以下是关于机器学习的task，要求复刻《Bond Risk Premiums with Machine Learning（RFS）》中的主要操作，并适当延伸

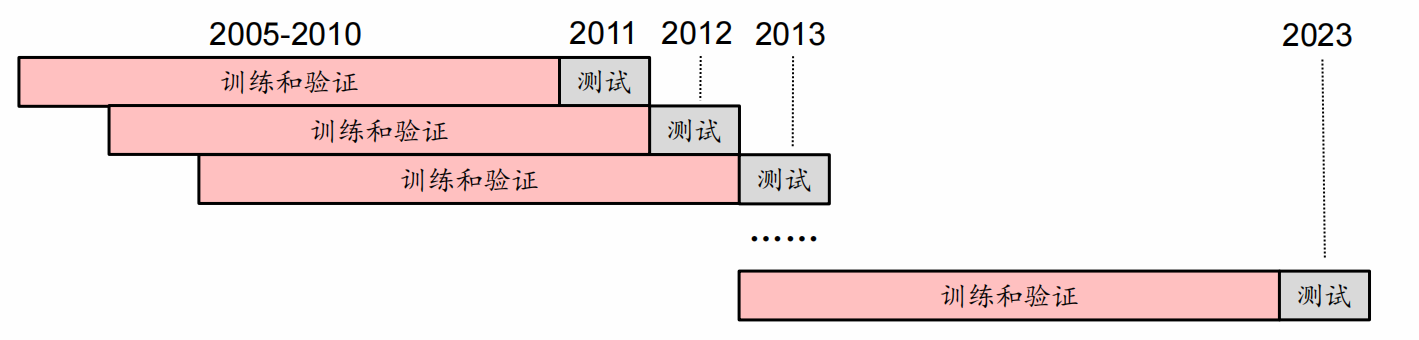
注：第1章描述的数据构建、和2.1的数据来源均与压缩包中同源，时间跨度有所差异。

在2.2章描述了作者用到的机器学习方法，请……

1. 额外使用一个OLS方法，最普通的那个；
2. 在2.2.1 - 2.2.3的机器学习方法，按文中提到的来；
3. 在2.2.4神经网络部分，它使用的层数和节点数的意义，我不是很能理解。如果你准备不按文献中设置的层数和节点数来也可以，比如弄一个NN4，或者减少节点数量，向我解释一下这样做的好处，我需要写实验报告。可以的话，给我推几篇支持上述修改的（机器学习领域神经网络相关）的参考文献。

