

## **Model Sistem Cerdas untuk Rekomendasi Market Matching Negara Tujuan Ekspor bagi UMKM**

### ***Smart System Model for Market Matching Recommendations for Export Destination Countries for MSMEs***

**Eka Yuniar<sup>1\*</sup>**  
**Weda Adistianaya Dewa<sup>2</sup>**  
**Fakhruddin Arozi<sup>3</sup>**  
**Ahmad Mihdan Advani<sup>4</sup>**

<sup>1,2</sup>Sistem Informasi, STMIK PPKIA Pradnya Paramita

<sup>3,4</sup>Teknologi Informasi, STMIK PPKIA Pradnya Paramita

<sup>1</sup>eka@stimata.ac.id, <sup>2</sup>weda@stimata.ac.id

<sup>3</sup>fakhrudddin\_21520023@stimata.ac.id, <sup>4</sup>ahmad\_21520005@stimata.ac.id

#### **\*Penulis Korespondensi:**

Eka Yuniar  
eka@stimata.ac.id

#### **Riwayat Artikel:**

Diterima : 8 September 2022

Direview : 10 November 2022

Disetujui : 5 Desember 2022

Terbit : 19 Desember 2022

#### **Abstrak**

Ekspor merupakan salah satu cara untuk dapat melakukan peningkatan kinerja dan pertumbuhan bisnis UMKM. Namun, memasuki pasar ekspor tidak semudah yang dibayangkan. UMKM harus mempertimbangkan banyak faktor seperti persyaratan peraturan negara tujuan, preferensi konsumen, dan persaingan dengan produsen lokal di negara tujuan. Selain itu, beberapa kendala ekspor dari produk UMKM adalah pada saat perputaran produk yang cenderung lama di negara importir yang disebabkan karena stok produk tidak habis terjual di Negara tersebut dan terkadang sampai mencapai batas kadaluarsa. Jumlah UMKM Indonesia mencapai 65,5 juta yang merupakan terbanyak di negara ASEAN. Selain itu, kontribusi UMKM tercatat mencapai kisaran 61% terhadap PDP nasional dan menyerap 97% dari total tenaga kerja pada tahun 2021. Untuk membantu UMKM dalam memilih negara tujuan ekspor, dibutuhkan sistem cerdas yang dapat merekomendasikan negara tujuan yang tepat berdasarkan karakteristik produk UMKM. Adapun cara menentukan negara Tujuan Ekspor dengan menggunakan salah satunya berdasarkan situs Trademap. Urgensi penelitian ini bertujuan untuk membangun model sistem cerdas berbasis machine learning dengan menggunakan metode KNN untuk merekomendasikan negara tujuan ekspor yang tepat bagi UMKM. Model sistem cerdas berbasis machine learning ini ditujukan agar menjadi solusi bagi UMKM untuk memperluas pasar mereka ke luar negeri dengan efektif. Penelitian ini juga diharapkan dapat memberikan kontribusi untuk pengembangan sistem cerdas yang dapat membantu UMKM dalam memilih negara tujuan ekspor yang tepat dan meningkatkan efektivitas ekspor. Hasil penelitian menunjukkan tingkat akurasi sebesar 99%, dengan tingkat precision 100%, recall 100%, f1-score 100%.

**Kata Kunci:** ekspor; Key Nearest Neighbour; KNN; Sistem Cerdas; UMKM.

*Cite: Yuniar, E, dkk. (2023). Model Sistem Cerdas untuk Rekomendasi Market Matching Negara Tujuan Ekspor bagi UMKM. SMATIKA : STIKI Informatika Jurnal, 13(2). doi: <https://doi.org/10.32664/smatika.v13i02.943>*

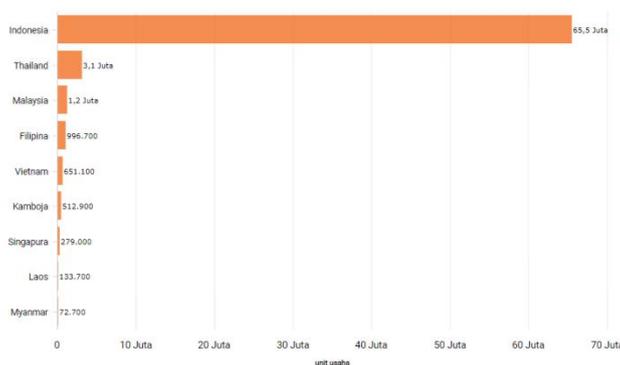
### Abstract

Export is one way to improve the performance and growth of the MSME business. However, entering the export market is not as easy as one might think. MSMEs have to consider many factors such as regulatory requirements of the destination country, consumer preferences, and competition with local producers in the destination country. In addition, some of the obstacles to the export of MSME products are slow product turnover in importing countries because product stocks do not sell out immediately in these countries and sometimes reach the expiration date. The number of Indonesian MSMEs has reached 65.5 million, which is the largest in ASEAN countries. In addition, the contribution of MSMEs is recorded at around 61% of the national PDP and absorbs 97% of the total workforce in 2021. To assist MSMEs in choosing export destination countries, an intelligent system is needed that can recommend the right destination countries based on the characteristics of MSME products. As for how to determine the Export Destination country by using one of them based on the Trademap site. The urgency of this research aims to build an intelligent system model based on machine learning using the KNN method to recommend the right export destination countries for MSMEs. This machine learning-based intelligent system model is expected to be one of the solutions for MSMEs to effectively expand their markets abroad. This research is also expected to contribute to the development of an intelligent system that can assist MSMEs in choosing the right export destination countries and increasing export effectiveness. The results showed an accuracy rate of 99%, with a precision level of 100%, 100% recall, 100% f1-score.

