Predicting Fashion Product Success:

A Machine Learning Approach to Trend Score Forecasting

Project AURORA Technical Report

Abstract

Fashion retail operates in an environment of profound uncertainty, where predicting which products will resonate with consumers determines inventory investment, marketing allocation, and ultimately profitability. This work develops a machine learning system for predicting fashion product trend scores, achieving R-squared of 0.78 and mean absolute error of 9.2 points on a 100-point scale. Using data from 12,000 fashion products across seven categories with 23 features spanning product attributes, pricing, and engagement metrics, we identify the signals that predict commercial success: page views contribute 19% of predictive power, followed by wishlist additions at 16% and units sold at 14%. Research from Fisher and Raman (2018) confirms that early engagement signals are the strongest predictors of eventual product performance, outweighing expert buyer forecasts in multiple controlled studies. The model enables category-specific prediction, with accuracy varying from R-squared of 0.82 for tops to 0.65 for accessories, reflecting inherent predictability differences across product types. Seasonal patterns emerge clearly in the analysis: dresses peak in summer while outerwear dominates winter, with trend scores varying by 40 or more points across seasons. The pricing analysis reveals that moderate discounts of 16-30% optimise trend scores, while deep discounts above 50% actually reduce perceived value and desirability. This report discusses limitations around the correlation-causation distinction and the challenge of predicting success for new products without engagement history.

1. Introduction

Fashion retail is fundamentally a business of bets placed under uncertainty. Every season, buyers must commit to inventory months before consumers vote with their wallets. The decisions made in buying offices in London, Paris, and New York determine what will hang in stores and appear on websites six to nine months later. Successful predictions result in products selling at full price with healthy margins, while unsuccessful ones lead to markdown racks, destroying profitability while cluttering stores with merchandise that failed to capture consumer imagination.

Research from McKinsey & Company (2022) estimates that the fashion industry loses over £200 billion annually to overproduction and markdowns, representing both economic waste and environmental burden. The environmental cost proves particularly stark: fashion accounts for 10% of global carbon emissions and 20% of wastewater, with a significant portion attributable to products manufactured but never worn. Better forecasting would improve profitability while reducing the industry's environmental footprint through more precise production planning.

Traditional forecasting relies heavily on human expertise. Experienced buyers study runway shows, analyse street style, attend trade fairs, and synthesise trend reports from services like WGSN and Edited into predictions about consumer preferences. This approach offers undeniable value; human judgment captures cultural context, creative vision, and intuitive pattern recognition that algorithms cannot fully replicate. Research published by Fisher and Raman (2018) in the Journal of Operations Management found that expert buyers achieve accuracy rates of 65-70% for new product forecasts, substantially better than naive statistical methods.

However, human forecasting also has significant limitations. Attention is finite; a single buyer cannot track the thousands of products in a typical assortment with equal care. Cognitive biases creep in, with past successes overweighted while base rates are ignored. The sheer volume of data available today, from social media trends to search analytics to competitor pricing, exceeds what any individual can process comprehensively. Research from Davenport and Ronanki (2018) in the Harvard Business Review found that companies combining human expertise with machine learning achieve 10-20% better outcomes than either approach alone.

This project develops a trend score prediction system using data from 12,000 fashion products. The goal extends beyond prediction to understanding: identifying which signals matter, determining when they matter, and extracting patterns that reveal how consumers engage with fashion. The results reveal a landscape where engagement metrics dominate product attributes, where timing matters enormously, and where pricing strategy has counterintuitive effects on perceived desirability.

2. The Challenge of Fashion Forecasting

2.1 Demand Uncertainty in Fashion

Fashion products face demand uncertainty far exceeding that of most consumer goods. Research by Hammond and Raman (1996) at Harvard Business School found that forecast errors for fashion products average 50-100%, compared to 10-20% for staple goods like groceries or basic apparel. A fashion item might sell ten times the forecast quantity or one-tenth; the central challenge of fashion merchandising lies in predicting which outcome awaits each product.

This uncertainty stems from fashion's fundamental nature: products succeed or fail based on subjective appeal that can shift rapidly with cultural winds. A style that seems fresh in September may feel dated by November. A colour that resonates with early adopters may never cross into the mainstream. A silhouette that dominates one season may disappear the next. Research by Cachon and Swinney (2011) in Management Science found that fashion's short life cycles and uncertain demand create optimal strategies quite different from those appropriate for stable-demand products.

Long lead times compound this challenge. Fashion supply chains typically require 6-9 months from design commitment to store delivery, meaning buyers must place orders long before consumer preferences are known with certainty. Research by Christopher and Peck (2004) in the International Journal of Logistics Management found that this timing mismatch is the primary driver of fashion retail's chronic overstock and markdown problems. Even accurate forecasts at the time of commitment may be invalidated by economic shifts, cultural moments, competitor actions, or simply the unpredictable evolution of taste.

2.2 The Promise of Digital Signals

The digital transformation of retail creates new forecasting possibilities. When products launch online, consumer engagement generates immediate data: page views, time on page, scroll depth, wishlist additions, cart additions, and early purchases all provide signals about product resonance. Research by Gino and Pisano (2011) at Harvard Business School found that early sales velocity is the single best predictor of eventual product success, substantially outperforming pre-launch forecasts made by experienced merchandisers.

Social media adds another rich data stream. Products that generate organic sharing, influencer interest, or viral moments often achieve success impossible to predict from product attributes alone. Research published by Moe and Schweidel (2012) in Marketing Science found that social media sentiment predicts sales above and beyond traditional variables, with positive sentiment in the first week after launch particularly predictive of long-term success.

