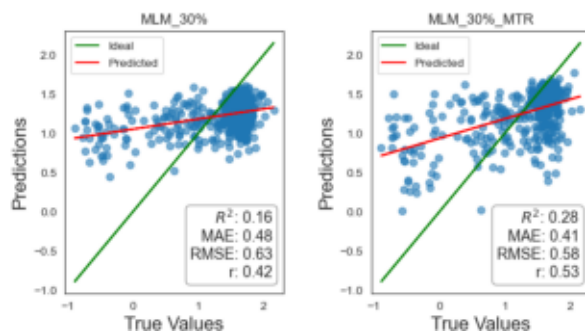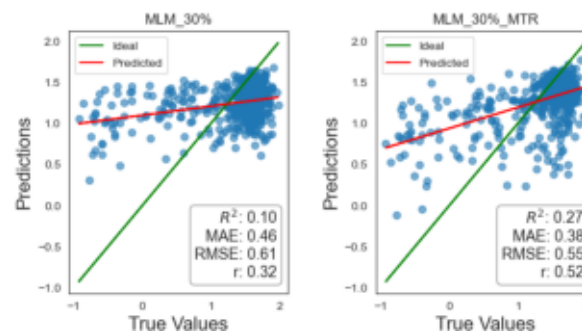# Dataset: Solubility / Pre-training objective: mlm / split: datasail