- Align preferences
- Supplement reasoning
- Elimination method
- Incomplete sentences
- Repetitive words

**ARES**

RL

Score

Teacher

Correction

SFT

Generate

Policy

Generate