IIT delhi

ASSIGNMENT-5

Q1 Perform PCA and transform the given data:

X	Y
2	1
3	4
5	0
7	6
9	2

Solution: To perform PCA, we need to calculate covariance matrix, eigen values, eigen vectors and transformation matrix.

Step 1: Calculate the covariance matrix for the given data:

$$C = \begin{bmatrix} cov(x,x) & cov(x,y) \\ cov(y,x) & cov(y,y) \end{bmatrix}$$

Where, $cov(x,x) = \Sigma = \frac{1}{n-1}((X-\bar{X})(X-\bar{X})^T)$

Now, mean of X and Y :-

$\bar{X} = 5.2$ $\quad$ $\bar{Y} = 2.6$

X	Y	$X-\bar{X}$	$Y-\bar{Y}$	$(X-\bar{X})^2$	$(Y-\bar{Y})^2$	$\{(X-\bar{X})(Y-\bar{Y})\}$
2	1	-3.2	-1.6	10.24	2.56	5.12
3	4	-2.2	1.4	4.84	1.96	-3.08
5	0	-0.2	-2.6	0.04	6.76	0.52
7	6	1.8	3.4	3.24	11.56	6.12
9	2	3.8	-0.6	14.44	0.36	-2.28
				32.8	23.2	6.4

Now, $Cov(x,x) = \frac{32.8}{5-1} = 8.2$

$Cov(y,y) = \frac{23.2}{5-1} = 5.8$

$Cov(x,y) = Cov(y,x) = \frac{6.4}{5-1} = 1.6$

$\therefore$ Covariance matrix $= \begin{bmatrix} 8.2 & 1.6 \\ 1.6 & 5.8 \end{bmatrix}$

Step 2: Calculate eigen values for the covariance matrix

Using the relation:

$C - \lambda I = 0$

where λ = eigen matrix

I = identity matrix

$$\begin{bmatrix} 8.2 & 1.6 \\ 1.6 & 5.8 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$$

$$\Rightarrow \begin{bmatrix} 8.2-\lambda & 1.6 \\ 1.6 & 5.8-\lambda \end{bmatrix} = 0$$

Solving the above matrix using determinant:-

$(8.2-\lambda)(5.8-\lambda) - 1.6^2 = 0$

$\lambda^2 - 14\lambda + 45 = 0$

Solving the quadratic equation, we have:

iit delhi

$\lambda_1 = 9 \qquad \lambda_2 = 5$

$\therefore$ Eigen values are 9 and 5.

Step 3: Calculate the eigen vectors from eigen values and covariance matrix

For $\lambda_1 = 9$

$$\begin{bmatrix} 8.2 & 1.6 \\ 1.6 & 5.8 \end{bmatrix} \begin{bmatrix} X_1 \\ Y_1 \end{bmatrix} = 9 \begin{bmatrix} X_1 \\ Y_1 \end{bmatrix}$$

$\Rightarrow \quad 8.2\, X_1 + 1.6\, Y_1 = 9 X_1$

and $\quad 1.6 X_1 + 5.8\, Y_1 = 9 Y_1$

solving above set of equation, we have:

$X_1 = 2 Y_1$

Now, the eigen vectors can be found by taking $\| w_2 \|$ of the X_1 and Y_1

$$\therefore \begin{bmatrix} X_1 \\ Y_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \Rightarrow \sqrt{2^2 + 1^2} = \sqrt{5} = 2.236$$

$$\therefore \begin{bmatrix} 2/2.236 \\ 1/2.236 \end{bmatrix} = \begin{bmatrix} 0.8944 \\ 0.447 \end{bmatrix}$$

$\therefore$ 1st eigen vector $e_1 = \begin{bmatrix} 0.8944 \\ 0.447 \end{bmatrix}$

Similarly the second eigen vector for $\lambda_2 = 5$

$$\Rightarrow \begin{bmatrix} 8.2 & 1.6 \\ 1.6 & 5.8 \end{bmatrix} \begin{bmatrix} X_2 \\ Y_2 \end{bmatrix} = 5 \begin{bmatrix} X_2 \\ Y_2 \end{bmatrix}$$

Solving as previous manner:

$$X_2 = -2Y_2$$

$$\therefore \begin{bmatrix} -2 \\ 1 \end{bmatrix} \Rightarrow \sqrt{2^2 + 1^2} = \sqrt{5} = 2.236$$

$$\Rightarrow \begin{bmatrix} -2/2.236 \\ 1/2.236 \end{bmatrix} = \begin{bmatrix} -0.8944 \\ 0.447 \end{bmatrix}$$

$\therefore$ 2nd eigen vector is $\begin{bmatrix} -0.8944 \\ 0.447 \end{bmatrix}$

$\therefore$ Eigen matrix becomes = $\begin{bmatrix} 0.8944 & -0.447 \\ 0.447 & 0.8944 \end{bmatrix}$

Now, Since $\lambda_1 = 9$ is greater than $\lambda_2 = 5$, λ_1 becomes our Principle component as it can explain $\left(\frac{9}{9+5}\right)\% = 64.2\%$ of data.

