
目录

前言： 2

一、 题目 2

二、 技术背景 2

三、 发明目的 2

四、 技术方案 3

五、 实施示例 5

六、 技术效果 10

七、 本发明的优点 10

八、 本发明的关键点和保护点 10

九、 摘要 12

前言：

专利撰写包括：题目、技术背景、发明目的（一般不出现在终稿）、技术方案（可不写，写清楚实施示例即可）、实施示例、技术效果（一般不出现在终稿）、本发明的优点（一般不出现在终稿）、本发明的关键点和保护点、摘要

由于发明人写完交底书（原稿）并提交后，公司会交给专门的代理机构修改、优化，再提交专利局评审，因此交底书核心是把技术内容说明清楚（即使不够清楚，代理机构也会联系发明人了解清楚），发明目的、技术效果等用于辅助代理机构理解本发明的价值和背景。

一、题目

一种基于人工智能的车辆首发故障判断方法和系统

凸显核心的创新点或解决的核心问题，命名一般是“一种 xxx 的方法/系统/装置”，代理机构会根据发明内容进行优化调整。

二、技术背景

根据车质网 2020 年统计，车辆质量有效投诉 97007 宗，再次刷新历史记录，连续三年实现阶梯式增长，较 2019 年上涨 10.03%。车辆可靠性已成为汽车制造商（OEM）品牌口碑的主要因素，是提高品牌竞争力的关键点。而能否及时发现车辆的首发故障（自身品牌车辆中首次出现的质量问题），是 OEM 进行产品的质量监控、制定维修策略、制定召回策略等环节的关键，及时对首发故障进行针对性处理，不仅能防止更大的安全事故发生，而且能防止品牌的负面信息在市场中扩散。因此，首发故障的判断是 OEM 提升品牌口碑的重要一环。

目前，OEM 对首发故障判断投入了大量的人力和物力，但是效果不佳。当前模式是结合以往的历史故障信息，人工对每天产生的维修报告进行检验，进而判断是否首发。但由于个人的主观性和数据量的庞大，且加上冗长的提交审核流程，使得目前的首发故障判断不仅准确率低，而且效率低。因此，研发更快速、更智能的判断方法，提升首发故障判断的准确率和效率是 OEM 当前亟待解决的问题。

模板：“解决问题很重要/有价值”+“目前方法不好/没有方法”+“简述方法可行性（可不写）”。

“很重要/有价值”把业务背景写上即可；

“方法不好”从效率低、高费用、不准确等方向写，“没有方法”直接写行业没有解决办法即可；

“方法可行性”说明客观条件具备，以本发明例子：“OEM 已有大量维修记录，记录中包含着故障部件和现象...人工智能的 NLP 算法，可以对海量非结构化的文本进行分析...”

三、发明目的

首发故障的判断是 OEM 提升品牌口碑的重要一环，而当前的人工判断模式不仅准确率

低，而且效率低。针对这些问题，本发明提供一种基于人工智能的车辆首发故障判断方法和系统，能快速、高效、准确的识别首发故障，减轻 OEM 人力物力的投入，提升首发故障识别的准确率和效率，为其进行产品的质量监控、制定维修策略等环节奠定基础。

模板：“一句话，当前问题”+“一句话，提出方法”+“一句话，获得价值”。

也可“问题”和“价值”二选一，如“针对这些问题，本发明提供一种基于人工智能的车辆首发故障判断方法和系统，能快速、高效、准确的识别首发故障，减轻 OEM 人力物力的投入，提升首发故障识别的准确率和效率，为其进行产品的质量监控、制定维修策略等环节奠定基础。”

注：让代理机构明白即可，一般不会在最终稿中出现。

四、技术方案

按自己的技术路线，划分步骤，再用“总-分”结构撰写。

撰写方式，分别用一句话总结各个大步骤，再分别用一句话总结大步骤中的小步骤，最后详细说明每个小步骤。

注：1）公开的方法不需要详细展开，如：“通过神经网络算法，对文本内容分类...”，不需要展开解释什么是神经网络算法，而是解释如何把文本处理好给算法计算，输出的结构要如何使用等。

