PEMANFAATAN DATA MINING CLUSTERING DALAM PENENTUAN MEDIA PROMOSI UMKM DI KOTA BATAM

SKRIPSI

OLEH
FITRINI
171510001

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2021

PEMANFAATAN DATA MINING CLUSTERING DALAM PENENTUAN MEDIA PROMOSI UMKM DI KOTA BATAM

SKRIPSI

Untuk memenuhi salah satu syarat
memperoleh gelar Sarjana

OLEH
FITRINI
171510001

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2021

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : FITRINI
NPM : 171510001
Fakultas : Teknik dan Komputer
Program Studi : Sistem Informasi

Menyatakan bahwa "**Skripsi**" yang saya buat dengan judul.:

PEMANFAATAN DATA MINING CLUSTERING DALAM PENENTUAN MEDIA PROMOSI UMKM DI KOTA BATAM

Adalah hasil karya sendiri dan bukan "duplikasi" dari karya orang lain. Sepengetahuan saya, didalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pemah ditulis atau diterbitkan oleh orang lain , kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila temyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur - unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun.

Batam, 25 Januari 2021

FITRINI
171510001

PEMANFAATAN DATA MINING CLUSTERING DALAM PENENTUAN MEDIA PROMOSI UMKM DI KOTA BATAM

SKRIPSI

Untuk memenuhi salah satu syarat

memperoleh gelar Sarjana

Oleh

FITRINI

171510001

Telah disetujui oleh Pembimbing pada tanggal

seperti tertera di bawah ini

Batam, 25 Januari 2021

Erlin Elisa, S.Kom., M.Kom.

Pembimbing

ABSTRAK

Pada perkembangan teknologi yang semakin besar ini, sangat menarik perhatian masyarakat untuk semakin mengetahui dan menggunakan teknologi tersebut. Teknologi yang sering digunakan saat ini adalah *gadget.* Kemajuan *gadget* sangat menarik perhatian masyarakat untuk mengetahui lebih dalam akan kecanggihannya. Selain dapat digunakan dalam berkomunikasi jarak jauh, bermain *games*, dan aktivitas lainnya. *Gadget* saat ini juga dapat digunakan untuk mengakses media sosial. Sejak *gadget* sudah dapat terhubung dengan jaringan internet, semakin banyak fitur yang menarik perhatian masyarakat diberbagai kalangan, salah satunya adalah media sosial. Kemunculan berbagai macam media sosial membuat *gadget* semakin dikenal oleh masyarakat di seluruh dunia. Dengan adanya media sosial, banyak kegunaan yang dapat digunakan oleh masyarakat seperti membagikan momen spesial di media sosial, membuat status dimedia sosial, bermain *games online*, dan bahkan dapat digunakan dalam berbisnis. Berbagai macam media sosial seperti *Facebook, Whatsapp, Instagram*, dan lainnya dapat digunakan masyarakat untuk berbisnis. Pada saat ini, banyak UMKM yang memulai mempromosikan bisnis mereka di media sosial. Akan tetapi, dengan banyaknya macam media sosial sehingga UMKM kesulitan dalam menentukan media sosial mana yang paling sering diakses oleh masyarakat. Sehingga, pada tugas akhir ini penulis tertarik untuk melakukan penelitian terhadap media promosi yang paling diminati dengan menggunakan salah satu metode dalam teknik *datamining* yaitu metode algoritma *clustering K-Means.* Pada algoritma *clustering K-Means* dapat membantu penulis dalam pengelompokkan media promosi yang paling diminati oleh masyarakat yang ada di Kota Batam khususnya yang berada pada kecamatan Batam Kota berdasarkan data responden masyarakat pada penyebaran kuesioner yang telah dilakukan oleh penulis. Pada hasil akhir akan terdapat 3 kategori yaitu media promosi yang paling diminati, kurang diminati, dan paling tidak diminati. Serta untuk mendapatkan hasil tersebut akan dilakukan perhitungan manual berdasarkan metode *clustering K-means* dan pembuktian menggunakan *software Rapidminer 5.3*.

Kata Kunci: *datamining, clustering, K-Means, Rapidminer 5.3*, Media sosial, UMKM

ABSTRACT

On this increasingly large technological development, it attracts people to know more and use that technology. The technology that is often used these days is a gadget. Technology advances attract people to know more about its sophistication. Besides being able to be used for long-distance communication, games and other activities. These days, the gadget can be used to access social media too. Since gadget can be connected with the internet, there comes more features that attract people from different circles, for example, is social media. The emergences of various social media have made gadgets known to people around the world. With social media, many functions can be used by people like sharing their special moment on social media, create their status on social media, playing online games, and even used for business. Some social media application like Facebook, Whatsapp, Instagram and others can be used for business. These days, there is a lot of UMKM promote their business on social media. However, with so many kinds of social media, UMKM have difficulty in determining which social media is the most frequently accessed by the public. So, in this final project, the writer is interested in researching the most demanded promotional media by using one of the methods from data mining technique specifically algorithm clustering K-means method. Algorithm clustering K-means can help the writer to group the promotional media that is most demanded by people in Batam City especially those in Batam Kota district based on respondent data from distributing questionnaires that have been done by the writer. There will be 3 categories in the final result that are the most demanded, less demanded and least demanded. And to get these result, manual calculations will be done by clustering K-means method and verification using Rapidminer 5.3 software.

Keywords: *datamining, clustering, K-Means, Rapidminer 5.3, social media, UMKM*

KATA PENGANTAR

Puji syukur penulis mengucapkan kepada Tuhan Yang Maha Esa yang melimpahkan segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Sistem Informasi Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan, penulis menyadari bahwa skripsi ini tidak akan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada :

1. Rektor Universitas Putera Batam Ibu Dr. Nur Elfi Husda, S.Kom., M.SI.;
2. Ketua Program Studi Sistem Informasi Bapak Muhammat Rasid Ridho, S.Kom., M.SI.;
3. Ibu Erlin Elisa, S.Kom., M.Kom. selaku pembimbing Skripsi pada Program Studi Sistem Informasi Universitas Putera Batam;
4. Dosen dan Staff Universitas Putera Batam;
5. Kedua orang tua penulis yang selalu mendoakan dan menyemangati penulis hingga penulisan skripsi ini selesai;
6. Semua pihak yang telah bersedia meluangkan waktu, tenaga dan pikirannya dalam memberikan data dan informasi selama penulis membuat skripsi yang tidak dapat penulis sebutkan satu-persatu.

Semoga Tuhan membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Amin.

