

Sales Forecasting using Time Series and Neural Networks

Carol Steffi Dcunha	171046001
Sushmitha Shetty	171046037

2nd Sem BDA

Guide: B.V. Ravindra

Contents

- Introduction
- Objective
- Requirements
- Workflow
- Data Collection
- Model - ARIMA
- Model - Neural Network
- Assumptions
- Steps
- Error Analysis
- Conclusion

Introduction

- Time series modelling and forecasting has fundamental importance in various practical domains.
- Aims to:
 - Collect and study the past observations of a time series
 - Develop an appropriate model which describes the inherent structure of the series
 - Forecast future values for the series

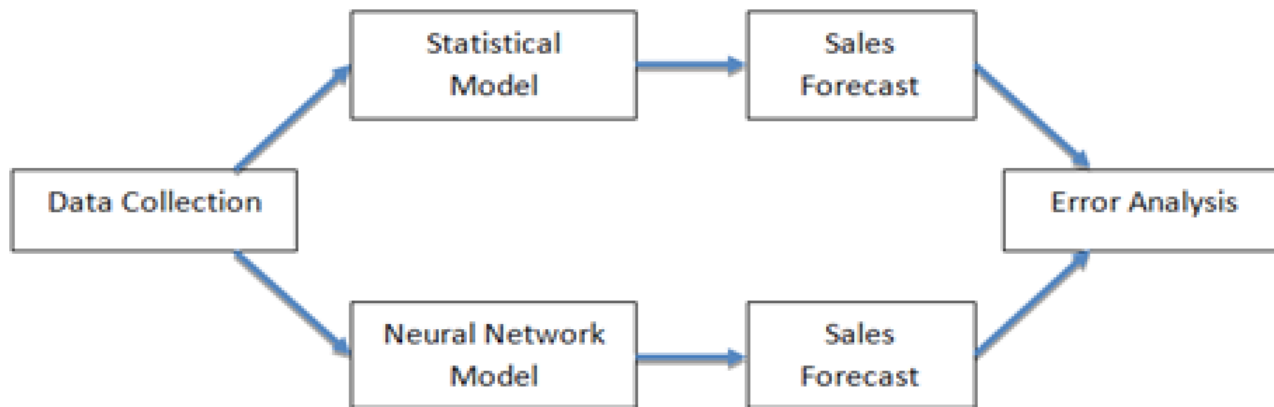
Objective

- Time series model and neural network back-propagation model in analyzing the behavior of sales in a medium size enterprise.
- To compare the accuracy of sales forecast among the two models.

Requirements

- OS: Ubuntu 16.06
- RAM: 2GB
- Hard disk capacity: 1GB
- Software: R

Workflow



Data Collection

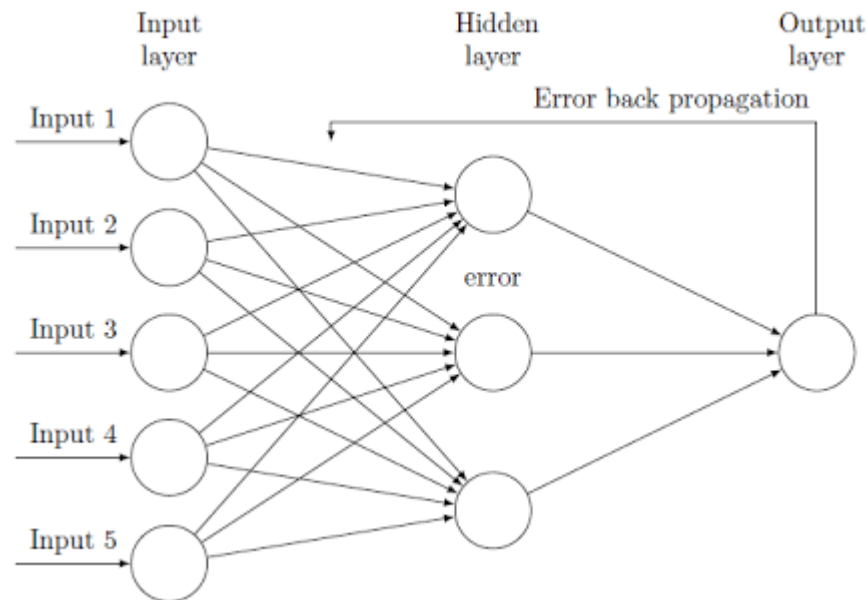
- Collected a sales dataset of Rossmann store
- <https://www.kaggle.com/c/rossmann-store-sales/data>
- Rossmann Sales Dataset
 - Store ID
 - Date
 - Sales
 - Open
 - Promo
 - State Holiday
 - School Holiday

Model - ARIMA

- ARIMA(p,d,q)
- Autoregressive (p) - finds the influence of the values of previous periods on the current period
- Integrated (d) - the original series is subtracted with its lagged series to extract trends from the data
- Moving Average (q) - finds relationships between the previous periods' error terms on the current period's error term.

Model - Back Propagation Neural Network

- Back-propagation is a supervised process where each epoch has a forward activation flow of outputs and backward error propagation of weight adjustments



Assumptions

- StateHoliday isn't included in the model:
 - Only 0.11% of stores are open on public holidays
 - Model requires regressor matrix to have full rank

Steps

- Select stores that are open
- Group the data based on store IDs
- For each store:
 - Split the data into train:test data (80:20)
 - Fit arima/neuralnet model with train dataset, $\text{sales} \sim (\text{customer} + \text{promo} + \text{schoolholiday})$
 - Forecast sales of test dataset using the fitted model
 - Calculate NRMSE

Error Analysis

- $$\text{RMSE} = \sqrt{\sum \frac{(y_{pred} - y_{ref})^2}{N}}$$
- $\text{NRMSE} = \text{RMSE}/\text{S.D.}$
- Average NRMSE ARIMA = 0.327348
- Average NRMSE NN = 0.266387

Conclusion

- The model obtained by the neural network was superior to ARIMA model in forecasting for the data analyzed.
- Only three regressors - Customers, Promo and SchoolHoliday are considered in the models. As a part of future work, additional parameters such as store size, product quantity and weather details can be considered to develop a more robust and accurate forecasting model.

Thank you