

IMPLEMENTASI MACHINE LEARNING PADA SISTEM *PETS IDENTIFICATION* MENGUNAKAN PYTHON BERBASIS UBUNTU

Fahrizal¹, Faiz Octa Reynaldi², Noer Hikmah³

Program Studi Sistem Informasi, Universitas Bina Sarana Informatika¹
Program Studi Teknologi Komputer, Universitas Bina Sarana Informatika²
Program Studi Sistem Informasi, Universitas Bina Sarana Informatika³
fahrizal.fzl@bsi.ac.id¹, faizoctareynald@gmail.com², noer.nhh@bsi.ac.id³

Abstrak

Pets Identification adalah salah satu jenis sistem identifikasi “Biometrik” Sistem ini mengidentifikasi hewan dengan fitur khusus pada bagian tertentu yang membedakan satu hewan dengan hewan lainnya. Dalam penelitian ini diimplementasikan dengan Machine Learning. Machine Learning itu sendiri adalah salah satu bagian dari Artificial Intelligence (Kecerdasan Buatan) yang fokus pada pengembangan sebuah sistem yang mampu belajar secara otomatis tanpa harus diprogram oleh manusia. Dengan kata lain, Machine Learning akan belajar dari data yang terdapat dalam sebuah dataset. Proses training dilakukan oleh Machine Learning dengan algoritma Convolutional Neural Network (CNN). Hasil penelitian menunjukkan bahwa akurasi dari Pet’s Identification sangat tergantung dari algoritma Machine Learning dan kualitas dataset yang diproses oleh Machine Learning.

Kata Kunci: Python, Machine Learning, Pets Identification, Linux

I. PENDAHULUAN

Didalam sebuah Machine Learning terdapat beberapa parameter yang mempunyai karakter objek dengan metode klasifikasi ini. Ada dua tahap dalam pembuatan model klasifikasi ini yaitu yang pertama adalah melatih data dan membedakannya dengan kelas yang berbeda, dan yang kedua adalah klasifikasi mengevaluasi objek yang tidak dikenal. Seperti yang dikutip oleh (Matuska, Hudec, Kamencay, Benco, & Zachariasova, 2014)

“The collection of features or parameters characterizing the object by classifications methods to handle classification task are used. There are two phases of creation a classification model. First, training data collections are used to set up the classification model parameters to distinguish different classes. Then, the classifier is able to regarding to classification model parameters successfully evaluate an unknown objects to the appropriate class [4], [5]. In this work, for classification model combination bag of keypoints and Support Vector Machine methods are used.”

Komputer dapat berpikir seperti halnya manusia dan komputer dapat membantu kita untuk memahami bagaimana manusia berpikir seperti yang dikutip oleh (Warwick & Shah, 2016) “What we do is agree with Turing that engineering a machine to think can help us to understand how it is that we humans think”.

Machine Learning telah sukses dalam memprediksi aplikasi keuangan dan banyak penelitian dalam memprediksi kebangkrutan seperti yang dikatakan oleh (Barboza, Kimura, & Altman, 2017) “Machine learning models have been very successful in finance applications, and many studies examine their use in bankruptcy prediction”.

Maksud dari pembuatan program ini adalah untuk mengimplementasikan Kecerdasan Buatan pada Program Pets Identification agar komputer dapat mengidentifikasi suatu objek, dalam hal ini yaitu untuk mengidentifikasi hewan kucing dan anjing. Program implementasi ini akan dikembangkan secara lanjut untuk keperluan Internet of Things di masa yang akan datang.

