

# 面向微博用户的消费意图识别算法

贾云龙<sup>1</sup> 韩东红<sup>1,†</sup> 林海原<sup>1</sup> 王国仁<sup>2</sup> 夏利<sup>1</sup>

1. 东北大学计算机科学与工程学院, 沈阳 110819; 2. 北京理工大学计算机学院, 北京 100081

+通信作者, Email:handonghong@cse.neu.edu.cn

**摘要** 通过分析微博用户发布的大量带有购买意愿的文本数据, 可以识别其消费意图并挖掘用户的需求, 在众多应用领域都有重要意义。首先利用迁移学习的方法, 融合京东问答平台数据与少量已标注的微博数据以构建训练集, 提出一种基于注意力机制的双向长短期记忆神经网络(Attentional-Bi-LSTM)模型以识别用户的隐性消费意图; 针对显性意图识别问题, 提出一种结合 TF-IDF (Term Frequency-Inverse Document Frequency) 与句法分析中的动宾关系 (VOB) 的消费意图对象提取算法。实验结果表明, 通过迁移京东问答平台的数据与微博数据相融合可以有效扩充训练集, 在此基础上训练的神经网络分类模型具有较高的准确率和召回率; 融合 VOB 和 TF-IDF 的显性消费意图对象提取方法在准确率上达到了 78.8%。

**关键词** 消费意图识别; 意图对象提取; 迁移学习; 注意力机制

## Consumption Intent Recognition Algorithms for Weibo Users

Jia Yunlong<sup>1</sup> Han Donghong<sup>1,†</sup> Lin Haiyuan<sup>1</sup> Wang Guoren<sup>2</sup> Xia Li<sup>1</sup>

1. School of Computer Science and Engineering, Northeastern University, Shenyang 110819; 2. College of Computer, Beijing Institute of Technology, Beijing 100081

†Corresponding Author, E-mail: handonghong@cse.neu.edu.cn

**Abstract** By analyzing a large number of text data with purchasing intention published by Weibo users, we can identify their consumption intent and search their needs, which is great significant in many application fields. Firstly, the data set is constructed by the data of Jingdong Question Answer Platform and Weibo based on transfer learning method and a bi-directional long-term and short-term memory neural network model based on attention mechanism is proposed to identify users' implicit consumption intention. For the problem of explicit intention recognition, a new algorithm for extracting consumer intention objects is proposed, which combines TF-IDF (Term Frequency-Inverse Document Frequency) with the verb-object relationship (VOB) in parsing. The experimental results show that the training set can be effectively expanded by merging the data of Jingdong Question Answer Platform and Weibo. And then, the classification model has high accuracy and recall rate. The method of extracting explicit consumer intent objects by fusing VOB and TF-IDF achieves 78.8% accuracy.

**Keywords** Consumption Intention Detection; Intention Object Extraction; Transfer Learning; Attention Mechanism

微博作为一种新的社交网络平台, 人们能够方便快捷地参与热点事件讨论、交流和表达情感。由此产生了海量社交网络数据, 其中部分数据蕴含了用户对某种商品的购买愿望, 即消费意图。消费意图分析是指挖掘用户通过文本或行为表达出的对某一产品或服务产生的购买意愿, 是针对社交网络用户的消费行为进行识别分析的研究任务<sup>[1]</sup>, 可以应用于诸如广告推荐、市场营销等领域, 具有重大的商业应用潜力。

用户消费意图分析研究主要包括识别显性消费意图和隐性消费意图, 其中隐性消费意图识别、显性消费意图中的消费对象抽取等研究更具有挑战性, 具体情况见表 1。文本是承载情感及意愿的最好载体, 而已有用户隐性消费意图分析大多基于用户行为特征展开研究, 未考虑用户发布的文本信息内容从而影响分析结果的准确率。

从计算语言学的角度出发, 针对社交网络用户消费意图挖掘的研究才刚刚起步, 特别是对隐性消费意图的研究成果鲜有问世。鉴于已有算法的不足, 本文侧重于研究分析隐性消费意图及抽取显性消费意图中的消费对象, 主要贡献包括: (1) 面向微博用户的隐性消费意图

