

ABSTRAK

Alvin Lorenz Golton (01112180018)

PERBANDINGAN PERFORMA PEMODELAN INDEKS DXY ANTARA MODEL ARIMA DAN MODEL *DYNAMIC REGRESSION* DENGAN DATA SENTIMEN BERITA EKONOMI AMERIKA SERIKAT

Skripsi, Fakultas Sains dan Teknologi (2023)

(xiv + 78 halaman, 19 gambar, 30 tabel)

Telah dipelajari berbagai model matematika dalam memahami pola pergerakan dari instrumen finansial. Namun, tak jarang pergerakan dari instrumen finansial tersebut didorong oleh sentimen pasar. Sentimen pasar umumnya dikemas dalam bentuk berita ekonomi harian. Pada skripsi ini, akan dibandingkan dua macam model matematika untuk mempelajari dan memprediksi pola pergerakan dari sebuah instrumen finansial, yakni indeks DXY. Model matematika yang pertama adalah model deret waktu *Autoregressive Integrated Moving Average* (ARIMA) dengan menggunakan data historis indeks harian DXY. Model matematika yang kedua adalah *Dynamic Regression model* (DRM) dengan menggunakan data historis indeks harian DXY dan data sentimen berita harian. Data Indeks harian DXY diperoleh dari laman *Wall Street Journal*, sedangkan data sentimen harian diperoleh dari hasil analisis sentimen berdasarkan data berita ekonomi yang disediakan oleh *MarketWatch* dan leksikon *SentiWordNet*, *Hu Liu*, *Jockers*, *SenticNet*, dan *MPQA Subjectivity* serta tiga metode agregasi nilai sentimen harian, yakni: penjumlahan, rata-rata, dan persentase. Data-data yang digunakan pada skripsi ini adalah data tanggal 6 Februari 2017 hingga 28 Februari 2020. Performa masing-masing model akan dievaluasi dengan metode *sliding window*. Periode *fitting* yang akan digunakan pada metode *sliding window* adalah 1 (satu) tahun, 1,5 (satu setengah) tahun, dan 2 (dua) tahun, sedangkan periode *forecasting* yang akan digunakan pada metode *sliding window* adalah 5 (lima) hari. Hasil evaluasi dari kedua model adalah dengan menggunakan leksikon tertentu, DRM mampu menghasilkan jumlah keseluruhan MSE ramalan *sliding window* yang lebih rendah daripada model ARIMA.

Kata Kunci : DXY, sentimen, leksikon, *Autoregressive Integrated Moving Average*, *Dynamic Regression Model*, *Sliding Window*

Referensi : 23 (1965-2022)

ABSTRACT

Alvin Lorenz Golton (01112180018)

COMPARISON OF DXY INDEX MODELLING PERFORMANCE BETWEEN THE ARIMA MODEL AND DYNAMIC REGRESSION MODEL USING THE UNITED STATES ECONOMIC NEWS DATA

Thesis, Faculty of Science and Technology (2023)

(xiv + 78 pages, 19 figures, 30 tables)

Mathematical models have been studied to understand the movement of financial instruments. However, the movement of these financial instruments is usually driven by the market sentiment. Market sentiment is usually found on daily economic news. In this thesis, there are two kinds of mathematical models that will be compared to study and predict the movement of a financial instrument, the DXY index. The first mathematical model is the Autoregressive Integrated Moving Average (ARIMA) time series model using historical data of DXY daily index. The second mathematical model is the Dynamic Regression model (DRM) using historical data of DXY daily index and daily news sentiment data. DXY daily index data is obtained from the Wall Street Journal page, while daily sentiment data is obtained from sentiment analysis results based on economic news provided by MarketWatch and the SentiWordNet lexicon, Hu Liu lexicon, Jockers lexicon, SenticNet lexicon, and MPQA Subjectivity lexicon and three daily sentiment aggregation methods, namely: sum, average, and percentage. The data used in this thesis is data from February 6, 2017 to February 28, 2020. The performance of each model will be evaluated using the sliding window method. The fitting period to be used in sliding window method is 1 (one) year, 1.5 year (one year and six months), and 2 (two) years, while the forecasting period to be used in sliding window method is 5 (five) days. The results of the evaluation of the two models are that by using a certain lexicon, DRM is able to produce a lower overall sum of sliding window forecasting MSE than the ARIMA model.

Keywords : DXY, sentiment, lexicon, Autoregressive Integrated Moving Average, Dynamic Regression Model, Sliding Window

Reference : 23 (1965-2022)