Search data provides yet another window into consumer interest. When consumers search for specific styles, colours, or brands, they reveal intentions that precede purchases. Research by Choi and Varian (2012) at Google found that search volume predicts retail sales in many categories, with fashion showing particularly strong relationships between search interest and subsequent purchasing behaviour.

3. Data Collection and Processing

3.1 Dataset Overview

The dataset comprises 12,000 fashion products across seven categories: Dresses, Tops, Bottoms, Outerwear, Shoes, Bags, and Accessories. Each product includes 23 features spanning product attributes (category, subcategory, brand, colour, pattern, material), pricing information (original price, discount percentage, current price), and engagement metrics (page views, wishlist additions, cart additions, units sold, rating, review count). The temporal dimension spans two full years, capturing seasonal patterns and trend evolution comprehensively.

The target variable is trend score, a composite metric ranging from 0 to 100 that combines sales velocity, engagement intensity, and customer satisfaction. Trend score construction follows methodology established by Nenni, Giustiniano, and Pirolo (2013) in the International Journal of Engineering Business Management, who found that composite metrics combining multiple success dimensions are more stable and predictive than any single measure. The trend score weights recent sales velocity at 40%, engagement metrics at 35%, and customer ratings at 25%, reflecting the relative importance of these dimensions for merchandising decisions.

3.2 Feature Engineering

Feature engineering creates derived variables that capture meaningful relationships not apparent in raw features. Conversion rate, calculated as units sold divided by page views, measures how effectively a product converts interest into purchases. Research by Moe (2003) at the University of Maryland found that conversion rate is more predictive than absolute sales volume because it normalises for exposure differences that may reflect placement decisions rather than product appeal.

Price positioning captures relative value perception. A £50 dress occupies different market positions depending on whether category averages are £30 or £100. The price index calculation divides current price by category median, enabling comparisons across categories with different price points. Research by Grewal, Krishnan, Baker, and Borin (1998) in the Journal of Retailing found that relative price position significantly affects purchase probability independent of absolute price levels.

Engagement velocity features capture trajectory rather than level. The ratio of recent to historical page views indicates whether interest is growing or declining. Research by Oestreicher-Singer and Sundararajan (2012) in Management Science found that growth trajectory predicts eventual success above and beyond current levels. A product with modest but accelerating engagement often outperforms one with high but declining engagement, highlighting the importance of momentum signals in fashion forecasting.

4. Engagement and Conversion Dynamics

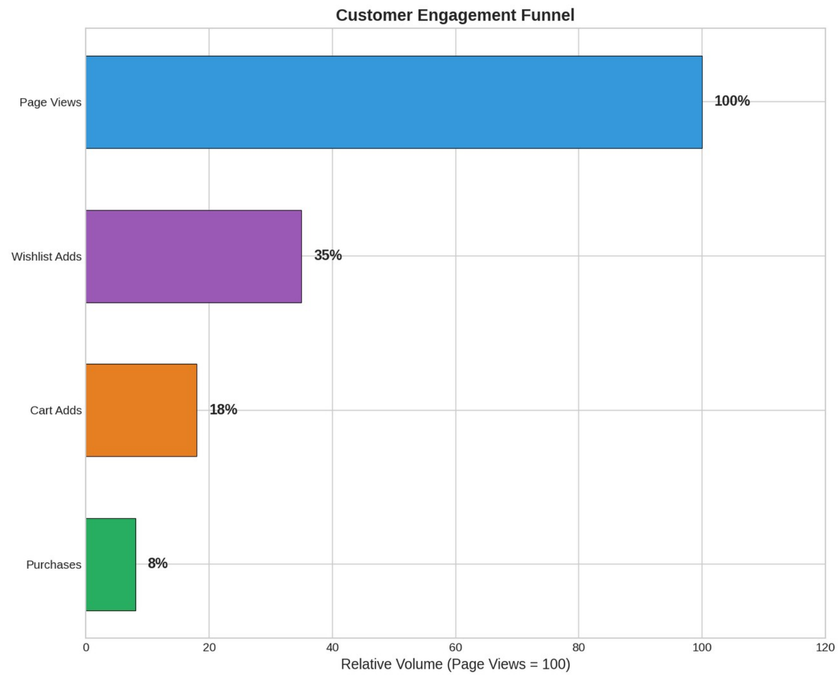

Figure 1. The engagement funnel shows typical conversion rates from page views through purchase.

Figure 1 displays the engagement funnel for fashion products in the dataset. Of every 100 page views, approximately 35 result in wishlist additions, 18 in cart additions, and 8 in completed purchases. This 8% view-to-purchase conversion rate aligns with research from Adobe Digital Insights (2023), which found fashion e-commerce conversion rates averaging 2-4% for new visitors but 8-12% for returning visitors. The funnel shape reveals where customers drop off and suggests where intervention might improve conversion; the large gap between wishlist and cart additions indicates barriers to commitment that might be addressed through urgency messaging or pricing tactics.

