AUS | الجامعة الأميركية في الشارقة American University of Sharjah

Print Demand Forecasting with Machine Learning at HP Inc.

Dr. Harshvardhan

Assistant Professor, Information Systems & Business Analytics

School of Business Administration, American University of Sharjah

IIM Indore – Research Seminar · 7 July 2026

About Me

- **IIM Indore alumnus** – IPM (2016–21):
  - B.A. Foundations of Management + M.B.A.
- **Ph.D. in Business Analytics**
  - University of Tennessee, Knoxville, USA (2021–25).
- **Assistant Professor of Information Systems & Business Analytics**
  - School of Business Administration, American University of Sharjah, U.A.E.
- **Research Areas** – machine learning, demand forecasting, and applied analytics for decision-making.
- **www.harsh17.in · harshvardhan@aus.edu**

Related Publications

Where this complete work is documented

Print Demand Forecasting with Machine Learning at HP Inc.

M. Harshvardhan, Cara Curtland, Jerry Hwang, Chuck VanDam, Adam Ghozeil, Pedro Neto, Frederic Marie & Charles Liu – INFORMS Journal on Applied Analytics, 55(6), 469–483 (2025).

Enterprise-Scale Machine Learning for Demand Forecasting

M. Harshvardhan, Cara Curtland, Adam Ghozeil & Chuanren Liu – Foresight: The International Journal of Applied Forecasting, Issue 79 (2025). Top-5 finalist worldwide, IIF Best Forecasting in Practice.

From Data to Decisions: Machine Learning for Enterprise Demand Forecasting

M. Harshvardhan – Ph.D. Dissertation, University of Tennessee, Knoxville (2025).

Outline

1. Introduction
2. Literature Review
3. Forecasting Demand with Machine Learning
4. MLOps: Machine Learning Operations
5. Operationalization & Human-in-the-loop
6. Beyond Prediction: Predictive Optimization
7. Concluding Remarks

Introduction

Evolution of Demand Forecasting

Background and Challenges

Contributions

Evolution of Demand Forecasting

Intuition → Structure → Intelligence

- Ancient Examples of Demand Forecasting:
  - *Arthashastra* (350 BCE, India) blends qualitative and quantitative data for forecasting
  - *Han Dynasty* (110 BCE, China) combines multiple sources like harvest records, weather to maintain “Ever-Normal Granary”
- Formal Forecasting in Modern Era:
  - Ford and General Motors – first structured production planning systems in 1920s
  - Box-Jenkins ARIMA methodology in 1970
- Machine Learning (ML)-based Demand Forecasting:
  - BIG data: volume, variety and velocity
  - Powerful algorithms like decision trees and neural networks

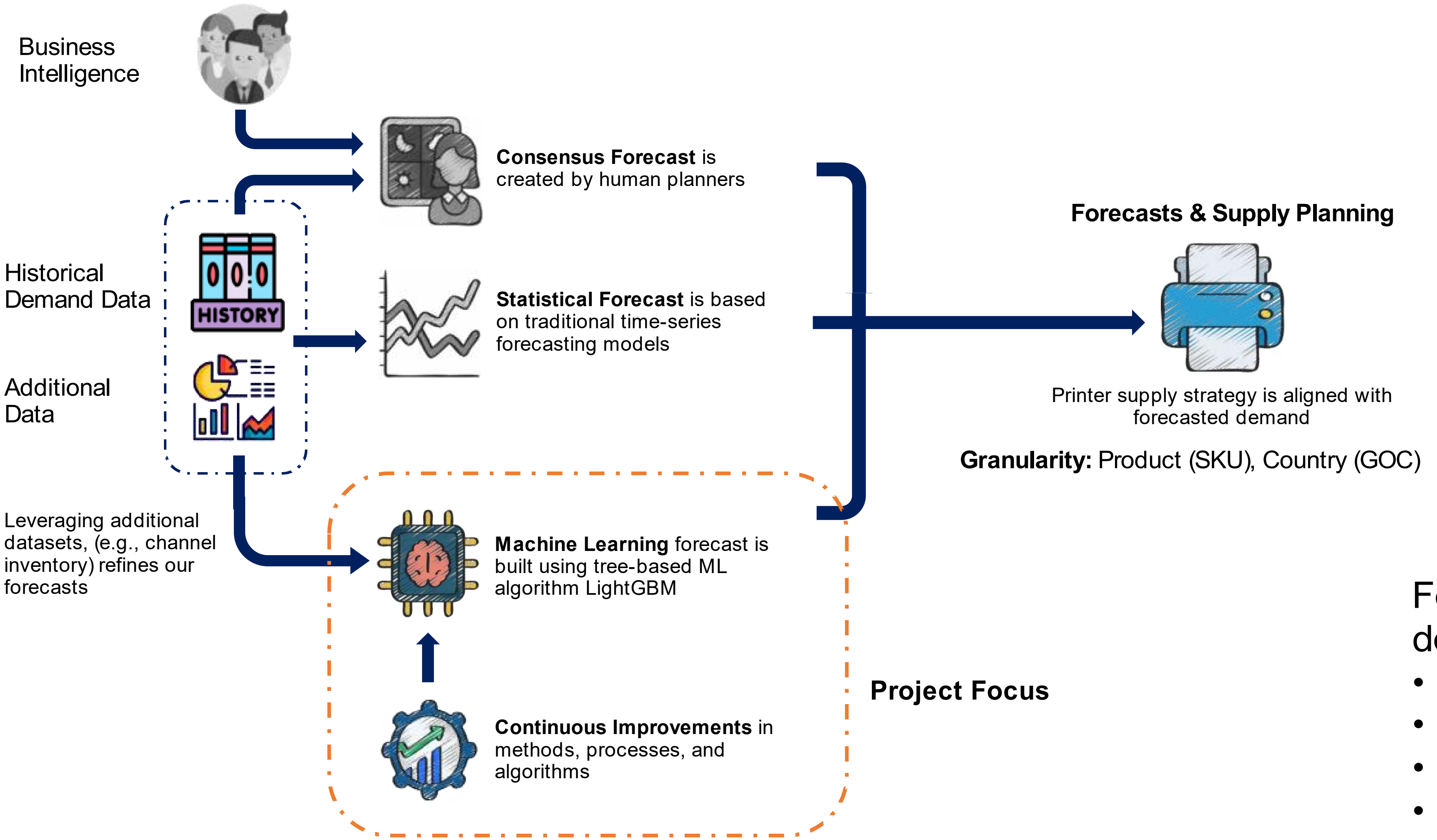
Background

About HP and Demand Forecasting at HP

- HP Inc. (HP, formerly Hewlett-Packard) is an American multinational information technology company
- Manufactures: personal computers (PCs), printers, and related supplies
- HP Print Division:
  - Products – 18,000 Stock Keeping Units (SKUs)
  - Market – 170 Countries
- Strategic Planning and Modelling Team (SPaM):
  - Internal consulting group at HP for supply chain planning

Print Demand Forecasting

Accurate forecasts increase product availability and profitability



Too much inventory leads to overage costs

Too little inventory leads to supply shortages

Forecast is used by multiple downstream processes:

- Inventory Sizing
- Purchasing
- Capacity sizing
- Contract negotiations
- Portfolio Planning

Research Challenges

Forecasting is hard due to scale and complexity of business

Business Complexity

- **Global market dynamics** vary across regions and product lines
- **Complex interplay** between economic conditions, seasonality, and local trends
- **Niche products** exhibit intermittent demand patterns

Technical Hurdles

- **High data variability** and inconsistency across markets
- **Automated preprocessing** and validation needed at scale
- **Model accuracy and generalization** capability

