

Analisis Pengaruh Seleksi Fitur Terhadap Pengukuran Korelasi Pada Pemetaan Penilaian Kinerja Karyawan Berdasarkan Data Kuesioner

Iqbal Ariefa Desyunda¹⁾, Dandy Pramana Hostiadi²⁾, Roy Rudolf Huizen³⁾

Program Studi Magister Sistem Informasi
Institut Teknologi dan Bisnis STIKOM Bali
Denpasar, Indonesia

e-mail: 222012033@stikom-bali.ac.id¹⁾, dandy@stikom-bali.ac.id²⁾, roy@stikom-bali.ac.id³⁾

Abstrak

Dalam era bisnis yang terus berkembang pesat, penilaian kinerja karyawan menjadi esensial dalam manajemen sumber daya manusia. Penggunaan kuesioner sebagai alat penilaian karyawan umum dilakukan untuk mendapatkan pandangan dari berbagai pihak terkait. Namun, mengelola data dari kuesioner tidaklah mudah karena beragamnya kualitas dan validitas data serta potensi kesalahan subjektif. Dalam mengatasi hal ini, proses ekstraksi fitur menjadi pengaruh dalam pemetaan penilaian kinerja karyawan. Ekstraksi fitur membantu menyederhanakan data dan mengekstrak informasi relevan, serta membersihkan data dari noise atau kesalahan. Proses ini melibatkan pembersihan data dan metode rank feature untuk menyeleksi fitur-fitur penting. Dengan memanfaatkan ekstraksi fitur, perusahaan dapat mengambil keputusan yang lebih tepat dalam manajemen sumber daya manusia, memastikan evaluasi kinerja karyawan menjadi dasar pengambilan keputusan strategis terkait pengembangan karir, pelatihan, dan penghargaan. Hasil proses seleksi fitur, berhasil mereduksi fitur sebesar 26.27% dari 15 fitur menjadi 11 fitur dan terdapat korelasi antara parameter Kerjasama dan Pendidikan terakhir yang memiliki nilai tertinggi. Manfaat dari penelitian adalah menunjukkan pengaruh teknik analisis seleksi fitur terhadap hasil pengukuran korelasi.

Kata kunci: Penilaian Kinerja Karyawan, Kuesioner, Ekstraksi Fitur, Korelasi

1. Pendahuluan

Dalam era bisnis yang terus berkembang pesat, penilaian kinerja karyawan menjadi salah satu aspek utama dalam manajemen sumber daya manusia[1]. Proses ini tidak hanya memberikan pandangan terhadap kontribusi individu terhadap tujuan perusahaan, tetapi juga menjadi dasar untuk pengembangan karir dan pengambilan keputusan. Salah satu teknik yang umum digunakan dalam proses penilaian kinerja karyawan adalah melalui kuesioner, yang memungkinkan pengumpulan informasi dari berbagai penjamin kepentingan, termasuk karyawan itu sendiri, atasan, dan rekan kerja.

Kuesioner instrumen yang digunakan untuk mengukur suatu peristiwa atau kejadian yang berisi kumpulan pertanyaan untuk memperoleh informasi terkait penelitian yang dilakukan[2]. Kuesioner berisi sejumlah pertanyaan kepada responden untuk memperoleh data penelitian dengan tujuan untuk mendeteksi minat (interest), sikap (attitude) dan kebiasaan (habits). Kuesioner berupa daftar pertanyaan untuk mendapatkan data penelitian yang dibagikan kepada responden untuk diisi dan kemudian dikembalikan pada peneliti[3].

Namun, data kuesioner dalam penilaian kinerja karyawan tidaklah mudah juga. Data yang diperoleh seringkali beragam dalam hal kualitas dan data yang *valid*, dan juga rentan terhadap kesalahan subjektif atau interpretasi yang tidak konsisten. Oleh karena itu, untuk menjalankan serangkaian proses sistem dalam mengolah data kuesioner, diperlukannya suatu proses yang perlu dilewati untuk menjaga kestabilan data dan keakuratan data yang dihasilkan serta mencegah terjadinya kesalahan pada proses pencocokan fitur.

