

PAPER NAME

Ghani-Template_Jurnal_Teknika.pdf

WORD COUNT

2714 Words

CHARACTER COUNT

16206 Characters

PAGE COUNT

5 Pages

FILE SIZE

206.3KB

SUBMISSION DATE

Jun 20, 2022 10:00 AM GMT+7

REPORT DATE

Jun 20, 2022 10:00 AM GMT+7

● 19% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 14% Internet database
- 7% Publications database
- Crossref database
- Crossref Posted Content database
- 11% Submitted Works database

● Excluded from Similarity Report

- Bibliographic material

Pendekatan Deep Learning untuk Prediksi Durasi Perjalanan

Nur Ghaniaviyanto Ramadhan^{1*}, Yohani Setiya Rafika Nur², Faisal Dharma Adhinata³

^{1*}Department of Software Engineering, Institut Teknologi Telkom Purwokerto, Purwokerto, Jawa Tengah

²Department of Informatic, Institut Teknologi Telkom Purwokerto, Purwokerto, Jawa Tengah

³Department of Software Engineering, Institut Teknologi Telkom Purwokerto, Purwokerto, Jawa Tengah

Email: ^{1*}ghani@ittelkom-pwt.ac.id, ²yohani@ittelkom-pwt.ac.id, ³faisal@ittelkom-pwt.ac.id

(Naskah masuk: dd mmm yyyy, direvisi: dd mmm yyyy, diterima: dd mmm yyyy)

Abstrak

Setiap orang dalam kehidupan memiliki kecenderungan untuk berpindah dari satu tempat ke tempat lainnya. Perpindahan tersebut dapat dilakukan dengan menggunakan berbagai macam cara seperti menggunakan transportasi pribadi atau umum (bis, taxi, pesawat, dan kereta api). Pada perkembangan teknologi saat ini mode transportasi sudah semakin canggih. Akan tetapi masih ada mode transportasi yang belum modern misalnya seperti taxi, yang dimana salah satunya tidak dapat memprediksi lama waktu perjalanan. Meskipun sudah ada taxi yang berbasis online seperti Uber, akan tetapi masih banyak taxi yang belum berbasis online sehingga tidak bisa dilakukan estimasi waktu dan jarak. Permasalahan diatas dapat diselesaikan dengan cara melakukan pendekatan berbasis pembelajaran mesin. Salah satu keuntungan yang didapatkan jika kita dapat mengetahui lama waktu estimasi perjalanan yaitu dapat mengatur waktu perjalanan sesuai dengan rutinitas yang sedang dikerjakan ataupun juga dapat menghemat biaya yang dikeluarkan dengan mengetahui jarak yang akan dijalankan. Pada penelitian ini bertujuan untuk memprediksi durasi perjalanan menggunakan pendekatan deep learning yaitu Long Short Term Memory Recurrent Neural Network (LSTM-RNN). Eksperimen dilakukan dengan melakukan tuning parameter terkait seperti epoch, nilai dropout, dan neurons. Pengukuran hasil menggunakan nilai Root Mean Square Error (RMSE) dan nilai loss. Hasil yang didapatkan menggunakan model LSTM-RNN sebesar 0.0012 untuk nilai loss dan RMSE 0.4.

Kata Kunci: Deep Learning, Durasi Perjalanan, RMSE, Fungsi Loss

Deep Learning Approach for Trip Duration Prediction

Abstract

Everyone in life tends to move from one place to another. The transfer can be done using various ways, such as using private or public transportation (buses, taxis, planes, and trains). However, there are still modes of transportation that are not yet modern, such as taxis, one of which cannot predict the length of the trip. Although there are already online-based taxis such as Uber, there are still many taxis that are not online-based so it is impossible to estimate time and distance. The above problems can be solved by using a machine learning-based approach. One of the advantages if we can know the estimated time of travel is that we can adjust the travel time according to the routine that is being worked on or can also save costs incurred by knowing the distance to be run. This study aims to predict the duration of the trip using a deep learning approach, namely Long Short Term Memory Recurrent Neural Network (LSTM-RNN). Experiments are carried out by tuning related parameters such as epochs, dropout values, and neurons. Measurement of results using the value of Root Mean Square Error (RMSE) and loss value. The results obtained using the LSTM-RNN model are 0.0012 for the loss value and RMSE 0.4.