识别研究缺少足够的标注语料库，针对这一问题我们利用迁移学习方法，将源域京东商城的问答数据迁移到目标域微博训练集中，然后提出一种基于 Attention 机制的 Bi-LSTM 模型预测微博用户的隐性消费意图，实验结果表明该模型具有较好的性能；(2) 我们提出了一种基于 TF-IDF 与句法分析的 TF-IDF-VOB 显性消费意图对象提取方法。

表 1 隐性消费意图与显性消费意图中的消费对象

Table 1 The consumption objects in the implicit and explicit consumption intentions

	隐性消费意图	显性消费意图中的消费对象
问题定义	文中蕴含或可以从中推断出的潜在购买意图，	以文本形式明确表达出的购买意图对象
例子	我不喜欢用有绳吸尘器	想求购一款触屏性能好的手机，求推荐
特点	无明确触发词	有明确触发词，如“求购”
面临挑战	大多基于用户行为特征展开研究，未考虑文本内容；且缺少足够的标注语料库	需要基于搜索引擎且需要联网使用，无法离线提取消费意图对象

## 1 相关工作

意图的概念已经在哲学和心理学层面研究过，但其研究人员通常不关心用来表达意图的语言，或者如何从书面语言中通过计算推断出意图。文献[2]首先从计算机语言学的角度定义了“购买意图”（显性消费意图），提出了基于二元图方法的自动抽取消费意图模板，同时加入文本作为消费意图识别的特征，从而提高了分类的准确率。Kroll 等<sup>[3]</sup>给出了“意图分析”的定义，认为意图分析和情感分析在一定程度上是正交的，例如推文“我很喜欢这款笔记本电脑”，同时表达出用户的正面情感和购买笔记本电脑的意图。他们定义了如果一条推文具有商业意图，则该推文需要至少包含一个动词即触发词。Wang 等人<sup>[4]</sup>提出具有显性消费意图的推文除了触发词外，还定义了意图客体即购买对象。陈浩辰<sup>[5]</sup>提出了基于深度学习的用户消费意图预测模型，并且提出了消费意图向消费行为的转化。Ding 等人<sup>[6]</sup>提出了基于领域自适应卷积神经网络的微博文本消费意图识别方法，并将此方法应用到了电影票房预测的任务上。Liu<sup>[7]</sup>等人用基于 Attention 机制的循环神经网络结构(RNN)进行意图识别，即 RNN 是基于序列的，句子中的每个词在 RNN 中都有隐藏状态，利用这些隐藏状态生成最后的意图类别。Ding<sup>[8]</sup>等人认为消费意图识别的任务具有领域相关的特性，为此构建了一个跨领域的基于深度学习的消费意图识别模型，其充分利用了深度神经网络和双样本检验的优势，使用基于树核的最大平均差异(TK-MMD)来提升模型的学习能力。文献[9]针对用户消费意图分类问题通过一系列实验比较了支持向量机(SVM)、朴素贝叶斯(NB)、长短期记忆网络(LSTM)等模型在意图分类上的效果，文中采用文献[10]提出的意图分类方法，将消费意图分成 6 类即 Food&Drink、Travel、Career&education、Goods&Service、Event&Activities 和 Trifle。文献[11]和文献[12]则分别利用基于 Convolutional-LSTM 模型和双向门控循环单元(BI-GRU)模型，进行显性消费意图预测。为解决源域标注数据不足的问题，Chen 等人<sup>[13]</sup>在挖掘“Intention Posts”的研究中提出了在不同领域表达购买意图的方式是相似的假设，利用迁移学习方法进行意图分类，即使用多个源域的标注数据辅助分类目标域内未标注数据，而 Song<sup>[14]</sup>等人则提出了一种基于多源域和多实例的迁移学习方法。

