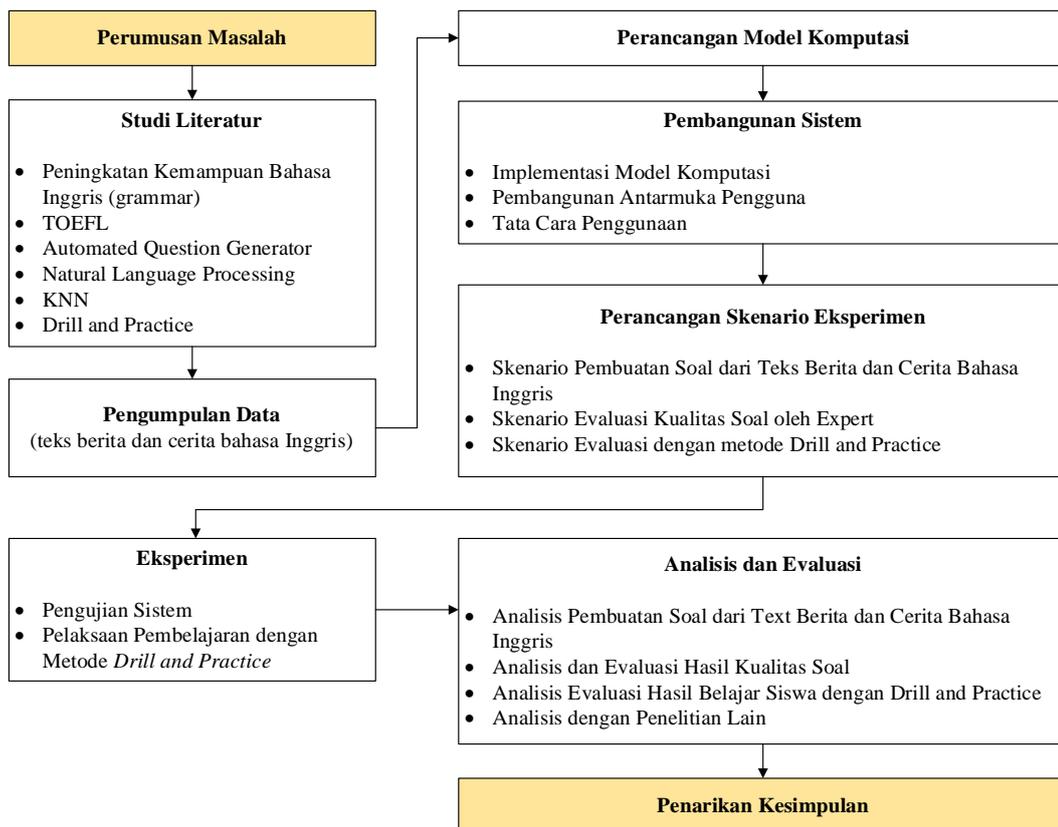


BAB III METODOLOGI PENELITIAN

3.1. Desain Penelitian

Desain penelitian adalah kerangka kerja yang disusun dalam pelaksanaan penelitian. Pada bagian ini peneliti menyusun desain penelitian dengan tujuan untuk menggambarkan tahapan sesuai dengan tujuan yang diinginkan.



Gambar 3. 1 Desain Penelitian

Gambar 3. 1 Menunjukkan tahapan yang dilakukan peneliti untuk membuat sistem *automated question generator* untuk tipe soal *sentence completion* dalam upaya peningkatan kemampuan bahasa Inggris pada tingkat SMP. Tahapan tersebut diawali dengan perumusan masalah yaitu merumuskan mengapa penelitian ini harus dilakukan dan solusi yang ditawarkan, tahap studi literatur yang berisi sumber teori, metode dan teknik penelitian yang telah dilakukan pada proses pembuatan sistem *automated question generator*. Tahap pengumpulan dilakukan dengan mengumpulkan teks berita bahasa Inggris dan teks cerita bahasa Inggris (*narrative teks*). Tahap perancangan model komputasi dan pembangunan sistem menjelaskan

alur pada saat pengerjaan sistem. Sistem *automated question generator* menerapkan teknik NLP dan KNN untuk menghasilkan yang sesuai. Tahap perancangan skenario eksperimen menjelaskan prosedur untuk melakukan eksperimen, tahap eksperimen merupakan tahapan utama dalam menentukan bagaimana sistem berjalan dengan baik. Tahap analisis dan evaluasi yang memaparkan cara peneliti menguji sistem yang dibuat dengan bantuan ahli serta bagaimana penerimaan calon pengguna terhadap sistem yang dibuat, dan yang terakhir adalah tahap penarikan kesimpulan. Tahapan ini dilakukan setelah semua tahapan berhasil dilakukan dan penulis melakukan penarikan kesimpulan dan saran yang dapat bermanfaat bagi penelitian selanjutnya dibidang ini.

1. Perumusan Masalah

Kemampuan komunikasi yang baik menjadi modal penting untuk dapat tetap bersaing baik dalam pendidikan dan pekerjaan. Di era sekarang bahasa Inggris merupakan alat komunikasi yang paling banyak digunakan karena hampir seluruh media informasi video maupun tulisan menggunakan bahasa Inggris sehingga kemampuan bahasa Inggris yang baik menjadi salah satu mata pelajar yang penting dipelajari di semua tingkat pendidikan. Latihan yang dilakukan secara berulang akan meningkatkan peluang siswa untuk meningkatkan kemampuan bahasa Inggris. Metode pembelajaran yang relevan dengan pemberian soal secara berulang untuk meningkatkan secara sistematis adalah *drill and practice*. Kebutuhan soal dalam jumlah besar mendorong banyak peneliti untuk mengembangkan sistem yang mampu membuat soal bahasa Inggris secara otomatis dalam jumlah tidak terbatas. Cerita tersebut menunjukkan adanya relevansi dengan penelitian yang telah dilakukan. Berdasarkan penelitian sebelumnya yang telah banyak dilakukan mendorong peneliti untuk berfokus pada sistem *automated question generator* pada tipe *sentence completion* dengan tema *grammar*.

