

Moonshot Mission

Theme 3: Social sentiment based models

by Steven Battilana, Noah Berner, Alexis Laignelet, Eric Léger

Datasets:

**End of the day stock price from Quandl,
Social media Analytics from Quandl,
Public Oil price.**

AWS not used.

Motivation & Challenges

Data: Use a sentiment signal to predict the stock price.

**Models: Use ensembling to make
a strong model out of weak models.**

Data Aggregation

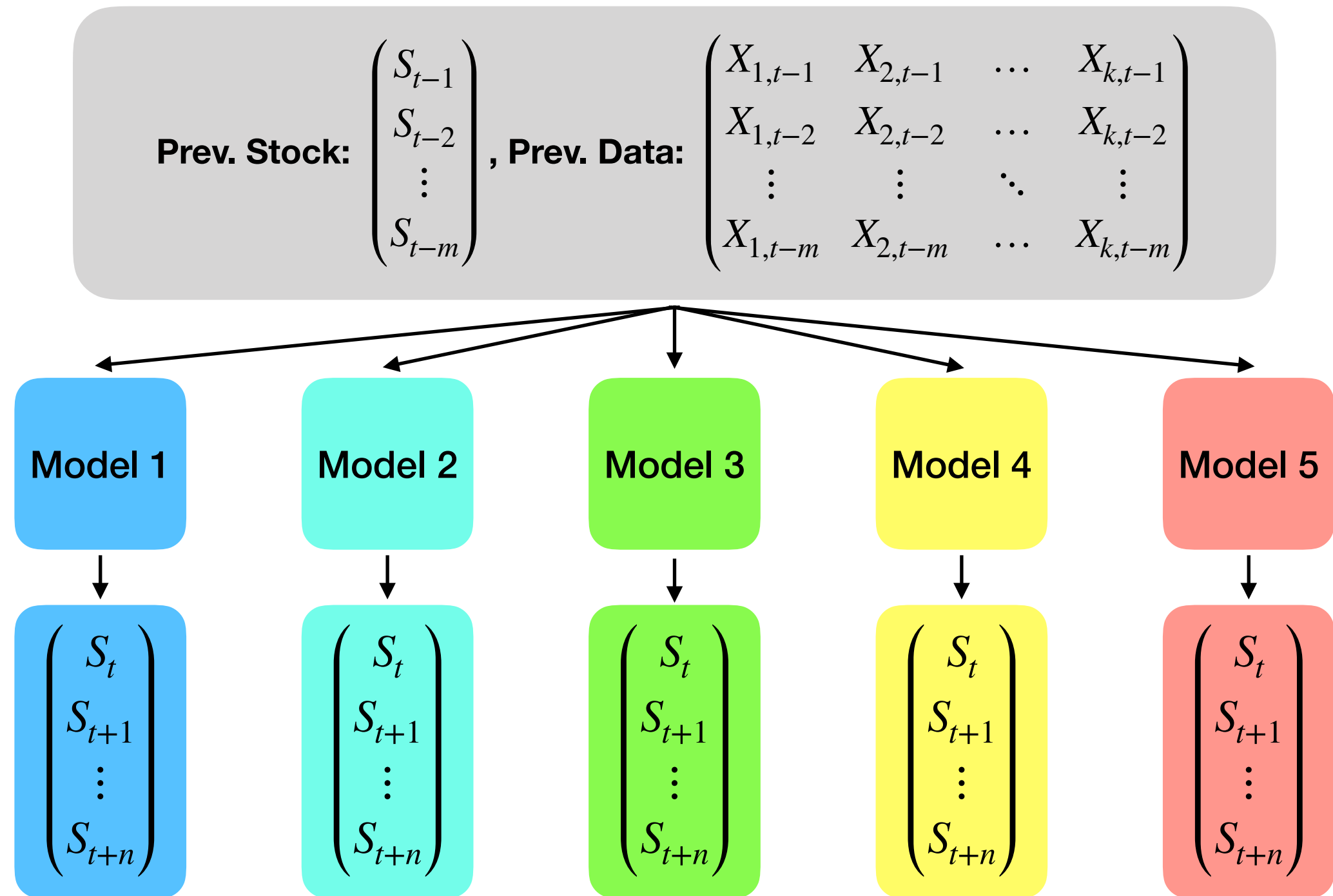
Multiple sources of data:

- Facebook
- Twitter
- Oil Price
- Stock Price (GM)

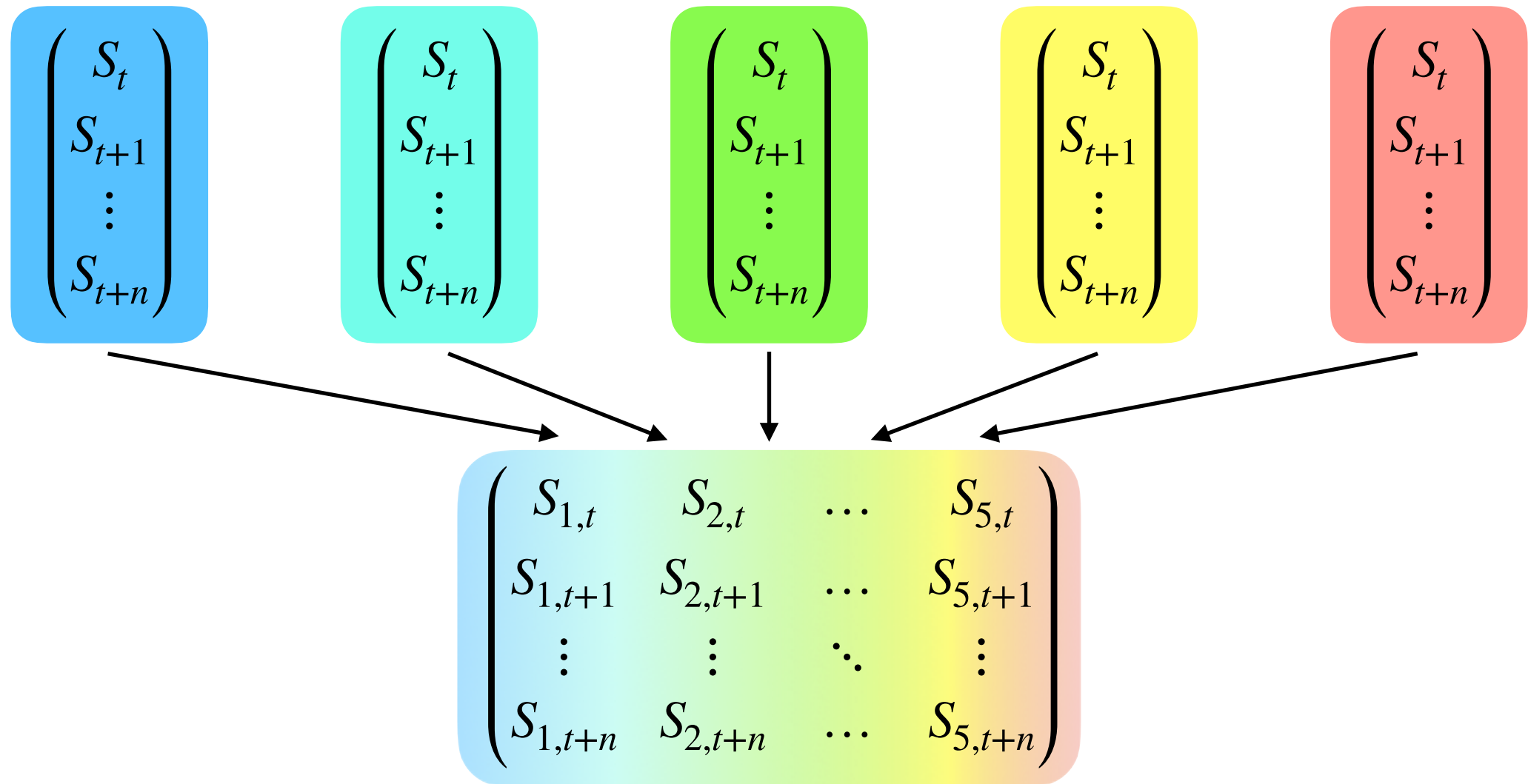
Challenges:

- Aggregation through summing of numerical data
- Merging intra-day data into one point

Ensembling



Ensembling



Ensembling

$$\begin{pmatrix} S_{1,t} & S_{2,t} & \dots & S_{5,t} \\ S_{1,t+1} & S_{2,t+1} & \dots & S_{5,t+1} \\ \vdots & \vdots & \ddots & \vdots \\ S_{1,t+n} & S_{2,t+n} & \dots & S_{5,t+n} \end{pmatrix}$$

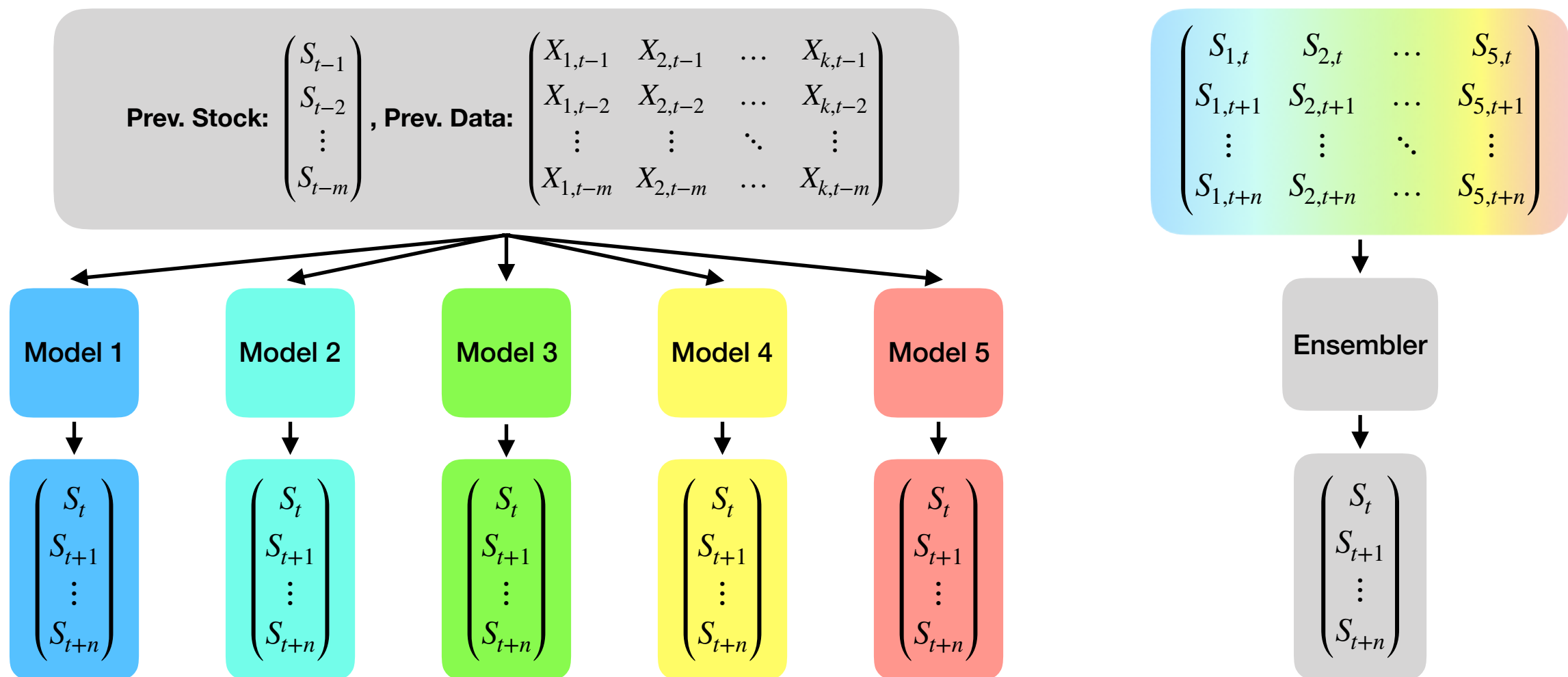


Ensembler

