AI for Time Series: Applications in Macroeconomic Forecasting

Bayes Business School

10th December 2025

Alexander Denev

AI/Machine Learning and Econometrics

The two fields use different tools and mindset, but they can benefit from each other

Econometrics	Machine Learning
Econometric models are based on economic theory, generally expressed through parametric models. Traditional statistical inference methods (such as maximum likelihood and the method of moments) are thus used to estimate the values of a vector of parameters θ, in a parametric model $m_\theta(.)$	In ML, by contrast, non-parametric models are often built based almost exclusively on data (i.e. no distribution hypothesis), very often Big Data

AI/Machine Learning and Econometrics

The two fields use different tools and mindset, but they can benefit from each other

Econometrics	Machine Learning
Econometrics focuses on studying the asymptotic properties of θ^* (viewed as a random variable, thanks to the underlying stochastic representation)	ML focuses to a greater extent on the properties of the optimal $m^*(\,)$ based on a criterion that has to be defined, or even simply $m^*(x_i)$ for observations i deemed to be of interest for example in a test population
Econometrics uses mostly in-sample goodness-of-fit measures e.g. AIC, BIC	In ML the meta-parameters used (tree-depth, penalty parameter, etc.) are optimised by cross-validation.

AI/Machine Learning and Econometrics

The two fields use different tools and mindset, but they can benefit from each other

Econometrics	Machine Learning
Econometric models are simple to interpret built with a significant amount of human input. Econometrics is a fundamentally a process of human learning. It is fundamentally a search for knowledge and *causal* understanding, and not just an ability to make better predictions.	ML focuses on prediction robustness while sometimes sacrificing transparency.

Econometrics helps us *explain*; ML helps us *predict*. Both perspectives are valuable.

A Big Data Approach to Forecasting

	How forecasting was done: Pre-Big Data age	How we do it: Big Data age
Models	• Approach based on descriptive models with **limited parameters** • Models contain many assumptions about the relationships between variables	• Models with **many more parameters** that contain very limited pre-conceived assumptions about the relationship between variables • Models exploit latest advances in **Machine Learning**
Data	• Data is mainly used to **estimate model parameters** • Limited data types used - only **few** exogenous and endogenous variables of similar nature	• Data is used to **estimate model parameters** and to **enhance predictive ability** (we use ~900 variables per country) ▪ **Significantly more data** used to fit models with **multiple data types**: numerical, text, images …

Forecasting is no longer only about choosing the “right” model—it’s about structuring and extracting signal from massive data.

Combining Alternative and Traditional Data

We use a variety of drivers which include traditional macroeconomic and alternative data variables, in total ~900 per country

Traditional Data Sources

Output, Production, Bus Activity
- Total Industrial Production
- Preliminary GDP
- Factory Orders, etc.

Retail
- Consumer Prices
- Personal Consumption Expenditure
- Retail Trade, etc.

Labor
- Wage Changes
- Unemployment
- Job Openings & Labor Turnover

Sentiment & Expectations
- Manufacturing Confidence
- Consumer Confidence
- Import & Export Forecasts , etc.

Money, Credit & Interest Rates
- Producer Prices
- Capital Account Statistics
- National Balance Sheets
- Money Supply
- Debt Statistics
- Government Balance Statistics
- CPI , etc.

Etc.

Alternative Data Sources

Output, Production & Bus Activity
- Air Pollutants & Air Quality & Rainfall
- Electricity Consumptions & Shortages
- Factory Capacity Utilisation
- Oil Rig Counts
- Natural Resources Stock levels (incl. injections & withdrawals), etc.

Retail
- Airport Arrivals and Departures
- Restaurant visits & reservations
- Ground, Sea & Air Traffic
- Entertainment ticket sales
- Overnight stays and expenditures
- Vehicle Registrations, etc.

Geopolitical
- Geopolitical Risk News
- World Trade Uncertainty, etc.

Labor
- Online Job Adverts
- Slack labor statistics, etc.

Sentiment & Expectations
- News Sentiment
  - Food
  - Energy
  - Communications
  - Entertainment
  - Housing, etc.

