

# Implementasi Multiple Layer Perceptron Untuk Klasifikasi Gender Berdasarkan Tinggi dan Berat Badan

## *Implementation of Multiple Layer Perceptron for Gender Classification Based on Height and Weight Of Body*

Ari Susanto

STMIK PONTIANAK, Jl. Merdeka No.374

e-mail: [daber.coding@gmail.com](mailto:daber.coding@gmail.com)

### **Abstrak**

*Klasifikasi gender merupakan suatu permasalahan yang umum didapati dengan standar yang berbeda – beda, baik dalam menentukan ciri fisik yang ideal pada setiap individu, ciri umum yang signifikan yang menjadi pembeda antara wanita dan pria adalah tinggi badan dan berat badan, akan tetapi tinggi badan berat badan pada suatu wilayah tertentu memiliki perbedaan yang merupakan keberagaman genetik, yang juga menjadi tantangan dalam melakukan klasifikasi terhadap perbedaan gender berdasarkan tinggi badan dan berat badan dengan tidak adanya standar yang pasti, dengan menggunakan Multiple Layer Perceptron (MLP) tentunya akan dapat mengatasi masalah dalam melakukan klasifikasi yaitu dengan membuat model yang dapat mempelajari data dengan optimisasi menggunakan backpropagation dan jumlah data pelatihan sebanyak 3000 data dan data uji sebanyak 205 menunjukkan bahwa model berhasil mempelajari data dengan akurasi 100%, model dibuat dengan menggunakan bahasa pemrograman C++ dengan menggunakan library Mlpack*

**Kata kunci**— Klasifikasi, Gender, Multiple Layer Perceptron(MLP), Mlpack

### **Abstrac**

*Gender classification is a problem that is commonly found with different standards, both in determining the ideal physical characteristics of each individual, a significant common feature that distinguishes between women and men is height and weight, but height and weight in a particular region have differences that are genetic diversity, which is also a challenge in classifying gender differences based on height and weight in the absence of definite standards. by using Multiple Layer Perceptron (MLP) will certainly be able to overcome the problem of classification by creating a model that can learn data with optimization using backpropagation and the amount of training data using 3000 data and test data of 205, by using Multiple Layer Perceptron (MLP) will certainly be able to overcome the problems in classifying, namely by creating a model that can learn data with optimization using backpropagation and the amount of training data as much as 3000 data and 205 test data shows that the model successfully learns data with 100% accuracy, the model is made using the C++ programming language using the Mlpack library.*

**Keywords**— Classification, Gender, Multiple Layer Perceptron (MLP), Mlpack

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Klasifikasi gender berdasarkan karakteristik dalam berbagai sektor seperti industri kecantikan, penerbangan, dan militer. Telah menjadi bagian yang cukup penting dan cukup meluas dalam pengaplikasiannya, baik dalam konteks bidang industri kecantikan dalam upaya melakukan pengembangan produk agar mencapai target yang diinginkan seperti pakaian dan peralatan yang digunakan dalam kehidupan sehari – hari, juga dalam upaya meningkatkan kenyamanan dan memperluas pasar dengan tujuan meningkatkan kepuasan konsumen.

Dalam sektor penerbangan penentuan gender dengan menggunakan karakteristik berdasarkan ciri fisik berat badan dan tinggi badan tentunya akan memaksimalkan dalam pengaturan beban dan pendistribusian penumpang di dalam pesawat dengan tujuan meningkatkan efisiensi operasional dan keamanan penumpang.

Sementara itu pada sektor militer dalam menentukan personel berdasarkan ciri fisik dan gender seperti tinggi badan dan berat badan yang dengan menggunakan standard yang telah ditentukan akan sangat membantu dalam mengembangkan kebijakan terkait sumber daya manusia dan penempatan penugasan personel dan pendistribusian kebutuhan fisik individu yang dapat mempengaruhi kinerja personel di lapangan.

