

# Transformasi Data pada Data Prediksi Nilai Tukar Rupiah

<sup>1</sup>Primandani Arsi, <sup>2</sup>Desty Rakhmawati

<sup>12</sup>Program Studi Teknik Informatika

<sup>12</sup>STMIK Amikom Purwokerto

<sup>12</sup>Purwokerto, Indonesia

<sup>1</sup>ukhti.prima@amikompurwokerto.ac.id, <sup>2</sup>desty@amikompurwokerto.ac.id,

**Abstrak**— Dewasa ini fenomena Big Data terus berkembang. Big Data adalah fenomena yang terjadi secara alami karena perkembangan volume data yang sangat besar, didukung oleh ketersediaan implementasi teknologi. Salah satu fenomena Big Data di bidang ekonomi adalah data nilai tukar rupiah terhadap dolar yang disajikan oleh [www.bi.go.id](http://www.bi.go.id). Untuk mendapatkan data input yang berkualitas, teknik preprocessing data dilakukan pada data prediksi nilai tukar rupiah terhadap dolar secara bertahap yaitu melalui proses reduksi, normalisasi dan windowing. Dari pengujian menunjukkan bahwa hasil prediksi terbaik adalah data yang telah diolah sebelumnya dan menghasilkan RMSE 0,010 +/- 0,001 sedangkan hasil dari data yang tidak diolah menghasilkan RMSE 168.275 +/- 41,884. Kesimpulannya adalah transformasi nilai tukar rupiah terhadap dolar dengan teknik reduksi, normalisasi dan windowing memiliki peran yang sangat penting dalam menentukan akurasi prediksi terbaik.

**Kata kunci**— Big data, preprocessing, reduksi, normalisasi, windowing dan RMSE

**Abstract**— Today the Big Data phenomenon continues to grow. Big Data is a natural phenomenon due to the development of a very large volume of data, supported by the availability of technology implementation. One of the Big Data phenomena in the field of economics is the exchange rate data of the rupiah against the dollar presented by page [www.bi.go.id](http://www.bi.go.id). In order to get quality input data, the data preprocessing technique is performed on prediction data on the rupiah exchange rate against the dollar by stages, reduction, normalization and windowing. From the test shows that the best prediction results are data that has been preprocessed which results in RMSE 0.010 +/- 0.001 while the results of the unprocessed data produce RMSE 168.275 +/- 41,884. The conclusion is the transformation of the rupiah exchange rate against the dollar with reduction, normalization and windowing techniques has a very important role in determining the accuracy of the best predictions.

**Keywords**—Big data, preprocessing, reduction, normalisation, windowing and RMSE

## I. PENDAHULUAN

Dewasa ini fenomena Big Data terus berkembang sangat pesat[1]. Big Data adalah fenomena yang terjadi secara alami akibat perkembangan volume data yang sangat besar, didukung oleh ketersediaan teknologi implementasinya[2]. Big data merupakan istilah lain dari dataset dalam jumlah yang sangat besar, rumit dan tak terstruktur[3]. Perkembangan Big Data sangat pesat pada semua bidang baik teknik, sains, ekonomi, ilmu fisik, biologi dan biomedis[4]. Big Data memiliki tiga karakteristik yang dikenal dengan istilah 3V: *Volume, Variety, Velocity*. Dalam hal ini, volume menggambarkan ukuran yang

super besar, variety menggambarkan jenis yang sangat beragam, dan velocity menggambarkan laju pertumbuhan maupun perubahannya[5]. Salah satu fenomena Big Data dalam bidang ekonomi adalah pada data nilai tukar rupiah terhadap dolar yang disajikan oleh laman [bi.go.id](http://bi.go.id). Setiap hari dan sepanjang tahun data tersebut terus berkembang dan disajikan dalam bentuk time series[6]. Data time series adalah data yang memiliki runtun waktu yang lebih dari satu tahun pada satu objek atau data yang dikumpulkan dari waktu ke waktu terhadap satu objek[7].

