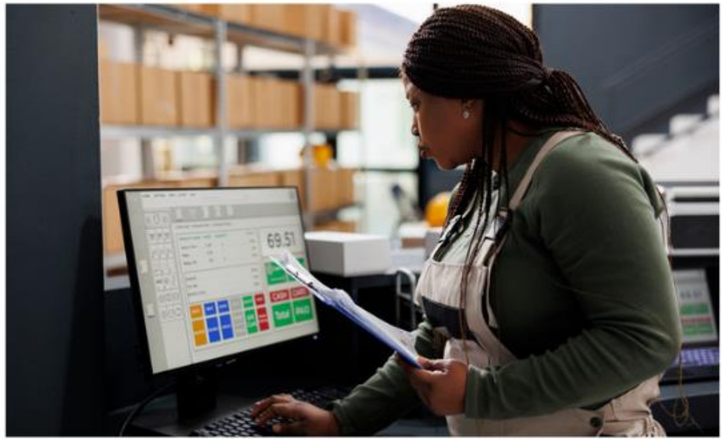
Inventory Management System.

Anish Sridharan
Jahnavi G Shankar
Japsahaj Kaur
Rishav Kumar

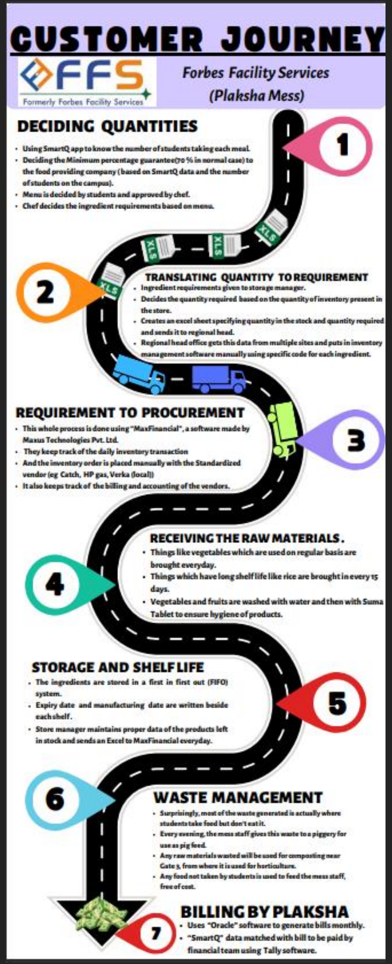
Intelligent inventory management system.



- Our "Intelligent Inventory Management System for MSMEs" revolutionizes inventory control for small and medium-sized enterprises.

- Through machine learning rigorous data analysis, and analyzing the market trends the system enables businesses to optimize product quantities, streamline operations, and improve decision-making.

- Our system also aims to implicate demand prediction of quantities and also error proofing to improve client and customer needs.



What did we do?

- From the primary market research, we identified the market segment, developed a comprehensive decision flow, and benchmarked the major players in the market.

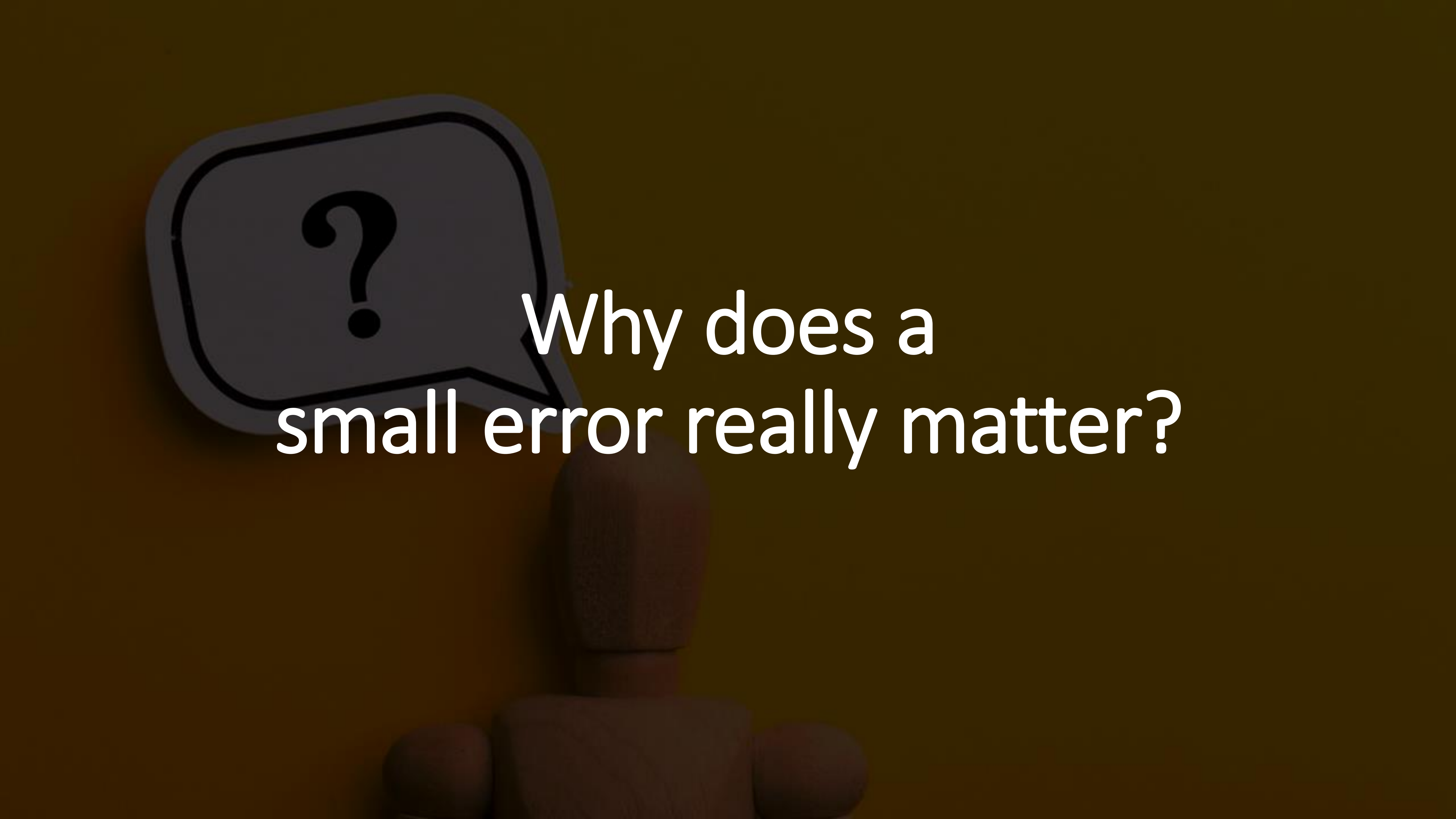
- We interviewed several retailers and Plaksha mess management to understand the need, of inventory management, the customer journey, and the challenges.

