MODEL KLASIFIKASI DAERAH PERDAGANGAN SATWA LIAR MENGGUNAKAN ALGORITMA *NAÏVE BAYES* DAN *RANDOM FOREST*

HENDRO UTOMO

PROGRAM STUDI ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024

IPB University
Bogor Indonesia
@Hak cipta milik IPB University
IPB University
Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.
Perpustakaan IPB University

PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul "Model Klasifikasi Daerah Perdagangan Satwa Liar Menggunakan Algoritma *Naïve Bayes* dan *Random Forest*" adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apapun kepada perguruan tinggi manapun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan ataupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2024

Hendro Utono
NIM G6501201019

IPB University
Bogor Indonesia
Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.
Perpustakaan IPB University
Dokumen milik IPB University
IPB University

IPB University
Bogor Indonesia

RINGKASAN

HENDRO UTOMO. Model Klasifikasi Daerah Perdagangan Satwa Liar Menggunakan Algoritma *Naive Bayes* dan *Random Forest*. Dibimbing oleh MEDRIA KUSUMA DEWI HARDHIENATA dan IRMAN HERMADI.

Perdagangan satwa liar merupakan sektor perdagangan yang kritikal namun rentan terhadap berbagai bentuk aktivitas yang mampu mengancam keanekaragaman hayati dan keseimbangan ekosistem. Oleh karenanya diperlukan regulasi dan pemantauan yang ketat terkait dengan perdagangan satwa liar. Penelitian ini bertujuan untuk mengembangkan model klasifikasi untuk mengkategorikan daerah-daerah yang rawan terhadap perdagangan satwa liar menggunakan pendekatan *machine learning*.

Data yang digunakan dalam penelitian ini adalah data perdagangan satwa liar yang berasal dari database CITES pada periode 2018–2020. Penelitian ini menggunakan teknik *K-Means* untuk pelabelan data serta *Naïve Bayes* dan *Random Forest* untuk mengkategorikan tingkat kerawanan daerah perdagangan satwa liar. Daerah perdagangan satwa liar pada penelitian ini dikategorikan ke dalam tiga kategori kerawanan, yaitu kategori rendah, sedang, dan tinggi. Cara melakukan klasifikasi daerah perdagangan satwa liar dapat dilakukan *dengan Naive Bayes dan Random Forest*, kemudian untuk mengimplentasikan dengan menggunakan bahasa pemrograman *python* sehingga memperoleh hasil sebaran risiko pada tiap negara yang terdapat pada *data base*. Data yang akan digunakan dalam penelitian ini merupakan data kasus perdagangan satwa liar yang terjadi dari tahun 2018 hingga 2020 di berbagai negara. Data tersebut diunduh dari situs *https://trade.cites.org/* dalam format csv (CITES, 2023). Data terdiri dari 16 kolom dan 179.524 baris. Pada tahap analisis data perdagangan satwa liar diantaranya menganalisis hewan yang banyak diperjualbelikan yaitu Acipenser dan terendah adalah Coralium. Bagian tubuh hewan yang diperjualbelikan paling tinggi adalah hati dan terendah adalah ekstrak, negara yang paling banyak melakukan impor perdagangan satwa liar adalah Korea dan paling rendah adalah Qatar, selanjutnya negara yang paling banyak ekspor adalah Perancis dan paling rendah adalah China. Tahun tertinggi yang melakukan perdagangan satwa liar terjadi pada tahun 2020 dan terendah pada 2018. Tahapan selanjutnya adalah tahap praproses data yang merupakan tahapan awal untuk mengubah data input menjadi data dengan format yang tepat dan siap untuk diproses. Praproses data yang dilakukan meliputi tahap data *cleaning*, data *reduction*, dan data *transformation*. Pada penelitian ini juga dilakukan seleksi fitur untuk memilih subset dari fitur yang optimal, dan mengeliminasi variabel yang tidak berhubungan dengan data penelitian. Tahapan selanjutnya adalah tahap *clustering* yang dilakukan menggunakan algoritma *K-Means*, yang bertujuan untuk mengelompokkan data yang mirip dan menemukan pola berdasarkan kemiripan (Eshimiakhe *et al.* 2022). Nilai k dalam *clustering* ditentukan dengan metode Elbow dan metode Silhouette, di mana nilai Silhouette mengukur kohesi serta separasi dalam kluster, dan semakin kecil nilai Silhouette maka semakin dekat nilai k dengan kluster yang tidak sesuai (Arnita *et al.* 2020). Setelah melalui tahap *clustering* selanjutnya dibangun model klasifikasi menggunakan algoritma *Naïve Bayes* dan *Random Forest*. Algoritma *Random Forest* adalah metode yang dapat meningkatkan hasil akurasi dalam membangkitkan atribut untuk setiap node yang dilakukan secara random, dan hyperparameter tuning untuk mendapatkan parameter terbaik dari RF sehingga RF dapat membentuk model klasifikasi yang baik.