Ekstraksi fitur adalah proses pengambilan ciri sebuah objek yang dapat menggambarkan karakteristik dari objek tersebut[4]. Feature Extraction atau ekstraksi fitur merupakan suatu pengambilan ciri/feature dari suatu bentuk atau tekstur yang nantinya nilai yang didapatkan akan dianalisis untuk proses klasifikasi[5]. Dengan mengurangi dimensi data atau menyoroti fitur-fitur yang penting, ekstraksi fitur membantu memperbaiki kinerja algoritma pembelajaran mesin dan menghasilkan hasil yang lebih baik[6]. Ekstraksi fitur ini meliputi pembersihan data (*data cleaning*), dan *rank feature*.

Data cleaning adalah proses mengidentifikasi, memperbaiki, dan menghilangkan kesalahan atau ketidaksesuaian dalam dataset untuk memastikan keakuratannya dan konsistensinya. Ini adalah langkah penting dalam analisis data karena dataset yang tidak bersih dapat mengarah pada kesalahan dalam interpretasi dan pengambilan keputusan yang salah[7]. *Data cleaning* juga bisa diartikan untuk menghilangkan noise data yang tidak konsisten[8], [9], [10], [11].

Selain itu, metode *rank feature* juga memainkan peran penting dalam pengelolaan data kuesioner untuk penilaian kinerja karyawan. Dengan menggunakan teknik ini, sistem ranking dapat diterapkan untuk menyeleksi dan mengambil fitur-fitur yang penting pada data untuk tujuan atau kriteria tertentu[12], [13]. Teknik ini memungkinkan perusahaan untuk mengidentifikasi atribut-atribut yang paling signifikan dalam menentukan kinerja karyawan dan mengalokasikan sumber daya secara efisien untuk pengembangan dan peningkatan.

Dengan menggunakan data kuesioner dengan memanfaatkan ekstraksi fitur yang digunakan, perusahaan dapat mengambil keputusan yang lebih tepat dan berorientasi pada data dalam manajemen sumber daya manusia, sehingga evaluasi kinerja karyawan bukan hanya menjadi refleksi dari kualitas dan kontribusi untuk individu, melainkan menjadi fondasi untuk pengambilan keputusan yang strategis terkait dengan pengembangan karir, pelatihan, bahkan penghargaan. Namun, tantangan setelah pengambilan data dari kuesioner adalah bagaimana cara mengolah dan menganalisis data untuk menjadi suatu informasi yang bermakna atau akurat.

2. Metode Penelitian

Berikut ini adalah alur penelitian yang dilakukan :



Gambar 1. Alur Penelitian

Pada tahap pertama dalam penelitian ini adalah menyiapkan data pertanyaan penilaian kinerja karyawan dalam bentuk data kuisisioner. Dari hasil pengisian kuisisioner, dilakukan proses data cleaning, yaitu proses analisis terhadap penilaian parameter yang memiliki nilai kosong, lengkap, tidak relevan, atau bermasalah lainnya untuk memastikan kualitas data yang baik. Setelah proses data cleaning, maka dilakukan proses data transformasi. Kemudian dilanjutkan dengan proses data transformasi, yaitu perubahan data yang bersifat kategorikal ke data numerikal, perubahan ini dilakukan agar nilai pada setiap parameter penilaian terstandarisasi. Hasil data transformasi menghasilkan nilai dari setiap parameter menjadi nilai numerik. Kemudian dilakukan pelabelan kelas pada setiap data responden yang dinilai berdasarkan label sangat berkualitas, cukup berkualitas dan berkualitas.

Setelah perubahan data dari kategorikal ke numerik dan memiliki kelas label, maka dilanjutkan dengan proses seleksi fitur, yaitu pemilihan fitur terbaik. Dalam penelitian ini digunakan metode gain ratio untuk melihat parameter atau fitur yang relevan berdasarkan kelas dalam pemetaan data. Di penelitian ini seleksi fitur juga bertujuan untuk mereduksi fitur dengan penentuan reduksi adalah sebesar 75%. Setelah didapatkan fitur terbaik, analisis dilakukan untuk mencari korelasi yang terbaik diantara pasangan fitur yang ada. Korelasi dilakukan dengan mengadopsi metode pearson korelasi. Persamaan pearson correlation ditunjukkan pada persamaan 1.

$$r_{x,y} = \frac{cov(x,y)}{s(x).s(y)} \quad (1)$$

Dimana x adalah parameter terbaik pertama dan y adalah parameter terbaik kedua dalam hasil seleksi fitur untuk menentukan korelasi fitur kinerja dosen. Dalam penelitian ini, korelasi terbaik pasangan fitur sebanyak 2 fitur yang didapatkan dari kombinasi terhadap 75% seleksi yang dilakukan dengan metode information gain ratio.

