NextGen Scientific Publication
American International Journal of Computer Science and Technology

Volume 7 Issue 1, Pg. No. 191-201, AIJCST-V7I1P116, 2025
https://doi.org/10.63282/3117-5481/AIJCST-V7I1P116

Original Article

Explainable Machine Learning for Market Entry Selection by US Consumer Technology Firms under Tariff, Currency, and Import Concentration Risks

***Krunal More**
Lawrence Technological University, Southfield, USA.

Abstract:

Selecting an overseas market for consumer technology requires a company to judge opportunity alongside costs that can change before launch. This paper develops an explainable decision framework that separates forecasts of product import growth from tariff and currency exposure and from concentration in import origins. The proposed application covers twelve destination markets and three consumer technology product groups. Historical decision years run from 2012 through 2018, with outcome measurement ending in 2020 and an information cutoff of December 2024. Release dated records prevent later revisions from entering earlier decisions. Regularized linear models provide an interpretable benchmark for shallow tree ensembles. Explanations describe predictive associations, while a separate accounting model translates specified tariff and exchange rate changes into unit contribution margins. A reproducible four market example shows how a destination with the highest baseline margin can be excluded by a concentration constraint, leaving a lower ranked destination as the preferred feasible option. These calculations are analytical illustrations, not observed company results or trained model performance. The contribution is an auditable connection between prediction, explanation, and explicit managerial constraints. The framework specifies temporal validation, country exclusion tests, uncertainty reporting, and rejection rules for missing customs evidence. It supports a defensible shortlist while reserving final entry decisions for company specific commercial and regulatory assessment.

Keywords:

International Market Selection, Explainable Machine Learning, Tariffs, Exchange Rates, Import Concentration, Consumer Technology.

Article History:

Received: 11.01.2025

Revised: 13.02.2025

Accepted: 25.02.2025

Published: 28.02.2025

1. Introduction

A US consumer technology company may compare several countries that appear attractive on income, import growth, and digital infrastructure. [1,2] Its actual launch economics also depend on customs treatment, the currency of receipts, and the concentration of existing import sources. [1,2] Country ranking methods provide a useful starting point, but combining these considerations in a single opaque score makes it difficult to identify why a destination was selected or rejected. [1,2]

The present framework concerns preliminary destination selection for a US headquartered firm selling imported physical products through local distributors. [3–5] It does not select factories, recommend securities, or estimate the value of a foreign subsidiary. [3–5] The candidate products are cellular telephones, portable computers, and headphones, represented by harmonized

Copyright by the Author(s). This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-NC-ND 4.0) License (https://creativecommons.org/licenses/by-sa/4.0/)

customs categories. [3–5] Existing research on internationalization and firm heterogeneity motivates evaluating opportunity relative to the company's resources rather than assuming that the same destination is best for every exporter. [3–5]

Three distinctions structure the analysis. [6–11] Company nationality differs from customs origin; exchange rate volatility differs from a specified depreciation shock; and import concentration by origin differs from concentration among competing brands. [6–11] Each distinction affects the interpretation of the resulting shortlist. [6–11] A US brand assembling products abroad must assess the importing destination's treatment of that origin, rather than automatically applying a tariff reported for goods originating in the United States. [6–11]

The paper contributes a modular architecture in which machine learning predicts an observable market proxy, explanations describe model behavior, and transparent constraints determine feasible destinations. [16–25,29,30] A worked example establishes the accounting implications of those constraints without presenting fabricated company observations. [16–25,29,30] The study therefore supplies a research design and an executed analytical demonstration; it does not claim an empirically validated market recommendation or superior predictive accuracy. [16–25,29,30]

2. Research Foundation and Specific Contribution

Systematic market selection research distinguishes country screening from detailed opportunity assessment. [1–5] Clustering and ranking can organize comparable destinations, but a high country score does not establish a firm's ability to serve local consumers profitably. [1–5] The proposed contribution preserves that separation: product import growth supports opportunity screening, while a company supplied contribution model addresses landed costs and available receipts. [1–5] A separate commercial review considers distribution capacity and product compliance. [1–5]