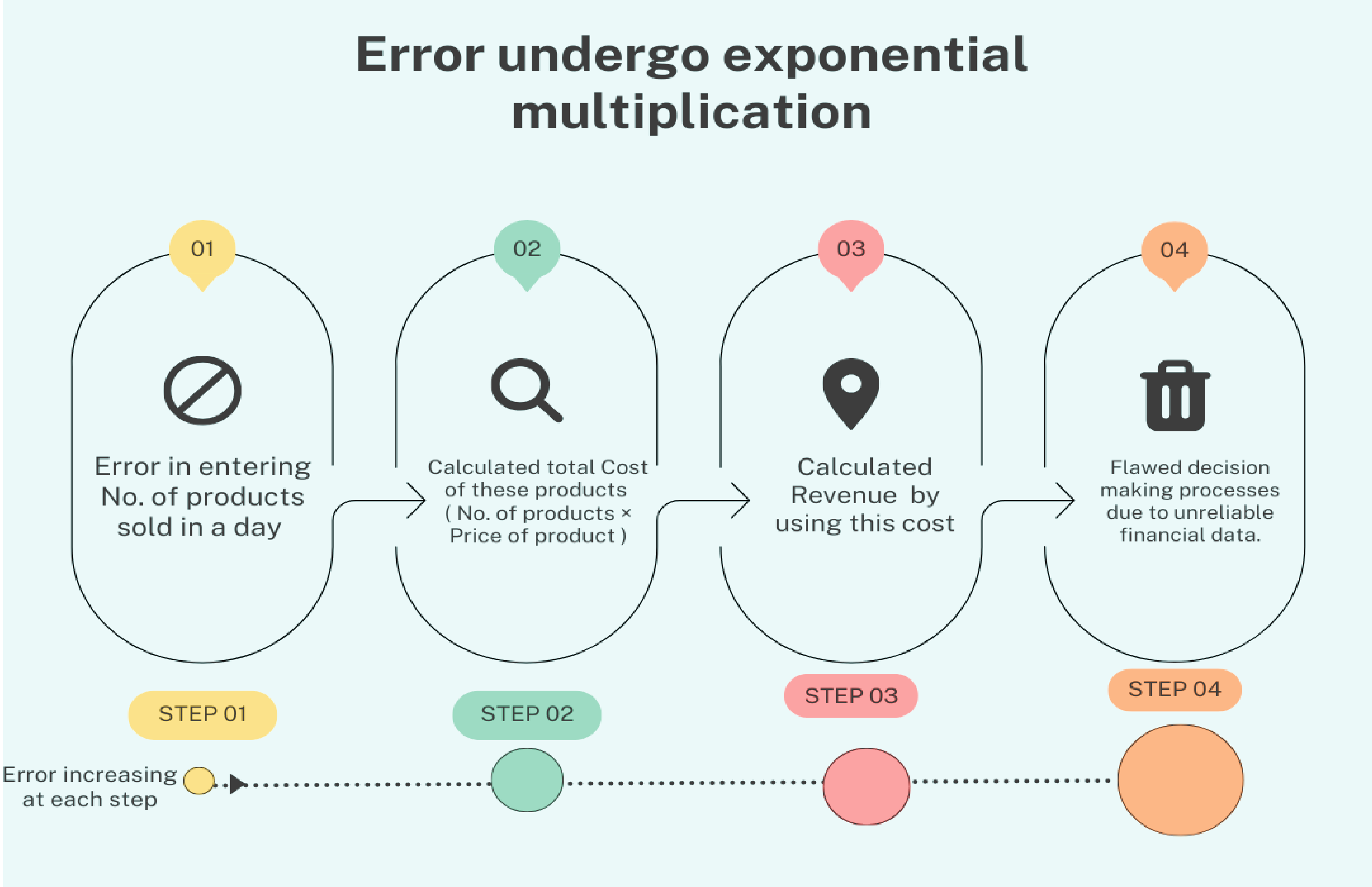
We deduced from our primary research:

Error proofing

Demand prediction

Are important aspects of Inventory management software.


Why does a
small error really matter?



Error undergo exponential multiplication
01
Error in entering No. of products sold in a day
02
Calculated total Cost of these products (No. of products × Price of product)
03
Calculated Revenue by using this cost
04
Flawed decision making processes due to unreliable financial data.
STEP 01
STEP 02
STEP 03
STEP 04
Error increasing at each step

Why Error proofing is required?

Small Error in the system

A small error occurs while entering data manually in an Excel file

1

2

Loss of Precision during data aggregation.

Data from several mess are compiled and shared with Chandigarh's central center. The error increases due to aggregation

Data shared at national level

This aggregated data is shared at national level for ordering food from different vendors

3

4

Discrepancies in final inventory calculations

Errors cause the company to order the wrong amount of inventory, resulting in huge losses

What happens if error-proofing is implemented?

Prevent Tax Miscalculations:
- Errors lead to discrepancies in tax calculations (e.g., GST).
- Underpayment results in penalties, while overpayment reduces profitability.

Efficient Operations:
- Inaccurate data causes inefficiencies in order management and decision-making.
- Poor inventory management hampers overall operational efficiency.

Reduced Costs:
- Errors lead to wasted resources, rework, and potential fines.
- Error-proofing prevents these issues, resulting in cost savings.

Improved Brand Reputation:
- Consistent accuracy builds trust with stakeholders.
- Error-proofing demonstrates a commitment to quality, enhancing brand image.

We looked at state-of-the-art algorithms that can help achieve our needs.

The main criteria here was speed, such that it can flag anomalies in real time.

