

# 浙江大学

## 暑期深度学习实验报告

实验名称: MNIST 数字图像分类

小组成员: 谢 集 322010xxxx

李浩浩 3220105930

赵一帆 322010xxxx

杨正宇 322010xxxx

课程名称: 课程综合实践 I

指导老师: 杨 易, 范鹤鹤

# 1. 数据清洗

## 1.1 清洗过程

使用python的pandas进行数据预处理，代码如下

```
import pandas as pd
import numpy as np

df = pd.read_excel('2021 Global Gas Flaring Volumes.xlsx', sheet_name=None)

data = df['2021 Global Gas Flaring Volumes']

# 除去缺失值
data.dropna()

# 除去重复值
data.drop_duplicates()

# 保存到新文件中
data.to_excel('washed_data.xlsx', index=False)
```

## 1.2 清洗结果