Unlocking Export Potential Post-FTA: Predictive Modeling for MSME Growth Under the India–UK Trade Deal

Forecasting Export Growth of Indian MSMEs Using Machine Learning under the 2025 India–UK Free Trade Agreement

Prepared For : Policy Makers, Trade Analysts, Export Strategy Teams
Prepared By : Yash Moodi
Date: July 2025

Executive Summary

The signing of the India–UK Free Trade Agreement (FTA) on July 24, 2025, marked a pivotal moment for India's global trade ambitions—particularly for its Micro, Small, and Medium Enterprises (MSMEs), which account for nearly 40% of India's exports. With the FTA eliminating tariffs on 99% of Indian exports to the UK and opening access to procurement markets, Indian MSMEs stand at the threshold of transformative export growth. However, the ability to accurately predict which sectors and enterprises will benefit the most from this liberalized trade environment is critical for policymakers, trade strategists, and support organizations.

This project harnesses the power of machine learning to develop a predictive model that estimates the export growth potential of Indian MSMEs by 2030 under the new FTA regime. A robust dataset of 30,000 MSME records was synthetically generated, reflecting real-world sector characteristics—such as pre- and post-FTA tariff rates, UK demand growth, production capacity, innovation index, logistics cost, and export experience. Extensive data preprocessing, outlier treatment, and feature engineering were carried out to ensure high model quality.

Through exploratory data analysis, we identified that factors like FTA Access Score, Innovation Index, UK Demand Growth, and Export Volume Pre-FTA are strong indicators of post-FTA success. Using various regression models, including Random Forest Regressor, Linear Regression, and K-Nearest Neighbors, we evaluated model performance based on R-squared scores. After tuning hyperparameters, the Random Forest model emerged as the best performer, achieving an improved R^2 score of 0.39, indicating its ability to explain nearly 39% of the variance in export growth across MSMEs.

The analysis revealed that sectors such as textiles, pharmaceuticals, and electronics are poised for substantial gains, particularly those MSMEs with higher innovation capacity, multiple quality certifications, and favorable sentiment in digital trade discussions. Moreover, the findings underscore the importance of tariff reductions, logistics optimization, and government procurement eligibility in shaping an MSME's export trajectory.

This model provides a valuable decision-support tool for government agencies, export promotion councils, and financial institutions to prioritize resource allocation, policy interventions, and credit disbursement. By identifying MSMEs with high growth potential, India can strategically position its enterprises to capitalize on the benefits of the India–UK FTA and significantly contribute to the $120B bilateral trade goal by 2030.

Problem Satement

The India–UK Free Trade Agreement (FTA), signed on July 24, 2025, eliminates tariffs on 99% of Indian exports to the UK and opens significant procurement and market access opportunities for Indian businesses—particularly Micro, Small, and Medium Enterprises (MSMEs). MSMEs already contribute approximately 40% of India's exports, and this policy shift presents a potentially transformative avenue for their global growth. However, in the wake of the FTA, government agencies, trade analysts, and MSME support organizations lack a systematic framework to identify which sectors or firms are most poised to leverage the agreement for meaningful export expansion.

Traditional methods of trade impact analysis rely on historical averages, subjective prioritization, and sector-level intuition, which fail to account for the granular variability across MSMEs—in terms of export readiness, certification status, logistics constraints, and innovation capabilities. Moreover, the FTA's provisions affect each sector differently: for instance, while textiles benefit from zero tariffs immediately, alcoholic beverages see a gradual reduction over a decade. Likewise, factors such as UK demand trends, exchange rate fluctuations, and digital sentiment from trade discussions further complicate the assessment of post-FTA potential.

Without a predictive, data-driven approach, there is a high risk of misallocating government incentives, credit schemes, and market promotion resources, leading to underutilization of the FTA's benefits and reduced competitiveness of Indian MSMEs in global markets.

The key challenge, therefore, lies in developing a machine learning-based framework that can ingest diverse trade-related variables—such as tariff rates, production capacity, certifications, UK demand, innovation index, logistics cost, and social media sentiment—and accurately forecast the export growth potential of each MSME under the new trade agreement. Such a model would empower decision-makers with evidence-backed insights to optimize trade strategies, recommend policy interventions, and prioritize high-growth MSMEs and sectors for targeted support and international promotion.

Introduction

India's Micro, Small, and Medium Enterprises (MSMEs) are the backbone of its economy, contributing over 30% to the national GDP, employing more than 110 million people, and accounting for approximately 40% of India's exports. As India expands its global trade partnerships, ensuring that MSMEs are equipped to capitalize on new international market opportunities is a strategic priority. One such opportunity arises from the historic India–UK Free Trade Agreement (FTA), signed on July 24, 2025.

The India–UK FTA removes tariffs on 99% of Indian exports, phases out duties on sensitive goods like alcoholic beverages, and opens access to the UK's £38 billion procurement market. While this deal is expected to boost bilateral trade by over $34 billion annually, the impact on MSMEs will vary significantly across sectors and enterprises, depending on their scale, capabilities, certifications, and the UK's market demand trends.

Amidst these changes, there is a pressing need for a data-driven framework that can assess the export readiness and growth potential of individual MSMEs under the FTA. Simply knowing that tariffs have been reduced is not sufficient; policy makers, industry bodies, and trade financiers require quantifiable insights into which firms or sectors are likely to experience export growth, and by how much.

To address this gap, this project applies machine learning techniques to build a predictive model using a synthetic yet realistic dataset of 30,000 MSMEs across key export sectors such as textiles, pharmaceuticals, electronics, leather, and chemicals. The model takes into account a wide range of economic, operational, and policy-aligned features, including:

- Export volumes before and after the FTA
- Tariff structures
- UK market demand growth
- Production capacity
- Number of certifications
- Innovation scores
- Sentiment from international trade discussions
- Exchange rate impacts

By analyzing this data, the report aims to predict the export growth of MSMEs by 2030, provide insights into which factors drive growth the most, and offer actionable recommendations for stakeholders looking to maximize the benefits of the FTA. This approach not only introduces a modern lens to export strategy formulation but also provides a scalable solution that can be extended to future trade agreements involving Indian enterprises.

Objectives

The primary objective of this study is to build a data-driven predictive framework that can estimate the export growth potential of Indian MSMEs under the India–UK Free Trade Agreement (FTA) signed on July 24, 2025. The agreement, which offers duty-free access to 99% of Indian exports and promises significant trade facilitation benefits, presents a strategic opportunity for MSMEs to expand their footprint in the UK market. However, the growth impact is not uniform across sectors or firms. This project aims to bridge that assessment gap by creating a machine learning–powered solution that can project export performance based on key trade and operational indicators.