Keywords: Deep Learning, Trip Duration, RMSE, Loss Function.

I. PENDAHULUAN

Setiap orang dalam kehidupan memiliki kecenderungan untuk berpindah dari satu tempat ke tempat lainnya. Perpindahan tersebut dapat dilakukan dengan menggunakan berbagai macam cara seperti menggunakan transportasi pribadi atau umum (bis, taxi, pesawat, dan kereta api), Pada perkembangan teknologi saat ini mode transportasi sudah semakin canggih. Akan tetapi masih ada mode transportasi yang belum modern misalnya seperti taxi, yang dimana salah satunya tidak dapat memprediksi lama waktu perjalanan. Meskipun sudah ada taxi yang berbasis online seperti Uber, akan tetapi masih banyak taxi yang belum berbasis online sehingga tidak bisa dilakukan estimasi waktu dan jarak.

Permasalahan diatas dapat diselesaikan dengan cara melakukan pendekatan berbasis pembelajaran mesin. Salah satu keuntungan yang didapatkan jika kita dapat mengetahui lama waktu estimasi perjalanan yaitu dapat mengatur waktu perjalanan sesuai dengan rutinitas yang sedang dikerjakan ataupun juga dapat menghemat biaya yang dikeluarkan dengan mengetahui jarak yang akan dijalanakan

Beberapa penelitian sebelumnya yang sudah dilakukan untuk memprediksi waktu tempuh perjalanan, seperti pada paper berikut. Penulis melakukan prediksi waktu perjalanan menggunakan model deep learning yaitu XGBoost dan MLP [1]. Pada penulis lain melakukan study tentang memperkirakan tujuan perjalanan berdasarkan lintasan awal dengan menggunakan model klasifikasi long short-term memory dan embedding technology [2]. Penulis lain melakukan prediksi waktu perjalanan secara statis dari tempat satu ke tempat lainnya dengan menerapkan model regresi XGBoost [3].

Penelitian lainnya membahas tentang prediksi waktu perjalanan taksi Brunswick berdasarkan lintasan data mobil dengan menggunakan model XGBoost dan random forest [4]. Penelitian ini menyajikan sistem model untuk prediksi tujuan perjalanan dengan data GPS multi-hari dengan membangun model rantai markov dan multinomial logit [5]. Pada penelitian ini melakukan prediksi waktu pengiriman untuk layanan pos yang mana dilakukan percobaan beberapa metode seperti regresi linier dan tree ensembles, bagging, dan boosting [6].

Penelitian lainnya melakukan prediksi jangka panjang dengan pandangan beberapa hari dan prediksi jangka pendek dengan pandangan satu jam. Kedua perspektif tersebut relevan untuk tugas-tugas perencanaan dalam konteks mobilitas perkotaan dan layanan transportasi [7]. Pada penelitian ini akan dilakukan pengembangan model prediksi tujuan perjalanan berdasarkan layanan pencarian dan penemuan berbasis lokasi online (khususnya, Google Places API) dan kumpulan data perjalanan terbatas yang biasanya tersedia setelah perjalanan selesai [8].