**Keywords:** export; Key Nearest Neighbors; KNN; Intelligent Systems; MSMEs

### 1. Pendahuluan

Perkembangan teknologi dan aksesibilitas yang semakin luas, usaha mikro, kecil, dan menengah (UMKM) semakin mudah untuk memperluas pasar mereka ke luar negeri. Ekspor merupakan salah satu cara untuk dapat meningkatkan pertumbuhan bisnis UMKM. Namun, memasuki pasar ekspor tidak semudah yang dibayangkan. UMKM harus mempertimbangkan banyak faktor seperti persyaratan peraturan negara tujuan, preferensi konsumen, dan persaingan dengan produsen lokal di negara tujuan impor. Selain itu, beberapa kendala ekspor produk UMKM adalah dalam siklus penjualan produk yang lambat di negara tujuan impor dikarenakan stok produk tidak habis terjual di Negara tersebut dan terkadang sampai mencapai batas kadaluarsa. Penyebabnya karena kurangnya informasi mengenai perkembangan lokasi pasar dan persaingan UMKM dengan produk yang sama[1].



**Gambar 1.** Jumlah UMKM di Negara ASEAN

(sumber : databooks.katadata.co.id)[2]

Pada Gambar 1. Terlihat jumlah UMKM Indonesia mencapai 65,5 Juta yang merupakan terbanyak di negara ASEAN. Selain itu, UMKM juga berkontribusi mencapai kisaran 61% terhadap PDP nasional dan menyerap 97% dari total tenaga kerja pada tahun 2021[3]. UMKM berperan penting dalam pertumbuhan ekonomi[4]. Masuknya UMKM ke dalam pasar ekspor yang lebih luas dan kompetitif membutuhkan strategi yang tepat dan informasi yang akurat.

Salah satu informasi penting adalah tentang negara tujuan ekspor yang tepat, yang dapat membantu UMKM untuk menemukan peluang dan mengambil keputusan yang tepat dalam pemilihan negara tujuan ekspor. Akan tetapi, minimnya informasi mengenai besaran permintaan dan tersedianya produk di masing-masing negara sangat mempengaruhi siklus perputaran barang. Sebagai contoh ada 5 industri dengan jenis produk yang sama mengirim ke negara Australia dimana permintaan produk hanya sedikit. Tanpa adanya informasi yang tepat kelima industri hanya melakukan ekspor produk tanpa mempertimbangkan ketersediaan produk di negara tersebut. Sehingga jumlah produk yang berada di negara Australia melebihi jumlah permintaan, hal ini mengakibatkan adanya penumpukan produk dan industri mengalami kerugian dikarenakan minimnya tingkat penjualan. Produk dengan perputaran cepat adalah barang yang habis terjual dalam waktu yang relative cepat[1].

Untuk membantu UMKM dalam memilih negara tujuan ekspor, dibutuhkan sistem cerdas yang dapat merekomendasikan negara tujuan yang tepat berdasarkan karakteristik produk UMKM. Teknologi yang dapat digunakan untuk membangun model sistem cerdas salah satunya adalah *machine learning*. *Machine Learning*/ Sistem Cerdas merupakan bidang ilmu komputer yang menggunakan teknik statistika untuk memberi kemampuan system computer agar dapat belajar dari data, tanpa diprogram secara eksplisit[5].

Metode *machine learning* yang sering digunakan untuk klasifikasi adalah K-Nearest Neighbour (KNN). K-Nearest Neighbour (KNN) merupakan algoritma supervised learning dan dapat melakukan classification data set berdasarkan k jumlah tetangga terdekat [5], Metode KNN digunakan untuk mengklasifikasikan data berdasarkan kemiripan dengan data yang sudah ada sebelumnya. KNN mempertimbangkan data yang memiliki karakteristik yang mirip dengan data yang dicari dan mengeluarkan rekomendasi berdasarkan data yang mirip tersebut. KNN dapat digunakan untuk merekomendasikan negara tujuan ekspor yang tepat bagi UMKM berdasarkan karakteristik produk UMKM dan negara tujuan ekspor.

Adapun cara menentukan negara Tujuan Ekspor dengan menggunakan salah satunya berdasarkan laman Trademap yaitu sebuah situs Peta Dagang menyediakan - dalam bentuk tabel, peta-indikator, gambar dan kinerja ekspor, pasar alternatif dan pasar kompetitif, permintaan internasional, serta direktori perusahaan impor dan ekspor. Peta Perdagangan mencakup 220 negara dan wilayah dan 5300 produk dari Sistem Harmonisasi. Arus perdagangan bulanan, triwulanan, dan tahunan tersedia dari tingkat paling agregat hingga tingkat garis tarif.