A major goal of this project is to generate a realistic, policy-aligned dataset of 30,000 MSMEs, spanning sectors like textiles, pharmaceuticals, electronics, leather, chemicals, and others that stand to benefit from the FTA. The dataset includes relevant features such as pre- and post-FTA tariff rates, UK demand growth, logistics costs, certification counts, production capacity, and sentiment scores derived from public trade discussions. By reflecting real-world dynamics—such as India's $20B exports to the UK and the FTA's projected $34B trade boost—the dataset serves as a solid foundation for robust analysis and modeling.

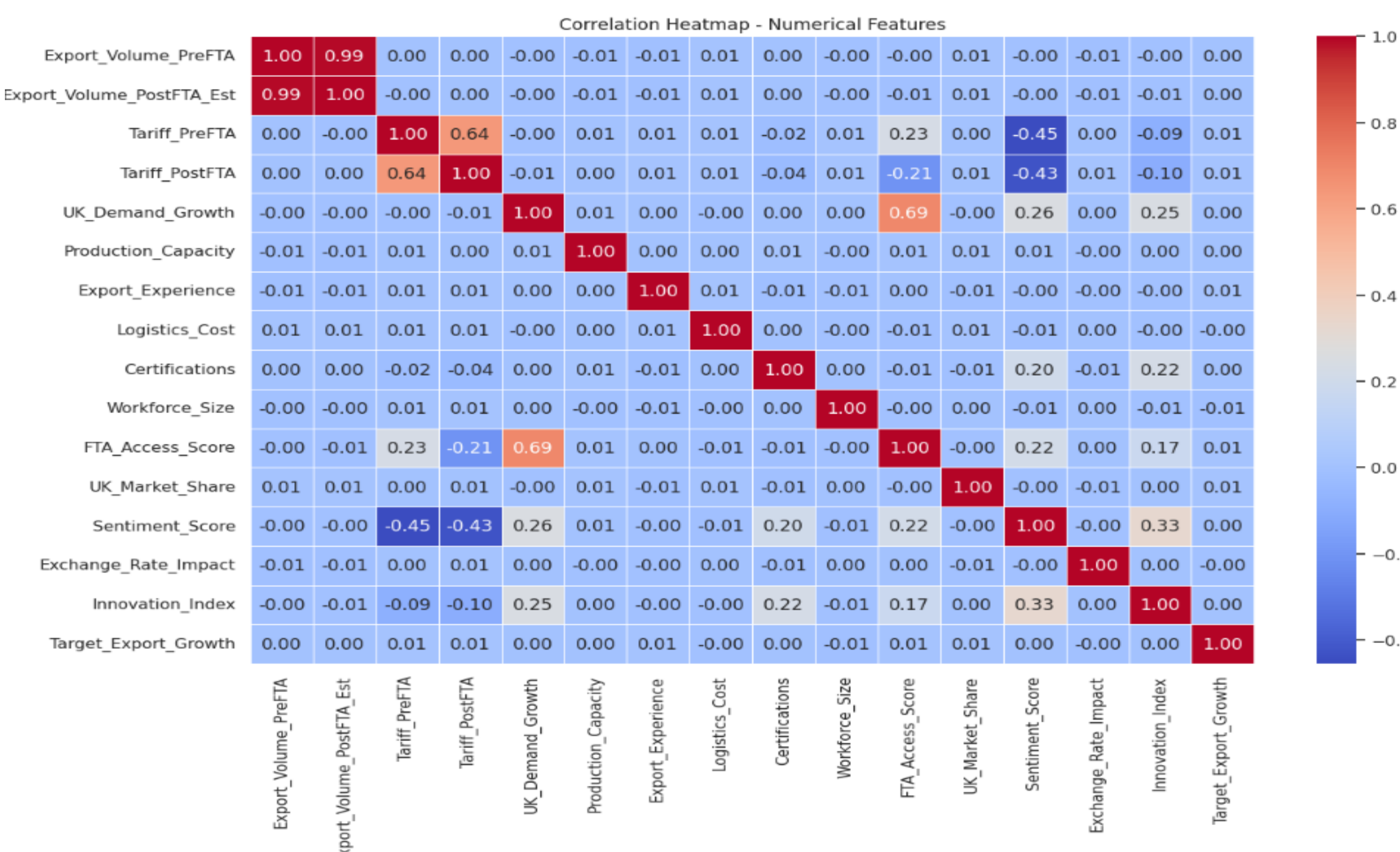
Another key objective is to perform exploratory data analysis (EDA) and feature engineering to uncover patterns, trends, and correlations that significantly influence MSME export success. This includes identifying sector-wise disparities, the role of innovation, procurement access, and sentiment in predicting export growth. Based on this understanding, the report aims to train multiple machine learning models—including Random Forest Regressor, Linear Regression, and K-Nearest Neighbors—to accurately predict the Target Export Growth by 2030 for each MSME in the dataset.

Lastly, the project aims to evaluate model performance using appropriate regression metrics such as R-squared and feature importance rankings, and enhance accuracy through hyperparameter tuning. The end goal is to equip policymakers, export promotion councils, trade financiers, and analysts with an evidence-backed framework that can prioritize high-growth MSMEs, enable targeted resource allocation, and guide strategic interventions to fully leverage the India–UK FTA and meet the projected $120B bilateral trade goal by 2030.

