KOSIME

Predicting Benchmark Product Potential through Technological

Phylogeny and Developmental Strategy Analysis

Youwei He

Ph.D. candidate, SNU

TEMEP

heyouwei2022@gmail.com

2024. 12. 20.

CONTENTS
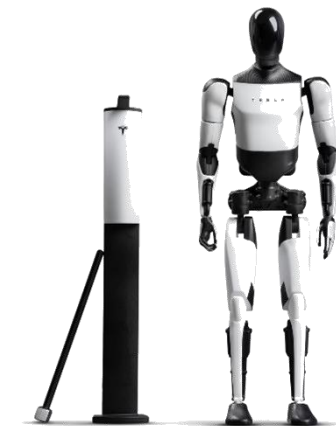
Research Background

- **What** is benchmark products?

- **Why** benchmark products are important?

- **How** to identify and predict benchmark products?

Introduction

Research Background

- **What** is benchmark products?

  A **benchmark product** is a product that sets the standard of excellence within its category, exemplifying superior quality, performance, and innovation (Bhutta & Huq, 1999; Camp, 2024).

- **Why** benchmark products are important?

  Benchmark products play a pivotal role in shaping technological trajectories by building on their predecessors, advancing existing innovations, and setting new standards for quality and functionality, thereby exerting profound **impacts** (Knudsen et al., 2023).

- **How** to identify and predict benchmark products?

  Benchmark products can be identified through technological phylogenetic approach (Lee et al., 2022; Park et al., 2024), which trace product evolution and highlight transformative influences and predict the benchmark products by machine learning methods.

Literature review

Theories of Technology forecasting

Technology forecasting categories in Council (2010)

Technology forecasting categories	Method	Literature
Judgmental or intuitive methods that rely on expert opinions	Delphi Method	(Dalkey & Helmer, 1963; Gordon & Helmer-Hirschberg, 1964; Linstone & Turoff, 1975; Rowe & Wright, 1999)
Extrapolation and trend analysis	Trend extrapolation	(Moore, 1998)
	Substitution analysis	(Fisher & Pry, 1971)
	S-curve model	(Kuznets, 2024)
	Analogies	(Armstrong, 2001; Green & Armstrong, 2007)
	Morphological Analysis and TRIZ	(Kucharavy & De Guio, 2011)
Theoretical frameworks for predicting technological evolution	Increasing returns	(Arthur, 1996; Kelly, 1999)
	Influence diagrams	(Howard & Matheson, 2005)
Scenarios and simulations	Hypothetical scenarios	(Kahn, 1962; Kahn, 2014)

Literature review

Literatures of Technology forecasting with machine learning methods

Data	Method	Results and implications	Time available	Literature
Patent	Supervised models	Emerging technologies	Early stage	(Kyebambe et al., 2017; Lee et al., 2018; Zhou et al., 2020)
Patent	Supervised models	Technology convergence	Middle stage	(Cho et al., 2021; Lee et al., 2021; San Kim & Sohn, 2020)
Product	Supervised models with phylogenetic approach	Benchmark product	Late stage	This study

Research Question
How can we define and identify the benchmark product? What makes a product to be a benchmark status? Can we predict the potential for a product to become a benchmark product?

Research gap

Despite extensive research on new product development and innovation, a significant gap exists in identifying and predicting benchmark products. While technological forecasting has been well studied, research applying machine learning methods to predict benchmark products remains notably absent from the literature.

Methods

| Data |
|---|

Data source: ZOL Mobile[1]

Cell phones in the Chinese market from 2001 to 2020. Out of 3,430 cell phone models, across 63 different brands and 11,006 technological features.

We use a technological phylogenetic tree to identify the benchmark products by years. We define the benchmark products as the top n% most impactive products within 5 years by the phylogenetic tree. The impacts can be measured by the descendant counts (Shi & Evans, 2023). We also use the XGBoost (Chen et al., 2016) to predict the benchmark products by years.

