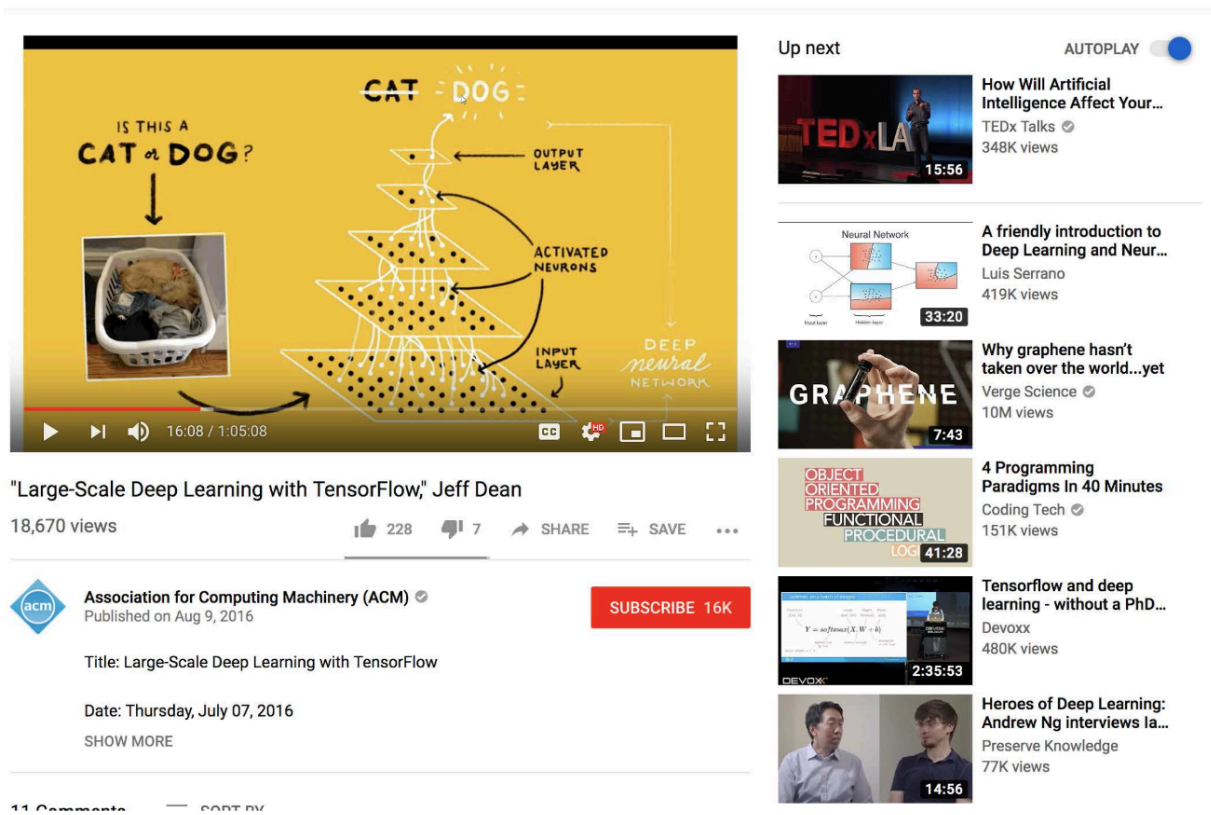


概述

在该论文中，YouTube提出了一种双塔结构的排序模型，新颖的点主要在于：

1. 模型是基于多目标进行训练的；
2. 借鉴广告CTR中的思想，引入了用于消除位置偏差的浅层网络；
3. 引入了MMoE，带门控的专家网络；

需要注意的一点是，该论文的应用场景与YouTube之前那篇经典论文"[Deep Neural Network for YouTube Recommendation System](#)"不同，之前的推荐场景是在首页推荐频道中，而本文的推荐场景是发生在，用户正在看或看完一个视频以后，紧接着YouTube进行推荐。



一、召回

召回部分，文中说得非常笼统，用了一小段话就给带过了，这里只能逐句逐句来进行解析了：

Our video recommendation system uses multiple candidate generation algorithms, each of which captures one aspect of similarity between query video and candidate video.