"MELFORD: Using Neural Networks to Find Spreadsheet Errors", Singh et al., 2017

	A	B	C	D	E	F
1		July	August	September	October	November
2	North	4	3	4	2	4
3	East	7	8	9	5	10
4	South	4	5	6	4	7
5	West	5	6	6	5	7
6						
7	Total	20	22	26	16	28

Figure 2: Result of MELFORD correctly identifying error at D7 in the example spreadsheet.

	A	B	C	D	E	F
1		July	August	September	October	November
2	North	4	3	4	2	4
3	East	7	8	9	5	10
4	South	4	5	6	4	7
5	West	5	6	6	5	7
6						
7	Total	=SUM(B2:B5)	=SUM(C2:C5)	26	=SUM(E2:E5)	=SUM(F2:F5)

(a) Formula view.

- The **first** Microsoft Research paper to implement neural networks to solving the spreadsheet anomaly detection problem.
- In a nutshell, the researchers use a classifier based on feedforward networks and Long Short Term Memory Networks (LSTMs) to predict whether a certain cell in a spreadsheet contains a number where it should be containing a formula.
- The paper also mentions some measures already in place for this problem, but doesn't divulge too many details (Microsoft can't go around detailing its proprietary architecture, you know).

"How effectively can spreadsheet anomalies be detected: An empirical study", Zhang et al., 2016

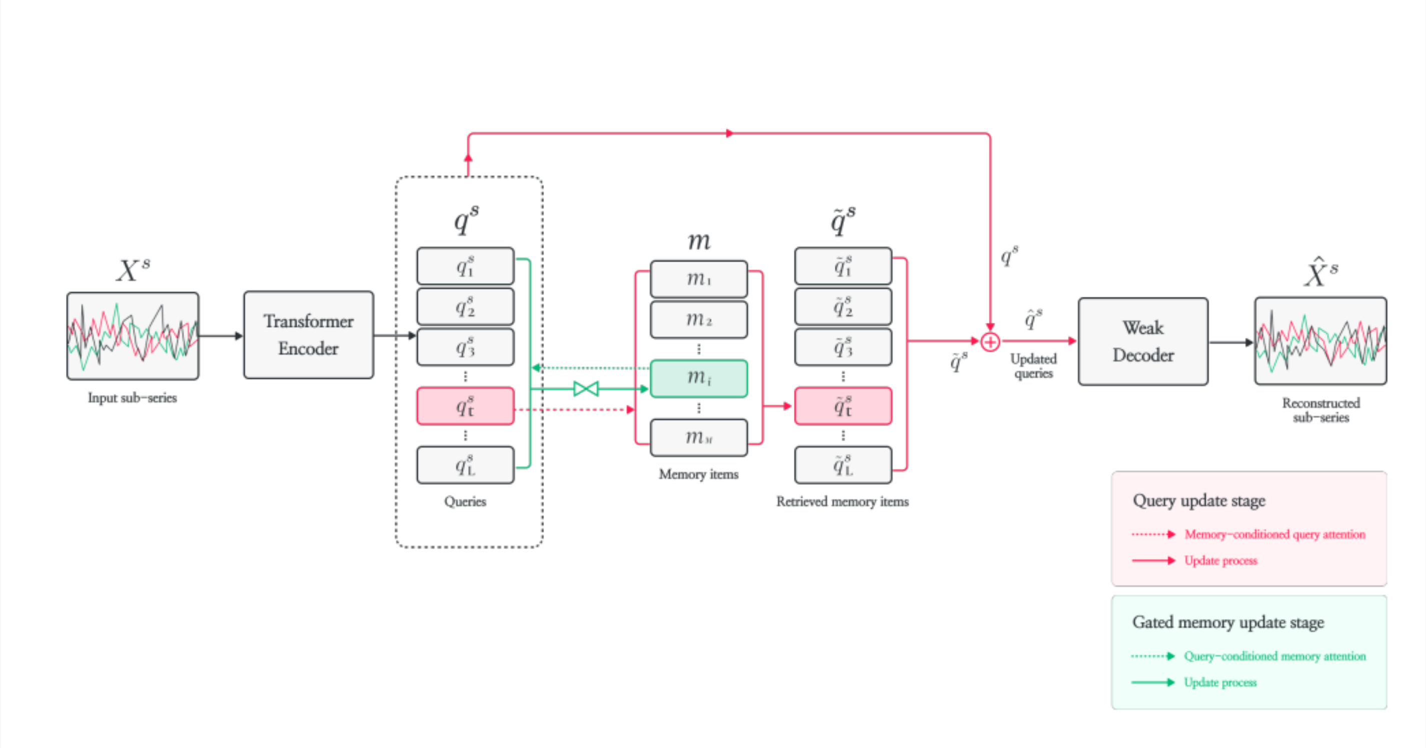
- Very useful for the breakdown it provided on the general types of anomaly detection in spreadsheets- broadly, preventing entry of erroneous data, detection of such data, and automatic fixing of it. Gave us a lens to view methods with in the future.
- The empirical study was on three algorithms- AmCheck (Dou et al., 2014), UCheck (Abraham and Erwig, 2007) and Dimension (Chambers and Erwig, 2009).
- It also referred to many other methods in the literature (Luckey et al., 2012; Cunha et al., 2014, Badame and Dig, 2012; Burnett and Erwig, 2002; Hermans et al., 2012; Hermans et al., 2013, Abraham and Erwig, 2008; Abraham and Erwig, 2009)

"Forecasting and Anomaly Detection approaches using LSTM and LSTM Autoencoder techniques with the applications in Supply Chain Management"- Nguyen et al., 2021

- Code available on GitHub.
- After initial training, it still takes "a few minutes" to retrain on the current user data. This works if the product is intended to be used just before audit, but we intend our product to be active all the time.
- For real-time usage, we should ideally be able to achieve this in a fraction of a second.

