

BAB 1

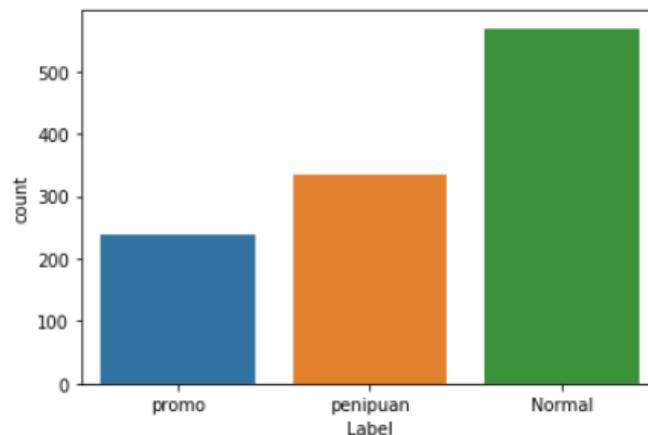
PENDAHULUAN

1.1 Latar Belakang

SMS (*Short Message Service*) adalah suatu fasilitas untuk mengirim dan menerima suatu pesan singkat berupa teks melalui telepon selular.

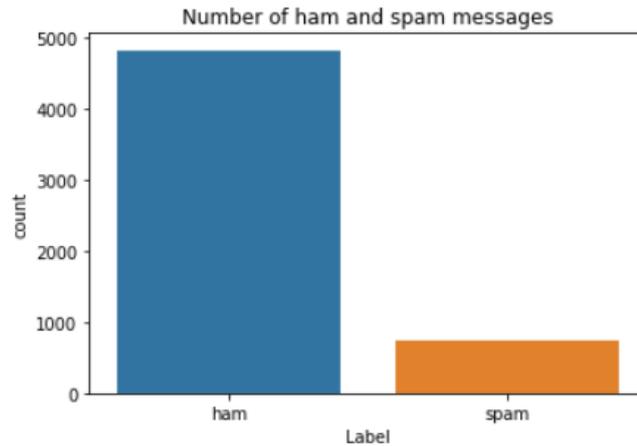
SMS tidak memiliki *header* pada setiap pesan yang dikirim seperti pada email. Pesan teks juga penuh dengan kata singkatan dan memiliki lebih sedikit Bahasa formal dari apa yang diharapkan. Semua faktor ini dapat mengakibatkan penurunan serius dalam kinerja banyak algoritma penyaringan *spam* terkenal apabila diterapkan.

Dalam penelitian ini pesan yang masuk akan diklasifikasikan, untuk dataset Bahasa Inggris memiliki 2 kategori yaitu *spam* atau bukan *spam (ham)* dan untuk dataset Bahasa Indonesia memiliki 3 kategori yaitu kategori normal, promo dan penipuan dengan menggunakan algoritma LSTM (*Long Short Term Memory*) dan RNN(*Recurrent Neural Network*).



Gambar 1. Label perbandingan bahasa Indonesia

Pada gambar 1 data yang digunakan adalah data SMS berbahasa Indonesia terlihat perbandingan SMS normal (sms yang bukan berisi tentang penipuan, promo) lebih tinggi dibandingkan dengan promo dan penipuan disini bisa disimpulkan bahwa SMS masih efektif penggunaannya, mungkin karena lebih murah dan juga tidak memerlukan koneksi internet



Gambar 2. Label perbandingan bahasa Inggris

Pada gambar 2 data yang digunakan adalah data SMS berbahasa Inggris dan terlihat perbandingan antara *ham* dan *spam*, *ham* lebih tinggi dibandingkan dengan *spam* dan bisa disimpulkan bahwa SMS di Inggris pun masih efektif penggunaannya.

Dalam penelitian ini metode LSTM dan RNN akan digunakan untuk mengklasifikasi jenis SMS berdasarkan dataset yang berisi pesan Bahasa Indonesia dan Bahasa Inggris. Diharapkan metode LSTM dan RNN ini dapat meningkatkan performa klasifikasi.

1.2 Perumusan Masalah

Adapun rumusan masalah yang diangkat pada penulisan proposal tugas akhir ini adalah sebagai berikut :

1. Menentukan metode eksperimen untuk mendapatkan parameter model LSTM dan RNN yang tepat.
2. Menentukan metode evaluasi untuk mengukur performa algoritma LSTM dan RNN dalam melakukan klasifikasi SMS.

1.3 Batasan Masalah

Ruang lingkup permasalahan pada penelitian ini hanya akan dibatasi pada beberapa hal seperti di bawah ini :

- Data yang digunakan merupakan data berbahasa Indonesia dan Bahasa Inggris. Data Bahasa Inggris berasal dari Kaggle dan data Bahasa Indonesia merupakan data yang dikumpulkan dari Yudi Wibiseno.

1.4 Tujuan Penelitian

Adapun tujuan dari penulisan tugas akhir ini adalah sebagai berikut:

1. Melakukan eksperimen parameter model LSTM dan RNN.
2. Melakukan evaluasi performa model algoritma LSTM dan RNN pada klasifikasi SMS.

1.5 Manfaat yang diharapkan

Manfaat dari penelitian ini adalah sebagai berikut :

1. Model klasifikasi yang dihasilkan pada penelitian ini dapat digunakan untuk melakukan klasifikasi SMS secara otomatis.