

Modeling Business Investment Attractiveness of Countries Using Time Series Approaches

1. Topic Definition

Business investment is a major pillar of economic growth and an important corporate decision. Lots of econometric research is devoted to assess the investment attractiveness of an economy^{[1][2]}. Our topic is to use a ML/DL time series techniques to turn it into a modeling pipeline in which we provide reliable insight.

2. Methods

To identify opportunities in structural transformation of sectors, we target Manufacturing and Service value added as a percentage of GDP of 9 countries, modeling it with indicators carefully selected from World Bank. 11 imputation techniques are experimented to deal with over 50% missingness, and 8 models are finally stacked to predict into future 5 years(2022-2026).

Step 1. Dataset Consolidation from World Bank

2 tasks
Manufacturing
Service

×

10+ features

17 for manufacturing
20 for service are carefully
selected from world bank
based on literature review
& correlation test

×

9 Countries

AUS, BRA, CAN
CHN, GBR, IND
JPN, SGP, USA

×

30 year of
data points

From 1980 or
later(earliest available
data for the country) to
2021

Table 1. Features of 2 tasks

Manufacturing Task			
No	Variable Name	Short Description	Unit
X ₁	Air_freight_million_ton_km	Air freight in million ton-km measured in metric tons times kilometer traveled.	Million ton-km
...
X ₁₇	Logistic_performance	The quality of the logistic related infrastructure. In this index, trade, and transport related infrastructure are evaluated.	Numerical index (1=lowest to 5=highest)
y	Manufacturing_value_added_%_of_GDP	Manufacturing refers to industries belonging to ISIC divisions 15-37.	Percentage
Service Task			
No	Variable Name	Short Description	Unit
X ₁	Educational_attainment	The percentage of people ages 25 and above who completed education in post-secondary non-tertiary level.	Percentage
...
X ₂₀	Population_labor_rate	The percentage of population ages 15-64	Percentage
y	Service_value_added_%_of_GDP	Services correspond to ISIC divisions 50-99 and they include value added in wholesale, retail trade, education ...	Percentage

Step 2. Data Preprocessing

Row Deletion

Applied for missing value of y before first valid y

Imputation

Use NOCB and Spline to impute, based on experiments upon a baseline LSTM model with lots of combinations of available imputation techniques:

- Last Observation Carried Forward (LOCF)
- Next Observation Carried Backward (NOCB)
- Rolling Statistics
- Linear interpolation
- Simple/Weighted/Exponential Moving Average
- Linear/Spline interpolation
- Multivariate Imputation by Chained Equation

Column Deletion

- Country-task-specific: imputation rate over a pre-determined threshold
- Model-specific: VAR requires stationary

