Vol. 1, No.10, Juni 2024, pp.147~155

JEIS (JOURNAL ENGINEERING IN INDUSTRIAL SYSTEMS)

Journal Homepage:http://journal-live.ithb.ac.id/EIS
e-ISSN: xxxx-xxxx, DOI: xxxxx

Studi Awal Penerapan Reinforcement Learning pada Penyelesaian Heterogeneous Vehicle Routing Problem with Soft Time Windows

Nadine Desyani[1], Sonna Kristina[2], Vina S. Yosephine[3].

[1]Program Studi Teknik Industri
Institut Teknologi Harapan Bangsa, Bandung, Indonesia
nadinecrss@gmail.com

[2]Program Studi Teknik Industri
Institut Teknologi Harapan Bangsa, Bandung, Indonesia
sonna@ithb.ac.id

[3]Program Studi Teknik Industri
Institut Teknologi Harapan Bangsa, Bandung, Indonesia
vinayosephine@ithb.ac.id

nadinecrss@gmail.com

INFO ARTIKEL

Sejarah artikel:
Diterbitkan 27 Maret 2024

Kata kunci:
vehicle routing problem with time windows;
heterogeneous vehicle routing problem with soft time windows; *reinforcement learning, machinelearning*.

ABSTRAK

Sistem distribusi yang efisien menjadi penting bagi perusahaan karena dapat meminimasi pengeluaran biaya dalam proses transportasi. Salah satu caranya adalah dengan menentukan rute transportasi atau dikenal dengan vehicle routing problem (VRP), sebuah ilmu optimasi yang paling banyak dipelajari. VRP biasanya dapat diselesaikan dengan linear programming lewat bantuan aplikasi LINGO. Penelitian ini akan menerapkan Reinforcement Learning (RL) ke dalam heterogeneous vehicle routing problem with soft time windows (HVRPSTW). Penggunaan RL dapat memperoleh insight dari agent yang berinteraksi dengan lingkungannya untuk mencapai suatu tujuan lalu mampu menangani data yang besar dan beragam, serta menarik kesimpulan antar kumpulan data bahkan dalam situasi yang kompleks untuk terus melakukan perbaikan berkelanjutan. Secara definisi RL merupakan bagian dari artificial intelligence dan machine learning, yang berfokus pada integrasi antar statistik, optimasi dan subjek matematika lainnya. Hasil penelitian ini bahwa model RL dapat menyelesaikan HVRPSTW.

*Nadinecrss@gmail.com

1. PENDAHULUAN

Sistem distribusi merupakan bagian penting pada perusahaan, khususnya dalam kegiatan penyaluran barang hasil produksi ke tangan konsumen. Menurut Philip dan Gary, 2011 [1], distribusi dapat diartikan sebagai rangkaian aktivitas yang dilakukan secara berulang yang berhubungan dengan pemasaran produk. Sistem distribusi menjadi penting untuk menjamin produk yang dipasarkan tersedia di pasar. Umumnya setiap perusahaan mempunyai divisi distribusi tersendiri atau dapat menggunakan jasa pihak ke-tiga untuk mendistribusikan produk mereka. Masalah yang sering dihadapi oleh perusahaan adalah efisiensi, sedangkan efisiensi dalam pengiriman dapat meminimasi biaya yang dikeluarkan pada saat proses, contohnya minimasi biaya bahan bakar. Untuk mencapai pengiriman yang efisien bergantung pada pemilihan rute.

Masalah penentuan rute dikenal dengan istilah Vehicle Routing Problem (VRP), biasa digunakan untuk melayani sejumlah pelanggan yang akan membantu memilih rute perjalanan sehingga jarak atau waktu yang dihasilkan adalah terpendek atau tercepat [2] VRP mempunyai banyak varian, salah satu contohnya adalah Capacitated VRP yaitu varian VRP dengan batasan kapasitas [3]. Contoh lainnya adalah Time Windows VRP, time windows bisa diartikan sebagai batasan waktu beroperasinya konsumen, atau batasan waktu suatu kendaraan bisa mengunjungi konsumen [4]. Secara sederhana, pelanggan hanya bisa dilayani pada interval waktu tertentu. Dengan time windows yang berbeda untuk tiap konsumen maka penentuan rute akan menjadi semakin kompleks.

