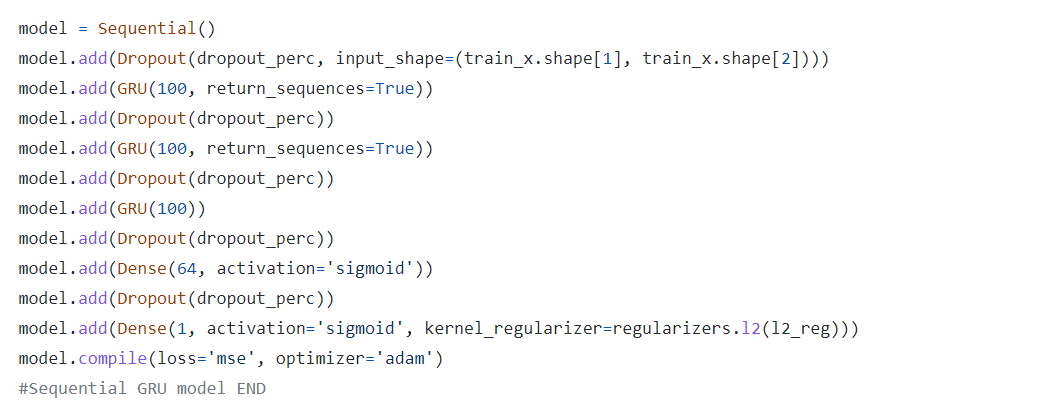
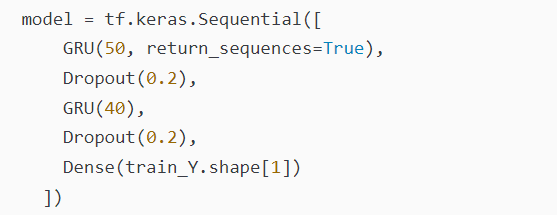
# multilayer原文作者提示的架构