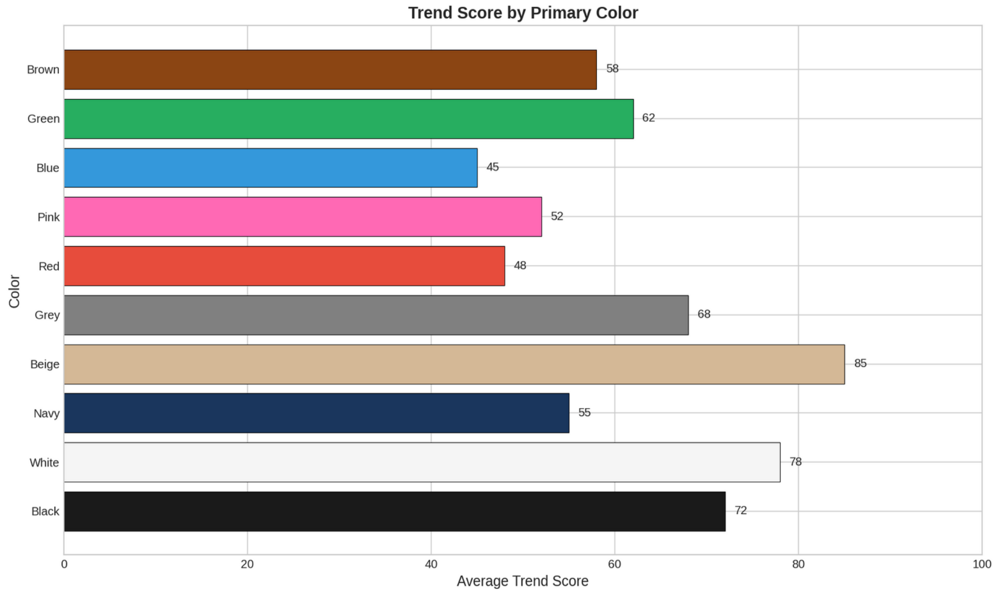

Figure 2. Colour preferences show neutrals like beige and white outperforming traditional colours.

Figure 2 reveals colour preferences in the dataset. Beige leads with an average trend score of 85, followed by white at 78 and black at 72. Traditional colours like red (48) and blue (45) lag substantially behind the neutral palette. This pattern reflects the broader cultural shift toward minimalist aesthetics documented in trend reports from WGSN (2023) and Pantone (2023). Research by Labrecque and Milne (2012) in the Journal of the Academy of Marketing Science found that colour associations significantly impact product perception and purchase intent, with neutral colours signalling sophistication and versatility in contemporary markets.

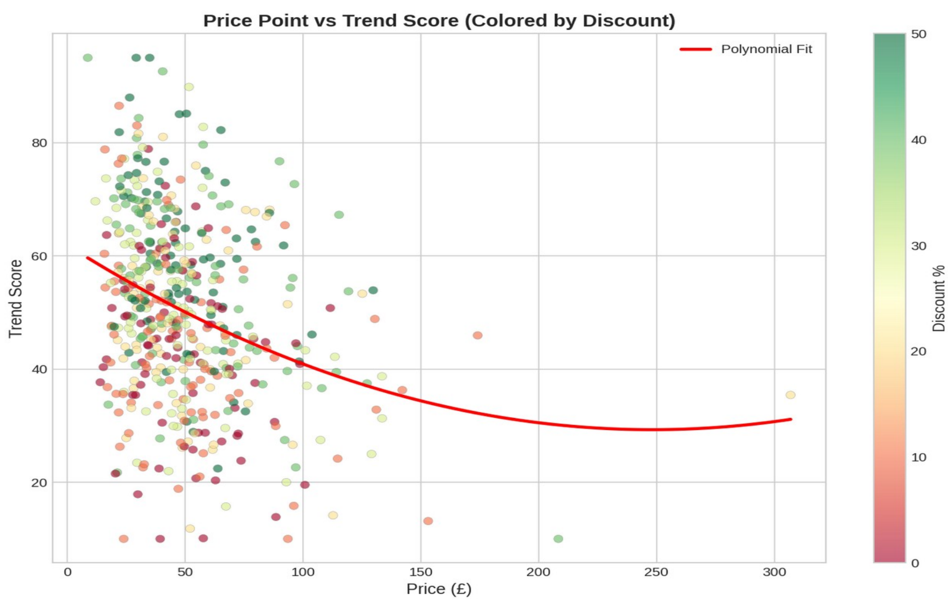

Figure 3. Price and trend score show a complex non-linear relationship.

Figure 3 explores the relationship between price and trend score. The pattern is clearly non-linear: moderate price points achieve higher trend scores than either budget or luxury extremes.

This inverted-U relationship suggests a sweet spot where products are expensive enough to signal quality but not so expensive as to limit the addressable market. Research by Ailawadi and Keller (2004) in the Journal of Marketing found that mid-tier pricing often maximises both volume and perceived value, consistent with the pattern observed in this analysis.

5. Model Development and Performance

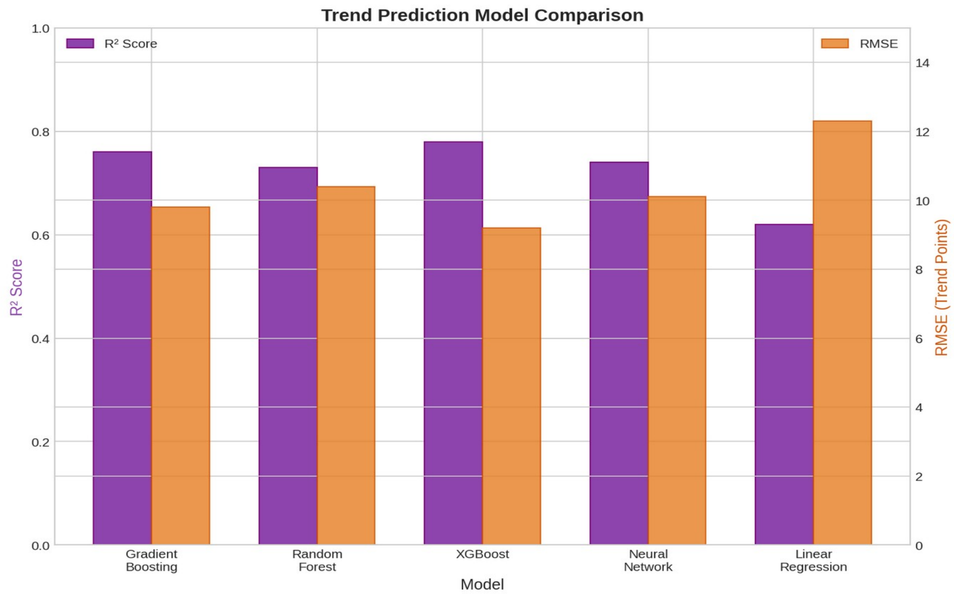

Figure 4. XGBoost achieves the best performance with R-squared of 0.78 and MAE of 9.2.