Pada kasus VRP yang paling sederhana sebuah kendaraan dengan kapasitas tertentu bertanggung jawab untuk mengirimkan barang kepada beberapa pelanggan, objektifnya untuk optimasi rute yang dimulai dan berakhir pada suatu node disebut depot, dengan tujuan untuk mendapatkan hasil maksimal, yang seringkali menghasilkan total jarak yang ditempuh atau total waktu pelayanan sebagai hasil yang diinginkan. Berkembangnya algoritma tanpa campur tangan manusia membuat reinforcement learning menjadi pilihan menarik yang berpotensi menjadi milestone dalam pendekatan penyelesaian VRP [5].

Salah satu varian VRP yang dapat diselesaikan dengan reinforcement learning adalah heterogeneous vehicle routing problem with soft time windows (HVRPSTW). Penelitian milik Wijaya [6] mengimplementasikan model matematis (HVRPSTW) ke dalam aplikasi LINGO.11 (trial version) untuk menentukan rute transportasi yang dapat meminimasi biaya transportasi dengan mengoptimalkan sumber daya kendaraan pada PT. XYZ yang memiliki dua jenis kendaraan, yaitu mobil dan motor Model matematis yang digunakan mempertimbangkan kapasitas kendaraan, time windows, fixed cost dan variabel cost. Variabel cost merupakan biaya yang akan berubah tergantung banyaknya produk atau jasa yang dihasilkan, seperti biaya penalti, biaya parkir, biaya bahan bakar, dan biaya tenaga kerja. Fixed cost merupakan biaya yang jumlahnya tetap tanpa dipengaruhi oleh banyak atau sedikit barang yang dijual, pada penelitian fixed cost didapatkan dari biaya depresiasi kendaraan dan biaya service rutin, lalu dibuat lebih sesuai dengan kondisi lapangan. Dengan hasil bahwa penelitian tersebut berhasil menerapkan model matematis HVRPSTW. Pada penelitian ini penulis menerapkan reinforcement learning (RL) untuk menyelesaikan masalah pada HVRPSTW dengan menggunakan konsep dasar dari reinforcement learning yaitu policy iteration dan value iteration. Pemilihan metode RL didasarkan pada kemampuannya untuk menemukan insight yang sebelumnya tidak diketahui, menangani data yang besar dan beragam, serta menarik kesimpulan antar kumpulan data bahkan dalam situasi yang kompleks untuk terus melakukan perbaikan berkelanjutan [7].

*Nadinecrss@gmail.com

2. DASAR TEORI

2.1. *Vehicle Routing Problem*

Vehicle Routing Problem (VRP) merupakan masalah umum yang terdapat pada bidang operational research yang dapat didefinisikan sebagai permasalahan pencarian rute pengiriman yang optimal dari satu atau beberapa depot dengan batasan yang telah ditentukan, objektifnya adalah untuk mencari rute perjalanan yang dapat memenuhi kebutuhan
dengan cost seminimal mungkin. Pengiriman bahan bakar, distribusi barang, pengiriman surat, merupakan contoh umum dari aplikasi VRP [8].

2.2. *Vehicle Routing Problem with Time Windows Vehicle Routing Problem with Time Windows*

Vehicle Routing Problem with Time Windows (VRPTW) pelanggan dilayani dalam rentang waktu tertentu. Objektifnya adalah untuk menentukan jumlah perjalanan minimal, dan minimasi total jarak tempuh atau durasi perjalanan. Setiap pelanggan hanya dapat dilayani oleh satu kendaraan dan pengiriman dimulai dari depot lalu akan berakhir kembali ke depot. Turunan dari VRP adalah Vehicle Routing Problem with Soft Time Windows (VRPSTW) merupakan relaksasi dari Vehicle Routing Problem with Hard Time Windows (VRPHTW), VRPSTW memiliki karakteristik dimana pelayanan terhadap pelanggan memiliki batas waktu yang fleksibel sehingga pelanggan dapat dilayani di luar batas time windows, tetapi perusahaan akan dikenakan denda jika pelanggan dilayani di luar batas time windows [9].

