# **BAB 1**

# **PENDAHULUAN**

## **Latar Belakang**

Setiap orang yang terlibat secara daring, baik di forum digital, kolom pesan, komentar atau jejaring sosial selalu ada risiko serius bahwa dirinya mungkin menjadi sasaran ejekan dan bahkan pelecehan dalam bahasa atau kalimat kasar. Kata – kata dan kalimat kasar yang diungkapkan dapat berdampak besar pada pengalaman pengguna. Seiring semakin banyak orang yang berkomunikasi secara daring, maka kebutuhan untuk menyaring komentar kasar secara otomatis yang memiliki akurasi tinggi menjadi sangat diperlukan. Dalam bahasa Indonesia, kata-kata kasar atau kotor biasanya berasal dari suatu kondisi, hewan, makhluk astral, benda, bagian tubuh, anggota keluarga, aktivitas, dan profesi (Tjahyanti:2020).

Kondisi.Kata-kata yang mengungkapkan kondisi yang tidak menyenangkan

dalam percakapan biasanya digunakan sebagai kata-kata kasar. Secara umum ada tiga hal yang dapat atau mungkin berhubungan dengan kondisi tidak menyenangkan ini, yaitu gangguan mental (misalnya: *gila, bego, goblok, idiot, sinting, bodoh, tolol, sontoloyo, geblek, sarap*), penyimpangan seksual (misalnya: *lesbi, homo, banci, waria*), kurangnya modernisasi (misalnya: *kampungan, udik, alay*), cacat fisik (misalnya: *buta, budek, bolot, bisu*), kondisi di mana seseorang tidak memiliki etika (misalnya: *brengsek, bejat, bajingan*) kondisi yang tidak disetujui oleh Tuhan atau

agama ( misalnya: *keparat, jahanam, terkutuk, kafir, najis*), dan kondisi yang terkait dengan keadaan yang tidak menguntungkan (misalnya: *celaka, mati, modar, sialan, pantek, mampus*) (Ibrahim dan Inda Budi, 2018).

Lapak belajar merupakan sebuah platform digital yang menyediakan layanan belajar bagi pelajar SMA dan SMK untuk lebih mudah memahami materi pembelajaran dengan efektif dan efisien. Lapak belajar memiliki beberapa fitur yaitu layanan kelas *one to one*, bimbingan privat belajar serta fitur diskusi yang mana seluruh pengguna lapak belajar bisa berdiskusi didalamnya membahas terkait materi yang tidak mereka pahami dan lain sebagai bagainya. Selain itu di dalam forum diskusi tersebut, pengguna lapak belajar bisa saling mengomentari setiap opini dan jawaban yang diberikan oleh pengguna lainnya ataupun oleh mentor. Namun pada fitur diskusi tersebut, diskusi tidak selalu berjalan dengan lancar sering kali terdapat pengguna yang tidak puas dengan jawaban atau komentar yang diberikan oleh pengguna dan mentor sehingga pengguna memberikan komentar yang kasar dan tidak seharusnya dikirimkan karena mengganggu kenyamanan dan pengalaman pengguna yang lainnya. Untuk mengatasi masalah ini pihak lapak belajar telah melakukan beberapa tindakan seperti menegur pengguna yang memberikan komentar kasar, menghapus komentar kasar secara manual hingga melakukan penangguhan akun pengguna tersebut. Dari semua tindakan yang telah dilakukan pihak lapak belajar tidak satupun yang berjalan secara efektif dan efisien, untuk tindakan pertama dengan melakukan peneguran itu hanya berdampak untuk 1 orang pengguna dan jika ada 10 orang pengguna maka pihak lapak belajar harus menegur ke 10 orang tersebut dan itu juga tidak menjamin mereka untuk tidak melakukannya kembali. Untuk tindakan ke 2 menghapusnya secara manual cukup efektif tapi memerlukan waktu karena harus di cek secara berkala dan komentar yang terdapat di dalam *database* lapak belajar sangat banyak sehingga tidak efektif jika dilakukan secara manual. Untuk tindakan ke 3 yaitu menangguhkan akun pengguna, itu akan membuat lapak belajar kehilangan pengguna dan itu berdampak buruk bagi keberlangsungan platform serta bisnisnya.

