

DAFTAR ISI

| | |
|---|-------------|
| HALAMAN SAMPUL | |
| HALAMAN JUDUL | |
| PERNYATAAN KEASLIAN DAN PERSETUJUAN UNGGAH KARYA TUGAS AKHIR | ii |
| PERSETUJUAN DOSEN PEMBIMBING TUGAS AKHIR | iii |
| PERSETUJUAN TIM PENGUJI TUGAS AKHIR..... | iv |
| ABSTRAK | iv |
| ABSTRACT | v |
| KATA PENGANTAR..... | vi |
| DAFTAR ISI..... | viii |
| DAFTAR GAMBAR..... | xi |
| DAFTAR TABEL | xiii |
| BAB I PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang Masalah..... | 1 |
| 1.2 Rumusan Masalah | 2 |
| 1.3 Tujuan Penelitian | 2 |
| 1.4 Manfaat Penelitian | 3 |
| 1.5 Batasan Masalah..... | 3 |
| 1.6 Sistematika Penulisan | 4 |
| BAB II Tinjauan Pustaka..... | 5 |
| 2.1 Landasan Teori..... | 5 |
| 2.1.1 Serangan Siber (<i>Cyber Attack</i>) | 5 |
| 2.1.2 Media Sosial Twitter | 8 |
| 2.1.3 <i>Web scraping</i> | 11 |

| | | |
|--|--|-----------|
| 2.1.4 | <i>Data Mining</i> | 13 |
| 2.1.4.1 | <i>Text Mining dan Data Mining</i> | 14 |
| 2.1.4.2 | <i>K Means Clustering (KMC)</i> | 16 |
| 2.1.4.3 | <i>Word2Vec</i> | 19 |
| 2.1.4.4 | <i>Dimensionality Reduction</i> | 21 |
| 2.1.5 | <i>Text Mining Cyber Attack</i> | 21 |
| 2.2 | Penelitian Terdahulu | 22 |
| BAB III METODOLOGI PENELITIAN | | 24 |
| 3.1 | Analisis Proses | 24 |
| 3.1.1 | Analisis Proses <i>Web scraping</i> | 25 |
| 3.1.2 | Analisis Proses <i>Text Preprocessing</i> | 26 |
| 3.1.3 | Analisis Proses Ekspor Hasil <i>Text Preprocessing</i> | 29 |
| 3.1.4 | Analisis Proses <i>Text clustering</i> | 29 |
| 3.1.5 | Analisis Proses Penentuan Topik | 33 |
| 3.2 | Analisis Sistem..... | 34 |
| 3.2.1 | <i>Use Case</i> Program Analisis <i>Trend</i> Menggunakan <i>K-Means</i> | 34 |
| 3.2.2 | <i>Activity Diagram</i> | 38 |
| 3.2.3 | Parameter Program | 42 |
| 3.2.4 | <i>State Diagram</i> | 45 |
| 3.2.5 | Rancangan <i>Database</i> | 47 |
| BAB IV HASIL DAN PEMBAHASAN | | 49 |
| 4.1 | Hasil Penelitian | 49 |
| 4.1.1 | Hasil <i>Text clustering</i> Menggunakan $k = 2$ | 53 |
| 4.1.2 | Hasil <i>Text clustering</i> Menggunakan $k = 3$ | 57 |
| 4.1.3 | Hasil <i>Text Clustering</i> Menggunakan $k = 4$ | 60 |

| | | |
|----------------------------------|------------------|-----------|
| 4.2 | Pembahasan..... | 62 |
| BAB V Kesimpulan dan Saran | | 68 |
| 5.1 | Kesimpulan | 68 |
| 5.2 | Saran..... | 68 |
| DAFTAR PUSTAKA | | 70 |



DAFTAR GAMBAR

| | |
|--|----|
| Gambar 2.1 <i>Tweet</i> dengan <i>Tag Ransomware</i> | 9 |
| Gambar 2.2 Balasan Pengguna Lain Pada <i>Twitter</i> | 9 |
| Gambar 2.3 <i>Tweet</i> Mengenai Potensi Serangan <i>Ransomware</i> | 10 |
| Gambar 2.4 <i>Tweet</i> Mengenai Kerentanan Pada <i>Jaringan Mobile</i> | 10 |
| Gambar 2.5 <i>Tweet</i> Melindungi dari Sebuah Kerentanan | 11 |
| Gambar 2.6 <i>Request</i> dan <i>Response</i> dari REST API <i>Twitter</i> | 13 |
| Gambar 2.7 <i>Flowchart</i> Menentukan Nilai <i>k</i> Optimal Menggunakan <i>Silhouette</i> Score..... | 17 |
| Gambar 2.8 <i>Flowchart</i> Algoritma <i>KMC</i> | 18 |
| Gambar 2.9 <i>Diagram</i> <i>Skip Gram</i> | 20 |
| Gambar 2.10 <i>Diagram</i> <i>CBOW</i> | 20 |
| Gambar 3.1 Alur Sistem Analisis Tren Menggunakan <i>KMC</i> | 24 |
| Gambar 3.2 Alur <i>Web scraping</i> | 25 |
| Gambar 3.3 Alur Ekspor Hasil <i>Text Preprocessing</i> | 29 |
| Gambar 3.5 Contoh Hasil Vektorisasi dari Kata " <i>attacks</i> " Menggunakan <i>word2vec</i> | 30 |
| Gambar 3.6 Relasi Kesamaan Kata dari Kata " <i>attacks</i> " | 31 |
| Gambar 3.7 Contoh Hasil <i>Dimensionality Reduction</i> Menggunakan <i>PCA</i> Pada Kata " <i>attacks</i> " | 31 |
| Gambar 3.4 Alur <i>Text clustering</i> | 32 |
| Gambar 3.8 Alur Penentuan Topik Menggunakan <i>Frekuensi Kata</i> | 33 |
| Gambar 3.9 <i>Use Case</i> Program Analisis Tren Menggunakan <i>K-Means</i> | 34 |
| Gambar 3.10 <i>Activity Diagram</i> <i>Web scraping</i> | 38 |
| Gambar 3.11 <i>Activity Diagram</i> <i>Text Preprocessing</i> | 39 |
| Gambar 3.12 <i>Activity Diagram</i> Ekspor <i>Dataset</i> | 40 |
| Gambar 3.13 <i>Activity Diagram</i> <i>Text clustering</i> | 41 |
| Gambar 3.14 <i>State Diagram</i> Program Analisis Tren Serangan Siber Pada <i>Twitter</i> | 46 |

