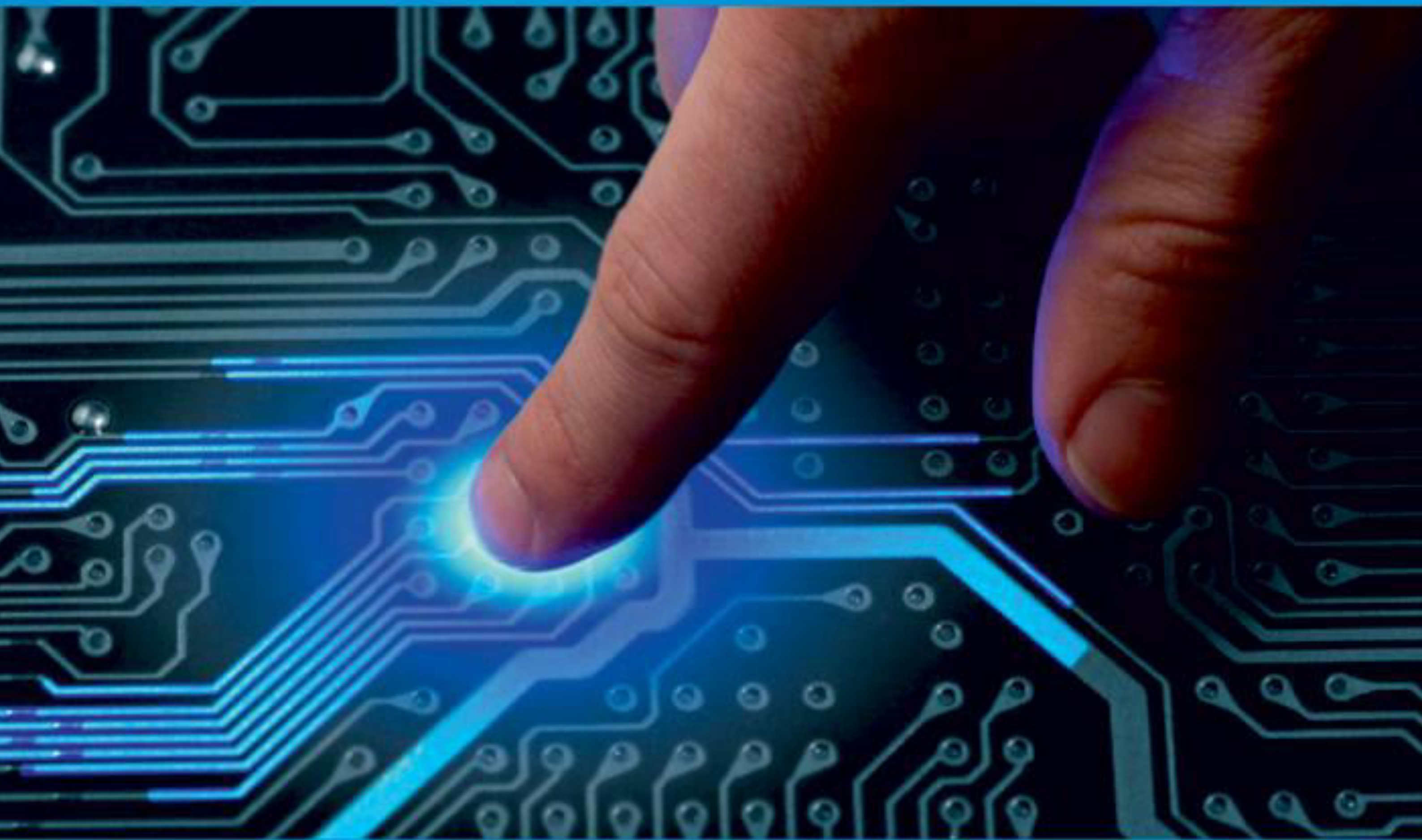
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005272 |

Demand Forecasting in Retail using Machine Learning on Azure Cloud

Madhuri S Chandane, Prof. Radha Shirbhate

Department of Computer Engineering JSPM'S Bhivarabai Sawant Institute of Technology and Research Pune, India

