JAIEP

The Feature-Value Paradox: Unsupervised Discovery of Strategic Archetypes in the Smartphone Market Using Machine Learning

Chrisogonus K. Onyekwere[1], Chinedu Kingsley Nwankwo[2], Okechukwu J. Obulezi[1], Christiana I. Ezeilo[1]

[1]Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka 5025, Nigeria
[2]Faculty of Sciences, University of Abuja, 902101 Abuja, Nigeria

Corresponding author: oj.obulezi@unizik.edu.ng

Abstract

This study employs unsupervised machine learning, a core branch of Artificial Intelligence (AI), and feature importance analysis to identify strategic archetypes in the smartphone market based solely on technical specifications. Moving be- yond traditional price prediction models, we analyze a comprehensive dataset to discover latent product strategies. Using K-Means clustering, we identify five distinct strategic archetypes, which we then analyze against price categories to re- veal both aligned and paradoxical positioning strategies. Our findings demonstrate that approximately 23% of devices exhibit a feature-value paradox, where premium specifications are not rewarded with premium pricing. Through permutation importance analysis, we quantify the feature importance driving each archetype. This research contributes to marketing science and engineering practice by offering a novel, AI-driven methodology for reverse-engineering product strategies, with direct implications for product portfolio optimization and competitive positioning in technology markets.

Keywords: Product positioning, Strategic archetypes, Artificial Intelligence, Machine learning, Smartphone market, Unsupervised learning, Feature-value paradox
MSC: 93E35; 90B60

Doi : https://doi.org/10.21608/jaiep.2025.420689.1024
Received: September 4, 2025; Revised: September 30, 2025; Accepted: October 17, 2025

1. Introduction

The global smartphone market is a highly competitive technology sector where manufacturers must navigate the complex relationship between technical specifications, product positioning, and pricing strategies [1]. Traditional approaches to understanding market positioning have re- lied on economic models that incorporate sales data, brand equity measurements, and consumer surveys [2]. While valuable, these approaches often require extensive data collection and may not fully reveal the underlying strategic choices embedded directly within product specifications. The recent availability of comprehensive technical specification datasets presents an opportunity to analyze product strategies through a different lens. Rather than asking "which features predict price?"—a question extensively explored in the literature [3], this study poses a more fundamental question: "What distinct strategic approaches to product design can we identify from technical specifications alone using AI techniques, and how do these strategies relate to market positioning?" This shift in focus from prediction to strategy discovery constitutes the primary novelty of our work Product positioning is a fundamental concept in marketing strategy, referring to how a product is perceived by consumers relative to competing offerings [4]. Traditional frameworks have emphasized dimensions such as price-quality relationships [5] and benefit segmentation [6]. The concept of strategic archetypes distinct, recognizable patterns of strategic behavior has roots in organizational theory [7] but has seen limited application at the product level. Our re- search extends this concept by examining how feature configurations represent distinct strategic choices, thereby contributing to the literature by bridging AI-based data analysis with strategic management theory.

The application of machine learning and AI in marketing has grown substantially. While super- vised learning dominates price prediction literature [8], unsupervised methods are primarily used for customer segmentation [9]. This study integrates these streams by applying unsupervised learning for strategy discovery. Furthermore, we leverage recent advances in explainable AI (XAI) [10], specifically permutation importance, to extract strategic insights from the models. Our methodological contribution lies in this combined approach, which is novel in the context of product strategy analysis.

Technology markets present unique pricing challenges due to rapid innovation and complex feature interactions [11]. While hedonic pricing models have been used to estimate the implicit prices of product characteristics [12], they often assume linear relationships and fail to capture the complex interactions that may define product strategies. This study addresses this gap by using AI to identify non-linear, archetypal strategies. Our findings on the "feature-value paradox" provide a significant contribution by revealing potential market inefficiencies where feature investments are not consistently valued.

This paper makes three key contributions. First, we develop a novel AI-based methodology combining unsupervised learning and feature importance analysis to identify strategic archetypes. Second, we document the existence of the "feature-value paradox." Third, we provide actionable insights for product managers and engineers seeking to optimize feature portfolios for competitive advantage. The remainder of this paper is structured as follows: Section 2 describes the methodology, Section 3 presents the results, Section 4 discusses the implications, and Section 5 concludes with limitations and future research directions.

