

Cross-Country Transfer Learning for FDI Forecasting in Small Macroeconomic Datasets

Khlyostkina Xeniya

Master's Student, Applied Statistics

Zhejiang University of Science and Technology, China

ABSTRACT: This study examines whether transfer learning can improve machine learning performance under data-scarce forecasting conditions. The empirical application focuses on forecasting foreign direct investment (FDI) inflows using CatBoostRegressor. The source domain is represented by a broad international panel of countries, while the target domain consists of Central Asian economies.

CatBoost is evaluated in two settings: without transfer learning, where the model is trained only on target-region data, and with transfer learning, where the model is first pretrained on the source domain and then fine-tuned on the target domain. The results are compared with simple baseline models, including a lagged-target naive model and a regional mean benchmark.

The findings show that transfer learning substantially improves predictive accuracy. Compared with the target-only CatBoost model, the transfer learning model reduces RMSE from 4.0421 to 3.0426 and MAE from 3.3504 to 2.3369. R^2 also improves from -5.8205 to -0.0761 . These results suggest that transfer learning can help stabilize machine learning forecasts when target-domain observations are limited.

KEYWORDS: Machine Learning, Transfer Learning, CatBoost, FDI, Data-Scarce Economies

I. INTRODUCTION

Machine learning methods are increasingly used for forecasting tasks because they can capture nonlinear relationships, interactions between predictors, and complex patterns that may be difficult to model using traditional statistical approaches. However, their practical use is often limited by the availability of data. Many real-world forecasting problems, especially in economics and country-level analysis, are characterised by small samples, short time series, missing observations, and heterogeneous data quality. In such settings, conventional machine learning models may fail to generalise well because they are trained on a limited number of target-domain observations.

Data scarcity, however, does not necessarily mean that machine learning methods should be rejected. Instead, it creates the need for approaches that can use information more efficiently. Transfer learning provides one such framework. Its main idea is to reuse knowledge learned from a related source domain or task in order to improve prediction in a target domain where direct training data are limited. Rather than relying only on the small target sample, a model can first learn general patterns from a broader dataset and then adapt this knowledge to the specific target setting.

This approach is especially relevant when the target domain is not fully isolated but shares some structural similarities with a larger source domain. In forecasting problems, this means that relationships observed in a broader sample may help stabilize prediction in a smaller sample, provided that the source and target domains are sufficiently related. At the same time, transfer learning does not assume that the two domains are identical. It explicitly allows for distributional differences and therefore offers a flexible methodological framework for working with small and heterogeneous datasets.

The present study applies this idea to the forecasting of foreign direct investment inflows. The empirical setting uses a broader international country panel as the source domain and a smaller regional sample as the target domain. The main objective is not only to forecast FDI, but also to examine whether transfer learning can improve machine learning performance under data-scarce conditions. For this purpose, the study compares a CatBoost model trained only on target-region data with a transfer learning version pretrained on the source domain and then adapted to the target domain.

By doing so, the study contributes to the growing literature on machine learning under limited-data conditions. It shows how transfer learning can serve as a practical strategy when the target dataset is too small for stable standalone model training, but related external



data are available. The results therefore provide evidence on the usefulness of transfer learning as a methodological tool for data-scarce forecasting problems.

II. LITERATURE REVIEW

Transfer learning (TL) is a machine learning approach that reuses knowledge from related source tasks or domains to improve performance on a target task, particularly when labelled data in the target domain are scarce, costly to obtain, or unavailable. Pan and Yang (2010) note that conventional machine learning typically assumes that training and testing data are drawn from the same feature space and follow the same probability distribution. When this assumption is violated, predictive performance may deteriorate because of distribution mismatch or covariate shift. Shimodaira (2000) specifically discusses this issue in the context of predictive inference under covariate shift. In such settings, TL becomes especially relevant because it provides a framework for using information from related domains when direct target-domain learning is constrained. Weiss et al. (2016) similarly emphasize that transfer learning is useful when target-domain data are limited, expensive, or difficult to collect.

