Journal of Technological Information, Management & Engineering Sciences (2020),1(2), 122-137

To create an AI-Driven Demand Forecasting Model for Perishable Goods in a Fragmented Retail Market. To develop an AI-Driven Demand Forecasting Model for Perishable Goods in a Fragmented Retail Market

[1] **Dr. Rehan Mustafa**
[2] **Eman Khalil**

How to cite this article:

Mustafa, R., & Khalil, E. (2020). Developing an AI-Driven Demand Forecasting Model for Perishable Goods in a Fragmented Retail Market. Journal of Technological Information, Management & Engineering Sciences, 1(2), 122-137.

Received	Accepted	Published
19 October 2020	23 November 2020	28 December 2020

Abstract

The operational problem of perishable goods is one of the most difficult demand forecasting problems because of the limited shelf life, the quick variation in demand, and the cost of having unsold stock go bad due to quality deterioration or the loss of sales and customer dissatisfaction due to stock outs. The situation is particularly difficult in fragmented retail markets, where small retailers, independent grocers, local distributors and informal supply chains commonly struggle to keep accurate records, lack effective digital systems, and have access to varying amounts of market information. In this paper, an artificial intelligence (AI) based demand forecasting framework is proposed which combines time-series learning, machine-learning features, hierarchical information sharing, and inventory-aware decision rules. The system is proposed to combine historical sales, price and promotion data, calendar effects, weather, local events, replenishment lead times, stock availability and the product shelf-life characteristics. It evaluates interpretable statistical baseline methods and machine learning techniques and a hybrid forecasting model, explicitly addressing censored demand due to stockout, data sparseness, new products, and store-level heterogeneity. The quality of forecasts is assessed by scaled error measures, probabilistic calibration and operational measures like waste, service level and lost sales. As no original retail data is provided, the paper does not include fabricated empirical results, but rather a research design and hypotheses. The framework states that the accuracy of the forecast is not an end in itself but should be considered within the framework of ordering decisions and perishability. A practical implementation ought to be simple to use, interpretable, respect privacy, and provide actionable forecasts even in the face of sparse data at the store level.

Keywords: *artificial intelligence, demand forecasting, perishable goods, inventory optimization, machine learning, probabilistic forecasting, supply chain, fragmented retail and food waste.*

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan

[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Technological Information, Management & Engineering Sciences (2020),1(2), 122-137

1. Introduction

Demand forecasting enables the retailer to determine the quantity to order, the timing of orders, and the distribution of the product and inventory across various retail locations. These choices are made in the context of time, for foods such as fresh vegetables, fruits, dairy products, meat, bakery items, ready to eat foods, and some pharmaceuticals. Having too much inventory can go bad or go stale, having too little inventory can lead to stockouts and lost sales. The amount of the forecast error will thus not only affect the economic cost of the forecast error, but also the direction of the error, the product margin, the product shelf life, the replenishment frequency, as well as the customer reaction.

Fragmented retail markets renders prediction more complex. There are numerous independent retailers with imperfections in their sales history, different product descriptions, irregular sales recording time periods and orders from a variety of wholesalers. Many retailers might not use integrated point of sale (POS) software systems, but instead handwritten records or messaging apps. Weather, paydays, holidays, school calendars, religious festivals, neighborhood events and changes in wholesale prices are other factors that can impact local demand. A model that is trained solely on the central data of a large chain might not generalize well to other settings in the chain.

However, traditional forecasting techniques such as moving averages, exponential smoothing and seasonal time-series models are still useful as they are easy to understand and work well when the pattern is stable. The machine-learning techniques can include nonlinear relationships and external variables, and deep-learning architectures can learn the temporal patterns of numerous related series. However, more complex models may not necessarily yield better results. Sparse data, inconsistent product identifiers, stockouts, price changes and structural breaks can impact performance. Such a practical AI system needs, however, to have data-quality controls, appropriate baselines, uncertainty estimates, and inventory-aware decisions.

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan

[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)

This work is licensed under a Creative Commons Attribution 4.0 International License.

This paper suggests an AI-enabled framework that is applicable in disintegrated retail market. It integrates store-item forecasts, shares information among related stores and products, includes external drivers when applicable, and transforms forecasts into recommendations for replenishment, taking into consideration shelf life and risk of stockout. It focuses on cost-effectiveness, interpretability and operating assessment. The contribution is not an actual reduction of waste or forecasting error that has been achieved for the proposed model, but rather the architecture and methodology that will enable this reduction.