如果没有留意到业务应用场景的话，可能会对"query video"这个词感到懵逼，毕竟一般的推荐场景中不会有"query"这种东西。

这里的推荐是在用户看某个视频的时候就在旁边给他出一个推荐清单。由于正在观看的视频极有可能是用户当前感兴趣的内容，我们认为与该视频相关的视频也更有可能会被点击观看，所以在推荐时需要考虑"watching video"，也就是"query video"。按这个思路，我们还能想到，其实可以用基于物品的协同过滤算法来作为baseline。

通过不同算法来捕捉"query video"和"candidate video"在不同方面的相似度。具体有哪些方面呢？

For example, one algorithm generates candidates by matching topics of query video.

Another algorithm retrieves candidate videos based on how often the video has been watched together with the query video.

作者举例说了两种，也是比较容易想到的：视频在主题上的相关性；两个视频间共现频率；

简单发散一下，我能想到的还有：

- 两个视频的UP主之间的相关性；
- 两个视频在封面、标题上的相关性；
- 两个视频在时长、观看人数上的相关性；
-

We construct a sequence model similar to [10] for generating personalized candidate given user history.

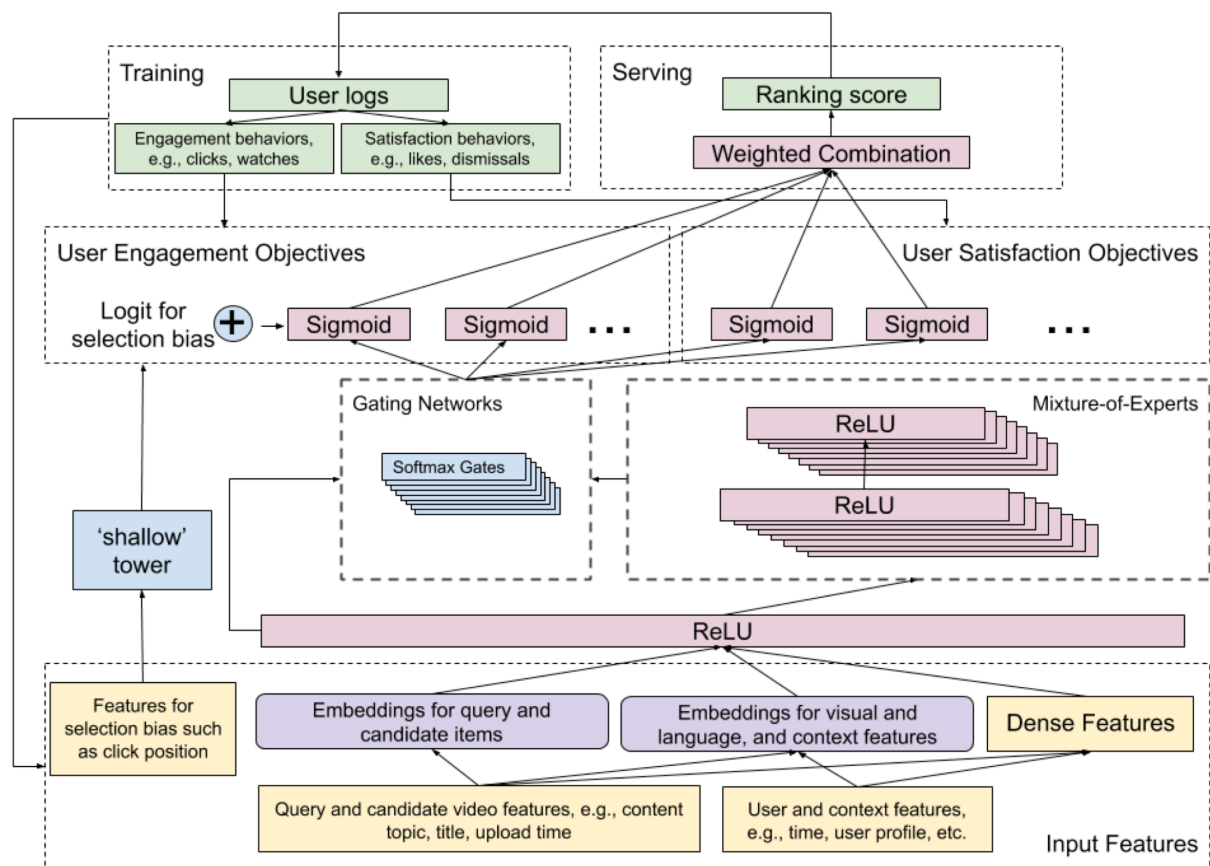
借鉴了论文"Deep Neural Network for YouTube Recommendation System" ([论文地址](#)、[博客地址](#)) 的做法，来基于用户观看历史构建序列模型（类似于Word2Vec），进而生成候选视频。

We also use techniques mentioned in [25] to generate context-aware high recall relevant candidates.

