Operational Entrepreneurship project final report

fashion and beauty industry demand forecasting using BMA by ensemble learning approach.

Guided by prof. Lohithaksha Maniraj Maiyar

By Ankith Jain S

(EM23MTECH14004)

Abstract

This study explores the application of Bayesian Model Averaging (BMA), in the beauty products supply chain to enhance forecasting precision, resilience and flexibility. By utilizing a cluster based methodology the data is segmented into clusters based on product categories, customer characteristics and revenue indicators. This segmentation enables strategies for market segments contributing to the advancement of modern forecasting techniques in complex industrial settings.

Introduction

The fashion and beauty sector requires demand forecasting to optimize supply chains and reduce costs effectively. Conventional methods often struggle with fluctuations and unpredictability resulting in inventory or stock shortages. This project introduces Bayesian Model Averaging (BMA) an learning method that revolutionizes demand prediction, in cosmetics supply chains. BMA enhances forecasting accuracy by incorporating model uncertainty and harnessing the strengths of models. Through a cluster based approach data is divided into clusters based on product types, customer profiles and revenue metrics. Each cluster is analyzed independently to develop tailored strategies aligned with market segments.

Literature Review

The fashion and beauty sector encounters hurdles when it comes to predicting demand because of shifts and consumer driven fluctuations. This review delves into sophisticated approaches to tackle these obstacles. The techniques encompass Bayesian Model Averaging, Cluster Based Modeling and Ensemble Learning all pertinent, within the realm of Supply Chain Management, in the fashion and beauty industry.

Bayesian Model Averaging (BMA) and Ensemble Learning in Fashion and Beauty Industry

Bayesian Model Averaging (BMA) makes predictions better by dealing with uncertainty when choosing models. In sectors like fashion and beauty, BMA combines the top models, based on their chances, to adjust to new data and interactions. Ensemble Learning adds to BMA by merging models to improve predictions. In predicting demand, ensemble methods mix different data and models to handle various influences. Cluster-Based Modeling splits big data into similar groups for focused analysis in supply chain management strategy.

Supply Chain Management in the Fashion and Beauty Industry

Fashion and beauty change a lot. People like different things at different times. The industry needs to plan and be ready to change quickly. They need to manage getting products to people without making too much or too little. The industry is hard. They need smart ways to predict what people will want.

Synthesizing Insights from Literature

1. Recognizing the Need for Advanced Forecasting Techniques: Insights from Christopher, Peck, and Towill (2006) on the dynamics of supply chain management underscore the necessity for advanced forecasting techniques that can cope with the high variability and rapid changes characteristic of the fashion and beauty industry. Traditional forecasting methods often fall short in accuracy and adaptability, leading to significant inventory mismanagement.

2. Adoption of Bayesian Model Averaging: Bayesian Model Averaging (BMA) was identified through the work of Hoeting et al. (1999) as an effective statistical technique that accounts for model uncertainty, providing a probabilistic framework that enhances prediction reliability. BMA's ability to incorporate multiple models and weight them based on their performance addresses the complex and nonlinear relationships that often exist in fashion and beauty product sales data.

3. Utilization of Cluster-Based Modeling: The strategic advantage of cluster-based modeling in managing diverse datasets by segmenting them into more homogeneous groups was highlighted by Kaufman and Rousseeuw (2009). This method aligns well with the segmented nature of consumer demographics and product types in the makeup industry, allowing for more customized and precise forecasting approaches.

4. Integrating Ensemble Learning: Ensemble learning's strength in improving the robustness and accuracy of predictive models through the aggregation of multiple model outputs was critically reviewed by Dietterich (2000). The literature supports the use of ensemble methods in scenarios where no single model can capture the entire spectrum of data behaviors, which is typical in industries driven by trends and consumer preferences.