Telah ada penelitian mengenai aplikasi kecerdasan buatan untuk memecahkan berbagai permasalahan diantaranya dalam bisnis untuk memprediksi kebangkrutan yang tujuannya untuk menilai tingkat risiko dan kegagalan, selain itu dalam dunia digital yang membuat sistem komputer seperti hidup, dapat berpikir dan memiliki intelegensi. Berdasarkan berdasarkan hal tersebut penulis bermaksud untuk menguji apakah machine learning dapat diaplikasikan dalam sistem pets identification untuk mengidentifikasi hewan kucing dan anjing dan disajikan dalam bentuk sebuah program aplikasi. Maksud dari pembuatan program ini adalah untuk mengimplementasikan Program kecerdasan buatan ke dalam sistem Pets Identification pada komputer agar komputer dapat mengidentifikasi suatu objek, dan

program implementasi ini akan dikembangkan secara lanjut untuk keperluan Internet of Things di masa yang akan datang.

Tujuan dari pembuatan program ini adalah untuk pengembangan berikutnya yang akan digabungkan dengan Internet of Things agar manusia dapat berkolaborasi dengan mesin sehingga pekerjaan Manusia lebih mudah, efektif, efisien dari sebelumnya.

Tujuan berikutnya dari pembuatan program ini adalah untuk mengembangkan penelitian yang sudah dilakukan sebelumnya oleh (Matuska, Hudec, Kamencay, Benco, & Zachariasova, 2014).

II. LITERATUR DAN METODE

Agar lebih terarah, metodologi yang digunakan dalam penelitian ini yaitu:

1. Metode Wawancara

Yaitu penulis melakukan wawancara langsung kepada yang berpengalaman di bidang kecerdasan buatan.

2. Metode Kepustakaan

Yaitu dengan mencari dan mempelajari buku-buku serta beberapa e-book dan jurnal-jurnal yang sesuai yang penulis cari di internet guna memberi pemahaman lebih baik terhadap topic penulisan dan memperkaya pengetahuan penulis tentang kecerdasan buatan.

3. Metode Observasi

Yaitu penulis melakukan pengamatan secara langsung pada yang ahli di bidang kecerdasan buatan serta mempraktekkan secara langsung dalam proses pembuatan sistem pet's identification ini.

Fungsi dibuat Program ini adalah untuk membuat mesin kecerdasan buatan yang dapat dilatih dengan cara menggunakan himpunan data dan diproses oleh mesin kecerdasan buatan ini untuk mengidentifikasi sebuah gambar berupa hewan Anjing dan hewan Kucing melalui Perangkat Smartphone Android menggunakan aplikasi Droidcam.

Sebagai referensi dalam penelitian ini memperoleh literatur dari beberapa buku dan penelitian sebelumnya yaitu:

1. Jurnal dengan judul "Integrasi Data Terstruktur dan Tidak Terstruktur dalam Sistem Inteligensi Bisnis" yang ditulis oleh Choerul Afianto, dalam jurnal tersebut disebutkan bahwa terdapat dua jenis data yaitu data yang terstruktur dan data yang tidak terstruktur. Data terstruktur dan tidak terstruktur apabila diintegrasikan dalam menganalisa suatu permasalahan akan memberikan pemahaman dan

solusi yang lebih lengkap dan tepat sasaran. dalam jurnal ini menggunakan teknik text tagging dan annotation dalam mengintegrasikan kedua jenis data tersebut. Teknik tersebut merupakan teknik yang populer dalam natural processing techniques untuk preprocessing data tidak terstruktur agar bisa dengan mudah diintegrasikan dengan data terstruktur.

