

# FORECASTING SALES THROUGH TIME SERIES CLUSTERING

Monica Sanwlani<sup>1</sup> and Prof.M.Vijayalakshmi<sup>2</sup>

<sup>1</sup>M.E.[Information Technology]student, Department of Information Technology,  
V.E.S.I.T,Chembur, Mumbai, Maharashtra

moni0308@gmail.com

<sup>2</sup>Asst. Professor And HOD,Department of Information Technology, V.E.S.I.T,  
Chembur,Mumbai, Maharashtra

viji.murli@gmail.com

## **ABSTRACT**

*We study the effect of decomposing time series into multiple components like trend, seasonal and irregular and performing the clustering on those components and generating the forecast values of each component separately. In this project we are working on sales data. Multiple forecast experts are used to forecast each component series. Statistical method ARIMA, Holt winter and exponential smoothing are used to forecast these components. We performed clustering for forecasting and discovered a set of best, good and bad forecasters. Selection of best, good and bad forecasters is performed on the basis of count and rank of expert id's generated. Since we have thousands of experts, we experiment with combining method to get better forecast. Finally absolute percentage error (APE) is used for comparing forecast.*

## **KEYWORDS**

*Time Series, Decomposition, Combining, Sales Forecasting, Clustering, APE*

## **1. INTRODUCTION**

The accuracy of a sales forecast has great impact on manufacturing and sales. Sales forecasting is an important part of supply chain management both at the retailer end and at the distributors, manufacturers and suppliers. Timely and accurate sales forecasts are crucial in bridging the gap between supply and demand, thereby decreasing holding cost while maintaining a negligible probability of stock-out.

Since market conditions are volatile and depend on too many factors, it is really hard for a traditional forecast system to be able to forecast demand with high accuracy. In this scenario Data Mining provides an alternative viable solution.

Forecasting is stated as predicting future values [4]. Forecasting methods can be qualitative and quantitative. Qualitative techniques are more useful in the earlier stages of the product life cycle. Quantitative forecasting techniques are generally more objective than their qualitative counterparts. Quantitative forecasts can be time-series forecasts (i.e., a projection of the past into the future) or forecasts based on associative models (i.e., based on one or more explanatory variables). Our research uses quantitative statistical forecasting methods such as Holt-Winter, ARIMA (Auto-Regressive Integrated Moving Average) and exponential smoothing, as base models to design our Forecasting Component [5].

Decomposition is applied on time series data [6] to calculate trend (T), seasonality (S) and irregular (I) components. Multiple models are used with each series for decomposition. We use a combination forecast, where each forecasting expert is a triplet generated by the Cartesian product of the sets of forecasters for the T, S and I components.

We have used multiple experts for our sales series and for individual components of the series obtained after applying basic decomposition. Since we generate a large number (about hundred thousand), we need a method to identify a subset of good models that can improve the forecasting accuracy. We propose the use of standard clustering techniques to solve this problem.

We are clustering the experts based on their performance. Clustering of Best, Good and Bad experts at each point of time is done for 50% and 70% of the series as test data.

Finally to check the performance of every series APE (Average Percentage Error) is calculated. As error measure plays an important role in calibrating and refining forecasting model/method. This calibration helps analyst to improve forecasting method. The choice of an error Measure may vary according to the situation, number of time series available and on whether the task is to select the most accurate method or to calibrate a given model.

The outline of paper is as follows. In section 2, we present literature survey. Section 3, specifies our proposal and implementation. Section 4, consist of graphical representation of original and forecast values of sample series. APE (Average Percentage Error) for each series is tabulated. Overall Experiments and results are represented in this section. We compared all the series APE values. A summary of APE Values of all the series and conclusion are included in section 5.

## 2. LITERATURE SURVEY

### 2.1. Time Series

A time series is a sequence of observations taken sequentially in time[6] An intrinsic features of time series it that, typically, adjacent observations are dependent. Time series analysis is concerned with the technique of analysis of this dependence [6].

The main objective of the time series analysis is to model a process, which is generating the data, to provide compact description and to understand the generating process. To allow for the possibly unpredictable nature of the future observations, selection of the probability model for the data is very important. We can define a time series model for the observed data {xt}, to be a specification of the joint distribution of a sequence of random variables {Xt} of which xt is postulated to be a realization.

Time series forecasting [6] refers to the use of a model to forecast future values of the series, based on known past values before they are measured.

Time	Observations									
1 - 10	4	16	12	25	13	12	4	8	9	14
11 - 20	3	14	14	20	7	9	6	11	3	11
21 - 30	8	7	2	8	8	10	7	16	9	4

Table1. Represents sample weekly demand for weeks 1 through 30

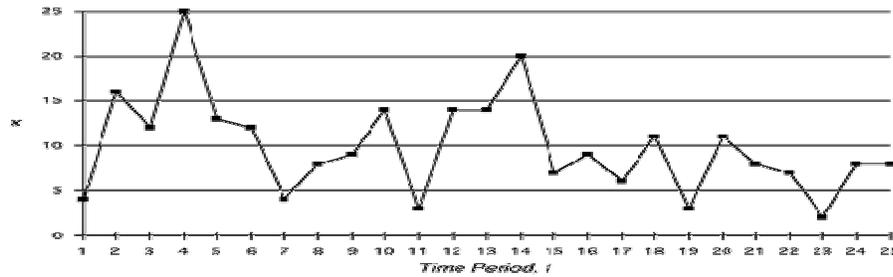


Figure.1 Graph representation of above time series data of weekly demand

## 2.2. Forecasting Models

Forecasting Models [3] are the mathematical model which make certain assumptions about the time series, captures certain features and then predicts the next unknown value.

### 2.2.1. Holt's Linear Exponential Smoothing (LES)

Holt's two parameter exponential smoothing method is an extension of simple exponential smoothing. This method is used when a series has no seasonality and exhibits some form of trend. This is an extension of exponential smoothing to take into account a possible linear trend.

Suppose that the series  $\{ y_t \}$  is non-seasonal but does display trend. Now we need to estimate both the current level and the current trend. Here we define the trend  $b_t$  at time  $t$  as the difference between the current and previous level. There are two smoothing constants  $\alpha$  and  $\beta$ .

The equations are:

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1})$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}$$

$$F_{t+m} = L_t + b_t m$$

Here  $L_t$  and  $b_t$  are respectively (exponentially smoothed) estimates of the level and linear trend of the series at time  $t$ , whilst  $F_{t+m}$  is the linear forecast from  $t$  onwards.

