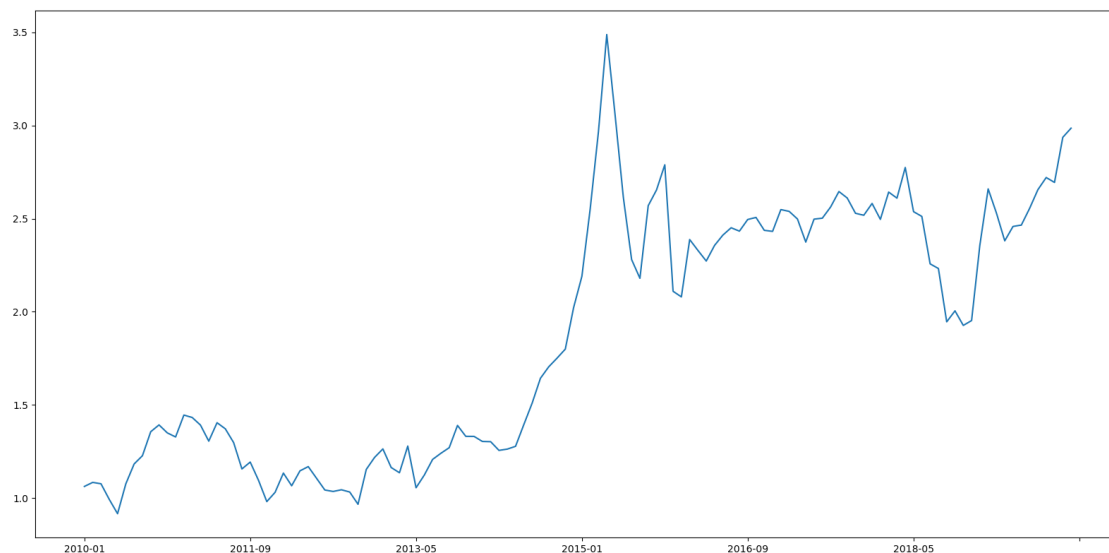


说明

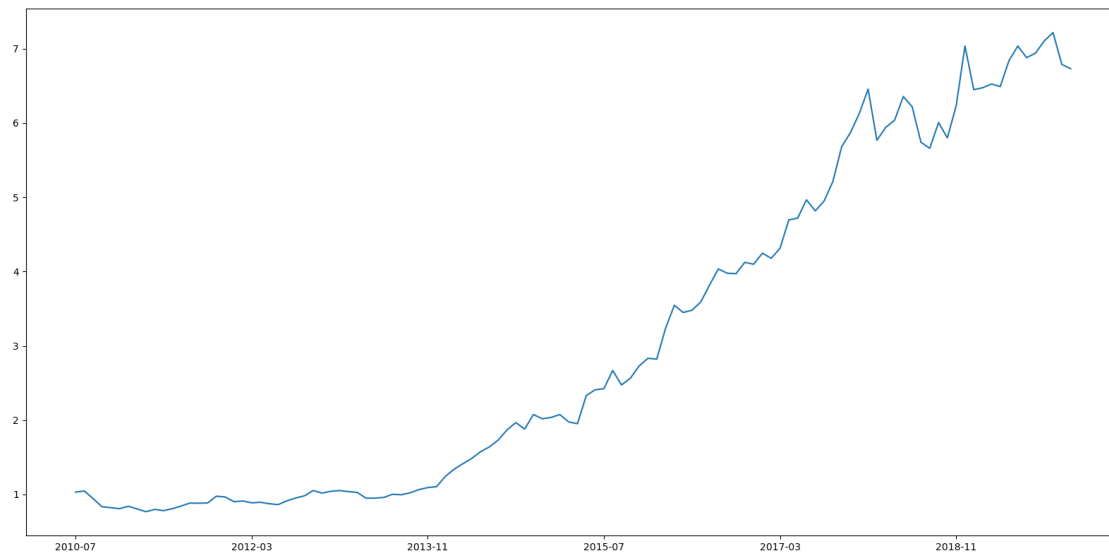
由于数据不一致，所以做出效果与研报不一样。我按照研报思路把研报的几个部分实现了一遍。

1. 基本面模型的构建



对应研报的图2

2. 最大化IC部分



对应研报的图10

此处就两张图进行说明：

1. 第一张图是做多选出的股票的累计收益。
2. 第二张图是做多按照ic因子排序前35只股票，做空ic因子后35只股票。

先解释图一中，与研报差异较大的原因。

1. 因子的计算方式，我选择的因子是研报中与优矿里面介绍对应的因子，可能不完全一致。
2. 由于数据选择的问题，选出来的股票数可能不一样。在选取月度因子的时候，对因子操作和容