Organizational Transformation

- **Data-driven forecasting** from human judgement
- Trust in **ML-based predictions**
- **Change management** and process integration

Research Contributions

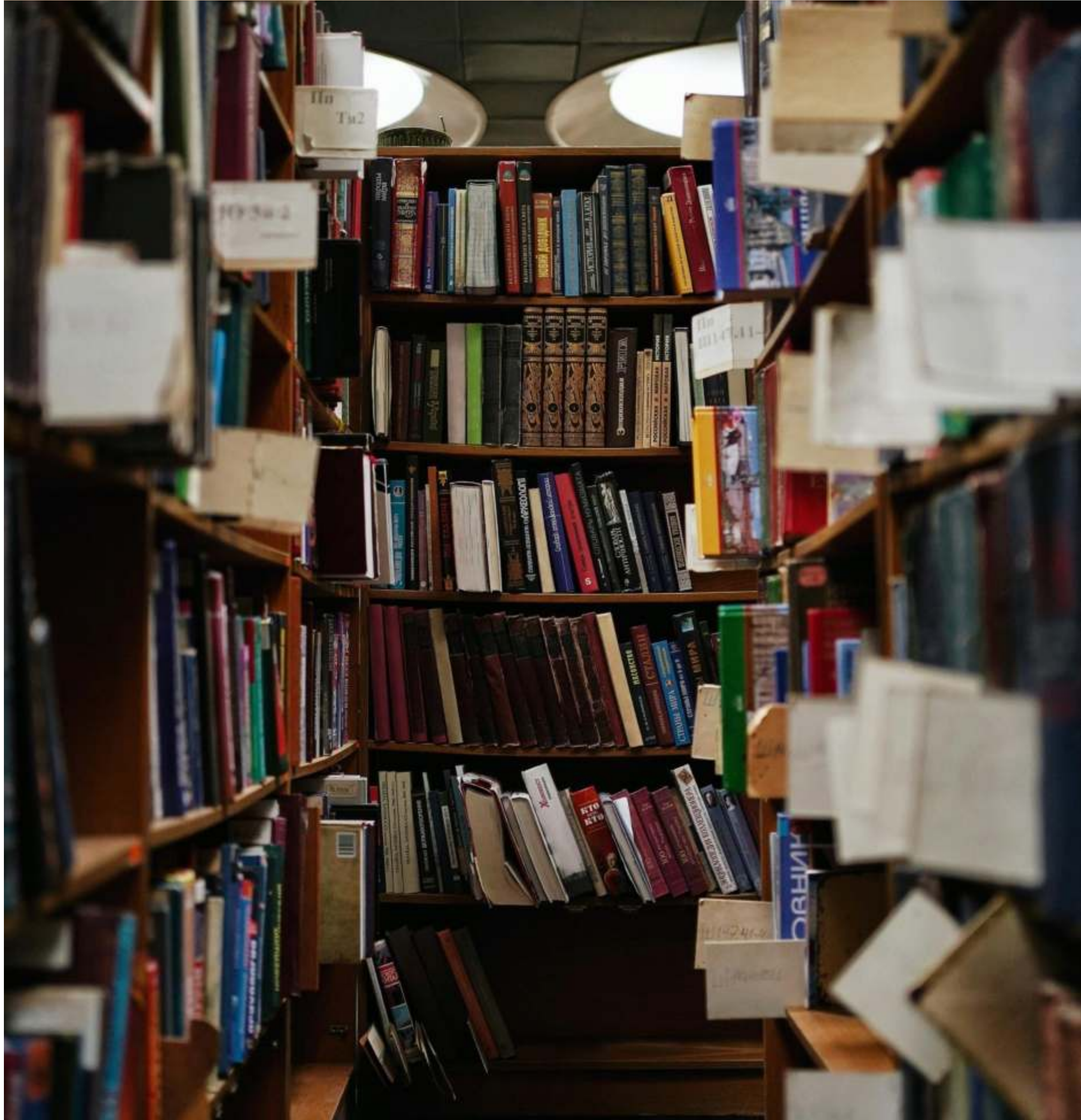
Key contributions to the academic and practitioner community

1. Scalable ML-based Forecasting Framework
2. MLOps for Enterprise Forecasting
3. Human-in-the-loop Architecture
4. Practical Contribution – Case Study
5. Proposed Predictive Optimization

Parts of this research have been published in

- *INFORMS Journal on Applied Analytics*
- *Foresight: The International Journal of Applied Forecasting*

Literature Review



Analytical Forecasting

Judgmental Forecasting

Human-in-the-loop Ensembling

ML-based Demand Forecasting

Machine Learning Operations (MLOps)

Analytical Forecasting Models

Traditional statistical methods of time-series forecasting

- **ARIMA** by Box and Jenkins (1970): Autoregressive (AR), Integrated/Differencing (I), Moving Average (MA)
- **Exponential Smoothing Model (ETS)** by Brown (1956) and Holt (1957)
- Severely limited due to their simplistic structures
  1. Linearity assumption
  2. Stationarity requirement
  3. Cannot handle exogenous variables
  4. Univariate models, cannot do multivariate forecasting
  5. Not scalable – hard to automate model selection and tuning
  6. Rigid structures – once defined by parameters, doesn't change
  7. No learning from similar series

Recently, ML methods got popular with more data and computation power!

Judgmental Forecasting Methods

Human decision making based on qualitative expert inputs

- Precede history of statistical models
- Most useful and tends to outperform analytical forecasts when (Hyndman and Athanasopoulos, 2018; Syntetos et al., 2016; Franses and Legerstee, 2011):
  - Historical data is limited (e.g. new product launch)
  - Issue is unprecedented (e.g. Australia forces plain packaging on cigarettes)
  - Human expert has ‘soft data’ (e.g. knowledge of upcoming promotions)
- However, judgmental forecasts are marred with cognitive biases (Hogarth and Makridakis, 1981; Lawrence et al., 2006):
  - Overconfidence
  - Illusion of control
  - Seeing patterns in randomness

Human-in-the-loop Ensembling

Machine Precision + Human Insight = Best Forecast?

- Lawrence et al. (1985, 1986) found that integrating judgmental inputs with statistical models yields improved forecast accuracy
- Blattberg and Hoch (1990) gave the golden ratio: 50/50 mixture of both
  - When humans are weak (strong), models are strong (weak)
- **However, most of these studies are behavioural lab experiments!**
- Integration strategies:
  i. Judgmental adjustments (Fildes and Petropoulos, 2015; Ibrahim et al., 2021)
  ii. Quantitative correction (Fildes, 1991)
  iii. Forecast combination (Blattberg and Hoch, 1990)
  iv. Input to analytical model (Sanders and Ritzman, 2004; Green and Armstrong, 2012)
- Effectiveness depends on forecaster's expertise and level of contextual information (Arvan et al., 2019)

Machine Learning-based Demand Forecasting

Advanced Analytics for Demand Forecasting

Makridakis (M-series) competitions has been key test-bed for algorithms

Direct forecasting method yields lower MSE than iterative but superiority in practice isn't guaranteed (McElroy, 2015)

Not enough evidence on use of **external features** from customers

Table 2.1: Summary of related research papers on ML-based demand forecasting.

