

Penggunaan *Machine Learning* Algoritma *Support Vector Machine* (SVM) untuk Mengidentifikasi Kadar Pasir Besi di Kabupaten Aceh Besar

Muhammad Rizki Kana^{1,*}, Nadhiratur Rahmi¹, Mulkal¹

¹ Program Studi Teknik Pertambangan, Fakultas Teknik, Universitas Syiah Kuala, Banda Aceh, Indonesia.

* Corresponding author: rizki820@mhs.usk.ac.id

Received: May 5, 2024; Accepted: Jun 1, 2024

DOI: doi.org/10.31764/jpl.v5i1.23216

Abstrak. Seiring dengan perkembangan zaman, teknologi eksplorasi berkembang dengan sangat pesat. Salah satunya ialah penerapan *machine learning* dalam kegiatan eksplorasi. Penggunaan *machine learning* memungkinkan untuk mendapatkan sebaran mineral pasir besi pada lokasi tertentu dengan menambahkan berbagai parameter yang berpengaruh sehingga mendapatkan *output* berupa keputusan terkait keterdapatan dan kadar mineral pasir besi pada daerah tersebut. Dalam hal ini, penelitian yang dilakukan hanya berfokus pada algoritma *Support Vector Machine* (SVM). SVM merupakan salah satu algoritma *machine learning* yang digunakan untuk tugas-tugas klasifikasi dan regresi. Oleh karena itu, penelitian ini dilakukan untuk mendapatkan model algoritma SVM yang dapat digunakan untuk mengidentifikasi kadar pasir besi dengan menambah beberapa parameter pendukung seperti data jarak titik sampel terhadap pantai, jarak titik sampel terhadap sungai, jarak titik sampel terhadap sesar, nilai pixel, data ketinggian, temperatur, data curah hujan, dan jenis batuan penyusun. Hasilnya grafik regresi linear menunjukkan hubungan nilai kadar Fe prediksi dari model SVM dan kadar Fe aktual, dimana nilai *Root Mean Square Error* (RMSE) adalah 0,076 dan nilai r^2 adalah 0,705. Artinya nilai model algoritma SVM yang dibuat memiliki tingkat kesalahan yang kecil dan korelasi antar data yang kuat sehingga algoritma tersebut dapat dijalankan untuk mengidentifikasi kadar pasir besi.

Kata Kunci: teknologi eksplorasi, SVM, machine learning, pasir besi.

Abstract. Along with the times, exploration technology is developing very rapidly. One of them is the application of machine learning in exploration activities. The use of machine learning makes it possible to obtain the distribution of iron sand minerals in a particular location by adding various influential parameters so as to get output in the form of decisions related to the presence and level of iron sand minerals in the area. In this case, the research conducted only focuses on the Support Vector Machine (SVM) algorithm. SVM is one of the machine learning algorithms used for classification and regression tasks. Therefore, this research was conducted to obtain an SVM algorithm model that can be used to identify iron sand levels by adding several supporting parameters such as data on the distance of the sample point to the beach, the distance of the sample point to the river, the distance of the sample point to the fault, the pixel value, altitude data, temperature, rainfall data, and the type of constituent rock. The result is a linear regression graph showing the relationship between the predicted Fe content value of the SVM model and the actual Fe content, where the Root Mean Square Error (RMSE) value is 0.076 and the r^2 value is 0.705. This means that the value of the SVM algorithm model created has a small error rate and a strong correlation between data so that the algorithm can be run to identify iron sand levels.

Keywords: exploration technology, SVM, machine learning, iron sand.

1. Pendahuluan

Dalam tahapan kegiatan pertambangan, untuk mengidentifikasi keberadaan endapan bahan galian terutama pasir besi, perlu dilakukan proses eksplorasi terlebih dahulu. Kegiatan eksplorasi terdiri dari kegiatan eksplorasi langsung dan kegiatan eksplorasi tidak langsung. Kegiatan eksplorasi langsung

ini merupakan metode eksplorasi yang dilakukan penggalian langsung pada endapan bahan galian yang akan di eksplorasi. Hal ini membuat kegiatan eksplorasi langsung memiliki risiko yang tinggi serta memerlukan biaya yang besar dengan waktu yang cukup lama, sehingga dilakukannya kegiatan eksplorasi tidak langsung bertujuan untuk mengurangi risiko (Shirmard, H, dkk, 2022). Pada kegiatan eksplorasi tidak langsung, pengamatan dilakukan dengan memanfaatkan bantuan teknologi untuk mempersingkat waktu serta menghemat biaya operasional.