Batam, 25 Januari 2021

FITRINI

DAFTAR ISI

Halaman

HALAMAN SAMPUL ..... **i**
HALAMAN JUDUL ..... **ii**
SURAT PERNYATAAN ORISINALITAS ..... **iii**
HALAMAN PENGESAHAN ..... **iv**
ABSTRAK ..... **v**
ABSTRACT ..... **vi**
KATA PENGANTAR ..... **vii**
DAFTAR ISI ..... **viii**
DAFTAR GAMBAR ..... **x**
DAFTAR TABEL ..... **xi**
DAFTAR RUMUS ..... **xii**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Identifikasi Masalah 5
1.3 Batasan Masalah 5
1.4 Rumusan Masalah 6
1.5 Tujuan Penelitian 6
1.6 Manfaat Penelitian 6
BAB II TINJAUAN PUSTAKA **8**
2.1 Teori 8
2.2 Kerangka Pemikiran 16
2.3 Hipotesis Penelitian 16
2.4 Penelitian Terdahulu 17
BAB III METODE PENELITIAN **22**
3.1 Desain Penelitian 22
3.2 Objek Penelitian 24
3.3 Populasi dan Sampel 25
3.4 Variabel Penelitian 25
3.5 Teknik Pengumpulan Data 26
3.6 Model Penelitian 28

BAB IV HASIL DAN PEMBAHASAN..... 29
4.1 Hasil Analisa dengan *datamining* Algoritma *clustering K-Means* 29
4.2 Melakukan Pra Proses Data Penelitian 29
4.3 Analisa Data 33
4.4 Pengujian Menggunakan *Software RapidMiner 5.3* 59
BAB V SIMPULAN DAN SARAN 72
5.1 Simpulan 72
5.2 Saran..... 73
DAFTAR PUSTAKA 74
LAMPIRAN
LAMPIRAN 1 PENDUKUNG PENELITIAN
LAMPIRAN 2 DAFTAR RIWAYAT HIDUP
LAMPIRAN 3 SURAT KETERANGAN PENELITIAN

DAFTAR GAMBAR

Halaman

Gambar 1.1 Statistik pengguna media sosial 2019 2
Gambar 1.2 Jumlah Unit UMKM 3
Gambar 2.1 Tahapan *Knowlegde Discovery in Database* (KDD) 8
Gambar 2.2 Kerangka Pemikiran 16
Gambar 3.1 Deain Penelitian 23
Gambar 3.2 Model Penelitian 28
Gambar 4.1 Data *Rapidminer* Skripsi.xlxs 59
Gambar 4.2 Tampilan Awal *Rapidminer 5.3* 60
Gambar 4.3 Menu *New Process* 60
Gambar 4.4 Menu *New Process* 60
Gambar 4.5 Mencari *Read Excel* 61
Gambar 4.6 *Read Excel* 61
Gambar 4.7 *Import Configuration Wizard* 62
Gambar 4.8 Data *Import Wizard Step* 1 62
Gambar 4.9 Data *Import Wizard Step* 2 63
Gambar 4.10 Data *Import Wizard Step* 3 63
Gambar 4.11 Data *Import Wizard Step* Terakhir 64
Gambar 4.12 *Clustering K-Means* 64
Gambar 4.13 Penentuan Titik Pusat 65
Gambar 4.14 Hubungkan *Read Excel* ke *Clustering K-Means* 66
Gambar 4.15 *Cluster Distance Performance* 67
Gambar 4.16 Proses *Run* 67
Gambar 4.17 *ExamplesSet(ReadExcel)* 68
Gambar 4.18 *Cluster Model* 69
Gambar 4.19 *Centroid Table* 69
Gambar 4.20 *Graph View* 70
Gambar 4.21 *Centroid Plot View* 70
Gambar 4.22 *Performance Vector* 70

DAFTAR TABEL

Halaman

Tabel 4.1 Data Sampel Menentukan Media Promosi UMKM di Kota Batam..... 30
Tabel 4.2 Lanjutan..... 31
Tabel 4.3 Lanjutan..... 32
Tabel 4.4 Lanjutan..... 33
Tabel 4.5 Transformasi Data Jenis Kelamin 33
Tabel 4.6 Transformasi Data Usia..... 34
Tabel 4.7 Transformasi Data Status 34
Tabel 4.8 Transformasi Media Sosial..... 34
Tabel 4.9 Transformasi *E-Commerce*..... 34
Tabel 4.10 Pra Proses Data Penelitian..... 35
Tabel 4.11 Lanjutan..... 36
Tabel 4.12 Lanjutan..... 37
Tabel 4.13 Perhitungan Iterasi Pertama..... 38
Tabel 4.14 Lanjutan..... 39
Tabel 4.15 Lanjutan..... 40
Tabel 4.16 Lanjutan..... 41
Tabel 4.17 Lanjutan..... 42
Tabel 4.18 Lanjutan..... 43
Tabel 4.19 Lanjutan..... 44
Tabel 4.20 Perhitungan Iterasi Kedua 45
Tabel 4.21 Lanjutan..... 46
Tabel 4.22 Lanjutan..... 47
Tabel 4.23 Lanjutan..... 48
Tabel 4.24 Perhitungan Iterasi Ketiga 49
Tabel 4.25 Lanjutan..... 50
Tabel 4.26 Lanjutan..... 51
Tabel 4.27 Lanjutan..... 52
Tabel 4.28 Hasil Akhir Perhitungan Algoritma *Clustering K-Means*..... 52
Tabel 4.29 Lanjutan..... 53
Tabel 4.30 Lanjutan..... 54
Tabel 4.31 Lanjutan..... 55
Tabel 4.32 Lanjutan..... 56
Tabel 4.33 Lanjutan..... 57
Tabel 4.34 Lanjutan..... 58
Tabel 4.35 Lanjutan..... 59

DAFTAR RUMUS

Halaman

Rumus 2.1 Teori Jarak *Euclidean* 12

BAB I

PENDAHULUAN

1.1 Latar Belakang

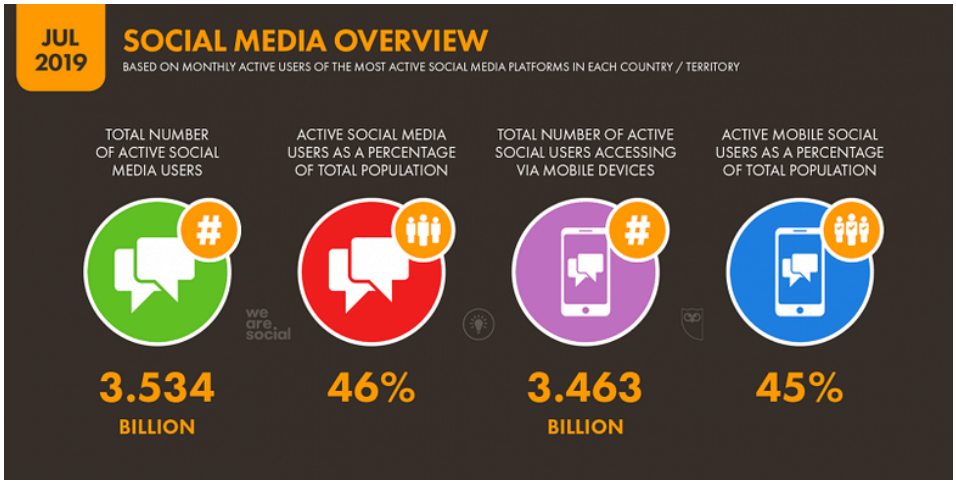
Di era saat ini, perkembangan teknologi menjadi perhatian publik. Dapat kita rasakan bahwa kemajuan teknologi diberbagai bidang seperti komunikasi, trasportasi, dan internet. Kemajuan teknologi yang sangat pesat saat ini adalah *gadget*. Masyarakat Indonesia sering menggunakan *gadget* untuk bermain game, komunikasi, dan berbagai keperluan lainnya bahkan mengakses berbagai media sosial.

