

AI算法在测试开发业务中的典型应用

QA中心-测试开发部

目录

背景介绍

GDBT在360文件发布系统上的应用

LSTM在自动化用例平台上的应用

总结与展望

背景

测试工具
测试框架开发



AI算法
(机器学习深度学习)



高效测试

集成学习算法的应用

文件发布系统

安全卫士文件更新和新上线发布管理系统

使用场景

QA、开发、产品、测试、审批组在该系统上对安全卫士文件发布管控

痛点

对新上线或者更新文件发布的对象以及量没法规范的给出
系统有很多无效的数据，需要QA再次审核修改，增加了发布工作量

解决思路

利用已有的历史数据和相关的机器学习算法对文件发布进行预测

集成学习算法的应用

360技术嘉年华

操作列表

创建申请

* 产品/开发收件人：

多个收件人请用分号分隔

+

* QA收件人：

多个收件人请用分号分隔

+

质量收件人：

多个收件人请用分号分隔

+

* 审批收件人：

多个收件人请用分号分隔

+

* 任务名称：

* 发布时间：

年 / 月 / 日

* 添加发布方式：

捆包正式

添加

提交申请

☒ 不需要质量组确认

放量比率及对象

设置灰度

V5新上线

删除

文件名	版本号	放量比率及对象	设置灰度	Flag	第一次捆包时间及安装包版本	修改内容	优先级	全量发布	释放路径	备注	所属模块	操作
-----	-----	---------	------	------	---------------	------	-----	------	------	----	------	----

添加文件

V5调量

删除

文件名	版本号	只调整比率或灰度	放量比率及对象	设置灰度	目前量级	目前比率	目前V5条件	修改内容	优先级	释放路径	备注	操作
-----	-----	----------	---------	------	------	------	--------	------	-----	------	----	----

添加文件

文件发布预测模型构建流程

360技术嘉年华

数据预处理

- 数据清洗
- 缺省值处理

特征工程

- 特征数值化
- 特征选择

预测模型

- 学习算法选择
- 模型构建

V5新上线

删除

文件名	版本号	放量比率及对象	设置灰度	Flag	第一次捆包时间及安装包版本	修改内容	优先级	全量发布	释放路径	备注	所属模块	操作
-----	-----	---------	------	------	---------------	------	-----	------	------	----	------	----

添加文件

V5调量

删除

文件名	版本号	只调整比率或灰度	放量比率及对象	设置灰度	目前量级	目前比率	目前V5条件	修改内容	优先级	释放路径	备注	操作
-----	-----	----------	---------	------	------	------	--------	------	-----	------	----	----

添加文件

数据预处理——缺省值

1

- 平均值、中位数、众数

2

- 其他变量预测缺失变量

3

- 该变量多一种取值——unknown

V5新上线

删除

文件名	版本号	放量比率及对象	设置灰度	Flag	第一次捆包时间及安装包版本	修改内容	优先级	全量发布	释放路径	备注	所属模块	操作
-----	-----	---------	------	------	---------------	------	-----	------	------	----	------	----

添加文件

V5调量

删除

文件名	版本号	只调整比率或灰度	放量比率及对象	设置灰度	目前量级	目前比率	目前V5条件	修改内容	优先级	释放路径	备注	操作
-----	-----	----------	---------	------	------	------	--------	------	-----	------	----	----

添加文件

特征工程

数值化

- 数值类特征：量纲统一、归一化
- 非数值类特征：数值化，唯一数字或者向量代替

选择

- 人工挑选、Pearson系数、随机森林

组合

- 特征A和特征B组合成一个特征

V5新上线

删除

文件名	版本号	放量比率及对象	设置灰度	Flag	第一次捆包时间及安装包版本	修改内容	优先级	全量发布	释放路径	备注	所属模块	操作
-----	-----	---------	------	------	---------------	------	-----	------	------	----	------	----

添加文件

V5调量

删除

文件名	版本号	只调整比率或灰度	放量比率及对象	设置灰度	目前量级	目前比率	目前V5条件	修改内容	优先级	释放路径	备注	操作
-----	-----	----------	---------	------	------	------	--------	------	-----	------	----	----

添加文件

预测模型

360技术嘉年华

基于
GDBT的
文件发布预测模型

SVM、决策树、集成学习算法

定位问题（分类）

文件发布预测服务搭建

文件

文件发布
预测模型

文件放量对象、
比率、灰度值

现存问题

新的未知文件
模型精度不佳
放量对象: 78%
放量比率: 63%
放量灰度值:

95%

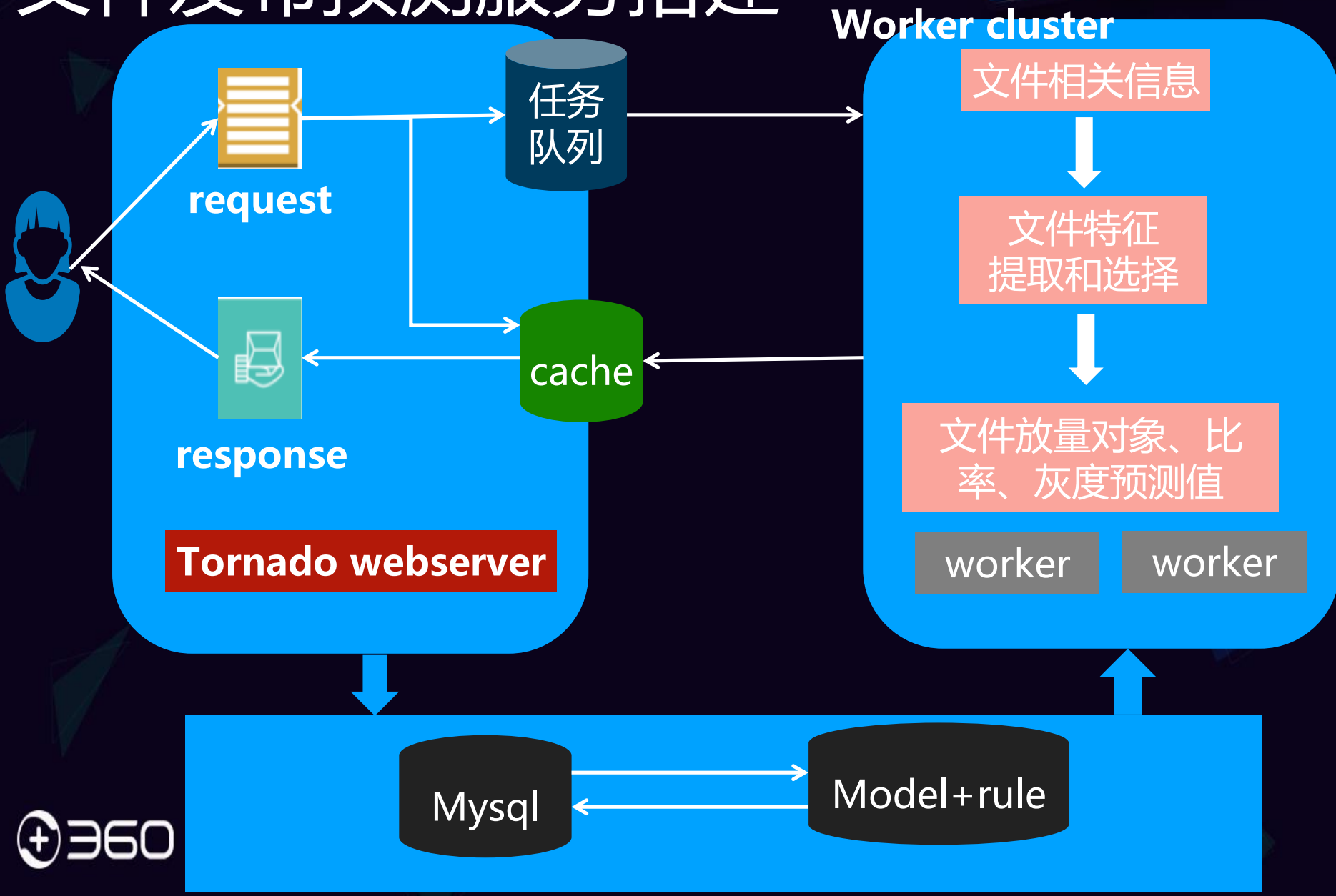
文件

文件发布
预测模型
+
规则

文件放量对象、
比率、灰度值

文件发布预测服务搭建

360技术嘉年华



LSTM模型的应用

自动化用例平台

帮助测试人员编写自动化测试用例，管理自动化用例

使用场景

编写测试用例、参考已有用例

痛点

已有函数大约有2000个
新的用例需要参考原有的用例

解决思路

用例是时间序列
N-gram, CBOW, RNN, LSTM

用例步骤推荐模型构建流程

360技术嘉年华

数据预处理

- 数据清洗
- 用例步骤标识

训练数据

- 词典构建
- 语料生成

预测模型

- 学习算法选择
- 模型构建

数据预处理

1

- 去除脏用例（废弃）

2

- 统计函数使用频率

3

- 去除低频函数（低频函数都认为是unknown）

4

- 函数数值化，用唯一数值代表具体函数

训练数据生成——词典构建

360技术嘉年华

原始函数集

[123,234,235,238,543,567,432,unknown]

排序后的函数集

[123,234,235,238, 432,543,567,unknown]

映射后词典

{123:0,234:1,235:3,238:4,432:5,543:6,'unknown':7}

训练数据生成——训练语料

360技术嘉年华

原始用例集

[25839, 25881, 26264, 25720, 25882, 25816, 25761, 25793, 25793,
25761, 25779, 26274]



映射用例集

[2541, 2559, 2601, 2487, 2560, 2529, 2502, 2520, 2520, 2502, 2514,
2609]



训练语料

[2541, 2559, 2601, 2487, 2560],[2559, 2601, 2487, 2560, 2529].....