Seiring dengan perkembangan zaman, teknologi eksplorasi berkembang dengan sangat pesat. Salah satunya ialah penerapan *machine learning* dalam kegiatan eksplorasi. Uniknya dalam penelitian yang dilakukan oleh Jung, D., & Choi, Y. (2021) sejak 2018 studi *machine learning* telah aktif dilakukan pada industri pertambangan yang sebagian besar untuk eksplorasi mineral.

Machine learning merupakan kecerdasan buatan pada komputer yang dapat mengekstraksi dan mempelajari pola dari suatu data masukan, kemudian dengan pola yang dipelajari tersebut komputer akan mampu mengenali dan memprediksi hasil atau kejadian (Fathurohman, A. 2021). Penggunaan *machine learning* memungkinkan untuk mendapatkan sebaran mineral pasir besi pada lokasi tertentu dengan menambahkan berbagai parameter yang berpengaruh sehingga mendapatkan *output* berupa keputusan terkait keterdapatan dan kadar mineral pasir besi pada daerah tersebut. Dalam hal ini, penelitian yang dilakukan hanya berfokus pada algoritma *Support vector machine* (SVM). SVM merupakan salah satu algoritma *machine learning* yang digunakan untuk tugas-tugas klasifikasi dan regresi. SVM bekerja dengan cara mencari *hyperplane* terbaik (sebuah permukaan pemisah) yang memisahkan dua kelas data dengan margin maksimal. Keunggulan algoritma SVM ini mampu menghasilkan model klasifikasi yang baik meskipun dilatih dengan kumpulan data yang relatif sedikit (Ichwan, M., & Dewi, I. A., 2018). Oleh karena itu, hal inilah yang mendasari penelitian ini menggunakan algoritma SVM untuk memprediksi kadar pasir besi.

Penelitian serupa pernah dilakukan oleh Zandiyyeh, F, dkk (2016) dan Abdulmaleki, M, dkk (2020) terkait penggunaan algoritma SVM pada eksplorasi *Iron Oxide Copper Gold* (IOCG), hanya saja penelitian tersebut terbatas pada pemetaan area deposit, sedangkan pada penelitian ini model algoritma SVM yang dibuat dapat digunakan untuk mengidentifikasi kadar pasir besi di wilayah Aceh Besar.

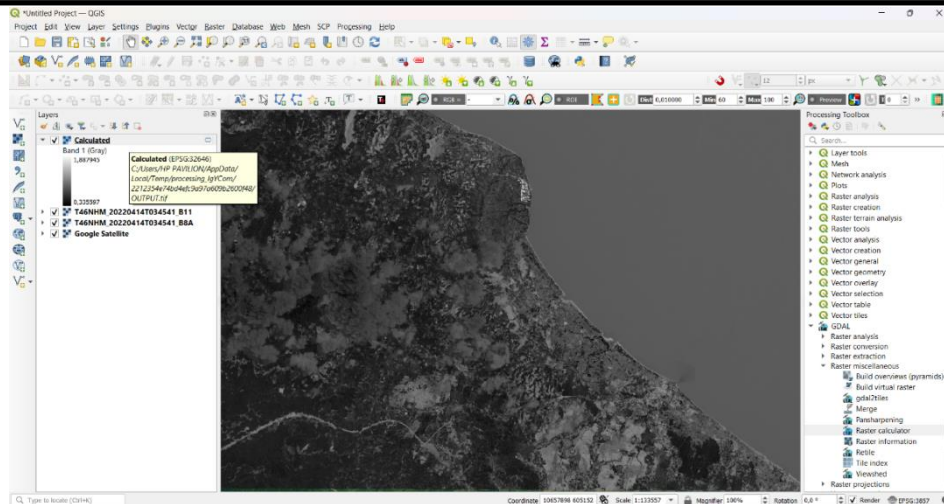
2. Metode Penelitian

Data yang diperlukan dalam penelitian ini terbagi atas data primer dan data sekunder. Data primer yang diperlukan yaitu data kadar pasir besi, data titik koordinat sampel yang ditentukan secara acak berdasarkan hasil identifikasi sebaran pasir besi, juga beberapa parameter pendukung algoritma SVM yang terdiri dari data jarak titik sampel terhadap pantai, jarak titik sampel terhadap sungai, jarak titik sampel terhadap sesar, dan nilai pixel. Data kadar pasir besi diperoleh dari hasil pengujian sampel tanah di laboratorium dengan metode *spektrofotometri UV/VIS*, untuk data terkait titik koordinat yang disebutkan sebelumnya didapat dari hasil pengolahan data menggunakan *software QGIS*, sedangkan data sekunder yang dibutuhkan pada penelitian ini terdiri dari data citra sentinel-2a, data ketinggian, temperatur, data curah hujan, dan jenis batuan penyusun.