2. Studi Literatur

Studi literatur mengacu pada proses mempelajari dan menganalisis topik *automated question generator*. Topik yang dipelajari terkait peningkatan kemampuan bahasa Inggris pada pemahaman *grammar*, TOEFL, *automated question generator*, *natural language processing*, KNN dan bagaimana penerapan

pada secara langsung dengan menggunakan metode *drill and practice*. studi literatur dilakukan dengan menggunakan berbagai sumber informasi, seperti jurnal ilmiah, buku, tesis, disertasi, laporan penelitian, dan dokumen lain yang relevan dengan topik *automated question generator*.

3. Pengumpulan Data

Pengumpulan data pada penelitian ini diawali dengan mencari informasi dari buku paket bahasa Inggris yang banyak digunakan di tingkat SMP. Sehingga menjadi dasar penggunaan teks bahasa Inggris yang relevan dengan kemampuan siswa. Terdapat beberapa jenis teks yang digunakan yaitu *narrative text* berupa cerita dan *information text* berupa teks berita bahasa Inggris.

Teks yang digunakan berasal dari berbagai sumber seperti buku dan *website* kurus bahasa Inggris yang telah diverifikasi oleh guru bahasa Inggris. Artikel yang diambil kemudian digunakan sebagai data uji terhadap sistem *automated question generator* yang sudah dikembangkan. Peneliti menguji apakah sistem yang telah dikembangkan dapat menghasilkan soal dengan tipe *sentence completion* dari berbagai tipe *narrative text*.

4. Perancangan Model Komputasi

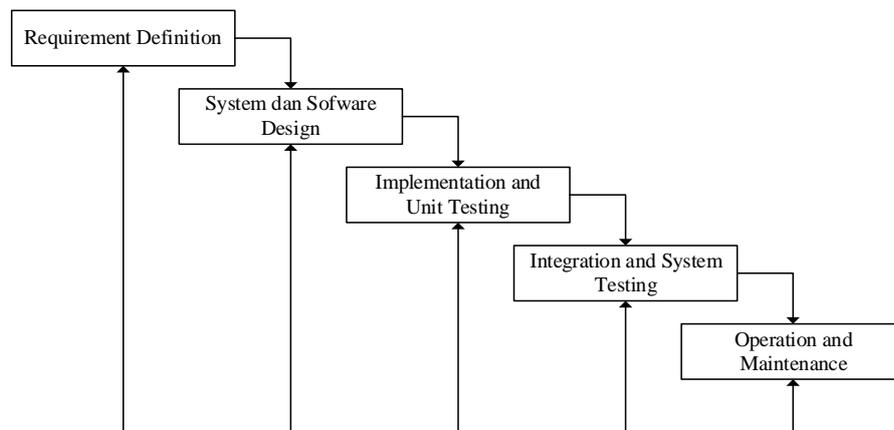
Tahap ini merupakan tahapan untuk merancang model komputasi yang akan digunakan dalam mengembangkan sistem *automated question generator*; perancangan model komputasi secara umum dibagi menjadi lima kelompok algoritma, algoritma pertama adalah *input* merupakan tahap memasukan data ke dalam sistem, data masukan terbagi menjadi dua yaitu data latih (*data training*) dan data uji (*data testing*). Algoritma kedua adalah NLP merupakan untuk mengolah teks sehingga dapat dikenali oleh sistem. Algoritma ketiga adalah pra-proses melakukan ekstraksi fitur dari setiap kata dan pos tagging dari pengolahan sebelumnya menjadi sebuah matrik. Algoritma keempat adalah proses. Ini merupakan algoritma utama mencari matriks tetangga dari data latih dan menentukan *blank position* setiap pos dengan label *true*. Pada algoritma ini juga data uji berupa kalimat yang sudah dipilih, kemudian dicari *distractor* atau pengecoh sesuai dengan soal yang akan dibuat. Terakhir merupakan hasil yaitu berupa soal bahasa Inggris dengan tipe *sentence completion*.

5. Pembangunan Sistem

Tahap ini dibagi menjadi tiga tahap, pertama adalah implementasi model komputasi dengan menerapkan metode *waterfalls*. Selanjutnya adalah pembangunan antar muka pengguna dengan melakukan rancangan antar muka atau *Graphical User Interface* (GUI) yang sesuai kebutuhan pengguna. Ketiga peneliti memberikan penjelasan kepada pengguna dalam penggunaan sistem. Secara keseluruhan sistem terbagi menjadi dua bagian yaitu bagian *backend* sebagai generator soal menggunakan bahasa Python dan tampilan antar muka menggunakan PHP Laravel.

a. Implementasi Model Komputasi

Implementasi perangkat lunak merujuk pada metode pengembangan sebuah sistem. Terdapat banyak metode implementasi yang dapat digunakan dalam pengembangan sebuah sistem. Pada penelitian ini peneliti menggunakan model *waterfall*. Model ini juga sering disebut model *sequential linier*. Adapun tahapan model *waterfall* dapat dilihat pada Gambar 3. 2.