Akan tetapi dalam melakukan klasifikasi berdasarkan berat badan dan tinggi badan tentunya menjadi kendala tersendiri, terutama dalam menentukan berat badan dan tinggi badan yang ideal pada sekelompok individual dengan gender tertentu, akan tetapi permasalahan ini dapat diselesaikan dengan menggunakan *Multiple Layer Perceptron* (MLP), yang merupakan salah satu pengembangan dari model jaringan perceptron.

*Multiple Layer Perceptron* (MLP) merupakan salah satu jenis jaringan syaraf tiruan yang dimodelkan dalam bentuk hubungan yang kompleks antar node – node yang terhubung, dengan tiga bagian utama yaitu input layer, hidden layer, dan output layer [1]. Umumnya pada hidden layer dapat memiliki sub-layer lebih dari satu, dan salah satu penggunaan MLP umumnya untuk melakukan klasifikasi berdasarkan data – data dalam bentuk vektor, dan dilatih dengan menggunakan algoritma *backpropagation*, dengan menggunakan MLP maka permasalahan dalam melakukan klasifikasi berdasarkan berat badan akan dapat diselesaikan dengan membuat sebuah model yang mampu mempelajari *dataset* yang ada dalam melakukan klasifikasi.

Desain model yang telah dibuat akan diimplementasikan dengan menggunakan bahasa pemrograman C++ dengan menggunakan library Mlpack dengan tujuan untuk memaksimalkan efisiensi waktu dan kecepatan dalam komputasi, dalam pembuatan dan pelatihan model akan ada beberapa hal yang harus diperhatikan yaitu pemilihan data pelatihan, pemilihan struktur atau model dari jaringan yang akan digunakan, *error function* yang digunakan dan algoritma optimisasi yang digunakan.

Mlpack merupakan pustaka atau library C++ yang dikembangkan oleh NumFocus, “*mlpack is an intuitive, fast, and flexible header-only C++ machine learning library with bindings to other languages. It is meant to be a machine learning analog to LAPACK, and aims to implement a wide array of machine learning methods and functions as a "swiss army knife" for machine learning researchers*” [2]

### 1.2 Rumusan Masalah

Tantangan utama dalam klasifikasi gender berdasarkan tinggi dan berat badan adalah memastikan akurasi dan konvergensi dalam proses pelatihan model. Variasi didalam data yang signifikan serta perubahan dalam waktu dapat menyulitkan model dalam melakukan klasifikasi yang

konsisten. Hal ini menuntut penggunaan pendekatan yang adaptif seperti Multiple Layer Perceptron (MLP), untuk mengidentifikasi pola-pola yang mungkin tidak terdeteksi oleh metode konvensional. Sehingga dalam melakukan pelatihan pada model akan ada beberapa faktor yang menjadi permasalahan yaitu:

- 1) Bagaimana meningkatkan model yang dibuat agar berhasil mencapai tingkat akurasi dan konsistensi dalam mengolah data yang diberikan, sehingga dapat membantu pengambilan keputusan yang lebih akurat dan efisien?

### 1.3 Tujuan

Tujuan dari penelitian ini adalah untuk mengembangkan model klasifikasi gender berdasarkan tinggi dan berat badan menggunakan Multiple Layer Perceptron (MLP) yang dapat:

- 1) Meningkatkan akurasi dan konsistensi prediksi gender berdasarkan data empiris sehingga dapat membantu pengambilan keputusan yang lebih akurat dan efisien.