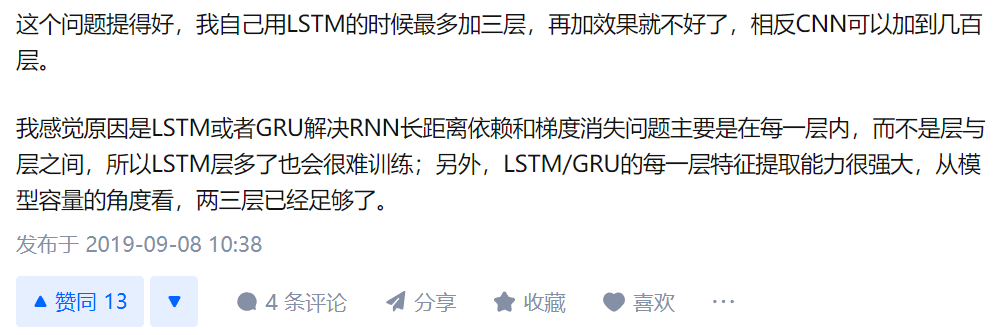


3层，且层层dropout，看来是很复杂的模型

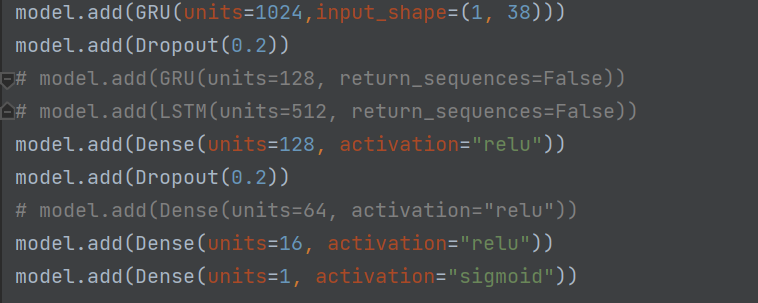


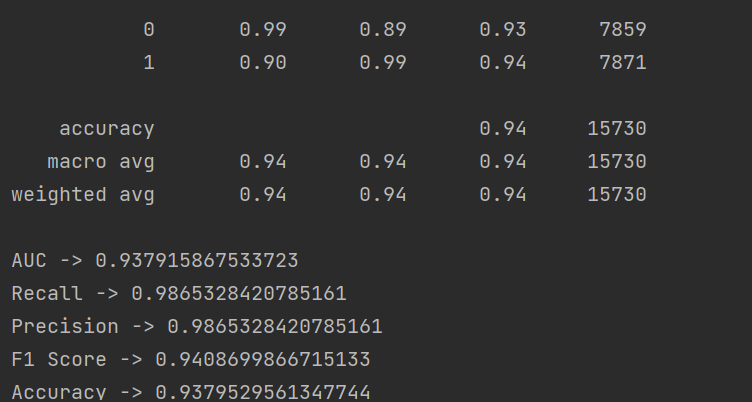
两层，且只对1个shape输出

# 知乎建议

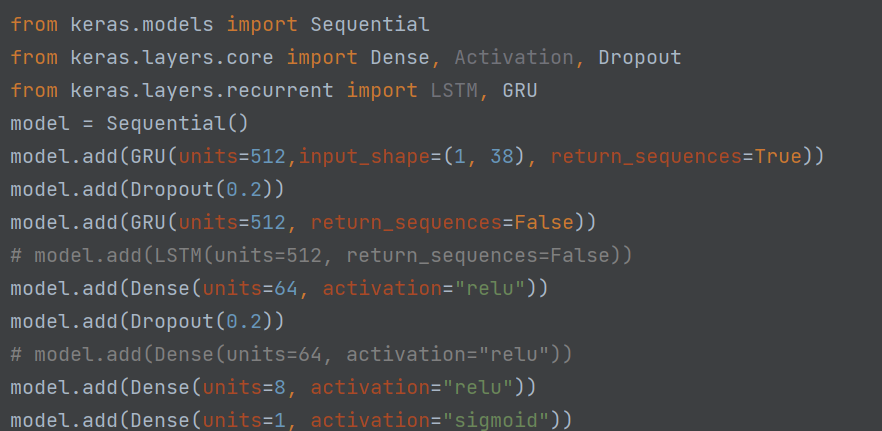


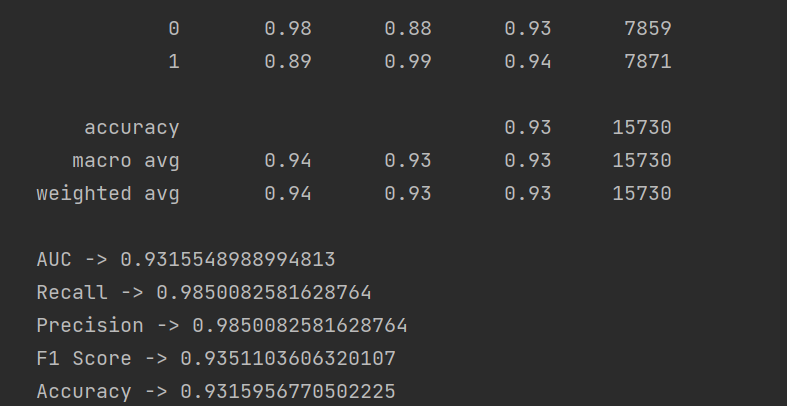
