École Polytechnique Fédérale de Lausanne

Financial Econometrics

FORECASTING INFLATION

June 19, 2022

Paul DE CALONNE 288053
Jérôme Hirschi (AUDIT-E: 16-218-398)
Barthélemy WELSCH 262474
Pierre-Benoît GRIMAUX 302257
Jean MOLIN 288296

Abstract

The aim of this report is to provide both theoretical and practical frameworks for inflation forecasting using different models and ranges of data. Inflation − measured by the consumer price index (CPI) − is a determining factor of the welfare of an economy. As policy makers aim to achieve price stability, they must be forward thinking in regards to price evolution. Thus, tools forecasting inflation in the short, medium or long term are widely studied and in great demand. In this paper, we put forth five different models divided in three sections based on the data they use. The first set of models utilizes the time series evolution of the CPI in the USA from 1947 to 2022. These are composed of an ARIMA model and a neural network. The second set of models are an extended application of the first set using a wider range of data (macroeconomic data correlated to inflation such as interest rate, wages or money supply). Finally, we focus on more "exotic data" as we explore the google trends words of the past fifteen years and their correlation with inflation.

Contents

List of Figures

List of Tables

1 Introduction

Inflation, or the change in prices of goods and services over time, is a significant factor that determines the welfare of an economy. If prices continuously increase over time, money will lose purchasing power. Therefore, inflation can induce serious societal costs.

In this section we will explain why inflation and its forecast gained in importance over the last century and its impact on financial markets. We will pursue by considering different measurements of inflation, and illustrate its evolution over time, with a focus on inflation volatility. Finally, we will review inflation forecasting techniques.

1.1 Inflation forecasting

In the following, we are presenting the motivations to forecast inflation. First, we are considering the historical change in the way monetary policy is conducted, then we will reflect on the motivation of corporations and the financial sector to predict future inflation.

1.1.1 Monetary policy

Generally, the primary objective of central banks is to achieve price stability. Until the 1930s, many central banks hoped to achieve this goal by pegging their currency to a fixed amount of gold. The appeal was that there is a certain amount of gold which only growths slowly, thereby limiting the growth of money supply and ultimately stabilizing inflation. However, the Great Depression forced the major currencies to suspend the gold standard in order to fight depression. After World War II (WWII), the major world economies adopted a new monetary system as a result of the 1944 Bretton Woods Agreement. The so called Bretton Woods system aimed to stabilize the exchange rate and promoting trade, by pegging the different currencies against the US dollar, the later being pegged to gold. By the 1970s, the amount of dollar, which was relatively small after the WWII, dramatically increased for two reasons: (i) As the world economy grew, more dollars were demanded. (ii) The expansionary pubic policy and public debt creation of the USA. As the monetary mass of US dollar grew much faster than the stock of gold, financial agents started to doubt the convertibility of US$ to gold. That's why in 1971 President Nixon ended the convertibility of the US dollar to gold, and in 1973 the system collapsed.

As a result, currencies were no longer linked to a commodity, which marked the start of fiat currency. Most currency could freely float, the exchange rates therefore being driven by market forces. However, the dollar remains at the center of the world economy, as most of the traded goods and services are denominated in US$ [Swo69], this is known as the "Dollar standard". During the 1980s and 1990s, the volatility of GDP dramatically declined, and the economy prospered from stability, therefore this period is often referred to as the Great Moderation [Tay13]. During this time, the European Monetary System was established, and aimed at stabilizing the relative value of currencies by linking them together, implicitly setting a exchange rate driven monetary policy.

Inflation forecasting has significantly gained in importance since the 1990s, when central banks began to transit to inflation targeting (see figure 39). The idea being, that they can achieve price stability (often their primarily objective), by defining a nominal anchor with a range in which inflation should remain. As a quantitative reference, the FED uses the Consumer Price Index (CPI). The difficulty of such a monetary policy lies in the fact that macroeconomic fundamentals are observed on a low frequency, usually quarterly or monthly. Thereby making the availability of up-to-date and accurate measurements more challenging, and hence the adoption of a correct monetary policy more uncertain. Moreover, the frequent reviews and discussions in central banks, usually bi-weekly [AF09], and the acute need of real-time information during crisis led to a growing need for short term inflation forecasting.

1.2 Definition of inflation

1.2.1 Different measures (mainly CPI)

There are different measures of inflation, the most common are Consumer Price Index (CPI), Harmonized Consumer Price Index (HCPI), and Retail Price Index (RPI). Harmonized Price Index is commonly used in Europe, it gives a possibility to compare the inflation on an international scale. Retail Price Index has roughly the same idea as Consumer Price Index. However, the former includes all costs related to housing. It might be the repayment of the mortgage, local taxes on housing, or its depreciation.

As Consumer Price Index is the most known measure of inflation, we will discuss it in greater detail. Consumer Price Index (CPI) is measured on a private household consumption basis. It includes different types of products and services, mainly used on a daily basis. Two points should be taken into account while covering CPI. The first one is just the simplicity of the method: we count the data that we can (relatively) easy to gather and interpret. Last, but not least, is the frequency of usefulness/importance of a particular product of a particular product/service to a household. For instance, household consumption has slightly changed during the lockdown due to Covid-19. Demand for leisure services has significantly decreased. Thus, the survey itself conducted among households had to be slightly changed.

Let's compare the current Consumer Price Index in the US with the one in Switzerland (see 35 and 36). Annual prices for commodity goods in the US increased significantly in March 2022, while in April 2022 they declined a little bit, even though this decrease was too little. Food prices, prices for vehicles as well as for accommodation kept going up since last month in the US. As a result, according to Trading Economics source, the inflation rate in the US in March 2022 was about 8.5 percent while in April it decreased by 20 basis points. Today's jump in prices for household goods and services is the most significant compared to the last ten years. Switzerland follows roughly the same trend as the US. However, the prices in Switzerland kept going up in April 2022, and an increase of a 40-point basis was pointed out. And according to Trading Economics, inflation in Switzerland has increased by 250 basis points compared to the previous year.

1.2.2 Evolution of consumer basket

A consumer basket is composed of different types of goods and services. The most diversified category is food. The interesting fact is that it only includes the products bought at grocery shops and does not include the restaurant services which are considered a service apart. Other important categories are clothing, utilities, health, and education. The consumption basket has experienced a huge evolution, especially during the last two years. Due to Covid-19, the lifestyle and the habits of consumers have considerably changed. Thus these changes were reflected in the consumer basket. For instant, face masks were added to it. Another important evolution involved in the lockdown is the weight change. The same products and services have stayed in the basket but some of them have become more/ less important than before. The weight of all leisure services as well as tourism services, hotels, and restaurants has significantly decreased in weight: it has dropped roughly by 50 percent. People travelled less during the pandemic. Instead, they started to spend much more time at home due to the permission for remote work and studies, and the increase in online streaming services. Thus, it is not surprising that 2021 food, health, and communication services experienced a jump in weight.

1.2.3 Volatility of inflation

The inflation volatility is impacted by two economic factors: demand and supply. Supply is characterized by the change in prices mainly of commodities such as oil. When the oil prices change a lot, it has a direct impact on inflation volatility. Demand factor has a more significant impact on developing countries because the food and other daily life products constitute a bigger percentage of their consumption basket. The evidence from Turkey proved it: Lopcu et al [Yil+10] have studied the impact of the food volatility on the inflation volatility and found out that indeed level of inflation volatility is driven by the volatility of food prices. Thus, the main recommendation is to authorities while stabilizing the inflation volatility strongly take into account food volatility.

1.3 History and quantitative summary

In this part, we will first analyze the general evolution of inflation over the past decades as well as its sources of variation. Then, we will focus on the past years and will provide the latest update on inflation.

Figure 38 in annexes shows us six major episodes in the world history that were the main cause of the high inflation after the Second World War. One can state that major episodes are either political ones or related to the increase in the price of commodity goods, probably some episodes combine both of them.

The situation with inflation changed at the beginning of 20 century. Food prices, as well as commodity prices, have started to increase. In Figure 1. Six episodes of post-WWII inflation, this is a sixth episode named Rising gas prices in 2008. However, the global crisis has frozen significantly the world economy. Prices have not only come back to the previous mark-up but the world, in particular developed countries, has even experienced a slight deflation just after the financial crisis.

The latest update on inflation shows that inflation has significantly increased. We can see it in Figure 2. Evolution of the inflation over the past 4 years: the inflation in March 2022 has doubled compared to the past three years due to COVID 19.

2 Forecasting inflation with the Consumer Price Index

This section will cover different models that have been implemented by us and based on the CPI index time series. In order, to create models that are able to forecast inflation different option exists. Indeed it exists many different tools and many different approach. By asking someone to forecast inflation one of the most probable answer is to look at previous data. Like Robert Kiosaki said “The best way to predict the future is to study the past, or prognosticate.”. Then this part will focus on the past inflation in order to predict futures inflation.

2.1 Litterature Review

N.B. : Should have been done by Tetiana Yashchenko but as discussed with the teacher she unfortunately left the group.

2.2 A first model to forecast inflation using Times Series

2.2.1 Data Analysis

The first model we are going to implement in this part of our study is a Time Series model. To study and forecast inflation, we have chosen to focus our work on inflation in the USA from January, 1st 1947 to February, 1st 2022. The data is a set of monthly price index on this range of time from 1947 to 2022. As it is defined on the website of the federal reserve bank : “The Consumer Price Index for All Urban Consumers: All Items (CPIAUCSL) is a price index of a basket of goods and services paid by urban consumers”. The data comes from the website of the Saint Louis federal reserve bank and more precisely from the US bureau of Labor Statistics.

We have chosen the United States of America as it was easy to find reliable data concerning inflation on a large period of time. In this part we will focus a little bit on the data before starting to implement a model to forecast inflation.

The following chart illustrates the way we transformed the data in our model into a panda data frame :

Date	CPI
1947-01-01	21.480
1947-02-01	21.620
1947-03-01	22.000
...	...
2021-12-01	280.126
2022-01-01	281.933
2022-02-01	284.182

Table 1: Overview of the data we use in our model

Below is a plot representing the evolution of the Consumer Price Index over the period 1947 (January)-2022 (February) on a monthly basis.

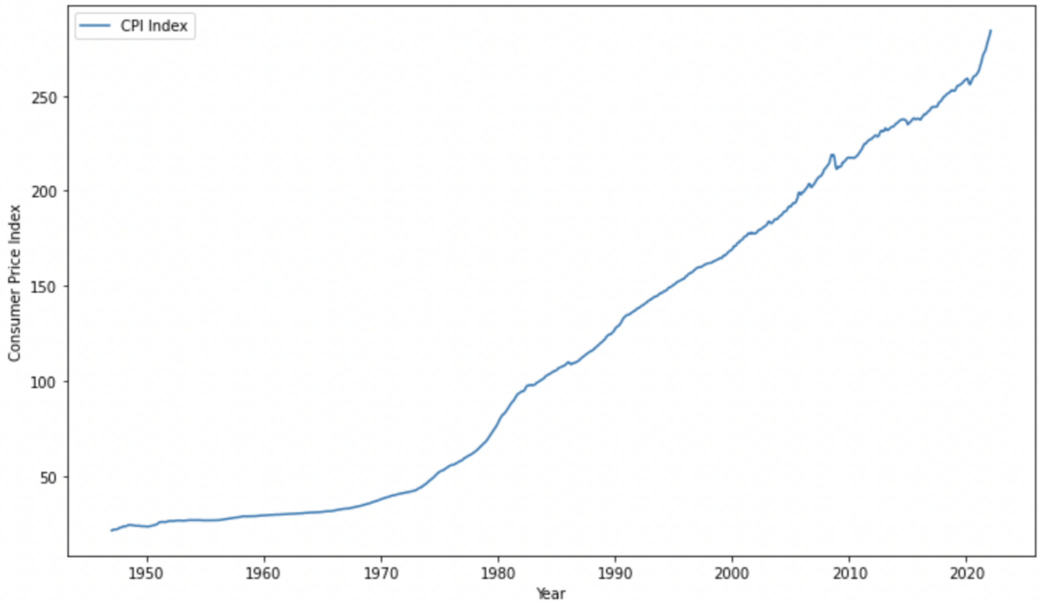


Figure 1: Evolution of the Consumer Price Index in the USA from 1947 to 2022 on a monthly basis

We can notice that inflation has kept increasing from 1947 to 2022 which is a famous result : inflation increases over time. There are some recognizable peaks on our graph such as the strong inflation in the 1970s due to the 1973 oil crisis or the peak followed by a fall during the Global Financial Crisis. Despite those different events - which were hard to foresee and had a strong impact on price evolution - it is consistent to assert that inflation is a growing function of time and thus we should expect our predictions to move in this direction.

Thereafter, we will focus on two different models to explain our Consumer Price Index evolution over time. The goal of those models is to explain inflation evolution in a quantitative way in order to forecast it on one year i.e. on twelve months as we use a monthly basis.

2.2.2 A model using time series

The first model we chose to implement is a time serie model using ARIMA. This has been computed using python and the different steps will be detailed below. First of all, we give some definitions and basis properties of ARIMA model which we will use thereafter.

An ARIMA model (Auto Regressive Integrated Moving Average Model) is the generalization of an ARMA (Auto Regressive Moving Average) model. This is used to explain, understand time series or forecast future points when knowing a certain number of them belonging to the past.

Some mathematical definitions of an ARIMA model

Let's consider X_t which is a time depending process meaning the process X_t depends on time and takes different values as time evolves. We will first focus on a discrete process. An ARMA(p,q) process can be defined as :

$$X_t - \alpha_1 \cdot X_{t-1} - ... - \alpha_p \cdot X_{t-p} = \varepsilon_t + \theta \cdot \varepsilon_{t-1} + ... + \theta_q \cdot \varepsilon_{t-q}$$

Which can be simplified using the Lag Operator L:

$$\left(1 - \sum_{i=1}^{p} \alpha_i L^i\right) X_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t$$

The left part is the Autoregressive (AR) part of the model and the right part corresponds to the Moving Average (MA) part of the model. Both sides combined together describe an ARMA model.

We assume that there is a unit root in our lag polynomial which means 1 is a root of our polynomial. Let's consider the multiplicity of our unit root is d. Using this, we factorize our polynomial and thus can write our model the following way :

$$\left(1 - \sum_{i=1}^{p'} \alpha_i L^i\right) \cdot (1-L)^d X_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t$$

This last equation caracterizes an ARIMA(p',d,q) process.

ARIMA model research

Firstly, we split our dataset in two different parts : a training set and a test set. The first one which is made of all the data except the last 12 will be used to find and train the model which fits the best our time series. The second set will be used to test and compare our model. Which chose a test set of twelve values, as we want to predict inflation over one year (to recall, we have access to monthly inflation).

Now, the goal is to identify which ARIMA model fits the best our time series representing inflation which means identifying the coefficients p', d and q which were defined above.