## BAB II

### METODE PENELITIAN

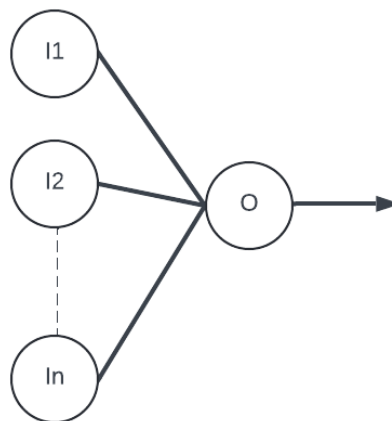
#### 2.1 Metode Penelitian

Metode penelitian yang digunakan adalah metode kuantitatif dengan menggunakan data sekunder dalam menjawab atau mencapai tujuan yang telah didefinisikan pada perumusan masalah, penelitian kuantitatif ini merupakan penelitian yang berfokus pada analisa dan pengumpulan data yang berbentuk numerik baik untuk menjelaskan, memprediksi, dan memahami fenomena yang sedang diamati

#### 2.2 Perceptron

Multiple Layer Perceptron merupakan pengembangan dari jaringan syaraf perceptron dengan membentuk model menjadi beberapa layer perceptron sendiri hanya menerima input dan menghasilkan output seperti gambar 1.1

Gambar 1.1 Perceptron



Dimana input mulai dari  $I_1, I_2, \dots, I_n$  merupakan input dan  $O$  merupakan output, dalam proses perhitungannya pada setiap edge atau penghubung antara node input ke node output akan terdapat bobot yang diinisialisasikan secara acak, dalam pemrosesan dari input hingga ke output vektor input  $I$  akan dikalikan dengan vektor bobot  $W$  dengan menggunakan dot product kemudian hasil perkalian menjadi input bagi fungsi aktivasi, proses input hingga output dapat diuraikan menjadi berikut.

$$\begin{bmatrix} I_1 \\ I_2 \\ \cdot \\ \cdot \\ I_n \end{bmatrix} \times [W_1 \quad W_2 \quad \dots \quad W_n] = y_{in} \quad (1)$$

Pada persamaan diatas vektor input di kalikan dengan vektor bobot, sehingga menghasilkan output yaitu  $y_{in}$  kemudian hasil dari perkalian disubstitusikan kepada fungsi aktivasi, pada jaringan perceptron sederhana fungsi aktivasi hanya menggunakan tresshold biasa seperti pada persamaan berikut.

$$f(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (2)$$

Kemudian dilakukan pengkondisian dimana jika hasil aktivasi sama dengan target dari input maka bobot tidak akan diperbarui, akan tetapi jika sebaliknya maka bobot akan di perbarui dengan persamaan sebagai berikut.

$$\begin{aligned} w_i &= w_{i-1} + \Delta w_i \\ \Delta w_i &= \eta(t - y)x_i \end{aligned} \quad (3)$$

Dimana variabel  $w_i$  merupakan bobot ke  $i$  dan  $\eta$  merupakan learning rate, sedangkan  $t$  merupakan target dari input dan  $y$  merupakan output dari jaringan dan  $x_i$  merupakan input data ke  $i$

### 2.3 Multiple Layer Perceptron

Multiple layer perceptron merupakan pengembangan dari single perceptron atau jaringan perceptron tunggal, yaitu dengan cara menyusun jaringan perceptron menjadi tiga layer yaitu input, hidden layer, dan output pada proses perhitungan tidak ada perubahan melainkan hasil output dari layer sebelumnya akan menjadi input bagi layer berikutnya dan ada penambahan pada saat melakukan perkalian pada vektor input dan vektor output yaitu bias, umumnya bias diinisialisasikan dengan nilai acak sama seperti bobot sehingga persamaan pada  $y_{in}$  menjadi sebagai berikut.

$$\begin{bmatrix} I_1 \\ I_2 \\ \cdot \\ \cdot \\ I_n \end{bmatrix} \times [W_1 \quad W_2 \quad \dots \quad W_n] + b = y_{in} \quad (4)$$

Pada multiple layer perceptron bobot diperbarui dengan metode yang berbeda yaitu dengan menggunakan algoritma backpropagation, yaitu dengan menghitung gradient dari bobot ataupun bias yang akan di perbaharui terhadap fungsi – fungsi yang dilaluinya dengan algoritma berikut:

- 1) Feed Forward pada jaringan
- 2) Menghitung error pada jaringan berdasarkan output, pada single output umumnya dapat digunakan Sum Square Residual (SSR), Means Square Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), dan Root Mean Square Error (RMSE). Sedangkan pada multiple output atau pada klasifikasi umumnya digunakan Cross-Entropy dan negative log-likelihood
- 3) Menghitung gradient pada bobot dan bias pada setiap layer
- 4) Melakukan pembaruan pada bobot dan bias
- 5) Lakukan langkah 1 – 5 hingga mencapai iterasi maksimum atau target loss yang diinginkan

### 2.4 Backpropagation

Pada tipe jaringan yang kompleks seperti MLP untuk mengoptimasi bobot dan bias digunakan algoritma backpropagation, algoritma ini memperbarui bobot dan bias dengan mengalikan  $\Delta w_i$  dengan gradient dari bobot atau bias yang akan di update, dengan algoritma sebagai berikut:

- 1) Menghitung gradient pada bobot dan bias yang akan diperbarui melalui persamaan berikut.

$$\frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial y_{out}} \frac{\partial y_{out}}{\partial y_{in}} \cdots \frac{dy_{in}}{dw_i} \quad (5)$$

$$\frac{\partial E}{\partial b_i} = \frac{\partial E}{\partial y_{out}} \frac{\partial y_{out}}{\partial y_{in}} \cdots \frac{dy_{in}}{db_i}$$

- 2) Hasil dari perhitungan pada gradient bobot dan bias kemudian dikalikan dengan  $\Delta w_i$  dan di kurangkan dengan bobot dan gradient yang ingin diperbarui dengan persamaan berikut.

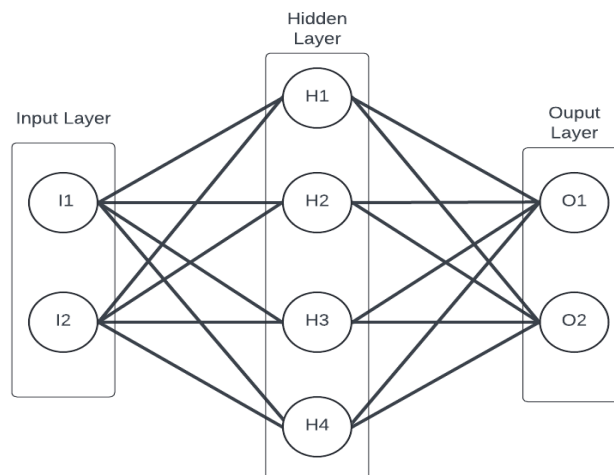
$$w_i = w_{i-1} - \Delta w_i \times \frac{\partial E}{\partial w_i} \times x_i \quad (6)$$

$$b_i = b_{i-1} - \Delta w_i \times \frac{\partial E}{\partial b_i}$$

## 2.5 Desain Model

Pada model untuk klasifikasi gender berdasarkan berat badan dan tinggi badan ini akan dibentuk dengan dua node input, empat hidden node, dan dua node output seperti pada gambar 1.2

Gambar 1.2 Model Jaringan



dan untuk error fungsi yang digunakan yaitu Cross-Entropy dengan fungsi aktivasi pada hidden layer yaitu sigmoid dan fungsi aktivasi pada output layer yaitu SoftMax. Model akan dibuat dengan menggunakan bahasa pemrograman C++ dengan menggunakan library Mlpack, dengan tujuan untuk memaksimalkan proses komputasi dan resource pada proses pelatihan.