Dalam penelitian ini, untuk meninjau keberhasilan model algoritma SVM dalam mengidentifikasi kadar pasir besi dilakukan perbandingan kesesuaian kadar prediksi yang dihasilkan oleh model algoritma SVM dengan kadar aktual. Kesesuaian tersebut diukur dengan metode analisis regresi linear.

2.1. Remote Sensing

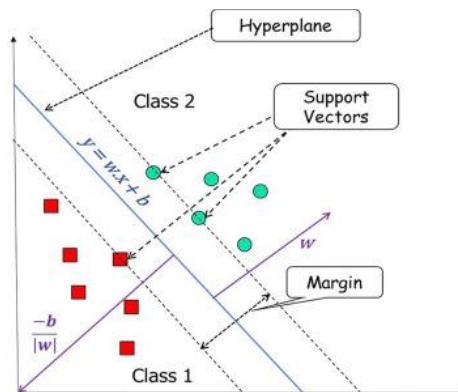
Penginderaan jarak jauh (*remote sensing*) adalah ilmu, teknik, dan seni untuk memperoleh informasi tentang suatu objek, daerah, atau fenomena melalui analisis data yang diperoleh dengan suatu alat tanpa kontak langsung dengan objek, daerah, atau fenomena yang dikaji (Lillesand dan Kieffer, 1994). Penginderaan jarak jauh dibedakan menjadi 2 (dua) sistem, yaitu sistem aktif dan pasif. Penginderaan jarak jauh dengan sistem aktif diperoleh dari satelit yang membawa sensor akti, sedangkan penginderaan jauh sistem pasif diperoleh dari satelit yang membawa sensor pasif. Satelit ini menggunakan energi cahaya matahari sebagai salah satu sumber dalam penyiaran (proses pengambilan gambar). Penelitian ini memanfaatkan citra satelit sentinel-2a band rasio 11/8a dengan persentase awan < 10%. Data citra satelit ini diolah dengan menggunakan *software QGIS*, untuk mendapatkan daerah potensi mineral pasir besi.



Gambar 1. Hasil penggabungan citra band 11 dan 8a.
Fig. 1. The result of merging images of band 11 and 8a.

2.2. Support Vector Machine (SVM)

Support Vector Machine adalah sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah ruang fitur (*feature space*) yang berdimensi tinggi dan mengimplementasikan *learning* bias yang berasal dari teori pembelajaran statistik yang dilatih dengan algoritma pembelajaran. SVM merupakan salah satu algoritma *machine learning* yang digunakan untuk tugas-tugas klasifikasi dan regresi. Jumeilah, F. S. (2017) dalam penelitiannya menjelaskan, secara sederhana konsep SVM adalah usaha mencari *hyperplane* “terbaik” yang berperan penting sebagai garis batas dua buah kelas (SVM mencari *hyperplane* ini berdasarkan *support vectors* dan *margin*. *Support vectors* adalah seluruh vektor data yang berjarak paling mendekati *hyperplane*, sedangkan *margin* menyatakan lebar dari *separating hyperplane*. *Linearly separable data* merupakan data yang dapat dipisahkan secara linier.



Gambar 2. Algoritma SVM
Fig. 2. SVM algorithm.

