KLSBench: All Models Performance Overview
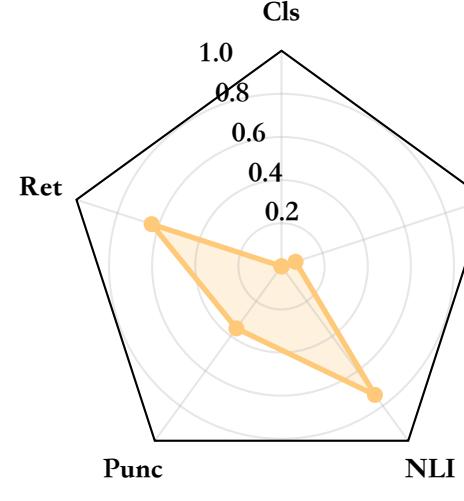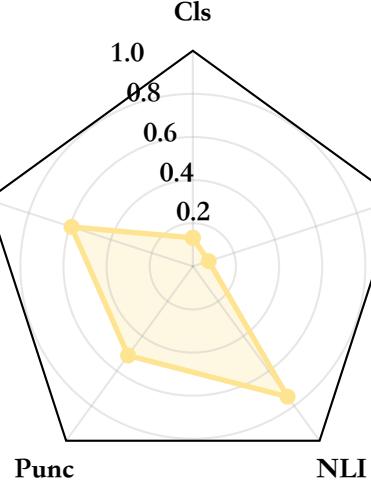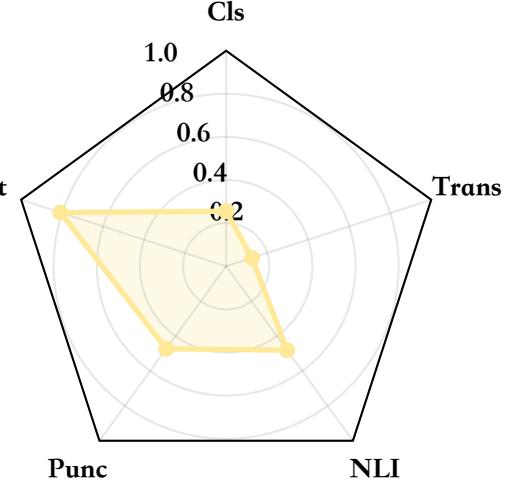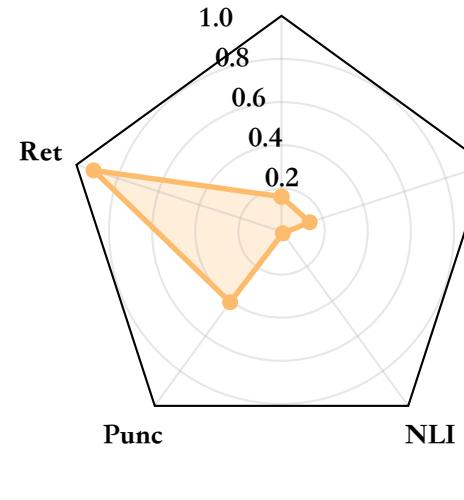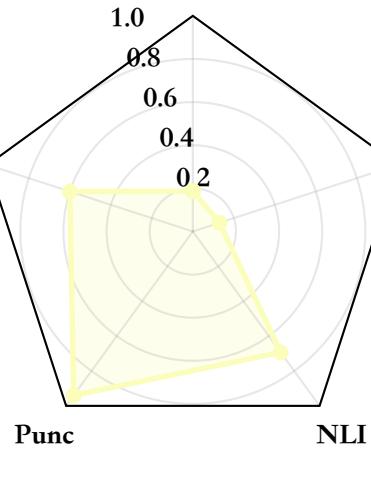

EXAONE-3.0-7.8B-I...
(0.359)

Qwen2.5-7B-Instruct
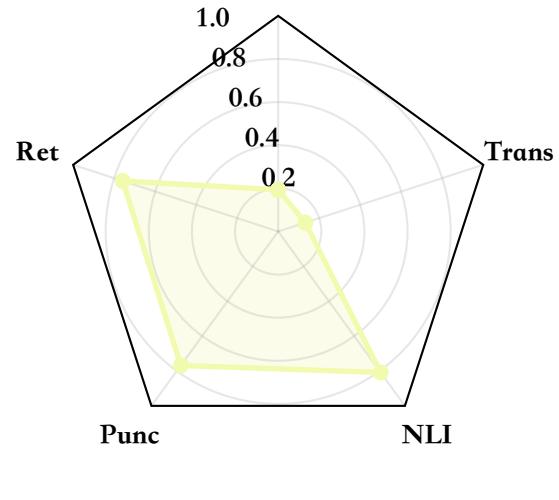(0.412)

claude-3-opus-202...
(0.429)

claude-sonnet-4-5...
(0.326)

gpt-3.5-turbo
(0.510)

gpt-4-turbo
(0.532)

Llama-3.1-8B-Inst...
(0.384)