## BAB III

### HASIL DAN PEMBAHASAN

#### 3.1 Hasil Pelatihan

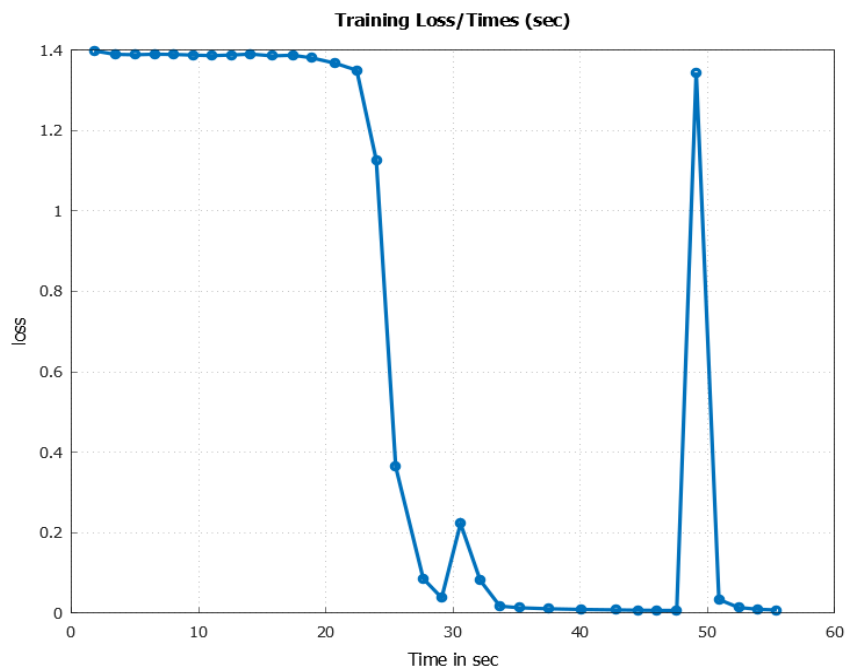
Pada pelatihan dibutuhkan waktu selama 52 detik bagi model untuk mencapai konvergensi dengan epoch sebanyak 34 epoch dengan detail pada tabel 1.0

Tabel 1.0 Detail Pelatihan Model

iter	loss	loss change	gradient	step size	total time
0	1.399	0.000	17.537	0.010	1.773
3	1.390	0.008	14.474	0.010	6.520
6	1.387	0.003	8.750	0.010	11.034
9	1.386	0.001	13.913	0.010	15.785
12	1.368	0.018	7.263	0.010	20.708
15	0.364	1.004	5.462	0.010	25.477
18	0.223	0.141	65.529	0.010	30.572
21	0.013	0.210	0.276	0.010	35.227
24	0.008	0.005	0.095	0.010	42.830
27	0.006	0.001	0.058	0.010	47.570
30	0.014	-0.008	0.544	0.010	52.429

dan untuk perubahan loss atau error dari model selama pelatihan sebagai berikut:

Gambar 1.3 Plot loss / error pada pelatihan model



dan pada tingkat akurasi model berhasil mencapai 100% tingkat akurasi pada dataset testing dengan confusion matrix sebagai berikut

Table 1.1 Confusion Matrix

	P	N
P	205	0
N	0	0

Accuracy	1
Precision	1
Recall	1

hasil dari perhitungan terhadap akurasi, presisi, dan recall yang sangat sempurna, sebelumnya telah dilakukan beberapa uji coba seperti mengganti fungsi aktivasi pada hidden layer dengan ReLU (Rectified Linear Unit) didapati bahwa model gagal dalam mencapai konvergensi dengan epoch 34, hal yang sama juga didapatkan ketika data pelatihan dinormalisasi dengan rentang 0 – 1, dan ketika jaringan dioptimasi menggunakan Adam Optimizer juga didapati hal yang serupa.

Model dibuat dengan menggunakan library Mlpack dengan bahasa pemrograman C++ dengan default parameter pada model yang sudah cukup mampu dalam mencapai konvergensi dengan default epoch yaitu 34 epoch.

