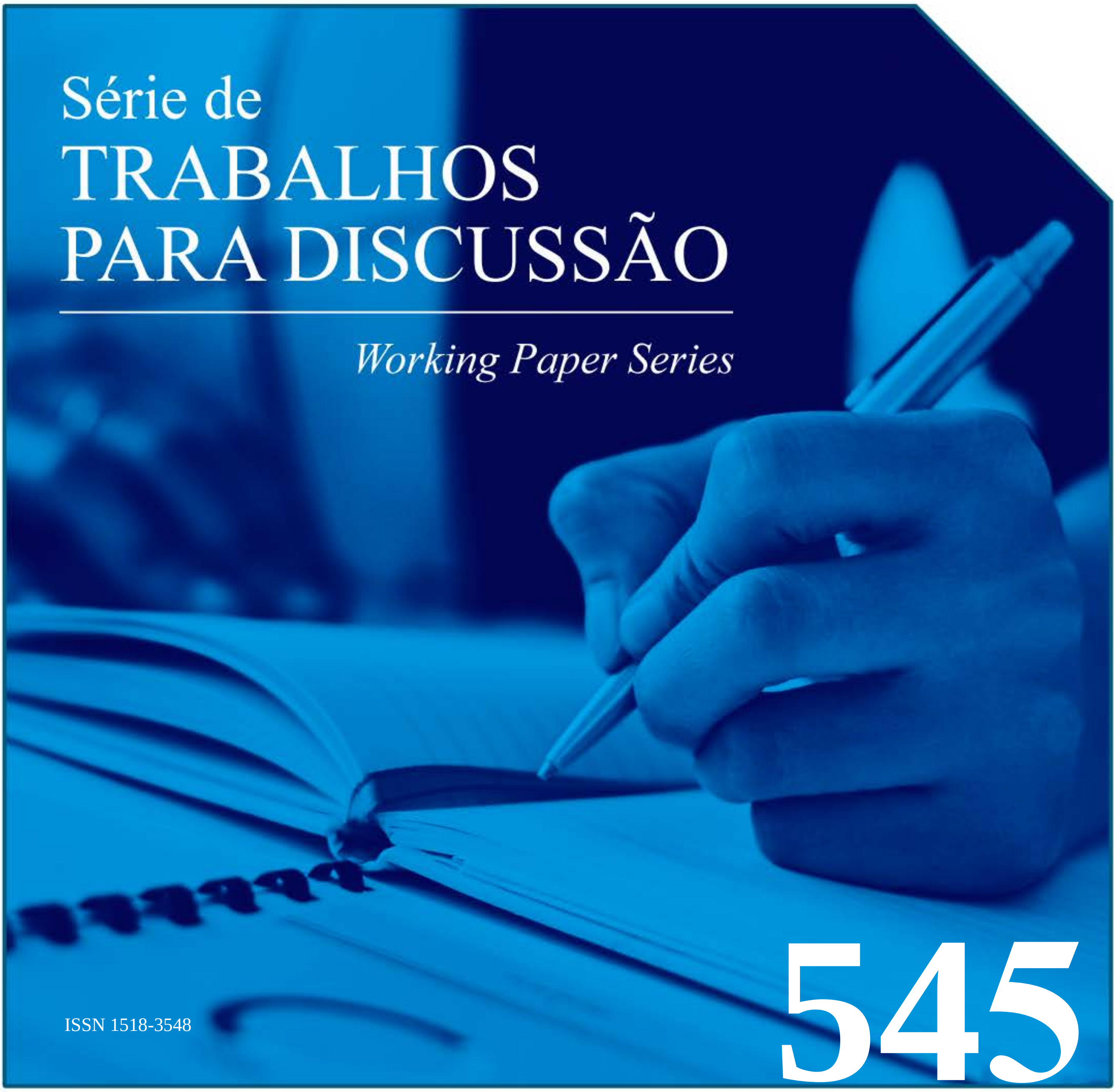

FEBRUARY 2021

Transfer Learning for Business Cycle Identification

Marcelle Chauvet, Rafael R. S. Guimaraes

BANCO CENTRAL
DO BRASIL

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	no. 545	February	2021	p. 1-25

Working Paper Series

Edited by the Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Francisco Marcos Rodrigues Figueiredo

Co-editor: José Valentim Machado Vicente

Head of the Research Department: André Minella

Deputy Governor for Economic Policy: Fabio Kanczuk

Citizen Service Division

Banco Central do Brasil

Deati/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Toll Free: 0800 9792345

Fax: +55 (61) 3414-2553

Internet: http//www.bcb.gov.br/?CONTACTUS

Non-technical Summary

Monitoring business cycle phases is a traditional task in applied macroeconomics. Progressive market integration has induced a worldwide interest in analyzing cyclical fluctuations using economic indicators. Changes in exchange rates, outputs, consumption, inflation, and interest rates in different parts of the world can influence the effectiveness of government policies and the competitive position of businesses, even those not directly related to international operations.

Can we use information from adult humans to train an intelligent system for diagnosing infant heart disease? Such a problem is known as domain adaptation or transfer learning. The population of interest is called the target domain, for which labels are usually not available, and training a classifier might not be possible. However, if data from a similar population is available, it could be used as a source of additional information.

In this work, we explore the transfer learning capability of artificial neural networks and propose a method that combines deep neural networks with transfer learning to identify business cycle phases when data is limited or there is no business cycle dating committee. The approach demonstrated excellent empirical performance with data from the US, Europe, and Brazil, emerging as a potential supplementary tool for governments and the private sector to conduct their activities in the light of national and international economic conditions.

Sumário Não Técnico

O monitoramento das fases do ciclo de negócios é uma tarefa tradicional em macroeconomia aplicada. A integração progressiva do mercado induziu um interesse mundial na análise das flutuações cíclicas por meio do uso de indicadores econômicos. Mudanças nas taxas de câmbio, PIB, consumo, inflação e taxas de juros em diferentes partes do mundo podem influenciar a eficácia das políticas governamentais e a posição competitiva das empresas, mesmo aquelas não diretamente relacionadas às operações internacionais.

Podemos usar informações de adultos para treinar um sistema inteligente para diagnosticar doenças cardíacas infantis? Esse problema é conhecido como adaptação de domínio ou aprendizagem por transferência. A população de interesse é chamada de domínio de destino, para a qual os exemplos com identificação, ou rótulos, geralmente não estão disponíveis, e treinar um classificador pode não ser possível. No entanto, se houver dados de uma população semelhante, eles podem ser usados como fonte de informações adicionais.

