International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 21 (2018) pp. 15192-15195

Forecasting: A Review and Its Application to Handwritten Signatures

Dr. Alpana Deka[1], Dr. Lipi B. Mahanta[2]

[1]*Assistant Professor, Department of Computer Science, NERIM Group of Institutions, Guwahati, Assam, India.*
[2]*Associate Professor, Department of CCNS, IASST, Guwahati, Assam, India.*

Abstract

Forecasting helps us to predict the pattern of future data on the basis of present and past or historical data. There are three terms, or periods, of forecasting viz. short-term, medium-term and long-term forecasting. Generally, forecast for short period is more accurate than that of the long period since, as we move further ahead, the probability of uncertainty increases. The concept of forecasting is widely applicable in different areas like business related predictions, governmental budgets, policies and planning, other areas like weather as well as temperature forecasting etc [1]. The concept of forecasting can also be applied in offline handwritten signatures since there is presence of intrapersonal variations among signatures of same person. In the present study, we have put forward an overview about forecasting techniques and subsequently the application of its concept on offline handwritten signatures.

INTRODUCTION

The application of forecasting has great importance to estimate future data in various fields. The length of the forecast horizon decides different forecasting terms. Usually, duration of 1 day to 3 months, 3 months to 24 months and more than two years are recognized as short-term, medium-term and long-term forecasting respectively. Depending on the factors as described below, different forecasting models like Time Series, Cause-And-Effect and Judgmental methods can be applied to describe the future estimated data [2].

(i) Purpose of forecasting: Purpose of forecasting can be considered as one of the main factors for selecting a forecasting model. For example, forecasting of every 5 years government based planning will be different from that of daily weather or temperature forecasting. As the purpose of forecasting changes, the requirement of desired time period for collecting recorded historical data will also vary.

(ii) Availability of historical data: Since some of the forecasting models require more historical data than the other or sometimes historical data may not be present, therefore availability of historical data are taken into consideration during the selection of an appropriate forecasting model.

(iii) Nature of available historical data: Usually, in time series forecasting model, the most common patterned observed in historical data are trend, seasonal, cyclical and horizontal. Similarly the cause-and-effect model presents a relationship between independent and dependent variables, e.g. during rainy season, usually number of purchasing umbrellas increases. Again in some cases, where historical data are not available then the judgmental method can be applied. Therefore, in choosing a suitable forecasting model, the nature of availability of historical data plays an important role.

(iv) Length of forecast horizon: Depending on the length of forecast horizon, forecasting time period can be considered as short-term, medium-term and long-term. All the three different forecasting models deal with these three different kinds of forecasting horizons.

DIFFERENT FORECASTING MODELS

Each of the forecasting models can again be divided, most common of them are given in Fig. 1.

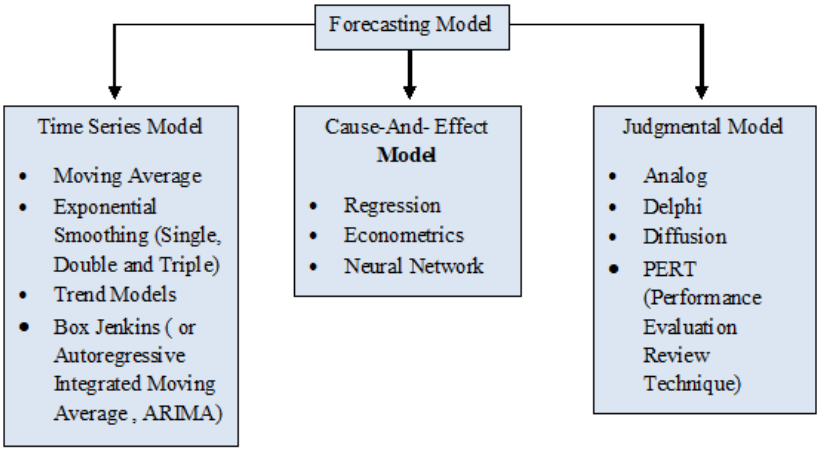

Figure 1: Overview of different forecasting models

In Time Series model, it is expected that the flow of data that exists in present and past, will also remain same in future. Again Cause-And-Effect model is a combination of independent and dependent variables (like relationship between sales and advertising of a product) which are known as cause and effect respectively. The time series along with cause-and-effect models are known as quantitative methods which are based on mathematical and statistical formulae. In Judgmental model, the opinions or judgments of experts help to estimate the future data and also termed as qualitative methods. [3].