对每次回归，简要写个报告说明回归的结果（如预测的效果等）

**注意表格中是样本为R^2**

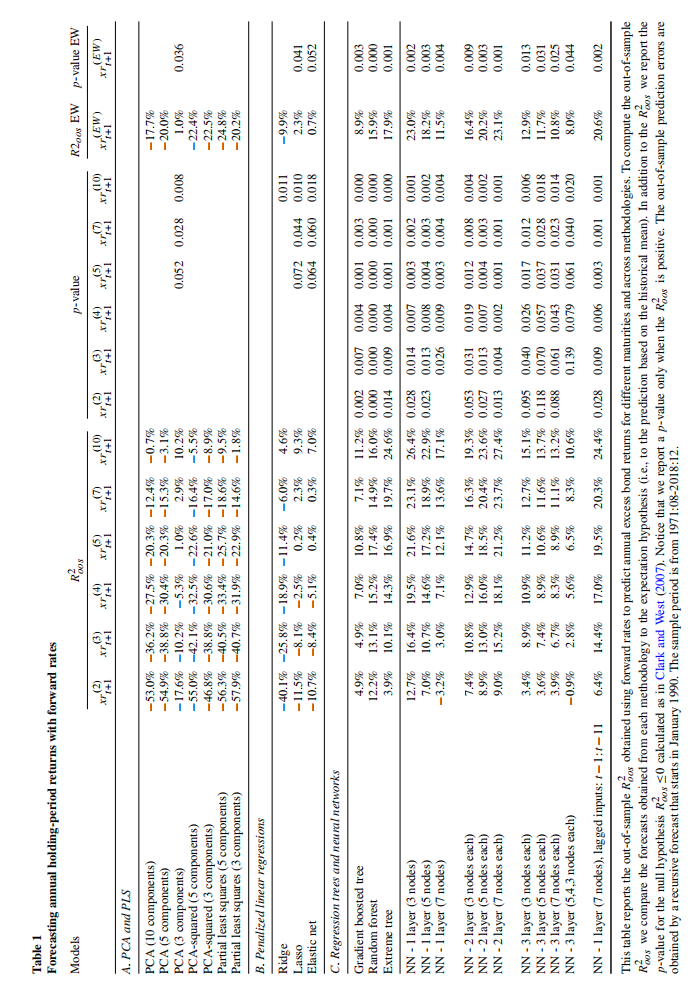


一、复刻3.1章的Table1

使用表格“f-rx(1973.01-2022.12)\_clear”的数据

y\_1, f\_2……,f\_10对应10个用于主成分分析的自变量

rx\_2等为因变量。和文献中跑回归用的那几个一样，即2/3/4/5/7/10，以及这6个的均值rx\_ew1



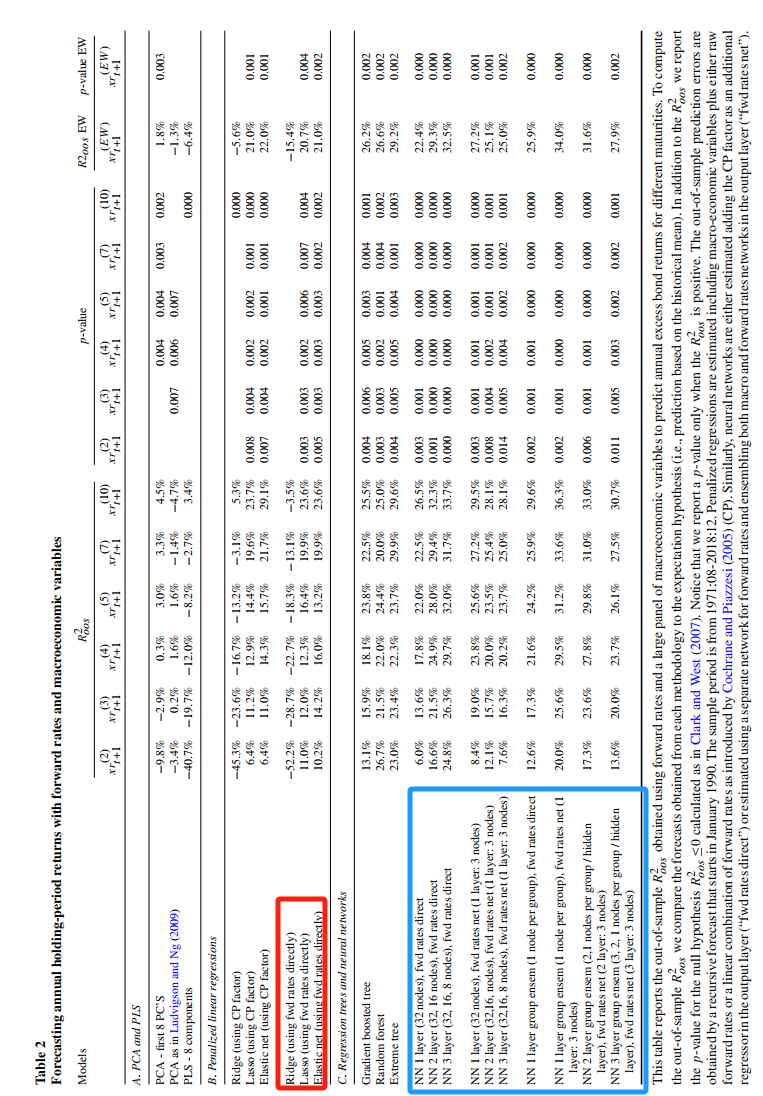