3. Hasil dan Pembahasan

Penelitian ini melakukan analisis terhadap pengaruh seleksi fitur terhadap pengukuran korelasi diantara fitur. Di penelitian ini menggunakan sampe data dari pengukuran kinerja karyawan yang melibatkan 35 responden. Pengolahan data dilakukan dengan menggunakan aplikasi orange data mining.

3.1 Data Pertanyaan

Di penelitian ini, menggunakan dataset private yang didapatkan dari proses penyebaran kuisisioner penilaian kinerja karyawan. Kuisisioner yang disebar memiliki pertanyaan dan standar penilaian, dimana contoh pertanyaan dan standar penilaian ditunjukkan pada Tabel 1.

Tabel 1. Data Pertanyaan

No	Pertanyaan	Score				
		1	2	3	4	5
1	Tingkat Pendidikan yang diselesaikan					
2	Pernah mengikuti sertifikasi yang relevan					
3	Tingkat presentase kehadiran dalam perhitungan 1 tahun					
4	Konsisten dalam mentaati peraturan yang ada					
5	Banyaknya prestasi yang pernah diraih selama bekerja					
6	Penghargaan yang diraih diluar instansi					
7	Pernah melakukan pelanggaran kerja dan menerima sanksi					
8	Jujur dan Amanah dalam melakukan pekerjaan					
9	Dapat bekerjasama dengan tim dengan baik					
10	Mau mendapatkan kerjaan diluar jam kerja					
11	Dapat menyelesaikan pekerjaan dengan baik					
12	Sikap yang ditunjukkan ke atasan, setara, atau bawahan					
13	Hasil pekerjaan yang sesuai dengan perintah					
14	Dapat berkomunikasi yang baik dengan Masyarakat dalam maupun luar					
15	Masa kerja di perusahaan					

Nilai Pengisian parameter Kedisiplinan, Etika, Tanggung Jawab, Kemampuan Berkomunikasi, Loyalitas, Kejujuran, Kerjasama, Kualitas Hasil Kerja :
1 = Sangat Kurang; 2 = Kurang; 3 = Cukup; 4 = Baik; 5 = Sangat Baik

Nilai Pengisian parameter Untuk Pendidikan Terakhir
1 = SMA/K Sederajat; 2 = D3; 3 = S1; 4 = S2; 5 = S3

Nilai Pengisian parameter Masa Kerja
1 = <1 Tahun; 2 = 1-3 Tahun; 3 = 3-5 Tahun; 4 = 5-7 Tahun; 5 = >7 Tahun

Nilai Pengisian parameter Absensi
1 = <50%; 2 = 50-60%; 3 = 61-70%; 4 = 71-80%; 5 = >80%

Nilai Pengisian parameter Sertifikasi relevan, prestasi, penghargaan (diluar instansi):
1 = Tidak Ada; 2 = 1 Sertifikasi relevan/prestasi/penghargaan; 3 = 2 Sertifikasi relevan/prestasi/penghargaan; 4 = 3-4 Sertifikasi relevan/prestasi/penghargaan; 5 = <5 Sertifikasi relevan/prestasi/penghargaan

Nilai Pengisian parameter Pelanggaran Kerja :
1 = 4 Kali; 2 = 3 Kali; 3 = 2 Kali; 4 = 1 Kali; 5 = Tidak Pernah

3.2 Data Cleaning

Berdasarkan hasil data kuesioner yang telah diperoleh dalam wawancara, proses ini bertujuan untuk mendapatkan kualitas data yang baik. Di penelitian ini tidak ditemukan data bermasalah seperti nilai kosong, lengkap, tidak relevan, atau bermasalah lainnya. Sehingga terdapat data sejumlah 35 instance (*no missing data*) dengan 15 fitur (*features*).

