TAP 3

Zane Hassoun

University of York

April 15, 2024

Overview

- Decision Science

Completed Work

- Paper Submitted to Management Science
  - Kairosis: A method for dynamical probability forecast aggregation informed by Bayesian change point detection
    - Awaiting review from fast-track submission process (5 Weeks)
  - We present a new method, informed by work on Bayesian change-point detection, for aggregating probability forecasts over time, which we call "kairosis". Our method begins by constructing for all points in time a posterior probability that this point partitions the forecasts into two sets which are distributed differently. The posterior probabilities are then integrated to give a cumulative mass function from which a weighted median forecast is calculated. Kairosis outperforms standard methods, and is especially suitable for geopolitical forecasting tournaments because it is observed to be robust across disparate questions and forecaster distributions.

Algorithm 1

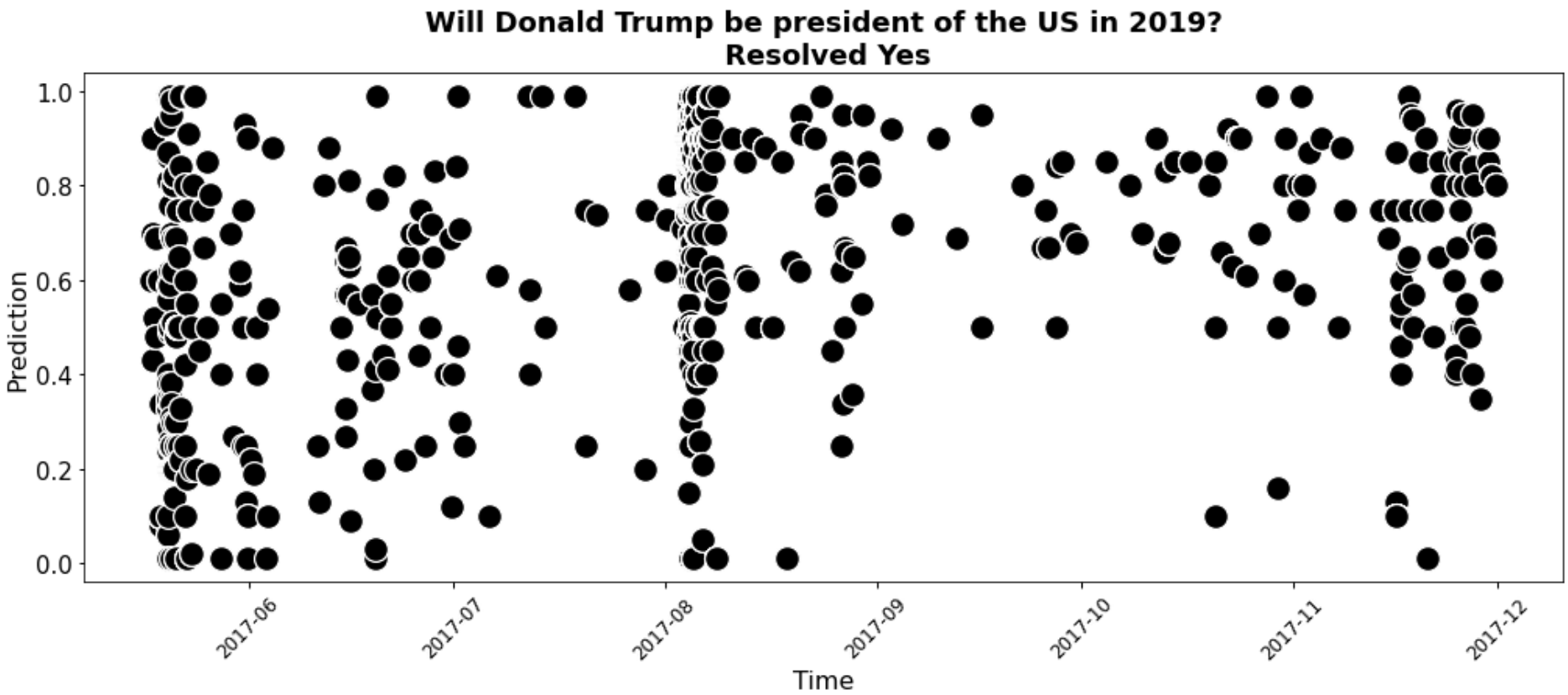


Figure: Trump presidency. Probability forecasts (vertical axis, 0 to 1) submitted to Metaculus by date (horizontal axis) in response to the title question