The core idea of TL is that knowledge acquired from previous tasks can help solve new but related problems more efficiently. Pan and Yang (2010) define transfer learning as the process of improving the target predictive function by using information from a related source domain or source task. This idea is closely connected to earlier concepts such as learning to learn, knowledge transfer, inductive transfer, meta-learning, multi-task learning, and domain adaptation. Thrun and Pratt (1998) discuss learning to learn as an early framework in which previous learning experience can improve performance on new tasks, while Pan and Yang (2010) later systematize these related concepts within the broader transfer learning literature.

Several survey studies provide theoretical support for the use of TL. Niu et al. (2020) review transfer learning research from 2010 to 2020 and argue that TL is particularly useful when conventional machine learning faces insufficient labeled data, limited computational resources, and distribution mismatch. Farahani et al. (2020) similarly emphasize that TL can address discrepancies between domains and data distributions that traditional ML methods often cannot handle. Zhuang et al. (2021) provide a broader systematic review of TL methods and classify existing approaches from the perspectives of data-based and model-based transfer. Hosna et al. (2022) also note that TL enables faster and efficient model development across multiple application domains, including image processing, speech recognition, and natural language processing.

At the same time, the literature stresses that transfer learning is not automatically beneficial. Niu et al. (2020) highlight the risk of negative transfer, which may occur when unrelated or weakly related source-domain knowledge is transferred to the target task. They identify source-domain selection, domain-distance measurement, and the stability of transfer across distant domains as key open challenges. These issues are directly relevant for cross-country macroeconomic forecasting, where countries may share some structural patterns but also differ in institutional, political, and economic conditions.

Recent literature on time-series transfer learning further supports the relevance of TL for forecasting tasks. Weber et al. (2021) conduct a systematic mapping study of transfer learning with time series data and show that TL has increasingly been applied to prediction problems where historical observations are limited. They emphasize that time-series TL can reduce the need to collect large amounts of data for each individual target domain and can support forecasting in settings with limited time-series observations. This is important for macroeconomic applications, where country-level annual data often produce relatively short panels.

Empirical evidence also suggests that TL can improve forecasting performance. Solís and Calvo-Valverde (2022) compare deep learning models with and without transfer learning for multiple-step-ahead monthly time-series forecasting using M3 and M4 competition data. Their results suggest that TCN, LSTM, and CNN models with transfer learning tend to outperform traditional forecasting methods in several settings. In a related macroeconomic application, Solís and Calvo-Valverde (2023) propose transfer learning for monthly macroeconomic time-series forecasting, arguing that pre-trained deep learning models can be useful when macroeconomic time series are short. Nguyen and Nguyen (2020) also directly apply transfer learning to macroeconomic forecasting, indicating that TL has begun to enter economic forecasting research, although the literature remains relatively limited. Quantitative evidence from other domains is also compelling. Mlinarič et al. (2025) report up to a 20% improvement in classification accuracy for industrial quality inspection tasks with small datasets of 100 instances by transferring knowledge from larger datasets of 10,000 instances. Laurer et al. (2023) find that BERT-based transfer learning improves text classification performance by 10.7–18.3 percentage points compared with classical models, and that models trained on 500 texts can match classical models trained on 5,000 texts. These findings support the broader argument that TL can be especially valuable when the target dataset is small.



Recent studies further summarize the practical advantages of TL. Topsakal and Akinici (2023) emphasize that TL can reduce training time, require less labeled data, improve model performance, and enhance generalization by reusing knowledge from related source tasks. They also distinguish several TL strategies, including adaptation of learned representations, multi-task learning, and pre-trained model adaptation. These ideas are relevant for data-scarce forecasting problems, where target-domain observations are limited and models may benefit from information transferred from related domains.