Trade research shows that tariffs affect prices, sourcing, and welfare through mechanisms that depend on incidence and adjustment. [6–9] Currency research further emphasizes invoicing and imported inputs. [6–9] These findings justify modeling tariff and currency channels explicitly; they do not supply universal pass through coefficients for headphones or portable computers in every destination. [6–9] Accordingly, the accounting demonstration states its pricing and invoicing assumptions, and the empirical design treats pass through as an unresolved firm specific input. [6–9]

An explanation method can decompose a prediction without identifying the effect of changing a country's policy. [19–25] Feature importance is especially difficult to interpret when tariff levels, income, and import structure covary. [19–25] The framework consequently compares an interpretable benchmark with tree ensembles and accompanies local attributions with checks for correlated predictors. [19–25] Explanations serve audit and model diagnosis; policy effects require a separate identification strategy. [19–25]

Table 1. Research Gaps and Proposed Responses [1,2,21,26,30]

Gap	Specific Response	Evidence Needed
Opaque aggregate ranking	Separate prediction and constraints	Audit complete shortlist
Historical data leakage	Require publication eligible vintages	Release ledger by fold
Unclear explanation meaning	Group correlated predictive contributions	Stability and support checks

The central research question asks whether an auditable combination of product opportunity forecasts and explicit risk constraints produces a more defensible shortlist than opportunity ranking alone. [16–28,30] H1 proposes lower out of sample prediction error for a shallow ensemble than for persistence. [16–28,30] H2 proposes greater explanation stability when correlated predictors are interpreted together. [16–28,30] H3 proposes that joint tariff and currency stress changes the feasible shortlist relative to baseline margins. [16–28,30] Only H3 is demonstrated analytically here; H1 and H2 require the specified data and validation. [16–28,30]



Figure 1. Architecture Separating Market Prediction Explanation Accounting and Managerial Feasibility Gates. [1,2,21,30]

3. Scope and Data Construction

The candidate universe contains Canada, Mexico, Brazil, the United Kingdom, Germany, France, India, Japan, South Korea, Australia, Indonesia, and South Africa. [10,15] These destinations provide a predefined comparative scope rather than a claimed optimal global sample. [10,15] Three HS 2012 categories are specified: 851712 for cellular or other wireless telephones, 847130 for portable automatic data processing machines, and 851830 for headphones and related sets. [10,15] Their definitions are broader than individual consumer brands. [10,15]

Decision years span 2012 through 2018, yielding at most 252 destination product year rows before exclusions. [10,12,14] The target is two year dollar import growth, measured as log import value in year d plus two minus log import value in year d. [10,12,14] Outcome years therefore end in 2020. [10,12,14] This is a market import proxy, not a label for actual entry, consumer adoption, or firm profit; exchange rates and product composition can change dollar values without equivalent changes in unit demand. [10,12,14]

Table 2. Study Scope and Observation Structure [10,12,15]

Element	Specification
Destinations	Twelve predefined markets
Products	HS 851712 847130 and 851830
Decision years	2012 through 2018
Outcome horizon	Two years ending by 2020
Maximum panel	12 × 3 × 7 = 252 rows
Information cutoff	31 December 2024

Each row represents information available by 31 December of decision year d. [10–13,26,28] A feature uses the latest eligible observation whose publication date is no later than that decision date. [10–13,26,28] Observational year alone is insufficient because trade and tariff records can be published late or revised. [10–13,26,28] Every stored series retains observation year, release date, extraction date, revision status, and source identifier; a subsequent download cannot recreate a historical information set unless the earlier vintage is documented. [10–13,26,28]

Training labels obey the same availability rule. [10,26–28] A training row is eligible only when both endpoint trade records used in its target had already been released at the relevant model fitting date. [10,26–28] A calendar embargo of two years alone cannot guarantee this condition. [10,26–28] If release histories cannot be recovered, the analysis is labeled a retrospective revised data exercise rather than a real time backtest, and its performance cannot support claims about historical decision readiness. [10,26–28]