Gambar 3. 2 Model Waterfall (Sommerville, 2011)

Pada Gambar 3. 2. Menunjukkan bahwa model *waterfall* terdapat lima tahap. Tahapan ini dapat membantu menyelesaikan dalam proses pengembangan sistem. Jika terdapat kesalahan, bisa mengacu pada tahap sebelumnya untuk melakukan perbaikan. Menurut Sommerville pada tahun 2011 menjelaskan bahwa setiap tahap adalah sebagai berikut:

1. *Requirements Definition (Analysis)*

Tahap analisis, sistem yang akan dibuat dikonsepsi sedemikian rupa. Pada tahapan analisis fitur-fitur yang akan dibuat dikonsultasikan juga dengan pengguna sistem. Hal ini akan ditetapkan sebagai spesifikasi sistem secara rinci.

2. *System and Software Design*

Software akan didesain mulai dari basis data, arsitektur, hingga *user interface*. Desain sistem dibuat berdasarkan spesifikasi yang telah ditetapkan.

3. *Implementation and Unit Testing*

Pada tahap implementasi ini bisa juga disebut tahap *coding*. Dalam tahap ini, desain perangkat lunak dibentuk sebagai satu program. Program ini kemudian akan diuji untuk memenuhi kebutuhan spesifikasi.

4. *Integration and System Testing*

Tahapan pengujian atau *testing* ini dilakukan untuk setiap bagian dalam program, apakah sudah sesuai kebutuhan spesifikasi atau belum. Selain itu, dalam tahap ini juga setiap bagian program diuji apakah sudah terintegrasi dengan baik atau belum. Jika secara keseluruhan sistem sudah terintegrasi dengan baik dan juga sesuai spesifikasi dan syarat, maka program atau sistem akan dikirim ke pengguna sistem.

5. *Operation and Maintenance*

Dalam tahap ini, sistem mulai diaplikasikan dan juga digunakan. Bisa juga pada tahap ini ditemukan beberapa *error* yang harus diperbaiki dan tidak ditemukan saat pembuatan. Pada tahap ini pun, aplikasi bisa ditambah dengan fitur baru yang sekiranya dibutuhkan pengguna.

b. Pembangunan Antar Muka Pengguna

User Interface atau disebut sebagai antar muka merupakan komponen penting dari sebuah perangkat lunak yang menjadi perantara antara mesin dengan manusia (T.Suratno, 2016). Antar muka pada penelitian ini dibangun menggunakan Laravel dengan menggunakan bahasa pemrograman PHP. Sistem dibangun dengan menerapkan beberapa bagian penting seperti, input link, input text dan juga pilihan

soal yang akan digunakan. Sisanya adalah untuk membantu membuat soal secara manual.

c. Tata Cara Penggunaan

Tahapan ini dilakukan untuk memberi arahan atau petunjuk bagi pengguna untuk menggunakan aplikasi ini. Pada sub bab 4.3 akan menjelaskan bagaimana cara menggunakan sistem secara keseluruhan, sampai sistem dapat digunakan oleh guru dan siswa. Cara berinteraksi dengan antar muka pengguna yang dibuat juga dijelaskan pada tahapan ini. Contohnya adalah cara memasukkan artikel berita kedalam *input box* dan cara memulai proses menghasilkan soal.

6. Perancangan Skenario Eksperimen

Sebelum melakukan eksperimen, hal yang pertama kali dilakukan adalah merancang skenario eksperimen yang nantinya akan dilakukan. Pada penelitian ini, penulis melakukan empat bentuk skenario eksperimen, skenario awal menggunakan 30 soal yang dihasilkan dari data testing. Dalam upaya penyesuaian materi level 9 SMP maka *judgement* dilakukan oleh guru level SMP. Maka dipilihlah dua *expert* Bahasa Inggris untuk memberikan evaluasi.

a. Skenario Menghasilkan Soal dari Teks Berita dan Teks Cerita

Skenario pertama yang dilakukan adalah menghasilkan soal dari data teks berita dan juga teks cerita. Tujuannya adalah untuk menguji sistem yang telah dibuat dapat menghasilkan soal dari text dengan dua jenis teks yang berbeda, yaitu *information text* (teks berita) dan *narrative text* (teks cerita). Pada skenario ini peneliti menyiapkan lima jenis topik bacaan dalam jenis *information text* yaitu: *lifestyle*, *economy*, *politic*, *music*, dan *healty*. Sedangkan untuk *narrative text* menggunakan lima jenis teks yaitu *fables*, *fairy tales*, *folktales*, *legend*, dan *historical fiction*.

Sistem yang dibuat menghasilkan tiga buah soal untuk jenis teks berita dengan masing-masing dua judul sehingga menghasilkan 30 soal, sedangkan untuk jenis teks cerita yaitu menghasilkan soal yang digunakan dalam setiap kegiatan yang dilakukan dalam proses pembelajaran, seperti *pre-test*, *treatment* dan *post-test*. Terlihat pada Tabel 3. 1.

Tabel 3. 1 Skenario Menghasilkan Soal (Narrative Text)

	Fables	Farity Tales	Folktales	Legend	Historical Fiction	Jumlah
Pre-Test		13	13	14		40
Treatment		5	5	10		20
Treatment		5	5	10		20
Treatment	5	5	5	5		20
Post-Test	15	5	5	10	5	40

Soal yang dihasilkan secara langsung digunakan dalam kelas. Soal yang telah berhasil dibuat oleh sistem kemudian dipilih untuk keperluan eksperimen di kelas yaitu: *pre-test* 20 soal, masing-masing *treatment* soal 10 dan *post-test* menggunakan 20 soal dengan total 70 soal.