2）经量使用数学、物理等工科语言表达技术方案，如公式、规则等。

3）方法可以尽量详细，代理机构明白方案后，会对文章结构、术语做调整，内容会模糊化。

4）术语、参数名要统一，避免自创术语，用通用术语。

本发明所要解决的技术问题是通过以下技术方案实现的：

提供一种基于人工智能的车辆首发故障判断方法，包括以下步骤：

步骤 1：从质量管理系统中随机抽取部分维修记录，以此作为样本数据，在辅助标注工具的基础上进行人工标注。具体包括三个步骤：

（1）数据获取与清洗。从质量管理系统中抽取部分维修记录，即客户投诉内容，技师检查结果，技师维修策略等文本数据。并过滤乱码、错位等数据。

所述质量管理系统是 OEM 的存储每次维修详细记录的平台，包括了客户投诉内容，相应的技师检查结果和维修策略等。

（2）开发辅助标注工具。针对步骤（1）的文本格式开发辅助标注工具，能实现文本的简明显示和文本打标签。

（3）数据标注。通过辅助标注工具，对每条维修记录，标注出部件词及其对应的现象词（车辆故障现象），及故障对。部件词如：发动机、变速箱等，现象词如：漏油、抖动、异响等，故障对如：发动机—漏油，变速箱—异响等。

步骤 2：基于每条数据标注出的部件、现象，结合 NLP 技术构建故障模式实体抽取器，基于该抽取器抽取故障对（部件词—现象词）。构建过程具体包括三个步骤：

所属的故障模式实体抽取器包括三个模型，部件现象抽取模型，故障对组合筛选模型，实

体标准化模型。

(1) 识别每条数据的部件和故障现象。

1) 根据步骤 1 获得标注后的维修数据, 对原始文本进行 BIO 序列标记, 通过开源的预训练模型将文本转换为词向量。2) 将 BIO 序列标记和词向量结合, 得到训练数据。3) 在此基础上, 训练深度学习实体识别模型 (部件和现象抽取模型)。4) 通过模型将剩余未标注的维修数据, 进行实体识别, 对每条维修数据识别出部件集合 $P_i = \{part_1, part_2, \dots, part_n\}$, 现象集合 $F_i = \{mal f_1, mal f_2, \dots, mal f_m\}$, i 表示第 i 条维修数据, $part_n$ 表示该条维修数据中第 n 个部件, $mal f_m$ 表示该条维修数据中第 m 个故障现象。

(2) 故障对组合和筛选。

1) 根据步骤 1 获得标注后的维修数据, 取其中的故障对作为正样本故障对, 再随机抽取部件词和现象词进行排列组合, 去除正样本中的故障对, 得到负样本故障对。2) 基于此训练深度学习语义判别模型, 用于判断故障对组合是否合理。3) 将生成的集合 P_i 和 F_i 中的元素排列组合, 得故障对集合 $PF_i = \{part_1_mal f_1, part_1_mal f_2, \dots, part_n_mal f_m\}$, 再由判别模型判断组合后的故障对是否合理, 保留组合合理的故障对, 作为该维修记录的故障对集合 $PF'_i = \{part_1_mal f_1, part_1_mal f_2, \dots\}$ 。

(3) 故障对标准化。

1) 专家预先定义部分标准的部件词库和现象词库。2) 根据步骤 (1) 中的预训练模型, 将词库中的标准词和集合 PF'_i 中的部件词和现象词向量化表示, 再计算其语义相似度 δ , 得到每个部件词和现象词最相近的标注词。3) 当最相近标准词的语义相似度 δ 大于预设的阈值时, PF'_i 中的部件词和现象词分别用最相近的标注词代替, 若最相近词的 δ 小于阈值, 则保留, 基于此得到标准化后的故障对集合 $PF''_i = \{part'_1_mal f'_1, part'_1_mal f'_2, part'_1_mal f'_3, \dots\}$ 。

步骤 3: 获取质量管理体系全部数据, 通过实体抽取器生成首发故障判断知识库, 基于此判断故障是否首发故障, 具体包括两个步骤:

(1) 首发故障判断知识库构建。1) 知识建模, 定义知识图谱实体、关系及属性, 实体包括故障对、部件、现象。关系包括归属、并发关系。属性包括名称、权重、出现次数。2) 知识抽取, 通过步骤 2 的故障模式实体抽取器, 从质量管理体系的文本中抽取图谱所需的实体、关系及属性。3) 知识检验与存储, 对抽取出来的实体、关系及属性进行人工抽样检验, 满足要求则存入库中, 形成图谱型的首发故障判断知识库。