### 2.2.2. Holt-Winter Method

The Holt-Winter's method is one of the best known forecasting technique for the time series that has both trend and seasonal components. There are two variants of this method, additive and multiplicative. The seasonality is multiplicative if the magnitude of the seasonal variation increases with an increase in the mean level of the time series. It is additive if the seasonal effect does not depend on the current mean level of the time series. The basic Holt-Winter forecasting method with multiplicative seasonality (exponential smoothing of level ( $S_t$ ), trend ( $T_t$ ) and seasonal index ( $I_t$ )) is described by

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1}$$

$$S_t = \alpha(D_t/I_{t-p}) + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad I_t = \gamma(D_t/S_t) + (1 - \gamma)I_{t-p}$$

Here  $p$  is the number of observation points in a cycle ( $p = 12$  for monthly data).  $\alpha$ ,  $\beta$ , and  $\gamma$  are the smoothing constants.

The forecast at time  $t$  for time  $t + i$  is:

$$D_{t+i} = (S_t + i \times T_t)I_{t-p+i}$$

### 2.2.3. Autoregressive Integrated Moving Average (ARIMA)

Models for time series data can have many forms. When modelling variations in the level of a process, three broad classes of practical importance are the *autoregressive* (AR) models, the *integrated* (I) models, and the *moving average* (MA) models. These three classes depend linearly on previous data points. Combinations of these ideas produce autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) models.

The Autoregressive Integrated Moving Average (ARIMA) models, or Box-Jenkins methodology, are a class of linear models that is capable of representing stationary as well as non stationary time series. ARIMA models rely heavily on autocorrelation patterns in data.

The Box- Jenkins ARIMA models are , most general class of models for forecasting a time series which can be done by transformations such as differencing and lagging.

ARIMA methodology of forecasting is different from most methods because it does not assume any particular pattern in the historical data of the series to be forecast. An ARMA model predicts the value of the target variable as a linear function of Lags of differenced series appearing in forecasting equation lag values (the auto-regressive part) plus an effect from recent random values (the moving average part). A time series which needs to be differenced to be made stationary is said to be an "integrated" version of a stationary series.

Autoregressive (AR) models can be coupled with moving average (MA) models to form a general and useful class of time series models called *Autoregressive Moving Average (ARMA)* models. These can be used when the data are stationary. This class of models can be extended to non-stationary series by allowing the differencing of the data series. These are called *Autoregressive Integrated Moving Average (ARIMA)* models.

Seasonal ARIMA models a pattern that repeats seasonally over time. It is classified as an **ARIMA(p,d,q)x(P,D,Q)** model,

P=number of seasonal autoregressive (SAR) terms,

D=number of seasonal differences,

Q=number of seasonal moving average (SMA) terms

Non Seasonal ARIMA Models are classified as an "ARIMA(p,d,q)" model, where:

**p** is the number of autoregressive terms,

**d** is the number of non seasonal differences, and

**q** is the number of lagged forecast errors in the prediction equation.

## 2.4. Decomposition

A time series is a composition of many individual underlying component time series. Some of these components are predictable whereas other components may be almost random which can be difficult to predict. Decomposing a series into such components enables us to analyze the behaviour of each component and this can help to improving the accuracy of the final forecast. Decomposition is applied on original series to calculate trend, seasonality and irregular. Decomposing a series into such components enables us to forecast better. A typical sales time series can be considered to be a combination of three components i.e. trend component (**T**), seasonal component (**S**), and irregular component (**I**).

- **The Trend** represents changes in the level of the series. Trend does not imply a monotonically increasing or decreasing series but simply the lack of a constant mean.
- The **Seasonal** component models patterns of change in a time series within a year. These patterns relate to periodic fluctuations of constant length, tend to repeat themselves each year. It refers to a cycle of fixed period (e.g. weekly, monthly, yearly etc.).
- The **Irregular** component refers to variations or random pattern.

Mathematical representation s for decomposition is

$$Y_t = f(T_t, S_t, I_t)$$

Where  $Y_t$  = Actual value at point  $t$

$T_t$  = Trend Component at point  $t$

$S_t$  = Seasonal Component at point  $t$

$I_t$  = Irregular Component at point  $t$

The decomposition could be additive if the magnitude of seasonal fluctuations do not vary with the level of the series.

$$Y_t = T_t + S_t + I_t$$

- We have a multiplicative decomposition if seasonality fluctuates and increases and decreases with the level of the series.

$$Y_t = T_t \times S_t \times I_t$$

- Multiplicative model is more prevalent with economic series since most seasonal economic series have seasonal variation which increases with the level of the series. In this project we shall thus use the multiplicative decomposition model.
- A method used in forecasting a single component is referred to as a forecasted value. A forecasted value for the original series is a triplet made up of the forecasted value for each component. The set of such triplets is the Cartesian product of the sets of forecasters for the T, S and I components. Each such triplet of forecasted value (T, S, I) is called an “expert”.

- Trend (T)

$$T_t = \begin{cases} \frac{D_1 + D_2 + \dots + D_{12}}{12} & \text{for } t = 1 \text{ to } 12 \\ \frac{D_{t-11} + D_{t-10} + \dots + D_{t-1} + D_t}{12} & \text{for } t > 12 \end{cases}$$

- Seasonality (S)

$$DT_t = \frac{D_t}{T_t}$$

$$S_t = \text{Average}(DT_t, DT_{t-12}, DT_{t-24}, \dots, DT_{t-k}) \text{ till } t - k > 0$$

- Irregular Component (IC)

$$IC_t = \frac{DT_t}{S_t}$$

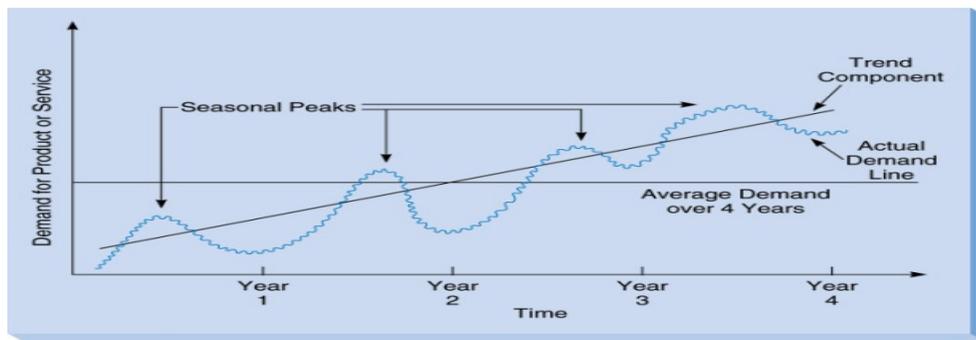


Figure 2. Components of Time Series

## 2.5. Motivation for Combining

There exist numerous series with different attributes. It is not possible for any one method to perform better on all kind of series. Some series can be forecasted better by application of one forecasting methods, whereas same methods will perform badly when applied to another kind of series. Different forecasting methods are designed to catch different attributes. As no single method can consider all important aspects of time series, we have used multiple forecasting experts. Now using multiple set of experts, we can try to consider many attributes of a series.

