

# 可解释推荐系统：身怀绝技，一招击中用户心理

原创 2017-09-14 王希廷、谢幸 微软研究院AI头条

▲ 点击蓝字关注微软研究院AI头条



Microsoft Research  
微软亚洲研究院

编者按：推荐系统被广泛应用在电商网站、社交网络、生活服务网站、搜索引擎等平台。一个好的推荐系统不仅需要准确地把握用户的需求，推荐给用户想要的内容，更需要学会与用户沟通，了解用户心理，以用户容易接受的方式给出适当的推荐。

本文中，微软亚洲研究院的研究员王希廷和谢幸将为大家详解可解释推荐系统的分类、推荐解释生成方法以及面临的机遇和挑战。

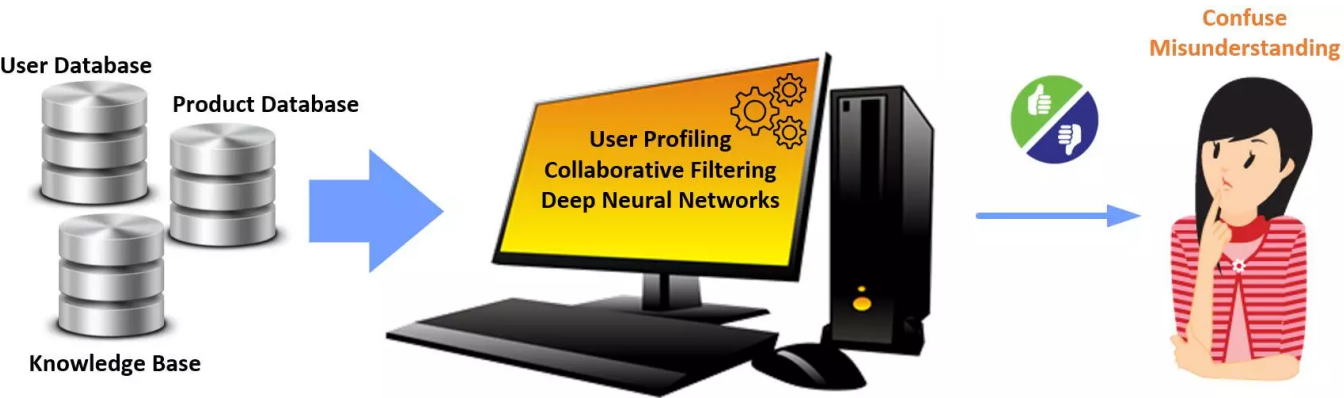
《战国策》中有个有趣的故事：当年赵国被入侵，只有送出质子才能求来救兵。但是赵太后特别疼爱小儿子，怎么也不愿意让他当质子。大臣们越是劝谏，赵太后就越是愤怒。这时候触龙出面，成功说服了赵太后，让她的爱子出质于齐，换来救兵。他到底是如何做到的呢？

这个故事其实跟我们最近进行的推荐系统的研究非常相关。如果把“送出质子”当作一个推荐给太后的策略，那么故事的关键不在于推荐内容的好坏，而在于对推荐内容的解释。如果能让赵太后这样顽固的用户接受这个推荐，关键点是什么？

