

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang

*Natural Language Processing* (NLP) adalah bagian dari kecerdasan buatan yang berfokus pada interpretasi komputer dalam bahasa Indonesia yang banyak diterapkan mulai dari sekolah hingga perguruan tinggi yang sedang melaksanakan pembelajaran online. Contoh dari penggunaan NLP yang telah diterapkan oleh dunia pendidikan adalah *Automated Essay Scoring* (AES) (Ramesh & Sanampudi, 2021).

*Automated Essay Scoring* (AES) sendiri telah digunakan oleh beberapa peneliti, salah satunya pada tahun 2019 (Maulidya, 2019), menerapkan sistem *Automated Essay Scoring* Berbasis Web, tujuannya untuk mengevaluasi sistem pembelajaran secara daring, adapun metode yang telah digunakan oleh (Maulidya, 2019) yaitu metode *Latent Semantic Analysis* (LSA), fungsi dari LSA sendiri disini untuk mengekstrak dan mempresentasikan suatu kalimat dengan perhitungan secara matematis.

(Thamrin et al., 2021) melakukan penelitian *Automated Essay Scoring* (AES) pada tahun 2021 di Universitas Muhammadiyah Kalimantan Timur yang berjudul "*Text Classification and Similarity Algorithms in Essay Grading*". Dengan tujuan untuk melakukan penilaian essay dengan teknik klasifikasi dan algoritma kesamaan menggunakan *Support Vector Machine* (SVM) dengan *Term Frequency - Inverse Document Frequency* (TF-IDF) dan LSA, hasil dari penelitiannya menggunakan SVM dan TF-IDF akan dihitung menggunakan *Root Mean Square Error* (RMSE).

Pada tahun yang sama (Verdikha et al., 2021) melakukan penelitian yang berjudul "*Regression and Oversampling Method for Indonesian Language Automated Essay Scoring*". Dengan tujuan mengurangi resiko kesalahan model pembelajaran yang terdapat pada penelitian Thamrin sebelumnya. Dalam penelitiannya, Verdikha menggunakan tiga metode *regresi*, yaitu *Support Vector*

*Regression (SVR) dengan kernel RBF, Logistic Regression (LR), MLP-Regression (MLP-R) dan mencoba untuk menggunakan oversample dataset menggunakan (SMOTE) agar dapat menyeimbangkan data-data penelitiannya.*

Dalam penelitian (Verdikha et al., 2021) untuk melakukan perbandingan sama seperti yang dilakukan oleh Thamrin, yaitu menggunakan nilai evaluasi *Root Mean Square Error (RMSE)*, yang mana nilai yang mendekati NOL adalah nilai yang terbaik. Dalam penelitian Verdikha dengan menggunakan metode SVR nilai yang dihasilkan RMSE lebih baik dibandingkan penelitian (Thamrin et al., 2021) yaitu 2,16. Hasil ini menunjukkan bahwa metode SVR menggunakan *kernel RBF* lebih baik dari pada klasifikasi metode untuk model pembelajaran yang sudah diteliti sejauh ini.

Metode SVR memiliki beberapa parameter *kernel*, yaitu *kernel Polynomial, kernel Sigmoid, kernel Radial Basis Function (RBF)*, dan *Kernel liniear*, pada penelitian ini peneliti memilih untuk menggunakan metode SVR dengan *Kernel liniear* karena pada penelitian sebelumnya yang dilakukan oleh (Pangarkar et al., 2020) terbukti bahwa dengan menggunakan *kernel liniear* mendapatkan nilai RMSE terbaik dibandingkan *kernel-kernel* lainnya.

Penulis akan melakukan metode komparasi untuk membandingkan nilai dari penelitian sebelumnya yang dilakukan oleh Verdhika dengan menggunakan metode ekstraksi fitur TF-IDF dan metode SVR dengan parameter *kernel liniear*. Dalam mengukur tingkat kesalahan hasil dari penelitian di sini menggunakan nilai evaluasi yang sama dengan penelitian Verdhika, yakni *Root Mean Square Error (RMSE)*.

## **1.2. Rumusan Masalah**

Berdasarkan latar belakang yang sudah disusun oleh penulis, maka disimpulkan rumusan masalah pada penelitian ini adalah:

1. Berapa hasil nilai evaluasi RMSE pada AES Bahasa Indonesia menggunakan metode SVR dengan *Kernel liniear*.
2. Bagaimana hasil evaluasi RMSE jika dibandingkan dengan penelitian sebelumnya yang dilakukan oleh (Verdikha et al., 2021).

### **1.3. Tujuan Penelitian**

Berdasarkan rumusan masalah yang sudah dibuat oleh penulis maka dapat disimpulkan tujuan dari penelitian ini adalah:

1. Mencari nilai evaluasi dari RMSE pada AES Bahasa Indonesia menggunakan metode SVR dengan *Kernel liniear*.
2. Membandingkan hasil nilai evaluasi dengan penelitian sebelumnya yang dilakukan oleh (Verdikha 2021).