Perkembangan teknologi di Indonesia semakin meningkat dari tahun ke tahun. Masyarakat semakin banyak menggunakan teknologi untuk berbagai kegiatan dalam sehari – hari. Berdasarkan data dari Badan Pusat Statistik (BPS) bahwa Indonesia memasuki peringkat 111 negara yang paling banyak menggunakan teknologi dari 176 negara. Penggunaan internet di Indonesia juga semakin meningkat seiring semakin majunya *gadget*. Berdasarkan data Kementrian Komunikasi dan Informatika, bahwa pengguna internet di Indonesia mencapai 175,5 juta jiwa. Sampai saat ini, Pemerintah sedang memperluas jangkauan frekuensi internet agar semakin berkembangnya berbagai kegiatan yang dapat kita lakukan dengan adanya internet, salah satunya adalah media sosial.

Media sosial saat ini menjadi perhatian para masyarakat di Indonesia. Media sosial dapat memberikan berbagai kemudahan bagi pengguna. Kemunculan media sosial tidak hanya untuk diakses saat waktu luang saja. Tetapi banyak sekali

masyarakat memanfaatkan media sosial untuk bermain games, dalam dunia pendidikan, berbelanja online, dan peluang dalam bisnis.

Kemunculan berbagai macam media sosial yang semakin pesat, memberikan berbagai kegunaan untuk masyarakat. Pada media sosial, kita dapat memunculkan berbagai ide kreasi berupa menulis, bercerita, dan sebagainya. Data yang di dapatkan oleh Kementrian Komunikasi dan Informatika, penggunaan internet sebanyak 63 juta orang dan sebanyak 95% digunakan unutk mengakses media sosial seperti *Facebook*, *Twitter*, dan sebagainya. Berikut grafik penggunaan media sosial secara global berdasarkan data dari *We are social and Hootsuit*.

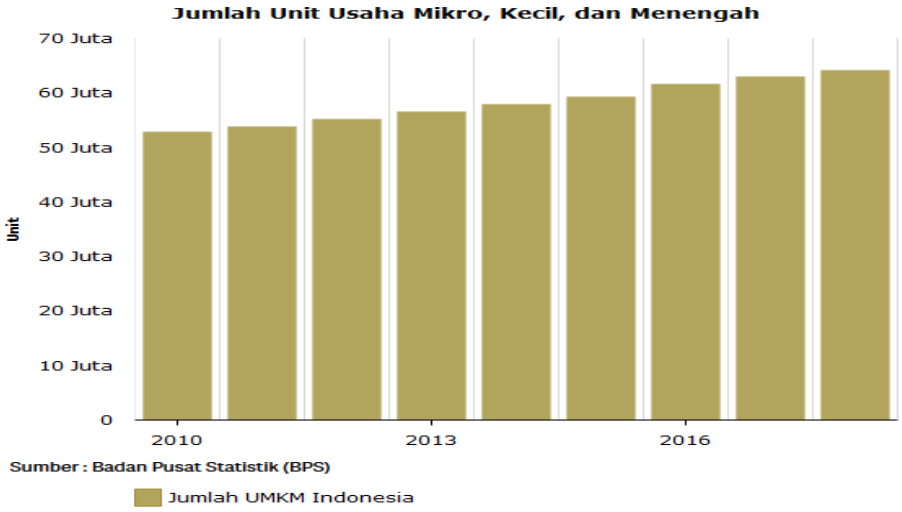


Gambar 1.1 Statistik pengguna media sosial 2019
Sumber : (*We Are Social and Hootsuite*)

Berdasarkan grafik di atas dapat kita lihat bahwa tidak hanya saja masyarakat di Indonesia yang banyak menggunakan media sosial. Tapi seluruh dunia sudah menggunakan media sosial dan sudah menjadi rutinitas hari – hari masyarakat. Dapat dilihat bahwa sudah 3,5 miliar lebih yang menggunakan media sosial.

Sebanyak 46% orang yang aktif menggunakan media sosial. Dan sebanyak 45% pengguna mengakses media sosial melalui *gadget*.

Dengan adanya media sosial maka banyak dimanfaatkan untuk kegiatan berbisnis bagi pembisnis kecil maupun besar. Berikut grafik statistik jumlah UMKM berdasarkan Badan Pusat Statistik.



Gambar 1. 2 Jumlah Unit UMKM
Sumber : (Badan Pusat Statistik, 2020)

Berdasarkan grafik di atas, dapat kita ketahui bertapa besarnya perkembangan UMKM di Indonesia, salah satunya di Kota Batam. Masyarakat Batam banyak membuka bisnis diberbagai bidang seperti, makanan, minuman, *fashion*, dan lainnya. Dalam melakukan transaksi bisnis, saat ini para pebisnis lebih menggunakan media sosial dalam mempromosikan barang dan jasa, bahkan ada yang mengunakan iklan berbayar. Namun perlu kita ketahui bahwa keuntungan yang kita dapatkan setelah mempromosikan melalui iklan berbayar tidak menjamin akan kesuksesan tersebut.

Selain menggunakan iklan berbayar, kita juga dapat menggunakan media sosial yang sudah terdapat berbagai pilihan dan *E-Commerce* serta bisa mempromosikan produk maupun bisnis tanpa perlu biaya yang besar. Hal tersebut sangat mengguntungkan dan memudahkan bagi para masyarakat yang masih pemula untuk memulai bisnis.

Dengan banyaknya masyarakat menggunakan media sosial untuk sebagai media promosi bisnisnya, penulis melakukan observasi pada lingkungan sekitar. Penulis mengetahui bahwa banyak masyarakat yang kebinggungan dalam memilih media sosial yang cocok untuk mempromosikan bisnisnya tersebut. Karena, terkadang masyarakat memilih media sosial tersebut tetapi tidak adanya perkembangan dalam bisnis tersebut.