---

1 mobile.zol.com.cn

Methods

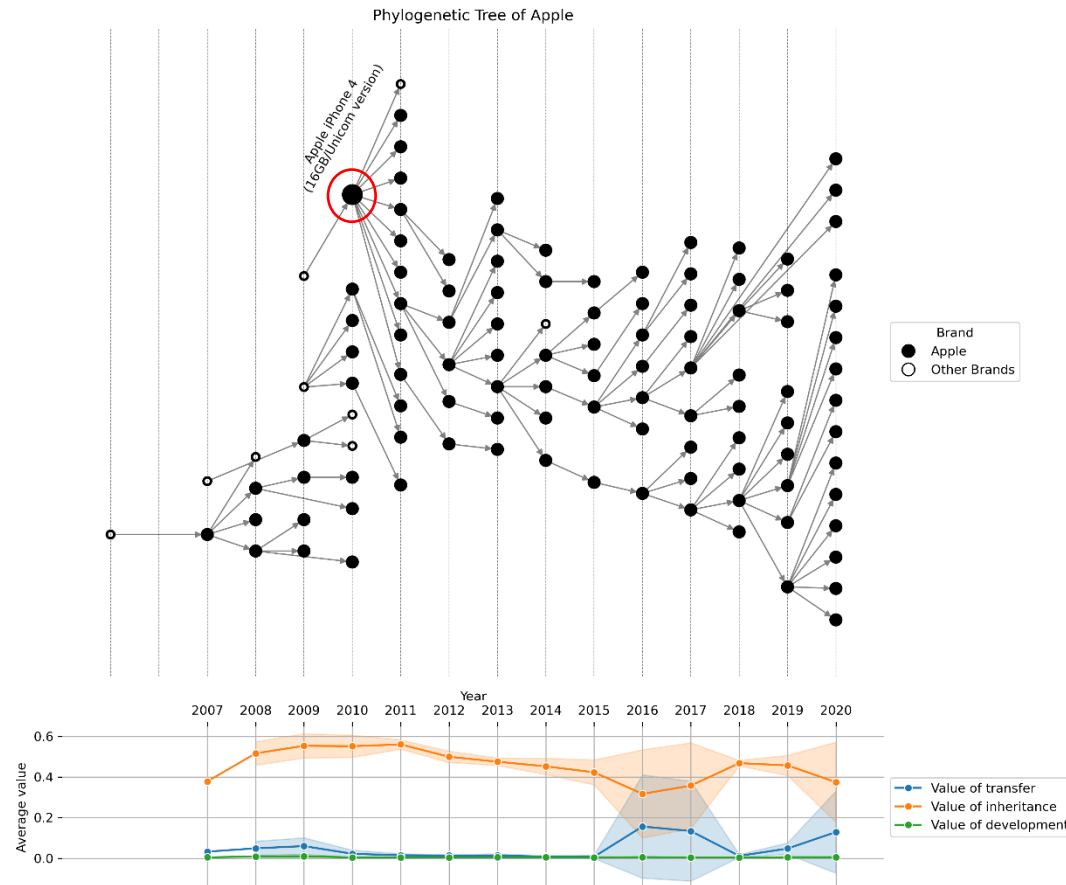## Conceptual framework

Technology Management,
Economics and Policy Program
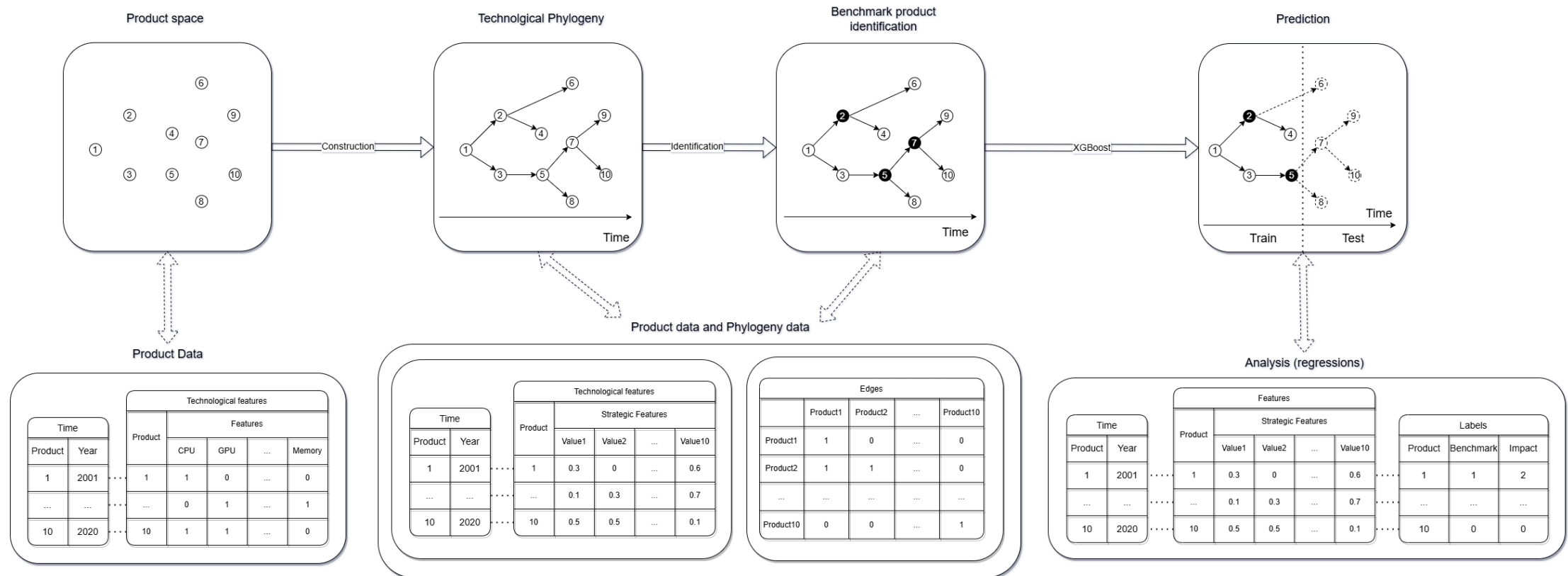
Identifying benchmark products
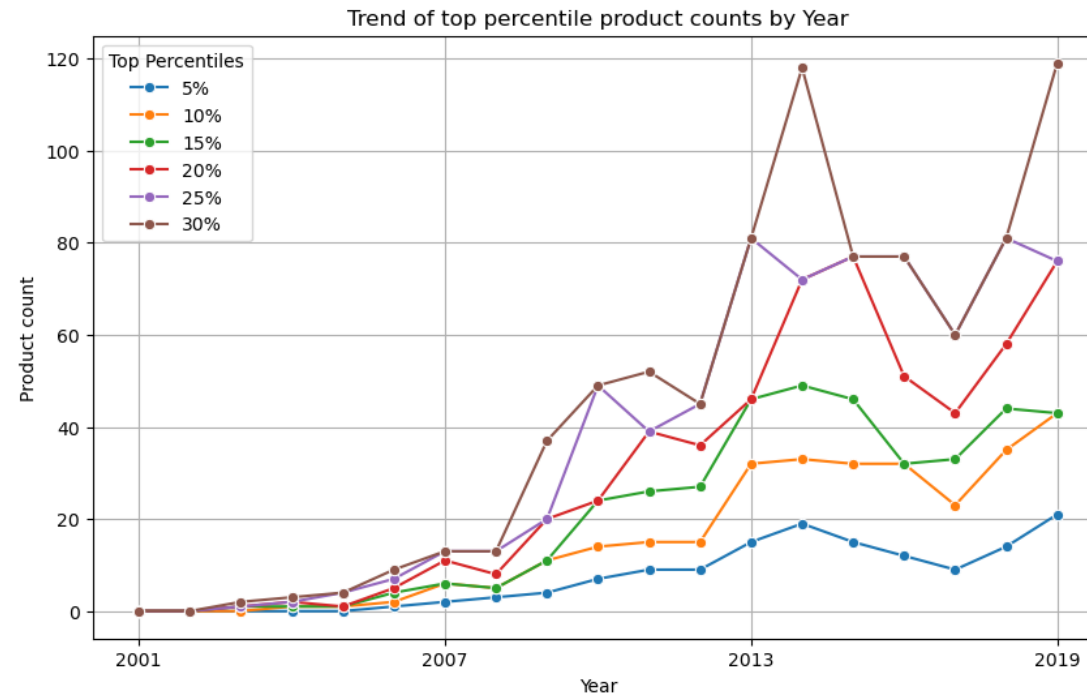

iPhone 4 (2010)