使用了"Efficient Training on Very Large Corpora via Gramian Estimation" ([论文地址](#)) 中的技术来生成与上下文相关的高召回候选集合。

二、排序

一图胜千言，先看一眼论文中所提出的模型框架：



虚线框标出了几个主要模块：

1. Training

虽然标着是Training，但其实是用户日志，标为Data会更合适一些。可以看到，把用户日志分了两类，一类是反映用户参与度的行为，如点击、观看；另一类是反映用户满意度的行为，如like、dismiss。为什么分了两类行为？往下看就看到了它的意图——对多目标进行建模学习，主要分为参与度和满意度两类目标。

2. User Engagement Objectives

3. User Satisfaction Objectives

通常情况下，我们的模型训练目标是点击CTR或视频观看时长，抽象来看，是对用户参与程度的反映。YouTube的该论文认为，除了参与度"engagement objectives"以外，推荐系统还应该考虑用户的满意度，也就是"satisfaction objectives"。满意度可以从用户对所推荐的视频点击like或者dismiss的操作来体现。

从参与度和满意度两类目标的框图中，我们看到了熟悉的"Sigmoid"方块。可以猜想，每个"Sigmoid"方块代表一个预测值，例如用户click的概率、用户like的概率、用户dismiss的概率。

这两个虚线框内，主要的区别在于，用户参与度的部分多了一个标为"Logit for selection bias"的小加号。从字面意思来看，这里是加入了“选择性偏差”到预测值中。什么是选择性偏差？这对了解广告CTR预估的同学来说应该是很熟悉了。

4. Input Features

特征包括三大块：视频特征、用户上下文特征、消偏用特征。

从视频特征中得到"query video"和"candidate video"的embedding；结合用户上下文特征中得到综合的embedding以及一些dense features；消偏用的特征则是通过一个所谓的"shallow tower"去生成前面所说的"Logit for selection bias"。

5. Mixture-of-Experts

6. Gating Networks

MoE 和 Gating Networks 组成了 MMoE，初次见到这种模块，对我来说是挺新鲜的。深入了解过后，感觉其实和CNN中的多channel有点类似。

7. Serving

在这里我们看到了它是如何把多个目标统一起来学习的，就是做加权。具体怎么加权？

The weights are manually tuned to achieve best performance on both user engagements and user satisfactions.

也就是调参调出来的.....

论文中重点展开的是"Selection Bias"和"MMoE"两部分，下面我们分别来看。

1. 模型

1.1 Selection Bias

选择性偏差指的是，我们看到的用户历史行为数据，有受到过往的推荐算法以及诸如位置、设备等环境的影响的。

假设之前的推荐算法特别简单粗暴，只推综艺类视频，并直接拉历史CTR作为打分进行排序、推荐。那么如果不采取特殊手段，我们的推荐模型会被“教坏”，模型的推荐结果严重偏好综艺类视频，且偏好封面党、标题党。