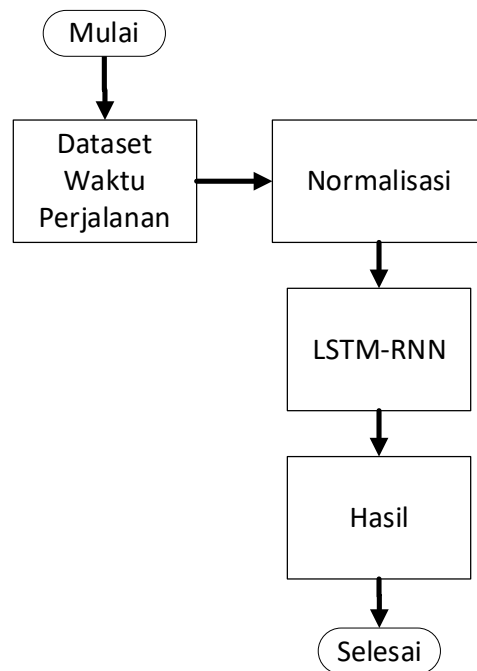
Pada makalah ini dilakukan prediksi round trip time berbasis pembelajaran mendalam menggunakan Recurrent Neural Networks (RNN) [9]. Dalam paper ini, penulis melakukan eksplorasi model pembelajaran mesin yang mendalam yaitu model jaringan saraf LSTM yang dilakukan untuk prediksi waktu perjalanan [10]. Studi ini mengevaluasi metode penyediaan informasi pra-perjalanan yang memprediksi waktu perjalanan dengan mencocokkan

akumulasi masa lalu dan data lalu lintas pada hari perjalanan dengan tetap memperhitungkan waktu yang dibutuhkan untuk mencapai tujuan [11].

Paper berikut ini melakukan penelitian tentang prediksi waktu kedatangan bus dan desain ulang jadwal dengan menggunakan penggabungan model Support Vector Regression (SVR) dan Kalman Filter [12]. Penelitian ini memiliki tujuan untuk menganalisis variasi waktu perjalanan bus dan memprediksi waktu perjalanan secara akurat dengan penerapan model SVR [13]. Penelitian ini mengusulkan suatu metode berdasarkan setiap zona bangunan, situasi penggunaan lahan, dan aksesibilitas bus untuk meramalkan arus perjalanan penumpang bus pada periode mendatang [14]. Metode yang digunakan yaitu Artificial Neural Network (ANN) [14]. Penelitian ini untuk memprediksi durasi perjalanan sewa sepeda di Seoul Bike sharing system [15]. Prediksi dilakukan dengan kombinasi data Seoul Bike dan data cuaca dengan beberapa teknik pada data mining yaitu regression, gradient, KNN, dan Random Forest [15]. Penelitian lain menerapkan model klasifikasi pohon keputusan pada kumpulan data perjalanan taxi [16].

Berdasarkan pemaparan permasalahan yang ada pada penelitian sebelumnya beserta metode-metode yang pernah digunakan, sehingga pada penelitian ini akan melakukan prediksi waktu perjalanan dari satu titik ke titik lainnya pada moda transportasi dengan menggunakan model deep learning Long-Short Term Memory Recurrent Neural Network (LSTM-RNN).

II. METODE PENELITIAN



Gambar 1. Desain Proses

A. Dataset

Penelitian ini menggunakan dataset yang berasal dari rekam waktu perjalanan menggunakan transportasi taxi [17]. Dataset terdiri dari 11 kolom dan 10.485.576 baris data.

Tabel 1. Karakteristik Dataset

No	Nama Kolom	Type Data
1	ID	Integer
2	Vendor ID	Integer
3	Pickup Datetime	Date
4	Dropoff Datetime	Date
5	Passenger Count	Integer
6	Pickup Longitude	Integer
7	Pickup Latitude	Integer
8	Dropoff Longitude	Integer
9	Dropoff Latitude	Integer
10	Store	String
11	Trip Duration	Integer

Pada dataset Tabel 1 memiliki fitur-fitur yang terkait dengan durasi perjalanan seperti *pickup datetime* dan *dropoff datetime*. Dataset tersebut berisikan informasi-informasi terkait perjalanan, jumlah penumpang, ketinggian lokasi penjemputan, dan ketinggian lokasi *dropoff* penumpang. Hal tersebut berguna untuk menentukan proses prediksi waktu yang ditempuh.