Product impact refers to the number of descendants a product has within 5 years.

A **benchmark product** is defined as a product that falls within the top percentile of product impact in the same year.

Technology Management,
Economics and Policy Program

Benchmark product count trend (by top percentiles)

Trend of top percentile product counts by Year

Methods

| Variable definition		
Variable Name	Description	Related theories
Count of development	The count of new technologies in the product	Knowledge-Based View (Grant, 1996), Creative destruction (Schumpeter, 2013), First-Mover Advantages (Lieberman & Montgomery, 1988; Kerin et al., 1992)
Value of development	The weighted value of the new technologies in the product	
Count of inheritance	The count of the existing technologies in the product in the same brand	Technological trajectories (Dosi, 1982), Path dependency (Arthur, 1989)
Value of inheritance	The weighted value of the existing technologies in the product in the same brand	
Count of transfer	The count of the existing technologies in the product in the other brands	Technology transfer (Alkhazaleh et al., 2022), Technology diffusion (Geroski, 2000; Loch & Huberman, 1999; Rogers et al., 2014)
Value of transfer	The weighted value of the existing technologies in the product in the other brands	
Count of technologies	The count of the total technologies in the product	Technology Sourcing and Combination (Arora et al., 2004; Fleming, 2001)
Value of technologies	The weighted value of the total technologies in the product	
Brand difference	The brand difference between the product and its ancestor	Open Innovation (Chesbrough, 2003)
Year difference	The year difference between the product and its ancestor	Technological discontinuities (Tushman & Anderson, 2018), Late mover advantage (Shankar et al., 1998; Zhou, 2002)
Product distance	The technological distance between the product and its ancestor	

The weight of the technology is calculated using frequency weight.

Results

Prediction results

Prediction: Training on year (from 2001 to T), Testing on year T+1,Features: around 20 product variables (strategic variables), Label: Benchmark product dummy

Benchmark products (Top percentile)	AUC of Validation	AUC of Test
5%	0.8732	0.8703
10%	0.8652	0.8803
15%	0.8907	0.8939
20%	0.9140	0.9044
25%	0.9365	0.9352
30%	0.9616	0.9512
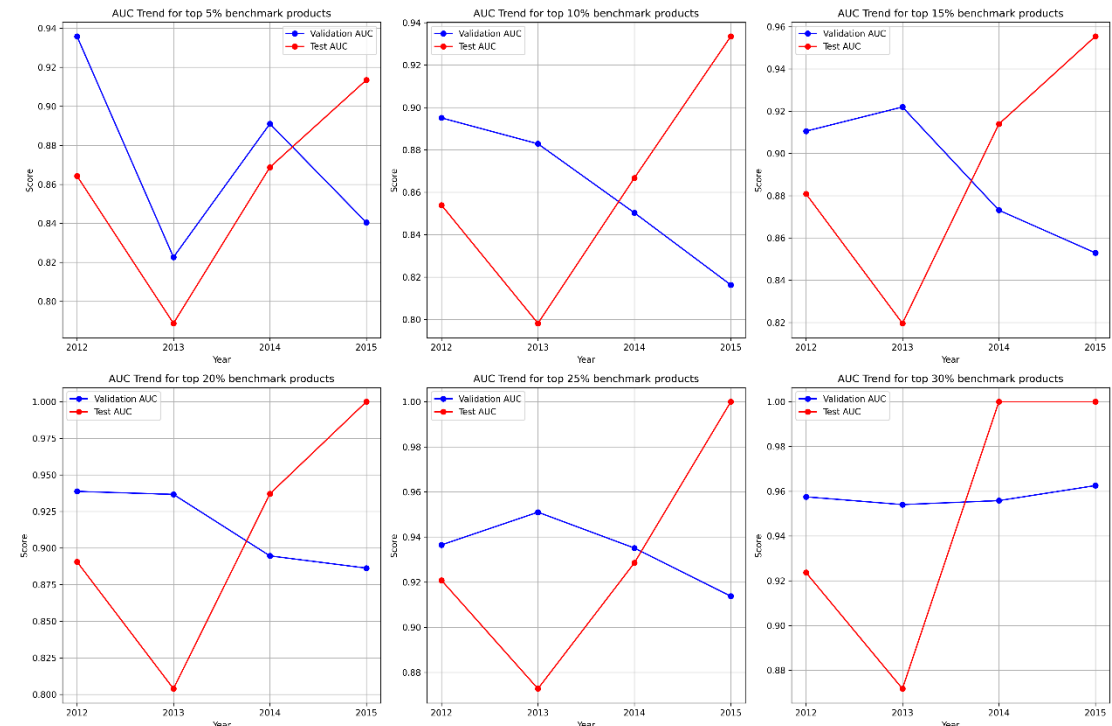
AUC stands for **Area Under the Curve**, specifically referring to the area under the **Receiver Operating Characteristic (ROC)** curve. It is a crucial metric for evaluating the performance of binary classification models.