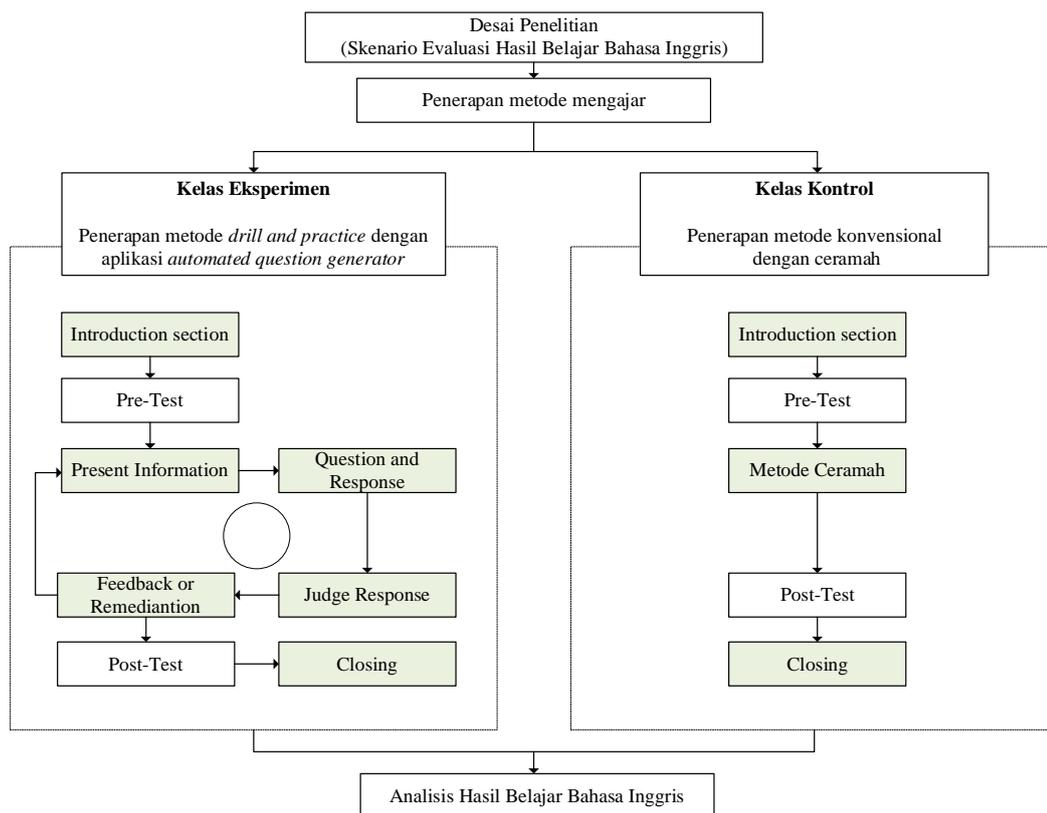
b. Skenario Evaluasi Kualitas Soal oleh Expert

Skenario ke dua yang dibuat adalah skenario evaluasi kualitas soal. Tujuan dari skenario ini adalah untuk mengetahui apakah soal yang dihasilkan oleh sistem sudah termasuk layak untuk digunakan. Pada penelitian *automatic question generation*, evaluasi yang melibatkan manusia adalah yang paling umum. Di mana, *expert* diminta untuk menilai kualitas dari soal berdasarkan kriteria tertentu (Amidei et al., 2018).

c. Skenario Evaluasi dengan Drill and Practice

Skenario ke tiga yang dibuat adalah untuk menguji kelayakan sistem secara langsung kepada siswa level SMP dengan menggunakan metode *drill and practice*. Metode *drill and practice* merupakan metode yang digunakan dalam proses pembelajaran untuk meningkatkan tingkat kedisiplinan dengan cara mengerjakan latihan secara berulang (Pacol & Patacsil, 2017). Secara umum tujuan metode *drill and practice* untuk memberikan pengetahuan dan keterampilan.

Skenario ini merupakan tahapan utama dalam seluruh rangkaian penelitian. Skenario menjadi tahap untuk menguji apakah sistem yang telah berhasil dikembangkan memberikan dampak yang baik atau tidak. Metode *Drill and practice* merupakan yang diterapkan dengan melatih secara terus menerus kepada dengan soal bahasa Inggris yang telah dihasilkan oleh sistem *automated question generator*.



Gambar 3. 3 Desain Skenario Drill and Practice

Pada penelitian ini skenario awal dilakukan dengan memisahkan kelas kontrol dan kelas eksperimen, penelitian dilakukan di MTs Al Amanah pada level 9, yaitu pada kelas 9A (20 siswa) dan kelas (19 siswa).

a. Kelas Kontrol

Kelas kontrol merupakan kelas pembanding tanpa menggunakan metode khusus dalam penelitian. Pada penelitian ini kelas kontrol diberikan materi *grammar* secara konvensional dengan metode ceramah dan hanya memberikan pre-test, materi dan pos-test.

b. Kelas Eksperimen

Kelas eksperimen menjadi kelas uji coba terhadap metode dan sistem yang dikembangkan dalam penelitian ini. Eksperimen dilakukan dengan cara menggunakan metode *drill and practice*. Menurut (Pacol & Patacsil, 2017) tahapan tahapannya adalah sebagai berikut:

1. Introduction Section

Tahap pengenalan dimulai dengan memperkenalkan materi atau konsep kepada siswa. Tahap ini menjelaskan secara keseluruhan setiap target dan

tujuan dari materi yang akan dipelajari secara singkat. Tahap ini juga memberikan gambaran tentang pemilihan materi yang akan dipelajari.