Reference	Input	Model	Evaluation Metric
Dodin et al. (2023)	Lagged demands, demand statistics, seasonality components, region and month index, average age of shipped products	Improved LightGBM, Elastic Net	RMSSE
Qi et al. (2023)	Lagged demand, inventory	End-to-end Model (Dynamic Programming, RNN, MLP)	Stockout rate, turnover rate, total inventory management, holding, and stockout costs
Deng et al. (2023)	Lagged demand, inventory, among others	DeepAR, N-BEATS, Prophet	WMAPE
Makridakis et al. (2018)	M-3 data	MLP, BNN, RBF, GRNN, KNN, CART, SVR, GP, RNN, LSTM, SES, ETS	sMAPE, MASE
Sagaert et al. (2018)	Lagged demand, macroeconomic indicators	LASSO Regression	MAPE
Hamzaçebi et al. (2009)	Lagged demand	Artificial Neural Networks (ANN)	SAE, SSE
Marcellino et al. (2006)	Lagged demand	Linear models	MSFE
Gardner (1990)	Lagged demand	Exponential-smoothing Model (ETS)	Investment and Delay Time

Forecasting Demand with Machine Learning

Problem Formulation

Iterative Forecasting Algorithm

Model Input Features

Performance Evaluation

Solution Overview

Enterprise-level Forecasting Pipeline from Data to Decisions

Prepare Data

Historical **Demand**

Product Hierarchy

Location Hierarchy

Channel Data

Product **Lifecycle**

Create Model

LightGBM: tree-based Learner

Dynamic features

Hyperparameters Tuning

Feature Engineering

Run Process

Train model to create predictions

Predict future demand

Back-test models for performance

Store data for future

Publish Results

Create and Share **Forecast Accuracy Dashboard**

Disaggregate demand for each SKU, every GOC

Publish forecasts into SAP planning tool

Problem Formulation

We describe our forecasting task as a supervised learning problem

- Our work addresses the problem of predicting demand for a product p in country c at time t
- $X_{t,c,p}$: Input features for period t, product p in country c
- $y_{t,c,p}$: Actual demand for period t, product p in country c
- Training Dataset
  - $\mathcal{D} = \{(X_{t,c,p}, y_{t+1,c,p}): \forall\ c, p, t_{first} \leq t \leq t_{now}\}$
  - where t_{first} is the first time period
- Training Loss Function:
  - $\ell(\not{f}|\mathcal{D}) = \sqrt{E_{x,y\in\mathcal{D}}(\not{f}(X) - y)^2}$ (Root Mean Squared Error, RMSE)

Problem Scale:

- 170+ Countries
- 18,000 SKUs

Iterative Forecasting Algorithm

LightGBM as predictive engine but adaptable to other models

Algorithm 1 Enhanced training and forecasting algorithm with LightGBM

1: **Preprocess the data:** Data cleaning and feature engineering.
2: **Determine optimal hyperparameters:** Use grid search or random search for the LightGBM model.
3: **Initialize forecast horizon** T (e.g., 7).
4: **for** t_α in $(t_{\text{first}} : t_{\text{now}})$ **do**
5: Create the training data:

$$D_\alpha = \{(X_{t,c,p}, y_{t,c,p}) : \forall c, p, t_{\text{first}} \leq t \leq t_\alpha\}$$

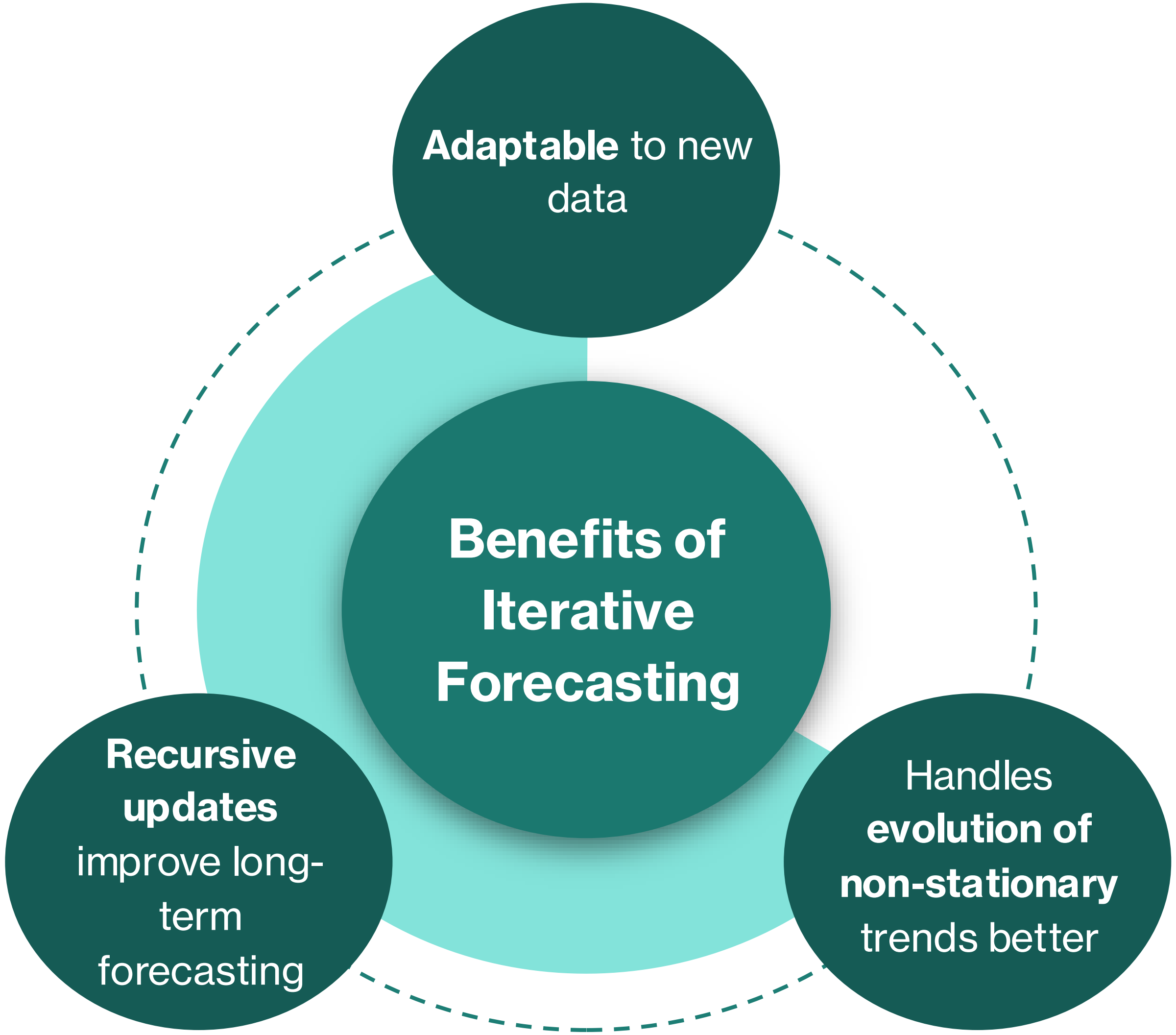
6: Perform time-series cross-validation on D_α and train the LightGBM model $f(\cdot)$ with optimal hyperparameters, minimizing loss (RMSE):

$$\ell(f|D) = \sqrt{\mathbb{E}_{X,y\in D}\,(f(X) - y)^2}$$

7: With the fitted model, create T forecasts for $t_\alpha + 1$ to $t_\alpha + T$:

$$F^T_{t_\alpha,c,p} = \left(f(\hat{X}_{t_\alpha+1,c,p}), f(\hat{X}_{t_\alpha+2,c,p}), \cdots, f(\hat{X}_{t_\alpha+T,c,p})\right)$$

8: Update the LightGBM model incrementally by warm starting from last month's best results if possible, or retrain it from scratch.
9: **end for**
10: **Perform Backtesting:** Apply the trained model to a historical dataset $D_{\text{historical}}$ to simulate past predictions. Evaluate its performance using appropriate metrics (e.g., RMSE, MAE).
11: **Store Forecasts:** Save the generated forecasts $F^T_{t_\alpha,c,p}$ to a dedicated database or file storage for future evaluation, comparison, or direct usage.
12: **Log Model:** Serialize the LightGBM model, hyperparameters, and performance metrics for future reference or retraining using MLFlow.