ABSTRACT: Demand forecasting is one of the main decision-making tasks of retail industry. For demand forecasting first raw sales data is collected from the market, then according to data, the future sale/product demands are forecasted. This prediction is based on collected data that compiles through different sources. The machine learning engine executes data from different modules and determines the weekly, monthly, and quarterly demands of goods/commodities. In demand forecasting, its perfect accuracy is non-compromising, the more accurate system model is more efficient.
Using machine learning, demand forecasting for product for next few weeks is possible. Implementing demand forecasting on azure cloud gives us scalability & availability of cloud features using pay as you go model.

KEYWORDS: Machine learning, Random Forest model, Feature engineering, Cloud computing, Virtual machine

I.INTRODUCTION

Almost every business needs to predict the future to make better decisions and allocate resources more effectively. As an example, accurately forecasting spikes in demand for products and services can give a company a competitive advantage. The better the forecasting, the more they can scale as demand increases, and the less they risk holding onto unneeded inventory. Use cases include predicting demand for a product in a retail/online store.

Traditionally, forecasting requires a deep level of knowledge and understanding of different models and methods. Most of the time, forecasting is done based on historical data which might not be easily available. Even when the data is available the process of predicting customer demand requires testing different methods, variables and distinguishing good from bad forecasts which can be time-consuming.

When companies face all these challenges, they often decide to take a more simplified approach which can lead to inaccurate predictions and misinformed decisions. For this reason, everyone wants to forecast demand using Azure Machine Learning cloud based on historical data.

We picked Azure Machine Learning due to the flexibility it provides since it is a fully managed cloud service that can easily scale up or down the resources to train Machine Learning (ML) models as needed. Using Azure cloud, we can use virtual machine Lift & Shift approach with vertical scaling or using Horizontal scaling depending upon the workload & as per your budget.

Using horizontal scaling, we can use Virtual Machine scale set using load balancer, so that multiple servers can execute parallel batches at the same time so that we can reduce execution time of forecasting service which includes:

1) Model Building service: model is trained using data

2) scoring service: forecasting for current scenario.

3) Prediction service: make the prediction on discount for price of product.

II. REVIEW OF LITERATURE

Author proposes [1] a new approach to predict demand forecasting of product for retail business using machine learning .

Author proposes [2] a best idea to execute machine learning project by migrating it on azure cloud using Virtual Machine scale set & Lift & Shift approach so that reduces execution time for demand forecasting of products.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005272 |

Author proposes [3] demand forecasting model to run maximum parallel batches so that prediction can be achieved more accurately.

Author proposes [4] a best approach to reduce upfront cost of project by using Azure Cloud as it provides pay-as-you go model means as demand increases, we can increase number of resources & as demand decreases, we can reduce number of resources.

Author proposes [5] proposed capabilities enable ML-based software to produce more accurate and reliable forecasts in complex scenarios.

Author proposes [6] demand forecasting for new products using machine learning approach is complex as it needs historical data for comparison because more data, more accuracy..

III. PROPOSED METHODOLOGY

A supermarket chain needs to forecast sales of products over the upcoming quarter. The forecast allows the company to manage its supply chain better and ensure it can meet demand for products at each of its stores. The company updates its forecasts every week as new sales data from the previous week becomes available and the product marketing strategy for next quarter is set. Quantile forecasts are generated to estimate the uncertainty of the individual sales forecasts.

Solving Approach:

• The forecasts from scoring are saved in a csv initially and another R service is used to store it into the DB post some transformations.

• The Prediction service is called by a recommendation service. Recommendation service is written in Java. The Prediction services provides the forecasts to the recommendation service and the recommendation service writes it into the oracle DB.

• Model Building and Scoring are triggered using batch files in windows.

• Prediction is run through recommendation which is in turn triggered by windows batch files.

Model building/Scoring has the following parts.

