Forecasting Seasonal Footwear Demand Using Machine Learning

By Majd Kharfan & Vicky Chan, SCM 2018
Advisor: Tugba Efendigil

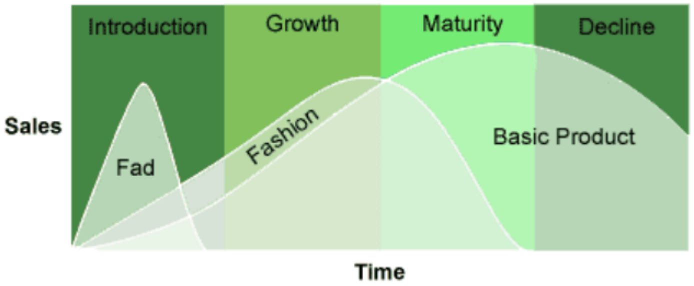
MIT Supply Chain
MANAGEMENT

Agenda

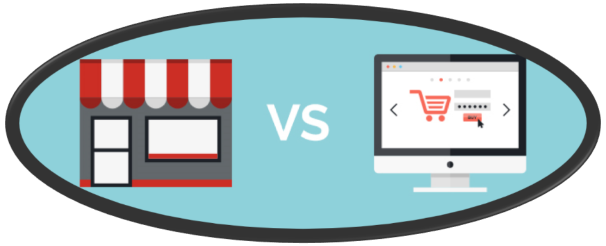
- The State Of Fashion Industry
- Research Objectives
- AI In the Fashion Industry
- Literature Review
- Methodology
- Results
- Conclusion

The State of Fashion Industry

Long Lead Times

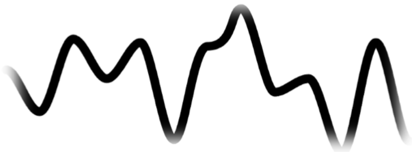


Short Product Lifecycle



System Shifts

Consumer Shifts



Volatility

Uncertainty

Research Objectives

How footwear and apparel companies can optimize their demand forecasting toward having an agile supply chain strategy that meets today's challenges?

1. *Leverage AI and machine learning technologies to recommend solutions that improve demand forecasting capabilities and prediction accuracy in the apparel and footwear industry*

2. *Maximize the utilization of POS data and help uncover new insights to be used in developing a demand forecasting framework that meets the today's strategic needs*

AI and the Fashion Industry

MIT Supply Chain
MANAGEMENT

"Many fashion executives regard AI as too mechanical to capture the creative core of fashion, and so are uncertain of what exactly it can do for them" [1]

Fashion industry lags behind other industries when it comes to AI

High forecast error on SKU level can be as high as 100% [2]

Large and diverse data sets

Advancement in ML algorithms and computing power

The Benefits of AI-enabled demand forecasting in retail: [3]

30% ~ 50%
Forecast Error Reduction

65%
Lost Sales Reduction

25% ~ 40%
Warehousing Cost Reduction

20% ~ 50%
Overall Inventory Reduction

1. *The Business of Fashion and Mckinsey & Company, The State of Fashion, 2017*
2. (Chase, 2009: 78)
3. Smartening up with artificial intelligence (AI): What's in it for Germany and its industrial sector? McKinsey & Company, 2017