Feature engineering

ML algorithms are needed in alternative data world to extract features from a variety of unstructured data sources

- **Unstructured data** – usually requires ML algorithms to extract
- **Structured data** – constructing appropriate features can be useful too

Image processing (e.g., VLMs)

Text/sound processing (e.g., LLMs)

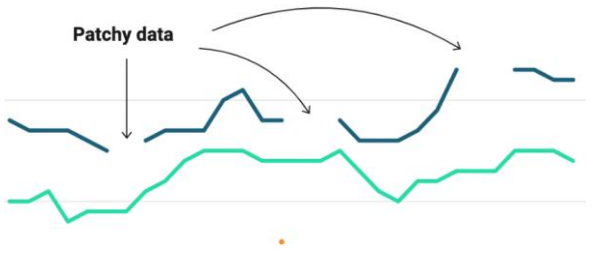
Geofencing for spatial datasets

Feature engineering is often more important than model choice

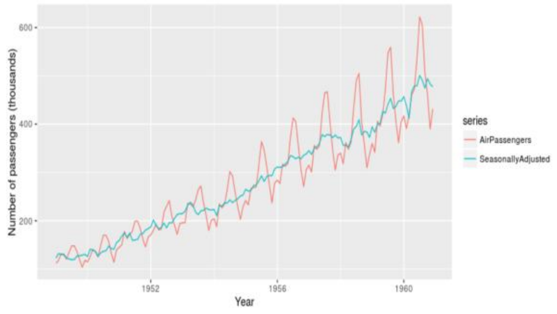
Algorithms - Data Transformations

Data needs to be curated to ensure data cleanliness, reliability and that it is ML ready

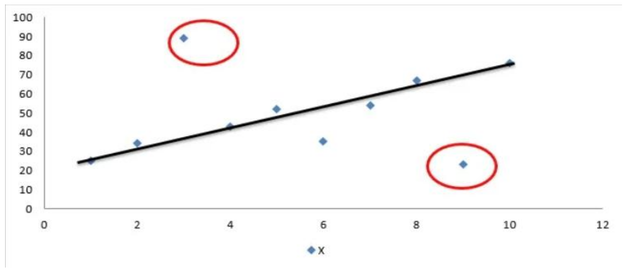
Missing Data (e.g., EM algorithm)



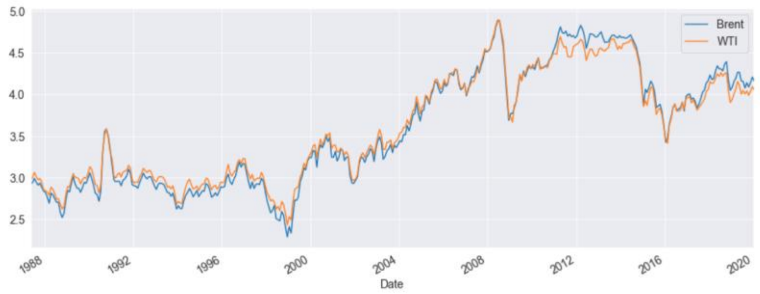
Making variables de-seasonalized (e.g., X-13)



Outlier removal (e.g., Mahalanobis distance)



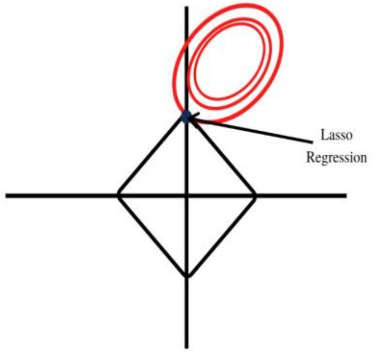
Search for cointegrating variables and stationarization



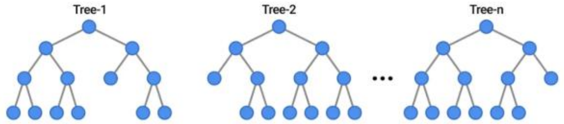
Algorithms – Forecasting

Novel algorithms are needed in alternative data world to deal with the variety of potential explanatory variables as well as their **(big) number** and **non-linearities**

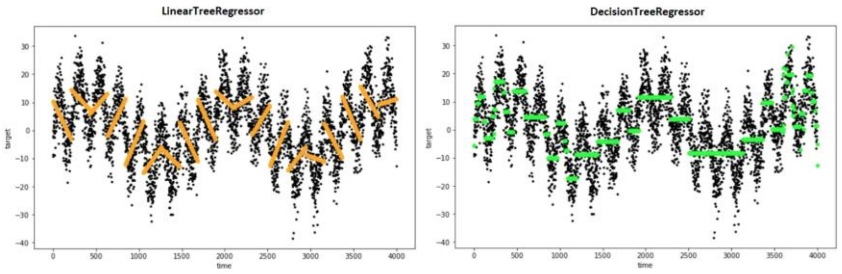
Linear (e.g., LASSO, Ridge, Elastic Nets)
able to extrapolate, bad with non-linearities