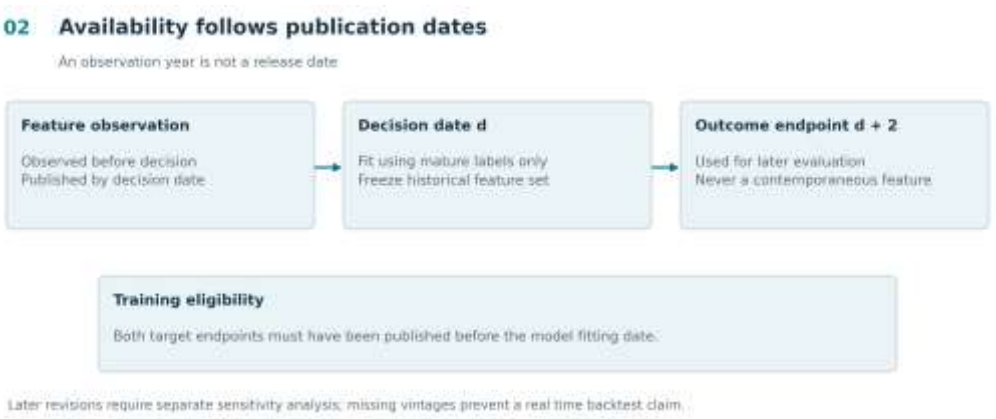


Figure 2. Publication Aware Timeline Showing How Feature and Outcome Availability Constrain Historical Training [10,26,28]

Primary trade records use destination reported imports, partner origins, and an explicit customs classification revision. [10,12,15] A BACI reconciliation provides a separately labeled robustness dataset, not silently substituted observations. [10,12,15] National totals and partner flows are checked for coverage and valuation consistency. [10,12,15] HS concordances require documented aggregation; one to many mappings without defensible value allocation are excluded rather than split equally to create artificial product histories. [10,12,15]

Tariff features distinguish the product's ordinary most favored nation rate, any eligible preferential rate, and applicable additional duties. [6,7,10,11] Eligibility must be verified for the planned production origin and product classification. [6,7,10,11] An effectively applied rate in a research database does not by itself demonstrate that a particular shipment meets preference rules. [6,7,10,11] The main feature uses a simple line average within the mapped category; alternative aggregation rules are reported because import weighted averages can reflect prior sourcing responses. [6,7,10,11]

Table 3. Data Provenance and Admissibility Rules [10–15]

Source	Required Fields	Exclusion Trigger
Trade records	Reporter origin product value release	Unresolved coverage or mapping
Tariff schedule	Destination origin eligibility duties	Unknown applicable treatment
Currency series	Units frequency observation release	Unverified quote convention
BACI sensitivity	Version and reconciled flows	Undocumented vintage

4. Risk Measurement and Decision Quantities

Currency exposure uses local currency units per US dollar, so an increase indicates local depreciation against the dollar. [8,13] Predictors include the preceding twelve month exchange rate change and the standard deviation of monthly log changes, calculated only from released observations. [8,13] A real effective exchange rate may describe broader competitiveness but cannot replace the bilateral conversion rate in the unit margin calculation. [8,13] Invoicing currency and hedging remain company supplied parameters. [8,13]

Import concentration is H = sum of squared origin shares, with shares calculated within destination, product, and year from a consistent set of partner flows. [10,12,14] The index lies between zero and one when shares sum to one. [10,12,14] Missing origins cannot be assigned zero without evidence, and residual partner categories remain visible. [10,12,14] High concentration signals dependence on a narrow origin base; it does not prove limited brand competition or predict a disruption probability. [10,12,14]

The outcome model uses lagged import value, recent import growth, origin concentration, origin count, verified tariff measures, bilateral currency change, and currency volatility, together with product indicators. [10–14,19,20,26] Feature construction remains parsimonious because 252 rows is only an upper bound and repeated observations share country conditions. [10–14,19,20,26] Income

and population may enter a prespecified sensitivity model, but additional predictors are not added merely because they improve in sample fit. [10–14,19,20,26]



Figure 3. Three Risk Channels with Distinct Measures Units and Limits of Interpretation. [6–14]

Opportunity forecasts and margin constraints remain in separate units. [1,2,29,30] Predicted import growth ranks candidates within a product and decision year. [1,2,29,30] The accounting module calculates dollar contribution per unit after specified import costs, using company supplied prices and costs. [1,2,29,30] A feasible destination must meet both a minimum stressed contribution and a concentration ceiling. [1,2,29,30] These are explicit managerial choices, not coefficients learned from aggregate trade data or regulatory thresholds implied by the paper. [1,2,29,30]

