

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pemerintah melakukan ujian yang dinamakan *Assessment* Nasional Berbasis Komputer (ANBK) pada setiap jenjang Pendidikan mulai dari tingkat SD, SMP, dan SMA. Hasil yang diperoleh dipublikasikan melalui rapor pendidikan. Rapor Pendidikan adalah platform yang menyediakan data laporan hasil evaluasi sistem pendidikan [1]. Informasi yang terdapat di dalam rapor pendidikan memberikan gambaran umum terkait dengan kemampuan numerasi siswa di sekolah.

Berdasarkan hasil rapor pendidikan secara nasional, pada tahun 2022 46,67% murid memiliki kompetensi numerasi di atas minimum, naik 16,01 dari 2021 (30,66%) [2]. Berdasarkan hasil tersebut, secara skala nasional hasil numerasi siswa pada jenjang SD masuk dalam kategori sedang yang artinya 40% - 70% peserta didik telah mencapai kompetensi minimum. Berikut hasil rapor pendidikan dari beberapa sekolah dari salah satu yayasan pendidikan di Tangerang.

**Tabel 1. 1 Hasil Rapor Pendidikan 2022**

<b>Unit Sekolah</b>	<b>Capaian Numerasi Siswa</b>
<b>SLH AMBON</b>	40% - 70% peserta didik telah mencapai kompetensi minimum untuk numerasi namun perlu upaya mendorong lebih banyak peserta didik dalam mencapai kompetensi minimum.
<b>SLH BANJAR AGUNG</b>	40% - 70% peserta didik telah mencapai kompetensi minimum untuk numerasi namun perlu upaya mendorong lebih banyak peserta didik dalam mencapai kompetensi minimum.
<b>SLH CURUG</b>	Sebagian besar peserta didik telah mencapai batas kompetensi minimum untuk numerasi.
<b>SLH GUNUNG AGUNG</b>	Kurang dari 40% peserta didik telah mencapai kompetensi minimum untuk numerasi perlu upaya mendorong peserta didik dalam mencapai kompetensi minimum.
<b>SLH JATI AGUNG</b>	40% - 70% peserta didik telah mencapai kompetensi minimum untuk numerasi namun perlu upaya mendorong lebih banyak peserta didik dalam mencapai kompetensi minimum.
<b>SLH KAMP. HARAPAN</b>	Kurang dari 40% peserta didik telah mencapai kompetensi minimum untuk numerasi perlu upaya mendorong peserta didik dalam mencapai kompetensi minimum.
<b>SLH KOJA</b>	Sebagian besar peserta didik telah mencapai batas kompetensi minimum untuk numerasi.
<b>SLH KUPANG</b>	40% - 70% peserta didik telah mencapai kompetensi minimum untuk numerasi namun perlu upaya mendorong lebih banyak peserta didik dalam mencapai kompetensi minimum.
<b>SLH LABUAN BAJO</b>	Sebagian besar peserta didik telah mencapai batas kompetensi minimum untuk numerasi.
<b>SLH MEDAN</b>	40% - 70% peserta didik telah mencapai kompetensi minimum untuk numerasi namun perlu upaya mendorong lebih banyak peserta didik dalam mencapai kompetensi minimum.
<b>SLH NIAS</b>	Sebagian besar peserta didik telah mencapai batas kompetensi minimum untuk numerasi.
<b>SLH PALOPO</b>	40% - 70% peserta didik telah mencapai kompetensi minimum untuk numerasi namun perlu upaya mendorong lebih banyak peserta didik dalam mencapai kompetensi minimum.
<b>SLH SANGIHE</b>	40% - 70% peserta didik telah mencapai kompetensi minimum untuk numerasi namun perlu upaya mendorong lebih banyak peserta didik dalam mencapai kompetensi minimum.
<b>SLH TOMOHON</b>	Sebagian besar peserta didik telah mencapai batas kompetensi minimum untuk numerasi.
<b>SLH TORAJA</b>	40% - 70% peserta didik telah mencapai kompetensi minimum untuk numerasi namun perlu upaya mendorong lebih banyak peserta didik dalam mencapai kompetensi minimum.

Terdapat kekurangan di dalam rapor pendidikan tersebut. Pada hasil pemaparan numerasi di jenjang pendidikan SD, pemerintah hanya menggunakan sampel yaitu siswa kelas 4 dari total keseluruhan yang terdapat di sekolah. Identifikasi numerasi siswa dapat dimanfaatkan oleh guru untuk melakukan pembelajaran yang dipersonalisasi. Pembelajaran yang dipersonalisasi adalah pendekatan pendidikan yang berpusat pada kebutuhan, minat, tujuan, dan

kemampuan setiap pelajar [3]. Untuk menerapkan pembelajaran yang dipersonalisasi, diperlukan informasi mengenai siswa, salah satunya adalah hasil identifikasi kemampuan numerasi. Oleh karena itu, tes diagnostik dibutuhkan untuk membantu guru dalam merancang pembelajaran yang sesuai. Tes diagnostik dapat memberikan informasi terkait kekurangan peserta didik secara individu [4]. Selain itu, tes diagnostik dapat memberikan gambaran terkait permasalahan dan eksplorasi kesulitan yang dihadapi oleh siswa pada saat proses pembelajaran [5]. Melalui tes diagnostik, diharapkan memberikan gambaran terkait kondisi pengetahuan siswa pada saat melaksanakan pembelajaran.