1.1 Research Objectives

The goals are to: (1) design a forecasting architecture that is appropriate for sparse, heterogeneous retail data; (2) integrate time-series and machine-learning techniques with commonality and similarities in stores and products; (3) create a validation design that avoids temporal leakage and that considers stockouts; (4) assess forecast accuracy in addition to waste, availability, and inventory cost; and (5) recommend practical governance and deployment strategies for small retailers.

1.2 Research Questions

The study aims to investigate if a hybrid AI model could be adopted in fragmented retail environments to enhance forecast accuracy for simple baselines, the role of information sharing for stores with limited sales history, how stockout and product expiry should be incorporated, and how best to measure the operational value of a demand forecast for perishable products.

2. Literature Review

Retail demand planning is based on classical time-series forecasting. Moving averages eliminate the short-term volatility, exponential smoothing assigns more weight to the latest observations, and seasonal models are used to capture the weekly or annual patterns in demand. The Box–Jenkins model of autoregressive integrated moving average models provides a systematic approach to the modelling of temporal dependence (Box et al., 2015). These are also good reference techniques when data is scarce and are relatively easy to interpret.

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

They might need extra feature engineering to consider promotions, weather, stock availability, complex interactions, etc.

Retail demand can be sporadic, particularly for slow moving goods or small retailers. Croston's method is applicable for intermittent series because it creates an estimate to the nonzero size of the demand and the interval between the demand events are separated (Croston, 1972). Perishables are different because requirements can be regular, but both fluctuating and unpredictable; and the impact of errors varies with the amount of shelf life remaining. Forecasting methods should thus be chosen according to the specific characteristics of the product, and not by an algorithm applied to all products.

Lagged sales, calendar features, prices, promotions, weather, and store attributes can be joined by machine-learning models like random forests and gradient-boosted trees. They can describe nonlinear relationships without having to fully specify a time-series model. They have problems with the sensitivity to feature leakage, extrapolating trends, and careful validation. The neural forecasting framework can capture shared representations from a multitude of time series, and could prove beneficial when several stores and product categories are available. For instance, DeepAR is a model that generates probabilistic forecasts for multiple related time series, and can generate prediction distributions instead of just point predictions (Salinas et al., 2020).

Global models trained on large number of related series are capable of benefiting data-poor places by bringing the statistical strength. Hierarchical forecasting methods aggregate and/or reconcile the forecasts on store, neighborhood, category and SKU level; consistency between the store and aggregate plans. In a fragmented market, however, stores can be different in terms of their customer base, their product mixes, hours of operation and buying habits. Pooling ought to be selective; information sharing could take place between similar stores and products, while the local patterns should be represented by store-specific features or adjustments.

One of the difficulties is that what is recorded as a sale may not reflect actual demand. Note that if stock drops to zero, new purchases might be missed or replaced, which will lead to understated sales data. Observations of stockout

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Technological Information, Management & Engineering Sciences (2020),1(2), 122-137

periods may not be true low-demand periods. A model that is not provided with the information on product stock availability when it learns directly from these records may predict lower demand just for those products that are consistently selling out. Studies on retail forecasting and stock management thus focus on the link between the demand forecast, the replenishment policy and availability.

A range of error metrics are needed for forecast evaluation. It is easy to interpret the mean absolute error, and the larger the error, the more it is penalized in the RMSE. When the actual demand is equal to zero or very close to zero, the mean absolute percentage error might not be stable. Scaled measures, like mean absolute scaled error, can be used to compare the performance of products with different units and scales (Hyndman & Koehler, 2006). Loss and interval scores are possible assessment measures for probabilistic forecasts, and service level measures, along with inventory related measures such as waste or stockout frequency as well as total cost, are all potential assessment measures for inventory policies.

Develop a problem formulation and establish a market context.

3.1 Demand, Sales, and Availability

Let y(s,i,t) be the observed sales in store s, item i and time period t. When the number of items available for sale is less than the number of items that customers want, the underlying demand for the item d(s,i,t) can be greater than the number of items sold. A simple relationship is y = min(d, a), where a is the amount available to be sold over that time period, and where no substitution, no returns, and no intraperiod replenishment occurs. Availability may vary during the day in practice; if possible, a flag for daily availability should be recorded in the data pipeline, and stockout times should be recorded. When stockouts are common, predicting the sales series may underestimate the demand.