Technology Management, Economics and Policy Program

Regression results

Fixed effects Logit regressions of Top 5% benchmark product

	(1)	(2)	(3)	(4)	(5)	(6)
	Benchmark product (Top 5%)					
Technology count	-0.0068	-0.0067	-0.0048	-0.0113	-0.013	-0.0775***
	(0.0064)	(0.0069)	(0.0072)	(0.008)	(0.0081)	(0.017)
Product distance	20.8173***	20.7458***	19.6974***	18.8911***	19.1054***	24.3384***
	(1.4746)	(2.1198)	(2.3087)	(2.2981)	(2.3062)	(3.8943)
Different year		0.0103	0.0242	0.0637	0.0281	0.5652
		(0.17)	(0.1688)	(0.1691)	(0.1698)	(0.365)
Brand difference			0.2598	0.2362	0.1532	-0.1819
			(0.322)	(0.3259)	(0.3308)	(0.3177)
Value of development				50.8458*	53.1544*	54.8671•
				(25.5566)	(26.9397)	(29.9711)
Value of transfer					2.4713•	29.7932***
					(1.3557)	(4.3081)
Value of inheritance						26.9801***
						(4.1997)
Brand	✓	✓	✓	✓	✓	✓
Obs.	1620	1620	1620	1620	1620	1620
Adj. Pseudo R2	0.1637	0.1607	0.1588	0.162	0.1638	0.2437

Note: Significance Indicators: '***' indicates highly significant ($p < 0.001$), '**' indicates significant ($p < 0.01$), and '*' indicates moderately significant ($p < 0.05$).

	Regression results

Fixed effects regressions of benchmark product bottom percentile

	(1)	(2)	(3)	(4)	(5)	(6)
	Benchmark (top percentile)					
Value of development	4.2947***	4.2652***	1.6713**	2.1195***	-0.5390	-0.5253
	(0.9299)	(0.9231)	(0.6566)	(0.6853)	(0.4598)	(0.4576)
Year difference	0.3058***	0.3057***	0.3078***	0.3077***	0.0595***	0.0597***
	(0.0360)	(0.0361)	(0.0317)	(0.0316)	(0.0187)	(0.0185)
Value of transfer		-0.0191	0.5598***	0.7114***	0.5351***	0.5248***
		(0.0456)	(0.1400)	(0.1647)	(0.0607)	(0.0612)
Value of inheritance			0.5829***	0.7281***	0.6116***	0.6068***
			(0.1351)	(0.1598)	(0.0570)	(0.0571)
Technology count				-0.0009***	-0.0019***	-0.0019***
				(0.0003)	(0.0002)	(0.0002)
Product distance					2.6139***	2.5804***
					(0.1286)	(0.1305)
Brand difference						0.0090
						(0.0112)
Year	✓	✓	✓	✓	✓	✓
Brand	✓	✓	✓	✓	✓	✓
Obs.	1771	1771	1771	1771	1771	1771
R^2	0.661	0.661	0.687	0.689	0.870	0.870

Note: Significance Indicators: '***' indicates highly significant ($p < 0.001$), '**' indicates significant ($p < 0.01$), and '*' indicates moderately significant ($p < 0.05$).
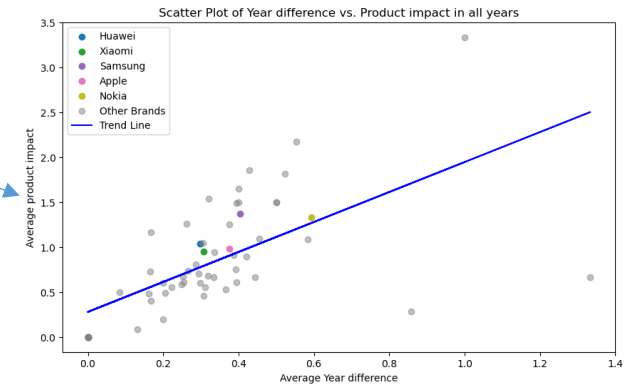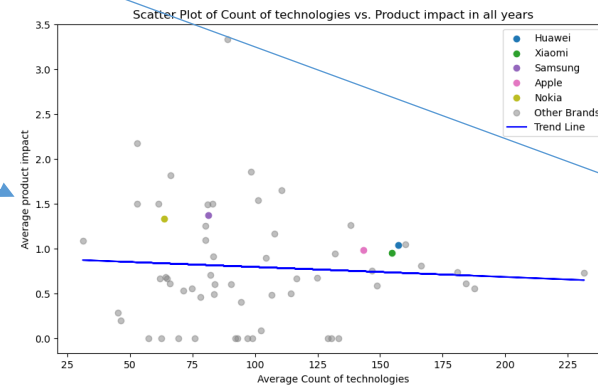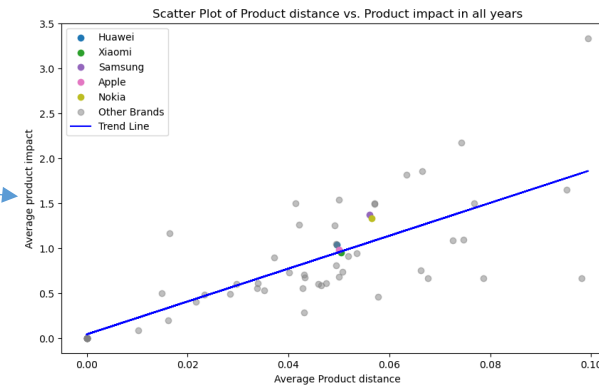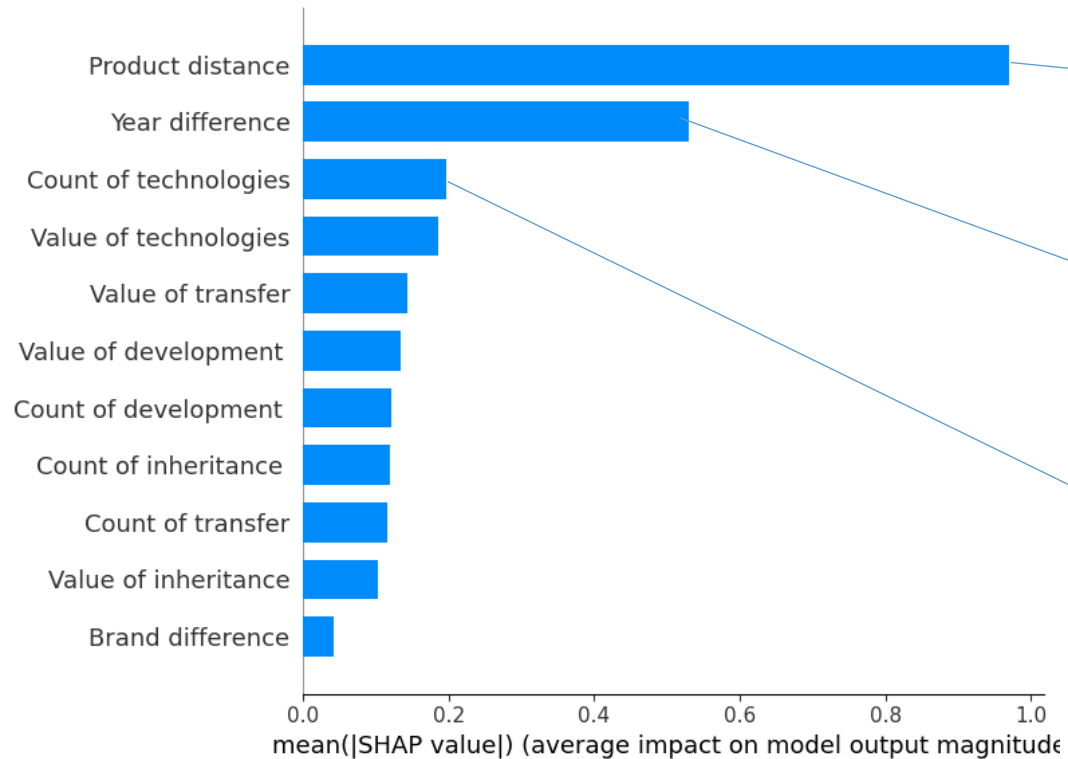
| **Robustness** | | | | | | |

