

PREDIKSI JUMLAH TENAGA KERJA ASING DI JAWA BARAT MENGGUNAKAN PERBANDINGAN ALGORITMA SUPPORT VECTOR REGRESSION DAN DECISION TREE REGRESSION

Farill Andika Wardana¹⁾, Safitri Juanita^{*2)}

1. Sistem Informasi, Fakultas Teknologi Informasi, Universitas Budi Luhur, Indonesia
2. Sistem Informasi, Fakultas Teknologi Informasi, Universitas Budi Luhur, Indonesia

Article Info

Kata Kunci: decision tree regression; jawa barat; prediksi tenaga kerja; support vector regression

Keywords: decision tree regression; labor prediction; support vector regression; west java

Article history:

Received 14 Oktober 2024

Revised 8 November 2024

Accepted 9 Desember 2024

Available online 15 March 2025

DOI :

<https://doi.org/10.29100/jupi.v10i2.6003>

* Corresponding author.

Safitri Juanita

E-mail address:

safitri.juanita@budiluhur.ac.id

ABSTRAK

Beberapa tahun ini, Indonesia sedang menghadapi masalah mengenai peningkatan jumlah tenaga kerja asing yang masuk ke Indonesia, salah satunya di provinsi Jawa Barat. Sehingga diperlukan model prediksi untuk memprediksi jumlah tenaga kerja asing yang masuk di provinsi Jawa Barat. Metode yang digunakan pada penelitian ini adalah CRISP-DM, dengan menggunakan dataset jumlah tenaga kerja asing di Jawa Barat periode 2014-2023, dan pada tahap pemodelan membandingkan 2 algoritma yaitu Decision Tree Regression (DTR) dan Support Vector Regression (SVR) dengan metode pengujian Cross-Validation (CV). Hasil pengujian performa kedua algoritma menggunakan Root Mean Squared Error (RMSE) dan Mean Absolute Error (MAE). Penelitian ini bertujuan untuk menemukan model peramalan untuk melakukan prediksi terhadap jumlah tenaga kerja asing yang masuk di provinsi Jawa Barat. Kesimpulan penelitian ini adalah model prediksi yang memiliki performa lebih unggul adalah Decision Tree Regression (DTR) dengan nilai RMSE sebesar 78.04% dan MAE sebesar 69.57%, sedangkan Support Vector Regression (SVR) hanya memiliki nilai RMSE sebesar 81.80% dan MAE sebesar 70.79%.

ABSTRACT

In recent years, Indonesia has been faced issues related to the increasing number of foreign workers entering the country, particularly in West Java province. Therefore, a predictive model is needed to forecast the number of foreign workers entering West Java province. The method used in this research is CRISP-DM, utilizing a dataset of the number of foreign workers in West Java from 2014 to 2023. In the modeling stage, two algorithms are compared: Decision Tree Regression (DTR) and Support Vector Regression (SVR) using Cross-Validation (CV) testing methods. The performance of both algorithms is evaluated using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). This research aims to find a forecasting model to predict the number of foreign workers entering West Java province. The conclusion of this study is that the superior predictive model is Decision Tree Regression (DTR), with an RMSE of 78.04% and an MAE of 69.57%, while Support Vector Regression (SVR) only has an RMSE of 81.80% and an MAE of 70.79%.

I. PENDAHULUAN

Perkembangan dunia global mendorong adanya peningkatan mobilitas penduduk dunia yang menimbulkan berbagai dampak dalam kepentingan serta kehidupan masyarakat Indonesia khususnya di provinsi Jawa Barat [1]. Indonesia saat ini sedang menghadapi masalah mengenai banyaknya tenaga kerja asing (TKA) yang masuk ke Indonesia dalam beberapa tahun terakhir [2]. Masuknya TKA di Indonesia juga memiliki dampak di sektor penanaman modal atau investasi karena biasanya ketika sebuah negara atau provinsi mendapatkan dana investasi dari negara lain, negara yang berinvestasi tersebut hampir pasti mengirimkan tenaga kerja dari negara yang berinvestasi dengan tujuan untuk memantau perkembangan investasi yang telah diberikan di negara yang diberikan investasi tersebut [3].

Salah satu provinsi di Indonesia yang mendapatkan dana investasi terbesar yaitu provinsi Jawa Barat. Jawa Barat mendapatkan investasi sebesar US\$8.28 miliar dari total US\$50.27 miliar investasi yang masuk ke Indonesia [4]. Selain itu, Jawa Barat merupakan provinsi terbesar ketiga yang memiliki distribusi tenaga kerja asing (TKA)

terbanyak di Indonesia pada bulan November 2022[5]. Masuknya TKA di Jawa Barat memiliki dampak seperti meningkatnya investasi modal karena adanya negara luar yang mengirimkan tenaga kerja berbarengan dengan modal investasi, lalu adanya transfer ilmu pengetahuan dan teknologi terbaru dari TKA menuju tenaga kerja lokal sehingga teknologi yang ditransfer tersebut bisa diadopsi dan digunakan oleh tenaga kerja lokal [6].

Dengan adanya peningkatan distribusi tenaga kerja asing (TKA) dengan investasi modal asing, maka diperlukan penelitian untuk menemukan model peramalan yang dapat melakukan prediksi terhadap jumlah TKA yang masuk di Jawa Barat. Sehingga dari hasil prediksi tersebut, instansi terkait dapat merumuskan kebijakan yang sesuai agar memberikan dampak yang bermanfaat bagi Masyarakat dan pertumbuhan ekonomi di provinsi Jawa Barat.

Beberapa penelitian sebelumnya yang menggunakan algoritma klasifikasi untuk peramalan atau prediksi beberapa studi kasus yang ada di masyarakat telah dilakukan, antara lain adalah perbandingan kinerja dua algoritma regresi decision tree dan regresi linear berganda untuk memprediksi indeks massa tubuh (BMI) berdasarkan dataset Asthma dengan variabel tambahan, penelitian ini membandingkan akurasi algoritma dalam bentuk MAPE, di mana algoritma regresi linear ganda lebih baik daripada algoritma regresi decision tree. Algoritma Regresi liner berganda mendapatkan nilai MAPE 12.73%; 12.76%; 12.89%; dan 12.99% untuk 1000 kali percobaan [7]. Kemudian peramalan harga saham INDF dan MYOR dengan menggunakan algoritma Pencarian Grid Research dan algoritma SVR selama satu periode ke depan, hasil penelitian menunjukkan bahwa saham INDF memiliki nilai MAPE dan R^2 sebesar 5.57% dan 79.90% pada data uji, sedangkan saham MYOR memiliki nilai MAPE dan R^2 sebesar 2.95% dan 96% pada data uji[8].

