

Condensed report

Max Austin - Mada2

2025-04-20

Introduction

This paper analyzes the United Kingdom's international trade activity from 1995 to 2023 using data from the CEPII BACI database. The aim is to test whether future UK trade values—total trade, product-level flows, and country-level trade patterns—can be accurately forecasted using current data-based techniques.

To answer this, the report focuses on three significant questions: Will UK volumes of trade rise or fall in coming years? What goods and trading partners will set future trends? And what is difficult when attempting to forecast long-term trade?

The original data comprises over 258 million observations of trade between 238 nations and 5,000 products. It is reduced to UK trade only as an importer or exporter. The work is conducted in four stages: initial exploration and visualization, correlation analysis, product and country clustering, and finally, deep learning models for prediction.

Additional exploratory plots and additional figures are provided in the supplementary PowerPoint appendix due to space constraints.

Exploration and Visualisation

Before analysis, the raw CEPII BACI dataset of more than 258 million records of international trade was pre-filtered to keep only rows where the UK imported or exported. This reduced the dataset to 13.4 million observations.

Product categories were originally 6-digit Harmonized System (HS) codes and country codes were numeric and both replaced with descriptive labels. Products were too broad in their detail (e.g., 21 varieties of meat, multiple varieties of coffee and eggs), so these were merged into more general product categories where feasible.

After cleansing and consolidation post, the dataset came down to 4.96 million observations around equal to 1.9233751% of the initial dataset. They consist of 3.39 million exports and 1.59 million imports.

In spite of additional export deals of quantity, this does not include greater trading value. There are greater total values of import over the period for UK compared to exports, with its net trading balance being -5.1058488^9 , which is equal to continuous deficit.

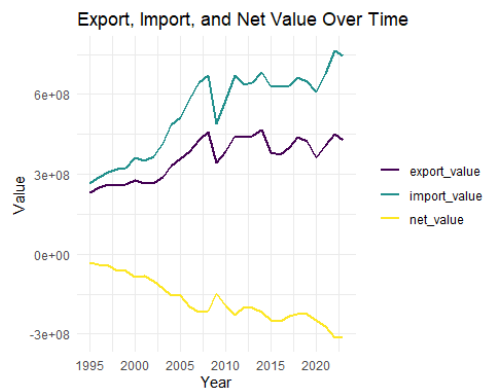
Visualisation

Trade value over time

Figure X illustrates the UK's net trade, import, and export totals between 1995 and 2022. Exports (purple) and imports (teal) exhibit steady growth up to 2008, when they varied depending on world economic turning points. Net trade (yellow), always in negative figures, shows a persistent and growing trade deficit—especially after 2015.

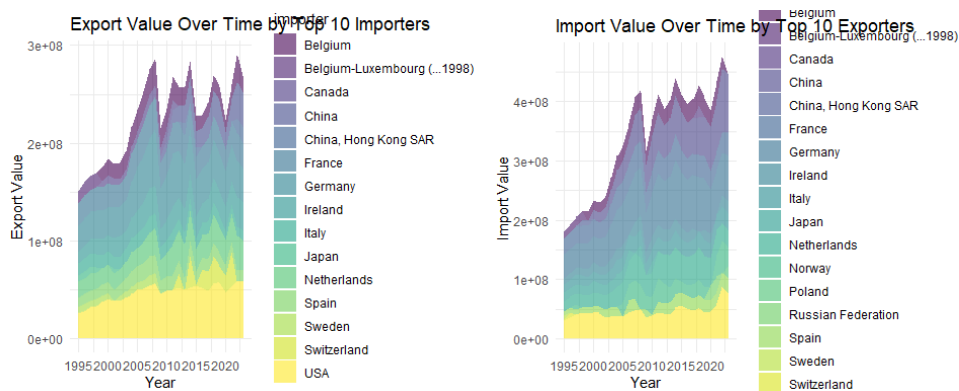
Severe economic shocks such as the 2008 financial crisis and pandemic 2020 triggered sharp declines in trade, then partial recoveries. Post-Brexit (2016) onwards, export growth plateaus while imports remain unchanged, which indicates emerging new trade frictions.

This historical trade deficit bears witness to UK dependence on imports and highlights the challenge of export competitiveness.