2. *Present Information*

Informasi atau materi disampaikan kepada siswa secara menyeluruh dengan jelas dan sistematis. Penyampaian materi bisa dilakukan dengan berbagai cara seperti presentasi langsung, bahan bacaan, atau sumber daya pembelajaran lainnya. Pada penelitian materi bahasa Inggris untuk materi *grammar* sebagai materi yang akan dipelajari.

3. *Questions and Response*

Siswa diajak untuk menjawab serangkaian pertanyaan atau melakukan latihan terkait dengan materi yang baru dipelajari. Tahap ini menjadi tahap utama dalam metode *drill and practice* yaitu melakukan dengan melakukan latihan berulang guna untuk meningkatkan sebuah kemampuan yang diharapkan. Tahap ini juga ini bertujuan untuk memastikan pemahaman dan penguasaan konsep oleh peserta didik.

4. *Judge Response*

Respons dari siswa terhadap pertanyaan yang atau latihan yang diberikan dievaluasi oleh guru sehingga guru dapat menyimpulkan tentang bagaimana hasil peningkatan yang diraih setiap siswa. Tahap ini juga guru akan mengukur sejauh mana siswa memahami dan mampu menerapkan materi yang dipelajari.

5. *Feedback or Remediation*

Siswa menerima *feedback* atau umpan balik terkait kinerja yang telah dilakukan, umpan balik dari guru juga termasuk poin-poin yang perlu diperbaiki dan ditingkatkan. Jika ditemukan kesalahan atau kekurangan, siswa dapat bimbingan tambahan atau remedial untuk memperbaiki konsep yang dianggap masih perlu ditingkatkan. Tahap ini merupakan tahap di mana guru akan menentukan apakah sebuah latihan akan diberikan kembali atau tidak.

6. *Closing*

Guru akan memberikan ringkasan dari setiap materi, merinci setiap hal penting atau memberikan arahan untuk tugas berikutnya.

Metode *drill and practice* sangat menekankan pada latihan yang berulang-ulang untuk memperkuat pemahaman dan keterampilan siswa. Umpan balik yang

diberikan pada tahap *feedback or remediation* akan membantu siswa meningkatkan performa siswa.

7. Eksperimen

Tahap selanjutnya adalah tahap eksperimen. Eksperimen melibatkan data yang sebelumnya telah dikumpulkan, yaitu artikel berita bahasa Inggris dan cerita bahasa Inggris. Pada tahap ini sistem yang telah dibuat diuji untuk dapat menghasilkan soal dari berbagai macam topik artikel yang telah ditentukan pada tahap perancangan skenario eksperimen. Selanjutnya, eksperimen yang dilakukan adalah melakukan evaluasi terhadap kualitas soal yang dihasilkan oleh sistem, eksperimen kekonsistenan jawaban.

Setelah sistem dianggap layak, selanjutnya sistem diuji secara langsung kepada siswa level 9 MTs. Al Amanah Kota Bandung. Eksperimen dilakukan dengan lima pertemuan terbagi menjadi *pre-test*, tiga kali *treatment* dan *post-test*. Metode yang digunakan pada tahap eksperimen adalah dengan menggunakan metode *drill and practice* pada dua kelompok eksperimen yaitu kelas *control* dan kelas *experiment*.

8. Analisis dan Evaluasi Hasil

Setelah sistem dibuat, dan dapat menghasilkan soal *sentence completion* beserta antar muka penggunaannya. Tahap selanjutnya adalah melakukan analisis dan validasi kualitas soal yang dihasilkan dan penerimaan pengguna terhadap sistem *automated question generator* yang telah dibuat. Analisis dan validasi untuk soal yang dihasilkan dilakukan dengan bantuan *expert*. Sedangkan analisis juga dilakukan dari hasil eksperimen di MTs. Al Amanah Kota Bandung.

a. Analisis dan Hasil Evaluasi Kualitas Soal

Analisis dan hasil evaluasi kualitas soal tentunya memerlukan sebuah instrumen penelitian. Instrumen penelitian digunakan untuk mengukur sejauh mana keberhasilan sebuah penelitian dilakukan. Dengan kata lain instrumen penelitian merupakan variabel yang akan diteliti untuk keperluan sebuah penelitian (Sugiyono, 2013). Instrumen yang akan digunakan pada penelitian ini merupakan instrumen validasi soal.

Validasi soal diperlukan sebagai acuan kelayakan suatu soal yang dihasilkan oleh sistem, sehingga dibutuhkan ahli untuk meninjau sebaik apa soal yang telah

dihasilkan. Pada penelitian ini dibutuhkan ahli dalam bidang bahasa Inggris. Pada penelitian *automated question generation*, evaluasi yang melibatkan manusia adalah yang paling umum. Dimana, *expert* diminta untuk menilai kualitas dari soal berdasarkan kriteria tertentu (Amidei et al., 2018). Pengukuran kualitas soal yang dihasilkan menggunakan metrik yang digunakan oleh (Araki et al., 2016) yaitu, *answer existence*, *distractor quality*, dan *difficulty indeks*.

1. **Answer Existence (AE)**

Answer existence, yaitu mengidentifikasi tersedianya jawaban yang sesuai dari 4 opsi jawaban yang disediakan. Poin ini penulis membagi menjadi 2 tingkatan nilai, antara lain:

- a. 1 (*No*). Menandakan bahwa jawaban tidak dapat ditemukan pada pilihan jawaban
- b. 2 (*Yes*). Menandakan bahwa jawaban ada pada pilihan jawaban.