Model Input Features

Table 3.1: Summary of Forecasting Model Input Features and Their Utility

Feature Name	Description	Granularity	Utility for Forecasting
Lagged Demand	Size of demand from previous m months, m varies per product group	Month (t)	Captures influence of past trends on future demand
Rolling Demand Features	Statistics of demand within an n-month rolling window (mean, coefficient of variation, outliers)	Month (t)	Assesses recent trend and variability
Product-based Statistics	Mean and coefficient of variation of lagged demand and rolling features, per product category	SKU (p)	Identifies category-specific demand trends
Geography-based Statistics	Mean and coefficient of variation of lagged demand and rolling features, per country	Country (c)	Captures location-specific demand patterns
Seasonal Fluctuation	Binary indicator for each fiscal quarter and integer month within a quarter	Month (t)	Accounts for seasonal variations in demand
Product Life Cycle	Proportion of product life cycle left, calculated as $(M - m)/M$	SKU, Country (p, c)	Determines stage of the product in its life cycle
Channel Inventory	Inventory levels reported by distribution channel partners	SKU, Country, Month (p, c, t)	Indicates potential reordering needs
Sell-through	Sales to distribution channel partners	SKU, Country, Month (p, c, t)	Reflects downstream demand at distribution level

100+ Input Features

- 1 thru 15 month lagged demand for trends
- Product Life Cycle for growth and promotions
- Channel data provides early demand sensing

Feature Selection and Engineering Methods

- Fast AI Method (Howard 2019)
- QPFS (Rodriguez-Lujan et al., 2010)

Trade-offs

- Stability vs Guarantee → Explainability
- Speed of execution

Hyperparameter Tuning

- Hyperopt Tree-Structured Prazen Estimator (TPE)
- Hyperparameters:
  - Learning Rate
  - Tree Depth
  - Regularization Parameters
  - And more…

Forecast Accuracy Metrics

RMSE, Bias and wMAPE

Bias measures the weighted percentage error in forecasts, signified by a positive or negative value indicating over or underforecasting, respectively. Bias is calculated using the formula:

$$Bias = \sum_{i=1}^{n} \frac{y_i - \hat{y}_i}{y_i}.$$

wMAPE represents the weighted mean of absolute percentage errors, a metric easily understood even by non-technical stakeholders as percentage deviation from actuals. It is expressed as:

$$wMAPE = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{\sum_{i=1}^{n} y_i}.$$

RMSE (Root Mean Squared Error) is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}}.$$

Bias:

- ❑ Easiest to address
- ❑ Consistently positive = overforecasting
- ❑ Consistently negative = underforecasting

wMAPE:

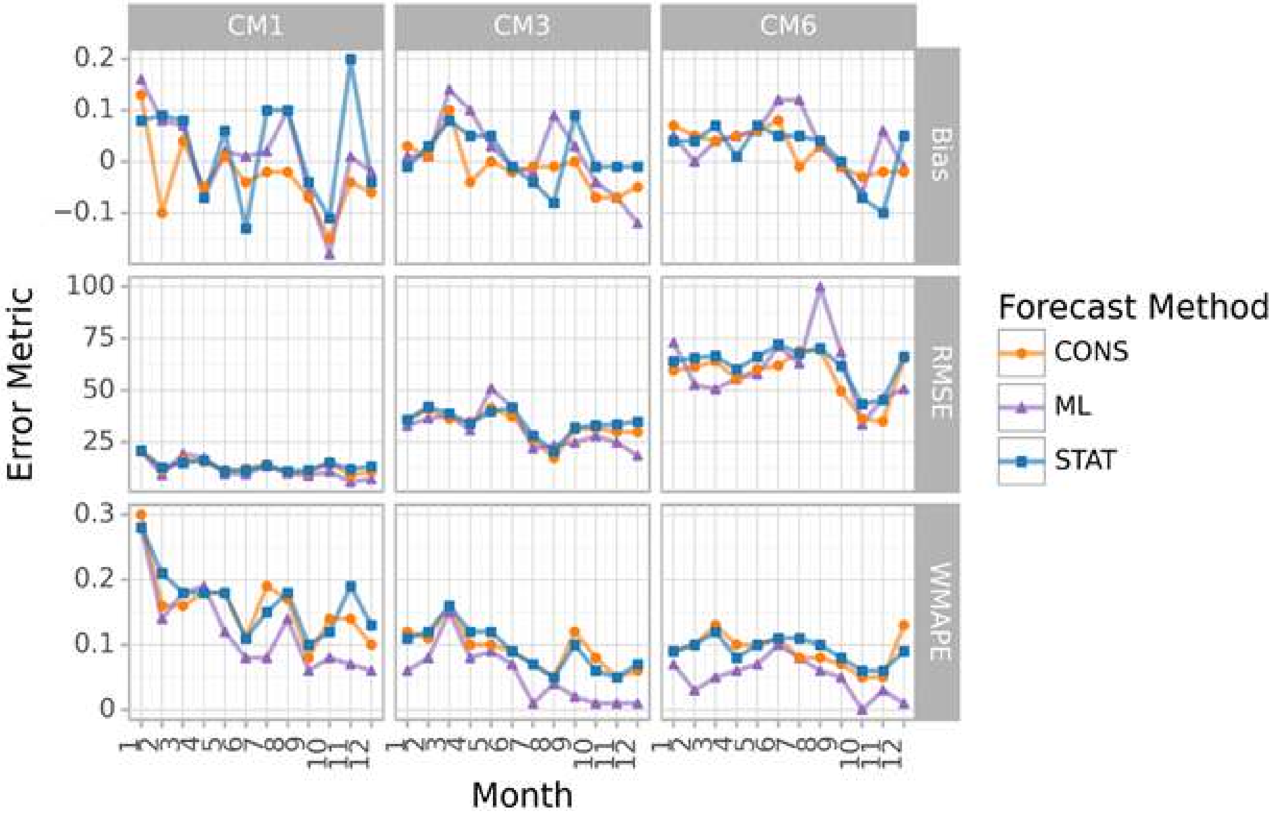
- ❑ Percentage error; easy to interpret
- ❑ Business KPI

RMSE:

- ❑ Used for ML model training
- ❑ Symmetric and always positive
- ❑ Continuously differentiable
- ❑ Sensitive to outliers

Performance Evaluation

ML Forecast Consistently Improves Business KPI



ML model wMAPE outperforms alternatives

Comparable performance in other metrics

Table 2.4: Comparative statistical analysis of forecasting accuracy metrics—namely Bias, wMAPE, and RMSE—is presented for consensus (CONS), machine learning (ML), and statistical (STAT) methods across cumulative forecast horizons (CM1, CM3, CM6). The table delineates t-statistics along with the corresponding p-values (enclosed in brackets) to facilitate a comprehensive evaluation.

