

Actividad

Poryecto Regresion

Table of Contents

Actividad.....	1
Poryecto Regresion.....	1
Fecha:.....	1
Objetivos:.....	1
Nombre:.....	2
Repository:.....	2
Librarys:.....	2
Paso 0: Descartar cualquier cambio realizado en el repositorio clonado.....	2
Paso 1: Limpiar variables y linea de comandos.....	3
Paso 2.- Configuración de carpeta ./src para librerias.....	3
Paso 3- Configuración de carpeta de ./data para datasets.....	3
Paso 4- Desde Open Energy Data Initiative, seleccionar un archivo CSV cualquiera de cualquier estado.....	3
Informacion del dataset.....	3
Paso 4- Buscar los nombres y Cargar los datos de todos los archivos dentro de la carpeta ./data.....	4
Paso 5- Extraer nombres de variables y crear datetime.....	7
Paso 6- Graficar todas las variables.....	8
Paso 7- Definir la variable de salida y las variables de entrada del sistema.....	9
Paso 9- Cambiar la resolucion temporal de los datos: Dias, Semanas, meses.....	10
Paso 8- Seleccionar Variables o caracteristicas.....	11
Paso 9- Seleccionar el algortimo de ML con un menos error de prediccion empleando el toolbox de Matlab Regression Learner.....	13
Paso 10- Dividir el dataset en 70% para entrenar y 30% validar.....	14
Paso 11- Usando el algoritmo de ML se entrena el modelo de regression (costo computacional).....	16
Paso 12- Cargar y validar el modelo entrenado.....	16
paso 13 - Graficar el valor predecido vs el valor real.....	17
paso 14 - Mejorar el modelo de prediccion.....	17

Fecha:

```
fecha = datetime('now', 'Format', 'dd-MM-yyyy');
disp(['Fecha actualizada: ', char(fecha)])
```

Fecha actualizada: 06-07-2024

Objetivos:

- Desde Open Energy Data Initiative, seleccionar un archivo CSV cualquiera de cualquier estado: https://openei.org/datasets/files/961/pub/COMMERCIAL_LOAD_DATA_E_PLUS_OUTPUT/
- Definir la variable de salida y las variables de entrada del sistema
- Definir basado en alguna aplicacion, Cambiar la resolucion temporal de los datos: n Dias (**Ajustar nhoras**)
- Seleccionar Variables o caracteristicas empleando la matriz de correlacion (**Ajustar el threshold**)
- Seleccionar el algortimo de ML con un menos error de prediccion empleando el toolbox de Matlab Regression Learner

- Dividir el dataset en 70% para entrenar y 30% validar (**Ajustar el PSplit**)
- Usando el algoritmo de ML se entrena el modelo de regression
- Cargar y validar el modelo entrenado
- Graficar el valor predecido vs el valor real
- **Analizar el resultado obtenido y tratar de reducir el error**

Nombre:

- sunombre

Repository:

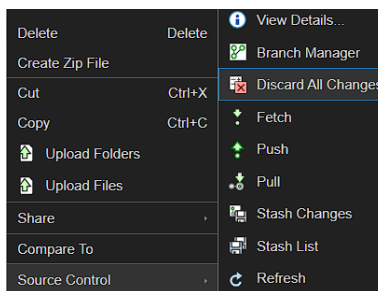
- <https://github.com/vasanza/SSE>

Librarys:

- https://github.com/vasanza/Matlab_Code
- <https://es.mathworks.com/help/matlab/ref/lis.htm>
- <https://es.mathworks.com/help/matlab/ref/matlab.git.gitrepository.discardchanges.html#d126e406558>

Paso 0: Descartar cualquier cambio realizado en el repositorio clonado

```
%Version Online, Opcion 1:
% Source Control -> Discard all changes
% Source Control -> git pull
```



```
%Version Online, Opcion 2:
% repo = gitrepo;
% discardChanges(repo,repo.ModifiedFiles);
% Source Control -> git pull
```

```
Command Window
>> repo = gitrepo;
discardChanges(repo,repo.ModifiedFiles);
>>
```

```
% Version para PC, en el Bash del Git:
% git status
```

```
% git reset --hard
% Git pull
```

```

MINGW64 ~/SSE/2024 (main)
$ git status
On branch main
Your branch is behind 'origin/main' by 5 commits, and can be fast-forwarded.
(use 'git pull' to update your local branch)

Changes not staged for commit:
  (use 'git add <file>' to update what will be committed)
  (use 'git restore <file>...' to discard changes in working directory)
        modified:   ACTIVIDAD10/main.mlx
        modified:   ACTIVIDAD09/main - Copy.mlx
        modified:   ACTIVIDAD09/main.mlx

no changes added to commit (use "git add" and/or "git commit -a")

MINGW64 ~/SSE/2024 (main)
$ git reset --hard
HEAD is now at Sae2883 Add files via upload

MINGW64 ~/SSE/2024 (main)
$ git pull
Updating Sae2883..c69ecb8
Fast forward
 2024/ACTIVIDAD10/data/Cliente1/2024-06-25.csv | 2814 ++++++
 2024/ACTIVIDAD10/data/Cliente1/2024-06-26.csv | 2820 ++++++
 2024/ACTIVIDAD10/data/Cliente1/2024-06-27.csv | 2820 ++++++
 2024/ACTIVIDAD10/data/Cliente1/2024-06-28.csv | 2820 ++++++
 2024/ACTIVIDAD10/data/Cliente2/2024-06-2.csv | 2820 ++++++
 2024/ACTIVIDAD10/data/Cliente2/2024-06-21.csv | 2814 ++++++
 2024/ACTIVIDAD10/data/Cliente2/2024-06-22.csv | 2820 ++++++
 2024/ACTIVIDAD10/data/Cliente2/2024-06-23.csv | 2820 ++++++
 2024/ACTIVIDAD10/data/Cliente3/2023-06-21.csv | 2814 ++++++
 2024/ACTIVIDAD10/data/Cliente3/2023-06-22.csv | 2820 ++++++
 2024/ACTIVIDAD10/data/Cliente3/2023-06-23.csv | 2820 ++++++
 .../data/Cliente3/Copy_of_2023-06-23.csv | 2820 ++++++
 .../data/Copy_2_of_Cliente1/2023-06-21.csv | 2814 ++++++
 .../data/Copy_2_of_Cliente1/2023-06-22.csv | 2820 ++++++
 .../data/Copy_2_of_Cliente1/2023-06-23.csv | 2820 ++++++
 .../data/Copy_2_of_Cliente3/Copy_of_2023-06-23.csv | 2820 ++++++
 2024/ACTIVIDAD10/main.mlx | Bin 235418 -> 53128 bytes
 2024/ACTIVIDAD10/src/FPPlot.m | 31 +
 18 files changed, 22579 insertions(+), 22548 deletions(-)
 create mode 100644 2024/ACTIVIDAD10/src/FPPlot.m

MINGW64 ~/SSE/2024 (main)
$

