Forecasting rice prices in the Philippines (initial results)

Vidal Marvin Gabriel

Data Analytics Specialist

April 21, 2021

Outline of the presentation

Objective

- ▶ Model rice prices in the Philippines
- ▶ Forecast rice prices in the Philippines

Proposed models

- ▶ Artificial Neural Networks (Can work with multivariate inputs)
- ▶ Support Vector Machines (Can work with multivariate inputs)
- ▶ ARIMA (univariate)
- ▶ VECM

Data to be used

- ▶ Total Stock
- ▶ Household Stock
- ▶ Commercial Stock
- ▶ NFA stock
- ▶ Rice inflation
  Datasource: PSA, 1995 to 2020

Methodology - ANN

▶ Inputs:
  ▶ For univariate, 3 previous data points = previous 3 months
  ▶ For multivariate, 3 previous data points and other possible time series (features)
    Features will be selected based on correlation coefficient.
▶ 1 to 2 hidden layers - will vary based on performance.
▶ Each hidden layer will have 5 to 50 nodes - will vary based on performance.
▶ Activation function - *tanh*; Optimizer - *gradient descent*
▶ 8 iterations/loops will be done to choose the most optimal regressor.

Methodology - SVM

▶ Inputs:
  ▶ For univariate, previous data point = previous month
  ▶ For multivariate, previous data point and other possible time series (features)
    Features will be selected based on correlation coefficient.

▶ Grid search to look for parameters - will look for different combinations of the parameters of SVM

▶ Kernel - radial base function

Methodology - ARIMA

▶ Univariate rice inflation
▶ Grid search to look for parameters - will look for different combinations of the parameters of ARIMA

Methodology - VECM

- Lag selection - automated
- Check for cointegration

Performance metric

- RMSE - 3 data points
- MAE - 3 data points
- Confusion matrix - proposed metric

Results

Results

Model	RMSE	MAE
ARIMA	0.9	0.9
VECM	0.4	0.4
SVM_uni	1.3	1.2
SVM_mul	0.3	0.3
ANN_uni	0.7	0.6
ANN_mul	0.5	0.5

Performance metrics score.

Confusion matrix

- ▶ Recall/True positive rate - proportion of actual positives was identified correctly.
- ▶ Precision - proportion of positive idenfications was actually correct.
- ▶ F1 score - weighted average of precision and recall.

Confusion matrix identifier

Dichotomous data: **1** - when the model predicted higher inflation than previous inflation and actual inflation is higher than previous inflation; **0** - when the model predicted lower inflation than previous inflation but in actual inflation is higher than previous inflation.

Model	Recall	Precision	F1
ANN_uni	0.8	0.6	0.7

Confusion matrix trial run.

What to do next?

▶ Include confusion matrix (upon approval)
▶ Documentation

Thank You!

Questions?