

Overconfidence vs. Macro-F1 Performance

