

Universidade do Minho

Escola de Engenharia

Inteligência Ambiente: Tecnologias e Aplicações

Questão de aula nº1 MiEI - 4^{o} Ano - 1^{o} Semestre

A84167 Susana Marques

 ${\rm Braga},$ 15 de dezembro de 2020

Conteúdo

1	Pré-Processamento	2
2	Resolução	3

1 Pré-Processamento

A documentação do código deste trabalho explica ao pormenor todas as decisões que foram tomadas ao longo do mesmo, por isso é aconselhável que durante a leitura deste relatório se preste atenção aos comentários do código mencionado neste.

Este trabalho foi realizado em Python e os dados do ficheiro anexo.csv foram lidos para um dataframe onde depois foram colocados adicionais dados e os outputs pedidos.

Optou-se por remover linhas duplicadas do dataset e alterar a percepção de como o ficheiro é lido, como se poderá ver pelo seguinte código:

```
#!/usr/bin/env python
# -*- coding: utf-8 -*-
import pandas as pd
import logging
import datetime

#leitura do ficheiro .csv para o dataframe df
df= pd.read_csv('anexo.csv')

#remove linhas duplicadas
df.drop_duplicates()

#altera o display para uma melhor percepcao
#de todo o ficheiro de logs
pd.set_option('display.max_rows', None)
pd.set_option('display.max_columns', None)
pd.set_option('display.width', 1000)
```

A seguir, no final do pré-processamento de dados criou-se o ficheiro de logs que posteriormente irá receber o dataframe criado com toda a informação:

```
#adicao de todo o dataframe (toda a informacao contida
#no csv, mais toda a informacao obtida neste SBR)
#ao ficheiro de logs
logging.info(df)
```

2 Resolução

O ficheiro de logs será constituído por todo o dataset que nos foi dado (anexo.csv) em conjunto com todos os outputs pedidos e informação intermédia de modo a uma mais fácil visualização dos resultados da aplicação das regras estabelecidas e uma fácil perceção de como os resultados dos outputs foram obtidos:

```
##CRIACAO DAS COLUNAS ADICIONAIS DO FICHEIRO DE LOGS##

#criacao de um array que determinara a estacao a
#que pertence cada linha
estacao =[]

#criacao de arrays que apresentam o resultado
#do calculo da temperatura de conforto e da humidade
#atrav s da esta o
temp_conforto = []
humid_conforto = []

#criacao de array que apresentam o output pedido
#da temperatura e da humidade consoante os valores calculados
output_temp =[]
output_humid =[]
```

Para se conseguir descobrir a estação do ano, recorreu-se à seguinte função que através do parsing da data que se encontra no dataset e do hemisfério (que é calculado pela latitude que também se encontra no dataset) consegue verificar a estação. O cálculo do hemisfério é útil uma vez que apesar do dataset apenas ter cidades de Portugal, se escalarmos é algo possível de se considerar:

```
# calcula a esta o atrav s da data e do hemisf rio em que se
                                  encontra
def estacaoo(dataa, hemisferio):
   md = dataa.month * 100 + dataa.day
   if ((md > 320) and (md < 621)):
        s = 0 \#primavera
    elif ((md > 620) and (md < 923)):
        s = 1 #verao
    elif ((md > 922) and (md < 1223)):
       s = 2 #outono
    else:
        s = 3 \#inverno
   if hemisferio != "norte":
       if s < 2:
           s += 2
        else:
   return s
for value, value1 in zip(df["dt_iso"],df["lat"]):
   datee = datetime.datetime.strptime(value,"%Y-%m-%d %H:%M:%S")
   #calcula o hemisf rio onde se encontra atrav s da
   #coordenada altitude
   if value1 >=0:
       hemisferio = "norte"
   else:
```

```
hemisferio ="sul"
#calcula a esta o
s=estacaoo(datee,hemisferio)
if s==0:
    est="Primavera"
elif s==1:
    est="Verao"
elif s==2:
    est="Outono"
else:
    est="Inverno"
```