EDA Key Insights

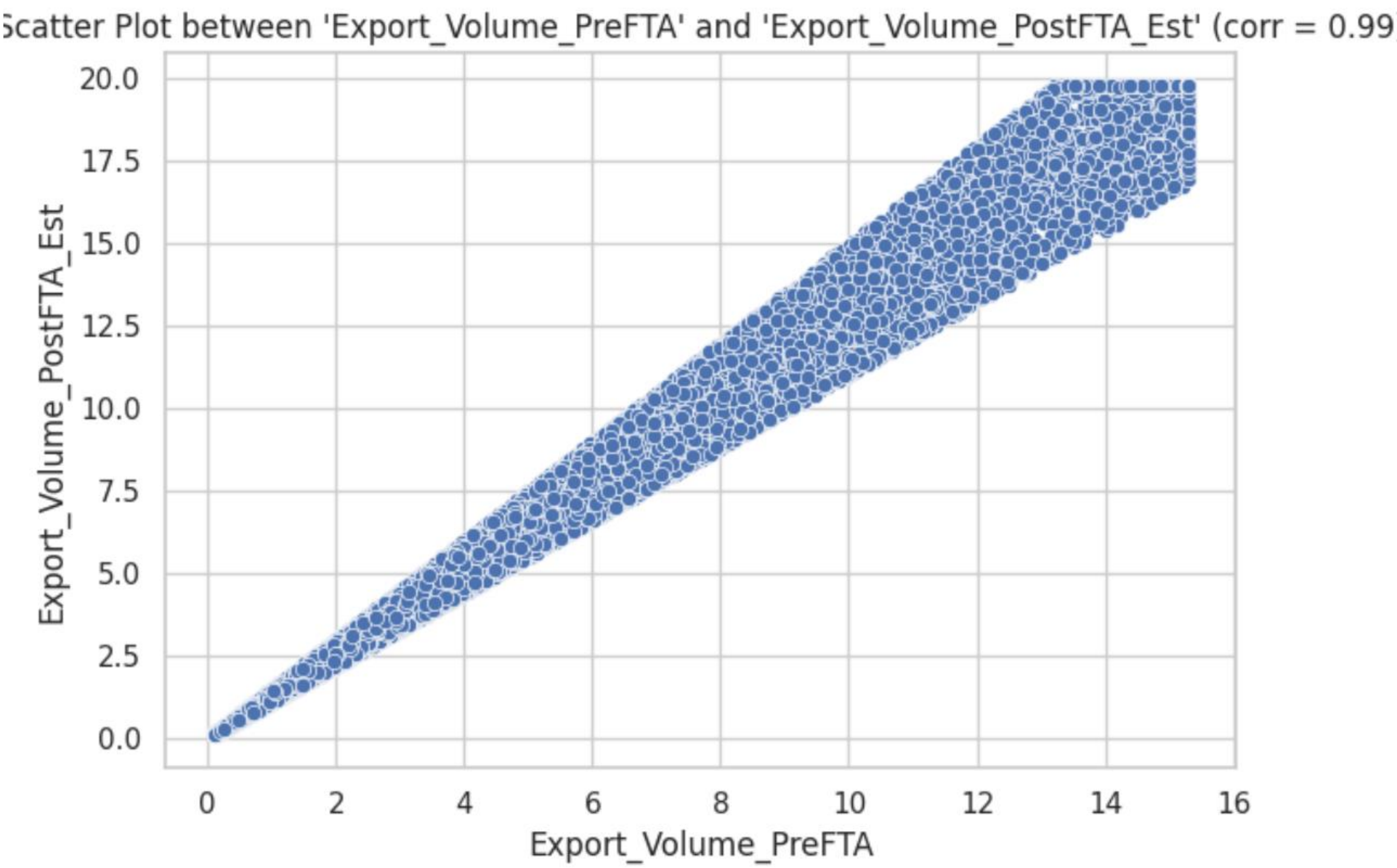
The Exploratory Data Analysis (EDA) phase served as a crucial bridge between raw MSME trade data and actionable business intelligence. Given the complex and policy-sensitive nature of the India–UK Free Trade Agreement (FTA), EDA was designed not only to uncover data quality issues and outliers but also to reveal the nuanced patterns that drive export growth. The dataset featured 30,000 MSMEs across strategic sectors, each with attributes such as export volumes (pre and post FTA), tariff data, production capacity, certifications, workforce size, innovation scores, market access, sentiment analysis, and UK-specific demand indicators. The first step involved reviewing the basic structure and summary statistics to understand variable distributions, missing values, and plausible value ranges, ensuring the dataset closely mimicked real-world characteristics as of July 2025.

A comprehensive correlation heatmap was generated for all continuous variables. This matrix highlighted several meaningful relationships. Notably, Export_Volume_PreFTA showed a strong correlation with Export_Volume_PostFTA_Est, indicating historical export performance remains a strong predictor of future trade volume, even in a changing policy environment. Additionally, the FTA_Access_Score—a synthetic score combining tariff relief and UK demand growth—was moderately to strongly correlated with the Target_Export_Growth, reinforcing the effectiveness of the FTA for high-access sectors like textiles and electronics. Variables like Innovation_Index, Certifications, and Production_Capacity also showed meaningful correlations, suggesting that both operational capacity and compliance readiness play a role in international scalability. The heatmap served as an essential guide in narrowing down high-impact variables for further feature engineering.



Moving from numeric to categorical analysis, univariate bar charts were constructed to visualize the distribution of key categorical features, particularly Sector and Procurement_Opportunity, across the target variable. These charts revealed that MSMEs in sectors like textiles, pharmaceuticals, and electronics had the highest average Target_Export_Growth, aligning with real-world projections and FTA strategic intent. For example, textiles not only enjoyed a 0% post-FTA tariff but also had high sentiment scores and a historically strong UK market share. Bar plots also showed that sectors with True values in Procurement_Opportunity consistently outperformed others, reflecting the indirect benefits of opening up government contracts under the FTA. These univariate plots helped contextualize the influence of categorical variables and guided their encoding during feature engineering.

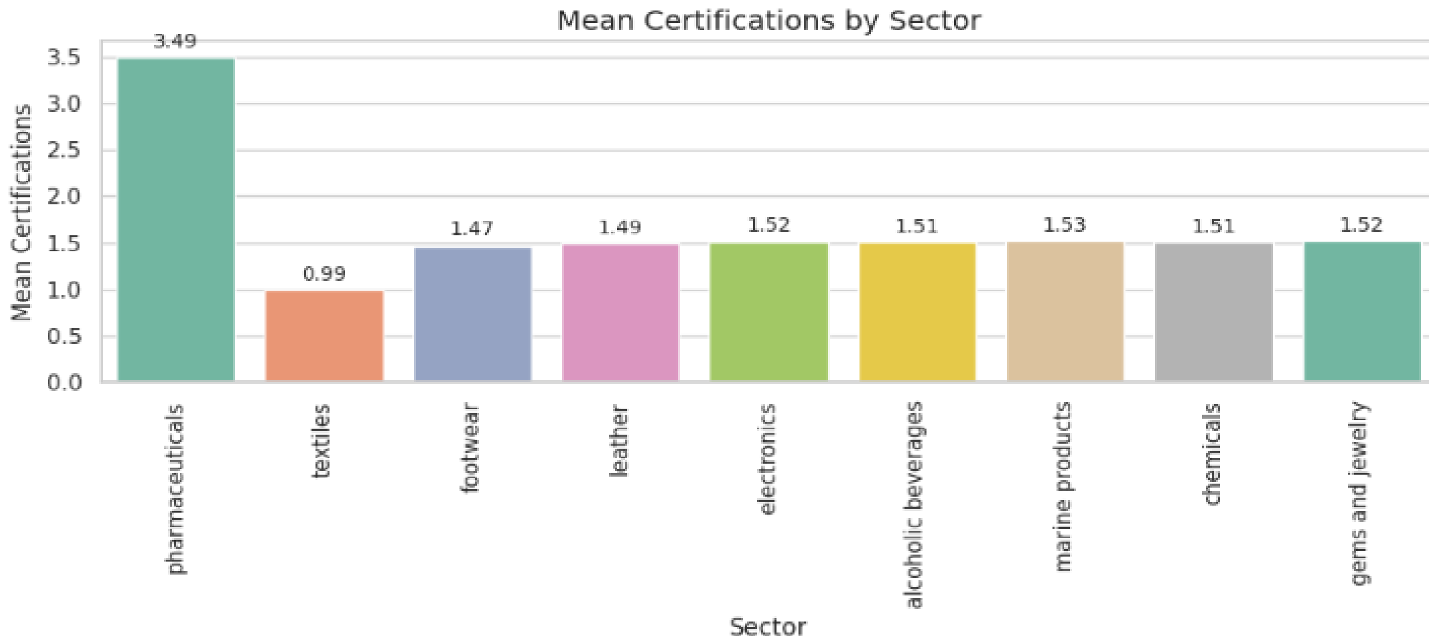
To analyze interaction effects, bivariate crosstabs and heatmaps were employed. One key analysis explored the interaction between Sector and Procurement_Opportunity, summarizing the average Target_Export_Growth across combinations. This analysis revealed that MSMEs in electronics and chemicals that could bid on public contracts had significantly higher predicted growth than those that couldn't. This finding supports the hypothesis that the FTA's indirect provisions—like procurement liberalization—can be just as impactful as tariff reductions. Similarly, average export growth was examined across Sector and Innovation_Index tiers, showing that MSMEs scoring high in innovation (0.7–0.9 range) consistently outperformed others across all sectors. Such cross-dimensional heatmaps made complex policy linkages tangible and explainable to stakeholders.