To get final forecast, a function can be applied on the forecast of various experts. There are many forecasting techniques used in time series forecasting. Typically one technique is selected based on a selection criterion, hypothesis testing, and/or graphical inspection, and selected model is used for the forecasting. However, model selection is often unstable. To overcome the above problems, combining techniques can be useful.

There are two most popular and simple methods to combine forecasts. One is taking mean and another is taking median of forecasts from different forecasting methods.

In combining, we combine the forecast from different forecasting experts to generate final forecast. This combining function can be as simple as taking mean to as sophisticated as applying dynamic function over dynamic set of forecasting experts. Combining forecasts can be used for risk minimization as it reduces the variance of the final forecast. That means worst MAPE will follow non-increasing improvement as the number of experts in combining increases. Combining techniques can also be used to improve forecast accuracy.

In this work, we use a total of 86 Trend models (atomic forecasters), 33 Seasonal models and 34 Irregular component models. These are mostly ARIMA and seasonal ARIMA models of different orders. The Cartesian product of the Trend, Seasonal and Irregular models gives rise to 96,492 expert forecasts per point. The Appendix A includes a list of atomic forecasters used in this work.

## 2.6. Clustering for Forecasting

Clustering could be “the process of organizing objects into groups whose members are similar in some way” [2]. A *cluster* is therefore a collection of objects which are “similar” between them and are “dissimilar” to the objects belonging to other clusters [1]. Clustering is making an impact in this area for selecting a model for a given series.

We choose experts to forecast each component series. We perform forecasts for each series separately which are then used to derive the forecast for the original series.

In our approach we are clustering the experts. Clustering of Best, Good and Bad experts at each point of time is done.

For performing clustering, we calculated count and rank. Where count is frequency of appearance of expert id And Rank is summations of each row number divided by number of rows in particular series.

Clustering is done for 2 years (last 24 points) data on 70% and 50% of the series as test data. Finally to check performance at each point of time t, we calculated APE (Absolute percentage Error) of each expert for comparing forecasts.

### 3. OUR APPROACH

Forecasting large numbers of time series is a costly and time-consuming exercise. Before forecasting a large number of series that are logically connected in some way, we can first cluster them into groups of similar series. Similar series are first grouped together based on count and rank. The series in each cluster are then pooled together and their forecasts are obtained. Finally absolute percentage error (APE) is calculated.

#### 3.1. Preliminaries

##### Data used

Original series is time series data. We have used 31 series from time series data library. The time series data used is monthly sales data. Most of the series represent sales of Fast Moving Consumer Goods. All the series are listed in Appendix B. The below figure shows original time series data represented by graph of Abraham series.

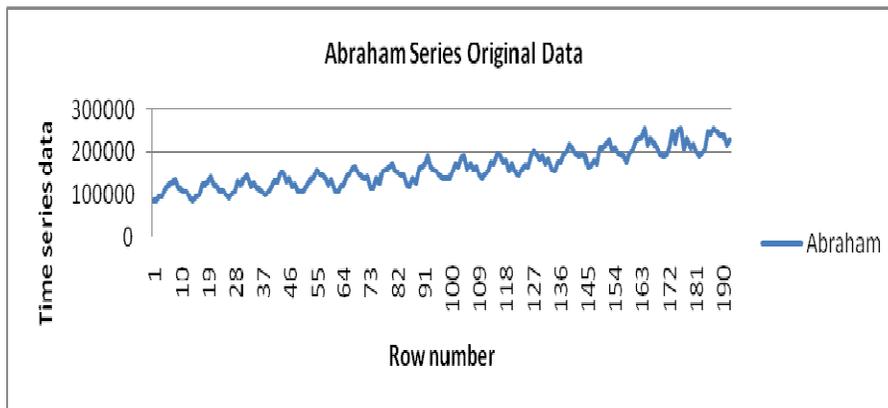


Figure 3. Graph represents Original Series of Abraham

##### Experts Used

The various experts are obtained from a Cartesian product of 86 Trend Experts, 33 Seasonal Experts and 34 Irregular Experts ( $96492 = 86 \times 33 \times 34$ ). All these individual experts belong to the ARIMA family. These experts are listed in Appendix A.

### **Calculated Predicted Values using ARIMA Models by using Experts**

We have used multiple experts (forecasters) for forecasting individual component series, which gives multiple forecasted values for each component. We have performed Cartesian product of individual component forecasts which ultimately gives us a large number of experts.

We have used multiple experts for our sales series and also multiple experts for individual components of the series obtained after applying basic decomposition and by using combining methods. The models considered for constructing the expert are ARIMA (including Seasonal ARIMA) models of various orders and various Exponential Smoothing Models such as Holt Winter Model, Holt Model, Seasonal Exponential Smoothing, etc in SPSS Tool. Some Appropriate transformation like logarithmic transformation is often used for time series that show exponential growth or variability proportional.

These predicted values are calculated in IBM SPSS tool which is used to analyze historic data and predict values. Original series are forecasted by this set of experts. Most of the experts belong to the ARIMA family.

In practice, we do not require higher order models such as ARIMA(4,0,0), so we were able to restrict the pool size to a reasonable limit.

### **3.2. Implementation**

#### **Model Collection and by Selecting Top (Best) 20000 and Bottom (Bad) 20000 samples after performing sorting operation.**

As we have used multiple experts so by performing Cartesian product we got 96,492 forecasted values. Individual component of the series are obtained by applying basic decomposition. We have selected and sorted top 20,000 and bottom 20,000 predicted values to perform clustering on the basis of best and bad. Further best models are clustered in best and good. A set of top K experts (24 months data) is chosen at this point t and are combined by taking average of trend, season and irregular values to give us a forecast at time t.