Cumulative	Comparison	Bias	RMSE	WMAPE
CM1	CONS vs ML	-1.295 (0.209)	0.716 (0.482)	1.421 (0.169)
	STAT vs ML	0.385 (0.704)	1.128 (0.272)	1.832 (0.080)
CM3	CONS vs ML	-0.929 (0.363)	0.518 (0.610)	2.528 (0.019)
	STAT vs ML	-0.065 (0.949)	1.089 (0.288)	2.541 (0.019)
CM6	CONS vs ML	-0.701 (0.490)	-0.507 (0.617)	3.526 (0.002)
	STAT vs ML	-0.767 (0.451)	0.399 (0.694)	4.074 (0.001)

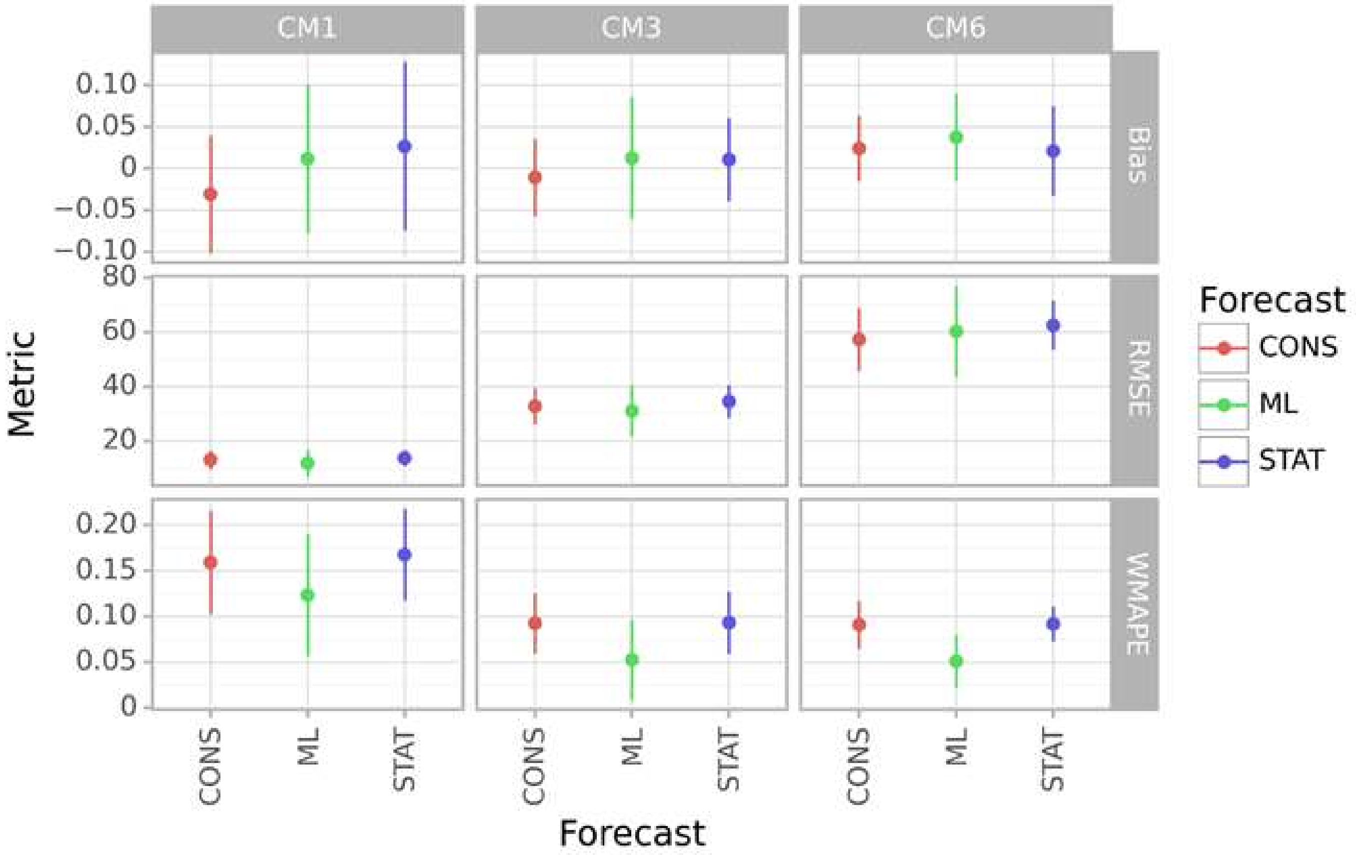
- Since CONS and STAT models are exposed to ML when making their forecasts, we are underselling ML's improvement
- ML has additional benefits: fewer human capital resources, less cognitive biases, automation is important

CM1 - immediate sales performance
CM3 - representative of the order to delivery cycle time for planning from global factories
CM6 - capacity planning and budgeting

Results for a select business segment of HP Print (1,484 products)

Comparing Performance between Methods

No model is universally better or worse; which is why human expertise is valued



Mean as dot and one standard deviation as whiskers.
Lower is better.

ML model has comparable accuracy (and sometimes better) vs. other methods

Table 2.3: Forecasting Accuracy Metrics. The table shows bias, RMSE, and wMAPE for cumulative forecast horizons (CM1, CM3, CM6) with Mean (Standard Deviation).

Model	CM1			CM3			CM6		
Metric	Bias	RMSE	wMAPE	Bias	RMSE	wMAPE	Bias	RMSE	wMAPE
Value (SD)	-3.08% (7.05%)	13.09 (3.38)	15.92% (5.62%)	-1.08% (4.68%)	32.76 (6.66)	9.25% (3.28%)	2.42% (3.96%)	57.29 (11.51)	9.08% (2.64%)
	1.17% (8.92%)	11.87 (4.87)	12.33% (6.69%)	1.25% (7.34%)	31.03 (9.43)	5.25% (4.39%)	3.75% (5.26%)	60.28 (16.83)	5.08% (2.91%)
	2.67% (10.14%)	13.71 (2.89)	16.75% (4.99%)	1.08% (5.00%)	34.55 (6.03)	9.33% (3.42%)	2.08% (5.38%)	62.47 (9.08)	9.17% (1.90%)

Results for a select business segment of HP Print (1,484 products)

MLOps: Machine Learning Operations

Scaling Machine Learning Models in Practice

Scaling Machine Learning Models

Execution of ML models in practice requires many additional machinery

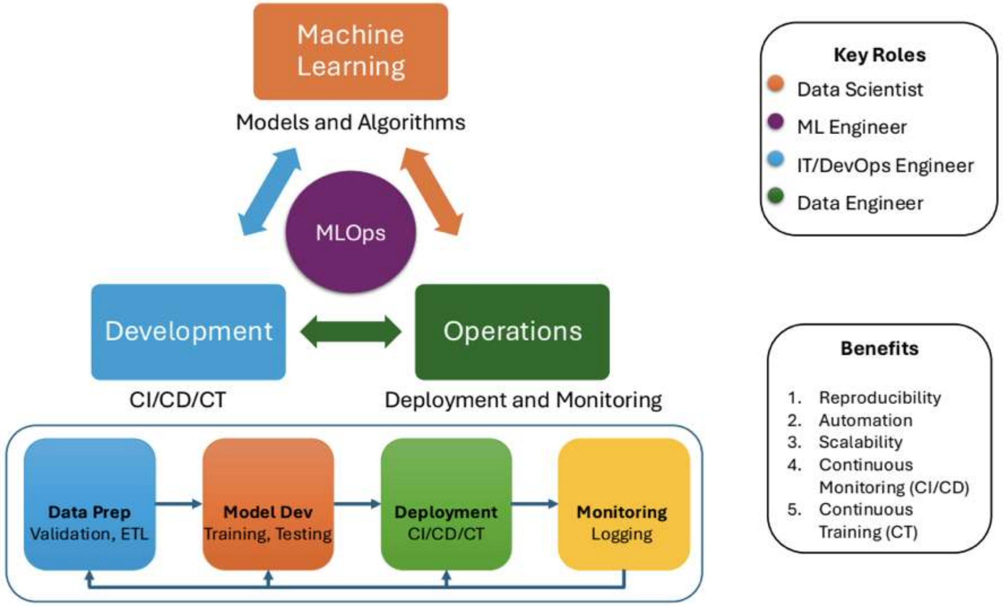


Figure 4.1: Figure of integrated MLOps framework showing the intersection of Machine Learning, Development, and Operations with a continuous workflow pipeline and key roles and benefits highlighted.