Stacked Area Plots

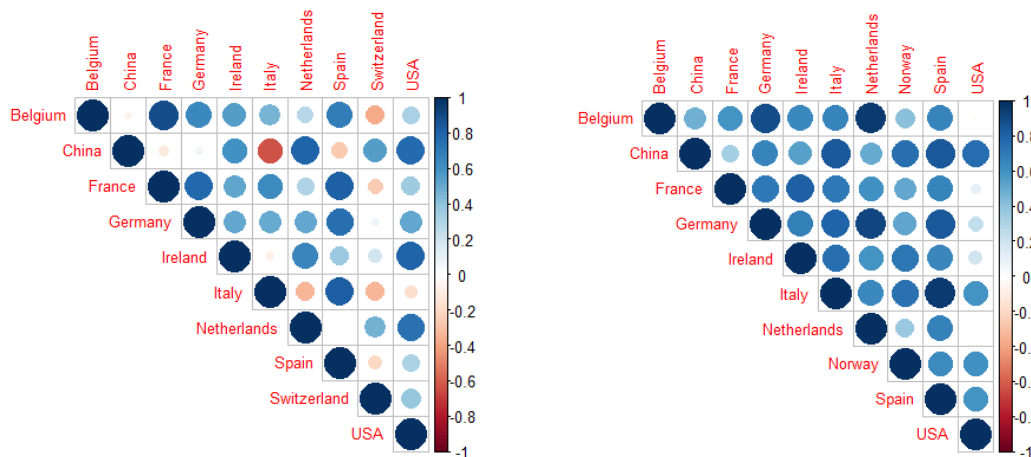
The stacked area plots indicate the UK's export and import values with its top ten trading partners between 1995 and 2023. Export values rose steadily until the late 2010s, led consistently by the US, Germany, France, and the Netherlands. Imports also rose, with Germany, China, and the US being leading suppliers. Notably, China's role expanded exponentially in the 2000s. Both graphs identify the UK's historic reliance on its key European and world partners, and the persistent import and export value deficit is a signal of a long-term structural trade deficit.



Correlations

Correlation between trade partners

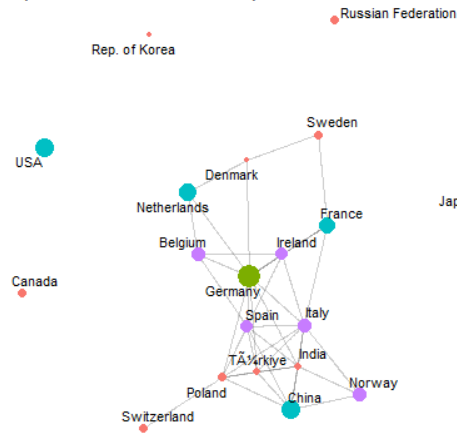
These two correlation matrices highlight the UK's prime trading partners segregated according to whether they are being exported to (left) or imported from (right). Exporting, the UK is extremely positively correlated with most of Europe's partners—Germany, France, Ireland, and the Netherlands—sitting in synchronised export trends into these countries. The USA, being an important partner, is less correlated, likely reflecting differences in economic cycles. On the import side, the relationships are closer still, especially between EU countries, and could reflect coordinated supply chains and integrated buying from the single market. On aggregate, the matrices reveal strong regional trade linkages, useful for identifying co-movements in trade volume that impact future projections.