An applied example is provided by Alolayan et al. (2021), who use transfer learning for shale gas production forecasting under data scarcity. They develop county-specific DNN models by transferring knowledge from models trained on nearby counties with more observations. Using data from more than 6,000 shale gas wells across 17 counties, their transfer learning models reduce forecasting error by 11-47% compared with the Arps decline curve benchmark. This study is especially relevant for cross-country macroeconomic forecasting because it demonstrates how information from related regions can improve local forecasting models when target-region data are limited.

The relevance of machine learning for macroeconomic forecasting is also supported by studies outside the TL literature. Medeiros et al. (2021) show that random forests can outperform traditional benchmarks in inflation forecasting, while Naghi et al. (2024) provide new evidence that several machine learning methods are competitive in inflation forecasting across the United States, Canada, and the United Kingdom. This literature suggests that nonlinear ML models can be useful for macroeconomic prediction, although performance may vary across countries, periods, and crisis episodes.

For FDI specifically, Islam et al. (2024) examine foreign direct investment inflow forecasting for Bangladesh using ARIMA and Random Forest methods. Although their study does not use transfer learning, it is relevant because it shows that FDI forecasting has begun to adopt machine-learning-based approaches. However, most existing FDI forecasting studies remain country-specific and do not explicitly address the problem of data scarcity through cross-country transfer. This creates a methodological gap for the present study: applying transfer learning to FDI inflow forecasting in Central Asia, where individual country datasets are small but related cross-country information may improve predictive performance.

III. METHODOLOGY

3.1 Transfer Learning

Transfer learning is a machine learning approach based on the idea of transferring knowledge between related tasks or domains. Unlike the classical approach, where a model is trained and applied within the same data distribution, transfer learning allows the source and target domains to differ. Its main objective is to use knowledge obtained in one context to improve learning in another, provided that there is a meaningful or statistical relationship between them.

The idea of knowledge transfer has been developing in machine learning since the 1990s within the broader research areas of learning to learn and lifelong learning. One of the early works in this field is the book by Thrun and Pratt (1998), which examined the ability of algorithms to use prior experience for more efficient learning of new tasks. Later, transfer learning received more systematic theoretical treatment in the works of Pan and Yang (2010), Torrey and Shavlik (2010), and Weiss et al. (2016).

According to Pan and Yang (2010), a domain is defined as:

D = {X, P(X)} (1)

where X denotes the feature space and P(X) is the marginal probability distribution of the features. A learning task is defined as:

T = {Y, f(.)} (2)

where Y is the output space and f(.) is the predictive function learned from the data.

In transfer learning, a distinction is made between the source domain D_s and the target domain D_t:

D_s = {X_s, P(X_s)} (3)

D_t = {X_t, P(X_t)} (4)

as well as between the source task T_s and the target task T_t:

T_s = {Y_s, f_s(.)} (5)

T_t = {Y_t, f_t(.)} (6)

Transfer learning is relevant when the source and target domains or tasks are not identical:

D_s ≠ D_t or T_s ≠ T_t (7)



At the same time, there must be a sufficient relationship between them that allows useful information to be transferred. Formally, the objective of transfer learning can be written as:

$$f_t(\cdot) = \text{learn}(D_t, T_t; D_s, T_s) \tag{8}$$

where $f_t(\cdot)$ is the target predictive function constructed with the use of information from the source domain or source task.

The literature commonly distinguishes three main forms of transfer learning: inductive, transductive, and unsupervised transfer learning. Inductive transfer learning is used when the source and target tasks differ, but labeled data are available for the target task. Transductive transfer learning is applied when the task remains the same, but the distributions of the source and target domains differ. Unsupervised transfer learning refers to cases where the target task has no labeled output variable and is related, for example, to clustering or dimensionality reduction.

Transfer learning can also be classified by the mechanism through which knowledge is transferred. Instance-based transfer learning transfers or reweights individual observations. Feature-based transfer learning constructs a shared feature representation for the source and target domains. Parameter-based transfer learning transfers model parameters or model structure. Relational transfer learning is used when relationships between objects are important.