Berdasarkan rekomendasi dari website ukmindonesia.id [6] untuk menentukan tujuan ekspor adalah dengan mempelajari performa negara tujuan melalui Trademap, yaitu analisis terhadap (1) Trade Balance, apabila negara dengan nilai trade balance negative artinya negara tersebut lebih banyak import daripada ekspor (2) Unit Value, adalah harga dari produk yang diekspor apabila mendekati atau lebih kecil dari unit value rata-rata maka produk layak untuk di ekspor di negara tersebut, (3) Tarif Masuk, semakin rendah nilai tarif masuk ke suatu negara maka biaya pengiriman dapat ditekan juga sehingga harga jual dapat bersaing di negara tujuan. Dengan banyaknya jumlah UMKM di Indonesia dan termasuk dalam yang tertinggi di negara ASEAN (Gambar 1), maka tidak menutup kemungkinan untuk UMKM Indonesia bersaing di pasar global. Hal ini juga mendukung dengan komitmen pemerintah untuk menjadi produsen Halal No. 1 di Dunia, sehingga membutuhkan penguatan dalam perkembangan produksi dan rekomendasi untuk tujuan ekspor. Urgensi penelitian ini bertujuan untuk membangun model sistem cerdas berbasis machine learning dengan menggunakan metode KNN untuk merekomendasikan negara tujuan ekspor yang tepat bagi UMKM. Model sistem cerdas berbasis machine learning ini diharapkan menjadi salah satu solusi bagi UMKM untuk memperluas pasar mereka ke luar negeri dengan efektif. Penelitian ini juga diharapkan dapat memberika

kontribusi untuk pengembangan system cerdas yang dapat membantu UMKM dalam memilih negara tujuan ekspor yang tepat dan meningkatkan efektivitas ekspor.

Variabel dalam pembuatan Sistem Cerdas Rekomendasi Sistem Market Matching negara tujuan ekspor untuk UMKM berbasis machine learning dengan metode KNN dengan menggunakan analisis analisis terhadap (1) Trade Balance, apabila negara dengan nilai trade balance negative artinya negara tersebut lebih banyak import daripada ekspor (2) Unit Value, adalah harga dari produk yang diekspor apabila mendekati atau lebih kecil dari unit value rata-rata maka produk layak untuk di ekspor di negara tersebut, (3) Tarif Masuk, semakin rendah nilai tarif masuk ke suatu negara maka biaya pengiriman dapat ditekan juga sehingga harga jual dapat bersaing di negara tujuan.

Penelitian sebelumnya yaitu Aplikasi Market Matching Berbasis Fuzzy sebagai Penunjang Keputusan Ekspor Produk UMKM berbasis Desktop dengan menggunakan Ms. Access, system ini dibuat berdasarkan data Stok, Permintaan, dan daya saing sebagai kategori penjunagan keputusan dan menghasilkan rekomendasi tujuan ekspor berdasarkan kategori produk dan tingkat kebutuhan importir yang dioperasikan[1]. Sedangkan sistem yang akan dibuat adalah Sistem rekomendasi market matching ini memanfaatkan teknologi machine learning dan big data analytics untuk mengidentifikasi pasar yang cocok dengan produk UMKM berdasarkan karakteristik produk, lokasi geografis, preferensi konsumen, dan faktor-faktor lainnya. Dengan menggunakan algoritma yang cerdas, sistem rekomendasi dapat menghasilkan rekomendasi pasar yang akurat dan relevan untuk setiap UMKM.

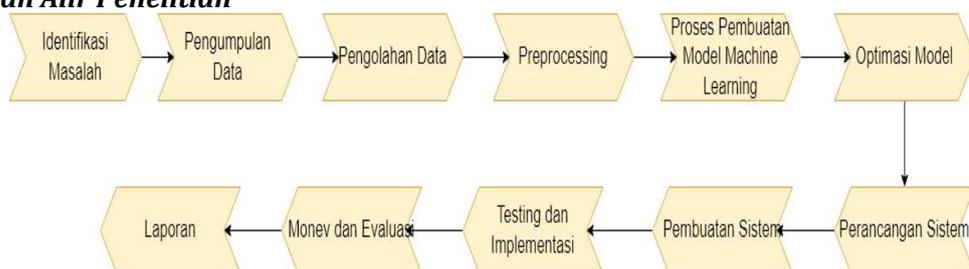
Keuntungan dari penggunaan sistem rekomendasi market matching ini adalah UMKM dapat menghemat waktu dan biaya dalam mencari pasar potensial, serta dapat meningkatkan peluang keberhasilan ekspor mereka. Selain itu, sistem ini juga dapat membantu pemerintah dalam mempromosikan produk UMKM secara efektif ke pasar global.

## 2. Metode Penelitian

### Subyek dan Obyek Penelitian

Penelitian ini memiliki subjek yang merupakan data scrapping yang diambil dalam situs trademap yang merupakan situs peta perdagangan internasional. Adapun variable yang menjadi rekomendasi yaitu : 1) Trade Balance, 2) Unit Value, 3) Tarif Masuk. Penelitian ini memiliki objek yaitu penerapan dari algoritma K-Nearest Neighbors dengan Teknik Cross Validation.

### Bagan Alir Penelitian



**Gambar 2.** Bagan Alir Penelitian

### Identifikasi Masalah

Identifikasi masalah dimulai dari tim peneliti melakukan identifikasi masalah bersama dengan UMKM pada saat pelaksanaan Workshop Kurasi Ekspor yang diadakan oleh Disperindag Kab. Malang pada Tahun 2022, setelah itu dilakukan proses pencarian studi literatur.