2. **Difficulty Index (DI)**

Difficulty index, yaitu menentukan seberapa sulit soal yang dihasilkan oleh sistem. Poin ini penulis membagi menjadi 3 tingkatan nilai, antara lain:

- a. 1 (*Easy*). Penilaian ketika pertanyaan yang dihasilkan dianggap mudah.
- b. 2 (*Medium*). Penilaian ketika pertanyaan yang dihasilkan dianggap cukup sulit.
- c. 3 (*Hard*). Penilaian ketika pertanyaan yang dihasilkan dianggap sangat sulit.

3. **Difficulty Quality (DQ)**

Distractor quality adalah sebuah penilaian untuk mengukur seberapa baik pilihan pengecoh dari opsi jawaban. Penulis membuat skala dua poin untuk penilaian ini sebagai berikut:

- a. 1 (*Worst*): Pilihan benar (jawaban sebenarnya) dapat mudah diidentifikasi sebagai jawaban benar.
- b. 2 (*Best*): Pilihan dengan sulit diperoleh dengan 3 pengecoh yang layak.

4. **Note**

Merupakan kolom tambahan untuk mengetahui saran pengecoh dari evaluator **jika**, pengecoh yang dibangkitkan dianggap tidak layak (*worst*).

Setiap poin yang diberikan oleh *expert* terhadap setiap parameter selanjutnya akan dijumlahkan dan dirata-ratakan. Setelah mendapatkan *score* rata-rata dari *expert* untuk keempat parameter. Kemudian, *score* tersebut dihitung persentasenya dengan menggunakan rumus *rating scale* di bawah ini.

$$P = \frac{\text{skor hasil penilaian}}{\text{skor maksimal}} \times 100\% \quad (3.1)$$

Keterangan P = Angka persentase
 Skor maksimal = Skor maksimal (*best*) pada setiap parameter

Selanjutnya hasil perhitungan menggunakan metode *rating scale* akan dikategori lima kategori dengan menggunakan skala seperti pada Tabel 3. 2.

Tabel 3. 2 Kategori Rating Scale

Skor Persentase (%)	Interpretasi
$P \leq 20 \%$	Sangat Kurang
$20 \% < P \leq 40 \%$	Kurang
$40 \% < P \leq 60 \%$	Cukup
$60\% < P \leq 80 \%$	Baik
$80\% < P \leq 100 \%$	Sangat Baik

b. Analisis Hasil Belajar Siswa dengan Drill and Practice

Analisis evaluasi hasil belajar siswa merupakan rangkaian pengujian terakhir yang dilakukan. Evaluasi ini dilakukan kepada siswa level 9 MTs Al Amanah dengan menggunakan metode *drill and practice*.

a. Analisis deskriptif

Analisis deskriptif digunakan untuk memperoleh gambaran umum mengenai data yang dihasilkan dari penelitian. Data diperoleh dari skor tes siswa akan disajikan dalam bentuk pengukuran *mean*, median dan modus dan standar deviasi untuk mendeskripsikan karakteristik variabel penelitian.

$$\bar{x} = \frac{\sum_{i=1}^n X_i}{n} \quad (3.2)$$

Keterangan \bar{x} = Skor rata-rata
 X_i = Skor perolehan

n = Jumlah data

Standar deviasi (Sugiyono, 2016) menggunakan persamaan berikut:

$$S = \sqrt{\frac{n \sum x_i^2 - (\sum x_i)^2}{n(n-1)}} \quad (3.3)$$

Adapun untuk menentukan koefisien varians menggunakan persamaan:

$$KV = \frac{S}{\bar{x}} \quad (3.4)$$

Keterangan KV = Koefisien varians

S = Standar deviasi

\bar{x} = Rara-rata sampel

Menurut Departemen Pendidikan dan Kebudayaan bahwa skor standar umum yang digunakan adalah skala lima tingkat pembagian penguasaan yang terbagi atas lima kategori. Adapun skala pembagian berdasarkan Peraturan Pemerintah Nomor 59 Tahun 2016 pada Tabel 3. 3.

Tabel 3. 3 Kategoriasi penilaian hasil belajar bahasa Inggris

No	Skor	Kategori
1	$0 \leq x \leq 52$	Sangat renda
2	$52 < x \leq 64$	Rendah
3	$64 < x \leq 76$	Sedang
4	$76 < x \leq 89$	Tinggi
5	$89 < x \leq 100$	Sangat Tinggi

b. Analisis inferensial parametrik atau non parametrik

Analisis ini digunakan untuk menguji populasi melalui data statistik. Pengujian ini dilakukan mengetahui pengaruh *treatment* terhadap kelas eksperimen dengan menggunakan metode *drill and practice*.

1. Uji Normalitas

Uji normalitas dilakukan untuk mengetahui apakah data yang diperoleh berasal dari populasi yang berdistribusi normal atau tidak (Rahmawati et al., 2013). Jika signifikan kurang dari 0,05 maka kesimpulannya data tidak berdistribusi normal, jika signifikan lebih dari 0,05 maka data berdistribusi normal. Dalam konteks uji normalitas untuk data *pre-test* dan *post-test* dengan ukuran sampel yang relatif kecil pilihan yang umumnya digunakan

adalah menggunakan uji Shapiro-Wilk. Persamaan uji normalitas dengan uji shapiro-wilk adalah sebagai berikut:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{(\sum_{i=1}^n (x_i - \bar{x})^2)} \quad (3.5)$$

