

# Self Learning Algorithm Based on Euclidean Distance for Imputation of Values

Mukta Agarwal

Vidya Bhawan Rural Institute, Udaipur, Rajasthan, India

## Abstract

The appearance and the development of computers made it possible to consider new techniques for data analysis, particularly in rapidly running some calculations such as matrix inversion or the diagonalization. This is the case of factor analysis for such analysis or principal components ACP- who responded to the need view the data. These methods realize indeed projections of said planes "main" (containing the maximum of inertia) which is best seen the cloud of points. But there is a need to supplement this by new methods which are capable of handling large data bases having a non-linear structure, and therefore poorly explained by traditional tools. Some neural networks, as multilayer perception, can replace conventional analysis. But they do not always offer graphical representations and interpretations are sometimes data analysis using the Kohonen maps difficult. The classification of Kohonen has, in turn, the double advantage to analyze on linear database and lend itself to graphical representations easily use.

## Keywords

Data Analysis, Matrix Inversion, Projections, Neural Networks, Multilayer Perception

## I. Formulation of Problem

The proposed method is based on exchanging absent feature. These strategies are incredibly much useful for statistical qualities and also accountable beneath the flag connected with chronological examination. Generally, most of these strategies are generally centralized about search connected with values that are quite near the middle propensity on the feature and also best towards the importance connected with simply just preceding and also doing well importance on the absent values.

## II. Algorithm

The process depends only on entries and does not require the intervention of a supervisor, one speaks in this case of self-organization .Its main function is to match the elements of the input space with units ordinate on a map - which is a graph where each unit is surrounded by its neighbors, neighborhoods have been defined a priori. The result is a function of the input space to all the units, such that the images of two neighboring elements in the sense of a certain distance in the inputs of the space are the same unit or neighboring units on the map.

1. Choose an integer  $k : 1 \leq k \leq n$ .
2. Calculate the distance  $d(Y_i; Y_j) ; i = 1; \dots; n$
3. Retain observations  $Y(i_1); \dots; Y(i_k)$  why these distances are the smallest.
4. Assign missing values average values of  $k$  neighbors:

$$(Y_{ij})_{\text{miss}} = Y_{i^*j^*} = \frac{1}{k} (Y_{(i_1)} + \dots + Y_{(i_k)})$$

In addition, notion of distance between individuals must be chosen carefully.

## III. Outcomes

Table-A offered inside appendix demonstrates the earth vast emission connected with co2 (CO2) from the consumption of Fat and also Natural gas respectively with the a long time 1960 in order to 09. Your necessarily mean emission connected with co2 (CO2) on account of fossil fuel, Fat and also Natural gas usually are 2109, 2260 and also 880 respectively.

Table-B demonstrates the actual factors using seen and also lacking valuations. It could be noted of which inside the organized means 20 % of the valuations usually are lacking inside the arbitrary fashion for all you factors from Table-A. Your implies measured from not whole facts pieces usually are 2259 for Fat and also 874 for Natural gas. It's seen which means that valuations connected with not whole facts pieces connected with Table-B usually are a little bit below the actual necessarily mean valuations from all of the several factors connected with Table-A.

Your offered straightforward can be used on the facts pieces connected with Table- M in order to fill the actual lacking valuations. Prices saved or generated with this strategy usually are demonstrated inside Table-C for equally factors which can be pointed out through underline. More, it really is seen that the necessarily mean valuations acquired after changing the actual lacking valuations with the best match valuations inside Table-C will be towards the true necessarily mean seeing that offered inside Table-A.

## IV. Conclusion

It's widely recognized that there's not really 100 % successful strategy of coping with lacking capability valuations. Your offered procedures are helpful for statistical capability, acquiring trivial change in the necessarily mean. This process is acceptable with the consolidated survey, in addition right and also suited to install specific lacking valuations. Below, the actual approximated price provides a likeness obtain in the former and also succeeding valuations. As a result, it really is seen of which processes for coping with connected with lacking capability valuations need to be preferred on their own or good characteristics and also form of facts.

## References

- [1] Ansley C.F., "An algorithm for the exact likelihood of a mixed autoregressive-moving average models", *Biometrika*, 66, pp. 59-65, 1979.
- [2] Atlas L., Connor J., Damborg M., "Comparisons or conventional techniques and neural networks in power system load forecasting and stability", In *Proceedings of the American Power Conference*, 1991.
- [3] Attali J. et Pagès G., "Approximation of functions by perceptrons: A new approach", *Neural Processing Letters*, 22, 5, 19, 1995.
- [4] Auriault J., Lesne H. Le modèle précoce (PRÉvion à COURT terme de la Consommation d'Énergie électrique), Note HR/25-1780, EDF5 Azencott R., Girard Y., astier R., Baudin M., Girard B., Jakubowicz P., et Martin M.

- [5] MANDRAKE: un progiciel expert en analyse de séries temporelles. Harcourt Brace and World, Paris. (1991).
- [6] Ball G.H, Hall D.J. A clustering technique for summarizing multivariate data. Behavioral Sciences, 12, pp. 153-155, (1967).
- [7] Barron A. Universal approximation bounds for superpositions of a sigmoidal function. I.E.E.E. Transactions on Information Theory, 39, pp. 390-945. (1993).
- [8] Benaïm M., Fort J.C, Pages G., Convergence of the one-dimensional Kohonen algorithm, advanced appl. Prob. 30, pp. 850-869
- [9] Bishop C.M., Svensén M., Williams C.K.I. GTM: a principled Alternative to the self-organizing map. In advances in Neural Information Processing Systems, Editor: MC Mozer and M.I Jordan and T. Pitche, pp. 354-360 (1997).
- [10] Bishop C. M., Tipping M.E., Jordan M. I. Hierarchical latent variable model for a data visualization. IEEE Transactions on Pattern Analysis and Machine Intelligence 20, pp. 281-293
- [11] Bolzern, P. and Fronza, G. Role of Weather Inputs in Short-Term Forecasting of Electric Load, Electric Power and Energy Systems, 8. 1, pp. 42-46. (1986).
- [12] Bouton C., Pagès G. Self-organization of the one-dimensional Kohonen algorithm with non uniformly distributed stimuli, Stochastic Process. Appl. 47, 1993, pp. 249-274.
- [13] Bouton C., Pagès G. Convergence in distribution of the one-dimensional Kohonen.