

Macro F1 Score: Validation vs OOD Test Set