Literature Review

MIT Supply Chain MANAGEMENT

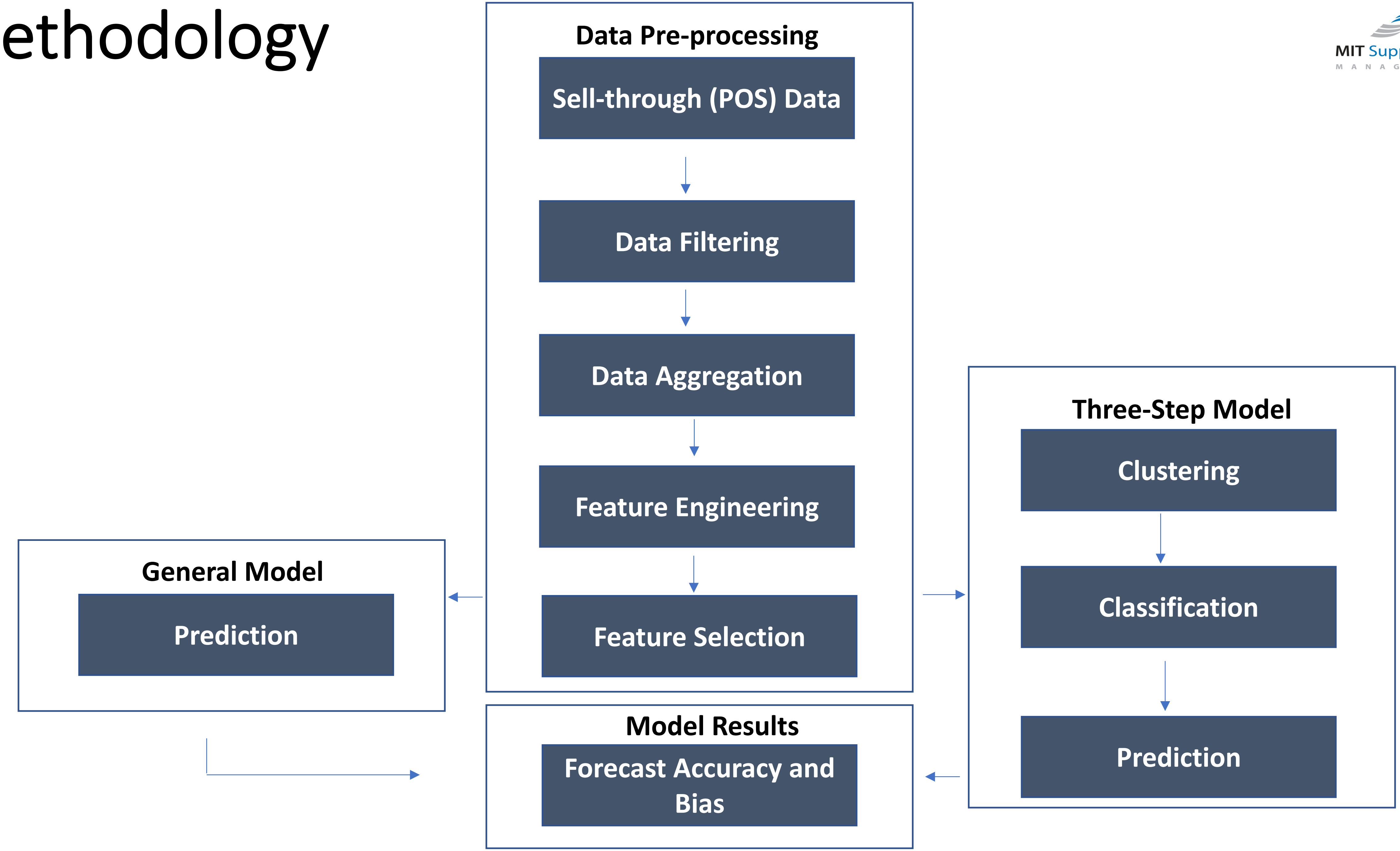
Types of demand forecasting techniques

	Traditional Approaches	Advanced Approaches
Techniques	Moving average, linear regression, Holt-Winters, exponential smoothing, ARIMA	Support vector machines (SVM), neural networks, decision trees, clustering, fuzzy inference system (FIS)
No. of predictor variables	Single or a few	Unlimited
Data source	Mainly demand history	Multiple
Data manipulation/ cleansing need	High	Low
Data requirements	Low	High
Technology requirements	Low	High

Lit. Review Findings & Our Contribution

- In general advanced or hybrid approaches perform better than traditional approaches
- Few studies on fashion industry
- Contradictory findings
- Identify the best mix of forecasting approaches for apparel and footwear companies
- A new forecasting approach for look-alike group of products

Methodology



Methodology - Data Pre-Processing

MIT Supply Chain MANAGEMENT

Scope and Granularity of Data

- **Dataset Collected:**
  - Sell-in (shipment)
  - Sell-through (POS) data from Jul 2013 – Mar 2018

