

**OPTIMASI ALGORITMA K-NEAREST NEIGHBOR MENGGUNAKAN
ALGORITMA GENETIC DALAM MENGLASIFIKASI BIDANG USAHA MIKRO
DAN MENENGAH DI DESA SENENAN**

***OPTIMIZATION OF THE K-NEAREST NEIGHBOR ALGORITHM USING GENETIC
ALGORITHM IN CLASSIFYING MICRO AND MEDIUM ENTERPRISE FIELDS IN
SENEAN VILLAGE***

ENGLISH ARTICLE TITLES (uppercase, times new roman, 12pt, italic, bold, and centered)

Retno Setyawati^{*1}, Teguh Tamrin², Nur Aeni Widiastuti³

^{1,2,3}Universitas Islam Nahdlatul Ulama Jepara

Email : ¹reajpr5@gmail.com, ²teguh@unisnu.ac.id, ³nuraeniwidiastuti@unisnu.ac.id

Abstrak

Pemberdayaan UMKM ditengah arus globalisasi dan tingginya persaingan membuat UMKM harus mampu mengadaptasi tantangan global. Seperti, inovasi produk, baik berupa barang dan jasa. Pelatihan UMKM yang diberikan kepada pelaku usaha merupakan hal yang penting agar dapat bersaing secara sehat dan berkembang pesat di wilayah Jepara. Dalam penelitian ini peneliti memilih metode Optimasi *k- nearest neighbor* menggunakan *algoritma genetic* (K-Nn+Ga) karena dapat mengatasi masalah klasifikasi dengan baik. Tujuan dari peneliti ini adalah menerapkan algoritma K-Nn+Ga untuk mendapatkan nilai *k* yang optimal sehingga dapat memberikan prediksi yang lebih akurat. Data yang doleh pada *tools Rapid Minner* dengan 6 atribut yaitu nama, jenis usaha, status usaha, bidang usaha, dan dilakukan pengujian model menggunakan *k- fold cross validation* serta *confusion matrix* untuk evaluasi dan validasi hasil akurasi sebesar 85,19%.

Kata Kunci: UMKM, *k- nearest neighbor*, *algoritma genetic*, *RapidMinner*

Abstract

Empowerment of MSMEs in the midst of global challenges. Like, product innovation, both in the form of goods and services. MSME training provided to business actors is important in order to compete in a healthy manner and develop rapidly in the Jepara region. In this study the researchers chose the k-nearest neighbor optimization method using the genetic algorithm (K-NN+GA) because it can solve classification problems well. The aim of this research is to apply the K-Nn+Ga algorithm to get the optimal value of k so that it can provide more accurate predictions. Data obtained on Rapid Minner tools with 6 attributes, namely name, type of business, business status, line of business, and model testing was carried out using k-fold cross validation and a confusion matrix for evaluation and validation of results with an accuracy of 85.19%.

Keywords: UMKM, *k- nearest neighbor*, *algoritma genetic*, *RapidMinner*

1. PENDAHULUAN (huruf besar, 10pt, tebal)

Usaha Mikro Kecil dan Menengah (Umkm) merupakan bidang usaha yang berkembang dan konsisten dalam perekonomian dalam sektor Nasional. Umkm menjadikan wadah baik bagi pelaku usaha yang menciptakan lapangan pekerjaan yang produktif dan berkelanjutan. Usaha Mikro Kecil dan Menengah disebut sebagai sektor ekonomi nasional yang strategis di mata masyarakat kita yang dapat menjadi tulang punggung serta peningkatan atau bahkan penurunan perekonomian nasional. Umkm juga merupakan sebuah kelompok pelaku usaha perekonomian terbesar di dalam perekonomian Indonesia yang telah terbukti menjadi pengaman perekonomian Nasional dalam masa krisis era revolusi pasca pandemi COVID[1].

Umkm sendiri memiliki arti usaha yang bersifat padat karya, yaitu tidak membutuhkan persyaratan tertentu seperti, keahlian (keterampilan) khusus, tingkat Pendidikan, dan modal usaha relatif sedikit tergantung sesuai apa target usaha tersebut, dan teknologi yang digunakan cenderung sederhana. Hal tersebut sangat diperlukan karena di era milenial ini dunia semakin kompetitif dalam persaingan baik secara digital ataupun secara Tradisional[2].