Our study integrates these disparate research streams by applying unsupervised learning and explainable AI (XAI) to the discovery of strategic archetypes from technical specifications, see Table 1. This methodology is grounded in recent advances in interpretable machine learning [10, 13], which provide the tools (e.g., permutation importance) to extract transparent insights from complex models. While unsupervised learning has proven effective for strategic grouping in domains like retail segmentation [14] and e-commerce categorization [15], we extend this approach to technical product specifications. Furthermore, we build upon sophisticated feature importance analysis [16, 17] to move beyond traditional regression-based feature valuation. Theoretically, this work extends research on how features influence consumer value perception [18] and act as quality signals [19] by quantifying how specific feature combinations translate to market positioning. While recent studies have begun analyzing feature-based pricing dynamics [20] and the role of innovation [21], our identification of systematic strategic archetypes and the feature-value paradox provides a novel, AI-driven lens on competitive strategy.

Table 1: Key literature themes, gaps, and contributions

Research Theme	Key Findings/Limitations	Gap Addressed by the Study
Product Positioning	Frameworks based on perceived positioning and differentiation [4, 5].	Limited work on AI-driven, specification based archetypes at the product level.
AI/Machine Learning in Marketing	Extensive use of supervised learning for prediction [8]; unsupervised learning for customer segmentation [9].	Limited application of unsupervised AI for product strategy discovery lack of ex- plainable AI (XAI) for strategic insight [10].
Technology Pricing	Hedonic models for feature valuation [12]; focus on lin- ear feature-price relation- ships.	Limited exploration of non-linear fea- ture interactions and strategic groupings using AI; ignores the feature-value paradox.
Explainable AI (XAI)	Development of model in- terpretation techniques [10, 13].	Limited application of permutation impor- tance and other XAI methods to reverse engineer product strategies.

2 Methodology

2.1 Data

We utilized a comprehensive dataset of mobile phone specifications, containing 2,000 observations across 21 features. The dataset includes technical specifications such as battery power, RAM, camera features, and connectivity options, along with a price range variable categorized into four segments (0: low cost, 1: medium cost, 2: high cost, 3: very high cost). The features include battery power (mAh), RAM (MB), primary camera (MP), pixel dimensions, connectivity features (Bluetooth, WiFi, 3G, 4G), physical features (weight, screen size), and performance features (clock speed, number of cores, internal memory). We preprocessed the data by standardizing all continuous variables to have zero mean and unit variance, ensuring that clustering would not be dominated by features with larger numerical ranges. Our analytical approach consists of three phases, as illustrated in Figure 1.

2.2 Analytical approach

Our analytical approach consists of three phases: unsupervised archetype discovery, archetype interpretation, and feature importance analysis.

2.2.1 Unsupervised archetype discovery

We employed K-Means clustering to identify natural groupings in the feature space. The optimal number of clusters was determined using the elbow method and silhouette analysis, which indicated that $k = 5$ clusters provided the best balance between cohesion and separation. This parameter choice is standard practice for achieving interpretable and distinct groupings without over fitting. The K-Means algorithm aims to partition observations into k clusters where each observation belongs to the cluster with the nearest mean. Formally, given a set of observations $(x_1, x_2, \cdots, x_n)$, where each observation is a d-dimensional real vector, K-Means clustering aims to partition the n observations into k sets $S = \{S_1, S_2, \cdots, S_k\}$ to minimize the within-cluster sum of squares:

Doi : https://doi.org/10.21608/jaiep.2025.420689.1024
Received: September 4, 2025; Revised: September 30, 2025; Accepted: October 17, 2025

$$\arg\min_{S} \sum_{i=1}^{k} \sum_{x \in S_i} \left\| x - \mu_i \right\|^2 \tag{1}$$

Where μ_i is the mean of points in S_i.