| | |
|---|----|
| Gambar 4.1 Linimasa <i>Tweet</i> dengan <i>Tag Cyberattack, Cyberattacks, dan Cybernews</i> Pada Rentang Waktu 01 September 2021 - 30 November 2021 | 49 |
| Gambar 4.2 Proses Melakukan <i>Web Scraping</i> Menggunakan Snsrape | 50 |
| Gambar 4.3 Proses Modul <i>Text Preporcessing</i> | 51 |
| Gambar 4.4 Hasil Text Preprocessing Dalam Format CSV..... | 52 |
| Gambar 4.5 Visualisasi Linimasa Jumlah <i>Tweet Cluster C01</i> Pada $k = 2$ | 54 |
| Gambar 4.6 Salah Satu <i>Tweet</i> dengan Topik “Kelompok Media Sosial Facebook Down” | 54 |
| Gambar 4.7 Salah Satu <i>Tweet</i> dengan Topik “Penyalahgunaan DMCA Twitter” | 55 |
| Gambar 4.8 Visualisasi Jumlah <i>Tweet Cluster C01</i> Pada $k = 2$ Menggunakan Pendekatan Frekuensi kata..... | 56 |
| Gambar 4.9 Visualisasi Linimasa Jumlah <i>Tweet Cluster C01</i> Pada $k = 3$ | 58 |
| Gambar 4.10 Visualisasi Linimasa Jumlah <i>Tweet Cluster C03</i> Pada $k = 3$ | 59 |
| Gambar 4.11 Visualisasi Linimasa Jumlah <i>Tweet Cluster C01</i> Pada $k = 4$ | 61 |
| Gambar 4.12 Visualisasi Linimasa Jumlah <i>Tweet Cluster C04</i> Pada $k = 4$ | 61 |
| Gambar 4.13 Perbandingan Jumlah Kosakata Antar <i>Cluster k = 2</i> | 64 |
| Gambar 4.14 Perbandingan Jumlah Kosakata Antar <i>Cluster k = 3</i> | 64 |
| Gambar 4.15 Perbandingan Jumlah Kosakata Antar <i>Cluster k = 4</i> | 64 |
| Gambar 4.16 Visualisasi Data <i>Cluster k = 2</i> | 65 |
| Gambar 4.17 Visualisasi Data <i>Cluster k = 3</i> | 65 |
| Gambar 4.18 Visualisasi Data <i>Cluster k = 4</i> | 66 |
| Gambar 4.19 Hasil Pencarian <i>Tweet</i> Menggunakan Kata " <i>ransom</i> " dan " <i>extortion</i> " | 67 |

DAFTAR TABEL

| | |
|---|----|
| Tabel 3.1 Parameter Library Snsrape | 26 |
| Tabel 3.2 Keluaran Snsrape..... | 26 |
| Tabel 3.3 Proses <i>Text Preprocessing</i> | 27 |
| Tabel 3.4 Use Case Description Melakukan <i>Web scraping</i> | 35 |
| Tabel 3.5 Use Case Description <i>Text Preprocessing</i> | 35 |
| Tabel 3.6 Use Case Description Ekspor Hasil <i>Text Preprocessing</i> | 36 |
| Tabel 3.7 Use Case Description Melakukan <i>Text Clustering</i> | 37 |
| Tabel 3.8 Deskripsi Parameter Utama | 42 |
| Tabel 3.9 Sub Parameter Untuk Action "web-scraping" | 42 |
| Tabel 3.10 Sub Parameter Untuk Action "text-preprocessing" | 43 |
| Tabel 3.11 Sub Parameter Untuk Action "export" | 43 |
| Tabel 3.12 Sub Parameter Untuk Action "text-clustering" | 44 |
| Tabel 3.13 Rancangan Tabel Posts | 47 |
| Tabel 4.1 Jumlah <i>Tweet</i> Hasil <i>Web Scraping</i> yang Dikelompokkan Berdasarkan <i>Tag</i> | 50 |
| Tabel 4.2 Nilai Silhoutte Score Menggunakan Sampel k 2, 3, dan 4..... | 52 |
| Tabel 4.3 Hasil Penentuan Topik Pada $k = 2$ | 57 |
| Tabel 4.4 Jumlah <i>Tweet</i> Pada Masing-Masing Topik..... | 57 |
| Tabel 4.5 Hasil Penentuan Topik Pada $k = 2$ | 59 |
| Tabel 4.6 Hasil Penentuan Topik Pada $k = 2$ | 62 |