



EFC1 - Exercise 1

Rafael Claro Ito

September 2019

1 Theoretical Activites

1.1 Exercise 1

1.1.1 a) Obtenha $P(X)$ e $P(Y)$.

i) $P_X(0)$:

$$P_X(0) = \sum_y P_{XY}(0, y)$$

$$P_X(0) = P_{XY}(0, 0) + P_{XY}(0, 1)$$

$$P_X(0) = 1/6 + 3/8$$

$$P_X(0) = 13/24$$

ii) $P_X(1)$:

$$P_X(1) = \sum_y P_{XY}(1, y)$$

$$P_X(1) = P_{XY}(1, 0) + P_{XY}(1, 1)$$

$$P_X(1) = 1/8 + 1/3$$

$$P_X(1) = 11/24$$

iii) $P_Y(0)$:

$$P_Y(0) = \sum_x P_{XY}(x, 0)$$

$$P_Y(0) = P_{XY}(0, 0) + P_{XY}(1, 0)$$

$$P_Y(0) = 1/6 + 1/8$$

$$P_Y(0) = 7/24$$

iv) $P_Y(1)$:

$$P_Y(1) = \sum_x P_{XY}(x, 1)$$

$$P_Y(1) = P_{XY}(0, 1) + P_{XY}(1, 1)$$

$$P_Y(1) = 3/8 + 1/3$$

$$P_Y(1) = 17/24$$

1.1.2 b) Calcule $P(X = 0 | Y = 0)$.

$$P_{X|Y}(x, y) = \frac{P_{XY}(x, y)}{P_Y(y)}$$

$$P_{X|Y}(0, 0) = \frac{P_{XY}(0, 0)}{P_Y(0)}$$

$$P_{X|Y}(0, 0) = \frac{1/6}{7/24}$$

$$P_{X|Y}(0, 0) = 4/7$$

1.1.3 c) Calcule $E[X]$ e $E[Y]$.

i) $E[X]$:

$$E[X] = \sum_k x_k P_X(x_k)$$

$$E[X] = 0 \cdot P_X(0) + 1 \cdot P_X(1)$$

$$\boxed{E[X] = 11/24}$$

ii) $E[Y]$:

$$E[Y] = \sum_k y_k P_Y(y_k)$$

$$E[Y] = 0 \cdot P_Y(0) + 1 \cdot P_Y(1)$$

$$\boxed{E[Y] = 17/24}$$

1.1.4 d) As variáveis são independentes? Por quê?

TO BE WRITTEN

1.2 Exercise 2

1.2.1 a) Calcule $H(X)$, $H(Y)$ e $H(X-Y)$.

$$P_X(0) = P_{XY}(0,0) + P_{XY}(0,1) = 1/4 \quad P_X(1) = P_{XY}(1,0) + P_{XY}(1,1) = 3/4$$

$$P_Y(0) = P_{XY}(0,0) + P_{XY}(1,0) = 3/8 \quad P_Y(1) = P_{XY}(0,1) + P_{XY}(1,1) = 5/8$$

i) $H(X)$:

$$H(X) = - \sum_x p(x) \log_2[p(x)]$$

$$H(X) = -P_X(0) \log_2[P_X(0)] - P_X(1) \log_2[P_X(1)]$$

$$H(X) = -\frac{1}{4} \log_2\left(\frac{1}{4}\right) - \frac{3}{4} \log_2\left(\frac{3}{4}\right)$$

$$H(X) = 0.5 + 0.3113$$

$$\boxed{H(X) = 0.8113}$$

ii) $H(Y)$:

$$H(Y) = - \sum_y p(y) \log_2[p(y)]$$

$$H(Y) = -P_Y(0) \log_2[P_Y(0)] - P_Y(1) \log_2[P_Y(1)]$$

$$H(Y) = -\frac{3}{8} \log_2\left(\frac{3}{8}\right) - \frac{5}{8} \log_2\left(\frac{5}{8}\right)$$

$$H(Y) = 0.5306 + 0.4238$$

$$\boxed{H(Y) = 0.9544}$$

iii) $H(X,Y)$:

$$H(X,Y) = - \sum_x \sum_y p(x,y) \log_2[p(x,y)]$$

$$\begin{aligned}
H(Y) &= -P_{XY}(0,0) \log_2[P_{XY}(0,0)] - P_{XY}(0,1) \log_2[P_{XY}(0,1)] - P_{XY}(1,0) \log_2[P_{XY}(1,0)] - \\
&P_{XY}(1,1) \log_2[P_{XY}(1,1)] \\
H(Y) &= -0 \cdot \log_2(0) - \frac{1}{4} \log_2\left(\frac{1}{4}\right) - \frac{3}{8} \log_2\left(\frac{3}{8}\right) - \frac{3}{8} \log_2\left(\frac{3}{8}\right) \\
H(Y) &= 0 + 0.5 + 0.5306 + 0.5306 \\
\boxed{H(Y) = 1.5613}
\end{aligned}$$

