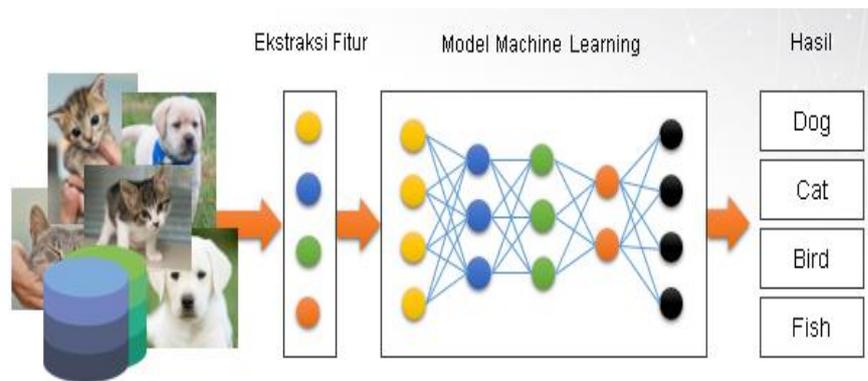


# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

*Machine learning* sebuah teknik bagaimana cara komputer mampu belajar seperti manusia dengan mengenal data yang diberikan [1]. Misal pada gambar 1.1 adalah gambar yang akan diberikan kepada komputer untuk membuat menjadi cerdas yang dapat mendeteksi jenis binatang seperti gambar 1.1 berikut:



**Gambar 1.1** *Machine Learning*

Permasalahan pada gambar 1.1 mempresentasikan prediksi hewan apa yang ada pada kumpulan data tersebut dan hasil adalah mempresentasikan komputer memprediksi setiap data yang dimasukkan. Hubungannya dengan pada gambar 1.1 adalah mempresentasikan dari tujuan tugas akhir ini yaitu menerapkan model *machine learning* dengan *titanic survivor problem*. *Titanic Survivor problem* yaitu sebuah dataset yang diberikan berisikan data-data penumpang selamat dan tidak selamat yang akan diekstraksi fitur untuk memodelkan *machine learning*. Tugas *machine learning* memprediksi keselamatan pada penumpang pada tenggelamnya kapal adalah hasil dari *machine learning*. *Titanic survivor problem* dalam memprediksi selamat atau tidak dengan menggunakan model *machine learning*

dengan melihat faktor-faktor apa saja penumpang selamat atau tidaknya dalam tenggelamnya kapal *titanic* [2]. Tugas akhir ini menggunakan *random forest* sebagai model *machine learning* dalam implementasi masalah yang dihadapi.

*Random forest* merupakan metode *tree* gabungan yang berasal dari pengembangan metode *classification and regression tree* (CART) atau algoritma *decision tree*, yaitu dengan menerapkan metode *bootstrap aggregating* (*bagging*) dan *random feature selection* [3].

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dari penelitian ini adalah bagaimana cara menerapkan model *machine learning Random forest* pada *Titanic survivor problem*.

## **1.3 Tujuan**

Tujuan tugas akhir ini dibuat adalah untuk dapat menerapkan metode *random forest* dalam memprediksi penumpang selamat atau tidak pada *titanic survivor problem* dengan bahasa pemrograman *python*.

## **1.4 Manfaat**

Manfaat penelitian ini adalah sebagai:

- 1) Memahami *Machine learning* secara umum
- 2) Memahami dan mampu menerapkan *Random forest* pada bahasa pemrograman *Python*.
- 3) Memahami *Titanic survivor problem*.

- 4) Memahami dan mampu menggunakan dataset terutama dataset *titanic survivor problem*.
- 5) Mampu menerapkan *random forest* pada *dataset titanic* dengan bahasa pemrograman *python*.
- 6) Memperoleh hasil prediksi *Titanic Survivor* dengan model *Random forest*.

### 1.5 Batasan Masalah

Agar pengerjaan tugas akhir ini menjadi lebih terarah dan mendapatkan hasil yang lebih spesifik, maka penelitian ini yang dibatasi sebagai berikut:

- 1) Dataset yang digunakan adalah *dataset titanic* dan dataset diperoleh melalui *Kaggle*.
- 2) *Tuning* parameter *random forest* *sckit-learn* hanya menggunakan *n\_estimator* dan tidak menggunakan *tuning hyperparameter*.

### 1.6 Metodologi Penelitian

Metodologi penelitian untuk analisa algoritma dari *random forest* pada dataset *titanic* terdiri atas:

- I. Pengumpulan Data
  - Studi literatur, yakni mempelajari referensi mengenai *Random forest* dan *Titanic survivor problem* dari jurnal, buku, internet, maupun media lainnya.
  - Mencari dataset dari *Kaggle*.
- II. Analisis dan Perancangan, yang meliputi:
  - a. Analisis algoritma dari model *Random forest*.

- b. Analisis *Titanic survivor problem*.
- c. Analisis dataset *Titanic survivor problem* yang diperoleh dari Kaggle.
- d. Perancangan alur penerapan model *Random forest* pada dataset *Titanic survivor problem*.

### III. Implementasi

Pelaksanaan implementasi *machine learning* yakni teknik klasifikasi *supervised* dengan model *Random forest* pada *dataset titanic* guna menyelesaikan *Titanic survivor problem* berupa prediksi *titanic survivor* dengan bahasa pemrograman *python*.

## 1.7 Sistematika Penulisan

Penulisan laporan Tugas Akhir ini terbagi dalam beberapa bab yang tersusun sebagai berikut:

### **BAB I PENDAHULUAN**

Bab ini berisi latar belakang, rumusan masalah, tujuan dan manfaat, batasan masalah, metodologi dan sistematika penulisan laporan.

### **BAB II LANDASAN TEORI**

Pada bab ini diuraikan tentang teori yang berhubungan dengan judul tugas akhir, seperti *machine learning*, *Random forest*, *Titanic survivor problem* dan lain-lain.

### **BAB III ANALISIS DAN PERANCANGAN**

Bab ini berisi uraian tentang analisis algoritma dari model *Random forest* melakukan analisis *titanic survivor problem* berisi analisis dataset *titanic survivor problem*. Proses alur penerapan *titanic survivor problem* ini juga ditampilkan dalam bentuk *flowchart* untuk memudahkan dalam memahami setiap proses yang dilakukan dan melakukan modeling *Random forest* dan *Titanic survivor problem* ini kedalam bahasa pemrograman *python* dengan *jupyter notebook*.

### **BAB IV IMPLEMENTASI**

Bab ini berisi uraian tentang tahapan implementasi bahasa pemrograman *python* pada *Titanic survivor problem* dan pengujian parameter tuning untuk *n\_estimator*.

### **BAB V PENUTUP**

Bab ini berisi kesimpulan dari keseluruhan tugas akhir yang telah dilakukan serta saran untuk kegiatan penelitian selanjutnya terkait dengan tugas akhir yang telah dilaksanakan.