Problem Description

In the fast-paced and ever-changing fashion and beauty industry, accurate demand forecasting is crucial for optimizing supply chain efficiency and minimizing operational costs. The industry's rapid changes in consumer preferences, seasonal influences, and high product variability make it challenging to forecast demand accurately. Traditional forecasting methods often struggle to adapt to these conditions, leading to either surplus inventory or product shortages, both of which are costly to businesses.

Scientific Context

The problem involves developing a robust forecasting model that can handle the complexities of the fashion and beauty supply chain. The primary scientific challenges addressed in this project include:

Non-linearity and Multimodality: The relationships between various factors influencing demand are non-linear and often multimodal, complicating the application of traditional linear forecasting models.

Data Segmentation and Heterogeneity: The varied nature of products within the fashion and beauty industry requires a segmented approach to demand forecasting. Each segment, defined by product type or consumer demographics, may exhibit unique behaviors and trends, requiring tailored forecasting strategies.

Model Uncertainty: No single statistical or machine learning model can sufficiently capture the complexities of the dataset or provide consistently accurate forecasts, posing significant challenges for traditional single-model approaches.

Dynamic Market Conditions: The fashion and beauty market is subject to rapid shifts due to emerging trends and external influences, such as social media. Forecasting models must be adaptable and responsive to these changes to remain effective.

Proposed Solution: Bayesian Model Averaging (BMA) and Cluster-Based Modeling

To address these challenges, this project proposes a novel approach combining Bayesian Model Averaging (BMA) with cluster-based modeling. BMA enhances forecasting accuracy and robustness by considering a combination of models and averaging their predictions based on calculated posterior probabilities, thus accommodating model uncertainty and non-linearity. Cluster-based modeling allows for the data to be segmented into meaningful groups (clusters) based on shared characteristics, facilitating more accurate and tailored forecasting by focusing on the specific dynamics of each group.

The integration of these two methods into an ensemble learning framework aims to leverage the strengths of multiple predictive models and data segmentation strategies to create a highly adaptive and accurate forecasting system. This system is designed to mitigate the issues of overfitting, enhance the flexibility of the forecasting model, and ultimately provide a more reliable tool for supply chain management in the fashion and beauty industry.

Problem Formulation

The aim of this project is to develop a segment-based demand forecasting model using Bayesian Model Averaging (BMA) within an ensemble learning framework. This approach addresses the non-linear, high-dimensional, and variable characteristics of the fashion and beauty industry's demand data. Below, we formalize this problem mathematically and outline the development of the corresponding hypothesis.

Mathematical Formulation

The primary objective is to predict future demand y based on a set of explanatory variables X, which include price, availability, stock levels, lead times, order quantities, and marketing costs. We define the demand forecasting problem as estimating the function f such that:

$$\boldsymbol{Y = f(X) + \epsilon}$$

where ϵ represents the error term, capturing unmodeled influences or random fluctuations.

Bayesian Model Averaging (BMA)

In BMA, the forecast **y** is the weighted average of forecasts from multiple models Mi where the weights wi are proportional to the posterior probabilities of the models given the data. This can be expressed as:

$$\boldsymbol{Y = \sum wi\, fi(X)}$$

Posterior probability

This formula converts the BIC scores into probabilities that can be used to compare different models statistically. The weights wi are computed based on the Bayesian Information Criterion (BIC) for each model:

$$Wi = \frac{\sum \exp(-0.5 \times \mathrm{BIC}j)}{\exp(-0.5 \times \mathrm{BIC}i)}$$

Bayesian Information Criterion (BIC)

When it comes to choosing the best model from a set of options, the Bayesian Information Criterion (BIC) is a powerful tool that helps us make an informed decision. This formula is based on the likelihood function, which measures how well a model fits the data. But BIC doesn't just stop at the data - it also takes into account the complexity of each model, penalizing those with more parameters.

$$\mathrm{BIC} = -2 * \ln(L) + k * \ln(n)$$

Where:

- L is the maximum likelihood of the model
- k is the number of parameters in the model
- n is the sample size

Cluster-Based Modeling

Data is segmented into clusters Ck based on product type, customer demographics, and other relevant factors. Each cluster k is modeled separately:

$$Yk = fk(Xk) + \epsilon k$$

where Xk and yk are the predictors and response variable for cluster k.