Method & Refs.	RMSE
MTW-BLSTM ensemble (Xia et al., 2020)	12.61
LSTM- FW-CatBoost (Deng et al., 2020)	15.8
RBM-LSTM-FNN (Ellefsen et al., 2019)	12.56
Proposed method	**9.71**

Table 2: RMSE comparison with the literature on the C-MAPSS FD001 dataset



MEMTO: Memory-guided Transformer for Multivariate Time Series Anomaly Detection

- MEMTO is an unsupervised reconstruction-based model for multivariate time series anomaly detection.
- It utilizes a Gated memory module to adaptively capture normal patterns in response to input data.
- The model uses a two-phase training paradigm involving K-means clustering for memory item initialization enhancing stability and robustness.
- MEMTO introduces a bi-dimensional deviation-based anomaly criterion, considering both input and latent space for comprehensive anomaly detection.
- Experimental results show that MEMTO achieves good detection results compared to existing methods on real-world benchmark datasets

- The F1 score of MEMTO for anomaly detection is reported to be **95.74%**
- MEMTO is designed for multivariate time series anomaly detection, such as in computer vision applications
- MEMTO's Transformer encoder may not effectively utilize features or identify relevant queries in univariate datasets like inventory transactions and stocks required.
- The decoder tends to map inputs to the highest transaction value encountered, lacking contextual understanding of the inventory management system.
- MEMTO's architecture and functionality are not aligned with the objectives of error-proofing and accurately detecting anomalies in inventory management systems.

The algorithms which we tested are:

1)Isolation forest

2)MEMTO **(**Memory-guided Transformer for Multivariate Time Series Anomaly Detection)

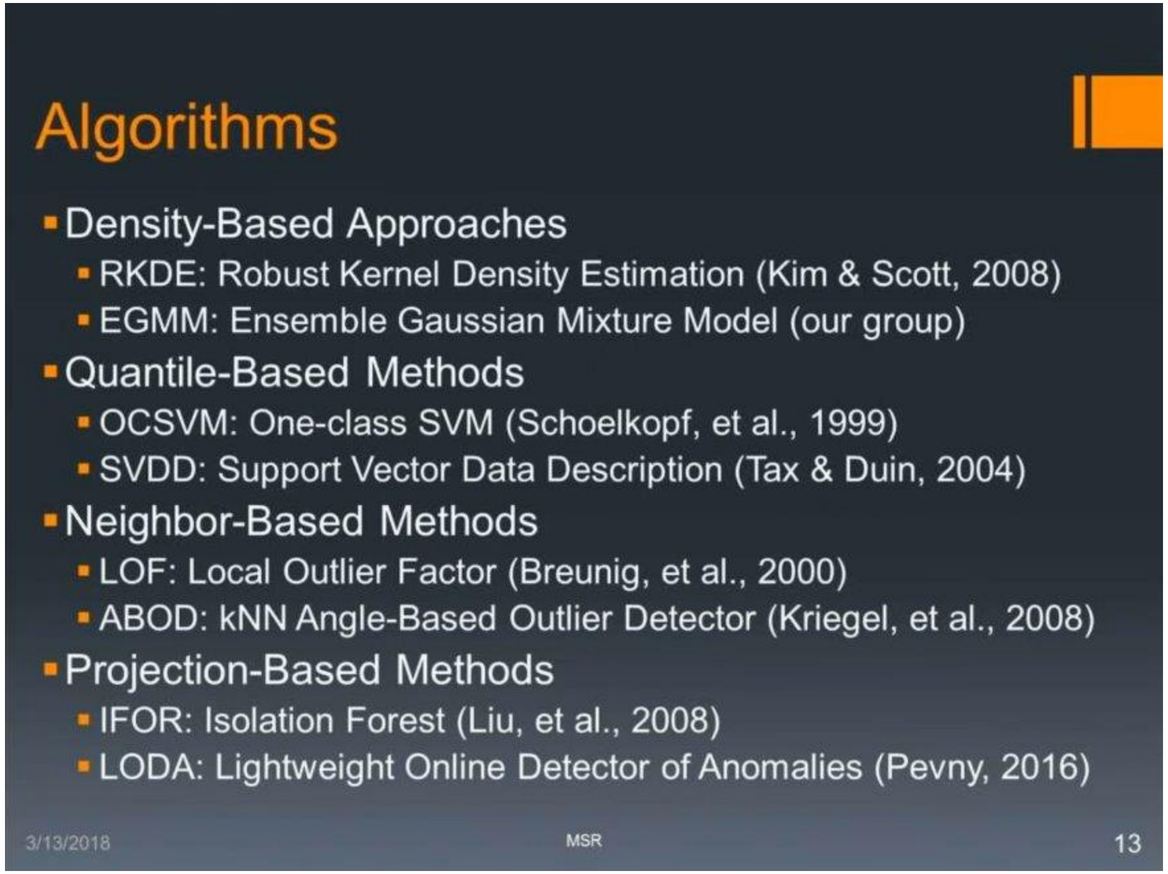
3)EGMM (Ensemble Gaussian mixture model)

4)RKDE (Robust Kernel Density estimation)

5)One class SVM

"Anomaly Detection: Algorithms, Explanations, Applications"- Talk by Dr. Thomas Dietterich at Microsoft Research, 2018 (available on YouTube)

- Among the many algorithms and approaches here, we implemented the following-
  - RKDE
  - EGMM
  - SVM
  - IFOR
- Among these algorithms, we found IFOR to outperform the others. However, its accuracy was not reliable for different orders of error.