B. Normalisasi

Proses ini dilakukan normalisasi menggunakan formula min-max (1) [18]. Normalisasi bertujuan untuk menyamakan nilai pada setiap fitur antara nilai 0-1.

$$N^* = \frac{N - \min(n)}{\max(n) - \min(n)} \quad (1)$$

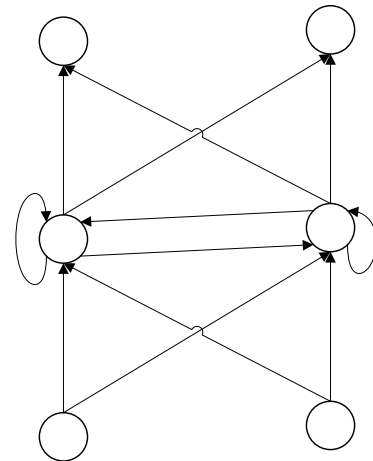
Dimana n yaitu data yang belum dinormalisasi. $\min(n)$ merupakan nilai minimum dari semua data dan $\max(n)$ adalah nilai maksimum dari semua data.

C. LSTM-RNN

Model LSTM-RNN merupakan model yang termasuk ke dalam pendekatan *deep learning*. LSTM-RNN dapat diaplikasikan ke berbagai permasalahan seperti *early learning tasks* dan *cognitive learning tasks* (*speech recognition*, *machine translation*, dan *image processing*). LSTM dapat mempelajari cara menjembatani jeda waktu minimal lebih dari 1.000 langkah waktu diskrit [18]. Solusinya menggunakan korsel kesalahan konstan (CEC) yang menegakkan aliran kesalahan konstan dalam sel khusus. Akses ke sel ditangani oleh unit gerbang multiplikasi yang mempelajari kapan harus memberikan akses.

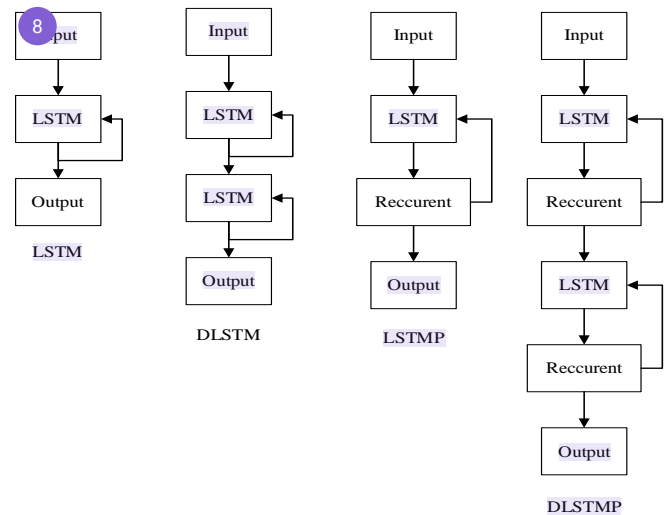
Recurrent Neural Networks (RNN) adalah sistem dinamis yang memiliki keadaan internal pada setiap langkah waktu klasifikasi [18]. Hal ini disebabkan lingkaran koneksi antara neuron lapisan yang lebih tinggi dan lebih rendah dan umpan

balik mandiri opsional koneksi. Koneksi umpan balik ini memungkinkan RNN untuk menyebarkan data dari peristiwa sebelumnya ke langkah pemrosesan saat ini. Jadi, RNN membangun memori waktu acara seri. Untuk arsitektur *basic* dari RNN dapat dilihat pada Gambar 2.



Gambar 2. RNN Arsitektur

Sedangkan untuk arsitektur model LSTM-RNN penelitian ini dapat dilihat pada Gambar 3 yang mana menggabungkan kedua arsitektur dari LSTM dan RNN [19].



