



M.Tech. Thesis Proposal

RICHA BANDLAS

Roll Number: 1913106

Month of Submission: June 2020

Supervisor: Dr. Clint P. George, Computer Science and Engineering

Co-supervisor: Dr. Sunil Paul, Humanities and Social Sciences

Predicting Recession in India^{1 2 3}

Richa Bandlas

June 21, 2020

A business cycle generally oscillates between expansion and contraction phases. Sometimes, the contraction phase, where economy goes down can lead to recession where people lose jobs, government takes more debt, investors lose confidence on market. Predicting this phase is one of the major issues in economy. This work intends to find various indicators of economy which gives early signals of this phase and includes literature available for solving this issue, showing different methodology. It also focus on how machine learning can be important in predicting recession as machine learning works faster on data and can compare various models simultaneously. At the end, this work proposes to find out better methodology with less forecasting errors to predict recession using different machine learning models.

Introduction and Overview

A common perception of working modern economies shows that they work around a trend rate of growth, with expansion and recession phases. The expansion periods bring economic growth, and usually increasing standards of living, such as purchasing power, increased salaries and the ability to afford good healthcare and education. As for the recession periods, the opposite tends to occur. As the economies halt their growth, salaries tend to stabilize, and with it, a reduction in purchasing power that sometimes can lead to difficulties in accessing basic healthcare and education. These fluctuations in growth affect not only individuals but also businesses, with reduced demands, reduction of profits and profitable economic opportunities. So it becomes obvious that people, businesses and government gives significant importance to knowing these shifts in the cycle.

There can be many relevant questions related to economy like when does expansion and contraction occurs in business cycle, also called turning point prediction, extent and time horizon for which this prediction can be done or identification of proper economic indicators that predict these changes. Each of these issues are dealt differently. The primary concern of this work is to predict recessions with enough lead times so that proper actions can be taken to recover from it. For predicting this, there should be a basic understanding of economic terms.

To begin with, a definition of economy and economics is needed. Economy represents the consumption and production of scarce resources in a certain area, relating to goods and services, which exist to satisfy its participant's needs. The study of different economies

¹ A report submitted to the Department of Computer Science and Engineering in partial fulfillment of the requirements for the M. Tech. thesis proposal at the INDIAN INSTITUTE OF TECHNOLOGY GOA.

² Supervisor: **Dr. Clint P. George**, Computer Science and Engineering

³ Co-supervisor: **Dr. Sunil Paul**, Humanities and Social Sciences

and their affecting factors are analyzed in a social science called Economics. Economics has two main subdivisions, *MicroEconomics* and *MacroEconomics*. This study discuss about *MacroEconomics* which deals with the economy as a whole. Macroeconomics introduces the concept of business cycle or economic cycle, which describe the fluctuations in economic growth through time, based on expansions and contractions(recession) of the economy.

Burns and Mitchell [1946] defined business cycle "*Business cycles are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises*". To be precise, business cycles has mainly two phases: Expansion, which indicates the prosperity in economy and Contraction, indicating downfall in economy leading to recession. Usually most economists tend to use more specific analysis, by measuring the trend of the GDP, the National Bureau of Economic Research (NBER) uses a more extensive range of parameters to define each cycle. So for predicting these cycles, proper examination of various macroeconomic indicators is required.

Recession According to the International Monetary Fund, there is no official definition of the term economic recession. In 1974, economist Julius Shiskin came up with a few rules of thumb to define a recession: The most popular was two consecutive quarters of declining GDP. But only GDP cannot show some changes happening in economy. Then, National Bureau of Economic Research(NBER) defined a recession "*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*". Because of recession, the unemployment rate increases and long term interest rate on bonds also decreases which in turn, make investors doubt on investing. Recession causes mess in entire economy and to recover from it, monetary policies are loosened so that money can come in the market. Most economists use the real Gross Domestic Product (GDP), a measure of total output, to determine if the economy of a nation is growing or shrinking. Other indicators like industrial growth, manufacturing of new orders, and retail sales volume growth can also be used for this purpose.

When it comes to prediction, standard econometric models tends to "over-fit" the samples and therefore the outcome might be misleading. Also, the more complex model is, higher is variance and the lower is bias which leads to forecasting error that can be big sometimes. Machine Learning gives more accurate results by minimizing forecasting errors and do the forecast much faster and with the usage of more data. Also, machine learning algorithms can analyze many alternative models at the same time, when in traditional econometric

one model can be analyzed at a time.

The approach proposed in this study is data driven approach, using leading indicators to get the early signals of recession. This approach places indicators in specific categories. For instance, variables including industrial production, output, and employment levels are used to capture a snapshot of the current state of the economy. These are characterised as “*coincident variables*”. The indicators which identify ups and downs before coincident activity are called *Leading indicators*. Similarly, elements which show delayed affects are called *Lagging indicators*.

This work will try to find these indicators and show their relevance using past recession data of Indian economy and develop machine learning models that analyze the selected indicators and gives possibility of recession .

Literature review

Types of economic model

Filardo [1999] examined the reliability of five popular recession prediction models and explains their strength and weakness in predicting a recession. Each model has its own way of predicting recession.

Simple rules of thumb using Composite Leading Indicator (CLI) model takes two month consecutive fall in CLI as an indicator for recession. Similarly, **Neftci model** checks if the estimated probability of monthly observations of CLI exceeds a threshold value or not. **Probit Model** takes multiple indicators and find probability of future recession. **GDP forecasting model** uses consecutive fall in GDP as indicator and **Stock-Watson Model** uses Coincident matrix and Experimental recession index to indicate future recession. All models are tested on basis of timeliness and accuracy. However, in terms of their ability to forecast an imminent recession, he favored three models – Probit model, GDP Forecasting model, and Stock-Watson model. Plakandaras et al. [2017] compared Probit model with Support Vector Machines (SVM) and concluded that the probit models can foresee the U.S. recession periods more closely than SVM models for up to 6 months ahead while the SVM models are more accurate at longer horizons.

Economists such as Estrella and Mishkin [1998] shows relevant use of Probit models to forecast recessions of the limited nature of the dependent variable, that is, whether the economy at any point is either in recession or not in recession as said by Chauvet and Potter [2005]. In these models the dependent variable takes the value 1 if there is a recession in the current period, and 0 if not. However,

Filardo [1999] believed the probit model might miss a few recessions that exhibit unusual lead times.

Significance of various indicators

As stated by NBER, *real GDP, real income, employment, industrial production, and wholesale-retail sales* gives good set of indicator for defining recession. Plakandaras et al. [2017] suggested *oil prices, stock returns and the term spread* as the leading indicators for forecasting.

The yield curve has been used to forecast recessions in many papers such as Zhang [2019], Estrella and Hardouvelis [1991], Chauvet and Potter [2005], Qi [2001], Shaaf [2000], Estrella and Mishkin [1998], Estrella and Trubin [2006], Rachidi Kotchoni [2020]. Yield curve is the difference between long term and short term Treasury bond rate. The economists suggest a flatter curve (smaller spread) indicates a weaker growth while a steeper one (bigger spread) indicates a stronger growth. However, many economists suggest that if yield curve is used with some other indicator such as *stock price*, then it may give better results and their dynamic pattern can be used to know about possibilities of turning points in business cycle. All models considered in the study done by Chauvet and Potter [2005] indicated that the yield curve was signaling weak future economic activity in 2001–2002 and also shows parameter instability of models.

Filardo [1999] used composite index of leading indicators (CLI) in his models. CLI includes 10 elements: manufacturing hours worked, consumer expectations, stock price, initial unemployment claims, building permits, money supply, difference of long-term and short-term interest rate in government securities, vendor performance, manufacturing orders for capital goods, and manufacturing orders for consumer goods.

Zhang [2019] suggested that other than yield curve, *Non-farm Payrolls, Fed Funds rate, Consumer price index, 10-year Treasury Rate, S&P 500 index* can also be used to predict recession using machine learning models.

A real GDP growth is mentioned in the definition of recession and is recommended by various economists such as Anderson and Vahid [2001], Shaaf [2000], Estrella and Mishkin [1998] and Rachidi Kotchoni [2020]. The negative value of this indicator indicates recession sometimes. Other than GDP growth, Anderson and Vahid [2001] developed nonlinear leading-indicator models for *interest-rate spread and growth in M2* as leading indicators, where M2 is the money supply which is mostly in the form of savings and time deposits, certificates of deposits, and money market funds. However, they find that the marginal contribution of M2 growth in preceding recessions, condi-

tional on the spread, is negligible. Rachidi Kotchoni [2020] also used *treasury rate, credit score and term structure* to perform multi-horizon, real-time and nonlinear forecasting.

Estrella and Mishkin [1998] focused on Testing whether simple financial variables maybe useful predictors for a future recession. Four types of financial indicators are used which are *interest rate and spirits, stock prices, currency, monetary aggregates*. Results show that if the stock price is used for getting predicting recession it is more useful than yield curve but if both the stock prices and yield curve are used it gives the best result. However, if any macroeconomic variable like *GDP* is used with these two, then the prediction probability may increase for one quarter, but it will not be true for every other horizon. They concluded that the *yield spread and stock price indexes* emerge as the most useful financial indicators and can be used for quick reliability checks.

Use of Machine Learning Models

Zhang [2019] considered 6 individual machine learning models viz. K Nearest neighbour, Regularized Linear regression, Naive Bayes, Support vector machines, Gaussian process and XGBoost. Author used 2 ensemble models viz. Grand average model and Weighted average model as testing models on above models. Author finally concluded with Support vector machines as final deployment model for predicting recession in U.S.

Chauvet and Potter [2005] and Shaaf [2000] used Bayesian methods and neural networks respectively , Estrella and Hardouvelis [1991] applied regression and calculated root mean squared error and Qi [2001] used a Three-layer feed-forward neural network. Levenberg–Marquardt algorithm to know predictive power for Yield curve.

Some economists such as Anderson and Vahid [2001] and Rachidi Kotchoni [2020] used an extension of auto-regressive models like Logistic Smooth Transitions Auto-regressive models and Augmented Auto regressive (AAR) model to find probability of recession using GDP, interest-rate spread, growth in M2, credit score etc. as indicators.

Estrella and Mishkin [1998] proposed to use statistical regression techniques to predict recession using financial variables. Probit equation was used in the model and the variables were assumed to have only 2 values either 1 or 0. A function was applied which gives the probability of recession as a result. If the result was close to 1 then there were strong chances of recession.

Data and Exploratory Analysis

The data of various indicators of Indian Economy are not available at a single source. The indicators selected for exploratory analysis are divided into three categories as Leading, Lagging and Coincidental indicators.

Leading Indicators: These are the indicators which gives early signals of recession. If seen in terms of business cycle, they start showing changes before actual economy slowdown. Stock market indicators are said to be the best leading indicators. Following indicators chosen as leading:

1. *Bond Yields:* 10-year Treasury bond and 3 month treasury bill interest rates are used as they decline before whole economy declines.
2. *Bank credit growth:* This indicator shows that what banks has credited to commercial sector. When economy slowdowns, commercial sector people don't take loans because production is less, which make the value of this indicator go down. It's changes with year can be seen in 1
3. *S&P BSE 500 index:* Standard & Poor's rating for 500 goods and services is the stock market index of Bombay Exchange. If its value decreases, investor loses confidence in market and hence, money supply is less.
4. *CLI:* Composite index of leading indicator, developed by OECD in 1995, based on growth cycle. As defined by OECD, "The composite leading indicator (CLI) is designed to provide early signals of turning points in business cycles showing fluctuation of the economic activity around its long term potential level. CLIs show short-term economic movements in qualitative rather than quantitative terms." OECD computed CLI by using 8 leading indicators viz. , Business confidence index, IIP, Money supply, Exchange rate, Deposit interest rate, share stock prices, imports.
5. *Housing price index:* Rising housing prices encourage consumer spending and gives higher economic growth. Low index may indicate recession.

Lagging Indicators: They can be know only after the event, but that doesn't make them useless. They can clarify can confirm a pattern that is occurring over time.

1. *Unemployment rate:* If it increases last month and before that, it means economy is doing poorly over time and can continue to do so. unemployment rate of 2019-2020 raise to 27.2
2. *Corporate earning profit:* This is the contribution of profits of corporate sector in GDP. As economy goes down, this profit also decreases and vice versa.

Bank credit to commercial sector vs. Year

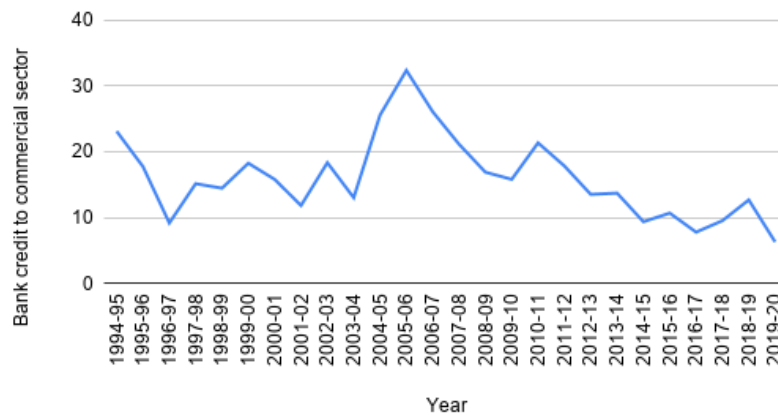


Figure 1: Changes in Credits given by banks to commercial sector(annually)

3. *Total liability (% of Gdp)*: To recover from economic slowdown, government takes loans either from inside country called internal debt or from outside the country called external debt. If this debt increases, it leads to economic slowdown.

4. *Real GDP growth*: Changes in Gross Domestic Product captured at constant prices. It measures the overall output of a country. As economy declines, it also declines. Refer to figure 2

GDP Growth (%) vs. date

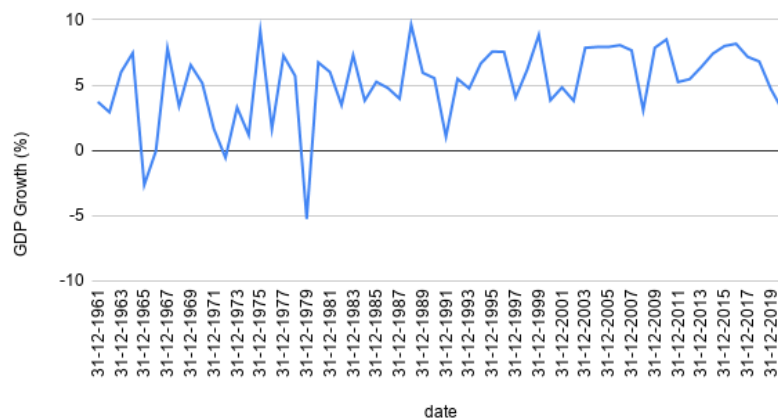


Figure 2: Annual changes in real GDP

Coincident Indicators: These are analyzed and used as they occur. These are key numbers that have a substantial impact on the overall economy.

1. *Repo rate*: Repo rate is the rate at which the central bank of a

country (Reserve Bank of India in case of India) lends money to commercial banks in the event of any shortfall of funds.

2. *IIP General Index*: Index of industrial production falls down if recession hits as demand goes down, hence production is slow.

3. *Inflation rate*: Recession and Inflation rate are very much related. Inflation rate generally falls in recession due to lower demand and output. It increases the money supply in the market for which repo rate is also reduced so that more loans can be taken by people. Once money is sufficient in the market, economy starts growing and inflation rate increases. Refer to figure 3

4. *Consumer Price index*: CPI falls during recession as consumer cut out from all the extra expenses.

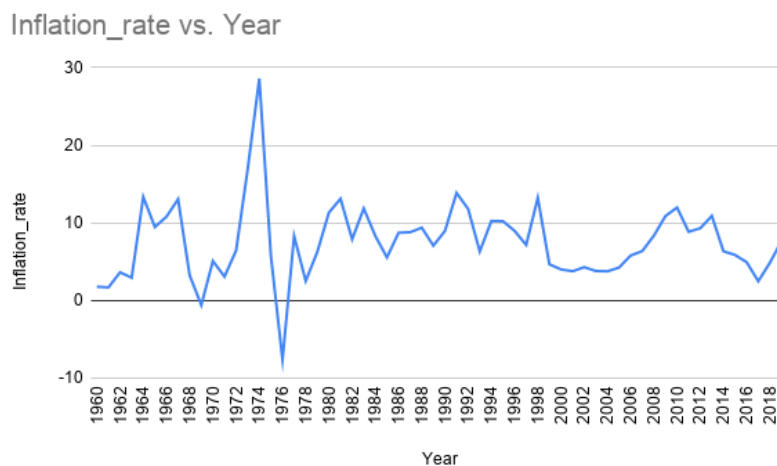


Figure 3: Rate of changes of Inflation per year

Proposed work

Conclusion of literature review

1. It is not explicitly defined which factors indicate recession.
2. Yield Curve is one of the most researched indicator and works as leading indicator.
3. CLI is base indicator of the economic models proposed.
4. Probit model is simplest approach to predict recession.
5. Auto-regressive models can be significantly used as Machine learning models as indicators are time series data.
6. Neural networks are good with only 2-3 indicators.

NOTE: As 1979 was recession year in India, and if seen carefully, GDP growth rate declined sharply at that time. Similarly, Inflation rate went sharply down in 1976 but then it increased and again declined in 1979.

For 2019-20, GDP curve is declining, but bank credit growth indicator starting declining beforehand. They may show some signals of recession in year 2020-21.

Work Proposal: As it is a data driven approach, so the contribution of this work will be in increasing the availability of data of various macroeconomic and financial indicator from different sources and different frequencies at a single place. Analysing the cross-correlation between the indicators during exploratory data analysis will be the next step. Once data is ready and available to work on, some machine learning techniques will be used which can give early signals about potential recession.

The results of various machine learning models will be compared to see which model gives better accuracy and least forecasting error during prediction.

References

- Heather M. Anderson and Farshid Vahid. Predicting the probability of recession with nonlinear autoregressive models. *Macroeconomic Dynamics*, 2001.
- Arthur Burns and Wesley Mitchell. *Measuring Business Cycles*. 1946.
- Marcelle Chauvet and Simon Potter. Forecasting recessions using the yield curve. *Journal of Forecasting*, 2005.
- Arturo Estrella and Gikas Hardouvelis. The term structure as a predictor of real economic activity. *Journal of Finance*, 1991.
- Arturo Estrella and Frederic S. Mishkin. Predicting u.s. recessions: Financial variables as leading indicators. *The Review of Economics and Statistics*, 1998.
- Arturo Estrella and Mary R. Trubin. The yield curve as a leading indicator: Some practical issues. *Current Issues in Economics and Finance*, 2006.
- Andrew Filardo. How reliable are recession prediction models? *Economic Review*, 1999.
- Bill Plakandaras, Juncal Cunado, Rangan Gupta, and Mark Wohar. Do leading indicators forecast u.s. recessions? a nonlinear re-evaluation using historical data. *International Finance*, 2017.
- Min Qi. Predicting us recessions with leading indicators via neural network models. *International Journal of Forecasting*, 2001.

Dalibor Stevanovic Rachidi Kotchoni. Gdp forecast accuracy during recessions. *EconomiX-CNRS, Universit'e Paris Nanterre.*, 2020.

Mohamad Shaaf. Predicting recession using the yield curve: An artificial intelligence and econometric comparison. *Eastern Economic Journal*, 2000.

Terrence Zhang. Recession prediction using machine learning. 2019.