与显性消费意图识别问题相比，隐性消费意图识别的相关研究成果很少。刘挺等<sup>[1]</sup>提出了隐性意图识别的方法，该方法将隐性消费意图识别视作多标记分类问题，并综合使用基于用户关注行为、意图关注行为、意图转发行为以及个人信息的多种特征，并利用京东商城中的评论信息作为意图行为转化的依据。Park 等<sup>[15]</sup>构建了平行语料库，即隐性意图的表达方式与相应的显性意图表达方式对。例如“我好饿”对应的就是“我想吃东西”，前者就是隐性表达方式，后者是前者的相应的显性表达方式。Ding 等<sup>[6]</sup>提出的基于领域自适应卷积神经

网络的微博文本消费意图识别方法也可以适用于隐性意图识别，利用 Word Embedding 挖掘词的表示信息，通过卷积和池化操作提取局部词汇信息表示并得到句子级的表示，作者认为这种句子表示可以在不同的领域间有着相似性。

相比而言，很少研究学者消费意图对象的抽取问题，而用户的购买对象正是广告商所关心的。付博等<sup>[16]</sup>比较了几种基于句法依存关系的消费意图对象提取，并且提出了一种意图对象提取算法，即首先分析具有消费意图微博中的关键词，将其视为查询提交给搜索引擎进行搜索，最后从搜索结果中提取消费意图对象。Duan 等<sup>[17]</sup>先根据依赖性解析器自动提取候选产品，然后利用搭配抽取模型，从候选集中识别出真实的意向相关产品。Park<sup>[18]</sup>通过构建包含隐含意图文本和相应的显性意图文本的并行语料库方式完成移动应用检索任务。Wang<sup>[19]</sup>提出从微博中挖掘趋势相关产品的课题，将“趋势”定义为被用户热烈讨论的话题。

针对本研究领域面临的问题，本文提出利用迁移学习方法将京东商城数据用于预测微博用户的隐性消费意图，为微博用户的隐性消费意图识别提供了一种新的解决思路。在显性消费对象抽取的研究中，本文提出的基于 TF-IDF 与句法分析的用户消费意图对象的提取方法，在准确率方面也有较好的表现。

## 2.面向微博用户的消费意图识别

本文采用的微博用户消费意图识别的整体处理流程如图 1 所示。用户的消费意图识别可以分为显性消费意图识别和隐性消费识别，识别其消费意图后，又可以根据其属于显性消费意图或隐性消费意图继续提取其消费意图对象。

本文将消费意图识别视作二分类问题，并做如下定义：源域  $D_s = \{D_s^1, D_s^2 \dots D_s^n\}$  代表  $n$  条京东商品问答数据，其中  $D_s^i$  代表第  $i$  条京东问答数据；目标域  $D_t = \{(D_t^1, y^1), (D_t^2, y^2) \dots (D_t^m, y^m)\}$  代表少量具有标签的微博数据，其中  $(D_t^i, y^i)$  代表第  $i$  条数据，其对应的标签为  $y^i$ ， $y^i \in Y$ ， $Y = (-1, +1)$ 。