Demonstration of Solution Method

This section outlines a structured approach to implementing a cluster-based demand forecasting system using Bayesian Model Averaging (BMA) and Analytic Hierarchy Process (AHP) within an ensemble learning framework. The process is divided into sequential steps, each contributing to the development of a robust forecasting model for the fashion and beauty industry's supply chain.

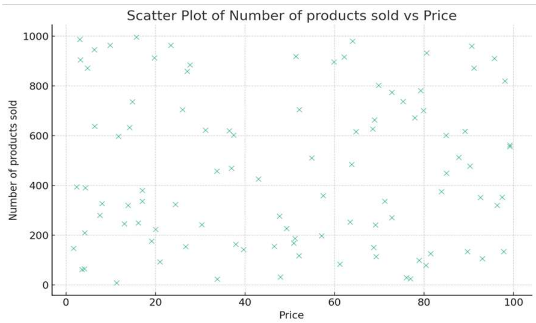
Step 1: Selecting Independent Features Affecting the Demand

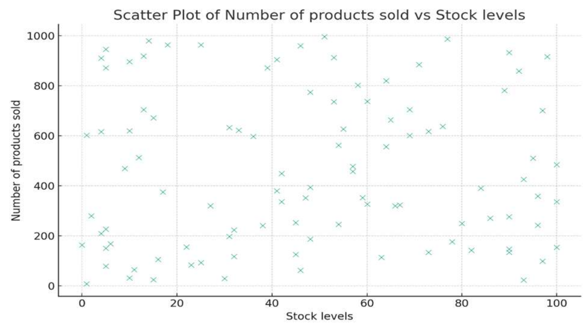
By reading research papers and discussing with PhD students

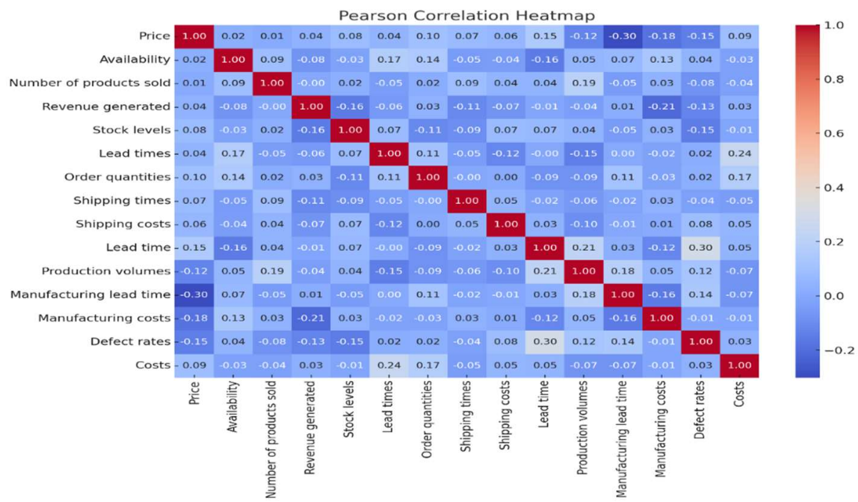
Dependent variables or input features: Price, Availability, Stock levels, Lead times, Order quantities, advertising cost.

Independent variables: Number of products sold.