Penelitian lain adalah Perbandingan algoritma Simple Linear Regression (SLR) dan SVR dalam Prediksi Jumlah Penduduk di Sulawesi Tenggara, dengan menggunakan evaluasi matriks RMSE dan MAPE, hasil penelitian menunjukkan bahwa algoritma SLR unggul untuk memprediksi lebih banyak kabupaten/kota sedangkan algoritma SVR unggul dalam memprediksi hanya beberapa kota dengan jumlah yang sedikit. [9]. Kemudian penelitian tentang prediksi penyakit stroke dengan mengidentifikasi pada variabel yang ditemukan dan menganalisisnya untuk membuat prediksi dengan menggunakan algoritma Decision Tree C4.5 dan SVR, hasilnya menemukan bahwa algoritma Decision Tree C4.5 menghasilkan nilai Root Mean Square Error (RMSE) sebesar 0.23 lebih rendah daripada yang dihasilkan oleh SVR yaitu 0.39[10]. Penelitian lainnya, komparasi Algoritma Long Short-Term Memory (LSTM) dengan SVR pada prediksi harga saham di Indonesia menemukan bahwa algoritma LSTM mampu memprediksi harga saham dari tahun 2017 hingga 2019 dengan hasil yang lebih akurat namun memerlukan waktu proses lebih lama dari pada algoritma SVR [11]. Penelitian lainnya, pada perbandingan Algoritma Linear Regression (ALR), SVR, dan Artificial Neural Network (ANN) untuk Prediksi Data Obat menemukan algoritma SVR unggul jika dibandingkan algoritma lainnya untuk memprediksi tiga jenis obat yang diuji[12].

Berdasarkan penelitian sebelumnya, maka dapat disimpulkan penelitian ini memiliki kemiripan studi kasus pada penelitian [9] namun perbedaannya adalah pada penelitian [9] memprediksi jumlah penduduk di Sulawesi Tenggara, sedangkan penelitian ini memprediksi jumlah tenaga kerja asing (TKA) di Jawa Barat. Selain itu perbedaan lainnya jika dilihat dari kontribusi penelitian, penelitian [9] memiliki kontribusi untuk memberikan informasi tentang performa SLR dan SVR dalam memprediksi jumlah penduduk di Sulawesi Tenggara, yang dapat membantu pemerintah dalam membuat kebijakan terkait kependudukan. Sedangkan penelitian ini memiliki kontribusi untuk memberikan informasi tentang efektivitas SVR dan DTR dalam memprediksi tenaga kerja asing di Jawa Barat, yang dapat membantu pemerintah dalam membuat kebijakan terkait tenaga kerja asing.

Selain itu berdasarkan penelitian sebelumnya, algoritma SVR terbukti unggul dalam melakukan prediksi [8] [12] dibandingkan algoritma Grid Research [8] dan algoritma ALR[12]. Sedangkan algoritma DTR dipilih pada penelitian ini karena mampu melakukan prediksi dengan baik seperti pada penelitian Perbandingan Algoritma Multilinear Regression (MLR) dan Decision Tree Regressor (MLR) dalam Memprediksi Efisiensi Penghambatan Korosi Piridazin, dimana algoritma DTR unggul dan mampu memprediksi dengan baik dibandingkan algoritma MLR [13].

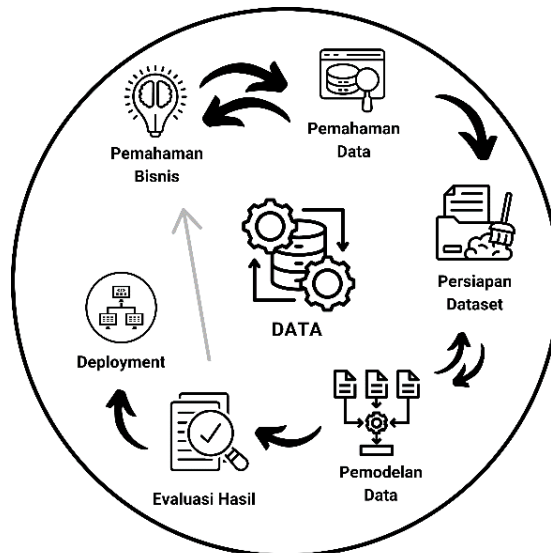
Pemilihan kedua algoritma SVR dan DTR juga merujuk pada karakteristik dataset penelitian ini, dimana Algoritma SVR cocok dengan dataset yang memiliki dimensi data tinggi serta memiliki karakteristik untuk memecahkan persamaan linear serta memberikan model pelatihan yang lebih cepat [14]. Sedangkan algoritma DTR memiliki keunggulan karena tidak perlunya normalisasi pada dataset dan hubungan non-linear antara fitur dan target [15]. Selain itu, kedua algoritma ini juga dipilih karena mewakili model regresi dalam supervised machine learning untuk data deret waktu.

Berdasarkan masalah dan studi literatur penelitian sebelumnya, maka kontribusi pada penelitian ini adalah melakukan peramalan atau memprediksi distribusi tenaga kerja asing (TKA) di provinsi Jawa Barat dengan membandingkan algoritma DTR dan SVR. Tujuan penelitian ini adalah untuk menemukan algoritma terbaik dengan membandingkan 2 algoritma (DTR dan SVR) yang dapat meramalkan atau memprediksi jumlah TKA di

Provinsi Jawa Barat agar dapat membantu pemerintah atau institusi terkait untuk membuat program atau kebijakan yang berkaitan dengan tenaga kerja supaya memiliki dampak yang positif kepada masyarakat dan perekonomian daerah Jawa Barat.

II. METODE PENELITIAN

Kerangka Penelitian yang berisi tahapan-tahapan penelitian menggunakan metodologi CRISP-DM [16]. Proses eksperimen menggunakan bahasa pemrograman Python yang diproses dengan perangkat Google Collab. Menggunakan pustaka `sklearn.svm` untuk metode SVR, pustaka `sklearn.tree` untuk metode DTR, pustaka `sklearn.model_selection` untuk cross-validation, dan pustaka `sklearn.metrics` untuk RMSE dan MAE. Pada Gambar 1 adalah kerangka penelitian.