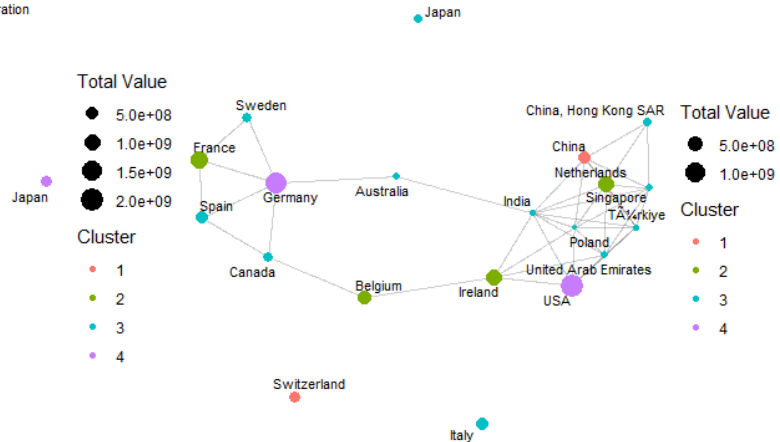
Clustering

The clustering diagrams reveal how UK trade relationships group naturally based on trade similarity. In the import clusters (left), the majority of European exporters—including Germany, Italy, Spain, and Belgium—form a dense, interconnected network, suggesting similar trade patterns and co-movement in supply. The USA and Asian partners appear more isolated, indicating distinct dynamics. The export clusters (right) show a comparable trend, with Western nations such as France, Germany, and Canada forming cohesive groups, while emerging markets like the UAE, India, and Singapore form their own separate hub. These clusters highlight patterns not obvious through raw trade values alone, revealing hidden structure that supports the motivation for using deep learning models to detect and learn such implicit trade relationships. Identifying these latent clusters helps inform model design by emphasizing the importance of modeling interaction effects between countries and products—precisely the kind of complex structure deep learning is well-suited to capture.

Exporter Clusters - UK Imports



Importer Clusters - UK Exports



Deep Learning

Methods

Two deep learning models were created to forecast UK trade values, one for when the United Kingdom is an importer and one for when it is an exporter. These models were implemented using the Keras API in TensorFlow (version 2.10.0) under the R environment. To prepare the data for modeling, all trade transactions worth less than 5,000 were removed in an effort to reduce noise and prevent the likelihood of overfitting. Categorical columns such as trade partners and product types were embedded to integers and passed through embedding layers to learn their internal representation. Numerical features such as trade year and trade value were normalized to have zero mean and unit variance.

Both models were trained using 30 epochs and a batch size of 216, selected after trying different combinations of hyperparameters. This configuration provided consistent stable performance and avoided convergence issues associated with larger or smaller batch sizes. Mixed precision training (with the mixed_float16 policy) was also enabled to improve training efficiency without sacrificing numerical accuracy.

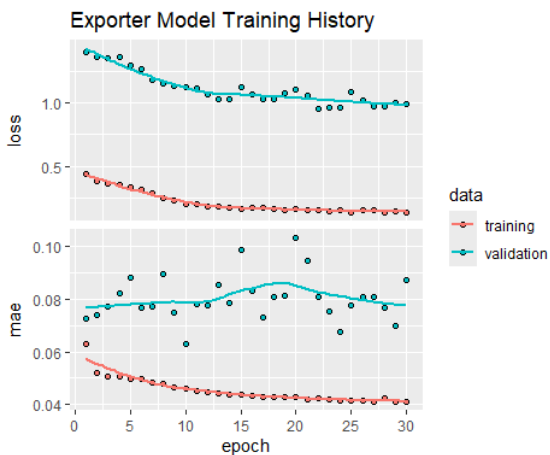
The loss function used was mean squared error (MSE), with mean absolute error (MAE) tracked as a secondary performance measure. A validation split of 20% was used in training, and shuffling was enabled to aid generalization. Model performance was monitored while training via training history plots. Both models displayed steady convergence and minimal divergence between training and validation error curves, indicating high generalization and no overfitting evidence.

Training

Exporter Model Training Performance

The exporter model shows a steady decrease in training loss and MAE, indicating consistent learning over the 30 epochs. While validation loss is decreasing early on as well, it begins to stabilize and fluctuate from roughly epoch 15, which means that while the model generalizes

reasonably well, its capacity for continued improvement decreases. Validation MAE remains close to training MAE across the board with no clear sign of overfitting but with a hint of limited predictive variability acquired for unseen observations.



Importer Model Training Performance

The importer model also displays a comparable trend of shrinking training loss and MAE, and flattening validation metrics at around midpoint training. The gap between training and validation curves is reasonably small, giving credence to the model's stability. Validation MAE's slight decline in later epochs also indicates soft late-stage generalization improvement, but the occasional jitter indicates some batch composition sensitivity or minimal underfitting in more complex import relations.