最常见的一种选择性偏差是位置偏差——热搜榜第一条跟第20条，显然你更容易点击第一条。

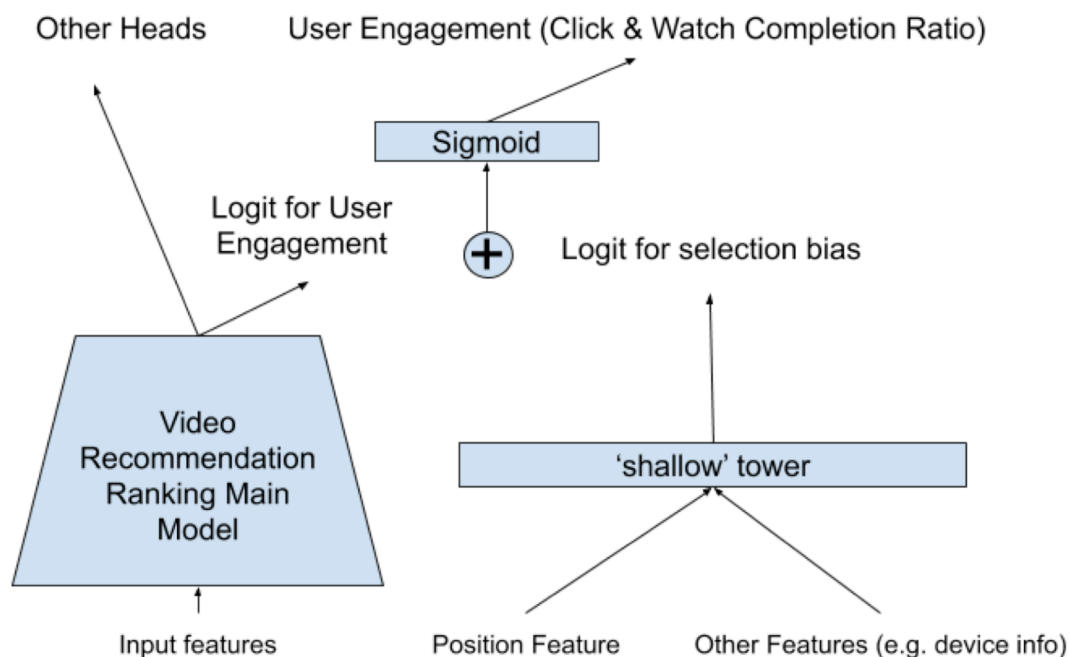


Figure 3: Adding a shallow side tower to learn selection bias (e.g., position bias).

这张图里稍微详细了一点，但依旧没有解释"shallow tower"到底是什么。

作者提到：

Our proposed model architecture is similar to the Wide & Deep model architecture

所以可能"shallow tower"就只是一两层ReLU.....

还提到了一个小细节：

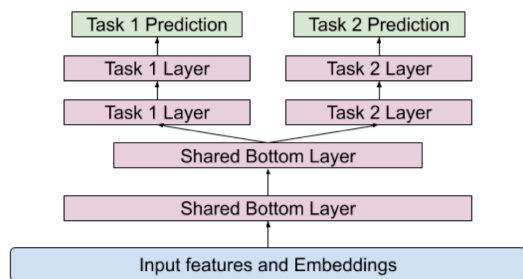
In training, the positions of all impressions are used, with a 10% feature drop-out rate to prevent our model from over-relying on the position feature. At serving time, position feature is treated as missing.

这种training和serving的略有不同，想必也都是工程师们的宝贵经验吧。

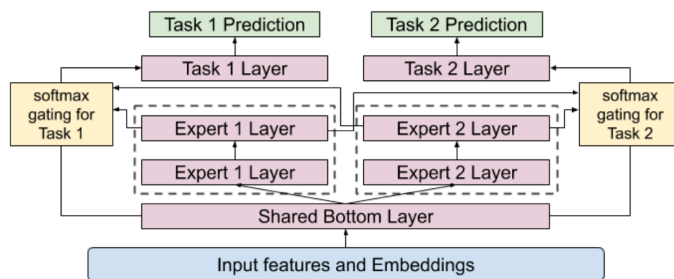
仔细看这幅图，还会有个小发现："Click & Watch Completion Ratio"。点击自然不用说，我觉得有意思的是对于观看行为，这里不是采用预测具体的观看时间（像之前那篇YouTube论文那样），而是采用了视频的“完读率”，即用户看了某个视频中的多大比例。具体实践时，可能还得考虑：如果用户拉了进度条呢？我看了几秒，然后直接拉到50%，再拉到70%，那么整体怎么算？也就是说，在计算比例时，分母是视频时长，分子是什么？

1.2 MMoE: Multi-gate Mixture-of-Experts

对于MMoE模块，文中给出了这么一张对比图：



(a) Shared-Bottom Model with shared bottom hidden layers and separate towers for two tasks.



(b) Multi-gate Mixture-of-Expert Model with one shared bottom layer and separate hidden layers for two tasks.

