

数据建模笔试题

根据某客户2006年用电量数据（见附件），用机器学习模型来预测未来一小时的电量和未来6小时的用电量。数据文件中“time”列是时间戳（日期及相应的小时、分钟），“KWH”是每个时间戳对应的那个小时里的用电量（单位：度）。具体来讲，在每一个小时 t ，我们需要预测（ $t+1$ ）小时的用电量 $q(t+1)$ 。可以使用任意开源的Python packages, such as sklearn, tensorflow, Keras, 等等。使用Mean Absolute Percentage Error来展示所建模型的预测效果。

1. 使用XGBoost模型来建立预测模型。
2. 使用Long Short Term Memory模型来建立预测模型。
3. 描述一下模型所用到的feature set。上述两个模型哪个更适合解决这个问题？
 - 如果有兴趣，还可以解决在小时 t 预测一个6个元素的向量 $[q(t+1), q(t+2), q(t+3), q(t+4), q(t+5), q(t+6)]$ 的问题。