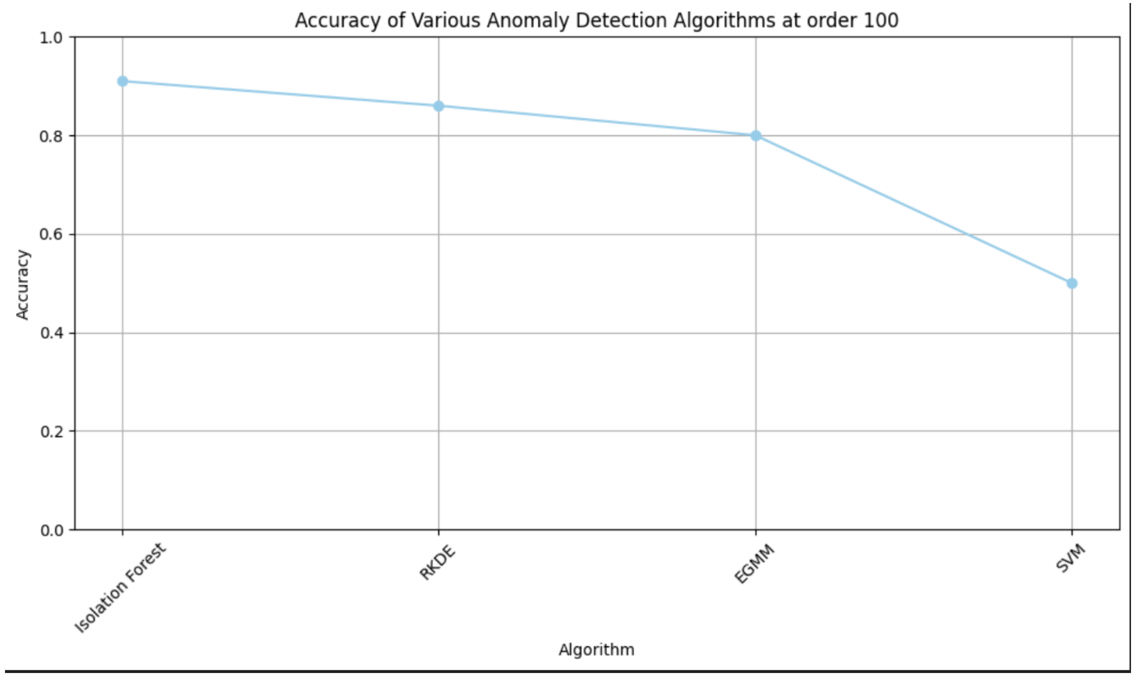


Data which we used...

- Data Source: Kaggle
- Description: Sales data of 15 stores from Favorita stores in Ecuador
- Training Data: Includes dates, store numbers, and transactions
- Entries in Training Dataset: 3,008,280

Algorithm which worked the Best!

- Order of Accuracy given by different models gives the same order as the order given by ml expert paper. Isolation forest performed the best. Recall was very high close to 1.



Problems :

- Not good accuracy on average varied from 0.8 to 0.95
- Took more time than others.

ISOLATION FOREST ALGORITHM FOR

ANOMALY DETECTION

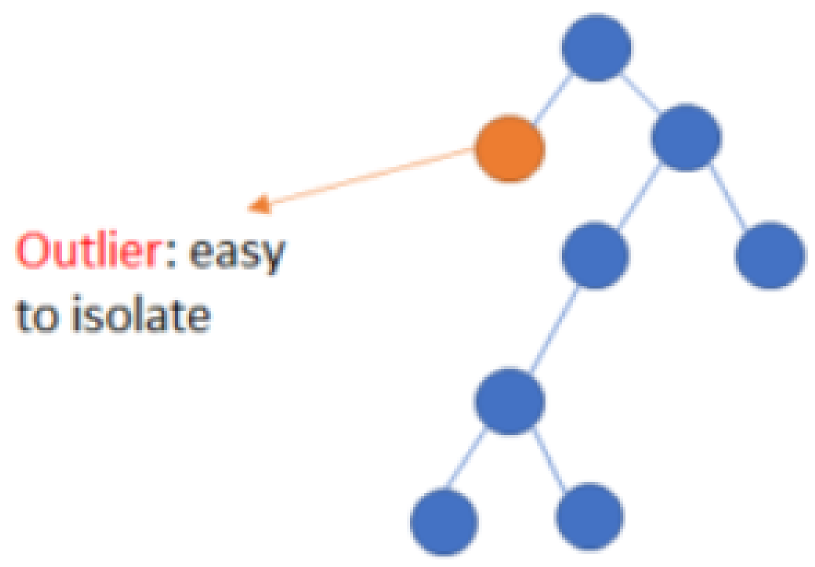
How isolation forest algorithm works?

1. **Random Partitioning:**

   Dataset is recursively partitioned randomly until all points are isolated or a stopping criterion is met.

2. **Isolation:**

   Anomalies are isolated quicker than normal points, aiding in their identification as anomalies are located in less dense regions of the feature space.



4) Path Length Calculation:

- Computes average path length of each point to isolation.

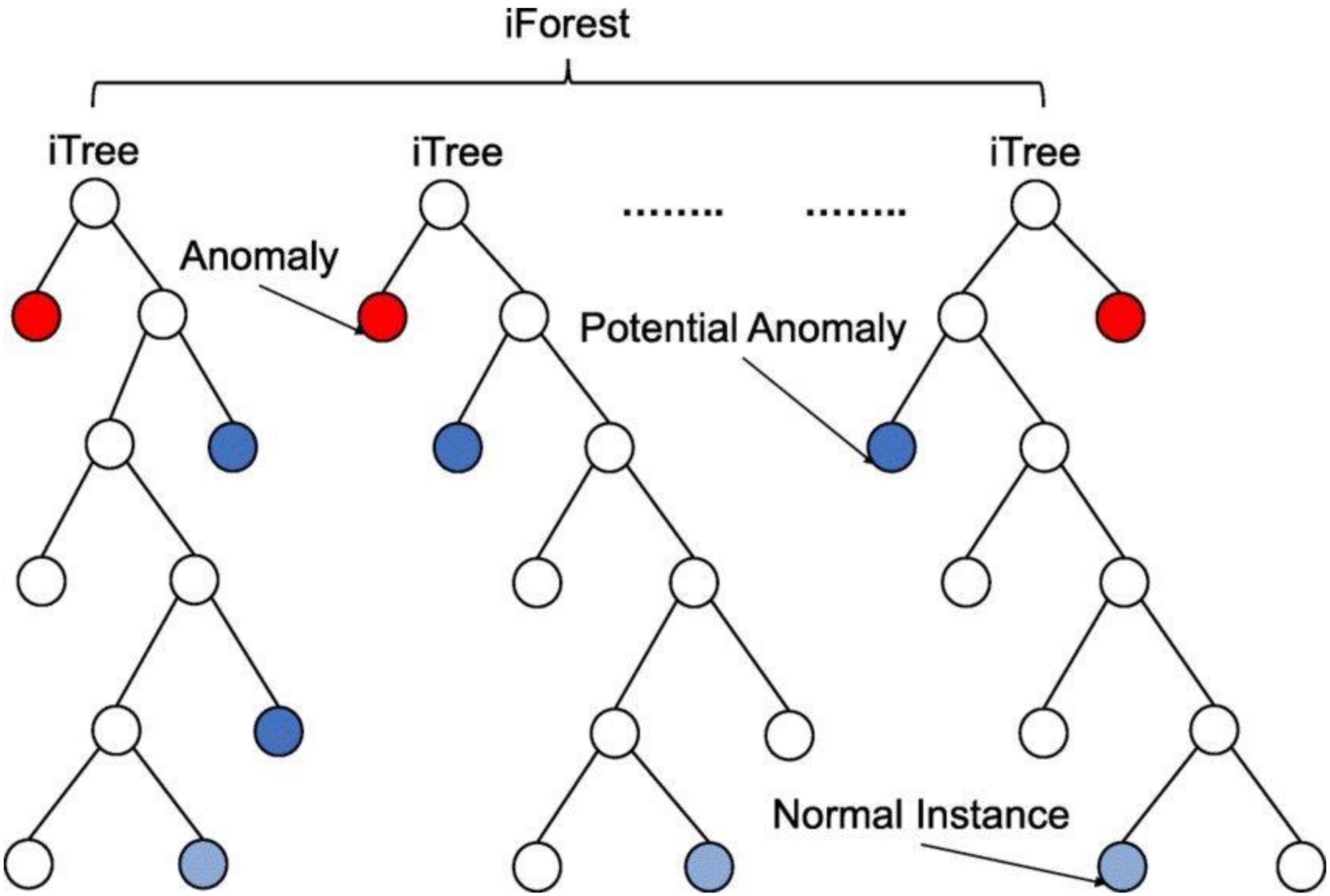
5) Anomaly Score:

- Each data point gets a score based on path length calculations.

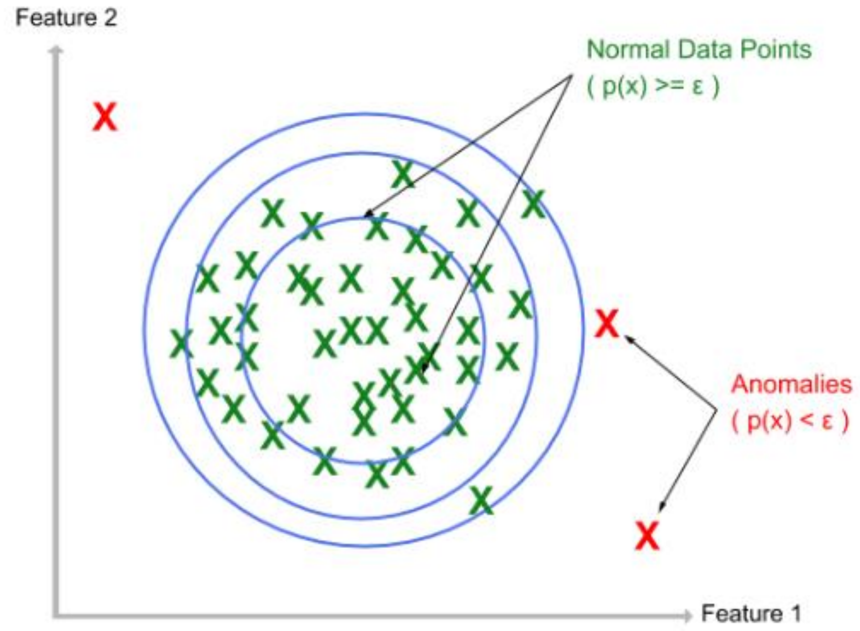
6) Decision Threshold:

- Anomalies are classified using a predefined threshold for scores. Generally a score near 1 means is likely an anomaly, while score around 0.5 means it's a normal point.

To tackle the issue of low accuracy. We looked at the algorithm closely.



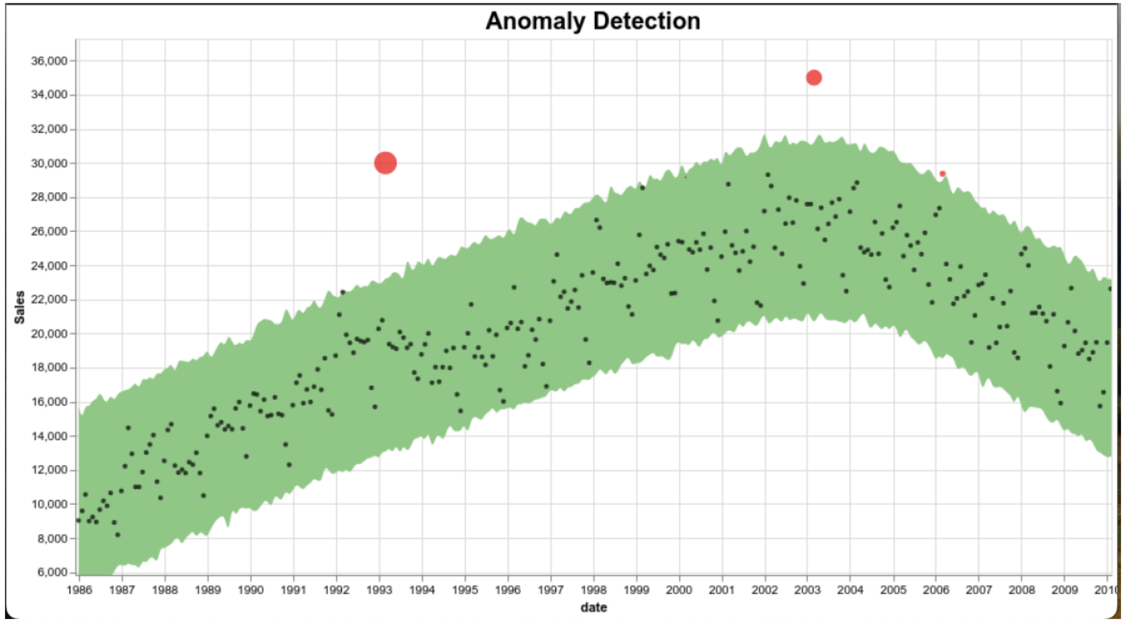
Random forest randomly selects nodes and builds trees. A data point is considered an anomaly if it is separated earlier in most of the instances.



- The algorithm currently classifies a data point as an error if it falls outside the distribution of historical data.
- This approach overlooks the natural drift present in time series data caused by factors such as the expansion of shops and other microeconomic variables.
- Consequently, it may incorrectly label changing values over time as anomalies, even when they are not anomalies in reality.
- This tendency is evident in the algorithm's high recall but comparatively low accuracy values.
- To mitigate this limitation, we propose a range-based approach that accounts for the drift within the dataset.
- By incorporating this range-based approach, we aim to provide flexibility that accommodates the inherent drift in the data, thereby enhancing the accuracy of anomaly detection.

