Attention Mechanism: Where the Model Looks
When Generating Each Word