(2) 新维修单首发故障自动判断。1) 新维修单故障对抽取, 步骤 2 的故障模式实体抽取器, 对新的维修工单内容 (客户投诉内容, 技师检查结果, 技师维修策略等) 进行故障对抽取。2) 知识库检验新故障对, 将新故障对在知识库进行所以, 如发现不存在则判断为首发故障, 否则判断为非首发故障。

步骤 4: 将判别结果反馈首发故障判断知识库, 实现对其的动态可持续更新。在持续的监

测过程中，出现首发故障时，则通过步骤 3 的知识检验与存储，将新的故障信息更新至首发故障判断知识库，为后续判断奠定基础。

本发明的技术方案还包括一种基于人工智能的车辆首发故障判断系统，其包括两个部分，知识库生成部分和首发故障判断部分。其中，知识库生成部分包括：

（1）质量管理系统，OEM 本地维修数据存储及管理系统，与本发明系统框架相连通。

（2）数据抽取模块，用于从质量管理系统中抽取和处理维修文本数据；

（3）自动抽取器，用于从维修文本数据中提取出每条维修记录对应的故障对（部件及对应的故障现象）。整合了部件现象识别模块、组合筛选模块、实体标准化模块，具体功能与步骤 2 的各个模型对应；

（4）首发故障判断知识库，用于存储从质量管理系统中抽取的故障对，提供首发判断基础；

（5）更新维护模块，用于将每次产生的新故障信息存入知识库，包括故障部件、现象等。

首发故障判断部分包括：

（1）自动抽取器，和知识库生成部分的抽取器相同，此处不再赘述；

（2）首发判断模块，结合首发故障判断知识库，用于判断新的故障对是否是首发故障；

（3）系统交互模块，用于系统结果的显示和与用户的交互。

本发明所达到的有益效果是：本发明能快速、高效、准确的识别首发故障，减轻 OEM 人力物力的投入，提升首发故障识别的准确率和效率，为其进行产品的质量监控、制定维修策略等环节奠定基础。

五、实施示例

按自己的技术路线，划分步骤，再用“总-分”结构撰写。

撰写方式，分别用一句话总结各个大步骤，再分别用一句话总结大步骤中的小步骤，最后详细说明每个小步骤。

注：1）公开的方法不需要详细展开，如：“通过神经网络算法，对文本内容分类...”，不需要展开解释什么是神经网络算法，而是解释如何把文本处理好给算法计算，输出的结构要如何使用等。

2）经量使用数学、物理等工科语言表达技术方案，如公式、规则等。

3）方法可以尽量详细，代理机构明白方案后，会对文章结构、术语做调整，内容会模糊化。

4）术语、参数名要统一，避免自创术语，用通用术语。

注：以图、表、文结合等形式，通过例子对方案进行说明。只要能说清楚，“三、技术方案”也可不写，因为内容基本是重复的。

为了进一步描述本发明的技术特点和效果，以下结合附图和具体实施方式对本发明做进一步描述。

图 1 是本发明一种基于人工智能的车辆首发故障判断方法流程图，如图 1 所示，该方法包括如下步骤：

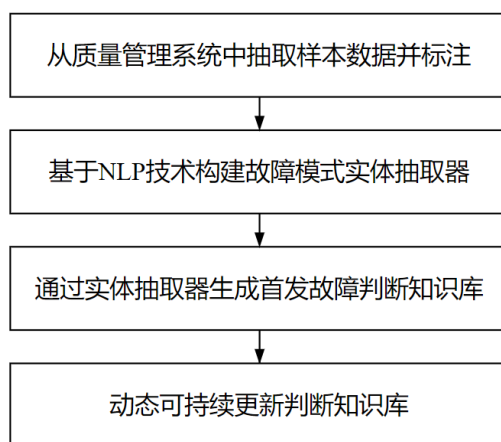


图 1

步骤 1：从质量管理系统中随机抽取部分维修记录，以此作为样本数据，在辅助标注工具的基础上进行人工标注。具体包括三个步骤：

（1）数据获取与清洗。从质量管理系统中抽取部分维修记录，即客户投诉内容，技师检查结果，技师维修策略等文本数据。并过滤乱码、错位等数据。

所述质量管理系统是 OEM 的存储每次维修详细记录的平台，包括了客户投诉内容，相应的技师检查结果和维修策略等。

（2）开发辅助标注工具。针对步骤（1）的文本格式开发辅助标注工具，能实现文本的简明显示和文本打标签。

（3）数据标注。通过辅助标注工具，对每条维修记录，标注出部件词及其对应的现象词（车辆故障现象），及故障对。部件词如：发动机、变速箱等，现象词如：漏油、抖动、异响等，故障对如：发动机—漏油，变速箱—异响等。图 2 是本实施例开发的标注工具和标注示例。