1.2.2 b) Calcule $H(X|Y)$ e $H(Y|X)$.

$$\begin{aligned}
i) H(X|Y) : \\
H(X,Y) &= H(Y) + H(X|Y) \\
H(X|Y) &= H(X,Y) - H(Y) \\
H(X|Y) &= 1.5613 - 0.9544 \\
\boxed{H(X|Y) = 0.6068}
\end{aligned}$$

$$\begin{aligned}
ii) H(Y|X) : \\
H(X,Y) &= H(X) + H(Y|X) \\
H(Y|X) &= H(X,Y) - H(X) \\
H(Y|X) &= 1.5613 - 0.8113 \\
\boxed{H(Y|X) = 0.75}
\end{aligned}$$

1.2.3 c) Calcule $I(X,Y)$.

$$\begin{aligned}
I(X,Y) &= H(X) - H(X|Y) \quad I(X,Y) = 0.8113 - 0.6068 \quad \boxed{I(X,Y) = 0.2044} \\
I(X,Y) &= H(Y) - H(Y|X) \quad I(X,Y) = 0.9544 - 0.75 \quad \boxed{I(X,Y) = 0.2044}
\end{aligned}$$

1.3 Exercise 3

$$\begin{aligned}
C_1 &\sim \mathcal{N}(-1,1) \quad \mu_1 = -1 \quad \sigma_1 = 1 \\
C_2 &\sim \mathcal{N}(1,1) \quad \mu_2 = 1 \quad \sigma_2 = 1
\end{aligned}$$

1.3.1 a) Através do critério ML, temos que x pertence a classe C_1 sep $p(x|C_1) > p(x|C_2)$ and

$$p(x|C_1) > p(x|C_2)$$

$$\begin{aligned}
\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right) &> \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(x-\mu_2)^2}{2\sigma_2^2}\right) \\
\frac{1}{\sqrt{2\pi \cdot 1^2}} \exp\left(-\frac{(x-(-1))^2}{2 \cdot 1^2}\right) &> \frac{1}{\sqrt{2\pi \cdot 1^2}} \exp\left(-\frac{(x-1)^2}{2 \cdot 1^2}\right) \\
\exp\left(-\frac{(x+1)^2}{2}\right) &> \exp\left(-\frac{(x-1)^2}{2}\right) \\
-(x+1)^2 &> -(x-1)^2
\end{aligned}$$

$$-x^2 - 2x - 1^2 > -x^2 + 2x - 1$$

$$-2x > 2x$$

$$\boxed{x < 0}$$

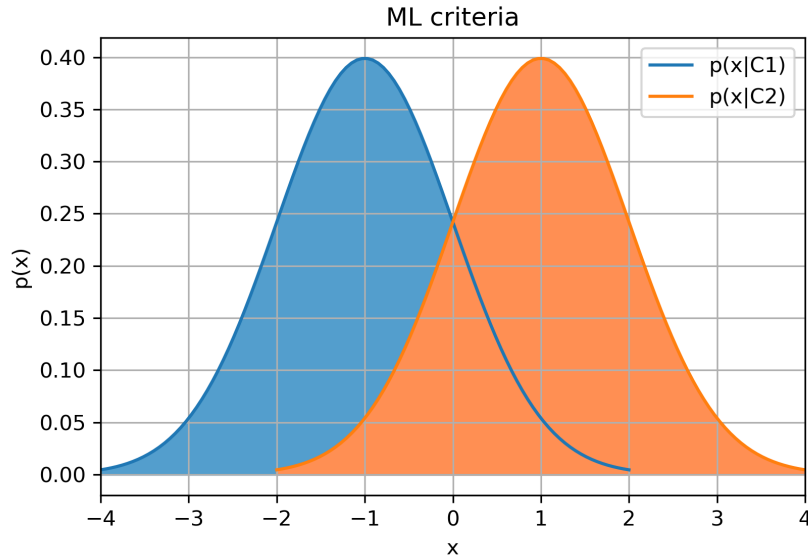


Figure 1: Visualization of C1 and C2 distributions under the ML criteria

As we can see in the 1

1.3.2 b) Calcule $P(X = 0 | Y = 0)$.