- **POS data:**
  - Daily style-location
  - 115 retail outlet stores

- **Types of attributes:**
  - Product attributes
  - Calendar attributes
  - Store attributes
  - Price and promotion attributes

Data Filtering and Aggregation

- **Aggregated data:**
  - All stores
  - Monthly level

- **Filtered data:**
  - Outlet exclusive products
  - Full price status
  - Lifecycle of 1 – 4 months

Feature Selection and Engineering

- **Additional attributes**:
  - Lifecycle
  - Store count
  - Average sales

- **Feature selection**:
  - Recursive feature elimination
  - Decision trees

Dataset Partitioning

- Training set (67%)
- Validation set (25%)
- Test set (8%)

Methodology - General Model

MIT Supply Chain MANAGEMENT

Individual Models

- Regression Trees
- Random Forests
- *k*-Nearest Neighbors
- Neural Networks

Ensemble Models

- Average of the four individual models
- Median of the four individual models

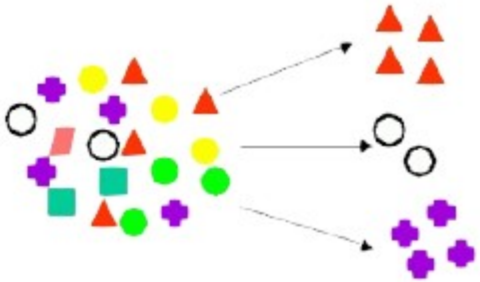
Methodology - Three-Step Model

MIT Supply Chain MANAGEMENT

Model Objective: Leverage POS data to identify look-alike group of products and use their average sales as a proxy to forecast the sales for brand-new products

Clustering

- t-distributed Stochastic
- Neighbor Embedding (t-SNE)
- *K*-Means



Classification

- SVM (Support Vector Machine)
- Regression Trees
- Random Forests

Prediction

- Regression Trees
- Random Forests
- Neural Networks
- K Nearest Neighbor
- Linear Regression
- Median + Average