Setelah melakukan 3 tindakan dan tidak ada satupun yang membuahkan hasil memuaskan maka penulis merumuskan hipotesis jika membuat sebuah program otomatis yang dapat menyaring komentar kasar secara otomatis dan cepat maka akan mengurangi komentar kasar secara drastis dan juga menjadi solusi efektif dan efisien dalam menanggulangi komentar kasar.

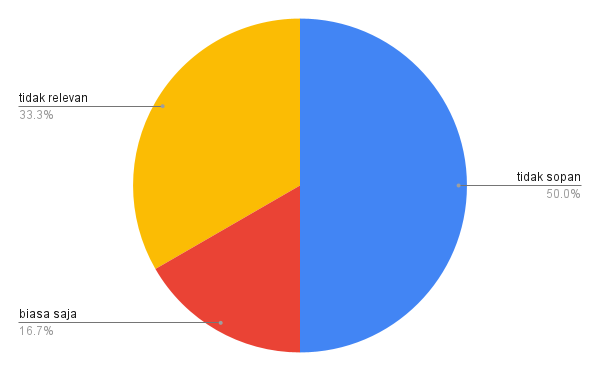
Oleh karena itu penulis melakukan survey kepada pihak internal lapak belajar serta kepada beberapa pengguna nya untuk mengetahui seberapa besar dampak komentar kasar yang ditimbulkan baik untuk platform lapak belajar baik bagi penggunanya. Survey yang disebarkan kepada pengguna lapak belajar meliputi, penulis menanyakan apakah pengguna lapak belajar pernah mendapatkan komentar kasar pada saat berdiskusi secara daring, 66,7% mengatakan pernah dan 36,3% mengatakan tidak pernah

Chart, pie chart

Description automatically generated

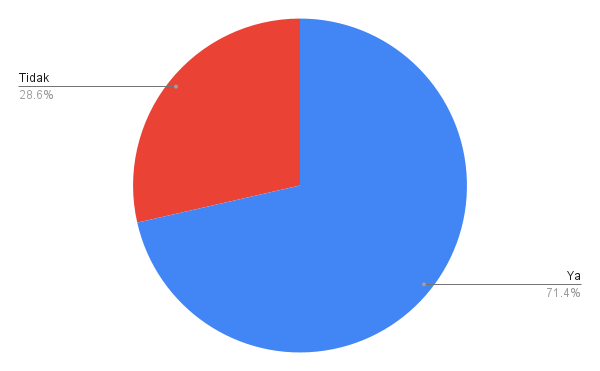
**Gambar 1.1** *Persentase pengguna yang mendapatkan komentar kasar*

penulis penanyakan tanggapan pengguna ketika mendapatkan komentar kasar dan hasilya sebagai berikut, 33% mengatakan tidak relevan, 16,7% mengatakan biasa saja dan 50% mengatakan tidak sopan



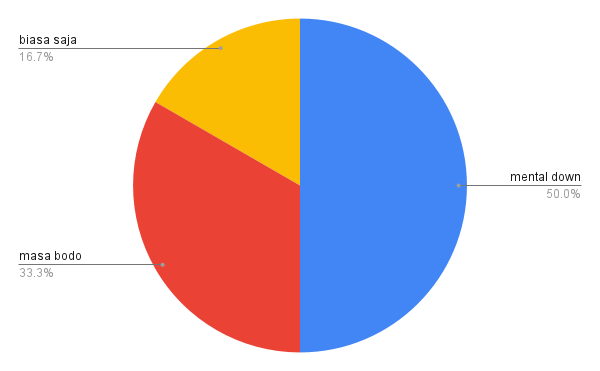
**Gambar 1.2** *Tanggapan pengguna terhadap komentar kasar*