在面板A前面在加一行用OLS预测的结果

可以的话加个NN4

面板C的最后一行，即含滞后项的这个，不用做

二、复刻3.2章的Table2

宏观因子在表格“宏观数据1973.01-2022.12”中（先不用管Group那一行）



一样的，加上OLS

红框中的三个回归删掉

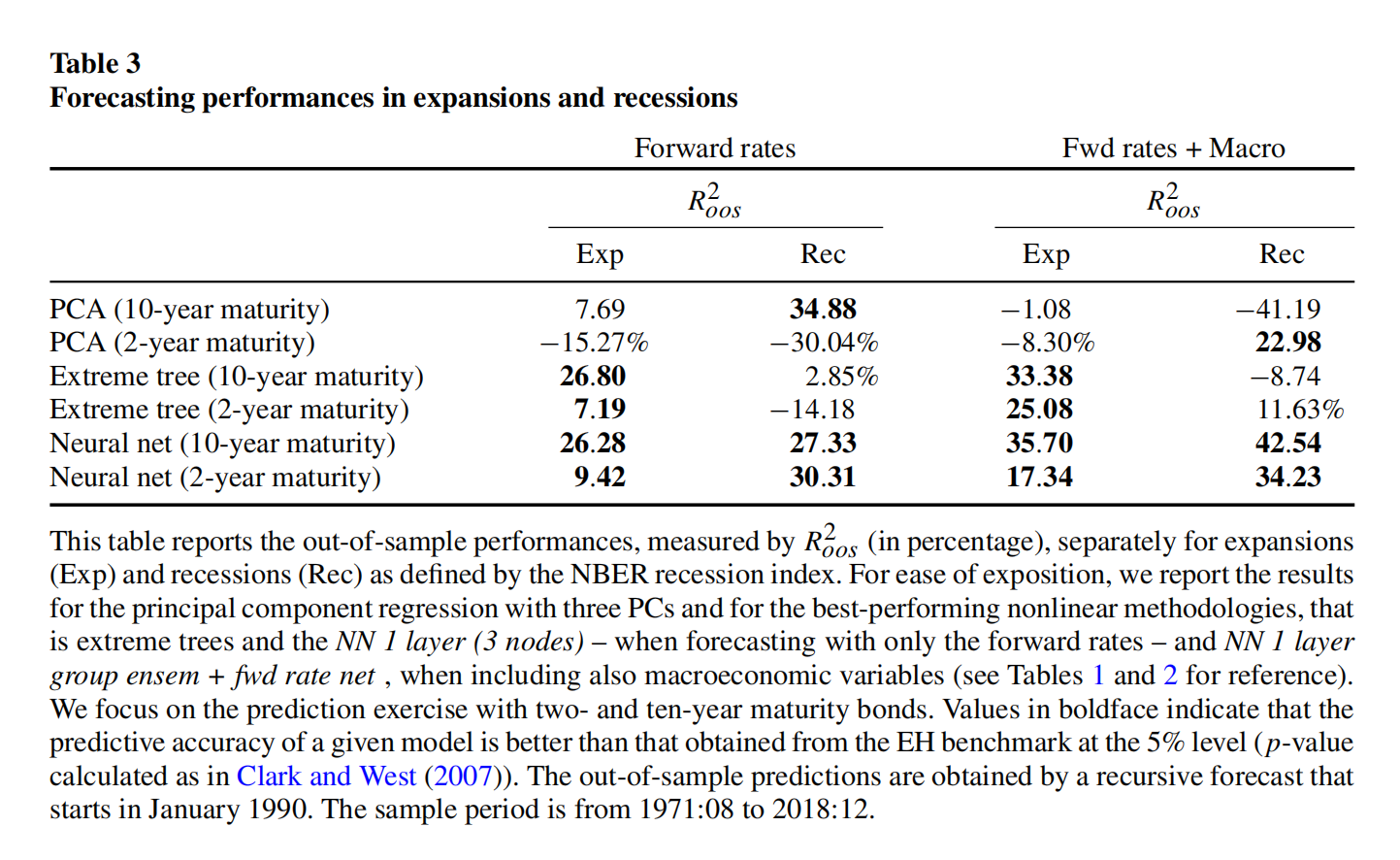
蓝框中的NN，不必完全按文献中的来，可以的话加个NN4。

不用做滞后项的回归

三、复刻4.1章的Table3

用表格“NBER based US\_Recession Indicators”中的虚拟变量，进行分组回归

