

一种客服舆情预警方法、装置、电子设备及存储介质

相关申请的交叉引用

本申请要求享有于 2024 年 01 月 26 日提交的名称为“一种客服舆情预警方法、装置、电子设备及存储介质”的中国专利申请 CN202410115982.2 的优先权，该申请的全部内容通过引用并入本文中。

技术领域

本申请涉及人工智能技术领域，具体而言，涉及一种客服舆情预警方法、装置、电子设备及存储介质。

背景技术

随着互联网的快速发展，电子商务行业迅速崛起，电商平台成为了消费者在线购物的主要渠道。然而，随之而来的是大量的咨询和投诉需求，电商平台客服部门面临着巨大的压力。

然而由于客服人员很难及时处理所有的咨询和投诉，并且有些用户的表达方式可能不够理性，容易引发舆情，相关技术也提出了用于舆情预测的预测方法，但其在进行舆情预测时所考虑的维度较少，导致预测准确率较低。

发明内容

本申请实施例的目的在于提供一种客服舆情预警方法、装置、电子设备及存储介质，用以改善客服舆情预警准确率。

第一方面，本申请实施例提供一种客服舆情预警方法，该方法包括：根据进线客服工单，构建工单画像；其中，所述工单画像包括工单维度舆情发生概率、用户维度舆情发生概率以及订单维度舆情发生概率中至少一项；根据所述工单画像，预测所述进线客服工单的舆情发生概率；根据所述舆情发生概率以及舆情预警阈值，对所述进线客服工单进行舆情预警处理。

在上述方案的实现过程中，从工单维度、用户维度以及订单维度中至少一个维度来构建工单画像，从而根据工单画像预测工单的舆情发生概率，该方案采用工单维度、用户维度以及订单维度中至少一个维度来构建工单画像，有利于改善舆情预警准确率。

在第一方面的一种实现方式中，所述根据所述工单画像，预测所述进线客服工单的舆情发生概率，包括：若所述工单画像包含工单维度舆情发生概率、用户维度舆情发生概率以及订单维度舆情发生概率中至少两项，则利用训练好的集成学习模型，根据所述工单画像，预测所述进线客服工单的舆情发生概率。

在上述方案的实现过程中，在工单画像包含多个维度的舆情发生概率

时，采用训练好的集成学习模型，结合多个维度的舆情发生概率，共同预测进线客服工单的舆情发生概率，有利于改善上述客服舆情预警方法的舆情预警准确率。

在第一方面的一种实现方式中，所述根据进线客服工单，构建工单画像，包括：确定进线客服工单的工单主题；根据历史工单主题中对应工单主题的舆情发生概率，确定当前所述进线客服工单的工单维度舆情发生概率。

在上述方案的实现过程中，通过进线客服工单的工单主题即可快速确定当前进线客服工单的工单维度舆情发生概率，有利于改善上述客服舆情预警方法舆情预警效率。

在第一方面的一种实现方式中，所述确定进线客服工单的工单主题，包括：对所述进线客服工单进行分词处理；确定所述进线客服工单中单词的单词主题；根据所述进线客服工单中单词的所述单词主题，确定所述进线客服工单的工单主题。

在上述方案的实现过程中，通过对进线客服工单进行分词处理并确定单词的单词主题，从而确定进线客服工单的工单主题，有利于改善工单主题的主题确定准确率，进而有利于改善上述客服舆情预警方法的舆情预警准确率。

在第一方面的一种实现方式中，所述确定所述进线客服工单中单词的单词主题，包括：对所述进线客服工单中单词的单词主题进行随机初始化处理；针对所述进线客服工单中各所述单词及其所述单词主题，计算第一占比和第二占比；其中，所述第一占比用于指示所述进线客服工单中指定给当前单词主题的单词占比，所述第二占比用于指示历史客服工单中包含当前单词的客服工单的工单主题被指定为当前单词主题的工单占比；根据所述第一占比和所述第二占比，计算单词主题分配概率；根据所述单词主题分配概率，更新所述单词主题；重复迭代上述所述第一占比和所述第二占比的计算步骤、所述单词主题分配概率计算步骤以及所述单词主题更新步骤，直至达到预设迭代停止条件。