Dibutuhkan metode yang tepat untuk dapat melakukan identifikasi kemampuan numerasi siswa khususnya pada jenjang pendidikan SD. Dengan *machine learning* dapat membantu dalam pengelompokan siswa dalam proses pembelajaran. *Machine learning* dapat membantu mengklasifikasikan atau mengelompokkan siswa dalam kemampuan akademik [6]. Pemanfaatan *machine learning (ML)* memiliki potensi besar dalam mengembangkan sistem peringatan dini untuk mengidentifikasi siswa yang berisiko dalam perkembangan akademik [7]. Hal sesuai dengan pandangan AWS yang menyatakan penggunaan *ML* dapat dimanfaatkan untuk melakukan identifikasi siswa yang mengalami kendala dalam pembelajaran dan menetapkan tujuan intervensi yang tepat [8]. Selain itu, *ML* adalah aspek kecerdasan buatan yang beroperasi dalam sistem komputasi dapat belajar dari data dan membuat kesimpulan sehingga memudahkan pekerjaan dalam melakukan identifikasi dari informasi yang dimiliki [9]. Dalam sebuah penelitian *ML* dapat melakukan identifikasi hasil belajar siswa untuk mengurangi risiko kegagalan siswa dalam mengikuti proses pembelajaran [10]. Fokus

penelitian tersebut adalah menciptakan model hibrida dengan metode *ensemble stacking*, menggabungkan algoritma seperti *Naïve Bayes*, *Random Forest*, dan lainnya. Kinerja model diukur dengan berbagai metrik, dan hasilnya dievaluasi menggunakan kumpulan data demografis dan akademik siswa. Metode termasuk *stratified k-fold cross-validation* dan optimasi *hyperparameter* untuk meningkatkan kinerja model. Uji coba dengan dua kombinasi data menunjukkan akurasi model hibrida mencapai 94,8% dengan data demografi dan akademik, serta 98,4% hanya dengan data akademik. Hasilnya memberikan landasan untuk prediksi kinerja siswa yang terhadap identifikasi berisiko lebih dini, memungkinkan guru memberikan bantuan tambahan kepada siswa yang memperlihatkan kinerja rendah.

Selain itu, terdapat model *machine learning* yang dapat melakukan prediksi dengan tingkat akurasi yang baik yaitu dengan model *k-Nearest Neighbor* (k-NN). Penelitian Kurniadi menunjukkan bahwa variasi dan jumlah data pelatihan yang tepat dapat meningkatkan kinerja algoritma k-NN, dengan akurasi terbaik sebesar 95,83 persen dalam memprediksi mahasiswa yang berpotensi mendapatkan beasiswa [11]. Sejalan dengan hasil yang diperoleh, Alban dan Mauricio memperoleh hasil yang baik dengan menggunakan model *Neural Network*. Hasil penelitian tersebut menunjukkan bahwa tidak ada perbedaan signifikan dalam tingkat akurasi model yang diusulkan, yaitu 96,3% untuk *multilayer perceptron* dan 96,8% untuk fungsi basis radial [12].

Dibutuhkan metode yang efektif untuk melakukan identifikasi. Hasil penelitian Sanzana menunjukkan penggunaan *Random Forest (RF)*, *Classification and Regression Tree (CART)* dapat digunakan untuk mengidentifikasi kelompok

siswa [13]. Selain itu, penelitian oleh Theephoowiang dan Chaowicharat yang menggunakan *Naïve Bayes*, *Neural Network*, regresi, dan *Support Vector Machine* (SVM) menemukan bahwa model regresi memberikan hasil prediktor terbaik, dengan rata-rata kesalahan absolut sekitar 0.57 dari skala 1-5 antara nilai prediksi *machine learning* dan label dari pakar. [14]. Hal ini dapat dipergunakan untuk melakukan identifikasi terkait dengan hasil yang akan diperoleh oleh siswa.

Dalam penelitian yang dilakukan oleh Dangi menjelaskan kinerja algoritma *Naïve Bayes* dalam mengklasifikasikan data ke dalam kategori tertentu [15]. Data yang diperoleh dapat digunakan untuk menemukan pola yang bermakna bagi siswa untuk melakukan identifikasi terkait dengan hasil belajar yang diperoleh siswa [15]. Berdasarkan penelitian lainnya yang telah dilakukan, menunjukkan hasil eksperimen evaluasi akademik mahasiswa dengan menggunakan Klasifikasi *Naïve Bayes* (NBC) tersebut mencapai tingkat akurasi sebesar 76.79% [16]. Namun terdapat beberapa penelitian lain yang mengungkapkan bahwa model algoritma pembelajaran ensemble *AdaBoost* lebih unggul dibandingkan *single classifier algorithm* [17].