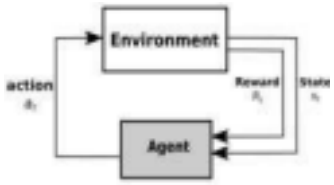
2.3. *Heterogeneous Vehicle Routing Problem with Soft Time Windows*

Heterogeneous Vehicle Routing Problem with Soft Time Windows (HVRPSTW) merupakan pengembangan dari VRP yang berupa gabungan dari HVRP dan VRPSTW, dengan mempertimbangkan kapasitas kendaraan yang bervariasi serta adanya fixed cost dan variabel cost dengan tujuan agar dapat mengetahui penggunaan kendaraan yang tepat sesuai dengan rute dan permintaan pelanggan dengan biaya transportasi yang minimal [10].

2.4. *Reinforcement Learning*

Reinforcement Learning (RL) adalah sebuah teknik yang diterapkan oleh smart agent untuk memecahkan masalah dalam lingkungan baru. RL mampu menemukan cara untuk menyelesaikan masalah dengan menghasilkan reward secara maksimal, caranya agent akan melakukan exploitation yang artinya pengambilan keputusan untuk melakukan sesuatu berdasarkan informasi yang tersedia, dan exploration untuk pengambilan keputusan dengan melakukan sesuatu yang baru [11]. Dalam RL ada beberapa istilah yang sering digunakan diantaranya adalah agent yaitu sebuah entitas yang melakukan action, sedangkan action adalah aksi yang dilakukan, lalu ada environment yang merupakan skenario atau lingkungan yang dihadapi oleh agent, reward yaitu umpan balik yang diberikan agent ketika melakukan action dan terakhir state yaitu keadaan saat ini. Konsep dari RL yang digunakan untuk melakukan penelitian ini adalah policy iteration yang dapat didefinisikan sebagai cara untuk mencari kebijakan yang optimal, yaitu memilih aksi terbaik pada proses iterasi dimulai dari mencari value function dan value iteration yang diartikan sebagai sebuah perhitungan yang bertujuan untuk memaksimalkan nilai pada setiap iterasi, Ada beberapa istilah yang sering dipakai dalam reinforcement learning, Agent yaitu sebuah entitas yang melakukan action, Action yaitu tindakan/aksi yang akan dilakukan oleh agent, Environment yaitu skenario atau lingkungan yang akan dihadapi oleh agent, Reward yaitu umpan balik yang diberikan pada agent ketika melakukan action, State (S) yaitu keadaan saat ini, Policy (π) yaitu strategi yang diterapkan oleh agent untuk memutuskan action selanjutnya, berdasarkan keadaan terkini, Value function yaitu nilai dari sebuah state yang akan menentukan total jumlah reward, Environment Model yaitu perilaku sebuah lingkungan/situasi, yang dapat membantu membuat kesimpulan.

*nadinecrss@gmail.com



Gambar 1 – Ilustrasi RL [12]

2.5. *Bellman Equation*

Tujuan dari penggunaan Bellman Equation adalah untuk memaksimasi atau minimasi suatu fungsi dari sebuah masalah, dengan cara memecah masalah menjadi sub-masalah dan meninjau bagaimana membuat keputusan terbaik. Bellman Equation akan menghasilkan solusi berupa policy, peluang sebuah agent memilih action dalam state tertentu untuk memilih value, seberapa bagus action yang dipilih oleh agent dalam state tertentu [13].