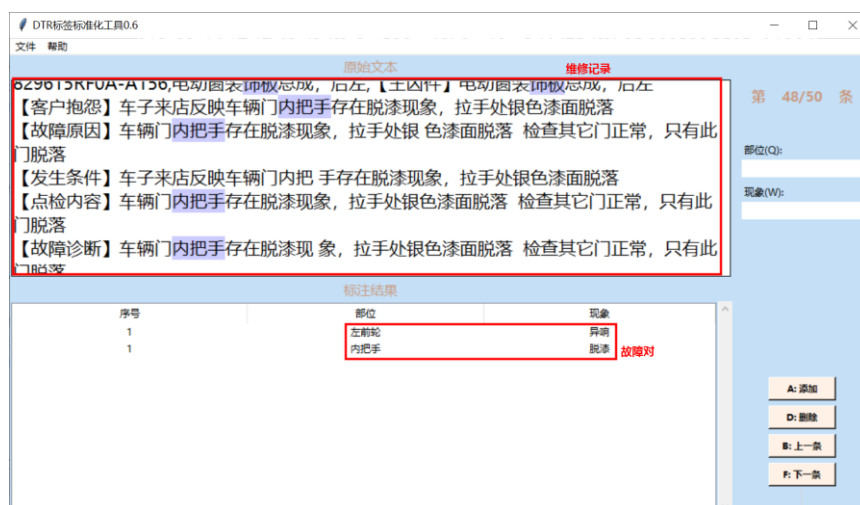


图 2

步骤 2：基于每条数据标注出的部件、现象，结合 NLP 技术构建故障模式实体抽取器，基于该抽取器抽取故障对（部件词—现象词）。构建过程具体包括三个步骤：

所属的故障模式实体抽取器包括三个模型，部件现象抽取模型，故障对组合筛选模型，实体标准化模型。

（1）识别每条数据的部件和故障现象。

1) 根据步骤 1 获得标注后的维修数据，对原始文本进行 BIO 序列标记，通过开源的预训练模型将文本转换为词向量。2) 将 BIO 序列标记和词向量结合，得到训练数据。实施例选择用 bert 预训练模型。3) 在此基础上，训练深度学习实体识别模型（部件和现象抽取模型），本实施例采用了 LSTM+CRF 模型进行构建。4) 通过模型将剩余未标注的维修数据，进行实体识别，对每条维修数据识别出部件集合 $P_i = \{part_1, part_2, \dots, part_n\}$ ，现象集合 $F_i = \{mal f_1, mal f_2, \dots, mal f_m\}$ ， i 表示第 i 条维修数据， $part_n$ 表示该条维修数据中第 n 个部件， $mal f_m$ 表示该条维修数据中第 m 个故障现象，当输入语句“外部空调低压管表面有结冰现象”，从中抽取实体 $P_i = \{\text{空调低压管}, \dots\}$ 和 $F_i = \{\text{结冰}, \dots\}$ 的过程如图 3 所示。