1. Get Data: Getting Data from different table in Oracle into R.
2. Merge Data: Merge data from different sources into a single data frame.
3. Data Aggregation: Aggregate the data into the desired level. In our case grouped-item level.
4. Feature Creation: Create the features which will then be passed to the Model.
5. Model Training

Migrating Machine learning code on Azure cloud using Virtual Machine Lift & shift approach is more flexible & scalable

Horizontal scaling, also called scaling out and in, means adding or removing instances of a resource. The application continues running without interruption as new resources are provisioned. When the provisioning process is complete, the solution is deployed on these additional resources. If demand drops, the additional resources can be shut down cleanly and deallocated.

VM Scale sets are used to deploy identical sets of VM Instances. i.e., when you enable scaling in azure, VM Instances deployed will have the same Configuration as memory and storage. Scale sets will also increase (Scale-Out) VM’s or decrease (Scale In) VM’s as and when the demand increases or decreases. A scale set can be automatic, manual, or both.

Scale-Out: It’s a term used for increasing VM’s while scaling in azure.

Scale In: It’s a term used for decreasing VM’s while scaling in azure.

You can define rules to automatically adjust the capacity.

- You can Scale out (increase) or Scale In (decrease) the number of VM’s in the set.
- You can schedule events to increase or decrease VM’s at a specific time period.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005272 |

- It reduces the monitoring time and optimizes the performance of the application.
- Minimum no. of 2 instances of VM are required for auto scaling.

IV. MATHEMATICAL MODEL

In Demand Forecasting for product, it is directly proportional to Impurity function i(t) which depends on the task being solved. For regression problems a common goal is to minimize the mean squared error or the residual sum of squares given that

$$i(t) = \frac{1}{N_t} \sum_{x,y \in \mathcal{L}_t} (y - \widehat{y_t})^2$$

where Nt: is the number of node samples at node t,
Lt is the subset of learning samples falling into the node t,
y^t is the prediction for that node and y is the actual value of the
output variable Y
After each tree is constructed and trained, the random forest algorithm uses each decision tree to calculate a prediction for the output variable y. These predictions are then combined into a final prediction value by taking an average of them.

V. SIMULATION RESULTS

Model building/Scoring has the following parts.
1. Get Data: Getting Data from different table in Oracle into R.
2. Merge Data: Merge data from different sources into a single data frame.
3. Data Aggregation: Aggregate the data into the desired level. In our case grouped-item level.
4. Feature Creation: Create the features which will then be passed to the Model.
5. Model Training