预测模型

360技术嘉年华

基于
LSTM的
用例步骤推荐模型

N-gram,cbow,rnn,lstm.....

定位问题（分类、时序类问题）

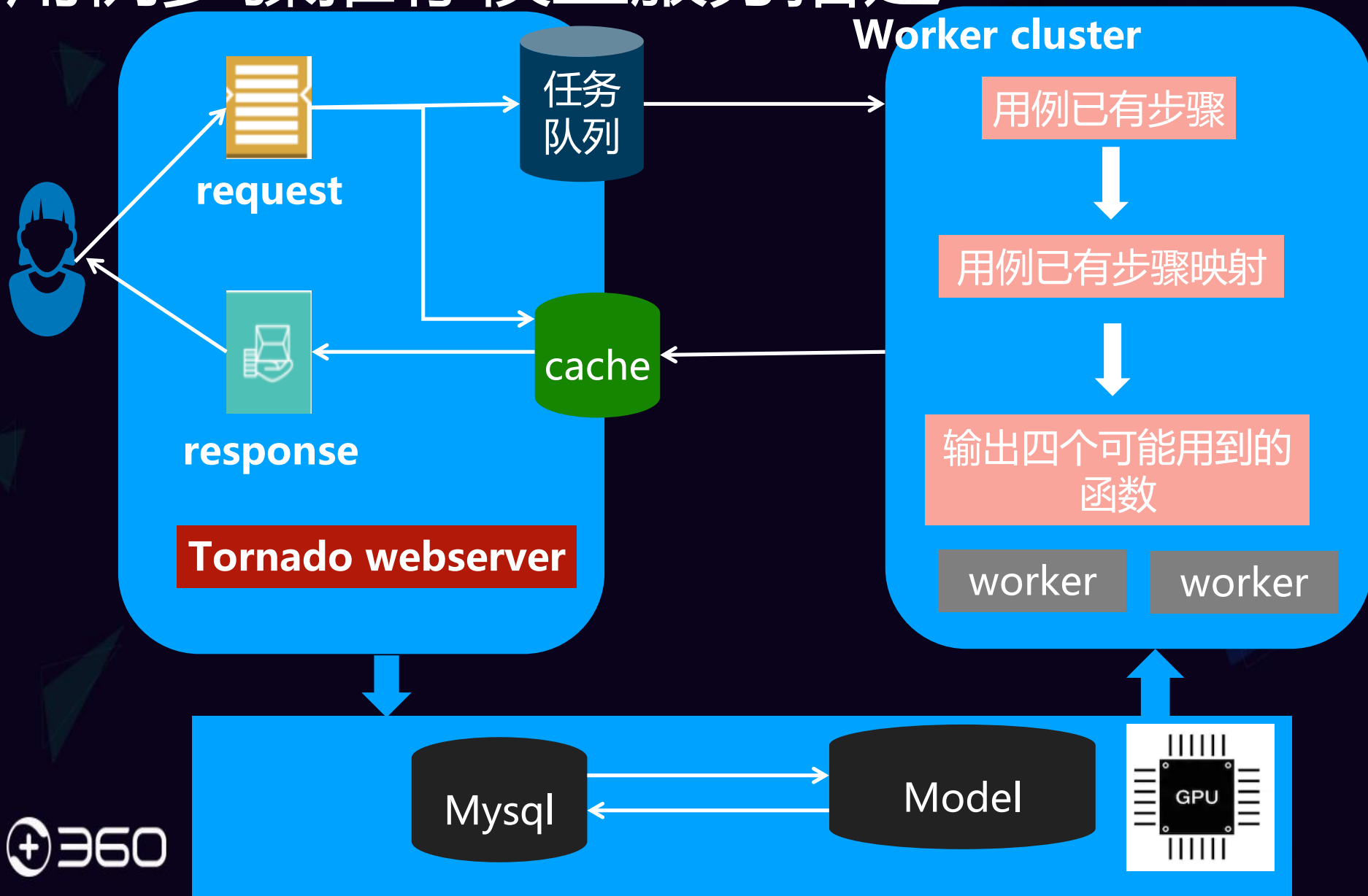
模型精度

数据集	训练集	测试集
样本数	202970	40594

推荐函数个数	N-gram	CBOW	LSTM
1	0.5437	0.6964	0.6768
2	0.6728	0.7721	0.7564
3	0.7509	0.8497	0.8319
4	0.7868	0.8862	0.8757

用例步骤推荐模型服务搭建

360技术嘉年华



总结

360技术嘉年华

- AI算法使用顺序：迁移学习、机器学习、深度学习
- AI算法+规则
- AI算法搭建的服务形成闭环，需要依赖线上数据不断训练和优化

展望

360技术嘉年华

- 不断优化文件发布预测和用例步骤推荐模型
- 基于知识图谱的自动化用例生成

谢谢！