在上述方案的实现过程中，通过不断迭代更新进线客服工单中的单词主题，从而根据迭代更新后的单词主题确定进线客服工单的工单主题，有利于改善工单主题的主题确定准确率，进而有利于改善上述客服舆情预警方法的舆情预警准确率。

在第一方面的一种实现方式中，所述根据进线客服工单，构建工单画像，包括：获取与所述进线客服工单所关联的用户数据；在所述用户数据中提取用户特征；采用训练好的用户维度舆情发生概率预测模型，根据所述用户特征，预测用户维度舆情发生概率。

在上述方案的实现过程中，通过在与进线客服工单所关联的用户数据中所提取的用户特征，对用户维度舆情发生概率进行预测，从而丰富工单

画像，有利于改善上述客服舆情预警方法的舆情预警准确率。

在第一方面的一种实现方式中，所述根据进线客服工单，构建工单画像，包括：获取与所述进线客服工单所关联的订单数据；在所述订单数据中提取订单特征；采用训练好的订单维度舆情发生概率预测模型，根据所述订单特征，预测订单维度舆情发生概率。

在上述方案的实现过程中，通过在与进线客服工单所关联的订单数据中所提取的订单特征，对订单维度舆情发生概率进行预测，从而丰富工单画像，有利于改善上述客服舆情预警方法的舆情预警准确率。

第二方面，本申请实施例提供一种客服舆情预警装置，该装置包括：

画像构建模块，用于根据进线客服工单，构建工单画像；其中，所述工单画像包括工单维度舆情发生概率、用户维度舆情发生概率以及订单维度舆情发生概率中至少一项；

舆情发生概率预测模块，用于根据所述工单画像，预测所述进线客服工单的舆情发生概率；

预警处理模块，用于根据所述舆情发生概率以及舆情预警阈值，对所述进线客服工单进行舆情预警处理。

第三方面，本申请实施例提供一种电子设备，包括：处理器、存储器和通信总线，其中所述处理器和所述存储器通过所述通信总线完成相互间的通信；所述存储器中存储有可被所述处理器执行的计算机程序指令，所述计算机程序指令被所述处理器读取并运行时，执行第一方面或第一方面的任意一种可能的实现方式提供的方法。

第四方面，本申请实施例提供一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机程序指令，所述计算机程序指令被处理器读取并运行时，执行第一方面或第一方面的任意一种可能的实现方式提供的方法。

本申请的其他特征和优点将在随后的说明书阐述，并且，部分地从说明书中变得显而易见，或者通过实施本申请实施例了解。本申请的目的和其他优点可通过在所写的说明书、权利要求书、以及附图中所特别指出的结构来实现和获得。

附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例中所需要使用的附图作简单地介绍，应当理解，以下附图仅示出了本申请的某些实施例，因此不应被看作是对范围的限定，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他相关的附图。

图1为本申请实施例提供的客服舆情预警方法的流程示意图；

图2为本申请实施例提供的客服舆情预警装置的结构示意图；

图3为本申请实施例提供的电子设备的结构示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行描述。以下实施例仅用于更加清楚地说明本申请的技术方案，因此只作为示例，而不能以此来限制本申请的保护范围。

除非另有定义，本文所使用的所有的技术和科学术语与属于本申请的技术领域的技术人员通常理解的含义相同；本文中所使用的术语只是为了描述具体的实施例的目的，不是旨在于限制本申请；本申请的说明书和权利要求书及上述附图说明中的术语“包括”和“具有”以及它们的任何变形，意图在于覆盖不排他的包含。

在本申请实施例的描述中，技术术语“第一”“第二”等仅用于区别不同对象，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量、特定顺序或主次关系。在本申请实施例的描述中，“多个”的含义是两个以上，除非另有明确具体的限定。

在本文中提及“实施例”意味着，结合实施例描述的特定特征、结构或特性可以包含在本申请的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例，也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是，本文所描述的实施例可以与其它实施例相结合。

在本申请实施例的描述中，术语“和/或”仅仅是一种描述关联对象的关联关系，表示可以存在三种关系，例如A和/或B，可以表示：单独存在A，同时存在A和B，单独存在B这三种情况。另外，本文中字符“/”，一般表示前后关联对象是一种“或”的关系。