Struktur perekonomian Kabupaten Jepara menurut lapangan usaha tahun 2021 didominasi oleh empat lapangan usaha utama, yaitu: Industri Pengolahan (35,11 persen), Perdagangan Besar dan Eceran; Reparasi dan Perawatan Mobil dan Sepeda Motor (16,58 persen), Pertanian, Kehutanan dan Perikanan (13,05 persen), dan Konstruksi (7,57 persen)[3]. Artinya dalam hal ini meningkatnya usaha mikro kecil dan menengah dalam sektor Umkm dapat bersaing secara sehat dan berkembang pesat di wilayah Jepara, mengingat Jepara dalam masa peralihan semenjak tahun sebelumnya hingga akhir tahun 2021 setelah pandemic berkepanjangan terdapat peningkatan yang signifikan dalam sektor Umkm atau lapangan pekerjaan, tanpa pengurangan ketenaga kerjaan atau PHK yang drastis di Jepara.

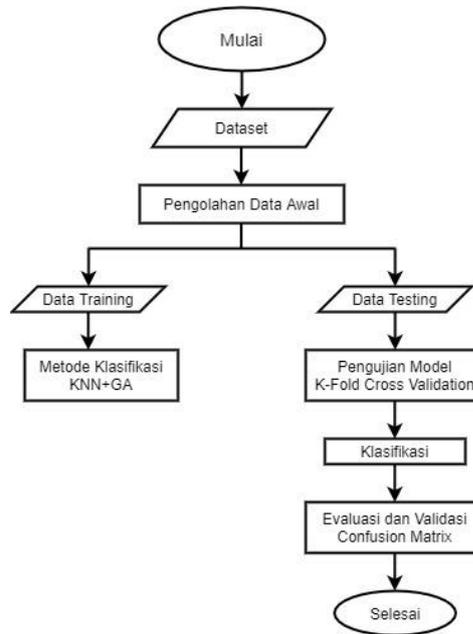
Pelatihan Umkm adalah pelatihan yang diberikan kepada pelaku usaha dan tiap tahunnya dinas Diskopukmnakertrans melakukan kegiatan yang salah satunya mengadakan kegiatan pelatihan. Beberapa tema pelatihan yang dibuat Diskopukmnakertrans. Diantaranya: pelatihan penyusunan paket wisata, pelatihan tata rias pengantin, pelatihan packaing kopi, pelatihan pemeliharaan dan perbaikan Ac untuk rumah tangga, pelatihan menjahit, pelatihan mode pakaian dewasa, dan berbagai pelatihan lainnya demi berkembangnya Umkm. Dalam penelitian ini memfokuskan pada pihak petugas Diskopukmnakertrans yang memiliki permasalahan dalam menggelar pelatihan karena waktu yang singkat, persiapan kurang maksimal, sehingga kurang tepat sasaran dan belum adanya data pengolahan Umkm karena data masih campuran, juga belum adanya kategori dibidang usaha pada data Umkm di Desa Senenan. Dari data Umkm agar bisa tepat sasaran data akan di olah menggunakan data mining guna meminimalisir dan beberapa algoritma khususnya k-nn untuk mengatasi permasalahan karena merupakan metode klasifikasi dengan prinsip kerja mencari jarak terdekat pada data training dan data testing. sehingga mengeluarkan sebuah pengetahuan. Serta mampu memberikan pelayanan yang sebaik - baiknya terhadap masyarakat khususnya masyarakat pelaku usaha ekonomi.

Penelitian terkait k-nearest neighbor telah banyak dilakukan, salah satunya penelitian menjelaskan bahwa Metode klasifikasi *K-Nearest Neighbor* memiliki kelebihan diantaranya pelatihan cepat, sederhana, tahan pada data pelatihan yang memiliki noise (gangguan) dan lebih efektif apabila data latih yang berjumlah besar. Namun K-Nn perlu menentukan nilai dari parameter K (jumlah dari tetangga terdekat), untuk mendapatkan hasil terbaik, Berdasarkan penjelasan diatas, penulis tertarik melakukan penelitian yang berjudul " Optimasi Algoritma K-Nearest Neighbor Menggunakan Algoritma Genetic dalam Mengklasifikasi Bidang Usaha Mikro dan Menengan di Desa Senenan " dengan menggunakan algoritma *K-Nearst Neighbour*. dengan bantuan penegasan *Genetic* Algoritma sebagai pembobotan atribut pada data Umkm sehingga diharapkan dapat menjadi rekomendasi, untuk meresume, menganalisa atau sebagai bahan dalam perkembangan yang akan datang bagi pihak diskopukmnakertrans dalam melaksanakan kegiatan

atau pelatihan Umkm. Sehingga mengeluarkan sebuah pengetahuan yaitu klasifikasi dibidang usaha tersebut dengan evaluasi hasil algoritma menggunakan *confusion matrix*.