Berdasarkan permasalahan tersebut, maka penulis akan meneliti pemanfaatan *datamining* dengan teknik *clustering* algoritma *K-means* dalam menentukan media promosi UMKM di Kota Batam dengan acuan pemanfaatan media sosial yang banyak digunakan oleh masyarakat di Batam. Dimana *datamining clustering* ini pada penelitian sebelumnya telah di uji untuk mengukur Penentuan Startegi Promosi Usaha Mikro Kecil dan Menengah (UMKM) di Purwokerto oleh (Astuti, 2019). Pada penelitian tersebut, peneliti melakukan penelitian strategi dalam mempromosikan usaha. Peneliti melakukannya dengan mengfokuskan pada tiga aspek yaitu, jenis usaha, omset dan aset. Peneliti juga melakukan pengelompokkan berdasarkkan jenis usaha agar memudahkan UMKM dalam melakukan prmosi usaha tersebut.

Berdasarkan latar belakang diatas maka penulis tertarik untuk menganalisa permasalahan tersebut dalam bentuk tugas akhir yang berjudul "**Pemanfaatan Data Mining Clustering Dalam Penentuan Media Promosi UMKM Di Kota Batam**" agar dapat membantu masyarakat UMKM yang berada di Kota Batam dapat mengembangkan bisnisnya terutama untuk strategi pemasaran dengan lebih baik lagi.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah dijelaskan diatas, penulis mengidentifikasikan masalah – masalah yang ada dalam penelitian sebagai berikut.

1. Masyarakat UMKM belum mengetahui strategi pemasaran atau promosi di media sosial yang banyak digunakan masyarakat.
2. Kebanyakkan dari UMKM belum mengetahui media sosial yang banyak digunakan bagi calon pelanggan guna untuk meningkatkan promosinya disana.

1.3 Batasan Masalah

Untuk menghindari adanya penyimpangan maupun pelebaran pokok masalah dalam penyusunan penelitian ini maka peneliti memberikan batasan masalah, yaitu :

1. Penelitian ini dilaksanakan pada UMKM yang ada di Kota Batam pada kecamatan Batam Kota.
2. Metode yang digunakan untuk mengelompokkan media promosi untuk UMKM tersebut adalah *clustering* dengan algoritma *K-Means* sebagai

perhitungan analisisnya, kemudian akan diuji dengan *Rapidminer5.3* untuk pengujian hasilnya.

1.4 Rumusan Masalah

Berdasarkan identifikasi masalah, maka rumusan masalah dalam penelitian ini adalah :

1. Bagaimana *clustering* dengan algoritma *K-Means* menganalisa data pemanfaatan media sosial untuk menentukan media promosi pada UMKM yang ada di Kota Batam ?
2. Bagaimana hasil analisa dari *datamining clustering* dengan algoritma *K-Means* dalam menentukan pengelompokkan media promosi ?

1.5 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan sebelumnya, tujuan yang hendak dicapai dari penelitian ini adalah :

1. Untuk mengetahui analisa *clustering* dengan algoritma *K-Means* dalam menentukan media promosi pada UMKM yang ada di Kota Batam.
2. Untuk mengetahui hasil analisa dari *datamining clustering* dengan algoritma *K-Means* dalam menentukan pengelompokkan media promosi.

1.6 Manfaat Penelitian

Demikian penjelasan penulis mengenai manfaat penelitian yang diharapkan dapat diterima oleh pihak lain atau peneliti lainya, yaitu :

1.6.1 Bagi Penulis

Memberikan pengetahuan lebih mendalam pada penulis dalam bidang penelitian yang berkaitan dengan pemanfaatan *datamining clustering* dalam penentuan media promosi UMKM di Kota Batam, dalam menerapkan wawasan teori yang diperoleh dari Universitas Putera Batam dan berbagai referensi pendukung lainnya.

1.6.2 Bagi Mahasiswa

Dapat menjadi referensi dan pedoman penelitian bagi mahasiswa yang akan melakukan penelitian khususnya yang menggunakan teknik *datamining* dan perhitungan *clustering K-Means* dan membahas penentuan media promosi UMKM di Kota Batam.

1.6.3 Bagi Objek Penelitian

Penelitian ini diharapkan menambah kajian baru yang dapat menjadi pertimbangan dalam memilih media sosial yang akan digunakan untuk mempromosikan bisnis tersebut. Agar bisnis yang di promosikan akan tepat pada sasaran yang di harapkan.

BAB II

TINJAUAN PUSTAKA

2.1 Teori

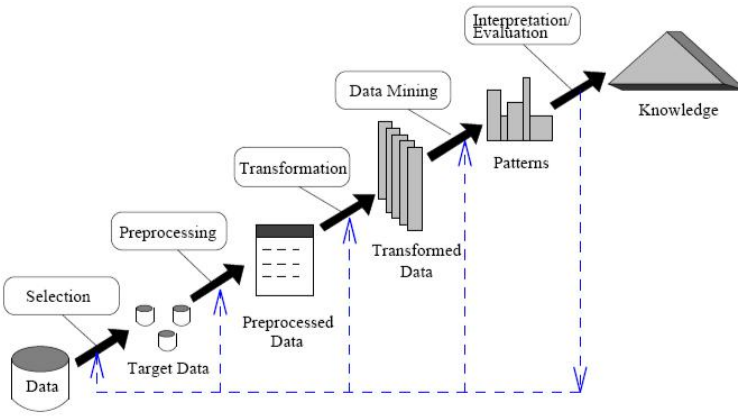
Penulis akan membahas beberapa teori yang akan menjadi pedoman dasar dalam melaksanakan penelitian, yaitu dengan menjelaskan mengenai pengertian dan istilah yang berkaitan dengan judul penelitian.

2.1.1 Teori Umum

Penulis membahas beberapa teori umum yang berkaitan dengan pelaksanaan penelitian yang terdiri dari pengetian dan istilah secara umum.

2.1.1.1 *Knowledge Discovery in Database* (KDD)

KDD merupakan sebuah aktifitas yang memiliki beberapa tujuan dan tahapan dalam mengumpulkan data yang berjumlah besar. Berikut tahapan dari *Knowlegde Discovery in Database* secara detail (Arnol Holman Tua, 2020).



Gambar 2.1 Tahapan *Knowlegde Discovery in Database* (KDD)

Dari gambar 2.1 diatas maka tahapan KDD hingga menjadi suatu *Knowledge* atau pengetahuan baru adalah :

a. Selection

Selection digunakan untuk mimilih variable yang akan diambil agar tidak adanya kesamaan dalam variable dan terjadinya perulangan dalam pengolahan data.

b. Processing

Processing terdapat 2 tahap, yaitu:

a. Untuk menghilangkan data yang tidak diperlukan yaitu *Data Cleaning*.
b. Data integration untuk mengidentifikasikan data yang unik.

c. Transformation

Transformation adalah sebuah proses perubahan data, akan tetapi data tidak akan beda dari format existensi yang telah ditetapakan dengan prosedur pengolahan *datamining*.

d. Datamining

Proses utama dalam mendapatkan informasi dan data baru kemudian dilakukannya pemrosesan yang sesuai dengan prosedur yang sudah diberlakukan di *datamining*.

e. Interpretation / Evaluation

Untuk mengidentifikasikan pola yang menarik kedalam data yang akan diidentifikasikan.

f. Knowledge

Sebuah data yang telah didaptka hasil untuk dijadikan sebuah informasi baru agar dapat digunakan dalam pengambilan sebuah keputusan.

Selain itu, KDD juga merupakan suatu proses dimana sebuah data yang mentah diolah menjadi sebuah informasi yang bermanfaat (Anita, M.Kom, Yulya, M., M.Kom, Sri, T., 2018).

2.1.1.2 *Datamining*

Datamining dapat dikatakan sebagai metode pengolahan data yang berskala besar, sehingga sangat berperan penting dalam berbagai bidang (Rofiqo et al., 2018). Selain itu, *datamining* data digunakan dalam pengambilan keputusan. *Datamining* juga sering disebut dengan *Knowledge Discovery in Database* (KDD) atau *Pattern Recognition* (Sadewo et al., 2019). *Datamining* juga merupakan sebuah metode yang digunakan untuk pengolahan data yang berjumlah banyak (Maulida, 2018). Selain itu, *Datamining* juga digunakan dalam mengelompokkan penemuan pengetahuan di dalam sebuah basis data, serta *Datamining* merupakan sebuah proses yang menggunakan statistik, kecerdasan buatan, matematikan, dan *machine learning* yang nantinya akan menghasilkan kelompok data yang bermanfaat (Elisa, 2018).

2.1.1.3 *Clustering*

Sebuah proses pembagian data kebeberapa kelompok dengan kesamaan data yang sama. Terdapat dua teknik metode *cluster* yaitu (Sadewo et al., 2019):

a. *Hierarchical Clustering*

   Suatu metode pengelompokkan data yang cara kerjanya menggunakan 2 data atau lebih yang memiliki kecenderungan mirip, kemudian dilajutkan proses objek lain yang mempunyai kedekatan, proses terus berlangsung sampai cluster mencapai sebuah bentuk *tree*.

b. *Non Hierarchical Clustering*

   Pada proses ini dimulai dengan memilih cluster yang diinginkan, kemudia proses *cluster* dilakukan tanpa harus mengikuti proses hirarki.

Clustering juga merupakan metode yang dalam prosesnya tidak tentu arah, yang berarti dalam metode ini tidak adanya training maupun pelatihan khusus dan tidak perlu adanya output (Suriani, 2020). Kelemahan dari metode *clustering* adalah metode tersebut dapat dikatakan sangat sensitif akan saat pemilihan *cluster* awal dan perhitungan dalam mencapai hasil yang optimal (Handoko, 2016).

2.1.1.4 *K-Means*

K-Means merupakan suatu metode pengelompokkan data dengan *Non Hierarchical Clustering* yang dilakukan dengan memasukkan data ke dalam bentuk dua atau lebih kelompok (Annur, 2019). Langkah – langkah dalam melakukan *clustering* dengan metode *K-Means* adalah sebagai berikut (Triyansyah & Fitrianah, 2018) :

1. Tentukan jumlah *cluster* k.
2. Menentukan titik pusat awal dari setiap *cluster*.
3. Kelompokkan semua data/objek ke *cluster* terdekat dan untuk menghitung jarak semua data ke setiap titik pusat *cluster* dapat menggunakan teori jarak *Euclidean* yang dirumuskan sebagai berikut :

$$D(x, y) = \sqrt{(x1 - y1)^2 + (x2 - y2)^2} \quad (1)$$

Rumus 2.1 Teori Jarak *Euclidean*

4. Hitung kembali pusat *cluster* dengan kelompok *cluster* yang sekarang. Pusat *cluster* adalah rata – rata dari semua data/objek dalam *cluster* tertentu.
5. Lakukan perhitungan lagi setiap objek memakai pusat *cluster* yang baru. Jika pusat *cluster* tidak berubah lagi maka proses clustering selesai. Dan jika adanya perbedaan makan akan kembali ke langkah nomor 3 sampai pusat *cluster* tidak berubah lagi.

K-Means juga merupakan metode yang membagikan data menjadi beberapa kelompok yang memiliki karakteristik yang sama dan jika ada yang berbeda maka akan di masukkan kedalam kelompok lainnya (Bena Siti Ashari, Steven Christ Otniel, 2019).

2.1.1.5 *Rapidminer*

Rapidminer merupakan perangkat lunak yang berupa ilmu data yang dikembangkan oleh perusahaan bernama sama dengan yang menyediakan

lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam penambangan teks, dan analisis prediktif (Dewi et al., 2019).

Rapidminer merupakan aplikasi *Open Source* dan dapat menjadi solusi dalam mengerjakan *dataming*, *text mining* , dan analisa prediksi . *Rapidminer* memiliki 500 operator *datamining*. *Rapidminer* juga merupakan aplikasi yang dapat berdiri sendiri dalam menganalisis sebuah data (Ardiansyah & Walim, 2018).

2.1.2 Teori Khusus

2.1.2.1 UMKM

UMKM merupakan sebuah usaha yang untuk mengembangkan perekonomian disetiap Negara, khususnya pada negera yang masih berkembang yang berdasarkan Undang – undang no. 20 pasal 1 Tahun 2008. Yang mengartikan UMKM berdasarkan usaha kecil, usaha menengah, dan usaha besar yang dimana memenuhi kriterial usaha berdasarkan undang – undang yang memiliki kekayaan bersih 50.000.000 sampai 10 milyar , sedangkan hasil penjualan mencapai 300 juta sampai 500 milyar per tahun (Sari & Santoso, 2019).

UMKM sudah menjadi peran yang sangat penting di Indonesia. Hal tersebut dikarenakan UMKM merupakan usaha yang berjumlah lebih banyak dibandingkan dengan usaha di bidang industri (Suci et al., 2017).

Selain itu, UMKM sudah menjadi perhatian bagi pemerintah untuk terus dikembangkan. UMKM juga sudah menjadi sector dalam menyediakan peluang besar bagi masyarakat yang ingin mencari pekerjaan dan terbukti dapat membantu masyarakat dalam mengatasi ekonomi yang krisis (Ilmaniati & Putro, 2019).

2.1.2.2 Internet

Singkatan dari internet adalah *Interconnection Networking*. Internet merupakan sebuah jaringan yang dapat menghubungkan jaringan satu dengan lainnya agar komputer dapat melakukan komunikasi dengan komputer lainnya (Irawan, 2018).

Internet merupakan sebuah jaringan komputer yang sangat besar dan memiliki jutaan perangkat komputer yang menghubungkan jaringan satu ke jaringan komputer lainnya yang dapat digunakan unutk melakukan pertukaran informasi (Putri, 2020).

2.1.2.3 Media Sosial

Sebelum adanya perkembangan dalam dunia teknologi, kita menonton dengan televisi yang hanya terdapat warna hitam putih. Dengan adanya perkembangan teknologi maka munculnya televisi digital yang dapat kita sambungkan ke internet. Selain itu, juga terdapat media sosial yang dulu kita gunakan media cetak dan sekarang kita sudah dapat menggunakan sambungan internet dan menggunakan berbagai macam media sosial yang semakin canggih dan mudah digunakan, seperti *Facebook*, *Instagram*, *Whatsapp*, dan sebagainya (Gani, 2020).

Media sosial merupakan komunikasi yang dilakukan melalui daring menggunakan sebuah situs *websites* yang mengubah komunikasi menjadi dialog interaktif . Dulu kita menggunakan media tradisional dengan media cetak, sekarang kita mengakses media sosial dengan internet (Wisti Aristika.., 2019).

Media sosial merupakan media yang berbasis internet untuk digunakan dalam berbagai kegiatan seperti untuk melakukan komunikasi dengan pengguna lainnya dan untuk sebagai wawasan luas bagi pengguna dalam berinteraksi dengan pengguna media sosial lainnya (Budi Mahardhika & Sunariani, 2019).

Media sosial juga merupakan komunikasi alternatif yang sangat mudah digunakan oleh masyarakat, sehingga dapat mempersingkat waktu dalam komunikasi jarak jauh serta dapat dilakukan kapapun dan dimana saja (Siti Nurul Intan & Hutabarat, 2020).

2.1.2.4 Media online

E-commerce saat ini sudah terbukti dapat membantu dalam kita mempromosikan usaha ataupun produk yang akan dijual. Dapat disebutkan bahwa proses jual dan pembelian sebuah produk secara internet atau *online* adalah *e-commerce. E-commerce* juga merupakan sebuah proses dalam membuat, mengelola sebuah hubungan komersial secara *onlin*e (Arifianto & Coiri, 2018).

2.2 Kerangka Pemikiran

Kerangka pemikiran dari penelitian sebagai berikut :

Gambar 2.2 Kerangka Pemikiran
Sumber : (Penulis, 2020)

2.3 Hipotesis Penelitian

Demikian penjelasan dari hipotesis yang menjadi atau dugaan penulis sementara terhadap penelitian, yaitu :

1. Diduga *datamining* dengan algoritma *K-means* dapat menentukan media sosial yang paling diminati oleh masyarakat di Kota Batam tepatnya pada kecamatan Batam Kota.
2. Diduga hasil pengolahan *datamining* dengan algoritma *K-means* dapat memberikan solusi bagi UMKM dalam menentukan strategi promosi di media sosial.

2.4 Penelitian Terdahulu

Penulis menemuan beberapa penelitian terdahulu yang dapat membantu dalam melakukan penelitian ini yang berkaitan dengan metode *datamining* dengan algoritma *clustering K-Means*.

1. Penelitian tentang "Penentuan Tingkat Minat Belanja Online Melalui Media Sosial Menggunakan Metode *Clustering K-Means*" (Anita, M.Kom, Yulya, M., M.Kom, Sri, T., 2018). Dari hasil penelitian tersebut menggunakan perhitungan *clustering K-means* dapat mengetahui media sosial yang paling sering digunakan dalam melakukan belanja online.
2. Penelitian tentang "Implementasi *Rapidminer* Dengan Metode *K-Means* (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi)" (Agus Perdana,. Riyani Wulan Sari, 2018). Dari hasil penelitian tersebut diharapkan pemerintah dapat lebih memperhatikan imunisasi campak pada balita khusus di provinsi yang keterbatasan dalam imunisasi campak pada balita.
3. Penelitian tentang "Penerapan *Datamining* Dalam Meningkatkan Mutu Pembelajaran Pada Instansi Perguruan Tinggi Menggunakan Metode *K-Means Clustering* (Studi Kasus di Program Studi TKJ Akademi Komunitas Solok Selatan)" (Handoko, 2016). Dari hasil penelitian tersebut dapat kita lihat akademik mahasiswa tersebut tergolong sedang karena IP rata – rata sekitar 2,50 s/d 3,00 dan jarak ke kampus terlampau jauh sekitar 5 M serta jumlah kehadiran berjumlah 85 s/d 90 kehadiran.
4. Penelitian tentang "Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode *Crisp –DM* dengan Algoritma *K-*

Means Clustering" (Astuti, 2019). Dari hasil penelitian adalah kita harus memiliki strategi dalam melakukan promosi berdasarkan omset masing – masing dalam meningkatkan penjualan.

5. Penelitian tentang "Penerapan *Clustering K-Means* Untuk Menentukan Pengaruh Media Sosial *Facebook* Terhadap Usaha Mikro , Kecil, dan Menengah (UMKM) Di Kecamatan Pekanbaru Kota" (Wisti Aristika.., 2019). Dari hasil penelitian tersebut dapat membantu dalam menentukan seberapa besarnya pengaruh media sosial *Facebook* terhadap UMKM di Kecamatan Pekanbaru.
6. Penelitian tentang "Algoritma *K-Means* Dalam Mengelompokkan Desa/Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Berdasarkan Provinsi" (Sadewo et al., 2019). Dari hasil penelitian ini dapat menjadi sebuah acuan bagi Pemerintah agar lebih memperhatikan dalam pembangunan lampu di jalan utama.
7. Penelitian tentang "Penerapan *Clustering* Pada Penduduk Yang Mempunyai Keluhan Kesehatan Dengan *Datamining K-Means*" (Rofiqo et al., 2018). Dari hasil penelitian tersebut dapat diketahui terdapat 6 provinsi yang memiliki keluhan kesehatan yang tinggi, 18 provinsi dengan keluhan kesehatan yang sedang dan 9 provinsi yang memiliki keluhan kesehatan yang rendah.
8. Penelitian tentang "Penerapan *Datamining* Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan di Prov. DKI Jakarta Dengan *K-Means*" (Maulida, 2018). Dari hasil penelitian tersebut dapat menjadi perhatian pemerintah untuk lebih memperhatikan tempat wisata yang

perlu untuk di perbaikan ulang kembali agar pengunjung dapat berwisata dengan nyaman.

9. Penelitian tentang “Penerapan Metode *K-Means Cluster Analysis* Pada Sistem Pendukung Keputusan Pemilihan Konsentrasi Untuk Mahasiswa Internasional Class STMIK AMIKOM Yogyakarta” (Aranda & Natasya, 2016). Dari hasil penelitian tersebut bahwa 4 mahasiswa diarahkan untuk mengambil pemograman, 4 mahasiswa mengambil multimedia dan 3 – 5 mahasiswa mengambil jaringan komputer.
10. Penelitian tentang “Pengelompokan Data Kriminal Pada Poldasu Menentukan Pola Daerah Rawan Tindak Kriminal Menggunakan *Datamining* Algoritma *K-Means Clustering*” (Suriani, 2020). Dari hasil penelitian tersebut dapat diketahui bahwa masih belum efektik karena metode yang digunakan masih belum mendukung. Dan kategori dicari berdasarkan daerah yang sangat rawan, sedang, dan tidak rawan.
11. Penelitian tentang “Penerapan Teknik *Clustering* Sebagai Startegi Pemasaran Pada Penjualan Buku Di Tokopedia dan Shopee” (Rachman & Wasiyanti, 2019). Dari hasil penelitian tersebut dapat diketrahui bahwa *clustering* dapat dilakukan untuk menganalisis pemasaran.
12. Penelitian tentang “Analisis *Datamining* Menggunakan Algoritma *K-Means Clustering* Untuk Menentukan Strategi Marketing” (Triyansyah & Fitrianah, 2018) . Dari hasil penelitian tersebut dapat diketahui bahwa terdapat 3 tipe sepatu yang paling diminati.

13. Penelitian tetang *"Analysis Of The K-Means Algorithm On Clean Water Customers Based On The Province" (Windarto et al., 2019). To cluster information on the number of cleaned water clients by k-means strategy utilizing tools RapidMiner, the result of clustering with 3 clusters: C1: 6 areas: North Sumatera, South Sumatera, Bali, South Kalimantan, East Kalimantan and South Sulawesi, C2: 4 territories : DKI Jakarta, West Java, Central Java and East Java, C3: 24 territories: Aceh, West Sumatra, Riau, Jambi, Bengkulu, Lampung, Bangka Island, Riau Island, Yogyakarta, Banten, West Nusa Tenggara, Central Nusa Tenggara, West Kalimantan, Central Kalimantan, North Sulawesi, North Sulawesi, Central Sulawesi, Southwest Sulawesi, Gorontalo, West Sulawesi, Maluku, North Maluku, West Papua and Papua.*
14. Penelitian tentang *"Using The K-Means Method For Diagnosing Cancer Stage Using The Pandas Library"(Meniailov et al., 2019). The consider analyzed the measurable reliance between the factors that determine the condition of patients. It is appeared how it is conceivable to decide the patient's having a place to a certain lesson on the premise of factual information - enlisted state variables. The created program arrangement, based on the Python Pandas library, makes it possible to classify on the premise of preparing tests, which guarantees a tall percentage of acknowledgment, as well as to distinguish the point of reference having a place to one of a few clusters. It ought to be famous that the depicted*

approach is widespread and can be utilized not only for biomedical frameworks, but too specialized, financial, etc.

15. Penelitian tentang "*Implementation Data Mining With K-Means Algorithm For Clustering Distribution Rabies Case Area In Palembang City "(Rahayu et al., 2020). The result of information preparing utilizing RapidMiner program with the strategy of K-Means can be known from 16 sub-districts in Palembang, seven sub-districts counting the territorial cluster is very inclined to rabies (C0), whereas the four sub-districts included within the cluster Rabies-prone areas (C1), and five sub-districts included within the territorial cluster isn't inclined to rabies (C2). With the scourge bunch can be mediated to a program of Palembang wellbeing benefit that is fast-responsive and on target.*

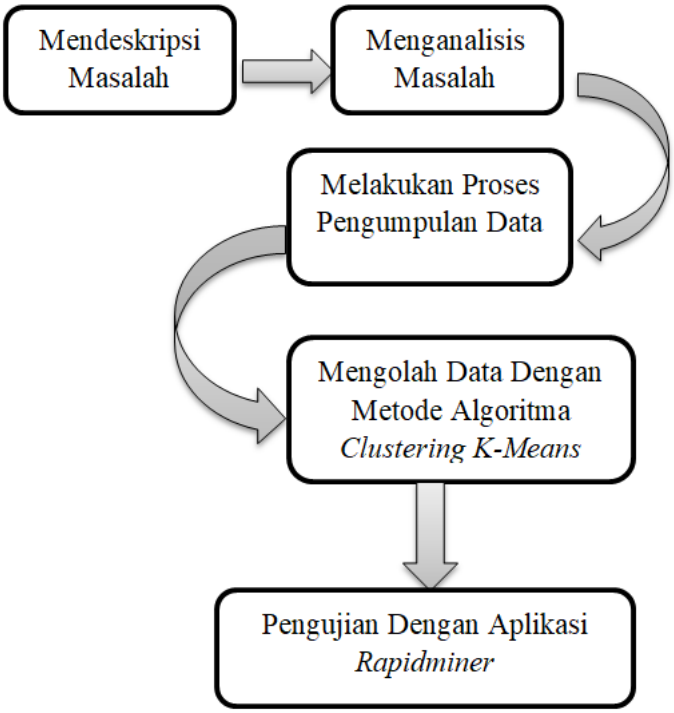
BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Penelitian ini adalah penelitian kualittif. Dimana penelitian kualitatif merupakan sebuah hasil penelitian yang tidak dapat menggunakan prosedur statistik atau biasanya disebut dengan kuantitatif. Selain itu, penelitian kualitatif biasanya dapat meneliti kehidupan masyarakat, sejarah, tingkah laku, fungsionalisme organisasi, pergerakkan sosial dan hubungan kekerabatan. Menurut Denzin dan Lincoln bahwa penelitian kualitatif merupakan sebuah penelitian yang dilakukan berdasarkan sebuah latar belakang dan dengan menafsirkan suatu kejadian yang telah terjadi serta dilakukan berdasarkan metode yang ada pada penelitian kualitatif (Shidiq & Choiri, 2019).

Pada penelitian khususnya dalam laporan ini adalah mengenai "**Pemanfaatan Data Mining Clustering Dalam Penentuan Media Promosi UMKM Di Kota Batam"**. Model dan atau gambar desain dari penelitian yang dilakukan oleh penulis sebagai berikut.



Gambar 3.1 Deain Penelitian

Sumber : (Penulis,2020)

Berdasarkan diagram diatas maka penjelasan desain dari penelitian sebagai berikut:

1. Mendeskripsi Masalah

   Pada tahap ini, harus lebih terdahulu menentukan batasan dari permasalahan yang akan diteliti agar tidak terjadinya penyimpangan yang tidak diinginkan dan proses penelitian pun akan lebih teratur. Walaupun ini merupakan tahap awal dari sebuah penelitian, tetapi tahapan awal ini sangatlah penting untuk melanjutkan pada tahap selanjutnya.

2. Menganalisis Masalah

   Pada tahap ini, penulis akan menganalisis beberapa masalah yang telah ditentukan pada deskripsi masalah agar informasi yang disampaikan lebih mudah diapahami.

3. Melakukan Proses Pengumpulan Data

   Setelah melakukan analisis masalah, penulis akan melakukan pengumpulan data sesuai masalah yang akan diteliti. Disini, penulis akan menggunakan teknik kuesioner dalam pengumpulan data. Selain itu, penulis juga melakukan observasi dan mewanwacarai lingkungan di sekitar objek penelitian dan melakukan studi literatur dengan mempelajari data yang telah dilakukan oleh peneliti terdahulu.

4. Mengolah Data Dengan Metode *Clustering K-Means*

   Pada tahap ini, penulis akan melakukan pengolahan data berdasarkan aturan dari perhitungan dengan algoritma *clustering K-Means*.

5. Pengujian Hasil Menggunakan *Rapidminer 5.3*

   Pada tahap ini, dari hasil perhitungan menggunakan algoritma *clustering K-Means* penulis akan melakukan pengujian menggunakan bantuan dari *software Rapidminer* 5.3.

3.2 Objek Penelitian

3.2.1 Lokasi Penelitian

Penelitian dilakukan pada UMKM yang Berada di kecamatan Batam Kota di Kota Batam.

3.3 Populasi dan Sampel

3.3.1 Populasi

Populasi penelitian ini adalah seluruh masyarakat yang bergerak pada usaha UMKM di Kota Batam khususnya yang berada di kawasan kecamatan Batam Kota.

3.3.2 Sampel

Untuk sampel analisis penulis menggunakan teknik *accidental sampling*, dimana *accidental sampling* merupakan pegambilan sampel yang secara kebetulan maupun yang dianggap dapat memenuhi kriterial yang dibutuhkan dalam penelitian tersebut (Cahyani & Sitohang, 2016).

Dalam penelitian ini karena keterbatasan waktu yang dimiliki penulis, maka akan diambil sebanyak 133 populasi yang akan dijadikan sampel dengan teknik accidental sampling berdasarkan kriteria masyarakat UMKM yang sudah melakukan usaha tersebut selama 1 tahun. Yang dimana data didapatkan dari teknik pengumpulan data observasi, wawancara dan penyebaran kuesioner.

3.4 Variabel Penelitian

Pengolah data sesuai dengan hasil kuesioner yang di respon oleh masyarakat di Kota Batam yang berada di kawasan kecamatan Batam Kota.

Dimana variabelnya adalah :

1. Usia
2. Jenis Kelamin

3. Status
4. Media Sosial yang digunakan
5. *E-Commerce* Yang digunakan

Untuk hasil akhir adalah mengetahui media sosial yang paling sering digunakan oleh masyarakat dalam berbelanja dan dapat digunakan untuk pembisnis dalam menentukan media sosial yang cocok untuk melakukan promosi usaha tersebut serta dapat memiliki startegi dalam media promosi.

3.5 Teknik Pengumpulan Data

Proses pengumpulan data yang menjadi kebutuhan bahan penelitian dilakukan agar penulis dapat mencapai tujuan penelitian, yaitu :

1. Observasi

   Teknik ini merupakan teknik pengumpulan data berdasarkan pengamatan pada lingkungan sekitar. Sehingga pada tahap ini penulis melakukan pengamatan pada lingkungan kerja dan tempat tinggal sekitar untuk mendapatkan informasi yang berdasarkan fakta.

2. Wawancara

   Teknik ini merupakan pengumpulan data secara langsung bertatapan muka dengan objek yang akan kita teliti. Pada tahap ini, penulis melakukan pengumpulan data dengan langsung kepada objek di lingkungan tertentu yang dekat dengan penulis. Sehingga penulis dapat melakukan kesimpulan data sementara.

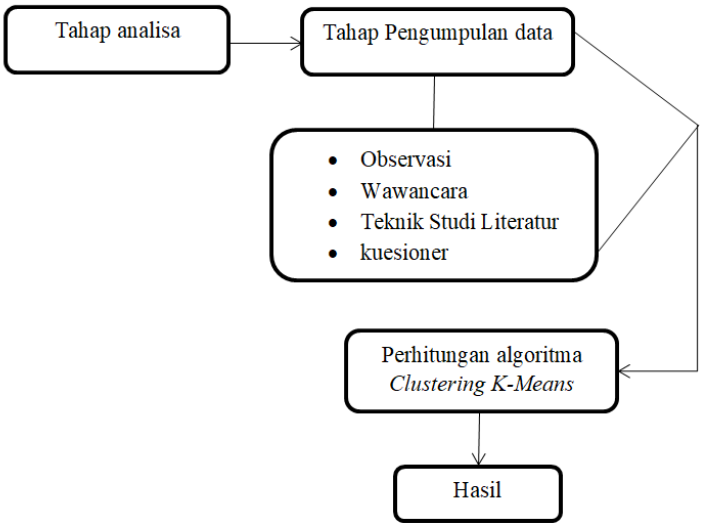
3. Teknik Studi Literatur

   Teknik ini merupakan sebuah pembahasan pada bidang penelitian tertentu. Kita juga dapat mendapatkan gambaran singkat yang dapat dipelajari , agurmentasi, dan berdasarkan masalah untuk permasalahan yang akan diteliti. Sehingga pada penelitian ini, penulis akan mencari berbagai informasi pendukung dari jurnal maupun buku. Penulis akan mempelajarinya dan menarik kesimpulan yang dapat membantu dalam melakukan penelitian tersebut.

4. Kuesioner

   Teknik ini merupakan pengumpulan data berdasarkan pengisian dari beberapa pertanyaan yang telah dibuat oleh penulis. Dan data tersebut akan dijadikan sebagai hasil peneliitian pada penelitian tersebut. Sehingga penulis akan membuat kuesioner sesuai dengan data yang dibutuhkan dan menyebarkannya kepada masyarakat yang berada di Kota Batam khususnya di kawasan Kecamatan Batam Kota.

3.6 Model Penelitian



Gambar 3. 2 Model Penelitian

Sumber : (Penulis,2020)

Berdasarkan gambar diatas, penulis akan melakukan analisa masalah menggunakan *datamining* dengan algoritma *clustering K-Means*. Tahapan awal akan dilakukan pengumpulan data menggunakan observasi, wawancara, teknik studi literature dan kuesioner. Selanjutnya, data yang terkumpul akan dilakukan perhitungan menggunakan algoritma *clustering K-Means* untuk mendapatkan *cluster* yang tertinggi. Hasil perhitungan akan dihitung sampai tidak adanya *cluster* yang berbeda. Selanjutnya, dari hasil perhitungan algoritma *clustering K-Means* didapatkan pola berupa aturan dasar penilaian mengenai **"Pemanfaatan Data Mining Clustering Dalam Penentuan Media Promosi UMKM Di Kota Batam"**.