目前，电商平台成为了消费者在线购物的主要渠道，电商平台客服部门需要处理大量的咨询和投诉需求，在这样的环境下，用户对平台的满意度以及对电商平台的评价成为了重要的课题。用户对电商平台的满意度体现在一个重要的方面，即是否发生舆情，因此，可以通过对舆情进行预警的方式来提前做好准备以应对即将发生的舆情，或者在发生舆情前及时对客服进行处理，从而改善用户对平台的满意度。

相关技术中的舆情预测方法大都是针对用户进行画像，并判断该用户发生舆情的概率从而进行舆情预警，但这种方案由于仅考虑了用户维度，导致舆情预警准确率较低。

基于此，本申请实施例提供一种客服舆情预警方法，不同于相关技术中利用用户画像来进行舆情预警的方式，上述方案可以从工单维度、用户维度以及订单维度中至少一个维度来构建工单画像，从而根据工单画像预测工单的舆情发生概率，该方案采用工单维度、用户维度以及订单维度中至少一个维度来构建工单画像，有利于改善舆情预警准确率。

下面对上述客服舆情预警方法进行详细介绍，请参见图 1，本申请实施例提供一种客服舆情预警方法，该方法包括：

步骤 S110：根据进线客服工单，构建工单画像；

其中，工单画像包括工单维度舆情发生概率、用户维度舆情发生概率以及订单维度舆情发生概率中至少一项；

步骤 S120：根据工单画像，预测进线客服工单的舆情发生概率；

步骤 S130：根据舆情发生概率以及舆情预警阈值，对进线客服工单进行舆情预警处理。

上述步骤 S110 中的工单是一种记录和跟踪客户问题和需求的文档和电子表格。客服工单则是指用户向平台提出问题或需求时，人工或自动创建的工单，该工单记录有用户的基本信息、问题描述、需求描述以及处理进度等内容中的一项或多项，以供客服系统使用。

上述客服舆情是指针对企业或组织的客户服务部门的公众舆论情况，这种舆论通常涉及客户对企业服务质量、服务态度、服务效率等方面的评价和看法，是客户对服务部门综合态度的体现。

上述步骤 S110 中的工单维度舆情发生概率是指从工单自身维度所获取的单个维度下的舆情发生概率，工单维度舆情发生概率的获取方案如下：

作为上述客服舆情预警方法的一种可选实施方式，上述步骤 S110 包括：确定进线客服工单的工单主题；根据历史工单主题中对应工单主题的舆情发生概率，确定当前进线客服工单的工单维度舆情发生概率。

可以理解的是，上述当前进线客服工单的工单维度舆情发生概率可以通过对应工单主题下发生舆情的工单占对应主题下的历史工单总数的占比来确定历史工单主题中对应工单主题的舆情发生概率。

可以理解的是，上述方案可以采用基于规则的主题提取方法来提取进线客服工单的工单主题，基于规则的主题提取方法例如关键词匹配、正则表达式等，规则可以是人为设定的提取规则。

上述方案通过进线客服工单的工单主题即可快速确定当前进线客服工单的工单维度舆情发生概率，有利于改善上述客服舆情预警方法舆情预警效率。

另外，可以理解的是，上述工单主题还可以采用基于强化学习的方法来确定，下面对基于强化学习的主题提取方法的一种可选实施方式进行介绍：

作为上述客服舆情预警方法的一种可选实施方式，上述确定进线客服工单的工单主题，包括：对进线客服工单进行分词处理；确定进线客服工单中单词的单词主题；根据进线客服工单中单词的单词主题，确定进线客服工单的工单主题。

可以理解的是，可以采用成熟的分词模型对进线客服工单进行分词处理。

上述进线客服工单中单词的单词主题以及进线客服工单的工单主题可以通过 NMF (Non-negative Matrix Factorization , 非负矩阵分解) 算法、LSA (Latent Semantic Analysis , 奇异值分解) 算法、HDP (Hierarchical Dirichlet Process , 层次狄利克雷过程) 算法等分词算法来获取。上述 NMF 算法、LSA 算法和 HDP 算法均为相关技术中较为成熟的算法，本申请实施不再赘述。

上述方案通过对进线客服工单进行分词处理并确定单词的单词主题，从而确定进线客服工单的工单主题，有利于改善工单主题的主体确定准确率，进而有利于改善上述客服舆情预警方法的舆情预警准确率。