Algorithm 2

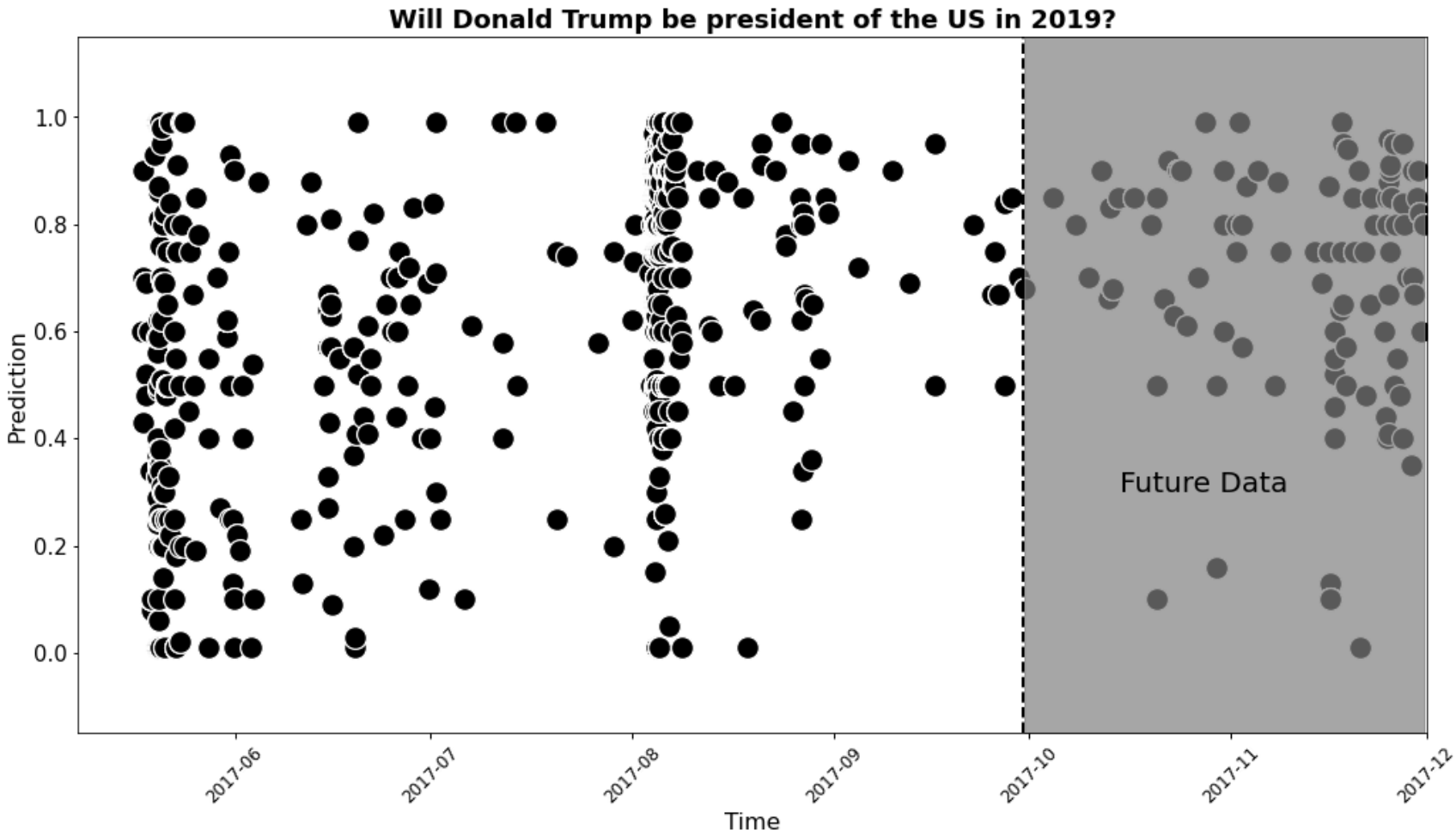


Figure: Trump presidency question at Oct 10, 2017. We seek a method to aggregate forecasts received before this date.