Data mining adalah suatu teknik menggali informasi berharga yang terpendam atau tersembunyi pada suatu koleksi data yang sangat besar sehingga ditemukan suatu pola yang menarik yang sebelumnya tidak diketahui[8]. Tahap-tahap data mining secara umum adalah pembersihan data, integrasi data, transformasi data, aplikasi teknik data mining, evaluasi pola yang ditemukan dan presentasi pengetahuan[9]. Tahap awal pada saat sebelum dilakukan aplikasi teknik data mining adalah preparasi data.

Preparasi data atau sering disebut juga dengan data preprocessing adalah suatu proses yang dilakukan untuk membuat data mentah menjadi data yang berkualitas[10]. Proses ini merupakan langkah pertama dalam proses knowledge discovery[11]. Preprocessing data membutuhkan banyak usaha dan waktu lebih dari 50% dari keseluruhan proses[12]. Biasanya, data mentah memiliki banyak ketidaksempurnaan seperti missing values, noise, redundansi dan inkonsistensi. Jika data yang berkualitas rendah tersebut tetap digunakan maka akan mengakibatkan penurunan performa pada algoritma pembelajaran yang diterapkan. Oleh sebab itu diperlukan tahapan preprocessing data yang tepat guna meningkatkan kualitas dan kehandalan hasil mining[13].

Teknik preprocessing data diantaranya adalah data cleaning, data integration, data reduction dan data transformation. Tahap ini merupakan proses yang dianggap wajib dan penting[14]. Hal ini dikarenakan hasil dari proses tersebut bertujuan untuk mengubah inputan menjadi lebih berkualitas untuk dapat memenuhi proses teknik penambangan data[15].

Teknik penambangan data dapat dilakukan dengan syarat data telah dilakukan pengolahan awal. Penelitian yang dilakukan oleh Xiaorong Cheng dan Hui Liu melakukan teknik preprocessing data dengan menggunakan data sistem peralatan keselamatan dimana teknik data cleaning dilakukan dengan 4 tahapan yaitu *log data collection, data cleaning, data merging*

dan *data conversion*[16]. Penelitian selanjutnya menggunakan data struktur web yang heterogen dan tidak terstruktur. Dengan menerapkan algoritma preprocessing yang terdiri dari 4 tahap yaitu data cleaning, data integration, data transformation dan data reduction[17]. Dwivedi dan rawat melakukan preprocessing data dengan tahapan; pengumpulan data, pembersihan data, dilanjutkan dengan indentifikasi sesi, identifikasi pengguna dan tahap penyelesaian[18].

Pada penelitian ini data nilai tukar rupiah yang diperoleh memiliki kompleksitas sehingga apabila tidak dilakukan preprocessing data maka menyebabkan hasil dari proses data mining selanjutnya menjadi kurang akurat[19]. Oleh karena itu diperlukan suatu teknik preprocessing data yang tepat guna mengurangi kompleksitas. Pengurangan kompleksitas menjadi langkah wajib yang harus dilakukan sebelum data diolah ketahap selanjutnya. Penyederhanaan dari kompleksitas data dapat dilakukan dengan teknik reduksi. Proses reduksi dapat mempertahankan konteks asli data namun dengan ukuran yang lebih mudah untuk dikelola[20]. Dengan demikian untuk mendapatkan data masukan yang berkualitas maka dilakukan teknik preprocessing data pada data prediksi nilai tukar rupiah terhadap dollar dengan tahapan, reduksi data, normalisasi dan windowing.

II. METODE PENELITIAN

A. Reduksi Data

Data nilai tukar rupiah terhadap dolar dengan kurun waktu selama 6 tahun yakni 1 januari 2012 hingga 31 Agustus 2018. Diperoleh dari halaman resmi Bank Indonesia dengan jumlah 1470 record, terdiri dari 4 atribut yakni nilai, kurs beli, kurs jual dan tanggal. Data kemudian direduksi untuk mendapatkan struktur yang diinginkan untuk mempermudah proses mining yang akan dilakukan.