```

Paso 1: Limpiar variables y linea de comandos

```
clear % Para borrar el workspace y liberar memoria RAM
clc % Limpiar el command window
```

Paso 2.- Configuración de carpeta ./src para librerias

```
%nombre de la carpeta donde estan los codigos
addpath(genpath('./src'));
```

Paso 3- Configuración de carpeta de ./data para datasets

```
%Nombre de la carpeta donde estan los archivos csv
datapath=fullfile('./data/');
```

Paso 4- Desde Open Energy Data Initiative, seleccionar un archivo CSV cualquiera de cualquier estado

Informacion del dataset

- USA_AK_FAIRBANKS.csv
- USA_AK_Anchorage.Intl.AP.702730_TMY3/

```
cd data %Comando linux para entrar una carpeta
httpsUrl = "https://openei.org/datasets/files/961/pub/COMMERCIAL_LOAD_DATA_E_PLUS_OUTPUT/USA_AK_FAIRBANKS.csv";
dataUrl = strcat(httpsUrl, "/RefBldgSmallHotelNew2004_v1.3_7.1_8A_USA_AK_FAIRBANKS.csv");
DataFile = "RefBldgSmallHotelNew2004_v1.3_7.1_8A_USA_AK_FAIRBANKS.csv";
DataFileFullPath = websave(DataFile,dataUrl);
```

```
cd .. %Comando linux para salir de la carpeta
clear httpsUrl dataUrl DataFile DataFileFullPath
```

Paso 4- Buscar los nombres y Cargar los datos de todos los archivos dentro de la carpeta ./data

```
%Leer un archivo csv y lo carga como una tabla
filename = FindCSV(datapath);
maxnames=size(filename,1);
% Dataset es una tabla donde cada columna es una variable con su
% respectivo nombre
index=1; % El archivo que quiero que lea desde la carpeta data
filename(index).name
```

```
ans =
'RefBldgSmallHotelNew2004_v1.3_7.1_8A_USA_AK_FAIRBANKS.csv'
```

```
Dataset=fLoadTableCSV_index(filename,datapath,index)
```

Warning: Column headers from the file were modified to make them valid MATLAB identifiers before creating variable names for the table. The original column headers are saved in the VariableDescriptions property. Set 'VariableNamingRule' to 'preserve' to use the original column headers as table variable names.
Dataset = 8760×11 table

	Date_Time	Electricity_Facility_kW__Hourly__	Fans_Electricity_kW__Hourly__
1	'01/01 01:0...	49.7663	4.9317
2	'01/01 02:0...	49.0528	4.9225
3	'01/01 03:0...	45.6644	4.9503
4	'01/01 04:0...	46.4704	4.9665
5	'01/01 05:0...	48.4044	4.9731
6	'01/01 06:0...	52.7617	4.9778
7	'01/01 07:0...	70.7867	4.9323
8	'01/01 08:0...	83.3511	4.8494
9	'01/01 09:0...	103.3132	4.7682
10	'01/01 10:0...	92.3172	4.8234
11	'01/01 11:0...	67.0967	4.9117
12	'01/01 12:0...	63.2558	4.9205
13	'01/01 13:0...	63.0341	4.9156
14	'01/01 14:0...	62.7386	4.9091
15	'01/01 15:0...	63.2817	4.9199
16	'01/01 16:0...	67.8483	4.9440
17	'01/01 17:0...	76.1231	4.9644
18	'01/01 18:0...	82.4558	4.9259

	Date_Time	Electricity_Facility_kW_Hourly_	Fans_Electricity_kW_Hourly_
19	'01/01 19:0...	93.7888	4.8671
20	'01/01 20:0...	91.5739	4.8169
21	'01/01 21:0...	94.8901	4.7803
22	'01/01 22:0...	93.0372	4.7903
23	'01/01 23:0...	75.6534	4.8633
24	'01/01 24:0...	57.3961	4.9514
25	'01/02 01:0...	52.3303	5.0023
26	'01/02 02:0...	52.2822	5.0345
27	'01/02 03:0...	52.2930	5.0724
28	'01/02 04:0...	53.4376	5.0965
29	'01/02 05:0...	55.5334	5.1071
30	'01/02 06:0...	62.8768	5.1003
31	'01/02 07:0...	86.3975	4.9614
32	'01/02 08:0...	101.8616	4.8937
33	'01/02 09:0...	95.1096	4.9931
34	'01/02 10:0...	92.2926	5.0178
35	'01/02 11:0...	72.2894	5.0903
36	'01/02 12:0...	70.3084	5.0902
37	'01/02 13:0...	68.6029	5.0545
38	'01/02 14:0...	68.8431	5.0533
39	'01/02 15:0...	68.4585	5.0467
40	'01/02 16:0...	72.0661	5.0511
41	'01/02 17:0...	79.6772	5.0558
42	'01/02 18:0...	87.6466	4.9736
43	'01/02 19:0...	92.1161	4.9279
44	'01/02 20:0...	94.1571	4.9163
45	'01/02 21:0...	100.7682	4.8789
46	'01/02 22:0...	94.1325	4.8557
47	'01/02 23:0...	74.4350	4.9378
48	'01/02 24:0...	58.8025	5.0255
49	'01/03 01:0...	54.1345	5.0405
50	'01/03 02:0...	54.2916	5.0761
51	'01/03 03:0...	54.0627	5.1100