Penulis menanyakan apakah komentar kasar tersebut mempengaruhi mereka dalam berdiskusi secara daring dan hasilnya 71,4% mengatakan iya dan 28,6% mengatakan tidak



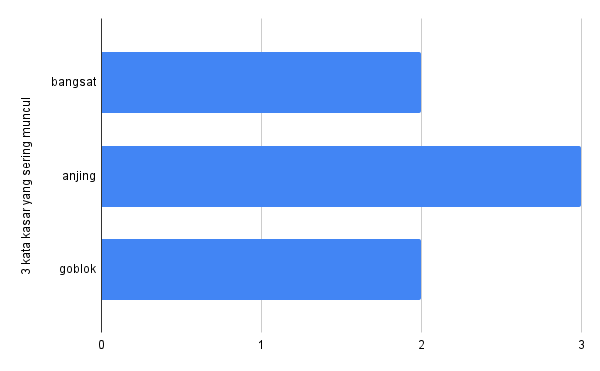
**Gambar 1.3** *persentase komentar kasar mempengaruhi diskusi secara daring*

Penulis menanyakan pengaruh apa saja yang didapatkan dari komentar kasar yang mereka dapatkan dan hasilnya 16,7% mengatakan biasa saja, 33,3% mengatkan masa bodo dan 50% mengatakan mental *down*



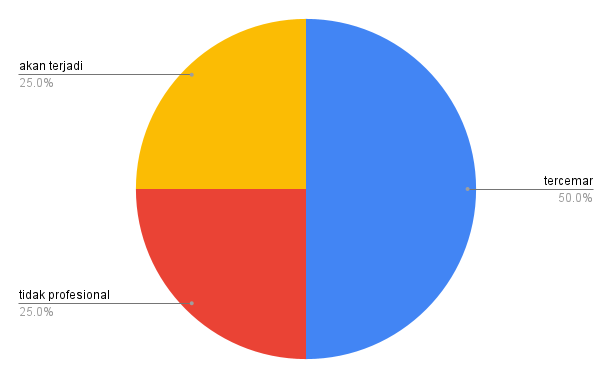
**Gambar 1.4** *persentase pengaruh yang didapatkan dari komentar kasar*

Penulis menanyakan tentang seperti apa contoh komentar kasar yang mereka dapatkan dan hasilnya sebagai berikut



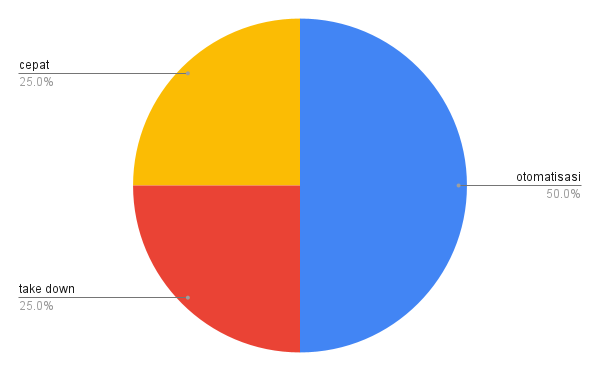
**Gambar ­­1.5** *3**contoh kata kasar yang paling sering didapatkan*

Selain grafik diatas penulis juga mengumpulkan data dari pihak internal lapak belajar untuk lebih meyakinkan serta untuk mengetahui keinginan dari pihak belajar seperti apa dalam menganggulangi komentar kasar, pertanyaan yang penulis ajukan meliputi, menanyakan dampak komentar kasar bagi platform lapak belajar dan hasilnya 25% mengatakan akan terjadi pertengkaran, 25% mengatakan lapak belajar akan menjadi kelihatan tidak profesional dan 50% mengatakan lapak belajar menjadi tercemar.



***Gambar 1.6*** *dampak komentar kasar bagi lapak belajar*

Selain itu penulis menanyakan tentang bagaimana harapan mereka dalam menanggulangi dan mencegah komentar kasar dan hasilnya 25% mengatakan ingin cepat, 25% mengiginkan akun pengguna di tangguhkan dan 50% otomatisasi.