3.3 Energy and Maintenance Services

The inventory decision should consider the shelf life left. Exceeding orders can cause spoilage, marking down or disposing of items; underselling orders can result in lost sales and loss of customer confidence. The cost of overforecasting is not necessarily equal to the cost of underforecasting. If the product has a short life, overforecasting can create waste, and for a high demand product, an

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan

[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)

This work is licensed under a Creative Commons Attribution 4.0 International License.

underforecast can be very expensive when replenishment lead time is long. These asymmetric costs must be considered, along with the service level objective, and the desired quantile.

3.3 Fragmentation and Data Heterogeneity

Retail break-up can lead to non-standardized product identification and unsynchronized transaction systems. Local names, varying package sizes or units of measure are used to describe similar products. Some stores may report daily totals, some stores may provide transaction level information. The proposed framework includes a canonical product dictionary, normalization of units, data-quality flags, and flexible aggregation to produce comparable features. The missingness should be considered informative instead of the missing value being set to zero when an explicit reason for missingness is known.

4. Proposed AI-Driven Forecasting Framework

The suggested structure comprises seven modules: data ingestion, data cleaning and entity resolution, feature engineering, baseline forecasting, AI forecasting and information-sharing, probabilistic calibration, and inventory decision support. The modules are deployable in the central location or via a light-weight cloud-service, and have simple deployment interfaces for the stores, that are not able to support advanced local infrastructure.

4.1 Data Sources

Historical sales data, product identifiers, store identifiers, prices, promotions, inventory, replenishment quantities, suppliers' lead times and expiry/shelf life data are core data. Additional external data may be available, such as the wholesale market price, public holidays, and local events, as well as rainfall and temperature. The features should be dated as to when they were added. For instance, weather forecasts are represented by the forecast available at the time of ordering, not the weather that actually happens after ordering. This removes the possibility of having information from future evaluation seeping into historical evaluation.

4.3 Data Wrangling and Data Quality Assessment

The pipeline will check for negative amounts, duplicate transactions, unit inconsistencies, missing dates, extreme price changes and product substitution.

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

It normalizes the units (kg, liters, pieces, packages counts), and maintains conversion rules. Features are lagged sales, rolling means, rolling variability, day of week, day of month, proximity to holidays, price, promotion status, last delivery time, store size, product category and remaining shelf life. Only information available at each forecast origin should be used when computing lags and rolling features.

When the absence of a value is informative, indicators should be used with the missing value. The investigation of outliers should be considered before removing them as they might be a genuine festival, bulk usage or promotion. Product mapping can be achieved through a mixture of supplier codes, text normalisation, packaging attributes, and human review. It might be more realistic for small retailers to have a semi-automated catalog matching workflow rather than having to have data that's perfect before deployment.

4.3 Baseline Models

You first obtain forecasts using a seasonal naive model, a moving-average model and an exponential-smoothing model. For a seasonal naive forecast, the most recent observation in the same time of year (e.g. the same weekday of last week) is used. These baselines are crucial, because they're cheap, and give a reference that a complex model has to outperform. Croston-type methods could be added for intermittent demand. Model choice may differ if there are different product types and different density of data.

4.4 Machine-Learning Forecasting Model

The most straightforward machine-learning contender is a gradient-boosted decision tree model using store-item-time examples. Historical lags, rolling statistics, price, promotions, product information, calendar information, weather forecasts, inventory availability, supplier lead time, store attributes and product attributes. Depending on ordering practice, the target might be demand for the next day, several days or the lead-time horizon. By learning on multiple stores, the model can capture common effects, and local differences can be captured in the form of categorical or encoded store and product information.

A generic model: $\hat{y}(s,i,t+h) = f\theta(X(s,i,t), E(t), C(s,i))$, where X contains past and operational features, E contains external features, and C contains store and item

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

context. A multi-horizon model can generate forecasts for multiple horizons. The model should be trained using the loss which is relevant for the purpose of operation, such as absolute error for robust forecasting or quantile loss for inventory decisions.

4.5 Cross-Store Information Sharing

When there is a short history of each individual store information sharing is crucial. It trains a global model on several stores and products, and adjusts predictions using store-specific features. Geography, store format, store traffic, the product mix, and historical demand patterns are all areas that can be used to create similarity groups. Neighborhood or category-level signals can be obtained through hierarchical aggregation, for new stores and sparse items. Avoid indiscriminate pooling as a high volume urban store may not provide an appropriate model for a small rural store.