Multiple Dimensional Fixed Effect Regression results of impact (in 5 years) of products

	(1)	(2)	(3)	(4)	(5)	(6)
			Product impacts			
Value of development	58.1585***	54.5920***	35.0140**	43.6371***	24.9516*	24.7457*
	(14.4951)	(14.7029)	(14.5171)	(15.0404)	(13.4659)	(13.3499)
Year difference	1.8685***	1.8475***	1.8640***	1.8615***	0.1168	0.1138
	(0.2501)	(0.2484)	(0.2185)	(0.2156)	(0.1313)	(0.1325)
Value of transfer		2.3138*	6.3945***	9.3110***	8.0721***	8.2277***
		(1.3698)	(1.6188)	(1.9926)	(1.6791)	(1.7336)
Value of inheritance			4.3992***	7.1943***	6.3749***	6.4471***
			(0.9976)	(1.3206)	(0.8979)	(0.9386)
Technology count				-0.0171***	-0.0243***	-0.0249***
				(0.0045)	(0.0043)	(0.0047)
Product distance					18.3722***	18.8768***
					(1.4367)	(2.0457)
Brand difference						-0.1361
						(0.3272)
Year	√	√	√	√	√	√
Brand	√	√	√	√	√	√
Observations	1771	1771	1771	1771	1771	1771
R2	0.253	0.258	0.270	0.275	0.347	0.347

Note: Significance Indicators: '***' indicates highly significant ($p < 0.001$), '**' indicates significant ($p < 0.01$), and '*' indicates moderately significant ($p < 0.05$).

Technology Management,
Economics and Policy Program

Robustness

SHAP Values

Empirical application

Phylogenetic tree of Apple products

Average strategic values of Apple products

Year	Value of transfer	Value of inheritance	Value of development
2007	0.0328	0.3784	0.0049
2008	0.0493	0.5164	0.0089
2009	0.0602	0.5536	0.0103
2010	0.0224	0.5517	0.0041
2011	0.015	0.5603	0.0037
2012	0.0141	0.4995	0.0046
2013	0.0141	0.4754	0.0049
2014	0.0082	0.4526	0.0052
2015	0.0079	0.4231	0.0037
2016	0.1566	0.3171	0.006
2017	0.1341	0.3577	0.0044
2018	0.012	0.4682	0.0048
2019	0.0481	0.4573	0.0056
2020	0.1291	0.3747	0.0055