3.3 Data Transformasi

Data transformasi ini bertujuan untuk mengubah struktur data menjadi bentuk numerik. Dalam penelitian ini menggunakan teknik *one hot encode*. Hasil dari tahap ini adalah perubahan data kuesioner menjadi *numeric* dan mendefinisikan pertanyaan menjadi fitur. Hasil dari perubahan data transformasi ditunjukkan pada Tabel 2.

Tabel 2. Contoh hasil data *Transformation*

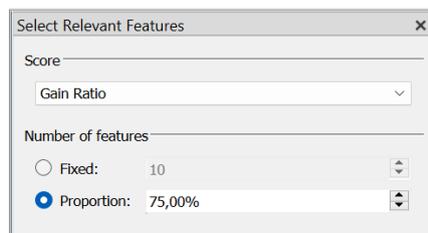
Nama Parameter	Data Responden Penilaian Kinerja Karyawan						
	R1	R2	R2	R4	R5	R6	Ri
Pendidikan Terakhir	3	4	4	3	3	4	...
Absensi	4	3	4	3	4	4	...
Kedisiplinan	5	5	4	5	5	4	...
Etika	4	5	3	4	3	5	...
Tanggung Jawab	4	4	4	5	5	5	...
Kemampuan berkomunikasi	3	3	4	4	3	3	...
Loyalitas	5	2	3	5	5	5	...
Sertifikasi Relevan	2	3	4	3	3	2	...
Kejujuran	5	3	4	3	5	3	...
Prestasi	3	4	1	2	1	3	...
Kerjasama	5	4	3	5	4	4	...
Penghargaan (diluar instansi)	2	3	2	1	2	2	...
Pelanggaran Kerja	4	4	4	5	4	5	...
Kualitas Hasil Kerja	4	4	4	4	5	4	...
Masa Kerja	5	5	5	5	4	5	...

3.4 Menentukan Kelas Label

Dalam proses ini dilakukan pelabelan kelas. Dimana pelabelan dilakukan dengan melibatkan pakar dan mengisi data dengan label sesuai penilaian yang dilakukan oleh pakar. Hasil pelabelan mengisi setiap data kinerja karyawan kedalam kelas sangat berkualitas, berkualitas dan cukup berkualitas.

3.5 Seleksi Fitur

Pada tahapan seleksi fitur dengan menggunakan metode *Gain Ratio*, yang menggunakan reduksi (proportion) sebesar 75% untuk memilih subset fitur yang paling relevan dan informatif untuk digunakan. *Gain Ratio* adalah metode evaluasi fitur yang berguna untuk menentukan seberapa baik suatu fitur dapat membagi dataset menjadi kelas-kelas yang homogen. Penentuan dan hasil pengukuran *Gain Ratio* ditunjukkan pada Gambar 3.



(a)



(b)

Gambar 3. (a) Metode Information Gain Ratio, (b) Hasil Seleksi Fitur.

Dari Gambar 3, menunjukkan bahwa metode information *gain ratio* mampu melakukan seleksi fitur. Hasil seleksi fitur mereduksi dari 15 fitur menjadi 11 fitur yang relevan, atau berhasil mereduksi fitur sejumlah 26.67%. Fitur terbaik yang dihasilkan menghasilkan rank feature yang dimana rank 1 jatuh ke fitur kedisiplinan dengan perolehan rasio 0,111, sedangkan fitur yang paling rendah jatuh kepada pendidikan terakhir dengan perolehan rasio 0,052. Hasil detail pengukuran *gain ratio* ditunjukkan pada Tabel 3.

Tabel 3. Hasil pengukuran korelasi

Parameter	Nilai Gain Ratio	Parameter	Nilai Gain Ratio
Kedisiplinan	0.111	Kemampuan Berkomunikasi	0.062
Masa Kerja	0.107	Prestasi	0.06
Loyalitas	0.107	Kualitas Hasil Kerja	0.059
Sertifikasi Relevan	0.104	Penghargaan (diluar instansi)	0.054
Kerjasama	0.096	Pendidikan terakhir	0.052
Absensi	0.076		

3.6 Hasil Korelasi

Dalam penelitian ini, pengukuran korelasi dilakukan dengan mengadopsi korelasi pearson. Hasil pengukuran korelasi ditunjukkan pada Gambar 4.