	Date_Time	Electricity_Facility_kW_Hourly_	Fans_Electricity_kW_Hourly_
52	'01/03 04:0...	55.3790	5.1367
53	'01/03 05:0...	58.3840	5.1635
54	'01/03 06:0...	65.5551	5.1533
55	'01/03 07:0...	87.3740	4.9801
56	'01/03 08:0...	101.8606	4.8951
57	'01/03 09:0...	96.2311	5.0159
58	'01/03 10:0...	94.6562	5.0658
59	'01/03 11:0...	73.6436	5.1208
60	'01/03 12:0...	71.7096	5.1220
61	'01/03 13:0...	71.6814	5.1240
62	'01/03 14:0...	71.4229	5.1147
63	'01/03 15:0...	71.2283	5.1121
64	'01/03 16:0...	74.2087	5.1025
65	'01/03 17:0...	80.9397	5.0877
66	'01/03 18:0...	90.7177	5.0376
67	'01/03 19:0...	94.2199	4.9713
68	'01/03 20:0...	93.2583	4.9050
69	'01/03 21:0...	98.3870	4.8429
70	'01/03 22:0...	92.6045	4.8333
71	'01/03 23:0...	74.7434	4.9462
72	'01/03 24:0...	60.7027	5.0635
73	'01/04 01:0...	58.1475	5.1168
74	'01/04 02:0...	58.0668	5.1488
75	'01/04 03:0...	56.0463	5.1503
76	'01/04 04:0...	55.7792	5.1479
77	'01/04 05:0...	57.4643	5.1506
78	'01/04 06:0...	64.5023	5.1379
79	'01/04 07:0...	87.5384	4.9854
80	'01/04 08:0...	102.0176	4.8993
81	'01/04 09:0...	95.5422	5.0068
82	'01/04 10:0...	93.0896	5.0381
83	'01/04 11:0...	73.9903	5.1289
84	'01/04 12:0...	72.9964	5.1468

	Date_Time	Electricity_Facility_kW__Hourly__	Fans_Electricity_kW__Hourly__
85	'01/04 13:0...	73.1282	5.1523
86	'01/04 14:0...	72.6056	5.1378
87	'01/04 15:0...	71.5413	5.1153
88	'01/04 16:0...	73.9028	5.1302
89	'01/04 17:0...	83.7824	5.1435
90	'01/04 18:0...	92.2774	5.0648
91	'01/04 19:0...	96.6051	5.0115
92	'01/04 20:0...	97.5994	4.9782
93	'01/04 21:0...	103.7455	4.9310
94	'01/04 22:0...	96.9002	4.9043
95	'01/04 23:0...	77.2206	4.9885
96	'01/04 24:0...	62.5694	5.0982
97	'01/05 01:0...	58.4620	5.1245
98	'01/05 02:0...	56.0691	5.1149
99	'01/05 03:0...	54.4357	5.1233
100	'01/05 04:0...	54.7291	5.1310

⋮

Paso 5- Extraer nombres de variables y crear datetime

```
% Extraer todos los nombres de variables de la tabla
% Se hace el cast de cell a string
varnames=string(Dataset.Properties.VariableNames);
% Eliminar el primero nombbre de variable
varnames=varnames(2:end)';

% Esto es para eliminar el warning de los legend en el plot
%LegendNames=char(varnames);
%LegendNames=LegendNames(:,1:15);
%LegendNames=[LegendNames char(65*ones([size(varnames,1),1]))];
%LegendNames=string(LegendNames);

% Crear datetime con una frecuencia de muestreo de un dato por hora segun el
% dataset
%Time = Start Time: Step Time: End Time
time = datetime(2004, 1, 1):hours(1):datetime(2004, 12, 31);
% Se elimina el primer valor
time=time(1,2:end)';

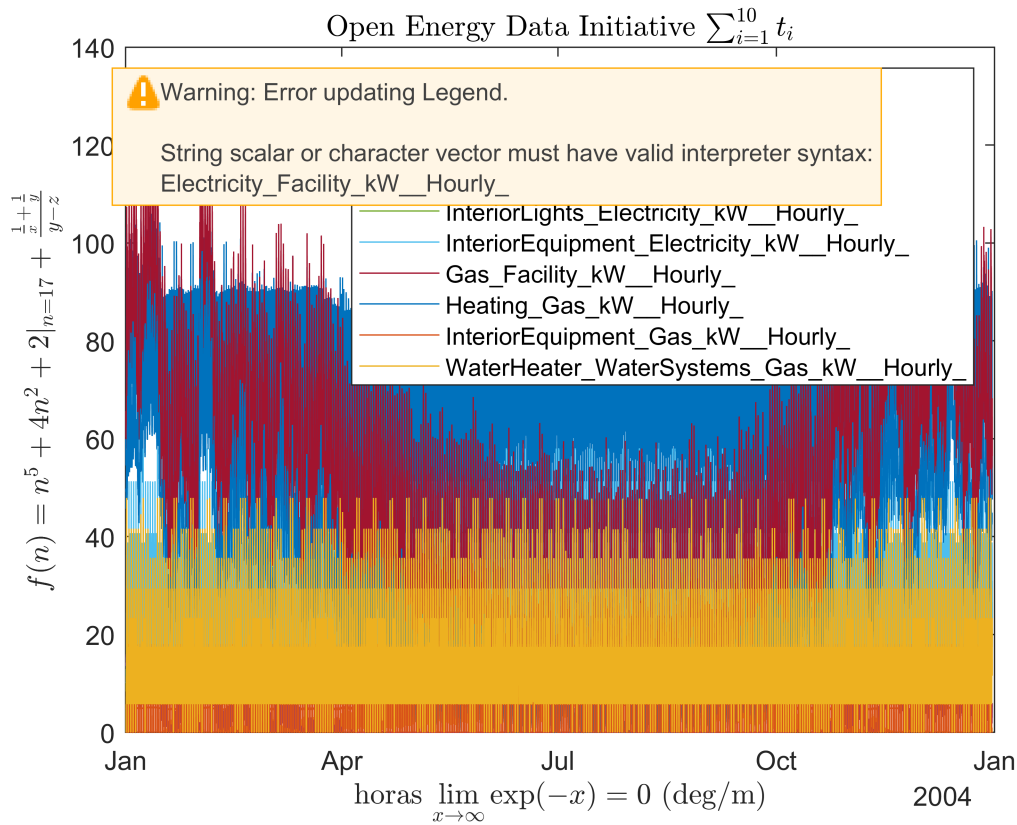
% para agregar una nueva variables en la tabla
%Dataset.('Time_Stamp')=time;
```