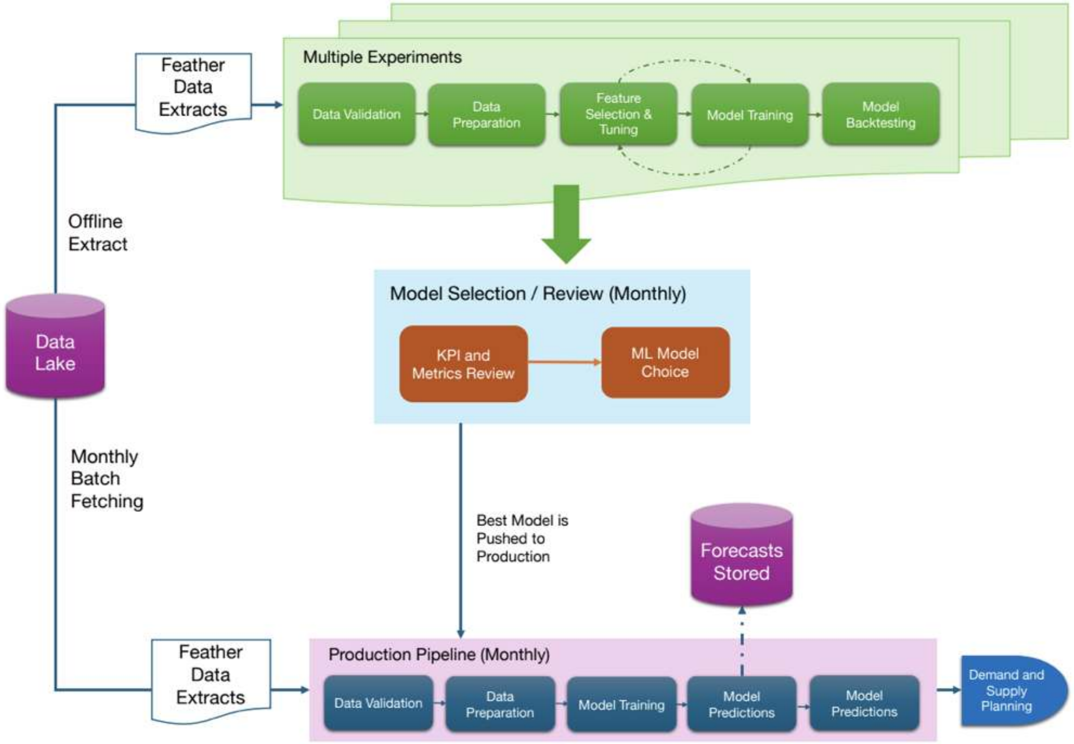


Figure 4.2: Project management for continuous deployment pipeline of our ML forecasting efforts.

Operationalizing Forecasting and Human-in-the-loop Framework

Implementation Journey

Human-in-the-loop Ensembling

Operational Benefits

Lessons for Practitioners

Implementation Journey

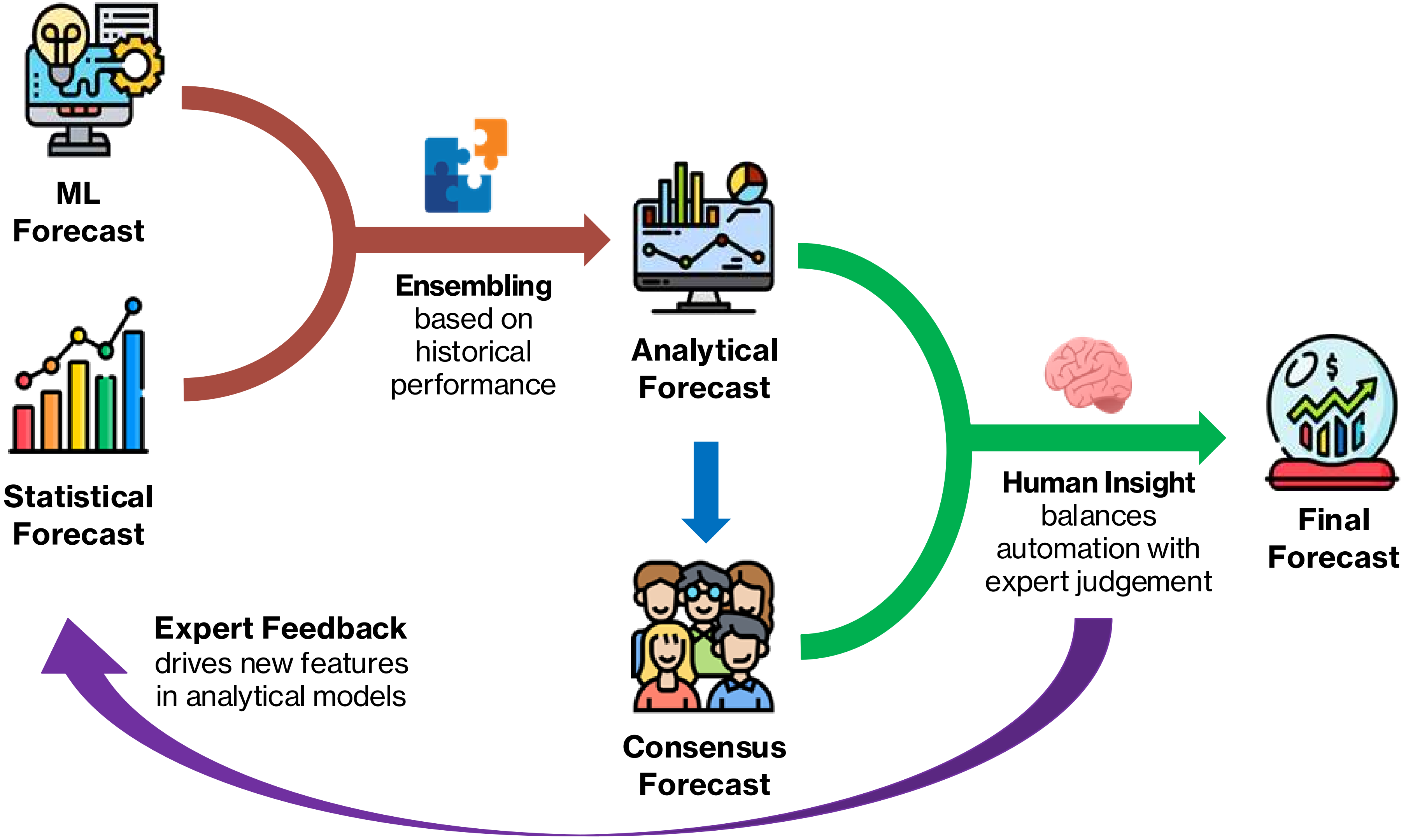
Business value from integration of ML forecasts into decision making