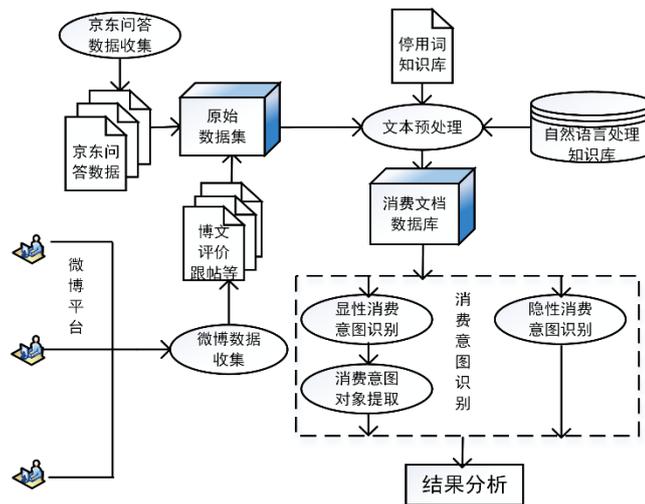


图 1 面向微博用户的消费意图识别处理框架

Fig.1 Processing framework of consumer intention recognition for Weibo users

### 2.1 基于迁移学习的目标域标注数据集构成方法

经统计微博中具有购买意图的数量占博文数的 3% 左右<sup>[5]</sup>，对微博数据全部进行人工标注以构建训练集，是非常耗费人力的工作。我们认为京东问答系统的文本数据经过预处理后均具有隐性购买意图，又基于不同应用领域中对购买意图的表达方式非常接近这一事实，本文提出利用京东问答数据作为源域，利用迁移学习方法将置信度较高的京东问答数据迁移到

---

具有少量标注数据的目标域微博训练集中，具体伪代码如算法 1 所示。

---

#### 算法 1: 基于迁移学习的目标域标注数据集构成算法

---

**Input:** 京东商品问答数据集  $D_s$ (源域), 少量标注的微博数据集  $D_t$ , 集合分类器中基分类器的数量  $N$ , 迭代次数  $M$ , 阈值  $K$

**Output:** 迁移源域数据的微博标注数据集  $D_t'$  (目标域)

**Begin**

1  $D_t' = D_t$

2 对  $D_t'$  和  $D_s$  进行文本预处理

3 **For**  $i = 1, \dots, M$  **Do**

4 基于 word2vec 词嵌入方法和 SVM 模型, 训练集成分类器中的  $N$  个基分类器

5 **For**  $j = 1, \dots, \text{Len}(D_s)$  **Do**

6 利用集成分类器中的  $N$  个基分类器分别预测  $D_s^j$  的类标签

7 **If**  $\text{Num}(\text{对 } D_s^j \text{ 极性预测一致的基分类器}) > K$

8  $D_t' = D_t' \cup D_s^j$

9 **EndIf**

10 **EndFor**

11 **EndFor**

12 **Return**  $D_t'$

---

算法的第 1 行将集合  $D_t$  赋予  $D_t'$ , 第 2 行对源域  $D_s$  和具有少量标注的微博数据集  $D_t$  进行预处理。之后第 4 行将  $D_t'$  中的博文转换成词嵌入的形式, 利用 SVM 模型训练  $N$  个基分类器。接着使用  $N$  个基分类器对源域  $D_s$  中的每个博文进行极性预测, 如果对第  $j$  个博文  $D_s^j$  极性的预测值一致的基分类器个数超过阈值  $K$ , 则认为该数据为置信度高的数据, 将其并入到微博标注数据集  $D_t'$  中 (7~9 行)。再使用  $D_t'$  重新训练分类器, 重复 3~9 行, 直至达到迭代次数  $M$ 。最后, 输出迁移源域数据的微博标注数据集  $D_t'$  (第 12 行)。

## 2.2 基于 Attentional-Bi-LSTM 模型的隐性消费意图识别算法

本文使用双向 LSTM 模型即同一个输入序列向前和向后各自训练一个 LSTM 模型, 并将输出进行线性化表示, 使得每一个 LSTM 单元都可以得到上下文信息, 然后将 Bi-LSTM 的输出向量集合构成注意力的权重矩阵。模型通过学习可以增加具有隐性消费意图的词语的权重, 降低无关词语的权重, 使得分类更加准确, 该模型如图 2 所示。

## 2.3 基于 TF-IDF-VOB 的消费意图对象提取算法

目前存在 2 种比较简单的基于经验性的分析消费意图对象在文本中的句法关系: 即修饰副词/名词+消费对象(ATT)和意图触发词+消费对象(VOB)的方式。而上述两种方式存在一定不足, 例如当博文较短不存在 ATT 或者 VOB 关系时, 则无法通过这种方式提取到意图对象。如果只是通过提取博文中的名词, 那么当名词数很多时无法识别哪个名词是真正的意图对象。通过分析具有显性消费意图对象的微博文本, 可以得出以下结论:

- (1) 消费意图对象通常是以名词为代表;
- (2) 消费意图对象通常与意图触发词构成动宾关系;
- (3) 微博文本中如果包含的某个商品链接, 那么这个商品则可能是用户的消费意图对象。

根据以上结论, 本文采用增加(2)和(3)中的意图对象所占的权重, 降低(1)中其他名词所占权重的方式提取意图对象。

消费意图对象通常与意图触发词构成 VOB 关系, 且为 VOB 关系中的名词部分, 故本

文通过 LTP 依存句法分析器提取与意图触发词搭配的动宾关系中的名词，如图 3 所示。

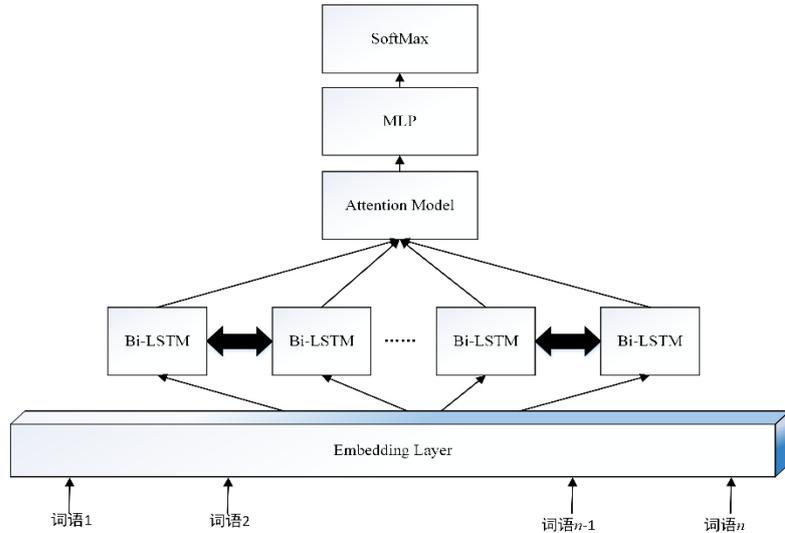


图 2 Attentional-Bi-LSTM 模型图

Fig.2 Attentional-Bi-LSTM Model

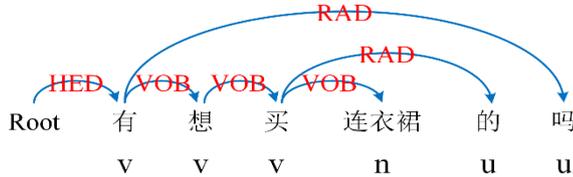


图 3 LTP 平台依存句法分析实例

Fig.3 An example of platform dependency syntax analysis

本文提出了一种基于 TF-IDF-VOB 的消费意图对象提取方法，其中用于衡量微博文本中每个名词重要程度的公式如(1)~(4)所示：

$$Score_{ij} = TF_{ij} * IDF_{ij} \quad (1)$$

$$TF_{ij} = \frac{n_{ij}}{\sum_k n_{kj}} \quad (2)$$

$$n_{ij} = \alpha \times n_{vob_{ij}} + \beta \times n_{others_{ij}} + \gamma \times n_{url_{ij}} \quad (3)$$

$$IDF_{ij} = \log \frac{N}{n_j + 1} \quad (4)$$

公式(2)中,  $TF_{ij}$ 表示第  $i$ 篇微博中第  $j$ 个名词的词频,  $k$ 表示微博的数量,  $n_{ij}$ 表示第  $j$ 个名词在该微博文本  $i$  相关的各部分文字中出现的次数的加权和。公式(3)中,  $n_{vob_{ij}}$ 表示第  $i$ 篇微博中第  $j$ 个名词, 并且是与意图指示词呈现 VOB 关系的名词出现的数量;  $n_{others_{ij}}$ 表示第  $i$ 篇微博中第  $j$ 个名词不是与意图指示词呈现 VOB 关系的名词的数量;  $n_{url_{ij}}$ 表示第  $i$ 篇微博中第  $j$ 个名词在所指向的 URL 链接中出现的次数;  $\alpha$ 、 $\beta$ 、 $\gamma$ 代表各自的权重。公式(4)中,  $N$ 表示微博数据集和京东问答数据集中博主的总和,  $n_j$ 表示其中包含单词  $j$ 的数量。对于每一篇微博的消费意图对象, 可以看作是通过式(1)计算得到的 Score 最大的单词。