Paso 6- Graficar todas las variables

```
figure;
% Dataset
Variable1=Dataset.(varnames(1));
plot(time,Variable1)
hold on
Variable2=Dataset.(varnames(2));
plot(time,Variable2)
Variable3=Dataset.(varnames(3));
plot(time,Variable3)
Variable4=Dataset.(varnames(4));
plot(time,Variable4)
Variable5=Dataset.(varnames(5));
plot(time,Variable5)
Variable6=Dataset.(varnames(6));
plot(time,Variable6)
Variable7=Dataset.(varnames(7));
plot(time,Variable7)
Variable8=Dataset.(varnames(8));
plot(time,Variable8)
Variable9=Dataset.(varnames(9));
plot(time,Variable9)
Variable10=Dataset.(varnames(10));
plot(time,Variable10)

hold off
%legend(LegendNames);
legend(varnames);

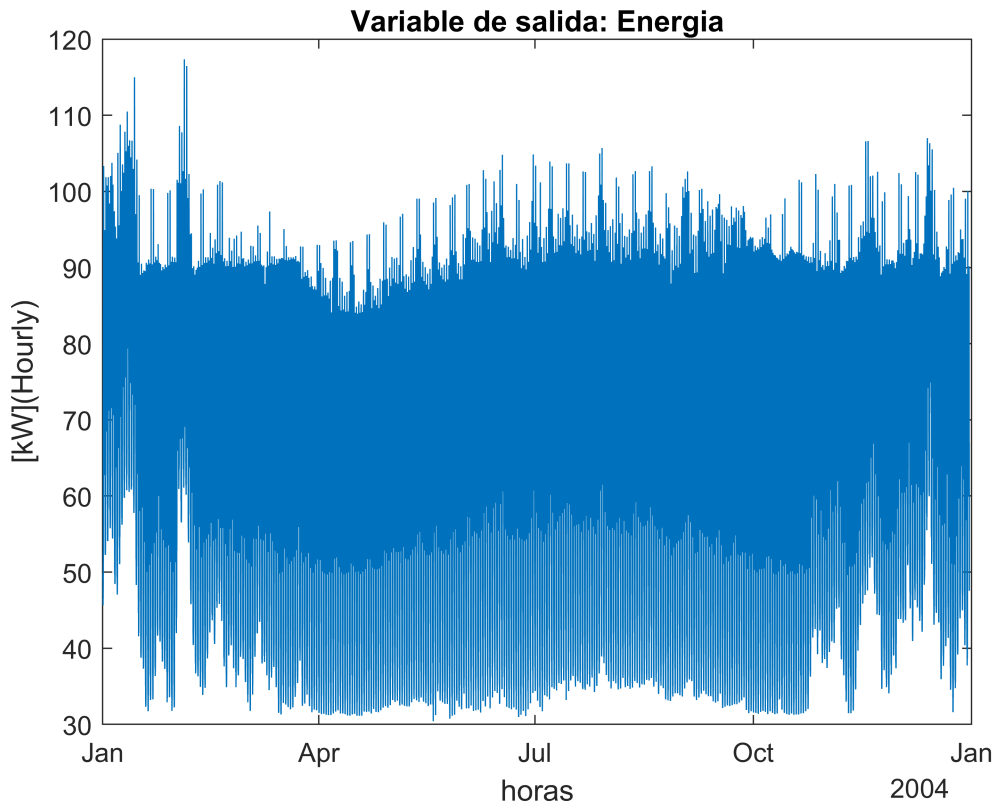
% tambien se puede usar latex para los lables y titles
% https://en.wikibooks.org/wiki/LaTeX/Mathematics
title('Open Energy Data Initiative $ \sum_{i=1}^{10} t_i $', 'interpreter', 'latex');
xlabel('horas $ \lim\limits_{x \to \infty} \exp(-x) = 0 $ (deg/m)', 'interpreter', 'latex')
ylabel(' $ f(n) = n^5 + 4n^2 + 2 \mid_{n=17} + \frac{\frac{1}{x}+\frac{1}{y}}{y-z} $', 'interpreter', 'latex')
```

```
clear filename maxnames index
clear Variable1 Variable2 Variable3 Variable4 Variable5 Variable6 ...
Variable7 Variable8 Variable9 Variable10
```

Paso 7- Definir la variable de salida y las variables de entrada del sistema

```
figure;
% Dataset
output=Dataset.(varnames(1));
plot(time,output)
title('Variable de salida: Energia');
xlabel('horas')
ylabel('[kW](Hourly)')
```



```
% Estamos usando las variables como características
input=[Dataset.(varnames(2)) Dataset.(varnames(3))...
       Dataset.(varnames(4)) Dataset.(varnames(5))...
       Dataset.(varnames(6)) Dataset.(varnames(7))...
       Dataset.(varnames(8)) Dataset.(varnames(9))...
       Dataset.(varnames(10))];
```