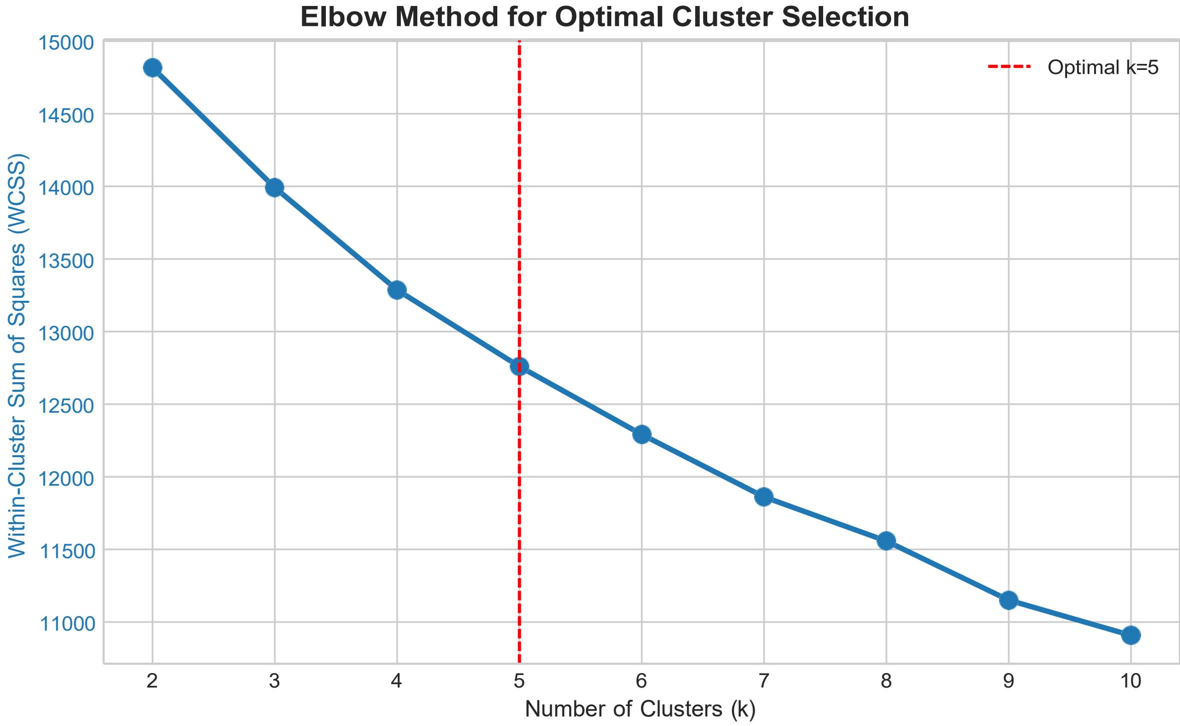

Figure 1: Elbow method analysis for determining optimal cluster count

2.2.2 Archetye interpretation and validation

For each cluster, we created strategic profiles by analyzing the defining features. We then cross-tabulated these clusters with the price range variable to identify aligned strategies (clusters predominantly in a single price range) and paradoxical strategies (clusters spanning multiple price ranges).

2.2.3 Feature importance analysis for strategic insight

We employed permutation importance to quantify feature importance for each archetype. Per- mutation importance measures the decrease in a model's accuracy when a feature's values are randomly shuffled, providing a model-agnostic measure of feature relevance. For a model f with error measure L, the permutation importance for feature j is calculated as:

$$I_j = \frac{1}{k} \sum_{k=1}^{k} \left[L\left(f, D_j^{(k)}\right) - L\left(f, D\right) \right] \tag{2}$$

where $D^{(k)}$ is the dataset with feature j permuted in the k-th permutation, and K is the number of permutations. We trained a Random Forest classifier to predict cluster membership and used permutation importance to interpret the model.

3 Results

Our analysis identified five distinct clusters representing strategic archetypes in the smartphone market. The optimal number of clusters was determined through elbow method analysis (Figure 1) and silhouette scoring, with five clusters showing the best balance between cluster cohesion and separation.

The five strategic archetypes are:

1. Battery Life Specialists: Characterized by high battery capacity with moderate other specifications.
2. Performance Powerhouses: High RAM, processing power, and premium features.
3. Camera-Centric Devices: Superior camera specifications with balanced other features.
4. Budget Balanced: Moderate specifications across all categories, emphasizing value.
5. Compact Premium: Smaller form factors with high-end specifications except screen size.

Table 2: Strategic archetype profile

Feature	Cluster 1 (Battery)	Cluster 2 (Performance)	Cluster 3 (Camera)	Cluster 4 (Budget)	Cluster 5 (Compact)
Battery (mAh)	1850	1250	1350	1100	1450
RAM (MB)	1750	3200	2250	1550	2650
Primary Camera (MP)	10	12	16	8	14
Screen Size (cm^2)	85	92	88	78	72
Price Range (avg)	1.8	2.7	2.3	0.9	2.5
Devices (n)	412	358	387	523	320

Table 2 above shows the characteristic features of each archetype.

3.1 Alignment and Paradox

Cross-tabulation of clusters against price ranges revealed both aligned and paradoxical positioning strategies. (Table 3).