Step 4 : Data transformation along Principal Component for dimensionality reduction :

Transformation matrix $V = \begin{bmatrix} 0.8944 \\ 0.447 \end{bmatrix}$

$\therefore$ New data, $Y = X \cdot V$

$$\therefore Y = \begin{bmatrix} 2 & 1 \\ 3 & 4 \\ 5 & 0 \\ 7 & 6 \\ 9 & 2 \end{bmatrix} \begin{bmatrix} 0.8944 \\ 0.447 \end{bmatrix}$$

$$Y = \begin{bmatrix} 2.2358 \\ 4.4712 \\ 4.472 \\ 8.9428 \\ 8.9436 \end{bmatrix}$$

Hence Y is the transformed form of original data X.

Assignment – 5

Question 2

Review Report

Deep learning for multi-year ENSO forecasts

Yoo-Geun Ham, Jeong-Hwan Kim & Jing-Jia Luo

Summary

In this paper, the authors have used a convolutional neural network with a transfer learning approach to forecast El-Nino/Southern Oscillation (ENSO) events with a lead time greater than 1 year. The authors have compared the CNN model with state of the art SINTEX-F model and demonstrated that CNN model can produce fairly skillful forecasts at a correlation coefficient of 0.5 and above for a lead time up to 17 months. The authors have also found that CNN model performs better at predicting the type of El Nino event with an accuracy of 66.7% in the validation period. Also, the CNN model performs better in targeted seasons, especially May-June-July with correlation skill exceeding 0.5 for a lead of up to eleven months. In comparison, the SINTEX-F has a lead of only four months for the same period. The convolution process allows CNN to train on a relatively small number of climate samples while identifying a basic shape that allows the precursors to properly affect the predictand even if the spatial distribution has been shifted or deformed.

Methodology

The variables considered for predicting the events are Sea Surface Temperature(SST) and oceanic Heat Content (HC) from time T-2 to T (T in months). The transfer learning technique has been employed to optimally train the model. Due to less number of samples available, the CNN model is first trained using CMIP5 (Coupled Model Intercomparison Project phase5) output. The trained weights are then used as initial weights to develop the final CNN model with the help of reanalysis. The observation data available from 1871 to 1973 is used for reanalysis after the transfer of weights. The validation period is from 1984 to 2017 with a gap of 10 years to eliminate any influence of oceanic memory in the training period on the validation period.

The CNN model used in the paper contains 3 convolutional layers and 2 max-pooling layers in between the first 2 convolutional layers. The max-pooling layers extract the largest value from a 2x2 grid. A total of 4 combinations of CNN models are used and the averaged value of outputs is considered. This reduces individual errors in the forecasts.

In the convolutional process, the local characteristics are extracted from global maps, and dot products are calculated between the convolutional filter and values in the input layer. The minimization of the cost function is used to iteratively decide the values of the convolutional filter. The heat map analysis is done using an output variable defining the output neuron which has a tanH activation function. By determining the contribution of each grid cell to the output, the heat map value is calculated. To define and train the model on the type of El Nino event, Nino3 and Nino4 index were used into a UCEI (unified complex ENSO index).

Critical Appraisal

The tropical climate variations like El Niño/Southern Oscillation (ENSO) often read to global variation and extremes. It would, therefore, be desirable to have a skillful year to year forecasts of such events. But, skillful forecast of ENSO events with more than one year of lead times remains a big challenge. The authors were available to achieve up to 17 months of lead time with skill forecasts of ENSO events using deep learning. This clearly shows the untapped potential of deep learning methods in climate forecasts.

Moreover, deep learning approaches such as Recurrent Neural Network (RNN) and Long Short Term Memory (LSTM) show a lot of promise to forecast global climate and hydrological events. With physically informed machine learning and hybrid models, there is potential to further improve the performance of statistical models both in terms of accuracy and lead times.