Step 2: Selecting the Preferred Models by Analyzing the Relationship Between the Features





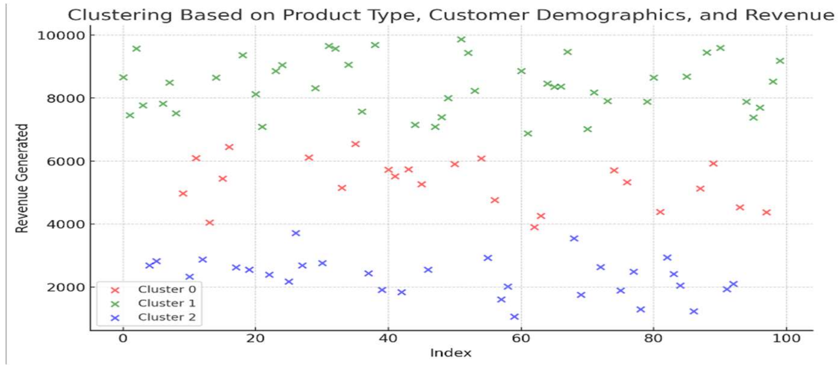


By this it is Conduct exploratory data analysis (EDA) by Pearson correlation using python, it should be close to 1 or -1.

From this it concludes it is **nonlinear relationship**. hence linear regression won't hold good. Random forest and SVM can be used for forecasting. So **linear regression, random forest, SVM** is selected. We are using liner regression even though its nonlinear because if in feature they become linear then it would be useful. Models contain both for linear and nonlinear to make the forecasting robust.

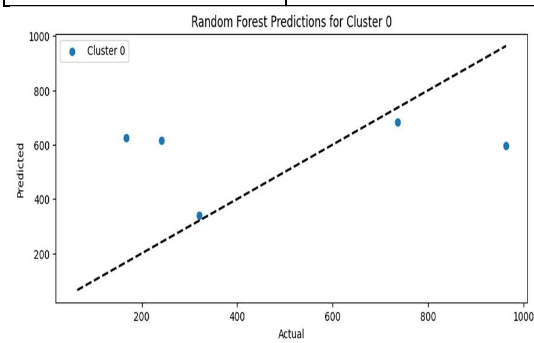
Step 3: Clustering with Product Type and Customer Demography and Revenue Using KMeans.

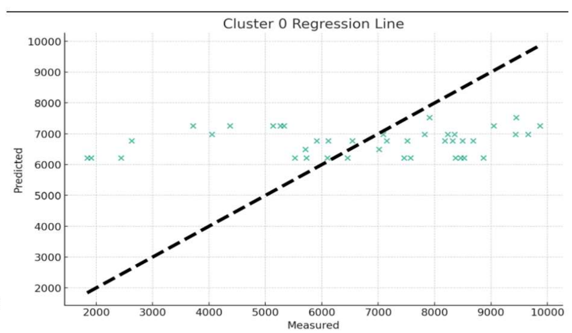
This would help us in segmentation according to revenue high, medium, low. So that we can focus on particular segment for forecasting.

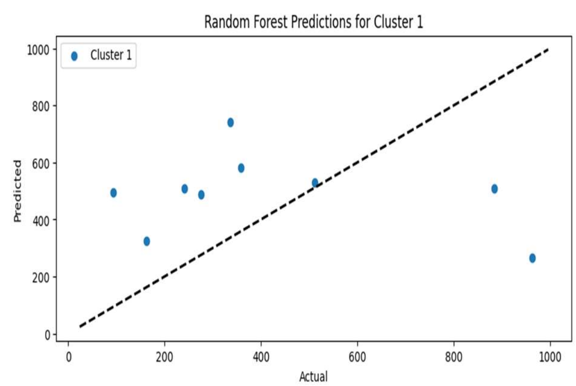


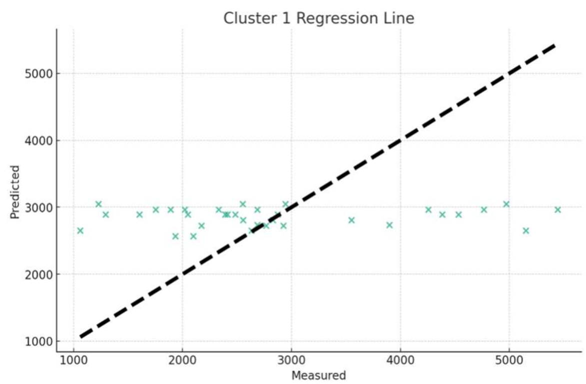
Step 4: Fitting Each Cluster with 3 Different Models (Regression, SVM, Random Forest) and calculating BIC

