

083 | 基于深度学习的推荐模型之三：利用深度学习来扩展推荐系统

2018-04-13 洪亮劼

AI技术内参

[进入课程 >](#)



讲述：初明明

时长 06:36 大小 3.03M



这周，我们主要讨论如何利用深度学习来提升推荐系统的精度。我们分别介绍了使用RBM（Restricted Boltzmann Machines，受限波兹曼机）来对推荐系统建模，和RNN（Recurrent Neural Network，递归神经网络）在推荐系统中的应用。

今天，我们最后再来看一看，还有哪些深度学习的思潮在影响着推荐系统。这些思想都是一些值得关注和学习的方向，虽然这些方法能否在短时间内成为主流还有待检验。

多层神经网络

深度学习开始从计算机视觉这一人工智能子领域慢慢向其他领域扩展，在这个大背景下，很多学者和实践者总结出了深度学习获得成功的一大法宝，那就是**能够从众多的信息中学习到高维度的特性**（Feature）。比如，在计算机视觉里，通过多层神经网络建立的深度学习模

型往往能够识别出类似“边”、“角”、“形状”等视觉概念。于是，大家就形成了这样一个普遍认识：**多层神经网络是一种提取特性的利器。**

和很多其他领域的思路一样，在推荐领域的学者也开始尝试利用多层神经网络对用户和物品的信息进行提取。这里面比较经典的思路是：直接利用**多层神经网络**，来尝试我们之前已经讲过的基于矩阵分解的推荐模型。

在矩阵分解的例子中，我们学习到的是用户的隐变量和物品的隐变量。这两者的点积成为评分的预测结果。在这个模型框架里可以认为，用户的隐变量和物品的隐变量是我们从数据中提取的特性信息。只不过，这种特性的提取是一种**线性变换**，而深度学习模型寻求的是**多层次的非线性变换**。

利用多层神经网络提取用户和物品特性的基本思路是这样的。

首先，我们的输入信息是用户的 ID 以及物品的 ID。这里我们可以认为 **ID 是高维的离散输入**。那么，根据这个输入信息，我们分别对用户和物品构建多层神经网络。这里比较常见的是至少有一层的神经网络，可以把离散的输入转换成为**连续的数据层**。

通常我们把这一层叫作“**嵌入层**”（Embedding）。嵌入层的基本思想是希望能够把离散信息转换成为连续信息，并且这些连续信息携带着“语义”（Semantic）上面的相似。什么意思呢？就是说如果两个用户是相似的，或者两个物品是相似的，那么我们会寄希望于他们的嵌入层在数值上相近，这样嵌入层就能够很好地捕捉到用户和物品的相似度。

当用户和物品两边分别得到了嵌入层以后，输入信息已经完全变成了连续数据信息，通常的做法是把两边的嵌入层拼接在一起（也有把两个嵌入层取加权平均的做法），形成一个大的嵌入层。这个新的拼接层就成为连结了用户信息和物品信息的输入。

然后，接下来要做的，就是对这个新的输入进行多层的神经网络变换，最后输出我们对评分的预测。

总而言之，**直接利用多层神经网络对用户和物品的建模可以简单归纳为两步**：首先，把离散的 ID 信息转换成为连续的信息，形成嵌入层；然后，利用多层神经网络对嵌入层进行变换，并最后输出预测结果。

虽然利用多层神经网络在推荐系统上可以算是非常直观，然而在实际的效果中，我们其实并没有发现利用多层神经网络比矩阵分解要更好。如何能够更加有效地利用多层神经网络目前还是一个研究方向。

其他深度学习模型

除了直接利用多层神经网络来对推荐系统进行建模以外，最近也有不少把深度学习中其他思潮应用到推荐系统中的尝试。这里我也简单提及一些。

首先，在深度学习领域有一个技术叫作“**自动编码器**”（Autoencoder）。这个方法的核心思想就是希望能够无监督化地学习到某种信号的编码后的信息，并且编码后的信息能够再次通过解码尽可能地还原最初的信号。如果我们能够找到这样完美的过程，那么，在这个方法的假设里，我们就获取了更有价值的特性信息。

于是，把自动编码器思想应用到推荐系统的尝试，主要就是对用户以及物品进行编码。需要指出的是，我们刚才提到的直接使用多层神经网络的方法，其实也是某种意义上的编码，但我们这里提到的自动编码器的思想，还强调能够从学习到的隐含信息中还原最初的信息，也就是尽可能保持学习到的隐含变量的**可还原性**。这一点是多层神经网络所不具备的。

另外一个就是 **CNN（卷积神经网络）** 的应用，学者和工程人员希望借助 CNN 来对用户和物品进行建模，从而扩展推荐系统的表现力。目前 CNN 在计算机视觉领域占据着举足轻重的地位，因此，很多人都希望能够直接把 CNN 的成功借鉴到推荐系统中。

和刚才我们介绍的直接使用多层神经网络建模的思路非常类似。一种常见的做法也是**直接利用 CNN 来提取用户和物品的信息**。通常的做法就是在我们刚才所说的嵌入层之后，不直接使用多层神经网络了，而是使用 CNN 来对嵌入层之后的信息进行变换。这里其实并没有太多创新的地方，基本上就是把视觉里面的模型拿来直接使用。

关于其他深度学习模型我们就说到这里，不做太多的介绍。原因是从总体上来看，深度学习在推荐系统上的应用，还是技术大于实际效果，目前并没有在太多的应用上有真正令人信服的成功案例。大多数情况是技术人员希望尝鲜，然后把一些成功的模型照搬到推荐任务上，虽然效果也不差，但从大的贡献来说，并没有太大的创新。

小结

今天我们继续讨论了如何利用深度学习模型来对推荐系统进行建模，我们聊了如何把深度学习的思路嫁接到推荐系统的开发中来。

一起来回顾下要点：第一，我们聊了一种最简单的应用，就是直接利用多层神经网络来对推荐系统进行扩展；第二，我们聊了一些其他的深度学习的思路，包括自动编码器和 CNN，简单讨论了如何把这些深度学习模型应用到推荐系统中。

最后，给你留一个思考题，深度学习模型在推荐领域应用的最大问题，或者说最大挑战是什么？

欢迎你给我留言，和我一起讨论。

推荐阅读

[深度学习 | 空竹里的秘密：自编码器](#)

 极客时间

AI 技术内参

你的360度人工智能信息助理

洪亮劼
Etsy 数据科学主管
前雅虎研究院资深科学家



新版升级：点击「 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 082 | 基于深度学习的推荐模型之二：基于RNN的推荐系统

一 篇 内 容 概 要

精选留言 (2)

写留言



永夜

2018-04-18

1

推荐的时候有一个很重要的影响效果的点是“用户可接受的推荐理由”，例如：你的哪个朋友喜欢，根据你的历史推荐，喜欢买某东西的还买什么，当前热门购买是什么，用户能理解的他就更买账一些，深度学习出来的推荐结果未必能给出理由来吧？



林彦

2018-04-13

1

深度学习在推荐系统的应用我的理解一个是开始数据不多(冷启动状况)时的问题。

还有一种可能是这些数据特征(包括嵌入的)的学习方法，可解释性。新特征的学习方法较多依赖于语义分析。是否这些输入特征已包含足够多的信息可以用来作推荐？和旧模型如何合理地比较提升水平？深度学习在推荐系统如何使用更原始的特征而不借助于其他模型...

展开