忍度不一样，导致不同的股票选择。

关于图2的解释：

IC的方法研报中并没有介绍如何操作（比如不同因子加权是如何加权的，综合打分是如何打分的）。我这边操作时计算历史6个月IC值最大的因子作为下一个月主要的选股因子，通过动态的选择因子来构建组合。

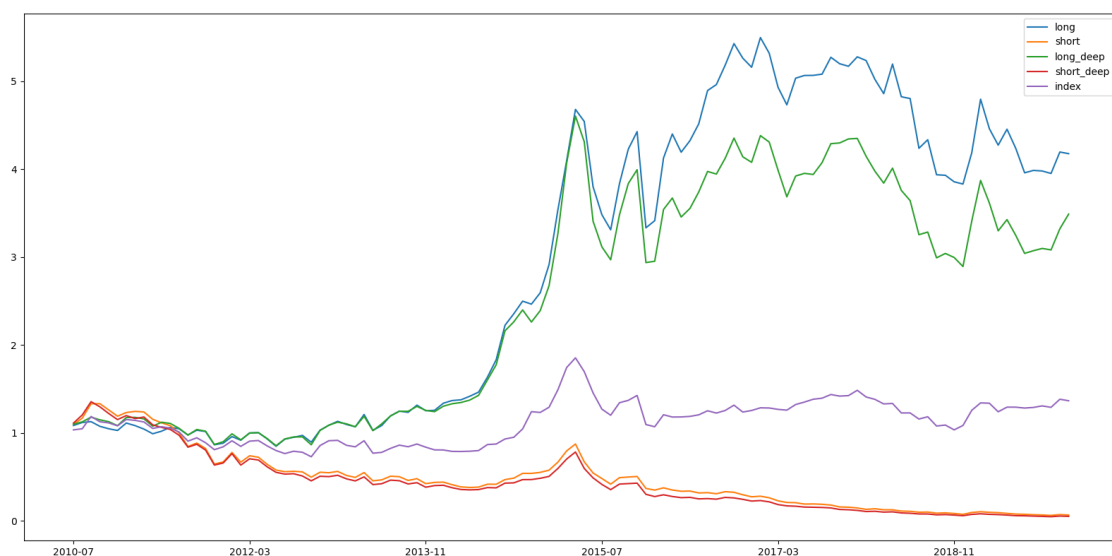
图二与研报的对冲方式不同，研报是对冲的指数。图二是对top-bottom的对冲方法。在做实验的时候发现，针对这一套数据集，使用top-bot的对冲方法是比较合适的。这套数据集中，因子普遍对跌的股票比较有效。因子偏向于选出跌的股票，就使得bottom部分的计算比较准确。top-bottom的方法就可以得到更好的效果。数据的选择与处理没有对错之分，只是之前得到因子数据的代码构成的这套数据比较适合使用top-bottom策略，感觉研报他们自己构建的数据集与计算的因子比较适合做多并对冲掉指数。

后面我们对深度学习的方法依然使用top-bottom的对冲方法，并且，修改了增强的方案原因如下：

1. 因为因子不一样，在计算时发现这套数据集对跌的股票的计算比较准确。deep的方法就会更加偏向于将跌的股票预测准一点，而涨的股票计算没有那么准。
2. 研报中使用基本面因子对算法增强，但是，文中对应的策略一只选择了做多的股票，选不出来做空的股票。再加上，选择最大IC的方法效果要好于基本面的方法。所以后面面对深度学习的增强算法使用的是IC增强的方法。

针对深度学习的结果：

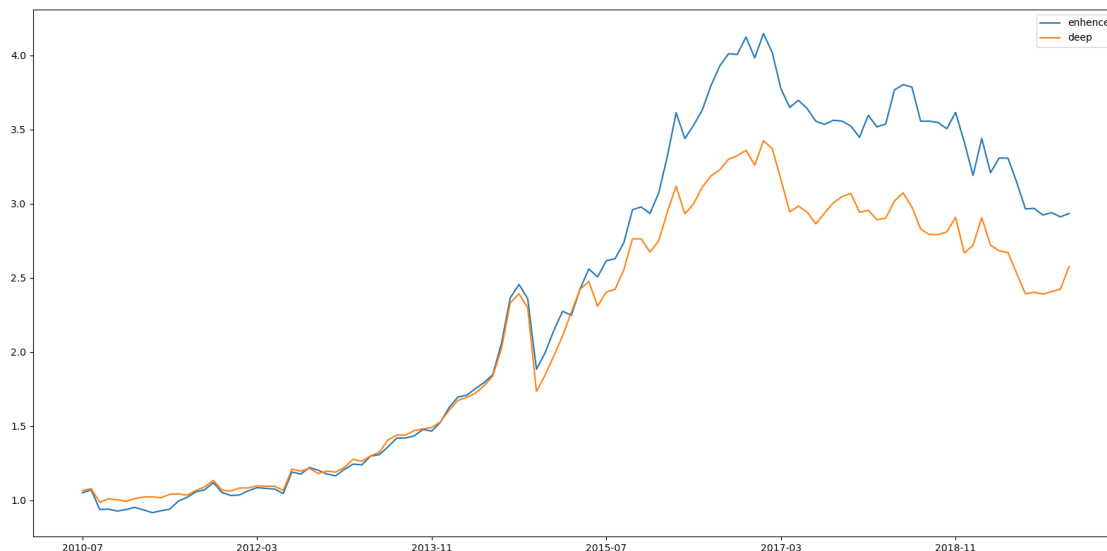
1. 文中没有给出深度学习训练的超参数，我这边使用的研报中完全一样的网络结构。使用了0.5的dropout与0.0001的学习率和10次早停。
2. 使用了5个月训练，一个月验证。滑动训练的方法。
3. 其深度学习增强方法与研报思想类似。先通过IC对股票进行删选，在使用deeplearning的方法进行进一步选择。
4. 选择35只股票的原因是，使用基本面的方法大概只能选出35只股票，研报中是75只。保证结构的统一性，我们一直用35来验证不同方法的效果。