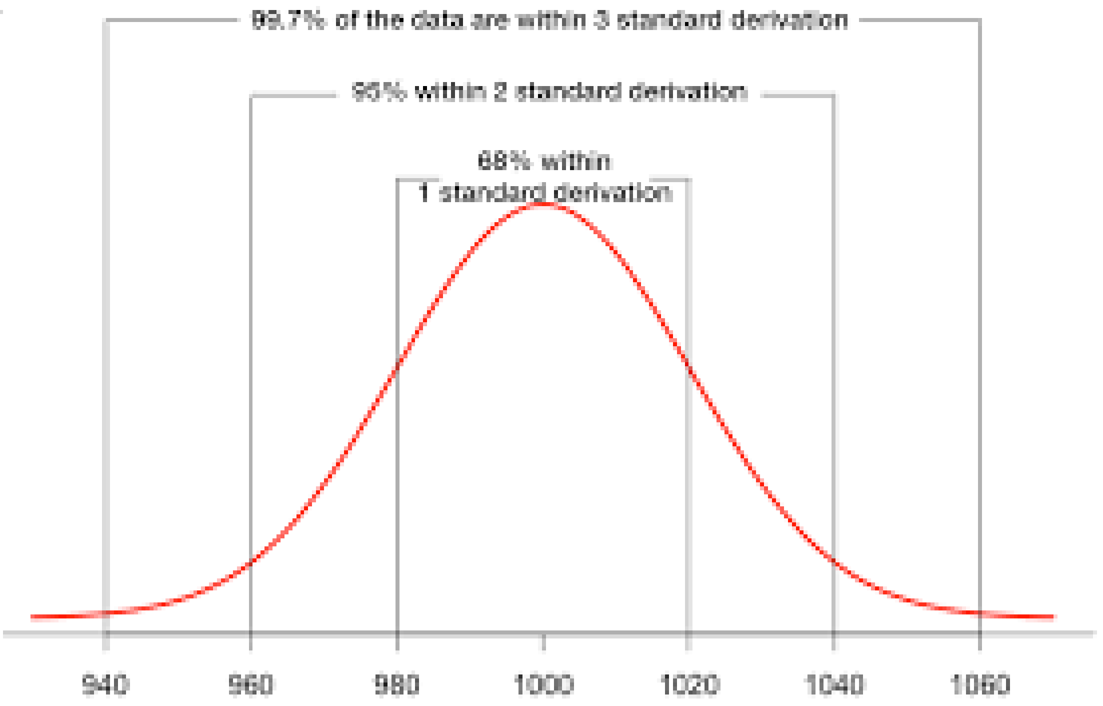
KNN Based Heuristic

- **Steps of the Heuristic:**
- Find K nearest Neighbour.
- Calculate the mean of the nearest neighbors.
- Calculate variance of the dataset.
- Calculate the range which is "mean of neighbors -standard deviation" to "mean of neighbors +standard deviation."
- Classify as error if data point is not in the range.



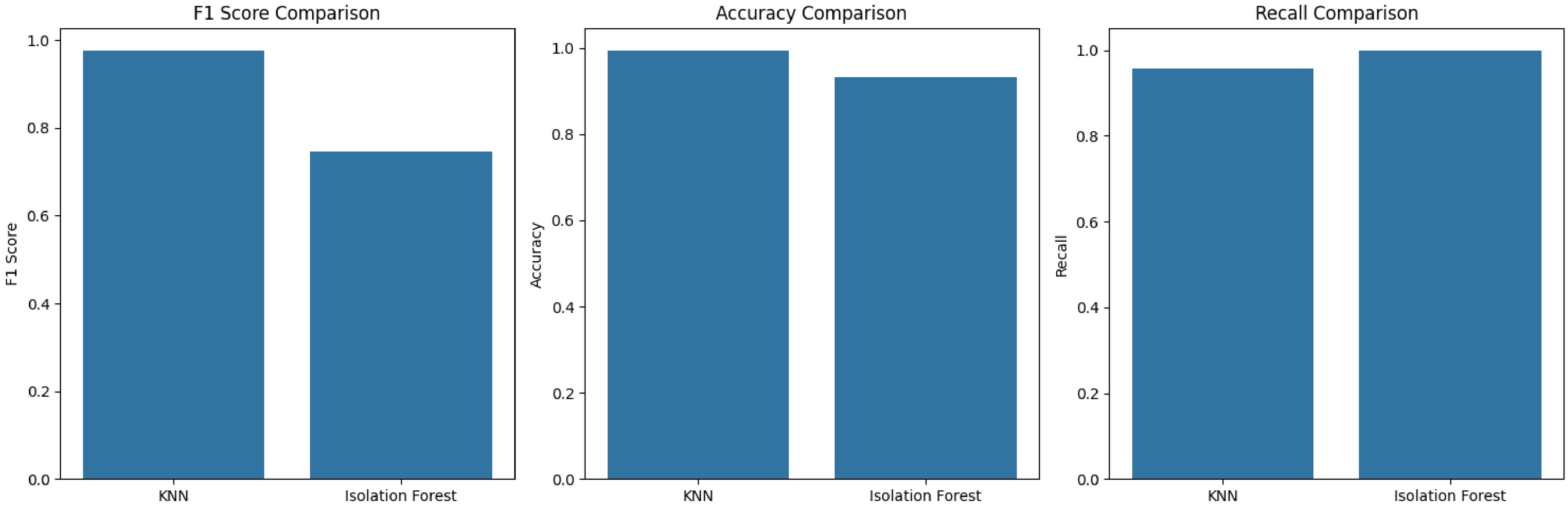
Intuition behind the use of Standard Deviation to find the range is based on statistics.

This choice stems from the intuition that in a normal distribution, approximately 68% of the data falls within one standard deviation of the mean, capturing the typical variability in transaction values.



Let's see a small demo on the working of heuristic.

- Using the heuristic we get better result in all the parameters and recall remains high comparable to isolation forest.