From Fig. 1, it is clear that the time series model have different forecasting models. Each of these models have different criteria as revealed in Table 1 [4-6].

Table 1: Different time series forecasting models

Time series forecasting model	Number of historical data required	Behavior of historical data	Forecast Horizon
Moving Average	2 to 30 observations	Stationary	Short
Single Exponential Smoothing	5 to 10 observations	Stationary	Short
Double Exponential Smoothing	10 to 15 observations	Trend	Short
Triple Exponential Smoothing (Holt- Winters method)	10 to 15 observations	Trend and seasonality	Short
Trend Model	10 to 20 observations	Trend	Short
Box Jenkins (or Autoregressive Integrated Moving Average Method, ARIMA)	Minimum of 50 observations	Stationary	Short, Medium and Long

LITERATURE REVIEW OF PREDICTIVE MODELING SYSTEM

Hejase et al. [7] explained ARIMA (Auto Regressive Integrated Moving Average) modeling to predict the mean daily global solar radiation of AI-Ain, United Arab Emirates (UAE). Here, to build and validate the model, they applied 10 (1995 – 2004) and 3 years (2005 – 2007) data respectively.

Lim et al. [8] presented an hourly basis forecasting system for solar irradiance and load demand with single exponential smoothing method with two different sets of data. The first approach related with current day data and the second one related with previous day data. Comparative result showed that the second approach achieved higher accuracy than the first one.

Bakari et al. [9] designed a predicted model for production and utilization of gas from Nigeria National Petroleum Company for the time period 1970-2004. Here to build the predicted model, they applied ARIMA model with Box-Jenkins methodology. In the proposed model, they forecasted the expected amount of gas production and utilization for four years.

Biswab et al. [10] presented ARIMA based modeling to forecast area and production of rice in West Bangal. Here, the data for 1947 to 2004 were utilized to develop the model and another three years data applied as testing data.

Gupta et al. [11] discussed an artificial neural network (ANN) based forecasting model for Indian rainfall. Here, out of the total collected data, they considered 130 years data for training and another 10 years data to test the model.

Muhammad et al. [12] proposed production based forecasting of paddy in Kelantan, Malaysia, using the data from 1970 - 2009. Here, they applied Brown's and linear's double exponential smoothing, Winter's multiplicative exponential smoothing as well as damped trend and random walk technique. In the proposed system, they achieved the combined model of double and damped trend as the most accurate method.

Sahu et al. [13] evaluated demand forecasting for sales of milk product (paneer) in Chhattisgarh on the basis of weekly data from October 2011 – October 2012. Here, they employed single and double moving average methods, single exponential smoothing, semi average method and naive method where the single exponential smoothing method attained as the most accurate one among them.

Popale et al. [14] described a modeling technique with ARIMA model for generation and forecasting of weekly rainfall for Rahuri region of Ahmednagar. Data were collected from the region for time period of 1982-2012. Here, to develop the model they employed the data for 1982-2011 and remaining one year used to validate the model.

Choudhury et al. [15] analyzed time series models for prediction of crop yield in Ghana. They employed the simple and double exponential smoothing, damped-trend linear exponential smoothing, and ARMA models by utilizing 17 years data (1992 – 2008) where ARMA model exhibited as the best among them.

Singh et al. [16] described a time series model for forecasting of boot in shoe industry. Here, they collected the data on monthly basis from April 2012 to March 2015. They applied moving average, single exponential smoothing method, double exponential smoothing method (Holt's method) and Winter's method where single exponential smoothing method found to be as the best one.