Technology Management, Economics and Policy Program
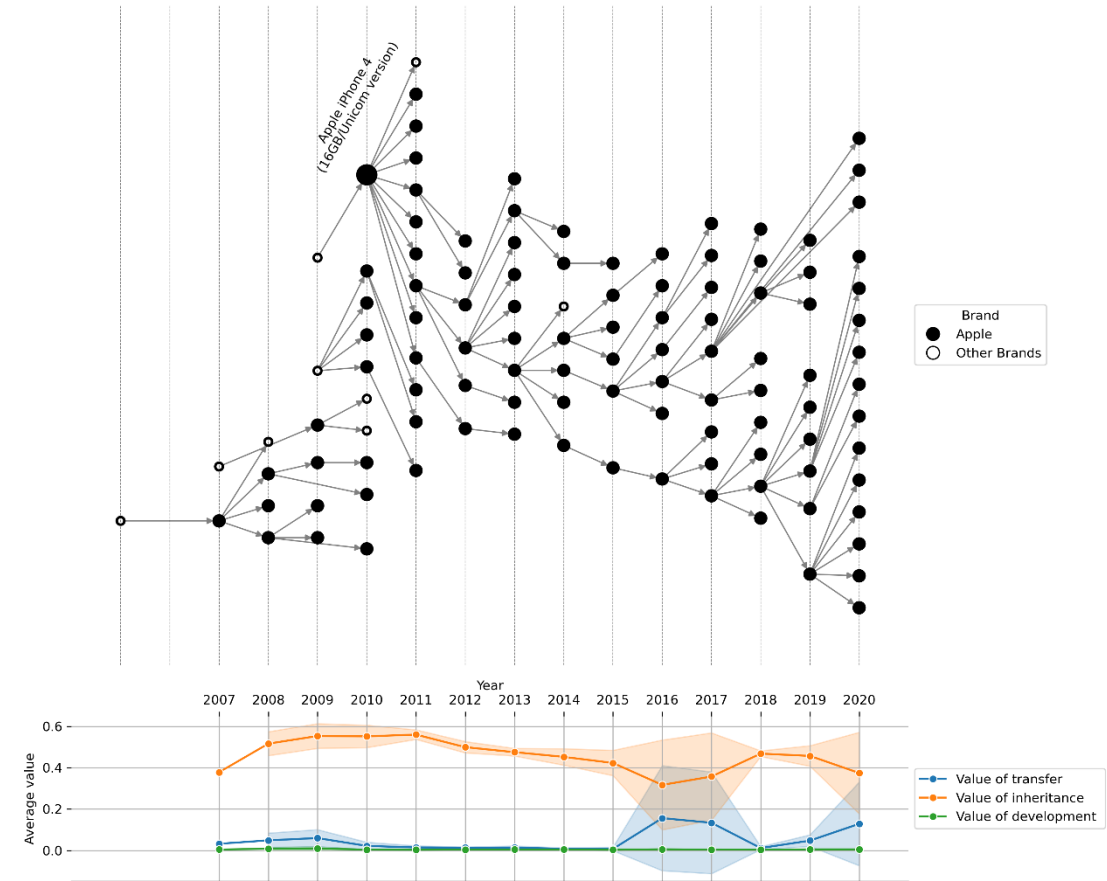
Empirical application

Apple iPhone4

Product	Apple iPhone 4 (16GB/Unicom version)
Brand	Apple
Year	2010
Count of development	21
Value of development	0.0118
Count of Inheritance	40
Value of Inheritance	0.4877
Count of transfer	16
Value of transfer	0.0863
Technology count	77
Technology value	0.5858
Top 5%	Yes
Top 10%	Yes
Top 15%	Yes
Top 20%	Yes
Top 25%	Yes
Top 30%	Yes

$f(x) = 9.657$

0.1953040977637001 = Product distance	+2.79
1.0 = Year difference	+1.27
77.0 = Count of technologies	+1.04
16.0 = Count of transfer	+0.76
0.0862718247175624 = Value of transfer	+0.69
0.5857583019513865 = Value of technologies	+0.59
0.0117539655369165 = Value of development	+0.55
40.0 = Count of inheritance	+0.38
21.0 = Count of development	+0.35
2 other features	+0.23

$E[f(X)] = 1$
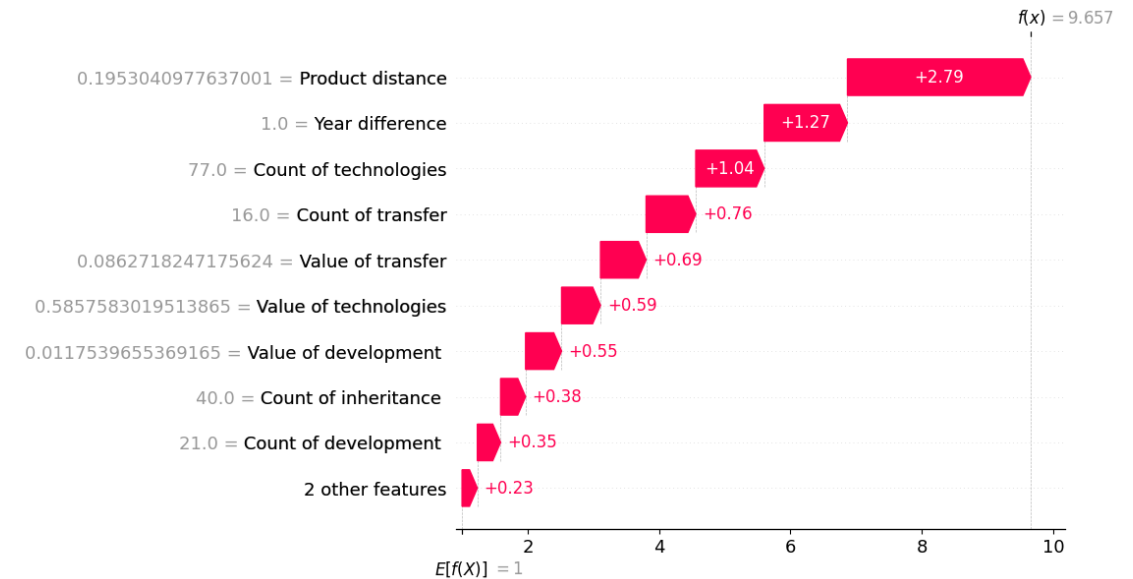
Conclusion

Based on technological phylogenetic approaches and machine learning methods, the potential for a product to become a benchmark product can be predicted with high performance.

Through regression analysis, we have identified key principles that explain what makes a product a benchmark. These include innovating beyond its key ancestor, adopting mature technologies, incorporating key technologies, and decreasing the number of technologies embedded in the product.

Conclusion

Implications

Theoretical Implications

Benchmark products set industry standards for quality, functionality, and innovation, shaping future product development. For companies, creating a benchmark product establishes market leadership, drives innovation, and reinforces long-term competitiveness by influencing industry trends and expectations.