Gambar 3. Arsitektur LSTM-RNN

Untuk formula dari LSTM dapat dilihat pada formula (2) (3) (4) dan (5).

$$cf_t = \sigma_1 (W_{cf} \cdot [O_{t-1}, x_t] + b_{cf}) \quad (2)$$

$$I_t = \sigma_2 (W_t \cdot [O_{t-1}, x_t] + b_t) \quad (3)$$

$$S_t = \tanh (W_s \cdot [O_{t-1}, x_t] + b_s) \quad (4)$$

$$S_t = cf_t \times S_{t-1} + I_t \times S_t \quad (5)$$

Dimana informasi baru yang akan disimpan dalam keadaan sel dihitung menggunakan dua lapisan jaringan. Lapisan sigmoid (σ) memutuskan nilai yang akan diperbarui (I_t) (3) dan lapisan tanh 1 yang mengembangkan vektor dari nilai kandidat baru (S_t) seperti yang ditunjukkan pada (4). Kombinasi untuk menambahkan status. Akhirnya, status sel diperbarui menggunakan (5).

Untuk menghitung hasil dari model LSTM-RNN menggunakan nilai *Root Mean Square Error* (RMSE) dan fungsi *loss*. Formula RMSE dapat dilihat pada (6) [18] dan fungsi *loss* (7) [18].

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_{ii} - y_i)^2}{n}} \quad (6)$$

Dimana y_{ii} dan y_i masing-masing adalah nilai prediksi dan aktual dari beban kerja pada saat sampel i dan N adalah jumlah data sampel.

$$Loss(p, q) = \frac{1}{|M|} \sum_{i=1}^{i=M} -q_i \log p_i \quad (7)$$

Di mana p dan q mewakili label yang diprediksi dan label yang sebenarnya, dan M menunjukkan ukuran batch.

III. HASIL DAN ANALISA

Eksperimen yang dilakukan pada penelitian ini yaitu dengan melakukan tuning beberapa parameter seperti epoch dan nilai dropout. Hasil eksperimen dapat dilihat pada Tabel 2.

Tabel 2. Hasil Eksperimen

Epoch	Neurons	Dropout	RMSE	Loss
100	25	0.1	0.4	0.0013
200	15	0.2	0.4	0.0013
500	35	0.3	0.4	0.0012
1000	55	0.7	0.4	0.0012

Berdasarkan hasil eksperimen Tabel 2 memperlihatkan bahwa nilai RMSE dan loss pada penelitian ini sangat stabil di angka 0.4 untuk nilai RMSE dan 0.12-0.13 untuk loss function. Proses tuning parameter dilakukan seperti pada nilai epoch awalnya bernilai 100 sampai dengan epoch 1000. Untuk nilai neurons dan nilai dropout juga dilakukan perubahan nilai semakin naik guna melihat hasil RMSE dan loss apakah ada perubahan. Namun, berdasarkan hasil tuning parameter untuk model LSTM-RNN pada permasalahan prediksi waktu perjalanan sangat tepat digunakan.

Pada Tabel 3 merupakan contoh hasil dari prediksi durasi perjalanan yang dihasilkan.

Tabel 3. Hasil Prediksi

Trip Duration Test (Minutes)	Trip Duration Predict (Minutes)
455	436
663	647
2124	2009
429	399
435	404
443	431
341	337
1551	1479

Berdasarkan hasil prediksi Tabel 3 terlihat bahwa nilai prediksi durasi yang dihasilkan mendekati waktu tempuh aslinya. Misalnya contoh pada trip duration test dengan waktu 435, hasil prediksinya yaitu yaitu 404. Maka hal tersebut juga berdampak pada nilai RMSE dan loss yang dihasilkan. Jika hasil prediksi sangat jauh dari nilai testing maka untuk nilai RMSE dan loss akan bernilai cukup besar, Nilai RMSE dan loss dikatakan baik jika yang dihasilkan nilainya semakin kecil.