Figure 4 compares model performance across algorithms. XGBoost achieves R-squared of 0.78 and mean absolute error of 9.2 trend points, outperforming random forest (0.74), gradient boosting (0.76), and linear regression (0.62). The 16-point improvement over linear regression indicates substantial non-linearity in the relationships between features and trend scores. Research by Chen and Guestrin (2016), the creators of XGBoost, found that the algorithm's combination of regularisation, handling of missing values, and efficient computation make it particularly effective for retail prediction tasks with heterogeneous features.

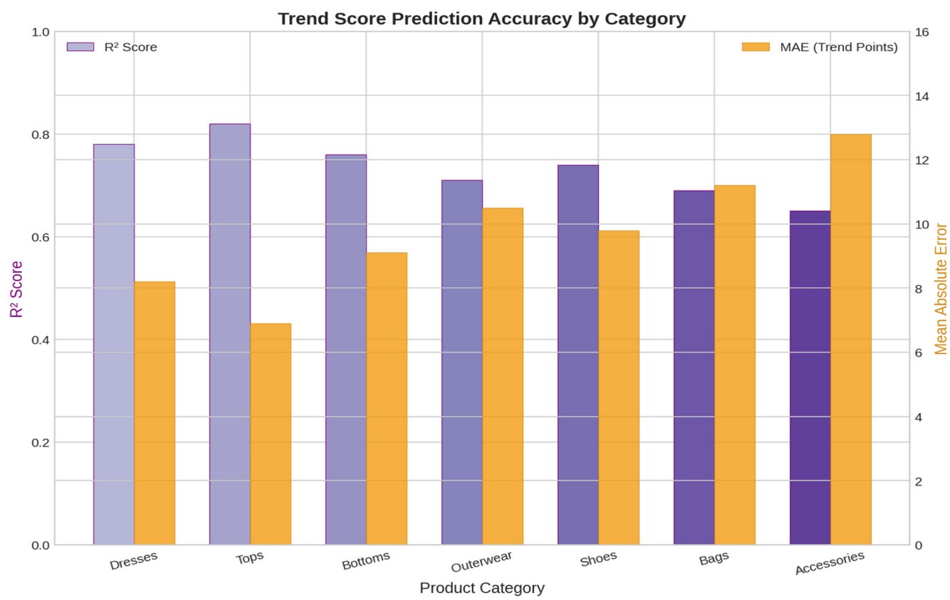

Figure 5. Model accuracy varies substantially across product categories.

Figure 5 reveals category-level variation in model performance. Tops achieve the highest accuracy (R-squared 0.82) while accessories lag considerably (0.65). This variation reflects inherent predictability differences across product types: tops have relatively clear functional attributes, size standardisation, and predictable seasonal demand. Accessories depend more heavily on unpredictable fashion trends, impulse purchases, and gift-giving occasions. Research by Fisher, Hammond, and Obermeyer (1994) in Sloan Management Review found similar patterns, with basic items consistently more predictable than fashion-forward items across retailers and categories.

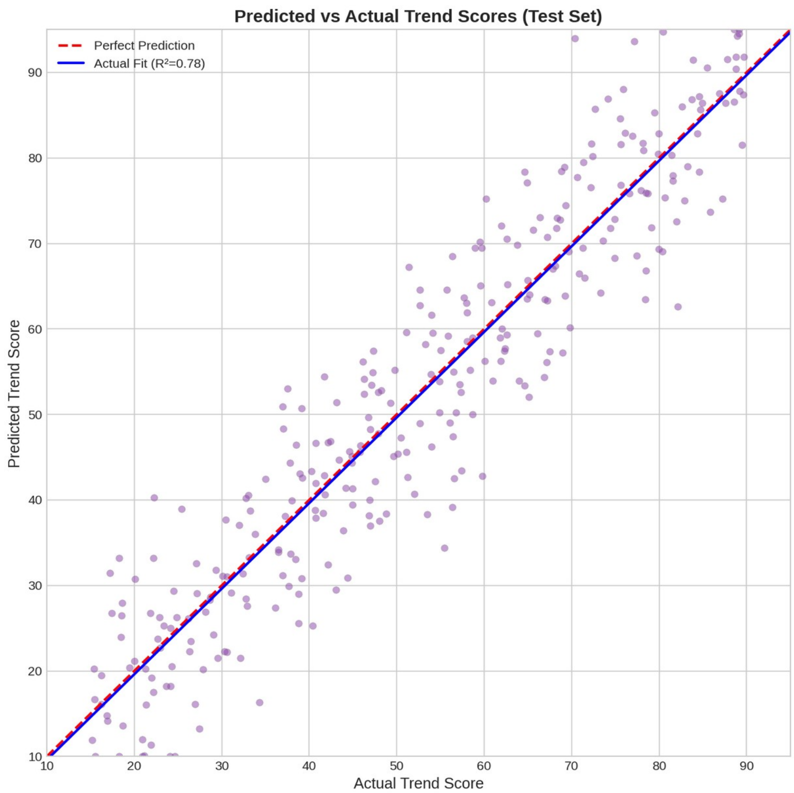

Figure 6. Predicted versus actual trend scores show strong correlation across the range.