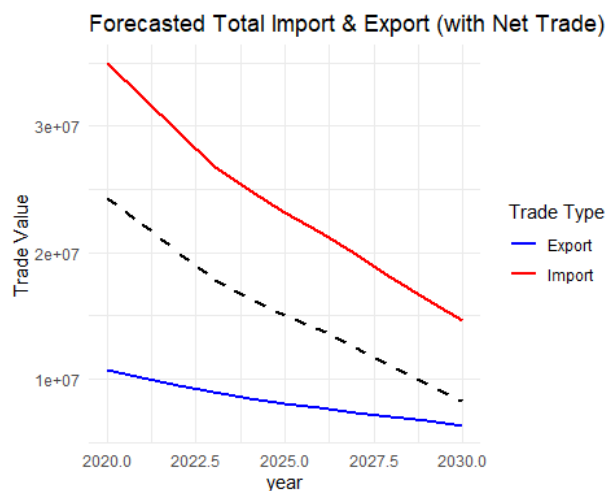


Summary of training

Both exporter and importer models register consistent decreases in training loss and mean absolute error (MAE), with validation curves plateauing halfway through training. The comparatively narrow and consistent gap between training and validation metrics is an expression of high generalization and minimum risk of overfitting. Whereas the exporter model is consistently performing well across, the importer model experiences moderate declines in late-stage

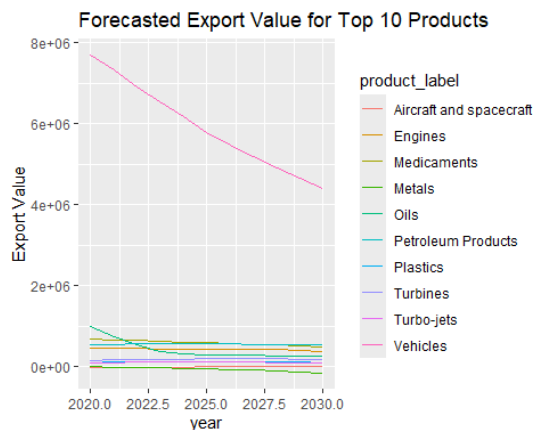
performance and occasional fluctuations, presumably as a result of batch variation or more intricate import dynamics. In general, both models converged and are suitable for prediction.

#Results ## Forecasted Total Trade Overview Figure X illustrates estimated overall UK imports, exports, and net trade from 2023 to 2030. Imports (red) and exports (blue) are both estimated to gradually decline, with imports always in excess, once again maintaining a chronic rising trade deficit (black dashed line). This may reflect increasing dependence on imported foreign goods and poorer export performance.



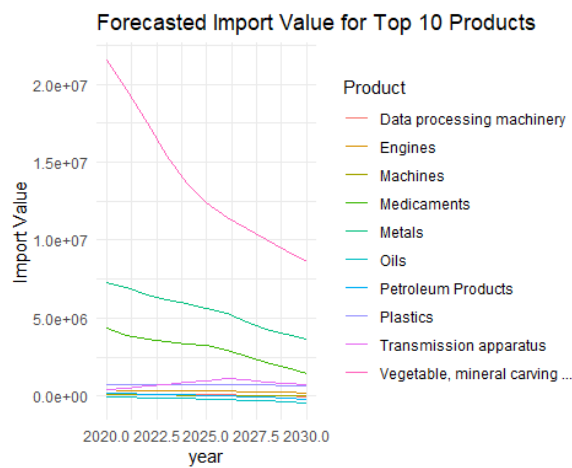
Export top 10 products

Figure X displays export forecasts by commodity. Vehicles remain the biggest but are projected to plummet by a huge margin. Other huge groups such as medicaments, engines, and petroleum products also fall, though at a more moderate pace. The simultaneous tendencies across goods are explained by the autoregressive characteristics of the model, which has a tendency to yield conservative and smoothened forecasts.



