1. 系统概念模型

说明：

1. WebCrawler为数据爬取子系统的数据爬取类，主要属性有：

（1）UrlString用于匹配输入URL的字符串的值

（2）RegExCode系统内置的对应URL的正则表达式

2． Comment、Restaurant、Category分别为数据处理子系统将原始数据处理完毕后，存储到数据库中的三个类。ComUser类为评论者类，由以上几个类计算出每个评论者发布过的评论数量。

（1）Comment类为某一餐馆用户评论类，主要属性有：评论者的ID、餐馆ID、评论ID、评论者给出的星级评价、评论者给出的口味评分、环境评分、服务评分、人均花费和评论内容。

（2）Reestaurant类为餐馆类，主要属性有：餐馆ID、餐馆名称、餐馆分类ID、总体星级评价、总体口味评分、总体环境评分、总体服务评分、总体人均消费、地址与评论数。

（3）Caategory类为餐馆类型类，主要属性有：类别名称和分类ID。

（4）ComUser类为评论者类，主要属性有：评论者的ID和该评论者发布过的评论数。

1. 类的职责

|  |  |
| --- | --- |
| **类名** | **职责** |
| WebCrawler | 1. 匹配输入的URL连接并返回对应的提示 2. 匹配正确，则爬取对应网页的数据 3. 获取系统内对应URL的正则表达式RegExCode并处理原始数据，在数据库中储存成若干个表（对应以下几个类） |
| Comment | 该类主要储存某一餐馆评论相关数据，主要属性有：评论者的ID、餐馆ID、评论ID、评论者给出的星级评价、评论者给出的口味评分、环境评分、服务评分、人均花费和评论内容。 |
| Restaurant | 该类主要储存某一餐馆相关数据，主要属性有：餐馆ID、餐馆名称、餐馆分类ID、总体星级评价、总体口味评分、总体环境评分、总体服务评分、总体人均消费、地址与评论数。 |
| Category | 该类主要储存餐馆类型，主要属性有：类别名称和分类ID。 |
| ComUser | 该类主要提供方法计算某一评论者发布的评论数，以供数据分析使用，主要属性有：评论者的ID和该评论者发布过的评论数。 |

1. 系统流程定义



1. 系统行为定义

