

Part 3/7:

Multi-Head Attention

- 8. Why is Multi-Head Necessary?
- 9. Multi-Head Attention Architecture
- 10. Multi-Head Operation Example
- 11. Multi-Head Attention Implementation Points