$$p(C_1) = 0.7$$

$$p(C_2) = 0.3$$

$$p(C_1|x) > p(C_2|x)$$

Aplicando a regra de Bayes, temos:

$$\frac{p(x|C_1) \cdot p(C_1)}{p(x)} > \frac{p(x|C_2) \cdot p(C_2)}{p(x)}$$

$$\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2}\right) \cdot p(C_1) > \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right) \cdot p(C_2)$$

$$\frac{1}{\sqrt{2\pi \cdot 1^2}} \exp\left(-\frac{(x - (-1))^2}{2 \cdot 1^2}\right) (0.7) > \frac{1}{\sqrt{2\pi \cdot 1^2}} \exp\left(-\frac{(x - 1)^2}{2 \cdot 1^2}\right) (0.3)$$

$$\exp\left(-\frac{(x + 1)^2}{2}\right) (0.7) > \exp\left(-\frac{(x - 1)^2}{2}\right) (0.3)$$

Tirando o \ln de ambos os lados:

$$\begin{aligned}
 -\frac{(x+1)^2}{2} + \ln(0.7) &> -\frac{(x-1)^2}{2} + \ln(0.3) \\
 -x^2 - 2x - 1 + 2\ln(0.7) &> -x^2 + 2x - 1 + 2\ln(0.3) \\
 -2x + 2\ln(0.7) &> 2x + 2\ln(0.3) \\
 -x + \ln(0.7) &> x + \ln(0.3) \\
 2x &< \ln(0.7) - \ln(0.3) \\
 x &< \frac{\ln(0.7) - \ln(0.3)}{2} \\
 \boxed{x < 0.4236}
 \end{aligned}$$

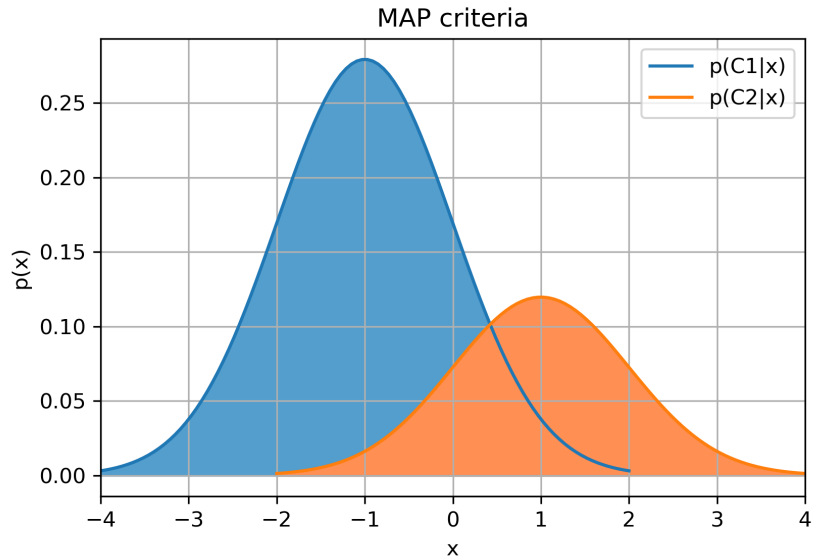


Figure 2: Vizualization of C1 and C2 distributions under the MAP criteria

As we can see in the 2

2 Computational Activites

All code presented and all figures showed here can be found at the following GitHub repository:

<https://github.com/ito-rafael/IA006C-MachineLearning>

In this repository, one can found the following files:

- Jupyter Notebook
 - theoretical-ex.ipynb

- pre-ex1.ipynb
- pre-ex2.ipynb
- ex1.ipynb
- ex2.ipynb
- Python code
 - ex1.py
 - ex2.py
- L^AT_EX
 - efc1.tex

The notebook "theoretical-ex.ipynb" is used to generate the figures 1 and 2. The notebook "pre-ex1" is used only for data visualization. It shows how the dataset is split in order to use the k-fold cross-validation method. The notebook "pre-ex2" has a similar used than the previous one. It is used to visualize the data, but besides the raw data, it also plots the dataset linearly transformed before and after the application of the hyperbolic tangent function. The two last notebooks, "ex1" and "ex2" actually implements all the functions to train the models and that answers the exercises 1 and 2.

The Python codes presented contain the same code of the notebooks. Finally, the L^AT_EX file is the source code to generate this pdf.

2.1 Exercise 1

The first thing we did before starting this exercise was to plot all the data, as can be seen in figure 3.

The next step was to decide the number of folds used in the k-fold cross-validation method. Since we are treating a time series forecasting problem, it was decided to split the dataset equally to the number of the years of the dataset excluding the data used to test the model: nine. When creating the folds, a method like the one that is showed in figure 4 was used.

In this method, only past data is used to predict forward-looking data. The first fold is formed by only the first year of the dataset, the year 1981, while the year used at the validation step is the next year, 1982. The second fold uses all the data of the previous fold (train plus validation) as the data to be trained, and the next year, 1983, as the validation data. Doing this for all the folds we get the configuration showed in figure ??.

