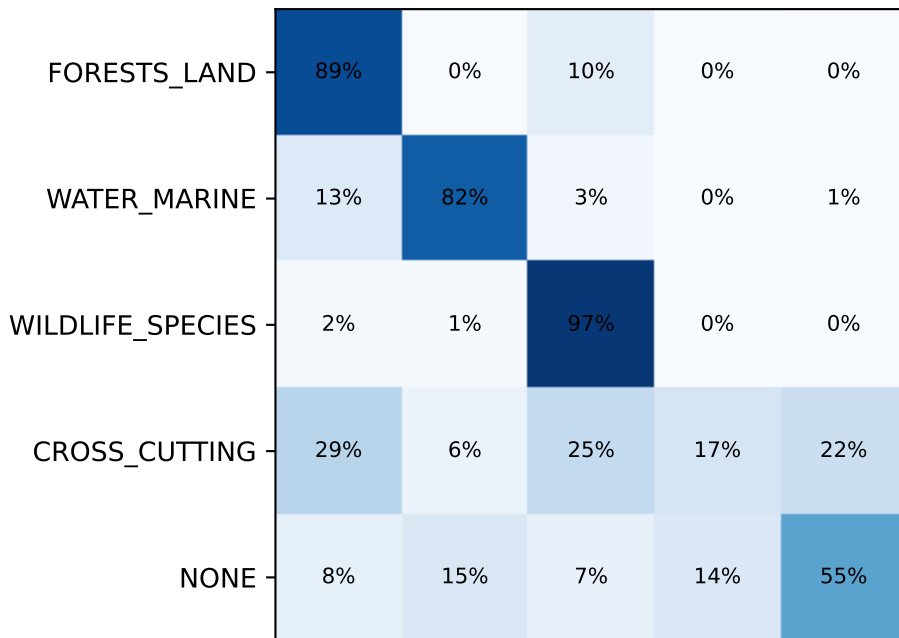


Confusion heatmap: GPT-4o vs GPT-3.5 (row %)

GPT-4o prediction



Row-normalized share (%)

FORESTS_LAND

WATER_MARINE

WILDLIFE_SPECIES

CROSS_CUTTING

NONE

GPT-3.5 prediction