#### **Forecasting technique for Clustering (Performed on the basis of count & rank)**

Based on frequency of model appearance we calculated count and rank. Where Count is total count of appearance of expert id And Rank is summations of each row number divided by number of rows in particular series.

Clustering is done on expert id's. Clustering of Best, Good and Bad experts at each point of time is done. Clustering is done for 2 years data on 70% and 50% data. Finally we calculated APE (Absolute percentage Error) of each expert id.

In the same way we obtained forecast for all time instances for a particular series.

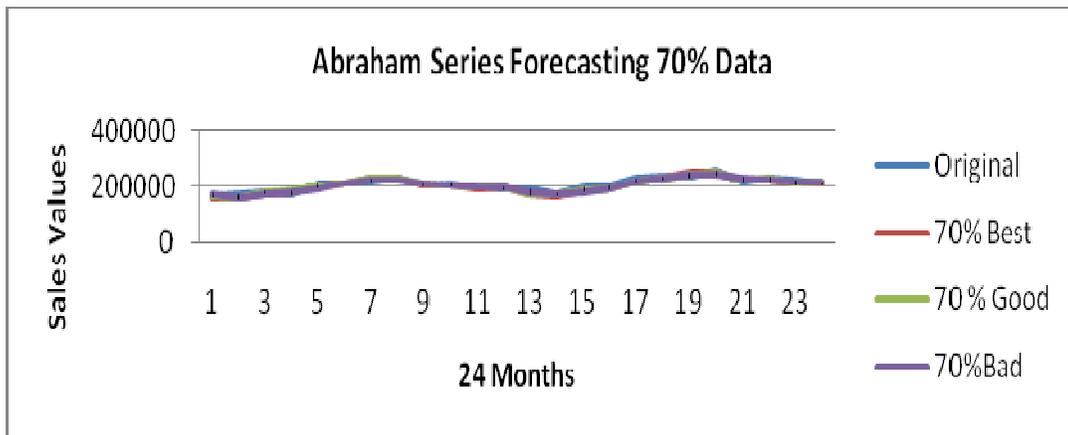


Figure 4. Sample line graph shows original values, best, good and bad forecasted values of 2 years 70% data

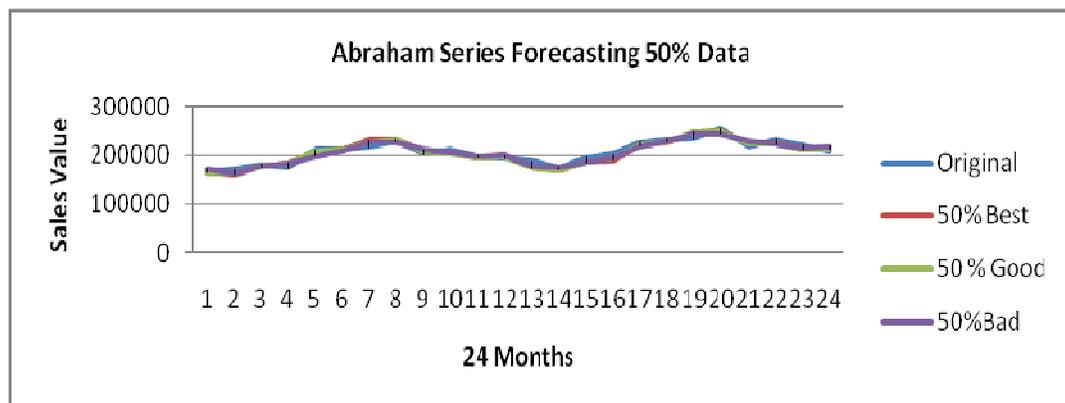


Figure 5. Sample line graph shows original values, best, good and bad forecasted values of 2 years 50% data

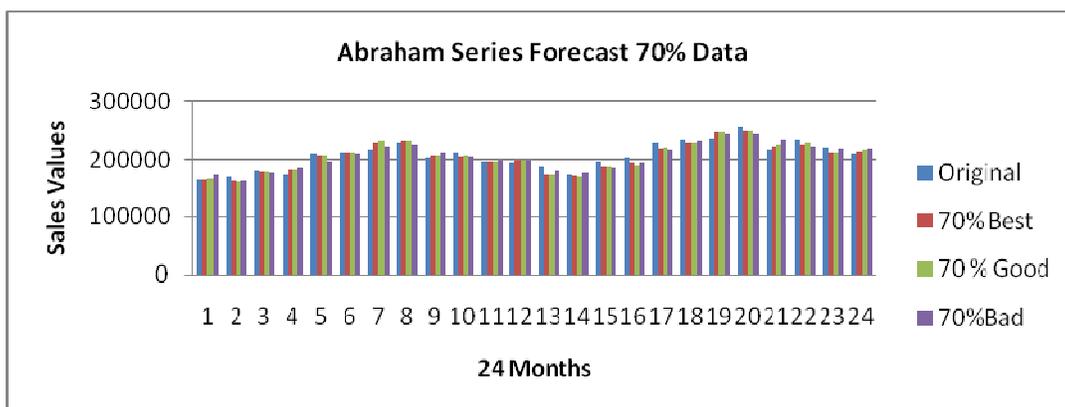


Figure 6. Sample bar graph shows original values, best, good and bad forecasted values of 2 years 70% data