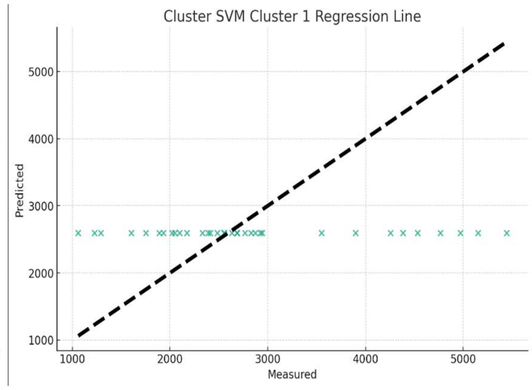
	Random forest BIC	Linear regression BIC	SVM BIC
Cluster 0	274.99	634.95	638.03
Cluster 1	276.90	485.94	488.15
Cluster 2	274.861	356.76	BIC: 359.7

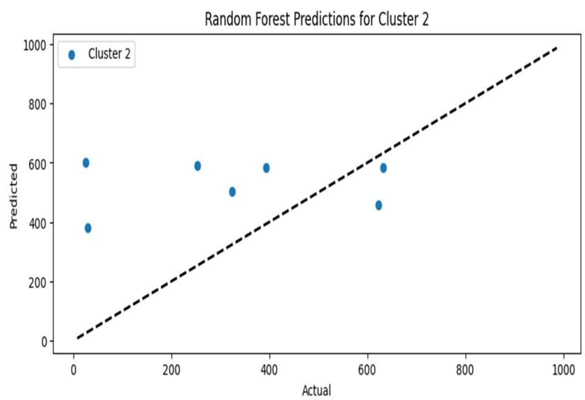


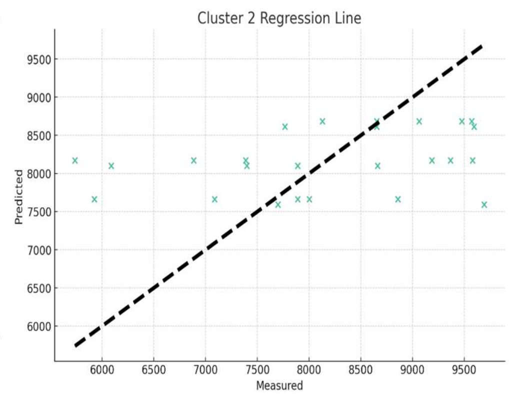












Step 5: Posterior probability of a model given the BIC scores

	Random forest ***Wr***	Linear regression ***Wl***	SVM ***Ws***
Cluster 0	**1.00**	**0.00**	**0.00**
Cluster 1	**1.00**	**0.00**	**0.00**
Cluster 2	**1.00**	**0.00**	**0.00**

```
Result
{'C0': {'Random Forest': 1.0,
  'Linear Regression': 6.849819808019933e-79,
  'SVM': 1.468471915019697e-79},
 'C1': {'Random Forest': 1.0,
  'Linear Regression': 4.050799535171125e-46,
  'SVM': 1.3416688878294486e-46},
 'C2': {'Random Forest': 1.0,
  'Linear Regression': 1.6438345845359408e-18,
  'SVM': 3.779594644161433e-19}}
```

From this we can understand that this problem is well fitted to random forest hence it has got all the weight, but in other cases all models will have different weightage.

step 6: applying the BMA using the weights.