作为上述客服舆情预警方法的一种可选实施方式，上述确定进线客服工单中单词的单词主题，包括：对进线客服工单中单词的单词主题进行随机初始化处理；针对进线客服工单中各单词及其单词主题，计算第一占比和第二占比；其中，第一占比用于指示进线客服工单中指定给当前单词主题的单词占比，第二占比用于指示历史客服工单中包含当前单词的客服工单的工单主题被指定为当前单词主题的工单占比；根据第一占比和第二占比，计算单词主题分配概率；根据单词主题分配概率，更新单词主题；重复迭代上述第一占比和第二占比的计算步骤、单词主题分配概率计算步骤以及单词主题更新步骤，直至达到预设迭代停止条件。该实施方式例如：

对于进行分词处理后的进线客服工单，将单词的单词主题随机初始化为 K 个主题中的一个 (K 个主题可以为事先选定的所有可选的工单主题)；

对于进线客服工单 D 以及进线客服工单中的单词 W，计算第一占比和第二占比，其中，第一占比用于指示进线客服工单 D 中指定给当前单词 W 的单词主题的单词占比，第二占比用于指示所有包含单词 W 的历史客服工单中，指定给工单主题 T 的概率；

第一占比的计算方法为：

$$P_1(T|D)$$

第二占比的计算方法为：

$$P_2(W|T)$$

根据第一占比和第二占比，计算单词主题分配概率 P_3 的方法为：

$$P_3 = P_1(T|D) * P_2(W|T)$$

根据上述单词主题分配概率，更新单词主题；

重新迭代上述第一占比计算步骤、第二占比计算步骤、单词主题分配概率计算步骤以及单词主题更新步骤，直至单词主题不再变化，或者直至重新迭代预设次数。

可以理解的是，上述方案所计算的工单主题可以为 LDA (Latent Dirichlet Allocations , 隐含狄利克雷分布) 主题。

上述方案通过不断迭代更新进线客服工单中的单词主题，从而根据迭代更新后的单词主题确定进线客服工单的工单主题，有利于改善工单主题的主题确定准确率，进而有利于改善上述客服舆情预警方法的舆情预警准

确率。

可以理解的是，在对进线客服工单进行分词处理之前，还可以对进线客服工单进行内容清洗、剔除特殊字符以及标点符号等预处理。

下面对用户维度舆情发生概率的获取方案进行详细介绍：

作为上述客服舆情预警方法的一种可选实施方式，上述步骤 S110 包括：获取与进线客服工单所关联的用户数据；在用户数据中提取用户特征；采用训练好的用户维度舆情发生概率预测模型，根据用户特征，预测用户维度舆情发生概率。

可以理解的是，进线客服工单会与用户所关联，因此可以根据与进线客服工单所关联的用户来获取用户数据，用户数据包括年龄、地域、历史交易行为，其中历史交易行为包括但不限于下单、售后、进线、投诉等。

另外，可以理解的是，在获取与进线客服工单所关联的用户数据后，还可以对用户数据进行缺失值处理以及特征分箱处理等预处理。

上述用户维度舆情发生概率预测模型可以采用 LightGBM 模型，LightGBM 模型是一种基于决策树算法的梯度提升框架模型，可以基于已有的用户层面舆情黑样本，训练 LightGBM 模型，从而使得 LightGBM 模型可以根据用户特征预测用户维度舆情发生概率。

上述方案通过在与进线客服工单所关联的用户数据中所提取的用户特征，对用户维度舆情发生概率进行预测，从而丰富工单画像，有利于改善上述客服舆情预警方法的舆情预警准确率。

下面对订单维度舆情概率的获取方案进行介绍：

作为上述客服舆情预警方法的一种可选实施方式，上述步骤 S110 包括：获取与进线客服工单所关联的订单数据；在订单数据中提取订单特征；采用训练好的订单维度舆情发生概率预测模型，根据订单特征，预测订单维度舆情发生概率。

可以理解的是，进线客服工单还会与订单相关联，因此可以根据与进线客服工单所关联的订单来获取订单数据，订单数据包括但不限于当前交易订单、历史交易订单、商品类型、商品价格、物流时间等。

另外，可以理解的是，在获取与进线客服工单所关联的订单数据后，还可以对订单数据进行缺失值填充以及特征分箱等预处理。

上述订单维度舆情发生概率预测模型也可以采用 LightGBM 模型，可以基于已有的订单层面舆情黑样本，训练 LightGBM 模型，从而使得 LightGBM 模型可以根据订单特征预测用户维度舆情发生概率。

