

**PENERAPAN MACHINE LEARNING DALAM SISTEM KLASIFIKASI
PENYAKIT MANUSIA DENGAN MODEL *DECISION TREE* DAN *NEURAL
NETWORK***

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Sains Bidang Studi Fisika



Oleh:

**JENNY ALANNA ENGKA
NIM. 08021281722037**

**JURUSAN FISIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA
2021**

LEMBAR PENGESAHAN

PENERAPAN MACHINE LEARNING DALAM SISTEM KLASIFIKASI PENYAKIT MANUSIA DENGAN MODEL DECISION TREE DAN NEURAL NETWORK

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Sains Bidang Studi Fisika

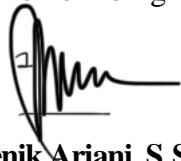
Oleh:

JENNY ALANNA ENGKA
NIM. 08021281722037

Indralaya, November 2021

Menyetujui,

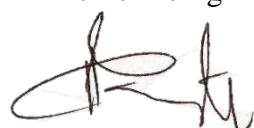
Pembimbing II



Dr. Menik Ariani, S.Si., M.Si

NIP: 197211252000122001

Pembimbing I



Netty Kurniawati, S.Si., M.Si

NIP: 197201031997022002

Mengetahui,



**PENERAPAN MACHINE LEARNING DALAM SISTEM KLASIFIKASI
PENYAKIT MANUSIA DENGAN MODEL *DECISION TREE* DAN *NEURAL
NETWORK***

Oleh:

JENNY ALANNA ENGKA

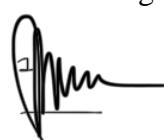
NIM.08021281722037

ABSTRAK

Dengan menggunakan machine learning, sebuah sistem dapat mengklasifikasi penyakit manusia berdasarkan gejala yang dialami seseorang. Tujuan penelitian ini adalah mendapatkan model *machine learning* terbaik untuk mengklasifikasi penyakit menggunakan *decision tree* dan *neural network* dan menggunakan model tersebut untuk dikonversi menjadi *tensorflow lite*. Dataset yang digunakan dalam penelitian ini adalah dataset Kaggle *Disease Prediction Using Machine Learning*. Setelah mendapatkan dataset, dilakukan *data preprocessing* dengan model *decision tree*. Model *neural network* digunakan untuk memprediksi penyakit manusia. Hasil akurasi model akhir dalam proses pelatihan dan validasi adalah 100%, sedangkan akurasi pengujian model adalah 97,6%. Model akan dikonversikan ke tensorflow lite menggunakan method TFLiteConverter agar dapat diimplementasikan dalam sistem klasifikasi penyakit manusia.

Kata Kunci : *Machine Learning*, Klasifikasi Penyakit, *Decision Tree*, *Neural Network*, *Tensorflow Lite*.

Pembimbing II



Dr. Menik Ariani, S.Si., M.Si

NIP: 197211252000122001

Pembimbing I



Netty Kurniawati, S.Si., M.Si

NIP: 197201031997022002

Mengetahui,



**APPLICATION OF MACHINE LEARNING IN HUMAN DISEASE
CLASSIFICATION SYSTEM WITH DECISION TREE AND NEURAL
NETWORK MODELS**

By:

JENNY ALANNA ENGKA

NIM.08021281722037

ABSTRACT

By using machine learning, a system can classify human diseases based on the symptoms experienced by a person. The purpose of this study is to obtain the best machine learning model for classifying diseases using decision trees and neural networks and use these models to be converted into TensorFlow lite. The dataset used in this study is the Kaggle Disease Prediction Using Machine Learning dataset. After getting the dataset, data preprocessing is carried out using the decision tree model. Neural network models are used to predict human disease. The results of the final model accuracy in the training and validation process are 100%, while the accuracy of model testing is 97.6%. The model will be converted to TensorFlow lite using the TFLiteConverter method so that it can be implemented in the human disease classification system.

Key Words : Machine Learning, Disease Classification, Decision Tree, Neural Network, Tensorflow Lite.