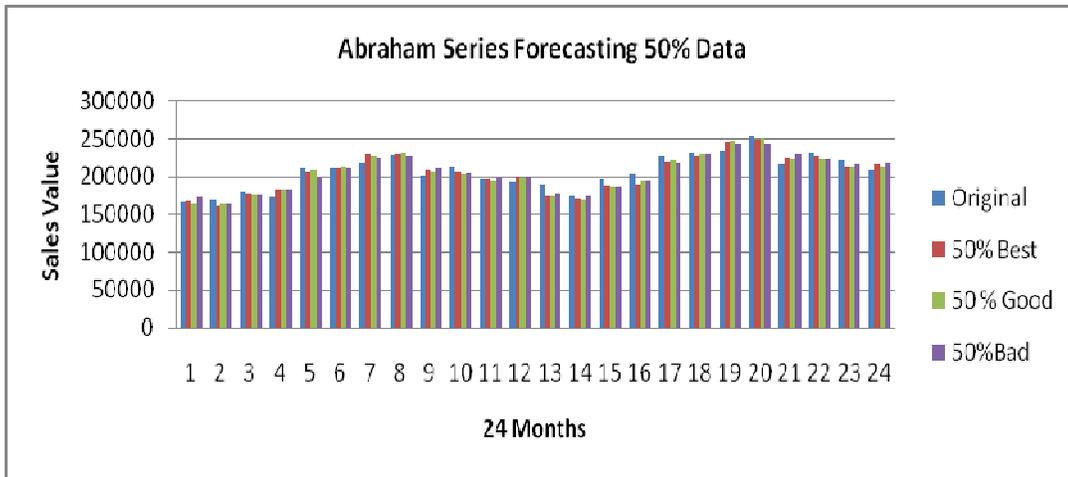


Figure 7. Sample bar graph shows original values, best, good and bad forecasted values of 2years 50% data

Figure 4, Figure 5, Figure 6 and Figure 7 are line and bar graphs. This represents forecasted values of last 24 months data on 70% and 50% data. They represent clusters as Best, Good and bad.

**Error Measure:**

At each point of time t, we calculate the APE (Absolute Percentage Error) of each expert, which is obtained by looking at the full past performance (up to time t-1). A set of top K experts are then chosen at this point t and are combined by taking simple mean of them, to give us a forecast at time t. In this way, we obtain forecasts for all time instants. APE is used to compare forecast.

N=Total Number of experts

L=Training length used by each expert i(i=1,N)

R<sub>t</sub>(i)= The ith top expert based on performance in [L+1,t]

F<sub>t</sub>(i)=Forecast of expert i at time t.

APE<sub>t</sub>(F)= Absolute percentage error of forecast F at point t.

MAPE<sub>t</sub>(F)= Mean Absolute Percentage Error for forecast F at point t.

$$APE = \frac{\sum_{i=0}^N (|forecasted_i - actual_i|) / actual_i}{N}$$

$$MAPE = 100 \times \frac{\sum_{i=0}^N (|forecasted_i - actual_i|) / actual_i}{N}$$

#### 4. EXPERIMENTS AND RESULTS

Sr No	Series Name	50% Best	50% good	70% Best	70% Good	50% Bad	70% Bad
1	Abraham	0.0299863	0.03219599	0.0299099	0.03224231	0.0347322	0.036618
2	Beer	0.024747	0.02465168	0.0246616	0.02458841	0.0250413	0.0250734
3	Clothing	0.0202779	0.02018791	0.0212176	0.01945408	0.0326225	0.0325787
4	Dry	0.094656	0.07549753	0.0934565	0.07548959	0.0838144	0.0941902
5	Equip	0.0197278	0.00985504	0.0132304	0.00995827	0.0143344	0.0099582
6	fortif	0.0661231	0.068624	0.065888	0.0686132	0.0662962	0.069088
7	Furniture	0.0258229	0.02344734	0.0256853	0.02356077	0.025832	0.0258869
8	gasoline	0.0364762	0.03554662	0.0363849	0.03559799	0.0366672	0.0360093
9	grocery	0.0162927	0.01711382	0.0159556	0.01724681	0.0184654	0.0172904
10	Hsales	0.0611422	0.05981831	0.0604853	0.05975369	0.0622628	0.0622817
11	Jewellery	0.030303	0.04552385	0.0301646	0.04527499	0.0576095	0.1029838
12	Merchandise	0.0093163	0.00892435	0.0148031	0.00870464	0.0128422	0.012621
13	Motorparts	0.0157009	0.01771644	0.0173973	0.01748443	0.0390029	0.039042
14	Newcar	0.0567432	0.05569477	0.0562089	0.0562089	0.0587326	0.0560556
15	paper	0.0652174	0.14107241	0.0650062	0.15207712	0.1398417	0.1491924
16	red	0.1482687	0.10319634	0.1524638	0.10354454	0.1719871	0.1721796
17	rose	0.1148993	0.12604517	0.1119353	0.12527912	0.1090942	0.1091095
18	Shoe	0.0311165	0.03497442	0.0311439	0.03483914	0.0353004	0.0351081
19	Software	0.0529116	0.04125589	0.0447838	0.04106998	0.0510674	0.0562282
20	spaper	0.0665146	0.06895758	0.066538	0.06915086	0.0839641	0.0888004
21	spark	0.0970054	0.08917429	0.0948307	0.08962883	0.0934534	0.0952212
22	stores	0.0208209	0.01324628	0.0180337	0.01325274	0.0245755	0.0255036
23	sweet	0.1166388	0.09805478	0.1152108	0.09795621	0.1008301	0.1030735
24	total	0.0067185	0.00875485	0.0070169	0.00912463	0.0196169	0.0195977
25	wine	0.0804485	0.07397853	0.0742661	0.07443381	0.0793084	0.0795037
26	syna1	0.000587	0.00486907	0.0006657	0.00425974	0.0048914	0.0049114
27	synb1	0.009258	0.0143174	0.0100376	0.01400302	0.0227773	0.0241911
28	sync1	0.0059299	0.01368244	0.005303	0.01349034	0.0217301	0.0229744
29	sync2	0.0186944	0.02676955	0.0187322	0.02673309	0.0245397	0.0259439
30	synd1	0.0063868	0.01676686	0.0064137	0.03017137	0.0493578	0.0491411
31	synd3	0.0544567	0.04131638	0.050418	0.0422888	0.0576887	0.0630482

Table 2. Represents average APE (Absolute Percentage Error) of last 24 months data for all the series

Each series described in Appendix B contains 150 to 200 points (transactions), each consist of 20,000 items (expert ids). The test data are the points in the last two years (24 months data). Table 2 represents average APE of last 24 points for all series. We worked on 50% and 70% data to find Best, Good and bad cluster. In almost all the series APE values are nearly identical for 50% and 70% data.