Algorithm 3

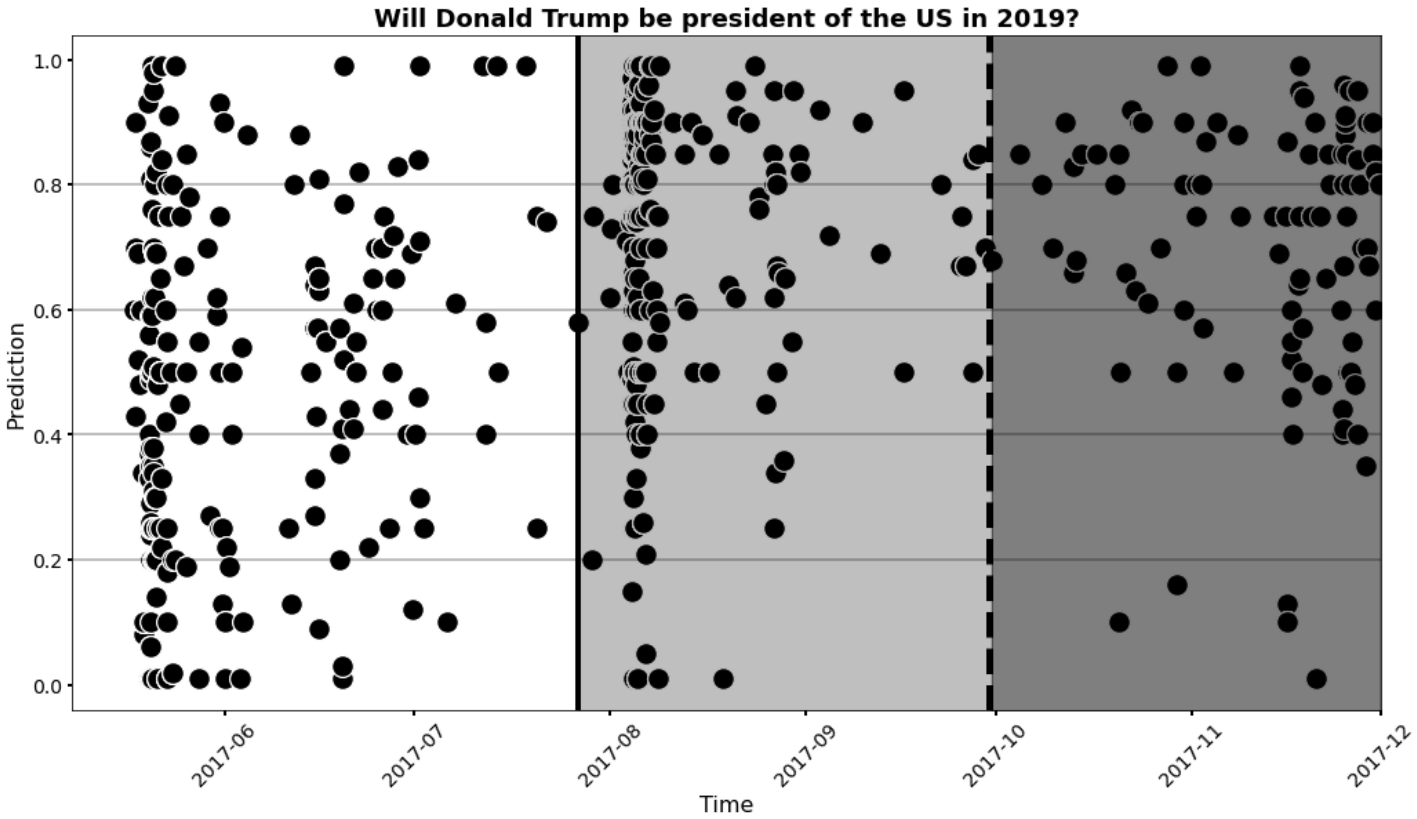