4.6 Probabilistic Forecasting

A "point forecast" is not a statement of uncertainty. In the proposed model, therefore, several quantiles, or a predictive distribution, are estimated. Models can be trained for levels like the upper demand quantile and the median, and the calibration evaluated on the held-out period of the validation set. If the stockout cost is high, then an upper quantile can be useful, and on the other hand if the expiry cost is high, then a lower quantile can be considered. The selection of quantiles should be based on the economics and operational factors of the product being processed.

4.7 Inventory Decision Layer

Forecasts are used to drive an inventory policy that takes into account current inventory, on-order inventory, anticipated delivery dates, lead time, shelf life and target service level. Simplified order quantity is $Q = \max(0, \hat{D}L + SS - IP)$, where $\hat{D}L$ is the forecast demand during the lead time, SS is the safety stock as a result of uncertainty and service targets, and IP is the inventory position, which includes on-hand inventory and confirmed orders. For perishables, the policy should also account for the age-bucket inventory and expiry dates, not all as equally fresh.

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

A practical system might suggest a number, and display a brief explanation, like recent sales pattern, holiday coming up, weather forecast, or delay in delivery. Recommendations should be overruled by users, and the reason documented. Although overrides are useful for feedback, they are not always accurate labels and should be viewed as such; they could be due to local knowledge, supplier limitations, or human error. The system should have to consider the reduction of total cost and waste in the system while maintaining an acceptable service level when making recommendations.

5. Research Methodology

The paper discusses a quantitative research design used to evaluate the proposed model. No original retailer data was provided so there are no claimed measured forecast improvements. Implementing in the future should be based on a documented set of data from the retailers involved or a suitable, relevant and licensed public benchmark. The research should be pre-registered or otherwise define the hypotheses, data exclusions, validation scheme, and primary metrics used in the research prior to the final test of the hypotheses.

5.1 Dataset and Sampling

The data set should contain several stores, product categories, sales time, product price, product promotion, product stock, product replenishment, and shelf-life attributes if available. The sampling period should include sufficient cycles of seasons to determine repeating effects, as well as normal and abnormal demand periods. Stores should be sorted by size, location, format and data quality to see if the benefit extends to the larger, or more digitally mature retailers.

5.2 Temporal Validation

If you randomly split your data by day, then it may contain future patterns in the training set. Rather, the study employs a rolling-origin evaluation, namely, train on a first historical window, forecast the next window, then shift the origin of the forecast to the right and repeat. All features should be made using the information contained at the origin. The hyperparameters are tuned on validation windows, and the last test window is left unchanged until the model

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

selection is done. If a model is run at multiple horizons, then each horizon should be reported separately.

5.3 Stockouts and Missing Data

The study should indicate when stocks-out occurred and separate zero sales at the time of the stock-out from the zero sales during the unavailable period. In the case of stock-out data, when the data is known, demand in periods of stock-out can be reported and the uncertainty of the data point can be included. The model(s) to be used should be compared to the ones trained without/with the known stockout periods (sensitivity analysis). If there are gaps in inventory records, the constraint should be expressed as a percentage and not as actual sales.

5.4 Evaluation Metrics

The metrics used for forecasting are MAE, RMSE, MASE and quantile loss. The MAPE is not meaningful or reliable for zero and small actual values so it should not be used as the sole measure. Forecast bias also needs to be calculated to identify if the model overestimates or underestimates demand. A probabilistic forecast should be assessed in terms of the quantile coverage and the width of the interval. Some operational measures are waste rate, stockout rate, fill rate, service level, gross margin, markdowns, and total inventory cost.

A simple waste rate is the number of expired or discarded units divided by the number of received or available units over a period of time, with the denominator being the same each period. Fill rate is the percentage of inventory units shipped from available units. The measures should be interpreted by category as the shelf life, packaging and replenishment cadence vary. A forecast model can have a positive impact on MAE without also having a positive impact on waste, and, vice versa, better ordering logic can have a positive impact even if point forecast gains are small.