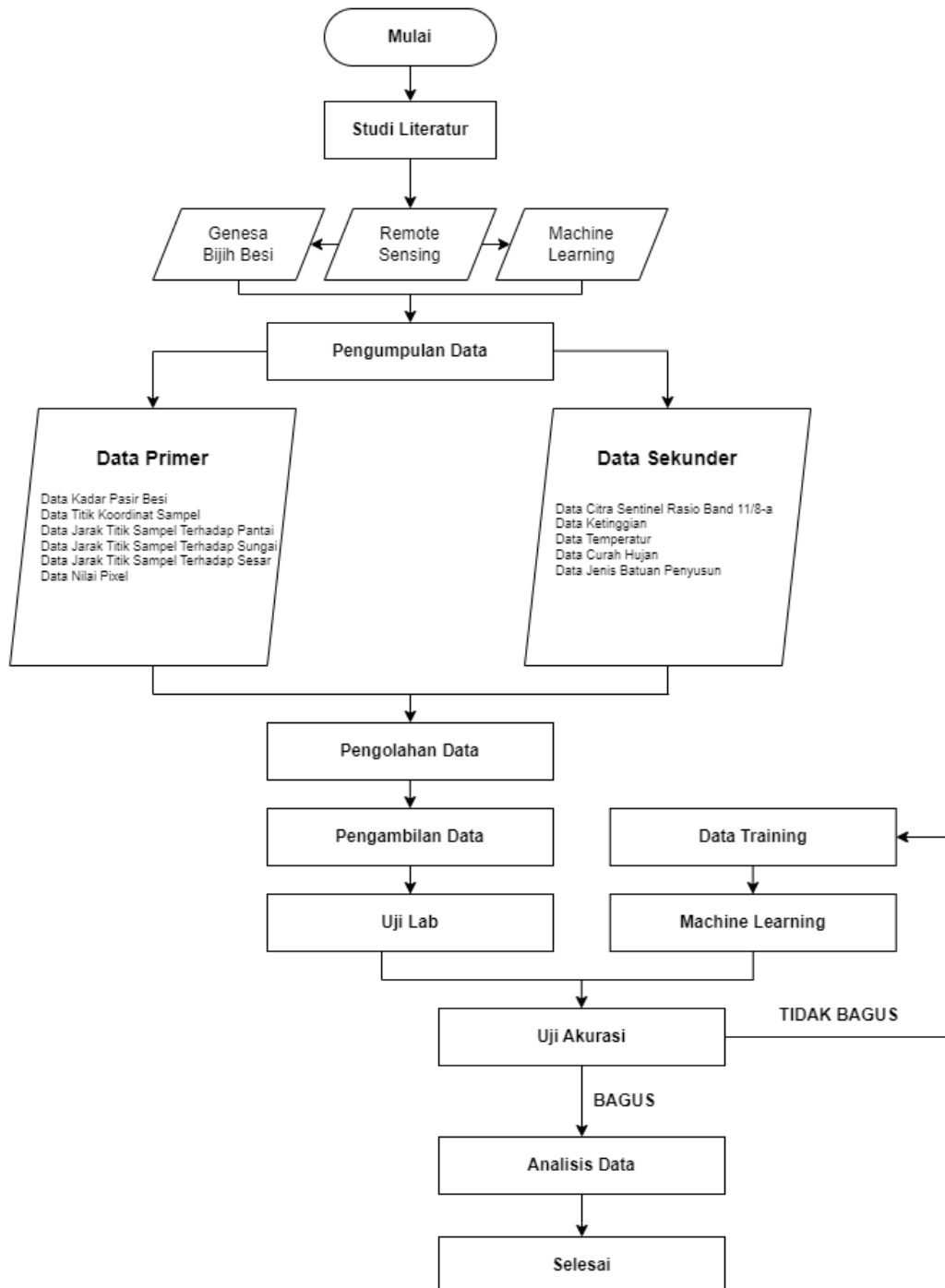
2.3. Regresi Linear

Untuk mengukur kekuatan hubungan antar *variable predictor* X dan *response* Y, dilakukan analisis korelasi yang hasilnya dinyatakan oleh suatu bilangan yang dikenal dengan koefisien korelasi (Yuliara, 2016). Persamaan koefisien korelasi (R) diekspresikan oleh:

$$R = \frac{n(\sum XiYi) - (\sum Xi)(\sum Yi)}{\sqrt{[n(\sum Xi^2) - (\sum Xi)^2][n(\sum Yi^2) - (\sum Yi)^2]}} \quad (1)$$

Korelasi dilambangkan dengan huruf R, yang nilainya berada direntang -1 sampai +1. Nilai R yang mendekati -1 dan +1 menunjukkan hubungan yang kuat antar dua variable, sementara nilai R yang mendekati 0 mengindikasikan hubungan yang lemah (Harlan, J., 2018). Kriteria hasil perhitungan interpretasi korelasi antar dua variable yaitu:

- 0 : Tidak ada korelasi antar dua variabel
- 0 - 0,25 : Korelasi sangat lemah
- 0,25 - 0,5 : Korelasi cukup baik
- 0,5 - 0,75 : Korelasi kuat
- 0,75 - 0,99 : Korelasi sangat kuat



Gambar 3. Diagram alir penelitian.

Fig. 3. Research flow chart.