Gambar. 1. Kerangka Penelitian Perbandingan Algoritma Support Vector Regression dan Decision Tree Regression Pada Prediksi Jumlah Tenaga Kerja Asing di Jawa Barat

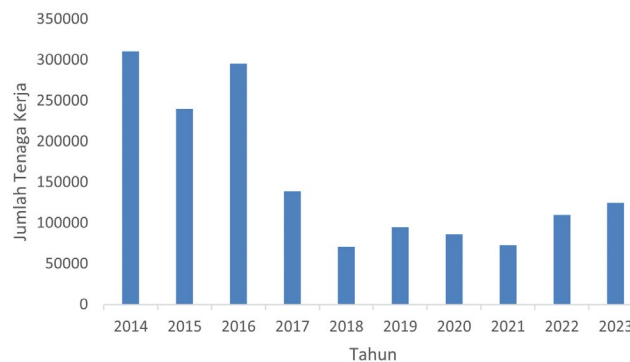
Berikut ini adalah penjelasan setiap langkah yang terdapat pada gambar 1.

A. Pemahaman Bisnis

Pada tahap ini, memahami tujuan penelitian dan kebutuhan penelitian dengan membaca beberapa referensi terkait dan mempelajari teori-teori berkaitan dengan penyerapan tenaga kerja asing.

B. Pemahaman Data

Pada tahap ini, penelitian menggunakan dataset yang bersumber dari laman pemerintah Jawa Barat, yaitu [Open Data Jawa Barat](#) berupa [Jumlah Penyerapan Tenaga Kerja Penanaman Modal Asing Berdasarkan Asal Negara di Jawa Barat](#). Dataset berisi kumpulan data jumlah tenaga kerja asing (TKA) di provinsi Jawa Barat yang berasal dari beberapa negara dengan total jumlah tenaga kerja 1,542,928 pada tahun 2014 hingga 2023 (10 tahun). Adapun total dataset yang digunakan pada penelitian ini adalah 890 baris dengan 3 atribut yang terdapat pada dataset mentah yaitu `id`, `kode_provinsi`, `negara_asal`, `jumlah_tenaga_kerja`, `satuan` dan `tahun`. Pada Gambar 2 adalah detail jumlah dataset yang berisi TKA yang ada di Jawa Barat.



Gambar. 2. Data Jumlah Tenaga Kerja Asing (TKA) di Provinsi Jawa Barat

Berdasarkan Tabel 1, dataset penelitian memiliki persebaran data TKA dari negara-negara yang berbeda setiap tahunnya, dimana Tabel 1 menampilkan data yang berisi tahun dengan jumlah negara yang mengirimkan tenaga kerja pada tahun tersebut.

TABEL I
 PERSEBARAN DATASET BERDASARKAN NEGARA ASAL TENAGA KERJA

Tahun	Jumlah Negara
2014	36
2015	45
2016	46
2017	49
2018	42
2019	39
2020	35
2021	35
2022	46
2023	52

Selain itu, dataset penelitian mengandung *missing value* yaitu beberapa data negara yang tidak mengirimkan tenaga kerja sama sekali pada tahun 2014-2023 dan ditulis dengan angka 0 pada dataset. Sedangkan anomali data pada penelitian ini dapat dilihat pada selisih antara negara dengan jumlah tenaga kerja terbesar dan jumlah tenaga kerja terkecil dalam satu tahun.

Berdasarkan Gambar 2 diketahui bahwa data tenaga kerja asing (TKA) di Jawa Barat pada tahun 2014 (309,668), 2015 (240,018), 2016 (295,012), 2017 (138,847), 2018 (70,685), 2019 (94,613), 2020 (85,922), 2021 (72,554), 2022 (109,788) dan 2023 (124,634). Pada Tabel 2 menyajikan detail atribut yang digunakan pada penelitian.

TABEL II
 DETIL ATRIBUT PENELITIAN

Atribut	Tipe	Keterangan
negara_asal	Nominal	Menyatakan asal negara penanaman modal asing
Tahun	Numerik	Menyatakan tahun data dengan tipe data numerik
Jumlah tenaga kerja	Numerik	Menyatakan jumlah penyerapan tenaga kerja penanaman modal asing

Setelah mengetahui detail atribut penelitian, maka selanjutnya disajikan sampel dataset yang digunakan untuk memprediksi jumlah tenaga asing di Jawa Barat pada Tabel 3.

TABEL III
 SAMPEL DATASET YANG DIGUNAKAN UNTUK MEMPREDIKSI JUMLAH TENAGA KERJA ASING DI PROVINSI JAWA BARAT.

No	Negara Asal	Tahun	Jumlah tenaga kerja
1	Afghanistan	2014	0
2	Amerika Serikat	2014	2,109
3	Anguilla	2014	0
4	Australia	2014	421
5	Austria	2014	0
...
890	Yunani	2023	0

C. Persiapan Dataset

Pada tahap ini dilakukan proses persiapan dataset penelitian yang bersumber dari [Jumlah Penyerapan Tenaga Kerja Asing di Provinsi Jawa Barat](#). Tahap persiapan data adalah tahap pra-pemrosesan data yang melalui beberapa tahap [17] yaitu: menghilangkan data yang duplikasi (*remove duplicate*), menghapus *missing value* (data yang bernilai 0), dan melakukan seleksi data. Secara garis besar terdapat 2 langkah pada tahap persiapan dataset, yaitu:

1) Pembersihan Dataset

Pada tahap ini menggunakan aplikasi Microsoft Excel untuk melakukan pembersihan data dengan memanfaatkan fitur *sorting* dan *delete*. Data yang dihapus adalah data negara yang jumlah tenaganya adalah 0, tujuannya adalah menghindari overfitting dan membuat akurasi dari pemodelan prediksi menjadi lebih optimal. Setelah dilakukan penghapusan data, maka dataset tersisa 426 baris dan 72 negara. Pada tabel 4 adalah gambaran dataset setelah pembersihan data.