$$BMAcluster\ k = Wr \times RFk + Wl \times LRk + Ws \times SVMk$$

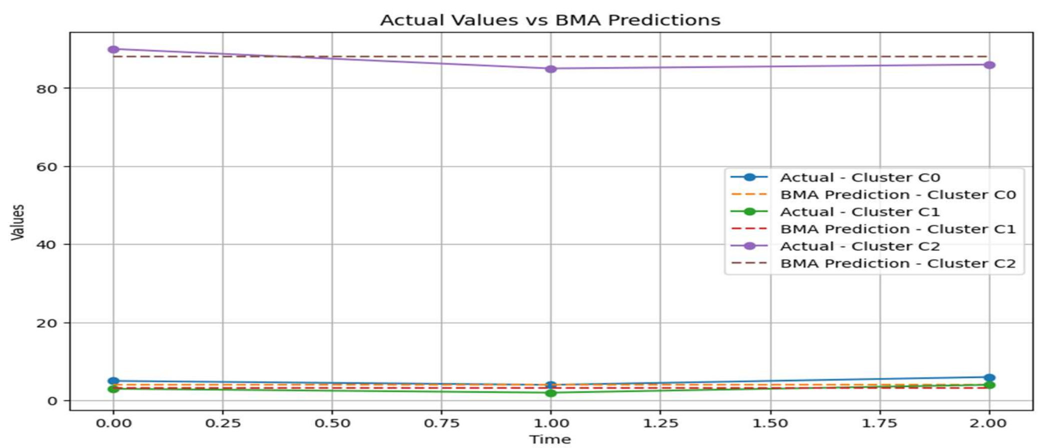
$$BMAcluster\ 0 = 1 \times RF0 + 0 \times LR0 + 0 \times SVM0 = RF0$$

$$BMAcluster\ 1 = 1 \times RF1 + 0 \times LR1 + 0 \times SVM1 = RF1$$

$$BMAcluster\ 2 = 1 \times RF2 + 0 \times LR2 + 0 \times SVM02 = RF2$$

	BMA Prediction
Cluster 0	**0.608**
Cluster 1	**0.314**
Cluster 2	**0.805**

Bayesian Model Averaging (BMA) confidence levels range from 0 to 1, where 1 indicates high confidence in the ensemble prediction, combining multiple models, 0 suggests greater uncertainty or variability among the models.



step 7: weighting the clusters by AHP ranking it by performing matrix of hypothesis on each cluster.

we need to create and solve the pairwise comparison matrices for each criterion, then aggregate these weights. hypothetical example based on the criteria previously mentioned: Sales Volume, Profit Margin, and Market Growth Potential.

Step 1: Define Pairwise Comparison Matrices Let's assume some values for pairwise comparisons:

Sales Volume (V):

	Cluster 0	Cluster 1	Cluster 2
Cluster 0	**1**	**2**	**3**
Cluster 1	**1/2**	**1**	2
Cluster 2	**1/3**	**1/2**	**1**

Profit Margin (P):

	Cluster 0	Cluster 1	Cluster 2
Cluster 0	**1**	**1/3**	**1/5**
Cluster 1	**3**	**1**	**1/2**
Cluster 2	**5**	**2**	**1**

Market Growth Potential (G)

	Cluster 0	Cluster 1	Cluster 2
Cluster 0	**1**	**1/4**	**1/7**
Cluster 1	**4**	**1**	**1/3**
Cluster 2	**7**	**3**	**1**

Step 2: Calculate Normalized Matrices and Priority Vectors

Calculate column sums and then normalize each entry by its column total. Compute the priority vector by averaging across rows.