	Kedisiplinan	Masa Kerja	Loyalitas	Sertifikasi Relevan	Kerjasama	Absensi	Kemampuan berkomunikasi	Prestasi	Kualitas Hasil Kerja	Penghargaan (diluar instansi)	Pendidikan Terakhir
Kedisiplinan		0.392	0.509	0.479	0.463	0.507	0.601	0.436	0.439	0.363	0.488
Masa Kerja	0.392		0.353	0.399	0.546	0.642	0.645	0.448	0.446	0.487	0.293
Loyalitas	0.509	0.353		0.536	0.534	0.481	0.532	0.492	0.389	0.615	0.419
Sertifikasi Relevan	0.479	0.399	0.536		0.688	0.478	0.569	0.441	0.524	0.593	0.353
Kerjasama	0.463	0.546	0.534	0.688		0.653	0.505	0.343	0.566	0.611	0.719
Absensi	0.507	0.642	0.481	0.478	0.653		0.478	0.491	0.651	0.486	0.421
Kemampuan berkomunikasi	0.601	0.645	0.532	0.569	0.505	0.478		0.574	0.452	0.622	0.501
Prestasi	0.436	0.448	0.492	0.441	0.343	0.491	0.574		0.574	0.570	0.502
Kualitas Hasil Kerja	0.439	0.446	0.389	0.524	0.566	0.651	0.452	0.574		0.498	0.379
Penghargaan (diluar instansi)	0.363	0.487	0.615	0.593	0.611	0.486	0.622	0.570	0.498		0.441
Pendidikan Terakhir	0.488	0.293	0.419	0.353	0.719	0.421	0.501	0.502	0.379	0.441	

Gambar 4. Hasil Korelasi menggunakan *Pearson Correlation*

Dari hasil pengukuran korelasi, didapatkan bahwa parameter pengukuran kinerja terbaik karyawan adalah pada pasangan parameter berkorelasi Kerjasama dan Pendidikan Terakhir. Ini menunjukkan bahwa dalam pengukuran kinerja, dua parameter ini dapat digunakan sebagai dasar pembobotan dalam pengukuran kualitas kinerja. Selain itu dilakukan analisis statistik terhadap hasil pengukuran korelasi. Hasil pengukuran ditunjukkan pada Tabel 4.

Tabel 4. Hasil Analisis Statistik Korelasi

Analisis	Nilai	Parameter
Nilai Maksimal	0.719	Kerjasama - Pendidikan Terakhir
Nilai Minimum	0.293	Masa Kerja-Pendidikan Terakhir
Nilai rata rata	0.502	
Median	0.492	

4. Kesimpulan

Dalam penelitian ini, mengusulkan analisis pengaruh proses seleksi fitur terhadap pengukuran korelasi kinerja karyawan. Dalam penelitian ini dilakukan pengukuran dengan menggunakan data sebanyak 35 responden. Proses penelitian dilakukan dengan beberapa tahap yaitu, data cleaning, data transformasi, menentukan kelas label, seleksi fitur dan evaluasi pengukuran korelasi. Hasil proses data cleaning menunjukkan bahwa data yang digunakan tidak memiliki nilai kosong. Setiap penilaian pertanyaan berhasil dirubah dalam bentuk numerik dengan skala penilaian adalah 1 sampai dengan 5. Hasil proses seleksi fitur, berhasil mereduksi fitur sebesar 26.27% dari 15 fitur menjadi 11 fitur. Hasil pengukuran korelasi menghasilkan korelasi terbaik yang ditemukan adalah parameter kerjasama-pendidikan terakhir dengan nilai korelasi sebesar 0.719. Sedangkan nilai terendah adalah 0.293 dengan parameter Masa Kerja – Pendidikan terakhir. Hasil yang didapatkan dapat dimanfaatkan untuk pembobotan setiap parameter dalam mengukur kualitas kinerja karyawan.

Penelitian selanjutnya dapat dikembangkan dari sisi penggunaan metode peramalan seperti metode AHP atau SAW untuk memprediksi kinerja berdasarkan analisis kinerja karyawan. Selain itu dapat

dikembangkan pertanyaan yang lebih banyak untuk mengembangkan dimensi fitur dalam pemodelan prediksi.