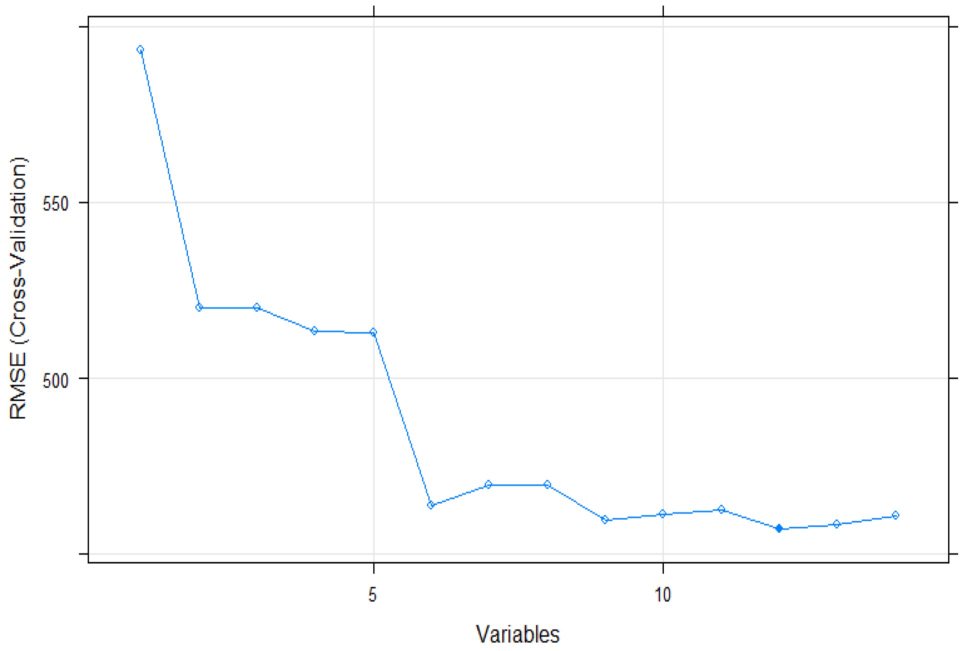
Results – Feature Selection

MIT Supply Chain MANAGEMENT

List of Attributes Selected for Model Building

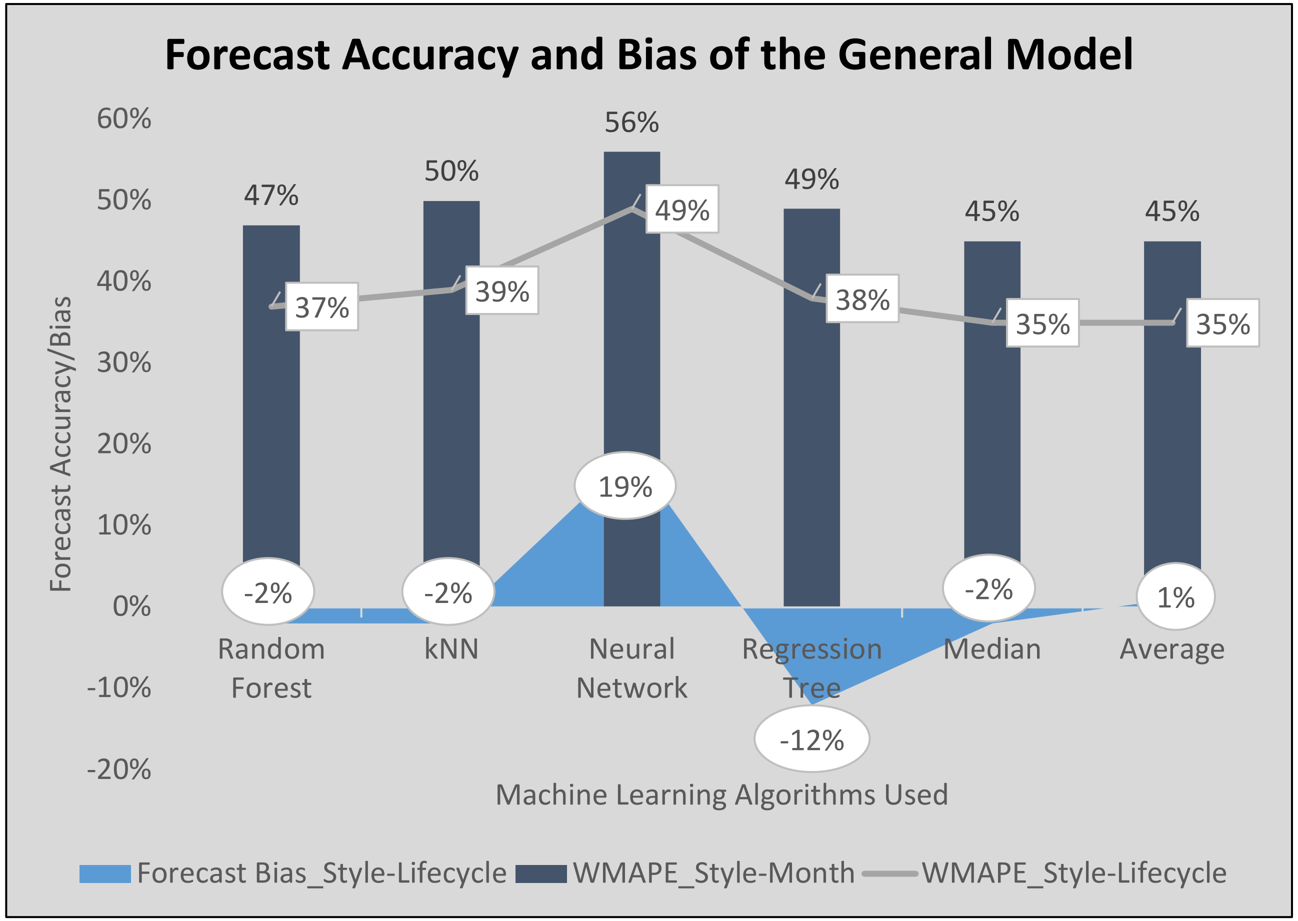
Importance Rank	Attribute	Attribute Category
1.	Store Count	Store
2.	Month	Calendar
3.	Lifecycle Month	Lifecycle
4.	Gender Desc	Product
5.	AUR	Price and Promotion
6.	Year	Calendar
7.	Basic Material	Product
8.	MSRP	Price and Promotion
9.	Color Group	Product
10.	Lifecycle	Lifecycle
11.	Cut Desc	Product
12.	Product Class Desc	Product