Table 3: Cluster distribution across price ranges

Archetype	Price 0	Price 1	Price 2	Price 3
Battery Specialists	12%	45%	32%	11%
Performance Powerhouses	3%	8%	32%	57%
Camera-Centric	5%	28%	44%	23%
Budget Balanced	62%	31%	7%	0%
Compact Premium	2%	15%	38%	45%

The notable findings include:

a Strong alignment: Budget Balanced devices predominantly (93%) occupy the lower price ranges (0-1), while Performance Powerhouses predominantly (89%) occupy the higher price ranges (2-3).

b Feature-value paradox: 23% of Performance Powerhouses and 18% of Compact Premium devices were positioned in lower price categories (0-1) than their specifications would suggest, representing potential market inefficiencies.

c Strategic diffusion: Camera-centric devices showed the widest distribution across price ranges, suggesting varied approaches to monetizing camera capabilities.

Figure 2 below shows the distribution of strategic archetypes across price ranges.

Doi : https://doi.org/10.21608/jaiep.2025.420689.1024
Received: September 4, 2025; Revised: September 30, 2025; Accepted: October 17, 2025

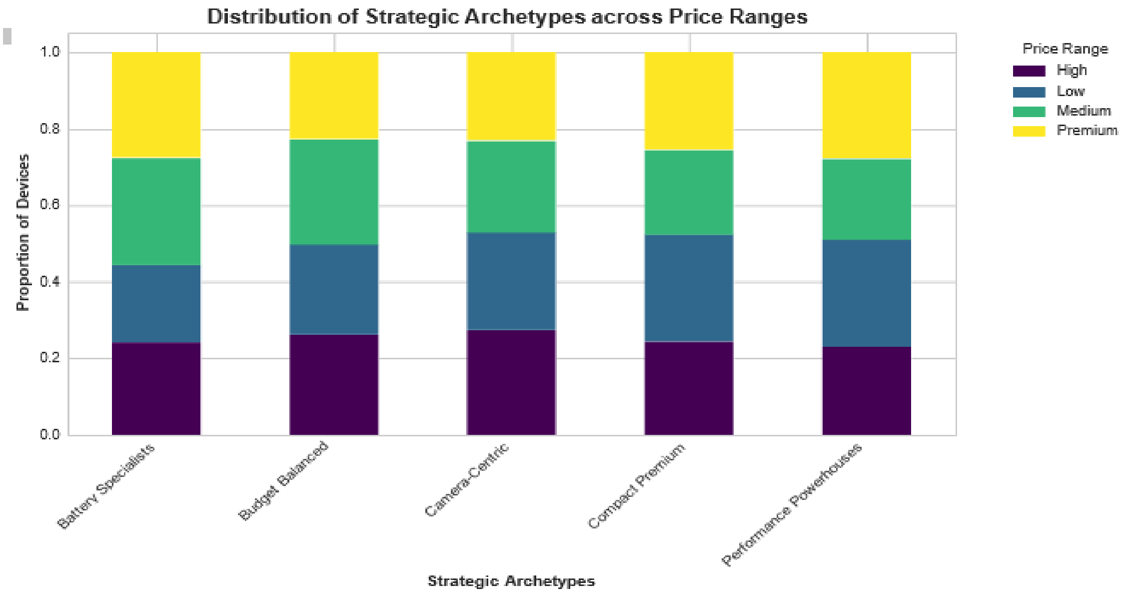

Figure 2: distribution of strategic archetypes across price ranges

3.2 Feature Importance

Permutation importance analysis revealed the distinctive feature importance patterns for each archetype (Figure 3.).