Daftar Pustaka

- [1] A. B. Candra, "Penilaian Kinerja Karyawan Dalam Unsur Pemetaan SDM Dan Organisasi; Studi Kasus PT XYZ (Tema Performance)," 2022.
- [2] R. Nur Amalia, R. Setia Dianingati, and E. Annisaa, "Pengaruh Jumlah Responden Terhadap Hasil Uji Validitas Dan Reliabilitas Kuesioner Pengetahuan Dan Perilaku Swamedikasi," *Journal of Research in Pharmacy*, vol. 2, no. 1, pp. 9–15, 2022.
- [3] A. R. Djajanegara, "Teknik Analisis Data (Analisis Kualitatif Pada Hasil Kuesioner)," *Jurnal Ilmu Pendidikan dan Dakwah*, vol. 1, no. 1, pp. 1–11, 2019.
- [4] D. Satria and Mushthofa, "Perbandingan Metode Ekstraksi Ciri Histogram dan PCA untuk Mendeteksi Stoma pada Citra Penampang Daun Freycinetia," *Jurnal Ilmu Komputer Agri-Informatika*, vol. 2, no. 1, pp. 20–28, 2019, [Online]. Available: <http://journal.ipb.ac.id/index.php.jika>
- [5] J. Angelina Widiars, H. Santoso Pakpahan, E. Budiman, and M. Soleha, "Klasifikasi Jenis Bawang Menggunakan Metode K-Nearest Neighbor Berdasarkan Ekstraksi Fitur Bentuk dan Tekstur," *JURTI*, vol. 3, no. 2, pp. 139–146, 2019.
- [6] P. M. Prihatini, "Implementasi Ekstraksi Fitur Pada Pengolahan Dokumen Berbahasa Indonesia," *JURNAL MATRIX*, vol. 6, no. 3, pp. 174–178, 2016.
- [7] I. Romli, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Klasifikasi Penyakit Ispa," *Indonesian Journal of Business Intelligence (IJUBI)*, vol. 4, no. 1, pp. 10–15, Jun. 2021, doi: 10.21927/ijubi.v4i1.1727.
- [8] N. Ailmi, Z. Saharuna, and E. Tungadi, "Metode Klasifikasi Pada Aplikasi Pendukung Keputusan Untuk Pemilihan Unit Kegiatan Mahasiswa," in *Prosiding Seminar Nasional Teknik Elektro dan Informatika (SNTEI)*, 2020, pp. 142–147.
- [9] S. B. Mallampati and S. Hari, "Fusion of Feature Ranking Methods for an Effective Intrusion Detection System," *Computers, Materials and Continua*, vol. 76, no. 2, pp. 1721–1744, Aug. 2023, doi: 10.32604/cmc.2023.040567.
- [10] H. Suroyo, "Penerapan Machine Learning dengan Aplikasi Orange Data Mining Untuk Menentukan Jenis Buah Mangga," *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, pp. 342–347, 2019, [Online]. Available: <https://seminar-id.com/semnas-sainteks2019.html>
- [11] N. L. P. M. Putu, Ahmad Zuli Amrullah, and Ismarmiaty, "Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 1, pp. 123–131, Feb. 2021, doi: 10.29207/resti.v5i1.2587.
- [12] R. Chairunisa, Adiwijaya, and W. Astuti, "Perbandingan CART dan Random Forest untuk Deteksi Kanker berbasis Klasifikasi Data Microarray," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 5, pp. 805–812, 2019.
- [13] S. Mohdiwale, M. Sahu, G. R. Sinha, and H. Nisar, "Investigating Feature Ranking Methods for Sub-Band and Relative Power Features in Motor Imagery Task Classification," *J Healthc Eng*, vol. 2021, pp. 1–11, 2021, doi: 10.1155/2021/3928470.
- [14] P. Tahiri, S. Sonia, P. Jain, G. Gupta, W. Salehi, S. Tadjour, "An Estimation of Machine Learning Approaches for Intrusion Detection System", 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2021, doi: 10.1109/ICACITE51222.2021.9404643