

BAB IV METODOLOGI PENELITIAN

4.3. Populasi dan Sampel Penelitian

Populasi yang digunakan dalam penelitian ini adalah data jejak akademik mahasiswa aktif Statistika Universitas Islam Indonesia.

Sedangkan data sampel diambil dari data mahasiswa aktif Statistika angkatan 2015 yang berjumlah 125 mahasiswa. Data diambil berdasarkan rekam mesin *fingerprint* berdasarkan setiap mata kuliah yang diambil mahasiswa aktif Statistika angkatan 2015 yang berada didalam *website* SIPRIMA. Setiap mahasiswa rata – rata mengambil 8 – 10 mata kuliah dalam satu semester.

4.4. Tempat dan Waktu Penelitian

Penelitian ini dilakukan dengan pengambilan data mahasiswa aktif Statistika angkatan 2015 berdasarkan presensi mahasiswa secara langsung di *website* SIPRIMA FMIPA dari tahun ajaran Genap 2015/ 2016 sampai Genap 2017/ 2018. Pengambilan data dilakukan dari tanggal 21 Januari 2018 – 28 Januari 2018.

4.5. Variabel Penelitian

Pada penelitian ini peneliti menggunakan beberapa variabel dengan beberapa dikategorikan dengan angka yaitu:

Tabel 4.1 Variabel Pada Data

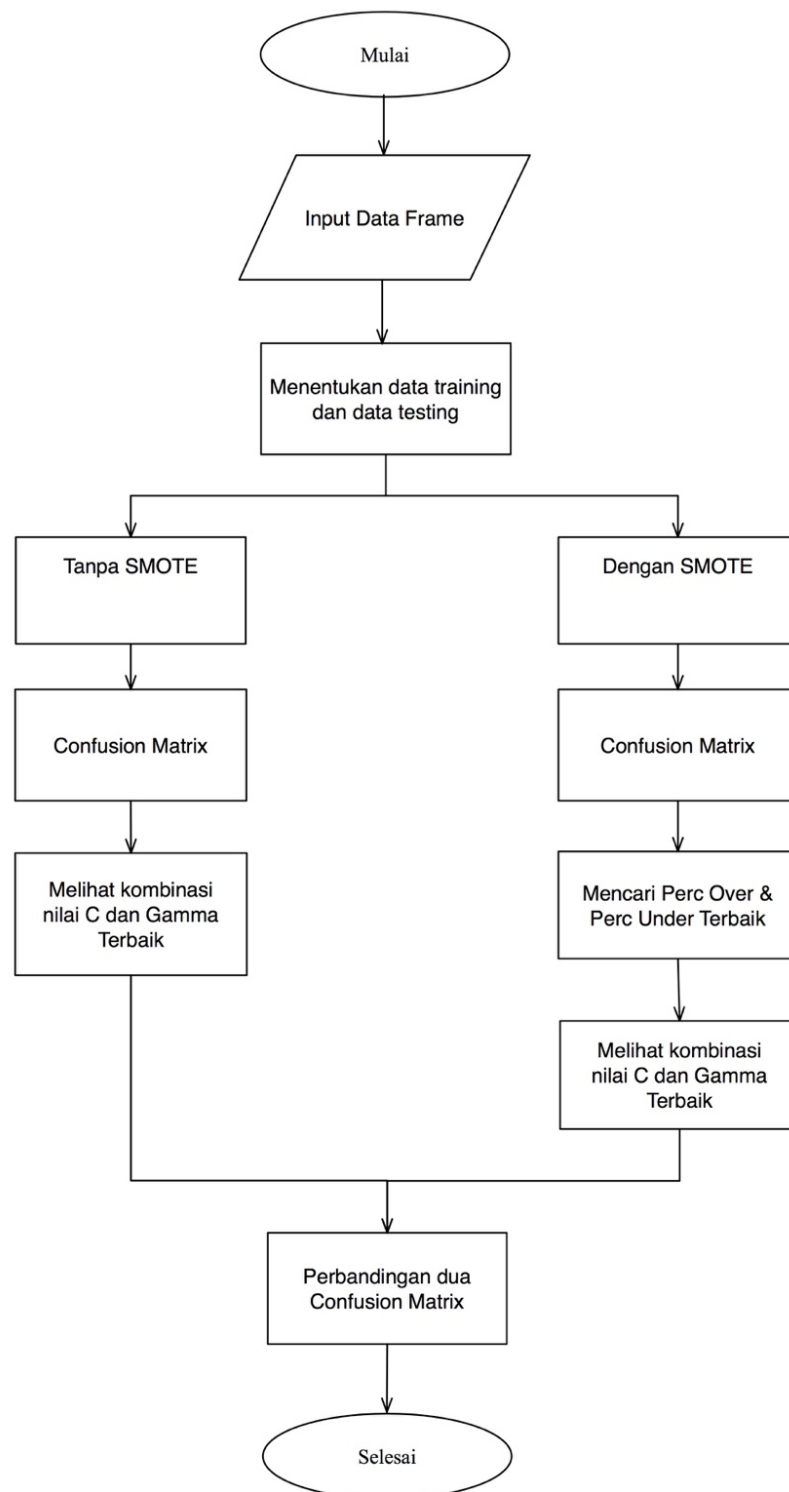
Variabel	Keterangan Variabel
Jenis Mata kuliah	<p>Berupa variabel yang berisi jenis mata kuliah yang diambil yaitu</p> <ul style="list-style-type: none">• Mata Kuliah Wajib Universitas, dikategorikan sebagai satu• Mata Kuliah Inti, dikategorikan sebagai dua• Mata Kuliah Pilihan Wajib, dikategorikan sebagai tiga

