1. Overall Architecture of the Attention Based Transformer Model