Notasi:

- a adalah action
- s adalah state
- R adalah reward
- γ adalah nilai gamma

1) Perumusan Bellman Equation

$$V(s) = \text{Max}a\,(R\,(s, a) + \gamma\, V(s')) \quad (1)$$

2) Perumusan Reward

$$\{(((xijk / Lk) \times BBk) + BPk)\} \quad (2)$$

3. METODOLOGI PENELITIAN

3.1. Model Acuan

Model acuan yang digunakan untuk menyelesaikan masalah pada penelitian ini adalah Heterogeneous Vehicle Routing Problem with Soft Time Windows (HVRPSTW). Pemilihan model ini didasarkan pada sistem distribusi PT. XYZ yang memiliki jenis kendaraan lebih dari satu dengan waktu pelayanan pelanggan yang berbeda namun fleksibel. Jadi pelanggan tetap bisa dilayani diluar waktu yang telah ditentukan namun ada biaya penalti yang dibebankan pada perusahaan. Penyelesaian permasalahan dilakukan pada HVRPSTW dengan bantuan RL, artinya akan terjadi proses trial and error sampai agent mendapatkan keputusan terbaik dalam menyelesaikan masalah. Berikut merupakan parameter dan variabel keputusan HVRPSTW [6].

1) Perumusan Model HVRPSTW

Notasi:

- i, j, p : Indeks Lokasi
- k : Indeks Kendaraan Pengiriman
- N : Seluruh lokasi
- Di : Demand permintaan pengiriman pada lokasi i
- Si : Waktu unloading pada lokasi i
- Fi : Batas waktu awal permintaan pengiriman lokasi i

*nadinecrss@gmail.com

- Yi : Batas waktu akhir permintaan pengiriman lokasi i
- Qk : Kapasitas maksimum kendaraan k
- K : Kendaraan pengiriman
- BBk : Biaya bensin untuk kendaraan k
- Lk : Kecepatan rata-rata kendaraan k
- CPk : Biaya penalti kendaraan k
- WWk : fixed cost untuk kendaraan k
- BPk : Biaya Parkir
- BT : Biaya Tenaga Kerja
- EPik : Lama waktu early penalti, lokasi I kendaraan k
- DPik : Lama waktu delay penalti, lokasi I kendaraan k
- Rij : Jarak antar lokasi i ke lokasi j
- Tijk : Waktu perjalanan lokasi i ke j, kendaraan k
- Aik : Waktu kendaraan k datang ke lokasi i
- Bik : Waktu keberangkatan kendaraan k dari lokasi i
- Cijk : Biaya kendaraan lokasi i ke lokasi j, kendaraan k
- TBPik : Lama kendaraan k parkir di lokasi i
- Uij : Variabel bantu eliminasi subtur

2) Fungsi Objektif

$$\text{Min } Z = \sum igN \sum j\in N \sum k\in K \text{ Xijk} * \text{Rij} * \text{BBk} + \sum i\in N \sum j\in N \sum k\in K \text{ Xijk} *((\text{DPjk} + \text{EPjk}) * \text{CPk}) + \sum k\in K \sum i\in N,i\leq 1 \sum j\in N,j\geq 1 \text{ Xijk} * \text{BT} + \sum k\in K \sum i\in N,i\leq 1 \sum j\in N,j\geq 1 \text{ Xijk} * \text{WWk} + \sum k\in K \sum i\in N,i\neq 1 \sum j\in N \text{ Xijk} * \text{TBPik*BPk} \quad (3)$$

3) Kendala

1. Pembatas jumlah maksimal kendaraan yang dapat melayani satu lokasi

$$\sum i\in N \sum k\in K \text{ Xijk} = 1, \forall j \in N, j \neq 1 \quad (4)$$

$$\sum j\in N \sum k\in K \text{ Xijk} = 1, \forall i \in N, i \neq 1 \quad (5)$$

2. Pembatas kekontinuan rute

$$\sum i\in N, \neq p \text{ Xipk} - \sum j\in N, j\neq p \text{ Xpjk} = 0, \forall k \in K, p \in N, p \neq 1 \quad (6)$$

3. Eliminasi subrute

$$\text{Uik-Ujk+N. Xijk} \leq \text{N-1}, \forall k \in K, i \in N, i \neq 1, j \in N, j \neq i \quad (7)$$

4. Pembatas jumlah muatan kendaraan

$$\sum i\in N, \neq 1 \sum j\in N \text{ Xijk} * \text{Di} \leq \text{Qk}, \forall k \in K \quad (8)$$