Penelitian ini menggunakan *Logistic Regression*, *Naïve Bayes*, *Neural Network*, dan *Random Forest* untuk melakukan identifikasi permasalahan numerasi dari hasil tes diagnostik siswa. Identifikasi yang diperoleh diharapkan memperoleh hasil yang detail melalui metode visualisasi nomogram berbasis *Logistic Regression* dan *Naïve Bayes* melalui aplikasi Orange.

## **1.2 Identifikasi Permasalahan**

Mempertimbangkan latar belakang di atas maka dapat dinyatakan definisi masalah adalah menganalisis hasil tes diagnostik yang mengukur kemampuan

numerasi siswa (baik atau buruk) dari hasil tes diagnostik yang dilakukan pada periode Mei – Juni 2023. Evaluasi kemampuan numerasi siswa telah dilakukan dengan 2 cara berikut:

- a. Melakukan perbandingan model *Logistic Regression*, *Naïve Bayes*, *Neural Network*, dan *Random Forest* terhadap kelompok data berdasarkan parameter.
- b. Perbandingan visualisasi data dengan menggunakan nomogram antara *Naïve Bayes* dan *Logistic Regression* dalam melakukan identifikasi kemampuan numerasi siswa.

### 1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah :

- a. Sumber data penelitian ini menggunakan hasil tes diagnostik kemampuan numerasi.
- b. Analisis yang dilakukan menggunakan aplikasi orange.
- c. Pada kasus ini menggunakan metode *Logistic Regression* dan *Naïve Bayes* serta nomogram untuk visualisasi dalam pengelompokan data.
- d. Periode data yang digunakan dari bulan Mei – Juni 2023
- e. Sumber data berasal dari <https://pemantik.or.id/>

### 1.4 Rumusan Masalah

Pada bagian ini dijelaskan secara lebih terinci dari apa yang akan diselesaikan pada batasan masalah sehingga dikatakan masalah selesai dikerjakan.

- a. Bagaimana menerapkan *Logistic Regression*, *Naïve Bayes*, *Neural Network*, dan *Random Forest* dalam melakukan prediksi hasil tes diagnostik siswa?



- b. Bagaimana metode nomogram dapat melakukan pengelompokan kemampuan numerasi siswa berdasarkan hasil tes diagnostik.

## 1.5 Tujuan Penelitian

Tujuan dari penelitian ini adalah

1. Menerapkan model *machine learning Logistic Regression, Naïve Bayes, Neural Network, dan Random Forest* untuk melakukan prediksi hasil tes diagnostik siswa sekolah dasar.
2. Membandingkan hasil kinerja model *machine learning* yang efektif antara *Logistic Regression, Naïve Bayes, Neural Network, dan Random Forest*.
3. Menerapkan visual *nomogram* dalam melakukan pengelompokan kemampuan numerasi siswa berdasarkan hasil tes diagnostik.

## 1.6 Sistematika Penulisan

Penulisan dalam penelitian ini terbagi menjadi sedikitnya dalam lima bab. Setiap bab mempunyai bahasan mengenai tujuan dan isi yang berbeda-beda. Adapun sistematikanya sebagai berikut:

Bab I Pendahuluan. Bab ini membahas tentang gambaran secara singkat mengenai latar belakang masalah mengapa penelitian ini dilakukan sampai pada tujuan penelitian

Bab II Kajian Teori. Bab ini membahas tentang teori-teori yang akan digunakan atau penelitian yang sudah dilakukan terkait dengan rumusan permasalahan yang dibicarakan pada Bab 1. Bagian ini merupakan bagian kunci untuk menentukan metode yang akan dipakai pada bagian selanjutnya.

Bab III Metodologi Penelitian. Bab ini berisi tentang rancangan penelitian dan atau rancangan pengujian

Bab IV Hasil dan Pembahasan. Menguraikan hasil dari penelitian yang telah dilakukan dan melakukan argumentasi atas apa yang dihasilkan dengan melampirkan karya ilmiah yang sudah atau akan dipublikasi.

Bab V Kesimpulan dan Saran. Pada bab ini menjelaskan tentang kesimpulan berdasarkan hasil dari penelitian yang diperoleh, serta saran-saran konstruktif yang perlu dikembangkan untuk penelitian berikutnya sehingga penelitian berikutnya menjadi lebih baik.

Di bagian akhir dari penulisan ini dilampirkan daftar Pustaka, lampiran-lampiran serta daftar riwayat hidup peneliti.

