

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang

Kemajuan teknologi yang telah berkembang secara pesat mengakibatkan persaingan secara kompetitif dalam bisnis. Perusahaan yang memegang peran penting dalam hal ini dituntut untuk dapat beroperasi secara efektif dan efisien agar mampu memenangkan persaingan dan memaksimalkan keuntungan. Salah satu hal terpenting dalam bisnis yaitu dengan mengetahui jenis barang yang paling banyak dan kurang diminati yang dijual pada periode tertentu. Barang yang banyak diminati dapat diartikan sebagai barang yang bisa digunakan atau dijual pada masa atau periode yang akan datang. Dengan menerapkan hal tersebut, aktivitas dan pelayanan perusahaan akan berjalan lancar sehingga dapat memberikan kepuasan terhadap para pelanggan serta dapat menjamin kelangsungan hidup perusahaan.

*Aliexpress* adalah layanan ritel *online* yang berbasis di *china* yang dimiliki oleh *Alibaba* yang diluncurkan pada tahun 2010 oleh *Jack Ma*. *Aliexpress* merupakan situs web *e-commerce* yang memfasilitasi usaha kecil untuk menjual kepada pelanggan diseluruh dunia, Penjual di *Aliexpress* dapat berupa perusahaan atau perorangan. *Aliexpress.com* setiap harinya mencatat berbagai macam transaksi penjualan yang banyak. Hal tersebut berdampak pada pertumbuhan jumlah data yang sangat pesat sehingga menimbulkan tumpukan data yang berjumlah besar, Gaun merupakan salah satu hal yang paling dibutuhkan manusia dalam kehidupan sehari-hari. dalam memilih gaun yang akan di beli proses pembelian gaun dimulai dari pemilihan model yang sesuai dengan selera. Namun banyaknya jenis gaun yang ada pada Toko *Online Aliexpress* membuat Pembeli mengalami kesulitan dalam memilih gaun yang akan dibeli, dengan mengklasifikasi jenis gaun yang maka akan membuat pembeli semakin mudah dalam memilih gaun sesuai apa yang diinginkan dengan cepat.

Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas. Klasifikasi melibatkan

proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya. Pada penelitian ini metode yang diusulkan untuk klasifikasi rekomendasi penjualan adalah metode *Decision Tree* yaitu membandingkan kinerja algoritma *C4.5*, dengan *Random Forest*. Algoritma tersebut nantinya akan dibandingkan untuk klasifikasi dan tingkat akurasi. Algoritma *C4.5* merupakan Algoritma yang di gunakan untuk membentuk sebuah pohon keputusan (*Decision Tree*), pohon keputusan berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi.

Pada penelitian sebelumnya yaitu komparasi metode *C4.5*, *Naïve Bayes*, dan *Random Forest* yang dilakukan oleh V.H.Valentino (2018), dalam penelitiannya mereka melakukan perbandingan nilai akurasi untuk menentukan kelulusan mata kuliah. Hasil yang didapat ketika menggunakan metode *C4.5* sebesar 98.89%, *Naïve Bayes* sebesar 96.67%, sedangkan nilai akurasi *Random Forest* sebesar 95.56%.

Pada penelitian yang dilakukan, komparasi algoritma klasifikasi *C4.5* dan *Random Forest* digunakan sebagai algoritma yang membentuk rekomendasi penjualan gaun dalam bentuk pohon keputusan. Hal ini didasarkan atas keunggulan algoritma *C4.5* yang mampu memangkas struktur hierarki pohon keputusan terhadap parameter yang dimiliki dalam proses klasifikasi sehingga memudahkan dalam implementasi pengambilan keputusan. *Random Forest* merupakan pengembangan dari *Decision Tree*, dimana setiap *Decision Tree* telah dilakukan training menggunakan sampel individu dan setiap atribut dipecah pada *tree* yang dipilih antara atribut subset yang bersifat acak. Dan pada proses klasifikasi, individunya didasarkan pada vote dari suara terbanyak pada kumpulan populasi *Tree*. Berdasarkan uraian diatas penulis mencoba menguji dan membandingkan dua algoritma yaitu *C4.5* dengan *Random Forest* pada data penjualan gaun *Aliexpress.com*. Dengan ini penulis membuat judul penilitan “ KOMPARASI ALGORITMA C4.5 DENGAN RANDOM FOREST UNTUK REKOMENDASI PENJUALAN GAUN ALIEXPRESS.COM ”.

## 1.2. Rumusan Masalah

Rumusan masalah yang didapat dari latar belakang adalah sebagai berikut:

1. Berapa tingkat Akurasi pada hasil klasifikasi Rekomendasi Penjualan menggunakan Algoritma *C45* dan *Random Forest*?
2. Berapa tingkat Presisi pada hasil klasifikasi Rekomendasi Penjualan menggunakan Algoritma *C45* dan *Random Forest*?

## 1.3. Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah sebagai berikut :

1. Dataset yang digunakan dalam penelitian ini adalah 500 data yang berasal dari penjualan Toko *Online Aliexpress.com*.
2. Data yang digunakan merupakan data penjualan gaun wanita.
3. Dataset yang di gunakan di unduh dari *UCI Machine Learning* :  
<https://archive.ics.uci.edu/ml/machine-learning-databases/00289/>.
4. Dataset yang digunakan pada penelitian ini yaitu tahun 2015.

## 1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang diuraikan diatas, maka tujuan dari penelitian ini adalah :

1. Mengetahui tingkat akurasi Metode *C4.5 & Random Forest*
2. Mengetahui tingkat presisi Metode *C4.5 & Random Forest*

## 1.5. Manfaat Penelitian

Adapun Manfaat yang diharapkan dalam penelitian ini adalah :

1. Bagi Peneliti Pengujian mengetahui hasil implementasi berdasarkan parameter akurasi yang dihasilkan serta membandingkannya dengan algoritma klasifikasi lain yakni Algoritma *C4.5* dan *Random Forest*.
2. Dengan adanya penelitian ini akan membuka peluang diadakannya penelitian perbandingan dari metode-metode yang ada pada Statistika dan *Machine Learning* lainnya.