### Pengumpulan Data

Proses mengumpulkan data dilakukan dengan teknik scrapping. Web scraping merupakan teknik pengambilan sebuah dokumen semi-terstruktur dari halaman web, dengan mengambil

data tertentu saja. Web scraping juga adalah teknik untuk mengekstraksi data dari internet dan disimpan ke file atau data base untuk kebutuhan analisis data[7]. Peneliti melakukan pengumpulan data dengan bersumber dari OSINT (Open Source Intellegent)[8] yaitu setiap orang berhak mengakses data tersebut yang kemudian akan diolah menjadi sebuah dataset. Proses pengumpulan data dilakukan pada tahap pra-pengolahan data. Tahap pra-pengolahan data dapat dibagi menjadi empat tahap utama yaitu tahap pengumpulan dan pemilihan data, tahap penentuan kelas data, tahap feature selection serta tahap resampling data[9]. Selain itu, pengumpulan data juga dilakukan dengan pencarian data ke Dinas terkait seperti Disperindag dan Bea Cukai di Malang Raya.

#### *Pengolahan Data*

Pengolahan Data dimulai dari proses cleaning yaitu, menghapus noise, membuang duplikasi data, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak[10].

#### *Pre-Processing*

*Preprocessing* dilakukan dengan cara membuat feature selection serta tahap resampling data yang digunakan melalui dataset yang dihasilkan dari proses pengolahan data awal[9].

#### *Proses Pembuatan Machine Learning*

Tahap Pembuatan Machine Learning dimulai dari pembuatan Training Dataset merupakan tahapan dimana untuk mendapatkan nilai akurasi yang tinggi terhadap algoritma yang dipilih. Pembuatan Training dataset digunakan untuk membuat model yang nantinya akan digunakan sebagai pengujian dalam prosesnya [11]. Model Machine Learning dengan Menggunakan Algoritma KNN. Algoritma k-Nearest Neighbor (kNN) adalah sebuah algoritma untuk klasifikasi objek berdasarkan data pembelajaran dengan jarak terdekat dengan objek tersebut. KNN merupakan algoritma supervised learning, yaitu hasil dari *query instance* terbaru dilakukan klasifikasi berdasarkan mayoritas dari kategori KNN. *Class* yang paling sering muncul merupakan *class* hasil klasifikasi. KNN memiliki tujuan yaitu, melakukan proses klasifikasi objek baru berdasarkan atribut dan training sample. Algoritma k-Nearest Neighbor menggunakan klasifikasi ketetanggaan (*neighbor*) dengan tujuan mencari nilai prediksi dari *query instance* yang baru. Algoritma KNN, bekerja berdasarkan jarak terdekat dari *query instance* ke training sample untuk menentukan ketetangannya. Adapun urutan dari cara menghitung metode KNN antara lain sebagai berikut :

- a) Melakukan identifikasi parameter k
- b) Melakukan perhitungan jarak antar data yang akan dievaluasi ke semua data *training*
- c) Melakukan pengurutan jarak yang telah terbentuk
- d) Melakukan pengurutan jarak yang paling dekat sampai dengan urutan k.
- e) Melakukan pemasangan *class* yang sesuai dengan k
- f) Melakukan perhitungan jumlah *class* dari tetangga yang paling dekat dan menetapkan *class* sebagai data yang akan dilakukan evaluasi [12]

#### *Optimasi Model*

Optimasi Model menggunakan teknik cross validation. Merupakan metode statistik yang digunakan untuk mengevaluasi kinerja model atau algoritma dimana data dipisahkan menjadi dua subset yaitu data proses pembelajaran dan data validasi /evaluasi Cross validation adalah membagi dataset menjadi dua bagian dengan satu bagian dijadikan data training dan bagian yang lain dijadikan data testing[13].

Dalam *k-fold cross validation*, data asli dipartisi ke dalam subset/bagian. Model ini kemudian dibangun menggunakan data dari subset K-1 (2,3,4,5, dst), dan terdapat bagian lain didalam subset yang digunakan untuk settes. Bagian subset(dataset) harus lebih banyak daripada settes,

dilakukan secara iteratif sampai memiliki model yang berbeda. Hasil dari masing-masing model K(akurasi) kemudian digabungkan menggunakan rata-rata untuk mendapatkan hasil akurasi dari keseluruhan data. K-fold cross-validation yang sering digunakan adalah 10-fold cross validation[14].

#### Perancangan Sistem

Perancangan sistem digunakan untuk membuat desain interface dari UI/UX. Perancangan merupakan tahap sesudah analisis dari siklus perancangan sistem yang dapat berupa penggambaran, perencanaan dan pembuatan model atau pengaturan dari beberapa elemen yang terpisah ke dalam satu kesatuan yang utuh dan berfungsi, termasuk menyangkut mengkonfigurasi dari komponen-komponen perangkat lunak dan perangkat keras dari suatu sistem[15] [16].

#### Pembuatan Sistem

Pembuatan sistem cerdas menggunakan bahasa pemrograman Python dengan menggunakan *framework streamlit*. *Streamlit* merupakan *framework* yang menggunakan Bahasa pemrograman Python dan bersifat open-source yang dibuat untuk memudahkan dalam membangun aplikasi web di bidang sains data dan *machine learning* yang interaktif [17]

#### Testing dan Implementasi

Kegiatan *testing* dan Implementasi sistem yang menggunakan metode *Blackbox testing* yaitu pengujian yang dilakukan dengan menguji hasil eksekusi melalui data uji dan pemeriksaan fungsional dari sistem[18]

#### Monev dan Evaluasi

Tahap selanjutnya yaitu periode monitoring dan evaluasi atas hasil dari kegiatan penelitian yang telah dilakukan. Pada tahapan ini juga dilakukan pembuatan drafting Jurnal dan pendaftaran Hak Cipta program komputer

#### Laporan

Tahap akhir yaitu pembuatan laporan akhir penelitian sebagai bentuk tanggung jawab peneliti terhadap pemberi dana yaitu Kemendikbud-DIKTI.