Pembimbing II



Dr. Menik Ariani, S.Si., M.Si

NIP: 197211252000122001

Pembimbing I



Netty Kurniawati, S.Si., M.Si

NIP: 197201031997022002

Mengetahui,



KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa karena berkat rahmat dan karunia-Nya skripsi ini dapat dibuat untuk melengkapi persyaratan untuk memperoleh gelar Sarjana Sains Bidang Studi Fisika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya. Saya menyadari bahwa dalam penyusunan skripsi ini masih terdapat banyak kekurangan dan jauh dari kesempurnaan, hal ini disebabkan oleh keterbatasan pengetahuan yang dimiliki oleh penulis. Oleh karena itu, saya sangat memerlukan dan mengharapkan bantuan berupa kritik dan saran yang sifatnya mendidik dan membangun.

Selanjutnya saya sangat mengharapkan agar kiranya skripsi ini dapat diterima oleh pihak terkait dan tidak lupa saya mengucapkan banyak terimakasih atas izin dan kesempatan yang diberikan. Saya mengucapkan banyak terima kasih kepada pihak yang telah banyak membantu selama proses penyusunan skripsi ini mulai dari awal hingga akhir. Terimakasih juga untuk:

1. keluarga tercinta (Mami, Papi, (almh.) Pho-pho, Kung-Kung, Thaiji, dan Thajichong) yang telah memberikan doa dan perhatian yang lebih,
2. Ketua Jurusan Fisika, Bapak Frinsyah Virgo dan dosen Jurusan Fisika serta staff terkait,
3. Ibu Netty Kurniawati selaku Dosen Pembimbing akademik sekaligus Pembimbing I pada skripsi ini,
4. Ibu Menik Ariani selaku Dosen Pembimbing II pada skripsi ini,
5. Ibu Erni sebagai Dosen Pembimbing dalam kegiatan Kampus Merdeka,
6. sahabat terbaik yang menemani saya selama kuliah ini, yaitu FRAKSI PERINDU: Nadia Nur Anisa, Elda Astrita, Rachma Puspa Wardani, Isfi Aprillia, Rendy Malikulmulki Wahid, Evan Kurniadi Wardana, Ryan Apratama dan Amin Rusli yang dengan tulus mendukung dan melengkapi setiap kekurangan saya,
7. *special thanks* juga untuk Dwi Aji Sobarna yang selalu mendukung dan memberi dorongan untuk semakin baik lagi,
8. *a lot of thanks* untuk Ko Ganda dan Mimi Anas sebagai pemimpin rohani saya yang selalu mendoakan dan mendukung saya,
9. *all of my precious family* di AOG GMS Palembang, dan
10. teman-teman satu angkatan Fisika 2017.

Akhir kata, saya mengucapkan banyak terima kasih bagi semua pembaca dan semoga hasil skripsi ini dapat bermanfaat.

Indralaya, September 2021
Penulis,

Jenny Alanna Engka
NIM.08021281722037

DAFTAR ISI

LEMBAR PENGESAHAN	i
ABSTRAK	ii
ABSTRACT	iii
KATA PENGANTAR	iv
DAFTAR ISI	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL	ix
BAB I PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah	2
1.4. Tujuan Penelitian.....	2
1.5. Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA	4
2.1. Kecerdasan Buatan	4
2.2. <i>Machine Learning</i>	5
2.2.1. <i>Supervised Learning</i>	6
2.2.2. <i>Semi-Supervised Learning</i>	6
2.2.3. <i>Unsupervised Learning</i>	7
2.2.4. <i>Reinforcement Learning</i>	7
2.3. <i>Deep Learning</i>	7
2.4. <i>Training, Validation, dan Testing</i>	8
2.5. <i>Generalization, Overfitting, dan Underfitting</i>	9
2.6. Algoritma <i>Decision Tree</i>	10
2.7. Algoritma <i>Neural Network</i>	11
2.7.1. <i>Perceptron</i>	12
2.7.2. <i>Multi Layer Perceptron</i> dan <i>Backpropagation</i>	13
2.7.3. Fungsi Aktivasi.....	14
2.7.4. <i>Learning Rate</i>	15
2.7.5. <i>Loss Function</i>	16
2.7.6. <i>Optimizers</i>	17
2.8. <i>Tensorflow Lite</i>	18

