**网页用http解析不了的时候，用jsoup解析**

**尽量使用select选择标签**

**Jsoup访问的时候遇到网页js后期跳转的见ieee**

1. Document document\_keywords = Jsoup.connect(url\_keywords).
2. header("Accept", " text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,\*/\*;q=0.8").
3. header("Accept-Encoding", "gzip, deflate, br").
4. header("Accept-Language", "zh-CN,zh;q=0.9,en;q=0.8").
5. header("Cache-Control", "max-age=0").
6. header("Connection", "keep-alive").
7. header("Host", "ieeexplore.ieee.org").
8. header("Upgrade-Insecure-Requests", "1").
9. header("User-Agent", " Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/67.0.3396.99 Safari/537.36").
10. timeout(10000).get();

**原因：首次进入的页面不加载具体信息，在文件中查找到包含信息的文件，通过request url 和 设置 header 进行访问。**

**Git的问题（还没解决）**

**Python Django（周末解决Django）**