	Sales volume	Profit margin	Market growth potential
Cluster 0	**0.55**	**0.10**	**0.08**
Cluster 1	**0.30**	**0.30**	**0.25**
Cluster 2	**0.15**	**0.60**	**0.67**

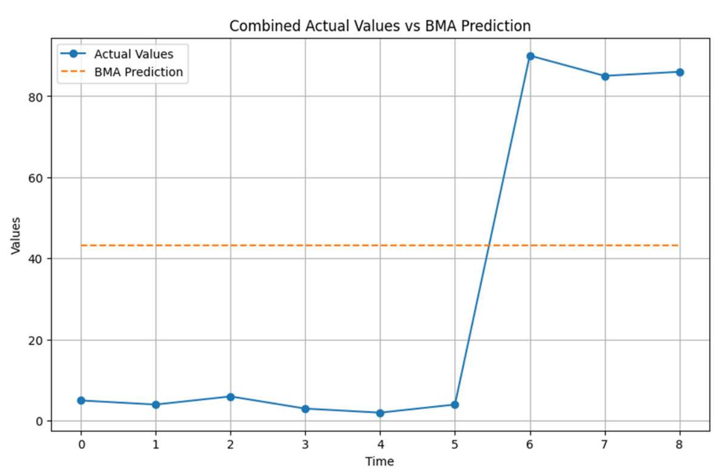
Step 3: Aggregate weights Across Criteria.

C0: (0.55 + 0.10 + 0.08)/3 = 0.24

C1: (0.30 + 0.30 + 0.25)/3 = 0.28

C2: (0.15 + 0.60 + 0.67)/3 = 0.47

Step 7: applying BMA for clusters using the weights from AHP .