For a prespecified scenario set, feasibility requires the minimum contribution across all included scenarios to reach the margin floor while origin concentration remains below its ceiling. [30] The firm first eliminates infeasible destinations and then ranks survivors by predicted opportunity; stressed contribution serves as a separately reported tie breaker. [1,2,30] The numerical example compares feasible margins only and therefore does not imply a combined score with invented opportunity predictions. [19,30]

Conditional value at risk is appropriate only in an extension with defensible scenario probabilities and a defined monetary loss. [29] The nine cell sensitivity grid supplies neither a probability distribution nor a forecast of shock frequency. [29,30] Consequently, its minimum margin is a worst case accounting quantity, not an estimated tail expectation. [29,30] The scenario owner must approve the range before observing which destination survives. [30]

Each artifact should retain an immutable identifier, source checksum, classification mapping, and transformation log. [10,19] Exact arithmetic tests verify the exchange rate direction, tariff percentage point conversion, concentration sums, and boundary equality. [10,11,30] Changing the concentration ceiling or margin floor requires a new recorded scenario version so that reviewers can reproduce the shortlist from its stated assumptions. [30]. Report threshold changes when comparing alternative shortlisted destinations. [30].

Table 4. Risk Measures and Decision Thresholds [6–14,30]

Quantity	Unit	Role
Tariff	Decimal ad valorem rate	Landed cost
Currency change	Local units per dollar	Receipt conversion
Currency volatility	Monthly log change deviation	Predictive feature
Origin HHI	Zero to one	Concentration constraint
Margin floor	25 dollars per unit	Demonstration policy
HHI ceiling	0.40	Demonstration policy

5. Explainable Learning and Validation

The benchmark predicts that the most recently observable two year import growth persists. [16–20] An elastic net regression supplies a regularized, interpretable comparison with country grouped dependence retained in evaluation. [16–20] Random forests and gradient boosted trees are challenger models. [16–20] Candidate boosted trees use depths one or two, twenty to one hundred trees, learning rates of 0.03 or 0.1, and minimum leaf sizes of ten or twenty; these are tuning options, not an asserted optimal configuration. [16–20]

All transformations are fitted inside the training fold. [19,20,26–28] Continuous predictors are scaled for regression, missing numeric values receive training medians plus missingness indicators, and target values are never imputed. [19,20,26–28] Structural zeros are distinguished from absent reports; logarithmic targets require strictly positive documented endpoints. [19,20,26–28] Winsorization, if used, takes thresholds from training observations only, and the original untrimmed outcomes remain the primary evaluation target. [19,20,26–28]

Outer evaluation uses decision origins 2016, 2017, and 2018, each with only previously available labels. [26–28] Inner selection follows earlier chronological origins and never accesses outer outcomes. [26–28] A country exclusion test removes all observations from a destination during training and evaluates transfer to that destination separately. [26–28] This secondary test asks a different question from forecasting another year in an already represented market, so its metrics are not pooled with the temporal test. [26–28]

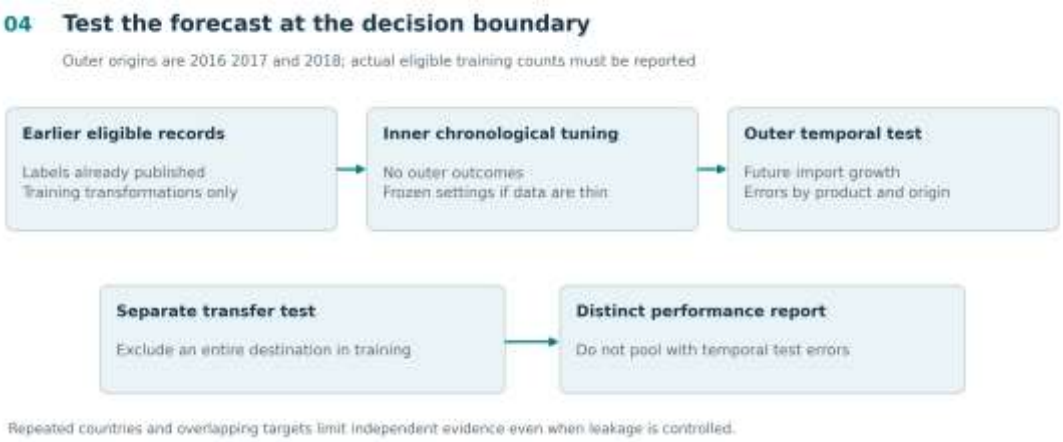


Figure 4. Nested Evaluation Distinguishing Historical Forecast Testing From Transfer to an Unseen Destination [26–28]