BAB III METODE PENELITIAN	19
3.1. Waktu dan Tempat Penelitian	19
3.2. Alat dan Bahan Penelitian.....	19
3.2.1. Alat	19
3.2.2. Bahan	19
3.3. Tahapan Penelitian.....	19
3.3.1. Menentukan Topik dan Rumusan Masalah.....	19
3.3.2. Studi Pustaka.....	19
3.3.3. Memilih <i>Dataset</i>	19
3.3.4. Membuat <i>Coding</i>	20
3.3.5. Analisa Hasil Model Machine Learning	21
3.3.6. Menarik Kesimpulan.....	21
BAB IV HASIL DAN PEMBAHASAN	22
4.1. <i>Data Preprocessing</i>	22
4.1.1. Pembersihan Data	22
4.1.2. <i>Data Mining</i>	22
4.1.3. Transformasi Data	25
4.2. Perancangan Model <i>Neural Network</i>	26
4.2.1. Penentuan Banyak <i>Neuron</i> pada Model.....	26
4.2.2. Pemilihan <i>Loss Function</i> pada Model	28
4.2.3. Pemilihan <i>Optimizer</i> pada Model	29
4.2.4. Hasil Model Akhir	30
4.3. Pengujian Model <i>Neural Network</i>	31
4.4. Konversi Model Menjadi <i>Tensorflow Lite</i>	32
BAB V PENUTUP	33
5.1. Kesimpulan.....	33
5.2. Saran	33
DAFTAR PUSTAKA	34
LAMPIRAN	36

DAFTAR GAMBAR

Gambar 2.1 Perbedaan pemrograman klasik dan pemrograman pembelajaran mesin	5
Gambar 2.2 Hubungan <i>artificial intelligence, machine learning, dan deep learning</i>	8
Gambar 2.3 Plot area <i>overfitting</i> dan <i>underfitting</i>	9
Gambar 2.4 Pohon keputusan.....	10
Gambar 2.5 Perbedaan neuron biologis dan neuron buatan.....	12
Gambar 2.6 Persamaan TLU	13
Gambar 2.7 Diagram perceptron	13
Gambar 2.8 Multi Layer Perceptron	14
Gambar 2.9 Grafik berbagai fungsi aktivasi	15
Gambar 2.10 Jebakan Gradient Descent	16
Gambar 2.11 Grafik learning rate terlalu kecil, terlalu besar, dan tepat.....	16
Gambar 3.1 Flowchart Coding	20
Gambar 4.1 Hasil Pohon Keputusan (lima node teratas)	23
Gambar 4.2 Akurasi dan loss model 1, model 2, dan model 3.....	27
Gambar 4.3 Akurasi dan loss model dengan optimizer SGD, RMSprop, dan Adam....	29
Gambar 4.4 Hasil akurasi dan loss model akhir	31
Gambar 4.5 Hasil konversi model ke tensorflow lite	32

DAFTAR TABEL

Tabel 2.1 Fungsi aktivasi beserta persamaan matematisnya	14
Tabel 4.1 Hasil precision, recall dan f1-score algoritma decision tree	23
Tabel 4.2 Perbandingan hasil prediksi dan nilai asli algoritma decision tree	24
Tabel 4.3 Lima fitur yang memiliki nilai importance terbesar.....	25
Tabel 4.4 Rincian ketiga model percobaan neural network	26
Tabel 4.5 Perbandingan hasil sebenarnya dan hasil prediksi neural network	31

BAB I

PENDAHULUAN

1.1. Latar Belakang

Dalam kehidupan manusia, kesehatan secara jasmani sangatlah penting. Kesehatan yang baik akan memberikan kualitas hidup yang baik pula dalam setiap pribadi. Tak hanya itu, individu yang sehat dapat berkontribusi nyata dalam pembangunan nasional Indonesia. Maka dari itu, pemerintah Indonesia berupaya meningkatkan kualitas kesehatan setiap masyarakat melalui menyediakan pelayanan kesehatan sebaik mungkin di setiap rumah sakit di Indonesia.