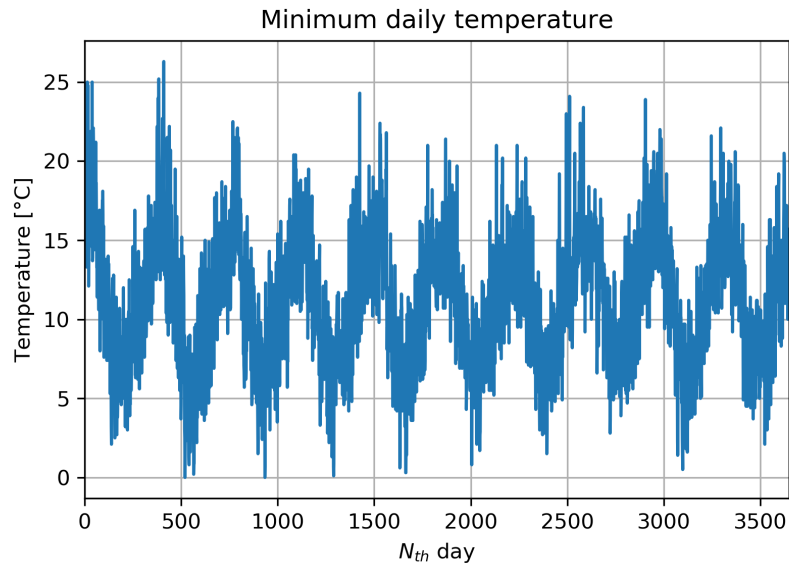


Figure 3: All data

The goal of using this method of cross-validation is to find the best value for the hyperparameter K , the number of previous days used as input for the model.

After choosing the hyperparameter, we no more use the folds division. Instead, all data before used for training and validation is now used only for training. This can be considered as the last training of the model and the one which will return the weights desired in this problem of linear regression. This can be seen in figure 6.

2.1.1 a) Progression of RMSE measured in validation set in function of hyperparameter K (from 1 to 31):

The model was trained for values of the hyperparameter K from 1 to 31 days. The progression of the root mean squared error (RMSE) measured in the validation set can be seen in figure 7.

asd

2.1.2 b) Plot of predicted and real temperatures of test set

2.2 Exercise 2

2.2.1 a)

2.2.2 b)

2.2.3 c)