2. METODE PENELITIAN

Penelitian ini dilakukan di desa Senenan Kecamatan Tahunan Kabupaten Jepara. Dengan tahapan penelitian sebagai berikut



Gambar 2.1 Tahapan Penelitian

1. Dataset Penelitian

Data set yang digunakan adalah data UMKM tahun 2020-2021 pada Desa Senenan Jepara berjumlah 871 data. Memiliki 5 atribut yaitu nama, jenis usaha, usaha, status usaha, dan bidang usaha.

2. Pengolahan Data Awal

Tahap ini dilakukan tahap awal dan tahap penting dalam penelitian. Metode yang digunakan dalam penelitian ini adalah transformasi data menggunakan normalisasi

$$V' = \frac{V - \min A}{\max A - \min A} \cdot (\text{new max } A - \text{new min } A) + \text{new min } A$$

Gambar 22 Rumus Normalisasi min-max

min-max yang merupakan transformasi linier terhadap data asli sehingga menghasilkan nilai keseimbangan nilai perbandingan antar data saat sebelum dan sesudah proses. Metode ini dapat menggunakan rumus sebagai berikut:

1. Klasifikasi

Tahap klasifikasi merupakan tahap untuk mengklasifikasikan kualitas pada dataset Umkm. Tahap klasifikasi sebagai berikut:

- a. K-fold Cross Validation: merupakan sebuah metode *statistic* yang berfungsi untuk mengevaluasi kinerja suatu algoritma atau model[8]. Pembagian data dilakukan menggunakan K-fold Cross Validation dengan nilai K sama dengan 10.
- b. K-Nn + Ga: setelah pembagian data uji dan latih, dilanjutkan proses klasifikasi untuk meningkatkan nilai akurasi

3. Evaluasi

Tahap evaluasi dilakukan dengan cara menganalisa akurasi dataset. Perhitungan akurasi dilakukan dengan cara membagi jumlah data uji yang benar (true positif dan true negative) dengan jumlah data uji keseluruhan kemudian dikalikan dengan 100%

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

Gambar 3.3 Confusion Matrix

3. HASIL DAN PEMBAHASAN

Tahap ini merupakan penguraian hasil penelitian yang diperoleh beserta penjelasannya.

1. Praproses

Dalam tahap ini dilakukan praproses terhadap dataset umkm. Menggunakan min-max normalization.

Tabel 3.1 Dataset UKM

| NO | NAMA | A1 | A2 | A3 | KET |
|-----|--------------|-----|-----|-----|----------------|
| 1 | Abdul Haris | 1 | 1 | 1 | Furniture |
| 2 | Abdul Kholik | 2 | 2 | 1 | Toko Kelontong |
| 3 | Abdul Majid | 3 | 3 | 1 | Kuliner |
| 4 | Abdul Rohman | 4 | 4 | 2 | Furniture |
| 5 | Abdul Zaed | 5 | 1 | 1 | Jasa |
| ... | ... | ... | ... | ... | ... |
| 871 | Zuningsih | 35 | 8 | 1 | Kuliner |

Min-Max Normalization: pada table 1 dapat dilihat nilai asli dari dataset sebelum di praproses. Banyak data yang memiliki rentang berbeda sehingga harus dilakukan normalisasi. Dataset tersebut harus ditransformasi menggunakan metode *Min-Max Normalization* dengan mengolah nilai minimum dan maximum dari setiap atribut. Hasil perubahan nilai dengan metode ini dapat dilihat pada table 2. Dataset ukm setelah dilakukan *Min-Max Normalization*

| NO | NAMA | A1 | A2 | A3 | KET |
|-----|--------------|-------|-------|-----|----------------|
| 1 | Abdul Haris | 1 | 1 | 1 | Furniture |
| 2 | Abdul Kholik | 1,006 | 1,004 | 1 | Toko Kelontong |
| 3 | Abdul Majid | 1,013 | 1,008 | 1 | Kuliner |
| 4 | Abdul Rohman | 1,019 | 1,013 | 1 | Furniture |
| 5 | Abdul Zaed | 1,026 | 1,038 | 1 | Jasa |
| ... | ... | ... | ... | ... | ... |
| 87 | ... | 1,218 | ... | ... | ... |
| 1 | Zuningsih | 8 | 1,029 | 1 | Kuliner |

Dari permodelan normalisasi data diatas dilihat. Pada tahapan ini peneliti menentukan batas nilai minimum 1, sedangkan batas nilai maksimumnya adalah 2.

