

# **BAB 1**

## **PENDAHULUAN**

### **1.1 Latar Belakang Masalah**

Deposito merupakan salah satu tabungan berjangka yang model pengambilannya berdasarkan pada kesepakatan dari pihak bank dengan nasabah deposito [1]. Suku bunga pertahunnya yang dapat diperoleh nasabah dari melakukan simpanan deposito cukup tinggi dibandingkan suku bunga yang diperoleh dari tabungan biasa ataupun dari suku bunga giro [2]. Hal ini menjadi daya tarik tersendiri dari masyarakat untuk lebih memilih melakukan simpanan deposito. Menurut Lembaga Penjamin Simpanan pada perss realise tanggal 29 Juli 2016 menunjukkan bahwa data pemilik rekening mengalami peningkatan yaitu pada bulan April 2016 sampai dengan bulan Mei 2016 berjumlah 180.940.127 menjadi 182.436.133 [3].

Bank adalah suatu badan usaha yang menghimpun dana dari masyarakat baik dalam bentuk simpanan, kredit, ataupun bentuk lain guna meningkatkan taraf hidup bagi rakyat banyak [1]. Penjelasan tentang definisi bank tersebut dapat disimpulkan bahwa bank adalah sebagai wadah penghimpun dana, yang harus tetap mempertahankan para nasabahnya agar dapat membantu kelangsungan hidup bagi rakyat banyak. Oleh karena itu, bank juga harus mengelola dengan baik para nasabahnya, khususnya para nasabah deposito yang memiliki loyalitas tinggi dalam simpanan deposito. Bagi bank untuk tetap mengelola dengan baik para nasabah deposito, bank membutuhkan strategi pemasaran yang dapat digunakan untuk mengolah data dalam jumlah besar dari data nasabah deposito. Beberapa cara strategi pemasaran dapat dilakukan, salah satunya strategi pemasaran yang dapat digunakan untuk mengolah data dalam jumlah besar berdasarkan dari data nasabah deposito sebelumnya yaitu dapat dilakukan dengan memprediksi. Prediksi tersebut dapat menggunakan data-data nasabah deposito

yang sudah ada kemudian diproses sehingga akan menemukan informasi yang sangat penting dan bernilai.

Dalam pemodelan Data Mining terbukti dapat mengidentifikasi informasi yang sangat penting dan berguna yang dapat diperoleh dari data besar yang tersimpan dalam gudang basis data [4]. Dewasa ini proses data mining sangatlah diandalkan beberapa pihak baik perusahaan atau instansi lainnya untuk pengelolaan dan pengolahan data dengan jumlah yang besar. Peran utama dari data mining itu sendiri adalah untuk klasifikasi, prediksi, klustering, estimasi, dan asosiasi. Peran utama tersebut sekaligus menjadi manfaat dari kegunaan data mining dalam mengelola dan mengolah jumlah data yang sangat besar. Dari peran utama tersebut masing- masing memiliki algoritma atau cara tersendiri untuk melakukan proses data mining, antara lain untuk model *Clustering* dengan algoritma *K-Means*, *Fuzzy C-Means*, dan lain-lain, untuk model *Assosiation Rules* dengan menggunakan algoritma *FP-Growth*, *A Priori*, dan lain-lain, untuk model *Classification* dengan menggunakan algoritma *K-Nearest Neighbor*, *ID3*, *C4.5*, *Naive Bayes Classifier* dan lainnya [5]. Pada perkembangannya proses prediksi dan klasifikasi data-data mempunyai berbagai tujuan termasuk untuk memudahkan beberapa pihak untuk pengambilan keputusan. Metode untuk memprediksi dan mengklasifikasi salah satunya adalah metode *Naive Bayes Classifier* yaitu metode untuk mengklasifikasikan data yang dapat digunakan sebagai probabilitas suatu kelompok [6].

Sejumlah penelitian telah banyak dilakukan, seperti penelitian oleh Alvino Dwi Rachman Prabowo [7] yaitu menerapkan algoritma *Naive Bayes Classifier* untuk penentu nasabah deposito dengan atribut yang dapat dijadikan sebagai prediktor meliputi faktor-faktor yang berasal dari data nasabah deposito antara lain : *age*, *job*, *marital*, *education*, *default*, *balance*, *haousing*, *loan*, *contact*, *day*, *month*, *duration*, *campaign*, *pdays*, *previous*, *poutcome*, dalam penelitiannya PSO (Particle Swarm Optimization) digunakan sebagai *feature selection* atribut dengan pembobotan atribut untuk meningkatkan akurasi. Hasil akurasi sebelum