### 3.实验与分析

由于包含消费意图的微博非常少, 所以本文采用搜索消费意图触发词的方式爬取一定量的新浪微博数据。本文采用人工标注的方式标注了 10000 条无消费意图的微博文本和 10000 条具有显性消费意图的微博文本, 并在 10000 条显性消费意图的微博中挑选出 1000 条, 标注其意图对象, 作为提取意图对象的数据集。同时, 我们将一部分包含显性消费意图的句子拆解成 2 部分, 即包含意图触发词的句子和不包含意图触发词的句子, 分别放入标注的显性

消费意图和隐性消费意图数据集。例如：“我的手机坏了，最近打算入手 p30”。其中“我的手机坏了”就是不含触发词，但是有一定的隐性消费意图。因此可以将其作为隐性消费意图识别的语料。

而对于京东商品问答数据集，我们假定在社区发表过商品问题咨询的用户均具有消费意图，并爬取京东问答数据 200000 条。经观察分析其中极少文本具有消费意图触发词，因此在预处理阶段将有购买触发词的文本删除，其他问答数据被视为具有隐性消费意图的文本加入京东问答数据集。

### 3.1 评价方法

本实验采用准确率  $P$ 、召回率  $R$  以及  $F1-Score$  值  $F1$  三个评价指标评价模型的性能，如公式(5)~(7)所示。

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (7)$$

### 3.2 性能分析

#### 3.2.1 基于迁移学习的目标域标注数据集构成

本文首先单独使用标注的微博数据集分别训练 SVM、LSTM 及 Bi-LSTM 模型，其结果如表 2 所示。可以发现在数据量不多的情况下，LSTM 模型和 SVM 模型的效果相当。

表 2 基于微博数据集的不同模型性能对比

Table 2 Performance comparison of different models based on Weibo dataset

	Precision	Recall	F1-Score
SVM	0.815	0.785	0.799718
LSTM	0.844	0.763	0.801458
Bi-LSTM	0.832	0.801	0.816205

其次，本文分别使用标注微博数据集、标注微博数据集+原始京东数据集和标注微博数据集+置信度高的京东数据集分别训练 LSTM 模型，其结果如表 3 所示。可以看出直接迁移京东数据到训练集中训练出的模型，其性能低于只使用微博数据集训练出的 LSTM 模型。而使用迁移置信度高的京东数据训练出的模型，性能比直接使用微博数据训练出的模型有了一定的提升。同时本文复现了 Chen<sup>[13]</sup>提出的 Co-Class 算法，并进行了对比。

表 3 基于不同数据组合的模型性能对比

Table 3 Performance comparison of models based on different data combinations

	Precision	Recall	F1-Score
LSTM(微博数据)	0.844	0.763	0.801458
LSTM(微博数据+直接迁移京东数据)	0.805	0.759	0.781323
Co-Class(微博数据+直接迁移京东数据)	0.829	0.780	0.803753
<b>LSTM(微博数据+迁移置信度高的京东数据)</b>	<b>0.929</b>	<b>0.891</b>	<b>0.909603</b>

### 3.2.2 基于 Attentional -Bi-LSTM 模型的隐性消费意图识别

本文在使用上述基于迁移学习方法生成目标域标注数据集的基础上，分别训练 LSTM、LSTM+Attention、Bi-LSTM 等模型，其结果如表 4 所示。可以看出本文首次提出的将京东平台的问答数据迁移到微博数据集中，可以提高对于隐性消费意图识别的准确率和召回率。