### 3. Hasil dan Pembahasan

#### Pengumpulan Data

Pengumpulan data menggunakan proses scrapping. Web scraping adalah sebuah cara pengambilan dokumen semi-terstruktur yang terdapat di halaman web, dengan mengambil beberapa data dari website yang di scrap. Web scraping juga merupakan teknik untuk mengekstraksi data dari internet dan disimpan ke file atau data base untuk kebutuhan analisis data[7]. Peneliti melakukan pengumpulan data dengan bersumber dari OSINT(Open Source Intellegent)[8] yaitu setiap orang berhak mengakses data tersebut yang kemudian akan diolah menjadi sebuah dataset. Proses pengumpulan data dilakukan pada tahap pra-pengolahan data. Tahap pra-pengolahan data dapat dibagi menjadi empat tahap utama yaitu tahap pengumpulan dan pemilihan data, tahap penentuan kelas data, tahap featureselection serta tahap resampling data[9].Komoditi yang diambil datanya merupakan komoditi unggulan ekspor untuk produk-produk UMKM di Indonesia, yaitu sebagai berikut :

**Tabel 1. Komoditi Unggulan Ekspor**

Kode	Keterangan
1	Briket/Bahan Bakar padat serupa yang dibuat dari batu bara
2	Gandum dan Meslin, Biji-Bijian yang Mengandung Karbohidrat
3	Barang dari plastik dan produk turunannya
4	Setelan jas pria atau anak laki-laki, jaket, blazer, turunannya
5	Serat optik dan bundel serat optik Kabel serat optik

Kode	Keterangan
6	Pupuk nitrogen mineral atau kimia (tidak termasuk yang berbentuk tablet atau bentuk serupa
7	Pakaian dan aksesoris pakaian dari kulit atau kulit komposisi
8	Sediaan kecantikan atau tata rias dan sediaan perawatan kulit termasuk tabir surya
9	Barang keramik
10	Kopi, digongseng atau dihilangkan kafeinnya, sekam dan kulit kopi pengganti kopi
11	Bumbu makanan
12	Alas kaki dengan sol luar dari kulit karet plastik atau kulit komposisi dan bagian atasnya
13	Perhiasan imitasi
14	Aksesoris pakaian jadi dan bagian dari garmen atau segala jenis aksesoris pakaian
15	Barang lainnya dari kayu
16	Kertas dan kertas karton lainnya tidak dilapisi dalam gulungan dengan lebar 36 cm atau berbentuk persegi atau persegi panjang
17	Anyaman dan produk semacam itu dari bahan anyaman dirakit menjadi anyaman strip maupun tidak
18	Makanan olahan diperoleh dengan cara menggembungkan atau memanggang sereal atau produk sereal
19	Peralatan makan, peralatan dapur, perlengkapan rumah tangga lainnya, dan perlengkapan toilet dari plastic
20	Peralatan makan, peralatan dapur, perlengkapan rumah tangga lainnya, dan perlengkapan toilet dari porselen atau keramik
21	Kulit kecokelatan atau kerak dan kulit sapi termasuk. binatang kerbau atau kuda yang tidak berbulu
22	Produk tekstil dan barang untuk keperluan teknis
23	Kain tenunan dari sutera atau dari sisa sutera
24	Minyak kelapa sawit dan fraksinya, dimurnikan maupun tidak (tidak termasuk yang dimodifikasi secara kimia)
25	Biji kakao

Pengumpulan Data yang dilakukan sejumlah 4916 dataset, dengan data yang diolah adalah sebagai berikut :

**Tabel 2. Dataset Impor Komoditi**

Kode	Importer	Trade Balance	Unit Value	Tax	Rekom
1	United States of America	-2.282.964		1.9	1
1	Germany	-42.900	4783	0.9	1
1	Mexico	-2.107.314		0	1
1	Thailand	-2.318.377	5488	9.9	1
1	France	-503.113	4707	0.9	1
:	:	:	:	:	:
25	Angola	0	1	0	0
25	Faroe Islands	0	9.000	0	0
25	Andorra	0	1	0	0
25	Iraq	0	1	0	0
25	Saint Helena	0	667	0	0

### **Pengolahan Data**

Pengolahan Data dimulai dari proses cleaning yaitu, menghapus *noise*, menghapus data duplikat, melakukan pemeriksaan data yang tidak konsisten, serta melakukan perbaikan dari kesalahan

data, seperti kesalahan cetak [10]. Proses pengolahan data dilakukan dengan melakukan pengecekan dari data yang ada dengan melakukan pre-processing. Proses pengolahan data antara lain sebagai berikut :

*Identifikasi Jenis Data*

Proses identifikasi jenis data dilakukan dengan menggunakan script berikut ini

```
>>trade_df.info()
```

maka hasil *output* adalah sebagai berikut :