Meskipun pemerintah berusaha menyediakan layanan kesehatan dengan sebaik mungkin, namun dalam pelaksanaannya terdapat berbagai kendala yang terjadi pada rumah sakit. Penanganan rumah sakit sering dianggap membutuhkan waktu yang lama. Kondisi pandemik Covid yang terjadi pada tahun 2019 juga membuat perubahan perilaku masyarakat yang membutuhkan kegiatan tanpa tatap muka. Ini menyebabkan diperlukannya perbaikan sistem dalam menangani setiap pasien dalam rumah sakit.

Seiring perkembangan zaman, teknologi telah berkembang pesat. Penggunaan teknologi untuk memperbaiki sistem ini adalah solusi terbaik yang dapat dilakukan. Gagasan untuk mempermudah penanganan pasien dapat dilakukan dengan adanya aplikasi android layanan kesehatan dengan fitur mengklasifikasi penyakit pasien. Pasien dapat mengetahui kemungkinan penyakit yang diderita berdasarkan gejala yang diderita. Hal ini pun akan membuat kerja dokter lebih efektif dengan menggunakan data.

Untuk mengklasifikasi suatu penyakit manusia dapat digunakan *machine learning*. *Machine learning* atau pembelajaran mesin adalah sub-bidang dari kecerdasan buatan (*artificial intelligence*). Tujuan *machine learning* adalah untuk mengajarkan mesin atau *software* untuk menyelesaikan permasalahan dengan menyediakan beberapa contoh (Richert and Coelho, 2013). Contoh yang diberikan pada *machine learning* berupa kumpulan data. Pada kasus ini, data yang diberikan merupakan data penyakit manusia beserta beberapa gejala yang mengindikasinya.

Dari berbagai model *machine learning* yang ada, penelitian ini akan berfokus menggunakan model *decision tree* dan *neural network*. Menurut Geron (2019), *decision tree* atau pohon keputusan adalah sebuah model kuat, yang mampu mencari hubungan

kompleks nonlinier dalam data. Sedangkan, *neural network* adalah inti dari *deep learning* dimana serbaguna, kuat, dan dapat diskalakan, menjadikannya ideal untuk menangani permasalahan besar dan kompleks. Dari kelebihan dua model inilah yang akan digunakan untuk menangani permasalahan klasifikasi penyakit manusia.

Penjabaran mengenai bagaimana dua model yang telah dijelaskan dapat mengklasifikasi penyakit manusia akan ada dalam penelitian ini. Dengan menggunakan bahasa pemrograman Python, akan diperoleh pemodelan *machine learning* terbaik dalam menangani kasus. Setelah mendapatkan model terbaik, model tersebut akan diimplementasikan ke dalam aplikasi android dengan mengkonversinya menjadi *TensorFlow Lite*. Penelitian ini pun akan menjelaskan bagaimana pengkonversian model.

1.2. Rumusan Masalah

Rumusan masalah yang bisa ditarik sesuai dengan latar belakang adalah sebagai berikut.

1. Bagaimana perancangan model *machine learning* untuk mengklasifikasi penyakit manusia menggunakan model *decision tree* dan *neural network*?
2. Bagaimana mengkonversi model *machine learning* menjadi *Tensorflow Lite* untuk dapat diimplementasikan ke dalam sistem klasifikasi penyakit manusia?

1.3. Batasan Masalah

Dalam penelitian ini terdapat beberapa batasan masalah yang menjadi fokus persoalan:

1. dataset yang digunakan adalah dataset Kaggle *Disease Prediction Using Machine Learning*,
2. merancang model *machine learning* untuk mengklasifikasi penyakit manusia,
3. model yang digunakan adalah *decision tree* dan *neural network*, dan
4. mengkonversi model *machine learning* menjadi *Tensorflow Lite*.

4.3. Tujuan Penelitian

Adapun berikut beberapa tujuan dari penelitian ini.

1. Mendapatkan model *machine learning* terbaik untuk mengklasifikasi penyakit manusia menggunakan *decision tree* dan *neural network*.
2. Menggunakan model *machine learning* yang telah dikonversi menjadi *Tensorflow Lite* ke dalam sistem klasifikasi penyakit manusia.