The first task to perform is to study the stationarity of the series. This is done using a Dickey Fuller augmented test. This test tests the null hypothesis that the serie is stationary using that a serie is stationary if it doesn't have any unit root. We compute this test on python and below are given the result of the test :

ADF Statisctic	p-value	1‰ Critical Value	5‰ Critical Value	10‰ Critical Value
2.11	0.999	-3.4	-2.9	-2.6

Table 2: Result of a Dickey Fuller Augmented test on the data before transformation

As we can see, the p-value is very high which means we can't reject the null hypothesis and therefore we can't reject the fact that the serie has a unit root. This means we have to transform our serie to make is stationary over time.

Transformation of the time series

The transformation of the serie will be done using a log shifting : we calculate the logarithm of each value of the data set and then make the difference between two consecutive values : $log(X_t) - log(X_{t-1})$. The first value is set by convention to zero.

DATE	CPIAUCSL	Log Inflation	Log shifting
1947-01-01	21.480	3.067122	0.000000
1947-02-01	21.620	3.073619	0.006497
1947-03-01	22.000	3.091042	0.017424
1947-04-01	22.000	3.091042	0.000000
1947-05-01	21.950	3.088767	-0.002275
...	...	...	...
2020-10-01	260.352	5.562035	0.000622
2020-11-01	260.721	5.563451	0.001416
2020-12-01	261.564	5.566679	0.003228
2021-01-01	262.200	5.569108	0.002429
2021-02-01	263.346	5.573469	0.004361

Table 3: Overview of the transformed data with the log shifting method

We perform then a Dickey Fuller augmented test on this new panel of data to check whether the series is stationary or not. Here are the results of the test :

ADF Statisctic	p-value	1‰ Critical Value	5‰ Critical Value	10‰ Critical Value
-4.17	0.00074	-3.4	-2.9	-2.6

Table 4: Result of a Dickey Fuller Augmented test on the transformed series

Now, we have a low p-value : below 5%. Regarding the critical values of our statistic, we can reject the null hypothesis of a unit root at a strong level of confidence (we are below the 1 % critical value of the statistic used in the model). The serie is now stationary.

We illustrate the stationarity by plotting the rolling mean, the rolling standard deviation (both on a rolling window of twelve months) and the log shifting:

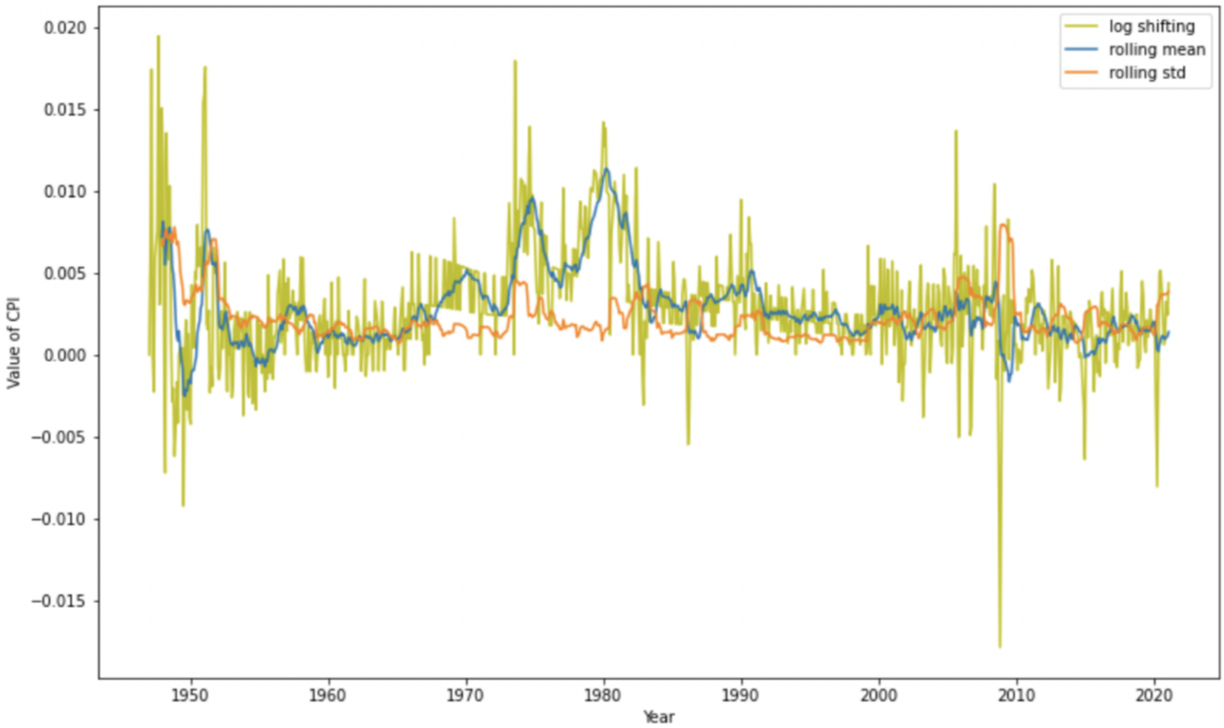


Figure 2: Plot of the log shifting, the rolling mean and the rolling standard deviation on a twelve months rolling window

Even though we can observe some peaks for the rolling mean, the oscillations remain around a certain value which is the average rolling mean. By consequence, the mathematical features of our model remain constant over time : our serie is stationary after transformation.

Study of the ACF and the PACF to find our model coefficients

We plot the Autocorelation function (ACF) and the Partial Autocorrelation function to find out the coefficients of our model to fit data in inflation.
The ACF has the following expression :

$$\rho_k = \frac{Cov(X_t, X_{t-k})}{V(X_t)}$$

The partial autocorrelation measures the additional correlation between X_t and X_{t-k} after adjusting for the intermediate values $Y_{t-1}, \cdots, Y_{t-k+1}$. We provide below the plot of the ACF and the PACF of our model.

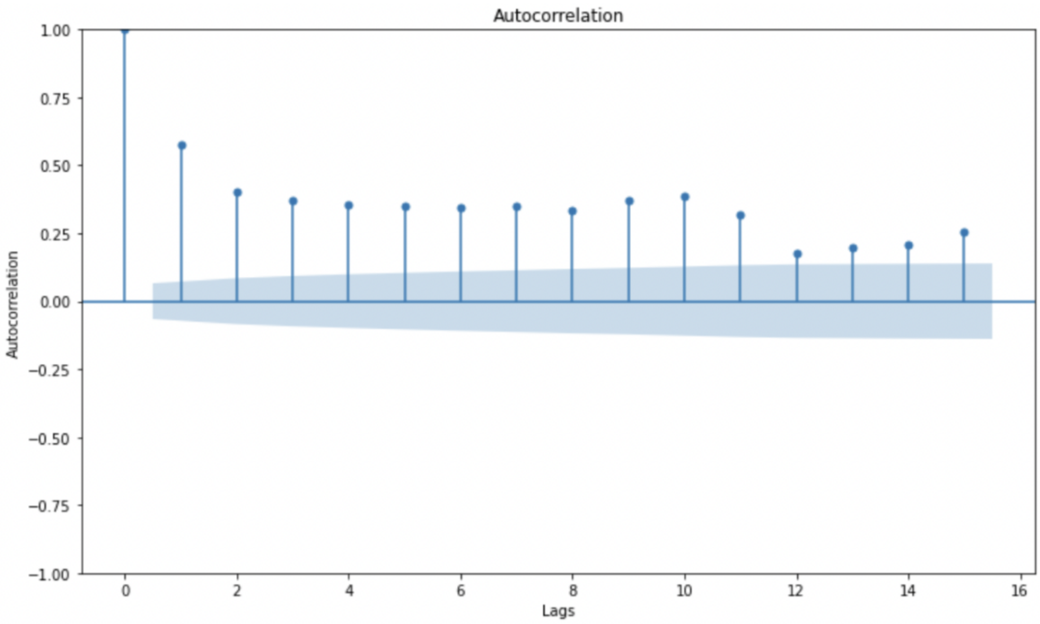


Figure 3: Plot of the Autocorrelation function

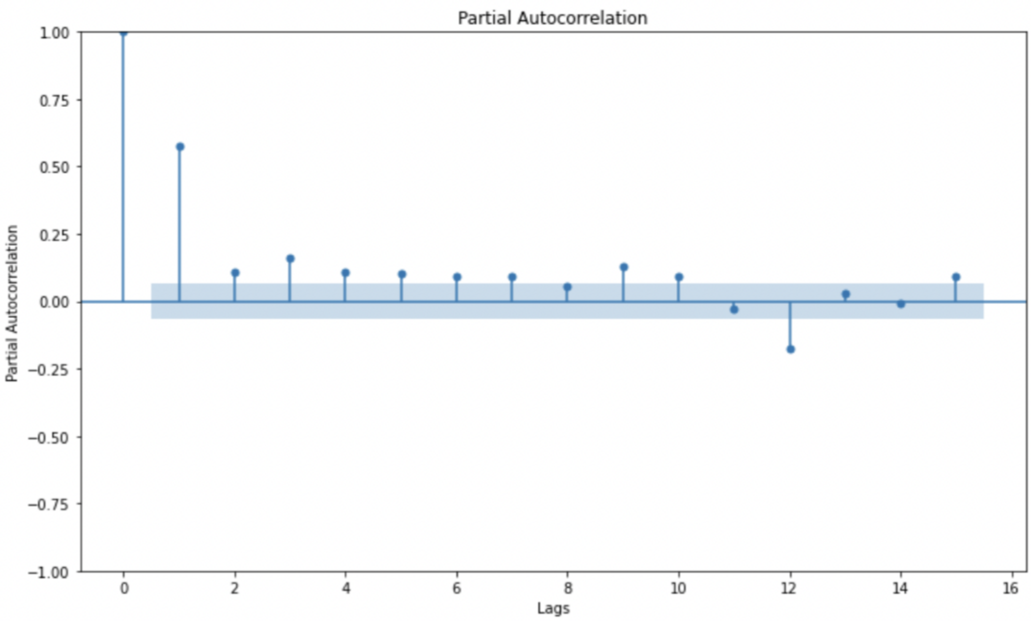


Figure 4: Plot of the partial autocorrelation function

The autocorrelation function drops after eleven lags and the partial autocorrelation function drops after 2 lags. By consequence, we choose to consider an ARIMA(2,1,11) model on the logarithm of our time series which would be in a sens equivalent to a an ARIMA(2,0,11) on the log shifting of our model.

The ARIMA(2,1,11) model

Using stats model on Python we apply our ARIMA model to find out the coefficients and the characteristics of the model.

Below is the chart with our different results :

	coef	std err	z	$P > \|z\|$	[0.025	0.975]
const	0.0028	0.000	7.391	0.000	0.002	0.004
ar.L1.D.Log Inflation	0.1127	0.096	1.177	0.239	-0.075	0.301
ar.L2.D.Log Inflation	0.6016	0.096	6.281	0.000	0.414	0.789
ma.L1.D.Log Inflation	0.3444	0.097	3.535	0.000	0.153	0.535
ma.L2.D.Log Inflation	-0.4151	0.097	-4.279	0.000	-0.605	-0.225
ma.L3.D.Log Inflation	-0.1013	0.056	-1.817	0.069	-0.211	0.008
ma.L4.D.Log Inflation	-0.0129	0.043	-0.304	0.761	-0.096	0.071
ma.L5.D.Log Inflation	-0.0281	0.040	-0.706	0.480	-0.106	0.050
ma.L6.D.Log Inflation	-0.0306	0.047	-0.649	0.516	-0.123	0.062
ma.L7.D.Log Inflation	0.0122	0.039	0.312	0.755	-0.065	0.089
ma.L8.D.Log Inflation	0.0360	0.037	0.964	0.335	-0.037	0.109
ma.L9.D.Log Inflation	0.1036	0.045	2.297	0.022	0.015	0.192
ma.L10.D.Log Inflation	0.1660	0.052	3.195	0.001	0.064	0.268
ma.L11.D.Log Inflation	0.1748	0.047	3.749	0.000	0.083	0.266

Table 5: Features of the ARIMA(2,1,11) model

What we can first notice is that some coefficients are not very significant such as the first five coefficients of the moving average part of the model. However, some are very significant, especially the second coefficient of the auto regressive part and the last three of the moving average part of the model (p-value below 5%). This explains why the model needs so many lags for the moving average part, the significant coefficients are located very far in the model.

Concerning the results of this ARIMA model, we obtain for our 887 observations an Akaike Information Criteria (AIC) of -7983,495 and a Bayes information criteria (BIC) of -7911.678. Those values must be as low as possible for the best model. Those numbers obtained were really convenient as the auto arima model applied to our time series gave similar value but worst predictions as we will see below.

We can now write the mathematical formula which describes our model where $y(t)$ represents the logarithm on the consumer price index that describes inflation :

$$\boxed{y_t = 0.0028 + MA(t) + AR(t)}$$

With :

$$MA(t) = \varepsilon_t + 0.3444 \cdot \varepsilon_{t-1} - 0.4151 \cdot \varepsilon_{t-2} - 0.1013 \cdot \varepsilon_{t-3} - 0.0129 \cdot \varepsilon_{t-4} - 0.0281 \cdot \varepsilon_{t-5} - 0.0306 \cdot \varepsilon_{t-6} + 0.0122 \cdot \varepsilon_{t-7} +$$

$$0.0360 \cdot \varepsilon_{t-8} + 0.1036 \cdot \varepsilon_{t-9} + 0.1660 \cdot \varepsilon_{t-10} + 0.1748 \cdot \varepsilon_{t-11}$$

And

$$AR(t) = 0.1127 \cdot y_{t-1} + 0.6016 \cdot y_{t-2}$$

Motivations of such a highly parameterized model

Our model is highly parameterized, especially for the moving average part of the process with 11 lags. The auto regressive part of the model only has 2 lags which remains a rather usual value for ARIMA models. The first explanation is the interpretation of the auto correlation and partial auto correlation functions of our time series. As it is hard to identify a drop in the lags of the Auto correlation function, the number of parameters are hard to determine precisely. This is a first argument to justify that an ARIMA might not be the best model to study forecast inflation. Moreover, considering those high values regarding the number of parameters, we had decided to verify whether this made sens or not using the auto ARIMA function of python which immediately provides the number of lags and their value. The result gave less parameters with an ARIMA(2,1,2) model. We interpret this result the following way : the software gives the result which has the best criteria (BIC and AIC)f or instance , which was indeed the case for this low parameterized model (but not by much). However, regarding the prediction, the highly parameterized ARIMA model gave much more satisfactory results. This can be explained that a highly parameterized model is more sensitive to fluctuations. Finally, the ninth, tenth and eleventh parameters of the moving average part very significant (p-value well below 5% even less than 0.1%) which motivated us to consider more lags and go further in the number of lags but increased our questions about the reliability and accuracy of this model.

To conclude on this question, the question of the parameters we faced was one of the major issue we encountered considering this model. This was a first indicator that the use of an ARIMA model might be reconsidered and thus wouldn't be sufficient for our study.

Verification of the models

We test other models and compare the AIC and BIC to check whether we can find a better model in term of accuracy. Our first model which is ARIMA(2,1,11) has an AIC of -7983 and a BIC of -7911.

We test the following models : ARIMA(3,1,11), ARIMA(2,1,12) and finally ARIMA(3,1,12). The two models where we increase q which corresponds to the moving average part of our model don't work and can't provide coefficients.

The model ARIMA(3,1,11) has an AIC of -7990 and a BIC of -7960. As we seek for the model which gives the smaller value of AIC and BIC, we should prefer this model to the previous one. However, the difference remains insignificant and the continuation of the project illustrated hat it was more relevant to use an ARIMA(2,1,12) to forecast inflation.

Residuals of the model

The plot below represents the residuals of our ARIMA(2,1,11) model. We can easily notice that the residuals of the model have a zero mean. However, the shape of the graph suggests that residuals are correlated. This is observed at the beginning of the graph where there is a high volatility. Then the volatility decreases and knows other peaks and oscillations between 2000 and 2010. This can be explained in a macro economic way. For instance the Global Financial crises of 2008 has had a huge impact on inflation As well as in the 1980s where lots of countries faced a very strong inflation due to energy prices in 1973 and 1979 especially (when oil producers suddenly increased oil prices which had a strong impact on western economies).

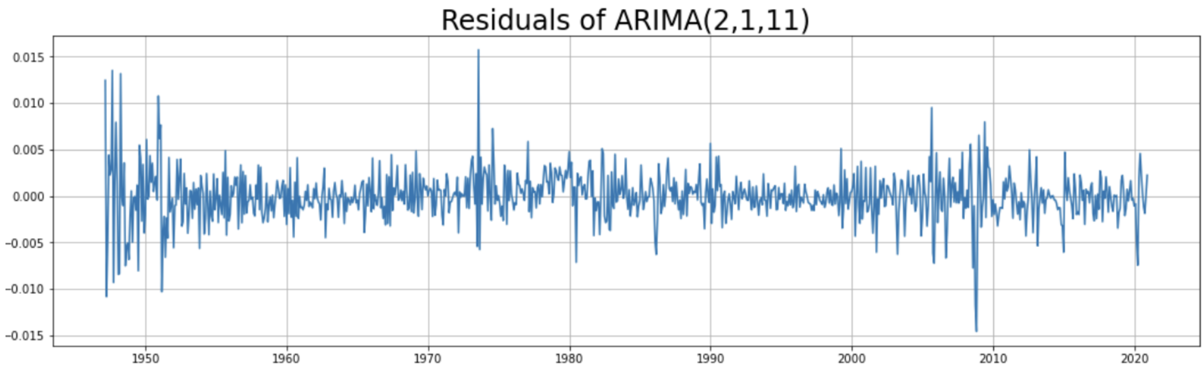


Figure 5: Plot of the residuals of our model

Forecasting inflation

We now compute our time series model in order to forecast inflation on a twelve months period which seems to be the best predictive range in term of accuracy. Moreover, politicians often want short term inflation prediction to be able to take measure to face it if needed. Finally, trying to predict on a larger period of time would probably give less accurate results.

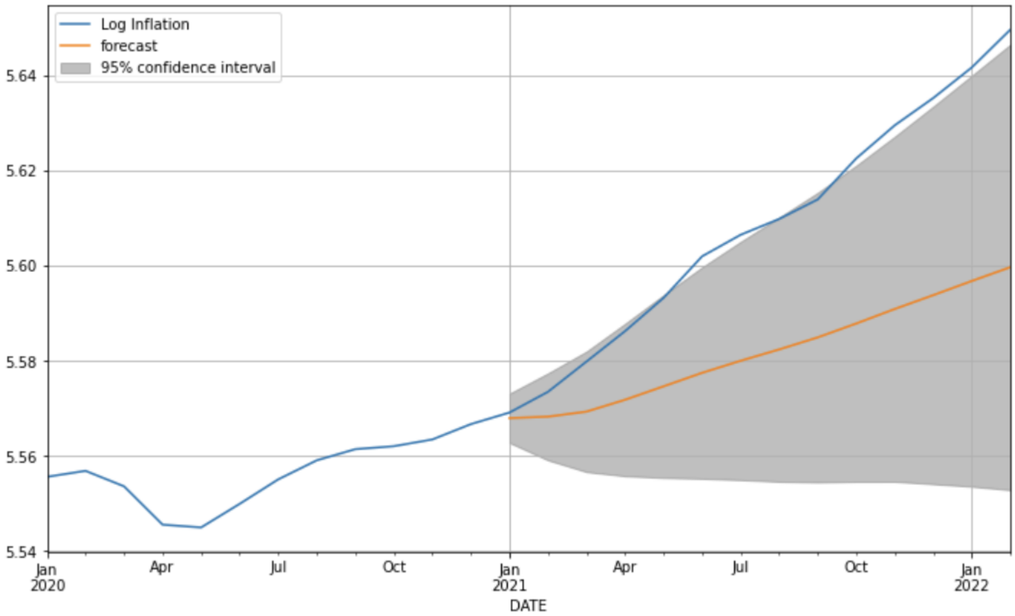


Figure 6: Prediction of log inflation and the 95% prediction interval associated

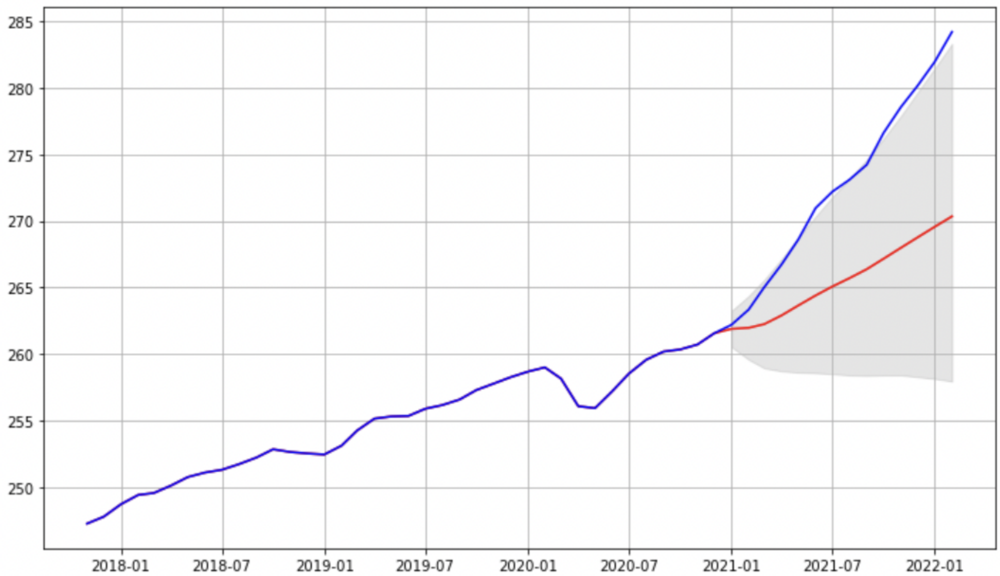


Figure 7: Prediction of inflation and the 95% prediction interval associated. Red: prediction. Blue: real data.

Interpretation of the results

What can be said regarding those results is that the model predicts inflation in a relatively encouraging way. The model predicts an increase in inflation with a regular trend and is located in the 95% interval prediction. This is the case for log inflation prediction and inflation prediction. However, our model only forecasts a trend, the model prediction remains far from real data and doesn't seem to be very sensible to change. The oscillations and fluctuations of inflation are not reproduced by our ARIMA model. As governments and central bank target a 2% inflation and conduct monetary and fiscal policy in that sense, it is quite normal to obtain such a prediction. However, we would like to get a model which is much more sensitive to fluctuations over time and doesn't only predicts a growth which is something well known in economics and not very useful in terms of interpretations.

Different factors can explain this difference between prediction and observation. First of all, inflation is a very complex phenomenon which takes into account lots of parameters and fluctuates a lot over time (even though globally it increases over time). Moreover, history has seen periods of extremely high inflation and others much lower, which makes it even more difficult (it makes inflation prediction even more difficult). Finally, as seen above, the residuals seem to be correlated which is one another source of error.

Now, it wouldn't be possible to use a Time Series model as an inflation predictor. The model remains too far from reality and can't be consider reliable enough to be used as a macro economic predictor. It seems that a time series model is not appropriated enough to study inflation. The model ARIMA(2,1,11) is one of the best as the auto arima model predictor of the software gave us similar predictions. This model was a first step to study and forecast inflation and can't be thrown away. However we need to set up and implement other more efficient models to study and forecast inflation.

2.3 Forecasting inflation with Machine Learning on the CPI time series

2.3.1 A Recurrent Neural Network Model

The first model implemented was the ARIMA model which was based on the CPI time series. The first model using time series was not accurate enough since the forecast fails to predict the high inflation growth of the post Covid-19 crisis. However, the model was simple and results obtained are not sufficient to say that the CPI time series can not be at the beginning of a model that try to forecast inflation. The goal of this subsection is to present a more complex model which would also take into input the CPI time series. This model will be based on Machine Learning and Recurrent Neural Network (RNN).

Introduction to Recurrent Neural Network

Recurrent Neural Network (RNN) is a type of supervised deep learning. Indeed RNN algorithm uses multiple layers of neural networks to perform in processing data and computations on a large amount of data. The principle of RNN algorithm is that output from the previous step is fed as an input to the current step. This means that the output will be restored to the input and used when the next input appears, in the next step. The advantages of using an RNN model are that an RNN algorithm computation depends on historical sequence data and the model size does not increase with input size. It converts the independent activation into dependent, thus reducing the complexity of increasing parameters and remembering each previous output by giving each output as input to the next hidden layer. In addition to classic hidden layers, the algorithm will be composed of LSTM layers which stands for Long-Short Term Memory. LSTM are a unique kind of Recurrent Neural Networks, capable of learning lengthy-time period dependencies.

Creation of two Models

Using a RNN model, two approaches can be implemented. The two approaches are quite similar and they must give equivalent results. The both have their strengths and their inconvenient. An RN model could be implemented directly on the CPI time series. The principal advantage is that the shape of the evolution of the CPI is not complex. However, in order to have a performing model, data may need to be normalized. Normalized the index that is already normalized does not make a lot a sense. By implementing a model on the growth of the CPI time series, this interrogation on the normalization will drop since it totally makes sense to normalize the growth. The output will be a new time series that will face smaller variation around its mean. However, the growth of the CPI index has a more complex shape to predict. Because it is difficult to determine which strategy will have the best result, both models are going to be implemented.

The two implemented models have the same conception. They only differ because the model on the CPI growth is more complex due to the shape of the time series. In order to get the prediction of 1 month, the last 72 months are going to be used in both model. 72 months correspond to the 6 previous years. Take the 6 previous years seems to be a good trade off between taking less data, where a short period of huge inflation in the input will have a strong impact on the output, and taking more data which doesn't seems to make sense since inflation older than 6 years have just a small correlation with the actual inflation. For all the layers, the ReLU function has been used as the activation function. To recall, $ReLU(x) = \max(0, x)$. These two models will have for the first layer an LSTM layer with 100 neurons. Then they both have two classics layers with 50 neurons. The difference between both model appears for the last layers. The model based on the CPI index has then one classic layer with 15 neurons. The model based on the growth of the index has then four classic layers with 15 neurons. Both models have been optimized using an Adam optimizer with a learning rate equal to 0.01 and a mean squared error function as the loss function. They have a different Epoch because the model based on the CPI index has an Epoch equal to 50 and the model based on the growth of the CPI index has an Epoch equal to 100. This was for the conception of the two models based only on the CPI index in input on using Machine Learning. Both model are going to be train on data from January 1947 to December 2019. The result of these two models will be discuss in the following subsection

2.3.2 Analysis of the results

With the two models implemented as explained above, two different type of predictions have been made. First of all, a long term forecasting has been performed. This forecasting is going to be useful because it could be directly compare to the ARIMA model since it used the same data set. Then, a a short term forecasting has been performed. This one is going to be more accurate than the long term forecasting because for each prediction the model is going to use in input real inflation data instead of predicted one. Results are going to show predicted values from January 2020 to February 2022 which was the last available data.

It appears that both models give similar results. They both have their highlight and their weakness which will be discussed. For a redundancy consideration only one model will be discussed for each type of forecasting. For each forecasting, the result of the model that is not discussed will be available in appendices [5]

Long Term Forecasting

In this subsection, results from the CPI index model are going to be shown. It is to remind the reader that with this forecast, the model has the data until December 2019 and gives predictions from January 2020 to February 2022. So, except for January 2020, all the predicted values used other predicted values. The obtained predictions are shown on figure [8].

It is possible to see on figure [8] that the obtained results are closer to CPI index time series than with the ARIMA model on figure [7]. With this model, the period of high inflation is predicted by the model. Moreover, better results are obtained with a larger forecast. Indeed, this forecast is made on 26 months while the ARIMA model performed a forecast on 14 months. Hence, it is possible to say that the RNN model is more efficient than the ARIMA model. This first RNN model gives a trend close to the actual one but it completely fails in giving the variation of the index.

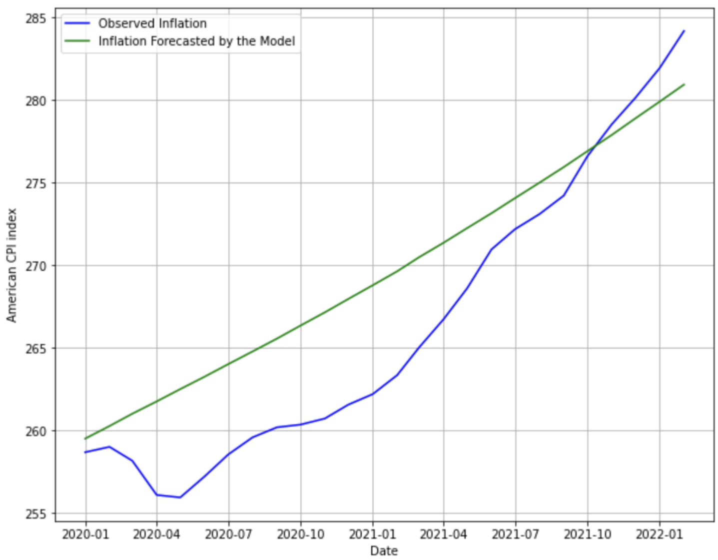


Figure 8: 2 years and 2 months predictions of inflation with RNN model based on CPI index

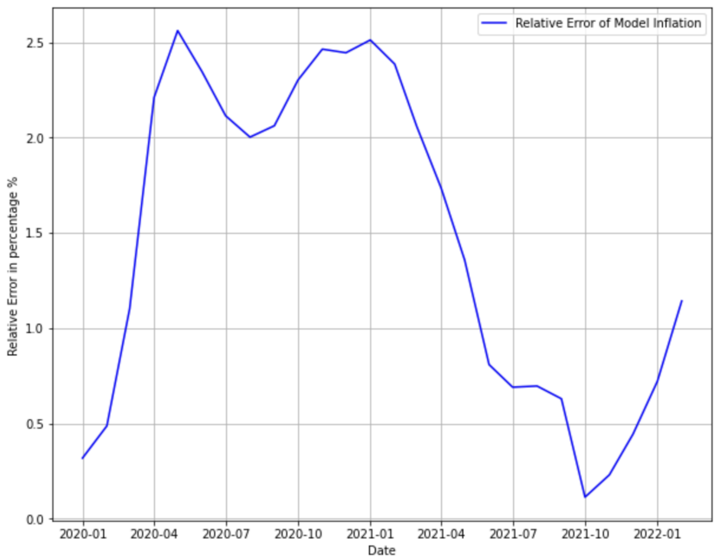


Figure 9: Evolution of the relative error between data and prediction from the RNN model based on CPI index

From figure [9], it is possible to see that the relative error is composed of three steps. The first one, where the relative increases in order to reach 2.5%. Then it decreases until the predictions line crosses the true data line. And finally after this intersection, it increases again. Thus it is possible to conclude that the trend given by the prediction is better that the one obtained with the ARIMA model but it not sufficient since the trends is different from the real one. It seems that the error will continue to grow exponentially if the forecasting is on a longer period of time since the shape at the beginning of 2022 are really different for the two curves.

For the model based on the CPI index growth, prediction's curve has a steeper slope. The conclusion is the same as for the model based on CPI index but the difference between the prediction and true data are amplified by this steeper trend. One explanation will be that for long term forecasting it is more difficult to keep good prediction for the index growth.

Short Term Forecasting

In this subsection, results from the CPI index growth model are going to be shown. It is remind to the reader that with this forecast, the idea is to predict only 1 month using the true data of the last 72 months. With this strategy, predictions are made for every month from January 2020 to February 2022. Results from the RNN based on RNN CPI index growth are the following:

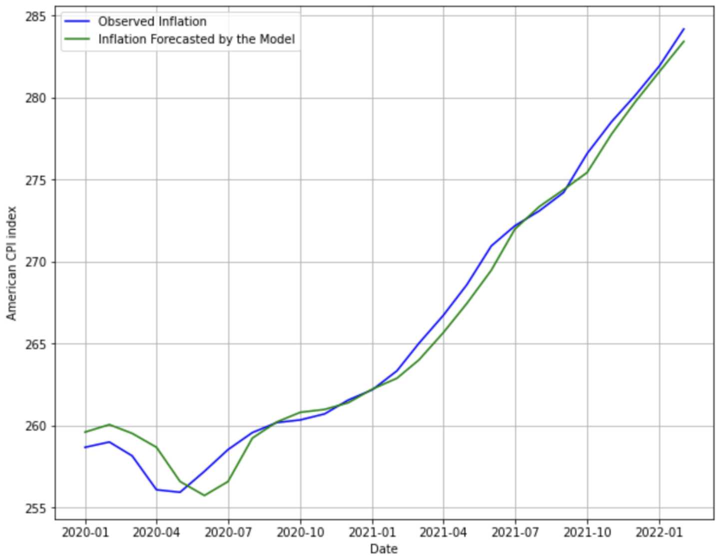


Figure 10: Monthly Forecasting from January 2020 to February 2022 with RNN model based on CPI index growth

Results shown on figure [10] are really close to the real data. However, it is not possible to compare this type of forecasting with all the other forecast that have been made using CPI index in this report. Indeed, the input of the model is corrected each month in order to predict one month with the last 72 monthly inflation measured by the FRED. But it is not because it is impossible to compare this result with previous result that this forecasting result is not interesting. Indeed, this forecast will be useful for the following forecast of the report. The obtained predictions are really efficient all along the data range, the evolution of relative error is shown in figure [11]:

As the figure [11] shows, the relative error of this forecasting is always lower than 1% and the mean is around 0.5%. These results are impressive and could be useful. Machine Learning allows us to know with a huge accuracy where the inflation will be next month. With a short term forecast, it appears that RNN model succeed well in adapting itself and predicting period of huge inflation.

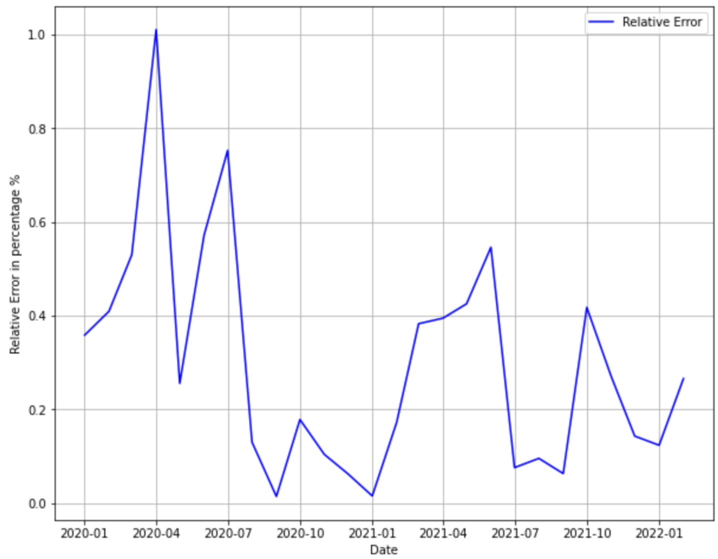


Figure 11: Evolution of the relative error between data and prediction from the RNN model based on CPI index growth for short term forecasting

For the model based on the CPI index, relative error curves had higher values but it was not so far from this model. Both model could have been presented in this part and the conclusion would have stay the same. One explanation of why the CPI index growth model is working better than the model based on CPI index will be that fort short term forecasting a small error on the growth has less impact on the prediction than a small error on the index itself.

Conclusion on this model

The use of Machine Learning and Recurrent Neural Network has given good results but it could be improved even more by adding more macroeconomics data for example. Both RNN model were more efficient than the ARIMA model but it still fails in reproducing the real trend of inflation. However, the use of RNN was a huge step forward in the forecasting of inflation with very efficient prediction for short term analysis. As we only used the CPI inde to forecast test and train our model, it might be relevant to use macroeconomic parameters which are correlated to inflation to describe and forecast inflation more accurately.

3 Forecasting inflation with macroeconomic data

In this section, we will forecast inflation using a wider range of macroeconomic data. First, we will estimate a traditional time series model, the Vector Autoregression model (VAR). Then, we will use Machine Learning technique to estimate a Neural Network.

3.1 Collecting the data

Dependent variable

On the FRED (Financial Reserve Economic Data) website, we downloaded the excel files containing the data since 1991 in the US of the inflation which is the dependent variable. We take the evolution of the CPI. Namely, we take (data of the month of year N divided by data of the month in the year N-1) - 1. In a second step, we wanted to test the efficiency of our model by taking inflation without energy and food.

Independent variables

We have also included 5 other independent variables for which we will motivate our choice here.

Firstly, we have selected the interest rates which is a predictor. It is a monetary policy instrument. We have divided the data by 100 because the data are percentages.

Secondly, we believed that it was relevant to include the money supply. For this predictor, which is also a monetary policy instrument, we take the evolution of the M2 seasonally adjusted. A seasonal adjustment is a statistical technique designed to even out periodic swings in statistics or movements in supply and demand related to changing seasons. It can, therefore, eliminate misleading seasonal components of an economic time series. According to the FRED website, M1 consists of:

- Currency outside the U.S. Treasury, Federal Reserve Banks, and the vaults of depository institutions.
- Demand deposits at commercial banks (excluding those amounts held by depository institutions, the U.S. government, and foreign banks and official institutions) less cash items in the process of collection and Federal Reserve float.
- Other liquid deposits, consisting of other checkable deposits (or OCDs, which comprise negotiable order of withdrawal, or NOW, and automatic transfer service, or ATS, accounts at depository institutions, share draft accounts at credit unions, and demand deposits at thrift institutions) and savings deposits (including money market deposit accounts).

Seasonally adjusted M1 is constructed by summing currency, demand deposits, and other liquid deposits, each seasonally adjusted separately.

M2 consists in M1 plus:

- Small-denomination time deposits less individual retirement account (IRA) and Keogh balances at depository institutions

- Balances in retail money market funds (MMFs) less IRA and Keogh balances at MMFs.

Seasonally adjusted M2 is constructed by summing small-denomination time deposits and retail MMFs, each seasonally adjusted separately, and adding the result to seasonally adjusted M1. Therefore, we choose the M2 seasonally adjusted for our dataframe. Thirdly, we thought that it might be appropriated to add the evolution of the unemployment rates to our dataframe. We have divided the data of this monetary policy instrument by 100 because FRED was giving the numbers in percentages. Fourthly, we made the choice to include the Real exchange rate of the dollar. Here, we also take the evolution of the data as before. Finally, we finished to build our dataframe by adding the evolution of the US wages. For the US wages, we took the nominal wage and subtracted by inflation to get the real wage.

Just to give a small insight of the data we have collected on the Fred's website, here are the graph showing the evolution of inflation taking into account all items and without taking into account food and energy.

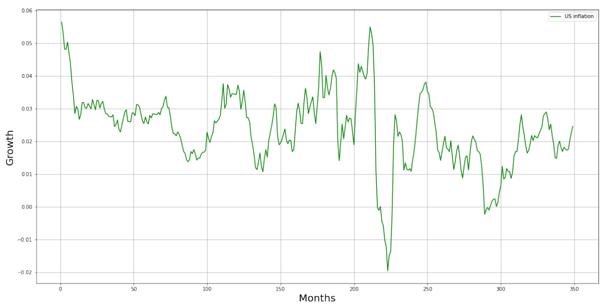


Figure 12: US Inflation

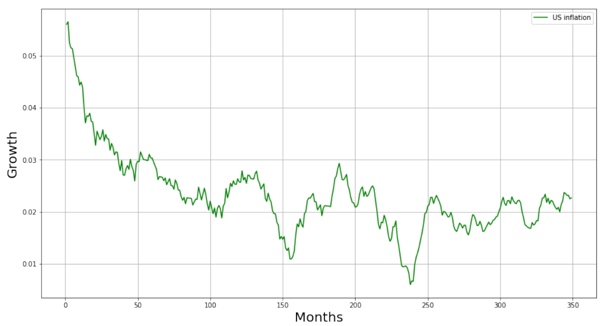


Figure 13: US Inflation less food and energy

Concatenation

To merge these columns we used the RStudio programming software named R. This allowed us to obtain a single csv file named dfmacrodata.

3.2 Vector Autoregression Model

In this section, we shall investigate the predictive ability of a Vector Autoregression model (VAR). VAR models are well established in economics. In their seminal paper, "What does monetary policy do?", Leeper et al. (1996) analysed the impact of monetary policy using such a model. Furthermore, we are interested in comparing this traditional method to the more novel machine learning techniques.

Theory

A VAR is an extension of the autoregression process (AR) to multiple dimension. In fact, a AR process is a special case of VAR, whit only one variable. Therefore, a VAR of order p estimates the following relationship:

$$\mathbf{y_t} = C + A_1\mathbf{y_{t-1}} + \ldots + A_p\mathbf{y_{t-p}} + \mathbf{u_t}, \tag{1}$$

where $\mathbf{y_t}$ is K-vector of K endogenous variables. In our case, K is equal to 6, which corresponds to the six macroeconomic variables considered: inflation, interest rate, M2, unemployment rate, real exchange rate of the dollar, and wages. A_p is the $K \times K$ coefficient matrix of lag p, and $\mathbf{u_t}$ is the K-vector of innovations.
From equation (1), it becomes obvious that the main advantage of VAR over more basic time series model, is that it allows for economical factors to affect each other over time.

Stationarity

In order to estimate a VAR model, it is essential to use stationary series. We therefore perform an augmented Dickey-Fuller test on each variable. In the cases where we cannot reject the null hypothesis of non-stationarity, we take the first difference and perform a second test on the transform variable. The results are show in table 6. After satisfying the stationarity condition, we proceed to the lag order selection.

Variable	Levels	1^{st} Difference
Inflation	0.007*	
Interest Rate	0.290	0*
M2	0.191	0*
Unemployment rate	0.157	0.003*
Real ER USD	0.002*	
Wages	0.005*	

Table 6: Augmented Dickey-Fuller Test − P-Values

Lag Selection

We use a set of four information criteria: the Akaike information criterion (AIC), the Bayesian information criterion (BIC), the final prediction error (FPE), and the Hannan-Quinn information criterion (HQIC). Furthermore, we aim to choose the lag order such that is informative while remaining parsimonious. As shown in table 7, when we allow the maximum lag order to be 14, AIC and FPE suggest 13 lags, while BIC and HQIC use 2 lags. However, as soon as the maximum lag order is limited to 11, AIC and FPE suggest 6, respectively 3 lags. We judge a lag order of 6 to be parsimonious while still capturing the dynamic in the process. Furthermore, we prefer the AIC over the BIC, because the latter assumes that the true model is among the set of candidates. We therefore proceed to estimated a VAR(6).

	AIC	BIC	FPE	HQIC
Lag Max 14:				
	13	2	13	2
Lag Max 11:				
	6	2	3	2

Table 7: Lag selection statistics

Normality

To further assess the estimation quality of our VAR(6), we check if the normality condition is met. While we reject the null hypothesis of the Normality test and the Portmanteau test, the Durbin-Watson test suggests that there is no serial correlation of the residuals. Because it is known that those tests have lower power, we finalize our analysis by investigating the auto-correlation of the residuals. As shown in figure 14, we see that all auto-correlations lie within the 95% confidence interval. This observation confirms that we can continue with this parameterisation.

Figure 14: VAR(6) Auto-correlation of residuals

Notes: ACF plots for residuals with $2/\sqrt{T}$ bounds. The rows and columns are ordered in the following way: inflation rate, real ER USD, real wage, interest rate, M2, unemployment rate.

Interpretation

To gather further insides of the model, and in particular how inflation reacts after a shock to itself or one of the other 5 variables, we consider the Impulse response function (IRF). As figure 15 shows, after a shock to inflation, all variables revert to zero, expect for the first difference of the US interest rate, which oscillates around 0 before stabilizing at a positive level.

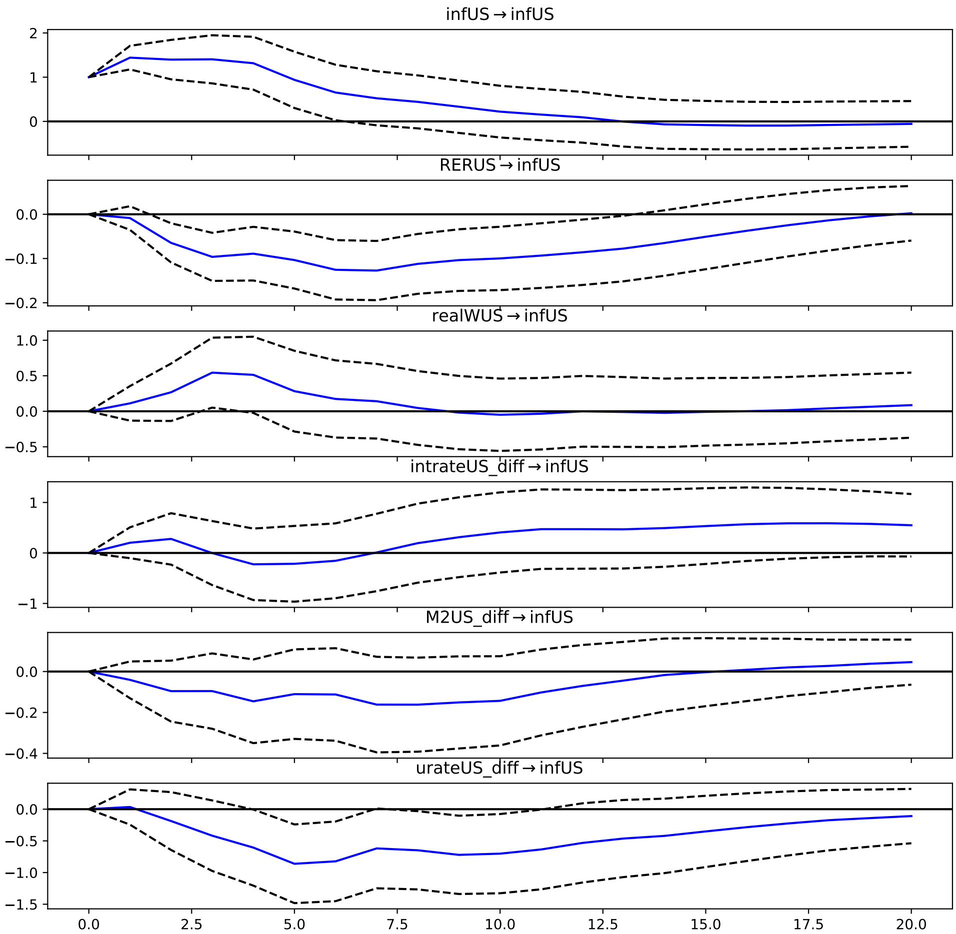


Figure 15: IRF, response of inflation.

Those dynamics become even more explicit when we consider the Cumulative IRF (CIRF). From figure 16 it becomes apparent that a shock on inflation has a permanent impact on all variables, ie. they never return to their original level.

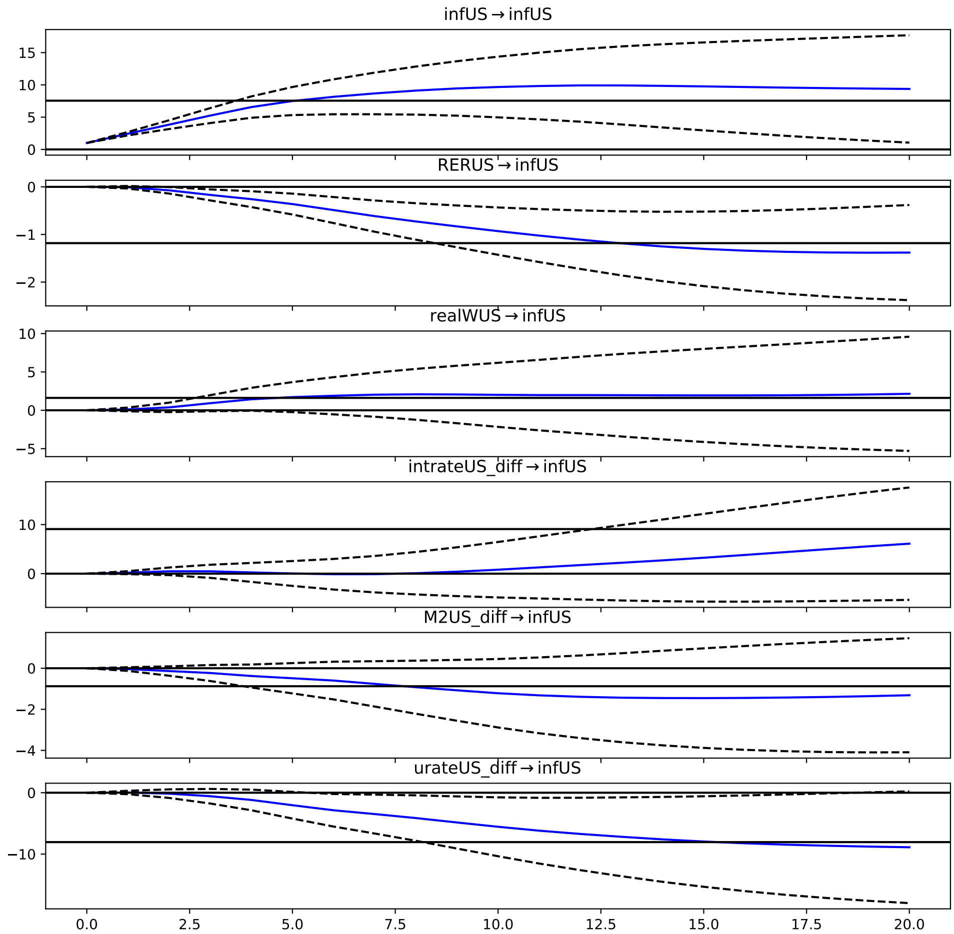


Figure 16: Cumulative IRF, response of inflation.

Inflation forecast of VAR(6)

Finally, we aim to evaluate our forecast, which is obtain after back transforming our first differenced variables. First we will consider the forecast of inflation, and second the forecast of the CPI.

Inflation

Figure 17 depicts our forecast for 2019 against the actual inflation data. We observe that in the first half, our struggles to capture the dynamic of inflation. However, in second half, the VAR(6) correctly forecasts the increase in inflation rate. The forecast performance of the model also reflects this fact. The relative error of the forecast (figure 18) ranges between −20% and +20%, with the strongest jumps at the beginning.

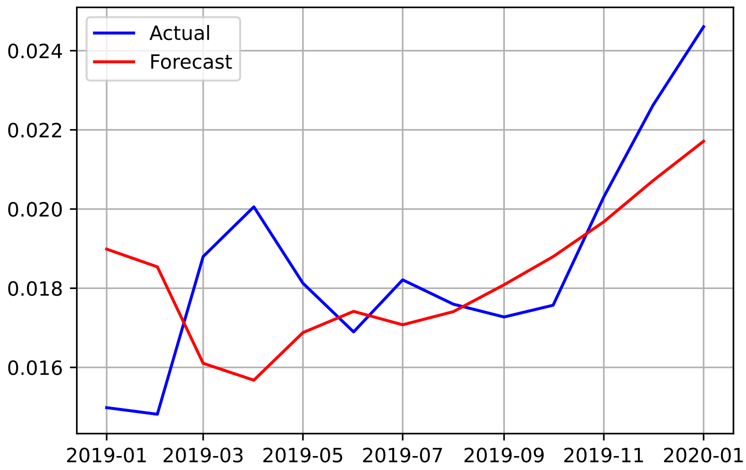


Figure 17: VAR(6) Inflation forecast

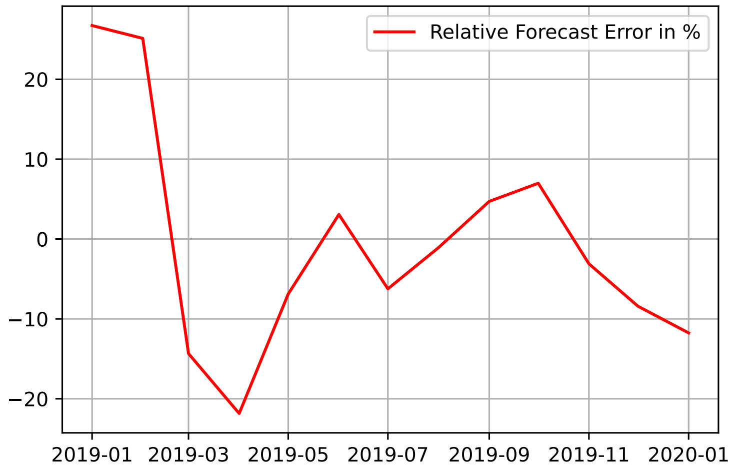


Figure 18: Relative error of the forecast of inflation

Consumer Price index

However the VAR(6) performs much better when forecasting the CPI. Figure 19 shows that our forecast follows closely the development of the CPI. Moreover, the relative forecast errors are much smaller, they approximately range within $[-0.4, 0.4]$ (see figure 20). This is mainly due to the nature of the two time series: the CPI is the level of the price, whereas the inflation is the change of of the price level compared to the last year's month. As a result, deviations from the actual CPI do not appear as severe as deviations from the inflation.

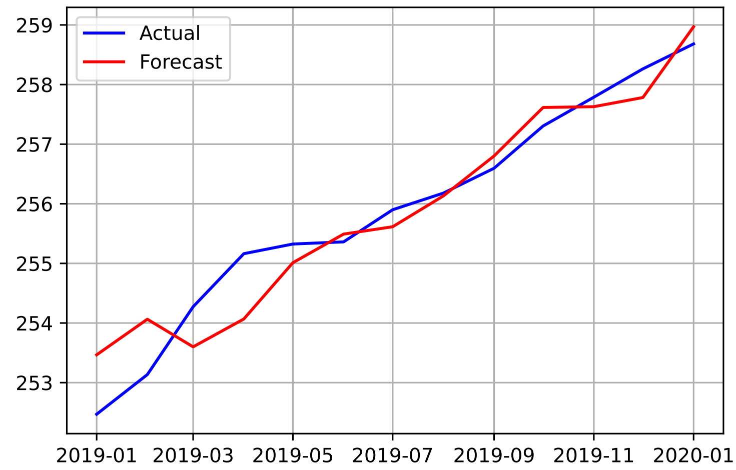


Figure 19: VAR(6) CPI forecast

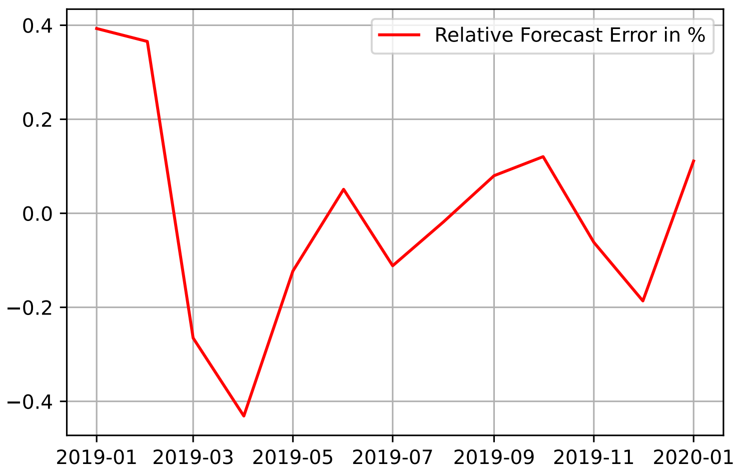


Figure 20: Relative error of the forecast of the CPI

We conclude that the likelihood of capturing the correct dynamics of inflation dynamics with a VAR model is low, but higher when predicting the CPI. Objectively speaking, the Relative Root Mean Square Error (RRMSE), given equation (2), confirms this. The RRMSE for predicting inflation is 0.1345, versus 0.0022 for the CPI prediction. In a further study, we could extend the set of macroeconomic variables, as well as impose structural restrictions based on economic theory, ie. estimate a structural VAR.

$$\mathrm{RRMSE} = \sqrt{\frac{1}{n}\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(\hat{y}_i)^2}} \tag{2}$$

3.3 Nowcasting inflation with Machine Learning

3.3.1 Creation of the Neural Network

Structure

To build the neural network, we were inspired by lecture W9b course entitled Feed-Forward Networks and in particular by the following figure 74. We have used the tensorflow and keras library.

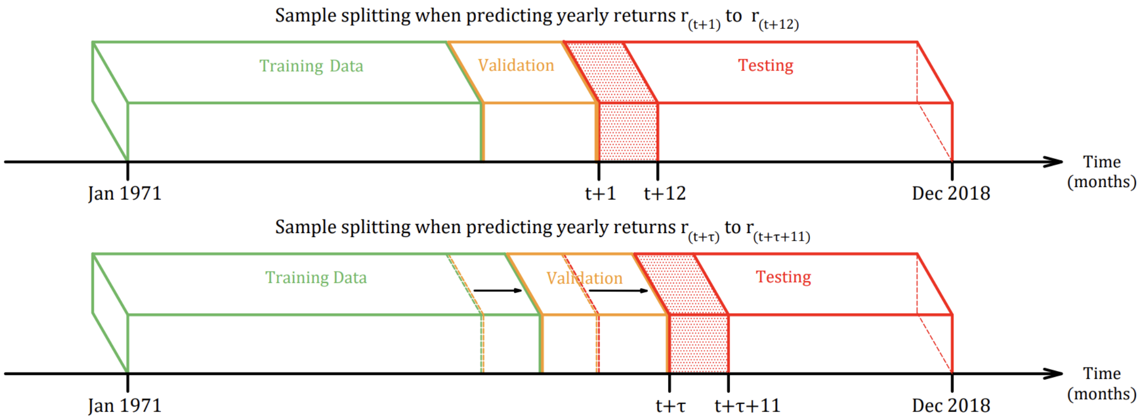


Figure 21: Data and sample splitting

To start with, we constructed a training set (data from 1991 to 2015) which is named dataset. We also created a validation set (data from 2015 to 2018) that we have named valset. For the test set, we have taken the data from 2018. The goal is to predict 2019 and to see if it is close to the reality. Thus, we will be able to know if our neural network is efficient. Then we have reshaped the data so that they can fit to the model. We have finally implemented the split function that splits a multivariate sequence into samples. We have chosen 12 observations (12 months) to predict 12 observations (12 months).

Parameters

For the activation function (that will determine the output layer), we performed different tests. We have first chosen the hyperbolic tangent because we are doing a regression with bounded outputs. Finally, we decided to use the ReLU function because it is the easiest one and above all, because it gives the best results.

For the first 2 hidden layers, we used 100 Long Short-Term Memory cells. It is a type of recurrent neural network. The advantage is that outputs could become input. As a matter of fact, there is a cell in the node named “state”. This cell is divided in 3 parts : the forget gate, where we decided what can be forgotten (0 or 1), the input gate, where we choose what new information should be added (0 or 1) and the output gate in which we say which part should be output (0 or 1).

We also chose to use a dropout. Thanks to this parameter, when you go from one stage to another, the program drops out a certain percentage of data. For instance if we set the dropout equal to 0.5, half of the data is dropped. For this project, we chose a dropout and a recurrent dropout equal to 0.25. Then we created 3 other classic hidden layers with 50 neurons. To finish, our output layer has 12 neurons (for the 12 next months). The learning rate is set at 0.0005. We have included the early stopping tool. Early stopping stop training before the model is suffering from overfitting. Every end of epoch, we check and stop training if validation accuracy decreases steadily, while train accuracy increases. In our model, the early stopping appears at the 51^{th} epoch.

As seen in class, the epoch is the number of time the learning algorithm works through the entire training set. We define the epoch equal to 1500. The reason for that is that it is approximately the one that allow us the reproduce the same variations. The mindelta was fixed at 0.01. It's the threshold for measuring the new optimum, to only focus on significant changes. The patience is set at 30. It's the number of epochs with no improvement after which learning rate will be reduced. The verbose is set at 0. By setting verbose 0, 1 or 2 you just say how do you want to 'see' the training progress for each epoch. We then obtain our $\hat{Y}$ that is a list of 12 elements representing the growth for the next 12 months.

Adjustments

As you can see on figure 22, at the beginning, the validation loss was first below and then above the training loss. In theory, the training loss should be lower. Here the validation was lower at the beginning because during the training process, we were using a dropout of 0.25 so 25% of the neurons were set to zero at each time.

Then, we have modified several parameters such as the learning rate, the number of layers, the number of neurons and the nature of the activation functions. Thus, model during testing is more robust as you can see on figure 23. Both training and validation errors seem small. We have done a good choice of dropout and now, we don't face overfitting.

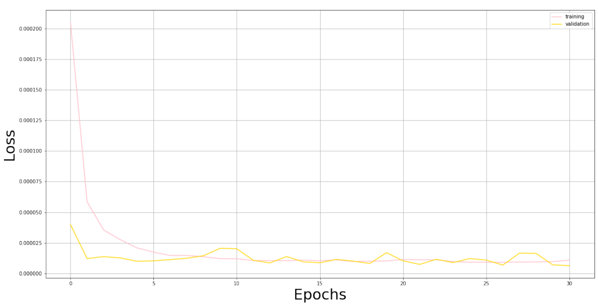


Figure 22: Loss Curve before adjustments

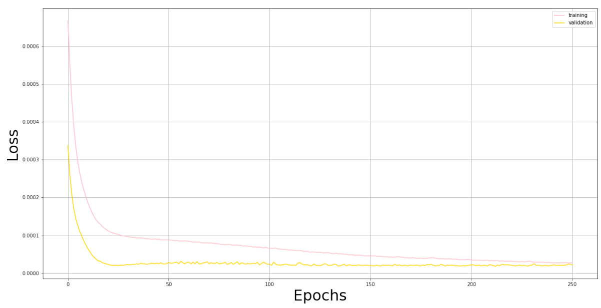


Figure 23: Loss Curve after adjustments

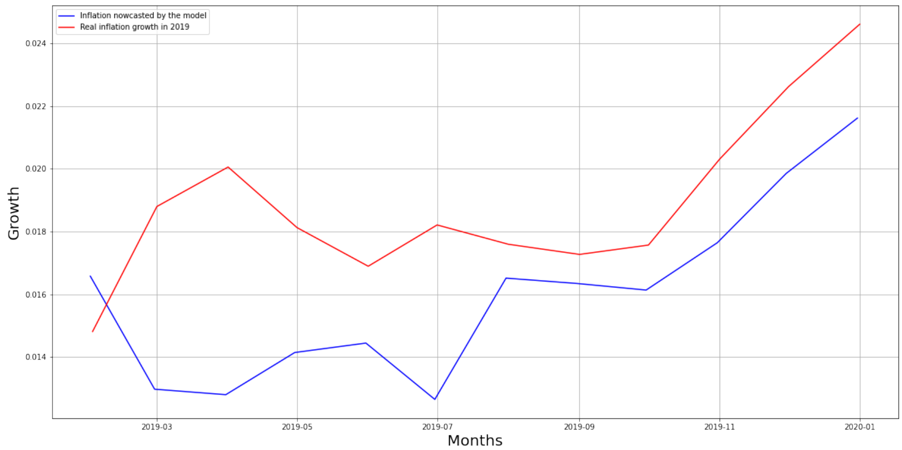


Figure 24: Prediction vs Observation for the Growth

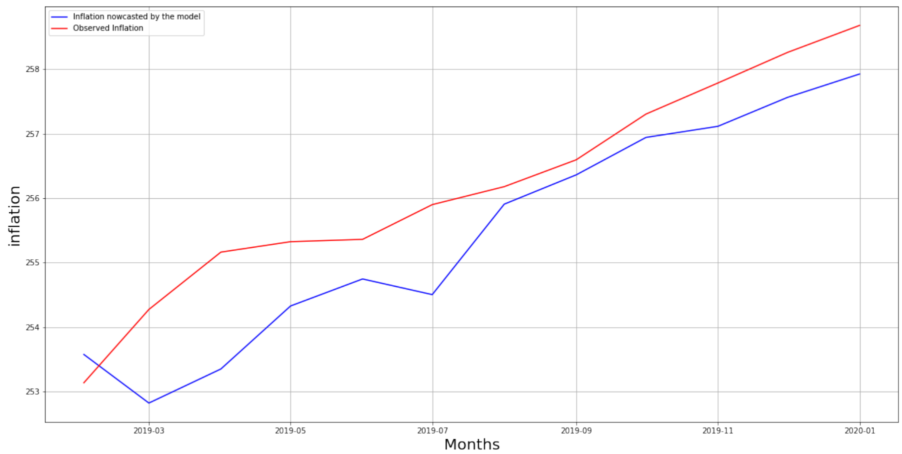


Figure 25: Prediction vs Observation for the CPI Index

3.3.2 Analysis of the results of the Machine Learning Nowcasting

Analysis of the error

Just to recall, the output of our model is the growth of the CPI for 12 months from February 1, 2019 to January 1, 2020. We have computed the absolute error. The absolute error obtained on the last try is equal to 0.322249.

This small example allows us to understand to what extent the modification of a parameter can have an impact on the error.

- 0.00619 with the hyperbolic tangent (in absolute value)
- 0.000605 with the ReLU function (in absolute value)

Here, just by changing the activation function, we obtained an error ten times smaller.

Following the teacher advice, we have also computed the relative error thanks to the following formula:

$$error = \frac{\mid predicted - observed \mid}{observed} * 100$$

This represents the mean of the difference between observed and predicted data. For the growth, the relative error obtained on the last trial is equal to about 14.94%.

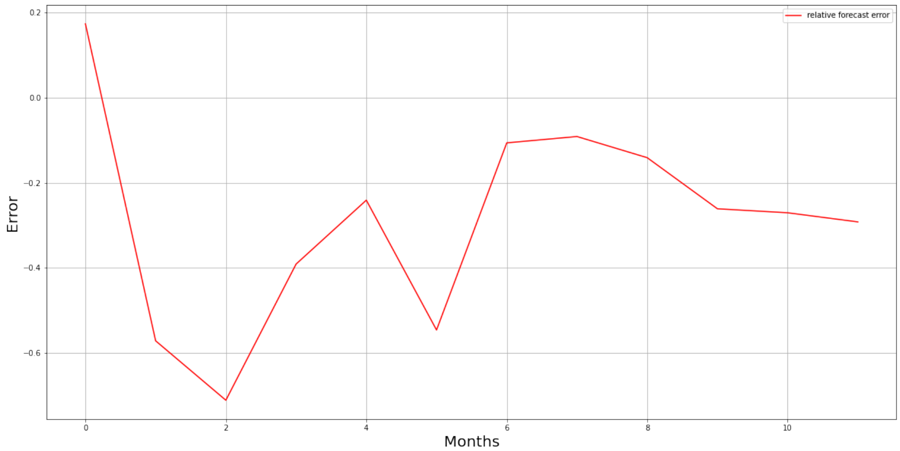


Figure 26: Relative Forecast Error

Output Processing

As the output of our model is the growth of the CPI for 12 months, we thought it was clearer to use the growth predictions to get the predictions of the CPI evolution.

Once we have included the evolution of the CPI, we again used the formula below to calculate the relative error.

$$error = \frac{\mid predicted - observed \mid}{observed} * 100$$

For the evolution of the CPI, the relative error obtained on the last trial is equal to about 0.29%.

3.3.3 Limits of the model

We believe that the limitation of our final model is that we could have created an even greater dataframe. Indeed, the more inputs there are, the more efficient the neural network can be. We could have added for example the gross domestic product. We could also have tried to process the data further. For example, we could have tried to take the growth in the US interest rate instead of just the US interest rate. To finish another original but efficient idea might have been to add the google trends data to the dataframe.

We also tried to test our model by taking the inflation without food and energy and as we can see below, the results are even more satisfactory.

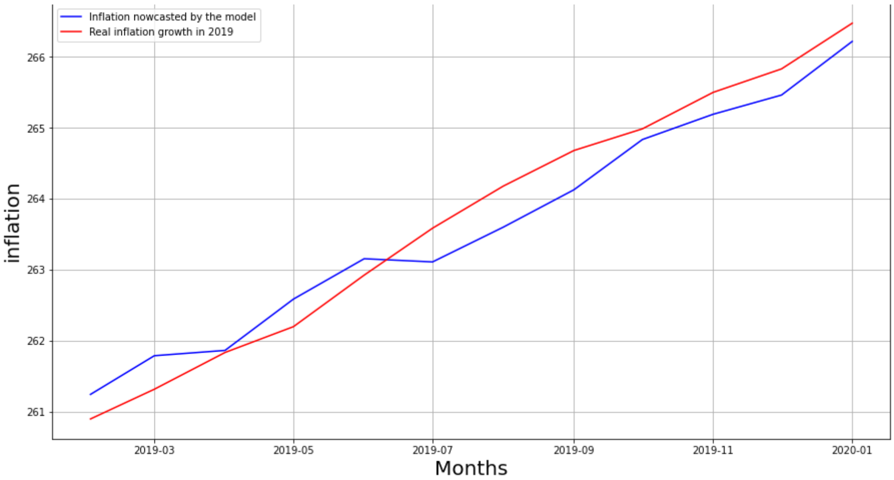


Figure 27: Prediction vs Observation

Conclusion of the model

Despite these limitations, we are proud of this final model. As a matter of fact, the relative error is very small and above all, over 12 months, we managed to predict if the growth of the CPI will increase or decrease more than half of the time. Namely, an investor who is following our recommendations will be profitable more than half of the time. This investor can take a long or a short position on the “Lyxor EUR 2-10Y Inflation Expectations UCITS ETF - Acc” depending on the forecast of our neural network. The first two parts of the report studied both time series and machine learning methods on tho different data set. We will know focus on a radically different approach to study and forecast inflation using google trends.

4 Forecasting inflation using Google Trends

4.1 A brief overview

This part provides what can be qualified as a more "exotic" way to forecast inflation. We decided to look at the forecasting problem through another angle. Indeed, we had the idea of observing inflation by looking at the main group impacted by it : the full population in the US. Having a sufficiently large sample of data to represent the population through surveys or other methods to ask people about they changing habits in consumption is very likely to express bias (people lying or do not remember, etc.). Hence we decided to use a neutral informative tool : Google trends.

According to its web-page, Google trends is a tool from *Google Labs* allowing the user to see how often specific keywords, subjects and sentences have been queried over a specific period of time in a specific geographic region.

Here it is straightforward that we will focus on the USA, moreover we decided to look at monthly data from 01/2004 (the oldest we can get from Google) to 02/2022. We also had to chose the words to look at in order to model the inflation with the less error possible. We had the idea that if global prices decrease, then an individual will have more buying power and then is more likely to search for high investment and high price goods. This led us to select the following set of 17 words to query in Google trends engine: amazon, booking, car, crisis, flight, ford, fridge, glassdoor, holidays, housing, inflation, las vegas, linkedin, oil, unemployment, unemployment benefits, usajob.

It is worth to mention that *Glassdoor* and *USAjobs* are both websites helping to find jobs in the USA.

4.2 Collecting the data

The main challenge we faced downloading the data is the fact that it does not provide a raw level of queries but scaled numbers from 0 to 100, playing the role of an index. To obtain this query index, a query share is first calculated by dividing the query volume of the searched term by the total number of searches we asked in the specified region and the given time frame. Then, the month with the highest relative search volume is normalized to 100, so a number of 50 indicates that the proportion of queries at this time is half the time where we get 100 [1].

This way of giving the information is a problem if we directly try to make a model on it because the different queries are not comparable. Indeed, a 100 index for the word "car" does not mean the same as a 100 index for the word "housing" since both words certainly do not have the same number of searches per months.

A solution to this problem would be to enter all the words in one request to Google trends so the share will be calculated over the whole amount of queries of all the words and then the data will be comparable.

Unfortunately, Google trends website allows us to do a maximum of 5 researches per request, which is way less than our 17 words. To manage this problem, the idea is to

[1] "The resulting numbers are then scaled on a range of 0 to 100 based on a topic's proportion to all searches on all topics." (Google Trends FAQs)

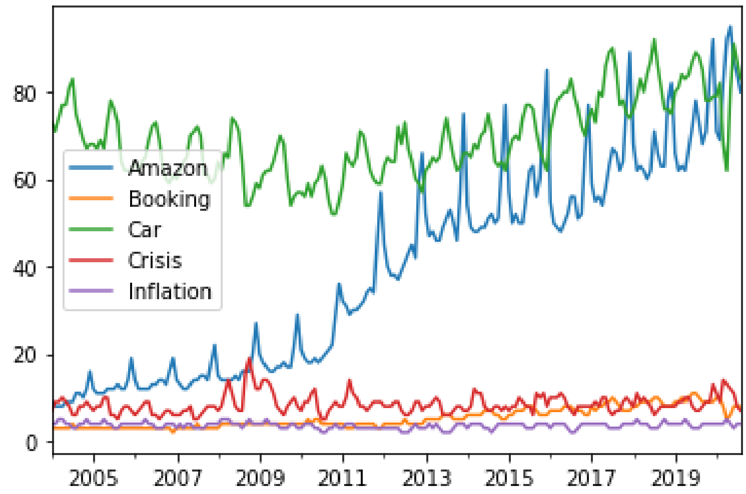


Figure 28: Google trends of few words of interest

make one request per word but using a "control" term each time to make all the data comparable. We have to use a word that we expect to not vary too much over time, to do so we looked at country names since we do not expect them to be too much impacted by interest variations of the population. We finally chose to use the term "Canada".

The figure 28 provides a quick overview of the shape of the data we get for some words thanks to Google trends[2].

4.3 Performing the OLS

The idea in this section is to seek for the best possible model through Ordinary Least Squares method (OLS) to approximate the inflation represented again by the CPI.

Training and test sets

Before going deeper in a model creation, we had to split the data into a training and a test set (respectively from 01/2004 to 01/2021 and from 02/2021 to 02/2022) in order to first select a model and then test the goodness of fit on unknown new data with the test set. At first, only one query was made to get all the data from 01/2004 to 02/2022, but it appeared that the training set and the test set were sharing the same scale so the training set was scaled knowing the test set, creating a look ahead bias (knowing the future has an impact on the model selection).

To illustrate this bias, lets take an example. Suppose for one word the maximum spike of queries is hit at a time t in the test set. Then the score will be of 100, and because of the *Google trends* scaling system, the highest number of queries for this same word within the training set (which is reached at a time denoted by t^*) will have a score bellow 100. This shows that at time t^* the score is influenced by data from the test set, and then the

[2] It is important to note that Google Trends data is computed using a sampling method, and the results therefore vary a few per cent from day to day.

test set and training set are not independents.

To overcome this problem, we decided to split the two queries for the two sets. However, doing so led to another problem due to frequency. Indeed, *Google trends* gives data with a frequency depending on the time laps we give. Here the training set data was given by month, which is what we want regarding that the fact that the CPI data is also monthly. But since the test set is of 1 year, the data is returned weekly, creating a problem of matching frequency with the CPI data.

Since the main problem we want to avoid is the look ahead bias in the training set, we proceeded by first downloading separately the training set (data from 01/2004 to 01/2021) and then to get the test set the full period data was downloaded and we took only the last year of interest (from 02/2021 to 02/2022). This way of collecting the data allows then to clear the bias coming from knowing the future in past data.

First approach

To begin with, a model containing all words has been sought in order to explore the global impact of Google trends' data on the inflation (we call it *model 1*).

As we can see on the Figure 29 presenting the fitted versus true values, the model estimate well the trend of the CPI. It is corroborated by looking at the goodness of fit through the R^2 which is of 0.963, indicating a good fit of the model (adjusted $R^2 = 0.960$).

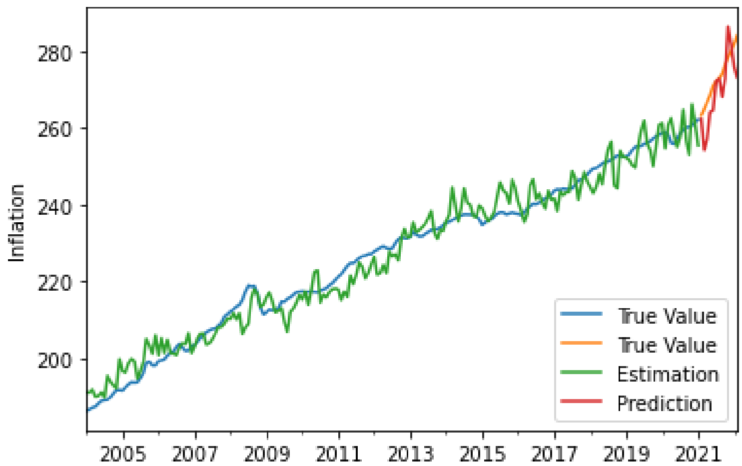


Figure 29: Inflation true value and modeled over time with *model 1*

This led to the following Table 8 showing the coefficients of the OLS regression. We also have the Akaike Information Criterion (AIC) equal to 1766 and a Bayesian Information Criterion (BIC) equal to 1773.

Then, we wanted to know how good is the fit thanks to the evaluation with the test set in greater details. The results are shown is the Figure 30 and the absolute error with respect to the true CPI value is depicted in Figure 31. We can see on the Figure 30 that the increasing trend if fairly well approximated but a bit underestimated.

Name	coef	std err	t	P> |t|	[0.025	0.975]
const	211.8477	6.616	32.020	0.000	198.796	224.899
Amazon	0.0714	0.061	1.175	0.242	-0.049	0.191
Booking	1.2325	0.572	2.155	0.032	0.104	2.361
Car	-0.4257	0.076	-5.596	0.000	-0.576	-0.276
Crisis	0.1770	0.192	0.922	0.358	-0.202	0.556
Flight	0.2682	0.053	5.022	0.000	0.163	0.374
Ford	-0.3192	0.087	-3.686	0.000	-0.490	-0.148
Fridge	2.8368	0.318	8.930	0.000	2.210	3.463
Glassdoor	-0.9261	0.339	-2.729	0.007	-1.595	-0.257
Holidays	-0.0243	0.047	-0.518	0.605	-0.117	0.068
Housing	0.3134	0.116	2.706	0.007	0.085	0.542
Inflation	0.5585	0.614	0.909	0.365	-0.654	1.771
Las Vegas	0.1011	0.065	1.550	0.123	-0.028	0.230
Linkedin	0.3230	0.042	7.718	0.000	0.240	0.406
Oil	0.0898	0.044	2.045	0.042	0.003	0.176
Unemployment Benefits	0.0528	0.276	0.192	0.848	-0.491	0.596
Unemployment Insurance	-1.7800	0.944	-1.886	0.061	-3.642	0.082

Table 8: Coefficients of the *model 1* with all words

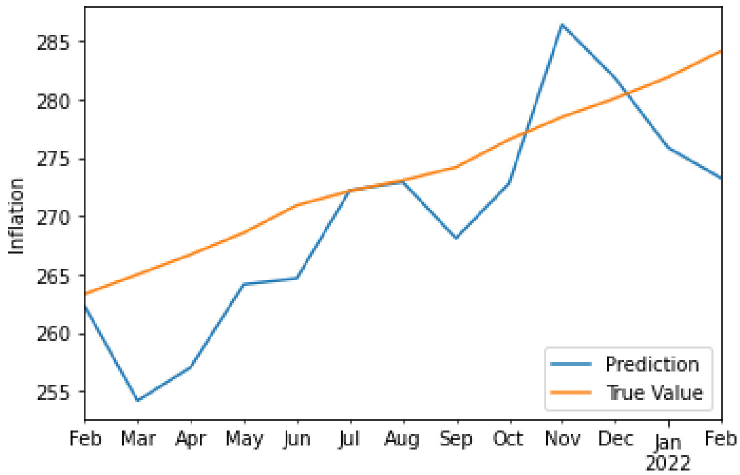


Figure 30: Focus on the test set for *model 1*

Looking at the absolute error in Figure 31, we can see that it does not go above 4%, which implies that the model fits pretty well the inflation.

However, despite having good results in terms of fitting, the *model 1* has to be improved. Looking at the Table 8 we can see that many p-values are above the usual threshold of 0.05, meaning that the hypothesis H_0 "The coefficient is different from 0" can not be rejected. As an example, the word "Amazon" has a p-value of approximately 0.242 and the its confidence interval at 0.05 is $[-0.049, 0.191]$ and then we we can not reject the possibility of a 0 coefficient at a 0.05 threshold.

We need to be cautious with selecting just with respect to the p-value since it is looking

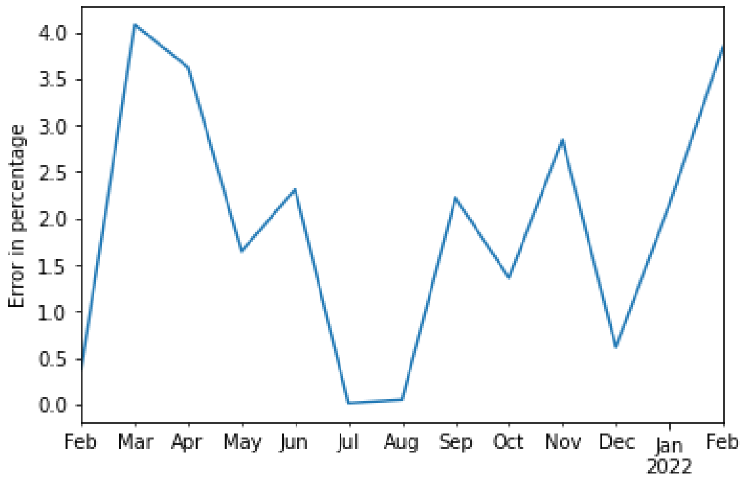


Figure 31: Absolute error of fitting within the test set for *model 1*

at the coefficient alone, a way to assess the deletion of a coefficient from the regression is then by looking at the error of fit in the test set.

It is also important to mention that there are extremely high Variance Inflation Factors among the coefficient with the word "Ford" as an example with a VIF of 402.90. This shows the presence of high multicollinearity among the explanatory variables. An usual rule for the VIF value is to accept it if VIF$<$ 10 (rule of thumb), which is not the case here for many coefficients as shown in the Table 9.

Features	**VIF**
Ford	402.90
Car	400.77
Housing	159.41
Booking	132.76
Flight	106.87
Amazon	91.06
Fridge	89.11
Oil	80.69
Las Vegas	80.25
Inflation	56.16
Crisis	34.42
Glassdoor	34.40
Linkedin	22.79
Unemployment Insurance	14.89
Unemployment Benefits	13.88
Holidays	1.62

Table 9: VIF of the *model 1* with all words

Refinement

To overcome those problems we decided in a first place to perform a backward elimination. It means that we start with the full model and at each step we delete the highest coefficient's p-value, above 0.05, and then recalculate a new OLS model with previous word but the one deleted. After performing this, we finally obtained a list of words with p-values below than the required threshold of 0.05. The new words in our model are now: Car, Flight, Ford, Fridge, Housing, Linkedin and Oil. We call this new model as *model 2*.

Then we looked at the VIF again of the explanatory variables of *model 2* which values are show in the Table 10. As we can see, we still have very high factors, so we had to find a way to manage multicollinearity.

Features	**VIF**
Car	338.31
Ford	290.19
Housing	90.87
Oil	60.98
Flight	43.45
Fridge	14.08
Linkedin	8.81

Table 10: VIF of the *model 2* after the backward elimination

To do so, we first thought that the two words Ford and Car must be very collinear because explaining the same field : cars industry. To remediate this problem we decided to remove the word Ford.

After this, we thought that Car and Housing were both explaining a wish of settlement from the people making the queries on Google. So we decided to create a new variable called Settlement being the addition of Car and Housing.

All those manipulations led to a final model called *model 3* with coefficients shown in the Table 11. We can see that all p-values are very small indicating a meaningfull impact on the regression. Moreover the R^2 indicator of the goodness of fit is equal to 0.96 (and adjusted $R^2 = 0.959$) which shows that we lost only very little part of the explanations induced by the explanatory variables with respect to *model 1*.

Looking at the AIC we can see that its value is of 1176 which is less than the one of *model 1* (and same for BIC = 1203) which shows that the *model 3* is not loosing much likelihood while being way more parsimonious.

Looking at the Variance Inflation Factor values in Table 12, we can see that they decreased greatly from the first and second model we made. However three of them are still above the rule of thumb so the multicollinearity is still strong.

On the graph of Figure 32, we can see that the fit of the training set is reliable and we can next focus on the test set approximation.

Looking at the approximation of the CPI from March 2022 to February 2022 in the Figure 33 we can see that the trend is fitting well but the model underestimate the true value of the CPI. Looking at the absolute error of this estimation depicted in the Figure

Name	coef	std err	t	P> |t|	[0.025	0.975]
const	216.1945	3.760	57.505	0.000	208.781	223.608
Flight	0.2945	0.040	7.379	0.000	0.216	0.373
Fridge	2.3648	0.124	19.069	0.000	2.120	2.609
Linkedin	0.2521	0.028	9.050	0.000	0.197	0.307
Oil	0.1309	0.044	2.953	0.004	0.043	0.218
Settlement	-0.3290	0.040	-8.244	0.000	-0.408	-0.250

Table 11: Coefficients of the *model 3* with selected words

Features	VIF
Oil	57.55
Flight	41.72
Settlement	36.07
Fridge	8.58
Linkedin	7.12

Table 12: VIF of the *model 3* with selected words

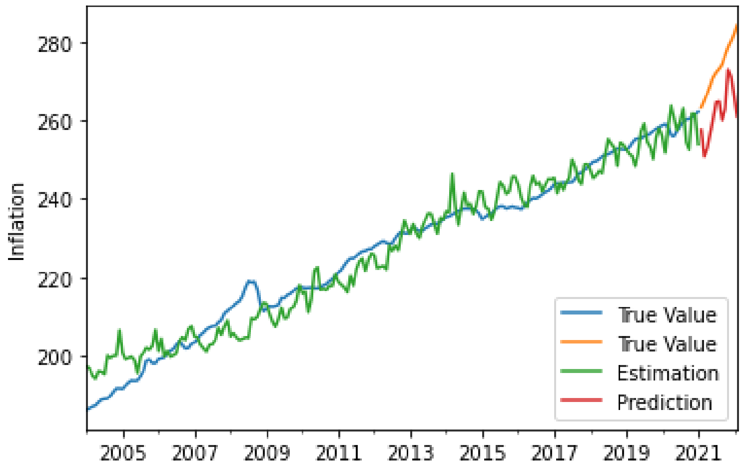


Figure 32: Inflation over time with *model 3*

34, we can see that it is higher than the one of the *model 1* but under the threshold of 8% which is fairly acceptable.

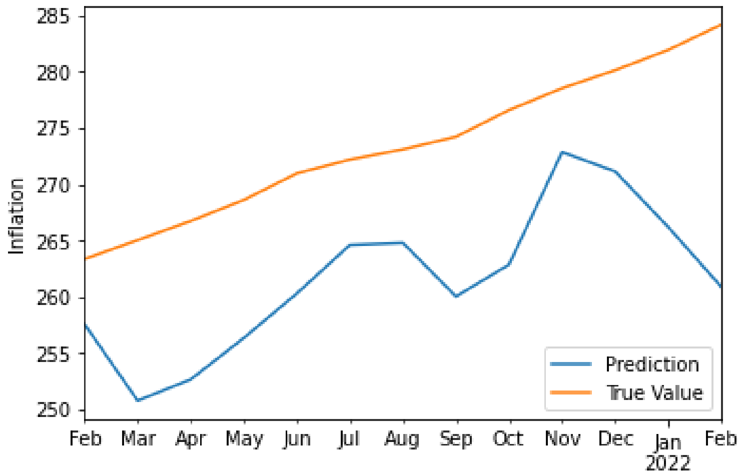


Figure 33: Focus on the test set for *model 3*

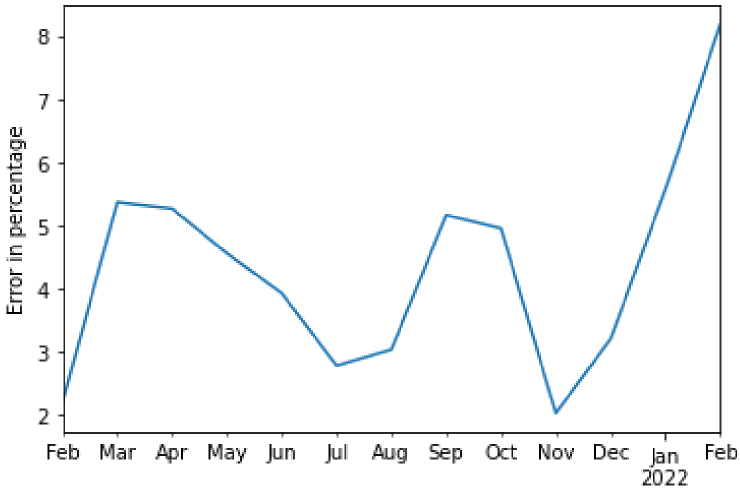


Figure 34: Absolute error of fitting within the test set for *model 3*

4.4 Interpretation of the results

It is difficult to interpret the coefficients' value because of multicollinearity, but at least we can take a look at the sign of those. It appears that the variable Settlement has a negative value, meaning that the terms related to it (Car and Housing) are expected to be less queried in times of high inflation. This seems straightforward to analyse, indeed consumers are less likely to make big investment in times of high growth of prices.

Looking at other terms we can see that Linkedin term is positive, so it might indicate that people are more likely to search jobs or at least opportunities when the CPI is high. Nevertheless, it is important to temper the analysis by noticing that the CPI globally always increased (in trending term) through the past years, and the it does not mean that Linkedin search queries impacts inflation but just that there exists a positive correlation between both.

Conclusion of the model

To conclude on the Google trends data method, as we can see the final *model 3* is showing a good fit to inflation and this can be seen thanks to the R^2 and also the two information criteria AIC and BIC. Moreover, looking at fit with respect to the test data, we can see that the model is able to approximate the inflation with a fairly small error. However, despite those good news it can be seen through the VIF values that the multicollinearity is still strong and after trying to reduce it at the minimum it is still above the rule of thumb for some variables. Unfortunately we were not able to manage more variables, leading to a worse model in term of the approximation of the CPI in the test. There is also a problem of underestimation when trying to select a more parsimonious model raising the original error of the *model 1* from a maximum of 4% to 7% approximately. This last model on google trend revealed to be interesting to study as it illustrates a relative different way to study and forecast inflation than the usual methods.

Although the method used is accessible (OLS), the data required considerable work and the results were extremely satisfactory given the difficulties encountered.

5 General Conclusion and Opening

We provide here a global conclusion as the specific interpretations of the results obtained through the different models have been described in their dedicated parts. It is relevant to notice that all the models where linked and did not come up independently in the project.

The project was divided in three parts which were however strongly related as the results of any model lead to the introduction of another one in order to improve the results and obtain a different or new benchmark.

The first idea was to use data on the evolution of the Consumer Price Index (CPI) in the USA from 1947 to 2022. At first, these values were the only ones used to train models in order to forecast inflation. On this set of data, we implemented what can be consider as the most evident model, an ARIMA model. The results were certainly encouraging but not sufficient for our study. Only a global trend was predicted (growth of inflation) which would not be of much help to forecast inflation and evaluating its volatility. Thus, we decided to use machine learning and we came up with a recurrent neural network with the evolution of the CPI as data. The neural network has been used to predict directly CPI and also inflation growth. In the last case, the predicted growth was used to calculate CPI. This approach gave very successful outcome, inflation has been well predicted with a relative error doesn't exceeding 1% in absolute value. The first models used and especially the neural network were very encouraging as we only used one type of data (CPI index), the next step was to used more data relative to inflation which means observable data impacted in different ways the evolution of the CPI. We chose the following macroeconomic data: interest rate, exchange rate (real exchange rate of the dollar), money supply, unemployment rate and wage. Considering the new data, we performed – as in the first part – a multivariate time series model: Vector Autoregression model (VAR) and a neural network (more precisely a feed forward network). Similarly to the first part, the neural network gave very successful outcome in term of inflation prediction. Unlike the second part, where the time series turned out to be significantly less efficient than the machine learning one, the VAR leads to extremely interesting results in term of CPI forecasting with very small relative error. This illustrates how important it is to use a wide range of data which impact the variable of interest which is here the CPI.

The third part of this project can be considered as more "exotic" as we chose to use a completely different type of data. Indeed, until here, we had only focused on economic data and their impact on inflation. However, the last 20 years have been the scene of the emergence of new tools through the use of the internet and especially Google search queries. Considering this, we decided to study a correlation between google trends and inflation since 2004 (the oldest data available). We selected some words which might have an impact on inflation in a positive or negative way (for instance car, unemployment, holidays). The idea was to use an Optimal Least Square regression between the volume of research of those words and inflation. The result were interesting (better than the first time series model but way less significant than the one obtained through Machine learning). However, this last model is a very promising one. It might be well improved using more data (not enough data are available at the moment, only from 2004) and a model different than an Optimal Least Square: stronger and less sensitive to fluctuations.

This project has been a success in term of inflation forecasting, machine learning recurrent neural networks as well as the VAR methods provided excellent results for both data sets used with a small relative error (one per cent for the worst case). Indeed, inflation is a volatile macroeconomic factor which is correlated with may others which makes it hard to predict and to anticipate. Even if the google trends method did not give perfect results, this idea has proven to be extremely promising and its improvement could give very satisfactory results.

References

[AF09] Marlene Amstad and Andreas M Fischer. "Are weekly inflation forecasts informative?" In: *Oxford Bulletin of Economics and Statistics* 71.2 (2009), pp. 237–252.

[Lee+96] Eric M Leeper et al. "What does monetary policy do?" In: *Brookings papers on economic activity* 1996.2 (1996), pp. 1–78.

[Swo69] Alexander Swoboda. "Vehicle currencies and the foreign exchange market: the case of the dollar". In: *The international market for foreign exchange* (1969), pp. 30–40.

[Tay13] John B Taylor. "Monetary policy during the past 30 years with lessons for the next 30 years". In: *Cato J.* 33 (2013), p. 333.

[Yil+10] Erhan Yildirim et al. "Yeni Keynesyen Makro Ekonomik Bir Model: Türkiye Uygulamasi1/A New Keynesian Macro Economic Model: Application to Turkish Economy". In: *Ege Akademik Bakis* 10.4 (2010), p. 1269.

Appendices

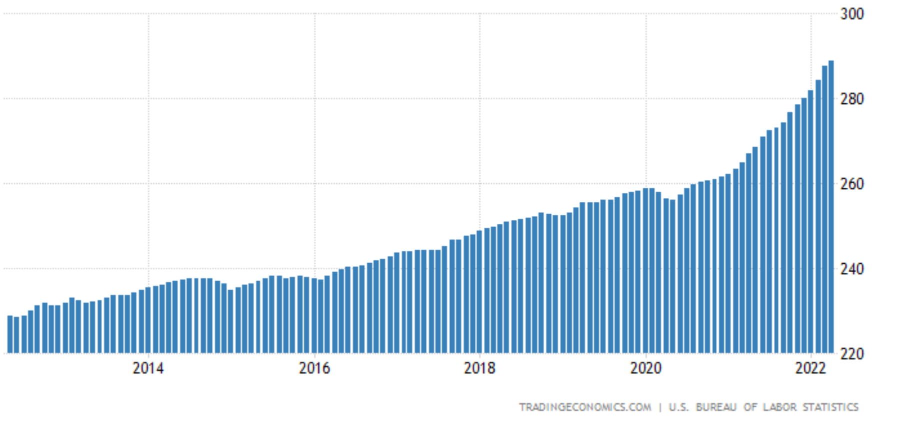


Figure 35: Consumer Price Index in the US

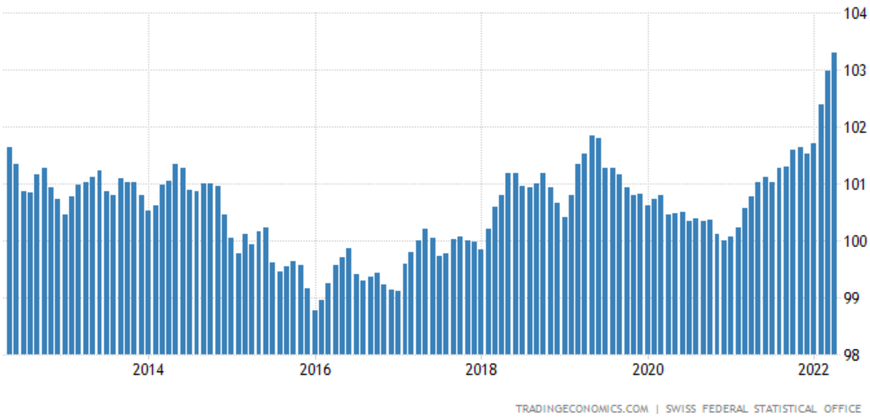


Figure 36: Consumer Price Index in Switzerland

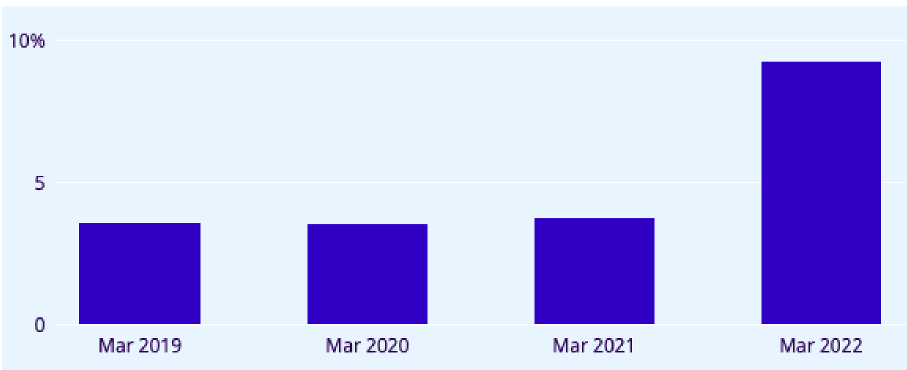
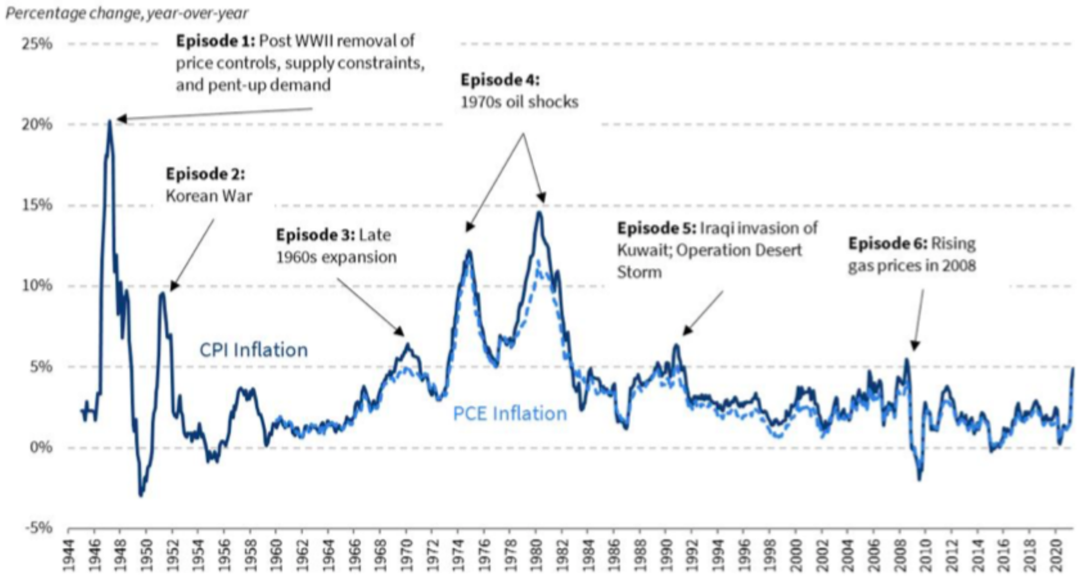


Figure 37: Evolution of the world inflation over the past 4 years



Source: Federal Reserve Economic Data (FRED), Haver Analytics, CEA Calculations.

Figure 38: Six episodes of post-WWII inflation

TARGETING INFLATION

Countries across the world have adopted inflation targeting irrespective of their income level.

COUNTRY	INFLATION TARGETING ADOPTION DATE	TARGET INFLATION RATE AT TIME OF ADOPTION	COUNTRY	INFLATION TARGETING ADOPTION DATE	TARGET INFLATION RATE AT TIME OF ADOPTION
New Zealand	1990	1 – 3	Philippines	2002	4 +/– 1
Canada	1991	2 +/– 1	Guatemala	2005	5 +/– 1
United Kingdom	1992	2 (point target)	Indonesia	2005	5 +/– 1
Australia	1993	2 – 3	Romania	2005	3 +/– 1
Sweden	1993	2 (point target)	Serbia, Republic of	2006	4 – 8
Czech Republic	1997	3 +/– 1	Turkey	2006	5.5 +/–2
Israel	1997	2 +/– 1	Armenia	2006	4.5 +/–1.5
Poland	1998	2.5 +/– 1	Ghana	2007	8.5 +/–2
Brazil	1999	4.5 +/– 2	Uruguay[1]	2007	3–7
Chile	1999	3 +/– 1	Albania	2009	3 +/– 1
Colombia	1999	2 – 4	Georgia	2009	3
South Africa	2000	3 – 6	Paraguay	2011	4.5
Thailand	2000	0.5 – 3	Uganda	2011	5
Hungary	2001	3 +/– 1	Dominican Republic	2012	3–5
Mexico	2001	3 +/– 1	Japan	2013	2
Iceland	2001	2.5 +/– 1.5	Moldova	2013	3.5–6.5
Korea, Republic of	2001	3 +/– 1	India	2015	2–6
Norway	2001	2.5 +/– 1	Kazakhstan	2015	4
Peru	2002	2 +/– 1	Russia	2015	4

Sources: Hammond 2011; Roger 2010; and IMF staff calculations.
Note: Countries are classified as inflation targeters based on the IMF's Annual Report on Exchange Arrangements and Exchange Restrictions (AREAER) database.
[1]Adoption date is based on the starting point when the interest rate became the main monetary policy instrument.

Figure 39: Adoption of inflation targeting across the world

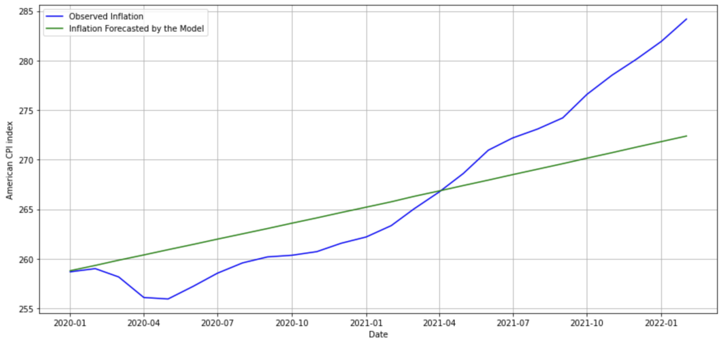


Figure 40: 2 years and 2 months long-term predictions of inflation with RNN model based on CPI index growth

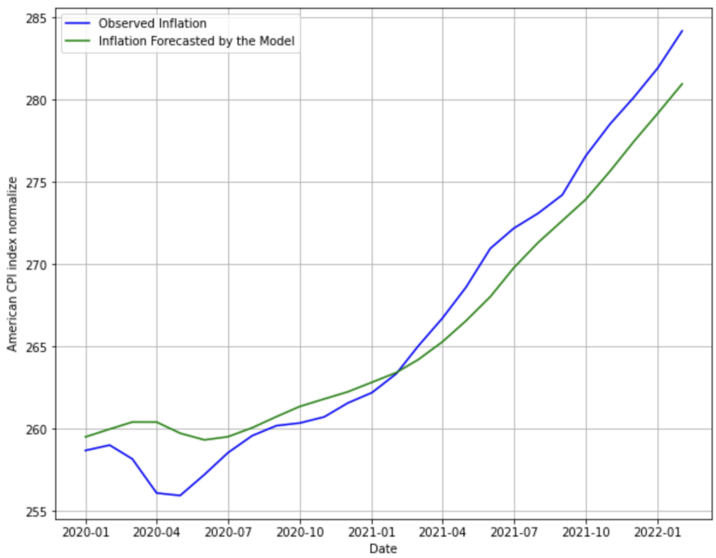


Figure 41: Monthly Forecasting from January 2020 to February 2022 with RNN model based on CPI index