## 功能实现：

通过使用request模块，使用os模块进行本地操作

**功能逻辑：**

对列表页进行遍历，爬取所有帖子的链接，并对所有链接内的帖子进行下一步爬取。

对于帖子内：

使用正则表达式re.findall('<img src="/upload/([\s\S]\*?)"', r.text)， 得到文中图片的路径，进行下载。使用getImg 函数 进行保存图片。

## 最终效果：

**运行窗口效果：**

文本

描述已自动生成

**爬取图片数据：**

图形用户界面, 应用程序

描述已自动生成

## 肉眼分析得出结论：

根据帖子标签得出结论：文学占比 20%, IT技术为18%，科幻为9% 其他为

53%.