**Gambar 1.7** *persentase harapan menanggulangi komentar kasar*

Berdasarkan hasil survey tersebut maka penulis akan membuat sistem penyaring komentar kasar dengan menerapkan *machine learning* menggunakan pendekatan *supervised learning*. *Machine Learning* merupakan salah satu bidang yang termasuk dalam kecerdasan buatan yang dapat mempengaruhi berbagai aspek yang lain yaitu statiska, matematika dan berbagai aspek teoritikal dari komputer sains. Pada dasarnya Pembelajaran mesin (*machine learning*) pada prinsipnya bertujuan membangun model matematika untuk membantu mesin memahami data. Model matematika ini terdiri dari berbagai variabel yang bisa berubah-ubah, sehingga program yang dibentuk dengan model matematika tersebut memungkinkan program beradaptasi dengan perubahan data (Aritonang, 2021) sedangkan *supervised learning* merupakan suatu pendekatan dimana sistem dilatih terlebih dahulu agar dapat melakukan prediksi atau melakukan klasifikasi (Haumahu, 2019). *Supervised Learning* merupakan metode yang memiliki data latih yang terdiri dari pasangan input dan output yang diinginkan dan bertujuan mempelajari pemetaan antara ruang input dan output (Simeone, 2018). Sedangkan menurut (Purwanto & Darmadi 2018) Supervised Learning bertujuan untuk menemukan pola baru dalam data dengan cara menghubungkan pola data yang sudah ada dengan data yang baru.

*Supervised learning* memiliki beberapa algoritma yang bisa digunakan seperti *k – nearest neighbor*, *naïve byes*, dan *support vector machine*. Penulis telah melakukan pengujian terhadap masing – masing algoritma tersebut namun yang memiliki akurasi paling tinggi adalah *support vector machine*, untuk *k – nearest neighbor* memiliki akurasi sebear 68%, *naïve byes* 57% dan *support vector machnine* 96%.Oleh karena itu maka penulis dalam penelitian ini akan menggunakan algoritma *support vector machine* sebagai algoritma untuk mengklasifikasi jenis komentar apakah termasuk kasar atau tidak.

Data yang diperlukan untuk melatih algoritma *support vector machine* merupakan data teks yang harus dikonversikan kedalam bentuk angka sehingga algoritma tersebut bisa mengonsumsi data yang telah penulis kumpulkan. Untuk mengubah data tersebut kedalam bentuk angka penulis menggunakan melakukan pemisahan kata dari kalimatnya berdasarkan spasi seperti contoh berikut “terimakasih atas jawabannya” menjadi kumpulan larik teks seperti berikut ini [“terimakasih”, “atas, “jawabannya”] kemudian setelah itu mengubah kata dari setiap elemen tersebut menjadi angka dari contoh diatas maka jika diubah menjadi angka akan menjadi seperti berikut ini [1, 2, 3].

Berdasarkan uraian diatas maka penulis dalam menyelesaikan masalah yang dialami oleh platform lapak belajar akan menerapkan *machine learning* dengan pendekatan *supervised learning* menggunakan algoritma *support vector machine*. Sedangkan untuk mengubah data teks menjadi angka dengan menggunakan cara merubah setiap kata dengan angka berdasarkan urutan katanya tersebut.

## **Perumusan masalah**

Berdasarkan uraian latar belakang di atas, maka dapat disimpulkan sebuah rumusan masalah dalam permasalahan ini yaitu, “bagaimana menghasilkan program penyaring komentar kasar secara otomatis menggunakan algoritma *support vector machine*”

## **Batasan Masalah**

Dalam perancangan program penyaring komentar kasar ini diperlukan batasan masalah agar menghindari pelebaran masalah yang diuraikan dalam laporan kerja praktek dan supaya sesuai dengan judul yang penulis sajikan. Penelitian yang akan dilakukan hanya dapat mengetahui jenis komentar yang dikirimkan apakah termasuk komentar kasar atau tidak