Business KPI Dashboard	ML Forecast Pilot	ML Forecast Adoption	SKU-level ML Forecast	Automated Ensembling
Single integrated KPI dashboard for the entire Print Business • Executive KPI alignment • Monthly review process • Forecast Value Add (FVA) informs decisions	ML visibility in dashboard led to **directional guidance** in forecast Bias reduction	**Manual ML Forecast use based on superior FVA** performance relative to Statistical Forecast	**Full integration into decision making pipeline** within Integrated Business Planning (SAP)	**Auto-Ensembling of ML and Statistical Forecasts** passed as Analytical Forecast to planners

Dashboard, ML Forecast, and Pipeline integration increasingly empower Human-in-loop Decisions

Human-in-the-loop Ensembling

Final forecast balances machine precision with human insight



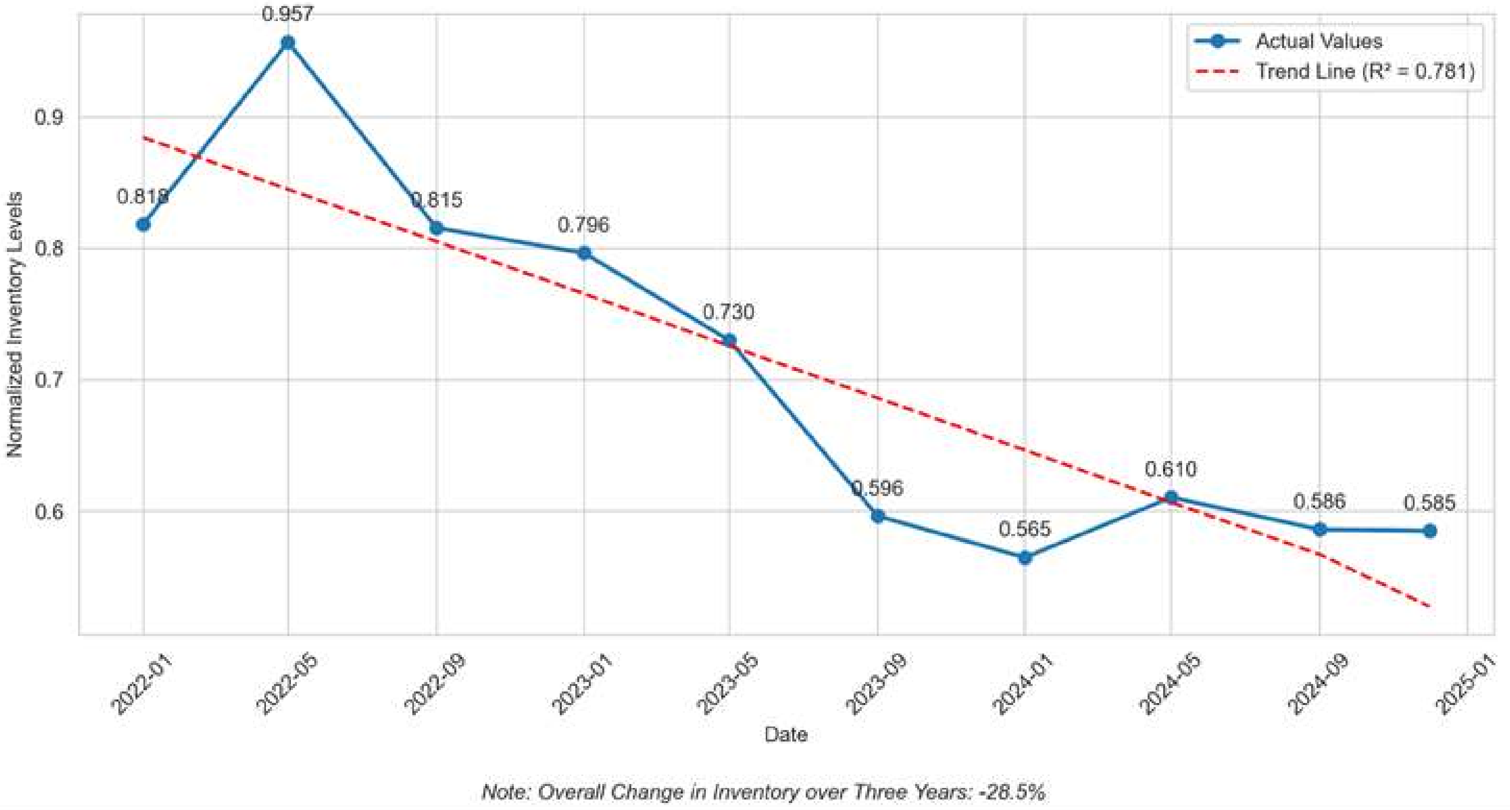
Forecast Value Add drives model choice

- Analytical
- Consensus
- Hybrid

Closed-loop feedback guides new feature development for ML and Statistical models

Inventory Reduction by 28% Over Three-year Period

Operational Benefits and Impacts



Sustained inventory decline demonstrates the effectiveness of our integrated approach

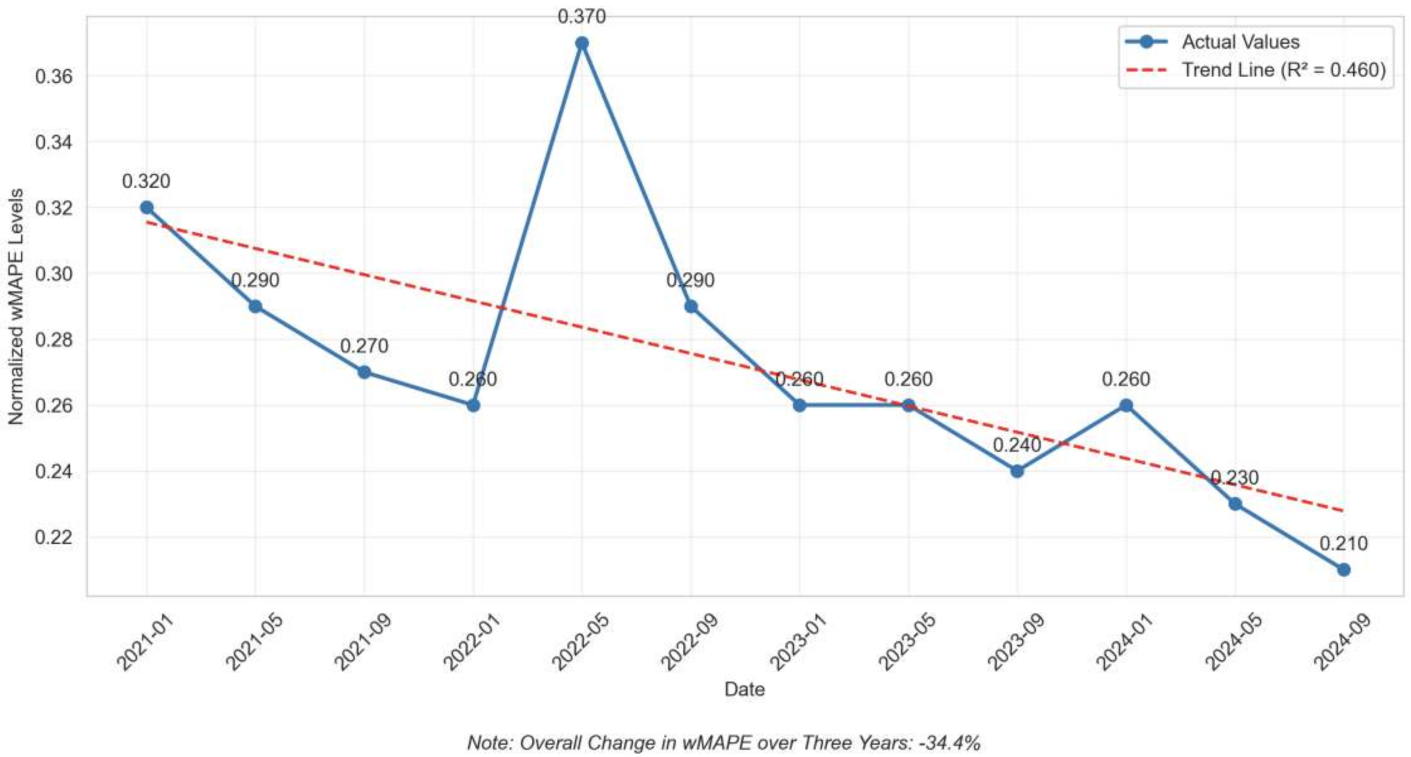
- ❑ 28.5% 3-yr reduction
- ❑ $R^2 = 0.781$
- ❑ Customer Service Levels maintained