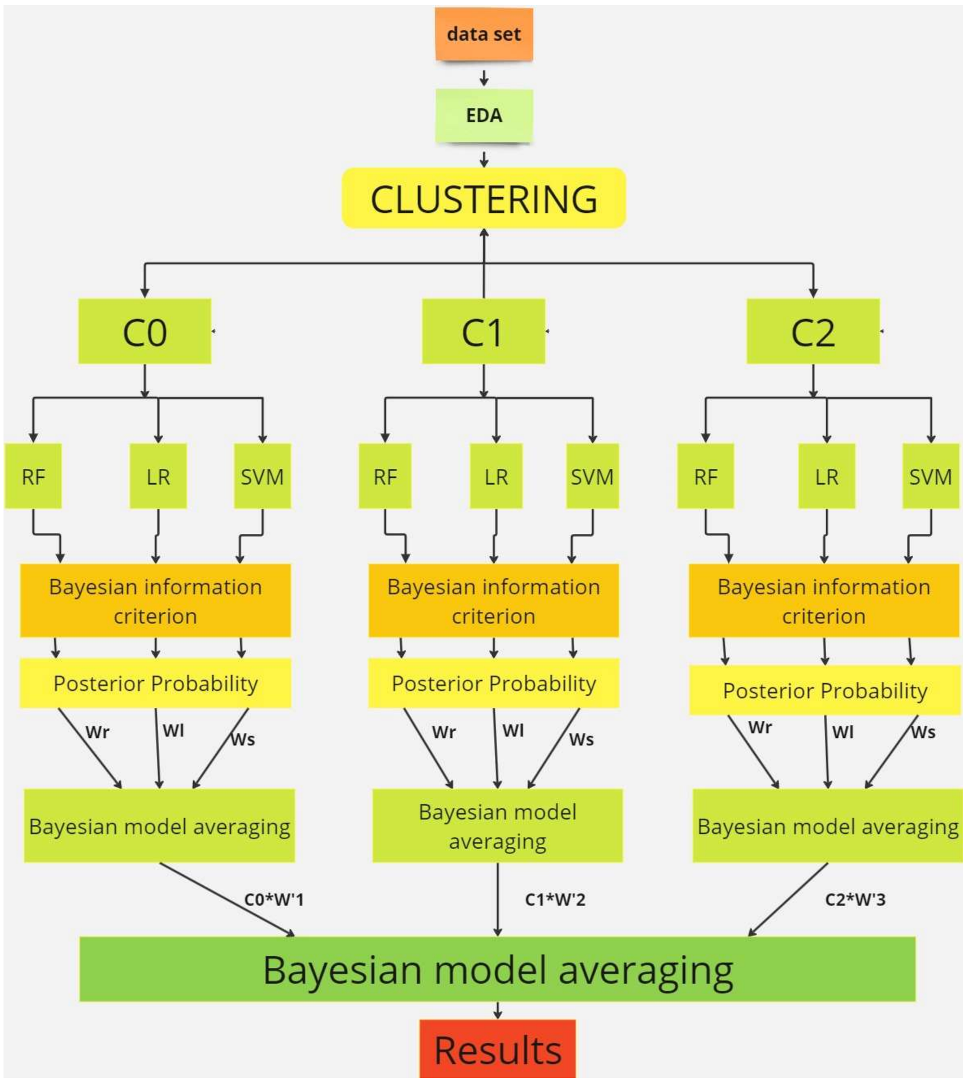
RMSE vale for combined clusters is 0.45

RMSE value of each cluster is C0 = 0.8, C1 = 0.6, C2 = 0.2

RMSE value must be near to zero.

We can observe that the performance is not upto the mark because we have taken assumed value for AHP but we can get lower values with real data and by optimising the code but we still see slight improvement in combined cluster.

Block diagram of model architecture:



key contributions and novelty of the project " fashion and beauty industry demand forecasting using BMA by ensemble learning approach " are:

Novel Integration of Methodologies: The project combines cluster-based modeling with Bayesian Model Averaging (BMA) and employs the Analytic Hierarchy Process (AHP) for an innovative ensemble learning framework. This integration tailors the forecasting process to accommodate the specific dynamics and segmentation of the fashion and beauty market.

Enhancing Decision-Making through Systematic Weighting the use of AHP to systematically weight clusters based on strategic business criteria is a novel approach within demand forecasting frameworks. This method quantifies decision-making factors that are often subjective, such as the importance of specific market segments or customer demographics.

References

1. Bayesian Model Averaging: A statistical technique that improves forecasting by considering the uncertainty in the choice of predictive models. https://www.sciencedirect.com/topics/mathematics/bayesian-model-averaging#:~:text=BMA%20accounts%20for%20uncertainty%20in,uncertainty%20into%20conclusions%20about%20parameters.

2. Cluster-Based Modeling: A method that segments data into meaningful clusters based on product type, customer demographics, and revenue metrics, allowing for tailored strategies that align closely with the specific characteristics and dynamics of different market segments. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8757628/

3. Ensemble Learning: A machine learning approach that combines the predictions of multiple models to improve the accuracy and robustness of the forecasting process. https://www.sciencedirect.com/science/article/pii/S0968090X21003594

4. Supply Chain Management: The process of planning, coordinating, and controlling the flow of goods, services, and information from raw materials to end customers. https://d1wqtxts1xzle7.cloudfront.net/106852190/61a742276f3234.26425361-libre.pdf?1698043717=&response-content-disposition=inline%3B+filename%3DConcept_of_Supply_Chain_Management.pdf&Expires=1714819686&Signature=Pl5FscpBkRUzhBlzjFNl3BPOVku6zT3Bi90KqrThWVywOEaleuD1uY8wpaZNQWxAz5FxwfudFz3zs0KJtveee2XU99pjE8~d7Wa1cY~FT18Kp-F5zL3kdceQVtwA6VdNsWXNQzacagRkkIbH0m3VQWbDWizBc29z~8oK0OIJsCDmtPET779xyOAMIsYyivOqg7IbdFq480Xr2K~XIp343aAeRF32lmMz5yrMpR55cmjO86zjljKnSMcUHgMrMBPQ8iJhz8sufh2WMRoZLUtyM9fJd8Ltol5TcU4FI1JHrcpZ96Gxii3MXOE7ZYlVvsw32UG65HXbAO8LAWdPIvPnPw__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA
5. Fashion and Beauty Industry: A dynamic and ever-changing landscape where the ability to accurately forecast demand is crucial for optimizing supply chains and minimizing costs. https://link.springer.com/article/10.1007/s12063-014-0097-6