## 单层架构也不差



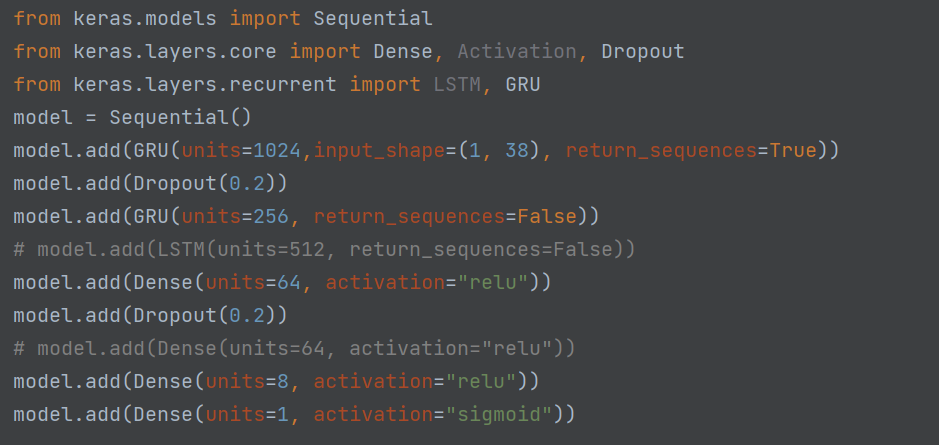


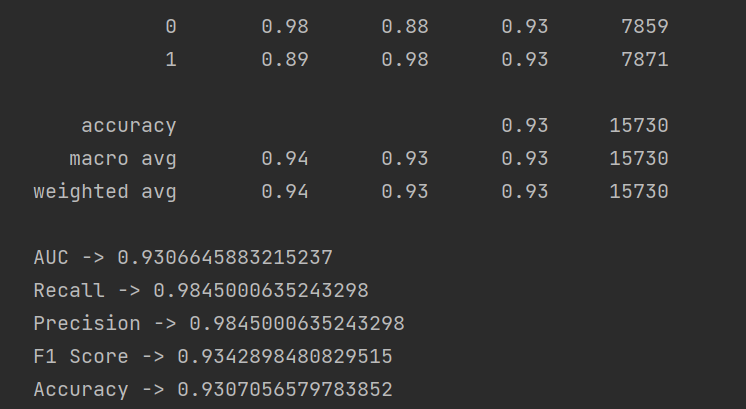
## 注意这才练了5000轮

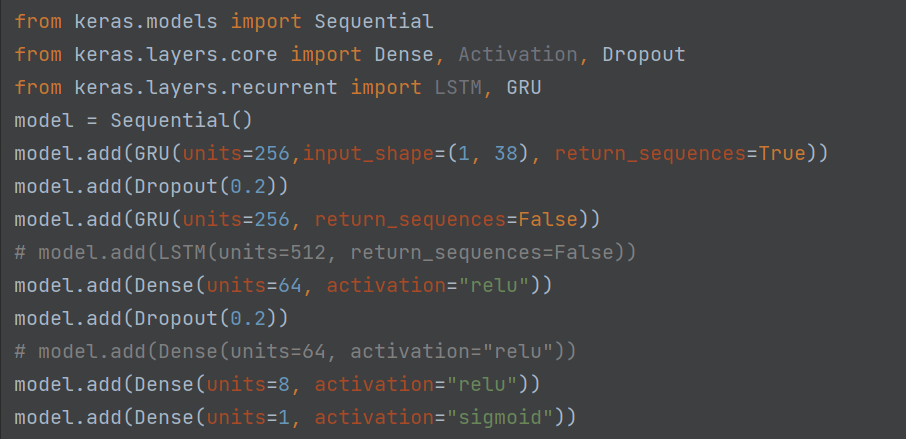


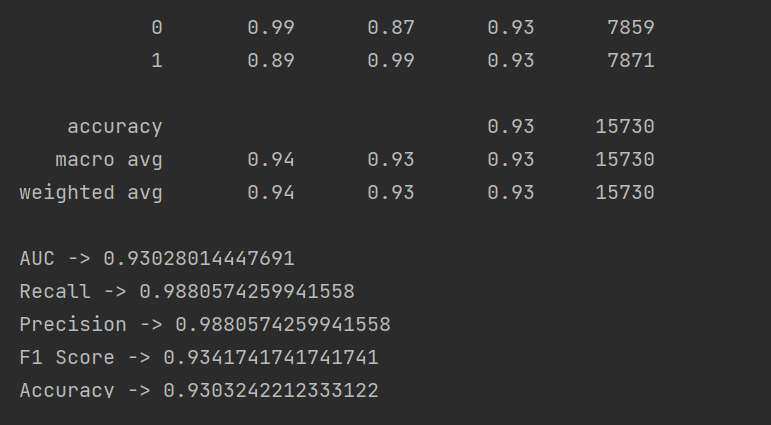


## 不整齐的架构效果更差

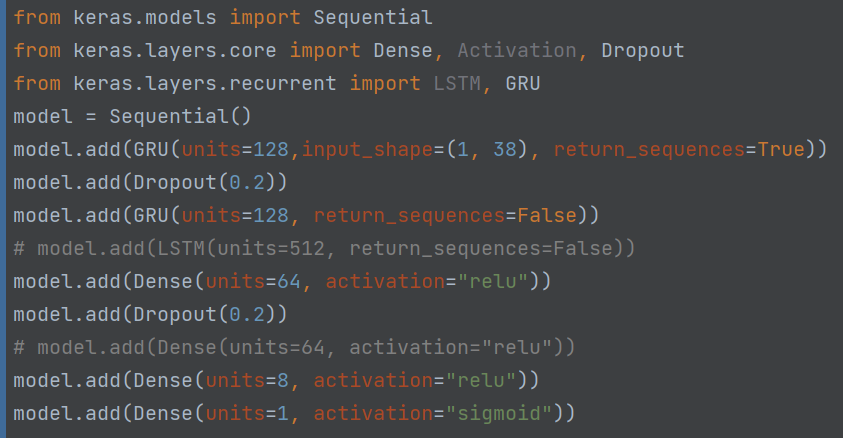