Selanjutnya dilakukan uji koefisien determinasi yang bertujuan untuk mengetahui seberapa besar garis regresi mendekati sebaran data yang dapat diindikasikan oleh nilai R square (R^2). Semakin besar nilai R^2 semakin baik pula model regresi karena sebagian besar variasi nilai aktual dijelaskan oleh model regresi (Kumar, 2023).

Setelah diketahui koefisien determinasi, selanjutnya dilakukan pengukuran *Root Mean Squared Error* (RMSE) yang merupakan salah satu cara untuk mengevaluasi model regresi linier dengan mengukur tingkat akurasi hasil perkiraan suatu model. Keakuratan metode estimasi dan kesalahan pengukuran ditandai dengan nilai RMSE yang kecil (mendekati nol). Secara matematis RMSE dapat dihitung menggunakan persamaan:

$$RMSE : \sqrt{\frac{\sum(y-y')^2}{n}} \quad (2)$$

Dimana:

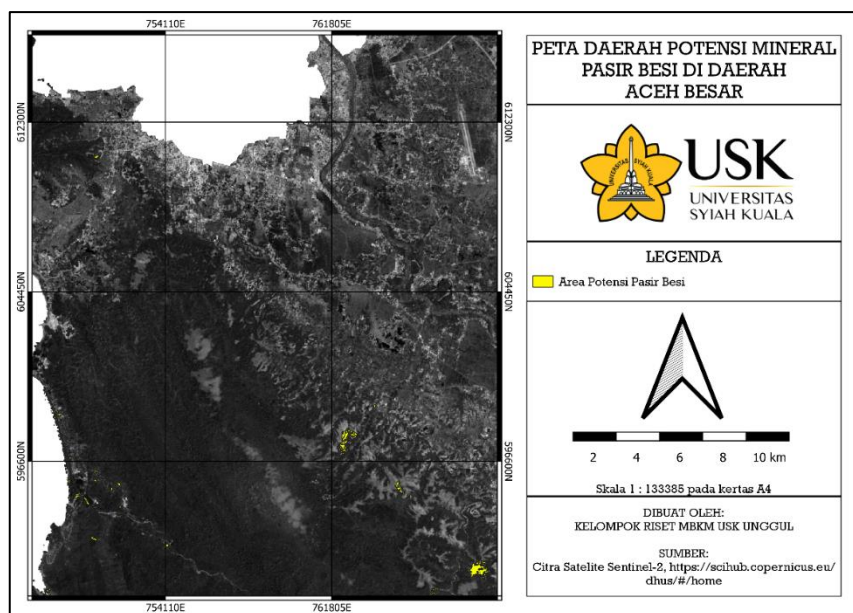
n = jumlah data

y = nilai aktual

y' = nilai prediksi

3. Hasil dan Pembahasan

Pada tahap awal dilakukan pengolahan data citra satelit sentinel-2a dengan menggunakan metode *remote sensing* pada *software* QGIS sehingga didapatkan daerah potensi mineral pasir besi seperti pada gambar 4 berikut. Daerah potensi mineral pasir besi diperoleh dengan melakukan kalkulasi band 11 dan 8a pada data citra satelit sentinel-2a. Pada daerah potensi mineral pasir besi ini selanjutnya ditentukan 30 titik pengambilan sampel untuk memvalidasi hasil metode *remote sensing*. Titik pengambilan sampel dapat dilihat pada gambar 4.