5. Pembatas bilangan biner

$$\text{Xiik} = 0, \forall k \in K, i \in N \quad (9)$$

*nadinecrss@gmail.com

$$Xijk = \{0,1\}, \forall k \in K, i \in N, j \in N \quad (10)$$

6. Pembatas waktu perjalanan

$$Tijk = Rij * 1/L, \forall k \in K, i \in N, j \in N \quad (11)$$

7. Pembatas waktu tiba dan keberangkatan kendaraan

$$Ajk \geq Bik + Tijk - (1-Xijk), \forall j \in N, j > 1, i \in N, k \in K \quad (12)$$

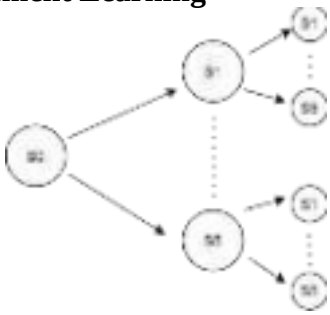
$$Ajk \leq Bik + Tijk + (1-Xijk), \forall j \in N, j > 1, i \in N, k \in K \quad (13)$$

8. Pembatas jendela waktu

$$Fi \leq Bik - S, \forall i \in N, i \geq 1, k \in K \quad (14)$$

$$Bik - Si \leq Y, \forall i \in N, i \geq 1, k \in K \quad (15)$$

3.2. Pendefinisian Model Reinforcement Learning



Gambar 2 – Ilustrasi Proses Perhitungan

RL bekerja dengan cara trial and error untuk mendapatkan sebuah solusi dengan cara mendapatkan reward untuk setiap action yang dilakukan. Reward didapatkan denganmengikuti suatu policy, yang menjadi dasar agent dalam memilih action. Pemilihan action ini dilakukan dengan policy iteration dan value iteration, policy iteration bertugas untuk mencari aksi terbaik yang dimulai dari value iteration. Keputusan pada policy iteration dan value iteration bergantung pada hasil yang akan didapatkan dari Bellman Equation, pada rumus tersebut terdapat empat notasi yang dapat didefinisikan.

RL a yaitu action dapat didefinisikan sebagai aksi dari agent dalam memilih rute atau pilihan pelanggan yang akan dikunjungi seperti pada gambar 2 dalam hal ini agent dapat memilih action ke pelanggan/node S1 yaitu a1 atau S2 yaitu a2, S3 yaitu a3, S4 yaitu a4 dan S5 yaitu a5. bermula dari S0 yaitu depot dengan kemungkinan pilihan pelanggan yang dapat dikunjungi S1-S5. Lalu S sebagai state adalah pelanggan pada kondisi terkini posisi kendaraan saat itu. Misalnya kendaraan yang bermula dari S0, memilih untuk mengunjungi S1, maka S1 itulah yang disebut state. Reward dimodelkan sebagai total cost, dengan memperhitungkan jarak dari i ke j, kecepatan kendaraan k, biaya bensin kendaraan k dan biaya parkir kendaraan k. asumsi nilai γ pada penelitian ini adalah 1 agar mendapatkan reward yang lebih maksimal.

3.3. Proses perhitungan menggunakan Reinforcement Learning

Proses perhitungan diilustrasikan pada gambar 3, perjalanan dimulai dari S0, lalu agent dapat memilih a=1, a=2, a=3…a=n, dengan state atau titik tujuan yang dituju adalah S1, S2, S3…Sn.

