Inflation Forecasting with T5

Joanne Im

May 2021

Abstract

The U.S. Bureau of Labor Statistics publishes monthly inflation statistics with an accompanying text summary that puts recent dynamics in context. This paper quantifies the marginal gain to inflation forecasting provided by this context. Text is interpreted using the state of the art T5 natural language processing model.

1 Introduction

1. When building a forecasting model, one first selects the data.

2. A human typically encounters information in numerical or textual format. The former is often through statistical release. The latter can be through news releases by the statistic agency that describe notable features of the release, or by third party analysts who put statistical releases in context using human reasoning and judgement.

3. The paper's asks: what is the marginal contribution of human context to statistics to inflation forecasts?

   To build up the machinery, we will ask this question using the context provided by the easy to retrieve, and standardized news releases published by the Bureau of Labor Statistics (BLS) when it publishes inflation statistics. These news releases provide context about what the agency considered notable and thus reportable changes, or stability in the index, and its underlying components; it neither makes explicit judgements about the data, nor forecasts.

   Notes:
   (1) I am not particularly interested in this use case, but think it will be a helpful base case for building up and applying the machinery.
   (2) Existing paper have attempted to leverage NLP models for use in forecasting financial time series data. T5 is both new, and a significant advance and to date. Its use seems largely restricted to the machine learning community. Its sister model BERT has appeared in a couple finance papers last year.

2 Background

T5 falls into the class of Transformer models, which unlike sequence-aligned models such as RNNs, do not process data in an ordered sequence manner. Instead, they process the entire sequence of data at once, using self-attention mechanisms to learn dependencies in the sequences. Thus, Transformer-based models have the potential to model complex dynamics of time series data that are challenging for sequence models.

On October 2019, Google published its Text-to-Text Transformer model, also known as T5, one of the most advanced natural language models to date. The text-to-text nature is possible because T5 contains an encoder and encoder-decoder block, distinguishing it from the popular Transformer model BERT, which only has an encoder block.

3 Data

In the United States, the Bureau of Labor Statistics (BLS) releases price indices, and an accompanying news release. We will focus on the price index series "Consumer Price Index for All Urban Consumers: All Items in U.S. City Average, NSA" because its ubiquity in policy discussions. We download the series data via FRED, and the text and pdf news releases from the BLS website: https://www.bls.gov/bls/news-release/cpi.htm. Our balanced dataset contains 326 monthly observations spanning February, 1994 to March, 2021.

We split the data into a training, and testing dataset. The training data set contains observations between February, 1994 to December, 2006, inclusive; the testing data set contains observations between December, 2006 to March, 2021. We chose this data split because we are interested in the ability of the model to forecast inflation well during episodes of high volatility in the series such as the Great Financial Crisis, and the Covid-19 Crisis.

4 T5 formatted dataset

The strength of T5 is that it is a text-to-text model, allowing it to encompass virtually all machine learning tasks. In our setting, the input text will be the prefix/query "What will be the year-over-year change in the Consumer Price Index for All Urban Consumers this month?" appended to text that gives last month's year-over-year change, and last month's BLS news release.

For example, the input text for March, 1994 looked like:

> Last month, the year-over-year change in the Consumer Price Index for All Urban Consumers was 2.52% CONSUMER PRICE INDEX–MARCH 1994 The Consumer Price Index for All Urban Consumers (CPI-U) rose 0.3 percent before seasonal adjustment in March to a level of 147.2 (1982-84=100), the Bureau of Labor Statistics of the U.S. Department of Labor reported today. For the 12-

> month period ended in March, the CPI-U increased 2.5 percent. The Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W) also increased 0.3 percent in March, prior to seasonal adjustment. The March 1994 CPI-W level of 144.4 was 2.3 percent higher than the index in March 1993. On a seasonally adjusted basis, the CPI-U rose 0.3 percent in March, the same as in February. The food index increased 0.1 percent in March, following declines in each of the preceding 2 months. The index for fruits and veg...

The output text–the "ground truth", or label–is the true year-over-year change in the consumer price index realized that month. For example, the output text for March, 1994 was

> 2.51

in string form.

5 Results

As a first pass, we are interested to see if T5 can forecast at least as well as an AR(1) model.

We plot (on the last page) realized inflation with the inflation forecasts of three models:

- T5: Base – the T5 model fed last month's inflation print
- T5: Base + BLS Text – the T5 modell fed last month's inflation print and the BLS text
- AR(1) – an OLS fitted AR(1) model

$$\pi_t = \alpha + \beta \pi_{t-1} + \epsilon_t$$

The mean squared error, average prediction error, and absolute error are reported below:

Model	Mean squared error (%)	Mean error (%)	Maximum error (%)
T5: Base	0.62	0.56	2.63
T5: Base + BLS Text	0.45	0.42	2.58
AR(1)	0.16	0.28	1.73

The forecasting performance of the T5 models are strictly dominated by those of an AR(1), but of the same order of magnitude. The T5 was only fed a news release about the prior month's statistical release. That is, we did not give it any information about the inflation process, or about what components of the release were relevant or not to forecasting.

We suspect that T5 was unable to replicate the AR(1) because of differences in the loss function and technical difficulties finding the optimum (that is, our parameters for optimization are not the best and warrant more experimentation). Given an AR(1) model, the data, and OLS assumptions, the optimal estimates can be computed using an analytic formula. Training a T5 model, however, requires us to deploy numerical optimization and uses cross entropy as its loss function, whereas OLS uses mean squared error. [1]

We also note that optimization is truly an art. For example, the BLS augmentation was done with the first 400 words of the BLS text. We found that increasing the text input led to the model doing worse out of sample. We also found that increasing epochs led to a marked drop in out of sample testing performance, suggesting that overfitting may be a large problem.

6 Plots

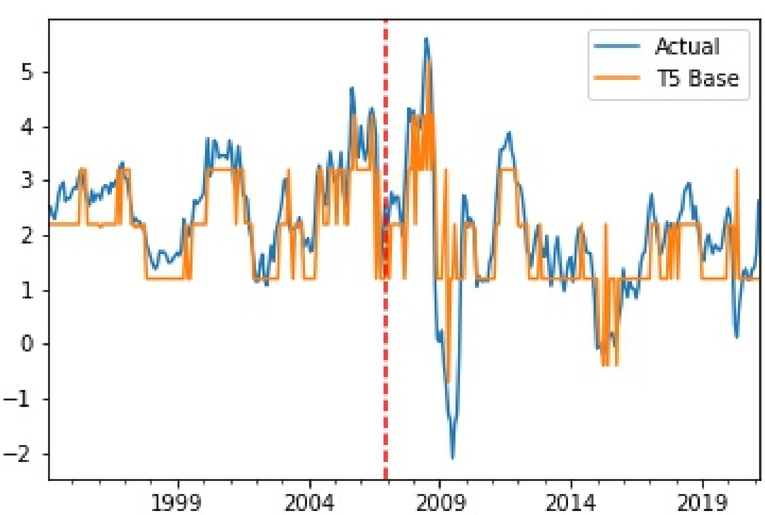

Figure 1: T5 Base vs. Actual
2

[1] Cross Entropy Loss: This is the basic loss used to compare the generated tokens with the reference tokens. Hence, it could be seen as a classification problem with number of classes being equal to vocabulary size; https://www.aclweb.org/anthology/2020.coling-main.1.pdf

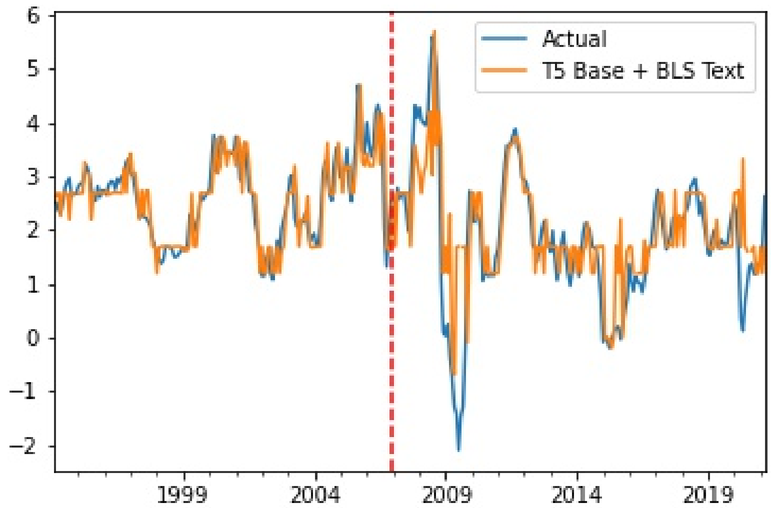

Figure 2: T5 Base + BLS Text vs. Actual

3

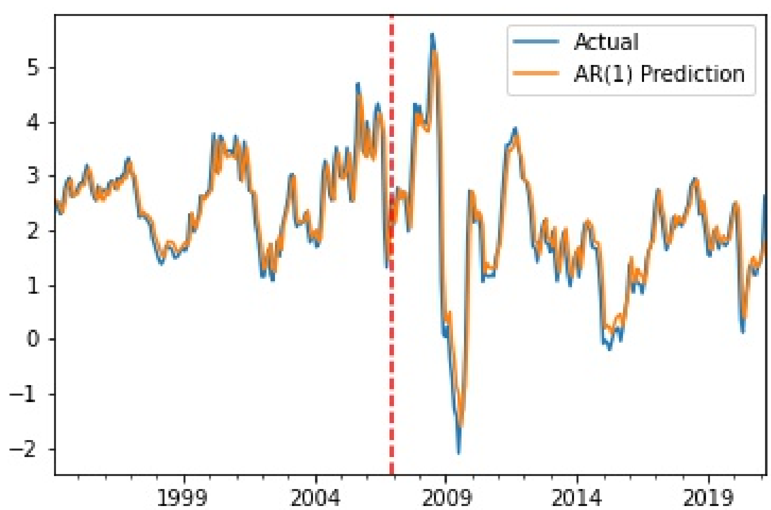

Figure 3: AR(1) vs. Actual

4