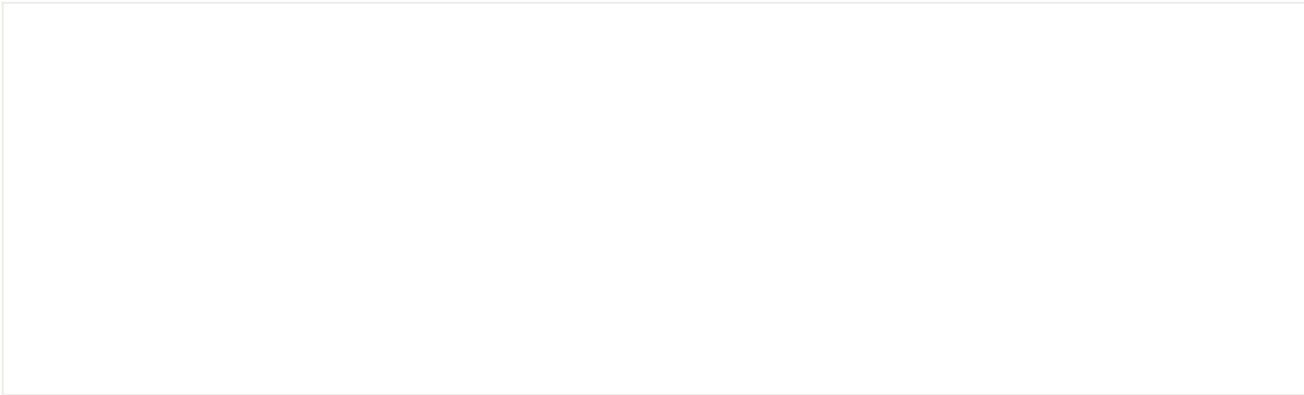
触龙给出的答案是：**沟通方式、用户心理、合适例证**。让我们详细看看他是怎么做的。

触龙刚见到赵太后就开始话家常，说：“我老啦，有个小儿子年龄很小，真是担心他的未来啊。”赵太后问：“你们男人也很宠爱小儿子吗？”触龙说：“比女人更宠呀。我认为赵太后更喜欢您的女儿，而不是小儿子。”赵太后表示不认同。触龙解释道：“您送公主出嫁的时候，尽管伤心，但也希望她不要被赶回来。这还不是为她的长远未来打算？”赵太后同意。触龙又说：“但是您却没怎么为长安君计划未来。您看赵国刚建立时的诸侯王，他们的子孙还有享受爵位的吗？其他国家的呢？”赵太后想了想，表示没有。触龙又说：“那是因为他们荣华富贵太甚，又没有什么功劳的缘故。您现在让长安君有如此高的地位，却不让他有功于国，一旦您驾崩了，长安君凭什么在赵国立足呢？”听了这番话后，赵太后同意了将长安君作为人质。

我们可以在这个故事里看到两类推荐系统。一类是朝廷大臣这种**传统的推荐系统**（下图(a)），将重心放在提高推荐准确性上，与推荐对象的沟通考虑得不够；而另一类是触龙这种**可解释的推荐系统**（下图(b)），能够以用户容易接受的方式（话家常），充分抓住用户心理（父母爱自己的小孩，应该为他们长远未来考虑），给出适当的例子（赵太后送公主出嫁）与用户沟通。研究发现，这样的系统不仅能够提升系统透明度，还能够提高用户对系统的信任和接受程度、用户选择体验推荐产品的概率以及用户满意程度等等。



