

ABSTRAK

Pendidikan di tingkat formal membutuhkan aspek pengujian untuk memastikan ilmu pengetahuan diserap oleh siswa. Sebagai upaya untuk menjamin siswa memahami materi yang diajarkan, para guru mengadakan uji kompetensi yang dibagi dalam beberapa termin setiap semesternya. Uji kompetensi terdiri dari berbagai macam tipe soal salah satunya yaitu pilihan ganda. Soal yang dihimpun oleh guru dalam platform *e-learning* membutuhkan pemilahan sehingga dapat mengukur tingkat intelektual siswa dan mencakup keterampilan berpikir.

Evaluasi dalam pemilahan soal didapati banyaknya soal yang belum tersaring sesuai dengan capaian kompetensinya. Pada penelitian ini, proses penyaringan yang melibatkan algoritma *Revised Bloom's Taxonomy* (RBT) diimplementasikan secara otomatis dengan menggunakan salah satu metode *machine learning* yaitu *Stochastic Gradient Descent* (SGD) pada soal sejarah tingkat SMA. Dikarenakan terdapat *imbalance* pada *dataset*, penelitian ini juga akan membandingkan antara *dataset* yang menerapkan dan yang tidak menerapkan metode *oversampling* SMOTE.

Hasil klasifikasi dari penelitian ini diperoleh melalui implementasi algoritma SGD dengan tools *python*. Pada *dataset* tanpa SMOTE, algoritma SGD memiliki skor rata-rata *K-Fold* yaitu 62%. Pada *dataset* dengan SMOTE, algoritma SGD memiliki skor rata-rata 93%. Adapun hasil akurasi dari *confusion matrix* menunjukkan algoritma SGD pada *dataset* tanpa SMOTE memiliki performa 62%, sedangkan algoritma SGD pada *dataset* dengan SMOTE mendapatkan performa 94%.

Hasil yang diperoleh melalui proses *K-Fold Cross Validation* dan *confusion matrix* tersebut menunjukkan bahwa *dataset* dengan menggunakan *oversampling* memiliki hasil yang lebih baik dibanding dengan *dataset* tanpa menggunakan *oversampling*.

Kata kunci—**klasifikasi pertanyaan, RBT, sejarah SMA, SGD, SMOTE**