In practice, transfer learning is widely applied in computer vision, natural language processing, medicine, bioinformatics, financial forecasting, and economic modelling. In many modern applications, it follows a two-stage procedure: first, a model is trained on a broader source domain, and then it is adapted to the target domain:

$$f_s = \text{train}(D_s, T_s) \tag{9}$$

$$f_t^L = \text{adapt}(f_s, D_t, T_t) \tag{10}$$

where f_s is the model obtained from the source domain, and f_t^L is the model after knowledge transfer and adaptation to the target domain.

An important limitation of transfer learning is the risk of negative transfer. It occurs when the source and target domains are insufficiently related, so transferred knowledge worsens rather than improves model performance. Positive transfer can be defined by comparing the loss of a transfer learning model with that of a model trained without transfer:

$$L_t(f_t^L) < L_t(f_t^{\text{NoTL}}) \tag{11}$$

If the opposite relationship holds:

$$L_t(f_t^L) > L_t(f_t^{\text{NoTL}}) \tag{12}$$

then the transfer can be interpreted as negative transfer. Therefore, empirical applications of transfer learning should be evaluated against a control model trained only on target-domain data.

In this study, transfer learning is applied to the forecasting of foreign direct investment. The source domain is represented by an international panel of countries outside Central Asia, while the target domain consists of Central Asian countries. The target variable is the same in both domains: foreign direct investment, net inflows as a percentage of GDP. Therefore, the difference between the domains is mainly related to the distribution of macroeconomic characteristics rather than to the forecasting task itself.

Thus, transfer learning serves as a methodological framework for transferring general patterns from a broader international context to a target regional setting. In the case of FDI analysis, this means using information about the relationship between macroeconomic indicators and investment inflows across countries and then adapting this knowledge to the Central Asian context.

3.2 ML Algorithms and Baseline Models

The empirical part of the study uses CatBoostRegressor as the main machine learning algorithm and compares its performance with two simple baseline models. CatBoost is applied in two settings: without transfer learning, where the model is trained only on the target-region data, and with transfer learning, where the model is first pretrained on the source domain and then fine-tuned on the target domain.

CatBoost is a gradient boosting algorithm based on decision trees. The general idea of boosting is to build an ensemble of weak learners sequentially, where each new tree attempts to correct the errors made by the previous trees. CatBoost extends this framework by using ordered boosting and efficient handling of categorical features, which helps reduce prediction bias and improve model stability (Prokhorenkova et al., 2018). In this study, CatBoost is suitable because it can capture nonlinear relationships and interactions between macroeconomic indicators without requiring a very large target-region sample.

The first baseline model is the lagged-target naive model. It assumes that the current value of FDI can be approximated by its previous observed value:



$$\hat{y}_t = y_{t-1} \tag{13}$$

where \hat{y}_t is the predicted value and y_{t-1} is the lagged value of FDI. This baseline is useful because FDI flows often exhibit persistence over time. Therefore, a machine learning model should outperform this simple rule to demonstrate added predictive value.

The second baseline model is the regional mean model. It predicts the same value for all observations, equal to the average FDI value in the available training sample:

$$\hat{y}_i = \underline{y}_{train} \tag{14}$$

where \underline{y}_{train} is the mean FDI value in the training data. This model serves as a minimal benchmark. If a more complex model performs worse than the mean baseline, it indicates weak predictive ability.

The two baseline models were included to provide simple and interpretable reference points for evaluating the predictive contribution of the machine learning model and the transfer learning procedure. The lagged-target baseline was selected because FDI inflows may exhibit temporal persistence. Therefore, the previous observed value of FDI represents a natural benchmark for forecasting. If a machine learning model cannot outperform this rule, its additional explanatory value is limited.