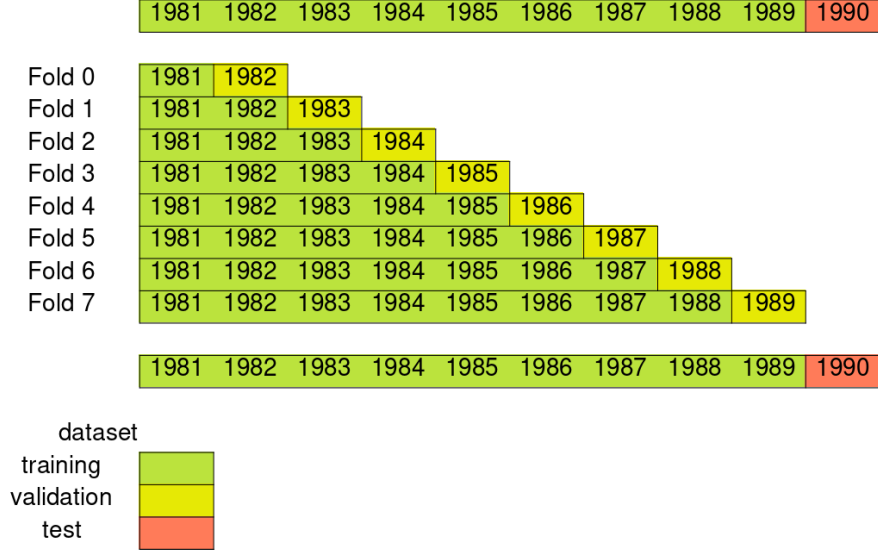


Figure 4: Folds used to cross-validate the model

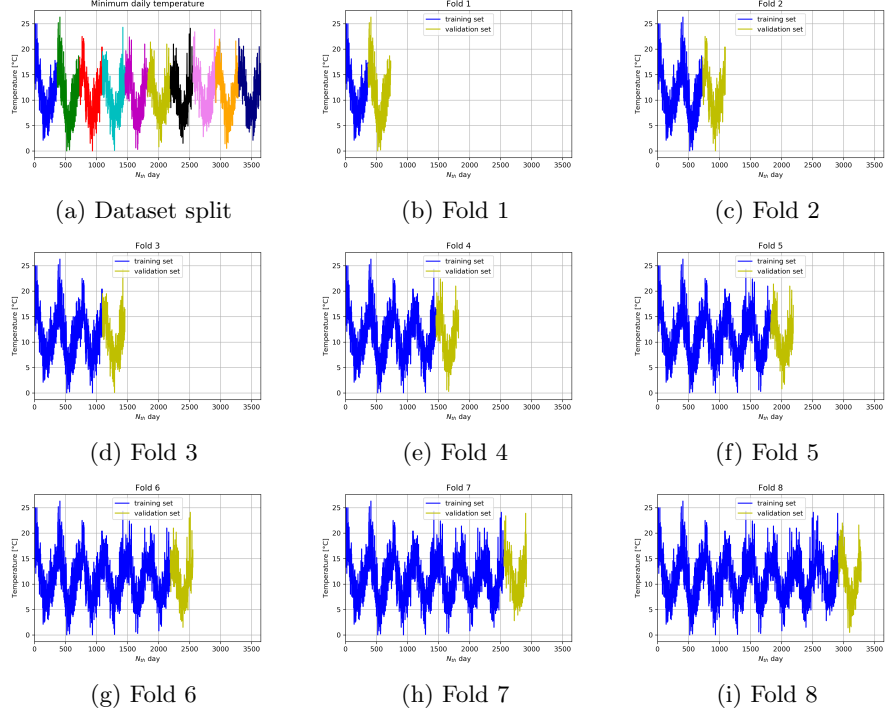


Figure 5: Folds used in cross-validation