## **Ruang lingkup penelitian**

Ruang lingkup penelitian dan pembahasan dalam penelitian ini membahas tentang pembuatan program penyaring komentar kasar yang meliputi : *text preprocessing* yaitu mengubah data teks menjadi angka,cara kerja algoritma *support vector machine,* mengimplementasikan algoritma ke dalam kode, melakukan pelatihan, melakukan testing dan yang terakhir evaluasi performa algoritma dengan menggunakan metrik nilai akurasi, nilai *recall*, nilai *precision* dan nilai *f1.*

Data latih yang akan digunakan merupakan kumpulan komentar kasar dan tidak kasar dalam bahasa indonesia yang telah diberikan label secara manual 1 untuk komentar kasar dan 0 untuk komentar tidak kasar. Bahasa pemrograman yang akan digunakan adalah bahasa pemrograman *python* versi ke 3.10, selain itu juga penulis menggunakan beberapa librari seperti *scikit-learn*, *scikit - learn* adalah sebuah library python yang memberikan tampilan standar untuk mengimplementasikan algoritma pembelajaran mesin (Ekaba Bisong, 2019), selain itu menggunakan librari *pandas*, *pandas* di definisikan sebagai librari perangkat lunak yang digunakan untuk menganalisis data dalam bahasa pemrograman *python* (Ahmed dkk, 2022)

Setelah melakukan pelatihan pada algoritma support vector machine maka setelah itu penulis akan mengunggah algoritma support vector machine yang telah dilatih kedalam server. Algoritma *support vector machine* yang sudah berada didalam server akan dikonsumsi oleh client yang mana client dalam penelitian ini merupakan sebuah website yang dihubungkan dengan menggunakan *REST API*, *REST (Representional State Transfer )* merupakan seperangkat prinsip arsitektur yang melakukan transmisi data melalui antarmuka yang terstandarisasi seperti *HTTP. REST* bekerja layaknya seperti aplikasi website biasa. *Client* dapat mengirimkan permintaan kepada server melalui protokol *HTTP* dan kemudian server memberikan respons balik kepada client, dan *API* adalah antarmuka yang digunakan untuk mengakses aplikasi atau layanan dari sebuah program. *API* memungkinkan pengembang untuk memakai fungsi yang sudah ada dari aplikasi lain sehingga tidak perlu membuat ulang dari awal. Pada konteks website, *API* merupakan pemanggilan fungsi melalui *Hyper Text Transfer Protocol* (HTTP) dan mendapatkan respon berupa *Extensible Markup Language (XML) atau JavaScript Object Notation (JSON)* (Hasanuddin dkk, 2022).

Pada tahap terakhir penulis akan melakukan pengujian menggunakan metode *blackbox*, hasil yang harus dicapai adalah ketika pengguna memasukan komentar kasar data tersebut akan dikirimkan ke dalam server melaui *REST API* kemudian didalam server akan diperiksa apakah data yang dikirimkan termasuk komentar kasar atau tidak jika terindikasi sebagai komentar kasar maka *REST API* akan mengembalikan kode status http 403 (*forbidden*) yang mengindikasikan bahwa komentar tersebut dilarang tidak boleh dikirimkan setelah itu di dalam tampilan antar muka website lapak belajar akan muncul peringatan dan menampilkan pesan bahwa komentar tersebut seharusnya tidak dikirimkan, yang terakhir jika komentar yang dikirimkan tidak terindikasi sebagai komentar kasar maka *REST API* akan mengembalikan kode status *http* 200 (sukses) kemudian komentar akan disimpan ke dalam *database*.