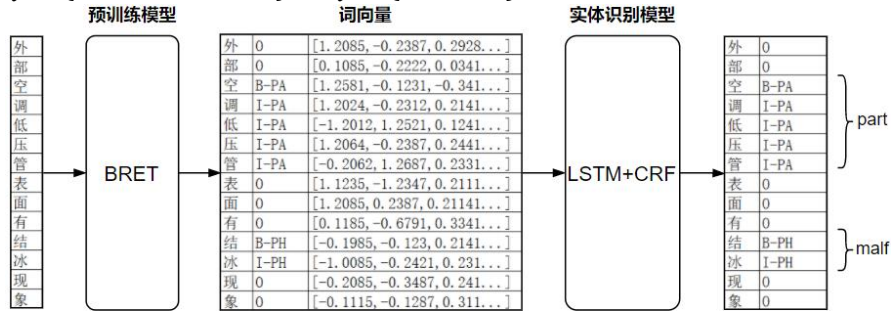


图 3

（2）故障对组合和筛选。

1) 根据步骤 1 获得标注后的维修数据，取其中的故障对作为正样本故障对，再随机抽取部件词和现象词进行排列组合，去除正样本中的故障对，得到负样本故障对。2) 将文本及其产生的组合故障作为训练数据，基于此训练深度学习语义判别模型，用语义判断故障对是否合理属于原始文本，基于此判断故障对组合是否正确。3) 将生成的集合 P_i 和 F_i 中的元素排列组合，得故障对集合 $PF_i = \{part_1_mal f_1, part_1_mal f_2, \dots, part_n_mal f_m\}$ ，再由判别模型判断组合后的故障对是否合理，保留组合合理的故障对，作为该维修记录的故障对集合 $PF'_i = \{part_1_mal f_1, part_1_mal f_2, \dots\}$ 。当输入语句“外部空调低压管表面有结冰现象而且发动机漏油”及其组合故障对： $PF_i = \{\text{空调低压管_结冰}, \text{空调低压管_漏油}, \text{发动机_结冰}, \text{发动机_漏油}\}$ ，从中识别合理对 $PF'_i = \{\text{空调低压管_结冰}, \text{发动机_漏油}\}$ 如图 4 所示。

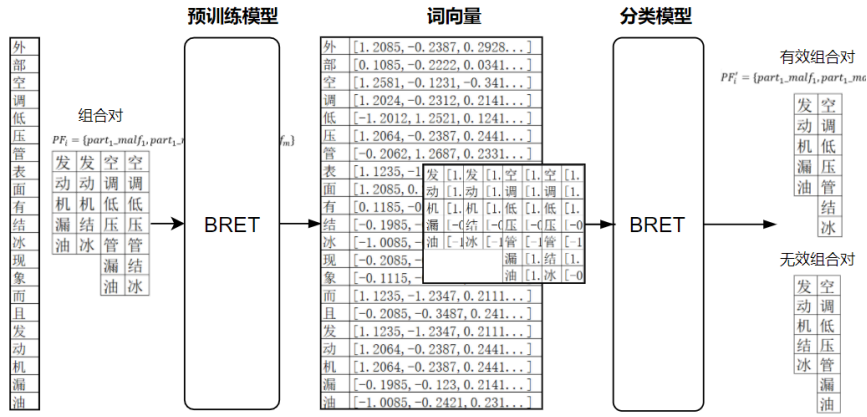


图 4

(3) 故障对标准化。

1) 专家预先定义部分标准的部件词库和现象词库。2) 将词库中的标准词和集合 PF_i' 中的部件词和现象词向量化表示，本实施例采用 word2vec 进行向量话，再通过 $\delta(X,Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$ 计算其语义相似度 δ ，得到每个部件词和现象词最相近的标注词。其中 X, Y 分布表示待比较的两个词向量， x_i 表示 X 向量第 i 个元素， y_i 表示 Y 向量第 i 个元素， n 表示向量长度。3) 当最相近标准词的语义相似度 δ 大于预设的阈值时， PF_i' 中的部件词和现象词分别用最相近的标注词代替，若最相近词的 δ 小于阈值，则保留，基于此得到标准化后的故障对集合 $PF_i'' = \{part'_1_mal f'_1, part'_1_mal f'_2, part'_1_mal f'_3, \dots\}$ 。

步骤 3: 获取质量管理体系全部数据，通过实体抽取器生成首发故障判断知识库，基于此判断故障是否首发故障，具体包括两个步骤：

(1) 首发故障判断知识库构建。1) 知识建模，定义知识图谱实体、关系及属性，实体包括故障对、部件、现象。关系包括归属、并发关系。属性包括名称、出现次数。如图 5 所示。2) 知识抽取，通过步骤 2 的故障模式实体抽取器，从质量管理体系的文本中抽取图谱所需的实体、关系及属性。3) 知识检验与存储，对抽取出来的实体、关系及属性进行人工抽样检验，满足要求则存入库中，形成图谱型的首发故障判断知识库。