Neste trabalho, exploramos a capacidade de aprendizagem por transferência de redes neurais artificiais e propomos um método que combina redes neurais profundas com aprendizagem por transferência para identificar as fases do ciclo de negócios quando os dados são limitados ou na ausência de um comitê de datação do ciclo de negócios. A abordagem demonstrou excelente desempenho empírico com dados dos Estados Unidos, Europa e Brasil, emergindo como potencial ferramenta complementar para governos e o setor privado conduzirem suas atividades à luz das condições econômicas nacionais e internacionais.

Transfer Learning for Business Cycle Identification*

Marcelle Chauvet**

Rafael R. S. Guimaraes***

Abstract

A transfer learning strategy is proposed to identify business cycles phases when data are limited or there is no business cycle dating committee. The approach integrates the idea of storing knowledge gained from one region's economics experts and applying it to other geographic areas. The first is captured with a supervised deep neural network model, and the second by applying it to another dataset, a domain adaptation procedure. The results indicate the method proposed leads to successful business cycle identification.

Keywords: neural networks, business cycle, transfer learning, deep learning.
JEL Classification: C45, E32, E37.

* This article is part of the Guimaraes' dissertation Ph.D. at Universidade Federal do Rio Grande do Sul. The authors would like to thank the organizers of the V Workshop of the Research Network in Economics and Finance of the Central Bank of Brazil, for the comments received from the participants and for the contributions of the debater Wagner P. Gaglianone. Remaining errors and omissions are the responsibility of the authors.

** University of California Riverside. Email: chauvet@ucr.edu

*** Banco Central do Brasil. Email: rafael.guimaraes@bcb.gov.br

1. Introduction

Monitoring business cycle phases is a traditional task in applied macroeconomics. Progressive market integration has induced a worldwide interest in analyzing cyclical fluctuations through the use of economic indicators (Chauvet, 2001, p.21). Changes in exchange rates, outputs, consumption, inflation, and interest rates in different parts of the world can influence the effectiveness of government policies and the competitive position of businesses, even those not directly related to international operations (Chauvet and Yu, 2006, p.43). As a result, a wide range of techniques has been developed since the seminal work by Burns and Mitchell (1946). Recently, new approaches have emerged due to the progress in machine learning (ML) research, centering on building models that achieve better forecasting performance than the non-ML models or that identify turning points more timely (Piger, 2020).

In this work, we contribute to the literature by exploring the transfer learning capability of artificial neural networks (Pratt et al., 1991). This characteristic has not yet been evaluated to monitor business cycle phases to the best of our knowledge. Our goal follows Chauvet and Yu (2006) in providing additional tools for governments and the private sector to conduct their activities in light of both national and international economic conditions. For that, we adopt a combined strategy of deep neural network and transfer learning to address the practical problem of identifying the business cycle phases when data is limited or a business cycle dating committee is absent. Deep learning is a sub-field within machine learning that is based on algorithms for learning multiple levels of representation in order to model complex relationships among data (Deng and Yu, 2014). Transfer learning improves learning in a new task by transferring knowledge from a related task that has already been learned (Torrey and Shavlik, 2009). More specifically, domain adaptation can be considered a special set of transfer learning that aims at transferring shared knowledge across different but related tasks or domains. A good feature representation should be able to reduce the difference in distributions between domains as much as possible, while at the same time preserving essential properties of the original data (Pan et al., 2011).

The rest of the paper is organized as follows. Section 2 presents a literature review, and in Section 3, we discuss the methodology. Section 4 presents our empirical findings, and Section 5, the final remarks.

2. Literature

Business cycles are recurrent sequences of alternating phases of expansion and contraction among many economic activities (Burns and Mitchell, 1946). According to Harding and Pagan (2005), there are three ways in the literature to describe what we mean by a cycle, depending on whether the main focus is on the fluctuation of the level of economic activity, the level of economic activity less a permanent component, or the growth rate of economic activity.

In the United States, the National Bureau of Economic Research (NBER) Business Cycle Dating Committee provides a chronology of business cycle expansion and recession dates. According to Piger (2020), because the NBER methodology is not explicitly formalized, literature has worked to develop and evaluate formal statistical methods for establishing the historical dates of economic recessions and expansions in both the U.S. and international data. Estrella and Mishkin (1998), Estrella et al. (2000), Kauppi and Saikkonen (2008), Rudebusch and Williams (2009) and Fossati (2016) use an available historical indicator of the class, such as the NBER dates, to estimate the parameters of models such as logit or probit ones. This strategy is called a supervised classifier in the statistical classification literature, in contrast to unsupervised classifiers, which endogenously determine the classes. Unsupervised classifiers have also been used, with the primary example being the Markov-switching (MS) framework of Hamilton (1989), which become a relevant tool for applied work in economics. Chauvet (1998) proposes a dynamic factor model with Markov-switching (DFMS) to identify expansion and recession phases from a group of coincident indicators and Chauvet and Hamilton (2005), Chauvet and Piger (2008) and Camacho et al. (2018) evaluate the performance of variants of this DFMS model to identify NBER turning points in real time. See Piger (2020) for a comprehensive review.

Recently, artificial intelligence (AI) has gained considerable prominence due to performances in autonomous vehicles, intelligent robots, image and speech recognition, automatic translations, medical and law usage (Makridakis, 2017). In Economics, the application of machine learning (ML) methods, an AI technique, is not new, and in a way it has followed the phases of use in other areas. This has extended from the earliest attempts in the 1940s, followed by the rising expectations and the results in the 1960s, through the period of frustration in the 1970s, to the continuity of its use by a small group of researchers in the 1980s, and the resurgence in the 1990s (Stergiou and Siganos, 2011).

Finally, from the beginning of the 21st century, significant progress has been observed in many areas, attracting attention and funding for research.

