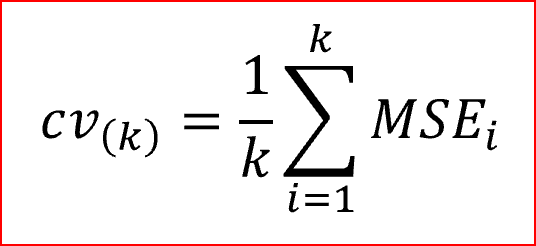
**Cross-validation (CV) techniques for forecasting financial time series**

Three cross-validation (CV) techniques researched include:

1. K fold Cross Validation: This technique involves randomly dividing the dataset into k groups or folds of approximately equal size. The first fold is kept for testing and the model is trained on k-1 folds. The process is repeated K times and each time different fold or a different group of data points are used for validation. As we repeat the process k times, we get k times Mean Square Error (MSE). MSE\_1, MSE\_2, …MSE\_K, so k-Fold CV error is computed by taking average of the MSE over K folds.



The advantages include a reduced bias. Also every data points get to be tested exactly once and is used in training k-1 times and the variance of the resulting estimate is reduced as k increases. However, the training algorithm is computationally intensive as the algorithm has to be rerun from scratch k times. The error rate could be improved by using stratification technique.

1. Stratified Cross Validation: This technique rearranges the data in a way that each fold has a good representation of the whole dataset. It forces each fold to have at least m instances of each class. This approach ensures that one class of data is not overrepresented especially when the target variable is unbalanced. It helps reduce bias and variance.
2. Rolling Origin Cross-Validation: Here the origin at which the forecast is based rolls forward in time. It is also known as time series cross-validation. Each day is a test data and we consider the previous day’s data as the training set. The model is trained with a minimum number of observations and the next day's data is added to test the model and so on moving through the data set. This ensures that the time series aspect of the data is taken into consideration for prediction.

References:

Prado Marcos López de. (2018). *Advances in financial machine learning*. New Jersey: Wiley.

Khandelwal, R. (2018, November 03). K fold and other cross-validation techniques – Data Driven Investor – Medium. Retrieved November 4, 2018, from <https://medium.com/datadriveninvestor/k-fold-and-other-cross-validation-techniques-6c03a2563f1e>

Kumar, A. (2018, February 12). Machine Learning: Validation Techniques - DZone AI. Retrieved November 4, 2018, from https://dzone.com/articles/machine-learning-validation-techniques