Lalu agent akan mendapatkan reward sebesar R1, R2, R3…Rn. Dalam setiap state terdapat V1, V2, V3… n. Perhitungan hasil V1… n didapatkan dari persamaan Bellman Equation (1), dengan menambahkan reward yang dilakukan action pada state tertentu dan menambahkan nilai gamma yang dikalikan dengan value state selanjutnya, reward dipilih yang paling maksimal sehingga akan menghasilkan value yang juga maksimal, asumsi agent dari S0 memilih action menuju S1 karena reward dan value yang dihasilkan adalah maksimal, agent dari S1 lalu memiliki pilihan action lainnya untuk melanjutkan ke titik selanjutnya atau kembali ke S0 yang berarti perhitungan berhenti. Lalu asumsikan kembali jika agent dari S1 memilih action menuju S2, dari S2 lalu memilih action menuju S3 hingga akhirnya kembali ke S0. Proses ini diulangi berkali-kali hingga agent mengerti action apa

*nadinecrss@gmail.com

yang harus dipilih demi mencapai hasil maksimal, tentu proses pengulangan ini bergantung pada jumlah titik yang dimiliki, jika ada tiga titik maka akan ada 6 kemungkinan, yang didapatkan dari 3! Sebagai contoh kemungkinan 1-2-3/1-3- 2/2-1-3/2-3-1/3-2-1/3-1-2. Setiap kemungkinan tersebut dilakukan dan diobservasi oleh agent mana yang dapat memaksimalkan reward yang diinginkan. Proses ini dilakukan hingga semua titik terkunjungi.

4. HASIL DAN PEMBAHASAN

Hasil berupa biaya transportasi yang dikeluarkan selama satu bulan dan berisi moda kendaraan yang akan digunakan. Data tersebut berisi nama pelanggan, demand tiap pelanggan, time windows setiap pelanggan, jarak antara pelanggan, dan kapasitas kendaraan. jumlah demand dalam satu bulan adalah 5263 box.

4.1. Parameter Kendaraan

PT. XYZ memiliki dua macam opsi kendaraan yaitu dua mobil dan empat motor, dengan kapasitas angkut motor 135 box dan mobil 946 box. Mobil diasumsikan berjalan konstan dengan kecepatan 30km/jam dan konsumsi bahan bakar 8km/liter, sedangkan motor berjalan dengan kecepatan 40km/jam dan konsumsi bahan bakar 28km/jam. Setiap kendaraan menggunakan bahan bakar jenis pertalite dengan harga Rp 7.650/liter. Waktu unloading adalah 0,022 menit untuk setiap satu box, biaya parkir tiap kali melakukan pengantaran, Rp 3.000/jam untuk mobil dan Rp 1.500/jam untuk motor. Lalu untuk fixed cost mobil menghabiskan Rp 82.366,66/pemakaian dan motor menghabiskan Rp 9.238/pemakaian, biaya penalti untuk mobul Rp 342.2/mnt dan untuk motor 38.7/mnt.

4.2. Proses Perhitungan

Pada proses perhitungan menggunakan dua kemungkinan aksi yang bertujuan untuk memetakan pilihan pelanggan yang dapat dikunjungi dengan cost terendah. Dua kemungkinan aksi ini berawal dari depot, lalu akan dihadapkan dengan opsi pelanggan pada hari itu, dan kemungkinan pelanggan yang dapat dikunjungi setelah pelanggan pertama selesai. Mobil akan bermula di S0 lalu mobil akan dihadapkan dengan berbagai kemungkinan pelanggan pada hari tersebut. menghitung cost dengan menggunakan rumus (8) setelah cost dihitung proyeksikan kemungkinan pelanggan dariawal. Ternyata penggunaan metode RL tidak se-optimal Lingo yang menyebabkan biaya lebih mahal karena pemilihan reward hanya berdasar dari total cost terendah, hasil yang didapatkan akan beda jika kebijakan penentuan reward juga diubah dengan berbagai kemungkinan.

Tabel 1 – Total Biaya Transportasi Selama Satu Bulan

	Total Jarak	Total Biaya Transportasi
RL	1.336.88 km	Rp 3.720.040
Lingo	1.442.5 km	Rp 3.540.119