Baseline comparisons and ablation (5.5 baseline comparisons and ablation)

The proposed model should be compared to a seasonal naive model, an exponential smoothing model and Croston-type model (where applicable) and a machine-learning model with no external variables. Ablation experiments are conducted by sequentially turning off the weather, promotions, store similarity,

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

stockout adjustments and probabilistic outputs. These comparisons indicate if added complexity to the signal will be helpful. The study should also report on the runtime and data requirements, as well as accuracy, since small retailers might not be able to keep a costly data pipeline.

6. Proposed Algorithmic Workflow

Step 1: Establish the forecast horizon, the intended measure of demand, the product range and inventory goals.

Step 2: Gather sales, inventory, price, promotion, supplier, shelf-life and (optionally) external data, all on a time-stamped basis.

Step 3: Normalize product identifiers and product units; validate data quality - missing, duplicate and unusual data.

Step 4: Identify observed sales vs. possible demand by identifying stockout and unavailable periods.

Step 5: Create lagged, rolling, calendar, price, promotion, weather, store and product features free of temporal leakage.

Step 6: Establish seasonal naive and statistical baseline for each appropriate product group.

Step 7: Train a global machine-learning model across stores and products and with local context features.

Step 8: Compute quantile forecasts and/or predictive intervals, and validate these on the validation windows.

Step 9: Transform forecasts into replenishment quantities based on inventory position, lead time, shelf life and service levels.

Step 10: Examine forecasts and inventory results on rolling-origin test windows.

Step 11: Conduct subgroup and ablation tests and quantify the uncertainty of performance differences.

Step 12: Deploy slowly, gather user overrides, track drift, and retrain with documented controls.

10. Analysis Plan and Expected Results

The main hypothesis is that a store and product context global machine-learning model will achieve better performance than simple baselines for many high-volume, promotion-sensitive products. A second hypothesis is that cross-store

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

information sharing will be particularly beneficial for stores and products with limited history. The third hypothesis is that probabilistic forecasts and shelf-life-aware ordering will lower the total cost of waste and stockouts than will fixed-quantity replenishment. These hypotheses must be tested and cannot be said to be findings until the experiment is performed.

Results should be provided by product category, store group, forecast time period and data availability. A model can work exceptionally well in dairy, or be very accurate in stores with many items but less useful in rural areas with few items. Confidence intervals or repeated rolling-origin results should be included in the report, as should be the proportion of the store-item series on which the AI model outperforms the baseline. These differences can be masked by using average aggregates.

Model: Good strength and major restriction (L1):Model: Good strength and major restriction (L1):

Seasonal naive / moving average: Simple, low cost and audit friendly; limited response to external drivers.

Exponential smoothing / seasonal models Stable structure and careful specification required Suitable for trend and recurring seasonality

Gradient-boosted trees Uses nonlinear drivers and interactions Needs feature quality and temporal validation

Local AI model with global context Limited information sharing among sparse stores and items Can overfit strongly local patterns

Uncertainty and ordering/service goals Requires reliable inventory, lead time, shelf-life data Probabilistic AI + inventory policy

7.1 Business impact and cost analysis

The cost of purchasing, carrying, expiration of stock, markdowns, stockouts and emergency ordering should be estimated in the business analysis. The costs should be assumed and then sensitivity analysed. For instance, a retail outlet with a high frequency of deliveries may have a lower safety stock than a retail outlet that has a long supplier lead time. A model should not be seen as successful just because it improves average error if it results in high waste or significant stockout rates of important product categories.

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

7.2 Robustness and Generalization

The robustness tests ought to consider the holidays, extreme weather, supplier delays, price shocks, and periods of unusual demand. If possible, models should be tested on stores that were not used in training to assess transfer to new stores. New products should be evaluated on their own, with the aid of metadata and category-level priors. Drift monitoring should measure sales distribution, prices, promotion usage, forecast error and inventory results.

8. Discussion

The proposed approach merges flexibility of predictions and operational constraints. Machine learning can be trained to identify relationships between demand and external drivers and global training can provide information across siloed stores. Probabilistic forecasts include uncertainty and can be used to provide a range of services. But value of the system is based on data quality and integration of processes. If inventory records are not accurate, lead times are unknown, and employees do not have access to inventory recommendations, then the sophistication of the forecast is not useful.

So, a staged deployment is recommended. The first phase should focus on the standardisation of product and sales information and define baseline forecasts. The second phase can incorporate global machine-learning forecasts for categories for which there is enough data. The third phase is for probabilistic forecasts and shelf-life informed inventory decisions. This gradual progression gives retailers time to prove their worth before adding to their technical complexity. It also helps to minimize the misconceptions that model errors are confused with operational errors or the other way round.