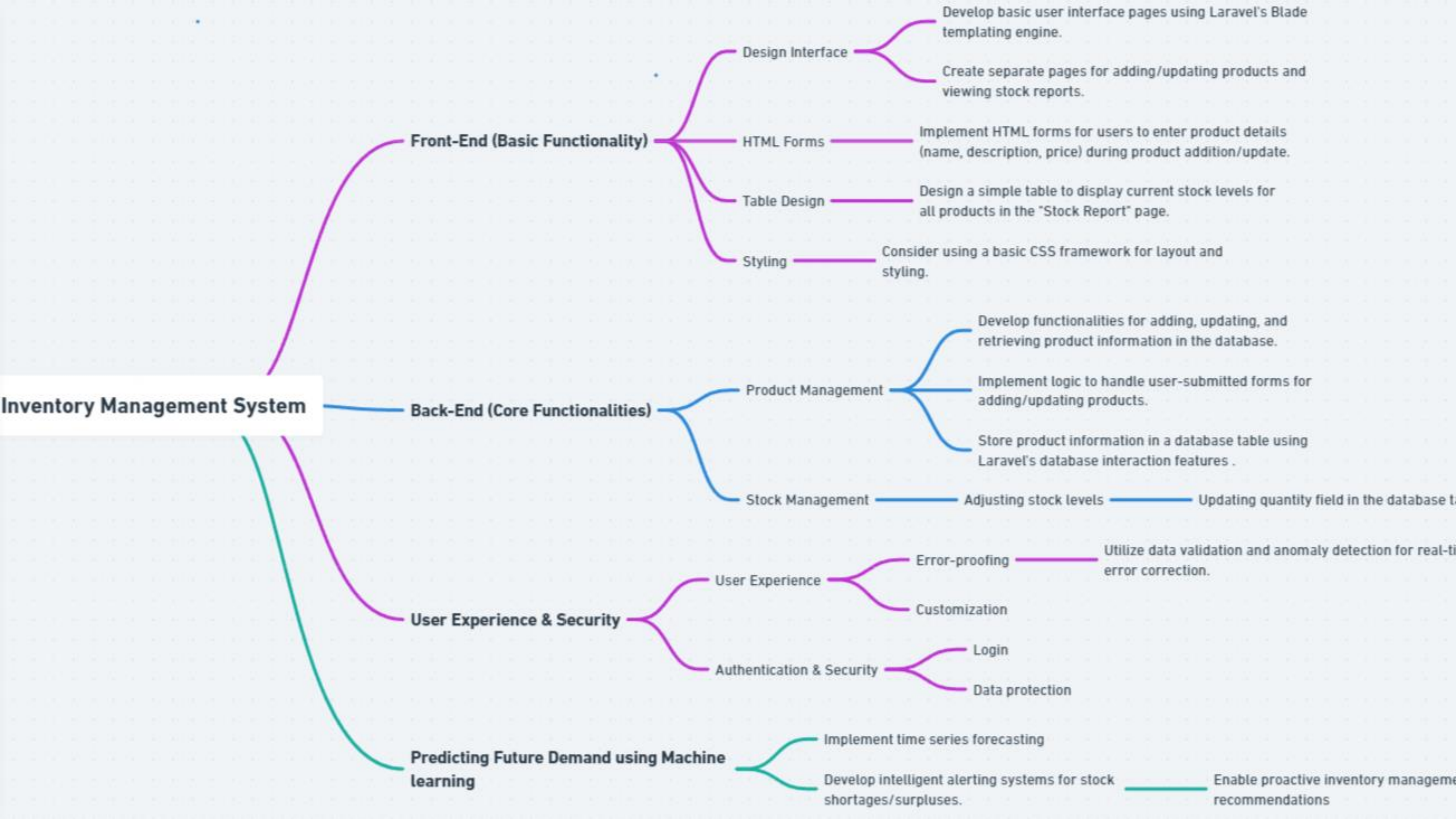



QuickTime Player
File
Edit
View
Window
Help
Fri 3 May 3:13 PM

Future tasks:

- Develop a demand prediction tool for the Plaksha Mess Integrate demand predication with the inventory management system

- To integrate the error-proofing algorithm with our developed inventory managing software.


Inventory Management System
Front-End (Basic Functionality)
Design Interface
Develop basic user interface pages using Laravel's Blade templating engine.
Create separate pages for adding/updating products and viewing stock reports.
HTML Forms
Implement HTML forms for users to enter product details (name, description, price) during product addition/update.
Table Design
Design a simple table to display current stock levels for all products in the "Stock Report" page.
Styling
Consider using a basic CSS framework for layout and styling.
Back-End (Core Functionalities)
Product Management
Develop functionalities for adding, updating, and retrieving product information in the database.
Implement logic to handle user-submitted forms for adding/updating products.
Store product information in a database table using Laravel's database interaction features .
Stock Management
Adjusting stock levels
Updating quantity field in the database
User Experience & Security
User Experience
Error-proofing
Utilize data validation and anomaly detection for
error correction.
Customization
Authentication & Security
Login
Data protection
Predicting Future Demand using Machine learning
Implement time series forecasting
Develop intelligent alerting systems for stock shortages/surpluses.
Enable proactive inventory
recommendations

References:

- Song, J., Kim, K., Oh, J., & Cho, S. (2024). MEMTO: Memory-guided transformer for multivariate time series anomaly detection. *Advances in Neural Information Processing Systems*, *36*.

- Microsoft Research (Director). (2018, April 6). *Anomaly Detection: Algorithms, Explanations, Applications*. https://www.youtube.com/watch?v=12Xq9OLdQwQ

- Singh, R., Livshits, B., & Zorn, B. (2017). MELFORD: Using neural networks to find spreadsheet errors. *Microsoft Research, Microsoft Tech Report Number MSR-TR-2017-5, Tech. Rep*.

- Zhang, R., Xu, C., Cheung, S. C., Yu, P., Ma, X., & Lu, J. (2017). How effectively can spreadsheet anomalies be detected: An empirical study. *Journal of Systems and Software*, *126*, 87-100.

- Nguyen, H. D., Tran, K. P., Thomassey, S., & Hamad, M. (2021). Forecasting and Anomaly Detection approaches using LSTM and LSTM Autoencoder techniques with the applications in supply chain management. *International Journal of Information Management*, *57*, 102282.