Values are normalized relative to peak inventory level (1.0 represents the maximum monthly inventory across all periods) to protect sensitive data while preserving trend patterns

Results for a select business segment of HP Print (1,484 products)

wMAPE Reduction by 34% Over Three-year Period

As ML forecast gained acceptance, overall forecast accuracy reduced



(a) wMAPE KPI from 2022 to 2025 showing reduction by 34.4%.

Sustained final forecast wMAPE decline demonstrates the effectiveness of our integrated approach

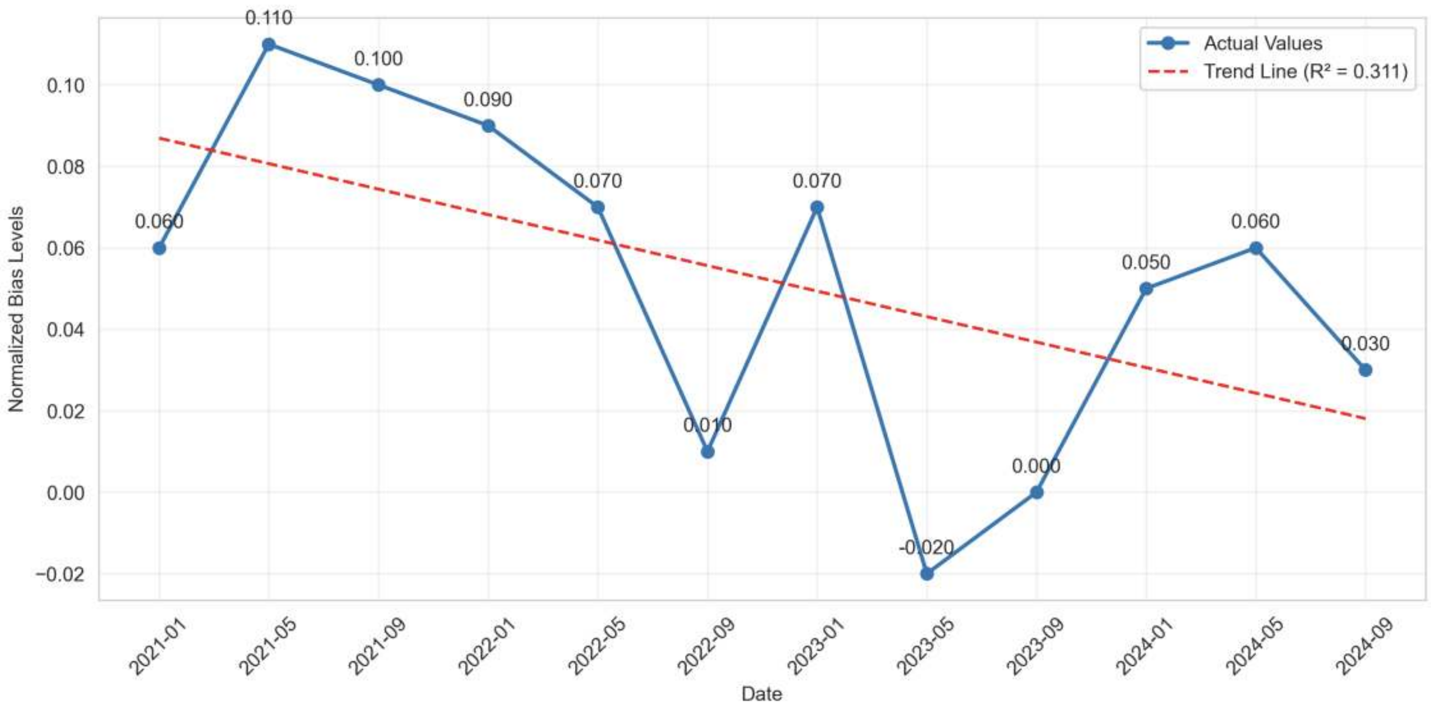
- ❑ 34.4% 3-yr reduction
- ❑ $R^2 = 0.46$

Values are normalized relative to peak wMAPE (1.0 represents the maximum wMAPE over all months) to protect sensitive data while preserving trend patterns

Results for a select business segment of HP Print (1,484 products)

Bias Reduction by 50% Over Three-year Period

As ML forecast gained acceptance, overall forecast accuracy reduced



Note: Overall Change in Bias over Three Years: -50.0%

(b) Bias KPI from 2022 to 2025 showing reduction by 50%.

Sustained final forecast Bias decline demonstrates the effectiveness of our integrated approach

- ❑ 50% 3-yr reduction
- ❑ R^2 = 0.33

Values are normalized relative to peak Bias (1.0 represents the maximum Bias over all months) to protect sensitive data while preserving trend patterns

Results for a select business segment of HP Print (1,484 products)

Managerial Insights

Trust, not accuracy, was the binding constraint on deployment

Visual Dashboards

- Interactive dashboards where planners can visualize and adjust decisions
- Include all products and geographies

Analytics as Guidance

- Goodwin et al. (2023) suggest analytics for baseline forecasts
- Planners retain control over decisions

Integration into Planning Tools

- Forecasts should be “importable” into IBP
- Else, they’re only for directional guidance

Planners should be included in model building

Scalability is important from start

Accuracy is not the only metric to chase

Planner's trust can be gained through “secondary analytics”

Secondary Analytics: Visual Dashboards, Data Engineering, etc.

Concluding Remarks

Key Takeaways

Future Work

Key Takeaways and Contributions

Analytics that think ahead, not just look back

- Developed **enterprise-scale** ML-based demand forecasting pipeline
- Achieved **34.4% improvement** in forecast accuracy (wMAPE) and **28.5% inventory reduction** while maintaining customer service levels over three years*
- Designed robust **MLOps Stack** for project management of ML in production
- **Human-in-the-loop Ensemble** → Practical adoption and trust by planners
- Introduced Predictive Optimization framework: forecasts aligned with decisions
- Case study of forecasting project implementation for 18,000 products sold in 170+ countries

*Results for a select business segment of HP Print (1,484 products)

Final Insight + Future Work

From Forecasting to Actionable Intelligence

🔮 What This Work Shows

- ML forecasting is not just technical – it's socio-technical
- Forecast error ≠ decision error
- Human-machine synergy beats automation alone

🚀 Future Directions

- Tree-DL hybrid models, hierarchical ensembling for new models
- End-to-end predictive optimization in supply chains
  - "Decisions, not just predictions."

Good forecasts don't make good decisions – good systems do.

Thank you!

Any questions?

Dr. Harshvardhan · harshvardhan@aus.edu · www.harsh17.in

Assistant Professor, Information Systems & Business Analytics · American University of Sharjah, U.A.E.

With sincere thanks to IIM Indore for the kind invitation.

Scan QR code for slides or visit www.harsh17.in/iimi2026