Small sample rules prevent overclaiming. [19,20,26–28] If an outer training set contains fewer than one hundred eligible rows, ensembles remain exploratory and the regularized benchmark is primary. [19,20,26–28] If an inner split contains too few mature outcomes for reliable tuning, the model uses frozen settings instead of optimizing on the outer test. [19,20,26–28] The report prints actual row counts, unique destinations, product coverage, and outcome availability for every fold. [19,20,26–28]

Table 5. Validation Outputs and Interpretation [19–28]

Check	Reported quantity	Interpretation
Temporal test	MAE by product and origin	Forecast performance
Country exclusion	Separate destination errors	Geographic transfer
Shortlist overlap	Shared selections among three	Ranking agreement
Explanation stability	Signs and leading contributions	Model dependence
Data maturity	Eligible rows in each fold	Historical feasibility

Primary accuracy is mean absolute error in log growth, averaged first within product and decision origin and then across origins. [19,26–28] Secondary diagnostics include root mean square error, directional accuracy, and Spearman rank correlation across

destinations. [19,26–28] Top three overlap measures shortlist agreement; its denominator is fixed at three only when three eligible candidates exist. [19,26–28] Financial regret is not inferred from import growth because the dataset does not observe each company's counterfactual profits. [19,26–28]

Model explanations use training only reference observations and identify the explained output as predicted log import growth. [21–25] For tree models, SHAP values satisfy prediction equals baseline plus feature contributions under the declared background and dependence convention. [21–25] Correlated tariff and sourcing variables are also summarized as groups. [21–25] Accumulated local effects provide a complementary diagnostic within supported predictor ranges, and explanations outside those ranges are flagged rather than treated as reliable marginal responses. [21–25]

Explanation stability is assessed across fitted folds and country resamples using sign agreement and overlap among the three largest absolute contributions. [21–23,26–28] Because countries share global shocks, a country bootstrap alone cannot capture all uncertainty. [21–23,26–28] The report therefore presents year specific performance and sensitivity to common shock years instead of interpreting a narrow interval as proof of generalizability. [21–23,26–28] With only three outer origins, comparisons remain descriptive unless additional independent evaluation periods are available. [21–23,26–28]

6. Reproducible Tariff and Currency Demonstration

The numerical demonstration considers four anonymous destinations, labeled A through D, with one identical product and a dollar import cost of 100 per unit. [6–8,30] Net local receipts, converted at the initial exchange rate, equal 160, 155, 150, and 145 dollars respectively. [6–8,30] Baseline tariff rates are five, two, zero, and ten percent. [6–8,30] These quantities are declared scenario inputs; they are neither observed national tariffs nor confidential firm outcomes. [6–8,30]

Let q denote those baseline dollar equivalent receipts, c the dollar customs cost, tau the tariff rate, and delta the proportional increase in local currency units per dollar. [6–8,30] The stress margin is m = q divided by one plus delta minus c multiplied by one plus tau plus a tariff increment. [6–8,30] Net local receipts are fixed, dollar costs are unhedged, and tariffs apply to the stated cost base; other taxes and fixed launch expenses are excluded. [6–8,30]

For clarity, the implemented expression is m = q/(1 + delta) − c(1 + tau + a), where a is an additive tariff change in decimal units. [6–8,30] A five percentage point increase means a equals 0.05; it does not mean multiplying tau by 1.05. [6–8,30] A ten percent local depreciation means delta equals 0.10 under this exchange rate convention. [6–8,30] The formula gives a conditional accounting result, not an econometric estimate of tariff incidence or consumer price adjustment. [6–8,30]

Table 6. Declared Numerical Inputs for Analytical Demonstration [6–8,30]

Destination	Receipts Q Dollars	Cost C Dollars	Tariff Tau
A	160	100	0.05
B	155	100	0.02
C	150	100	0.00
D	145	100	0.10

Origin share vectors are A: 0.75, 0.20, 0.05; B: 0.40, 0.30, 0.20, 0.10; C: 0.50, 0.25, 0.25; and D: 0.40, 0.30, 0.15, 0.15. [10,12,14] Squared share sums equal 0.605, 0.300, 0.375, and 0.295. [10,12,14] These examples describe origin exposure only. [10,12,14] No brand shares, disruption probabilities, or estimates of market power are inferred from them. [10,12,14]

