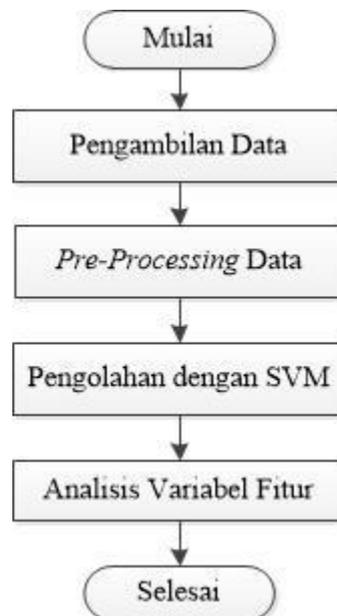


BAB III METODE PENELITIAN

3.1. Desain dan Prosedur Penelitian

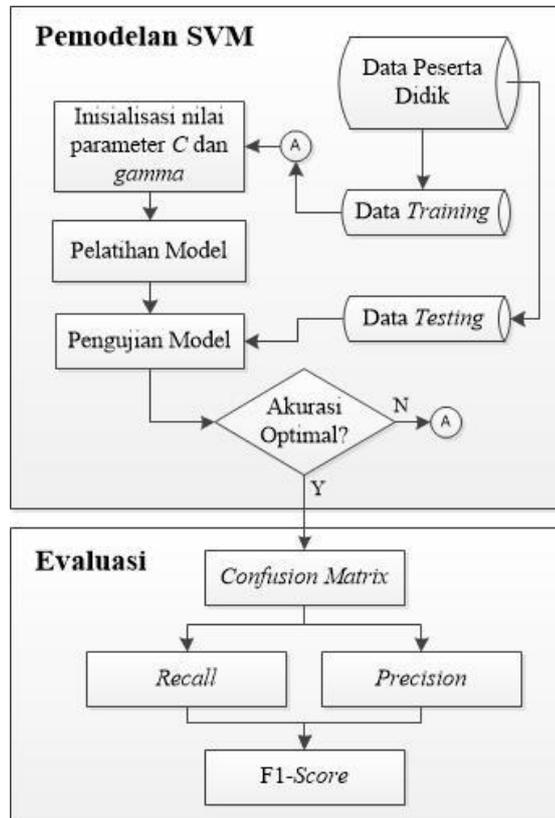
Eksperimen dilakukan untuk mengimplementasikan algoritma *Support Vector Machine* (SVM), sehingga model memiliki kinerja yang optimal dalam melakukan prediksi tingkat pemahaman materi peserta didik dalam pembelajaran *online*. Proses implementasi tersebut akan disimulasikan menggunakan aplikasi Jupyter serta bantuan *library* scikit-learn yang terdapat dalam bahasa pemrograman Python 3.5. Adapun prosedur penelitian yang dilakukan oleh penulis dapat dilihat pada **Gambar 3.1** berikut.



Gambar 3. 1 Flowchart Prosedur Penelitian

Seperti yang terlihat pada **Gambar 3.1** tahap pertama yang dilakukan penulis yaitu pengambilan data yang terdapat pada situs Kaggle. Setelah itu, merupakan tahap data *preprocessing*. *Preprocessing* yaitu mengubah data asli menjadi data yang sesuai, sehingga dapat diolah pada algoritma *Machine Learning* tertentu (Amrieh et al., 2016). Selanjutnya, yaitu tahap pengolahan data menggunakan algoritma *Support Vector Machine* (SVM) dan tahap terakhir yaitu menganalisis setiap fitur pada data penelitian yang berpengaruh terhadap hasil prediksi.

Proses pengolahan data yang lebih detail sehingga dapat memperoleh hasil terhadap tingkat pemahaman materi peserta didik dalam pembelajaran *online* menggunakan algoritma *Support Vector Machine* (SVM) dapat dilihat pada **Gambar 3.2**.



Gambar 3. 2 Flowchart Pengolahan Data

Setelah data melalui tahap *preprocessing*, data dibagi menjadi 2 bagian yaitu data *training* dan data *testing*. Telah dibahas pada BAB II bahwa data *training* digunakan untuk proses pelatihan model dengan cara mempelajari pola pada data tersebut, sedangkan data *testing* digunakan sebagai evaluasi terhadap model yang telah dilatih. Proses pelatihan dilakukan berberapa kali menggunakan kombinasi nilai dari parameter *C* dan *gamma* yang berbeda-beda untuk mendapatkan model dengan tingkat akurasi terbaik. Selanjutnya, dilakukan evaluasi terhadap model yang telah dilatih menggunakan metode *Confusion Matrix*, sehingga didapatkan *recall*, *precision*, dan *F1-Score* untuk masing-masing kelas sebagai representasi dari kinerja model *Support Vector Machine* (SVM) yang telah dibangun.

