**PowerAI人工智能马拉松编程大赛活动详介**

**活动名称**

PowerAI 人工智能马拉松编程大赛 —— 探索金融语音大数据的商业价值！

**时间地点**

2017年6月17日，上海浦东杨高南路729号1号楼41F

**参赛要求**

- 本次活劢针对泛金融领域工作者，包括银行、保险、金融、基金、证券、投资等企业内部开发者，以 及为上述企业提供软件外包不系统集成服务的企业或个人

- 从事深度学习，机器学习，人工智能等相关技术领域的研究或研发工作

参赛者需要

- 个人或团队形式参不本次活劢 （最好是团队形式，建议3-5个人一个组）

- 自行携带笔记本电脑

- 12小时现场编写程序代码

- 基亍大赛组委会提供的统一训练样本数据进行分类

大赛组委会为你准备

- 免费的开发环境，ML/DL的架构和库

- 在比赛现场可免费使用的训练样本数据

**比赛整体流程**

报名参赛>>赛前集训>>现场编程>>提交作品>>评审颁奖

**奖项设置**

-一等奖1名 奖金10000 RMB

-二等奖2名 奖金各8000 RMB

-三等奖3名 奖金各5000 RMB

**注意事项**

1. 通过Caffe、Tensorflow、Torch and Theano 进行模型训练， 鼓励发挥GPU on Power硬件特性；

2. 要求参赛者编写神经网络模型实现考题需求。参赛者可以修改Caffe，Torch, Theano或者Tensorflow源代码， 但是必须在修改后，提供代码修改说明以确定没有违规部分。参赛者也可以使用github上开源的不同发行版的Caffe，Torch，Theano或者Tensorflow， 但是必须自己解决在Power上的依赖和编译问题；

3. 本次比赛涉及金融领域对话文本分类问题。希望选手通过深度神经网络能够对客户感兴趣购买的保险或者理财产品的类型和意愿进行分类。语料中会有“肯定”，“疑问”，“否定或者双重否定”语句。传统的词频或者简单的SVD方式很难在测试数据集中有良好的表现；

4．组委会统一准备PowerAI环境以及sklearn，XGBoost，CRF++，Keras；

5. 作品提交内容包括：

a) 设计说明: 设计的神经网络说明

b) 修改代码说明(如果有)

c) 设计的训练模型文件, 部署模型文件和训练好的权重文件

d) 如果修改了代码或者使用了其他发行版, 需要提供完整的编译好的目标代码和源代码

e) 测试数据集的输出结果文件.格式为Jason (任何形式的手动修改输出结果都视为违规)

f) 测试数据集的总消耗时间日志文件(使用time 来运行test程序, 以real时间为准)

**题目说明**

1.大赛会提供训练样本数据.数据文件分为两部分, 训练和测试部分

a) 对话内容文件（文件格式可能存在转码需求，选手需要考虑）对话内容分为10+类

b) 对话内容标签文件，包括理财产品分类和意愿分类（肯定，否定，疑问）

c) 语料文件-文件格式为txt格式。内容包括中文和标点符号。文章内容是金融相关的文章。文件总量在50M左右。参赛团队可以利用这些文件来建立自己的语言模型，分词模型等

d) 训练数据：每个分类的样本个数不少于150条。数据样本间有一定区分度

2. 参赛者需要根据训练文件进行购买产品分类和意愿分类；

3. 大赛鼓励参赛选手自己构建分词和语言模型，如果选手选择使用第三方分词或者语言模型需要自己解决平台依赖；

**评选标准**

1. 实现对话内容分类（金融产品和意愿）功能（25分）

2. 正确率评分标准 （Precision）(40分）

金融产品分类对测试Top1正确率：分为5档，档位为95%以上正确率20分，85%以上15分，80%以上10，70%以上5，其他0分；

客户意愿分类测试Top1正确率：分为5档，档位为95%以上正确率20分，85%以上15分，80%以上10，70%以上5，其他0分；

3.性能评分标准: 25分

取前5名, 档位为第一名 25分, 第二名 20分, 第三名15分,第四名10 第五名 5分, 其他0分

4. 附加分(10分)

a) 选手根据大赛提供语料库训练分词系统并应用于后面的文本分类5分；

b) 选手使用深度学习框架生成词向量病应用于后面的文本分类5分；

|  |  |
| --- | --- |
| **比赛现场流程** |  |
| 8:00-8:30 | 参赛选手入场，签到 |
| 8:30-8:40 | 嘉宾致辞 |
| 8:40-9:00 | 大赛规则说明，赛前讲解，Q&A |
| 9:00 – 21:00 | 编程比赛 |
| 21:00-21:30 | 作品评选  团队上台轮流展示 |
| 21:30-22:00 | 评选结果公布，评委点评，嘉宾颁奖 |
| 22:00-22:20 | 合影留念，颁发证书 |
|  |  |
|  |  |
|  |  |