Keterangan w = Nilai Uji Shapiro-Wilk
 a_i = Koefisien yang bergantung pada urutan data
 $x_{(i)}$ = Nilai data yang diurutkan ke-i
 \bar{x} = Rata-rata sampel

2. Uji Homogenitas

Uji homogenitas digunakan untuk mengetahui populasi data antara dua kelompok atau lebih data memiliki varian yang sama. Penelitian ini memiliki empat kelompok data yaitu (pre-test kelas kontrol, pre-test kelas eksperimen, post-test kelas kontrol, dan post-test kelas eksperimen) namun tidak jumlah dianggap terbatas bahkan dianggap kecil, dan berpotensi tidak memenuhi standar normalitas, sehingga untuk memenuhi kebutuhan tersebut, uji homogenitas yang digunakan adalah uji Levene.

$$w = \frac{(N - k) \sum_{j=1}^k N_j (\bar{Z}_{.j} - \bar{Z}_{..})^2}{(k - 1) \sum_{i=1}^N \sum_{j=1}^k (Z_{ij} - \bar{Z}_{.j})^2} \quad (3.6)$$

Keterangan w = Nilai Uji Levene
 N = Total data
 k = Jumlah Kelompok
 N_j = Jumlah data dalam kelompok ke-j
 $\bar{Z}_{.j}$ = Rata-rata dalam kelompok ke-j
 $\bar{Z}_{..}$ = Rata-rata umum seluruh data
 \bar{Z}_{ij} = Standar deviasi dari rata-rata kelompok dalam kelompok ke-j

Kriteria pengujian hipotesis menurut (Uyanto, 2006) jika nilai signifikansi $\geq 0,05$ maka kedua kelas mempunyai varians yang sama

atau homogen. Namun jika nilai signifikansi $< 0,05$ maka kedua kelas memiliki nilai varians yang tidak sama atau tidak homogen.

3. Uji Hipotesis

Uji hipotesis dilakukan dengan dua kondisi jika data populasi berdistribusi normal dan populasi homogen maka dilakukan dengan pengujian statistik parametrik yaitu dengan uji-t (*independent sample T test*).

a. Uji statistik parametrik (uji t)

Uji t merupakan uji perbandingan dua sampel digunakan untuk membandingkan (membedakan) apakah kedua data (variabel) sama atau berbeda. Menurut (Sugiyono, 2016) “jika jumlah anggota sampel $n_1 = n_2$ dan varians homogen, maka dapat digunakan rumus *t-test* baik *seperated varians* maupun *polled varians*. Uji t dengan persamaan *seperated varians* yaitu:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (3.7)$$

Keterangan	n_1	=	Jumlah data pertama (kelas eksperimen A)
	n_2	=	Jumlah data kedua (kelas eksperimen B)
	\bar{x}_1	=	Nilai rata-rata hitung (data pertama)
	\bar{x}_2	=	Nilai rata-rata hitung (data kedua)
	s_1^2	=	Varians data pertama
	s_2^2	=	Varians data kedua

Pengambilan keputusan dengan membandingkan nilai t_{hitung} dengan t_{tabel} dengan taraf signifikansi $\alpha = 0,05$ dengan $df (n_1 + n_2 - 2)$. Jika $t_{hitung} \leq t_{tabel} \leq t_{hitung}$ maka, H_0 diterima dan H_a ditolak artinya tidak ada perbedaan dari kedua sampel. Sebaliknya jika H_0 ditolak dan H_a diterima maka terdapat perbedaan antara kedua sampel.

b. Uji statistik non parametrik (uji Mann-Whitney)

Pada kondisi lain data yang dihasilkan tidak berdistribusi normal maka pengujian statistik dilakukan uji statistik non parametrik dengan

Mann-Whitney (Quraissy & Madya, 2021). Menurut Sugiono, uji mann-whitney atau uji-U berfungsi sebagai alternatif uji-t jika prasyarat parametrik tidak terpenuhi. Persamaan umum uji-U adalah:

$$U = \min (U_1, U_2) \quad (3.8)$$

Keterangan U_1 = Jumlah peringkat 1
 U_2 = Jumlah peringkat 2
 U = Hasil uji Mann-Whitney

Untuk menghitung U_1 dan U_2 pada masing-masing sampel, yaitu pada sampel pertama dengan N_1 menggunakan persamaan:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \quad (3.9)$$

Sedangkan dari sampel kedua dengan N_2 menggunakan persamaan:

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 \quad (3.10)$$

Keterangan n_1 = Jumlah sampel 1
 n_2 = Jumlah sampel 2
 R_1 = Jumlah rangking pada sampel n_1
 R_2 = Jumlah rangking pada sampel n_2

Uji hipotesis adalah H_0 tidak terdapat perbedaan yang signifikan antara kelompok pertama dan kedua, sedangkan H_a adalah terdapat perbedaan yang signifikan antara kelompok pertama dan kelompok kedua.

Pengambilan keputusan dilakukan dengan cara mencari nilai p-value dari tabel uji Mann whitney. Jika nilai p-value < dari tingkat signifikansi (0,05) maka H_0 ditolak dan H_a diterima, dapat disimpulkan bahwa terdapat perbedaan antara kelompok pertama dan kelompok kedua, jika nilai p-value > dari tingkat signifikansi (0,05) maka H_0 diterima dan H_a ditolak, dapat disimpulkan bahwa tidak terdapat perbedaan yang signifikan antara kelompok pertama dan kelompok kedua.