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 4916 entries, 0 to 4915
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Kode             4916 non-null   int64
1   Importer         4916 non-null   object
2   Trade Balance    4916 non-null   int64
3   Unit Value       4916 non-null   int64
4   Tax              4916 non-null   int64
5   rekom            4916 non-null   int64
dtypes: int64(5), object(1)
memory usage: 268.8+ KB
```

**Gambar 3.** Output Identifikasi Jenis Data

Dalam Gambar 3 tersebut menunjukkan bahwa terdapat 6 kolom yang memiliki jumlah data di keseluruhan kolom sama yaitu 4916 dan tidak terdapat data kosong, sehingga layak untuk digunakan secara keseluruhan. Selain itu pada kolom Importer/ nama Negara tujuan Impor memiliki *type data Object*.

*Proses Cleaning*

Pada tahap ini dilakukan pemilahan data yang nantinya dipakai untuk pembuatan *machine learning* atau tidak. Pada Kolom Kode, dan Importer tidak dilakukan perhitungan untuk proses Algoritma karena data tersebut merupakan identitas dari komoditi dan Negara Tujuan. Sehingga yang akan dipakai sebagai *x\_label* adalah kolom *Trade Balance*, *Unit Value* dan *Tax*, yaitu sebagai berikut :

	Trade Balance	Unit Value	Tax
0	0	26	2
1	2	410	4
2	0	1755	2
3	0	500	10
4	0	1500	0

**Gambar 4.** Dataset *x\_label*

Proses selanjutnya yaitu melakukan labelisasi dari *y\_label* yang diambil dari kolom *rekom*. Kolom *rekom* berisi kode 0 untuk tidak layak dijadikan tujuan ekspor dan 1 untuk Layak dijadikan negara tujuan ekspor seperti dalam Gambar 5 berikut ini.