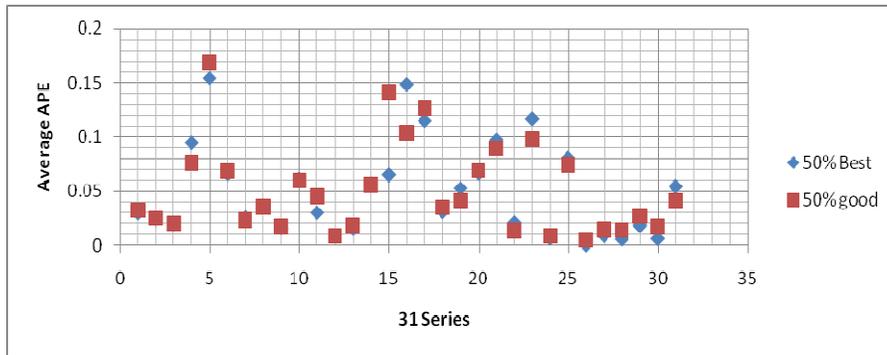


Figure 8. Sample scatter graph of 50% Best and Good Data

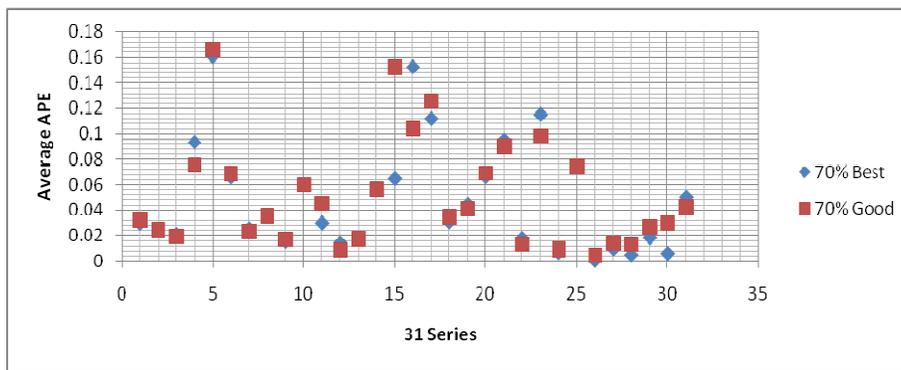


Figure 9. Sample scatter graph of 70% Best and Good Data

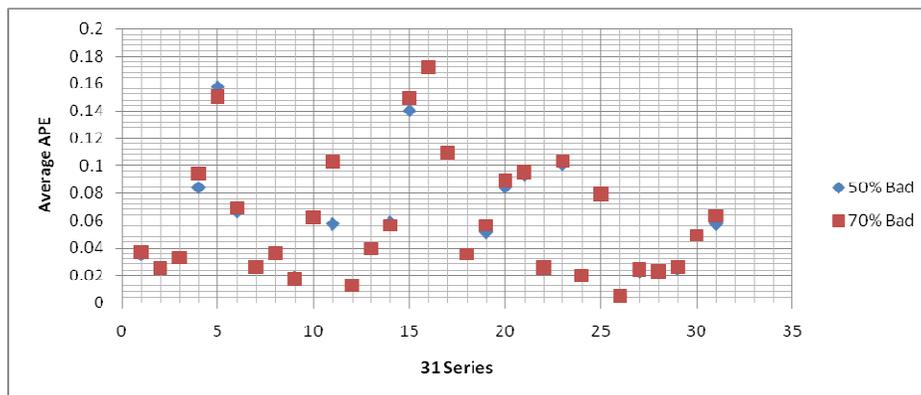


Figure 10. Sample scatter graph of 50% Bad and 70% Bad Data

## 5. SUMMARY AND CONCLUSION

Our strategy of series decomposition followed by combining results in huge set of experts. Decomposing a series into simpler components, predicting each component individually and then combining improves forecasts greatly. The challenge is to select a set of consistently good models for particular series.

First we identified set of experts and used those forecast series for last 24 points of series. We zoomed in T,S, and I Components models comprising each “good expert” at a particular point. We identified the frequency occurring of T,S and I. We clustered good sets of for each component series. The products of the average of the forecasts of models were used as final forecast. Then we clustered set of “bad experts” at a particular point.

As before, filtering of “bad” experts exhibits superior performance compared to retention of only the “good” experts. Simple forecasting methods such as “best model from previous point” or mean forecast are often employed due to their simplicity.

Using multiple decomposition methods, we are able to capture more features of the time series. Hence we got improvement in the forecasting accuracy and its performance. Table 3 represents that our approach has given a significant average improvement over Holt winter method.

Series Name	70% Best MAPE	Holt Winter
Abraham	2.99099	3.53
Beer	2.466156	2.64
Clothing	2.12176	2.1
Dry	9.34565	8.3
Equip	1.323	0.97
Fortif	6.5888	7.86
Furniture	2.56853	2.39
Gasoline	3.638489	3.3
Grocery	1.595565	1.18
Hsales	6.04853	7.87
Jewellery	3.01646	3.07
Merchandise	1.48031	0.87
Motorparts	1.73973	2.87
Newcar	5.62089	4.45
Paper	6.50062	4.62
Red	15.24638	9.79
Rose	11.19353	12.66
Shoe	3.11439	2.86
Software	4.47838	3.25
Spaper	6.6538	9.67
Spark	9.48307	11.98
Stores	1.80337	0.92
Sweet	11.52108	15.6
Total	0.70169	0.87
Wine	7.42661	7.05
sync2	1.87322	3.64
synd1	0.641374	0.62
synd3	5.041805	6.15

Table 3. Represents Comparison of ARIMA and Holt Winter on basis of Average MAPE

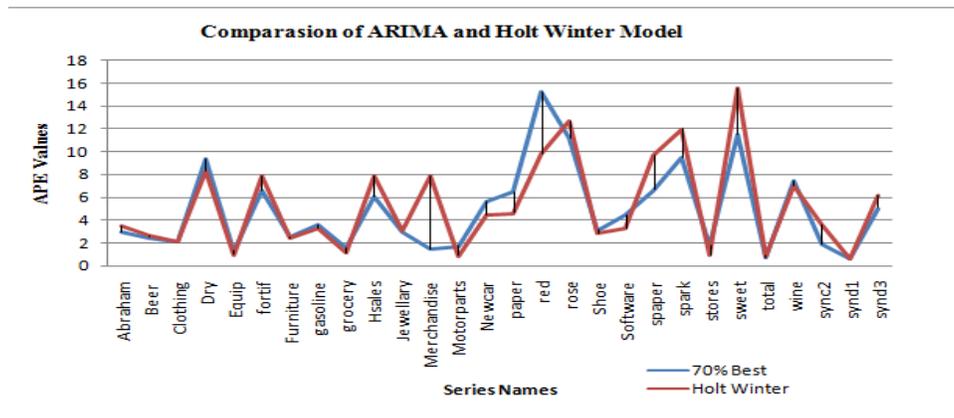


Figure 11. Line graph of comparison

### Appendix A

These below mentioned tables describe the various Trend, Seasonal and Irregular experts used in this research.