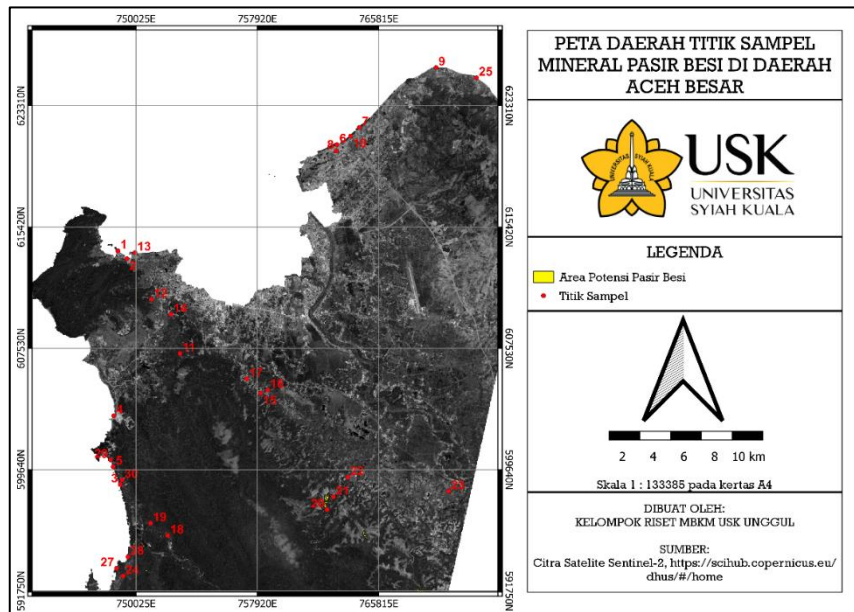


Gambar 4. Daerah potensi mineral pasir besi.

Fig. 4. Iron sand mineral potential area.

Kemudian dilakukan analisis sampel tanah dari titik koordinat yang telah ditentukan sebelumnya dengan menggunakan metode *Spectrofotometri* UV/VIS untuk mengetahui kadar pasir besi di daerah penelitian. Unsur yang dianalisis yaitu besi (Fe) pada sampel tanah yang diambil. Data hasil analisis dapat dilihat pada Tabel 1.

Berdasarkan titik sampel yang sudah ditentukan, dihitung juga beberapa data pendukung untuk algoritma SVM terdiri dari data jarak titik sampel terhadap pantai, jarak titik sampel terhadap sungai, jarak titik sampel terhadap sesar, data ketinggian, temperatur, data curah hujan, dan jenis batuan penyusun. Perhitungan data pendukung dilakukan dengan menggunakan *software* QGIS. Ringkasan informasi data yang telah didapatkan dapat dilihat pada Tabel 1 dan Tabel 2.



Gambar 5. Daerah titik sampel mineral pasir besi.

Fig. 5. Iron sand mineral sample point area.

Tabel 1. Data input parameter *machine learning* bagian 1.

Table 1. Machine learning parameter input data part 1.

No sampel	Temperatur	Curah hujan (mm)	Nilai pixel 11/8a	Kadar besi (Fe) (%)	Simbol geologi
1	25,9	262	1,189911142	0,017363	Qh
2	25,6	276	1,211868546	0,017047	Qh
3	25,7	246	0,930390568	0,014055	Qh
4	25,9	252	1,161739611	0,019335	Murlr
5	25,9	247	1,010453841	0,021616	Qh
6	26,4	266	1,100252784	0,019834	Qh
7	26,4	266	1,261100873	0,023016	Qh
8	26,4	269	1,187602544	0,018043	Qh
9	26,2	269	1,2182064	0,021343	QTvt
10	26,4	268	1,203543121	0,017615	Qh
11	25,8	261	1,224031394	0,306864	Mul
12	26,1	258	1,276501638	0,218682	Murlr
13	26,1	262	1,169132445	0,015686	Qh
14	25,9	268	1,182942559	0,247487	Murlr
15	25,8	251	1,212584169	0,358227	Murlr
16	26,0	242	1,310901404	0,312668	Murlr
17	25,9	248	1,203169669	0,303213	Murlr
18	25,4	259	0,707993044	0,995180	Qh
19	25,9	243	0,817114646	0,317751	Qh
20	24,9	276	0,974413572	0,376702	Murl
21	25,0	274	0,742344516	0,319987	Murl
22	25,7	246	1,139239493	0,336084	Murl
23	26,3	230	1,19287442	0,353309	Qpin
24	25,6	251	0,762419522	0,320404	Qh
25	26,2	268	1,211666356	0,717385	Qh

26	26,1	262	1,128477656	0,760111	Qh
27	26,0	241	0,925429376	0,276780	Qh
28	26,0	242	0,857244979	0,725492	Qh
29	25,9	247	1,197589971	0,630530	Qh
30	25,9	255	0,968839928	0,702072	Qh

Tabel 2. Data input parameter *machine learning* bagian 2.**Table 2.** *Machine learning* parameter input data part 2.

