Lecture 17: Multi-head Attention-part 1

The term "Multi-head" refers to dividing the attention mechanism into multiple heads. Each head operates independently.