2. Jurnal berjudul "Machine learning models and bankruptcy prediction" yang ditulis oleh Flavio Barboza, Herbert Kimura dan Edward Altman. Dalam jurnal tersebut, dilakukan pengujian model machine learning (support vector machines, bagging, boosting, dan random forest) untuk memprediksi kebangkrutan satu tahun sebelumnya, dan membandingkan kinerjanya dengan hasil dari analisis diskriminan, regresi logistik, dan jaringan saraf. mereka menggunakan data dari 1985 hingga 2013 pada perusahaan-perusahaan Amerika Utara, mengintegrasikan informasi dari database Salomon Center dan Compustat, menganalisis lebih dari 10.000 pengamatan tahun-perusahaan. Hasil utama dari penelitian ini adalah peningkatan substansial dalam akurasi prediksi menggunakan teknik machine learning terutama ketika, di samping variabel Z-skor Altman asli, mereka menyertakan enam indikator keuangan yang saling melengkapi. Berdasarkan Carton dan Hofer (2006), mereka menggunakan variabel baru, seperti margin operasi, perubahan laba atas ekuitas, perubahan harga menurut buku, dan ukuran pertumbuhan yang terkait dengan aset, penjualan, dan jumlah karyawan, sebagai variabel prediktif. Model pembelajaran mesin menunjukkan, rata-rata, sekitar 10% lebih akurat dalam kaitannya dengan model tradisional. Membandingkan model terbaik, dengan semua variabel prediktif, teknik pembelajaran mesin yang terkait dengan random forest menghasilkan akurasi 87%, sedangkan regresi logistik dan analisis diskriminan linear masing-masing menghasilkan akurasi 69% dan 50%, dalam sampel pengujian. mereka menemukan bahwa model bagging, boosting, dan random forest mengungguli teknik lainnya, dan bahwa semua akurasi prediksi dalam sampel pengujian meningkat ketika variabel tambahan dimasukkan.

3. Jurnal berjudul "Classification of Wild Animals Based on SVM and Local Descriptors". yang ditulis oleh Slavomir Matuska, Robert Hudec, Patrik Kamencay, Miroslav Benco, Martina Zachariasova. Dalam tulisan ini disajikan metode baru untuk pengenalan objek berdasarkan deskriptor lokal hybrid. Metode ini menggunakan kombinasi dari beberapa pendekatan yang terdiri atas dua bagian (SIFT - Scale-

invariant feature transform, SURF - Speeded Up Robust Features). Penerapan metode hybrid yang disajikan ditunjukkan pada beberapa gambar dari dataset. Kelas Dataset mewakili binatang besar yang terletak di negara Slovakia, yaitu serigala, rubah, beruang coklat, rusa dan babi hutan. Metode yang disajikan dapat juga digunakan di area lain dari klasifikasi gambar dan ekstraksi fitur. Hasil percobaan menunjukkan, bahwa kombinasi deskriptor lokal memiliki efek positif untuk pengenalan objek.

4. Buku berjudul “Applied Natural Language Processing with Python. Implementing Machine Learning and Deep Learning Algorithms for Natural Language Processing yang disusun oleh Beysolow II, T. dalam buku tersebut membahas tentang pemanfaatan kemampuan AI untuk natural language processing (NLP), yang dapat melakukan tugas-tugas seperti pemeriksa ejaan, peringkasan teks, klasifikasi dokumen, dan generasi bahasa alami. Dalam buku tersebut, mempelajari keterampilan untuk menerapkan metode ini dalam infrastruktur yang lebih besar untuk mengganti kode yang ada atau membuat algoritma baru. Penerapan Natural language processing menggunakan Python dimulai dengan meninjau konsep-konsep machine learning yang diperlukan sebelum membahas berbagai masalah NLP. Setelah membaca buku tersebut, Pembacanya akan memiliki keterampilan untuk menerapkan konsep-konsep ini di lingkungan sendiri.

Dalam penelitian ini adalah mengembangkan dan mengimplementasikan penemuan seperti yang sudah dilakukan sebelumnya oleh (Matuska, Hudec, Kamencay, Benco, & Zachariasova, 2014) dengan paper yang berjudul Classification of Wild Animals based on SVM and Local Descriptors, dari pengembangan penelitian paper internasional yang sudah dilakukan sebelumnya adalah pendeteksian objek secara realtime dan terintegrasi dengan smartphone berbasis operasi sistem Android. Dalam peninjauan kasus ini hal yang perlu diperhatikan yaitu bagaimana, suatu himpunan data berupa gambar masukan yang telah di input, disimpan, dan diproses dapat dikenali kembali dengan akurat.