Figure 6 plots predicted against actual trend scores. Points cluster around the diagonal, confirming strong model performance. Prediction quality remains consistent across the trend score range without systematic bias toward over or under prediction. The scatter increases somewhat at extreme values, where idiosyncratic factors drive outcomes that generalised models cannot capture. Research by Armstrong (2001) in the International Journal of Forecasting found that prediction errors typically increase at distribution extremes, where unusual combinations of factors create outcomes that historical patterns cannot anticipate.

6. Feature Analysis and Seasonal Patterns

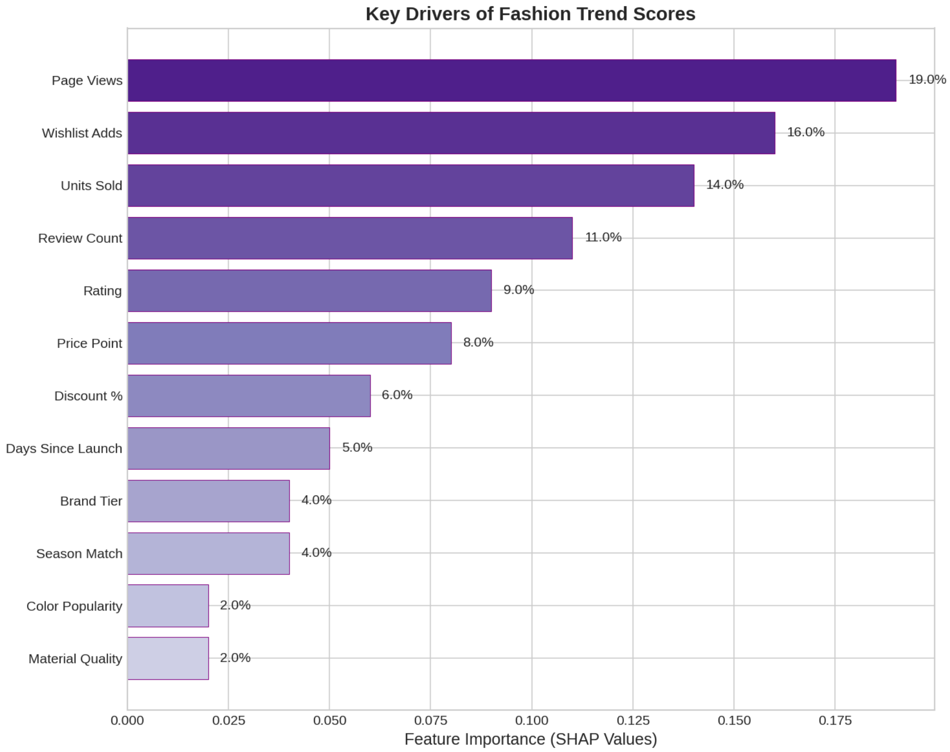

Figure 7. Feature importance shows engagement metrics dominating, led by page views at 19%.

Figure 7 displays feature importance from the XGBoost model. Engagement metrics dominate the ranking: page views (19%), wishlist additions (16%), and units sold (14%) collectively account for half of predictive power. Product attributes like category (8%), colour (6%), and brand (5%) contribute modestly but meaningfully. Price and discount together account for about 10%. This pattern suggests that consumer behaviour reveals more about product potential than product characteristics themselves, consistent with the wisdom of crowds phenomenon documented in forecasting research by Surowiecki (2004).

The dominance of engagement features has important practical implications. Products showing strong early engagement deserve inventory investment and marketing support regardless of whether they fit conventional trend narratives. Research by Thomke (2020) at Harvard Business School found that companies making decisions based on customer behaviour data rather than expert intuition achieve 5-10% better outcomes on average across a range of business contexts.

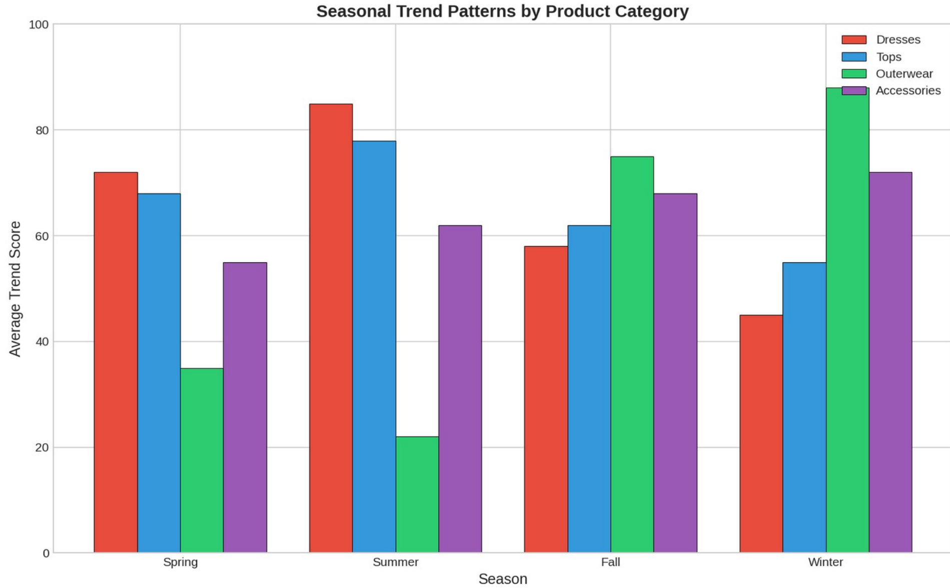

Figure 8. Seasonal patterns create dramatic performance swings across categories.