Expert ID	Expert Name	Expert ID	Expert Name	Expert ID	Expert Name	Expert ID	Expert Name
1	ARIMA (0,1,0)(0,0,1)s	11	ARIMA (0,2,1) NOINT	21	ARIMA (1,1,1) NOINT	31	ARIMA (2,1,0)
2	ARIMA (0,1,0)(1,0,0)s	12	ARIMA (1,0,1)	22	ARIMA (1,1,2)	32	ARIMA (2,1,0)(1,0,0)s
3	ARIMA (0,1,0)(1,0,0)s NOINT	13	ARIMA (1,1,0)	23	ARIMA (1,1,2)(0,0,1)s	33	ARIMA (2,1,0)(1,0,0)s NOINT
4	ARIMA (0,1,0)(1,0,1)s	14	ARIMA (1,1,0)(0,0,1)s	24	ARIMA (1,1,2)(1,0,0)s	34	ARIMA (2,1,0) NOINT
5	ARIMA (0,1,1)	15	ARIMA (1,1,0)(1,0,0)s	25	ARIMA (1,1,2) NOINT	35	ARIMA (2,1,1)
6	ARIMA (0,1,1)(1,0,0)s NOINT	16	ARIMA (1,1,0)(1,0,0)s NOINT	26	ARIMA (1,2,0)	36	ARIMA (2,1,1) NOINT
7	ARIMA (0,1,1) NOINT	17	ARIMA (1,1,0)(1,0,1)s	27	ARIMA (1,2,0) NOINT	37	ARIMA (2,1,2)
8	ARIMA (0,1,2)	18	ARIMA (1,1,0) NOINT	28	ARIMA (1,2,1)	38	ARIMA (2,1,2) NOINT
9	ARIMA (0,1,2) NOINT	19	ARIMA (1,1,1)	29	ARIMA (1,2,1) NOINT	39	ARIMA (2,2,1)
10	ARIMA (0,2,1)	20	ARIMA (1,1,1)(0,0,1)s	30	ARIMA (2,0,1)	40	ARIMA (2,2,1) NOINT

Table 4: Trend Experts

Expert ID	Expert Name	Expert ID	Expert Name	Expert ID	Expert Name	Expert ID	Expert Name
41	ARIMA (3,1,0)	53	Log ARIMA (0,1,2)	65	Log ARIMA (1,1,1) NOINT	77	Log ARIMA (2,1,0) NOINT
42	ARIMA (3,1,0)(0,0,1)s	54	Log ARIMA (0,1,2) NOINT	66	Log ARIMA (1,1,2)	78	Log ARIMA (2,1,1)
43	ARIMA (3,1,0)(1,0,0)s	55	Log ARIMA (0,2,1)	67	Log ARIMA (1,1,2)(0,0,1)s	79	Log ARIMA (2,1,1) NOINT
44	ARIMA (3,1,0) NOINT	5 6	Log ARIMA (0,2,1) NOINT	68	Log ARIMA (1,1,2)(1,0,0)s	80	Log ARIMA (2,1,2)
45	Holt	57	Log ARIMA (1,1,0)	69	Log ARIMA (1,1,2) NOINT	81	Log ARIMA (2,1,2) NOINT
46	Log ARIMA (0,1,0)(0,0,1)s	58	Log ARIMA (1,1,0)(0,0,1)s	70	Log ARIMA (1,2,0)	82	Log ARIMA (2,2,1)
47	Log ARIMA (0,1,0)(1,0,0)s	59	Log ARIMA (1,1,0)(1,0,0)s	71	Log ARIMA (1,2,0) NOINT	83	Log ARIMA (2,2,1) NOINT
48	Log ARIMA (0,1,0)(1,0,0)s NOINT	60	Log ARIMA (1,1,0)(1,0,0)s NOINT	72	Log ARIMA (1,2,1)	84	Log ARIMA (3,1,0)
49	Log ARIMA (0,1,0)(1,0,1)s	61	Log ARIMA (1,1,0)(1,0,1)s	73	Log ARIMA (1,2,1) NOINT	85	Log ARIMA (3,1,0)(0,0,1)s
50	Log ARIMA (0,1,1)	62	Log ARIMA (1,1,0) NOINT	74	Log ARIMA (2,1,0)	86	Log ARIMA (3,1,0)(1,0,0)s
51	Log ARIMA (0,1,1)(1,0,0)s NOINT	63	Log ARIMA (1,1,1)	75	Log ARIMA (2,1,0)(1,0,0)s		
52	Log ARIMA (0,1,1) NOINT	64	Log ARIMA (1,1,1)(0,0,1)s	76	Log ARIMA (2,1,0)(1,0,0)s NOINT		

Table 5: Trend Experts Contd.

Expert ID	Expert Name	Expert ID	Expert Name	Expert ID	Expert Name	Expert ID	Expert Name
1	ARIMA (0,0,1)(0,1,1)s	11	ARIMA (2,0,0)(0,1,1)s	21	Log ARIMA (0,1,2)(0,1,1)s	31	Log ARIMA (3,0,0)(0,1,1)s
2	ARIMA (0,0,2)(0,1,1)s	12	ARIMA (2,1,0)(0,1,1)s	22	Log ARIMA (1,0,0)(0,1,1)s	32	Log ARIMA (3,1,0)(0,1,1)s
3	ARIMA (0,1,1)(0,1,1)s	13	ARIMA (2,1,1)(0,1,1)s	23	Log ARIMA (1,0,1)(0,1,1)s	33	Holt Winter
4	ARIMA (0,1,1)s	14	ARIMA (2,1,2)(0,1,1)s	24	Log ARIMA (1,1,0)(0,1,1)s		
5	ARIMA (0,1,2)(0,1,1)s	15	ARIMA (3,0,0)(0,1,1)s	25	Log ARIMA (1,1,1)(0,1,1)s		
6	ARIMA (1,0,0)(0,1,1)s	16	ARIMA (3,1,0)(0,1,1)s	26	Log ARIMA (1,1,2)(0,1,1)s		
7	ARIMA (1,0,1)(0,1,1)s	17	Log ARIMA (0,0,1)(0,1,1)s	27	Log ARIMA (2,0,0)(0,1,1)s		
8	ARIMA (1,1,0)(0,1,1)s	18	Log ARIMA (0,0,2)(0,1,1)s	28	Log ARIMA (2,1,0)(0,1,1)s		
9	ARIMA (1,1,1)(0,1,1)s	19	Log ARIMA (0,1,1)(0,1,1)s	29	Log ARIMA (2,1,1)(0,1,1)s		
10	ARIMA (1,1,2)(0,1,1)s	20	Log ARIMA (0,1,1)s	30	Log ARIMA (2,1,2)(0,1,1)s		