4. Uji N-Gain

Pada Uji N-Gain, data skor tes diperoleh dari tes yang dilakukan sebanyak dua kali yaitu *pre-test* dan *post-test* untuk kelompok eksperimen maupun kontrol.

$$g = \frac{\text{post_test} - \text{pre_test}}{\text{skor_maks} - \text{pre_test}} \quad (3.5)$$

Uji N-Gain (Normalized Gain) adalah metode statistik yang digunakan untuk mengukur peningkatan atau perubahan skor di antara dua waktu pengukuran. Untuk mengetahui kualitas peningkatan kemampuan pemecahan masalah siswa dilakukan interpretasi terhadap indeks gain. Menurut (Nurdiana, 2011) kriteria indeks gain adalah sebagai berikut:

Tabel 3. 4 Interpretasi N-Gain

Gain	Interpretasi
$g > 0,7$	Tinggi
$0,3 \leq g \leq 0,7$	Sedang
$g \leq 0,3$	Rendang

Menurut Hake pada tahun 1999 melakukan kategori untuk nilai tafsir tentang efektivitas nilai N-gain.

Tabel 3. 5 Skala Efektivitas N-Gain

Persentase (%)	Keterangan
< 40	Tidak Efektif
40 – 55	Kurang Efektif
56 – 75	Cukup Efektif
>76	Efektif

c. Analisis Perbandingan dengan Penelitian Sebelumnya

Analisis ini bertujuan untuk melihat perbedaan dan hubungan dengan penelitian tentang AQG dan *dill and practice*. Analisis ini dibuat untuk menjadi tolak ukur sejauh mana penelitian terkait telah banyak dikembangkan. Penelitian terkait juga menjadi *literature* dalam penelitian ini. Maka dari itu, penting untuk

membuat perbandingan untuk memperjelas keterlibatan penulis dengan penelitian lain yang serupa.

9. Penarikan Kesimpulan

Pada tahap ini, setelah semua tahapan berhasil dilakukan penulis melakukan penarikan kesimpulan dan saran yang dapat bermanfaat bagi penelitian selanjutnya pada bidang ini. Tahapan ini terdiri dari butir-butir yang mewakili isi dari penelitian ini dari awal hingga akhir.

Terdapat butir yang menjelaskan mengenai hasil akhir dari penelitian ini, lalu pemaparan mengenai hasil analisis dan evaluasi. Selain itu, terdapat butir yang menyatakan mengenai batasan-batasan yang terdapat pada penelitian ini. Penjelasan mengenai tahapan ini dapat dilihat pada subbab 5.1.

Penarikan kesimpulan dalam penelitian ini memiliki tujuan untuk menyimpulkan hasil penelitian dan mengidentifikasi apakah hipotesis awal terbukti atau tidak. Selain itu, penarikan kesimpulan membantu dalam menyajikan temuan dan implikasi penelitian secara jelas.

Kesimpulan juga dapat memberikan sumbangan terhadap pemahaman umum dalam bidang penelitian yang telah dilakukan, dengan merangkum hasil dan temuan, penelitian dapat memberikan kontribusi pada literatur ilmiah dan membantu pengembangan pengetahuan lebih lanjut.

3.2. Metode Penelitian

Metode yang dilakukan pada penelitian ini terbagi dalam dua kelompok besar, pertama adalah metode pengembangan sistem. Metode pengembangan sistem dilakukan adalah dengan metode *waterfalls* yang telah dijelaskan pada sub bab sebelumnya. Kedua adalah metode pelaksanaan pembelajaran dengan metode *drill and practice*. Pelaksanaan pembelajaran dilakukan dengan melakukan identifikasi beberapa *point* penting yaitu:

1. Jenis Penelitian

Jenis penelitian ini termasuk penelitian kuantitatif eksperimen menggunakan pendekatan quasi eksperimen. Quasi Eksperimen merupakan pengembangan dari *true experiment design*. Jenis penelitian kuantitatif banyak menggunakan angka dari segi pengumpulan data, penafsiran dan penarikan kesimpulan.

Penelitian ini memiliki dua pembagian kelompok yaitu kelompok eksperimen dan kelompok kontrol. Perlakuan dari penelitian ini ditujukan untuk mengetahui hasil belajar siswa dengan menggunakan metode *drill and practice* pada kelas eksperimen dan metode konvensional pada kelas kontrol.

2. Populasi dan Sampel

Kebutuhan sampel yang akan diuji diambil dari populasi yang berupa siswa tingkat SMP dan MTs. Populasi sendiri merupakan suatu wilayah generalisasi yang terdiri dari subjek yang memiliki karakteristik serta kualitas tertentu yang ditetapkan oleh penulis (Sugiyoni, 2013).

Populasi dan sampel pada penelitian ini berjumlah 39 siswa level 9 dari MTs Al Amanah kota Bandung.

3. Desain Penelitian

Desain penelitian yang dilakukan pada penelitian ini yaitu *pre-test* dan *post-test non equivalent control group design*. Desain penelitian ini dapat dilihat pada

Tabel 3. 6 *Quasy Experiment Non Equivalent Control Group Design*

Kelas	Pre-Test	Treatment	Post-Test
Kontrol	P ₁	-	O ₁
Eksperimen	P ₂	X	O ₂

Keterangan P = Hasil *pre-test*

X = *Treatment*

O = Hasil *post-test*

4. Instrumen Penelitian

Instrumen penelitian yang digunakan adalah tes objektif dengan tipe pilihan ganda (*sentence completion*) sebanyak 20 soal pada saat *pre-test* dan 20 soal dengan empat pilihan jawaban pada saat *post-test*. Instrumen ini diberikan pada awal perlakuan (*pre-test*) dan akhir perlakuan (*post-test*) pada kelas eksperimen maupun kelas kontrol.