menggunakan PSO sebagai *feature selection* atribut yaitu sebesar 82,19%, sedangkan hasil akurasi dengan menggunakan PSO meningkat menjadi 89,70%. Namun, pada penelitian Alvino mempunyai kekurangan yaitu dalam pengolahan datanya tidak dilakukan terlebih dahulu tahap pembersihan data yang tidak lengkap (*missing value*). Padahal sangat penting dalam pengolahan data mining harus melalui beberapa tahapan, seperti pembersihan data yang tidak lengkap, penghapusan atribut yang tidak berpengaruh, penghapusan data yang sama (*duplicat data*), dan tahapan lainnya, karena data yang berkualitas baik akan mempengaruhi hasil akurasi [5]. Sehingga untuk mendapatkan nilai akurasi yang baik pada penerapan suatu algoritma data mining pada pemrosesan datanya tidak hanya berdasarkan dari 1 faktor tahapan pemrosesan data, tetapi harus berdasarkan dari beberapa faktor tahapan pemrosesan data seperti yang disebutkan diatas [8].

Berdasarkan uraian masalah diatas maka dalam penelitian ini akan menggunakan data dari bagian marketing sebuah bank di portugal, data tersebut didonasi pada tanggal 14 Februari 2012 oleh S. Moro, P.Cortez dan P.Rita dengan jumlah 4521 *record*, 16 atribut dan 1 label kelas. Data tersebut adalah data public yang diperoleh dari UCI *Repository Mechine Learning*. Dengan data tersebut akan digunakan untuk menerapkan algoritma *Naive Bayes Classifier* sebagai metode klasifikasi yang diharapkan mampu mendapatkan pola prediksi yang akurat guna menentukan nasabah yang berpotensi untuk simpanan deposito. Hasil penelitian yang sistematis diharapkan mampu memudahkan pihak bank untuk memprediksi nasabah yang berpotensi melakukan simpanan deposito lebih dini. Hal ini juga sangat diperlukan sehingga dapat membantu bagian marketing untuk salah satu cara strategi pemasaran.

## 1.2 Rumusan Masalah

Dalam penelitian ini rumusan masalah yang akan diuraikan, diantaranya yaitu:

1. Bagaimana memprediksi nasabah yang berpotensi untuk melakukan deposito secara lebih akurat dengan menggunakan metode *Naive Bayes Classifier*?
2. Bagaimana membuat model klasifikasi untuk mengetahui profile nasabah yang berpotensi melakukan simpanan deposito?
3. Bagaimana menerapkan metode *Naive Bayes Classifier* untuk mengetahui nasabah yang berpotensi untuk simpanan deposito?

### 1.3 Batasan Masalah

Agar menyusun Tugas Akhir ini tidak keluar dari pokok batasan yang dirumuskan, maka lingkup batasan masalahnya, sebagai berikut:

1. Data yang digunakan dalam penelitian ini adalah data dari bagian marketing sebuah bank di Portugal, data sekunder tersebut didonasi pada tanggal 14 Februari 2012 oleh S. Moro, P.Cortez dan P.Rita dengan jumlah 4521 *record*, 16 atribut dan 1 label kelas, data ini merupakan data public yang diperoleh dari *UCI Repository Machine Learning*.
2. Penelitian ini menggunakan algoritma *Naive Bayes Classifier* untuk memprediksi nasabah yang berpotensi untuk simpanan deposito pada bank.
3. Tools yang digunakan dalam penelitian ini untuk menemukan pola yaitu dengan menggunakan *Rapidminer* dan tools untuk mengembangkan *prototype* adalah dengan menggunakan *web*.

### 1.4 Tujuan Penelitian

Tujuan dari usulan penelitian ini berdasarkan latarbelakang dan permasalahan yang telah dibahas, yaitu sebagai berikut:

1. Memprediksi nasabah yang berpotensi melakukan deposito secara lebih akurat dengan menggunakan metode *Naive Bayes Classifier*.
2. Membuat model klasifikasi untuk mengetahui profile nasabah yang berpotensi melakukan simpanan deposito.
3. Menerapkan metode *Naive Bayes Classifier* untuk mengetahui nasabah yang berpotensi untuk simpanan deposito.

### **1.5 Manfaat Penelitian**

Kegunaan dalam penulisan dari penelitian ini, yaitu :

1. Dapat membantu pihak bank yaitu pada bagian marketing untuk menentukan nasabah yang memiliki potensi untuk melakukan simpanan deposito.
2. Diharapkan dengan prediksi nasabah yang berpotensi melakukan simpanan deposito ini dapat menjadi acuan bagi bagian marketing dan mendukung pengambilan keputusan penentuan nasabah yang berpotensi untuk simpanan deposito.