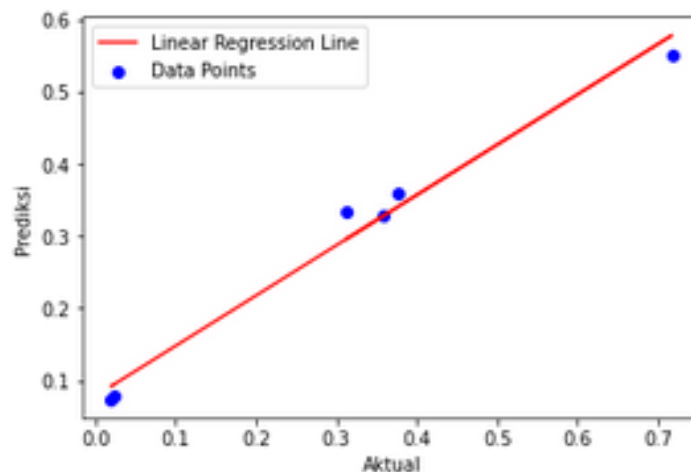
No Sampel	X	Y	Jarak dari sesar (m)	Jarak dari Pantai (m)	Jarak dari Sungai (m)	Ketinggian (m)
1	748847,0723	613874,5111	1000,614	5,698	190,488	3,558714628
2	749431,8986	613370,2794	886,188	562,992	198,648	2,15089345
3	749005,8703	598666,2486	1153,238	35,218	72,617	5,48543644
4	748585,818	603145,9469	995,01	19,012	26,505	12,95004368
5	748551,5595	599825,2493	1178,655	34,345	93,667	8,334195137
6	763099,8869	620786	5859,263	29,483	468,695	1,514180541
7	764556,7861	621897,7898	4701,417	60,244	384,037	3,590601683
8	763062,6895	620372,6952	5618,891	359,335	728,938	2,559281349
9	769547,7136	625792,7756	1367,384	1,466	960,665	4,215833187
10	763986,4256	621334,3176	5370,09	56,47	44,921	2,803016663
11	752891,6422	607192,6727	83,375	5848,978	192,483	55,2128067
12	751031,541	610716,4653	3057,727	3179,524	2254,387	32,54456329
13	749953,9902	613756,4502	1534,921	6,798	21,827	3,044605255
14	752298,268	609739,5477	1841,536	4588,005	2056,666	63,89227676
15	758127,6864	604607,4568	229,089	9626,724	129,362	62,88961029
16	758596,1908	604819,4756	253,909	10121,942	155,585	29,70429993
17	757229,1808	605544,6734	63,663	8935,813	313,304	68,84286499
18	752088,7242	595344,4518	1790,419	2459,616	116,783	15,8749752
19	750973,7621	596143,1899	1736,234	1422,352	274,214	6,306734085
20	762453,8957	597036,2953	452,318	12927,404	121,116	270,8105774
21	762857,8991	597862,2941	41,102	13429,778	279,878	203,7782593
22	763818,0204	599138,0942	340,518	14581,525	234,609	87,48858643
23	770362,2206	598228,6777	2529,531	20919,375	807,943	41,4730072
24	749175,319	592714,2985	145,883	599,236	350,175	25,20996475
25	772172,7568	625132,2817	1643,994	1,08	246,536	9,209646225
26	775024,0024	621728,1385	1750,321	10,241	357,247	3,773314476
27	748765,1078	593208,7653	759,623	7,575	675,313	6,550606728
28	749503,3679	593961,8851	558,287	30,078	385,454	5,931954384
29	748365,6871	600272,5631	994,799	65,402	164,083	7,075970173
30	749117,1139	598969,0736	865,643	229,942	102,39	6,702882767

Selanjutnya, data tersebut dijadikan sebagai parameter analisis yang di input pada algoritma SVM untuk membuat model yang bisa menentukan nilai kadar Fe dari koordinat yang di tentukan. Berdasarkan data yang di input pada algoritma SVM diperoleh prediksi kadar pasir besi yang dapat dilihat pada Tabel 3.

Tabel 3. Data prediksi kadar Fe.
Table 3. Predicted Fe content data.

No. sampel	Kadar Fe prediksi (%)	Kadar Fe aktual (%)
14	0,32963268	0,247487
15	0,33456074	0,358227
19	0,36031314	0,317751
24	0,55132607	0,320404
3	0,07369387	0,014055
6	0,07982548	0,019834

Dalam pelaksanaannya pada algoritma SVM dengan menggunakan 30 data titik koordinat, sebanyak 24 data digunakan untuk data *training* dan 6 data untuk data *testing*. Data *testing* yang di hasilkan dapat dilihat pada Tabel 3 kolom kadar Fe prediksi. Data *testing* inilah yang akan menjadi pembandingan kesesuaian dengan kadar aktual. Dari hasil tabel diatas, dapat dilihat perbedaan antara kadar Fe prediksi dan aktual tidak menunjukkan perbedaan yang signifikan. Dari enam data yang dimunculkan, pada data nomor sampel 15 dan 19 menunjukkan hasil yang nyaris sama antara kadar Fe prediksi dan aktual. Setelah berhasil membuat model algoritma SVM untuk menentukan kadar Fe. Hasil dari model algoritma SVM tersebut di uji untuk melihat akurasi hasil yang didapatkan dengan data aktual menggunakan analisis regresi linear.



