

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kecelakaan tambang menjadi suatu peristiwa yang berdampak serius terhadap keselamatan manusia dan lingkungan. Operasional pertambangan memiliki risiko kecelakaan dikarenakan kesalahan manusia, kegagalan dari berbagai faktor seperti peralatan atau kondisi geologis yang tidak terduga. Kecelakaan tambang dapat mengakibatkan kerugian nyawa, cedera serius, dan kerusakan lingkungan yang signifikan. Jumlah kecelakaan tambang pada tahun 2023 terdapat kecelakaan ringan sebanyak 104 kejadian, kecelakaan berat sebanyak 65 kejadian, dan kematian sebanyak 48 kejadian (Kementerian ESDM, 2023).

PT ANTAM Unit Bisnis Pertambangan Emas (UBPE) Pongkor merupakan salah satu unit perusahaan milik PT ANTAM yang berlokasi di Pongkor, Jawa Barat. Keselamatan kerja di tambang emas Pongkor dikelola oleh Departemen *Safety*, Biro *Health, Safety, Security, dan Environment* (HSSE). Pengelolaan dijalankan dengan berbagai program terkait keselamatan dan kesehatan kerja (K3). Salah satu program yang dibuat yaitu program laporan kondisi tidak aman dan tindakan tidak aman (KTA/TTA). Program pelaporan KTA/TTA dilaksanakan sebagai sistem pelaporan bahaya di lingkungan kerja PT ANTAM Pongkor yang dapat menerima laporan 24 jam. Setiap pekerja PT ANTAM dan perusahaan mitra diwajibkan melaporkan setiap kejadian bahaya yang berpotensi kecelakaan baik disebabkan oleh lingkungan ataupun tindakan pekerja yang tidak sesuai standar operasional pekerja.

Laporan KTA/TTA berisikan deskripsi kejadian dari suatu kecelakaan atau kejadian yang berpotensi kecelakaan. Setiap laporan yang masuk dilakukan analisis dan diklasifikasikan secara manual menjadi kategori “Kondisi Tidak Aman” dan “Tindakan Tidak Aman” oleh staff Departemen *Safety*. Jumlah pelaporan di tahun 2023 yang diterima Departemen *Safety* sebanyak 3507 laporan yang terdiri dari 2503 laporan kondisi tidak aman dan 594 laporan tindakan tidak aman. Pelaporan dapat dilakukan setiap hari selama 24 jam sehingga proses analisis harus dilakukan secara cepat agar tidak menambah beban kerja staff Departemen *Safety*. *Text mining* dapat digunakan sebagai pendekatan untuk memudahkan proses kerja staff Departemen *Safety* dalam mengklasifikasikan data teks seperti data di dalam Program Laporan KTA/TTA.

Text mining sebagai teknik dapat digunakan untuk menganalisis dan mengklasifikasikan data yang terkandung dalam laporan keselamatan (*safety report*) (Tanguy et al., 2016). *Text mining* atau *text data mining* yaitu proses mengubah data teks menjadi data numerik sehingga dapat dianalisis menggunakan algoritma *data mining* (Miner et al., 2012). *Support Vector Machine* (SVM), *Naive Bayes* (NB) dan *K-Nearest Neighbor* (KNN) dengan pemrosesan bahasa alami (NLP) sebagai algoritma data mining. Ada pendekatan *deep learning* yang populer untuk menganalisis data sekuensial seperti data teks yaitu arsitektur *Recurrent Neural Network* yang memiliki versi modifikasi algoritma yaitu *Long Short-Term Memory*. Algoritma LSTM mempelajari ketergantungan jangka panjang yang tidak dapat dilakukan oleh RNN dengan dua komponen penting yaitu *cell state* sebagai jalur informasi dan *gate* sebagai regulator yang menghapus dan menambahkan

informasi ke dalam *cell state* sehingga mampu mengingat informasi jangka panjang (Hochreiter, 1997).

Klasifikasi teks menggunakan algoritma LSTM sudah banyak dilakukan dalam penelitian sebelumnya seperti penelitian yang dilakukan Zhang et al. (2021) mengenai klasifikasi laporan investigasi kecelakaan transportasi penerbangan sipil kelas A pada tahun 2019 hingga 2021 di Amerika Serikat. Laporan investigasi diklasifikasikan *incident* atau *accident* oleh Zhang et al. (2021) menggunakan LSTM dan menghasilkan akurasi sebesar 73%. Penelitian yang dilakukan oleh Bahad et al. (2019) tentang pendeteksian berita bohong pada 2 *dataset* yang berbeda menggunakan LSTM dan optimasi RMSprop memperoleh akurasi 91,48% pada *dataset* pertama dan akurasi 98,63% pada *dataset* kedua. Algoritma LSTM juga diterapkan dalam penelitian Witanto et al. (2022) yang melakukan analisis sentimen *review* film terhadap *IMDb Largest Review* dengan membandingkan metode optimasi Adam dan RMSProp. Hasil penelitian tersebut diperoleh akurasi yang dihasilkan RMSProp lebih tinggi dibandingkan Adam dengan *hyperparameter batch size* 128 dan *epoch* 5.