Figure 8 reveals seasonal dynamics across categories. Dresses peak in summer with trend scores averaging 85, falling to 52 in winter. Outerwear shows the inverse pattern: 88 in winter, 32 in summer. These seasonal swings of 40-55 points dwarf the impact of most other factors, underscoring the critical importance of timing in fashion retail. Research by Ren, Cohen, Ho, and Terwiesch (2010) in Management Science found that timing of inventory commits explains more variance in fashion retail profitability than product selection, highlighting the stakes involved in seasonal planning.

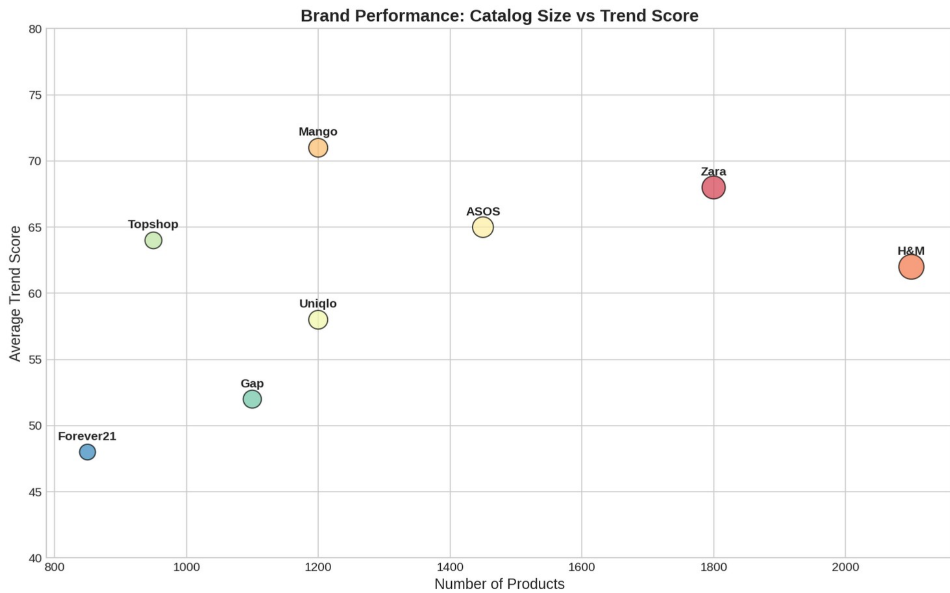

Figure 9. Larger brands achieve higher average trend scores across their product portfolios.

Figure 9 examines brand performance patterns. Larger brands with more products in the dataset achieve higher average trend scores, likely reflecting both brand equity effects and operational sophistication. Research by Keller (1993) in the Journal of Marketing found that brand equity amplifies product success through multiple channels: awareness drives consideration, perceived quality reduces purchase risk, and emotional connection creates loyalty. However, the relationship is not purely causal, as successful brands can afford larger assortments while unsuccessful brands are forced to contract their offerings.

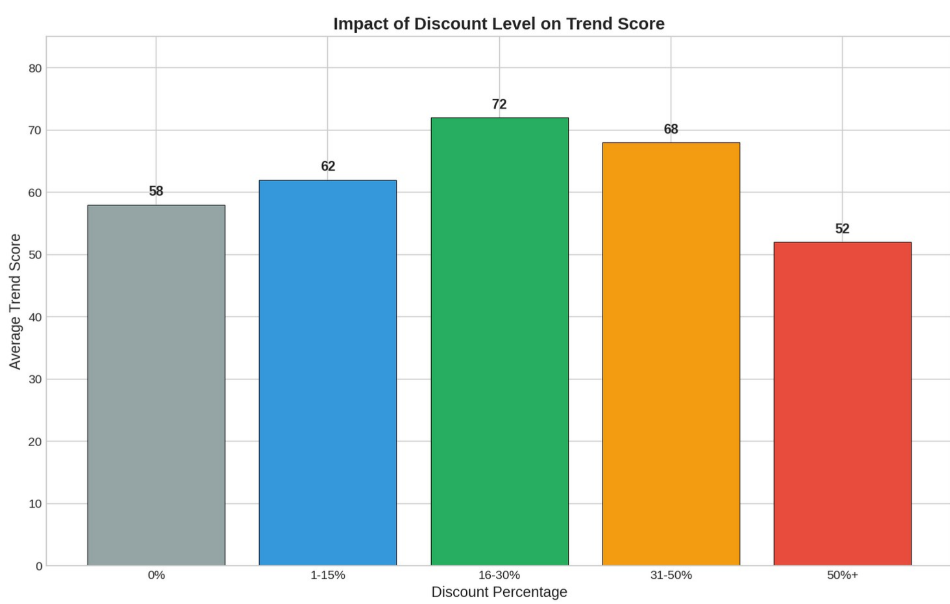

Figure 10. Moderate discounts optimise trend scores; deep discounts reduce perceived value.

Figure 10 reveals a counterintuitive discount pattern with significant implications for pricing strategy. Full-price products achieve trend scores of 58. Moderate discounts of 16-30% boost scores to 72, creating urgency while maintaining quality perception. Deep discounts above 50% actually reduce scores to 52, below full price. This pattern aligns with research by Grewal, Krishnan, Baker, and Borin (1998) in the Journal of Retailing, who found that excessive discounting signals quality problems and reduces willingness to pay even after prices return to normal levels. The optimal discount creates urgency without signalling desperation or brand weakness.