Applied ML papers related to business cycles can be separated depending on whether the main focus is *predicting* or *identifying* turning points and phases. For example, Hoptroff et al. (1991), Qi (2001), Klinkenberg (2003), Nasr et al. (2007), Berge (2013), Ma (2015), Garbellano (2016), Nyman and Ormerod (2017), and James et al. (2019) have applied machine learning techniques such as artificial neural networks, support vector machines, boosting, k-nearest neighbor, and random forest to *forecasting* turning points, recessions, or business cycles phases mainly in the US, but also other countries[1]. These studies have generally reported some improvements over non-ML strategies. The other set of papers is concerned about *identifying* the turning points for real-time classification. Morik and Ruping (2002), Giusto and Piger (2017), Soybilgen (2018), Raffinot and Benoit (2019) and Jackson and Rege (2019) have applied inductive logic programming, learning vector quantization, random forest, boosting, k-nearest neighbor and artificial neural networks fed with dynamic factors. Piger (2020), in a comprehensive analysis, compares five ML techniques with DFMS. These studies have reported quickly and accurately turning points identification.

Lastly, some literature is dedicated to the study of business cycles worldwide, as in Chauvet and Yu (2006), Cuba-Borda et al. (2018), Abberger et al. (2020), and the reference turning points of the OECD Composite Leading Indicators[2].

3. Methodology

Due to the availability of various ML methods and considering that we are, especially in economics, in the explanatory era of its applications, works often apply several ML approaches to a specific dataset to compare their performances, a strategy known as horse-race, as in Tiffin (2016), Cook and Hall (2017), Garcia et al. (2017), Gu et al. (2018), and Piger (2020). Makridakis et al. (2018) go further to compare various non-ML and ML forecasting methods. Another usual approach, our choice, is to select in advance a suitable strategy for the specific task. Deep learning seems well suited to transfer learning because it focuses on learning representations and, in particular, on

[1] United Kingdom, Japan, West Germany and Lebanon.

[2] Available at https://www.oecd.org/sdd/leading-indicators.

abstract representations, which ideally disentangle the factors of variation present in the input (Bengio, 2012). Learning representations of the data make it easier to extract useful information when building classifiers or other predictors, and, in the case of probabilistic models, a good representation is often one that captures the posterior distribution of the underlying explanatory factors for the observed input (Bengio et al., 2013).

3.1 Deep neural network

A deep neural network, also known as deep learning (DL), is an artificial neural network (ANN) with multiple layers hidden between the input and output layers (Bengio, 2009). The analytical function corresponding to one of the simplest forms of an ANN, the feed-forward network, can be written as follows (Bishop, 1994, 118-9). In a feed-forward network having two layers, there are d inputs, M hidden units and c output units. The output of the jth hidden unit is obtained by first forming a weighted linear combination of the d input values, and adding a bias, to give

$$a_j = \sum_{i=1}^{d} w_{ji}^{(1)} x_i + w_{j0}^{(1)}. \tag{1}$$

Here $w_{ji}^{(1)}$ denotes a weight in the first layer, going *from* input i *to* hidden unit j, and $w_{j0}^{(1)}$ denotes the bias for hidden unit j. The bias term for the hidden units is made explicit by the inclusion of an extra input variable x_0 whose value is permanently set at x_0 = 1. This can be represented analytically by rewriting (1) in the form

$$a_j = \sum_{i=0}^{d} w_{ji}^{(1)} x_i. \tag{2}$$

The activation of hidden unit j is then obtained by transforming the linear sum in (2) using an activation function $g(\cdot)$ to give

$$z_j = g(a_j). \tag{3}$$

The outputs of the network are obtained by transforming the activations of the hidden units using a second layer of processing elements. Thus, for each output unit k, we construct a linear combination of the outputs of the hidden units of the form

$$a_k = \sum_{j=1}^{M} w_{kj}^{(2)} z_j + w_{k0}^{(2)}. \tag{4}$$

Again, we can absorb the bias into the weights to give

$$a_k = \sum_{j=0}^{M} w_{kj}^{(2)} z_j, \tag{5}$$

which can be represented by including an extra hidden unit with activation $z_0 = 1$. The activation of the *k*th output unit is then obtained by transforming this linear combination, using a non-linear activation function, to yield

$$y_k = \tilde{g}(a_k). \tag{6}$$

Here we have used the notation $\tilde{g}(\cdot)$ for the activation function of the output units to emphasize that this need not be the same function as used for the hidden units. If we combine (2), (3), (5) and (6), we obtain an explicit expression for the complete function in the form

$$y_k = \tilde{g}\left(\sum_{j=0}^{M} w_{kj}^{(2)}\, g\left(\sum_{i=0}^{d} w_{ji}^{(1)}\, x_i\right)\right). \tag{7}$$

These models are called feed-forward because information flows through the function being evaluated from inputs, through the intermediate computations used to define the function, and finally to the output target. There are no feedback connections in which the outputs of the models are fed back into itself. When processing sequential data is required, there are alternatives, like recurrent neural networks (RNN), including the successful long short-term memory (LSTM) model. RNN process an input sequence one element at a time, maintaining in their hidden units a *state vector* that implicitly contains information about the history of all past elements of the sequence (LeCun et al., 2015). For technical details, see (Goodfellow et al., 2016, 367-415).

In essence, almost all DL algorithms can be described as a combined specification of a data set, a cost function, an optimization procedure, and a model (Goodfellow et al., 2016). We describe ours in subsections 3.4 and 3.5.

3.2 Transfer learning

Can we use information from *adult* humans to train an intelligent system for diagnosing *infant* heart disease? Such a problem is known as *domain adaptation* or *transfer learning*. The population of interest is called the target domain, for which labels are usually not available and training a classifier might be not possible. However, if data from a similar population is available, it could be used as a source of additional information (Kouw and Loog, 2018, 2). Thus, transfer learning (TL) refers to the situation where what has been learned in one setting (e.g., distribution P1) is exploited to improve generalization in another setting (say, distribution P2). The learner must perform two or more different tasks, but it is assumed that many of the factors that explain the variations in P1 are relevant to the variations that need to be captured for learning P2 (Goodfellow et al., 2016, 534).