Hasil penelitian menunjukkan bahwa hasil klasifikasi menggunakan algoritma *Naive Bayes* mencapai akurasi 92,45 persen, sementara *Random Forest* mencapai

@Hak cipta milik IPB University
IPB University
Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.
Perpustakaan IPB University

akurasi 98,11 persen. Selain itu penelitian ini juga menampikan visualisasi peta sebaran tingkat kerawanan perdagangan satwa liar. Dari tampilan dibawah ini menggunakan bahasa pemrograman Python dan menggunakan software *Visual Studio Code*, tahapannya seperti pada alur penelitian dan data dibawah ini 3 kelas yaitu rendah, sedang, tinggi dan diketahui tingkat *importer* tertinggi oleh negara Korea dan exporter tertinggi adalah Perancis diharapkan dengan adannya tampilan ini dapat memudahkan pembacaan data dan membatu *stakeholder* ditiap negara untuk bisa melakukan tindakan *preventif* dan *korektif* atas masalah ini.

Kata Kunci**:** Perdagangan Satwa Liar, *K-Means*, Klasifikasi, *Naïve Bayes*, *Random Forest*

IPB University
Bogor Indonesia
Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.
Perpustakaan IPB University
Hak cipta milik IPB University
IPB University

IPB University
Bogor Indonesia

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

@Hak cipta milik IPB University

IPB University

Perpustakaan IPB University

SUMMARY

HENDRO UTOMO. Wildlife Trade Area Classification Model Using Naive Bayes and Random Forest Algorithms. Supervised by MEDRIA KUSUMA DEWI HARDHIENATA and IRMAN HERMADI.

Wildlife trade is a critical trade sector but is vulnerable to various forms of activities that can threaten biodiversity and ecosystem balance. Therefore, strict regulation and monitoring are needed regarding wildlife trade. This study aims to develop a classification model to categorize areas vulnerable to wildlife trade using a machine learning approach.

The data used in this study are wildlife trade data from the CITES database for the period 2018–2020. This study uses the K-Means technique for data labeling and Naïve Bayes and Random Forest to categorize the level of vulnerability of wildlife trade areas. Wildlife trade areas in this study are categorized into three categories of vulnerability, namely low, medium, and high. The method for classifying wildlife trade areas can be done using Naive Bayes and Random Forest, then to implement it using the python programming language to obtain the results of the risk distribution in each country contained in the database. The data that will be used in this study is data on wildlife trade cases that occurred from 2018 to 2020 in various countries. The data was downloaded from the site https://trade.cites.org/ in csv format (CITES, 2023). The data consists of 16 columns and 179,524 rows. At the stage of wildlife trade data analysis, including analyzing the most traded animals, namely Acipenser and the lowest is Coralium. The highest traded animal body part is the liver and the lowest is the extract, the country that imports the most wildlife trade is Korea and the lowest is Qatar, then the country that exports the most is France and the lowest is China. The highest year for wildlife trade occurred in 2020 and the lowest in 2018. The next stage is the data preprocessing stage which is the initial stage to change input data into data with the right format and ready to be processed. The data preprocessing carried out includes the stages of data cleaning, data reduction, and data transformation. In this study, feature selection was also carried out to select a subset of optimal features, and eliminate variables that were not related to the research data. The next stage is the clustering stage which is carried out using the K-Means algorithm, which aims to group similar data and find patterns based on similarity (Eshimiakhe et al. 2022). The k value in clustering is determined by the Elbow method and the Silhouette method, where the Silhouette value measures cohesion and separation in the cluster, and the smaller the Silhouette value, the closer the k value is to the inappropriate cluster (Arnita et al. 2020). After going through the clustering stage, a classification model is built using the Naïve Bayes and Random Forest algorithms. The Random Forest algorithm is a method that can increase the accuracy of generating attributes for each node which is done randomly, and hyperparameter tuning to get the best parameters from RF so that RF can form a good classification model.

The results of the study showed that the classification results using the Naive Bayes algorithm achieved an accuracy of 92.45 percent, while Random Forest achieved an accuracy of 98.11 percent. In addition, this study also displays a visualization of the distribution map of the level of vulnerability to wildlife trade. From the display below using the Python programming language and using Visual

Studio Code software, the stages are as in the research flow and data below 3 classes, namely low, medium, high and it is known that the highest importer level is by Korea and the highest exporter is France. It is hoped that with this display it can facilitate data reading and help stakeholders in each country to be able to take preventive and corrective actions on this problem.

Keywords: Wildlife Trade, K-Means,Classification, Naïve Bayes, Random Forest.