2. Klasifikasi

Setelah praproses data dan sebelum dilakukan klasifikasi, data dibagi menjadi data uji dan data latih terlebih dahulu. Pembagian data dilakukan dengan k-fold cross validation. K yang digunakan pada k-fold sama dengan 10. Tahap selanjutnya adalah melakukan klasifikasi terhadap dataset umkm yang telah di praproses menggunakan algoritma K-Nn+Ga. K yang digunakan yaitu: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21. Perhitungan jarak pada algoritma ini dapat dilihat pada persamaan 3. Berikut hasil table penelitian. Hasil akurasi dengan metode K-Nn+Ga

| | K=3 | K=5 | K=7 | K=9 | K=11 | K=13 | K=15 | K=17 | K=19 | K=21 |
|----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1 | 83,36% | | | | | | | | | |
| 2 | | 85,19% | | | | | | | | |
| 3 | | | 84,50% | | | | | | | |
| 4 | | | | 83,01% | | | | | | |
| 5 | | | | | 80,30% | | | | | |
| 6 | | | | | | 78,88% | | | | |
| 7 | | | | | | | 77,84% | | | |
| 8 | | | | | | | | 77,27% | | |
| 9 | | | | | | | | | 77,74% | |
| 10 | | | | | | | | | | 77,27% |

Tahap akhir pada proses ini adalah penghitungan akurasi. Hasil akurasi dapat dilihat pada table III. Bahwa akurasi tertinggi ada di K=5 tingkat akurasi adalah 85,19%.

4. KESIMPULAN

Hasil penelitian menunjukkan bahwa model KNN+Ga menggunakan 5 atribut yang dapat mempengaruhi perhitungan berdasarkan kriteria. Atribut yang tidak mempunyai pengaruh pada akurasi akan dihilangkan sehingga menghasilkan nilai akurasi 85,19% dengan nilai $k=5$ dari pengujian model yang dilakukan menggunakan *K-fold validation* sebanyak 10 kali. Dibandingkan model KNN tanpa Ga hanya menghasilkan akurasi 74,28%. Dan KNN tanpa Normalisasi dan Ga hanya menghasilkan akurasi 75,09%. Dengan demikian Genetic Algoritma yang diterapkan pada pembobotan atribut KNN meningkatkan nilai akurasi.

Memberikan pernyataan bahwa apa yang diharapkan, seperti yang dinyatakan dalam "Pendahuluan" akhirnya dapat mengakibatkan "Hasil dan Diskusi", sehingga ada komparabilitas. Selain itu dapat juga ditambahkan prospek pengembangan hasil penelitian dan prospek penerapan studi lanjutan.

DAFTAR PUSTAKA

Referensi yang digunakan adalah jurnal internasional, jurnal nasional dan buku dengan

- T. Sudrartono *et al.*, *Kewirausahaan UMKM di era Digital*. Bandung: Widhia Bhakti Persada Bandung, 2022.
- Gunadi, G., & Sensuse, d. I. (2012). Penerapan metode data mining market basket analysis terhadap E. Suwarni and M. A. Handayani, "Strategi Pengembangan Bisnis Usaha Mikro Kecil Menengah Keripik Pisang Dengan Pendekatan Business Model Kanvas," *Manag. Business, Account.*, vol. 19, p. 3, 2020, doi: 10.33557/mbia.v19i3.1177.
- "Pertumbuhan Ekonomi Kabupaten Jepara 2021," 29 Maret, 2022.
<https://jeparakab.bps.go.id/pressrelease/2022/03/29/75/pertumbuhan-ekonomi-kabupaten-jepara-2021.html> (accessed Nov. 13, 2022).
- T. Khotiah, Y. Sugianela, D. Setiowati, and F. Arianto, "Model Klasifikasi Data Remisi Narapidana Menggunakan Algoritma K-Nearest Neighbors dengan Optimasi Algoritma Genetika," *Inf. Technol. J. UMUS*, vol. 4, no. 01, 2022, [Online]. Available: <http://jurnal.umus.ac.id/index/php/intech>.
- R. Primartha, *Algoritma Machine Learning*, Pertama. 2021.
- T. H. Fratiwi, M. Sudarma, and N. Pramaita, "Sistem Klasifikasi Musik Gamelan Angklung Bali terhadap suasana hati menggunakan Algoritma K-Nearest Neighbor Berbasis Algoritma Genetika," *Maj. Ilm. Teknol. Elektro*, vol. 20, no. 2, pp. 265–272, 2021, doi: 10.24843/MITE.2021.v20i02.P10.
- S. Harlina and M. O. Kadang, "Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi kelayakan Calon Nasabah Kredit Berbasis Web."
- "PERBANDINGAN NORMALISASI DATA UNTUK KLASIFIKASI WINE MENGGUNAKAN ALGORITMA K-NN," *J. Comput. Eng. Syst. Sci.*, vol. 4, no. 1, pp. 78–82, 2019, doi: 10.24114/cess.v4i1.11458.