Cross Validation Error by Number of Attributes



- 12 out of the 14 predictor variables were and 2 variables (category and sub-category) were dropped.
- Store count, month and lifecycle month are the top 3 numerical attributes, while gender, material and color are the top 3 categorical attributes

Results- General Model



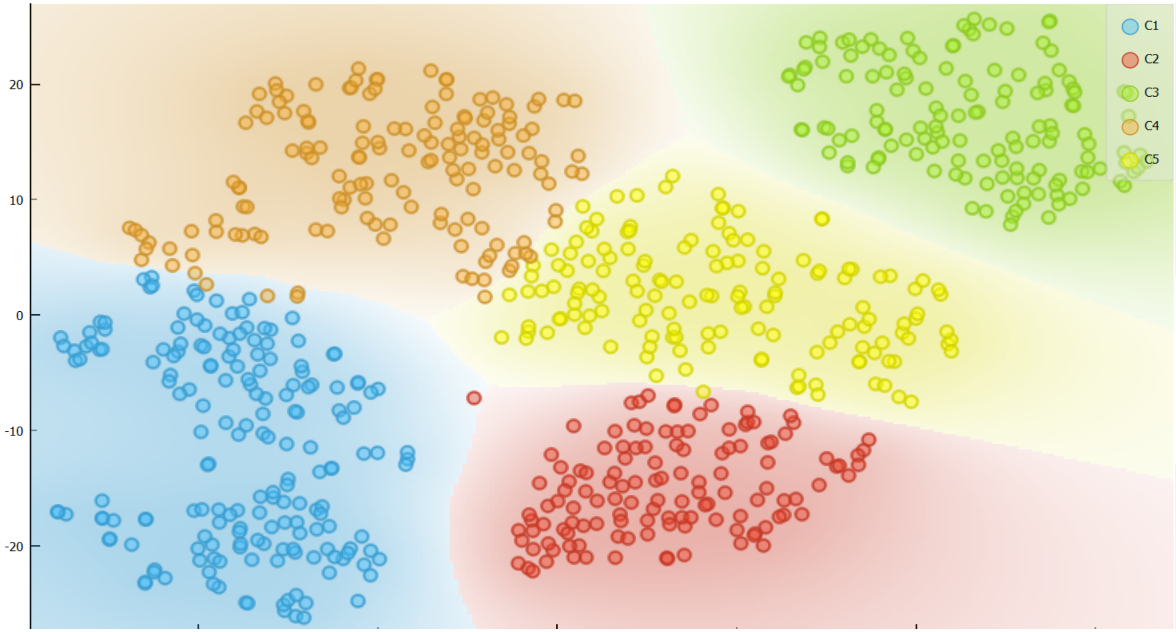
- **Individual Models:** Random forests gives the best predictive performance with the highest accuracy and lowest bias

- **Ensemble Models:** Median and Average yield similar results which are better than the individual models

- **Implication:** immediate implementation, outperforming the company's current forecasting model in terms of forecast accuracy and bias

