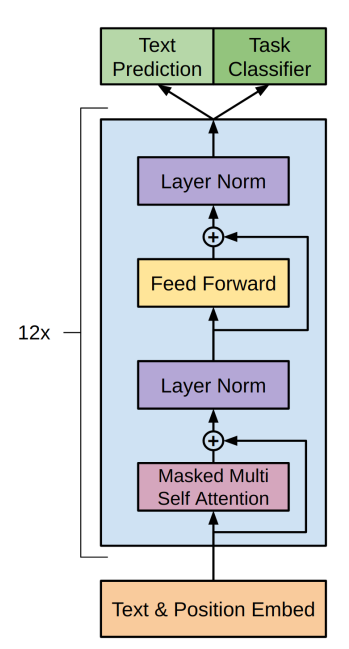
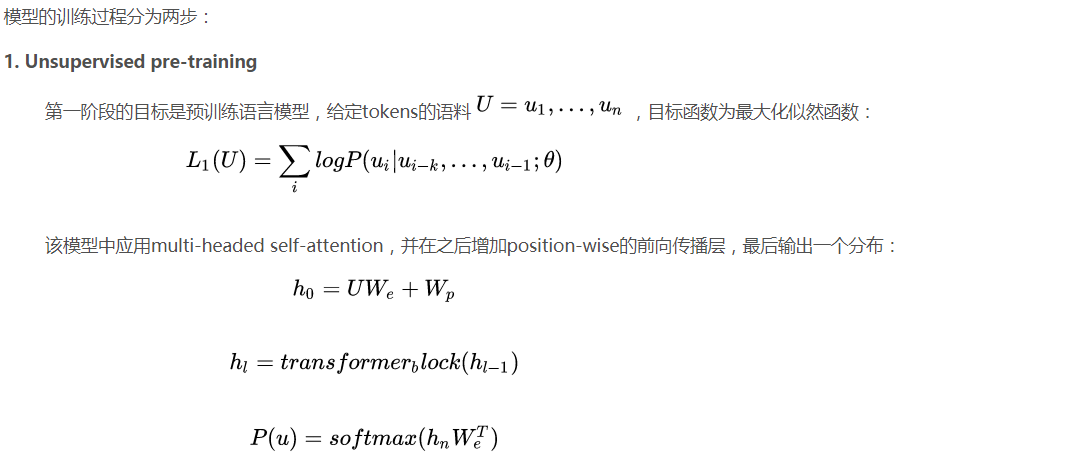
**OpenAI GPT**

**模型原理与架构**

OpenAI Transformer是一类可迁移到多种NLP任务的，基于Transformer的语言模型，其利用Transformer的结构来进行单向语言模型的训练。它的基本思想同ULMFiT相同，都是在尽量不改变模型结构的情况下将预训练的语言模型应用到各种任务。不同的是，OpenAI Transformer主张用Transformer结构，而ULMFiT中使用的是基于RNN的语言模型。



文中所用的网络结构如下：



其中k是上下文窗口的大小，条件概率P由神经网络参数θ建模。这些参数使用随机梯度下降训练。