Random Forest (RF)
Unable to extrapolate, but good with non-linearities



Linear Random Forest (LRF)
Able to extrapolate and deal with non-linearities

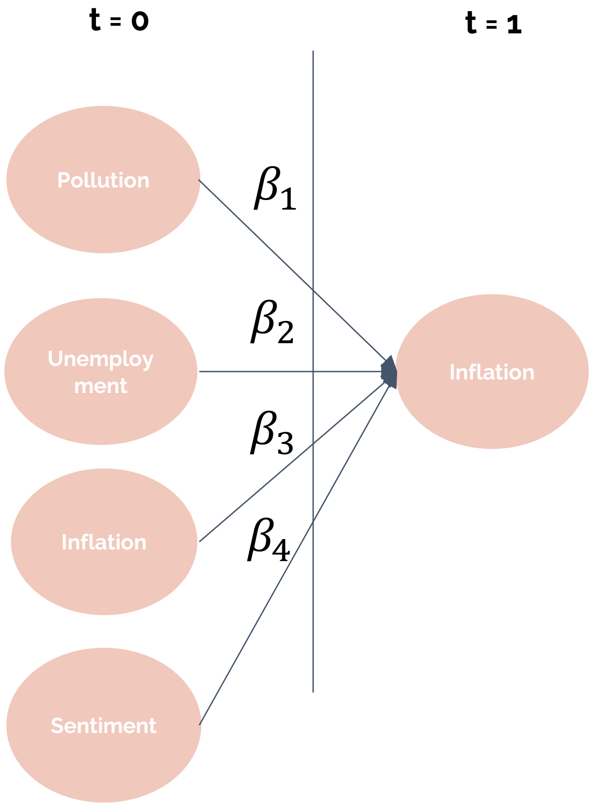


A Genetic Algorithm layer pre-filters the input variables (wrapper method)

