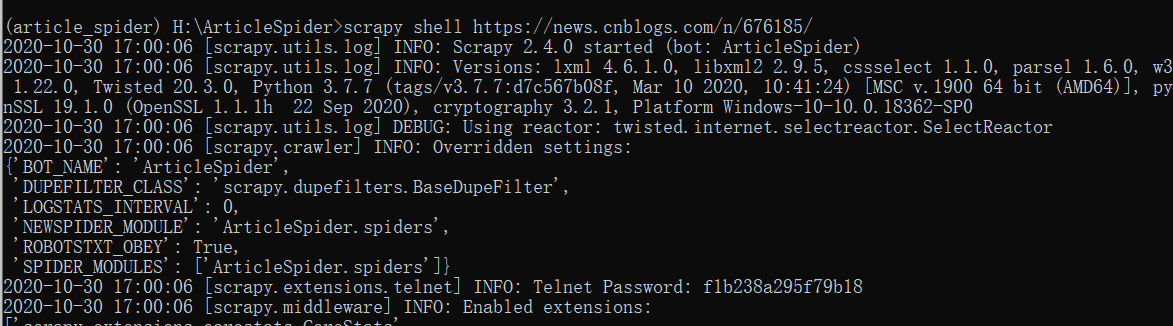
需要用到yield生成器将for里面的字段迅速交出去

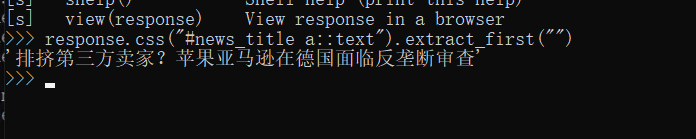
获取博客园的新闻字段

Scrapy是一个深度优先

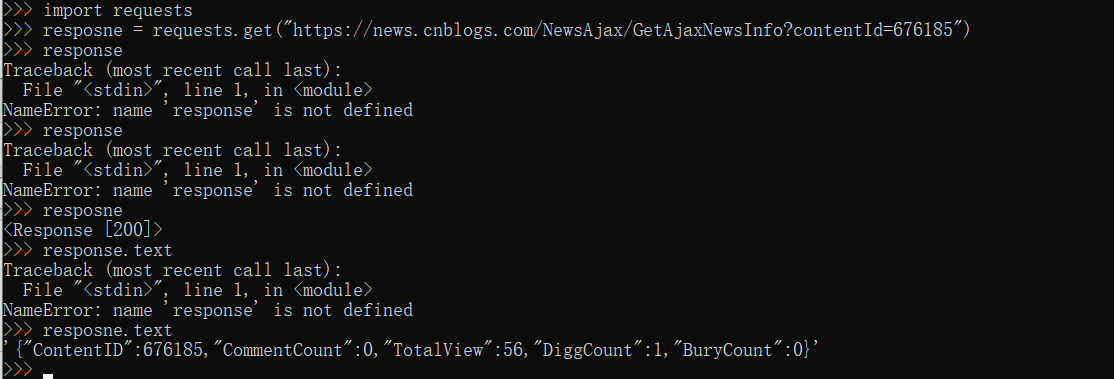
会使用scrapy shell www.

scrapy shell <https://news.cnblogs.com/n/676185/>





用Request来测试



用requests

post\_id = match\_re.group(1) # 此处提取post\_id来给下面的的id赋值  
html = requests.get(parse.urljoin(response.url, "NewsAjax/GetAjaxNewsInfo?contentId={}".format(post\_id)))

#注意因为requests是同步的而这里用的是一个异步的方法，所以换异步的方法实现

用yeild方法来吧这个同步的requests做出异步处理