## **Tujuan dan Manfaat Penelitian**

Berikut adalah tujuan dan manfaat penelitian yang dilakukan oleh penulis

### **Tujuan Penelitian**

Tujuan penelitian pada kerja praktek ini adalah untuk menghasilkan sebuah sistem yang mampu menyaring komentar kasar secara otomatis serta cepat agar memudahkan pihak lapak belajar dalam menyaring komentar kasar menjadikannya lebih efektif dan efisien serta membuat ruang diskusi di lapak belajar menjadi lebih baik

### **Manfaat Penelitian**

Manfaat penelitian dari penerapan algoritma support vector machine untuk menyaring komentar kasar adalah:

1. Bagi pengguna lapak belajar

Mengurangi komentar kasar serta menjadikan platform lapak belajar menjadi platform yang ramah untuk diskusi serta menghindari dari pengguna yang saling mejatuhkan.

1. Bagi pihak internal lapak belajar

Menghemat waktu yang diperlukan untuk menyaring komentar kasar serta menjadikan penyaringan komentar kasar menjadi lebih efektif dan efisien

1. Bagi penulis

Manfaat bagi penulis adalah menambah portofolio yang akan membantu pada saat melamar sebuah pekerjaan pekerjaan.

## **Metodologi Penelitian**

Berikut ini adalah bentuk penelitian yang digunakan dan metode pengumpulan data yang akan digunakan dalam penelitian

### **Bentuk Penelitian**

Dalam penelitian ini penulis menggunakan bentuk penelitian survey dimana survey yang dilakukan dalam bentuk kuisioner daring menggunakan google form. Dalam kuisioner ini terdapat beberapa pertanyaan baik untuk pihak internal lapak belajar maupun bagi penggunanya yang dimana pertanyaan didalamnya berhubungan dengan dampak, pengaruh, cara penanganan selama ini, serta harapan akan cara penganan komentar kasar dimasa yang akan datang. Dengan data tersebut penulis bisa merumuskan solusi yang tepat bagi pihak lapak belajar maupun bagi penggunanya.

Metode penelitan yang digunakan dalam penelitian ini yaitu metode penelitian *experiment research* ( penelitian eksperimen ). Menurut ibrahim dkk (2018:57) penelitian eksperimen terdiri dari enam tahapan sebagai berikut :

* 1. Memilih ide atau topik penelitian. Tahap ini merupakan hal penting dalam suatu penelitian, karena tahap ini untuk dapat menentukan keseluruhan penelitian yang akan dilakukan agar lebih terarah dan jelas.
  2. Merumuskan masalah penelitian. Selanjutnya penulis merumuskan masalah penelitian. Pada penelitian ini penulis perlu mengetahui permasalahan internal lapak belajar dalam menangani komentar kasar serta dampak komentar bagi lapak belajar serta bagi penggunanya
  3. Merumuskan hipotesis penelitian. Merupakan tahap pemecahan masalah, bagaimana cara mengatasi dan apa saja yang dibutuhkan untuk mengatasi kesulitan pihak lapak belajar dalam menangani komentar kasar
  4. Mengumpulkan data. Pada tahap ini, peneliti mengumpulkan data dengan cara survey dan observasi. Data yang didapatkan kemudian diolah dan disesuaikan dengan kebutuhan penelitian penulis
  5. Menganalisis hasil penelitian. Setelah mengumpulkan data, tahap selanjutnya adalah menganalisis hasil penelitian dengan cara mengelola data yang sudah dikumpulkan kemudian dianalisis untuk menghasilkan sebuah penelitian yang relevan
  6. Membuat kesimpulan, Pada tahap ini merupakan tahap akhir dalam metode penelitian, Penulis membuat kesimpulan berdasarkan data yang sudah dikumpulkan dan dianalisa pada tahap selanjutnya.