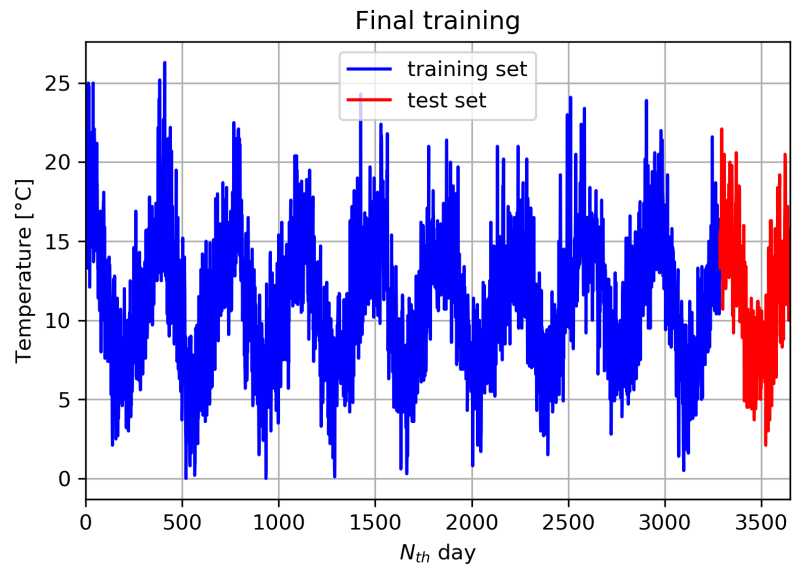


Figure 6

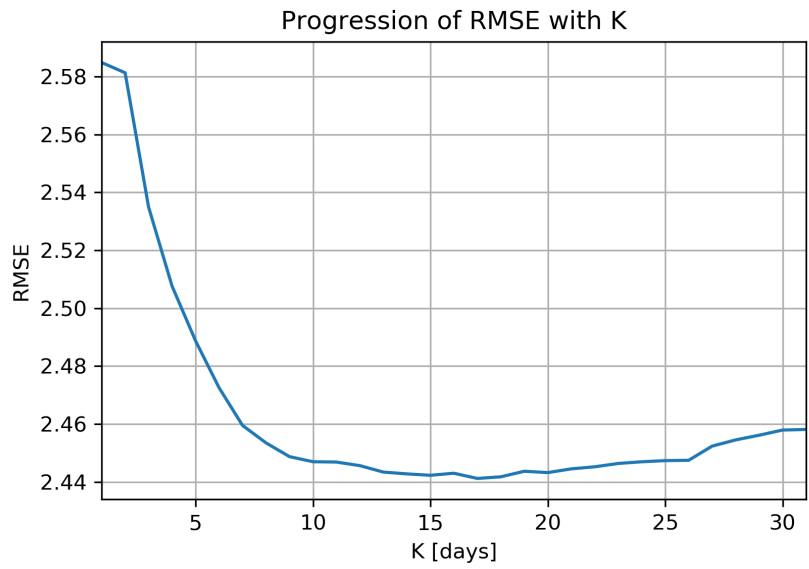


Figure 7

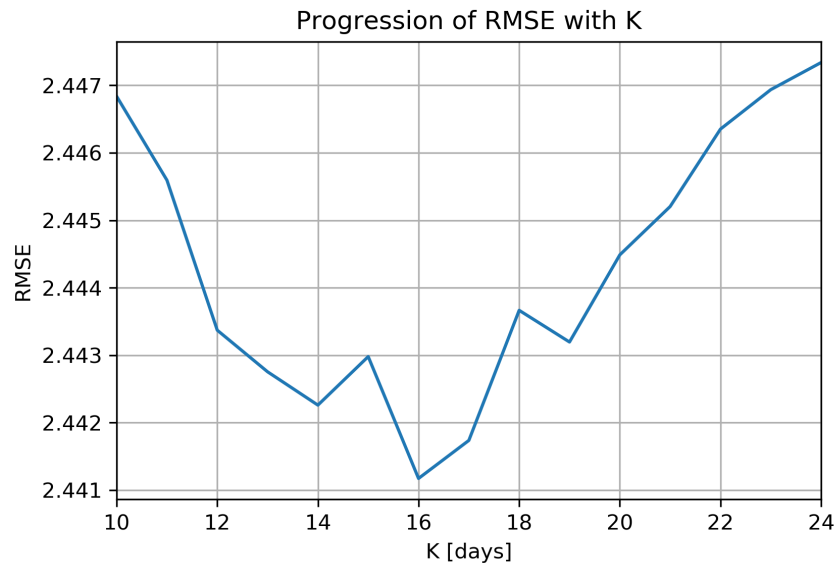


Figure 8

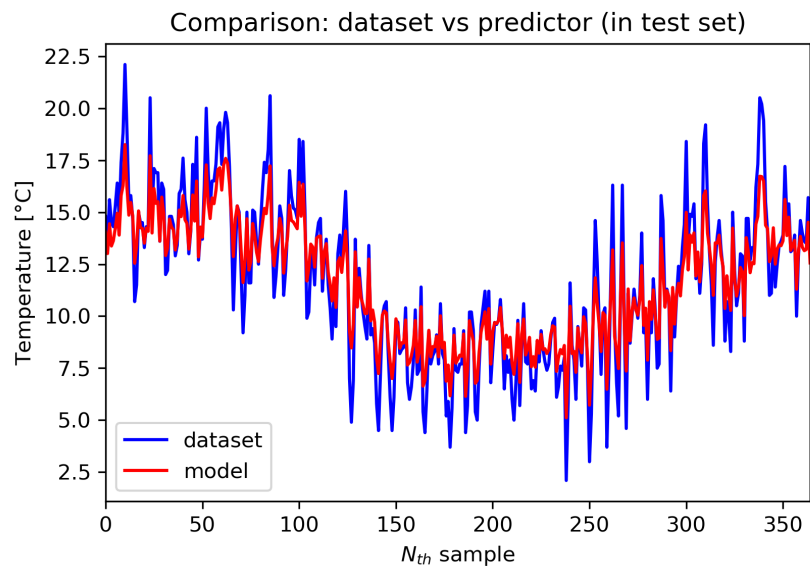