(a) 传统推荐系统

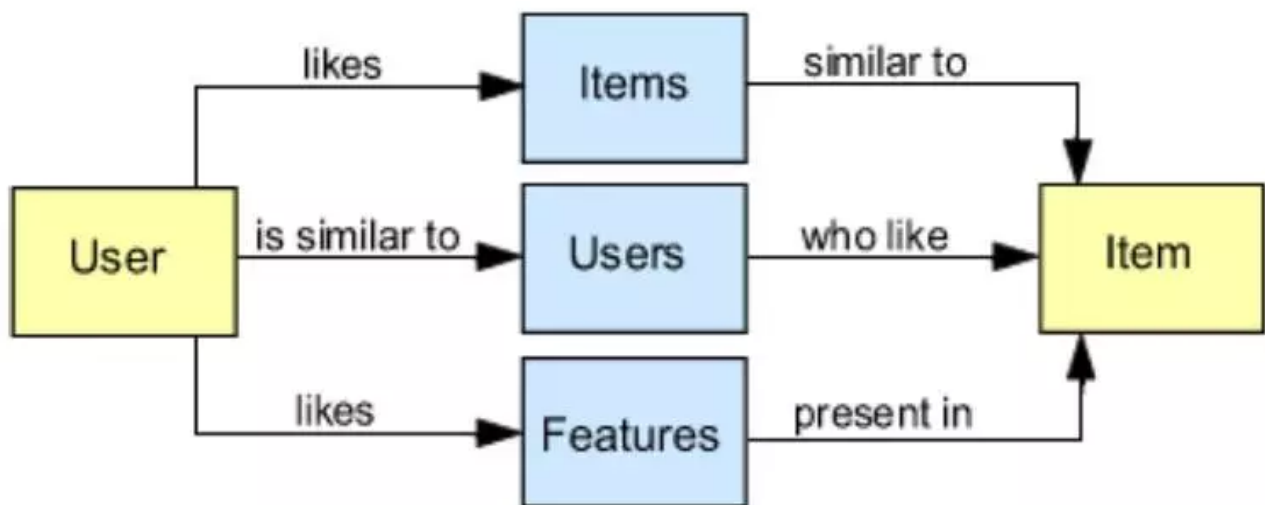


(b) 可解释推荐系统

设计一个像触龙这样可解释的推荐系统是我们的终极目标。目前，这个方向进展如何呢？未来，有什么比较好的研究课题呢？下面，我来介绍一下可解释推荐的分类、推荐解释生成方法以及可解释推荐面临的机遇和挑战。

可解释推荐的分类

一般来说，可解释的推荐分为下图所示的三大类：**以物品为媒介**、**以用户为媒介**、以及**以特征为媒介**。



三大类推荐解释：以物品、用户、特征为媒介

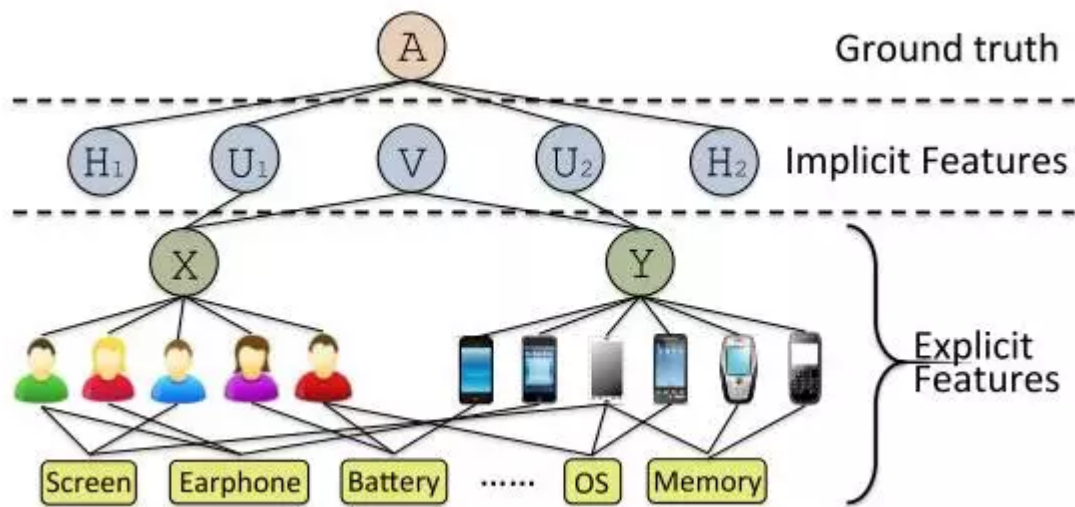
(图片来自文章：Tags planations: explaining recommendations using tags)

以**物品为媒介**的推荐解释用一句话表述是“**这个推荐物品和您喜欢的其他物品相似**”。在赵太后的故事中，触龙用了类似的推荐解释：“派长安君做人质”这个推荐策略与赵太后曾经采用过的策略（送公主出嫁）相似。这类解释在各大网站上如亚马逊、Netflix等有较为广泛的应用。它的展现方式常常是一个曾经购买、评论或者浏览物品的列表，列表中的物品和目前推荐的物品比较相似。研究表明这类方法能够增加系统的被接收程度。但是这类推荐解释的问题是用户有时可能难以找到物品之间的关联，另外用户有时可能希望看到一些自己没有购买过的、不一样的物品。