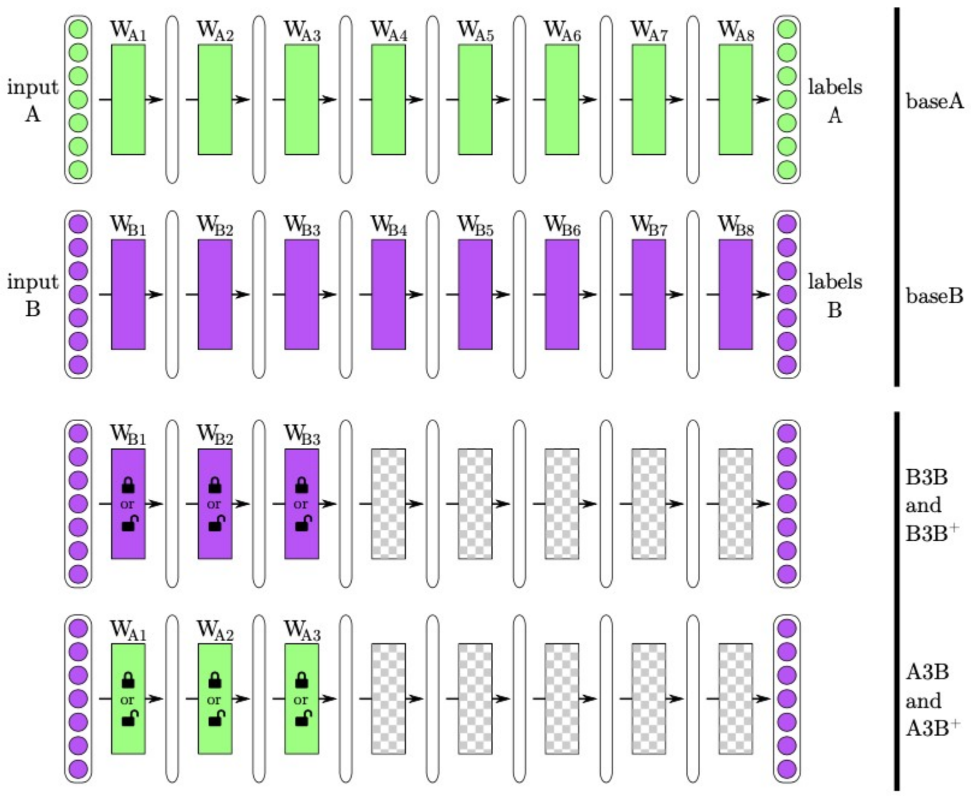

Figure 1: Transfer learning overview - reproduced from Yosinski et al. (2014)

In the conventional transfer learning approach, we first train a base network on a base dataset and task, and then we repurpose the learned features or transfer them to a second target network to be trained on a target dataset and task. This process will tend to work if the features are general, meaning suitable to both base and target tasks, instead of specific to the base task (Yosinski et al., 2014). Figure 1 illustrates these dynamics. The base networks (top two rows) are trained using standard deep learning procedures on datasets A and B. The labeled rectangles (e.g., WA1) represent the weight vector learned for that layer, with the color indicating which dataset the layer was originally trained on. The vertical, ellipsoidal bars between weight vectors represent the activations of the network at each layer. The target networks (bottom two rows) represent transfer learning strategies. The first n weight layers of the network (in this example, n = 3) are copied from a network trained on one dataset (e.g., A), and then the entire network is trained on the other dataset (e.g., B). Usually, the first n layers are either locked during training or allowed to learn.

3.3 Some assumptions

As in the dynamic factor model with Markov-switching introduced by Chauvet (1998), the deep neural network approach accounts for the idea of business cycles as the simultaneous, asymmetrical, and nonlinear movement of economic activity in various sectors. This data-driven framework is flexible enough to be training with different features, the independent variables, and targets, the dependent variables. A supervised deep learning model learns from the features-targets without the need for strong assumptions about its relation, which prevents the expert from choosing an underlying economics school of thought to set up a model, although the variable selection might represent it. The algorithm maps the input-output relation variables according to the training and validation data. For instance, once we choose NBER [3] turning points classification data as output label, the deep learning algorithm will learn from them how to classify the business cycle, implicitly following the same school of thought.

When we transfer learning, though, there are additional issues. According to Kouw and Loog (2018), the challenge is to overcome the differences between the domains so that a classifier trained on the source domain generalizes well to the target domain, but

[3] The most followed classification for U.S. business cycle.

generalizing across distributions can be difficult, and it is not clear which conditions have to be satisfied for a classifier to perform well. From a macroeconometrics perspective, we could assume that the input variables' marginal contributions in explaining the business cycle phases are similar. However, they are probably not equal for the different economic areas, as well as the respective input variables' sample covariance matrix. Our approach handles this well when the models successfully build a common representation space for the different domains, which can be empirically verified when there are labels for the target domain, allowing evaluation by some error criteria. Otherwise, it remains an open question.

3.4 Data

A deep learning model is capable of handling a large number of explanatory variables (features). However, there is a caveat to our strategy: the more features we use during learning, the more data preprocessing efforts will be required. For example: if we train for the United States business cycles classification using dozens of features, we will need the same quantity of features to transfer learning to other datasets or additional preprocessing strategies, as in Jackson and Rege (2019), which have fed an ANN with dynamic factors. Because it adds challenges, dealing with high-dimensional data is left for future work.

The feature selection comprises the coincident variables indicated by NBER[4] as the fundamental: gross domestic product (GDP), income, employment, industrial production, and wholesale-retail sales. Quarterly data is adopted because this is the frequency at which some relevant variables for the classification of the business cycle are available, and at this frequency, opposed to higher ones, the data usually carry less noise, what may facilitate the training and the transfer learning. We computed the first difference of the logarithm of the input features, capturing the growth rate (Harding and Pagan, 2005, 152-154). Alternatively, we run the model without this transformation, i.e., features in level. In both cases, the features are normalized.

[4] The NBER does not define a recession in terms of two consecutive quarters of decline in real GDP. Rather, a recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales. Source: https://www.nber.org/cycles.html.