变量和Table2中一样



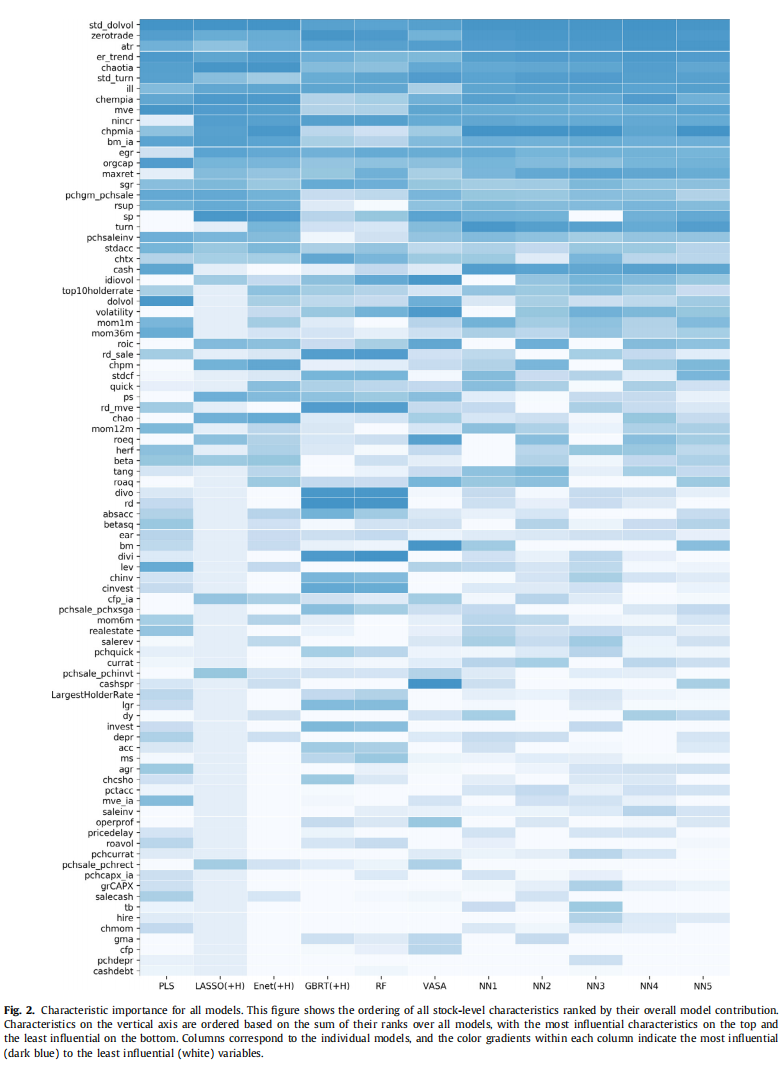
按他的方法来。此基础上加一组xr\_5做因变量的。

四、复刻4.3章的Figure5-a,b，但展示效果不和文献中一样

要求呈现模型中，各宏观因子的重要程度

最终效果类似下页，横轴为前面用到的机器学习方法（包括OLS），纵轴为各宏观因子，越靠上表示该宏观因子在模型中重要性越大

只做xr\_ew1做因变量的那一组，即这个图只用做一个



五、复刻第5章，以及表5

按文献中的公式来就行