```
<bound method NDFrame.head of 0      0
1      1
2      0
3      0
4      0
..
4911   0
4912   0
4913   0
4914   0
4915   0
Name: rekom, Length: 4916, dtype: int64>
```

**Gambar 5.** Dataset *y\_label*

*Pre-Processing*

*Preprocessing* dilakukan dengan cara membuat featureselection serta tahap resampling data yang digunakan melalui dataset yang dihasilkan dari proses pengolahan data awal[9].

*Encoding Trade Balance*

*Trade Balance* adalah nilai neraca perdagangan yang menjadi tolak ukur untuk melakukan kegiatan ekspor dan impor, semakin *negative* nilai *trade balance* maka negara tersebut memiliki potensi untuk menjadi negara tujuan ekspor, dengan anggapan bahwa negara tersebut kekurangan produk tersebut. Sedangkan Encoding adalah merupakan salah satu teknik mengubah data variable menjadi bentuk numerik agar dapat digunakan dalam algoritma pembelajaran mesin [19] teknik encoding yang digunakan dalam *trade balance* ini yaitu menggunakan label *encoding*. *Trade balance* akan dilakukan *encoding* sebagai berikut :

**Tabel 3.** Encoding Trade Balance

No	Keterangan	Encoding
1	-59.699.934 s/d -1.003.573	5
2	-991.518 s/d - 100.044	4
3	-99.933 s/d -10.000	3
4	-9.990 s/d - 1.002	2
5	-996 s/d -100	1
6	-99 s/d surplus	0

Selain itu proses *preprocessing* menggunakan metode *Standard Scaler*, yaitu merupakan metode *preprocessing* dimana metode tersebut akan melakukan standarisasi fitur dengan menghapus rata-rata dan menskalakan unit varian. Proses ini akan dilakukan pada setiap fitur pada sampel. *Preprocessing* ini dilakukan untuk mencegah adanya data yang memiliki nilai terlalu besar dibanding dengan nilai yang lain yang akan dapat mengakibatkan proses *training* tidak sesuai keinginan [21]. Rumus dari *Standard Scaler* ditunjukkan pada Persamaan (2), di mana *X* adalah rerataan nilai sampel dan  $\sigma$  adalah standar deviasi.

$$Z = \frac{X_i - \bar{x}}{\sigma} \tag{1}$$

Data yang diolah untuk proses *preprocessing* dengan menggunakan *Standard Scaler* menjadi seperti berikut :

**Tabel 4.** Hasil *Preprocessing* dengan menggunakan *Standard Scaler*

Index	Trade Balance	Unit Value	Tax
0	1.65756783	-0.26878761	-0.70689461
1	-0.45987713	-0.28733289	0.22030839
2	1.65756783	-0.09447689	-0.61417431
:			
4913	- 1.16569211	- 0.28733289	0.12758809
4914	-0.45987713	- 0.24422565	0.86935049
4915	0.95175284	- 0.17123303	- 0.70689461

### *Proses Pembuatan Machine Learning*

Tahap Pembuatan Machine Learning dimulai dari pembuatan Training Dataset merupakan tahapan dimana untuk mendapatkan nilai akurasi yang tinggi terhadap algoritma yang dipilih. Pembuatan Training dataset digunakan untuk membuat model yang nantinya akan digunakan sebagai proses pengujian[11]. Model Machine Learning dengan Menggunakan Algoritma KNN. Algoritma k-Nearest Neighbor (kNN) merupakan sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. kNN termasuk algoritma supervised learning dimana hasil dari query instance yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada kNN. Kelas yang paling banyak muncul itu yang akan menjadi kelas hasil klasifikasi. Tujuan dari algoritma ini adalah mengklasifikasikan objek baru berdasarkan atribut dan training sample. Algoritma k- Nearest Neighbor menggunakan klasifikasi ketetanggaan (neighbor) sebagai nilai prediksi dari query instance yang baru. Algoritma ini sederhana, bekerja berdasarkan jarak terpendek dari query instance ke training sample untuk menentukan ketetanggaannya. Langkah-langkah untuk menghitung metode k-Nearest Neighbor antara lain:

- a) Menentukan parameter k
- b) Menghitung jarak antara data yang akan dievaluasi dengan semua pelatihan
- c) Mengurutkan jarak yang terbentuk
- d) Menentukan jarak terdekat sampai urutan k
- e) Memasangkan class yang bersesuaian
- f) Mencari jumlah class dari tetangga yang terdekat dan tetapkan class tersebut sebagai class data yang akan dievaluasi [12].

Proses pembuatan *Machine Learning* dengan Bahasa pemrograman Python, adapun requirement dari kebutuhan sistem antara lain sebagai berikut :

- a. Image versi 1.5.33
- b. importlib-metadata versi 4.12.0
- c. importlib-resources versi 5.4.0
- d. matplotlib-inline versi 0.1.3
- e. numpy versi 1.17.3
- f. pandas versi 1.3.5
- g. pickle-mixin versi 1.0.2
- h. pickleshare versi 0.7.5
- i. scikit-learn versi 0.24.2
- j. scipy versi 1.8.0
- k. joblib versi 0.11
- l. streamlit versi 1.10.0

*Dataset* yang digunakan sejumlah 4.916 data yang dibagi menjadi data training dan data testing yaitu sebesar 3.441 untuk data training dan 1.475 untuk data testing atau dengan pembagian 70% data training dan 30% data testing.

### *Optimasi Model*

Optimasi Model menggunakan teknik confusion matrix. *Confusion matrix* adalah tabel yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah. Contoh *confusion matrix*. Contoh *confusion matrix* untuk klasifikasi biner ditunjukkan pada Tabel 5 berikut ini.

**Tabel 5. Confusion Matrix**

Kelas sebenarnya	Kelas Prediksi	
	1	0
1	TP	FN
0	FP	TN

Keterangan:

TP (True Positive) kelas 1 = jumlah dokumen dari kelas 1 yang benar diklasifikasikan sebagai kelas 1

TN (True Negative) kelas 0 = jumlah dokumen dari kelas 0 yang benardiklasifikasikan sebagai kelas 0

FP (False Positive) kelas 1 = jumlah dokumen dari kelas 0 yang salah diklasifikasikan sebagai kelas 1

FN (False Negative) kelas 0 = jumlah dokumen dari kelas 1 yang salah diklasifikasikan sebagai kelas 0

Rumus :

Rumus *confusion matrix* untuk menghitung *accuracy*, *precision*, dan *recall* seperti berikut:[20]

$$Accuracy = \frac{TP+TN}{Total} \tag{2}$$

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

**Pembuatan Sistem**

Pembuatan sistem cerdas menggunakan bahasa pemrograman Python dengan menggunakan *framework streamlit*. Proses pengkodean dengan menggunakan Bahasa pemrograman python. Sistem Rekomendasi yang saat ini dibuat dapat diakses melalui link <https://rekomendasi ekspor.streamlit.app/>

**Sistem Rekomendasi Ekspor Untuk UMKM**



Jenis Barang  
 Biji kakao

Input Negara Tujuan Ekspor

Skala Pengiriman Barang  
 skala besar mencapai puluhan juta unit

Input Harga Barang Yang Akan Dijual (Satuan Dollar)  
 0,00

Input Nilai Pajak Maksimal Negara Tujuan yang Diinginkan  
 0,00

### Tabel Inputan Data

	Skala	Harga Barang	Pajak
0	5	0.0000	0.0000

### Hasil Rekomendasi Negara Tujuan Ekspor

Submit

**Gambar 6.** Halaman Aplikasi Rekomendasi Ekspor

#### Testing dan Implementasi

Kegiatan *testing* dan implementasi sistem dengan menggunakan metode *Blackbox testing*. Berikut ini desain rancangan *testing* sistem rekomendasi

**Tabel 6.** Tabel Testing dan Implementasi Sistem

No	Input Jenis Barang	Input Negara Tujuan	Skala Pengiriman	Harga Barang yang Akan Dijual	Pajak	Hasil
1	Gandum	Italy	Skala Rendah	408	1	Tidak direkomendasikan
2	Biji Kakao	Belarus	Skala sedang mencapai	667	1	Tidak diekomendasikan

#### Pengujian Model

Hasil dari pengujian model KNN yang telah dibuat dengan menggunakan beberapa tes didapat sebagai berikut :

No	Train Test	Test Size	Random State	Hasil Akurasi	Tingkat
1	0,6	0,4	20	0.98830706659888	
2	0,7	0,3	20	0.99050847457627	
3	0,7	0,3	40	0.99322033898305	
4	0,6	0,4	40	0.99389933909506	

#### 4. Penutup

Hasil dari optimasi model menggunakan *classification report* adalah sebagai berikut :

```

Accuracy (KNN): 0.9938993390950687
Classification Report (KNN):
              precision    recall  f1-score   support

     0         0.99         0.99         0.99         930
     1         1.00         0.99         0.99        1037

 accuracy         0.99         0.99         0.99        1967
 macro avg         0.99         0.99         0.99        1967
weighted avg         0.99         0.99         0.99        1967

Confusion Matrix (KNN):
[[ 925   5]
 [   7 1030]]
    
