

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Bank Syariah Indonesia (BSI) diduga terkena serangan siber pada bulan Mei 2023 hingga memberi dampak gangguan layanan yang berkepanjangan. Layanan perbankan yang tidak dapat diakses, bocornya data pribadi, hingga hilangnya tabungan merupakan kerugian akibat gangguan tersebut[1]. Dampak dari keadaan ini yaitu berpengaruh terhadap kepercayaan nasabah pada bank. Kepercayaan nasabah dapat dilihat melalui tweet warganet pada sosial media X.

Sosial media X merupakan sosial media dengan pengguna aktif yang cukup banyak, oleh karena itu sosial media ini sering digunakan untuk menyampaikan gagasan serta opini mereka terhadap isu-isu yang terjadi. Isu kebocoran data pada Bank BSI juga menjadi perbincangan yang hangat pada cuitan para netizen di sosial media X. Netizen yang membuat cuitan sebagian besar adalah nasabah yang terdampak akibat kebocoran data tersebut. Sehingga, cuitan netizen pada sosial media merupakan hal yang penting untuk dilihat sebagai pertimbangan seberapa besar keresahan para nasabah terhadap layanan dari suatu lembaga atau instansi. Keresahan warganet yang dituangkan pada sosial media X dapat dijadikan sebuah Analisis Sentimen[2]. Analisis Sentimen adalah pendekatan yang dapat memberikan wawasan berharga tentang bagaimana masyarakat merespon peristiwa atau isu tertentu di media sosial. Dalam topik kebocoran data bank BSI, analisis sentimen pada sosial media X dapat membantu bank BSI untuk memahami perasaan dan pandangan nasabah terkait insiden ini[3].

Penelitian tentang analisis sentimen telah banyak dilakukan sebelumnya. Metode-metode yang telah digunakan diantaranya metode naïve Bayes, LSTM, Neural Network, SVM, BERT dan lainnya. Salah satu metode yang sangat efektif dalam analisis sentimen teks adalah BERT (*Bidirectional Encoder Representations from Transformers*). BERT adalah model bahasa yang mampu memahami konteks dan nuansa dalam teks dengan tingkat akurasi yang tinggi. Dengan menggunakan BERT, kita dapat mengurai pesan teks pada sosial media X yang sering kali singkat, adanya bahasa informal dan memahami perasaan dan pandangan yang terkandung dalamnya dengan lebih baik[4]. Keunggulan utama BERT dibandingkan dengan model lain yaitu BERT dapat memproses kalimat secara biarah, artinya model bisa melihat konteks dari kedua arah (kiri ke kanan dan kanan ke kiri), sehingga dapat menangkap nuansa sentimen yang kompleks yang terdapat didalam kalimat. Untuk melakukan tugas pemrosesan bahasa alami berbahasa Indonesia, juga terdapat IndoBERT yang merupakan model BERT yang dilatih untuk memahami bahasa Indonesia. IndoBERT dapat mengenali sentimen atau perasaan yang terkandung dalam teks. IndoBERT dilatih dengan lebih dari 220 juta kata bahasa Indonesia, diantaranya 74 juta dari

Wikipedia Indonesia, 55 juta dari artikel berita Kompas, tempo dan Liputan6 serta 90 juta dari Web Corpus Indonesia[5].

Penelitian yang menggunakan IndoBERT oleh Muhammad Fadhil Mubaraq serta Warih Maharani pada penelitian tersebut didapatkan hasil terbaik rata-rata F1-Score sebesar 95.6% dengan menggunakan learning rate  $2e-5(0.00002)$  dengan ukuran batch 16[6]. Juga penelitian oleh Priyan Fadhil Supriyadi serta Yuliant Sibaioni menghasilkan nilai terbaik rata-rata F1-Score 98% dengan menggunakan learning rate  $3e-5(0.00003)$  dan nilai epoch 4[7].

Oleh sebab itu, penelitian ini memiliki tujuan guna dianalisisnya sentimen terhadap komentar-komentar terkait dengan kebocoran data Bank Syariah Indonesia(BSI) di sosial media X menggunakan IndoBERT sebagai metode utama. Peneliti bertujuan untuk mengukur seberapa positif dan negatif perasaan masyarakat terhadap Bank BSI setelah terjadinya kebocoran data. Penelitian ini menggunakan dataset “Bank BSI *Ransomware Tweets – 24k Tweets*” dari kaggle yang berisi data yang diambil dari sosial media X dari tanggal 8 mei – 18 mei 2023 dengan total 24.000 data tweet dari pengguna sebagai tanggapan atas insiden kebocoran data Bank BSI.

## 1.2 Rumusan Masalah

Rumusan permasalahan yakni:

1. Berapa jumlah tweet positif, negatif dan netral pada sosial media X terhadap kebocoran data Bank BSI?
2. Berapa persen akurasi efektifitas algoritma IndoBERT dalam menganalisis sentimen tweet X masyarakat terhadap kebocoran data Bank BSI?

## 1.3 Tujuan Penelitian

Tujuan penelitian yakni:

1. Untuk mengukur sentimen masyarakat pada sosial media X terhadap kebocoran data Bank BSI.
2. Untuk mengetahui akurasi dan efektifitas algoritma IndoBERT dalam menganalisis sentimen pengguna sosial media X terkait kebocoran data Bank BSI.

## 1.4 Batasan Masalah

Batasan permasalahan yakni:

1. Bahasa pemrograman yang dipergunakan python
2. Platform yang dipergunakan adalah GoogleCollab Notebook
3. Algoritma yang dipergunakan adalah IndoBERT
4. Dataset yang digunakan adalah dataset “Bank BSI *Ransomware Tweets – 24k Tweets*” dari kaggle.