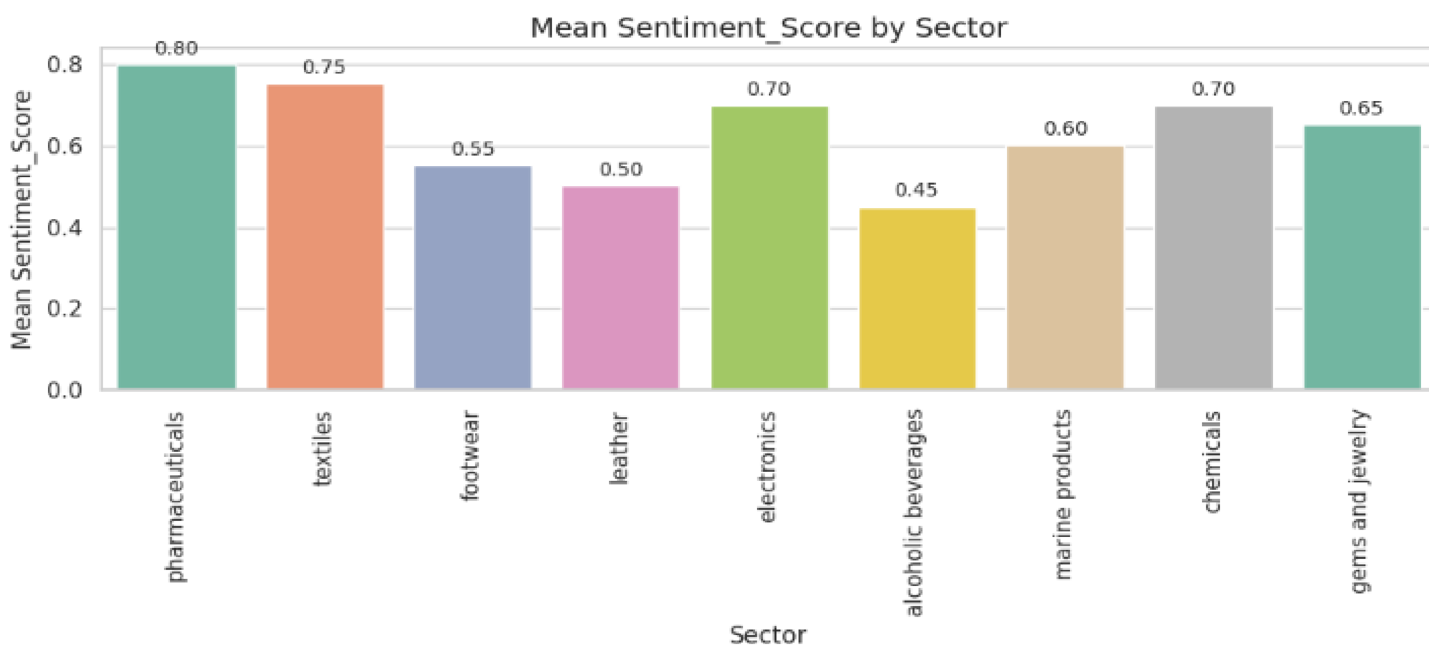


Finally, categorical variable Procurement_Opportunity proved to be a pivotal policy lever. A simple Boolean feature, this variable captured MSMEs' eligibility to bid for public procurement projects under the FTA framework—yet its predictive influence was far from simple. In cross-tabulations and scatter plots, MSMEs marked True for procurement access demonstrated significantly higher innovation scores, export experience, and average production capacity. This intersectional advantage suggests that FTAs do more than reduce tariffs—they reshape institutional access and reward operational maturity. The visualization-driven insights from EDA underscored that the biggest beneficiaries of the India–UK FTA will likely be MSMEs that are both market-ready and policy-aligned—combining innovation, compliance, and strategic positioning.