以**用户为媒介**的推荐解释用一句话表述是“**跟您相似的用户都喜欢该推荐物品**”。在赵太后的故事中，这个与赵太后相似的用户就是也宠爱小儿子的触龙。这类推荐解释展现形式多种多样，生成解释时也往往用到社交信息。微软研究员Amit Sharma等人对四种社交理由进行了用户测试，这四类理由是整体热度（2,612,211位Facebook 用户喜欢它）、朋友热度（您的朋友中有7人喜欢它）、好朋友（张三喜欢它）、好朋友及个数（张三和其他5位您的朋友喜欢它）。最终发现后两种解释方式效果对提升说服力效果最好。使用这两种解释方式时，选择一个正确的朋友非常重要。如果只是随机选择一个朋友，说服力反而会降低。以用户为媒介的推荐解释与以物品、特征为媒介的推荐解释相比，最能说服用户对推荐物品进行评估，但是对用户实际使用物品后喜爱程度影响比较小。如果推荐的物品用户不够喜欢，反而会降低用户的满意程度。

以**特征为媒介**的推荐解释用一句话表述是“**您可能喜欢推荐物品的这些特征**”。在赵太后的故事中，触龙抓住了赵太后最关心的特征——子孙的长远利益——进行推荐，从而使得赵太后接受了推荐。这类推荐解释常见的表现形式是将推荐物品中用户感兴趣的主要特征进行列举（如下图）。研究表明这类推荐解释有助于用户准确判断是否喜欢推荐物品，与前两类推荐解释相比，用户对这类推荐的满意度更高。这类推荐需要判断物品特征以及用户对不同特征的感兴趣程度，从而找到最适合用于解释的特征，因此建模粒度更细，也是目前的研究热点。