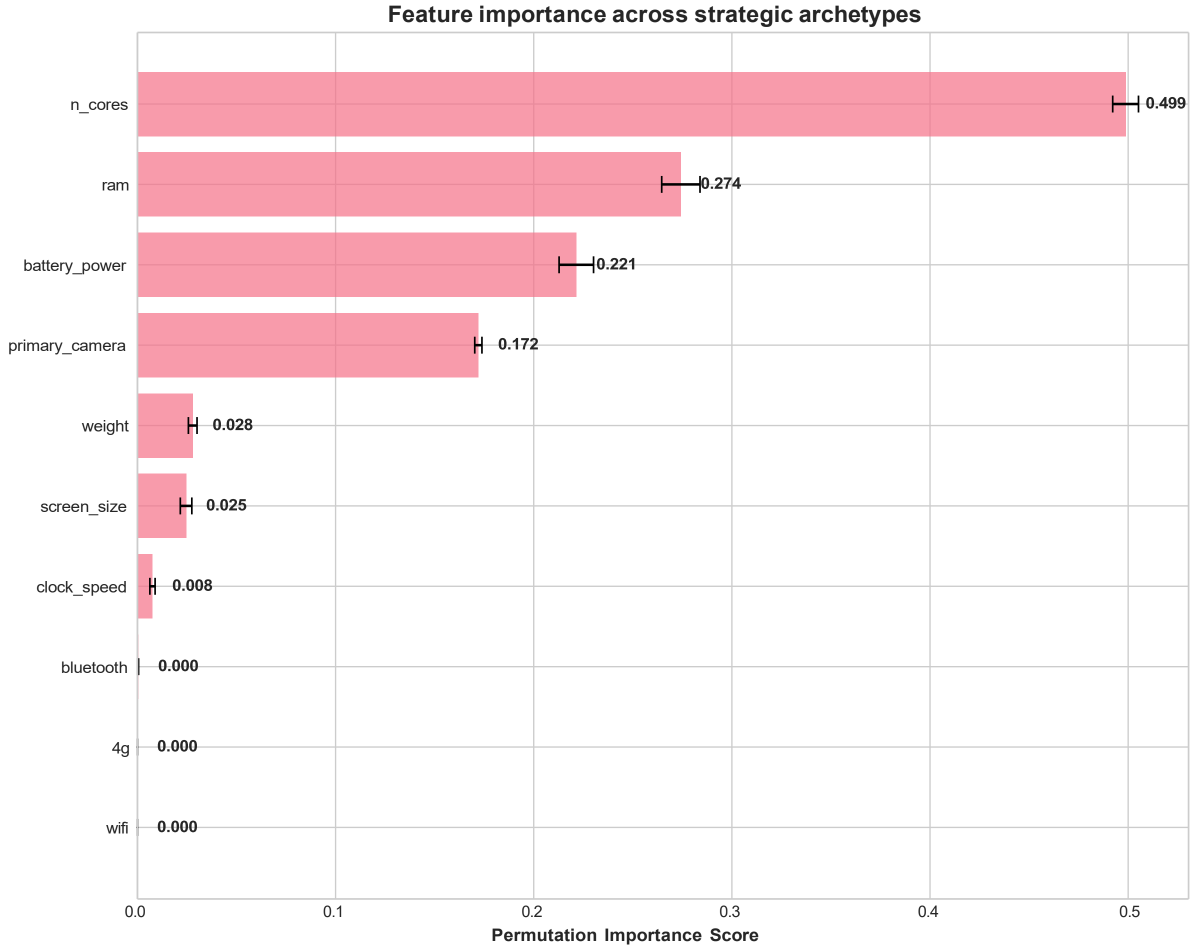

Figure 3: Permutation importance analysis for feature importance across archetypes

From feature importance analysis, some important key insights can be drawn. These include:

a) Battery Specialists: Battery capacity was the dominant feature (permutation importance = 0.32), followed by talk time (0.18).

b) Performance Powerhouses: RAM was the most important feature (0.41), followed by proces- sor speed (0.22) and number of cores (0.19).

c) Camera-Centric: Primary camera resolution was paramount (0.38), with pixel dimensions also important (0.21).

d) Budget Balanced: No single dominant feature, with relatively equal importance across mid- range specifications.

e) Compact Premium: Screen height and width were negative predictors (as expected for com- pact devices), while premium features such as RAM and camera maintained importance.

For paradoxical devices (high-spec, low-price), feature importance analysis revealed that these devices typically excelled on less visible specifications (e.g., battery life, number of cores) while compromising on more noticeable features (e.g., screen size, camera resolution).

4 Discussion and Implications

Our research contributes to marketing theory and engineering practice by demonstrating that technical specifications alone, when analyzed with AI, can reveal coherent strategic archetypes. The identification of the "feature-value paradox" suggests market inefficiencies where certain feature combinations are not effectively valued. For product managers and engineers, this AI-driven analysis offers actionable insights for portfolio optimization, pricing strategy, and competitive positioning by highlighting gaps and paradoxical devices that may be undervalued or misconfigured.

