

# 机器学习实验报告

## 一、实验基本信息

- 1、代码来源: github
- 2、网址: <https://github.com/pig6/wxrobot>
- 3、项目名称: 基于 python 的聊天机器人
- 4、作者: pig6

## 二、实验目的

### 1、研究问题

通过训练一个聊天机器人, 包括它的性格、年龄特征, 来实现人机交互自动回复消息, 其中机器人回复的消息应该语义通顺、逻辑正确, 甚至是幽默风趣而并非是无聊沉闷的类似于“好的、是吗”等安全回答。

### 2、研究意义

聊天机器人是自然语言处理的发展, 目前的聊天机器人可以做到在各行各业发挥作用, 比如情绪安抚、客服咨询, 这些功能都是依赖于神经网络、语义识别和深度学习等。未来的聊天机器人应该可以通过互联网进行自动学习和获取知识, 有语义化的理解, 并且逐步达到人脑的水平, 本项目是探究不同模式的训练会导致机器人回复消息的差异性。

## 三、基本原理

基于中文分词、开源的搜索引擎和模式识别技术等, 再加上一些语句筛选和过滤算法、命令模块, 提取词语并进行匹配。缺点就是语义资料库不够多。

## 四、实验步骤


- 1、下载代码的 zip 压缩包
- 2、在 pycharm 中载入 wxpy 等一系列工具包

```
E:\pycharm\wxrobot-master>pip3 install -U wxpy -i "https://pypi.doubanio.com/simple/"  
Looking in indexes: https://pypi.doubanio.com/simple/
```

### 3、进行消息测试


- 1) 首先使用未进行训练的机器人得到的消息回复如下图


 冷  
什么鬼

 知世  
这话我接不了呢

 冷  
这个检测不出来了

 知世  
聊点别的吧

 冷  
这不就是自动回复吗

 知世  
这话我接不了呢

可以看出此时机器人（知世）只会回答一些安全答案，并不能根据上下文进行分析

2) 将机器人训练为成年人 20 岁的性格

直接在 pycharm 中测试

测试 A、

```
data = {'key': apikey, 'info': '你好'}
```

我的好朋友你好。

测试 B、

```
data = {'key': apikey, 'info': '你的爱好是什么'}
```

我喜欢音乐、书籍、电影，是不是很有文艺范呢。

3) 将机器人训练为成年人 40 岁的性格

测试 A、

```
data = {'key': apikey, 'info': '你好'}
```

嗯...又见面了。

测试 B、

```
data = {'key': apikey, 'info': '你的爱好是什么'}
```

我喜欢音乐、书籍、电影，不知道你是不是同道中人呢。

#### 4) 将机器人训练为儿童 15 岁的性格

测试 A、

```
data = {'key': apikey, 'info': '你好'}
```

```
你好呀，我的好朋友。
```

测试 B、

```
data = {'key': apikey, 'info': '你的爱好是什么'}
```

```
嘿嘿嘿，当然是喜欢和宝宝在一起啦，这是我喜欢做的事呢。
```

#### 5) 结果分析

我们可以看到，不同的年龄性格的机器人对于同一个问题给出的回复是不一样的，40 岁的成年人明显比 20 岁要成熟稳重，但是在儿童模式下时候他说喜欢和宝宝在一起，有些偏向于母亲的视角，可见语义分析还不够完善。

### 五、实验结论与启发

该项目的 wxpy 工具包中聊天机器人已经有了一定的训练程度，但是还是有些时候欠缺语义和上下文的分析，在此基础上还可以增加语料库的规模，或者通过爬虫让其回答的消息更加人性化。

机器人其实是一个搜索引擎，但是如果可以自行通过搜索网络信息进行学习，它的功能会更加强大，能够达到思维近于人脑但是学习能力超过人脑的水平。