上述方案通过在与进线客服工单所关联的订单数据中所提取的订单特征，对订单维度舆情发生概率进行预测，从而丰富工单画像，有利于改善上述客服舆情预警方法的舆情预警准确率。

下面对步骤 S120 进行详细介绍：

由于工单画像可以包括工单维度舆情发生概率、用户维度舆情发生概

率以及订单维度舆情发生概率中至少一项，下面则针对工单画像仅包含一项舆情发生概率以及包含多项舆情发生概率这两种实施方式进行详细介绍：

第一种实施方式，工单画像仅包含一项舆情发生概率，此时，步骤 S120 包括：

将上述工单画像中所包含的一项舆情发生概率确定为进线客服工单的舆情发生概率。

第二种实施方式，工单画像包含多项舆情发生概率，此时，步骤 S120 包括：

利用训练好的集成学习模型，根据工单画像，预测进线客服工单的舆情发生概率。

集成学习是指将机器学习中的单个算法集合在一起进行建模的思路，集成学习有 Bagging(Bootstrap Aggregation) ,Boosting 和堆叠法三种方法，其中，Bagging 的核心思路是共同投票，每个基础模型都只有一票，投票最多的结果作为最终结果；Boosting 的核心思路是挑选精英，对基础模型进行考验和筛选，给予精英模型更多的投票权，给予其他模型较少的投票权，然后综合所有模型的投票获取最终结果；堆叠法由元估计器和基础估计器两部分构成，元估计器把基础估计器的预测结果作为输入变量，最终对被解释变量做出预测。基础估计器可以由不同的算法模型（如：逻辑回归和 KNN）进行组合，元估计器也可以选择完全不同的算法模型（如：决策树），以实现最终减小误差的更好效果。

本申请实施例中的集成学习模型可以采用 Boosting 的方式来进行学习，例如：结合历史舆情样本，分别确定工单维度舆情发生概率、用户维度舆情发生概率以及订单维度舆情发生概率对进线客服工单的舆情发生概率的影响关系，例如确定工单维度舆情发生概率对进线客服工单的舆情发生概率的影响程度较大，而用户维度舆情发生概率以及订单维度舆情发生概率对进线客服工单的舆情发生概率的影响程度较小，因此，可以给予工单维度舆情发生概率较大的投票权，而给予用户维度舆情发生概率以及订单维度舆情发生概率较小的投票权，基于此，可以采用下述内容所示的集成学习方式：

（1）在工单维度舆情发生概率较高时，若用户维度舆情发生概率高，或订单维度舆情发生概率高，或用户维度舆情发生概率中等且订单维度舆情发生概率中等，则集成学习模型赋予舆情发生概率较高数值；

（2）在工单维度舆情发生概率中等时，若用户维度舆情发生概率高且订单维度舆情发生概率高，则集成学习模型赋予舆情发生概率较高数值；

（3）在工单维度舆情发生概率较低时，则集成学习模型赋予舆情发生概率较低数值。

上述方案在工单画像包含多个维度的舆情发生概率时，采用训练好的集成学习模型，结合多个维度的舆情发生概率，共同预测进线客服工单的

舆情发生概率，有利于改善上述客服舆情预警方法的舆情预警准确率。

下面对步骤 S130 进行详细介绍：

可以理解的是，步骤 S130 中的舆情预警阈值可以为预设舆情预警阈值，也可以为以风险工单召回率和精确率为评估指标所定期选择或实时选择的最优舆情预警阈值。

本申请实施例还提供上述客服舆情预警方法在某场景下的一种具体实现方式，在该场景下，工单画像包括工单维度舆情发生概率、用户维度舆情发生概率以及订单维度舆情发生概率这三个维度的舆情发生概率，此时，上述客服舆情预警方法可以包括：