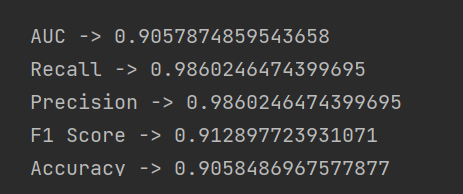




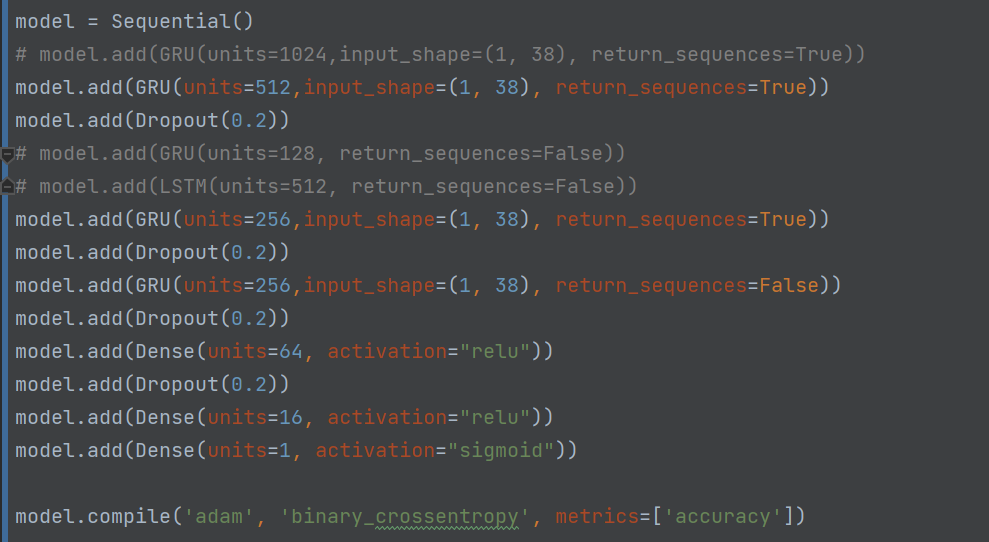


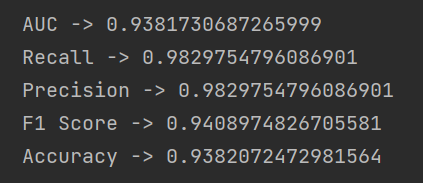
## 隐层少了，跟线性层效果差不多



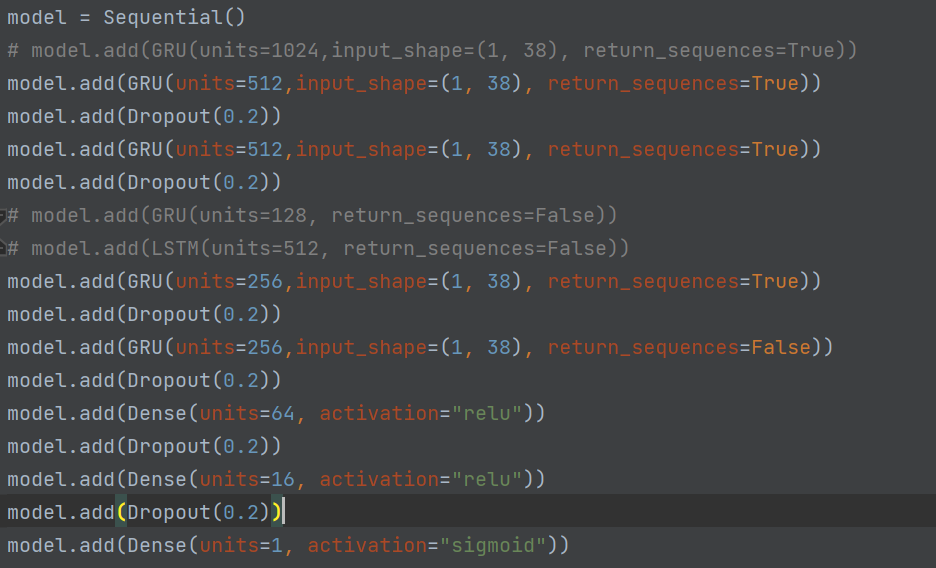


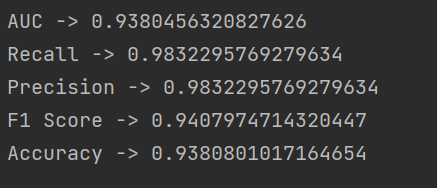
## 堆！往多了GRU堆



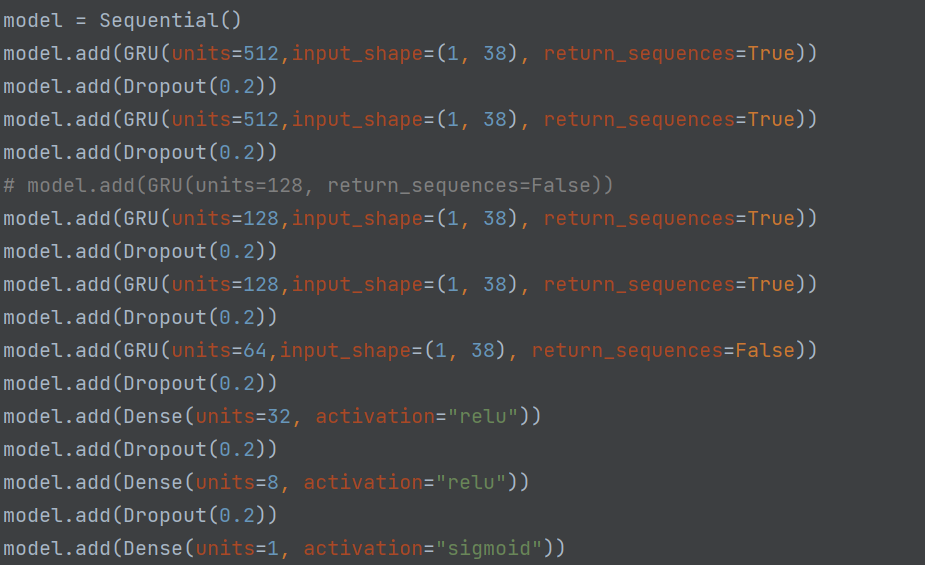


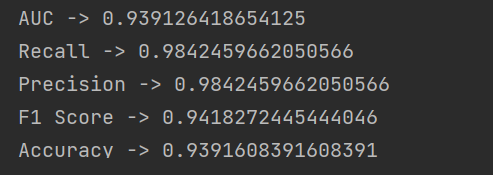
# 4层反而差了





## 5层有点效果





引用→多层GRU加入注意力有效果提升，但不加也有

1. S Jung, J Moon, S Park,等. An Attention-Based Multilayer GRU Model for Multistep-Ahead Short-Term Load Forecasting[J]. Sensors, 2021, 21(5):1639.

# 10层很捞