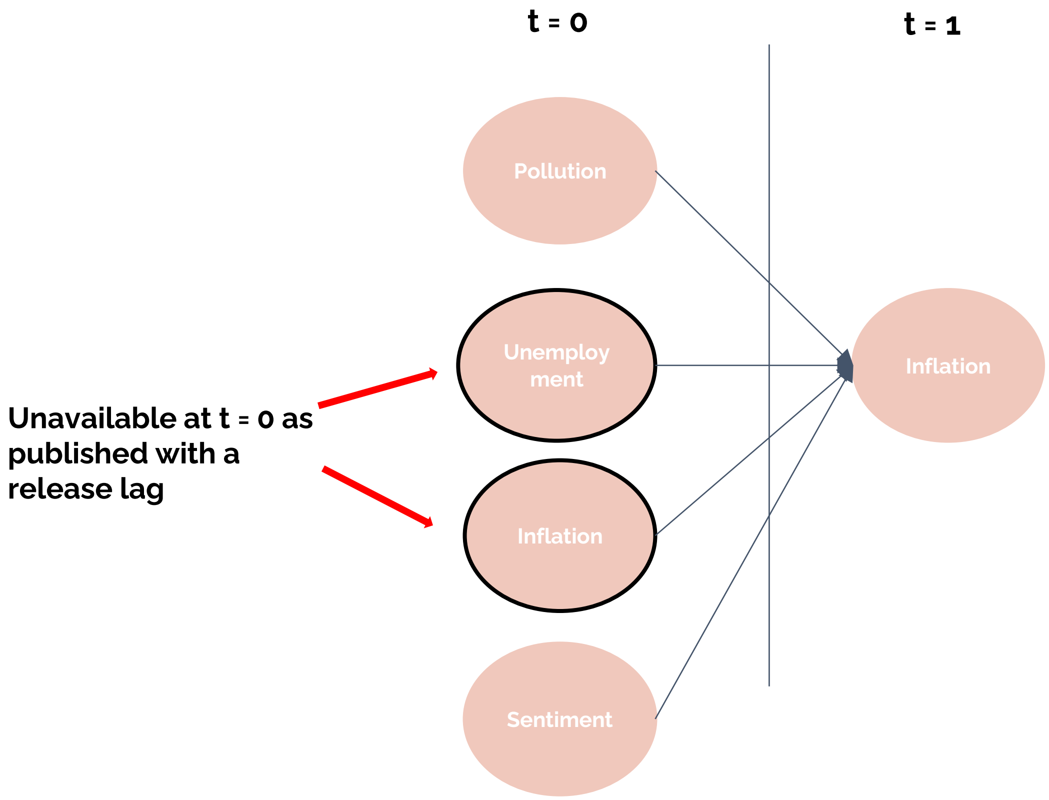
Forecasting & Bias

Since we aim to maximize predictive power sometimes by using a big number of variables, we can get biased coefficients that hide the true causal impact, but this does not impact predictive performance.

If this was the true causal model…



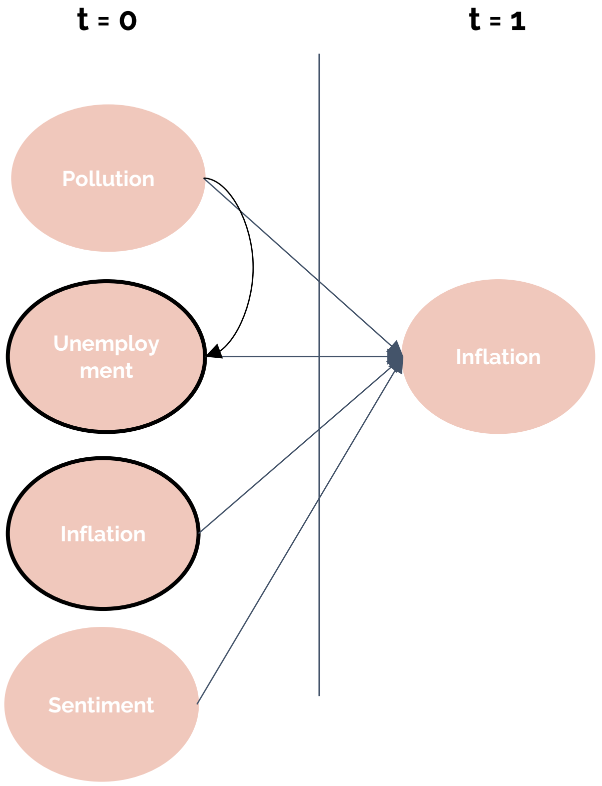
…it cannot be calibrated as many of the variables are non observed at time t=0



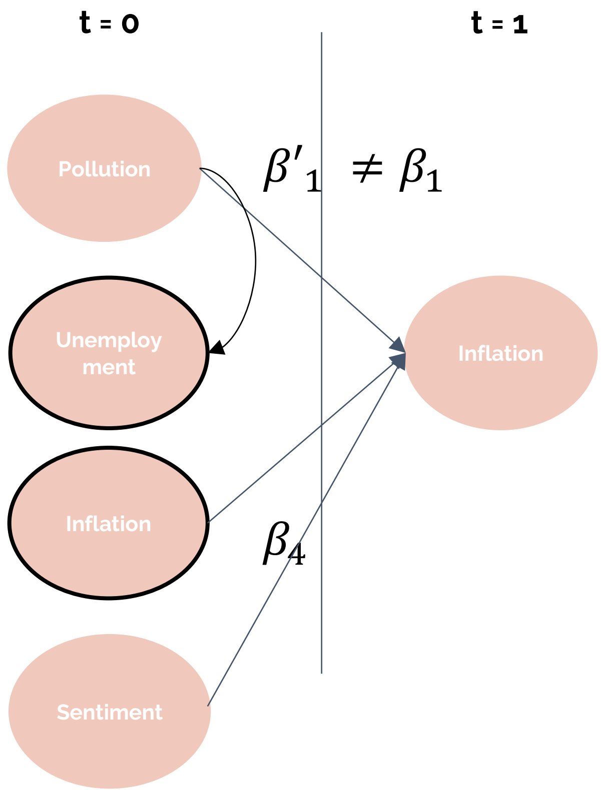
Forecasting & Bias

Since we aim to maximize predictive power sometimes by using a big number of variables, we can get biased sensitivities that hide the true causal impact, but this does not hinder predictive performance.