IV. KESIMPULAN

Berdasarkan hasil eksperimen yang telah dilakukan dan nilai yang dihasilkan. Pada penelitian ini yang awalnya memiliki tujuan untuk memprediksi durasi perjalanan dari satu titik ke titik lainnya dengan moda transportasi umum taxi menggunakan model LSTM-RNN disimpulkan sangat tepat menggunakan pendekatan model LSTM-RNN. Hal tersebut dapat dilihat dari nilai RMSE yang dihasilkan sebesar 0.4 dan nilai loss sebesar 0.12. Sehingga metode deep learning LSTM-RNN ini sangat cocok digunakan pada kasus permasalahan prediksi dengan data yang memiliki jumlah sangat besar, akan tetapi dari proses eksperimen untuk waktu nya jika nilai epoch semakin besar maka waktu yang dibutuhkan semakin lama. Perbedaan waktunya dapat mencapai 30 detik untuk kelipatan nilai epoch 200. Untuk parameter yang tidak berpengaruh signifikan yaitu neurons.

REFERENSI

- [1] Poongodi, M., Malviya, M., Kumar, C., Hamdi, M., Vijayakumar, V., Nebhen, J., & Alyamani, H. (2021). New York City taxi trip duration prediction using MLP and XGBoost. *International Journal of System Assurance Engineering and Management*, 1-12.
- [2] Tang, J., Liang, J., Yu, T., Xiong, Y., & Zeng, G. (2021). Trip destination prediction based on a deep integration network by fusing multiple features from taxi trajectories. *IET Intelligent Transport Systems*.
- [3] Kankanamge, K. D., Witharanage, Y. R., Withanage, C. S., Hansini, M., Lakmal, D., & Thayasivam, U. (2019, July). Taxi trip travel time prediction with isolated XGBoost regression. In 2019 Moratuwa Engineering Research Conference (MERCon) (pp. 54-59). IEEE.

-
- [4] Fiosina, J. (2021). Explainable Federated Learning for Taxi Travel Time Prediction. In VEHITS (pp. 670-677).
- [5] Zong, F., Tian, Y., He, Y., Tang, J., & Lv, J. (2019). Trip destination prediction based on multi-day GPS data. *Physica A: Statistical Mechanics and its Applications*, 515, 258-269.
- [6] Khiari, J., & Olaverri-Monreal, C. (2020, November). Boosting Algorithms for Delivery Time Prediction in Transportation Logistics. In 2020 International Conference on Data Mining Workshops (ICDMW) (pp. 251-258). IEEE.
- [7] Huang, H., Pouls, M., Meyer, A., & Pauly, M. (2020, September). Travel time prediction using tree-based ensembles. In International Conference on Computational Logistics (pp. 412-427). Springer, Cham.
- [8] Ermagun, A., Fan, Y., Wolfson, J., Adomavicius, G., & Das, K. (2017). Real-time trip purpose prediction using online location-based search and discovery services. *Transportation Research Part C: Emerging Technologies*, 77, 96-112.
- [9] Dong, A., Du, Z., & Yan, Z. (2019). Round trip time prediction using recurrent neural networks with minimal gated unit. *IEEE Communications Letters*, 23(4), 584-587.
- [10] Duan, Y., Yisheng, L. V., & Wang, F. Y. (2016, November). Travel time prediction with LSTM neural network. In 2016 IEEE 19th international conference on intelligent transportation systems (ITSC) (pp. 1053-1058). IEEE.
- [11] Warita, H., Morita, H., Kuwahara, M., Chung, E., & Tanaka, A. (2004). Travel time prediction for pre-trip information using latest traffic conditions on expressway. *International Journal of Intelligent Transportation Systems Research*, 2(1), 11-19.
- [12] Zhang, X., Yan, M., Xie, B., Yang, H., & Ma, H. (2021). An automatic real-time bus schedule redesign method based on bus arrival time prediction. *Advanced Engineering Informatics*, 48, 101295.
- [13] Bachu, A. K., Reddy, K. K., & Vanajakshi, L. (2021). Bus travel time prediction using support vector machines for high variance conditions. *Transport*, 36(3), 221-234.
- [14] Yu, S., Shang, C., Yu, Y., Zhang, S., & Yu, W. (2016). Prediction of bus passenger trip flow based on artificial neural network. *Advances in Mechanical Engineering*, 8(10), 1687814016675999.
- [15] Sathishkumar, V. E., Park, J., & Cho, Y. (2020). Seoul bike trip duration prediction using data mining techniques. *IET Intelligent Transport Systems*, 14(11), 1465-1474.
- [16] Ismaeil, Hadeer, Sherif Kholeif, and Manal A. Abdel-Fattah. "Using Decision Tree Classification Model to Predict Payment Type in NYC Yellow Taxi." *International Journal of Advanced Computer Science and Applications* 13.3 (2022).
- [17] Kaggle. New York Taxi Trip Duration. Available online: <https://www.kaggle.com/yasserh/nyc-taxi-trip-duration>, Accessed: 28 Desember 2021.
- [18] Staudemeyer, Ralf C., and Eric Rothstein Morris. "Understanding LSTM--a tutorial into long short-term memory recurrent neural networks." arXiv preprint arXiv:1909.09586 (2019).
- [19] Ramadhan, Nur Ghaniaviyanto, Nia Annisa Ferani Tanjung, and Faisal Dharma Adhinata. "Implementation of LSTM-RNN for Bitcoin Prediction." *Indonesia Journal on Computing (Indo-JC)* 6.3 (2021): 17-24.