III. METODE

Berisi tentang cara penelitian, metode pengambilan data, metode pengembangan, metode analisis, metode perancangan, metode pengujian atau metode lain yang di gunakan untuk pengolahan data serta arah konsep yang di kembangkan merujuk pada metode tertentu.

IV. HASIL DAN PEMBAHASAN (10 pt, 2 kolom)

Berisi tentang hasil penelitian dan atau pengembangan yang akan dilakukan dalam konsep tulisan berdasarkan apa yang di kemukakan pada pendahuluan.

Penelitian yang bersifat ini aplikatif dalam bentuk eksperimen. Langkah awal yang diperlukan adalah membuat sebuah aplikasi yang dapat membedakan sebuah objek. Dalam hal ini yang menjadi objek adalah sebuah objek hewan Anjing dan hewan Kucing. Pada Pembahasan ini akan diuraikan jenis dan cara mendapatkan data, serta cara analisis data dan kerangka penelitian. Dalam spesifikasi rancangan program ini memberikan suatu gambaran mengenai program yang dibuat, terdiri dari spesifikasi bentuk masukan, spesifikasi bentuk keluaran, spesifikasi file, spesifikasi program dan flowchart. Untuk penjelasan lebih lanjut tentang spesifikasi rancangan program ini akan dijelaskan dalam tabel dan pembahasan berikut:

Tabel 1. Spesifikasi Rancangan Program

No	Nama Directory	Fungsi	Media
1	<i>Tensorflow Model Research Object Detection</i>	Sebagai direktori untuk menyimpan suatu file dari algoritma CNN, program generate_tfrecord.py, program xml_to_csv.py, program export_inference_graph.py, program train.py, direktori Spesifikasi Bentuk Masukan, direktori training, direktori data, dan direktori images.	Nautilus file manager
2	training	Sebagai direktori	Nautilus

		untuk menyimpan file berformat ptxt, Checkpoint training, dan tensorboard graph loss function.	file manager
3	images	Sebagai direktori untuk menyimpan file gambar yang akan dilatih oleh mesin dan berisi folder test dan train juga terdapat file xml dan gambar hewan Kucing dan hewan Anjing	Nautilus file manager

Diagram HIPO

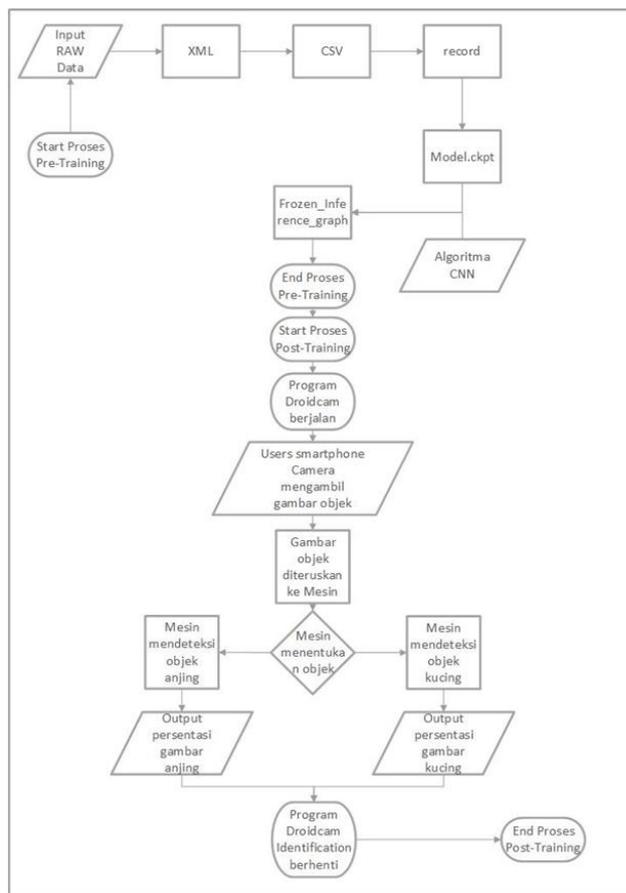
Dalam spesifikasi rancangan program dibutuhkan suatu struktur untuk memahami fungsi dari suatu sistem, input ke suatu program dan output yang dihasilkan oleh suatu program tersebut, berikut ini adalah gambar HIPO dibawah ini :