在对两个任务同时进行学习时，我们可以像左图那样，在底层对Embedding层提取信息，然后把这一个底层"Shared Bottom Layer"同时给到两个任务去用，两个任务分别基于这个"Shared Bottom"去做学习。

这种方案会带来什么问题呢？

However, such hard-parameter sharing techniques sometimes harm the learning of multiple objectives when correlation between tasks is low.

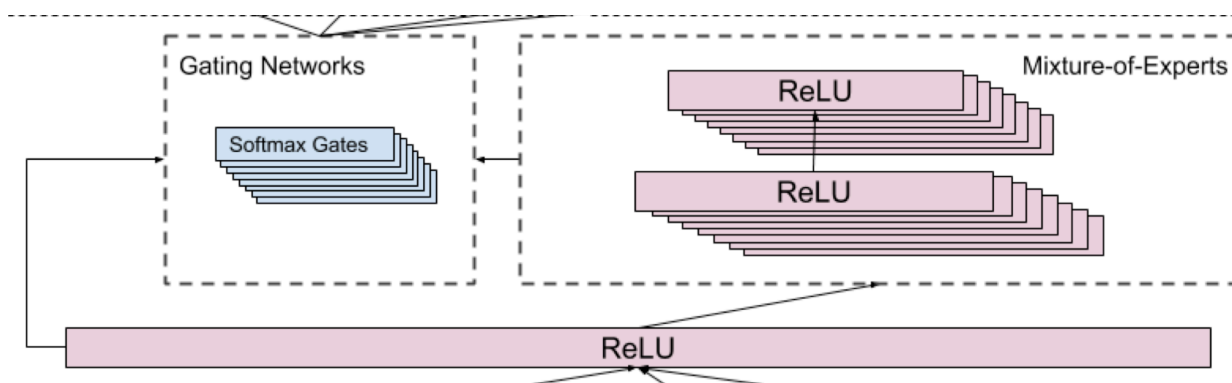
抽象的来理解，假如同时学习的两个任务之间的相关性较弱，那么用"Shared Bottom"强行将两者捆绑在一起，学习过后，我猜测"Shared Bottom"可能会出现：

1. 其中的每个神经元都处于“低活跃”的状态；
2. 一半的神经元只学习任务1而几乎不学习任务2，另一半的神经元只学习任务2而几乎不学习任务1；
3.

当然了，以上只是我个人猜测。后面肯定还是得安排着继续看论文了：

Jiaqi Ma, Zhe Zhao, Xinyang Yi, Jilin Chen, Lichan Hong, and Ed H Chi. 2018. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 1930–1939.

以上右图第一眼看着挺困惑的，但在我细看之后，发现其实和卷积层的多频道的概念有异曲同工之妙；我们也从前面的框架大图里面抠出来看，会更好理解一些。



从底层"Shared Bottom"同时输入给到专家网络和门控网络，门控网络同时还接收自专家网络的输入，然后再统一由门控网络输出。专家网络旨在分配若干个（不一定等于任务数）“专家”进行学习，每个“专家”是由两层ReLU层组成的网络；所有“专家”的最后一层输出，都会给到门控网络中的每个Softmax神经元中，每个Softmax神经元结合Shared Bottom再统一输出。

所以能看到，在这里分多个专家处理，就类似于卷积层中分多个频道处理。

文中采用了相对较少的专家数量：

we use a relatively small number of experts. This is set up to encourage sharing of experts by multiple gating networks and for training efficiency.

2. 实验

2.0 整体实验对比

Model Architecture	Number of Multiplications	Engagement Metric	Satisfaction Metric
Shared-Bottom	3.7M	/	/
Shared-Bottom	6.1M	+0.1%	+ 1.89%
MMoE (4 experts)	3.7M	+0.20%	+ 1.22%
MMoE (8 Experts)	6.1M	+0.45%	+ 3.07%

Table 1: YouTube live experiment results for MMoE.