*nadinecrss@gmail.com

5. KESIMPULAN

Pada penelitian ini sudah berhasil menerapkan model RL untuk menyelesaikan HVRPSTW dengan total biaya transportasi selama satu bulan Rp 3.724.876 dan total jarak tempuh 1.477,28 km. Hasil ini masih lebih tinggi Rp 184.757 dari segi biaya dan 34,78 km lebih tinggi dari segi jarak tempuh dibandingkan penelitian sebelumnya yaitu Rp 3.540.119 dan 1442,5 km dengan kata lain hasil belum mencapai optimal, ini disebabkan karena masih dilakukan satu kali iterasi yaitu pemilihan keputusan hanya berdasar pada total cost terendah, hasil akan mencapai optimal jika melakukan training pada RL. Lalu penggunaan RL pada HVRPSTW mampu menemukan insight yang didapatkan melalui proses pengulangan trial and error, insight tersebut berisi tentang kendaraan yang harus dipakai sesuai dengan demand dan rekomendasi action pada pelanggan yang sebaiknya dikunjungi untuk menghasilkan reward maksimal yang juga berdasarkan time windows dengan total cost terendah, sehingga dapat menghasilkan rute transportasi, dengan pilihan kendaraan yang tepat dan semua pelanggan terlayani dalam rentang waktu yang telah ditentukan. Usulan penelitian selanjutnya yang dapat dilakukan adalah melakukan training data pada RL agar hasil yang didapatkan mencapai optimum, proses trial and error harus lebih banyak dilakukan dengan opsi yang bervariatif,
sehingga insight yang didapatkan akan lebih banyak, disarankan untuk menerapkan RL menggunakan bantuan aplikasi lain seperti OR-Tools. Dan melakukan uji sensitivitas untuk hasil yang lebih baik dan mencoba menghitung dengan kemungkinan yang lebih banyak.

REFERENSI

[1] K. Philip dan A. Gary, Priciples of Marketing, 14th ed., New Jersey: Prentice Hall, 2011.
[2] P. Toth dan D. Vigo, The Vehicle Routing Problem, Philadelphia: Siam, 2019.
[3] A. Chandra, B. Setiawan, “Optimasi Jalur Distribusi dengan Metode Vehicle Routing Problem (VRP)”, Jurnal Manajemen Transportasi & Logistik Vol.5 No.2, 2018.
[4] S. Wahyuningsih, D. Satyananda, L. Octoviana, R. Nuhakiki, “Vehicle Routing Problem with Time Windows Variants and its Application in Distribution Optimization”, 2019.
[5] M. Nazari, A. Oroojlooy, L. Snyder, M. Takac, “Deep Reinforcement Learning for Solving the Vehicle Routing Problem”, 2018.
[6] H. Wijaya, Penyusunan Model Matematis Heterogeneous Vehicle Routing Problem with Soft Time Windows Untuk Menentukan Rute Transportasi Yang Dapat Meminimasi Biaya Transportasi, Bandung: ITHB, 2019.
[7] T. Wuest, D. Weimer, C. Irgens, K. Thoben, Machine Learning in Manufacturing: advantages, challenges, and applications ,2016.
[8] Y. Liong, K. Omar, “Vehicle Routing Problem: Models and Solutions”, JQMA Vol.4 No.1, 205-218, 2018.
[9] O. Braysy, M. Gendreau, “Vehicle Routing Problem with Time Windows, Part I: Route Construction and Local Search Algorithms”, Transportation Science Vol.39 No.1, 104-118, 2005.
[10] S. Kristina, R. Sianturi, V. J. Wijaya, Pengembangan Algoritma Ant Colony System pada Heterogeneous Vehicle Routing Problem with Soft Time Windows, Bandung: ITHB 2020.
[11] G. Xiaobing, W. Yan, L. Shuhai, N. Ben, “Vehicle Routing Problem with Time Windows and Simultaneous Delivery and Pick-Up Service Based on MCPSO”, Mathematical Problem in Engineering Vol 12, 1- 11, 2012.
[12] C. Koc, T. Bektas, O. Jabali, G. Laporte “Thirty years of heterogeneous vehicle routing”, European Journal of Operational Research, Vol.239 No.3, 2014.
[13] Jae Duk Seo, “Random notes for Bellman Equation,” medium, 2019. [Daring]. Tersedia: https://medium.com/@SeoJaeDuk/archived-post random-notes-for-bell-man-equation-b1d62a9038ec/. [Accessed 2021].

*nadinecrss@gmail.com