The independent variables can be also correlated (or better sometimes - causally related)…



…which biases the coefficients estimations in the case of omitted variables
which are now being proxied by other variables



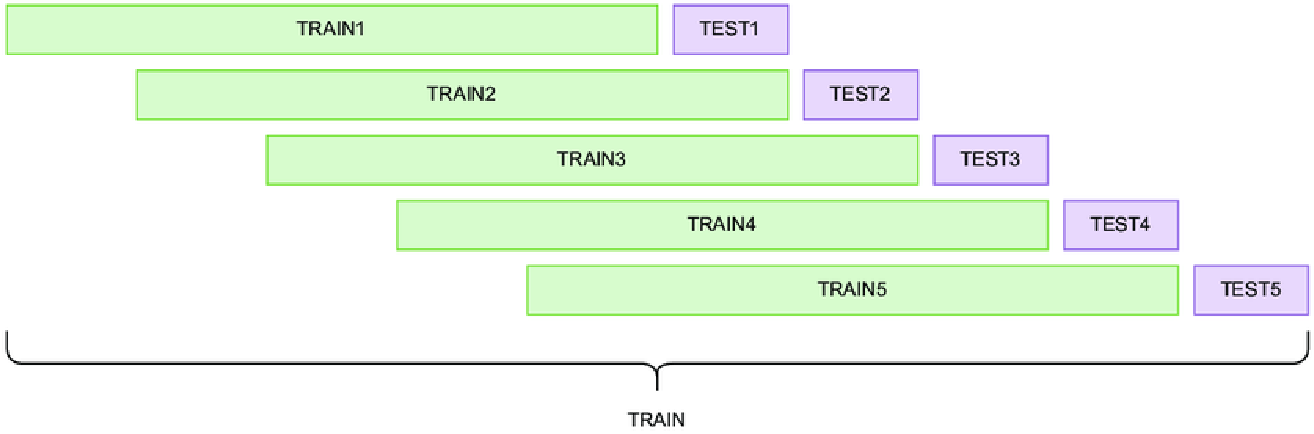
Causal vs Predictive Modeling

- Causal: Interpret coefficients; requires exogeneity.
- Predictive: Minimize error; bias in coefficients may be acceptable.

Training & Validation

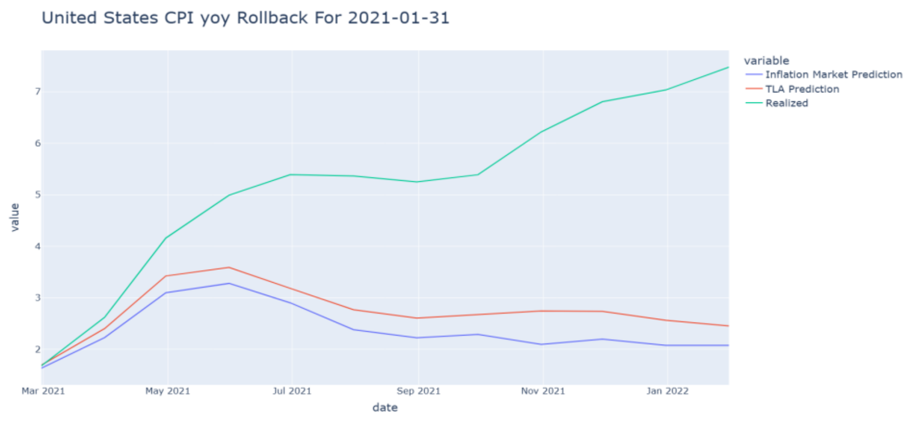
We do the training and validation through the rolling windows approach

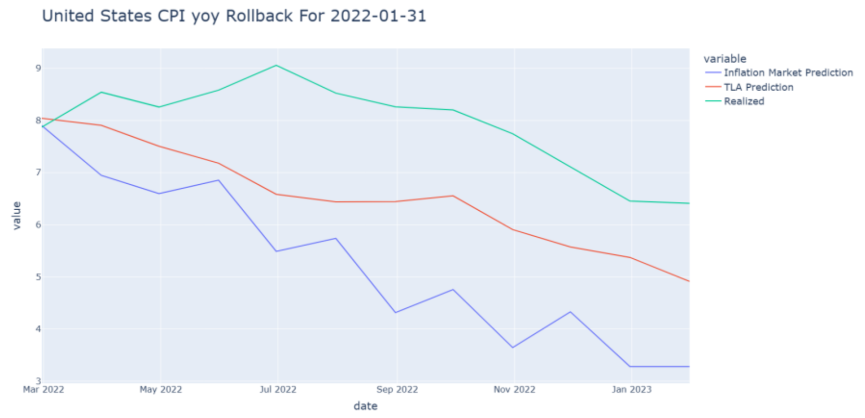
- All the macro and alternative data is snapped at each point in time, e.g. no future revisions are in the training sample at any given point in time
- All the data transformations are performed point-in-time in each training window e.g. deseasonalization, stationarization etc.

