

Predicting US Recessions with Gradient-boosted Regression Trees

Danny Cohen, Corporate Intelligence Analyst Intern

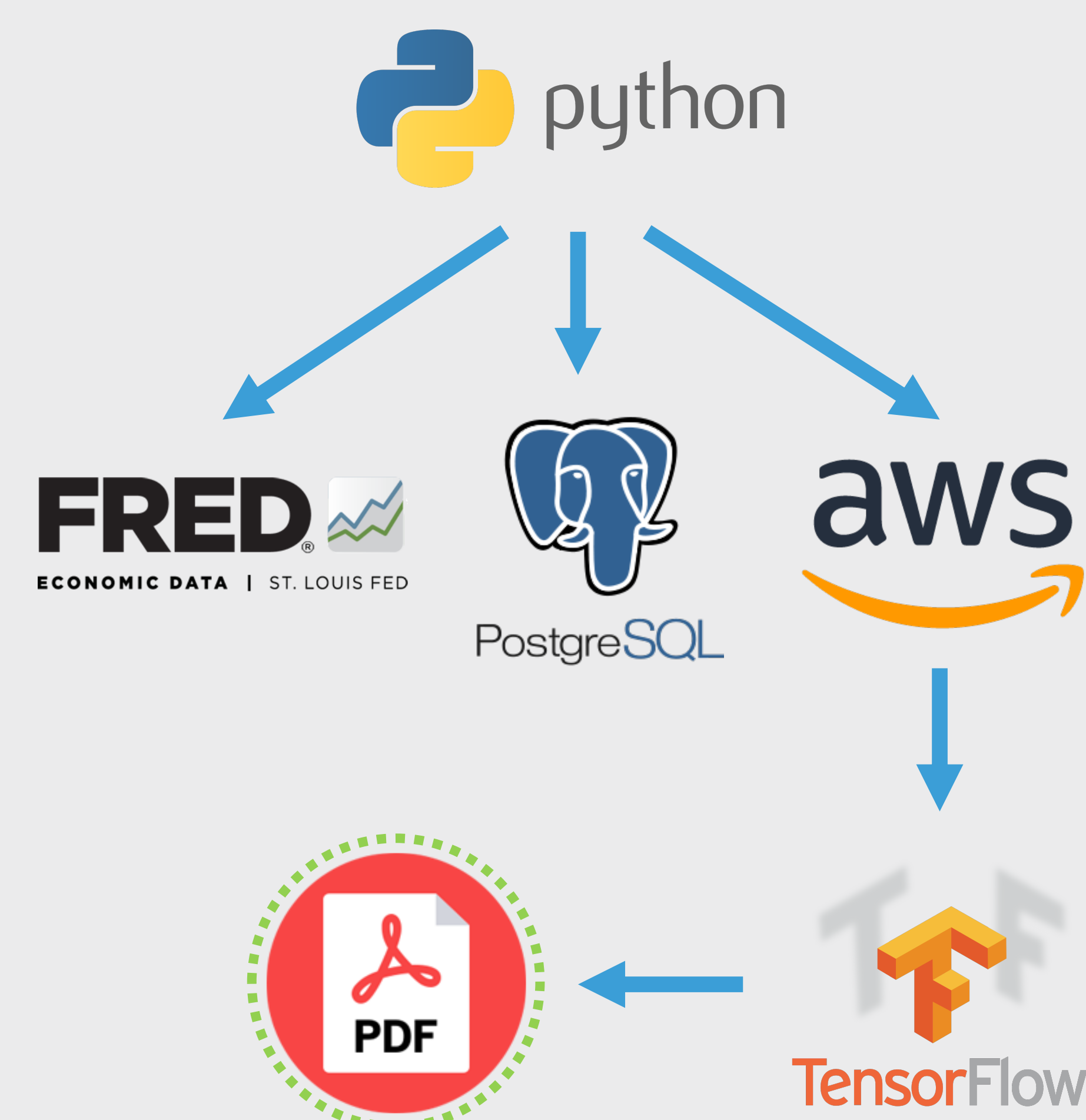
1 Project scoping

Problem: Macroeconomic downturns cause VMware customers to postpone or cancel their purchases, resulting in **unforeseen effects on revenue**.

Solution: Use federal data to make robust **macroeconomic predictions** that inform 1- to 3-month sales forecasts and help set realistic goals.

Deliverable: Monthly report to sales managers with macroeconomic trend predictions for coming 3 months.

2 Approach



Explanation: With **Python** as back end, use **FRED API** to gather federal economic data; process and upload to **PostgreSQL** data warehouse; use **TensorFlow** on **AWS** for ML predictions, and export as **PDF**.

Benefits: **modular**, readable codebase; privy to **570k+ economic data series**; **company-wide access** to underlying data; maximum cloud computing power for **quick, accurate algorithms**.