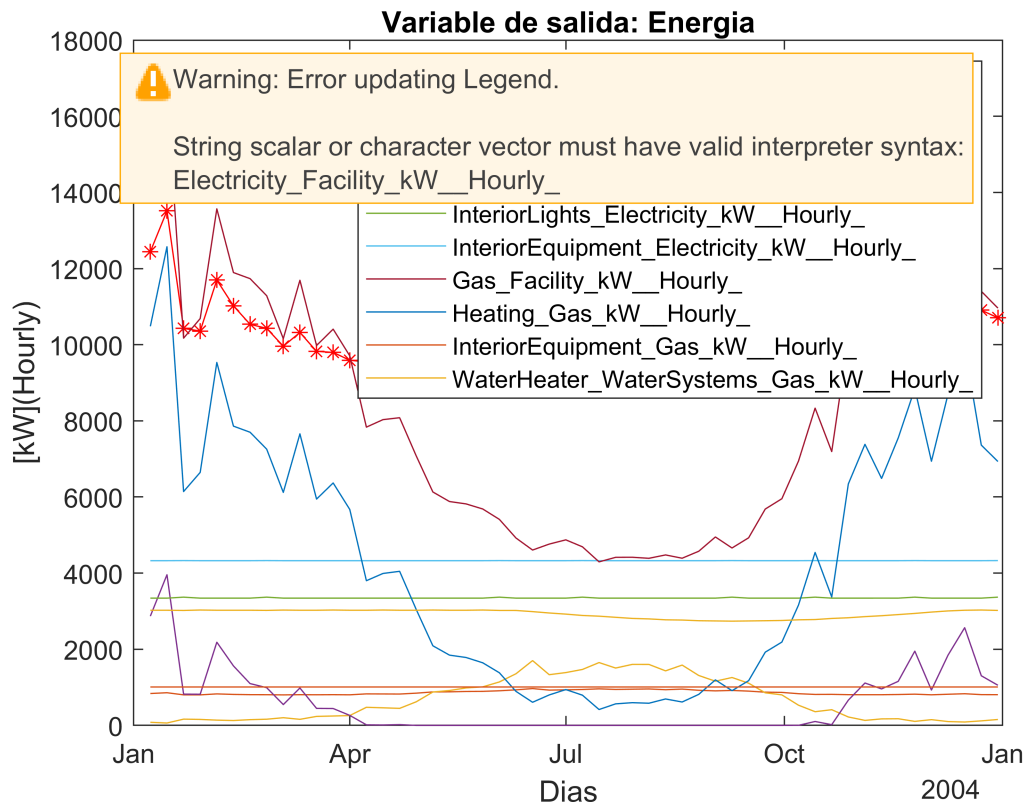
Paso 9- Cambiar la resolucion temporal de los datos: Dias, Semanas, meses

```
%Variable de conversion de horas a Dias, Semanas o meses
nhoras=24*7; %dias

outputDias=[];
inputDias=[];
for i=1:nhoras:size(output,1)-(nhoras-1)
    outputDias=[outputDias;sum(output(i:i+nhoras-1),1)];%1+23 =24
    inputDias=[inputDias;sum(input(i:i+nhoras-1,:),1)];%1+23 =24
end

% Crear datetime con una frecuencia de muestreo de un dato por hora segun el
% dataset
%Time = Start Time: Step Time: End Time
timeDias = datetime(2004, 1, 1):hours(nhoras):datetime(2004, 12, 31);
% Se elimina el primer valor
timeDias=timeDias(1,2:end)';
```

```
figure;
plot(timeDias,outputDias,'-*r')
hold on
plot(timeDias,inputDias)
hold off
title('Variable de salida: Energia');
xlabel('Dias')
ylabel(' [kW](Hourly)')
legend(varnames);
```



```
clear i nhoras;
```

Paso 8- Seleccionar Variables o características

```
%Maximum correlation value allowed
% Default 0.75
threshold = 0.75;

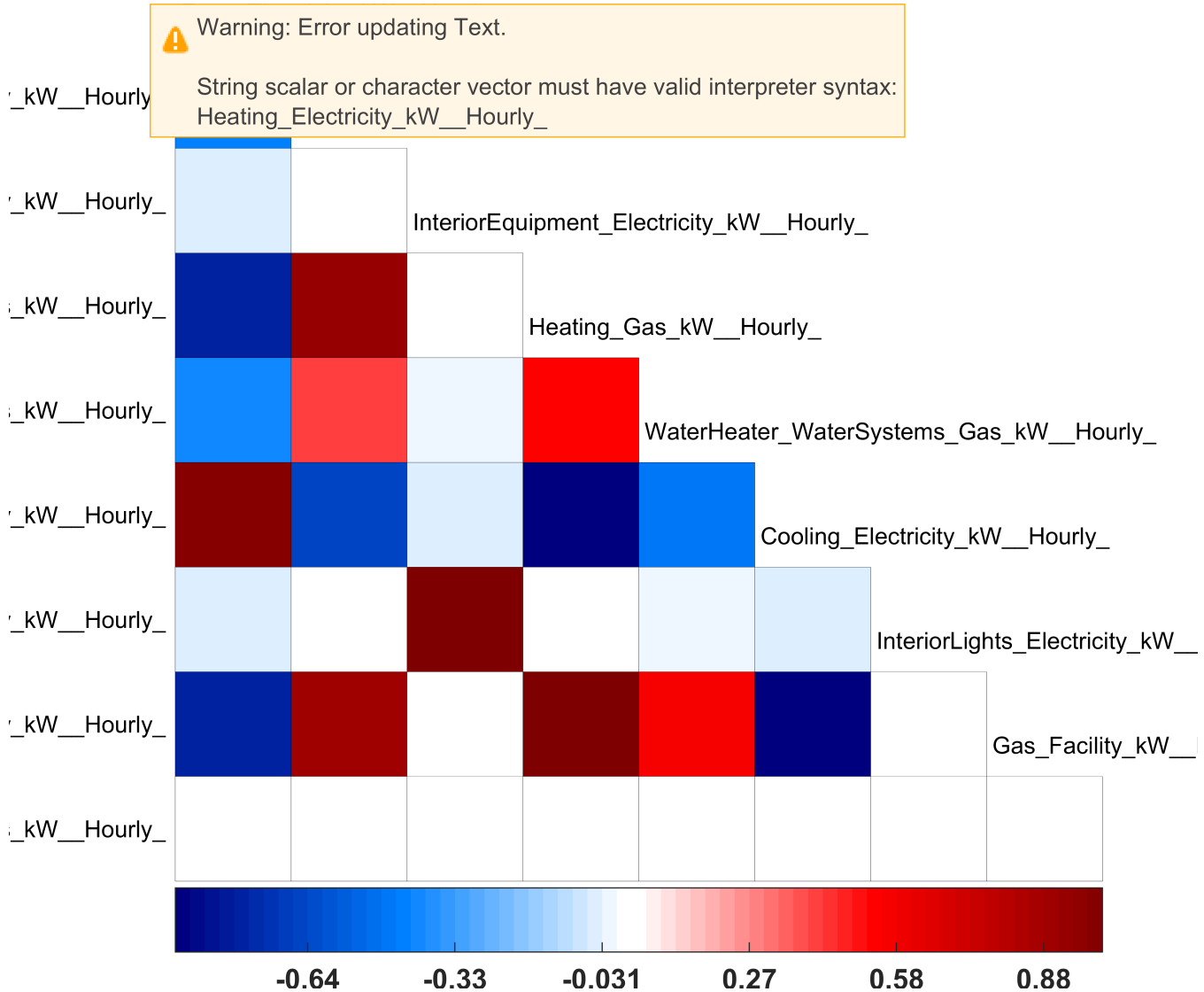
% El numero de variables a analizar debe ser igual al numero de nombres de
% variables
Features_labels=cellstr(varnames(2:end)');
%corrcoef(input)

% Example:
% a=[1:10];b=a+3;c=a.*b;
% corrcoef([a b c])
```

```
% 1.0000    1.0000    0.9816
% 1.0000    1.0000    0.9816
% 0.9816    0.9816    1.0000
```

```
[NewDataFeatures,NewFeaturesLabels,LabelsRemove] = Feature_Selection(inputDias,Features_labels,
```