The regional mean baseline was included as a minimal statistical benchmark. It predicts the same value for all observations, equal to the mean FDI value in the training sample. This baseline is important because it allows the evaluation of whether the model performs better than a simple average-based forecast.

In the standard definition, R^2 compares the model's prediction error with the variation of the observed values around their test-sample mean. Therefore, a negative R^2 indicates that the model performs worse than a simple mean-based benchmark for the evaluated sample.

Thus, these baseline models were not intended as complex alternatives to CatBoost, but as transparent reference points. They allow the study to assess whether CatBoost, and especially its transfer learning version, provides predictive value beyond simple persistence-based and mean-based forecasts.

IV. RESULTS

This section presents the predictive results of the transfer learning experiment for the target variable foreign direct investment, net inflows (% of GDP). After excluding Turkmenistan and removing observations with missing values after feature engineering, the final machine learning dataset includes 87 countries and 1,727 observations. The source domain consists of countries outside Central Asia, while the target domain includes Kazakhstan, Kyrgyzstan, Tajikistan, and Uzbekistan.

Table 1. Final machine learning sample

Sample component	Countries	Observations
Full ML dataset	87	1,727
Source domain: non-Central Asian countries	83	1,664
Target domain: Central Asian countries	4	63

The source model was first pretrained on the non-Central Asian sample. The best source-domain configuration achieved a validation RMSE of 2.1569. This model was then used as the starting point for fine-tuning on the Central Asian target domain under the Leave-One-Country-Out evaluation protocol.



Table 2. Best source-domain pretraining configuration

Parameter	Selected value
Source validation RMSE	2.1569
Learning rate	0.03
Tree depth	5
L2 regularisation	25
Random strength	4.0
Bagging temperature	2.0

The aggregate results show that transfer learning provides a clear improvement over the same model trained only on Central Asian data. The CatBoost model without transfer learning performs close to the simple regional mean benchmark and does not provide stable predictive gains. In contrast, the transfer learning model achieves the lowest RMSE and MAE among all evaluated approaches.

Table 3. Overall predictive performance across Central Asian countries

Model	Transfer learning	RMSE	MAE	R ²	Corr	Directional accuracy
CA mean baseline	-	4.0267	3.4517	- 6.0879	-	0.0000
Naive lagged-FDI baseline	-	4.0041	2.9822	- 0.6701	0.2489	0.4194
CatBoost	-	4.0421	3.3504	- 5.8205	0.3464	0.3537
CatBoost	+	3.0426	2.3369	- 0.0761	0.4169	0.3743

Compared with the CA-only CatBoost model, the transfer learning version reduces RMSE by 24.73% and MAE by 30.25%. The improvement in R² is also substantial, increasing from -5.8205 to -0.0761. Although the aggregate R² remains slightly negative, the transfer learning model is much closer to the mean benchmark and produces substantially smaller prediction errors.

Table 4. Improvement from transfer learning

Comparison	No transfer learning	Transfer learning	Improvement
RMSE	4.0421	3.0426	24.73%
MAE	3.3504	2.3369	30.25%
R ²	- 5.8205	- 0.0761	+5.7444 points



The comparison with baseline models is important for interpretation. The naive lagged-FDI baseline performs relatively well in directional accuracy, suggesting that FDI flows contain some temporal persistence. However, it does not achieve the lowest prediction errors. The transfer learning model performs best in terms of RMSE and MAE, indicating that it captures additional predictive information beyond both simple persistence and average regional behaviour.

Country-level LOCO results confirm that the effect of transfer learning is consistent across all four held-out countries in terms of RMSE and MAE. The strongest improvement is observed for Uzbekistan, while Kazakhstan also shows a clear shift from negative to positive R^2 .