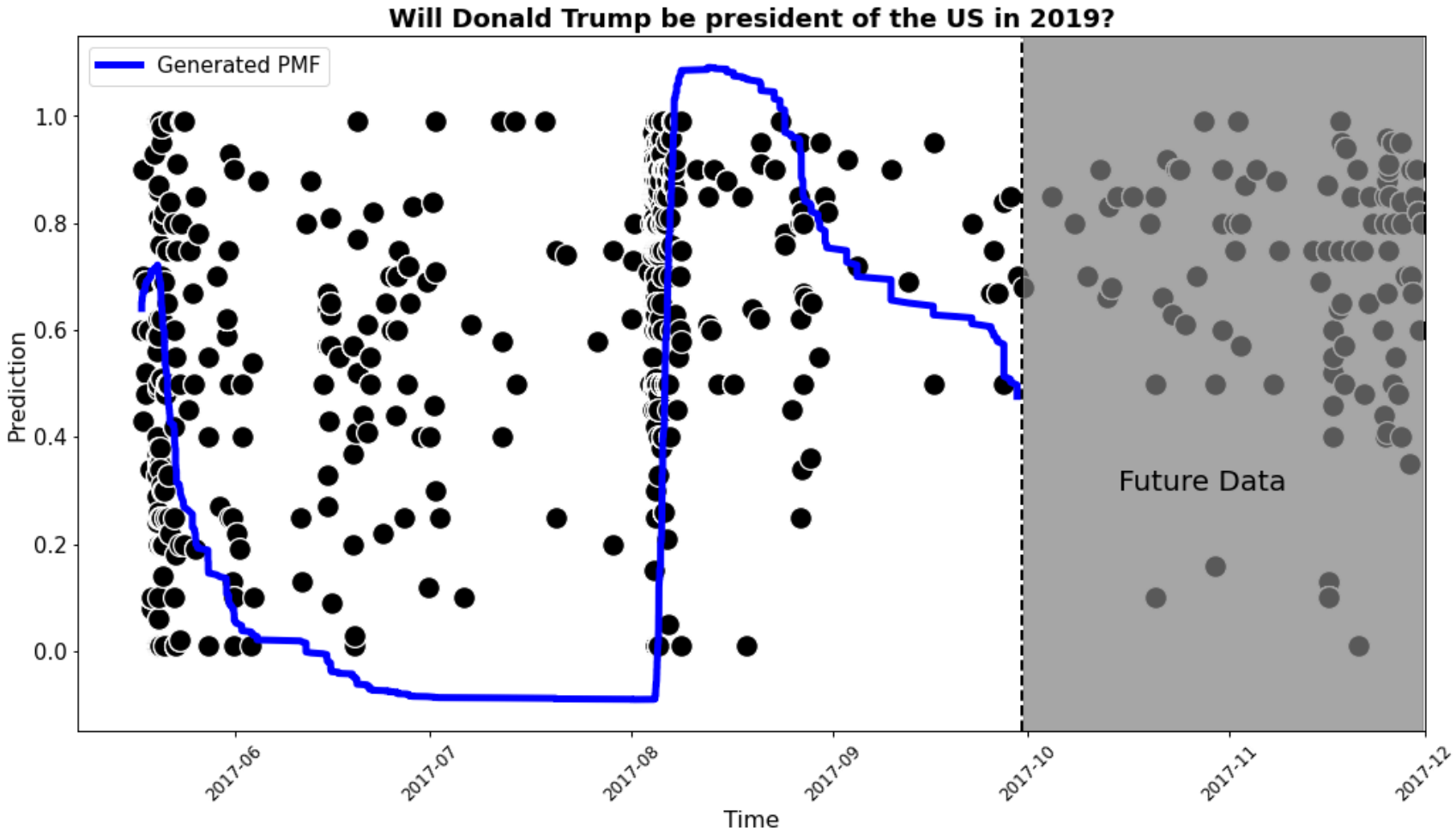