B. Normalisasi

Normalisasi data merupakan suatu pendekatan sistematis untuk meminimalkan redundansi data pada suatu database agar database tersebut dapat bekerja dengan optimal. Tujuan normalisasi data adalah untuk menghilangkan dan mengurangi redundansi data dan memastikan dependensi data (data berada pada tabel yang tepat). Data yang direkam dan dimasukkan secara mentah dalam suatu tabel pada bentuk ini sangat mungkin terjadi inkonsistensi dan anomali data. Dalam penelitian ini normalisasi data dilakukan menggunakan dua fungsi aktivasi yaitu fungsi *binary sigmoid* dan *bipolar sigmoid*. Fungsi *binary sigmoid*, menormalisasikan data dalam range 0 sampai 1 [21]. Berikut persamaan fungsi *binary sigmoid* (*Logsig*).

$$y' = \frac{x - x_{min}}{x_{max} - x_{min}} x (0.9 - 0.1) + (0.1) \tag{1}$$

Berikut persamaan fungsi *bipolar sigmoid* (*Tansig*).

$$y' = \frac{x - x_{min}}{x_{max} - x_{min}} x 2 - 1 \tag{2}$$

Keterangan:

- $y'$  = Hasil transformasi data
- $x$  = Nilai asli
- $x_{min}$  = Nilai min
- $x_{max}$  = Nilai max

C. Windowing

*Windowing* adalah sebuah teknik yang digunakan dalam menentukan dan menetapkan input dan output data pada prediksi time series jenis univariat. Data univariat merupakan data yang terdiri hanya satu atribut saja[22]. Dengan menggunakan teknik windowing data kurs rupiah akan dibagi menjadi beberapa data, dimana data terakhir merupakan output dari beberapa hari sebelumnya yang digunakan sebagai input. Dalam penelitian ini pembagian window yang dilakukan adalah 10 [23].

III. HASIL DAN PEMBAHASAN

A. Reduksi Data

Reduksi data pada data nilai tukar rupiah dilakukan dengan cara menyederhanakan jumlah atribut yang ada, namun tetap memiliki informasi yang sama. Data nilai tukar rupiah memiliki 4 atribut yaitu nilai, kurs beli, kurs jual dan tanggal, seperti tampak pada tabel 1 dibawah ini:

TABEL 1 DATASET REAL

No	Nilai	Kurs Beli	Kurs Jual	Tanggal
1	1	14.785,00	14.637,00	31 Agust 2018
2	1	14.728,00	14.582,00	30 Agust 2018
3	1	14.716,00	14.570,00	29 Agust 2018
4	1	14.687,00	14.541,00	28 Agust 2018
5	1	14.683,00	14.537,00	27 Agust 2018
6	1	14.728,00	14.582,00	24 Agust 2018
.....	.....	.....	.....	.....
.....	.....	.....	.....	.....
1467	1	9.787,00	9.689,00	7 Jan 2012
1468	1	9.723,00	9.627,00	4 Jan 2012
1469	1	9.718,00	9.622,00	3 Jan 2012
1470	1	9.733,00	9.637,00	2 Jan 2012

Setelah dilakukan teknik reduksi data dengan menyederhanakan beberapa atribut yang tidak diperlukan, maka didapatkan 2 atribut kurs jual dan tanggal. Seperti tampak pada tabel 2.

TABEL 2 DATASET SETELAH DIREDUKSI

Kurs Jual	Tanggal
14.637,00	31 Agust 2018
14.582,00	30 Agust 2018
14.570,00	29 Agust 2018
14.541,00	28 Agust 2018
14.537,00	27 Agust 2018
14.582,00	24 Agust 2018
.....	.....
.....	.....
9.689,00	7 Jan 2012
9.627,00	4 Jan 2012
9.622,00	3 Jan 2012
9.637,00	2 Jan 2012

### B. Normalisasi Data

Setelah melalui tahap reduksi data langkah selanjutnya adalah menormalisasi data, tujuannya adalah untuk menghilangkan dan mengurangi redundansi data. Dalam penelitian ini normalisasi dilakukan pada range 0 sampai 1. Berikut hasil data setelah dilakukan proses normalisasi.

TABEL 3 DATASET SETELAH DINORMALISASI

Row	Kurs	Date
1	0.798	30 Desember 2017
2	0.798	29 Desember 2017
3	0.800	28 Desember 2017
4	0.800	27 Desember 2017
5	0.800	22 Desember 2017
6	0.797	21 Desember 2017
7	0.803	20 Desember 2017
8	0.804	19 Desember 2017
9	0.804	18 Desember 2017
10	0.802	15 Desember 2017
.....	.....	.....
.....	.....	.....
1467	0.047	5 Januari 2012
1468	0.049	4 Januari 2012
1469	0.046	3 Januari 2012
1470	0.040	2 Januari 2012

### C. Eksperimen Window size

Dengan menggunakan teknik windowing data kurs rupiah dibagi menjadi beberapa data, dimana beberapa hari sebelumnya merupakan nilai input dan data terakhir merupakan output.