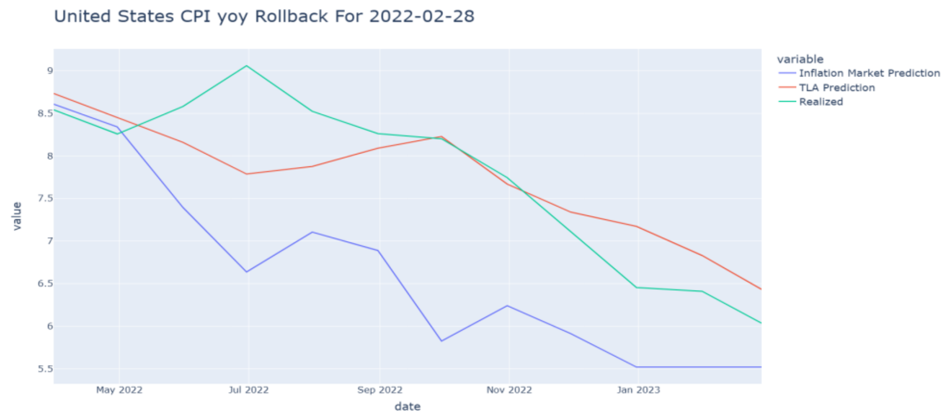


What has been predicted so far

~20,000 inflation forecasts published so far. Live performance for the US for some dates shown below







What has been predicted so far

Live performance since May 2022 (~20,000 forecasts published)

All Countries (36)													
	Tenor												
	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M	11M	12M	Total
Error Reduction	-0.413	-0.355	-0.333	-0.308	-0.265	-0.231	-0.197	-0.16	-0.129	-0.08	-0.017	0.09	-0.193
Directional Hit	0.765	0.754	0.776	0.778	0.762	0.782	0.783	0.794	0.804	0.798	0.801	0.804	0.782

EM (22 countries)													
	Tenor												
	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M	11M	12M	Total
Error Reduction	-0.357	-0.293	-0.265	-0.236	-0.188	-0.16	-0.122	-0.077	-0.043	-0.014	0.044	0.137	-0.121
Directional Hit	0.772	0.74	0.755	0.768	0.736	0.765	0.764	0.784	0.784	0.778	0.777	0.777	0.766

DM (14 countries)													
	Tenor												
	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M	11M	12M	Total
Error Reduction	-0.484	-0.441	-0.431	-0.415	-0.379	-0.337	-0.307	-0.283	-0.256	-0.181	-0.111	0.018	-0.299
Directional Hit	0.755	0.774	0.805	0.79	0.798	0.806	0.809	0.808	0.831	0.825	0.832	0.841	0.804

Metrics Introduced

- **Error Reduction:** MAE improvement over benchmark.
- **Directional Hit:** ability to predict ups/downs.

Interpretability

Lack of interpretability of the results can originate both from the use of non-interpretable models and the big number of variables being used

- If models are non-interpretable (e.g., Random Forests) interpretability methods can be used, such as:
  - SHAP
  - LIME
- Interpretable models (e.g., Lasso, EN)
  - Even if models are interpretable or made interpretable (e.g., through LIME) the big number of variables can still make interpretation obscure.
  - Word clouds facilitate understanding when the variables are too many

What a model can miss? Policy adjustments

There can be policies that can impact inflation, which may not always be reflected in the data we ingest. LLMs can be used to scan the news to help identify these. We adjust our forecasts using these estimates. This overlay has helped to improve our forecasts over the past year.