```

**Gambar 7.** Hasil Classification Report

Pada Gambar 7 tingkat akurasi sebesar 99, 38%, dengan tingkat *precision* 99%, *recall* 99%, *f1-score* 99%.

#### 5. Referensi

- [1] Bambang Nurdewanto, E. S. F. A. S. R., 2017. Aplikasi Market Matching Berbasis Fuzzy sebagai Penunjang Keputusan Ekspor Produk UMKM. *Jurnal Ilmu Komputer dan Teknologi Informasi*, 9(No. 2 (2017)), pp. 58-61
- [2] Databooks.katadata.co.id, 2022. Indonesia Punya UMKM Terbanyak di ASEAN. 2021, 10 (Dikutippada 12 April 2023 pukul 12.05 WIB). <https://databoks.katadata.co.id/datapublish/2022/10/11/indonesia-punya-umkm-terbanyak-di-asean-bagaimana-daya-saingnya>
- [3] Perekonomian, K. K. B., 2022. Pengembangan UMKM Menjadi Necessary Condition untuk Mendorong Pertumbuhan Ekonomi, Jakarta: Kementerian Koordinator Bidang Perekonomian. (Dikutippada 12 April 2023 pukul 12.10 WIB) <https://www.ekon.go.id/publikasi/detail/3301/nilai-ekspor-indonesia-catat-rekor-tertinggi-sepanjang-sejarah>
- [4] Primadewi, M. I. A. A., 2020. Analisa dan Permodelan Proses Bisnis Menggunakan Business Process Model and Notation (BPMN) Studi Kasus : LAZ DKD. *Proceeding of The 11th University Research Colloquium 2020: Bidang Sosial Humaniora dan Ekonomi*, Volume 1, pp. 401-409.
- [5] Bedy Purnama, S. M., 2019. Pengantar Machine Learning. Oktober 2019 ed. Bandung: Informatika Bandung.
- [6] Asih, H., 2023. Cara Menentukan Negara Tujuan Ekspor dengan Menggunakan Trade Map, Jakarta. (Dikutip pada Dikutip pada 12 April 2023 pukul 12.20 WIB ): <https://ukmindonesia.id/>.
- [7] Djufri, M., 2020. Penerapan Teknik Web Scraping Untuk Penggalan Potensi Pajak (Studi Kasus pada Online Market Place Tokopedia, Shopee dan Bukalapak). *Jurnal BPPK*, 13(2), pp. 65-75.
- [8] Dedy Hariyadi, F., 2019. Membangun Telegrambot Untuk Crawling Malware Osint Menggunakan Raspberry PI. *Indonesian Journal of Business Intelligence*, 2(1), pp. 18-24.
- [9] Bonifasius Yoga AdiPratama, H. A. Y., 2021. Perancangan Proses Implementasi Machine Learning dalam Maintenance Management Untuk Mencegah Derating. *JatiUndip :Jurnal Teknik Industri*, 16(2), pp. 134-142.

- [10] Ilsa Palingga Ninditama, I. P. N. W. C. M. A. D. A., 2021. Klasifikasi Keluarga Sejahtera Study Kasus :Kecamatan Kota Palembang. *Jurnal Tekno Kompak*, 15(2), pp. 37-49.
- [11] Ari Peryanto, A. Y. R. U., 2019. Rancang Bangun Klasifikasi Citra Dengan Teknologi Deep Learning Berbasis Metode Convolutional Neural Network. *Jurnal Format*, 8(2), pp. 138-147.
- [12] Mutiara Ayu Banjarsari, H. I. B. A. F., 2015. Penerapan K-Optimal Pada Algoritma Knn untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan IP Sampai Dengan Semester 4. *Kumpulan jurnaLIlmuKomputer (KLIK)*, 2(2), pp. 50-64.
- [13] Elkin Rilvani, dkk.2019. Penentuan Kelulusan Siswa Yayasan Cerdas Bakti Pertiwi Dengan Menggunakan Algoritma Naïve Bayes Dan Cross Validation. *Pelita Teknologi: Jurnal Ilmiah Informatika, Arsitektur dan Lingkungan* , 14(2), pp. 145-153.
- [14] Asep Muhidin, A. B., 2016. Klasifikasi Tingkat Produktivitas Tanaman Padi Di Kabupaten Karawang Menggunakan Metode Naive Bayes Dan K-Fold Cross Validation. *Jurnal TeknologiPelitaBangsa*, 7(2), pp. 135-140.
- [15] Jogiyanto, 2005. Analisis dan Desain. Yogyakarta: Andi Offset.
- [16] M. Agus Muhyidin, dkk, 2020. Perancangan UI/UX Aplikasi MY CIC Layanan Informasi Akademik Mahasiswa Menggunakan Aplikasi Figma. *Jurnal Digit :Universitas Catur Insan Cendekia Cirebon*, 10(2), pp. 208-219.
- [17] Ahmad Fathan Hidayatullah, S. M., 2021. Membuat Aplikasi Web Sains Data Dengan Mudah Menggunakan Streamlit. 2021/03/15, 15 03, p. 1
- [18] Astuti, P., 2018. Penggunaan Metode Black Box Testing (Boundary Value Analysis). *Factor Exacta*, 11(2), pp. 186-195.
- [19] Anak Agung Aryasatya Daniswara, I. K. (2023). Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru. *JINACS: Journal of Informatics and Computer Science*, 97-100.
- [20] Dwi Normawati, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)* , 697-711.
- [21] Vincentius Riandaru Prasetyo, M. M. (2022). Prediksi Rating Film Pada Website Imdb Menggunakan Metode Neural Network Film Rating Prediction On Imdb Website Using Neural Network. *Jurnal Ilmiah NERO*, 1-8.