TABEL IV
 GAMBARAN DATASET DISTRIBUSI TENAGA KERJA ASING (TKA) DI PROVINSI JAWA BARAT SETELAH DILAKUKAN PENGHAPUSAN DATA

No	Negara Asal	Tahun	Jumlah tenaga kerja
1	Amerika Serikat	2014	2,109
2	Australia	2014	421
3	Belanda	2014	2,582
4	Belgia	2014	845

5	British Virgin Islands	2014	2446
...
426	Yordania	2023	5

2) Seleksi Data

Langkah kedua adalah melakukan melakukan seleksi data, proses seleksi data menggunakan Microsoft Excel dengan fitur yaitu sorting dan filter pada atribut negara asal, jumlah tenaga kerja dan tahun. Pada proses seleksi data, menggunakan kriteria yang ditentukan pada penelitian ini yaitu:

(a) Menentukan standar minimal distribusi tenaga kerja asing.

Pada kriteria pertama, menentukan standar minimal jumlah tenaga kerja untuk proses seleksi data. Proses ini dilakukan dengan cara menghitung rata-rata tenaga kerja yang dikirimkan pertahun kemudian dibagi dengan jumlah periode selama 10 tahun (2014-2023). Pada Tabel 5 menyajikan proses penentuan standar minimal jumlah tenaga kerja untuk proses seleksi data pada tahap pertama. Berdasarkan proses perhitungan pada Tabel 5, maka rata-rata jumlah tenaga kerja yang dikirimkan ke Jawa Barat adalah 3136.75 atau dibulatkan menjadi 3136 tenaga kerja.

TABEL V
 PERHITUNGAN UNTUK SELEKSI DATA PERTAMA

Tahun	Rata-rata per tahun
2014	3593.17
2015	5334.40
2016	6295.98
2017	2835.12
2018	1683.55
2019	2426.77
2020	2455.09
2021	1959.69
2022	2386.72
2023	2396.98
Total rata-rata	3136.75

Menggunakan standar jumlah tenaga kerja yang dikirimkan minimal berjumlah 3136 per-tahun, maka pada tahap pertama dari proses seleksi data terdapat pada Tabel 6. Tabel 6 adalah hasil seleksi data menggunakan kriteria tahap pertama, dimana dataset penelitian mengalami penyusutan menjadi 56 baris dan terdiri dari 11 negara seperti yang ditampilkan pada Tabel 5.

TABEL VI
 HASIL SELEKSI DATA KRITERIA PERTAMA

No	Negara Asal	Tahun	Jumlah tenaga kerja
1	Belanda	2015	4,627
2	Belanda	2016	11,866
3	British Virgin Islands	2015	27,672
4	British Virgin Islands	2016	11,269
5	British Virgin Islands	2017	3,963
...
56	Thailand	2014	3,579

(b) Menentukan jumlah minimal pengiriman tenaga kerja asing (TKA) ke Jawa Barat selama 10 tahun.

Kriteria kedua adalah negara yang mengirimkan TKA ke Jawa Barat minimal 4 kali pada periode tahun 2014-2023 dengan mengabaikan informasi apakah dikirim pada tahun yang berurutan atau acak. Hasil seleksi data dari kriteria kedua, dataset menyusut kembali menjadi 46 baris data dan terdiri dari 6 negara. Pada Tabel 7 adalah tampilan dataset hasil dari proses seleksi data pada kedua kriteria.

TABEL VII
 HASIL SELEKSI DATA KESELURUHAN

No	Negara Asal	Tahun	Jumlah tenaga kerja
1	British Virgin Islands	2015	27,672
2	British Virgin Islands	2016	11,269
3	British Virgin Islands	2017	3,963
4	British Virgin Islands	2018	17,375
5	British Virgin Islands	2022	3,777
...
46	Singapura	2023	11,760

Berdasarkan proses seleksi data menggunakan kedua kriteria, maka dataset penelitian mengalami penyusutan dari 890 (pada Tabel 3), menjadi 46 baris (pada Tabel 7). Pada Tabel 8 merupakan total jumlah tenaga kerja asing (TKA) berdasarkan negara asal yang telah diseleksi pada Tabel 7. Tabel 8 menyajikan negara asal, tahun pengiriman TKA beserta jumlahnya, menggunakan kumpulan data yang terdapat pada tabel 8.

TABEL VIII
 TENAGA KERJA ASING PER NEGARA PERIODE 2014-2023

No	Negara	Tahun dan Jumlah Tenaga Kerja	Total
1	British Virgin Islands	2015 (27,672); 2016 (11,269); 2017 (3,963); 2018 (17,375); 2022 (3,777)	64,056
2	Hongkong, RRT	2015 (4,756); 2016 (8,148); 2017 (8,148); 2023 (3,390)	18,753
3	Jepang	2014 (83,710); 2015 (69,778); 2016 (101,805); 2017 (56,623); 2018 (20,034); 2019 (10,674); 2020 (17,057); 2021 (12,933); 2022 (22,068); 2023 (36,934)	452,865
4	Korea Selatan	2014 (58,325); 2015 (60,234); 2016 (93,806); 2017 (27,074); 2018 (15,719); 2019 (27,405); 2020 (39,784); 2021 (35,015); 2022 (52,959); 2023 (41,831)	451,494
5	R.R. Tiongkok	2014 (3,580); 2019 (33,465); 2020 (7,230); 2021 (6,380); 2022 (9,777); 2023 (12,065)	48,550
6	Singapura	2014 (22,514); 2015 (14,320); 2016 (27,552); 2017 (13,680); 2018 (5,403); 2019 (3,868); 2020 (9,536); 2021 (4,406); 2022 (9,058); 2023 (11,760)	138,661

D. Pemodelan Data

Pada tahap pemodelan data, proses training dan pengujian model algoritma untuk prediksi distribusi tenaga kerja asing (TKA) di Provinsi Jawa Barat menggunakan metode Cross Validation. Kedua algoritma ini menggunakan default parameter yang sudah dimiliki oleh kedua algoritma ini (SVR dan DTR). SVR menggunakan parameter Radial Basis Function (RBF) Kernel, karena kemampuan RBF Kernel untuk menangkap hubungan non-linear yang kompleks dengan memetakan data ke ruang dimensi tak terbatas [18]. Algoritma DTR menggunakan parameter random state, karena dalam cross-validation penentuan set pelatihan dan uji yang konsisten dengan random_state yang sama membantu dalam mendapatkan estimasi performa yang lebih stabil dan dapat diandalkan [19].

Pada tahap ini, kedua algoritma yang dibandingkan adalah algoritma Support Vector Regression (SVR) dan Decision Tree Regression (DTR). Algoritma SVR adalah salah satu algoritma yang termasuk dalam Machine Learning, dengan menggunakan fungsi kernel dalam proses kerja untuk meningkatkan akurasi regresi [8]. Sedangkan algoritma DTR merupakan salah satu teknik regresi yang paling mudah digunakan dan dapat digunakan untuk rangkaian data waktu. [20].