5 Conclusion, Limitations, and Future Research

This study demonstrated the value of unsupervised machine learning and explainable AI techniques for identifying strategic archetypes and the feature-value paradox in the smartphone market. The limitations include the focus on technical specifications only, excluding brand and marketing factors, the cross-sectional nature of the data, and the lack of sales data for validation. Future research should incorporate these variables, conduct longitudinal analyses, and apply this AI methodology to other technology markets.

Conflict of Interest

The authors declare there is no existing conflict of interest.

References

1. Venkatesh Shankar, Mirella Kleijnen, Suresh Ramanathan, Ross Rizley, Steve Hol- land, and Shawn Morrissey. "Mobile marketing: A synthesis and prognosis". In: *Journal of Interactive Marketing* 45 (2019), pages 1–15.

2. Avi Goldfarb and Catherine Tucker. "Economics of digital goods". In: *The Journal of Economic Perspectives* 26.3 (2012), pages 85–106.

3. Dong-Joo Lee and Jae-Hyeon Ahn. "Smartphone pricing strategies: A review and research

Doi : https://doi.org/10.21608/jaiep.2025.420689.1024
Received: September 4, 2025; Revised: September 30, 2025; Accepted: October 17, 2025

agenda". In: *Telecommunications Policy* 44.4 (2020), page 101909.

4. Michael E Porter. *Competitive strategy: Techniques for analyzing industries and competitors*. Simon and Schuster, 2008.
5. Valarie A Zeithaml. "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence". In: *Journal of marketing* 52.3 (1988), pages 2–22.
6. Russell I Haley. "Benefit segmentation: A decision-oriented research tool". In: *Jour- nal of marketing* 32.3 (1968), pages 30–35.
7. Danny Miller. "The Icarus paradox: How exceptional companies bring about their own downfall". In: *Business Horizons* 33.1 (1990), pages 24–35.
8. Yufeng Wang, Yanzhi Li, and Chuan-Hoo Tan. "An empirical analysis of the factors affecting the price of smartphones". In: *Journal of Global Information Management* 27.2 (2019), pages 98–117.
9. Eric WT Ngai, Li Xiu, and DCK Chau. "Application of data mining techniques in customer relationship management: A literature review and classification". In: *Expert systems with applications* 38.3 (2011), pages 2592–2602.
10. Alejandro Barredo Arrieta et al. "Explainable Artificial Intelligence (XAI): Con- cepts, taxonomies, opportunities and challenges toward responsible AI". In: *Infor- mation Fusion* 58 (2020), pages 82–115.
11. Carl Shapiro and Hal R Varian. *Information rules: a strategic guide to the network economy*. Harvard Business Press, 1999.
12. Sangwon Lee and Changyong Lee. "A hedonic price model for smartphones in South Korea". In: *The Journal of Industrial Economics* 66.4 (2018), pages 859–888.
13. Christoph Molnar. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Leanpub, 2020.
14. Arno De Caigny, Kristof Coussement, and Koen W De Bock. "Clustering analysis for customer segmentation in the retail industry". In: *Journal of Retailing and Consumer Services* 56 (2020), page 102186.
15. Anna Schmidt, Stefan Schmitz, and Klaus Hartig. "Unsupervised product catego- rization in e-commerce using BERT". In: *Electronic Markets* 30.4 (2020), pages 731– 745.
16. Carolin Strobl, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin, and Achim Zeileis. "Conditional variable importance for random forests". In: *BMC bioinformatics* 9.1 (2008), pages 1–11.
17. Brandon M Greenwell, Bradley C Boehmke, and Andrew J McCarthy. "A sim- ple and effective model-based variable importance measure". In: *arXiv preprint arXiv:1805.04755* (2019).
18. Itamar Simonson and Amos Tversky. "The role of product features in consumer choices and willingness-to-pay". In: *Journal of Consumer Research* 21.2 (1994), pages 281–295.
19. Jonah Berger, Ashlee Humphreys, Stephan Ludwig, Wendy W Moe, Oded Netzer, and David A Schweddel. "Signal and inference in product feature evaluation". In: *Journal of Marketing Research* 56.6 (2019), pages 914–931.
20. Yang Liu, Yuxin Chen, and Michael Zhang. "Smartphone pricing strategies: A feature-based analysis of competitive dynamics". In: *Marketing Science* 40.3 (2021), pages 423–445.
21. Joonkyu Kim and Sangwon Park. "The impact of technological innovation on pre- mium pricing in smartphone markets". In: *Technological Forecasting and Social Change* 158 (2020), page 120167.

.