In markets that are broken up into small pieces, humans still matter. A store manager might be aware of a local event, such as a festival, road closure, shortage of supplies or community event which is not included in the data. A helpful interface will provide forecasts along with uncertainty and a limited number of readable drivers, with documented overrides. Feedback should be systematically evaluated to identify missing data, model limitations, and inconsistent operating practices.

Ethical, Privacy, and Governance Considerations

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

Retail transaction data can include commercially sensitive information regarding supplier terms, local demand and pricing. Data sharing among separate stores must be done with explicit consent, with limits on purpose, access control, and ensuring that the businesses' performance does not become public knowledge. While aggregated learning can be beneficial, it is not a privacy guarantee because aggregation does not equate to privacy. Written agreements and relevant law should govern data retention, security and cross-store modeling.

The system must store training data, definitions of features, versions of models, time periods of forecasts, limitations, and business rules for ordering products. Human overrides need to be logged, and the model should default to a simple baseline if external data or serving of the model are unavailable. Forecast bias, error by category, staff adoption, waste and stockouts are among the elements that should be monitored. Only retrain a model when data-quality checks and validations have been performed, not because of an alert.

8.2 Limitations

This paper does not offer any new empirical findings. The proposed model's performance relies on the availability and reliability of transaction, inventory, expiry and external data. Some markets are very fragmented, and stock-out data and product identification might not be complete. Ordering rules, staff behavior, supplier constraints and consumer substitution are the mechanisms through which forecast accuracy relates to waste. Historical patterns may not be reliable in some areas, as weather conditions and event features could be unavailable or incorrect, and sudden economic or supply shocks can affect the reliability of historical patterns.

9. Conclusion

This paper introduced a demand forecasting model for perishables in a fragmented retail market based on artificial intelligence. The framework brings together data standardization, statistical baselines, global machine-learning models, cross-store information sharing, probabilistic forecasting and shelf-life-aware inventory decisions. It explicitly considers stockout-censored sales, heterogeneity at the local level, sparse histories, and the asymmetric cost of waste and lost sales.

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

The main finding is that forecasting needs to be assessed in an inventory decision system. Models should be validated using rolling origins, against appropriate baselines, and evaluated using statistical measures and operational results. Cross-store learning can be useful for data-poor retailers, but must maintain local differences and commercial confidentiality. Future empirical research could validate the framework with actual store data, measure the reduction of waste and availability and explore the value of data and model elements for retailers that lack digital infrastructure.

References

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control (5th Edition). Wiley.

Croston, J. D. (1972). Making a prediction and managing inventory in the case of sporadic orders. Operational Research Quarterly, 23(3), 289–303. https://doi.org/10.1057/jors.1972.50

Hyndman, R. J., & Koehler, A. B. (2006). Yet another discussion of measures of forecast accuracy. International Journal of Forecasting, 22(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

For instance, Makridakis, S., Spiliotis, E., & Assimakopoulos, V. 100,000 time series and 61 forecasting methods in the M4 Competition. International Journal of Forecasting, 36(1), 54–74. https://doi.org/10.1016/j.ijforecast.2019.04.014

Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). Probabilistic forecasting with autoregressive recurrent networks: DeepAR. International Journal of Forecasting, 36(3), 1181–1191. https://doi.org/10.1016/j.ijforecast.2019.07.001

Syntetos, A. A., & Boylan, J. E. (2005). The accuracy of estimates of intermittent demand. International Journal of Forecasting, 21(2), 303–314. https://doi.org/10.1016/j.ijforecast.2004.10.001

Taylor, J. W., & Letham, B. (2018). Forecasting at scale. The American Statistician, 72(1), 37–45. https://doi.org/10.1080/00031305.2017.1380080

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics and big data: A revolution that will change the way supply chain design and management are done.

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan

[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)

This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Technological Information, Management & Engineering Sciences (2020),1(2), 122-137

[1] Department of Artificial Intelligence, NED University of Engineering and Technology, Karachi, Pakistan

[2] Department of Computer Science, University of Karachi, Karachi, Pakistan

Copyright (c) 2020 Author(s)

This work is licensed under a Creative Commons Attribution 4.0 International License.