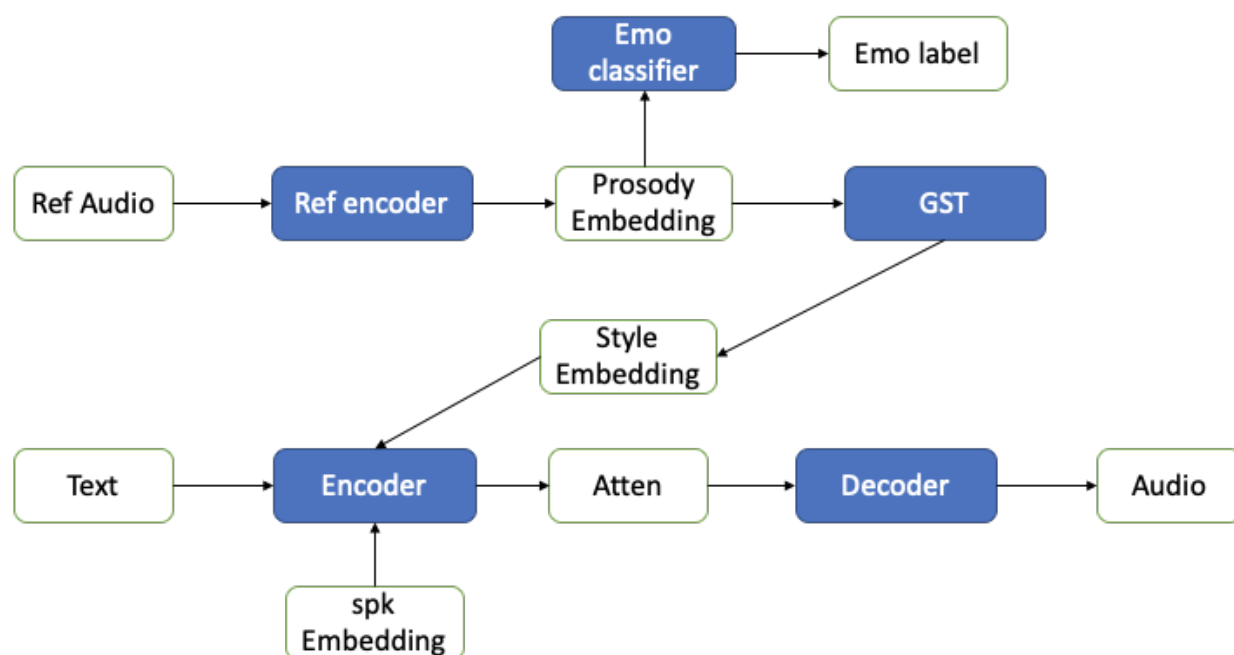


# 情感可控的个性化语音合成

## 模型结构



## 说明

多说话人数据，有很多说话人数据没有情感标签（不知道对应的是什么情感），对应的情感标签标注为 -1, 计算emo loss的时候要加mask。