E. Evaluasi

Evaluasi dilakukan untuk mengevaluasi akurasi dan efektivitas hasil modeling yang dilakukan. Ini dilakukan dengan menggunakan Mean Absolute Error (MAE), dan Root Mean Squared Error (RMSE). MAE adalah selisih rata-rata antara nilai aktual dan nilai perkiraan. [21]. Sedangkan RMSE adalah hasil dari kuadrat error, atau nilai selisih antara nilai sebenarnya dan nilai prediksi, membagi hasil penjumlahan yang diperoleh dengan banyaknya waktu prediksi, dan kemudian menarik akarnya [22]. Dengan nilai MAE dan RMSE yang lebih kecil, model algoritma akan lebih baik dalam melakukan peramalan atau prediksi. [23]

Pemilihan MAE dan RMSE ini dikarenakan kedua metrik ini umum digunakan untuk mengukur kesalahan prediksi model *machine learning* [24]. Selain itu, kedua metrik ini memberikan informasi tentang seberapa jauh prediksi model menyimpang dari nilai aktual [25]. Penelitian ini juga melakukan visualisasi antara data aktual dengan hasil prediksi menggunakan algoritma terbaik, agar bisa melihat sejauh mana perbedaan yang terjadi menggunakan tabel dengan perangkat lunak Microsoft Excel.

F. Deployment

Tahap berikutnya adalah *deployment*, yaitu proses menyusun laporan penelitian dan menampilkan visualisasi dari hasil pengolahan data yang telah dilakukan. Hasil visualisasi perbandingan kedua algoritma (SVR dan DTR) menggunakan model plotting. Selain itu, juga akan ditampilkan hasil prediksi jumlah tenaga kerja di tahun 2024 berdasarkan negara.

III. HASIL DAN PEMBAHASAN

A. Perbandingan Performa Algoritma SVR dan DTR

Pada tahap ini menyajikan hasil perbandingan performa dari algoritma SVR dan DTR menggunakan model pengujian Cross Validation dengan nilai $K=3$. Hasil pengujian dievaluasi menggunakan matriks pengujian RMSE dan MAE dan hasilnya ditampilkan dalam bentuk persentase (%). Proses pengujian model menggunakan dataset yang berisi negara-negara yang mengirimkan tenaga kerja asing (TKA) di Jawa Barat, sebagaimana disajikan pada Tabel 7. Hasil evaluasi performa kedua algoritma untuk prediksi jumlah TKA di provinsi Jawa Barat terdapat pada Tabel 9.

TABEL IX
 HASIL PERBANDINGAN PERFORMA ALGORITMA SVR DAN DTR UNTUK PREDIKSI JUMLAH TENAGA KERJA ASING (TKA) DI JAWA BARAT PADA TAHUN 2014-2023 (TEKS TEBAL UNTUK PERFORMA TERBAIK)

Negara_Asal	SVR		DTR	
	RMSE	MAE	RMSE	MAE
British Virgin Islands	133.98%	100.24%	157.88%	67.33%
Hongkong, Rrt	81.56%	62.97%	72.20%	94.63%
Jepang	78.94%	75.99%	62.21%	72.96%
Korea Selatan	55.49%	59.17%	57.74%	85.45%
R.R. Tiongkok	81.56%	62.97%	52.58%	28.25%
Singapura	59.24%	63.40%	65.63%	68.79%
Rata-rata	81.80%	70.79%	78.04%	69.57%

Berdasarkan hasil pengujian performa pada Tabel 9, menunjukkan bahwa performa algoritma DTR unggul dalam memprediksi negara R.R. Tiongkok, dan Jepang karena memiliki nilai lebih rendah jika dibandingkan algoritma SVR. Sedangkan algoritma SVR unggul dalam memprediksi negara Korea Selatan dan Singapura karena memiliki nilai MSE dan MAE lebih rendah jika dibandingkan algoritma DTR. Berdasarkan keseluruhan nilai MSE dan MAE dari 6 negara tersebut, maka algoritma DTR unggul karena memiliki rata-rata nilai MSE dan MAE yang lebih rendah yaitu 78.04% dan 69.57% dibandingkan algoritma SVR yang memiliki rata-rata nilai MSE dan MAE yaitu 81.80% dan 70.79%.

DTR unggul pada prediksi untuk negara Jepang dan R.R. Tiongkok karena data dari kedua negara tersebut memiliki data pengiriman tenaga kerja asing (TKA) lebih seimbang sebagaimana yang ditampilkan pada Tabel 8. Sedangkan SVR unggul pada prediksi untuk negara Korea Selatan dan Singapura karena data dari kedua negara memiliki data Pengiriman TKA tiap tahunnya fluktuatif atau memiliki perbedaan jumlah yang signifikan seperti yang ditampilkan pada Tabel 8.

B. Evaluasi Perbandingan Performa Model untuk Peramalan Jumlah Tenaga Kerja Asing (TKA) di Jawa Barat

Berdasarkan perbandingan performa kinerja dari kedua model (SVR dan DTR) pada Tabel 9, maka dihitung nilai rata-rata RMSE dan MAE dalam memprediksi jumlah TKA di Jawa Barat pada periode 2014-2023. Pada Tabel 10 menampilkan hasil perbandingan performa kedua model menggunakan nilai rata-rata perhitungan RMSE dan MAE.

TABEL X
 HASIL EVALUASI PERBANDINGAN ALGORITMA SVR DAN DTR BERDASARKAN PERFORMA MODEL (TEKS TEBAL UNTUK PERFORMA TERBAIK)

Keterangan	SVR		DTR	
	RMSE	MAE	RMSE	MAE
Rata-rata	81.80%	70.79%	78.04%	69.57%

Berdasarkan hasil evaluasi terhadap kedua algoritma, SVR dan DTR yang disajikan pada Tabel 9, maka dapat disimpulkan bahwa algoritma DTR memiliki performa lebih unggul dalam memprediksi jumlah tenaga kerja asing di provinsi Jawa Barat, dengan nilai RMSE 78.04% dan nilai MAE 69.57%, dibandingkan algoritma SVR dengan nilai MSE 81.80%, dan nilai MAE 70.79%. Pada Gambar 3 merupakan visualisasi hasil perbandingan antara kedua algoritma (SVR dan DTR) dengan menggunakan RMSE dan MAE. Algoritma DTR memiliki posisi yang lebih rendah dibandingkan dengan algoritma SVR di RMSE maupun MAE.