表 4 基于迁移后生成数据集的不同模型性能对比

Table 4 Performance comparison of different models based on migrated data sets

	Precision	Recall	F1-Score
LSTM	0.929	0.891	0.909603
LSTM+Attention	0.936	0.905	0.920239
Bi-LSTM	0.920	0.912	0.915982
CNN+LSTM	0.918	0.922	0.919996
<b>Attentional-Bi-LSTM</b>	<b>0.939</b>	<b>0.920</b>	<b>0.929403</b>

### 3.2.3 基于 TF-IDF-VOB 的显性意图对象提取

由于国内外没有公开的数据集，因此本文通过人工挑选出文本中具有消费意图对象的微博文本作为消费意图对象提取的数据 1000 条。由于样本较少，本文采用十折交叉验证方法，即将 1000 条数据平均分成 10 份，共做 10 轮实验，每轮选取不同的 1 份作为测试集，另外 9 份作为训练集。最终得到的评价指标是 10 轮实验的平均准确率。

本文在假定  $\alpha$ 、 $\beta$ 、 $\gamma$  在符合  $\alpha + \beta + \gamma = 1$  且  $\alpha$ 、 $\beta$ 、 $\gamma$  均大于 0 的约束条件下，通过调节各自的取值使得准确率最高。表 5 是通过分别限定  $\alpha$ 、 $\beta$ 、 $\gamma$  的值为 0，观察另外 2 个参数对于准确率的影响。通过适当的调节参数，将最佳的参数取值列出。

表 5 不同参数的对比

Table 5 Comparison of different parameters

	$\alpha$	$\beta$	$\gamma$	Precision(%)
$\alpha = 0$	-	0.35	0.65	50.5
$\beta = 0$	0.62	-	0.38	64.3
$\gamma = 0$	0.78	0.22	-	75.4

从结果中可以发现，通过提取和意图触发词构成 VOB 关系的名词所占的权重最大，和我们预期的消费意图对象通常与意图触发词构成动宾关系的结论保持一致。

另外，本文通过适当调节参数得到最优参数，并与传统的基于 ATT(方法 1)，VOB(方法 2)和只使用 TF-IDF 的方式(方法 3)三种基线方法做了比较，同时也复现了基于单词对齐的社交媒体检索方法<sup>[16]</sup>(方法 4)，性能对比如图 4 所示。可以发现本文提出方法(方法 5)的准确率高于其他方法。这是因为该方法通过分配权重的方式，提高了 VOB 关系中名词的权重，同时考虑到了其他名词的权重，从一定程度上解决了 VOB 或 ATT 方法的不足，也解决了在没有 VOB 或者 ATT 关系时，无法从名词中提取到意图对象的缺陷。

## 4.结论与展望

本文首先提出一种基于迁移学习的方法使用京东问答平台数据作为辅助数据，在此基础上提出一种基于 Attentional -Bi-LSTM 模型识别用户的隐性消费意图；同时提出一种基于 TF-IDF-VOB 的消费意图对象提取算法。实验结果表明，通过迁移京东问答平台的数据并与微博数据相融合可以有效扩充训练集，训练的分类模型具有较高的准确率和召回率；融合 VOB 和 TF-IDF 的显性消费意图对象提取方法在性能方面有较大提升。在未来研究工作中，我们

