对于在Math-500数据集上提高模型表现，我认为从两方面来说：

其一，训练策略，论文Let's Verify Step by Step中提出PRM的训练方法，加入了每一步step的标记，在自回归模型进行训练遇到结束位置标记时，训练PRMs去预测每一步step是否正确。这样的过程监督在大规模监督学习时可以得到更好的效果。在构建数据集时，将复杂问题的步骤拆解成多个简单问题，将拆解后的简单问题与原复杂问题混合训练。

使用数据增强方法，对数学问题进行同义替换、公式变形等操作，保留原始问题的核心信息和逻辑结构，使模型学习到更丰富的特征。

其二，模型策略，在本地部署时发现，模型的效果和推理时间是有一定关系的，让模型在遇到简单问题时进行较短时间的推理，面对难题时进行更长时间的思考，随着推理时计算量的增加，在数学难题上的效果应该会得到改进。另外，该方法也会有效减少模型的体量。

引入强化学习提升模型推理能力。可以采用如DeepSeek-R1所使用的GRPO等强化学习方法，结合准确性和格式奖励来增强模型的推理能力，这种方法也无需大量标注数据。