United Kingdom CPI Impact Report

August 2025–July 2026

Headline CPI Impact

Month–Year	Policy	Weight (%)	Pass-through (%)	Shock (%)	Impact (%MoM)
Sep-2025	ScotRail Peak Fare Removal[1]	0.96×10%	100	−15.0	−0.014
Oct-2025	University Tuition Fee Rise[2]	0.6192	100	+2.7	+0.017
Oct-2025	Ofgem Energy Cap (Oct '25)[3]	2.7	100	+2.0	+0.054
Jan-2026	Ofgem Energy Cap (Jan '26)[4]	2.7	100	−0.5	−0.014
Jan-2026	£3 Bus Fare Cap (Extended)[5]	0.30	100	0.0	0.00
Mar-2026	Regulated Rail Fare Rise[6]	0.96	100	+5.8	+0.056
Apr-2026	Fuel Duty Freeze Ends[7]	3.08	90	+4.4	+0.12
Apr-2026	Ofgem Energy Cap (Apr '26)[8]	2.7	100	+6.0	+0.16
Apr-2026	Air Passenger Duty Increase[9]	0.9	100	+3.0	+0.03
Apr-2026	Water Bills Rise[10]	1.2	100	+6.0	+0.07

Notes:

(1) ScotRail (Scotland) scrapped peak-time fares from 1 Sep 2025; for example, the Glasgow–Edinburgh return fare fell from £32.60 to £16.80 (about −48%).

(2) University tuition fees increase from £9,250 to £9,500 (+2.7%) in Oct 2025.

(3) Ofgem cap further reduction by 1.3% in Oct 2025.

(4) Ofgem energy price cap is projected to fall by 0.5% in Jan 2026.

(5) £3 bus fare cap extended to Mar-2027; previously anticipated fare increase removed.

(6) Regulated train fares in England are expected to rise by 5.8% in March 2026 (assuming RPI+1% formula).

(7) Fuel duty cut of 5p per litre ends March 2026, implying 4.4% price rise at pumps.

(8) Ofgem price cap is forecast to increase by 6% in April 2026 as wholesale energy costs rebound.

(9) APD increases by approximately 3% in April 2026.

(10) Water bills expected to rise by 6% in April 2026 due to infrastructure investments.

(11) Average Band D council tax bills in England will rise by 5.0% in April 2026.

Combining ML with expert knowledge improves real-time robustness

Challenges

Practical implementations are fraught with challenges that need to be overcome with the right processes' setup

- Complexity of managing diverse data sources
  - Data onboarding
  - Legal agreements
  - Data vendors risk management
  - Large software/technology stack (Apache Airflow, S3)
  - Internal processes setup
- Data variability
  - Accuracy of the data
  - Technical outliers
  - Interruption of service
  - Rebasing of data

MARCOS Architecture — Data → Models → Forecasts

Data Science

Data Ingestion
- Bespoke scraper library/LLMs for big data ingestion
- Automatic construction of indices tracking e.g. weather, food prices, sentiment
- Ingesting data of external data providers via API

Data Quality Assurance
- Data Normalization
- Time-Series Consistency Checks
- Missing Data Handling
- Correlation and Integration Layer

Point-In-Time Feature Processing
- Outlier treatment
- Time-Frequency Harmonization
- Decomposing the seasonal and trend dynamics
- Memory preserving stationarization & scaling

Features Selection
- Correlation and network-cluster screening of variables
- Scalable evolutionary algorithm search for the best variable combinations
- Mixed frequency methods for optimal handling of data on different time scales
- Insert domain knowledge with a human in the loop

Advanced Modeling
- Large suite of machine learning models
  - Tree-based methods
  - Penalized linear methods
  - Neural - Networks
- Bespoke combination of linear and non-linear models
- Unique methodology to forecast non-stationary time series

Comprehensive Results
- Detailed forecast driver breakdown
- Conformal forecast uncertainty bands
- Comprehensive overview of performance metrics and model diagnostics

Forecast Validation
- Comprehensive automated checks of model performance
- Automated review and verification of model outputs

Engineering

Data Ingestion	Data Quality Assurance	Point-In-Time Feature Processing	Features Selection / Advanced Modeling	Comprehensive Results	Forecast Validation
Automated workflow orchestration using Apache Airflow Extensive scraper framework capable of extracting data from web, image, and text data Efficient and scalable processing using Polars and Ray		Scalable feature processing library Creates cloud-based point-in-time feature stores Built on top of highly optimized in-house statistical library	Using a parallel cloud compute cluster	Forecasts delivered via a web dashboard, an API, as well as from AWS S3	

Data	Modelling	Forecasts

THANK YOU!