将尝试抽取微博中的多个意图对象，同时设计相应的推荐算法。

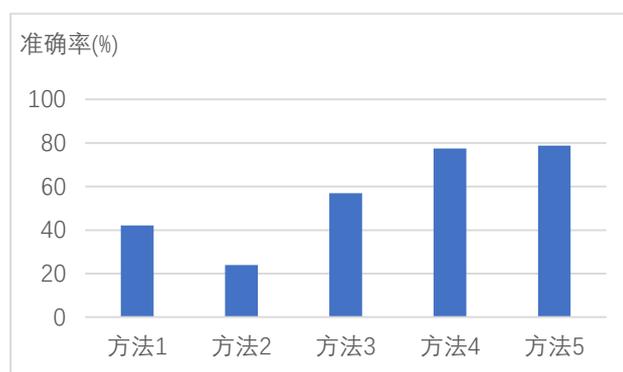


图 4 不同意图对象提取方法的性能对比

Fig.4 Performance comparison of different intent object extraction methods

### 参考文献

- [1] 付博, 刘挺. 社交媒体中用户的隐式消费意图识别[J]. 软件学报, 2016, 27(11):2843-2854.
- [2] Goldberg A B, Fillmore N, Andrzejewski D, et al. May all your wishes come true: A study of wishes and how to recognize them[C]//Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2009: 263-271
- [3] Mark Kröll, Strohmaier M. Analyzing human intentions in natural language text[C]// International Conference on Knowledge Capture. ACM, 2009:197-198.
- [4] Wang J, Cong G, Zhao W X, et al. Mining user intents in twitter: a semi-supervised approach to inferring intent categories for tweets[C]// Twenty-Ninth AAAI Conference on Artificial Intelligence. AAAI Press, 2015:318-324.
- [5] 陈浩辰. 基于微博的消费意图挖掘[D]. 哈尔滨工业大学, 2014.
- [6] Ding X, Liu T, Duan J, et al. Mining User Consumption Intention from Social Media Using Domain Adaptive Convolutional Neural Network[C]//AAAI. 2015, 15: 2389-2395.
- [7] Liu B, Lane I. Attention-based recurrent neural network models for joint intent detection and slot filling[J]. arXiv preprint arXiv:1609.01454, 2016.
- [8] Ding X, Cai B, Liu T, et al. Domain Adaptation via Tree Kernel Based Maximum Mean Discrepancy for User Consumption Intention Identification[C]//IJCAI. 2018: 4026-4032..
- [9] Yang M, Wang D, Feng S, et al. An Empirical Study on Learning Based Methods for User Consumption Intention Classification[C]//National CCF Conference on Natural Language Processing and Chinese Computing. Springer, Cham, 2017: 910-918.
- [10] Wang J, Cong G, Zhao W X, et al. Mining User Intents in Twitter: A Semi-Supervised Approach to Inferring Intent Categories for Tweets[C]//AAAI. 2015: 318-324.
- [11] 钱岳, 丁效, 刘挺, 等. 聊天机器人中用户出行消费意图识别方法[J]. 中国科学:信息科学, 2017(8):49-59.
- [12] 余慧, 冯旭鹏, 刘利军, et al. 聊天机器人中用户就医意图识别方法[J]. 计算机应用, 2018, v.38; No.336(08):36-40.
- [13] Chen Z, Liu B, Hsu M, et al. Identifying intention posts in discussion forums[C]//Proceedings of the 2013 conference of the North American chapter of the association for computational linguistics: human language technologies. 2013: 1041-1050.
- [14] Song H J, Park S B. Identifying intention posts in discussion forums using multi-instance learning and multiple sources transfer learning[J]. Soft Computing, 2017(4):1-12.
- [15] Park D H, Fang Y, Liu M, et al. Mobile app retrieval for social media users via inference of implicit intent in social media text[C]//Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. ACM, 2016: 959-968.
- [16] 付博, 刘挺. 基于跨社交媒体检索的微博消费对象识别[J]. 计算机科学与探索, 2015, 9(10):1247-1255.
- [17] Duan J, Chen Y, Liu T, Ding X, et al. Mining Intention-Related Products on Online Q&A Community. Journal of Computer Science and Technology. 2015.
- [18] Park D H, Fang Y, Liu M, et al. Mobile app retrieval for social media users via inference of implicit intent in social media text[C]//Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. ACM, 2016: 959-968.
- [19] Wang J, Zhao W X, Wei H, et al. Mining new business opportunities: Identifying trend related products by leveraging commercial intents from microblogs[C]//Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. 2013: 1337-1347.