Electrical Consumption Parameters



NewDataFeatures = 52x5

```
103 ×
0.8377    2.8704    3.3430    1.0094    3.0248
0.8583    3.9577    3.3430    1.0094    3.0246
0.8027    0.8216    3.3689    1.0094    3.0196
0.8022    0.8187    3.3430    1.0094    3.0309
0.8277    2.1854    3.3430    1.0094    3.0249
0.8140    1.5633    3.3430    1.0094    3.0246
0.8076    1.0988    3.3430    1.0094    3.0249
0.8064    0.9910    3.3689    1.0094    3.0208
0.8017    0.5473    3.3430    1.0094    3.0296
0.8071    0.9825    3.3430    1.0094    3.0246
⋮
```

```
NewFeaturesLabels = 1x5 cell
'Fans_Electricity_kW_Hourly_' 'Heating_Electricity_kW_Hourly_' 'InteriorLights_E...'
LabelsRemove = 1x4 cell
'Cooling_Electricity_kW_Hourly_' 'InteriorEquipment_Electricity_kW_Hourly_' 'Gas...'
Warning: Error updating Text.
```

```
String scalar or character vector must have valid interpreter syntax:
Gas_Facility_kW_Hourly_
```

```
%Variables eliminadas por superar el threshold
LabelsRemove'
```

```
ans = 4x1 cell
'Cooling_Electricity_kW_Hourly_'
'InteriorEquipment_Electricity_kW_Hourly_'
'Gas_Facility_kW_Hourly_'
'Heating_Gas_kW_Hourly_'
```

```
%variables que se quedan por no superan el threshold en coreelacion
NewFeaturesLabels'
```

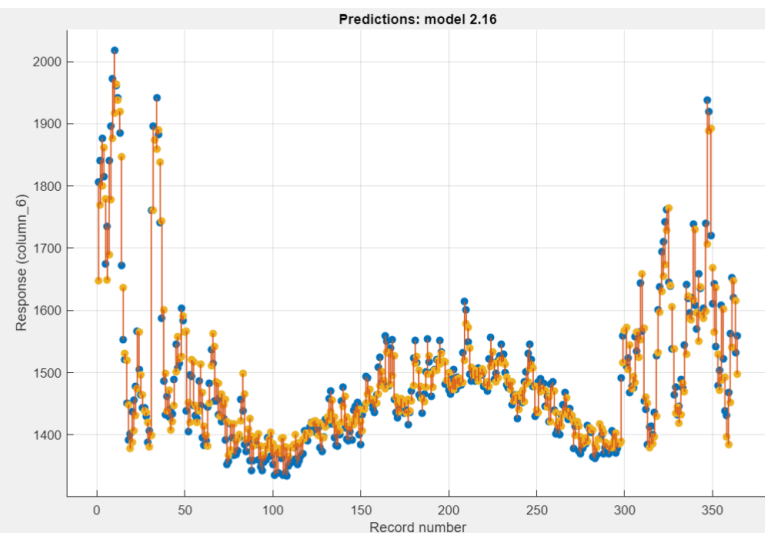
```
ans = 5x1 cell
'Fans_Electricity_kW_Hourly_'
'Heating_Electricity_kW_Hourly_'
'InteriorLights_Electricity_kW_Hourly_'
'InteriorEquipment_Gas_kW_Hourly_'
'WaterHeater_WaterSystems_Gas_kW_Hourly_'
```

```
clear threshold Features_labels LabelsRemove;
```

Paso 9- Seleccionar el algortimo de ML con un menos error de prediccion empleando el toolbox de Matlab Regression Learner

```
% Concatenando la variables de entrada actual con
% la salida al dia siguiente
DataRegression=[NewDataFeatures(1:end-1,:) outputDias(2:end,1)];
regressionLearner
```