IPB University
Bogor Indonesia
@Hak cipta milik IPB University
IPB University
Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.
Perpustakaan IPB University

© Hak Cipta milik IPB, tahun 2024
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

@Hak cipta milik IPB University

IPB University

IPB University
Bogor Indonesia

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

MODEL KLASIFIKASI DAERAH PERDAGANGAN SATWA LIAR MENGGUNAKAN ALGORITMA *NAÏVE BAYES* DAN *RANDOM FOREST*

HENDRO UTOMO

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister Sains
pada
Program Studi Ilmu Komputer

PROGRAM STUDI ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM INSTITUT PERTANIAN BOGOR
BOGOR
2024

IPB University
Bogor Indonesia
@Hak cipta milik IPB University
IPB University
Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.
Perpustakaan IPB University

Tim Penguji pada Ujian Tesis:

1. Dr. Toto Haryanto S.Kom., M.Si.

@Hak cipta milik IPB University

IPB University

IPB University
Bogor Indonesia

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

Judul : Model Klasifikasi Daerah Perdagangan Satwa Liar Menggunakan Algoritma *Naïve Bayes* dan *Random Forest*

Nama : Hendro Utomo

NIM : G6501201019

Disetujui oleh

Pembimbing 1:
Medria Kusuma Dewi Hardhienata., S.Komp., Ph.D.

Pembimbing 2:
Irman Hermadi, S.Kom., M.S., Ph.D.

Diketahui oleh

Ketua Program Studi:
Irman Hermadi, S.Kom, M.S, Ph.D.
NIP 19750311 200604 1 009

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam:
Dr. Berry Juliandi, S.Si, M.Si.
NIP 19780723 200701 1 001

Tanggal Ujian: 29/07/2024

Tanggal Lulus:

@Hak cipta milik IPB University

IPB University

IPB University — Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Judul yang dipilih dalam penelitian ini dilaksanakan sejak tahun 2022 sampai bulan Juli 2024 ialah Model Klasifikasi Daerah Perdagangan Satwa Liar Menggunakan Algoritma *Naïve Bayes* dan *Random Forest*.

Terima kasih penulis ucapkan kepada pada pembimbing, Medria Kusuma Dewi Hardhienata., S.Komp., Ph.D dan Irman Hermadi, S.Kom., M.S., Ph.D. yang telah membimbing penulis dengan sepenuh hati dan banyak memberi saran dan ilmu pengetahuan. Ucapan terima kasih juga disampaikan kepada Dr. Toto Haryanto S.Kom., M.Si. sebagai penguji luar komisi pembimbing. Kemudian tidak lupa disampaikan terima kasih kepada Alm. Totok Yudiarto A.Md. Akun. dan Aktif Suhartini S.Pd.I. yaitu kedua orang tua saya dan istri dr.Rani Anggraini serta anak saya Khayra Putri Sekar Arum yang senantiasa memberikan dukungan, doa, dan kasih sayangnya.

Penlis menyadari sepenuhnya bahwa dalam penulisan dan penyajian tesis ini masih jauh dari kata sempurna. Penulis berharap akan banyak menerima masukan dari setiap pembaca berupa saran dan kritik yang bersifat membangun. Akhir kata penulis berharap semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

IPB University — Bogor Indonesia —
Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.
Perpustakaan IPB University
©Hak cipta milik IPB University
IPB University

DAFTAR ISI

DAFTAR ISI 12

DAFTAR TABEL 13

DAFTAR GAMBAR 13

I PENDAHULUAN 14

1.1 Latar Belakang 14

1.2 Perumusan Masalah 15

1.3 Tujuan Penelitian 15

1.4 Manfaat Penelitian 16

1.5 Ruang Lingkup Penelitian 16

II TINJAUAN PUSTAKA 16

2.1 Perdagangan Satwa Liar 16

2.2 Machine Learning 16

2.3 *Clustering* K-Means 17

2.4 Algoritma Naïve Bayes 18

2.5 Algoritma *Random Forest* 19

III METODE 20

3.1 Tahapan Penelitian 20

IV HASIL DAN PEMBAHASAN 24

a) Praproses Data 24

V SIMPULAN DAN SARAN 30

5.1 Simpulan 30

5.2 Saran 30

DAFTAR PUSTAKA 31

IPB University — Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah

b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak cipta milik IPB University

IPB University

Perpustakaan IPB University

DAFTAR TABEL

1.Nama Fitur 20
2.Clustering Importer 26
3.Clustering Exporter 26
4.Hasil Test Akurasi 27

DAFTAR GAMBAR

1 Tahapan penelitian 19
2 Hewan yang banyak diperjualbelikan 23
3 Bagian tubuh hewan yang diperjualbelikan 24
4 Negara yang paling banyak impor 24
5 Negara yang paling banyak ekspor 24
6 Perdagangan hewan pertahun 25
7 Perdagangan hewan berdasarkan appendix 25
8 Metode Elbow 26
9 Peta Sebaran Negara yang paling banyak impor 27
10 Peta Sebaran Negara yang paling banyak ekspor 28

IPB University
Bogor Indonesia
@Hak cipta milik IPB University
IPB University
Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.
Perpustakaan IPB University