这里颇为神奇的是，拿运算量来比较.....也就是用来比较，相近模型复杂度的情况下，参与度指标和满意度指标有何变化。（用乘法运算量来作为模型复杂度的参考指标？这操作.....）

在相同模型复杂度之下，MMoE能显著提升参与度指标和满意度指标；

2.1 Selection Bias

对于位置偏差，作者给出了不同位置下的CTR情况：

CTR @ Position vs. Position

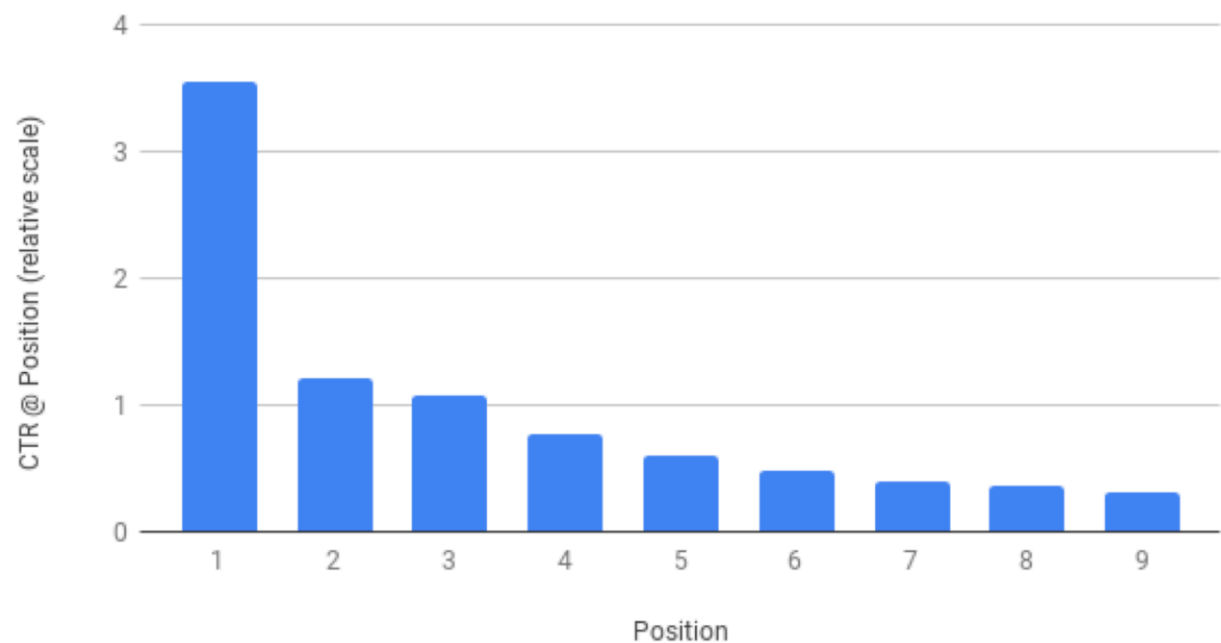


Figure 6: CTR for position 1 to 9.