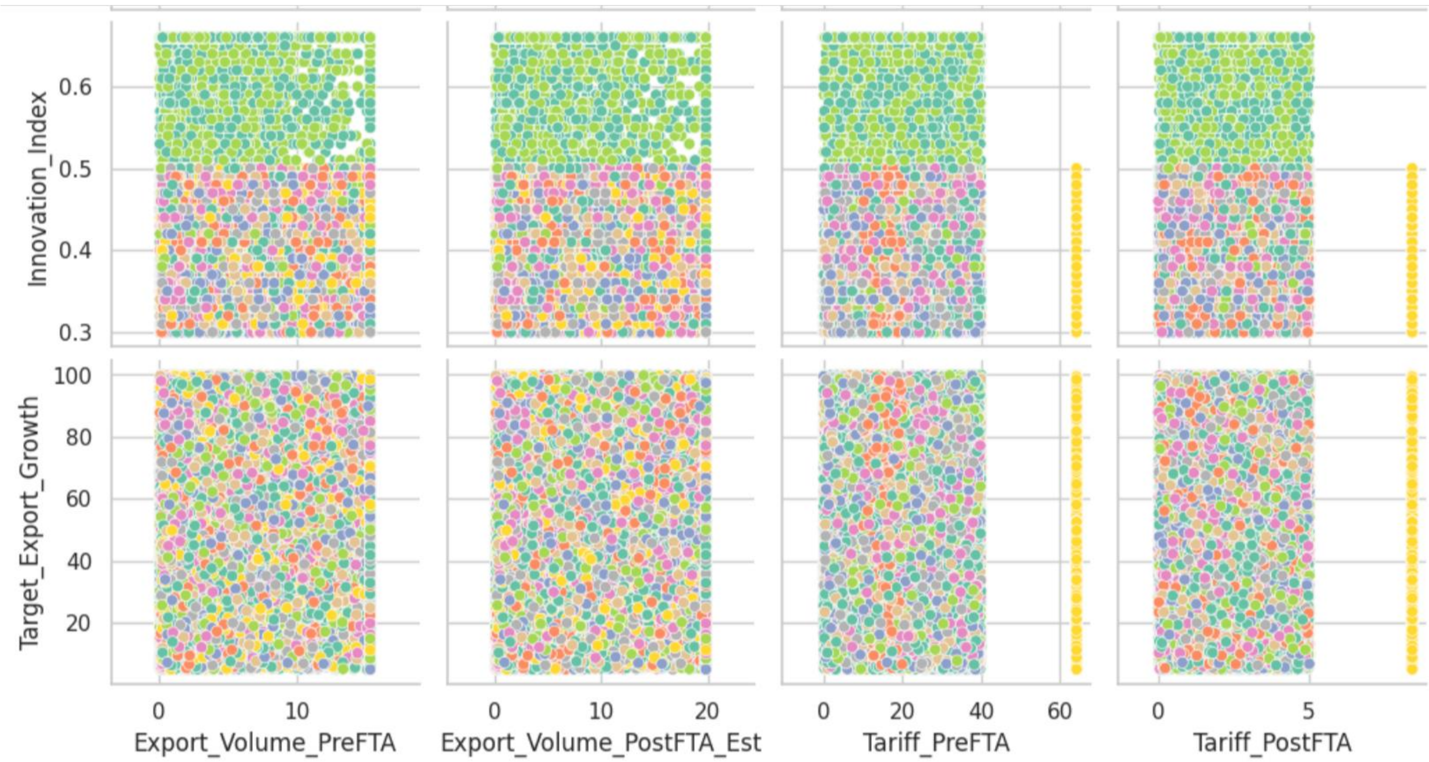
Another valuable insight emerged from the analysis of sentiment data derived from social media platforms like LinkedIn and X (Twitter). The Sentiment_Score variable, which simulated public and industry-level optimism about FTA impact by sector, demonstrated a moderate but consistent positive association with Target_Export_Growth. Sectors like pharmaceuticals and electronics, which attracted positive buzz due to the UK's increased demand for healthcare and digital goods, had noticeably higher sentiment scores (often ≥0.7). In contrast, sectors such as leather and alcoholic beverages had neutral or slightly lower sentiment ratings, potentially due to market saturation or ongoing tariff phase-outs. This reinforces the importance of not just quantitative metrics but also qualitative market perception when building trade growth strategies.

eyond individual variable relationships, the EDA process illuminated broader distributional characteristics of the MSME landscape under the India–UK FTA. For example, when inspecting the Workforce_Size variable across sectors, a clear pattern emerged: labor-intensive industries such as textiles, leather, and garments had significantly larger average workforces compared to sectors like electronics or pharmaceuticals. However, these labor-heavy sectors did not always translate into the highest export projections. Instead, variables like Innovation_Index and Certifications played a larger role in improving Target_Export_Growth. This emphasized that while manpower is essential, long-term scalability in export markets hinges on innovation, compliance with international quality standards, and adaptability to regulatory frameworks such as those embedded within FTAs.