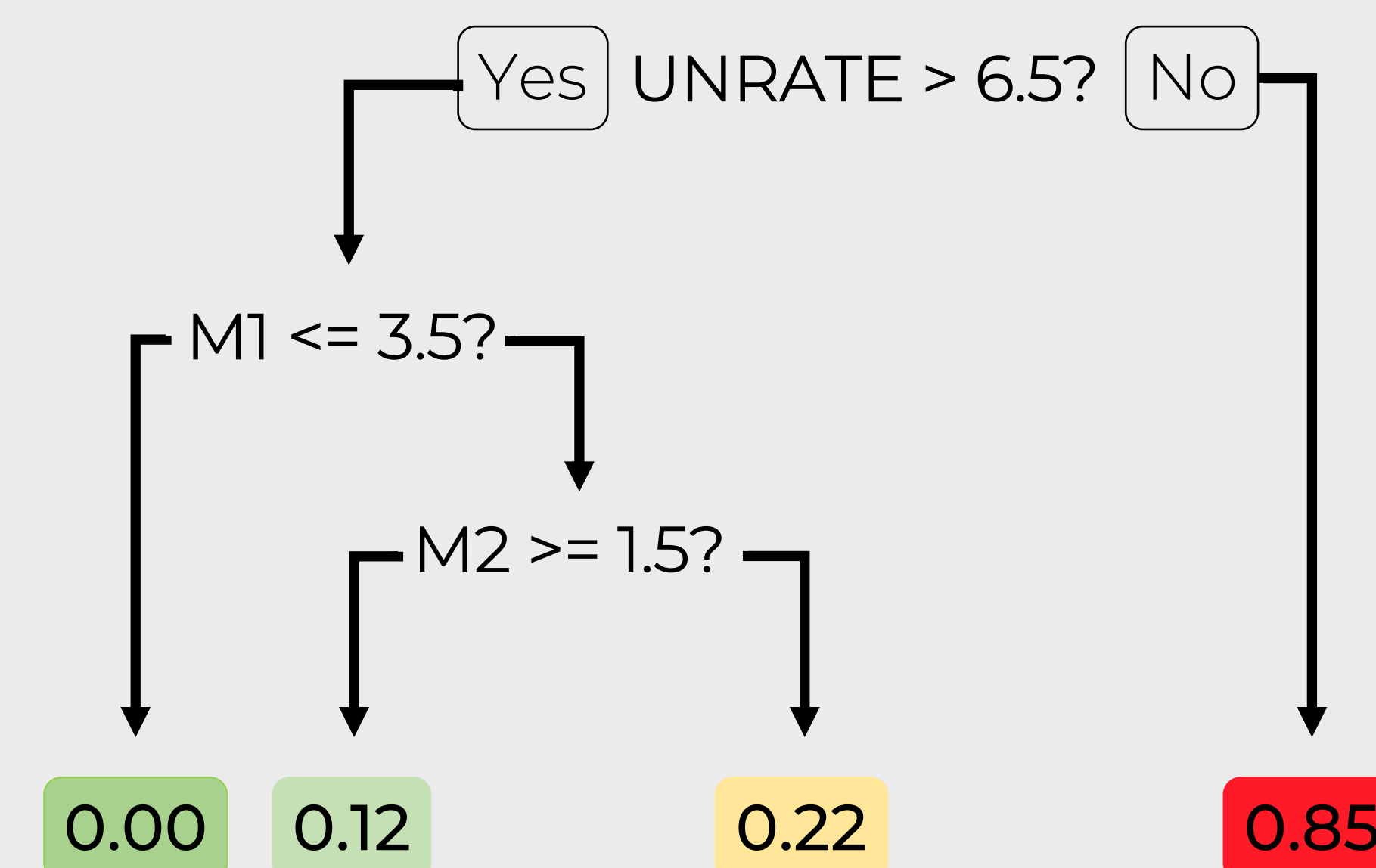
3 Model definition

Setup: Predict likelihood of recession (1) vs. not (0) in 1, 2, 3 months by **classifying** with various continuous, independent economic variables.

Index	Independent		Dependent
Date	UNRATE	M1	... USREC1
1971-10-01	6.74	3.25	... 1

Parameters (58): chosen based on literature review; revolve around **employment, money, stock markets, sentiment, housing, manufacturing, consumption, and business sales**.

Algorithm: **Gradient-boosted Regression Trees**: classify via ensemble of random binary decision trees



Modifications: **Bootstrapping** to reduce random error; **Five-fold Cross Validation** to verify algorithmic accuracy with unseen data