RELEVANCE OF FORECASTING IN HANDWRITTEN SIGNATURES

A handwritten signature represents a person's identity. Although the signature of a person may vary during his lifetime, but its basic pattern or style does not alter [17]. Usually a person follows a same signature to sign any documents, but due to security purpose some persons prefer two forms of signatures: one is formal and another informal. They apply the formal signatures in banking or financial transactions and any legal documents etc. Similarly, they sign with informal signatures in cases like autographs and any social documents etc. Between the two signatures, we are mostly required to verify the formal ones. Although overall patterns among the signatures of the same person remain unchanged, there can appear natural variations because two or more signatures of same person written within a moment (or keeping a sufficient time gap). These natural variations result as intrapersonal variations. During sample collections, some of the common natural variations that are mostly observed among the signatures of same person are listed below [18,19]:

(i) Written quality of letters: The quality of written letters within the signatures may be normal, irregular or shaky.
(ii) Spacing between the letters or words: During the signing process, usually spaces between the letters or words vary within a signature at each time.
(iii) Connectivity and continuity: Connectivity and continuity between the words can also create intrapersonal variation because sometimes these two characteristics are not uniform among the signatures.
(iv) Pen pressure: Every time, the pressure applied during the signing process may not be uniform.
(v) Skew angle: Since in each signing process, imaginary baselines of signatures of the same person do not remain constant, therefore the skew angle generated between the imaginary baseline and horizontal direction of a signature can vary. Based on the position of baselines, three different skewed angles viz. positive, zero and negative angles can generate in a signature.
(vi) Slant angle: Slant angle is the angle of measurement between vertical text strokes and vertical direction of a signature. The variation of slant angles also generate positive, zero and negative angles which result intrapersonal variation.
(vii) Misplacement of cross and dot marks: Usually, it is observed that the cross and dot marks tend to misplace among letters like E, F, t and i, j etc.
(viii) Stroke marks: Beginning and ending stroke marks within a signature are not uniform during each signing process. Mostly, a person completes his signature with stroke marks in unusual way.

These most common variations, as presented above, may appear among the signature samples of same person, signed within a moment or the signatures that are signed by keeping sufficient time gap. These variations occur due to several factors or conditions under which a signing process completes. Such possible common conditions are explained below [20,21]:

i) Body position: While signing a document, body position of a person can be considered as one of the remarkable factors. The pressure applied with pen on a paper will be different with different body positions e.g. during singing process, pressure applied by a sitting person will be different from that of a standing person. It is also important to notice whether a person's body is free from all perspectives or there is any burden on the signing hand.
ii) Writing material: Signature will look different on papers of different quality. Since the way of handling a pen or pencil is different to some extent, hence writing material may impact on signature. Usually the structural forms of ball or ink pens of different companies are different and hence their affects cannot be neglected.
iii) Purpose of signing: Signature is usually significantly different if taken in formal environment than in informal. Since in informal environment, a person does not have any mental pressure, therefore purpose of writing may be considered as a factor on which handwritten signature depend.
iv) Environmental factors: Mostly the surrounding environment, during each signing process does not remain uniform. This includes noise, calm and quiet, temperature, humidity, etc. Obviously an environment which is full of noises will be different from that of calm and quiet. Similarly, high temperature weather makes the atmosphere different from low temperature weather.
v) Physical and psychological state of the person: The shape and pattern of a signature may depend on factors of a person like his illness, injuries, fears, heart rate, person's age, calmness, goodwill etc. These factors have influence to increase the fluctuation of shape of a signature. Experiments reveal that person's age is also an important factor. The nature of signature may vary with respect to young age, middle age and old age.

Thus, it is clear that due to the presence of natural variations or intrapersonal variations, we may successfully forecast how far signatures of a person can vary during his life time. This would assist us in differentiating between actual fraud cases of a signature from personal variations of the same. If the case is that of just a personal variation it would predict the same as a genuine signature and relieve the concerned person from unnecessary harassment of having to go through all the paraphernalia of a fraud case.

CONCLUSION

The forecasting or predictive modeling system has wide application, from weather to governmental policy planning. Similarly, the time factor and intrapersonal variations are two most important factors, for which we can apply the concept of predictive modeling system to estimate future behavior of a signature. Because, during his/her life time, a signature may change to some extent by maintaining approximately the same pattern. The application of forecasting techniques to distinguish fraud cases of signature from genuine one, simply due to interpersonal variations, would be very useful as well as widely relevant.

REFERENCE

[1] Shah, J., 2009, "Supply Chain Management: Text and Cases", *Pearson Education India*, pp: 165-167.