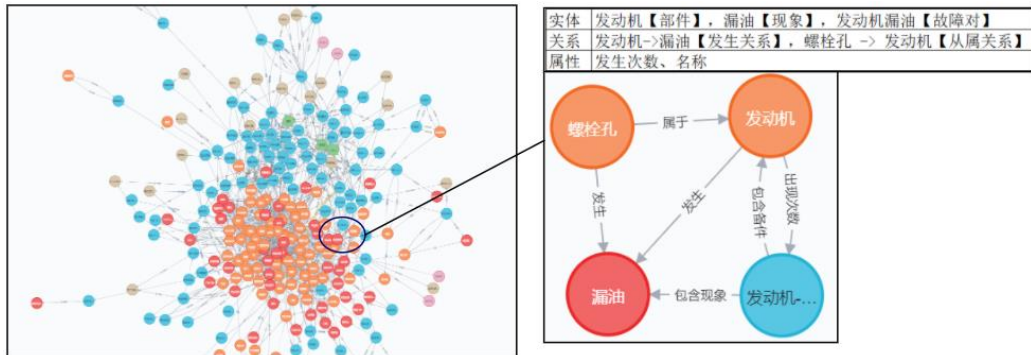


图 5

(2) 新维修单首发故障自动判断。1) 新维修单故障对抽取，步骤 2 的故障模式实体抽取器，对新的维修工单内容（客户投诉内容，技师检查结果，技师维修策略等）进行故障对抽取。

2) 知识库检验新故障对, 将新故障对在知识库进行所以, 如发现不存在则判断为首发故障, 否则判断为非首发故障。

步骤 4: 将判别结果反馈首发故障判断知识库, 实现对其的动态可持续更新。在持续的监测过程中, 出现首发故障时, 则通过步骤 3 的知识检验与存储, 将新的故障信息更新至首发故障判断知识库, 为后续判断奠定基础。

上述 4 个步骤的具体实施方案可在如图 6 所示的另一个实施例的装置中得以实现, 此系统包括两个部分, 知识库生成部分和首发故障判断部分。其中, 知识库生成部分包括:

(1) 质量管理系统, OEM 本地维修数据存储及管理系统, 与本发明系统框架相连通。

(2) 数据抽取模块, 用于从质量管理系统中抽取和处理维修文本数据;

(3) 自动抽取器, 用于从维修文本数据中提取出每条维修记录对应的故障对 (部件及对应的故障现象)。整合了部件现象识别模块、现象生成模块、组合筛选模块、实体标准化模块, 具体功能与步骤 2 的各个模型对应;

(4) 首发故障判断知识库, 用于存储从质量管理系统中抽取的故障对, 提供首发判断基础;

(5) 更新维护模块, 用于将每次产生的新故障信息存入知识库, 包括故障部件、现象等。

首发故障判断部分包括:

(1) 自动抽取器, 和知识库生成部分的抽取器相同, 此处不再赘述;

(2) 首发判断模块, 结合首发故障判断知识库, 用于判断新的故障对是否是首发故障;