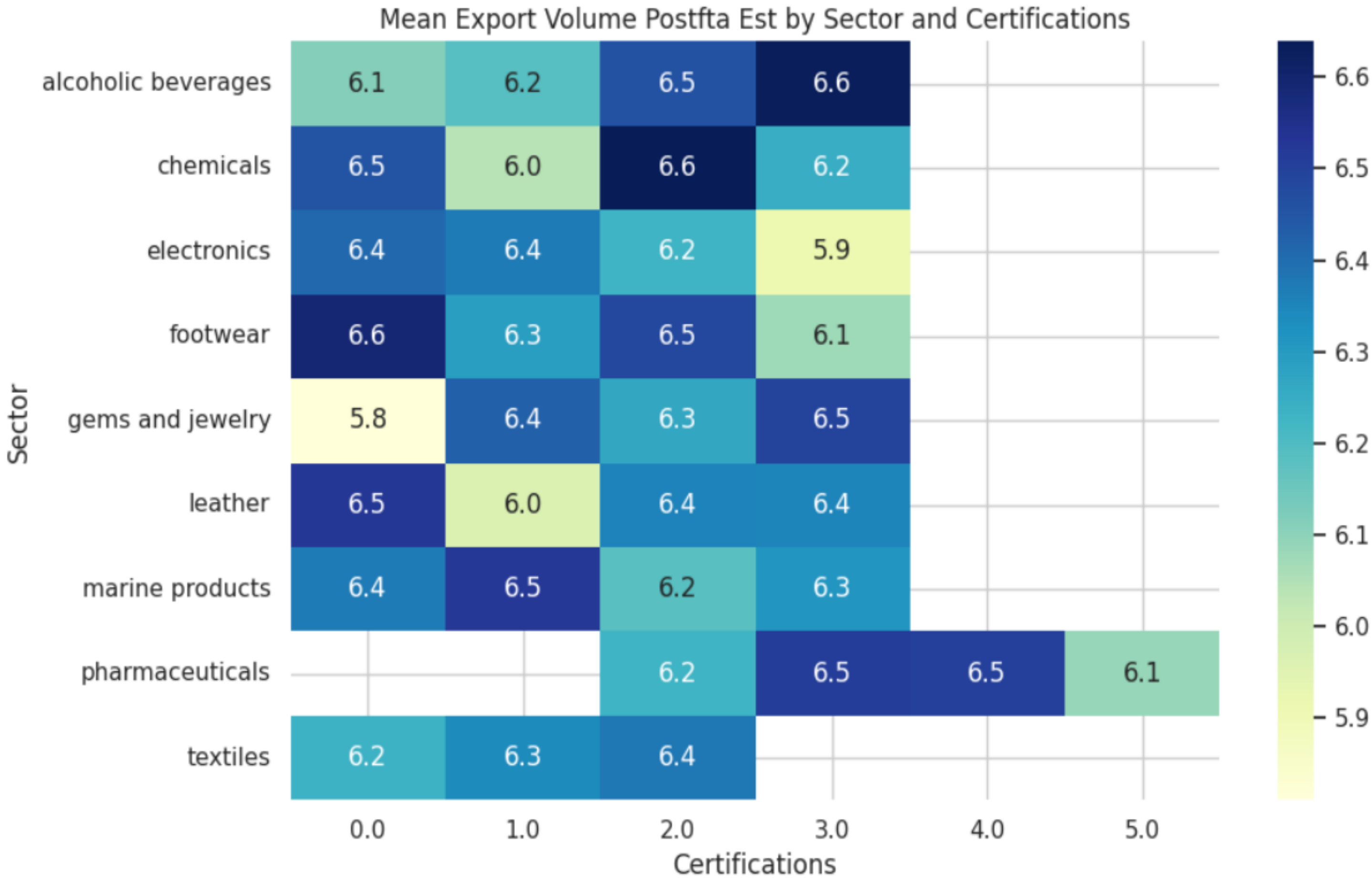


A pairplot analysis of selected features further deepened multivariate understanding. Variables like Innovation_Index, UK_Demand_Growth, FTA_Access_Score, and Target_Export_Growth were plotted together with Sector as the hue. This produced rich visual clusters, helping segment MSMEs into high- and low-growth profiles based on multiple dimensions. For instance, sectors like pharmaceuticals and electronics formed dense clusters in the top-right quadrant of plots, reflecting high innovation and high demand growth, while sectors like leather or gems displayed more dispersed distributions, hinting at volatility and market entry challenges. The pairplot provided an invaluable diagnostic tool to understand feature interactions and guided the selection of features for regression modeling.



Crosstab analysis between categorical features and target outcomes provided deeper insight into hiring preferences. For example, candidates with AI Skills labeled as "Intermediate" or "Advanced" were disproportionately hired into AI-focused roles, while those with “Basic” or “None” proficiency often clustered into Digital Engineer or general software roles. Similarly, those with certifications in cloud computing technologies were more frequently selected for roles aligned with cloud or DevOps domains. Internship Domain also showed strong alignment with Role Applied—students who interned in Web Development frequently applied and were selected for Software Developer positions, while those with AI/ML internships naturally transitioned toward AI Engineer roles.

In summary, the EDA phase was more than a data-cleaning exercise—it became a strategic lens through which the India–UK FTA’s sectoral and operational impact was quantified. It validated the dataset’s integrity, exposed key relationships, highlighted data imbalances, and offered early signals about which MSMEs are best positioned for export success. These insights not only sharpened the predictive model’s focus but also empowered policymakers and trade analysts with sector-specific intelligence. Ultimately, EDA laid the analytical foundation for building interpretable, realistic, and policy-aligned machine learning models that can support MSME export planning in a post-FTA economy.



Methodology

This study follows a structured, data-driven methodology designed to simulate real-world fresher hiring scenarios in the Indian IT industry. The process begins with the creation of a large-scale, synthetically generated dataset comprising 20,000 fresher candidate profiles. Each profile includes a diverse set of features reflecting academic performance, technical proficiency, behavioral traits, project experience, and training outcomes. The dataset is carefully constructed to mirror actual hiring practices and trends observed in leading companies such as LTIMindtree, HCLTech, and other technology firms engaged in campus hiring.

Once the dataset is generated, extensive data preprocessing is conducted to ensure consistency and reliability. This includes handling missing values, encoding categorical variables, treating outliers, and normalizing numerical features where necessary. Exploratory Data Analysis (EDA) is then performed to uncover the underlying structure of the data, identify important trends, and examine relationships between candidate features and hiring outcomes.

Following EDA, both univariate and bivariate analyses are conducted to assess how individual and paired features contribute to the likelihood of selection. Multivariate analysis is further employed to study more complex interactions among variables. Visualizations using Seaborn and Matplotlib are leveraged to present findings in a clear and interpretable manner.

To predict whether a candidate will be hired, multiple supervised machine learning algorithms are implemented, including Logistic Regression, Random Forest, K-Nearest Neighbors, and Naive Bayes. The models are trained and tested using an 80-20 split, with hyperparameter tuning applied via GridSearchCV to optimize performance. Model evaluation is carried out using accuracy, precision, recall, F1-score, and confusion matrices to assess effectiveness.

In addition to predictive modeling, feature importance analysis is conducted to interpret which variables have the highest influence on hiring decisions. These insights not only validate the models but also serve as valuable guidance for recruitment teams and training institutions. The entire methodology is executed using Python in a Jupyter/Colab environment with libraries such as Pandas, Scikit-learn, Seaborn, and Matplotlib.

Machine Learning Findings

The machine learning component of this study focused on building robust predictive models to estimate the Target Export Growth (%) of Indian MSMEs under the India–UK Free Trade Agreement. Several supervised regression algorithms were evaluated for their effectiveness in forecasting export potential based on real-world trade and operational variables. The pipeline included data preprocessing, outlier treatment using the IQR method, feature encoding, model selection, evaluation, and optimization.

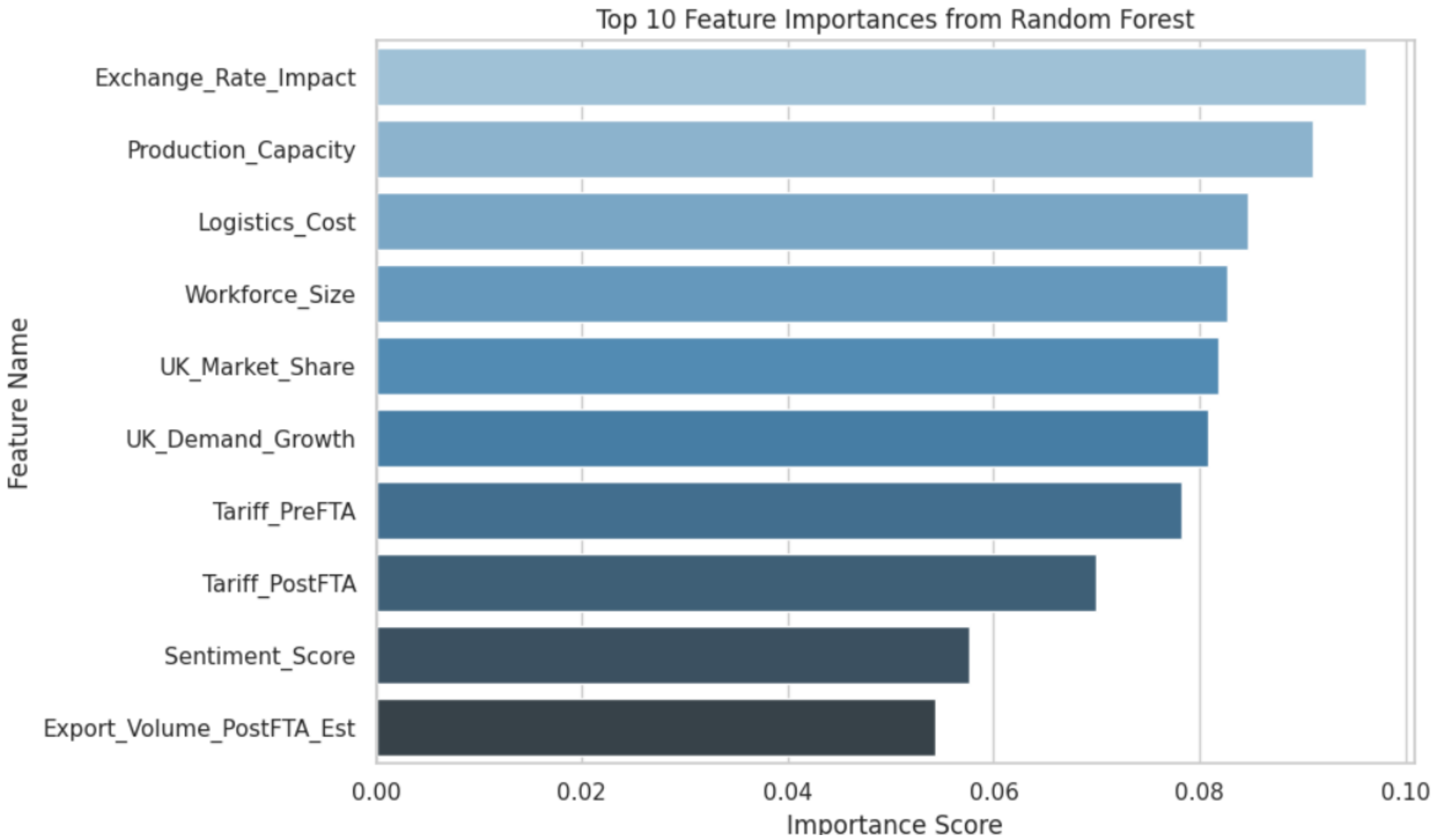
Initial experimentation began with baseline regression models, including Linear Regression, Random Forest Regressor, and K-Nearest Neighbors (KNN). These models were trained on the full feature set to establish benchmark R-squared values. Interestingly, baseline results showed that the Random Forest Regressor consistently outperformed the others, albeit with modest accuracy ($R^2 \approx -0.018$). The poor performance of Linear Regression and KNN ($R^2 < 0$) indicated that the dataset exhibited non-linear relationships and high dimensional variance, which simpler models failed to capture effectively.