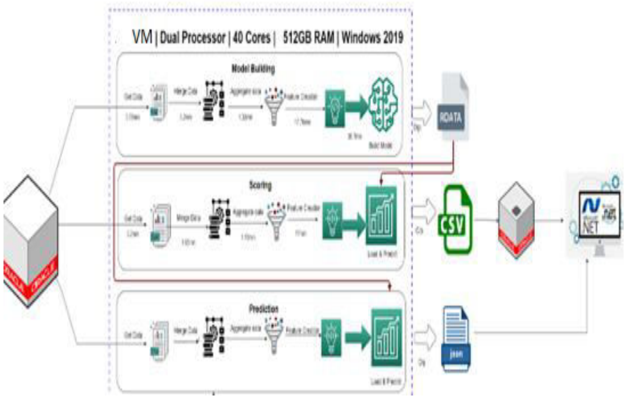


Fig.1. Solution Architecture

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005272 |

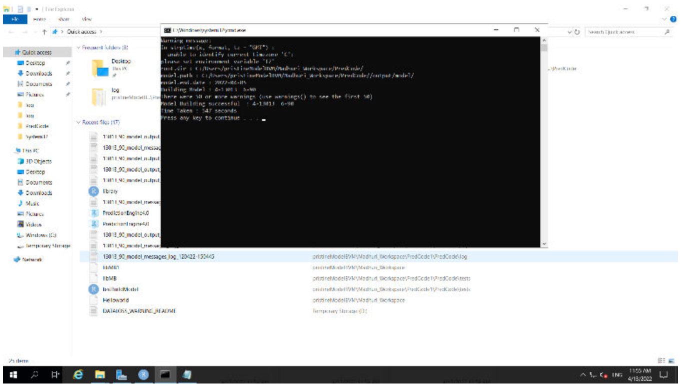


Fig.2 Batch Running output screen

```
@ECHO OFF
call initPaths.cmd
cd %root_path%\src\api
Rscript.exe .\buildModel_batch.R %root_path% %model_path% %model_end_date% 4 1 6 5
Rscript.exe .\buildModel_batch.R %root_path% %model_path% %model_end_date% 4 2 6 5
Rscript.exe .\buildModel_batch.R %root_path% %model_path% %model_end_date% 4 3 6 5
Rscript.exe .\buildModel_batch.R %root_path% %model_path% %model_end_date% 4 4 6 5
Rscript.exe .\buildModel_batch.R %root_path% %model_path% %model_end_date% 4 5 6 5
Rscript.exe .\buildModel_batch.R %root_path% %model_path% %model_end_date% 4 6 6 5
Rscript.exe .\buildModel_batch.R %root_path% %model_path% %model_end_date% 4 7 6 5
Rscript.exe .\buildModel_batch.R %root_path% %model_path% %model_end_date% 4 8 6 5
```

Fig.3 Batch Files

VI. Conclusion and Future Work

Demand forecasting model, which is used for retail product forecasting, by migrating code on azure cloud can reduce time to execute parallel batches. Here while migration to azure cloud, created virtual Machine Lift & Shit approach using vertical scaling and also using Horizontal scaling with Virtual Machine Scale Set (VMSS). Here using cloud migration, we can conclude that we can run a greater number of batches parallel so that we can give prediction about sales & product in time for customers.

Demand forecasting using machine learning can be well calculated using Azure native apps like Azure Databricks using cluster computing. Azure Databricks provides the latest versions of Apache Spark and allows seamlessly integrate with open-source libraries. Spin up clusters and build quickly in a fully managed Apache Spark environment with the global scale and availability of Azure. Clusters are set up, configured, and fine-tuned to ensure reliability and performance without the need for monitoring.

References

1. Planet Together. (2019). The Advantages of Effective Demand Forecasting
2. L. Luce, ‘‘Deep learning and demand forecasting,’’ in Artificial Intelligence for Fashion. Berkeley, CA, USA: Après, 2019, pp. 155–166.
3. H. Steller and H. Symington, ‘‘Evaluating qualitative forecasts: The FOMC minutes, 2006–2010,’’ Int. J. Forecasting, vol. 32, no. 2,pp. 559– 570, Apr. 2016.
4. V. M. Eguíluz, J. Fernández-Gracia, X. Irigoien, and C. M. Duarte,‘‘A quantitative assessment of arctic shipping in 2010–2014,’’ Sci. Rep.,vol. 6, no. 1, pp. 1–6, Aug. 2016.
5. C. Catal, K. Ece, B. Arslan, and A. Akbulut, ‘‘Benchmarking of regressionalgorithms and time series analysis techniques for sales forecasting,’’Balkan J. Electr. Comput. Eng., vol. 7, no. 1, pp. 20–26, 2019.
6. Gossain, S., Malhotra, A., & El Sawy, O. A. (2005). Coordinating for ∟exibility in e-business supply chains. Journal of Management Information Systems, 21(3), 7-45.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005272 |

7. Genov, R., & Cauwenberghs, G. (2001). Charge-mode parallel architecture for matrix-vec-tor multiplication. IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing, 48(10), 930-936.
8. Elman, J. L. (1990). Finding structure in time. Cognitive Science, 14(2), 179-211.
9. Makridakis, S., Chat∟eld, C., Hibon, M., Lawrence, M. et al. (1993). The M2-com-petition: A real-time judgmentally based forecasting study. International Journal of Forecasting, 9(1),
10. Raghunathan, S. (1999). Interorganizational collaborative forecasting and replenish-ment systems and supply chain implica-tions. Decision Sciences, 30(4), 1053.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING