Inflation Forecasting in Emerging Markets

A Machine Learning Approach

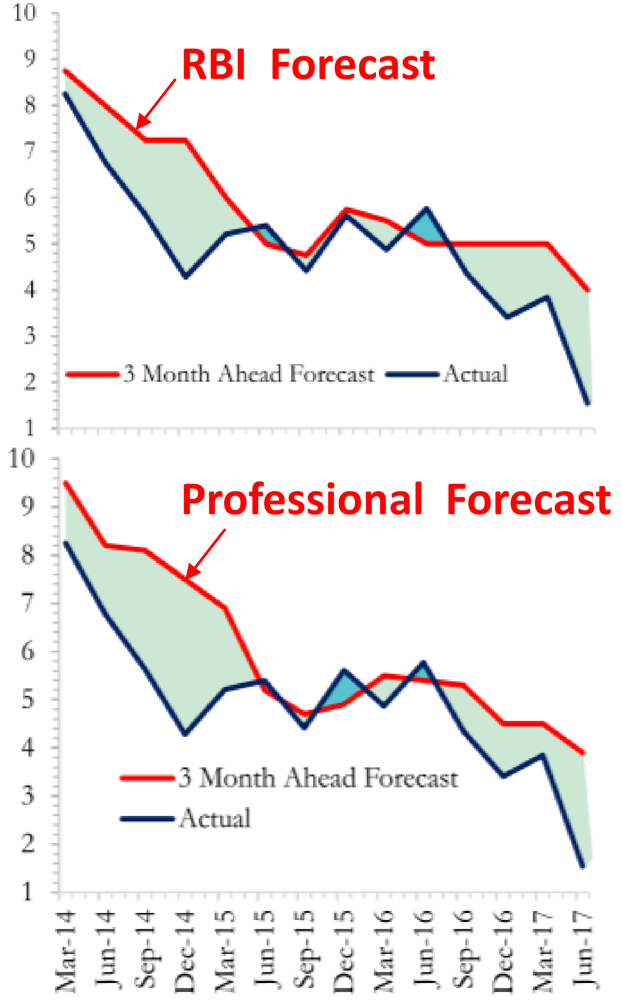
Kriti Mahajan & Anand Srinivasan

Motivation

Debate Over Accuracy of Inflation Forecasts

- “ Inflation has been overestimated by more than 100 basis points in six quarters… with an average error of 180 basis points” (Economic Survey of India, 2016-17)
- On the other hand, Raj et al (Mint Street Memo , 2019) argue that the large forecast errors were attributable to large unanticipated food price shocks, and that countries with high share of food prices in their CPI baskets tended to have higher forecast errors.

Figure 1: CPI Inflation Forecast vs Actual

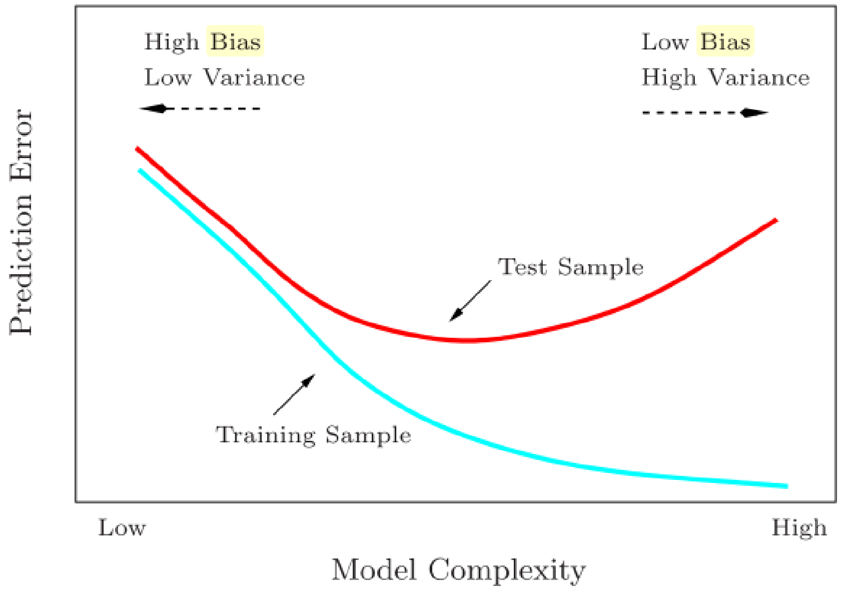


Source: Economic Survey of India, 2016-17

Why Machine Learning?

Model	Key Features vis-à-vis Forecasting
Traditional	• Based on BLUE and thus due to the bias variance trade off "do not yield the most accurate predictions"(Chalfin et al. 2016).
Machine Learning	• "Designed for prediction… use the data to …trade off bias and variance to maximize out-of-sample prediction accuracy." (Chalfin et al. 2016) • No assumptions regarding (Smalter Hall and Cook, 2017): a) the underlying data generating process b) the underlying relationship with the independent variables • Models non-linearities automatically • Improves forecast accuracy in a limited data environment (Jung, Patnam and Ter-Martirosyan, 2018).

Figure 2 : Bias Variance Trade Off



Source: Friedman, Hastie and Tibshirani (2001)

Expected prediction error = irreducible error + bias + variance

What We Do

- This paper uses a set of machine learning methods to explore if they can offer any improvements in forecast accuracy of CPI inflation in emerging economies, in particular India, China and South Africa.
- We compare the RMSE and MAE of each machine learning method against 2 benchmarks.
- We use three different classes of supervised machine learning methods:

Table 1 : Overview of Machine Learning Methods Used

Category	Method	Type	Customized for Time Series?
Shrinkage Model	Elastic Net	Linear	No
Tree Based Models	Random Forests	Non-Linear	No
	XG Boost	Non-Linear	No
Neural Networks	CNN	Non-Linear	Yes
	CNN+LSTM	Non-Linear	Yes
	Encoder Decoder	Non-Linear	Yes

Summary of Results

- Deep neural networks are the only class of machine learning methods to outperform the best performing benchmark in each horizon for each country.
- Notably, deep neural networks are able to forecast both the peaks and troughs in CPI inflation despite having been trained on small samples.
- Thus, there are gains to be made from adopting deep neural networks to inform policy decisions in India specifically and all emerging market economies generally.
- For CPI inflation in India, we find CPI and its sub-components, food, oil and bank related variables improve the forecast accuracy most significantly.

Overview of Deep Neural Nets (DNNs)

- In general, neural networks are composite functions which are universal function approximators.
- A neural network is a linear/non-linear transformation of the weighted linear combinations of the data.
- The structure of a neural network "follows the structure of a GLM model" but instead of using maximum likelihood estimation, it uses the feed forward mechanism and back propagation (a non-parametric algorithm) to determine the weights that result in the function for forecasting y_i (Smalter Hall and Cook, 2017).

Table 2 : Overview of DNNs

	CNN	CNN+LSTM	Encoder Decoder
Spatial	Yes	Yes	Yes
Sequential	No	Yes	Yes
Previous Predictions	No	No	Yes

Sample Construction and Variable Selection

- Broadly, the set of independent variables considered fall into the following categories : CPI/WPI and its Subcomponents, food, oil, monetary policy and finance and trade

Table 3: Sample Construction

Country	Full Period	Window Size	Training Obs. per Window	Validation Obs. per Window	Full Testing Period	No. of Independent Variables
India	1st Jan 03 – 1st Feb 19 (195 obs.)	153	138	14	1st Sep 16 – 1st Feb 19 (30 obs.)	48
China	1st Jan 03 – 1st July 19 (200 obs.)	158	142	15	1st Feb 17 – 1st July 19 (30 obs.)	30
South Africa	1st Jan 03 – 1st July 19 (200 obs.)	158	142	15	1st Feb 17 – 1st July 19 (30 obs.)	60

Results : Outline

- Forecast Accuracy For India
- Forecast Accuracy for Emerging Market Economies
- Neural Network Consensus Models
- Looking into the Black Box: Variable Importance

Forecast Accuracy (India)

Table 4: Comparison of RMSE and MAE Across All Models (India)

		1 Month		3 Months		6 Months		9 Month		12 Months	
Category	Method	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Best Benchmark	MA/Naive	**0.57**	**0.46**	1.33	1.13	1.81	1.64	1.75	1.54	1.46	1.16
Shrinkage Models	Elastic Net	0.62	0.51	2.42	2.10	1.52	1.19	2.21	1.79	2.61	2.31
Tree Based Models	Random Forests	1.41	1.24	2.21	1.93	1.74	1.39	3.18	2.85	3.00	2.80
	XG Boost	1.19	0.88	1.59	1.15	1.90	1.56	1.47	1.16	2.20	1.86
Neural Networks	CNN	0.96	0.77	0.80	0.61	**0.92**	**0.72**	1.03	0.83	0.89	0.69
	CNN+LSTM	0.77	0.59	**0.69**	**0.50**	1.07	0.84	1.11	0.79	**0.85**	**0.64**
	Encoder Decoder	0.72	0.57	0.85	0.68	1.07	0.87	**0.73**	**0.56**	1.09	0.85

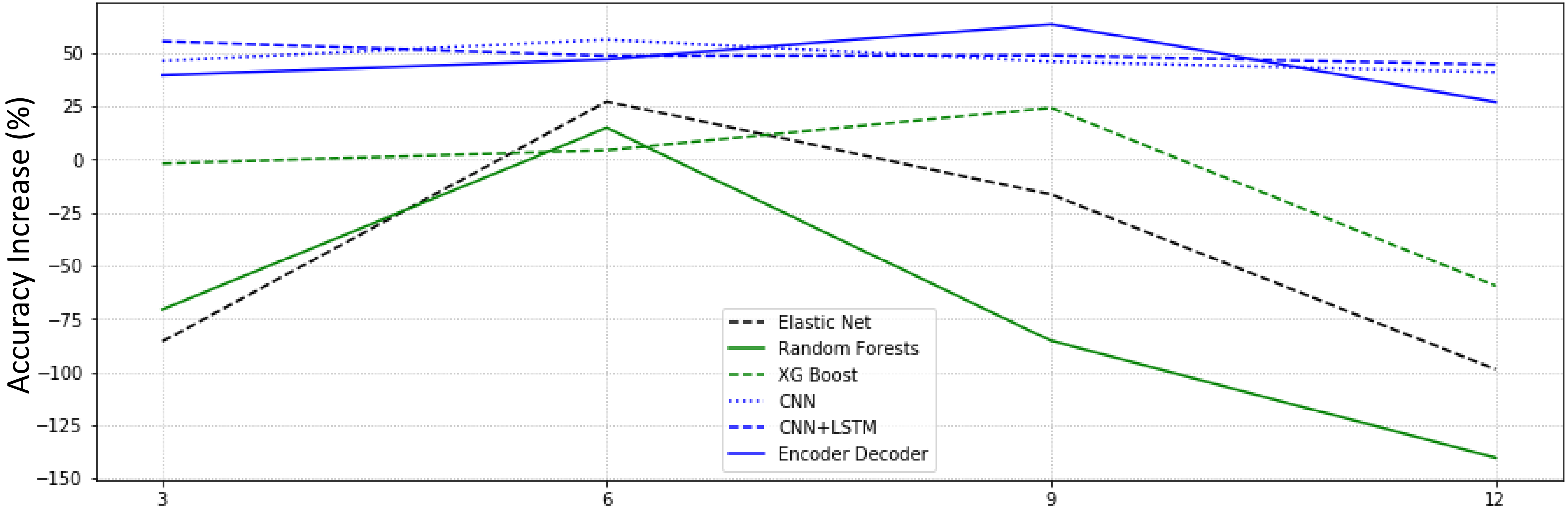
*Values in bold indicate the best performing model for a given horizon

Key Finding : Neural networks improve forecast accuracy at longer horizons.

- For the 3 month ahead forecast the reduction in forecasting error of the best neural net model (CNN+LSTM) relative to the best benchmark is **48.15% in terms of RMSE, and 55.54% if one considers MAE.**
- For the 12 month ahead forecast the reduction in forecasting error of the best neural net model (CNN+LSTM) relative to the best benchmark is **42.08% in terms of RMSE, and 44.59% if one considers MAE.**

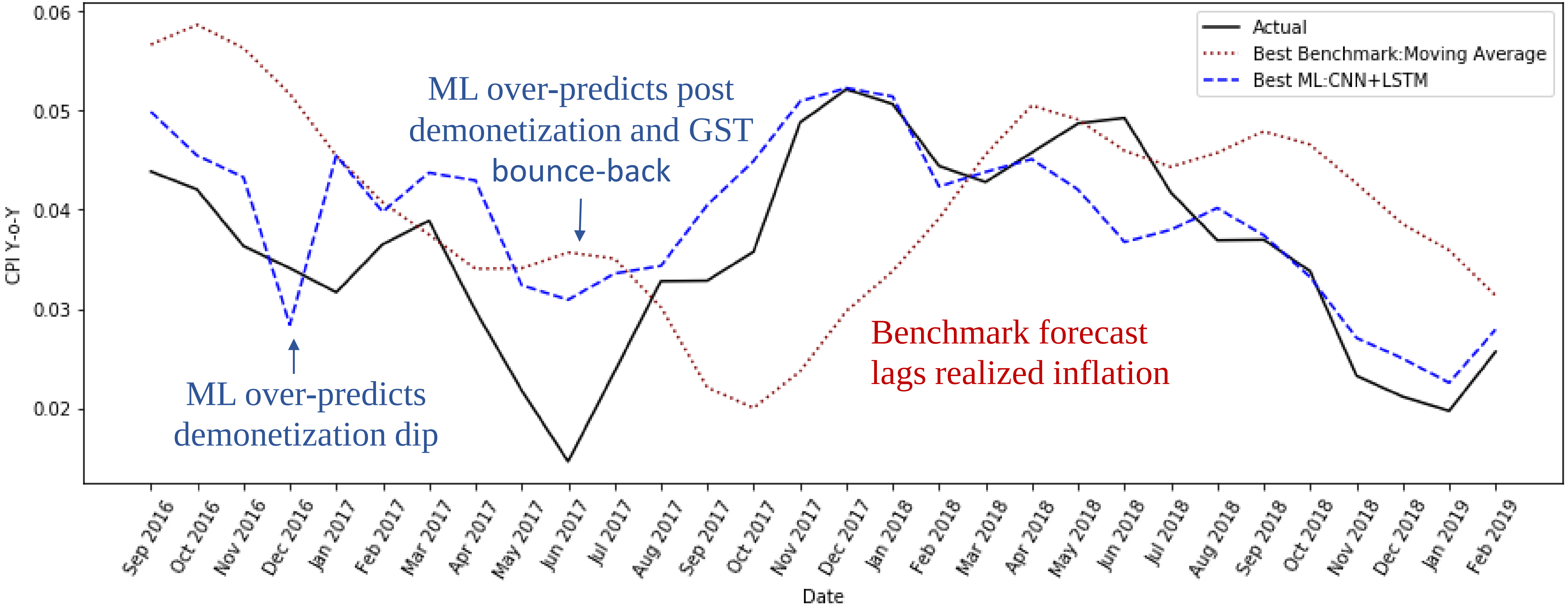
MAE Accuracy Increase Over Best Benchmark (India)

Figure 3: MAE Accuracy Increase (%)



India : 3 Months Ahead

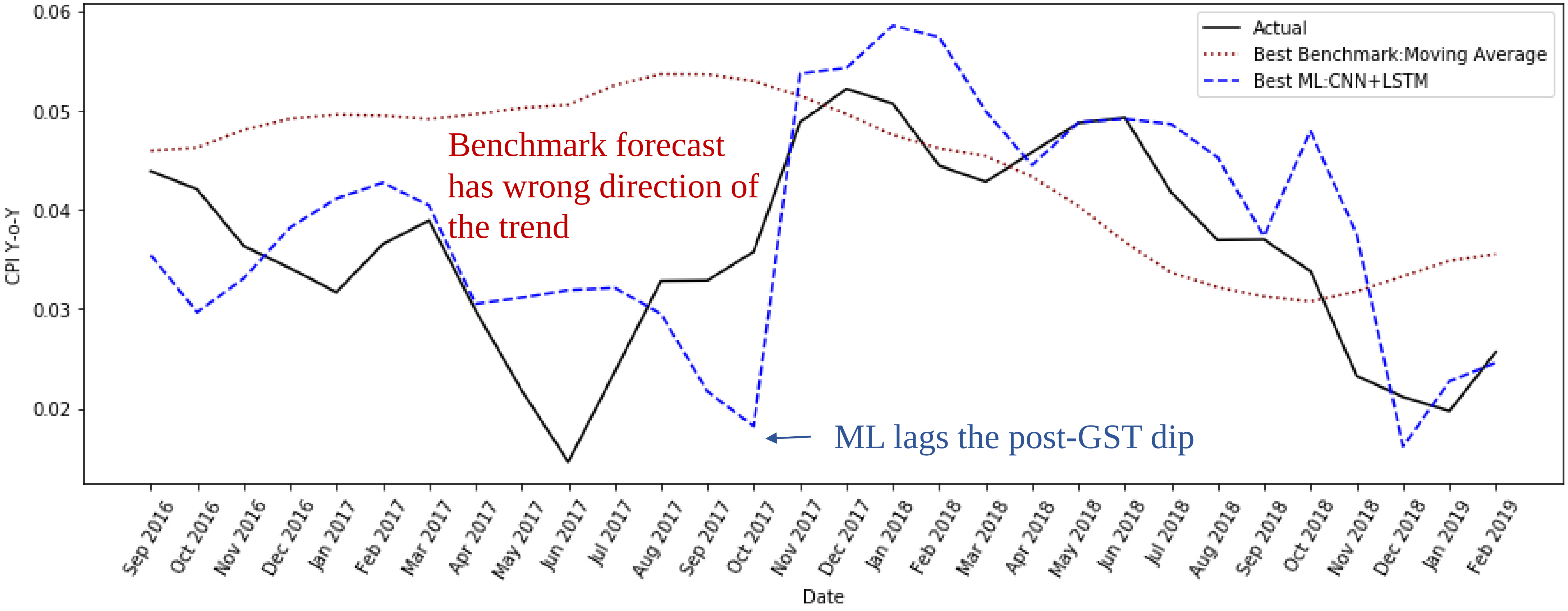
Figure 4: 3 Months Ahead Inflation Forecast (India)



From May 2018, the model tracks actual inflation well and predicts the turning points.

India : 12 Months Ahead

Figure 5: 12 Months Ahead Inflation Forecast (India)



Captures both the direction as well as magnitude of the actual inflation very well with the exception of post-GST dip.

Forecast Accuracy (Emerging Market Economies - 12 Months Ahead)

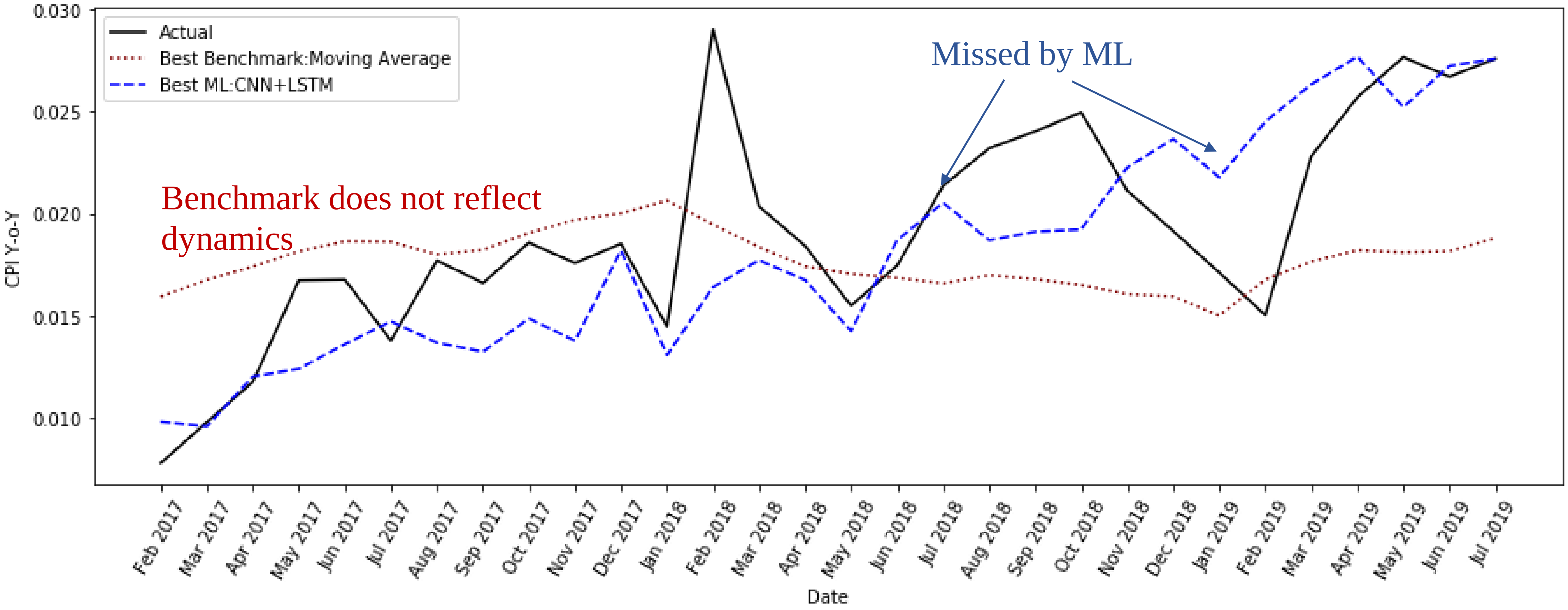
Table 5: Comparison of RMSE and MAE Across All Models (Emerging Market Economies)

Country	Category	Method	RMSE	MAE	RMSE Accuracy Increase (%)	MAE Accuracy Increase (%)
China	Best Benchmark	MA/Naive	0.540	0.447		
	Neural Networks	CNN	0.527	0.365	2.27	18.23
		CNN LSTM	**0.407**	**0.297**	**24.63**	**33.42**
		Encoder Decoder	0.526	0.420	2.49	5.85
South Africa	Best Benchmark	SARIMA	1.217	1.112		
	Neural Networks	CNN	0.635	0.518	47.85	53.400
		CNN LSTM	**0.590**	**0.469**	**51.52**	**57.849**
		Encoder Decoder	0.809	0.643	33.48	42.176

*Values in bold indicate the best performing model for a given horizon

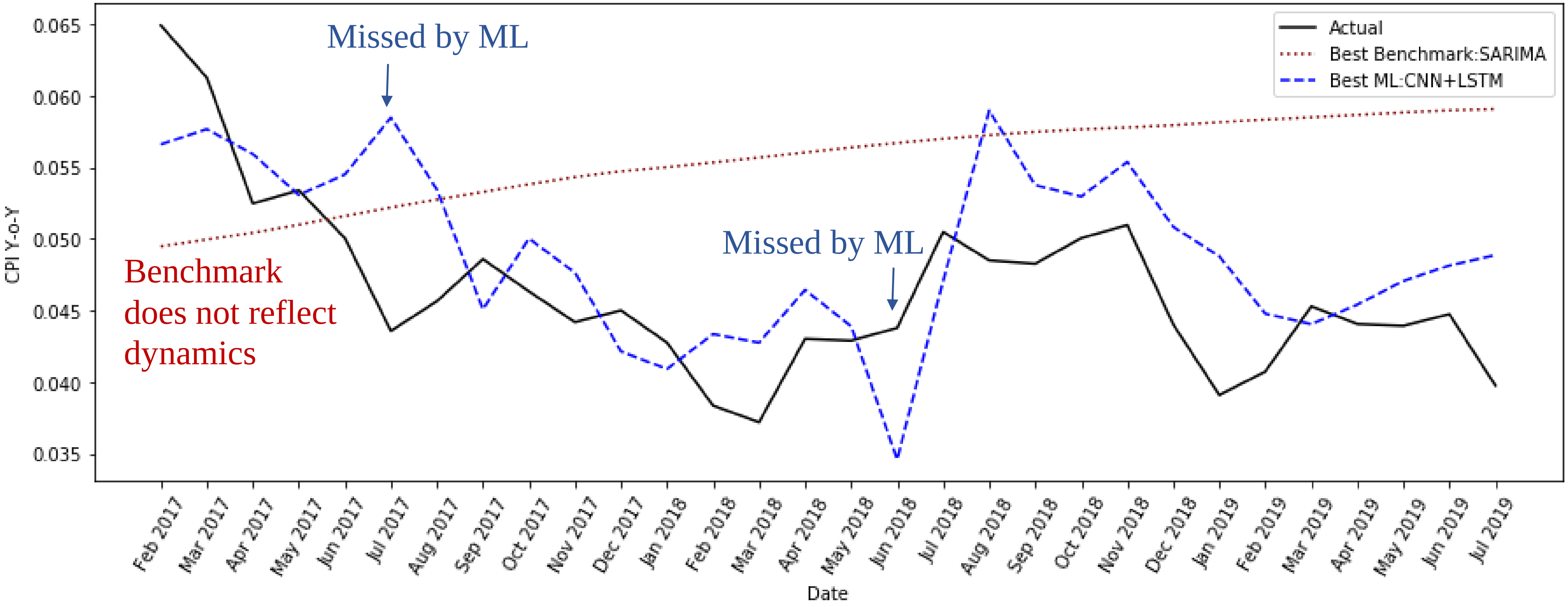
China: 12 Months Ahead Forecast

Figure 6: 12 Months Ahead Inflation Forecast (China)



South Africa :12 Months Ahead Forecast

Figure 6: 12 Months Ahead Inflation Forecast (South Africa)



Neural Network Consensus Model : Revisiting the Overestimation Debate

Table 6: Consensus Model – Forecast Accuracy and Bias

Country	Horizon	RMSE	MAE	RMSE Accuracy Increase (%)	MAE Accuracy Increase (%)	Over Estimate (%)
India	1 Month	0.613	0.517	-7.46	-11.35	56.67
	3 Month	0.604	0.458	54.55	59.52	50.00
	6 Month	0.877	0.725	51.45	55.66	36.67
	9 Month	0.699	0.582	60.17	62.13	56.67
	12 Month	0.710	0.559	51.35	51.93	53.33
China	12 Month	0.395	0.312	26.82	30.02	53.33
South Africa	12 Month	0.429	0.364	64.72	67.22	66.67

- In general, the RMSE of the consensus models is lower than that of individual neural networks for all horizons for all countries.
- The **overestimate (%) is close to 50%.**

Looking into the Black Box

Variable Importance (India)

- Variable importance describes “how much a prediction model’s accuracy depends on the information in each covariate”.

  - (Fisher, Rudin, and Dominici, 2018)

Rank	1 Month	3 Months	6 Months	9 Months	12 Months
1	CPI (1)	CPI Clothing (9)	Bank Rate (6)	Rainfall (12)	Commercial Bank Deposits (12)
2		CPI Clothing (7)	Wheat Stock (10)	Rainfall (9)	CPI-AL : Clothing (12)
3		WTI USD (4)	CPI-AL : Clothing (8)	Rainfall (11)	Rice Stock (12)
4		Bank Rate (4)	CPI-AL: Miscellaneous (9)	Rainfall (10)	RBI Total Liabilities (12)
5		CPI Fuel and Light (8)	Repo Rate (10)	Net FII and FPI (9)	Wheat Stock (12)
6		CRR (10)	Reverse Repo Rate (7)	Net FII and FPI (10)	Auto Production of Commercial Vehicles : 3 Wheelers (12)

*All variables are y-o-y

†Values in brackets give the lag.

Conclusion

- We conducted an analysis for three emerging markets – India, China and South Africa using a variety of machine learning methods.
- Out of the approaches used, deep neural networks were most effective in reducing forecast errors relative to SARIMA or moving average forecasts.
- Notably, deep neural networks are able to forecast both the peaks and troughs in CPI inflation despite having been trained on small samples.
- Thus, there are gains to be made from adopting deep neural networks to inform policy decisions in India specifically and all emerging market economies generally.
- Furthermore, preliminary results suggest that our best performing forecasts out perform:
  - the RBI 1 quarter ahead forecast by about 40%.
  - the Indian Professional Forecasters Survey at the 1, 2, 3 and 4 quarter ahead horizons by at least 25%.