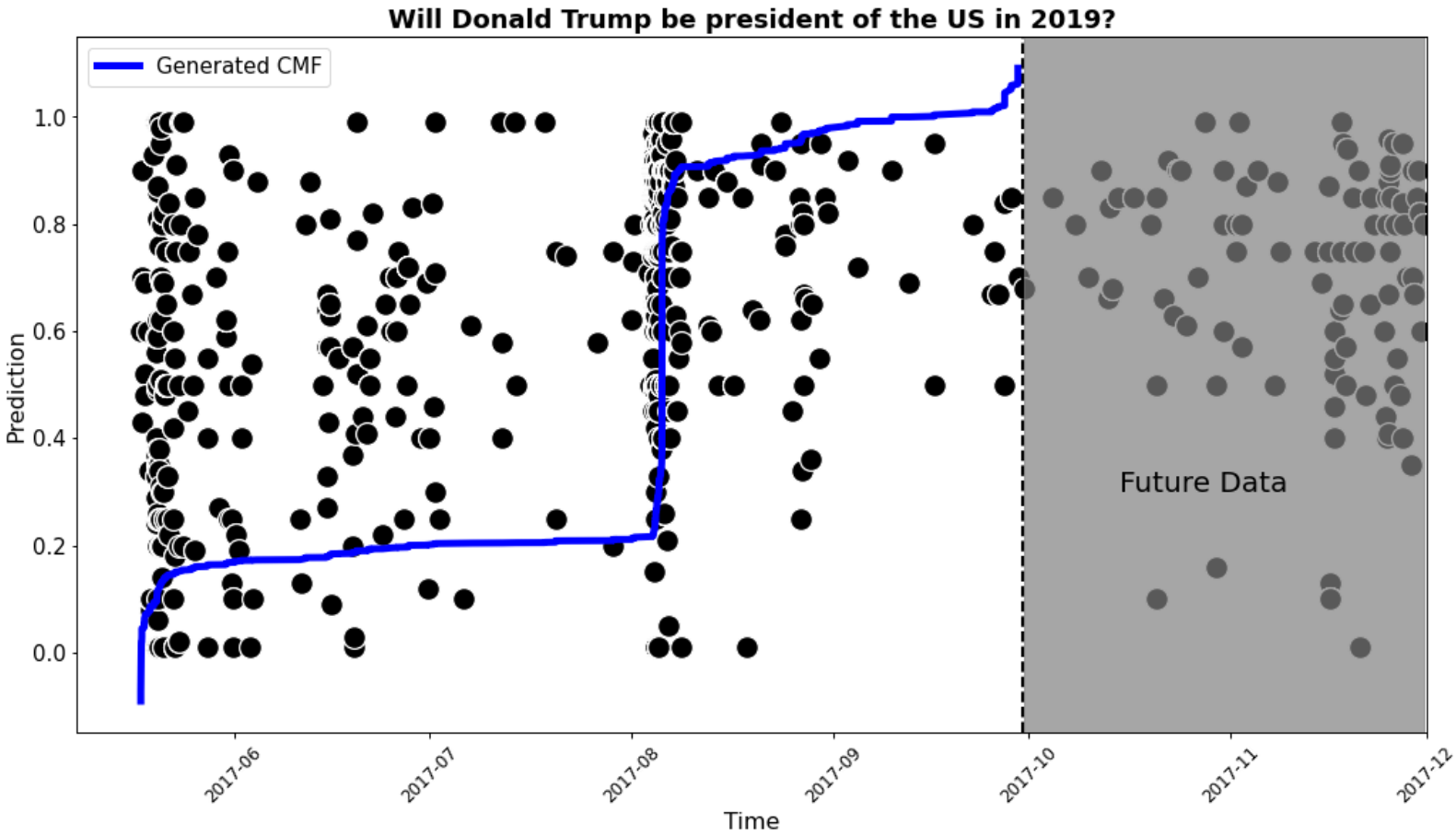