4.4. Manfaat Penelitian

Berikut manfaat yang didapatkan dari hasil penelitian ini.

1. Mengembangkan teknologi kesehatan yang dapat digunakan agar pelayanan di rumah sakit lebih efisien.
2. Dapat mengklasifikasi penyakit manusia berdasarkan gejala yang dialami pasien.

DAFTAR PUSTAKA

- Alsing, O. 2018. *Mobile Object Detection using Tensorflow Lite and Transfer Learning*. (Tesis Master in Computer Science, KTH Royal Institute Of Technology, 2018) Diakses dari <http://www.diva-portal.org/smash/get/diva2:1242627/FULLTEXT01>.
- Amrizal, V. dan Aini, Q. 2013. *Kecerdasan Buatan*. Jakarta: Halaman Moeka.
- Basheer, I. A. dan Hajmeer, M. 2000. *Artificial Neural Networks: Fundamentals, Computing, Design, and Application*. Journal of Microbiological Methods, 4(3): 5.
- Burkov, A. 2019. *The Hundred-Page Machine Learning*. Amerika: Andriy Burkov.
- Chollet, F. 2018. *Deep Learning with Python*. New York: Manning Publication.
- Geron, A. 2019. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (2nd ed)*. Sebastopol: O'Reilly Media.
- Jijo, B. T. dan Abdulazeez, A. M. 2021. *Classification Based on Decision Tree Algorithm for Machine Learning*. Journal of Applied Science and Technology Trends, 2(1): 21.
- Kabari, L. G. dan Nwachukwu, E. O. 2013. *Decision Support System Using Decision Tree and Neural Networks*. Journal Computer Engineering and Intelligent Systems, 4(7): 13.
- Kasih, P. 2019. *Pemodelan Data Mining Decision Tree dengan Classification Error untuk Seleksi Calon Anggota Tim Paduan Suara*. Journal Innovation in Research of Informatics, 1(2): 64-65.
- Muller, A. C. dan Guido, S. 2017. *Introduction to Machine Learning with Python*. Sebastopol: O'Reilly Media.
- Negnevitsky, M. 2005. *Artificial Intelligence: A Guide to Intelligent Systems (2nd ed)*. London: Pearson Education.
- Purnama, J. J dkk., 2019. *Analisis Algoritma Klasifikasi Neural Network Untuk Diagnosis Penyakit Diabetes*. Indonesian Journal on Computer and Information Technology, 5(1): 2-3.
- Putra, J. W. G. 2020. *Pengenalan Konsep Pembelajaran Mesin dan Deep Learning*. Tokyo: Jan Wira Gotama Putra.
- Ranjan, C. 2020. *Understanding Deep Learning Application in Rare Event Prediction*. Amerika: Connaissance Publishing.

- Richert, W. dan Coelho, L.P. 2013. *Building Machine Learning Systems with Python*. Birmingham: Packt Publishing.
- Russell, S. dan Norvig, P. 2010. *Artificial Intelligence: A Modern Approach (3rd ed)*. New Jersey: Pearson.
- Santoso, R. R., Megasari, R., dan Hambali, Y. A. 2020. *Implementasi Metode Machine Learning Menggunakan Algoritma Evolving Artificial Neural Network pada Kasus Prediksi Diagnosis Diabetes*. Jurnal Aplikasi dan Teori Ilmu Komputer, 3(2): 88-89.
- Shwart, S.S. and David, S. B. 2014. *Understanding Machine Learning From Theory to Algorithms*. New York: Cambridge University Press.
- Sulistiono, H. 2015. *Kajian Penerapan Algoritma C4.5 Neural Network dan Naïve Bayes untuk Klasifikasi Mahasiswa yang Bermasalah dalam Registrasi*. Jurnal Faktor Exacta, 8(5): 306.
- Kaushil. 2020. “Disease Prediction Using Machine Learning”, <https://www.kaggle.com/kaushil268/disease-prediction-using-machine-learning>, diakses pada 7 September 2021.