以及学习得到的位置偏差：

Position bias @ Position vs. Position

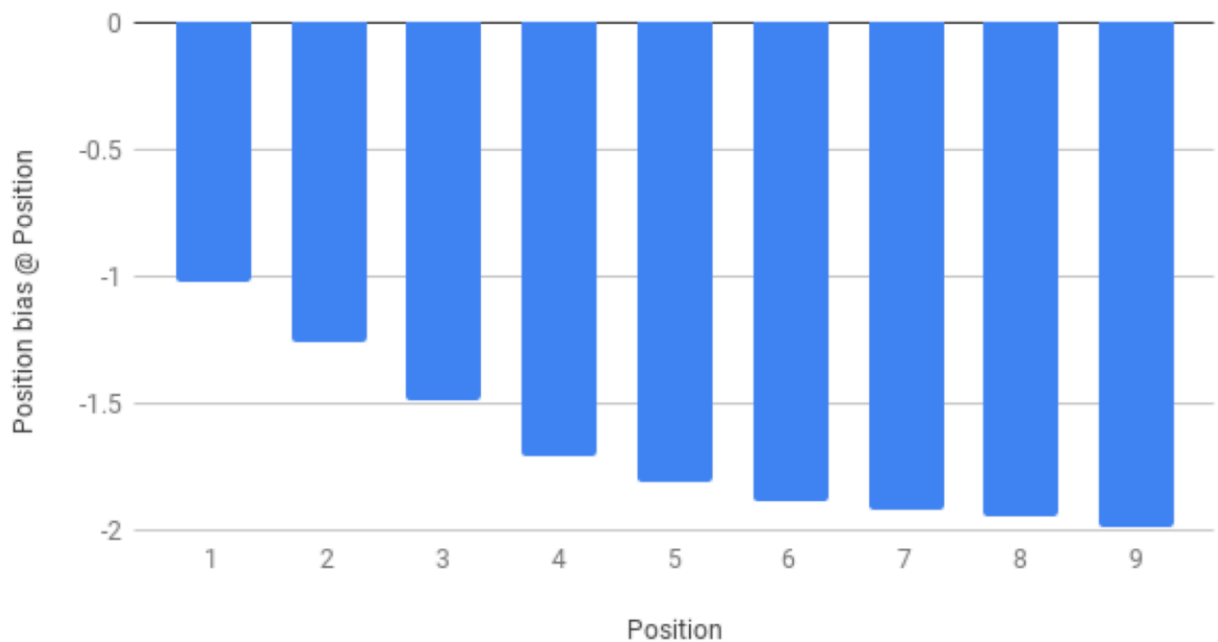
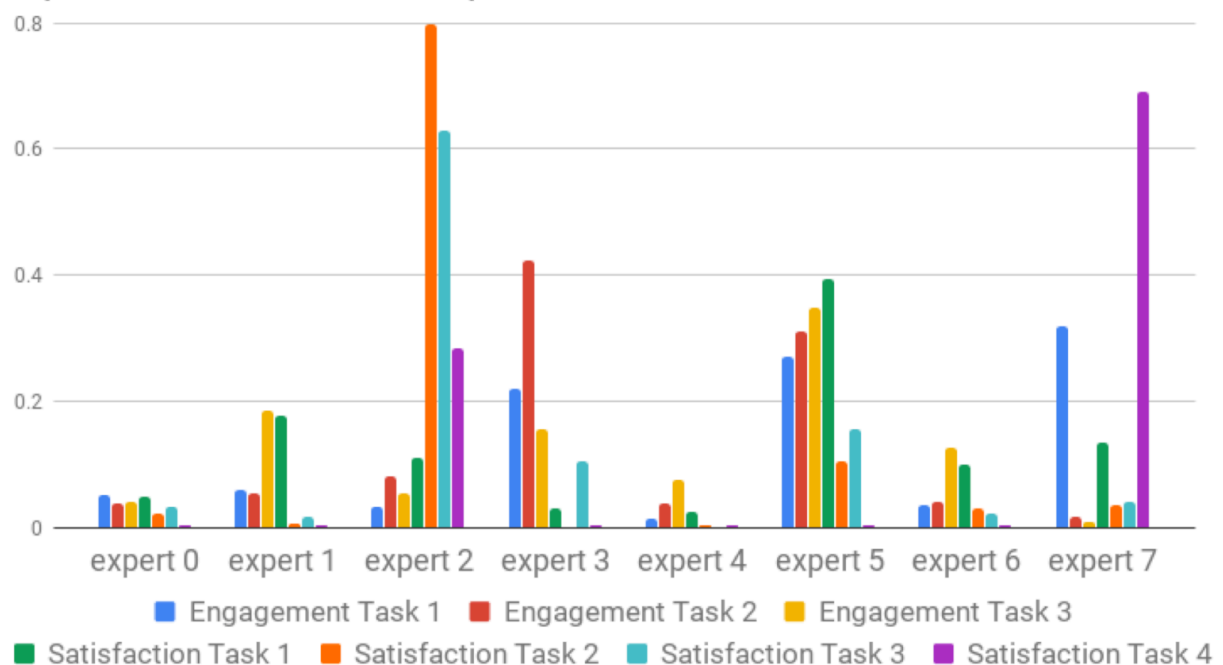


Figure 7: Learned position bias per position.

2.2 MMoE

对于MMoE的分析，作者还给出了下图：

Expert Utilization for Multiple Tasks



To further understand how MMoE helps multi-objective optimization, we plot the accumulated probability in the softmax gating network for each task on each expert

结语

关于这篇论文的解读就先写到这里了。总的来说，论文着重讲的是引入了消偏以及MMoE模块后的整体多目标学习的模型结构；对于排序模型结构以外的东西，如数据处理、特征工程等讲得特别少。很难说换个业务场景是不是也能生效。

后续可能需要看下MMoE的发布论文，以及如果有空的话，补充一下该论文在消偏部分可能遗漏了的一些细节。另外，多目标学习也是颇有意思的业务场景，值得继续深挖。