Figure 9

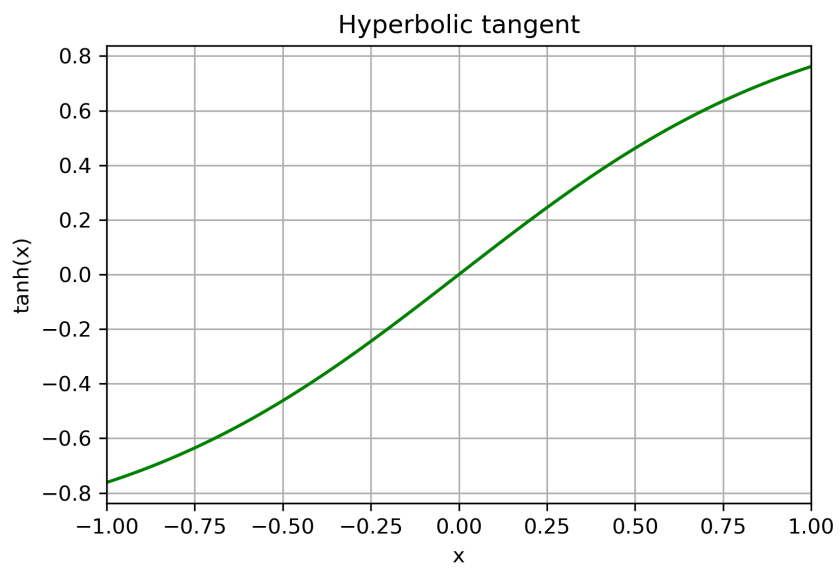


Figure 10

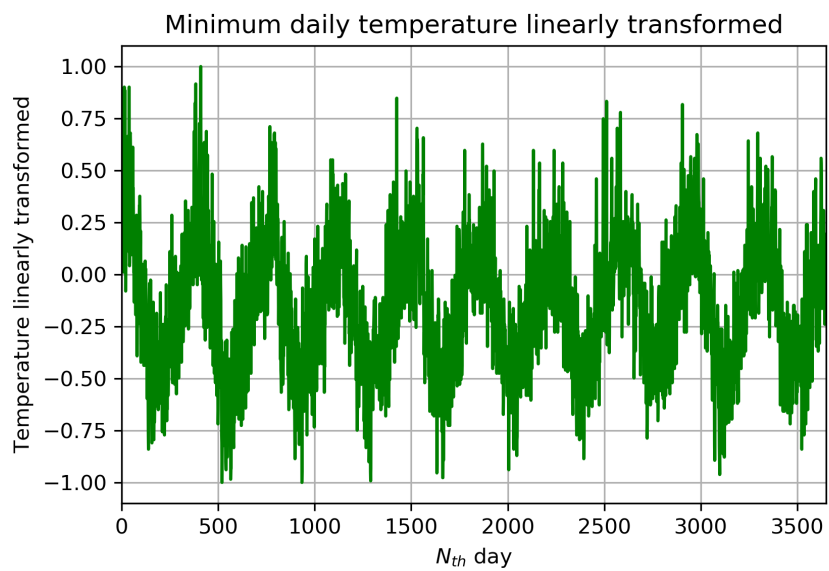


Figure 11

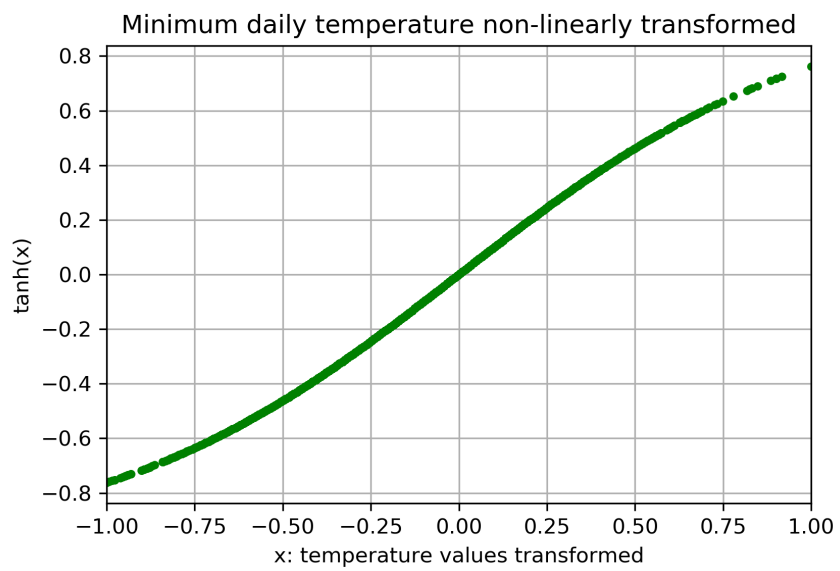


Figure 12

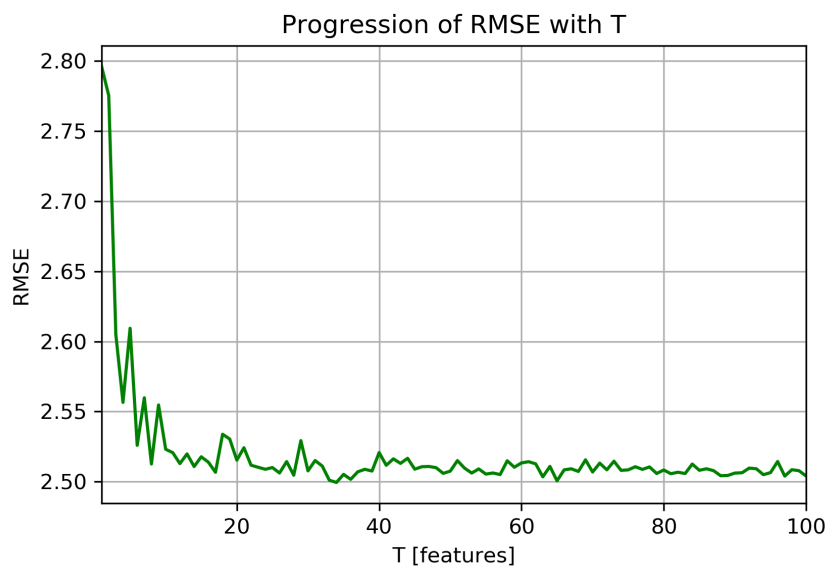


