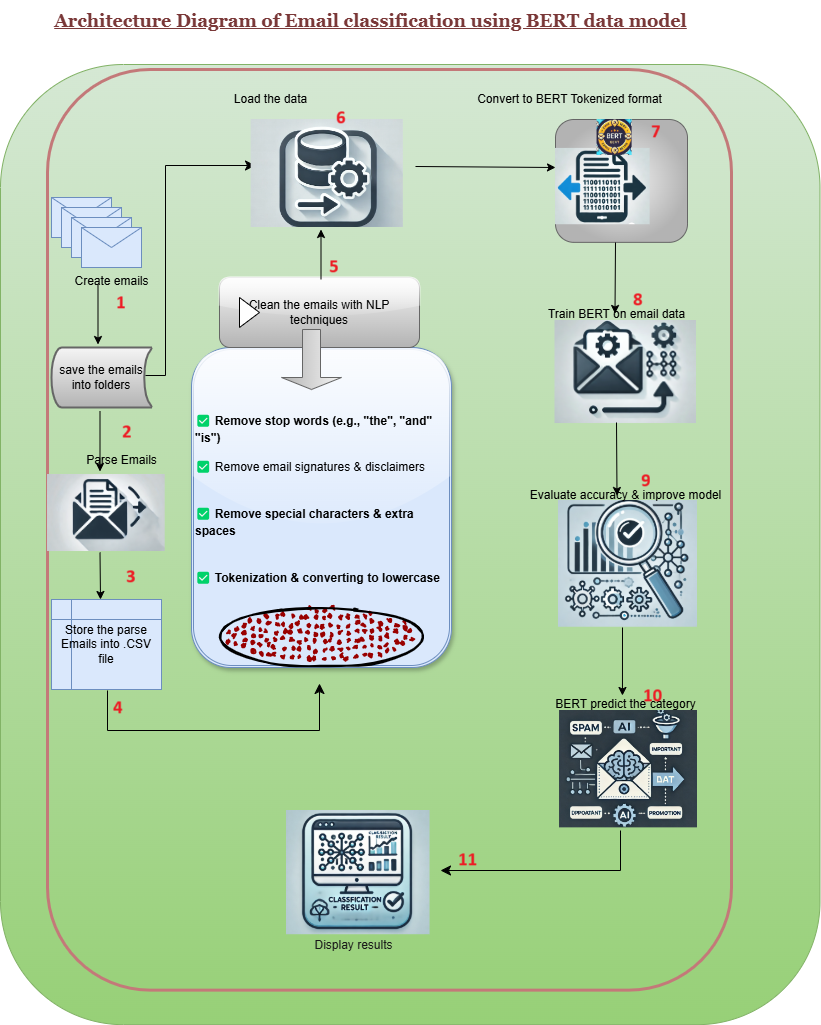
****

**Architecture Steps:**

1️. **Create Emails** – Generate or collect emails for training.  
2️. **Save Emails in a Folder** – Store emails in a structured dataset.  
3️. **Parse Emails** – Extract subject, body, and attachments from .eml or .msg files.  
4️. **Save Parsed Emails into CSV** – Store parsed email data (subject, body, labels) in a CSV file.  
5️. **Clean Email Text Using NLP** – Apply preprocessing (removing HTML, stopwords, special characters).  
6️. **Save Processed Dataset for BERT** – Save cleaned email text in a format compatible with BERT training.  
7️. **Train Model Using BERT** – Fine-tune a BERT model using labeled email data.  
8️. **Fine-Tune BERT** – Improve the model by training on domain-specific email data.  
9️. **Load the Fine-Tuned BERT Model** – Use the trained model for inference.  
10. **Tokenize New Email Text** – Convert incoming emails into BERT token embeddings for classification.

11. **Display the categorized output**: The trained BERT model predicts the category of the email.