Table 5. Country-level LOCO comparison: CatBoost without and with transfer learning

Held-out country	RMSE No TL	RMSE TL	MAE No TL	MAE TL	R^2 No TL	R^2 TL
Kazakhstan	4.5738	3.5523	3.2473	2.8078	- 0.1850	0.2852
Kyrgyzstan	4.8758	4.5799	3.8450	3.5909	- 0.0415	0.0811
Tajikistan	3.1642	3.0660	2.8278	2.1461	0.0790	0.1352
Uzbekistan	3.5546	0.9724	3.4815	0.8027	- 23.1345	- 0.8060

The Uzbekistan fold illustrates an important limitation of country-level R^2 in very small test samples. Even though the transfer learning model strongly reduces RMSE and MAE, R^2 remains negative because the statistic is highly sensitive to the variance of the dependent variable within the held-out country. For this reason, RMSE and MAE provide a more stable basis for interpreting forecasting improvement in the LOCO setting.

Overall, the results indicate that transfer learning improves predictive performance for FDI forecasting in Central Asian economies. The main gain is observed in magnitude-based error metrics, especially RMSE and MAE. The findings suggest that information transferred from the broader international panel helps stabilise model performance in the target regional domain and provides more reliable forecasts than both the CA-only model and simple baseline approaches.

V. CONCLUSION

This study examined the application of transfer learning to the forecasting of foreign direct investment inflows in Central Asian economies. The analysis focused on Kazakhstan, Kyrgyzstan, Tajikistan, and Uzbekistan, while a broader international panel of non-Central Asian countries was used as the source domain. The main objective was to evaluate whether knowledge learned from a wider cross-country dataset can improve predictive performance in a smaller regional target domain.

The results show that transfer learning provides a clear improvement over the same machine learning model trained only on Central Asian data. The transfer learning model reduced RMSE from 4.0421 to 3.0426 and MAE from 3.3504 to 2.3369, corresponding to improvements of 24.73% and 30.25%, respectively. The coefficient of determination also improved substantially, moving from a strongly negative value of - 5.8205 to - 0.0761. Although the aggregate R^2 remained slightly negative, the transfer learning model produced the lowest RMSE and MAE among all evaluated approaches, including both the CA-only model and simple baseline forecasts.

Country-level Leave-One-Country-Out results further confirm the stabilising role of transfer learning. Prediction errors decreased for all four held-out countries, with especially strong improvements for Uzbekistan and Kazakhstan. At the same time, the results also show that country-level R^2 should be interpreted cautiously in very small test samples, where low variation in the dependent variable can produce negative R^2 values even when absolute errors decline substantially. For this reason, RMSE and MAE provide a more reliable basis for evaluating forecasting gains in this setting.

The findings suggest that transfer learning is a useful methodological strategy for macroeconomic forecasting in regional contexts



where the target sample is naturally limited. Rather than replacing country-specific analysis, the approach allows the model to combine general international patterns with regional adaptation. In the case of FDI forecasting, this means that relationships learned from a broader cross-country panel can help improve predictions for Central Asian economies.

The study also has several limitations. First, the number of Central Asian observations remains small, which restricts the stability of country-level evaluation. Second, the results depend on the quality and availability of macroeconomic indicators across countries. Third, the transfer learning framework assumes that the source and target domains are sufficiently related; if this assumption is weak, there is a risk of negative transfer. Future research could extend the analysis by testing alternative source-domain definitions, adding institutional and geopolitical indicators, comparing several transfer learning algorithms, and combining predictive performance analysis with explainable AI methods to identify the economic drivers behind the forecasts.

Overall, the results demonstrate that transfer learning can improve the reliability of FDI forecasting in data-constrained regional settings. For Central Asian economies, this approach provides a practical way to use broader international evidence while still adapting the model to regional conditions.