Practical Implications

Predicting benchmark potential enables firms to optimize product design. By analyzing key ancestors, calculating product distance, and evaluating technology contributions, developers can iteratively adjust configurations to enhance benchmark likelihood. Integrating these insights with pricing ensures a balance between innovation and market competitiveness.

Conclusion

1. **Phylogenetic Tree Limitation:**

The use of phylogenetic trees excludes multiple ancestors, potentially introducing biases.

2. **Market and Scope Restriction:**

Focused on technical features and limited to the Chinese market, reducing generalizability.

3. **Forecasting Constraint:**

While effective for evaluating current products, predicting entirely novel technologies remains unattainable.

Future Directions

- Employ phylogenetic networks to capture multi-ancestor relationships.

- Expand datasets to include global markets for broader applicability.

- Develop methods to forecast novel technologies, enhancing prediction capabilities.

References

Alkhazaleh, R., Mykoniatis, K., & Alahmer, A. (2022). The success of technology transfer in the industry 4.0 era: A systematic literature review. *Journal of Open Innovation: Technology, Market, and Complexity*, *8*(4), 202.

Armstrong, J. S. (2001). *Principles of forecasting: a handbook for researchers and practitioners* (Vol. 30). Springer.

Arora, A., Fosfuri, A., & Gambardella, A. (2004). *Markets for technology: The economics of innovation and corporate strategy*. MIT press.

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The economic journal*, *99*(394), 116-131.

Arthur, W. B. (1996). Increasing returns and the new world of business.

Bhutta, K. S., & Huq, F. (1999). Benchmarking–best practices: an integrated approach. *Benchmarking: An International Journal*, *6*(3), 254-268.

Camp, R. C. (2024). *Benchmarking: the search for industry best practices that lead to superior performance*. CRC Press.

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,

Chesbrough, H. W. (2003). Open innovation: The new imperative for creating and profiting from technology. *Harvard Business School*.

Cho, J. H., Lee, J., & Sohn, S. Y. (2021). Predicting future technological convergence patterns based on machine learning using link prediction. *Scientometrics*, *126*, 5413-5429.

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management science*, *9*(3), 458-467.

References

Dosi, G. (1982). Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. *Research policy*, *11*(3), 147-162.

Fisher, J. C., & Pry, R. H. (1971). A simple substitution model of technological change. *Technological Forecasting and Social Change*, *3*, 75-88.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management science*, *47*(1), 117-132.

Geroski, P. A. (2000). Models of technology diffusion. *Research policy*, *29*(4-5), 603-625.

Gordon, T. J., & Helmer-Hirschberg, O. (1964). Report on a long-range forecasting study. *(No Title)*.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic management journal*, *17*(S2), 109-122.

Green, K. C., & Armstrong, J. S. (2007). The ombudsman: Value of expertise for forecasting decisions in conflicts. *Interfaces*, *37*(3), 287-299.

Howard, R. A., & Matheson, J. E. (2005). Influence diagrams. *Decision Analysis*, *2*(3), 127-143.

Kahn, H. (1962). Thinking About the Unthinkable. In: Horizon Press.

Kahn, K. B. (2014). *New product forecasting: an applied approach*. Routledge.

Kelly, K. (1999). *New rules for the new economy: 10 radical strategies for a connected world*. Penguin.

Kerin, R. A., Varadarajan, P. R., & Peterson, R. A. (1992). First-mover advantage: A synthesis, conceptual framework, and research propositions. *Journal of marketing*, *56*(4), 33-52.

References

Knudsen, M. P., Von Zedtwitz, M., Griffin, A., & Barczak, G. (2023). Best practices in new product development and innovation: Results from P DMA's 2021 global survey. *Journal of Product Innovation Management*, *40*(3), 257-275.

Kucharavy, D., & De Guio, R. (2011). Application of S-shaped curves. *Procedia Engineering*, *9*, 559-572.

Kuznets, S. (2024). Secular movements in production and prices. In *Business Cycle Theory, Part I Volume 1* (pp. 283-322). Routledge.

Kyebambe, M. N., Cheng, G., Huang, Y., He, C., & Zhang, Z. (2017). Forecasting emerging technologies: A supervised learning approach through patent analysis. *Technological Forecasting and Social Change*, *125*, 236-244.

Lee, C., Hong, S., & Kim, J. (2021). Anticipating multi-technology convergence: A machine learning approach using patent information. *Scientometrics*, *126*, 1867-1896.

Lee, C., Kwon, O., Kim, M., & Kwon, D. (2018). Early identification of emerging technologies: A machine learning approach using multiple patent indicators. *Technological Forecasting and Social Change*, *127*, 291-303.

Lee, J.-D., Jeong, D., Jung, E.-Y., Kim, Y., Kim, J., He, Y., & Choi, S. (2022). Mapping the evolutionary pattern of mobile products: a phylogenetic approach. *IEEE Transactions on Engineering Management*, *71*, 4776-4790.

Lieberman, M. B., & Montgomery, D. B. (1988). First-mover advantages. *Strategic management journal*, *9*(S1), 41-58.

Linstone, H. A., & Turoff, M. (1975). *The delphi method*. Addison-Wesley Reading, MA.

Loch, C. H., & Huberman, B. A. (1999). A punctuated-equilibrium model of technology diffusion. *Management science*, *45*(2), 160-177.

References

Moore, G. E. (1998). Cramming more components onto integrated circuits. *Proceedings of the IEEE*, *86*(1), 82-85.

Park, H., Jeong, D., & Lee, J.-D. (2024). Evolutionary mechanism for diversity dynamics in technology using a phylogenetic tree approach: directional suggestions for photovoltaic technology. *Industrial and Corporate Change*, dtae020.

Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In *An integrated approach to communication theory and research* (pp. 432-448). Routledge.

Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International journal of forecasting*, *15*(4), 353-375.