3.2. Objek Penelitian

Objek pada penelitian merupakan 480 data peserta didik yang melakukan kursus *online* (*Secondary School*) menggunakan *Learning Management System* (LMS) Kalboard 360. Para peserta didik ini berasal dari 14 Negara, namun didominasi dari Kuwait dan Jordan masing-masing sebesar 37.5% dan 36.67%, sedangkan negara lainnya dibawah 4.58%. Kemudian, topik kursus yang diambil oleh peserta didik juga berbeda-beda, diantaranya English 9.38%, Spanish 5.21%, French 13.54%, Arabic 12.29%, IT 19.79%, Math 4.38%, Chemistry 5%, Biology 6.25%, Science 10.62%, History 3.96%, Quran 4.58%, dan Geology 5%.

3.3. Data Penelitian

Data yang digunakan pada penelitian ini merupakan data sekunder, artinya data tidak dikumpulkan secara langsung oleh penulis, melainkan data yang dikumpulkan oleh (Amrieh et al., 2016), menggunakan alat pelacak aktivitas pelajar, disebut *Experient API* (xAPI) sehingga memungkinkan untuk memantau tindakan peserta didik dalam proses pembelajaran *online*. Kemudian data ini dimuat pada situs Kaggle (<https://www.kaggle.com/aljarah/xAPI-Edu-Data>), sehingga semua orang dapat mengaksesnya (*Open Data*) untuk keperluan lain, khususnya pada bidang *Machine Learning*. Beberapa bagian dari data ini dapat dilihat pada **Tabel 3.1** berikut.

Tabel 3. 1 Dataset Penelitian

No	gender	PlaceofBirth	VisitedResources	StudentAbsenceDays	Class
1	M	Kuwait	16	Under-7	M
2	M	Kuwait	50	Under-7	H
⋮	⋮	⋮	⋮	⋮	⋮
479	F	Jordan	17	Above-7	L
480	F	Jordan	14	Above-7	L

Terlihat pada **Tabel 3.1** target pada data ini adalah *Class* yang diklasifikasikan ke dalam tiga kategori berdasarkan nilai ujian akhir dari masing-masing peserta didik. Yaitu, *Low* (L) untuk rentang nilai 0 – 69, *Mid* (M) 70 – 89, dan *High* (H) 90 – 100. Hal ini berkaitan dengan permasalahan penelitian, karena penilaian/ujian dapat dijadikan sebagai gambaran terhadap

tingkat pemahaman materi peserta didik selama menempuh di suatu lembaga pendidikan, semakin besar nilai hasil ujian peserta didik, maka dapat diartikan bahwa peserta didik tersebut memahami sebagian besar materi yang telah diajarkan, begitupun sebaliknya semakin kecil nilai hasil ujian peserta didik, maka dapat diartikan bahwa peserta didik tersebut hanya memahami sebagian kecil materi yang telah diajarkan (Salamah, 2018).

Kemudian terdapat 15 fitur yang dikategorikan menjadi 4 bagian, yaitu: (1) fitur demografis, seperti jenis kelamin dan tempat lahir, (2) fitur akademik, (3) fitur perilaku peserta didik saat mengikuti pembelajaran, (4) fitur tambahan. Fitur-fitur ini disajikan pada **Tabel 3.2** berikut.

Tabel 3. 2 Fitur Data Penelitian

Kategori Fitur	Fitur	Deskripsi
Demografis	gender	Jenis kelamin peserta didik Pilihan: F atau M
	PlaceofBirth	Tempat lahir peserta didik Pilihan: Egypt, Iran, Iraq, Jordan, Kuwait, Lybia, Morocco, Palestine, SaudiArabia, Syria, Tunis, USA, Lebanon, atau Venezuela
Akademik	StageID	Tingkatan pendidikan peserta didik Pilihan: lowerlevel, MiddleSchool, atau HighSchool
	GradeID	Tingkatan kelas peserta didik (Kelas 2 sd Kelas 12) Numerik: 2-12
	SectionID	Ruangan Kelas peserta didik Pilihan: A, B, atau C
	Semester	Semester pembelajaran (Pertama atau Kedua) Pilihan: F atau S

	Topic	Topik pembelajaran yang diambil peserta didik Pilihan: English, Spanish, French, Arabic, IT, Math, Chemistry, Biology, Science, History, Quran, Geology
Perilaku Peserta Didik	raisedhands	Mengangkat tangan pada saat pembelajaran Numerik: 1-100
	VisitedResources	Berapa kali peserta didik mengunjungi konten pembelajaran Numerik: 1-100
	AnnouncementsView	Berapa kali peserta didik memeriksa pengumuman baru Numerik: 1-100
	Discussion	Berapa kali peserta didik berpartisipasi dalam kelompok diskusi Numerik: 1-100
Tambahan	ParentResponsible	Orang tua yang bertanggung jawab untuk peserta didik Pilihan: Mother atau Father
	ParentAnsweringSurvey	Orang tua peserta didik menjawab survei yang disediakan dari sekolah atau tidak Pilihan: Yes atau No
	ParentschoolSatisfaction	Tingkat kepuasan orang tua peserta didik kepada sekolah Pilihan: Good atau Bad
	StudentAbsenceDays	Jumlah ketidakhadiran peserta didik Pilihan: Under-7 atau Above-7