Gambar 1. Diagram HIPO

FlowChart

Flowchart adalah suatu bagan dengan simbol tertentu yang menggambarkan urutan proses secara mendetail dan hubungan suatu proses dan proses lainnya didalam suatu program. Berikut adalah Flowchart dari program *Pets identification* :



Gambar 2. Diagram Flowchart

Pengujian Black Box Testing

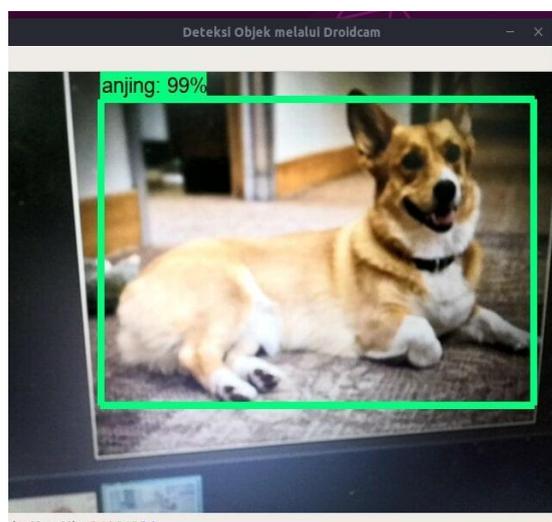
Black box testing dilakukan untuk memastikan apakah suatu masukan dapat menjelaskan proses yang tepat dan menghasilkan output yang sesuai dengan rancangan. Berikut adalah merupakan table pengujian black box testing.

Tabel 2. Pengujian BlackBox Testing

No	Skenari Test	Hasil	yang Ketera
----	--------------	-------	-------------

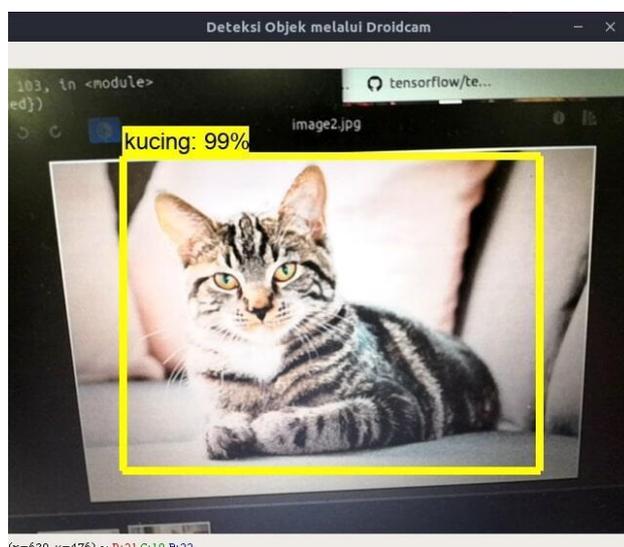
No	Case	diharapkan	ngan
1	Proses Input Gambar	Mena mpilk an Proses Input Gambar	Tampil Home Berhasil
2	Pengena lan Gambar Kucing Dengan Machin e Learnin g	Mena mpilk an Proses Inden tifikas i	Berhasil
3	Pengena lan Gambar Anjing Dengan Machin e Learnin g	Mena mpilk an Proses Inden tifikas i	Berhasil

Dari Hasi Pengujian Diatas maka didapatkan hasil dengan menggunakan machine learning untuk menentukan gambar tersebut 99 pesen adalah benar maka didapat tampilan dibawah ini.