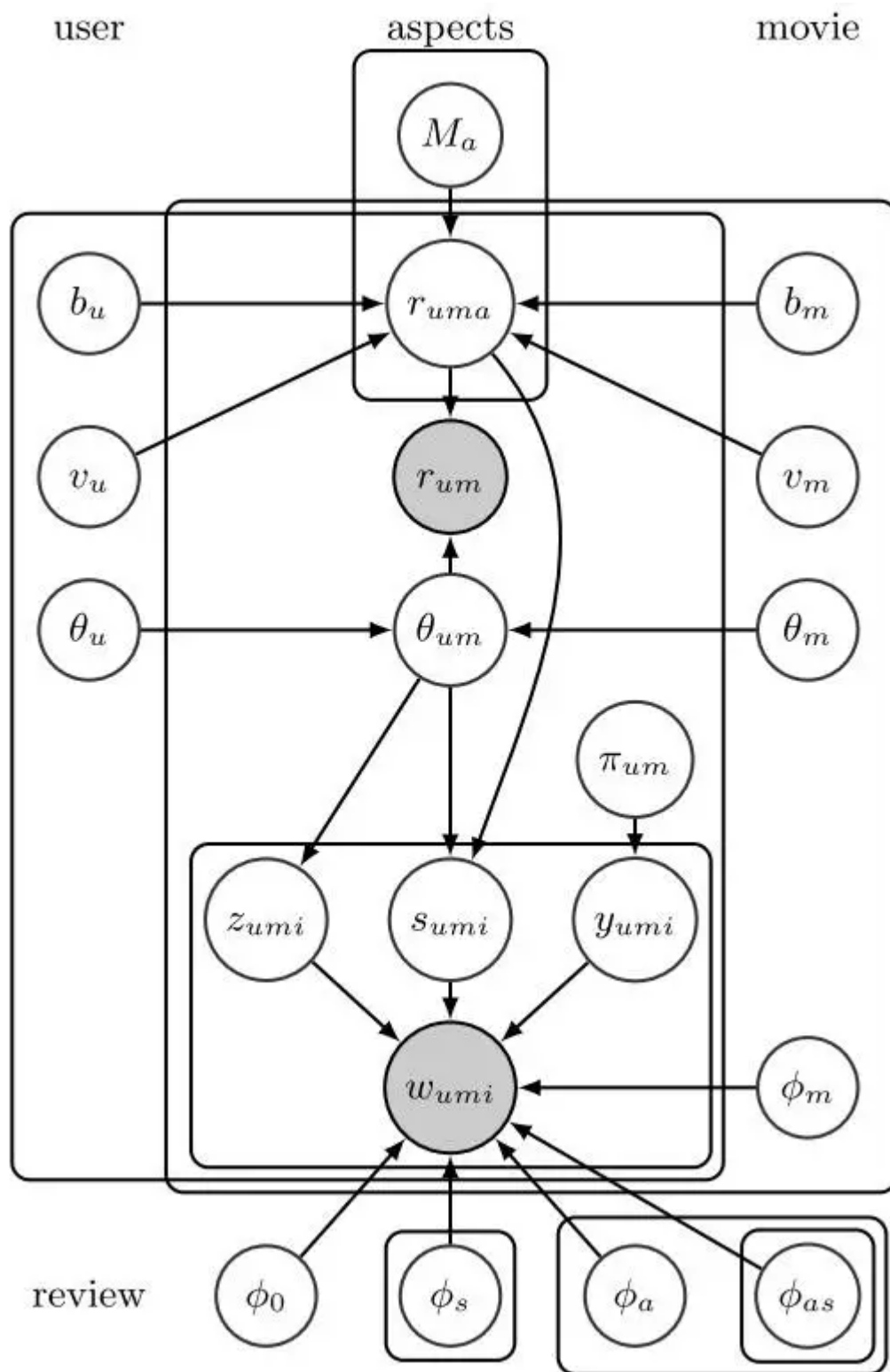
EFM (Explicit Factor Model) 框架

(图片来自文章: Explicit factor models for explainable recommendation based on phrase-level sentiment analysis)

为了解决上面的问题，近期一些工作开始利用网上丰富的评论文本进行可解释性的推荐。他们的主要思想是从评论中抽取用户最感兴趣的物品特征，并且将这些特征加入到目前的推荐系统中进行统一建模。其中一个较新的方法是Explicit Factor Model (EFM)。如上图所示，这个方法首先利用词级别的情感分析找出用户评论中提到物品的显式特征 (Explicit Features, 如Screen、Earphone、Battery) 及用户对这些特征的情感倾向 (正面或者负面)。利用这个信息，可以构建用户-特征矩阵 $X$ 以及物品-特征矩阵 $Y$ 。这两个稀疏矩阵可以用三个低维矩阵 $U_1$ 、 $U_2$ 、 $V$ 进行建模。这里 $V$ 中每一行记录了一个显式特征的隐式表达 (Latent Representations)。 $U_1$  ( $U_2$ ) 即记录用户 (物品) 与显式特征隐式表达之间的关系。除了显式特征以外，还有一些隐含特征 (Hidden Factors) 可能是评论中没有提到的，这些特征用额外的 $H_1$ 、 $H_2$ 矩阵表示。将 $U_1$ 、 $H_1$ 进行拼接就构成了用户与物品所有特征的隐式表达之间的关系， $U_2$ 、 $H_2$ 也是类似的。该模型希望能够通过这些矩阵拟合用户-物品的评价矩阵 $A$ 。

EFM在可解释的推荐系统上迈进了一大步，但是他们的方法将显式特征的提取和用户评价建模分为独立的两步，可能存在误差累积等问题。而JMARS模型能够对用户评分、评价以及物品内容进行统一建模，可以有效解决上述问题。该方法的概率图模型如下所示。这里面可观测量 $r_{um}$ 代表用户 $u$ 对电影 $m$ 的评分， $w_{umi}$ 代表用户 $u$ 对电影 $m$ 的评论中的第 $i$ 个词。可以看到，模型中，代表用户兴趣的隐变量 $b_u$ 、 $v_u$ 和 $\theta_u$ 以及代表电影不同特征的隐变量 $b_m$ 、 $v_m$ 和 $\theta_m$ 同时决定了评分与评论中的词，保证了评分与评论中词的一致性。这个工作还将评论中的词分成多类进行建模，包括在各个评论中都经常出现的Background Word ( $\phi_0$ ) 以及它对应的情感词 ( $\phi_s$ )、电影各个方面特征的Aspect-specific Word ( $\phi_a$ ) 及其对应的情感词 ( $\phi_{as}$ )。除了这些之外，还有每个电影中出现的Movie-specific Word ( $\phi_m$ )，即电影的标题等。这些不同类的词都可以直接由这一个图模型学习得到。