Ketidakseimbangan kelas (*class imbalance*) pada data menjadi permasalahan umum dalam analisis klasifikasi. Data yang tidak seimbang menjadi masalah karena jumlah data dari satu kelas melebihi data dari kelas yang lain (Padurariu & Breaban, 2019). Ketidakseimbangan kelas dapat dilakukan penanganan dengan *resampling* yaitu pengurangan jumlah data pada kelas mayoritas atau penambahan data pada kelas minoritas untuk menyeimbangkan data (Khushi et al., 2021). Metode *resampling* dengan pendekatan sederhana yaitu teknik *random sampling*,

teknik yang membangkitkan data secara acak yang terdiri dari *Random Undersampling* dan *Random Oversampling* (Alex et al., 2022).

Algoritma LSTM telah diterapkan dalam klasifikasi teks yang mengalami ketidakseimbangan kelas seperti penelitian yang dilakukan oleh Nadda et al. (2020) dengan algoritma LSTM untuk pendeteksian demam berdarah pada teks berbahasa Thailand yang terjadi ketidakseimbangan kelas. Pengukuran kinerja dalam penelitian Nadda et al. (2020) dilakukan berdasarkan nilai *Gmean* atau *geometric mean* yang diperoleh nilai *Gmean* tertinggi sebesar 77%. Penelitian Bagui et al. (2021) tentang klasifikasi email penipuan pada data email yang dikumpulkan dari beberapa industri di Amerika Serikat seperti asuransi, hukum, hotel, sekolahan, dan perbankan yang berjumlah 18.366 email yang terbagi atas 3.416 email penipuan dan 14.950 email biasa menunjukkan bahwa algoritma LSTM memiliki akurasi mencapai 95% dan penelitian tersebut menjelaskan bahwa penggunaan *dropout* mampu meningkatkan akurasi. Penelitian yang dilakukan oleh Humam dan Laksito (2023) menerapkan algoritma LSTM untuk melakukan analisis data teks dengan penambahan metode *Random Oversampling*. Hasil penelitian Humam dan Laksito (2023) dijelaskan bahwa metode *Random Oversampling* mampu meningkatkan kinerja model LSTM.

Penelitian ini menerapkan algoritma *Long-Short Term Memory* untuk mengklasifikasikan pelaporan bahaya terkait kondisi tidak aman dan tindakan tidak aman di PT ANTAM UBPE Pongkor dengan *resampling* untuk penanganan ketidakseimbangan kelas. Nilai *Gmean* yang dihasilkan oleh algoritma *Long-Short Term Memory* (LSTM) sebagai evaluasi kinerja model. Data yang digunakan data yang tidak seimbang dan data yang diseimbangkan dengan teknik *resampling*

Random Oversampling dan *Random Undersampling*. Peneliti memiliki ketertarikan untuk melakukan penelitian berjudul “Klasifikasi Teks pada Pelaporan Bahaya PT ANTAM Pongkor Menggunakan Algoritma *Long-Short Term Memory* (LSTM)”. Penelitian ini diharapkan dapat menghasilkan model terbaik untuk memudahkan klasifikasi laporan kondisi tidak aman dan tindakan tidak aman PT ANTAM UBPE Pongkor.

1.2 Rumusan Masalah

Rumusan masalah berdasarkan latar belakang tersebut pada penelitian ini sebagai berikut:

1. Bagaimana prosedur pemodelan *Long-Short Term Memory* yang optimal untuk klasifikasi pelaporan bahaya PT ANTAM UBPE Pongkor?
2. Bagaimana pengaruh klasifikasi yang dihasilkan algoritma *Long-Short Term Memory* dengan *Random Oversampling* dan *Random Undersampling* dalam klasifikasi pelaporan bahaya PT ANTAM UBPE Pongkor?
3. Bagaimana gambaran dari seluruh isi laporan kondisi tidak aman dan tindakan tidak aman PT ANTAM UBPE Pongkor periode 2 Januari 2023 hingga 31 Desember 2023?

1.3 Batasan Masalah

Batasan-batasan masalah pada penelitian ini sebagai berikut:

1. Data yang digunakan bersumber dari laporan kondisi tidak aman dan tindakan tidak aman Departemen *Safety* PT ANTAM UBPE Pongkor.
2. Data penelitian diambil dari laporan pada periode 2 Januari 2023 hingga 31 Desember 2023 yang berjumlah 3097 laporan dengan pembagian kelas 2503 kondisi tidak aman dan 594 tindakan tidak aman.

3. Klasifikasi teks menggunakan algoritma *Long-Short Term Memory* dengan optimasi RMSProp.
4. Penanganan ketidakseimbangan kelas dilakukan dengan metode *Random Oversampling* dan *Random Undersampling* pada data latih.
5. Penentuan kelas dilakukan oleh staf Departemen *safety* PT ANTAM UBPE Pongkor.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini diantaranya yaitu:

1. Memperoleh model *Long-Short Term Memory* yang optimal untuk klasifikasi teks pelaporan bahaya kondisi tidak aman dan tindakan tidak aman PT ANTAM UBPE Pongkor.
2. Meoptimalkan hasil pemodelan algoritma *Long-Short Term Memory* dalam mempelajari data yang tidak seimbang dengan *Random Oversampling* dan *Random Undersampling* dalam klasifikasi pelaporan bahaya PT ANTAM UBPE Pongkor.
3. Memperoleh informasi kondisi tidak aman dan tindakan tidak aman yang sering terjadi di PT ANTAM UBPE Pongkor periode 2 Januari 2023 hingga 31 Desember 2023.