步骤一、数据收集；

获取进线客服工单以及与进线客服工单所关联的订单数据和用户数据；

步骤二、工单维度数据梳理；

对进线客服工单进行内容清洗以及剔除特殊字符、标点符号等预处理，并对进线客服工单进行分词处理；

步骤三、单词主题随机初始化处理；

对于进线客服工单，随机将单词的单词主题初始化为 K 个主题中的一个（ K 个主题为预先选定的可选主题）；

步骤四、单词主题迭代；

对于进线客服工单 D ，浏览单词 W ，并计算 $P(T|D)$ 和 $P(W|T)$ ，其中， $P(T|D)$ 代表进线客服工单 D 中，指定给主题 T 的单词的比例； $P(W|T)$ 代表所有包含单词 W 的历史客服工单中，指定给主题 T 的比例；

考虑所有其他单词及其主题分配，以概率 $P(T|D)*P(W|T)$ 将单词 W 与主题 T 重新分配；

步骤五、LDA 主题模型建立；

重复迭代上述步骤四，直至单词主题稳定，并根据进线客服工单中单词的单词主题共同确定进线客服工单的工单主题；

步骤六、工单维度舆情发生概率计算；

基于进线客服工单的工单主题以及历史客服工单中对应工单主题的舆情发生概率，共同确定工单维度舆情发生概率；

步骤七、用户维度数据梳理；

用户数据包括但不限于：年龄、所处地域、历史交易行为；

其中，历史交易行为包括但不限于：下单、售后、进线、投诉等；

可以理解的是，梳理数据后还可以对数据进行缺失值填充以及特征分箱等预处理；

步骤八、计算用户维度舆情发生概率；

基于已有的用户维度舆情黑样本所训练的 LightGBM 监督模型，预测用户维度舆情发生概率；

步骤九、订单维度数据梳理；

订单数据包括但不限于：当前交易订单信息、历史交易订单信息、商品类型、商品价格、物流时间等；

可以理解的是，梳理数据后还可以对数据进行缺失值填充以及特征分箱等预处理；

步骤十、计算订单维度舆情发生概率；

基于已有的订单维度舆情黑样本所训练的 LightGBM 监督模型，预测订单维度舆情发生概率；

步骤十一、实时进线客服工单特征整合；

基于实时进线客服工单的唯一 ID，关联该进线客服工单的工单画像，并将其整个为集成学习特征；

步骤十二、实时模型训练；

基于已有样本数据，通过集成学习，将风险工单召回率与精确率作为评估指标，确定最优舆情预警阈值；

步骤十三、基于最优舆情预警阈值，实时针对进线客服工单，进行客服舆情预警，针对发生预警的客服工单，及时进行相关处理。

需要指出的是，上述内容中所指出的与用户相关的数据是在经过用户同意后所获取的。

请参见图 2，基于同一发明构思，本申请实施例中还提供一种客服舆情预警装置 200，该装置包括：

画像构建模块 210，用于根据进线客服工单，构建工单画像；其中，所述工单画像包括工单维度舆情发生概率、用户维度舆情发生概率以及订单维度舆情发生概率中至少一项；

舆情发生概率预测模块 220，用于根据所述工单画像，预测所述进线客服工单的舆情发生概率；

预警处理模块 230，用于根据所述舆情发生概率以及舆情预警阈值，对所述进线客服工单进行舆情预警处理。

作为上述客服舆情预警装置的一种可选实施方式，上述舆情发生概率预测模块 220 具体用于：在所述工单画像包含工单维度舆情发生概率、用户维度舆情发生概率以及订单维度舆情发生概率中至少两项时，利用训练好的集成学习模型，根据所述工单画像，预测所述进线客服工单的舆情发生概率。

作为上述客服舆情预警装置的一种可选实施方式，上述画像构建模块 210 包括：

工单主题确定单元，用于确定进线客服工单的工单主题；

工单维度舆情发生概率确定单元，根据历史工单主题中对应工单主题的舆情发生概率，确定当前所述进线客服工单的工单维度舆情发生概率。

作为上述客服舆情预警装置的一种可选实施方式，上述工单主题确定单元具体用于：对所述进线客服工单进行分词处理；确定所述进线客服工

单中单词的单词主题；根据所述进线客服工单中单词的所述单词主题，确定所述进线客服工单的工单主题。

作为上述客服舆情预警装置的一种可选实施方式，上述确定所述进线客服工单中单词的单词主题，包括：对所述进线客服工单中单词的单词主题进行随机初始化处理；针对所述进线客服工单中各所述单词及其所述单词主题，计算第一占比和第二占比；其中，所述第一占比用于指示所述进线客服工单中指定给当前单词主题的单词占比，所述第二占比用于指示历史客服工单中包含当前单词的客服工单的工单主题被指定为当前单词主题的工单占比；根据所述第一占比和所述第二占比，计算单词主题分配概率；根据所述单词主题分配概率，更新所述单词主题；重复迭代上述所述第一占比和所述第二占比的计算步骤、所述单词主题分配概率计算步骤以及所述单词主题更新步骤，直至达到预设迭代停止条件。