No final adicionamos a estação obtida ao dataframe criado:

```
#adiciona a estacao ao array de estacao criado antes
estacao.append(est)

#adiciona o array ao dataframe df como coluna
df["Estacao"]=estacao
```

A seguir, depois de uma exausta pesquisa calculou-se a temperatura de conforto para o dia e para a noite (considerando noite entre as 00:00:00 e as 06:00:00) uma vez que a nível de despezas monetária a temperatura de conforto também pode ser alterado durante a noite e a humidade através da estação que se obteve anteriormente:

Estação	Temperatura durante o dia (graus)	Temperatura durante a noite (graus)	Humidade (%)
Primavera	23	25	39
Verão	26	28	45
Outono	20	18	33
Inverno	17	15	27

Não faz sentido calcular a temperatura com referência apenas à humidade nem vice-versa uma vez que as duas variáveis se interligam e ao alterarmos um o outro também é alterado.

O código seguinte mostra a resolução para este problema:

```
#calcula o tempo de conforto e a temperatura de conforto
#atrav s da esta o
```

```
for val, estac in zip(df["dt_iso"],df["Estacao"]):
   time = datetime.datetime.strptime(val,"%Y-%m-%d %H:%M:%S")
   #tempo que ocorre naquela dia da data
   timee = datetime.time(time.hour,time.minute,time.second)
   start = datetime.time(23,0,0)
   end = datetime.time(6,0,0)
   #verifica qual estacao
   if estac == "Primavera":
        #verifica se
                       de noite
        if time_in_range(start, end, timee):
           temp_conforto.append(25)
           humid_conforto.append(39)
           temp_conforto.append(23)
           humid_conforto.append(39)
   elif estac == "Verao":
        if time_in_range(start, end, timee):
            temp_conforto.append(28)
           humid_conforto.append(45)
            temp_conforto.append(26)
           humid_conforto.append(45)
   elif estac == "Outono":
        if time_in_range(start, end, timee):
            temp_conforto.append(18)
           humid_conforto.append(33)
            temp_conforto.append(20)
           humid_conforto.append(33)
   elif estac == "Inverno":
        if time_in_range(start, end, timee):
            temp_conforto.append(15)
           humid_conforto.append(27)
            temp_conforto.append(17)
            humid_conforto.append(27)
# adicao dos arrays de temperatura de conforto e
#humidade de conforto j criados e preenchidoas
#acima ao dataframe que estamos a utilizar como colunas
df["Temp_conforto"] = temp_conforto
df["Humid_conforto"] = humid_conforto
```

Finalmente foi realizado o cálculo dos outputs tendo em conta a estação e se é noite ou dia para a humidade e para a temperatura:

```
#adicao do array de output da temperatura ao dataframe como coluna
df["Output_Temp"]= output_temp

#calculo do output que devera ser colocado na coluna
#de outputs de humidade do ficheiro logs
for val1, val2 in zip(df["humidity"],df["Humid_conforto"]):
    if val1 > val2:
        diferenca = val1-val2
        output_humidade = "HUMIDIFYING -{:.2f} %".format(diferenca)
    elif val1 < val2:
        diferenca = val2-val1
        output_humidade = "HUMIDIFYING +{:.2f} %".format(diferenca)
    else: output_humidade = "NO NEED TO CHANGE"
    output_humid.append(output_humidade)

#adi o do array de output da humidade ao dataframe como coluna
df["Output_Humid"]= output_humid</pre>
```

Finalmente, como já referido antes o ficheiro de logs recebe o dataframe com toda a informação:

```
#adi o de todo o dataframe (toda a informa o contida
#no csv, mais toda a informa o obtida neste SBR)
#ao ficheiro de logs
logging.info(df)
```

Podemos ver o resultado final num print de uma pequena porção do ficheiro de logs:

