

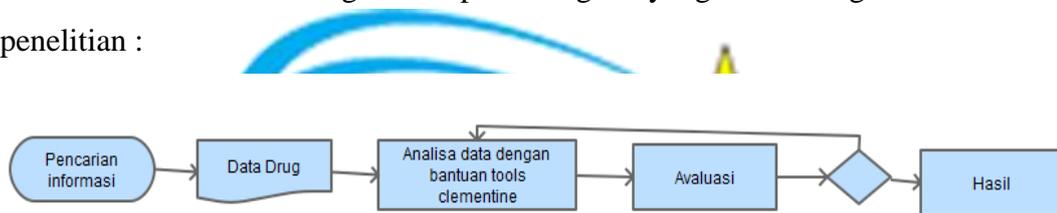
BAB III

METODELOGI PENELITIAN

Bab III menjelaskan tentang proses pengelompokan pada *Drug* data yang akan dianalisis oleh peneliti.

3.1 Tahap Penelitian

Membahas tentang alur perancangan yang akan digunakan dalam penelitian :

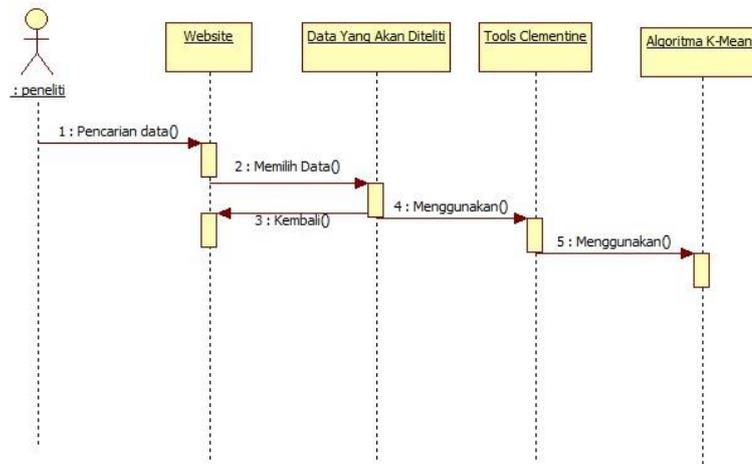


Gambar 3.1 Tahapan penelitian digambarkan dengan *flowchart*

1. Pencarian informasi yang berhubungan dengan penelitian.
2. Tahap pencarian data, pada tahap ini akan dilakukan pencarian data melalui *website Universal Centre and Intellegent system (UCI) Machine Learning Repository Education* merupakan sebuah data yang berjumlah sangat besar dan sudah diolah dengan menggunakan teknik klasifikasi.
3. Pengumpulan data, menganalisa data yang sudah didapatkan oleh peneliti dan mengimplementasikan dengan bantuan *tools Clementine*.

3.2 Gambaran Penelitian

Pada model analisis berisi *squens diagram* yang terdiri dari *preprocessing data*, dan *K-Means Clustering* sebagai berikut:



Gambar 3.2 Tahapan penelitian digambarkan dengan *sequence diagram*.

Pada gambar 3.2 Tahapan Pencarian Data mulai dengan tahapan pencarian data di website *Universal Centre And Intelligent system (UCI) Machine Learning Repository Education*, di website tersebut terdapat macam-macam data yang telah diteliti yang menggunakan berbagai teknik analisis data *mining*. Peneliti memilih data *drug* yang akan dianalisis menggunakan teknik *clustering* dengan *tools clementine*. Data *drug* merupakan data *pasien* yang mengalami depresi dan menyalahgunakan penggunaan obat keras. Berikut *sample data drug* yang berbentuk file dengan format *.xls*:

iddrug	drug Name	condition	review	rating	date	useful Count
165740	Mirtazapine	Depression	"I've tried a few antidepressants over the years (citalopram, fluoxetine, amitriptyline), but none of those helped with my depression, insomnia & anxiety. My doctor suggested and changed me onto 45mg mirtazapine and this medicine has saved my life. Thankfully I have had no side effects especially the most common - weight gain, I've actually lost a lot of weight. I still have suicidal thoughts but mirtazapine has saved me."	10.00	February 28, 2012	22
205473	Mesazamine	Croh's Disease, Maintenance	"My son has Crohn's disease and has done very well on the Asacol. He has no complaints and shows no side effects. He has taken as many as nine tablets per day at one time. I've been very happy with the results, reducing his bouts of diarrhea drastically. "Quick reduction of symptoms"	08.00	May 17, 2009	17
159672	Bactrim	Urinary Tract Infection	"Contrave combines drugs that were used for alcohol, smoking, and opioid cessation. People lose weight on it because it also helps control over-eating. I have no doubt that most obesity is caused from sugar/carb addiction, which is just as powerful as any drug. I have been taking it for five days, and the good news is, it seems to go to work immediately. I feel hungry before I want food now. I really don't care to eat, I don't just fill my stomach. Since I have only been on it a few days, I don't know if I've lost weight (I don't have a scale), but my clothes do feel a little looser, so maybe a pound or two. I'm hoping that after a few months on this medication, I will develop healthier habits that I can continue without the aid of contrave."	09.00	September 29, 2017	3
39293	Contrave	Weight Loss	"I have been on this birth control for one cycle. After reading some of the reviews on this type and similar birth controls I was a bit apprehensive to start. Im giving this birth control a 10 out of 10 as I have not been on it long enough for a 10. So far I love this birth control! My side effects have been so minimal its like Im not even on birth control! I have experienced mild headaches here and there and some nausea but other than that Ive been feeling great! I got my period on cue on the third day of the inactive pills and I had no idea it was coming because I had zero pms! My period was very light and I	09.00	March 5, 2017	35

Gambar 3.3 File *sample data drug* format excel.

Pada gambar 3.3 File *sampel* data *drug* format excel, data *drug* tersebut telah dianalisis dan diolah, peneliti akan mengolah kembali data tersebut menggunakan teknik *clustering algoritma K-Means* dengan bantuan *tools Spss Clementine*.

3.3 Analisa Data

Secara umum, system yang akan dibangun dalam penelitian ini adalah sebuah system dengan fungsi utama untuk melakukan pengelompokan data *Drug*. Peneliti menggunakan data yang diperoleh dari *website Universal Centre And Intellegent system (UCI) Machine Learning Repository Education* sebanyak 53767. Sehingga menghasilkan akan menganalisis perbandingan hasil Klasifikasi dengan *Clustering*.

3.4 Model Analisis

3.3.1 Tahap Pemahaman Bisnis (*Business Understanding Phase*)

Menyiapkan strategi awal dalam pencarian data di internet untuk mencapai tujuan penelitian. Data yang akan dianalisis diambil dari *website website Universal Centre And Intellegent system (UCI) Machine Learning Repository Education.s*

3.3.2 Pemahaman Data

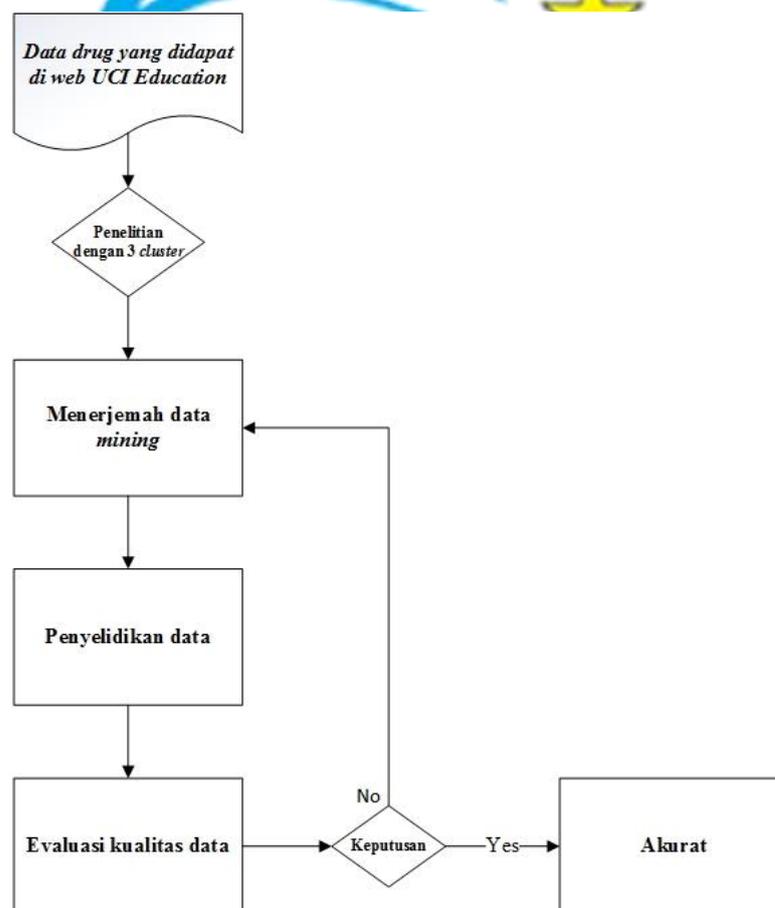
1. Pengumpulan data. Data *drug* adalah data yang diambil langsung dari *website Universal Centre And Intellegent system (UCI) Machine Learning Repository Education* sebanyak 53767 data yang sudah diolah.
2. Menentukan tujuan analisis secara detail dalam unit kesehatan secara keseluruhan. Bertujuan menganalisis data *drug* yang dikonsumsi oleh masyarakat dengan hasil perubahan mental setelah menggunakannya, akan dilakukan 3 *cluster* penggunaan berdasarkan variabel *cluster*. Data diperoleh dari dari *website Universal Centre And Intellegent system (UCI) Machine Learning Repository Education* :

<http://archive.ics.uci.edu/ml/datasets/Drug+Review+Dataset+%28Drugs.com%29#>. Jumlah *record* dari data tersebut adalah 53767 data dengan 7

variabel. Pada tahap ini dilakukan pemahaman terhadap kebutuhan penelitian secara *detail* dan menerjemahkannya kedalam tujuan data *mining*.

3. Penyelidikan data *drug* untuk mengenali lebih lanjut tujuan dataset tersebut dan memudahkan dalam pencarian di pengetahuan awal. *Dataset* tersebut terdiri dari : *id drug, drugname, review, rating, date, useful cont.*
4. Mengevaluasi kualitas data dengan hasil pengelohan data masih ditemukan banyak nilai yang kosong yang disebut dengan *missing value*. Menangani data yang tidak lengkap dengan melakukan *processing* dan *cleansing*.

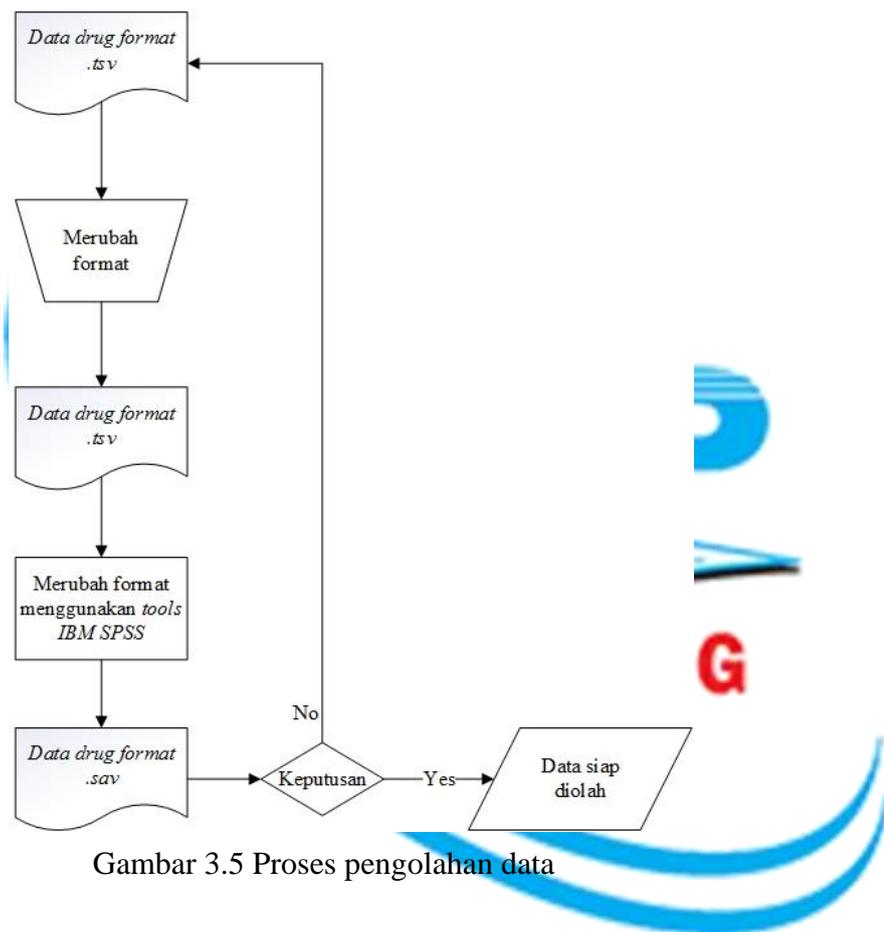
Tahapan pemahaman data dapat dilihat di *flowchart* Gambar 3.4