Last change: Fine Gaussian SVM	5/5 features
2.12 SVM	RMSE (Validation): 56.527
Last change: Medium Gaussian SVM	5/5 features
2.13 SVM	RMSE (Validation): 59.25
Last change: Coarse Gaussian SVM	5/5 features
2.14 Ensemble	RMSE (Validation): 86.693
Last change: Boosted Trees	5/5 features
2.15 Ensemble	RMSE (Validation): 61.292
Last change: Bagged Trees	5/5 features
2.16 Gaussian Process Regres...	RMSE (Validation): 53.81
Last change: Squared Exponential GPR	5/5 features
2.17 Gaussian Process Regre...	RMSE (Validation): 53.824
Last change: Matern 5/2 GPR	5/5 features
2.18 Gaussian Process Regre...	RMSE (Validation): 55.854
Last change: Exponential GPR	5/5 features
2.19 Gaussian Process Regres...	RMSE (Validation): 53.81
Last change: Rational Quadratic GPR	5/5 features
2.20 Neural Network	RMSE (Validation): 54.473
Last change: Narrow Neural Network	5/5 features
2.21 Neural Network	RMSE (Validation): 55.323
Last change: Medium Neural Network	5/5 features
2.22 Neural Network	RMSE (Validation): 66.714



Model 2.16: Gaussian Process Regression
Status: Trained

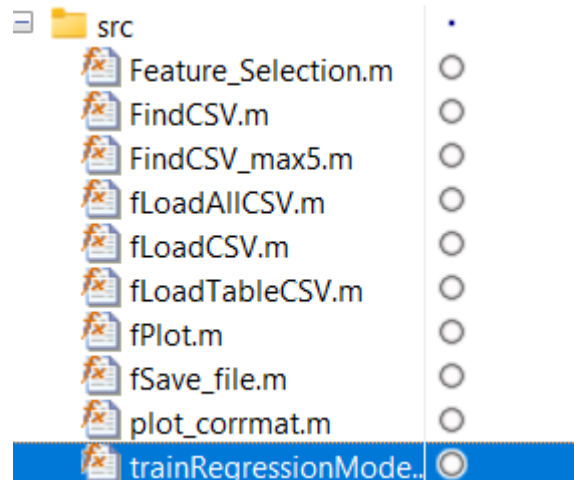
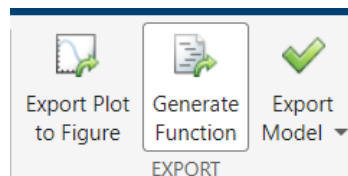
Training Results

RMSE (Validation)	53.81
R-Squared (Validation)	0.81
MSE (Validation)	2895.5
MAE (Validation)	37.747
Prediction speed	~5800 obs/sec
Training time	19.968 sec

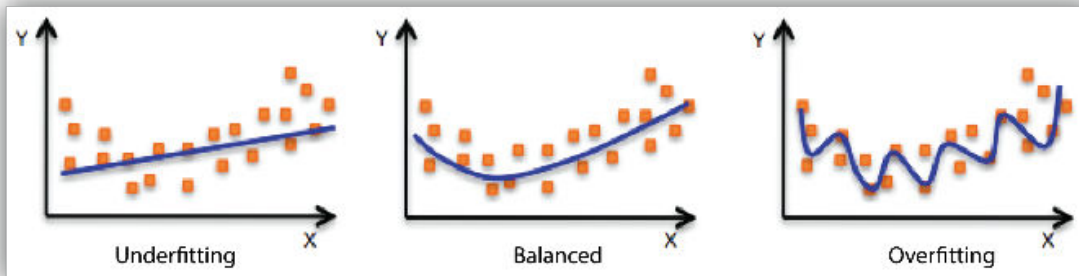
Model Hyperparameters

Preset: Squared Exponential GPR
Basis function: Constant
Kernel function: Squared Exponential
Use isotropic kernel: true
Kernel scale: Automatic
Signal standard deviation: Automatic
Sigma: Automatic
Standardize: true
Optimize numeric parameters: true

- ▶ **Feature Selection: 5/5 individual features selected**
- ▶ **PCA: Disabled**
- ▶ **Optimizer: Not applicable**



Paso 10- Dividir el dataset en 70% para entrenar y 30% validar

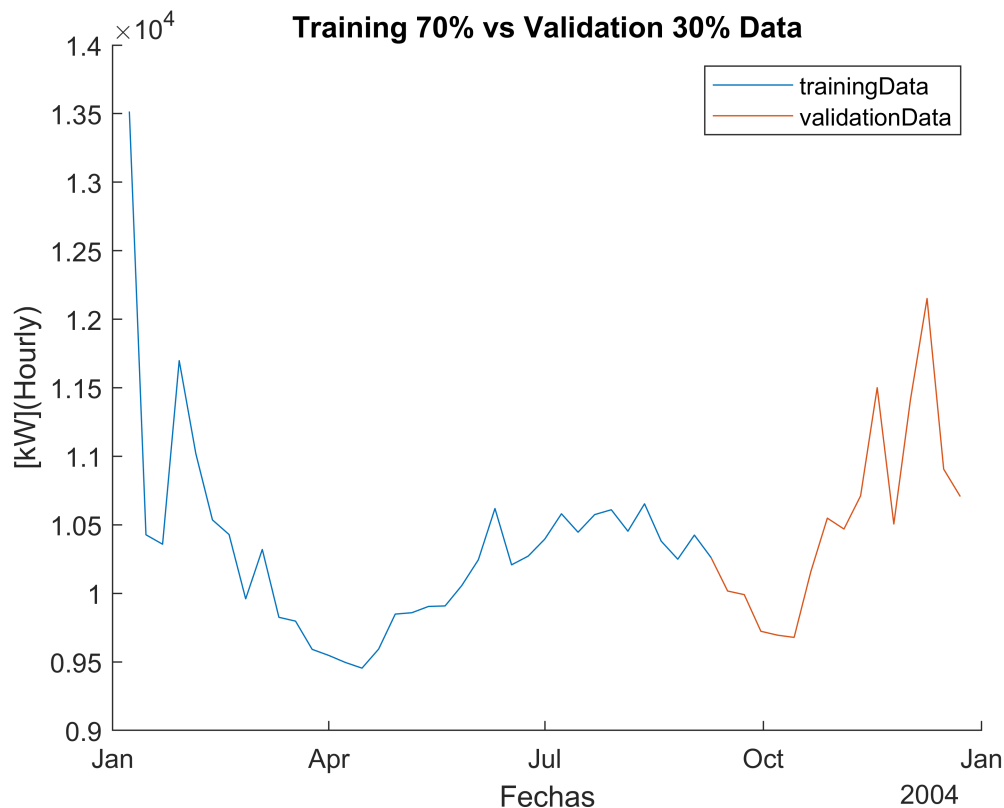


```
%Set de % de entrenamiento
% Dafault 0.70
PSplit = 0.7;
```

```
% Dataset de entrenamiento 70%
trainingData=DataRegression(1:round(end*PSplit),:);
traninTime=timeDias(1:round(end*PSplit));
%Calcular el minimo numero de filas entre Data y Time
nfilas=min([size(trainingData,1) size(traninTime,1)]);
%Tomamos el valor minimo de nfilas como maixmo de filas
trainingData=trainingData(1:nfilas,:);
traninTime=timeDias(1:nfilas,:);

% Dataset de validcion 30%
validationData=DataRegression(round(end*PSplit):end,:);
validationTime=timeDias(round(end*PSplit):end);
%Calcular el minimo numero de filas entre Data y Time
nfilas=min([size(validationData,1) size(validationTime,1)]);
%Tomamos el valor minimo de nfilas como maixmo de filas
validationData=validationData(1:nfilas,:);
validationTime=validationTime(1:nfilas,:);

figure;
% Dataset
hold on
plot(traninTime,trainingData(:,end))
plot(validationTime,validationData(:,end))
hold off
title(['Training ' num2str(PSplit*100) '% vs Validation '...
      num2str(100-PSplit*100) '% Data']);
xlabel('Fechas')
ylabel(['kW](Hourly)')
legend('trainingData','validationData')
```



```
clear PSplit;
```