Table 1 - Dataset

FRED Code	Series	Start	Characteristics of the raw data	Source
	UNITED STATES			
USRECQ	NBER based Recession Indicator	1967:Q1	Recession: 1 = true ; 0 = false	NBER
GDPC1	Real Gross Domestic Product	1967:Q1	Billions of Chained 2012 Dollars, s.a.	FRED
PIECTR	Real personal income ex current transfers	1967:Q1	Billions of Chained 2012 Dollars, s.a.	FRED
PRS85006013	Nonfarm Business Sector employment index	1967:Q1	Index 2012 = 100, s.a.	FRED
IPB50001SQ	Industrial production index	1967:Q1	Index 2012 = 100, s.a.	FRED
CQRMTSPL	Real manufacturing and trade ind. sales	1967:Q1	Millions of Chained 2012 Dollars, s.a.,	FRED
	EURO AREA			
N/A	CEPR based Recession Indicator	2005:Q1	Recession: 1 = true — 0 = false	CEPR
CLVMNACSCAB1GQEA19	Real Gross Domestic Product (19 countries)	2005:Q1	Millions of Chained 2010 Euros, s.a.	FRED
NAEXKP02EZQ189S	Private Final Consumption Expenditure	2005:Q1	Billions of Chained 2012 Dollars, s.a.	FRED
LFESEETTEZQ647S	Employees	2005:Q1	Persons, s.a.	FRED
PRMNTO01EZQ657S	Total Manufacturing Production	2005:Q1	Growth Rate Previous Period, s.a.	FRED
SLRTTO01EZQ657S	Volume of Total Retail Trade sales	2005:Q1	Growth Rate Previous Period	FRED
	BRAZIL			
N/A	CODACE based Recession Indicator	2000:Q1	Recession: 1 = true — 0 = false	CODACE
NAEXKP01BRQ652S	Total Gross Domestic Product	2000:Q1	Chained 2000 Real, s.a.	FRED
NAEXKP02BRQ189S	Private Final Consumption Expenditure	2000:Q1	Chained 2000 Real, s.a.	FRED
N/A	Registered Employees Index	2000:Q1	Index Dez-2001 = 100, s.a.	BCB
BRAPROINDQISMEI	Production of Total Industry	2000:Q1	Index 2015 = 100, s.a.	FRED
BRASARTQISMEI	Total Retail Trade	2000:Q1	Index 2015 =100, s.a.	FRED

To have a common starting point for each dataset, we restrict the series' start to eliminate missing values. For example, we do not acquire data for the Euro area before 2005 because we do not have employment data before this year for all the selected countries. We adopted a U.S. dataset for deep learning and two datasets, with data from Brazil and Europe, for transfer learning. The target values are the business cycle chronology provided by the NBER, the Brazilian Business Cycle Dating Committee (CODACE), and the CEPR-EABCN Euro Area Business Cycle Dating Committee (CEPR), respectively. Table 1 summarizes the information about all series, mostly from the FRED-MD[5] dataset, provided by the Federal Reserve Bank of St. Louis. Our data files are available at *https://github.com/rrsguim/PhD_Economics*.

3.5 Implementation Details

The models were built using TensorFlow[6], an interface for expressing machine learning algorithms, and an implementation for executing such algorithms. TensorFlow is flexible and can be used to express a wide variety of algorithms, including training and inference algorithms for deep neural network models. It has been used to conduct research and deploy machine learning systems into production across more than a dozen areas of computer science and other fields (Abadi et al., 2015).

The primary architectural considerations are choosing the depth of the network and the width of each layer. Deeper networks are often able to use far fewer units per layer and far fewer parameters, as well as frequently generalizing to the test set, but they also tend to be harder to optimize. The ideal network architecture for a task must be found via experimentation guided by monitoring the validation set error (Goodfellow et al., 2016, 194). The optimal set of hyperparameters was obtained using *Hyperband* (Li et al., 2018) from Keras Tuner[7]. Following Piger (2020), the objective function to maximize in the cross-validation step is the area under the ROC curve (AUC). This metric is desirable here for being scale-invariant, measuring how well predictions are ranked, rather than their absolute values, and classification-threshold-invariant, measuring the quality of the model's predictions irrespective of what classification threshold is chosen.

[5] https://fred.stlouisfed.org/.
[6] https://www.tensorflow.org/.
[7] https://www.tensorflow.org/tutorials/keras/keras_tuner.

Beginning with the deep learning step, we split the U.S. dataset into train, validation, and test sets. Then, we define a function that creates a neural network with hidden layers, ReLU as activation function, a dropout layer to reduce overfitting, and a sigmoid output layer that returns the probability of recession. Next, we retrain the model with the optimal hyperparameters, selected with *Hyperband*, to evaluate the results in both datasets, source and target, with binary cross-entropy as a loss function and Adam (Kingma and Ba, 2017) for optimization.

3.6 Robustness

Our baseline models are a logistic model (Logit) and deep learning models without the transfer learning approach.

The option for a feed-forward network as a deep learning model, which represents memoryless models, derives from the focus on contemporary movement between the selected variables and the business cycle. This implies disregarding the time dependence observed on the variables and shuffling the data to break it. The resulting model accounts just for coincident relations. To account for time dependency as an alternative strategy, we create additional models including a LSTM layer.

When transfer learning, the weight layers of the network for Euro and Brazillian data were copied from the network trained on the U.S. data, as in the last row of Figure 1, except that we do not retrain the parameters. It is as if these two datasets function as out-of-sample. Additionally, we unlock the last layers and retrain the parameters, a way of relaxing the macroeconometrics assumptions about input variables as mentioned in subsection 3.3.

Finally, we retrained the models with other loss function, the squared hinge, finding similar results.

4. Results

The deep learning models learned how to classify business cycles. Figures 2, 3, and 4 consolidate the results found. All codes used are available at *https://github.com/rrsguim/PhD_ Economics*. Results are slightly different each time the models are run due to different compositions in the selected data sets for training, validation and test, especially in the cross-sectional approach.

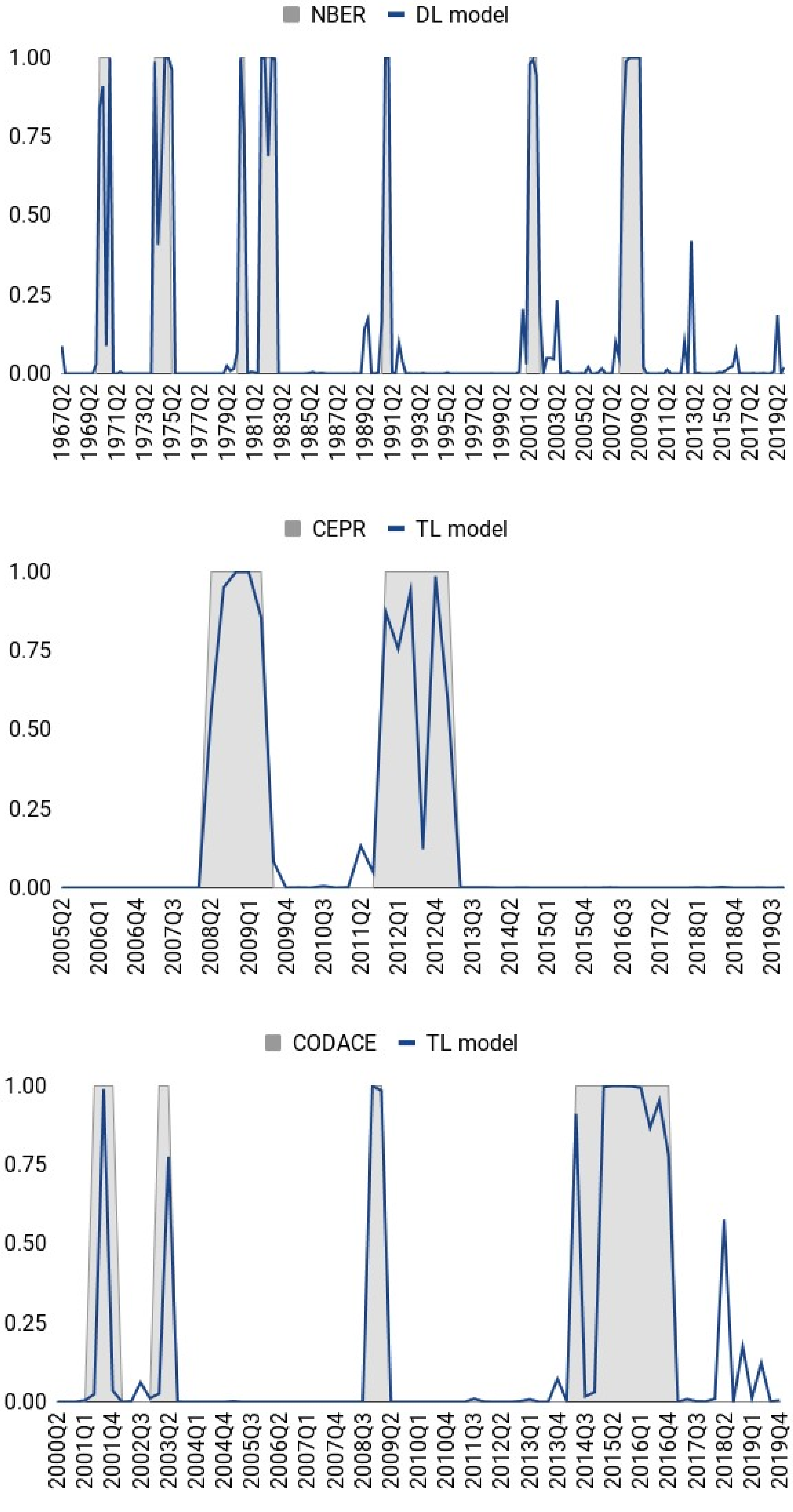

Figure 2: Deep Learning (U.S) and Transfer Learning (Euro and Brazil)

	Dropout	Hidden layers		*Hyperband* best hyperparameters		Target	Out-of-sample					
							Test size	Confusion matrix				AUC
		LSTM	Dense	units	learning rate			TN	FP	FN	TP	
Baseline models												
Logit_1df_US	-	-	-	-	-	NBER	77	36	0	32	9	0.610
DL_FNN_1df_US	0.5	0	3	208	0.01	NBER	64	56	0	2	6	0.875
DL_LSTM_1df_US	0.5	1	4	176	0.01	NBER	85	71	5	0	9	**0.967**
DL_FNN_level_US	0.5	0	4	176	0.01	NBER	64	54	0	10	0	0.500
DL_LSTM_level_US	0.5	1	4	176	0.01	NBER	85	76	0	9	0	0.500
DL_FNN_1df_EUR	0.5	0	3	112	0.01	CEPR	18	15	0	0	3	1.000
DL_LSTM_1df_EUR	0.5	1	4	48	0.0001	CEPR	42	36	0	6	0	0.500
DL_FNN_1df_BR	0.5	0	3	64	0.01	CODACE	24	18	1	3	2	0.674
DL_LSTM_1df_BR	0.5	1	4	240	0.001	CODACE	32	17	4	1	10	0.859
Transfer learning models												
TL_FNN_1df_EUR_locked	0.5	0	3	208	0.01	CEPR	60	48	1	2	9	0.899
TL_FNN_1df_EUR_unlocked	0.5	0	4	208	0.01	CEPR	42	36	0	6	0	0.500
TL_LSTM_1df_EUR_locked	0.5	1	4	176	0.01	CEPR	60	48	1	1	10	**0.944**
TL_LSTM_1df_EUR_unlocked	0.5	1	4	64	0.01	CEPR	42	36	0	6	0	0.500
TL_FNN_1df_BR_locked	0.5	0	3	208	0.01	CODACE	80	60	2	6	12	0.817
TL_FNN_1df_BR_unlocked	0.5	0	4	208	0.01	CODACE	24	16	3	1	4	0.821
TL_LSTM_1df_BR_locked	0.5	1	4	176	0.01	CODACE	80	61	1	3	15	**0.909**
TL_LSTM_1df_BR_unlocked	0.5	1	4	64	0.01	CODACE	48	31	4	2	11	0.866

TN - true negative | FP - false positive | FN - false negative | TP - true positive | AUC - area under the ROC curve.

Figure 3: Models specifications and results

Figure 2 shows the comparison between business cycles according to NBER, CEPR, and CODACE dating and those estimated by feed-forward neural networks (FNN) models with the cross-sectional approach. The graphics show the excellent performances of both the deep learning (NBER) and the transfer learning (CEPR and CODACE) steps. It should be noted that they refer to the locked models, meaning the estimates for EURO and Brazil operate as if they were out-of-sample because the parameters trained with U.S data are locked in the transfer learning phase when applied to target datasets (EURO and Brazil). Figure 3 presents the details of each model. Differences in the sizes of out-of-sample data sets reflect the need for adjustments according to time series, cross-sectional, locked and unlocked strategies. Concerning the baseline models with U.S. data, we observed a significantly superior performance, measured by the AUC with out-of-sample data, of the deep learning models with data in first difference (1df). The outcomes of the baseline models for EURO do not motivate confidence. The FNN model has a perfect classification (AUC = 1), while the alternative model that includes an LSTM layer results in a model unable to identify crises (AUC = 0.5). This point highlights one of the problems that the proposed methodology seeks to solve: identifying business cycles when data is limited.

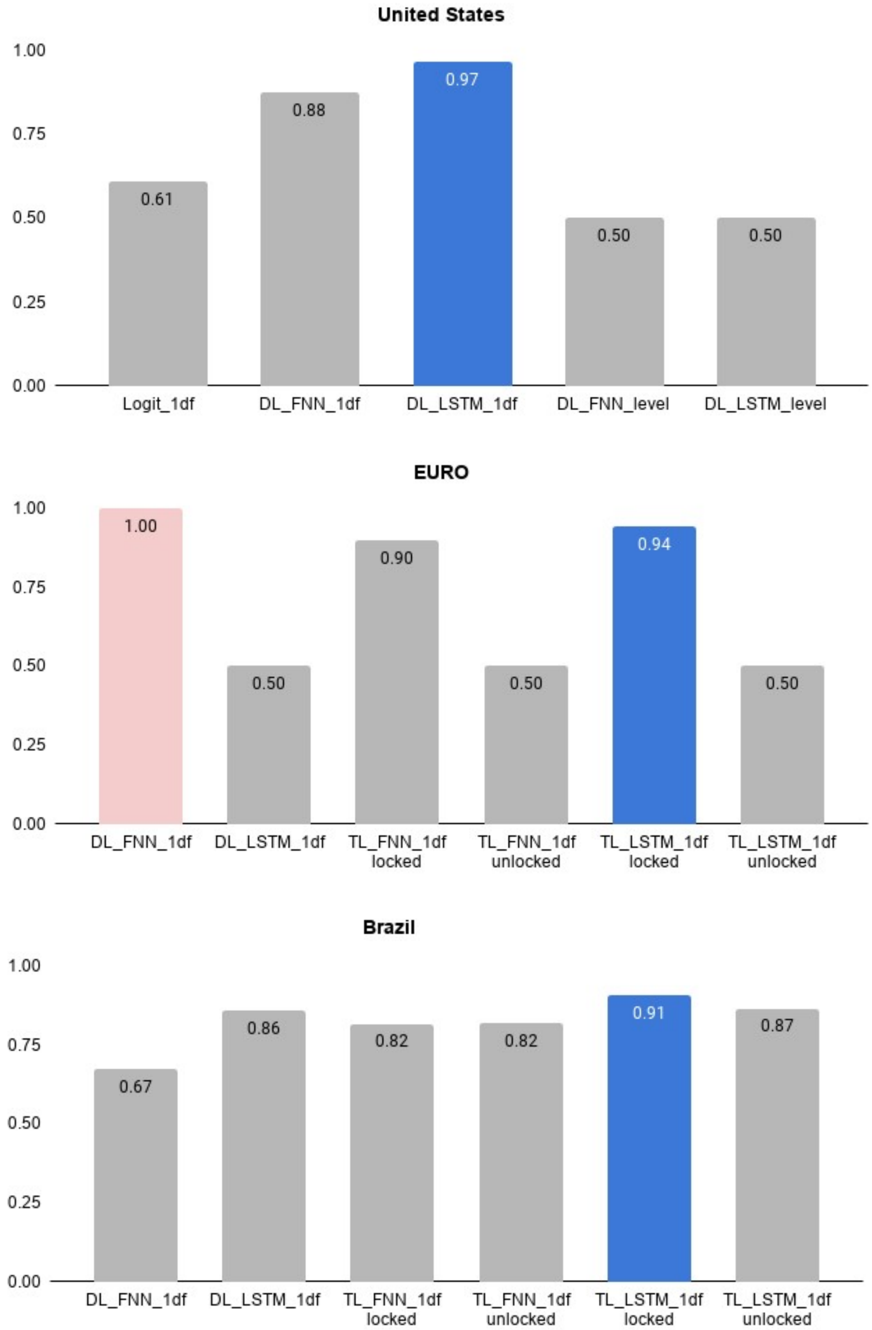

Figure 4: AUC out-of-sample

Note, in Figure 3, that there are only two recessions for the EURO in the period under analysis so that models trained only on these data show inconsistent results.

Regarding Brazil's outcomes, whose period contains more than one recession, the baseline models show more satisfactory performance. Finally, about transfer learning models, there is an improvement in the classification of business cycles for EURO and Brazil, compared to baseline models, with emphasis on the LSTM locked models. In addition to higher AUC, the fact that they are locked models allows more data in the out-of-sample set, increasing confidence in the results. Also, locked models can be applied when there is no business cycle dating committee. Figure 4 presents a summary comparison of the performance of each model.

5. Discussion

This paper has proposed a method that combines deep neural networks with transfer learning to identify business cycle phases when data is limited or in the absence of a business cycle dating committee. The approach demonstrated excellent empirical performance with data from the US, Europe, and Brazil, emerging as a potential supplementary tool for governments and the private sector to conduct their activities in the light of national and international economic conditions. To the best of our knowledge, the combined deep and transfer learning approach is underused for application to economic problems, indicating that there is plenty of room for research development.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Man´e, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Vi´egas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. URL: https://www.tensorflow.org/. software available from tensorflow.org.

Abberger, K., Graff, M., Campelo, A.J., Gouveia, A.C.L., Mu¨ller, O., Sturm, J.E., 2020. The Global Economic Barometers: Composite indicators for the world economy. Research Working Paper KOF WP 471-20. KOF Swiss Economic Institute.

Bengio, Y., 2009. Learning Deep Architectures for AI. Foundations and Trends in Machine Learning 2, 1–127.

Bengio, Y., 2012. Deep learning of representations for unsupervised and transfer learning, in: Guyon, I., Dror, G., Lemaire, V., Taylor, G., Silver, D. (Eds.), Proceedings of ICML Workshop on Unsupervised and Transfer Learning, JMLR Workshop and Conference Proceedings, Bellevue, Washington, USA. pp. 17–36.

Bengio, Y., Courville, A., Vincent, P., 2013. Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence 35, 1798– 1828.

Berge, T.J., 2013. Predicting recessions with leading indicators: model averaging and selection over the business cycle. Research Working Paper RWP 13-05. Federal Reserve Bank of Kansas City.

Bishop, C.M., 1994. Neural Networks for Pattern Recognition. Oxford University Press.

Burns, A.F., Mitchell, W.C., 1946. Measuring Business Cycles. National Bureau of Economic Research, Inc.

Camacho, M., Perez-Quiros, G., Poncela, P., 2018. Markov-switching dynamic factor models in real time. International Journal of Forecasting 34, 598–611.

Chauvet, M., 1998. An econometric characterization of business cycle dynamics with factor structure and regime switching. International Economic Review 39, 969–96.