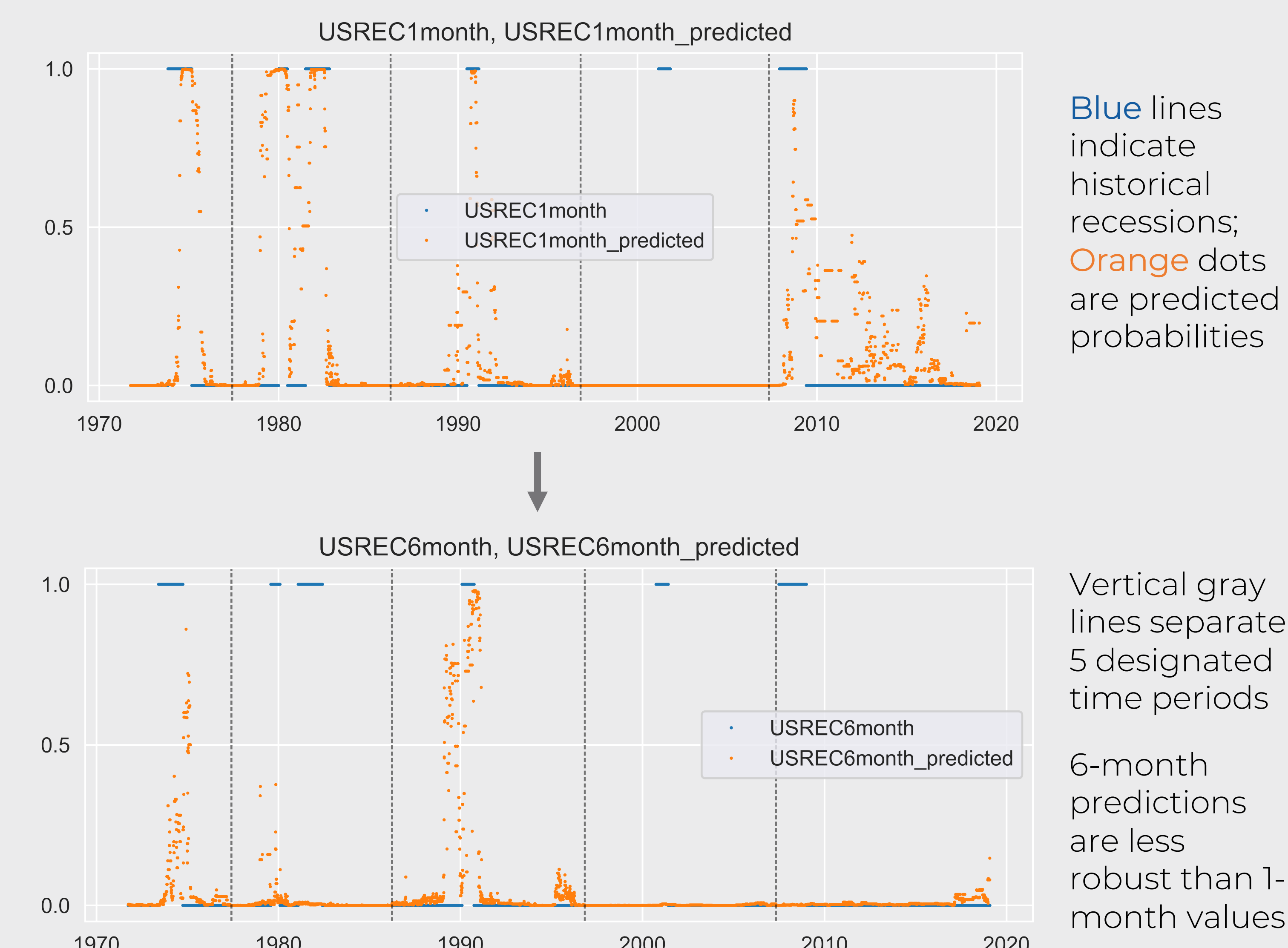
4 Benchmarks

Predictive accuracy: How well can each of the five recessions be predicted by the four others?

Speed: Can model be trained quickly?

Fit: Does model avoid under- or over-fitting?

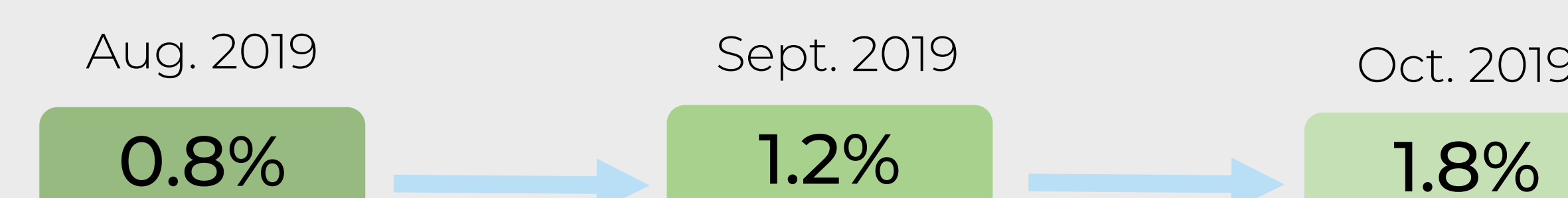
5 Preliminary results



Significance: Predictions for each period made without knowledge of macroeconomy during that period (**true prediction**); only one recession cannot be predicted by other downturns.

Use: Average the five algorithm results to predict future recessions.

6 Recession probability forecast



7 Challenges and next steps

Challenges: Ineffectiveness of initial algorithms; inconsistencies within FRED API; Python date formatting; interfacing with data warehouse

Next steps: Explore more advanced algorithms; redefine algorithmic accuracy; expand independent variable selection; develop new categorical dependent variables (e.g. downturn in particular market); widen model's geographic range; improve reporting format and customize for different consumers