Table 6: Seasonal Experts

Expert ID	Expert Name	Expert ID	Expert Name	Expert ID	Expert Name	Expert ID	Expert Name
1	ARIMA (0,0,1)s	11	ARIMA (2,0,0)	21	Log ARIMA (0,1,1)s NOINT	31	Log Linear Trend AR1
2	ARIMA (0,1,0)	12	ARIMA (2,0,0)(1,0,0)s	22	Log ARIMA (1,0,0)	32	Log Linear Trend AR2
3	ARIMA (0,1,1)	13	ARIMA (3,0,0)(1,0,0)s	23	Log ARIMA (1,0,0)s	33	Log Linear Trend AR3
4	ARIMA (0,1,1)(1,0,0)s NOINT	14	Linear Exponential	24	Log ARIMA (1,0,1)s	34	Random
5	ARIMA (0,1,1)s NOINT	15	Linear Trend AR1	25	Log ARIMA (1,1,0)		
6	ARIMA (1,0,0)	16	Linear Trend AR2	26	Log ARIMA (1,1,2)		
7	ARIMA (1,0,0)s	17	Linear Trend AR3	27	Log ARIMA (2,0,0)		
8	ARIMA (1,0,1)s	18	Log ARIMA (0,0,1)s	28	Log ARIMA (2,0,0)(1,0,0)s		
9	ARIMA (1,1,0)	19	Log ARIMA (0,1,0)	29	Log (3,1,1) NOINT		
10	ARIMA (1,1,2)	20	Log ARIMA (0,1,1)(1,0,0)s NOINT	30	Log Linear Exponential		

Table 7: Irregular Experts

## Appendix B

Sr. No	Series Name	Description
1	abraham	Monthly gasoline demand (Ontario gallon millions)
2	beer	US retail sales: beer, wine liquor stores(Million Dollars)
3	clothing	: US retail sales: Clothing Stores
4	dry	Monthly Australian sales of dry white wine (Thousands of liters)
5	equip	US retail inventories: Building materials, Garden equipment, supply stores (Million Dollars)
6	fortif	Monthly Australian sales of fortified wine(Thousands of liters)
7	furniture	US retail sales: Furniture Stores
8	gasoline	US retail sales: gasoline station(Million Dollars)
9	grocery	US retail sales: Grocery Stores
10	hsales	Monthly sales of new one-family houses sold in the USA.
11	jewelry	US retail sales: Jewelry Stores
12	merchandise	US retail inventories: General merchandise stores(Million Dollars)
13	motorparts	US retail inventories: motor vehicles and part dealers(Million Dollars)
14	newcar	US retail sales: New car dealers(Million Dollars)
15	paper	Monthly sale of printing and writing paper.(10 year monthly)
16	Red	Monthly Australian sales of red wine(Thousands of liters)
17	Rose	Monthly Australian sales of rose wine(Thousands of liters)
18	shoe	US retail sales: Shoe stores(Million Dollars)
19	software	US retail sales: Computer and software stores (Million Dollars)
20	spaper	CFE specialty writing papers monthly sales
21	spark	Monthly Australian sales of Sparkling wine(Thousands of liters)
22	stores	US retail inventories: Department stores (million dollars)
23	sweet	Monthly Australian sales of sweet wine(Thousands of liters)
24	total	US retail inventories: total(Million Dollars)
25	wine	Monthly Australian sales of wine(Thousands of liters)
26	synal	are synthetic series generated using formulas
27	synb1	are synthetic series generated using formulas
28	sync1	are synthetic series generated using formulas
29	sync2	are synthetic series generated using formulas
30	synd1	are synthetic series generated using formulas
31	synd3	are synthetic series generated using formulas

Table8: Describes the information of different series for sales forecasting

## REFERENCES

- [1] V.Kavitha and M.Punithavalli(2010),”Clustering Time Series Data Stream – A Literature Survey”, International Journal of Computer Science and Information Security, Vol. 8, No.1, ISSN 1947-5500.
- [2] Mahesh kumar(2010), “Using clustering to improve sales forecasting in Retail Merchandizing, Annals OR, 2010:33-46
- [3] Liu Gaojun(2009),”The Research on Combining Forecasting Model of the Automobile Sales Forecasting System”,IFCSTA,Vol. 3
- [4] Ville Hautamaki,Pekka Nykanen & Pasi Franti, (2008) “Time Series Clustering by Approximate prototypes”, IEEE.
- [5] Shaozhi Wu (2006),”An algorithm for time series data mining based on clustering”, ICCAS,Vol.3, No. 0-7803-9585-9.
- [6] Jason R .Chen(2005),”Making Subsequence Time Series Clustering Meaningful”, IEEE ICDM,Vol.5 , No.1550-4786/05.

**Authors**

Monica Sanwani

Prof.M.Vijayalakshmi

**Short Biography**

Name – Monica Sanwani

Position – Student M.E (Information Technology), VES Institute of Technology Mumbai

Profile – 2 Years teaching experience at Undergraduate level

Achievements –

1. State topper (second rank) in third year diploma
2. Secured 1<sup>st</sup> rank in M.E first year



Name - M.Vijayalakshmi

Position - Asst. Prof. HOD Dept. of Information Technology, VES Institute of Technology Mumbai

Profile - 20 yrs teaching experience both at Undergraduate and Post graduate level

Contributions

1. Served on syllabus board of Mumbai University for BE of Computer science and info tech. department.
2. Several contributions to conferences, national and international in field of Data Mining
3. Conducted several workshops on data mining related fields
4. Delivered expert lectures in data mining at several colleges in Mumbai
5. Submitted PhD thesis in data mining applications – forecasting sales at IIT, Mumbai