[2] Jain, T.R. and Aggarwal, S.C., 2009-10, "Business Mathematics and Statistics", *FK Publications*, pp: 342-344.

[3] Jain, C. L., and Malehorn, J., 2006, "Benchmarking Forecasting Practices: A Guide to Improving Forecasting Performance", *Institute of Business Forec*, pp: 46-50.

[4] Keating, B., and Wilson, J. H. 2003, "Managerial Economics: An Economic Foundation for Business Decisions", *Dreamtech Press*, pp: 313.

[5] Stevenson, W. J., 2009, "Operations Management", *Tata McGraw-Hill Education*, pp: 93-105.

[6] Wang, G. C. S., and Jain, C. L., 2003, "Regression Analysis: Modeling & Forecasting", *Institute of Business Forec*, pp: 151.

[7] Hejase, H. A. N., and Assi, A. H., 2012, "Time-Series Regression Model for Prediction of Mean Daily Global Solar Radiation in AI-Ain, UAE", *International Scholarly Research Network, ISRN Renewable Energy*, pp: 1-11.

[8] Lim, P. Y., and Nayar, C. V., 2012, "Solar Irradiance and Load Demand Forecasting based on Single Exponential Smoothing Method", *IACSIT International Journal of Engineering and Technology*, 4(4), pp: 451-455.

[9] Bakari, H. R., Chamalwa, H. A., and Mohammed, A. D., 2013, "Time Series Analysis Model for Production and Utilization of Gas(A Case Study of Nigeria National Petroleum Corporation "Nnpc") ", *IOSR Journal of Mathematics*, 9(1), pp: 17-23.

[10] Biswab, R., and Bhattacharyya, B., 2013, "ARIMA Modeling to Forecast Area and Production of Rice in West Bengal", *Journal of Crop and Weep*, 9(2), pp: 26-31.

[11] Gupta, A., Gautam, A., Jain, C., Prasad, H., and Verma, N., 2013, "Time Series Analysis of Forecasting Indian Rainfall", *International Journal of Inventive Engineering and Sciences*, 1(6), pp: 42-45.

[12] Muhammad, M., and Abdullah, M. H. H., 2013, "Modelling and Forecasting on Paddy Production in Kelantan under the Implementation of System of Rice Intensification (SRI)", *Academia Journal of Agricultural Research*, 1(7), pp: 106-113.

[13] Sahu, P. K., and Kumar, R., 2013, "Demand Forecasting for Sales of Milk Product (Paneer) in Chhattisgarh", *International Journal of Inventive Engineering and Sciences*, 1(9), pp: 10- 13.

[14] Popale, P. G., and Gorantiwar, S. D., 2014, "Stochastic Generation and Forecasting of Weekly Rainfall for Rahuri Region", *International Journal of Innovative Research in Science, Engineering and Technology*, 3(4), pp: 185-196.

[15] Choudhury, A., and Jones, J., 2014, "Crop Yield Prediction Using Time Series Models", *Journal of Economic and Economic Education Research*, 15(3), pp: 53-67.

[16] Singh, A. P., Gaur, M. K., KumarKasdekar, D., and Agrawal, S., 2015, "A Study of Time Series Model for Forecasting of Boot in Shoe Industry", *International Journal of Hybrid Information Technology*, 8(8), pp: 143-152.

[17] Vogel, D., 1999, "Financial Investigations: A Financial Approach to Detecting and Resolving Crimes", *DIANE Publishing*, pp: 285.

[18] Bertino, A., and Bertino, P., 2011, "Forensic Science: Fundamentals and Investigations 2012 Update", *Cengage Learning*, pp: 279-282.

[19] Houck, M. M., and Siegel, J. A., 2015, "Fundamentals of Forensic Science", *Academic Press*, pp: 525.

[20] Fotak, T., Baca, M., and Koruga, P., 2011, "Handwritten Signature Identification using Basic Concepts of Graph Theory", *Wseas Transactions on Signal Processing*, 7(4), pp: 117-129.

[21] Mahanta, L. B., and Deka, A., 2013, "A Study on Handwritten Signature", *International Journal of Computer Applications*, 79(2), pp: 1-5.