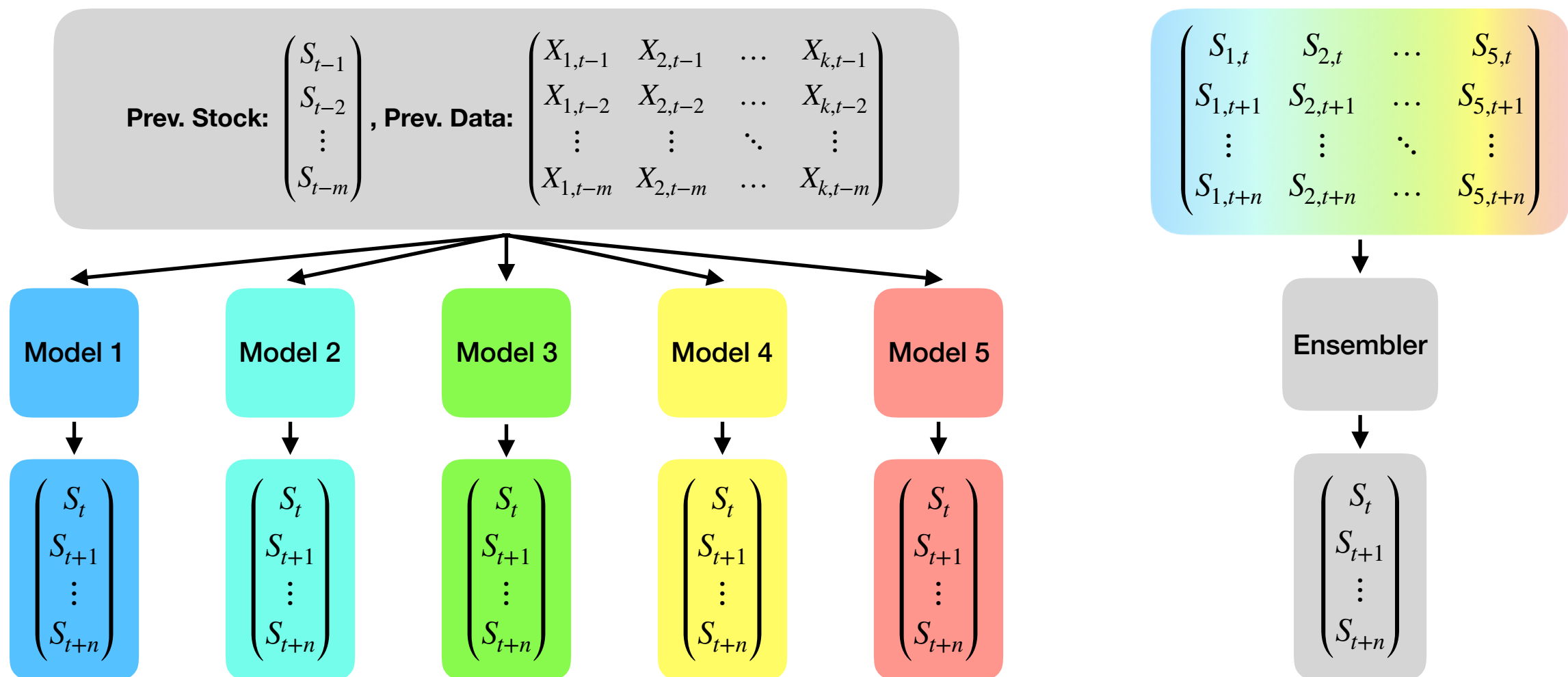


$$\begin{pmatrix} S_t \\ S_{t+1} \\ \vdots \\ S_{t+n} \end{pmatrix}$$

The Framework

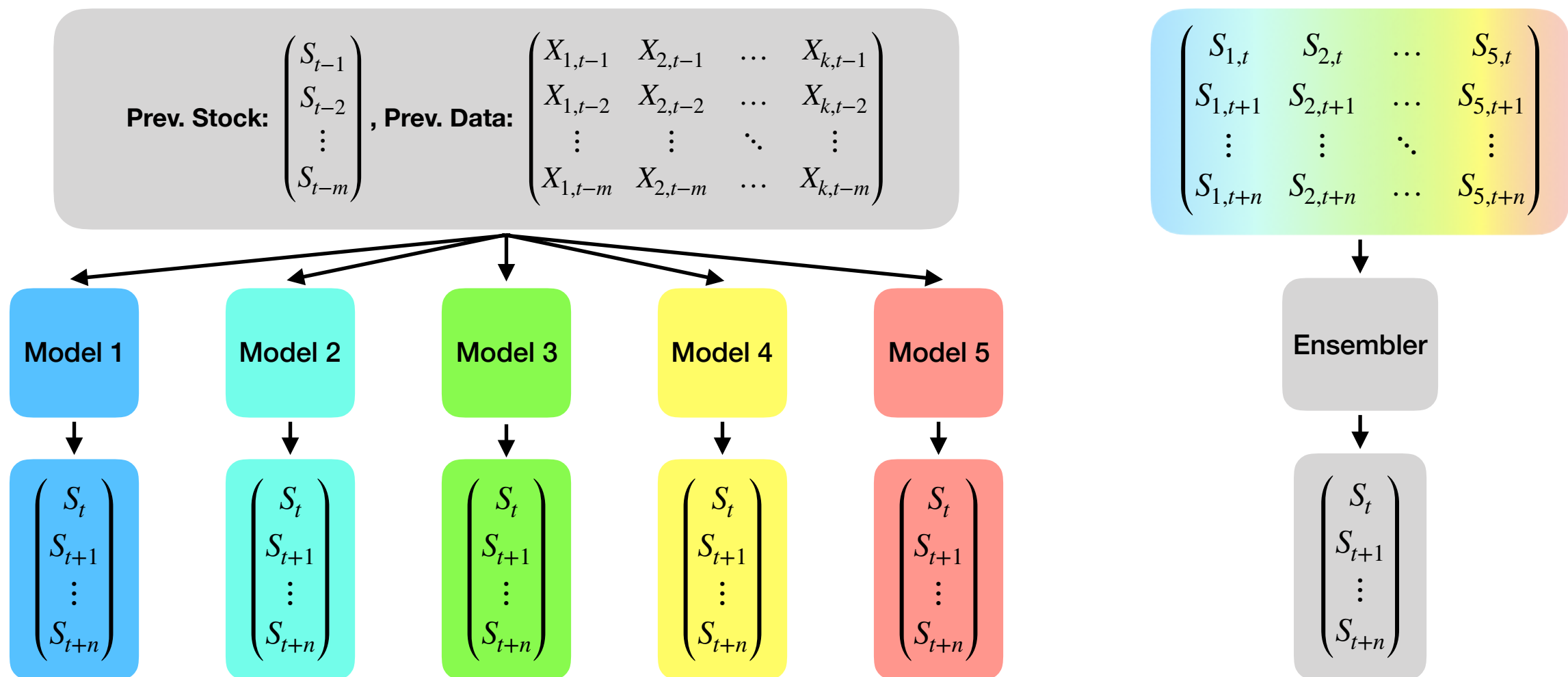


The Framework



Backtesting works!