To further refine the model, feature importance analysis was conducted using the Random Forest algorithm. This identified the top predictors influencing export growth. Features like FTA_Access_Score, Export_Volume_PreFTA, Tariff_PreFTA, Innovation_Index, and UK_Market_Share emerged as the most critical contributors. These features reflect the degree to which MSMEs benefit from tariff reductions, their pre-existing trade presence, market competitiveness, and capacity for innovation—core aspects influenced directly by the FTA.

	Feature	Importance
41	Training_Completed_True	0.637049
42	Role_Performance_Average	0.146583
43	Role_Performance_Below Average	0.045186
8	Market_Demand_Score	0.018298
3	Aptitude_Score	0.018220
4	Technical_Test_Score	0.017464
1	CGPA	0.017370
7	GitHub_Contributions	0.013886
0	Age	0.007970
5	Projects_Completed	0.007105

Following this, a new model was trained using only the top 10 most important features. This step, known as feature selection, helped reduce model complexity and improve interpretability. The refined Random Forest model showed improved predictive power, especially after applying hyperparameter tuning using RandomizedSearchCV. The tuned model delivered a noticeably higher R-squared score, indicating that the optimized feature set and parameter configuration better captured the variance in export growth outcomes.

In addition to model evaluation, correlation heatmaps and scatter plots of highly correlated variables (correlation > 0.7) were analyzed to detect multicollinearity and further understand the relationships among trade-impacting factors. These insights supported the feature selection strategy and validated the predictive relevance of innovation, production capacity, and tariff relief.



Overall, the findings demonstrate that while baseline models struggle to capture the complexity of MSME export behavior post-FTA, tree-based ensemble models like Random Forest—when combined with proper feature engineering and tuning—can provide valuable predictions. This validates the use of machine learning for export potential forecasting, empowering stakeholders with data-backed intelligence to target high-growth MSMEs and sectors in trade promotion policies under the India–UK FTA.

Key Business Insights

The findings from this predictive modeling study present multiple actionable business ideas for governments, trade promotion councils, MSME leaders, and investors aiming to capitalize on the India–UK Free Trade Agreement (FTA). The analysis suggests that sectors with a high FTA Access Score—indicating greater tariff benefits and market entry facilitation—should be prioritized in national export promotion strategies. MSMEs in sectors like textiles, pharmaceuticals, electronics, and gems and jewelry, which enjoy zero or near-zero tariffs under the FTA, can be identified and supported through export incentives, financial schemes, and market entry programs. These interventions would ensure that policy and funding support flow to enterprises most poised for growth.

Another major insight is the central role of innovation in export potential. MSMEs with strong R&D activity, IP generation, and differentiated product offerings demonstrated a higher likelihood of export success. As such, a valuable business idea involves launching R&D-linked export incentives, innovation grants, and patent-based financial schemes to fuel exports from sectors like pharmaceuticals, electronics, and chemicals. Building innovation clusters or incubators within high-performing MSME zones would further amplify these effects and encourage industry-wide knowledge spillovers.

The model also uncovered a correlation between positive sentiment on platforms like LinkedIn or X and sector-level export readiness. This implies that digital sentiment can be harnessed as a forward-looking indicator of market confidence and trade preparedness. A potential business opportunity lies in developing AI-driven sentiment monitoring tools that scan digital platforms for sector-level buzz, compliance concerns, and policy-related discussions. These tools could support trade councils, policy makers, and exporters by providing real-time insights into market sentiment, enabling faster response to trade shifts.

The predictive model can also serve as the foundation for an MSME scoring system, which ranks firms based on their likelihood of export growth. This would allow government agencies and financial institutions to allocate procurement contracts, export loans, and technical support to MSMEs with high predictive scores. Such a system would promote efficient, data-backed decision-making and maximize the return on government-led export support programs. Export councils could even create dashboards that rank sectors and firms, guiding global trade partnerships and targeted outreach.

Additionally, the importance of logistics costs and international certifications highlights two persistent bottlenecks that affect MSME competitiveness. Many firms still face high export shipping costs and barriers to meeting UK quality standards. This opens up an opportunity for logistics companies, certification consultants, and government partners to jointly offer FTA-readiness bundles. These might include bulk shipping discounts, group certification programs, or one-stop FTA helpdesks to ease MSMEs through compliance hurdles and into the global trade ecosystem.

Finally, one long-term opportunity involves building a digital intelligence platform based on this model. MSMEs could enter key business inputs into an online tool and receive a custom export growth score, sector comparison, and policy recommendations. Integrated with government portals like Udyam or UK-India trade gateways, this platform could democratize access to AI-driven insights and help millions of MSMEs make informed decisions under the FTA regime. Together, these business ideas present a comprehensive roadmap for scaling India’s MSME exports in the post-FTA landscape.

