

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Teknologi sangat berperan penting dalam mengelola informasi terhadap kehidupan. Salah satunya penggunaan teknologi dalam bidang pendidikan. Salah satu penggunaan teknologi sistem cerdas khususnya pembelajaran mesin untuk mempelajari dan mengevaluasi perkembangan nilai hasil pembelajaran berdasarkan tujuan adalah NLP. *Natural Language Processing* (NLP) adalah pembelajaran dari kecerdasan buatan yang memiliki hubungan interaksi antara komputer dengan manusia menggunakan bahasa alami. Dalam penerapannya, NLP dapat bekerja di beberapa bidang, salah satunya di bidang pendidikan yang menggunakan teknik penilaian AES. *Auto Essay Scoring* (AES) adalah teknologi komputer yang bekerja secara otomatis dalam melakukan penilaian jawaban esai (Shermis & Burstein, 2003).

Pada Penelitian AES yang dilakukan oleh Thamrin (2021), Peneliti telah mengumpulkan dataset berbentuk teks jawaban dan nilai masing-masing teks jawaban. Dataset diambil dari salah satu soal ujian online dari nilai Mata Kuliah Dasar Umum (MKDU) Bahasa Indonesia di Universitas Muhammadiyah Kalimantan Timur (UMKT) dengan hasil jawaban esai berjumlah 1648 data. Peneliti menggunakan *Feature Extraction* (TF-IDF) dan *Lantent Semantic Analysis* (LSA) dengan menggunakan jawaban esai yang telah di *preprocessing* dengan menggunakan teknik klasifikasi dan algoritma similaritas yang menunjukkan bahwa hasil metode klasifikasi *Support Machine Vector* (SVM) dengan TF-IDF menghasilkan nilai *Root Mean Square Error* (RMSE) sebesar 2,73 (Thamrin et al., 2021).

Pada tahun yang sama, Verdikha (2021) melanjutkan penelitian Thamrin (2021) dengan data dan metode yang sama bertujuan untuk mengurangi skala *error*. metode yang digunakan adalah *Support Vector Regression* (SVR), *Logistic Regression* (LR), dan *MLP Regression* (MLP-R). Hasilnya penggunaan SVR dengan

parameter kernel *Radial Basic Function* (RBF) mendapatkan nilai RMSE yang lebih baik dengan skor 2,16 (Verdikha et al., 2021).

Dari dua sumber data penelitian diatas, menghasilkan nilai evaluasi RMSE yang mungkin masih dapat ditingkatkan. Salah satu tindakan lain dalam meningkatkan performa model pembelajaran yakni menggunakan metode SVR dengan parameter *kernel* yang berbeda. Salah satunya yaitu menggunakan *kernel sigmoid*. Nambiar (2019), pada penelitiannya menggunakan dataset dari *Refrence Energy Disaggregation Data Set* (REDD) yang disediakan oleh MIT berupa data frekuensi rendah untuk kategori dua rumah dengan tipe 4 dan 5. Metode yang digunakan adalah SVR (sigmoid), kNN, DTR, Fully Connected NN, LSTM, dan SVR (rbf) menghasilkan nilai evaluasi RMSE dengan metode SVR (sigmoid) sebesar 49.23 lebih baik dari penggunaan SVR (rbf) (Nambiar et al., 2019).

Dengan demikian, penulis akan mengajukan penelitian dengan metode regresi SVR menggunakan parameter *kernel sigmoid* yang sudah dibuktikan pada penelitian Nambiar (2019) sebelumnya, lalu membandingkan hasil penelitian berdasarkan nilai RMSE yang diperoleh dari penelitian Verdikha (2021) untuk mengukur tingkat kesalahan nilai evaluasi tersebut.

## **1.2 Rumusan Masalah**

Bagaimana hasil evaluasi RMSE pada AES berbahasa Indonesia menggunakan metode SVR dengan parameter *kernel sigmoid* jika dibandingkan dengan hasil penelitian Verdikha (2021).

## **1.3 Tujuan Penelitian**

Mencari nilai evaluasi RMSE pada AES berbahasa Indonesia menggunakan parameter SVR dengan *kernel sigmoid* dan membandingkannya dengan penelitian Verdikha (2021) yang menggunakan *kernel* RBF.