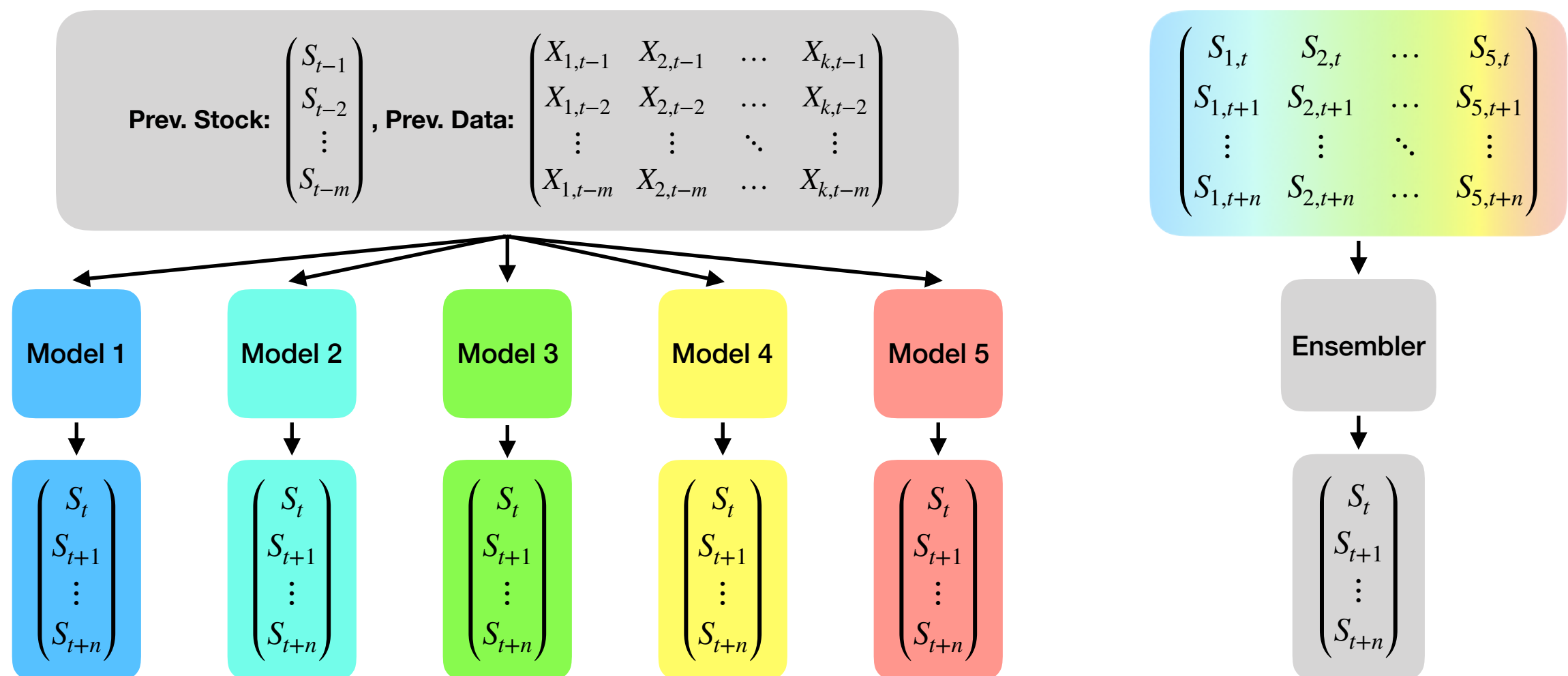
The Framework



Backtesting works!
Ensembling works!

Framework: Future Work

- Build a portfolio on top of backtesting
- Great scalability for any model with `train()` and `predict()`
- Develop multivariate models that adhere to framework requirements



Results

Univariate Models (sharpe-score):

- **Prophet: -22.7**
- **ARIMA: -17.4**
- **Ensemble of both: -1.7**