019 爬取在线课程并存入MySQL数据库

* 核心技术

**将爬取的网易云课程信息存入数据库，首先需要对课程数据进行分析，然后根据课程数据信息创建对应的数据表中的字段。最后，将爬取的课程数据写入到MySQL数据库。**

**在本例中，核心技术是使用PyMySQL模块操作数据库，主要流程如下：**

* 连接数据库

**使用PyMySQL的connect()方法连接数据库，示例代码如下：**

1. conn = pymysql.connect(host='localhost',
2. port=3306,
3. user='root',
4. passwd='root',
5. db='flask',
6. charset='utf8')
7. cur = conn.cursor()

* 查询数据

**使用PyMySQL的execute()方法可以执行SQL语句，例如**根据课程ID查询数据，示例代码如下：

1. # 根据course\_id查找course表中记录
2. sql = f'select course\_id from course where course\_id = **{**course\_id**}**'
3. # 执行SQL语句
4. cur.execute(sql)
5. # 查找一条记录
6. course = cur.fetchone()

* 批量写入数据库

**每次写入一条记录效率较低，使用批量写入数据库可以提高运行效率。PyMySQL的executemany()方法可以批量写入数据库，**示例代码如下：

1. sql\_course = """insert into course
2. values (%s,%s,%s,%s,%s,%s,%s,%s,%s,%s,%s,%s,%s,%s,%s,%s)
3. """
4. # exetemany方法插入多条记录
5. cur.executemany(sql\_course, course\_data)

* 提交操作

**写入操作完成后，需要将其提交到数据库，可以使用conn连接对象的commit方法提交数据。示例代码如下：**

conn.commit() # 提交到数据库

* 关闭连接

**操作完成以后，关闭游标，关闭连接，示例代码如下：**

1. cur.close() # 关闭游标
2. conn.close() # 关闭数据库连接