Chauvet, M., 2001. A monthly indicator of brazilian gdp. Brazilian Review of Econometrics 21, 1–47.

Chauvet, M., Hamilton, J.D., 2005. Dating Business Cycle Turning Points. NBER Working Papers. National Bureau of Economic Research, Inc.

Chauvet, M., Piger, J., 2008. A comparison of the real-time performance of business cycle dating methods. Journal of Business Economic Statistics 26, 42–49.

Chauvet, M., Yu, C., 2006. International business cycles: G7 and OECD countries. Economic Review 91, 43–54.

Cook, T.R., Hall, A., 2017. Macroeconomic Indicator Forecasting with Deep Neural Networks. Research Working Paper RWP 17-11. Federal Reserve Bank of Kansas City.

Cuba-Borda, P., Mechanick, A., Raffo, A., 2018. Monitoring the World Economy : A Global Conditions Index. IFDP Notes. Board of Governors of the Federal Reserve System (U.S.).

Deng, L., Yu, D., 2014. Deep Learning: Methods and Applications. Technical Report. Microsoft.

Estrella, A., Mishkin, F., 1998. Predicting u.s. recessions: Financial variables as leading indicators. The Review of Economics and Statistics 80, 45–61.

Estrella, A., Rodrigues, A.P., Schich, S., 2000. How stable is the predictive power of the yield curve? evidence from Germany and the United States. Staff Reports 113. Federal Reserve Bank of New York.

Fossati, S., 2016. Dating US business cycles with macro factors. Studies in Nonlinear Dynamics & Econometrics 20, 529–547.

Garbellano, J., 2016. Nowcasting recessions with machine learning: New tools for predicting the business cycle, in: Thesis.

Garcia, M.G.P., Medeiros, M.C., Vasconcelos, G.F.R., 2017. Real-time inflation forecasting with high-dimensional models: The case of brazil. International Journal of Forecasting 33, 679–693.

Giusto, A., Piger, J., 2017. Identifying business cycle turning points in real time with vector quantization. International Journal of Forecasting 33, 174–184.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.

Gu, S., Kelly, B., Xiu, D., 2018. Empirical asset pricing via machine learning. Chicago Booth Research Paper 18.

Hamilton, J.D., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. Econometrica 57, 357–384.

Harding, D., Pagan, A., 2005. A suggested framework for classifying the modes of cycle research. Journal of Applied Econometrics 20, 151–159.

Hoptroff, R.G., Bramson, M.J., Hall, T.J., 1991. Forecasting economic turning points with neural nets, in: IJCNN-91-Seattle International Joint Conference on Neural Networks, pp. 347–352.

Jackson, B., Rege, M., 2019. Machine learning for classification of economic recessions, in: 2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI), pp. 31–38.

James, A., Abu-Mostafa, Y.S., Qiao, X., 2019. Nowcasting recessions using the svm machine learning algorithm. arXiv:1903.03202.

Kauppi, H., Saikkonen, P., 2008. Predicting u.s. recessions with dynamic binary response models. The Review of Economics and Statistics 90, 777–791.

Kingma, D.P., Ba, J., 2017. Adam: A method for stochastic optimization. arXiv:1412.6980.

Klinkenberg, R., 2003. Predicting phases in business cycles under concept drift.

Kouw, W., Loog, M., 2018. An introduction to domain adaptation and transfer learning. CoRR abs/1812.11806. URL: http://arxiv.org/abs/1812.11806.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444.

Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., Talwalkar, A., 2018. Hyperband: A novel bandit-based approach to hyperparameter optimization. arXiv:1603.06560.

Ma, J.B., 2015. Applications of machine learning in forecasting recessions: boosting united states and japan.

Makridakis, S., 2017. The forthcoming artificial intelligence (ai) revolution: Its impact on society and firms. Futures 90, 46–60.

Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2018. Statistical and machine learning forecasting methods: Concerns and ways forward. PLoS ONE 13.

Morik, K., Ru¨ping, S., 2002. A multistrategy approach to the classification of phases in business cycles, in: Elomaa, T., Mannila, H., Toivonen, H. (Eds.), Machine Learning: ECML 2002, Springer Berlin Heidelberg. pp. 307–318.

Nasr, G.E., Dibeh, G., Achkar, A., 2007. Predicting business cycle turning points with neural networks in an information-poor economy, in: SCSC.

Nyman, R., Ormerod, P., 2017. Predicting economic recessions using machine learning algorithms. arXiv:1701.01428.

Pan, S.J., Tsang, I.W., Kwok, J.T., Yang, Q., 2011. Domain adaptation via transfer component analysis. IEEE Transactions on Neural Networks 22, 199–210.

Piger, J., 2020. Turning points and classification, in: Fuleky, P. (Ed.), Macroeconomic Forecasting in the Era of Big Data. Springer International Publishing, pp. 585–624.

Pratt, L.Y., Mostow, J., Kamm, C.A., Kamm, A.A., 1991. Direct transfer of learned information among neural networks, in: Proceedings of AAAI-91, pp. 584–589.

Qi, M.H., 2001. Predicting us recessions with leading indicators via neural network models, in: Princeton University Senior Theses.

Raffinot, T., Benoit, S., 2019. Investing Through Economic Cycles with Ensemble Machine Learning Algorithms. Working Papers. HAL.

Rudebusch, G., Williams, J., 2009. Forecasting recessions: The puzzle of the enduring power of the yield curve. Journal of Business Economic Statistics 27, 492–503.

Soybilgen, B., 2018. Identifying US business cycle regimes using dynamic factors and neural network models. MPRA Paper. University Library of Munich, Germany.

Stergiou, C., Siganos, D., 2011. Neural networks. Technical Report. Imperial College London.

Tiffin, A., 2016. Seeing in the dark: a machine-learning approach to nowcasting in lebanon. IMF Working Paper.

Torrey, L., Shavlik, J., 2009. Transfer learning, in: Olivas, E.S., Guerrero, J.D.M., Sober, M.M., Benedito, J.R.M., Lo´pez, A. (Eds.), Handbook Of Research On Machine Learning Applications and Trends: Algorithms, Methods and Techniques (2 Volumes). chapter 11.

Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep neural networks?, in: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 27. Curran Associates, Inc., pp. 3320–3328. URL: http://papers.nips.cc/paper/ 5347-how-transferable-are-features-in-deep-neural-networks.pdf.