

## **BAB II**

### **LANDASAN TEORI**

#### **2.1 Kredit**

Menurut Peraturan Otoritas Jasa Keuangan (PJOK) Nomor 32/PJOK.03/2018, kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam untuk melunasi utangnya setelah jangka waktu tertentu dengan pemberian bunga.

Menurut Muhamad Djumhana dalam bukunya Hukum Perbankan di Indonesia, terdapat unsur-unsur dalam kredit yaitu:

a) Kepercayaan

Kepercayaan adalah keyakinan dari pemberi kredit bahwa prestasi yang diberikannya dalam bentuk uang, barang, atau jasa, akan benar-benar diterimanya kembali dalam jangka waktu yang telah ditentukan.

b) Tenggang waktu

Tenggang waktu adalah suatu masa yang memisahkan antara pemberian prestasi dengan kontraprestasi yang diterima pada suatu masa yang akan datang. Unsur tenggang waktu terkandung pengertian agio dari uang, dimana uang yang ada sekarang nilainya lebih tinggi dari uang yang akan diterima di masa depan.

c) *Degree of risk*

*Degree of risk* merupakan tingkat risiko yang akan dihadapi. Semakin lama jangka waktu kredit diberikan maka semakin tinggi risikonya. Adanya unsur risiko ini membuat timbulnya jaminan pemberian kredit.

d) Prestasi

Prestasi yaitu berupa objek kredit tidak hanya diberikan dalam bentuk uang, tetapi juga dapat berbentuk barang atau jasa. Namun, karena kehidupan ekonomi dimasa sekarang didasarkan pada uang, maka transaksi-transaksi kredit yang menyangkut uang yang sering dijumpai dalam praktik perbankan.

## 2.2 Kredit Bermasalah

Kredit bermasalah atau *non performing loan* merupakan perbandingan antara kredit bermasalah dengan total kredit (Ismail, 2018). Ikatan Bankir Indonesia (2016) menjelaskan kredit bermasalah adalah kredit yang debiturnya menghadapi kesulitan untuk melunasi hutangnya kepada bank berupa pelunasan pokok, pembayaran bunga dan lain-lain. Karena masalah kredit ini nilai kerugian yang akan dialami mungkin sangat besar, oleh karena itu bank harus memiliki manajemen yang baik untuk mengelola risiko yang ditimbulkan. Adanya pemberian kredit tidak lepas dari kredit bermasalah (*non performing loans*). Kredit bermasalah selalu ada dalam aktivitas penyaluran kredit bank karena tidak adanya undang-undang untuk menghindari risiko kredit.

Berdasarkan Surat Edaran Bank Indonesia No. 15/28/DPNP tanggal 31 Juli 2013 tentang penilaian aktiva bank umum, tingkat kualitas kredit diklasifikasikan menjadi 5, yaitu kategori lancar, dalam perhatian khusus, kurang lancar, diragukan dan macet.

## 2.3 Data Mining

*Data mining* didefinisikan sebagai studi mengumpulkan, membersihkan, memproses, menganalisis, dan mendapatkan wawasan yang berguna dari data. Menurut Hendrian (2018), *data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan juga pengetahuan yang terkait dari berbagai *database* besar.

Istilah *data mining* seringkali digunakan bergantian dengan istilah *Knowledge Discovery in Database* (KDD). Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi saling berkaitan. *Knowledge Discovery in Database* (KDD) merupakan suatu metode untuk memperoleh pengetahuan dari *database* yang ada, yang di dalam *database* tersebut terdapat tabel-tabel yang saling berhubungan atau berelasi. Hasil pengetahuan yang diperoleh dalam proses KDD dapat digunakan sebagai basis pengetahuan (*knowledge base*) untuk

keperluan pengambilan keputusan. *Data mining* termasuk ke dalam salah satu tahapan dari keseluruhan proses KDD (Mardi, 2017).

## 2.4 Klasifikasi

Klasifikasi adalah suatu cara pengelompokan benda berdasarkan ciri-ciri yang dimiliki oleh objek klasifikasi. Sebuah pengklasifikasian dibuat dari sekumpulan data latih dengan kelas yang telah ditentukan sebelumnya. Klasifikasi merupakan salah satu proses di dalam *data mining*.

Klasifikasi merupakan pengelompokan fitur ke dalam kelas yang sesuai. Vektor fitur pelatihan tersedia dan telah diketahui kelas-kelasnya, kemudian vektor fitur pelatihan ini dimanfaatkan untuk merancang pemilah (Wibawa, Purnama, Akbar dan Dwiyanto, 2018). Klasifikasi masuk ke dalam metode *supervised learning*, metode yang digunakan untuk menemukan hubungan antara atribut masukan dan atribut target (Hendrian, 2018).