(3) 系统交互模块, 用于系统结果的显示和与用户的交互。

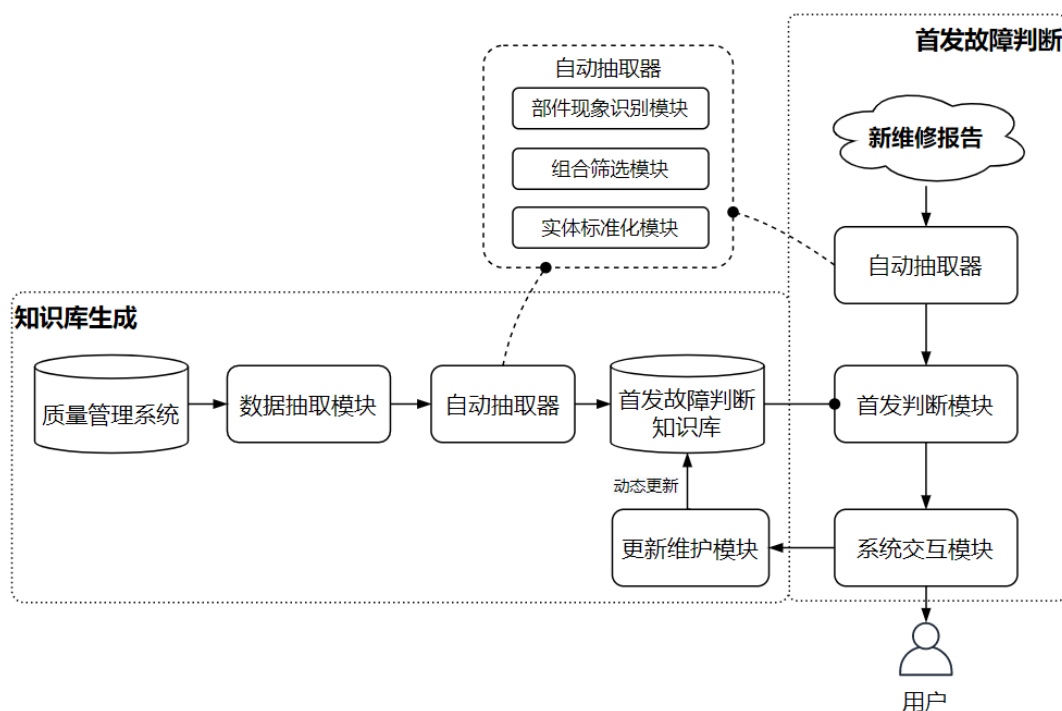


图 6

六、技术效果

同发明目的

本发明所达到的有益效果是：本发明通过引入 NLP 技术，帮助 OEM 实现高效、准确的自动化首发故障判断。充分利用非结构化的历史维修文本信息生成判断知识库，对新增维修信息实时处理，并判断是否为首发故障，为 OEM 进行产品的质量监控、制定维修策略等环节奠定基础。

七、本发明的优点

若行业有解决方案，则说明自己的技术优势；若无解决方案，则说明自己解决的问题

注：让代理机构明白即可，一般不会在最终稿中出现。

通过对比现有体系，本发明优势：

1. 基于 OEM 历史维修信息，引入 AI 领域的 NLP 技术判断首发故障，相比传统判断，不仅能实现智能决策，而且完全符合实际首发判断的工程逻辑；
2. 不仅能自动从非结构化文本中获取汽车部件和故障现象，且通过分别获取部件和故障现象再组合出故障对的形式，比传统 NLP 一次性获取方式有更高的精准率和召回率；
3. 判断方法和系统能动态、可持续运行，具有明确的反馈机制保障系统稳定性和准确性。

八、本发明的关键点和保护点

分点列举，简单说明即可，代理机构会根据内容总结出保护点

如：1.一种基于人工智能的车辆首发故障判断方法；2.的故障模式实体抽取器包括三个模型，部件现象抽取模型，故障对组合筛选模型，实体标准化模型；3.一种基于人工智能的车辆首发故障判断系统也可按本文以下例子：

1. 一种基于人工智能的车辆首发故障判断方法，其特征在于：

步骤 1：从质量管理系统中随机抽取部分维修记录，以此作为样本数据，在辅助标注工具的基础上进行人工标注。

步骤 2：基于每条数据标注出的部件、现象，结合 NLP 技术构建故障模式实体抽取器，基于该抽取器抽取故障对（部件词—现象词）。

步骤 3：获取质量管理系统全部数据，通过实体抽取器生成首发故障判断知识库，基于此判断故障是否首发故障。

步骤 4：将判别结果反馈首发故障判断知识库，实现对其的动态可持续更新。在持续的监

测过程中，出现首发故障时，则通过步骤 3 的知识检验与存储，将新的故障信息更新至首发故障判断知识库，为后续判断奠定基础。

所述步骤 1 具体包括三个步骤：

（1）数据获取与清洗。从质量管理系统中抽取部分维修记录，即客户投诉内容，技师检查结果，技师维修策略等文本数据。并过滤乱码、错位等数据。

所述质量管理系统是 OEM 的存储每次维修详细记录的平台，包括了客户投诉内容，相应的技师检查结果和维修策略等。

（2）开发辅助标注工具。针对步骤（1）的文本格式开发辅助标注工具，能实现文本的简明显示和文本打标签。

（3）数据标注。通过辅助标注工具，对每条维修记录，标注出部件词及其对应的现象词（车辆故障现象），及故障对。部件词如：发动机、变速箱等，现象词如：漏油、抖动、异响等，故障对如：发动机—漏油，变速箱—异响等。

2.所述步骤 2 所属的故障模式实体抽取器包括三个模型，部件现象抽取模型，故障对组合筛选模型，实体标准化模型。

（1）识别每条数据的部件和故障现象。

1）根据步骤 1 获得标注后的维修数据，对原始文本进行 BIO 序列标记，通过开源的预训练模型将文本转换为词向量。2）将 BIO 序列标记和词向量结合，得到训练数据。3）在此基础上，训练深度学习实体识别模型（部件和现象抽取模型）。4）通过模型将剩余未标注的维修数据，进行实体识别，对每条维修数据识别出部件集合 $P_i = \{part_1, part_2, \dots, part_n\}$ ，现象集合 $F_i = \{mal f_1, mal f_2, \dots, mal f_m\}$ ， i 表示第 i 条维修数据， $part_n$ 表示该条维修数据中第 n 个部件， $mal f_m$ 表示该条维修数据中第 m 个故障现象。