Dari hasil evaluasi perbandingan kedua algoritma maka perbedaan nilai *error* pada evaluasi matriks tersebut, tetap dapat membantu perencanaan kebijakan yaitu pemerintah Jawa Barat dalam mendapatkan tambahan devisa dengan bertambahnya tenaga kerja asing (TKA) yang masuk ke provinsi Jawa Barat. Karena berdasarkan data pada Tabel 11, data aktual TKA pada tahun 2014-2023 lebih banyak dibandingkan data prediksi menggunakan algoritma terbaik (DTR). Sehingga menggunakan algoritma DTR diasumsikan hasil prediksi TKA dapat melebihi data aktual TKA.



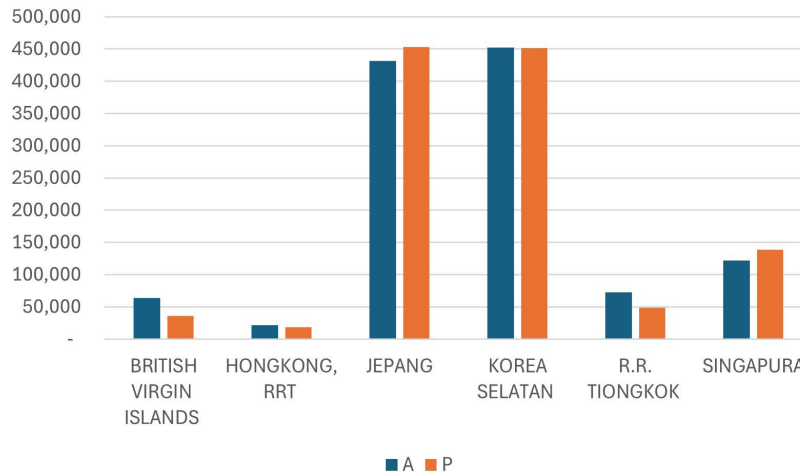
Gambar. 3. Hasil Perbandingan Performa Algoritma SVR dan DTR untuk Prediksi Jumlah Tenaga Kerja Asing di Jawa Barat

C. Prediksi Tenaga Kerja Asing (TKA) di Provinsi Jawa Barat

Berdasarkan perbandingan kedua algoritma (SVR dan DTR), maka pada bagian ini menyajikan prediksi menggunakan algoritma terbaik yaitu algoritma DTR. Berikut ini adalah hasil prediksi menggunakan Algoritma DTR.

(1) Prediksi Tenaga Kerja Asing (TKA) di Jawa Barat pada Tahun 2014-2023.

Pada Gambar 4 menampilkan hasil dari perbandingan data aktual (A) dan data prediksi (P) menggunakan algoritma terbaik yaitu DTR. Data aktual (A) menggunakan data pada Tabel 8, sedangkan data prediksi (P) hasil prediksi jumlah TKA di Provinsi Jawa Barat pada tahun 2014-2023 menggunakan algoritma DTR.



Gambar. 4. Visualisasi Perbandingan Data Aktual dan Prediksi Negara Asal Tenaga Kerja Asing di Provinsi Jawa Barat Periode 2014-2023

Gambar 4 menampilkan visualisasi perbandingan hasil prediksi menggunakan algoritma terbaik yaitu algoritma DTR dibandingkan dengan jumlah TKA di Jawa Barat pada periode 2014-2023 dengan negara-negara yang sudah dilakukan seleksi pada tahap persiapan dataset sebelumnya.

TABEL XI
 HASIL PREDIKSI TENAGA KERJA ASING DI JAWA BARAT PADA TAHUN 2024 MENGGUNAKAN ALGORITMA DTR

No	Negara Asal	Prediksi 2024
1	British Virgin Islands	942
2	Hongkong,RRT	5,297
3	Jepang	36,934
4	Korea Selatan	41,831
5	R.R Tiongkok	3,687
6	Singapura	11,760
Total Prediksi Tenaga Kerja 2024		100,451
Rata-rata Tenaga Kerja 2024		16,742

Berdasarkan Tabel 11, terdapat 6 negara yang terpilih dan diprediksi mengirimkan tenaga kerja ke Provinsi Jawa Barat pada tahun 2024 dengan total jumlah tenaga kerja asing (TKA) menurut prediksi adalah 100,451 dengan rata-rata setiap negara mengirimkan 16,742 tenaga kerja. Pada Gambar 5, menampilkan visualisasi data prediksi negara asal tenaga kerja asing (TKA) di Provinsi Jawa Barat pada tahun 2024.

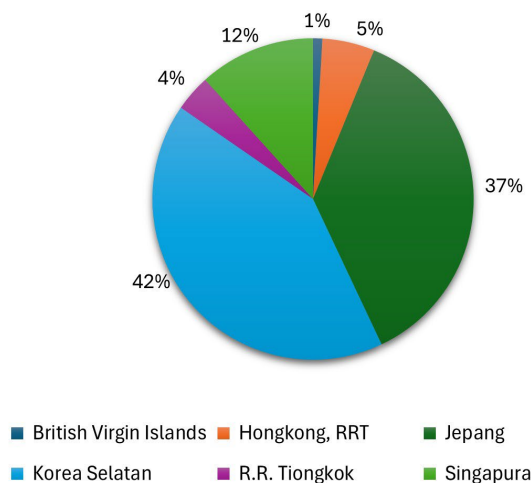
Dari data jumlah tenaga kerja asing provinsi Jawa Barat pada tahun 2024, otoritas terkait dalam hal ini pemerintah daerah dapat membuat beberapa kebijakan terkait untuk bisa mengelola TKA salah satunya dengan pembuatan regulasi untuk mengatur TKA sekaligus juga menyiapkan teknis penegakan serta pengawasan pelaksanaan regulasi tersebut. Seperti pada pasal 43 ayat 1 UU Ketenagakerjaan menyatakan bahwa pemberi kerja yang mempergunakan tenaga kerja asing harus memiliki rencana penggunaan tenaga kerja asing yang disahkan oleh Menteri atau Pejabat yang ditunjuk, sedangkan pada UU Cipta Kerja tepatnya Pasal 43 mengenai rencana penggunaan tenaga kerja asing dari pemberi kerja sebagai syarat mendapat izin kerja dimana dalam UU Cipta Kerja, informasi terkait periode penugasan ekspatriat, penunjukan tenaga kerja menjadi warga negara Indonesia sebagai mitra kerja ekspatriat dalam rencana penugasan ekspatriat dihapuskan [26].