TABEL 4 SAMPLE SET DENGAN WINDOW SIZE 10

Row Id	Window Id	Output	Input			
			9	8	N	0
1	0	0.801	0.798	0.798	.....	0.802
2	1	0.797	0.800	0.800	.....	0.798
3	2	0.794	0.803	0.804	.....	0.798
4	3	0.792	0.802	0.801	.....	0.794
5	4	0.794	0.798	0.797	.....	0.794
6	5	0.793	0.798	0.794	.....	0.790
7	6	0.797	0.794	0.792	.....	0.795
8	7	0.792	0.794	0.792	.....	0.796
9	8	0.795	0.790	0.794	.....	0.792
10	9	0.805	0.795	0.793	.....	0.790
.....	.....	.....	.....	.....	.....	.....
484	483	0.049	0.019	0.016	.....	0.054
485	484	0.053	0.018	0.022	.....	0.055
786	485	0.046	0.011	0.031	.....	0.051
487	486	0.046	0.054	0.049	.....	0.049

Teknik windowing yang telah dilakukan mendapatkan sample set sebanyak 487 dari 1470 dataset yang ada, seperti tampak pada tabel 4 diatas dengan nilai  $x_1 = 0.797$ ,  $x_2 = 0.797$ ,  $x_3 = 0.799$ ,  $x_4 = 0.800$ ,  $x_5 = 0.799$ ,  $x_6 = 0.797$ ,  $x_7 = 0.803$ ,  $x_8 = 0.804$ ,  $x_9 = 0.803$ ,  $x_{10} = 0.802$  dan nilai  $y = 0.801$ .

## IV. PENGUJIAN

Pengujian dilakukan dengan menerapkan algoritma neural network untuk diterapkan sebagai algoritma prediksi dari preprocessing data yang telah dilakukan.

TABEL 5 PENGUJIAN DATA TERHADAP ALGORITMA NEURAL NETWORK

Data	RMSE
Melalui tahap preprocessing	0.010 +/- 0.001
Tidak Melalui tahap preprocessing	168.275 +/- 41.884

Dari pengujian yang telah dilakukan tampak pada tabel 5 diatas menunjukkan bahwa hasil prediksi terbaik adalah pada data yang telah melalui tahap preprocessing data yakni menghasilkan RMSE 0.010 +/- 0.001. Pengolahan data awal memiliki peranan yang sangat penting dalam menentukan akurasi prediksi terbaik.

## V. KESIMPULAN

Pada penelitian ini dilakukan usulan tahapan preprocessing data dengan teknik reduksi, normalisasi dan teknik windowing pada transformasi data prediksi nilai tukar rupiah terhadap dolar. Pengujian yang dilakukan dengan menerapkan algoritma neural network pada data yang telah mengalami preprocessing menunjukkan perbedaan RMSE yang signifikan yakni sebesar 0.010 +/- 0.001 dan RMSE 168.275 +/- 41.884 pada data yang tidak mengalami preprocessing.

Dengan demikian dapat disimpulkan bahwa transformasi data pada data nilai tukar rupiah dengan teknik reduksi, normalisasi dan windowing berpengaruh pada akurasi prediksi yang dihasilkan.