REFERENCES

1. Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345-1359.
2. Shimodaira, H. (2000). Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of Statistical Planning and Inference*, 90(2), 227-244.
3. Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, 3(1), 1-40.
4. Thrun, S., & Pratt, L. (1998). Learning to learn: Introduction and overview. In S. Thrun & L. Pratt (Eds.), *Learning to Learn* (pp. 3-17). Springer.
5. Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345-1359.
6. Shimodaira, H. (2000). Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of Statistical Planning and Inference*, 90(2), 227-244.
7. Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, 3(1), Article 9.
8. Thrun, S., & Pratt, L. (1998). Learning to learn: Introduction and overview. In S. Thrun & L. Pratt (Eds.), *Learning to Learn* (pp. 3-17). Springer.
9. Caruana, R. (1997). Multitask learning. *Machine Learning*, 28, 41-75.
10. Niu, S., Liu, Y., Wang, J., & Song, H. (2020). A decade survey of transfer learning (2010-2020). *IEEE Transactions on Artificial Intelligence*, 1(2), 151-166.
11. Farahani, A., Voghoei, S., Rasheed, K., & Arabnia, H. R. (2021). A brief review of domain adaptation. In *Advances in Data Science and Information Engineering* (pp. 877-894). Springer.
12. Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43-76.
13. Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z., & Azim, M. A. (2022). Transfer learning: A friendly introduction. *Journal of Big Data*, 9(1), Article 102.
14. Weber, M., Auch, M. L., Doblander, C., Mandl, P., & Jacobsen, H.-A. (2021). Transfer learning with time series data: A systematic mapping study. *IEEE Access*, 9, 165409-165428.
15. Solís, M., & Calvo-Valverde, L.-A. (2022). Performance of deep learning models with transfer learning for multiple-step-ahead forecasts in monthly time series. *Inteligencia Artificial*, 25, 110-125.
16. Solís, M., & Calvo-Valverde, L.-A. (2023). A proposal of transfer learning for monthly macroeconomic time series forecast. *Engineering Proceedings*, 39, Article 58.
17. Nguyen, H. T., & Nguyen, D. T. (2020). Transfer learning for macroeconomic forecasting. In *2020 7th NAFOSTED Conference on Information and Computer Science (NICS)* (pp. 332-337). IEEE.
18. Mlinarič, J., Pregelj, B., Boškosi, P., Petrovčič, J., & Dolanc, G. (2025). A transfer learning approach to machine learning-based end-of-line quality inspection. *Advances in Production Engineering & Management*, 20(2), 277-290.



19. Laurer, M., van Atteveldt, W., Casas, A., & Welbers, K. (2024). Less annotating, more classifying: Addressing the data scarcity issue of supervised machine learning with deep transfer learning and BERT-NLI. *Political Analysis*, 32(1), 84-100.
20. Topsakal, O., & Akinci, T. C. (2023). A review of transfer learning: Advantages, strategies and types. In the *International Conference on Modern and Advanced Research* (pp. 390-394).
21. Alolayan, O. S., Raymond, S. J., Montgomery, J. B., & Williams, J. R. (2021). Towards better shale gas production forecasting using transfer learning. *arXiv preprint arXiv:2106.11051*.
22. Medeiros, M. C., Vasconcelos, G. F. R., Veiga, Á., & Zilberman, E. (2021). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *Journal of Business & Economic Statistics*, 39(1), 98-119.
23. Naghi, A. A., O'Neill, E., & Zaharieva, M. (2024). The benefits of forecasting inflation with machine learning: New evidence. *Journal of Applied Econometrics*, 39(7), 1321-1331.
24. Islam, M. M., Jannat, A., & Rahman, M. M. (2024). Forecasting foreign direct investment inflow to Bangladesh: Using an autoregressive integrated moving average and a machine learning-based random forest approach. *Journal of Risk and Financial Management*, 17(10), Article 451.
25. Thrun, S., & Pratt, L. (Eds.). (1998). *Learning to Learn*. Springer.
26. Torrey, L., & Shavlik, J. (2010). Transfer learning. In E. Soria Olivas, J. D. Martín Guerrero, M. Martínez Sober, J. R. Magdalena Benedito, & A. J. Serrano López (Eds.), *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques* (pp. 242–264). IGI Global.
27. Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, 3, Article 9.
28. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 31.