## 2.5 Imbalance Class

*Imbalance class* merupakan keadaan dimana rasio kelas pada setiap dataset tidak seimbang atau saat jumlah antara kelas mayoritas dan minoritas sangat jauh berbeda (Astuti dan Lenti, 2021). Apabila data tidak tersebar merata di setiap kelas maka kemungkinan terjadinya kesalahan klasifikasi cukup tinggi di kelas minoritas. *Imbalance class* pada proses klasifikasi dapat menyebabkan *classifier* belajar sedikit dari kelas minoritas sehingga prediksi akan lebih mengarah ke kelas mayoritas (Hamami dan Dahlan, 2022).

Penanganan *imbalance class* dapat menggunakan metode *resampling*. Metode *resampling* bekerja untuk menyeimbangkan data sehingga proporsi data yang dihasilkan akan sama besar (*balance*). Terdapat dua algoritma *resampling* yaitu, sebagai berikut (Saifudin dan Wahono, 2015):

### 2.5.1 Random Undersampling (RUS)

Teknik *undersampling* bekerja dengan menyeimbangkan dataset dengan mengurangi ukuran kelas mayoritas. Teknik ini digunakan ketika jumlah dataset

mencukupi. *Random Undersampling* secara acak menghapus data kelas mayoritas sehingga jumlah data antara kelas mayoritas dan minoritas sama.

### 2.5.2 *Random Oversampling* (ROS)

Teknik *random oversampling* digunakan jika jumlah dataset tidak mencukupi. Teknik ini menyeimbangkan dataset dengan memilih data kelas minoritas secara acak dan menambahkan ke dalam data latih. Proses pemilihan dan penambahan dilakukan berulang-ulang sampai jumlah data kelas minoritas sama dengan jumlah kelas mayoritas. Pertama dihitung selisih antara data kelas mayoritas dan kelas minoritas. Setelah itu dilakukan proses membaca data kelas minoritas secara acak dan ditambahkan ke dalam data *training*, sembari melakukan perulangan sebanyak hasil selisih antara kelas mayor dan minor (Chawla, Bowyer, Hall and Kegelmeyer, 2002).

## 2.6 *Confusion Matrix*

*Confusion matrix* adalah tabel yang menggambarkan performa dari suatu model atau algoritma secara spesifik. Setiap baris dari *matrix* tersebut mempresentasikan kelas aktual dari data dan pada tiap kolom mempresentasikan kelas prediksi dari data (atau sebaliknya) (Saputro dan Sari, 2019). *Confusion matrix* memberikan penilaian kinerja model klasifikasi berdasarkan objek yang diprediksi dengan benar dan salah. *Confusion matrix* dapat digunakan untuk mengukur performa pada permasalahan klasifikasi biner maupun *multiclass*, *confusion matrix* pada klasifikasi biner ditunjukkan pada tabel berikut:

**Tabel 2.1 *Confusion Matrix***

	<i>Predicted Negative</i>	<i>Predicted Positive</i>
<i>Actual Negative</i>	<i>True Negative</i> (TN)	<i>False Positive</i> (FP)
<i>Actual Positive</i>	<i>False Negative</i> (FN)	<i>True Positive</i> (TP)

- a. *True Negative* : seberapa banyak data yang aktual kelasnya negatif dan model memprediksi negatif.
- b. *True Positive* : seberapa banyak data yang aktual kelasnya positif dan model memprediksi positif.

- c. *False Negative* : seberapa banyak data yang aktual kelasnya positif, tetapi model memprediksi negatif.
- d. *False Positive* : seberapa banyak data yang aktual kelasnya negatif, tetapi model memprediksi positif.

Berdasarkan *confusion matrix* dilakukan perhitungan untuk memperoleh nilai akurasi, sensitivitas, spesifisitas dan *precision* yang berguna untuk mengukur performa suatu model, sebagai berikut:

- a. Akurasi merupakan total keseluruhan seberapa sering model mengklasifikasi benar. Persamaan akurasi ditulis pada Persamaan (2.1) berikut:

$$\frac{TP+TN}{TP+TN+FN+FP} \times 100\% \quad (2.1)$$

- b. Sensitivitas merupakan kemampuan model untuk memperhitungkan proporsi *true positive* dalam memprediksi nilai yang sebenarnya *positive*. Persamaan sensitivitas ditulis pada Persamaan (2.2) berikut:

$$\frac{TP}{TP+FN} \times 100\% \quad (2.2)$$

- c. Spesifisitas merupakan kemampuan model untuk memperhitungkan proporsi *true negative* dalam memprediksi nilai yang sebenarnya *negative*. Persamaan spesifisitas ditulis pada Persamaan (2.3) berikut:

$$\frac{TN}{TN+FP} \times 100\% \quad (2.3)$$

- d. *Precision* merupakan keadaan ketika model memprediksi *positive* dan seberapa sering prediksi itu benar. Persamaan *precision* ditulis pada Persamaan (2.4) berikut:

$$\frac{TP}{TP+FP} \times 100\% \quad (2.4)$$