

使用的模型结构：

修复模型可以分为两部分，一部分是使用大量程序语言数据训练出的GPT模型，学习编码特征和程序语言规范，用于输入数据的特征提取；二是一个CNN，根据提取出的特征，学习Bug行和正确补丁间的关系，生成补丁。

补丁评估模型是一个类似于修复模型的分类模型，同样是两层模型结构。

补丁生成的流程：

待修复程序输入后，首先转化为对应的格式，由GPT模型将其转化为向量形式，输入到CNN模型中，生成对应的补丁信息。在生成过程中，由于是逐字生成的，因此，在每一次生成后，都使用补丁评估模型判断其是否可能会存在漏洞，选择评分高的，继续生成。在所有补丁生成结束后，通过语法检查模型，只保留符合语法规则，也就是可以编译的代码，节省验证补丁过程的时间。