JMARS (Jointly Modeling Aspects, Ratings and Sentiments) 框架

(图片来自文章: Jointly modeling aspects, ratings and sentiments for movie recommendation (JMARS))

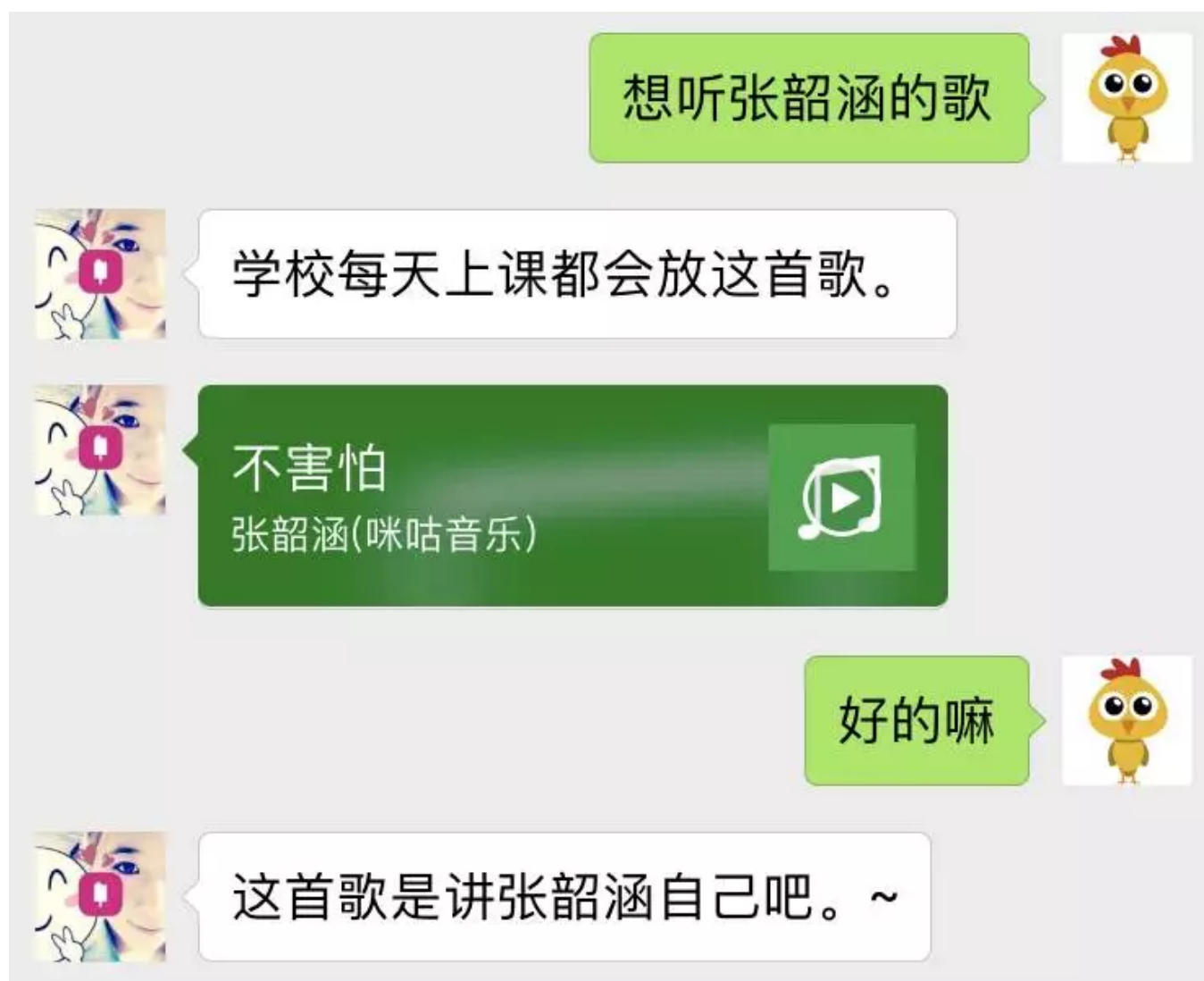
## 可解释推荐的机遇与挑战

作为推荐领域被探索得较少的一个方向，可解释推荐的很多方面值得研究与探索。目前，我们在考虑从下面三个方面进行研究。

**解释含有深度神经网络的推荐模型。**目前可解释推荐的系统往往是基于较浅层的图模型或者矩阵分解模型，对于含有深度神经网络的复杂、混合模型的解释能力还不够。我们在考虑借鉴对深层卷积神经网络进行解释、分析的技术，将复杂、高级的推荐系统升级为可解释的推荐。

**利用知识图谱打通不同媒介之间的关联。**现有的可解释推荐所生成的推荐解释往往只局限于以物品为媒介、以用户为媒介或者以特征为媒介中的某一种，对这三类媒介之间的关联挖掘得还不够。我们希望能够利用知识图谱，打通这三类媒介之间的关联，根据具体情况灵活选择其中最合适的媒介对用户进行推荐与解释。这样，我们还可以产生形式更丰富的推荐解释，例如将用户感兴趣的图片也进行展示。

**结合生成模型进行对话式推荐。**目前的推荐解释往往形式是预先设定、千篇一律的（如预先设定推荐解释是以用户为媒介的）。这样尽管也能根据用户心理举出一些例证，但是在沟通方式上还过于呆板，离触龙那样话家常式的推荐还有很大距离。如果能用生成模型让推荐系统“自创”一句通顺甚至高情商的话，就可以在与用户聊天的过程中进行灵活、多变地推荐解释了。我们团队已经与微软小冰合作，在这方面进行了一些尝试，为小冰生成音乐推荐解释。推荐过程如下图所示：



微软小冰的对话式音乐推荐