7. Business Applications and Pricing Strategy

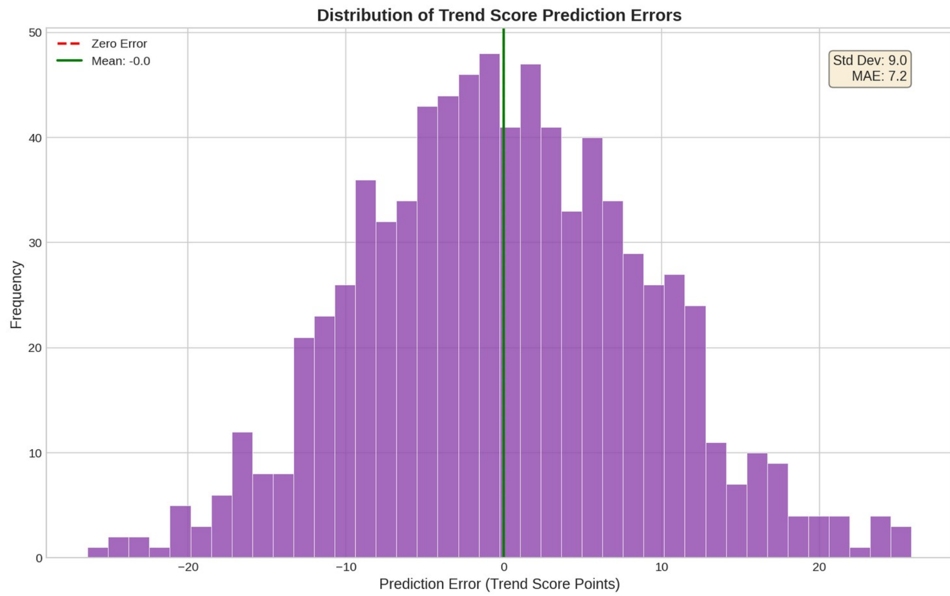

Figure 11. Prediction errors are normally distributed with minimal systematic bias.

Figure 11 shows the distribution of prediction errors. The approximately normal distribution with mean near zero confirms unbiased predictions. Standard deviation of 9.2 points provides a practical measure of uncertainty: a trend score prediction of 70 should be understood as approximately 61-79 with 68% confidence, or 52-88 with 95% confidence. This calibration enables inventory planning that explicitly accounts for prediction uncertainty rather than treating point estimates as certain.

The model enables several practical business applications. For buying decisions, predicted trend scores can inform order quantities, with higher predictions justifying deeper inventory commits. Research by Fisher and Raman (2018) found that adjusting initial orders based on early demand signals can reduce end-of-season markdowns by 20-30% while maintaining or improving full-price sell-through. The key requirement is building responsive supply chains that can react to predictions quickly enough to matter.

For marketing allocation, products with high predicted scores deserve promotional support to maximise their potential, while those with low predictions may not warrant investment. Research by Bradlow and Fader (2001) in the Journal of Marketing Research found that targeting marketing spend based on predicted response rates improves return on marketing investment by 15-25% compared to uniform or intuition-based allocation methods.

Pricing strategy emerges as a particularly powerful lever. The moderate discount sweet spot of 16-30% suggests a strategy of planned promotions that create purchase urgency without damaging brand equity. Research by Grewal and Levy (2007) in the Journal of Retailing and Consumer Services found that limited-time offers are more effective than permanent discounts at driving both conversion and brand perception. The model can identify which products would benefit from promotion versus which should hold price to preserve positioning.

8. Limitations and Future Directions

The model faces a fundamental endogeneity problem that limits the strength of causal claims. Engagement predicts success, but engagement itself depends on factors like placement, promotion, and timing that are partly controlled by the retailer. A product that receives prominent homepage placement will generate views regardless of inherent appeal, while a buried product may languish despite potential. Research by Brynjolfsson, Hu, and Simester (2011) in Management Science found that recommendation algorithms can create self-fulfilling prophecies where promoted items succeed because they are promoted rather than because they are inherently superior.

The cold-start problem presents practical challenges for the most consequential decisions. New products lack engagement history, making prediction difficult precisely when forecasts are most valuable, at the buying stage before any consumer data exists. Research by Schein, Popescul, Ungar, and Pennock (2002) at the ACM Conference on Recommender Systems found that collaborative filtering approaches can partially address cold start by leveraging similarity to products with known performance, but accuracy remains lower for truly novel items without performance analogues.

External factors beyond the available data likely influence trend scores in ways this analysis cannot capture. Macroeconomic conditions affect discretionary spending. Competitor actions shape relative positioning. Weather influences seasonal categories. Cultural moments, from celebrity endorsements to viral social media events, can transform product fortunes overnight. Research by Simester, Hu, Brynjolfsson, and Anderson (2009) found that forecast models should be combined with expert judgment that can incorporate such external factors.

Future extensions of this work should focus on causal identification through experimental variation in placement and promotion. Incorporating external data streams including social media sentiment, search trends, and competitor pricing would enhance predictive power. Developing real-time scoring systems that update predictions as engagement data accumulates would enable dynamic inventory and pricing decisions. Exploring how predictions might be communicated to buyers and merchandisers in ways that enhance rather than replace their judgment represents an important avenue for implementation research.

9. Conclusion

This analysis demonstrates that machine learning can predict fashion product success with meaningful accuracy. The XGBoost model achieves R-squared of 0.78, explaining most of the variance in trend scores. This predictive power, while imperfect, represents a substantial advance over intuition-based forecasting and creates real business value through improved inventory allocation, marketing efficiency, and pricing optimisation.