作为上述客服舆情预警装置的一种可选实施方式，上述画像构建模块 210 包括：

用户数据获取单元，用于获取与所述进线客服工单所关联的用户数据；

用户特征提取单元，用于在所述用户数据中提取用户特征；

用户维度舆情发生概率预测单元，用于采用训练好的用户维度舆情发生概率预测模型，根据所述用户特征，预测用户维度舆情发生概率。

作为上述客服舆情预警装置的一种可选实施方式，上述画像构建模块 210 包括：

订单数据获取单元，用于获取与所述进线客服工单所关联的订单数据；

订单特征提取单元，用于在所述订单数据中提取订单特征；

订单维度舆情发生概率预测单元，用于采用训练好的订单维度舆情发生概率预测模型，根据所述订单特征，预测订单维度舆情发生概率。

图 3 为本申请实施例提供的一种电子设备的示意图。参照图 3，电子设备 300 包括：处理器 310、存储器 320 以及通信接口 330，这些组件通过通信总线 340 和/或其他形式的连接机构（未示出）互连并相互通讯。

其中，存储器 320 包括一个或多个（图中仅示出一个），其可以是，但不限于，随机存取存储器（Random Access Memory，简称 RAM），只读存储器（Read Only Memory，简称 ROM），可编程只读存储器（Programmable Read-Only Memory，简称 PROM），可擦除可编程只读存储器（Erasable Programmable Read-Only Memory，简称 EPROM），电可擦除可编程只读存储器（Electric Erasable Programmable Read-Only Memory，简称 EEPROM）等。处理器 310 以及其他可能的组件可对存储器 320 进行访问，读和/或写其中的数据。

处理器 310 包括一个或多个（图中仅示出一个），其可以是一种集成电路芯片，具有信号的处理能力。上述的处理器 310 可以是通用处理器，包括中央处理器（Central Processing Unit，简称 CPU）、微控制单元（Micro

Controller Unit, 简称 MCU)、网络处理器(Network Processor, 简称 NP)或者其他常规处理器;还可以是专用处理器,包括数字信号处理器(Digital Signal Processor,简称 DSP)、专用集成电路(Application Specific Integrated Circuits, 简称 ASIC)、现场可编程门阵列(Field Programmable Gate Array, 简称 FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。

通信接口 330 包括一个或多个(图中仅示出一个),可以用于和其他设备进行直接或间接地通信,以便进行数据的交互。例如,通信接口 330 可以是以太网接口;可以是移动通信网络接口,例如 3G、4G、5G 网络的接口;还是可以是具有数据收发功能的其他类型的接口。

在存储器 320 中可以存储一个或多个计算机程序指令,处理器 310 可以读取并运行这些计算机程序指令,以实现本申请实施例提供的客服舆情预警方法以及其他期望的功能。

可以理解,图 3 所示的结构仅为示意,电子设备 300 还可以包括比图 3 中所示更多或者更少的组件,或者具有与图 3 所示不同的配置。图 3 中所示的各组件可以采用硬件、软件或其组合实现。例如,电子设备 300 可以是单台服务器(或其他具有运算处理能力的设备)、多台服务器的组合、大量服务器的集群等,并且,既可以是物理设备也可以是虚拟设备。

本申请实施例还提供一种计算机可读存储介质,该计算机可读存储介质上存储有计算机程序指令,所述计算机程序指令被计算机的处理器读取并运行时,执行本申请实施例提供的客服舆情预警方法。例如,计算机可读存储介质可以实现为图 3 中电子设备 300 中的存储器 320。

在本申请所提供的实施例中,应该理解到,所揭露装置和方法,可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,又例如,多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些通信接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