Baseline margins are 55, 53, 50, and 35 dollars. [29,30] Combining a five percentage point tariff increase with ten percent local depreciation produces margins of 35.45, 33.91, 31.36, and 16.82 dollars after rounding. [29,30] At a concentration ceiling of 0.40 and a stressed margin floor of 25 dollars, A fails concentration and D fails margin. [29,30] B and C remain feasible, with B providing the higher stressed contribution. [29,30]

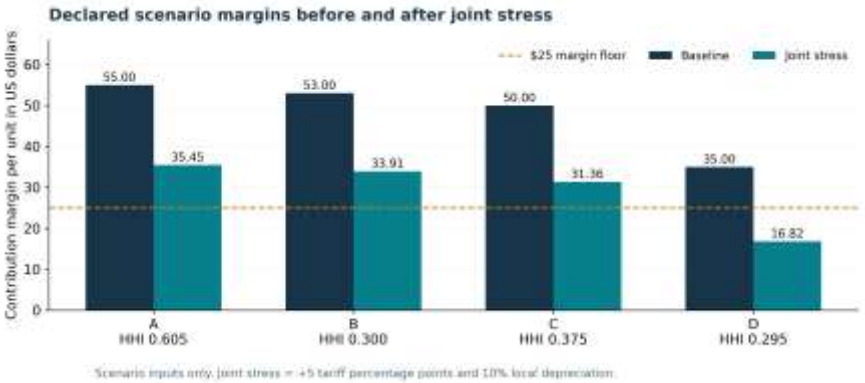


Figure 5. Calculated Baseline and Joint Stress Margins for the Four Declared Scenario Destinations [6–8,30]

The mechanical result shows why choosing the highest baseline margin is not equivalent to choosing the highest feasible stressed margin. [1,2,19,30] A leads before constraints, but B leads within the feasible set. [1,2,19,30] This result does not establish that B offers the strongest import opportunity or the greatest total profit. [1,2,19,30] Unit contribution omits launch scale and fixed costs, while the separate predictive module has not been fitted to empirical trade records in this manuscript. [1,2,19,30]

Table 7. Calculated Stress Results and Feasibility Under Declared Thresholds [29,30]

Destination	Baseline Dollars	Stress Dollars	HHI	Feasible
A	55.00	35.45	0.605	Fails concentration
B	53.00	33.91	0.300	Yes
C	50.00	31.36	0.375	Yes
D	35.00	16.82	0.295	Fails margin

For B, a three by three sensitivity grid varies tariff increments across zero, five, and ten percentage points and depreciation across zero, ten, and twenty percent. [8,30] Margins range from 53 dollars without either shock to 17.17 dollars at the largest combined shock. [8,30] At twenty percent depreciation and a five point tariff increase, B yields 22.17 dollars and fails the stated margin floor. [8,30] Feasibility therefore depends on the declared scenario set rather than a permanent destination label. [8,30]



Figure 6. Sensitivity Grid for Destination B with the 25 Dollar Feasibility Threshold Stated Explicitly [8,30]

The accounting model permits a further exact check: for a fixed tariff scenario, the largest tolerable depreciation solves delta = q/[c(1 + tau + a) + b] minus one, where b is the required contribution floor. [8,30] For B with a five point tariff increment and b equal to 25, the boundary is 17.42 percent. [8,30] This boundary assumes unchanged local receipts and does not predict the likelihood of that currency movement. [8,30]

7. Decision Governance and Reporting

A research implementation must publish the data dictionary, release ledger, excluded mappings, model settings, and numerical scenario code. [10,11,21,25,30] An explanation record names the destination, product, decision date, customs origin, data vintage, prediction baseline, leading predictive contributions, and binding constraints. [10,11,21,25,30] If customs eligibility is unresolved, the system returns "review required" instead of substituting a favorable tariff. [10,11,21,25,30] Missing information is a decision state rather than evidence that risk is absent. [10,11,21,25,30]

Table 8. Minimum Audit Record for a Reviewable Shortlist [10,11,21,25,30]