Figure 13

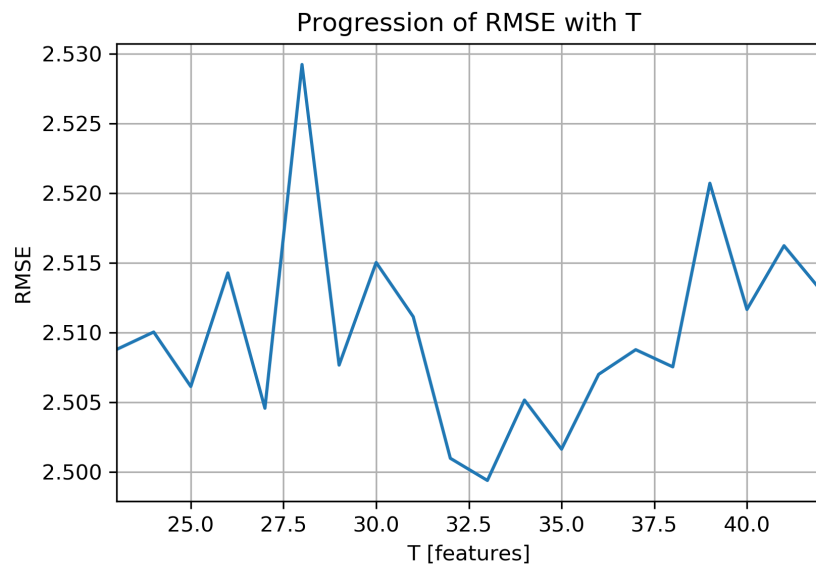


Figure 14

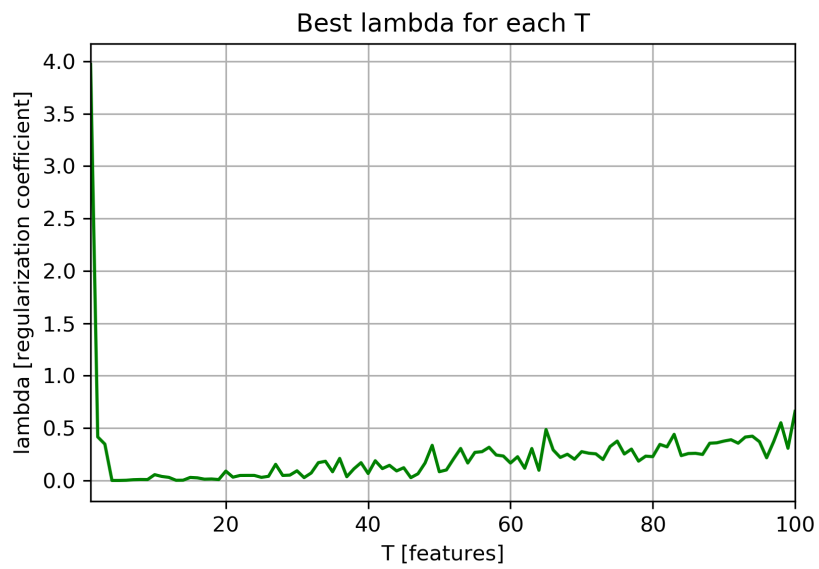


Figure 15

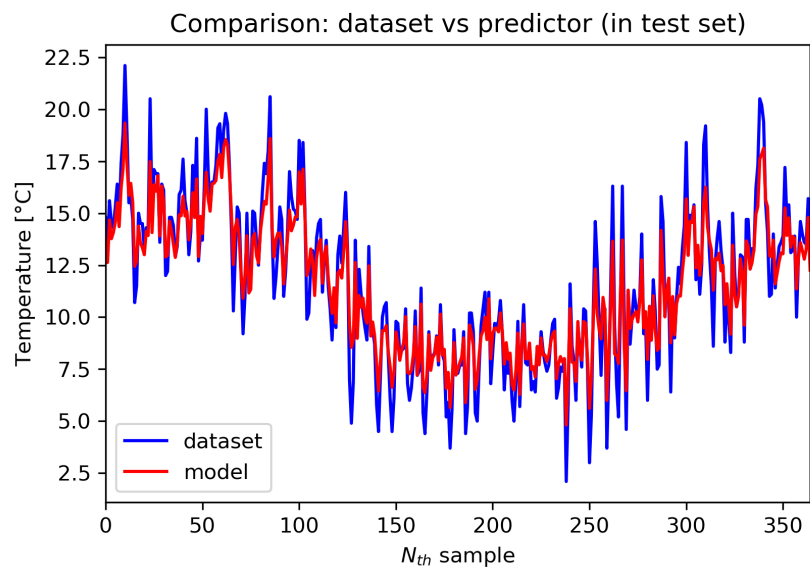


Figure 16