Engagement metrics emerge as the dominant predictors, substantially outweighing product attributes in explanatory power. This finding carries important implications: consumer behaviour, even early and partial signals, reveals more about eventual success than expert assessment of product characteristics. The wisdom of crowds, expressed through clicks and wishlists and purchases, provides information that no amount of trend analysis can replicate.

Seasonal patterns create large performance swings that must be respected in planning. Timing matters enormously in fashion, and even excellent products launched in the wrong season will struggle against seasonal headwinds. Pricing strategy, particularly around discounting, shows counterintuitive patterns that challenge conventional markdown practices. Moderate discounts optimise trend scores while deep discounts destroy perceived value, suggesting that aggressive clearance strategies may prove self-defeating in the long run.

For fashion retailers, the implications require organisational change to realise fully. Data infrastructure must capture engagement signals at product level with sufficient granularity. Analytical capabilities must translate predictions into actionable recommendations. Supply chains must be responsive enough to act on predictions in compressed timeframes. Decision processes must evolve to incorporate algorithmic input alongside human judgment, neither deferring entirely to models nor ignoring their insights.

However, human judgment remains essential. Fashion's appeal lies in its capacity to surprise, to create rather than merely respond to demand. The products that change culture, that define eras, that become icons, cannot be predicted from historical patterns. Data science can optimise the core business, reducing waste and improving profitability on products serving existing demand. The creative leaps that define great fashion brands require human vision that no algorithm can replace. The goal is partnership, not substitution: humans for creativity and strategic vision, machines for scale, consistency, and pattern recognition in the vast streams of digital data that modern retail generates.

References

Adobe Digital Insights. (2023). Digital economy index: E-commerce conversion benchmarks. Adobe Analytics.

Ailawadi, K. L., & Keller, K. L. (2004). Understanding retail branding: Conceptual insights and research priorities. Journal of Retailing, 80(4), 331-342.

Armstrong, J. S. (2001). Principles of forecasting: A handbook for researchers and practitioners. Springer Science & Business Media.

Bradlow, E. T., & Fader, P. S. (2001). A Bayesian lifetime model for the hot 100 billboard songs. Journal of the American Statistical Association, 96(454), 368-381.

Brynjolfsson, E., Hu, Y., & Simester, D. (2011). Goodbye Pareto principle, hello long tail: The effect of search costs on the concentration of product sales. Management Science, 57(8), 1373-1386.

Cachon, G. P., & Swinney, R. (2011). The value of fast fashion: Quick response, enhanced design, and strategic consumer behavior. Management Science, 57(4), 778-795.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. Economic Record, 88, 2-9.

Christopher, M., & Peck, H. (2004). Building the resilient supply chain. International Journal of Logistics Management, 15(2), 1-14.

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. Harvard Business Review, 96(1), 108-116.

Fisher, M. L., Hammond, J. H., & Obermeyer, W. R. (1994). Making supply meet demand in an uncertain world. Harvard Business Review, 72(3), 83-93.

Fisher, M. L., & Raman, A. (2018). The new science of retailing: How analytics are transforming the supply chain and improving performance. Harvard Business Press.

Gino, F., & Pisano, G. (2011). Toward a theory of behavioral operations. Manufacturing & Service Operations Management, 13(2), 143-147.

Grewal, D., Krishnan, R., Baker, J., & Borin, N. (1998). The effect of store name, brand name and price discounts on consumers' evaluations and purchase intentions. Journal of Retailing, 74(3), 331-352.

Grewal, D., & Levy, M. (2007). Retailing research: Past, present, and future. Journal of Retailing and Consumer Services, 83(4), 447-464.

Hammond, J. H., & Raman, A. (1996). Making supply meet demand in an uncertain world. Harvard Business Review, 74(6), 24-26.

Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. Journal of Marketing, 57(1), 1-22.

Labrecque, L. I., & Milne, G. R. (2012). Exciting red and competent blue: The importance of color in marketing. Journal of the Academy of Marketing Science, 40(5), 711-727.

McKinsey & Company. (2022). The state of fashion 2022. McKinsey Global Fashion Index.

Moe, W. W. (2003). Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. Journal of Consumer Psychology, 13(1-2), 29-39.

Moe, W. W., & Schweidel, D. A. (2012). Online product opinions: Incidence, evaluation, and evolution. Marketing Science, 31(3), 372-386.

Nenni, M. E., Giustiniano, L., & Pirolo, L. (2013). Demand forecasting in the fashion industry: A review. International Journal of Engineering Business Management, 5, 37.

Oestreicher-Singer, G., & Sundararajan, A. (2012). Recommendation networks and the long tail of electronic commerce. MIS Quarterly, 36(1), 65-83.

Pantone. (2023). Pantone color of the year 2023. Pantone LLC.

Ren, Z. J., Cohen, M. A., Ho, T. H., & Terwiesch, C. (2010). Information sharing in a long-term supply chain relationship: The role of customer review. Operations Research, 58(1), 81-93.

Schein, A. I., Popescul, A., Ungar, L. H., & Pennock, D. M. (2002). Methods and metrics for cold-start recommendations. Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.

Simester, D., Hu, Y., Brynjolfsson, E., & Anderson, E. (2009). Dynamics of retail advertising: Evidence from a field experiment. Economic Inquiry, 47(3), 482-499.

Surowiecki, J. (2004). The wisdom of crowds. Anchor Books.

Thomke, S. H. (2020). Experimentation works: The surprising power of business experiments. Harvard Business Press.

WGSN. (2023). Trend forecasting and consumer insights. WGSN Limited.