我们会从这几个方向出发，不断拓展可解释推荐的深度和广度。希望不久后的某一天，推荐系统都能像触龙那样厉害。

---

## 作者介绍

---



王希廷，微软亚洲研究院副研究员。2011年于清华大学获得工学学士学位。2017年于清华大学获得工学博士学位。她的研究成果发表在数据挖掘和可视化的顶级会议和期刊上，包括KDD、TKDE、AAAI、IJCAI、VAST和TVCG等。她的一篇一作论文被TVCG选为2016年12月的spotlight article。她曾担任TKDE、TVCG、InfoVis等顶级会议、期刊的审稿人。







谢幸博士于2001年7月加入微软亚洲研究院，现任社会计算组资深研究员，并任中国科技大学兼职博士生导师。他分别于1996年和2001年在中国科技大学获得计算机软件专业学士和博士学位。他在国际会议和学术期刊上发表了200余篇学术论文，并拥有50余项专利。他是ACM、IEEE高级会员和计算机学会杰出会员。他参与创立了ACM SIGSPATIAL中国分会，并曾担任ACM UbiComp 2011大会程序委员会共同主席。

#### 你也许还想看：

- 现实版“读心术”，读懂你性格的个性化推荐
- 讲堂I 谢幸：用户画像、性格分析与聊天机器人
- 演讲I 谢幸：如何让机器拥有像人一样的思维？

感谢你关注“微软研究院AI头条”，我们期待你的留言和投稿，共建交流平台。来稿请寄：  
msraai@microsoft.com。

微软

小冰进驻微软研究院微信啦！快去主页和她聊聊天吧。

# Microsoft Research

微软最前沿的科技信息



长按指纹  
识别二维码，  
一键关注黑科技！