Figure: Weighting a candidate change point. At Oct 10, 2017 (dashed vertical line) we evaluate a candidate change point (solid vertical line) by segregating previous and subsequent forecasts into five bins (separated by horizontal lines), and evaluating the joint likelihood using a Dirichlet-categorical distribution.

Algorithm 4


Will Donald Trump be president of the US in 2019?
Generated PMF
Future Data
Prediction
Time
1.0
0.8
0.6
0.4
0.2
0.0
2017-06
2017-07
2017-08
2017-09
2017-10
2017-11
2017-12

Algorithm 5


Will Donald Trump be president of the US in 2019?
Generated CMF
Future Data
Prediction
Time
1.0
0.8
0.6
0.4
0.2
0.0
2017-06
2017-07
2017-08
2017-09
2017-10
2017-11
2017-12

Algorithm 6

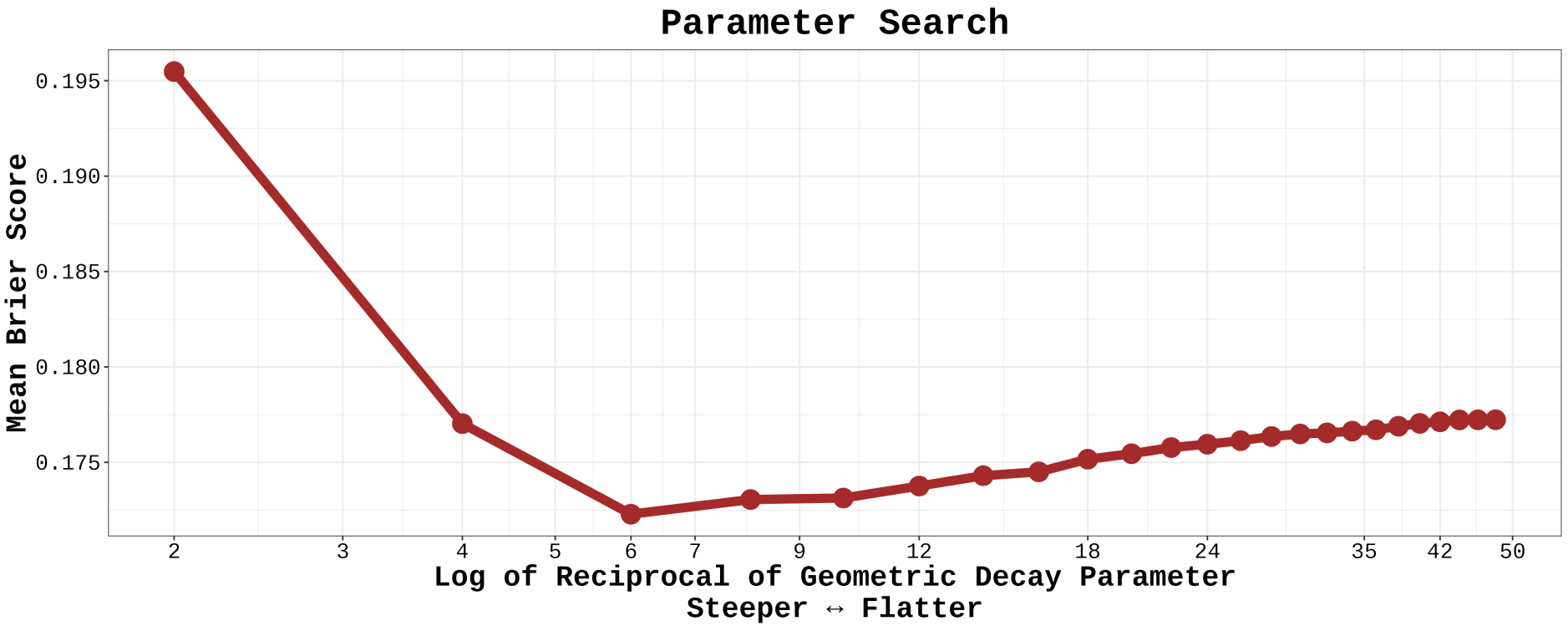


Figure: Geometric decay parameter. Mean Brier score (vertical axis; lower scores are better) against reciprocal of geometric decay parameter, $1/p$ (horizontal axis; log scale). Steeper decay (into the past) corresponds to larger p (on the left), flatter to smaller p (to the right).

Algorithm 7

Model	Aggregation	UBS	WBS	ULS	WLS
Unweighted	Median	0.000	0.000	0.000	0.000
	Mean	3.51	0.452	12.9	5.02
Kairosis	Median	−7.07	−7.37	−15.4	−16.6
	Mean	−4.46	−6.66	−6.31	−12.5
Most Recent 20%	Median	−4.53	13.4	−11.9	36.8
	Mean	−1.49	2.41	−0.57	7.06
Exponential Decay	Median	−4.83	−5.53	−12.2	−12.6
	Mean	−1.52	−5.98	1.09	−9.55

Table: Model performance comparison. Means and medians in four models (rows) against time-weighted and unweighted Brier and log scores (columns). Table entries are [score minus unweighted median score]$\times 10^3$.