Results - Clustering & Classification

Number of clusters with best classification match was 5



- Mono lifecycles
- Multiple lifecycles
- Variation by sales volume

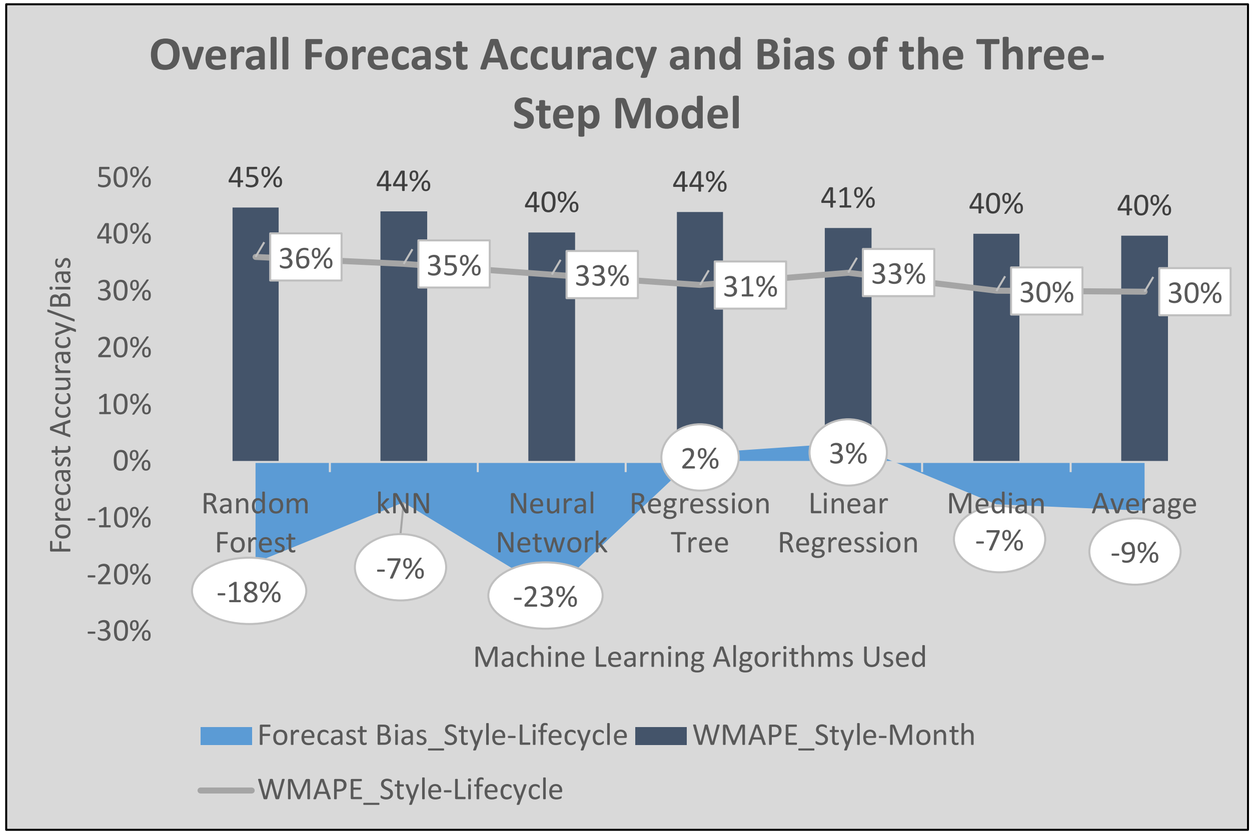
Confusion matrix for the overall classification accuracy

Predicted

Actual	C1	C2	C3	C4	C5	∑
C1	96.9 %	0.0 %	5.1 %	0.0 %	0.0 %	192
C2	0.0 %	90.5 %	0.0 %	8.4 %	0.0 %	78
C3	1.0 %	0.0 %	85.9 %	1.5 %	0.9 %	72
C4	2.1 %	9.5 %	9.0 %	90.1 %	3.5 %	140
C5	0.0 %	0.0 %	0.0 %	0.0 %	95.6 %	108
∑	194	74	78	131	113	590

- Overall matching accuracy: 93%
- Best performing algorithm: SVM

Results - Three-Step Model



Prediction results on a cluster-level:

Cluster Characteristics	Best Performing Algorithm	WMAPE (Style-Lifecycle)	WMPE (Style-Lifecycle)
Mono Lifecycle High Average Sales	K-NN Linear Regression	28%	+4% -11%
Mono Lifecycle Medium Average Sales	Random Forests	32%~37%	-11%~+6%
Multiple Lifecycle Low Average Sales	Regression Trees	39%~45%	-30%~0%

- **Ensemble Models**: highest forecast accuracy (30%) and low forecast bias (<10%)

- **Individual Models**: regression trees and linear regression, high forecast accuracy (>35%) with lowest bias (<5%)

Implication:

- Forecasting can be customized to deliver best possible results based on product characteristics

Conclusion

MIT Supply Chain Management

Key Insights

- Improved forecast accuracy
- Visibility into demand underlying factors and significance
- Make value out of categorical variables
- Forecast customization
- New approach to identify look-alike products

Future Opportunities

- Lost Sales
- Intended vs. Actual Lifecycle
- Higher Granular Data; Store and Weekly Level
- Price Optimization
- Size Curve Analysis

MIT Supply Chain MANAGEMENT

Q&A

Appendix

Efforts to Counter Current Industry Challenges

MIT Supply Chain Management

Adaptive Overarching Strategies

- Omni channel investments
- Brand experiments with direct-to-consumer
- Push the limits of time from design to shelf
- Proliferation of data

Agile Supply Chain Strategies

- Streamline manufacturing processes
- FG Inventory pooling and raw materials staging
- Digitize the supply chain for cost efficiencies
- Improve forecasting capabilities

Why Demand Forecasting?

MIT Supply Chain MANAGEMENT

*Agility: "The ability of an organisation to respond rapidly to changes in **demand** both in terms of volume and variety"*[4]

*Demand Forecasting is the art & science of predicting customer future **demand** for products.*

Why optimizing demand forecasting is one of the major initiatives for achieving agile supply chain?

- Input for planning across different supply chain and business functions (i.e., raw materials, sales, merchandising, etc.)
- Poor forecast results in:
  - Stock-outs i.e., lost revenue and consequently lost market share to competitors
  - Excessive inventory i.e., frozen net working capital and price mark-downs and both cause brand image deterioration

4. Martin Christopher, Industrial Marketing Management, 2000

Variables

MIT Supply Chain MANAGEMENT

List of Attributes from the Aggregated Data by Month at the Style Level

Variable category	Variable	Description
Meta Data	Style	Unique Style Number of Each Product
Meta Data	Style Description	Description of The Style
Calendar	Year	Fiscal Year
Calendar	Month	Fiscal Month
Product Attributes	Color	Color Code
Product Attributes	Basic Material	Type of Upper Material
Product Attributes	Gender	Gender or Age Group Description
Product Attributes	Category	Product Family
Product Attributes	Sub-category	Classic vs Modern
Product Attributes	Retail Outlet SubDept	Basic vs Seasonal
Product Attributes	Cut	Ankle Height
Product Attributes	Pillar	Product Sub-brand
Product Attributes	Product Class	Product Main Feature
Price and Promotion	Price Status	Full Price vs Mark-down
Price and Promotion	Manufacturer's Suggested Retail Price (MSRP)	Ticket price
Price and Promotion	Average Unit Retail (AUR)	Actual selling price
Sales Units	Retail Sales Units (Target variable)	Retail sales units

List of Attributes to be Considered for Feature Selection

Variable Category	Variable	Description
Meta Data	Style	Unique style number of each product
Meta Data	Style Description	Description of the style
Calendar	Year	Fiscal year
Calendar	Month	Fiscal month
Product Attributes	Color Group	Color code
Product Attributes	Basic Material	Type of material
Product Attributes	Gender	Gender or Age Group Description
Product Attributes	Category	Product Family
Product Attributes	Sub-category	Classic vs Modern
Product Attributes	Cut	Ankle Height
Product Attributes	Product Class	Product Main Feature
Price and Promotion	Manufacturer's Suggested Retail Price (MSRP)	Ticket price
Price and Promotion	Average Unit Retail (AUR)	Actual selling price
Lifecycle	Lifecycle	The total number of months in the lifecycle of a style
Lifecycle	Lifecycle Month	The number of months since product launch
Lifecycle	Lifecycle Start Month	The Month at which the Lifecycle has started
Store	Store Count	Number of Stores Selling a Style
Sales Units	Retail Sales Units (Target variable)	Retail sales units

Datasets

General Model:

Dataset	Months of sales	Number of Styles	Number of records
Training	36	578	1796
Validation	18	195	560

Three-Step Model:

Dataset	Months of sales	Number of Styles	Number of records
Training	35	539	1558
Validation	19	201	591
Testing	3	58	155