（2）故障对组合和筛选。

1）根据步骤 1 获得标注后的维修数据，取其中的故障对作为正样本故障对，再随机抽取部件词和现象词进行排列组合，去除正样本中的故障对，得到负样本故障对。2）基于此训练深度学习语义判别模型，用于判断故障对组合是否合理。3）将生成的集合 P_i 和 F_i 中的元素排列组合，得故障对集合 $PF_i = \{part_1_mal f_1, part_1_mal f_2, \dots, part_n_mal f_m\}$ ，再由判别模型判断组合后的故障对是否合理，保留组合合理的故障对，作为该维修记录的故障对集合 $PF'_i = \{part_1_mal f_1, part_1_mal f_2, \dots\}$ 。

（3）故障对标准化。

1）专家预先定义部分标准的部件词库和现象词库。2）根据步骤（1）中的预训练模型，将词库中的标准词和集合 PF'_i 中的部件词和现象词向量化表示，再计算其语义相似度 δ ，得到每个部件词和现象词最相近的标注词。3）当最相近标准词的语义相似度 δ 大于预设的阈值时， PF'_i 中的部件词和现象词分别用最相近的标注词代替，若最相近词的 δ 小于阈值，则保留，基

于此得到标准化后的故障对集合 $PF_i'' = \{part_1'malf_1', part_1'malf_2', part_1'malf_3', \dots\}$ 。

3.所述步骤 3 具体包括两个步骤：

(1) 首发故障判断知识库构建。1) 知识建模，定义知识图谱实体、关系及属性，实体包括故障对、部件、现象。关系包括归属、并发关系。属性包括名称、权重、出现次数。2) 知识抽取，通过步骤 2 的故障模式实体抽取器，从质量管理系统的文本中抽取图谱所需的实体、关系及属性。3) 知识检验与存储，对抽取出来的实体、关系及属性进行人工抽样检验，满足要求则存入库中，形成图谱型的首发故障判断知识库。

(2) 新维修单首发故障自动判断。1) 新维修单故障对抽取，步骤 2 的故障模式实体抽取器，对新的维修工单内容（客户投诉内容，技师检查结果，技师维修策略等）进行故障对抽取。2) 知识库检验新故障对，将新故障对在知识库进行所以，如发现不存在则判断为首发故障，否则判断为非首发故障。

4.本发明的技术方案还包括一种基于人工智能的车辆首发故障判断系统，其特征在于两个部分，知识库生成部分和首发故障判断部分。其中，知识库生成部分包括：

(1) 质量管理系统，OEM 本地维修数据存储及管理系统，与本发明系统框架相连通。

(2) 数据抽取模块，用于从质量管理系统中抽取和处理维修文本数据；

(3) 自动抽取器，用于从维修文本数据中提取出每条维修记录对应的故障对（部件及对应的故障现象）。整合了部件现象识别模块、组合筛选模块、实体标准化模块，具体功能与步骤 2 的各个模型对应；

(4) 首发故障判断知识库，用于存储从质量管理系统中抽取的故障对，提供首发判断基础；

(5) 更新维护模块，用于将每次产生的新故障信息存入知识库，包括故障部件、现象等。

首发故障判断部分包括：

(1) 自动抽取器，和知识库生成部分的抽取器相同，此处不再赘述；

(2) 首发判断模块，结合首发故障判断知识库，用于判断新的故障对是否是首发故障；

(3) 系统交互模块，用于系统结果的显示和与用户的交互。

九、摘要

模板：“列举技术方案的大步骤”+“解决的问题/带来的价值”

本发明公开了一种基于人工智能的车辆首发故障判断方法和系统。具体步骤包括：(1) 从质量管理系统中随机抽取部分维修记录，包括客户投诉内容，技师检查结果，技师维修策略等，以此作为样本数据，在辅助标注工具的基础上进行人工标注；(2) 由每条数据标注出的部件、现象（车辆故障现象），结合 NLP 技术构建故障模式实体抽取器；(3) 获取质量管理系统全部数据，通过实体抽取器生成首发故障判断知识库，结合业务规则判断故障是否首发故障；(4) 将判别结果首发故障判断知识库，实现对其的动态可持续更新。本发明能快速、高效、准确的

识别首发故障，减轻 OEM 人力物力的投入，提升首发故障的准确率和效率，为其进行产品的质量监控、制定维修策略等环节奠定基础。