Record	Required Content
Identity	Firm product destination and customs origin
Data	Decision date releases and exclusions
Prediction	Model version baseline and contributions
Scenario	Currency tariff pricing and cost assumptions
Constraints	Thresholds failures and binding gate
Review	Owner override rationale and approval

Managers should inspect the forecast and the accounting explanation separately. [6,7,21–25,30] A positive predictive attribution for concentration could reflect established sourcing networks or omitted market characteristics; it does not cancel a managerial concentration ceiling. [6,7,21–25,30] Likewise, a low historical tariff cannot override a scenario duty entered by the customs specialist. [6,7,21–25,30] The system records overrides, author, rationale, and affected assumptions so that an attractive recommendation cannot conceal an undocumented change in the decision rule. [6,7,21–25,30]



Figure 7. Audit Record Linking an Explained Forecast to Independently Documented Constraints and Review Responsibility [21–25,30]

The empirical acceptance rule requires improvement over persistence on the prespecified temporal metric with at most 5% deterioration in country exclusion MAE. [19,25–28,30] If the ensemble fails, the interpretable model remains the operational benchmark. [19,25–28,30] If both perform poorly or the feasible shortlist changes under modest supported assumptions, the output is a conditional comparison requiring further investigation. [19,25–28,30] Abstention is preferable to translating uncertain aggregate patterns into a confident firm level entry recommendation. [19,25–28,30]

8. Limitations and Conclusion

The principal limitation is the gap between trade aggregates and company outcomes. [4,5,10,14,15] Imports include intermediate and business demand, reexports, and changes in product mix; the selected customs categories cannot isolate every consumer use. [4,5,10,14,15] A subsequent firm study would require verified entry dates, planned origins, distribution costs, pricing, and postlaunch contribution records, including failed entries. [4,5,10,14,15] Without these data, the framework evaluates market opportunity and declared exposure, not realized entry success. [4,5,10,14,15]

A second limitation is the small and dependent panel. [19,21–28] The maximum sample contains repeated destinations, related products, and overlapping two year targets. [19,21–28] Chronological testing, release eligibility, and country exclusion address different leakage risks but do not manufacture independent evidence. [19,21–28] Finally, explanations remain conditional on feature definitions and the fitted model; neither a stable attribution nor an appealing stress chart proves a causal policy mechanism. [19,21–28]

This paper offers a specific route from observable product market signals to an auditable entry shortlist. [1,2,6–8,19,30] Its analytical demonstration establishes that tariff and currency stress combined with an origin concentration constraint can reverse the preferred feasible destination relative to baseline margins. [1,2,6–8,19,30] The broader predictive claims remain testable through the documented validation plan. [1,2,6–8,19,30] Final market entry requires company evidence and explicit acceptance of the assumptions that generated the shortlist. [1,2,6–8,19,30]

