**INTRODUCTION**

This Project focuses on the business application of forecasting techniques. The Project discusses forecasting retail demand for women’s dresses. Which helps the Fashion companies to reduce operating expenditure and rationalize capital expenditure in the future.

We have tried to forecast and model both the seasonality and the trend using the women. Sales and Attribute data gathered. Weka is used to build the Trend and Forecast Model. While the seasonality helps us reduce operating costs, the trend is used to forecast average sales thereby helping rationalize capital expenditure.

**DATA PREPERATION**

Data Sets

Data Set consist of two sheets namely: Dress Sales data (time series data) and Dress Attribute Data (cross section data)

|  |
| --- |
| **Dress Attribute Data** |
| 1.Dress Id |
| 2.Style |
| 3.Size |
| 4.Season |
| 5.Neckline |
| 6.Sleeve length |
| 7.Waistline |
| 8.Material |
| 9.Fabric Type |
| 10.Pattern Type |
| 11. Recommendation. |

|  |
| --- |
| **Dress Sales data** |
| 1.Dress ID |
| 2.Number of Dress sold |
| 3.Date of Dress Sold |

Pre-processing

The tasks in the process of data pre-processing include:

1. Converting the raw data into CSV format.

2. Data cleaning which in involved fill in missing values, smooth noisy data, identify or remove outliers, and resolve inconsistencies, normalizing the numerical attributes.

3. Data transformation: Converting the numeric data into nominal data.

Feature Selection

Some of the attributes may not be so relevant to the predictive power of the prediction model. Also having too many attributes and relatively too few the training instances may result in ‘under-fitting’ that means the prediction model is not generalized enough and prediction accuracy drops. Hence it is necessary to select only the significant attributes and eliminate those that are not so significant.

TRENDS ANALYSIS

After pre-processing of data, we use a classifiers for the trend analysis. To analyse the trend J48 classifier is used.

**J48 Classifier Result**

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2

Relation: Attribute DataSet-weka.filters.unsupervised.attribute.Remove-R3-5, 13-14-weka.filters.unsupervised.attribute.Remove-R4-9

Instances: 481

Attributes: 3

Dress ID

Style

Season

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree

------------------

: Casual (481.0/249.0)

Number of Leaves: 1

Size of the tree: 1

Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 226 46.9854 %

Incorrectly Classified Instances 255 53.0146 %

Kappa statistic 0.0046

Mean absolute error 0.1308

Root mean squared error 0.2594

Relative absolute error 99.2879 %

Root relative squared error 101.3834 %

Total Number of Instances 481

=== Detailed Accuracy by Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.029 0.01 0.333 0.029 0.053 0.493 Sexy

0.957 0.964 0.481 0.957 0.64 0.489 Casual

0 0.004 0 0 0 0.53 vintage

0 0.009 0 0 0 0.488 cute

0 0 0 0 0 0.446 bohemian

0 0 0 0 0 0.372 Novelty

0 0 0 0 0 0.261 Flare

0.039 0.009 0.333 0.039 0.07 0.539 party

0 0.002 0 0 0 0.433 sexy

0 0 0 0 0 0.422 work

0 0 0 0 0 0.149 OL

Weighted Avg. 0.47 0.468 0.315 0.47 0.324 0.488

=== Confusion Matrix ===

a b c d e f g h i j k <-- classified as

2 67 0 0 0 0 0 0 0 0 0 | a = Sexy

2 222 1 4 0 0 0 2 1 0 0 | b = Casual

0 25 0 0 0 0 0 0 0 0 0 | c = vintage

1 43 0 0 0 0 0 1 0 0 0 | d = cute

0 24 0 0 0 0 0 0 0 0 0 | e = bohemian

0 7 0 0 0 0 0 1 0 0 0 | f = Novelty

0 2 0 0 0 0 0 0 0 0 0 | g = Flare

1 48 0 0 0 0 0 2 0 0 0 | h = party

0 7 0 0 0 0 0 0 0 0 0 | i = sexy

0 16 1 0 0 0 0 0 0 0 0 | j = work

0 1 0 0 0 0 0 0 0 0 0 | k = OL

Trending Styles

It is clear from the Run information of J48 classifier that casual dresses are most dominant. The trends analyzed are as follows

**MOST TRENDING DRESSES**

s