Sementara untuk regulasi mengenai tenaga kerja yang ada di provinsi Jawa Barat diatur dalam di Jawa Barat mengaturnya dalam Peraturan Daerah (PERDA) Provinsi Jawa Barat Nomor 9 Tahun 2013, PERDA Provinsi Jawa Barat Nomor 6 Tahun 2014 dan Peraturan Gubernur (PERGUB) Jawa Barat Nomor 54 Tahun 2018[27]. Regulasi ini seharusnya dibuat agar dapat meningkatkan pendapatan asli daerah (PAD) serta dapat melakukan

transfer teknologi untuk tenaga kerja lokal. Selain itu, regulasi tersebut juga harus mengatur batasan-batasan terkait TKA agar tidak disalahgunakan oleh berbagai pihak.

(2) Visualisasi Prediksi Tenaga Kerja Asing (TKA) di Jawa Barat.

Pada Gambar 5 menampilkan visualisasi prediksi negara asal TKA di Provinsi Jawa Barat pada tahun 2024. Maka negara yang mengirimkan tenaga kerja di Jawa Barat terbanyak pada tahun 2024 diprediksi adalah Korea Selatan dengan nilai 42%, kemudian Jepang dengan nilai 37%, lalu Singapura 12%, Hongkong RRT sebesar 5%, R.R. Tiongkok sebesar 4% dan terakhir adalah British Virgin Islands yaitu 1%.



Gambar. 5. Visualisasi Prediksi Negara Asal Tenaga Kerja Asing di Provinsi Jawa Barat

D. Diskusi

Berdasarkan hasil eksperimen yang telah dilakukan, maka penelitian ini membuktikan bahwa algoritma DTR dan SVR mampu melakukan prediksi jumlah tenaga kerja asing di provinsi Jawa Barat. Dengan algoritma DTR lebih unggul dari algoritma SVR menggunakan pengukuran performa model RMSE dan MAE, dimana DTR mendapatkan nilai RMSE sebesar 78.04% dan MAE mendapatkan nilai 69.57%. Karakteristik dataset dan evaluasi metrik yang hampir serupa dengan penelitian [9] dimana penelitian ini menggunakan algoritma Algoritma Simple Linear Regression (SLR) dan SVR dalam Prediksi Jumlah Penduduk Di Sulawesi Tenggara dan menemukan bahwa algoritma SLR lebih unggul dari algoritma SVR, dengan SLR mendapatkan nilai RMSE 0.51% dan MAPE 1.89% sehingga dapat disimpulkan bahwa model dan evaluasi metrik yang digunakan terbukti sesuai.

IV. KESIMPULAN

Penelitian ini, membandingkan 2 (dua) algoritma klasifikasi, yaitu algoritma Support Vector Regression (SVR) dan Decision Tree Regression (DTR) untuk memprediksi jumlah tenaga kerja asing (TKA) di provinsi Jawa Barat. Berdasarkan hasil penelitian maka dapat disimpulkan bahwa metode Decision Tree Regression (DTR) terbukti memiliki performa lebih unggul dalam memprediksi jumlah TKA di Provinsi Jawa Barat dibandingkan dengan Support Vector Regression (SVR). Dimana DTR memiliki nilai RMSE sebesar 78.04% dan MAE sebesar 69.57%, sedangkan SVR hanya memiliki nilai RMSE sebesar 81.80%. dan MAE sebesar 70.79%. Pada penelitian berikutnya, kami merencanakan pendekatan deep learning untuk prediksi jumlah TKA di Jawa Barat dan Jawa Tengah.

V. DAFTAR PUSTAKA

- [1] R. Abduh, "Dampak sosial tenagakerja asing (TKA) di Indonesia," *SosekJurnal Sos. dan Ekon.*, vol. 1, no. 1, pp. 25–28, 2020, [Online]. Available: <https://jurnal.bundamedia grup.co.id/index.php/sosek/article/view/24/24>
- [2] R. I. Khalidun, G. A. Fita, A. N. F. Utami, and T. H. Tahawa, "Globalisasi, ancaman dan upaya peningkatan daya saing tenaga kerja," *J. Ilmu Hub. Intranasional*, vol. 01, no. 01, pp. 27–36, 2020, [Online]. Available: <https://ojs.unsulbar.ac.id/index.php/lino/article/view/827/444>
- [3] N. Aida, U. Ciptawaty, T. Gunarto, and S. Aini, "Analisis dampak penanaman modal asing dan tenaga kerja asing tiongkok terhadap perekonomian Indonesia," *J. Ekon. Pambang.*, vol. 10, no. 3, pp. 159–167, 2021, doi: 10.23960/jep.v10i3.301.
- [4] A. Ahdiat, "Sebaran Investasi Asing 2023, Terbanyak Masuk Jawa Barat," databooks. Accessed: Jun. 20, 2024. [Online]. Available: <https://databoks.katadata.co.id/datapublish/2024/03/07/sebaran-investasi-asing-2023-terbanyak-masuk-jawa-barat#:~:text=Menurut data Badan Koordinasi Penanaman Modal%28BKPM%29%2C sepanjang,nilai US%248%2C28 miliar atau 16%2C5%25 dari total PMA.>
- [5] A. Ahdiat, "Sebaran Tenaga Kerja Asing November 2022, Terbanyak di Sulteng," databooks. Accessed: Jun. 27, 2024. [Online]. Available: <https://databoks.katadata.co.id/datapublish/2023/01/18/sebaran-tenaga-kerja-asing-november-2022-terbanyak-di-sulteng>
- [6] D. Budiman, S. D. Nabella, and U. I. Sina, "Masuknya warga negara asing dalam perspektif ekonomi pertahanan," *BENING*, vol. 7, no. June, pp. 53–