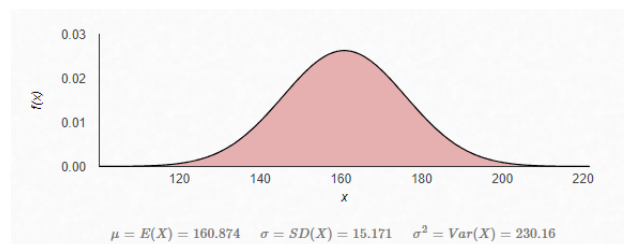
Input data pada pelatihan model memiliki hasil analisis statistik sebagai berikut:

Table 1.3 Mean dan Standard Deviation dari data

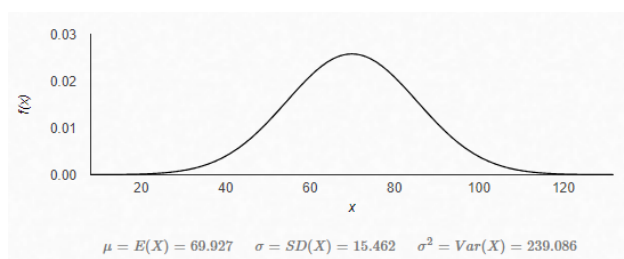
	Mean	Stdv
Tinggi	160,8741248	15,17104
Berat	69,92702849	15,4624

dengan distribusi normal pada tinggi badan dan berat badan sebagai berikut:

Gambar 1.3 Normal Distribution pada tinggi badan



Gambar 1.4 Normal Distribution pada berat badan





## BAB IV

### KESIMPULAN DAN SARAN

Berdasarkan hasil dari analisa pembelajaran model yang telah dibuat dapat disimpulkan, dalam melakukan pembuatan model parameter – parameter seperti learning rate, fungsi aktivasi, data input akan sangat berpengaruh pada hasil pembelajaran dan konvergenitas dari model, tentunya model yang dibuat masih memiliki banyak kekurangan baik dalam segi optimasi pelatihan dan pengenalan terhadap data yang cukup bervariasi pada pelatihan digunakan sebanyak 3000 data sekunder dengan 250 data pengujian. Maka, point – point utama pada pada hasil penelitian ini adalah:

Analisis dari pembuatan dan pelatihan model menunjukkan parameter seperti laju pembelajaran, fungsi aktivasi, dan data masukan secara signifikan mempengaruhi hasil pembelajaran dan konvergensi model.

Model yang dibangun masih memiliki banyak kekurangan dalam hal optimasi pelatihan dan pengenalan data yang beragam.

#### Poin-poin Utama dari Penelitian

- 1) Pengaruh Parameter Model: Parameter seperti laju pembelajaran, fungsi aktivasi, dan data input memiliki pengaruh yang besar terhadap hasil pembelajaran dan konvergensi model.
- 2) Kekurangan Model: Model yang dibuat masih memiliki banyak kekurangan dalam mengoptimalkan pelatihan dan mengenali data yang beragam.

#### Saran untuk peneliti berikutnya

- 1) Bereksperimen dengan laju pembelajaran yang berbeda untuk menemukan nilai optimal yang memungkinkan model untuk berkonvergensi secara efektif.
- 2) Mencoba model dengan berbagai fungsi aktivasi (misalnya, ReLU, Sigmoid, Tanh) untuk menentukan mana yang memberikan hasil terbaik untuk set data tertentu.
- 3) Augmentasi dan Prapemrosesan Data:
  - a) Tingkatkan keragaman dan volume data pelatihan melalui teknik augmentasi.
  - b) Melakukan preprocessing data yang lebih menyeluruh untuk memastikan data bersih dan diformat dengan tepat untuk pelatihan

DAFTAR PUSTAKA

- [1] J. W. G. Putra, *Pengenalan Konsep Pembelajaran Mesin dan Deep Learning*, 1.4. Tokyo, 2020.
- [2] R. R. Curtin *et al.*, “mlpack 4: a fast, header-only C++ machine learning library,” *J Open Source Softw*, vol. 8, no. 82, p. 5026, 2023, doi: 10.21105/joss.05026.