Gambar 3. Identifikasi Gambar Anjing

Saat menginput gambar kucing maka program akan memproses dan mengidentifikasi gambar tersebut. Tampilan diatas menunjukkan bahwa komputer mengenali 99 % gambar tersebut adalah hewan anjing. Begitu pula dengan gambar dibawah ini. Komputer dapat mengidentifikasi 99 % gambar tersebut adalah hewan kucing.



Gambar 4. Identifikasi Gambar Kucing

KESIMPULAN

Berdasarkan pembahasan dan pengujian Pemrograman *Pets Identification*, maka dapat menghasilkan kesimpulan bahwa *Machine Learning* adalah sebagai studi yang memberikan kecerdasan buatan memiliki kemampuan untuk belajar tanpa deprogram. Kemudian himpunan data yang tidak baik dapat menghasilkan ketidak akuratan dalam mendeteksi suatu objek.dan terakhir bahwa situasi pencahayaan, Kualitas perangkat kamera, dan sudut pengambilan gambar berperan penting untuk meningkatkan nilai akurasi dalam mengenal suatu objek.

SARAN

Program *Pets Identification* ini merupakan sebuah aplikasi alternatif bertujuan untuk mengenali suatu objek hewan anjing dan hewan kucing melalui perangkat kamera *Smartphone*, maka dari itu diperlukan komputer dengan spesifikasi yang memadai agar dapat berjalan dengan baik. Program ini menggunakan perangkat yang memiliki resolusi pixel gambar lebih besar dan lebih baik agar proses pendeteksian lebih akurat.

Berisi tentang hasil penelitian secara umum dicapai dan kontribusi bagi pengembangan ilmu juga untuk penelitian berikutnya.

REFERENASI

- [1] Afifanto, C. (2015). Integrasi Data Terstruktur dan Tidak Terstruktur dalam Sistem Inteligensi Bisnis Integrasi Data Terstruktur dan Tidak Terstruktur dalam Sistem. (July). <https://doi.org/10.13140/RG.2.1.3251.8242>
- [2] Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83, 405–417. <https://doi.org/10.1016/j.eswa.2017.04.006>
- [3] Beysolow II, T. (2018). Applied Natural Language Processing with Python. In *Applied Natural Language Processing with Python*. <https://doi.org/10.1007/978-1-4842-3733-5>

- [4] Ernst, M. D. (2017). Natural language is a programming language: Applying natural language processing to software development. *2nd Summit on Advances in Programming Languages (SNAPL'17)*, 71(4), 1–14. <https://doi.org/10.4230/LIPIcs.SNAPL.2017.4>
- [5] Gour, R. (2017). Train and Test Set in Python Machine Learning - How to Split. <https://medium.com/@rinu.gour123/train-and-test-set-in-python-machine-learning-how-to-split-58029a0d657b>
- [6] Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- [7] Machine Learning in Radiation Oncology. (2015). In I. El Naqa, R. Li, & M. J. Murphy (Eds.), *Machine Learning in Radiation Oncology*. <https://doi.org/10.1007/978-3-319-18305-3>
- [8] Nelli, F. (2018). Python Data Analytics. In *Python Data Analytics: With Pandas, NumPy, and Matplotlib: Second Edition*. <https://doi.org/10.1007/978-1-4842-3913-1>
- [9] Peartman, S. (2019). What is Data Preparation, <https://www.talend.com/resources/what-is-data-preparation/>
- [10] Rusk, N. (2016). Deep learning. *Nature Methods*, 13(1), 35–35. <https://doi.org/10.1038/nmeth.3707>
- [11] Warwick, K., & Shah, H. (2016). Can machines think? A report on Turing test experiments at the Royal Society. *Journal of Experimental & Theoretical Artificial Intelligence*, 28(6), 989–1007. <https://doi.org/10.1080/0952813X.2015.1055826>
- [12] Wibowo, A. (2017). 10 Fold-Cross Validation. <https://mti.binus.ac.id/2017/11/24/10-fold-cross-validation/>