Algorithm 8

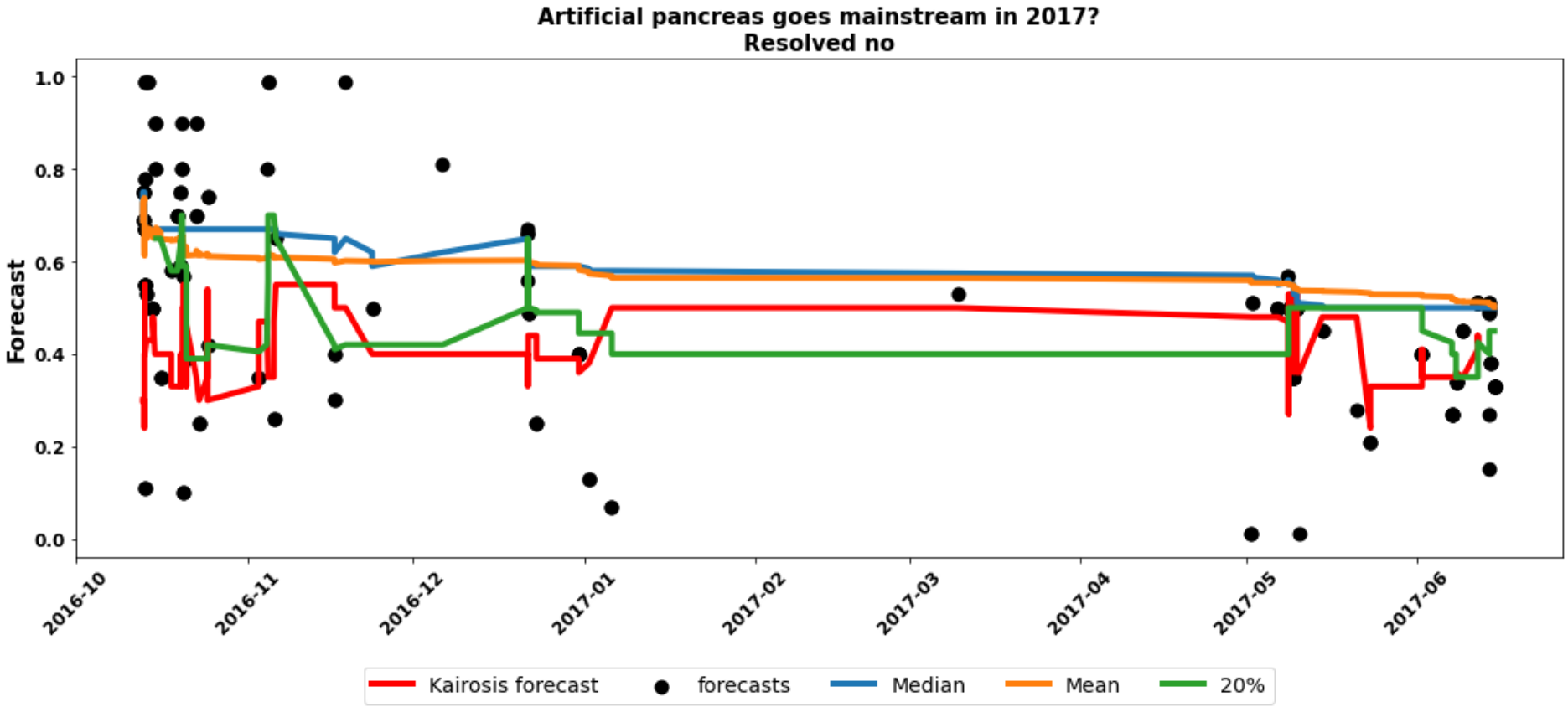


Figure: Comparison of Methods. For a further question resolved as "no", on the general availability of an artificial pancreas by 2017, the plot shows the kairosis forecast along with the mean, median, and median of most recent 20%.

Current Work

- Project 2
- Markov Decision Processes
- Womens Football
-

Click o visit Example website. Click o visit Example website.