另外,作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

再者,在本申请各个实施例中的各功能模块可以集成在一起形成一个独立的部分,也可以是各个模块单独存在,也可以两个或两个以上模块集成形成一个独立的部分。

以上所述仅为本申请的实施例而已,并不用于限制本申请的保护范围,对于本领域的技术人员来说,本申请可以有各种更改和变化。凡在本申请

的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

权利要求书

1. 一种客服舆情预警方法，其特征在于，所述方法包括：
根据进线客服工单，构建工单画像；其中，所述工单画像包括工单维度舆情发生概率、用户维度舆情发生概率以及订单维度舆情发生概率中至少一项；
根据所述工单画像，预测所述进线客服工单的舆情发生概率；
根据所述舆情发生概率以及舆情预警阈值，对所述进线客服工单进行舆情预警处理。
2. 根据权利要求 1 所述的客服舆情预警方法，其特征在于，所述根据所述工单画像，预测所述进线客服工单的舆情发生概率，包括：
若所述工单画像包含工单维度舆情发生概率、用户维度舆情发生概率以及订单维度舆情发生概率中至少两项，则利用训练好的集成学习模型，根据所述工单画像，预测所述进线客服工单的舆情发生概率。
3. 根据权利要求 1 所述的客服舆情预警方法，其特征在于，所述根据进线客服工单，构建工单画像，包括：
确定进线客服工单的工单主题；
根据历史工单主题中对应工单主题的舆情发生概率，确定当前所述进线客服工单的工单维度舆情发生概率。
4. 根据权利要求 3 所述的客服舆情预警方法，其特征在于，所述确定进线客服工单的工单主题，包括：
对所述进线客服工单进行分词处理；
确定所述进线客服工单中单词的单词主题；
根据所述进线客服工单中单词的所述单词主题，确定所述进线客服工单的工单主题。
5. 根据权利要求 4 所述的客服舆情预警方法，其特征在于，所述确定所述进线客服工单中单词的单词主题，包括：
对所述进线客服工单中单词的单词主题进行随机初始化处理；
针对所述进线客服工单中各所述单词及其所述单词主题，计算第一占比和第二占比；其中，所述第一占比用于指示所述进线客服工单中指定给当前单词主题的单词占比，所述第二占比用于指示历史客服工单中包含当前单词的客服工单的工单主题被指定为当前单词主题的工单占比；
根据所述第一占比和所述第二占比，计算单词主题分配概率；
根据所述单词主题分配概率，更新所述单词主题；
重复迭代上述所述第一占比和所述第二占比的计算步骤、所述单词主题分配概率计算步骤以及所述单词主题更新步骤，直至达到预设迭代停止条件。
6. 根据权利要求 1~5 中任一项所述的客服舆情预警方法，其特征在于，所述根据进线客服工单，构建工单画像，包括：

- 获取与所述进线客服工单所关联的用户数据；
在所述用户数据中提取用户特征；
采用训练好的用户维度舆情发生概率预测模型，根据所述用户特征，
预测用户维度舆情发生概率。
- 5 7. 根据权利要求 1~5 中任一项所述的客服舆情预警方法，其特征在于，所述根据进线客服工单，构建工单画像，包括：
获取与所述进线客服工单所关联的订单数据；
在所述订单数据中提取订单特征；
采用训练好的订单维度舆情发生概率预测模型，根据所述订单特征，
10 预测订单维度舆情发生概率。
8. 一种客服舆情预警装置，其特征在于，所述装置包括：
画像构建模块，用于根据进线客服工单，构建工单画像；其中，所述
工单画像包括工单维度舆情发生概率、用户维度舆情发生概率以及订单维
度舆情发生概率中至少一项；
15 舆情发生概率预测模块，用于根据所述工单画像，预测所述进线客服
工单的舆情发生概率；
预警处理模块，用于根据所述舆情发生概率以及舆情预警阈值，对所
述进线客服工单进行舆情预警处理。
9. 一种电子设备，其特征在于，包括：处理器、存储器和通信总线，
20 其中，所述处理器和所述存储器通过所述通信总线完成相互间的通信；所
述存储器存储有可被所述处理器执行的程序指令，所述处理器调用所述程
序指令能够执行如权利要求 1~7 中任一项所述的方法。
10. 一种计算机可读存储介质，其特征在于，所述计算机可读存储介
质存储计算机指令，所述计算机指令被计算机运行时，使所述计算机执行
25 如权利要求 1~7 中任一项所述的方法。