Paso 11- Usando el algoritmo de ML se entrena el modelo de regression (costo computacional)



```
%Esta funcion permite generar un modelo actualizado cada vez que se ejecuta
% Siempre que el numero de variables de entrada sea la misma y la cantidad
% de nuevos datos no sea muy alta
[trainedModel, validationRMSE] = trainRegressionModel(trainingData);
%Este es el error de entrenaamiento
validationRMSE
```

```
validationRMSE = 837.1932
```

```
%Permite guardar el modelo entrenado que se encuentra en el workspace
save("trainedModel.mat","trainedModel")
```

Paso 12- Cargar y validar el modelo entrenado

```
%Cargar el modelo entrenado y guardado
load("trainedModel.mat")
```



```

%Usar el modelo entrenado para predecir valores de consumo de energia
yest = trainedModel.predictFcn(validationData(:,1:end-1));

%El valor real de consumo de energia para comparar
yout = validationData(:,end);

>Error de prediccion con datos de validacion
validationRMSE = sqrt(mean((yest - yout).^2))

```

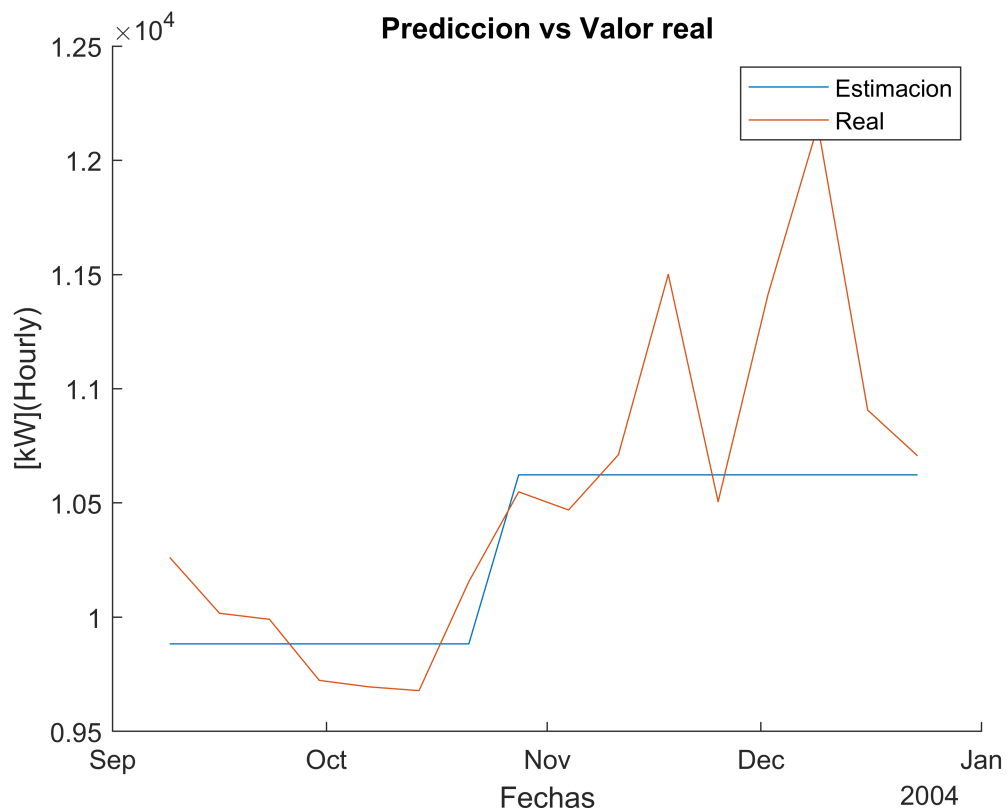
validationRMSE = 513.3857

paso 13 - Graficar el valor predecido vs el valor real

```

figure;
% Dataset
hold on
plot(validationTime,yest)
plot(validationTime,yout)
title('Prediccion vs Valor real');
xlabel('Fechas')
ylabel('[kW](Hourly)')
legend('Estimacion','Real')

```



paso 14 - Mejorar el modelo de prediccion