References

[1] N. Papadopoulos, J. E. Denis. Inventory taxonomy and assessment of methods for international market selection. International Marketing Review 5, 38–51, 1988. doi:10.1108/eb008357.
[2] S. T. Cavusgil, T. Kiyak, S. Yeniyurt. Complementary approaches to preliminary foreign market opportunity assessment. Industrial Marketing Management 33, 607–617, 2004. doi:10.1016/j.indmarman.2003.10.005.
[3] J. Johanson, J. E. Vahlne. The internationalization process of the firm. Journal of International Business Studies 8, 23–32, 1977. doi:10.1057/palgrave.jibs.8490676.
[4] M. J. Melitz. The impact of trade on intra industry reallocations and aggregate industry productivity. Econometrica 71, 1695–1725, 2003. doi:10.1111/1468-0262.00467.
[5] E. Helpman, M. Melitz, Y. Rubinstein. Estimating trade flows trading partners and trading volumes. Quarterly Journal of Economics 123, 441–487, 2008. doi:10.1162/qjec.2008.123.2.441.
[6] M. Amiti, S. J. Redding, D. E. Weinstein. The impact of the 2018 tariffs on prices and welfare. Journal of Economic Perspectives 33, 187–210, 2019. doi:10.1257/jep.33.4.187.
[7] P. D. Fajgelbaum, P. K. Goldberg, P. J. Kennedy, A. K. Khandelwal. The return to protectionism. Quarterly Journal of Economics 135, 1–55, 2020. doi:10.1093/qje/qjz036.
[8] G. Gopinath et al. Dominant currency paradigm. American Economic Review 110, 677–719, 2020. doi:10.1257/aer.20171201.
[9] K. Handley, N. Limão. Policy uncertainty trade and welfare theory and evidence for China and the United States. American Economic Review 107, 2731–2783, 2017. doi:10.1257/aer.20141419.
[10] United Nations. International Merchandise Trade Statistics Concepts and Definitions 2010. Statistical Papers Series M No. 52 Revision 3. New York, 2011. https://unstats.un.org/unsd/trade/eg-imts/IMTS%202010%20%28English%29.pdf.
[11] World Bank. World Integrated Trade Solution User Manual. Version 2.01. Washington DC, January 2011. https://wits.worldbank.org/data/public/WITS_User_Manual.pdf.
[12] G. Gaulier, S. Zignago. BACI International Trade Database at the Product Level The 1994–2007 Version. CEPII Working Paper 2010–23, 2010. https://www.cepii.fr/PDF_PUB/wp/2010/wp2010-23.pdf.
[13] M. Klau, S. S. Fung. The new BIS effective exchange rate indices. BIS Quarterly Review, March, 51–65, 2006. https://www.bis.org/publ/qtrpdf/r_qt0603e.pdf.
[14] World Bank. World Development Report 2020 Trading for Development in the Age of Global Value Chains. Washington DC, 2020. doi:10.1596/978-1-4648-1457-0.
[15] World Customs Organization. Harmonized Commodity Description and Coding System. 2012 edition. Brussels, 2012. https://www.wcoomd.org/en/topics/nomenclature/instrument-and-tools/hs_nomenclature_2012/hs_nomenclature_table_2012.aspx.
[16] L. Breiman. Random forests. Machine Learning 45, 5–32, 2001. doi:10.1023/A:1010933404324.
[17] J. H. Friedman. Greedy function approximation a gradient boosting machine. Annals of Statistics 29, 1189–1232, 2001. doi:10.1214/aos/1013203451.
[18] T. Chen, C. Guestrin. XGBoost a scalable tree boosting system. KDD, 785–794, 2016. doi:10.1145/2939672.2939785.

[19] T. Hastie, R. Tibshirani, J. Friedman. The Elements of Statistical Learning Data Mining Inference and Prediction. Second edition. Springer, 2009. doi:10.1007/978-0-387-84858-7.
[20] H. Zou, T. Hastie. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society B 67, 301–320, 2005. doi:10.1111/j.1467-9868.2005.00503.x.
[21] S. M. Lundberg, S. I. Lee. A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems 30, 4765–4774, 2017. https://arxiv.org/abs/1705.07874.
[22] S. M. Lundberg et al. From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence 2, 56–67, 2020. doi:10.1038/s42256-019-0138-9.
[23] D. W. Apley, J. Zhu. Visualizing the effects of predictor variables in black box supervised learning models. Journal of the Royal Statistical Society B 82, 1059–1086, 2020. doi:10.1111/rssb.12377.
[24] M. T. Ribeiro, S. Singh, C. Guestrin. Why should I trust you explaining the predictions of any classifier. KDD, 1135–1144, 2016. doi:10.1145/2939672.2939778.
[25] C. Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence 1, 206–215, 2019. doi:10.1038/s42256-019-0048-x.
[26] D. R. Roberts et al. Cross validation strategies for data with temporal spatial hierarchical or phylogenetic structure. Ecography 40, 913–929, 2017. doi:10.1111/ecog.02881.
[27] C. Bergmeir, R. J. Hyndman, B. Koo. A note on the validity of cross validation for evaluating autoregressive time series prediction. Computational Statistics and Data Analysis 120, 70–83, 2018. doi:10.1016/j.csda.2017.11.003.
[28] L. J. Tashman. Out of sample tests of forecasting accuracy an analysis and review. International Journal of Forecasting 16, 437–450, 2000. doi:10.1016/S0169-2070(00)00065-0.
[29] R. T. Rockafellar, S. Uryasev. Optimization of conditional value at risk. Journal of Risk 2, 21–41, 2000. doi:10.21314/jor.2000.038.
[30] A. Ben-Tal, L. El Ghaoui, A. Nemirovski. Robust Optimization. Princeton University Press, 2009. doi:10.1515/9781400831050.