Granger Causality Test

Step 3. Modeling & Evaluating

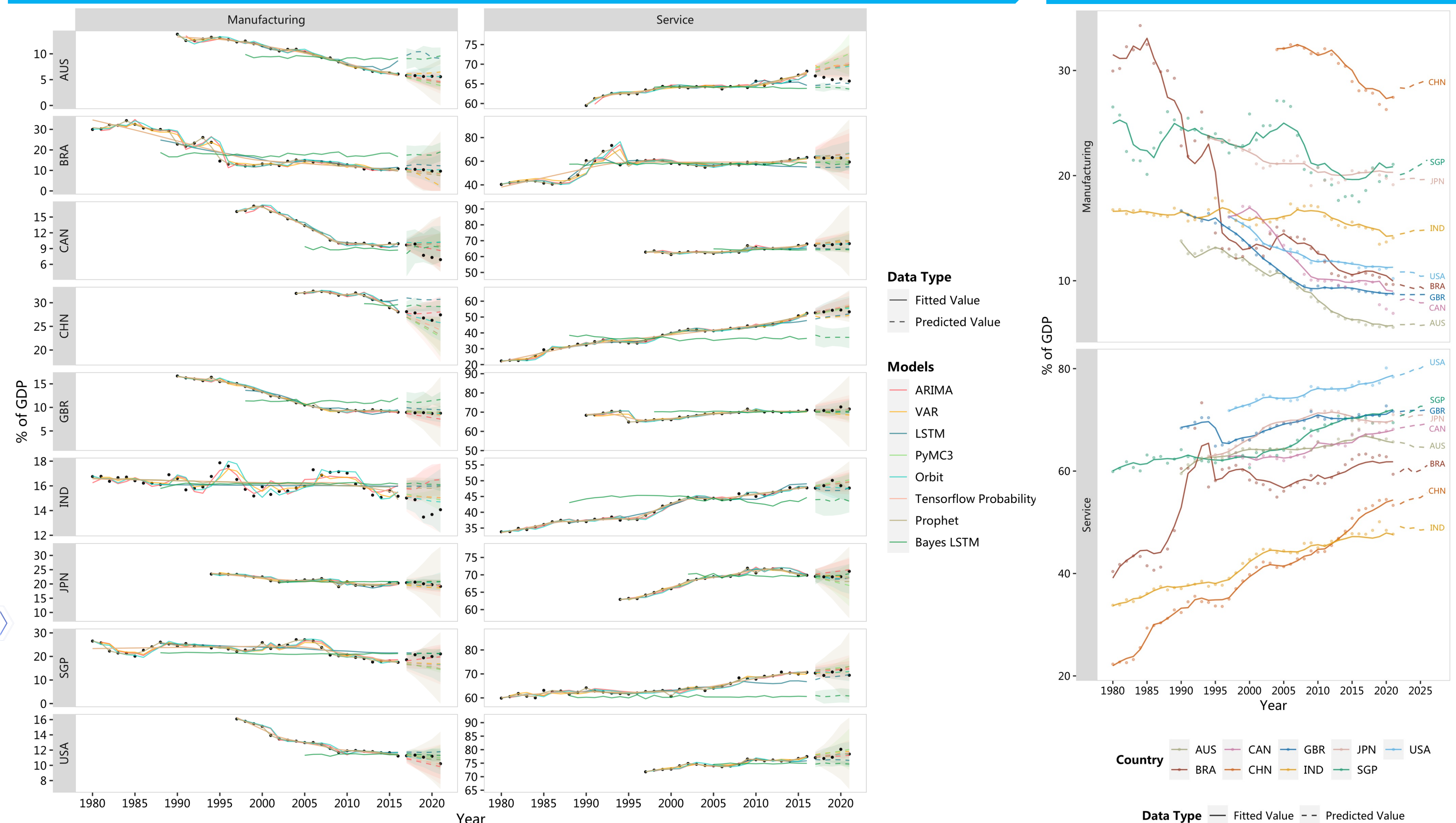


Figure 1. Model & Data pipelines and demo output for the project

3. Results

In the evaluation status, we found Orbit perform well in general in terms of MAPE. But it still vary from countries to countries and assessment metric.

Table 2. Average MAPE of all models in 2 tasks' test set (2017-2021)

Task	ARIMA	VAR	LSTM	PyMC3	Orbit	Tensorflow Probability	Prophet	Bayes LSTM
Manufacturing	0.075	0.225	0.138	0.114	0.062	0.088	0.118	0.170
Service	0.022	0.027	0.040	0.028	0.018	0.022	0.022	0.102

Inspired by Kurz, et al (2020)^[3], we stack the model country-task-wisely weighted on their performance, which is equally defined by 1/MAPE, 1/MAE and 1/RMSE.

Table 3. Average weight considered by final stacked model (averaged across countries)

Task	ARIMA	VAR	LSTM	PyMC3	Orbit	Tensorflow Probability	Prophet	Bayes LSTM
Manufacturing	14.56%	10.68%	9.26%	12.23%	24.17%	12.73%	10.63%	9.30%
Service	13.40%	13.42%	10.19%	11.58%	17.29%	13.91%	13.54%	8.16%

4. Conclusions

In this project, we carefully examine toolkits for infrequent aggregated time series with severe missing value issue. Base on predictions on the countries we selected, Singapore is the best for service sector investment and China best for manufacturing in the future 5 years.

Table 4. Difference of % of GDP compared between 2021 and 2026(forecasted)

Task	CHN	SGP	USA	CAN	IND	BRA	JPN	GBR	AUS
Service	2.458	3.196	2.546	1.096	0.949	1.727	-0.026	0.327	-1.117
Manufacturing	1.653	0.407	0.254	1.013	0.773	-0.586	0.511	-0.076	0.243

References

- [1] Ilyash, Olha, et al. "Evaluation of enterprise investment attractiveness under circumstances of economic development." Bulletin of Geography. Socio-economic Series 47.47 (2020): 95-113.
- [2] Bruneckiene, Jurgita, et al. "Assessment of investment attractiveness in European countries by artificial neural networks: What competences are needed to make a decision on collective well-being?." Sustainability 11.24 (2019): 6892.
- [3] Kurz, Christoph F., Werner Maier, and Christian Rink. "A greedy stacking algorithm for model ensembling and domain weighting." BMC research notes 13.1 (2020): 1-6.