Gambar 6. Grafik regresi linear hubungan nilai kadar Fe prediksi dan kadar Fe aktual.

Fig. 6. Linear regression graph of the relationship between predicted and actual Fe values.

Pada gambar grafik regresi linear di atas menunjukkan hubungan nilai kadar Fe prediksi dari model SVM dan kadar Fe aktual, dimana nilai *Root Mean Square Error* (RMSE) adalah 0,076 dan nilai r^2 adalah 0,705. Artinya nilai model algoritma SVM yang dibuat memiliki tingkat kesalahan yang kecil dan korelasi antar data yang kuat sehingga algoritma tersebut dapat dijalankan untuk mengidentifikasi kadar pasir besi. Namun disisi lain keterbatasan data membuat model algoritma SVM tidak optimal dan diperlukan kumpulan data yang lebih banyak serta lebih bervariasi sehingga model algoritma yang terbentuk mendapatkan referensi lebih banyak dalam menentukan kadar pasir besi untuk hasil yang lebih akurat.

4. Kesimpulan

Kesimpulan yang diperoleh dari penelitian ini antara lain:

1. Penggunaan *machine learning* dengan algoritma SVM untuk mengidentifikasi kadar pasir besi berhasil dilakukan.
2. Berdasarkan analisis regresi linear algoritma SVM yang terbentuk memiliki tingkat kesalahan yang kecil dan korelasi antar data yang kuat.
3. Untuk mendapatkan model algoritma SVM yang akurat dalam mengidentifikasi pasir besi diperlukan data yang lebih banyak dan bervariasi.

Ucapan Terima Kasih

Ucapan terima kasih kami sampaikan kepada pihak penyelenggara program Riset MBKM USK Unggul yang telah mendukung penelitian ini serta kepada Tim Penelitian MBKM USK Unggul.

Referensi

- Abdulmaleki, M., Rasmussen, T. M. & Pal, M. K., 2020. Exploration of IOCG Mineralization Using Integration of Space-Borne Remote Sensing Data with Airborne Geophysical Data. *Remote Sensing and Spatial Information Sciences*, Volume XLIII-B3-2020, pp. 9-16.
- Fathurohman, A. (2021). Machine Learning Untuk Pendidikan: Mengapa dan Bagaimana. *Jurnal Informatika Dan Teknologi Komputer (JITEK)*, 1(3), 57-62.
- Harlan, J., 2018. *Analisis Regresi Linear*, Depok: s.n.
- Ichwan, M. & Dewi, I. A., 2018. Klasifikasi Support Vector Machine (SVM) untuk Menentukan Tingkat Kemanisan Mangga Berdasarkan Fitur Warna. *MIND (Multimedia Intelligent Networking Database) Journal*, 3(2), pp. 16-23.
- Jumeilah, F. S., 2017. Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian. *Jurnal RESTI (Rekayasa Sistem dan Informasi Teknologi)*, 1(1), pp. 19-25.
- Jung, D. & Choi, Y., 2021. Systematic Review of Machine Learning Applications in Mining: Exploration, Exploitation, and Reclamation. *Minerals*, 11(2), p. 148.
- Kumar, A., 2023. *Mean Squared Error or R Square - Which One to Use?*. [Online] Available at: <https://vitalflux.com/mean-square-error-r-squared-which-one-to-use/> [Accessed 09 Februari 2023].
- Lillesand, T. M. & Kiefer, R. W., 1994. *Remote Sensing Interpretation*, Chichester: s.n.
- Shirmard, H., Farahbakhsh, E., Muller, R. D. & Chandra, R., 2022. A Review of Machine Learning in Processing Remote Sensing Data for Mineral Exploration. *Remote Sensing of Environment*, 268(112750).
- Zandiyyeh, F., Shayestefar, M. R., Ranjbar, H. & Saadat, S., 2016. Prospectivity Mapping of Iron Oxide-Copper-Gold (IOCG) Deposits using Support Vector Machine Method in Feyzaabad Area (East of Iran). *Journal of Himalayan Earth Sciences*, 49(2), pp. 50-62.
-