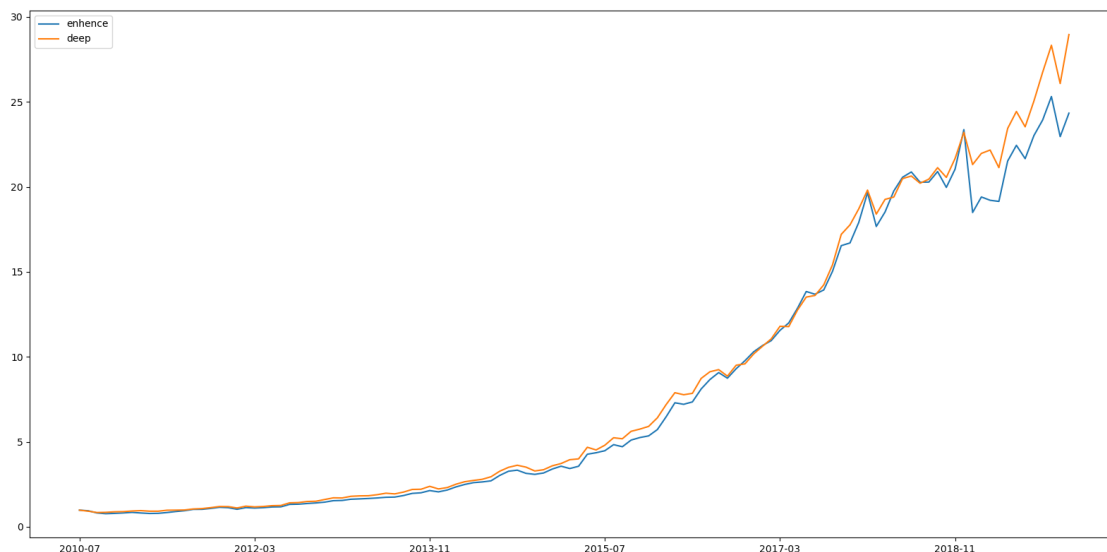
深度学习（有_deep后缀）结果与深度学习加强结果（long, short）如下所示。

其中，long表示做多的组合收益，short表示做空的收益。index表示800指数的收益。可以看到深度学习增强的方法增强了在深度学习方法在做多过程中的收益。但是并没有增强做空的收益。其原因还是因为数据集与研报不一致。在做空这方面，深度的方法已经可以得到较好的结果，加入额外信息反而会减弱其表现。

通过算法得到做多的结果明显是好于指数的。先使用研报的方法，对冲800指数，得到如下结果：



其对冲效果在17年左右就不太好，使用的数据集不一样的到的效果也不一样。其原因主要是后半段时间对跌的股票计算更为准确。我们回测了top-bottom，且双边千分之5手续费的结果。



从上面来看，深度学习的方法是有效的。只是由于数据集的原因导致算法的偏向将跌的预测更准。值得注意的是，通过IC没有对深度学习在top-bottom回测上有结果的加强，但是在做多的股票组合中有加强，说明，IC的信息影响了算法在算跌时的准确度。

总得来说：

1. 深度学习的方法算出来因子打分方案明显好于基本面与实力最大IC的方法。
2. IC的策略在做多的组合中对深度学习算法有加强，但是，在top-bottom构建组合时没有加强。
3. 优矿的因子数据集更适用来做top-bottom策略，而研报构建的因子数据集更适用对冲指数。

代码说明：

s1->s2->s3_deep->s3_eval_deep->s4

其中，s4是主要的评价程序。

代码中出现的month_list变量时，需要修改对应的数据集路径。这个数据集也是之前几个项目对应的数据集。