



Ilham Aulady Miftahurrizqy  
Kholilul Rohman  
Satrio Murdo Laras

202010370311462  
202010370311518  
202010370311522

- Machine Learning Course
- Supervised by Setio Basuki, Ph.D.

## Introduction

Media sosial merupakan salah satu sarana komunikasi antara satu pengguna dengan pengguna lain. Twitter merupakan salah satu platform media sosial yang bisa digunakan tidak hanya berkomunikasi, tetapi juga bisa digunakan sebagai promosi produk, sharing informasi, membuat komunitas, dan lain lain. Akan tetapi, tidak semua isi dari media sosial mengandung hal positif. Terdapat konten negatif yang berisi umpatan, cyberbullying, SARA, pornografi, dan sebagainya. Penelitian ini bertujuan untuk membandingkan ketiga algoritma Machine Learning, yaitu algoritma classic, Pre-trained Model, dan Deep Learning, dalam konteks klasifikasi kalimat toxic dan sara pada Twitter. Dengan memahami perbandingan kinerja dari ketiga algoritma ini, diharapkan penelitian ini dapat memberikan wawasan yang berharga dalam memilih algoritma yang paling efektif untuk tujuan klasifikasi kalimat toxic dan sara.

## Dataset

Dataset kami berasal dari scraping dari Twitter. dengan total 3430 instance. Label dataset kami berupa angka, yaitu 0 untuk Positif, 1 untuk Toxic, 2 untuk Kritik dan 3 untuk Others. Pelabelan dataset dilakukan secara manual oleh kelompok kami.

## Feature

- Kami menggunakan Bi-Gram, TF-IDF dan Word2Vec embedding untuk algoritma klasik.
- Untuk model Pre-Trained kami menggunakan model IndoBERTweet.
- Optimasi menggunakan Adam pada model LSTM.

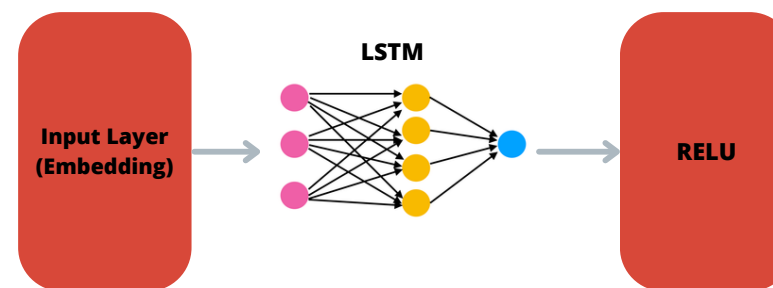
## Model

- **Decision Tree**

$$E(S) = \sum_{i=1}^c -p_i \log_2 p_i$$

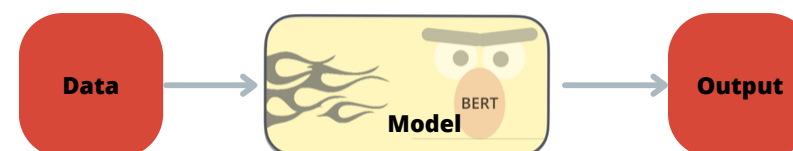
- **LSTM**

$$\begin{aligned} f_t &= \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \\ i_t &= \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \\ o_t &= \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \\ c_t &= f_t \circ c_{t-1} + i_t \circ \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \\ h_t &= o_t \circ \sigma_h(c_t) \end{aligned}$$



- **BERT (IndoBERTweet)**

BERT adalah model NLP dengan arsitektur Transformers yang dikemukakan oleh Google.



## Result

Algoritma		F1-score
Decision Tree	Bi-Gram	0.63
	TF-IDF	0.63
	Word2Vec	0.56
LSTM		0.65
Pre-Trained IndoBERTweet		0.71

## Discussion

- Model tidak begitu akurat karena data yang sangat imbalance.
- Model BERT sangat bagus dalam mengklasifikasi, tetapi memakan banyak waktu dalam training.
- Decision Tree dengan embedding Word2Vec memiliki akurasi terendah karena banyaknya slang dan kalimat typo yang menyebabkan model kurang memahami konteks dari teks

## Future Works

- Mempreprocess kata slang dan typo pada dataset.
- Melakukan oversampling atau undersampling pada data. Hal ini bisa dilakukan dengan cara EDA (Easy Data Augmentation)