### **Metode Pengumpulan data**

Pengumpulan data yang digunakan peneliti dalam pembuatan laporan kerja praktek ini adalah menggunakan data primer dan data sekunder :

1. Data primer

Data primer yang akan digunakan untuk melakukan pelatihan pada algoritma *support vector machine* merupakan data yang disimpan di database dan berasal dari riwayat komentar pengguna lapak belajar pada platform diskusinya yang secara manual diberikan label 1 untuk komentar kasar dan 0 untuk komentar tidak kasar

1. Data sekunder

Data sekunder merupakan data yang diperoleh secara tidak langsung berupa publikasi terhadap dokumen yang bersumber dari situs kaggle.com serta berdasarkan dari survey, yang mana datanya sudah diberikan label untuk komentar kasar 1 dan komentar tidak kasar 0. Data yang diambil dari kaggle merupakan hasil ekstraksi dari *tweet* di platform media sosial twitter

### **Teknik Pengumpulan Data**

Teknik pengumpulan data yang digunakan oleh penulis dalam penelitian ini yatitu sebagai berikut :

1. Observasi

Teknik pengumpulan data pada penelitian ini dengan melakukan observasi, dimana penulis melihat dan memberikan label pada data yang berada di dalam database lapak belajar kemudian serta menanyakan secara langsung kepada tim lapak belajar komentar seperti apa yang termasuk komentar kasar dan tidak kasar kemudian menggabungkannya dengan data sekunder

1. Survey

Penulis melakukan penyebaran kuisioner kepada pihak *internal* dan pengguna lapak belajar tentang seperti apa contoh komentar kasar yang biasanya mereka terima

### **Metode pengembangan**

Pada penelitian ini, metode pengembangan yang digunakan dalam penelitian ini adalah metode *prototype*, Metode Prototype adalah teknik pengembangan sistem yangmenggunakan prototype untuk menggambarkan sistem sehingga klien atau pemilik sistemmempunyai gambaran jelas pada sistem yang akan dibangun oleh tim pengembang (Kristina, 2022). Berikut merupakan langkah-langkah atau tahapan dalam metode prototype :

1. *Communication* atau komunikasi dan pengumpulan data awal, yaitu analisis terhadap kebutuhan pengguna.
2. *Quick plan*, yaitu tahapan perencanaan kebutuhan
3. *Modelling Quick Design*, yaitu tahapan pembuatan desain
4. Pembentukan *prototype*, yaitu pembuatan perangkat prototype termasuk pengujian danpenyempurnaan.

**1.6.5 Metode Pengujian**

Metode pengujian yang digunakan pada penulisan penelitian ini adalah pengujian black-box. Black box Testing (Pengujian kotak hitam) yaitu bertujuan untuk

menunjukan fungsi perangkat lunak tentang cara beroperasinya, apakah pemasukan data keluaran telah berjalan sebagaimana yang diharapkan dan apakah informasi yang disimpan secara eksternal selalu dijaga kemuktahirannya (Nurajizah & Aziz, 2019). Pengujian black-box berusaha menemukan kesalahan dalam kategori sebagai berikut (Suryawan, Prasetya, Budiawan, Nip, & Nipt, 2019):

1. Fungsi – fungsi yang tidak benar atau yang tidak ada
2. Kesalahan antar muka ( *interface error* )
3. Kesalahan dalam struktur data dan akses basis data
4. Kesalahan performa ( *performance error* )
5. Kesalahan inisialisasi dan terminasi

## **Sistematika Penulisan**

Sistematika penulisan Kerja Praktek ini terdiri dari 4 (empat) Bab, dimana tiap-tiap bab nya terdiri dari beberapa sub bab. Dengan mengenai materi dalam bab-bab yang terdapat di dalam laporan kerja praktek ini, dengan maksud agar pembaca mendapatkan gambaran yang jelas mengenai isi laporan kerja praktek :

**BAB 1 Pendahuluan**

Dalam bab ini dijelaskan tentang latar belakang penelitian, permasalahan, pembatasan masalah, ruang lingkup penelitian, tujuan dan manfaat penelitian serta sistematika penulisan.

**BAB 2 Landasan Teori**

Bab ini berisi landasan teori menjabarkan teori yang terkait dengan penelitian

**BAB 3 Hasil Penelitian**

Bab ini menguraikan semua proses yang dilalui dalam menghasilkan sebuah penelitian.

**BAB 4 Penutup**

Pada bab ini menjelaskan dari hasil penelitian dan saran dari penulis.