Gambar 3.4 Proses pemahaman data.

3.3.3 Pengolahan Data

1. Jenis data yang diambil adalah dari dokumen *website Universal Centre And Intellegent system (UCI) Machine Learning Repository Education* yang berupa format *.tsv* yang kemudian diubah kedalam bentuk dokumen excel kemudian diubah kembali kedalam bentuk dokumen yang berekstensi dengan format *file .sav*. Pengolahan data dapat dilihat di gambar 3.5



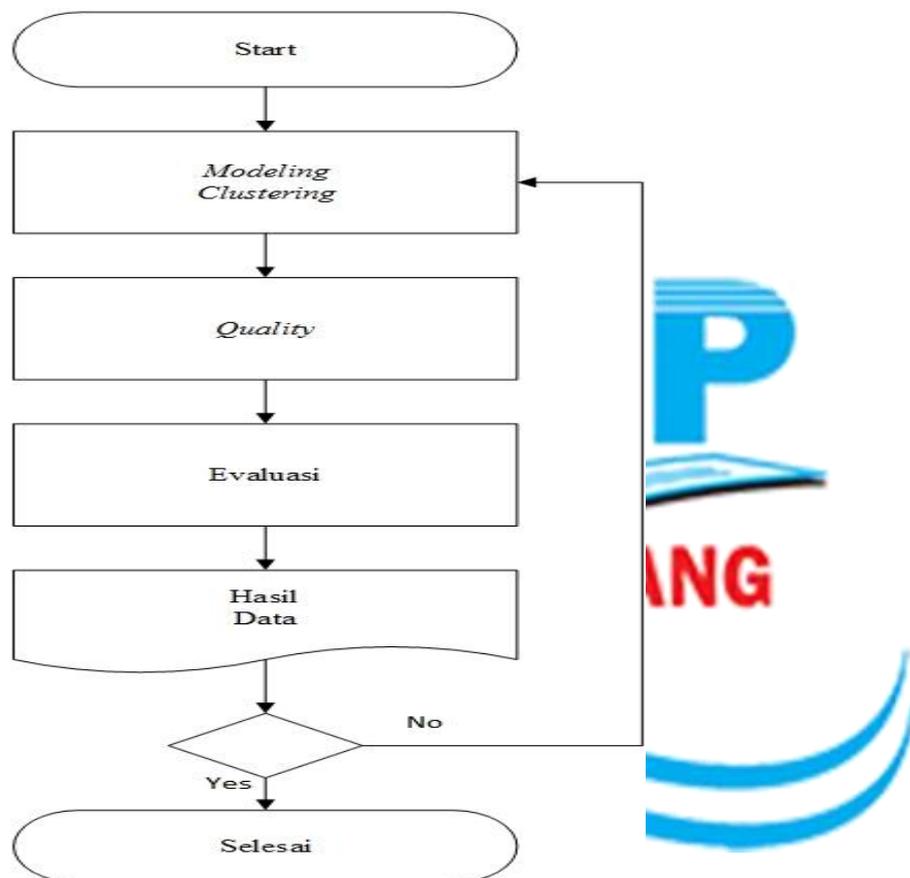
Gambar 3.5 Proses pengolahan data

Dapat dipahami proses pengolahan data, data awal *drug* dengan format *.tsv* dirubah keformat *excel* dengan proses manual. Data yang telah dirubah ke format *excel .xls* dirubah kembali ke bentuk format *.sav* dengan menggunakan *tools IBM SPSS*, Data siap diolah. Jika data dirubah ke format *.sav* dengan menggunakan *tools IBM SPSS* gagal akan kembali ketahap data awal.

2. Menentukan *variabel* dalam kasus yang akan diteliti sesuai dengan analisis yang akan dilakukan.

3.3.4 Tahap Pemodelan Data

Secara umum tahap pemodelan ini adalah tahap yang dilakukan secara langsung dengan melibatkan *datamining* dan menggunakan algoritma yang akan dilakukan. Dalam tahapan ini, peneliti menggunakan teknik *Clustering* dengan algoritma *K-Means*.



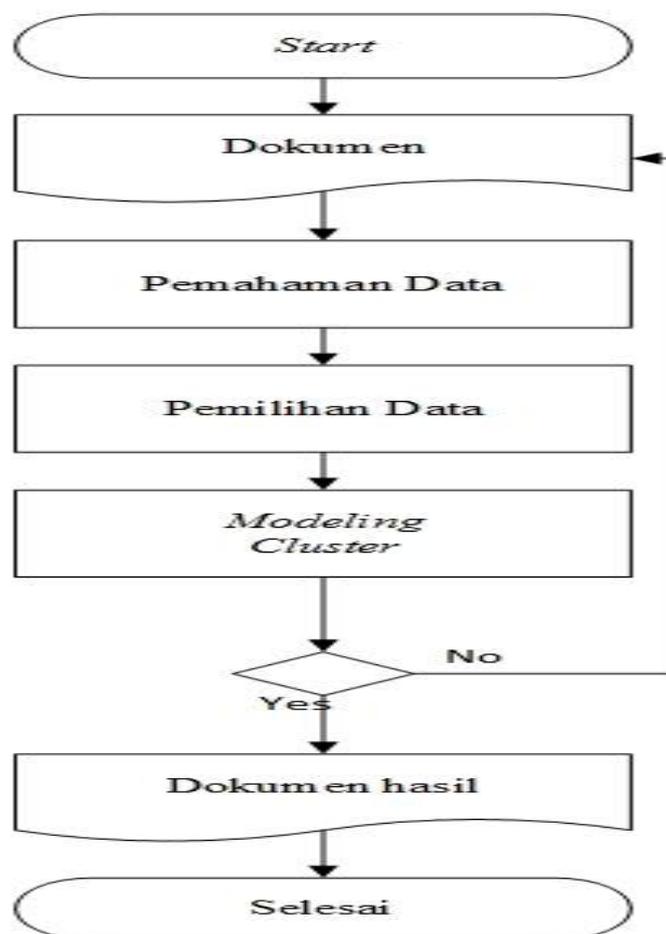
Gambar 3.6 Proses pemodelan data.

Dapat dipahami pada Gambar 3.6 proses pemodelan data, data yang siap untuk analisis akan menggunakan tahap pemodelan *clustering* dengan hasil memberikan informasi tentang kualitas data setelah data selesai dalam tahap *clustering*.

3.3.5 Tahap Evaluasi

Secara umum fasa tahap evaluasi ini, menentukan cara bagaimana mengevaluasi data yang telah di analisis sebelum masuk kedalam tahap fase penyebaran. Untuk tahapan evaluasi data tersebut sebagai berikut:

1. Menentukan model yang telah memenuhi tujuan pada fasa awal penelitian.
2. Menentukan permasalahan penting pada data penelitian yang sedang analisis



Gambar 3.7 Proses Evaluasi

Dapat dipahami pada Gambar 3.7 proses tahap evaluasi, data yang siap diolah dapat dipahami ketahap sebelumnya tentang pemahaman data keperluan untuk pemilihan data *variabel* yang akan diolah. Tahap berikut tentang *modeling*

clustering. Data yang telah siap diolah dengan *modeling cluster*. Jika berhasil dalam *metode clustering* akan mendapatkan berupa informasi. Tidak berhasil akan kembali ketahap persiapan dokumen.

3.5 Jenis Penelitian

Penelitian yang digunakan oleh penelitian berjenis penelitian kuantitatif. Penelitian ini berdasarkan hasil observasi pengamatan secara langsung terhadap data *drug* yang diambil dari *website Universal Centre and Intellegent system (UCI) Machine Learning Repository Education*.