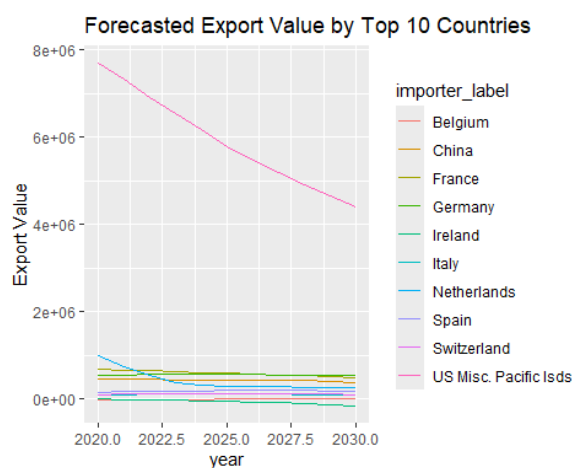
Import top 10 products

Figure X shows UK import projections of the country’s leading imported goods. Oils and cars lead in terms of number, with most categories—like metals, chemicals, and machinery—exhibiting flat or gradually declining trends. As indicated above, the smoothing bias of the model may downwardly bias volatility in sectoral demand.



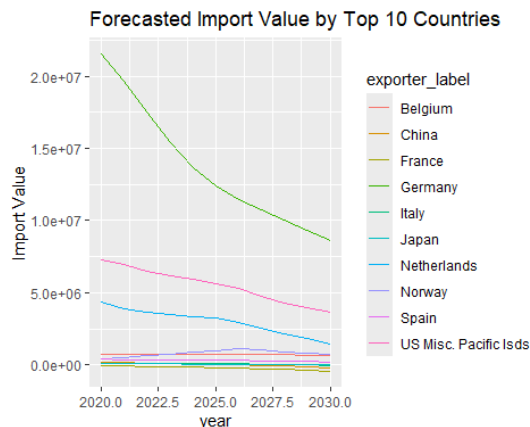
Export top 10 Countries

Figure X displays forecasted export values to the UK’s main trading partners. The USA remains the biggest recipient, followed by Germany and the Netherlands. All partners, though, show marginal decreases in export value, with no sign of growth in any nation, in support of the forecasted export contraction.



Import top 10 countries

Figure X shows UK imports from its top suppliers. Germany and China dominate, with other countries such as the USA, France, and the Netherlands showing similar falling trends. That none of these exporters is expected to increase indicates a general downturn in incoming trade.



Deep learning conclusion

The deep learning models developed for UK trade value prediction operated in a consistent way for both export and import scenarios. By encoding categorical variables and normalizing numerical values, the models picked up on general trends within historical trade data and made consistent predictions across countries and products.

However, due to the autoregressive nature of the forecasting process—where each future value is forecasted in relation to previously forecasted inputs—the results are smoothed and conservative. This is evidenced in the flat or gradually declining trends across all groups. Although reflecting the robustness of the model in identifying patterns, this limits the responsiveness of the model to abrupt economic shocks, policy shifts, or random changes in demand.

Lastly, the models provide a helpful high-level view of potential future trade patterns, highlighting the UK's persistent trade deficit and worsening export performance. However, for short-term forecasting or sectoral detail, complementary methods employing external variables or real-time indicators may be required to capture the complete richness of international trade.

Final conclusion

This report explored the potential for forecasting UK international trade flows using data-driven methods. The primary research questions—whether to expect expansion or a decline in UK trade, which partners and products will dominate future flows, and what are the pitfalls of forecasting long-term trade—have been addressed by a multi-stage investigation.

After data cleaning and visualization of CEPII BACI trade data, UK export and import trends were analyzed over time. Product and country level trends indicated underlying interdependencies and long-term dynamics, e.g., impacts of global events and shifting trade alliances. Cluster analysis

identified hidden groupings among goods and trading partners, offering insight into interdependent economic activities.

The core forecasting task was met with two deep learning models, one for UK imports and one for UK exports. These models captured general directional tendencies and produced stable forecasts but also demonstrated hallmark weaknesses of autoregressive deep learning, such as smoothing of uncertainty and subdued short-run volatility. For each of these, the models appropriately forecast future decreases in trade volumes and expanding trade deficits—results of considerable economic and policy significance.

Taken together, the analysis demonstrates how machine learning can provide valuable insight into the future of UK trade, along with the need for careful interpretation of long-term projections from data-driven models.

Reference

https://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=37

Appendix

All code and supporting documents are available on github in the following perma link
<https://github.com/maxaus2002/ASSESSMENT-2-for-MAST7220/tree/4ee2afca34a870aa6cfba57fd4bb3f3344b36c84>

Full version of report available on github Accompanying power point available on github
(Mainly contains condensed versions of data omitted from the condensed report but were present in the full report.)