

权利要求书

- 1、一种社交内容风险识别方法，包括：
获取待识别的社交内容数据；
提取所述社交内容数据中的社交行为特征和/或社交语料特征，并输入
5 树桩结构的机器学习模型进行处理，得到维度扩展的特征；
通过将所述维度扩展的特征输入深度机器学习模型进行处理，识别所述社交内容数据是否存在风险。
- 2、如权利要求 1 所述的方法，所述树桩结构的机器学习模型和所述深度机器学习模型预先利用存在风险的黑样本和不存在风险的白样本进行过
10 训练。
- 3、如权利要求 1 所述的方法，所述提取所述社交内容数据中的社交行为特征和/或社交语料特征，具体包括：
对所述社交内容数据进行数据清洗；
利用特征工程，从数据清洗后的所述社交内容数据中提取社交行为特
15 征和/或社交语料特征。
- 4、如权利要求 1 所述的方法，所述输入树桩结构的机器学习模型进行处理，得到维度扩展的特征，具体包括：
输入树桩结构的机器学习模型进行处理，得到多个叶节点的预测数据；
根据所述多个叶节点的预测数据，提取得到维度扩展的特征。
- 20 5、如权利要求 4 所述的方法，所述多个叶节点包括：所述树桩结构的机器学习模型处理后的最后一层叶节点。
- 6、如权利要求 1 所述的方法，所述识别所述社交内容数据是否存在风险后，所述方法还包括：
基于定时调度，定时获取并向风控平台提供社交内容数据的风险识别
25 结果。
- 7、如权利要求 1~6 任一项所述的方法，所述树桩结构的机器学习模型

包括梯度提升决策树 GBDT。

8、如权利要求 1~6 任一项所述的方法，所述深度机器学习模型包括深度神经网络 DNN。

9、如权利要求 1~6 任一项所述的方法，所述存在风险具体包括：包含
5 与指定领域相关的不当行为或者不当言论。

10、一种社交内容风险识别装置，包括：

获取模块，获取待识别的社交内容数据；

提取模块，提取所述社交内容数据中的社交行为特征和/或社交语料特征，并输入树桩结构的机器学习模型进行处理，得到维度扩展的特征；

10 识别模块，通过将所述维度扩展的特征输入深度机器学习模型进行处理，识别所述社交内容数据是否存在风险。

11、如权利要求 10 所述的装置，所述树桩结构的机器学习模型和所述深度机器学习模型预先利用存在风险的黑样本和不存在风险的白样本进行过训练。

15 12、如权利要求 10 所述的装置，所述提取模块提取所述社交内容数据中的社交行为特征和/或社交语料特征，具体包括：

所述提取模块对所述社交内容数据进行数据清洗；

利用特征工程，从数据清洗后的所述社交内容数据中提取社交行为特征和/或社交语料特征。

20 13、如权利要求 10 所述的装置，所述提取模块输入树桩结构的机器学习模型进行处理，得到维度扩展的特征，具体包括：

所述提取模块输入树桩结构的机器学习模型进行处理，得到多个叶节点的预测数据；

根据所述多个叶节点的预测数据，提取得到维度扩展的特征。

25 14、如权利要求 13 所述的装置，所述多个叶节点包括：所述树桩结构的机器学习模型处理后的最后一层叶节点。

15、如权利要求 10 所述的装置，所述装置还包括：

提供模块，在所述识别模块识别所述社交内容数据是否存在风险后，基于定时调度，定时获取并向风控平台提供社交内容数据的风险识别结果。

16、如权利要求 10~15 任一项所述的装置，所述树桩结构的机器学习模型包括梯度提升决策树 GBDT。

5 17、如权利要求 10~15 任一项所述的装置，所述深度机器学习模型包括深度神经网络 DNN。

18、如权利要求 10~15 任一项所述的装置，所述存在风险具体包括：包含与指定领域相关的不当行为或者不当言论。

19、一种社交内容风险识别设备，包括：

10 至少一个处理器；以及，

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够：

获取待识别的社交内容数据；

15 提取所述社交内容数据中的社交行为特征和/或社交语料特征，并输入树桩结构的机器学习模型进行处理，得到维度扩展的特征；

通过将所述维度扩展的特征输入深度机器学习模型进行处理，识别所述社交内容数据是否存在风险。