## DAFTAR PUSTAKA

- [1] M. A. Mayer, L. Fernández-Luque, and A. Leis, *Big Data For Health Through Social Media*. Elsevier Inc., 2016.
- [2] S. García, S. Ramírez-Gallego, J. Luengo, J. M. Benítez, and F. Herrera, “Big data preprocessing: methods and prospects,” *Big Data Anal.*, vol. 1, no. 1, p. 9, 2016.
- [3] K. Zhou, C. Fu, and S. Yang, “Big data driven smart energy management: From big data to big insights,” *Renew. Sustain. Energy Rev.*, vol. 56, pp. 215–225, 2016.
- [4] X. Wu, Fellow, G.-Q. Wu, and W. Ding, “Data Mining with Big Data,” *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 4, pp. 331–333, 2014.
- [5] L. N. Sanchez-pinto, Y. Luo, and M. M. Churpek, “Big Data and Data Science in Critical Care,” *Chest*, 2018.
- [6] “Bank Indonesia.” [Online]. Available: <https://www.bi.go.id/id/moneter/informasi-kurs/transaksi-bi/Default.aspx>.
- [7] V. Rajagopalan, A. Ray, R. Samsi, and J. Mayer, “Pattern identification in dynamical systems via symbolic time series analysis,” vol. 40, pp. 2897–2907, 2006.
- [8] J. Han, M. Kamber, and J. Pei, *Data mining: concepts and techniques*, Third Edit., vol. 5. USA: Elsevier, 2012.
- [9] D. T. Larose and C. D. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining*, Secon Edit., vol. 53, no. 9. Canada: John Wiley & Sons, Inc., Hoboken, New Jersey, 2014.
- [10] S. Garcia, J. Luengo, and F. Herrera, *Data Preprocessing in Data Mining*. Switzerland: Springer, 2015.
- [11] M. Rönkkö, J. Heikkinen, V. Kotovirta, and V. Chandrasekar, “Automated preprocessing of environmental data,” *Futur. Gener. Comput. Syst.*, vol. 45, pp. 13–24, 2015.
- [12] D. Pyle, *Data Preparation for Data Mining.doc*. San Fransisco, USA: Morgan Kaufmann, 1999.
- [13] S. Ramírez-Gallego, B. Krawczyk, S. García, M. Woźniak, and F. Herrera, “A survey on data preprocessing for data stream mining: Current status and future directions,” *Neurocomputing*, vol. 239, pp. 39–57, 2017.
- [14] D. T. Larose, *Discovering Knowledge In Data*, 1st ed. Ca: John Wiley & Sons, Inc., Hoboken, New Jersey, 2005.
- [15] P. P. Das, R. Bisoi, and P. K. Dash, “Data decomposition based fast reduced kernel extreme learning machine for currency exchange rate forecasting and trend analysis,” *Expert Syst. Appl.*, vol. 96, pp. 427–449, 2017.
- [16] X. Cheng and H. Liu, “Research on data preprocessing technology in safety equipment linkage system,” *Proc. - 2013 Int. Conf. Comput. Inf. Sci. ICCIS 2013*, pp. 1713–1716, 2013.
- [17] S. Sharma and A. Bhagat, “Data Preprocessing Algorithm for Web Structure Mining,” *Int. Conf. Eco-Friendly Comput. Commun. Syst.*, vol. Fifth, pp. 1–5, 2016.
- [18] S. K. Dwivedi and B. Rawat, “A review paper on data preprocessing: A critical phase in web usage mining process,” *Proc. 2015 Int. Conf. Green Comput. Internet Things*, pp. 506–510, 2015.
- [19] S. Piramuthu, “On preprocessing data for financial credit risk evaluation,” *Expert Syst. Appl.*, vol. 30, no. 3, pp. 489–497, 2006.
- [20] S. García, J. Luengo, and F. Herrera, “Tutorial on practical tips of the most influential data preprocessing algorithms in data mining,” *Knowledge-Based Syst.*, vol. 98, pp. 1–29, 2016.
- [21] I. Suryani and R. S. Wahono, “Penerapan Exponential Smoothing untuk Transformasi Data dalam Meningkatkan Akurasi Neural Network pada Prediksi Harga Emas,” *J. Intell. Syst.*, vol. 1, no. 2, pp. 67–75, 2015.
- [22] V. Steinwandter, M. Šišmiš, P. Sagmeister, U. Bodenhofer, and C. Herwig, “Multivariate analytics of chromatographic data: Visual computing based on moving window factor models,” *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.*, vol. 1092, pp. 179–190, 2018.
- [23] R. Dash, “Performance analysis of a higher order neural network with an improved shuffled frog leaping algorithm for currency exchange rate prediction,” *Appl. Soft Comput. J.*, vol. 67, pp. 215–231, 2018.