San Kim, T., & Sohn, S. Y. (2020). Machine-learning-based deep semantic analysis approach for forecasting new technology convergence. *Technological Forecasting and Social Change*, *157*, 120095.

Schumpeter, J. A. (2013). *Capitalism, socialism and democracy*. routledge.

Shankar, V., Carpenter, G. S., & Krishnamurthi, L. (1998). Late mover advantage: How innovative late entrants outsell pioneers. *Journal of Marketing research*, *35*(1), 54-70.

Shi, F., & Evans, J. (2023). Surprising combinations of research contents and contexts are related to impact and emerge with scientific outsiders from distant disciplines. *Nature Communications*, *14*(1), 1641.

Tushman, M. L., & Anderson, P. (2018). Technological discontinuities and organizational environments. In *Organizational innovation* (pp. 345-372). Routledge.

References

Zhou, Y., Dong, F., Liu, Y., Li, Z., Du, J., & Zhang, L. (2020). Forecasting emerging technologies using data augmentation and deep learning. *Scientometrics*, *123*, 1-29.

Zhou, Z. (2002). *Achieving late-mover advantage: The effects of enhancing and distinctive strategies*. Virginia Polytechnic Institute and State University.

Thank you☺

Appendices

The definition of technological phylogenetic tree and network

1) Tree: $T = G(V,E)$ and Network: $N = G(V,E)$, while T is a directed acyclic graph (DAG) with the limitation that one node can only has one ancestor, and N is a directed acyclic graph without any limitation.

2) Distance-based method for inferring the optimum technological phylogenetic tree and network (Occam's razor):

Definition: The <u>parsimony score</u> of a phylogenetic tree is the sum of the substitution costs along all the edges of the tree

$$PS(T,X) = \sum_{e_{i,j} \in E} d(e_{i,j}), subject\ to\ t_j < t_i, and\ e_{i,j} \in E, and\ i,j \in V,$$

Where:

$PS(T,X)$ is the parsimony length of the phylogeny.

X is the vector of the embedding or the chromosome of genes

e is an edge in the phylogenetic tree T.

$d(e_{i,j})$ is the distance of edge $e_{i,j}$.

i is the focal node's index

j is any node's index which appeared before the focal node i

The objective function of technological phylogenetic tree and network

The objective function of optimum technological phylogenetic tree refers to:

$$T^* = argmin_T[PS(T,X)] = argmin_T\left[\sum_{i=1}^{m} d(e_{i,j})\right], \ subject \ to \ t_j < t_i, and \ e_{i,j} \in E, and \ i,j \in V,$$

The objective function of the optimum technological phylogenetic network refers to:

$$N^* = argmin_N[PS(N,X)] = argmin_N\left[\sum_{i=1}^{m}\sum_{k=1}^{n} d(e_{i,j_k})\right], subject \ to \ t_{j_k} < t_i \ \forall k, and \ e_{i,j_k} \in E, and \ i,j_k \in V,$$

Where:

N^* represents the optimized network structure. m is the node count in V.

$j_1, j_2, \ldots, j_n$ are the indexes of the nodes corresponding to the n smallest distances for each focal node i.

The additional sum $\sum_{k=1}^{n}$ represents the selection of the n edges with the smallest distances for each focal node i.

$d(e_{i,j_k})$ is the distance associated with edge e_{i,j_k}.

Appendices

XGBoost and Hyperparameter Configuration

The **objective function** in XGBoost consists of the log loss (logistic loss) for binary classification and a regularization term to control model complexity:

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \ell(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

Where:

1. $\ell(y_i, \hat{y}_i) = -(y_i log(\hat{p}_i) + (1 - y_i)log(1 - \hat{p}_i))$: Logistic loss for binary classification.

   y_i: True label (0 or 1).

   $\hat{p}_i = \frac{1}{1+e^{-\hat{y}_i}}$: Predicted probability, calculated using the sigmoid function.

   $\hat{y}_i = \sum_k f_k(x_i)$: The predicted logit, which is the sum of the predictions from all the trees.

2. $\Omega(f_k) = \frac{1}{2}\lambda||w_k||^2 + \gamma T_k$: Regularization term for tree f_k.

   w_k: Leaf weights of the tree f_k.

   T_k: Number of leaves in tree f_k.

   λ, γ: Regularization parameters.

3. $f_k(x)$: A decision tree (weak learner), mapping an input x to a real-valued score.

Appendices

XGBoost and Hyperparameter Configuration

The **Activation Function** (Sigmoid for Binary Classification) is below:

After computing the raw prediction $\hat{y}_i = \sum_k f_k(x_i)$, the **sigmoid function** is applied to map the logit $\hat{y}_i$ to a probability:

$$\hat{p}_i = \frac{1}{1 + e^{-\hat{y}_i}}$$

XGBoost Hyperparameters

Hyperparameter	Values Explored	Description
max_depth	[6, 10]	Maximum depth of the trees. Controls the model's capacity to capture complex patterns. Deeper trees risk overfitting.
n_estimators	[100, 200]	Number of trees in the ensemble. More trees improve performance but increase computational cost.
learning_rate	[0.01, 0.1]	Step size for updating weights after each tree. Lower values lead to more precise learning but require more trees.
colsample_bytree	[0.3, 0.7]	Fraction of features used for training each tree. Reduces overfitting and speeds up training by introducing randomness.
subsample	[0.5, 1.0]	Fraction of training samples used for building each tree. Helps prevent overfitting by randomly sampling the training data.
objective	binary:logistic	Defines the loss function for binary classification tasks.
eval_metric	auc	Metric used for model evaluation during training. Area Under the Curve (AUC) measures the model's ability to distinguish between classes.
scale_pos_weight	Default: 1	Balances positive and negative samples in imbalanced datasets.
Default Parameters	gamma=0, reg_alpha=0, reg_lambda=1	Default regularization parameters to control model complexity and reduce overfitting.