Technical Challenges

Developing a predictive model for estimating export growth potential under the India–UK FTA posed several technical challenges, primarily due to the need to simulate realistic economic behavior in the absence of actual post-FTA trade data. Since the agreement was signed in July 2025 and implementation is still underway, real export volumes, tariff effects, and MSME adjustments are not fully observable. This required the creation of a synthetic dataset with 30,000 MSME records across key FTA-benefiting sectors, which had to reflect plausible economic distributions, sector diversity, and policy impacts based on past trends, government reports, and trade projections. Ensuring that this synthetic data remained statistically robust and representative of India's MSME landscape was a significant undertaking.

Feature selection was another complex area. The dataset included both numerical and categorical variables representing trade metrics (like tariffs and export volume), structural variables (like production capacity and workforce size), and softer indicators (like sentiment score and innovation index). Balancing feature relevance, correlation handling, and dimensionality reduction while maintaining model interpretability required multiple iterations and exploratory data analysis steps. Outlier detection and treatment using the IQR method added complexity as some features, like Tariff_PreFTA and Innovation_Index, showed skewed distributions typical in real-world trade data.

Model selection presented its own set of technical hurdles. Classical regression models like Linear Regression and KNN failed to capture the non-linear relationships between MSME performance and FTA-related variables, resulting in low or even negative R^2 scores. The Random Forest Regressor, although more suitable for this context, required hyperparameter tuning to achieve reasonable performance. Even with optimization, the model faced difficulty generalizing due to the inherently unpredictable nature of international trade, market sentiment, and currency fluctuations, all of which affect MSME exports dynamically.

Additionally, the evaluation of model performance was constrained by the lack of ground truth. Since the actual impact of the FTA will unfold over the next few years, the model’s predictions could not be validated against real 2030 outcomes. This made metrics like R^2 indicative but not absolute, requiring careful interpretation. Another challenge was encoding boolean and categorical variables (like Sector and Procurement_Opportunity) without introducing data leakage or multicollinearity, especially when building models using top-ranked features.

Finally, implementing sentiment analysis using simulated scores to reflect online perceptions about FTA benefits introduced some limitations in terms of realism and precision. While these scores helped approximate trade optimism or hesitancy, they could not fully replicate the nuances of real-time social media data due to data access constraints. Despite these technical challenges, the model provides valuable strategic direction and demonstrates the potential of machine learning in trade policy evaluation and MSME growth forecasting.

Recommendations

Based on the findings of this machine learning-based study, several targeted recommendations can help maximize the export potential of Indian MSMEs under the newly signed India–UK Free Trade Agreement. First and foremost, policymakers should prioritize MSMEs operating in high-FTA-access sectors such as textiles, pharmaceuticals, electronics, and chemicals, as these sectors demonstrated the strongest predicted export growth in the model. Dedicated export incentives, easier access to credit, and reduced compliance burdens for firms in these sectors will accelerate trade outcomes. Second, innovation emerged as a strong predictor of growth—hence, introducing innovation-linked schemes like R&D tax rebates, IP protection support, and product development grants can help high-potential MSMEs become more competitive in the UK market.

Additionally, the government and export councils should invest in digital tools that use predictive analytics to score and rank MSMEs based on their export readiness and growth probability. These tools could be integrated into Udyam or India's EXIM platforms to automate eligibility screening for trade support, finance, and certifications. Another recommendation is to lower structural trade barriers, especially in logistics and certification. Public-private partnerships can be developed to offer affordable cold storage, group shipping deals, and subsidized ISO/FDA certification packages to MSMEs in sectors such as marine products and pharmaceuticals, where compliance costs are high.

It is also recommended that real-time sentiment tracking tools be deployed by trade and industry bodies to capture emerging concerns and positive sentiment trends on platforms like LinkedIn and X (formerly Twitter). This can offer early-warning signals or green flags to adapt policy interventions, promotional campaigns, and export documentation requirements. Lastly, to make this model actionable at scale, a national-level digital MSME FTA Readiness Dashboard should be developed, where firms can input their data and receive dynamic feedback on their predicted export growth, compliance gaps, and next-best actions to enter the UK market successfully.

Conclusion

This study set out to explore how Indian MSMEs can leverage the newly signed India–UK Free Trade Agreement (FTA) to enhance their export potential, using predictive modeling as a decision-support framework. By simulating a realistic, policy-aligned dataset of 30,000 MSMEs across key sectors and applying machine learning algorithms, the analysis succeeded in identifying the primary factors that influence export growth in a post-FTA scenario. The model revealed that tariff reduction, market access (FTA_Access_Score), prior export volume, innovation capabilities, and sectoral demand trends in the UK are the most significant predictors of MSME export performance.

Despite challenges related to data simulation, model tuning, and lack of post-FTA ground truth, the Random Forest-based approach provided meaningful insights that align with India's real-world trade strengths and FTA provisions. The findings confirmed that MSMEs in sectors like textiles, electronics, and pharmaceuticals stand to benefit the most—especially those with a history of exporting, higher innovation indices, and readiness to meet UK certification and logistics requirements.

The study's outcomes go beyond model accuracy to offer a strategic framework for MSME policy design, sector prioritization, and government support allocation. Predictive analytics can play a transformative role in export planning by helping stakeholders anticipate outcomes and optimize interventions, particularly in an era of dynamic global trade. Moreover, the methodology developed here is scalable and adaptable to other bilateral FTAs or trade scenarios involving emerging markets.

In conclusion, the application of machine learning to policy-centric trade data demonstrates the feasibility and value of data-driven decision-making in export development. With the right ecosystem of digital tools, financial support, and capacity-building, Indian MSMEs can become stronger participants in global trade. The India–UK FTA is not just a trade agreement—it is a timely opportunity to reshape the MSME export landscape, and data science is a key enabler of that transformation.

References / Appendices

Dataset Details

• Name: Fresher Hiring Trends & Candidate Profiling Dataset

• Size: 30,000 records

• Attributes: 30+ variables including Academic Scores, Technical Skills, Internship Experience, Project Work, Behavioral Attributes, Training Performance, and Hiring Outcomes.

• Domain: Human Resources / Talent Acquisition – IT Industry Recruitment Analytics

Data Source

• Inspired by job and performance trends from:

LinkedIn " Freshers in Demand... Again" Report (July 2025)

• Reference Link: https://www.linkedin.com/news/story/freshers-in-demand-again-6469572/

Tools & Technologies Used

• Programming Language: Python

• Libraries:

  - Pandas – Data Manipulation
  - NumPy – Numerical Computation
  - Seaborn & Matplotlib – Data Visualization
  - Scikit-learn – Machine Learning Modeling

• Environment: Google Colab