● **19% Overall Similarity**

Top sources found in the following databases:

- 14% Internet database
- 7% Publications database
- Crossref database
- Crossref Posted Content database
- 11% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Surabaya University on 2020-12-11 Submitted works	2%
2	journals.ukitoraja.ac.id Internet	1%
3	ejurnal.stmik-budidarma.ac.id Internet	1%
4	Universitas Brawijaya on 2021-02-16 Submitted works	1%
5	Ridwan Pandiya, Widodo Widodo, Salmah, Irwan Endrayanto. "Non par... Crossref	1%
6	Universitas Amikom on 2021-01-23 Submitted works	<1%
7	academic-accelerator.com Internet	<1%
8	Suman Maria Tony, S. Sasikumar. "Chapter 43 Music Generation Using... Crossref	<1%

9	es.scribd.com Internet	<1%
10	garuda.kemdikbud.go.id Internet	<1%
11	pt.scribd.com Internet	<1%
12	doku.pub Internet	<1%
13	Sriwijaya University on 2021-03-22 Submitted works	<1%
14	Risa Eri Susanti, Arif Wirawan Muhammad, Wahyu Adi Prabowo. "Imple... Crossref	<1%
15	journal.ipm2kpe.or.id Internet	<1%
16	mes-journal.ru Internet	<1%
17	readbag.com Internet	<1%
18	Muhammad Lulu Latif Usman, Muhamad Azrino Gustalika. "Pengujian ... Crossref	<1%
19	Surabaya University on 2016-05-31 Submitted works	<1%
20	id.123dok.com Internet	<1%

21	id.scribd.com Internet	<1%
22	jurnal.ugm.ac.id Internet	<1%
23	link.springer.com Internet	<1%
24	publishing-widyagama.ac.id Internet	<1%
25	repository.its.ac.id Internet	<1%
26	Politeknik Negeri Bandung on 2020-02-17 Submitted works	<1%
27	(Jeff) Ban, X.. "Real time queue length estimation for signalized interse... Crossref	<1%
28	Jinjun Tang, Jian Liang, Tianjian Yu, Yong Xiong, Guoliang Zeng. "Trip ... Crossref	<1%