Variabel	Keterangan Variabel
	<ul style="list-style-type: none"> • Mata Kuliah Pilihan Bebas, dikategorikan sebagai empat
Jam Mata kuliah	<p>Berupa variabel yang berisi jam mata kuliah yang diambil dan pengkategorianya yaitu:</p> <ul style="list-style-type: none"> • pertama untuk jam mata kuliah yang dimulai dari jam 07.00 WIB, dikategorikan sebagai satu • kedua untuk jam mata kuliah yang dimulai dari jam 08.45 WIB, dikategorikan sebagai dua • ketiga untuk jam mata kuliah yang dimulai dari jam 13.00 WIB, dikategorikan sebagai tiga • keempat untuk jam mata kuliah yang dimulai dari jam 14.45 WIB, dikategorikan sebagai empat
Hari	<p>Berupa variabel yang berisi hari mata kuliah yang diambil dengan pengkategorian yaitu:</p> <ul style="list-style-type: none"> • senin, dikategorikan sebagai satu • selasa, dikategorikan sebagai dua • rabu, dikategorikan sebagai tiga • Kamis, dikategorikan sebagai empat • jumat, dikategorikan sebagai lima • sabtu, dikategorikan sebagai enam
SKS	<p>Berupa variabel yang berisi jumlah SKS mata kuliah yang diambil.</p> <ul style="list-style-type: none"> • 1 SKS, dikategorikan sebagai satu • 2 SKS, dikategorikan sebagai dua • 3 SKS, dikategorikan sebagai tiga

Variabel	Keterangan Variabel
Presentase Keterlambatan	Berupa variabel yang berisi presentase jumlah keterlambatan suatu mata kuliah dalam satu semester.
Predikat	<p>Berupa variabel yang berisi tingkatan klaster yang dimiliki oleh mahasiswa yaitu:</p> <ul style="list-style-type: none"> • Amat baik, dikategorikan sebagai satu • Baik, dikategorikan sebagai dua • Cukup, dikategorikan sebagai tiga • Kurang, dikategorikan sebagai empat • Sangat kurang, dikategorikan sebagai lima • Tidak memenuhi, dikategorikan sebagai enam

4.6. Alat dan Organisir Data

Proses analisis dalam penelitian ini menggunakan *software Excel* yang digunakan untuk menyiapkan *dataset* penelitian dan *R Studio* untuk memproses klasifikasi menggunakan metode SVM. Gambar 4.1 berikut adalah *flowchart* penelitiannya:



Gambar 4.1 *Flowcart* Penelitian

Keterangan:

1. *Input data frame* : yaitu memasukkan data kedalam *software R*, dan variabel disesuaikan pada tabel 4.1.

2. Menentukan data training dan testing : menentukan sampel yang diambil dari dataset dengan menggunakan sampel 66% dengan teknik *random sampling*, kemudian menggunakan data sampel ini sebagai data training. Sedangkan sisa dataset yaitu 33% akan digunakan sebagai data testing.
3. Dengan SMOTE :
 - a. klasifikasi SVM dengan menambahkan metode SMOTE yaitu dengan melakukan *over* dan *under sampling* pada data *training*. Pada metode ini untuk menemukan nilai *over* dan *under sampling* maka dilakukanlah beberapa percobaan dan dilihat *confusion matrix*-nya.
 - b. *Confusion Matrix* : *matrix* untuk melihat kinerja dari klasifikasi, berupa akurasi dan error.
 - c. Melihat kombinasi nilai C dan Gamma terbaik : selain melihat nilai *over* dan *under sampling*, maka perlu juga dilakukan percobaan untuk melihat nilai C dan Gamma terbaik dengan melihat *confusion matrix*-nya dengan nilai *over* dan *under sampling* terbaik yang sudah didapat sebelumnya.
4. Tanpa SMOTE :
 - a. Klasifikasi SVM tanpa menggunakan SMOTE.
 - b. *Confusion Matrix* : *matrix* untuk melihat kinerja dari klasifikasi, berupa akurasi dan error.
 - c. Melihat kombinasi nilai C dan Gamma terbaik : dilakukan beberapa percobaan untuk melihat nilai C dan Gamma terbaik dengan melihat *confusion matrix*-nya.
5. Perbandingan kedua *Confusion Matrix* : membandingkan *Confusion Matrix* hasil klasifikasi SVM dengan SMOTE dan tanpa SMOTE dengan melihat nilai akurasi, *error*, *recall* dan *precision*, untuk melihat seberapa bagus antara kedua metode SVM tanpa SMOTE atau dengan SMOTE.