- 65, 2020, doi: 10.33373/bening.v7i1.2354.
- [7] A. T. Nurani, A. Setiawan, and B. Susanto, "Perbandingan kinerja regresi decision tree dan regresi Linear berganda untuk prediksi BMI pada dataset asthma," *J. Sains dan Edukasi Sains*, vol. 6, no. 1, pp. 34–43, 2023, doi: 10.24246/juses.v6i1p34-43.
- [8] A. Hermawan, I. W. Mangku, N. K. K. Ardana, and H. Sumarno, "Analisis support vector regression dengan algoritma grid search untuk memprediksi harga saham," *J. Math. Its Appl.*, vol. 18, no. 1, pp. 41–60, 2022, doi: 10.29244/milang.18.1.41-60.
- [9] R. I. M. Chaidir, A. F. Ramadhan, H. Zaria, and R. A. Saputra, "Perbandingan Algoritma Simple Linear Regression Dan Support Vector Regression Dalam Prediksi Jumlah Penduduk Di Sulawesi Tenggara," *Method. J. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 1, pp. 27–31, 2024, doi: 10.46880/mtk.v10i1.2548.
- [10] R. A. Saputra *et al.*, "Implementasi algoritma decision tree C4.5 dan support vector regression untuk prediksi penyakit stroke," *J. Phys. Conf. Ser.*, vol. 1641, no. 1, pp. 61–67, 2020, doi: 10.1088/1742-6596/1641/1/012025.
- [11] A. Arfan and L. ETP, "Perbandingan Algoritma long short-term memory dengan SVR pada Prediksi harga saham di Indonesia," *Petir*, vol. 13, no. 1, pp. 33–43, 2020, doi: 10.33322/petir.v13i1.858.
- [12] S. M. Putri, R. Novita, and M. Afdal, "Perbandingan Algoritma Linear Regression, Support Vector Regression, dan Artificial Neural Network untuk Prediksi Data Obat," vol. 6, no. 1, pp. 54–63, 2024, doi: 10.47065/bits.v6i1.5184.
- [13] F. M. Haikal, M. Akrom, and G. A. Trisnapradika, "Perbandingan Algoritma Multilinear Regression dan Decision Tree Regressor dalam Memprediksi Efisiensi Penghambatan Korosi Piridazin," *Edumatic J. Pendidik. Inform.*, vol. 7, no. 2, pp. 307–315, 2023, doi: 10.29408/edumatic.v7i2.22127.
- [14] S. Saadah, F. Z. Z., and H. H. Z., "Support Vector Regression (SVR) dalam memprediksi harga minyak kelapa sawit di Indonesia dan nilai tukar Mata uang EUR/USD," *J. Comput. Sci. Informatics Eng.*, vol. 5, no. 1, pp. 85–92, 2021, doi: 10.29303/jcosine.v5i1.403.
- [15] A. B. Raharjo, A. Ardianto, and D. Purwitasari, "Random Forest Regression Untuk Prediksi Produksi Daya Pembangkit Listrik Tenaga Surya," *Briliant J. Ris. dan Konseptual*, vol. 7, no. 4, p. 1058, 2022, doi: 10.28926/briliant.v7i4.1036.
- [16] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," *Procedia Comput. Sci.*, vol. 181, no. 2019, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.
- [17] R. Oktafiani and R. Rianto, "Perbandingan Algoritma Support Vector Machine (SVM) dan Decision Tree untuk Sistem Rekomendasi Tempat Wisata," *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 2, pp. 113–121, 2023, doi: 10.25077/teknosi.v9i2.2023.113-121.
- [18] A. A. Fauzi, A. M. Soleh, and A. Djuraidah, "Kajian Simulasi Perbandingan Metode Regresi Kuadrat Terkecil Parsial, Support Vector Machine, Dan Random Forest," *Indones. J. Stat. Its Appl.*, vol. 4, no. 1, p. 204, 2020, doi: 10.29244/ijsa.v4i1.610.
- [19] S. Joses, D. Yulvida, and S. Rochimah, "Pendekatan Metode Ensemble Learning untuk Prakiraan Cuaca menggunakan Soft Voting Classifier," vol. 5, no. 1, p. 77, 2024.
- [20] A. Setiawan, R. F. Waleska, M. A. Purnama, and L. Efrizoni, "Komparasi algoritma k-Nearest neighbor (K-NN), support vector machine (SVM), dan decision tree dalam klasifikasi penyakit stroke," *J. Inform. Rekayasa Elektron.*, vol. 7, no. 1, pp. 107–114, 2024.
- [21] M. Aditya Pratama, M. Munawaroh, W. Joko Pranoto, P. Studi Teknik Informatika, F. Sains dan Teknologi, and U. Muhammadiyah Kalimantan Timur, "Perbandingan Performa Algoritma Linear Regresi dan Random Forest untuk Prediksi Harga Bawang Merah di Kota Samarinda," *J. Ilmu Tek.*, vol. 1, no. 2, pp. 172–182, 2024, [Online]. Available: <https://doi.org/10.62017/teknik>
- [22] F. N. Iman and D. Wulandari, "Prediksi harga saham menggunakan metode long short term memory," *Log. J. Ilmu Komput. dan Pendidik.*, vol. 1, no. 3, pp. 601–616, 2023, doi: 2985-4172.
- [23] A. Satria, R. M. Badri, and I. Safitri, "Prediksi hasil panen tanaman pangan sumatera dengan metode machine learning," *Digit. Transform. Technol.*, vol. 3, no. 2, pp. 389–398, 2023, doi: 10.47709/digitech.v3i2.2852.
- [24] M. Hariyanto, M. Kholiq, A. Yani, and Narti, "Optimasi Kinerja Linear Regression, Random Forest Regression Dan Multilayer Perceptron Pada Prediksi Hasil Panen," *Inti Nusa Mandiri*, vol. 14, no. 2, pp. 133–138, 2024, doi: <https://doi.org/10.33480/inti.v18i2.5269>.
- [25] A. Abdullah and P. Y. Utami, "Penerapan Metode Regresi Linier Berganda Untuk Memprediksi Panen Kelapa Sawit," *Progresif J. Ilm. Komput.*, vol. 20, no. 01, pp. 540–554, 2024.
- [26] I. Hanifah, "Peluang Tenaga Kerja Asing Untuk Bekerja Di Indonesia Berdasarkan Rancangan Undang-Undang Cipta Kerja," *LEGA LATA J. Ilmu Huk.*, vol. 6, no. 1, pp. 158–173, 2021.
- [27] L. Hakim, M. Luthfie, and B. Kartakusumah, "Pola Kerjasama Dalam Pengawasan Dan Pembinaan Tenaga Kerja Asing," *Adm. J. Adm. Publik*, vol. 3, no. 1, pp. 36–45, 2021, doi: : 2715-2499.