BAB II

TINJAUAN PUSTAKA

**2.1 Emosi**

Emosi adalah kunci perasaan dan pikiran orang (Jabreel & Moreno, 2019). Perasaan intens yang diarahkan pada sesuatu atau seseorang sebagai respon terhadap peristiwa internal maupun eksternal bagi individu digambarkan sebagai emosi (Gaind et al., 2019). Emosi bersifat umum dan penting dalam semua aspek kehidupan manusia (Fanesya et al., 2019). Model emosi yang sangat populer adalah model emosi Ekman. Model Ekman membagi emosi menjadi 6 label emosi yaitu bahagia, marah, takut, jijik, sedih, dan terkejut (Kowalska & Wróbel, 2017), label-label ini bersifat universal pada budaya yang berbeda.

**2.2 Deteksi Emosi**

Deteksi Emosi merupakan bagian dari area yang lebih luas dari komputasi afektif bertujuan untuk memungkinkan komputer mengenali dan mengekspresikan emosi (Picard, 1997). Manfaat deteksi emosi dapat digunakan untuk mengambil keputusan seperti di bidang pendidikan, bisnis, politik, psikologi, sehingga menunjukkan pentingnya deteksi emosi (Fanesya et al., 2019).

**2.3 Media Sosial**

Media sosial merupakan platform komunikasi, interaksi, dan berbagi informasi yang populer melalui internet (Wilson et al., 2012). Media sosial dapat digunakan oleh penggunanya untuk berekspreksi (Sari & Haranto, 2019). Pada media sosial, setiap acara, berita atau aktivitas di seluruh dunia dapat dibagikan, didiskusikan, diposting dan dikomentari oleh jutaan orang (Gaind et al., 2019). Contoh media sosial seperti *facebook*, *twitter*, *instagram*, *line*.

**2.4 Twitter**

Twitter adalah salah satu media sosial dimana pengguna dapat membaca dan berbagi pesan singkat dengan jumlah maksimal 280 karakter, *twitter* termasuk media sosial dengan kategori *microblogging* (Setiawan & Setyohadi, 2017). *Twitter* telah berdiri sejak 2006 dan telah berkembang menjadi media sosial populer (Juwiantho et al., 2020). *Twitter* memiliki laju pertumbuhan pengguna tertinggi di indonesia adalah *twitter*, pengguna aktif *twitter* di Indonesia menempati posisi ketiga di Asia Pasifik dari tahun 2012 sampai tahun 2018 (Saputri et al., 2019).

**2.5 Preprocessing**

*Preprocessing* data seringkali mempengaruhi kinerja dari *machine learning*, sehingga lebih efektif (S. B. Kotsiantis, 2006). *Preprocessing* berguna untuk membersihkan data dari derau dan bagian tidak informatif yang tidak dibutuhkan, sehingga text siap untuk diklasifikasi (Haddi et al., 2013). Dengan demikian, *preprocessing* akan menjadikan data siap digunakan, Adapun proses *preprocessing* yang akan digunakan adalah *case folding*, *remove punction*, *remove number*, *tokenizing*, *stop removal*, *stemming*.

**2.5.1 Case Folding**

*Case Folding* adalah mengubah seluruh huruf dalam teks tersebut menjadi *lower case* atau huruf kecil semua (Salam et al., 2018), dengan huruf yang diterima dari “a” sampai “z”. Tujuannya agar terdapat standarisasi dalam penulisan . Contoh dapat dilihat pada **tabel 2.1**.

Tabel **2.1** Case Folding

|  |  |
| --- | --- |
| Data sebelum *case folding* | “Suaranya bagus, apalagi kalau dIiringi gitaR” |
| Data setelah *Case Folding* | “suaranya bagus, apalagi kalau diiringi gitar” |

**2.5.2 Remove Punctuation**

*Remove punctuation* adalah menghapus tanda baca pada teks dengan tujuan mengurangi beban pemprosesan klasifikasi karena dianggap tidak penting. Beberapa contoh tanda baca yang yang dihapus titik (.), koma(,), tanda tanya (?), *slash* (/), *hastag* (#), tanda seru (!). Contoh dapat dilihat pada **tabel 2.2**.

Tabel **2.2**

|  |  |
| --- | --- |
| Data sebelum *remove punctuation* | “suaranya bagus !!!, apalagi kalau diiringi gitar” |
| Data setelah *remove punctuation* | “suaranya bagus apalagi kalau diiringi gitar” |

**2.5.3 Remove Number**

*Remove number* adalah menghapus angka pada suatu teks, penghapusan ini sebab angka dianggap tidak memiliki arti dan termasuk *delimiter*. Contoh dapat dilihat pada **tabel 2.3**.

Tabel **2.3**

|  |  |
| --- | --- |
| Data sebelum *remove number* | “barang yang dijual bagus2” |
| Data setelah *remove number* | “barang yang dijual bagus” |

**2.5.4 Tokenizing**

*Tokenizing* adalah pemotongan kalimat menjadi tiap-tiap kata penyusunnya, biasanya pemotongan berdasarkan *whitespace* seperti spasi, tab, dan enter.Tiap-tiap kata hasil *tokenizing* disebut token (Juwiantho et al., 2020). Contoh dapat dilihat pada **tabel 2.4**

Tabel **2.4** Tokenizing

|  |  |
| --- | --- |
| Sebelum proses *Tokenizing* | konten yang bermutu namun sayang suaranya kurang jelas |
| Setelah proses *Tokenizing* | [‘konten’,’yang’,’bermutu’,’namun’,’sayang’,’suaranya’,  ’kurang’,’jelas’] |

**2.5.5 Stopword Removal**

Stopword Removal merupakan proses untuk menghilangkan kata-kata yang tidak memiliki arti atau makna, namun tidak akan mengubah makna dari komentar tersebut (Juwiantho et al., 2020).Daftar *stopword* yang digunakan diperoleh dari Librian (2017). Pada penelitian ini menggunakan *stopword* dari *library* sastrawi.. Contoh dapat dilihat pada **tabel 2.5**

Tabel **2.5** Stopword Removal

|  |  |
| --- | --- |
| Sebelum Stop Removal | [‘konten’,’yang’,’bermutu’,’namun’,’sayang’,’suaranya’,  ’kurang’,’jelas’] |
| Setelah Stop Removal | [‘konten’,’bermutu’,’sayang’,’suaranya’,’kurang’,’jelas’] |

**2.5.6 Stemming**

Stemming merupakan proses untuk mentransformasi kata-kata yang terdapat pada suatu dokumen menjadi kata dasar, dengan menghapus awalan, akhiran maupun sisipan. Hal ini bertujuan untuk mengurangi varian kata dengan makna yang hampir sama pada suatu dokumen, serta meningkatkan performa pada tahap *information retrieval*. Pada penelitian ini akan menggunakan *library* sastrawi. Contoh dapat dilihat pada **tabel 2.6**

Tabel **2.6** Stemming

|  |  |
| --- | --- |
| Sebelum Stemming | [‘konten’,’bermutu’,’sayang’,’suaranya’,’kurang’,’jelas’] |
| Setelah Stemming | [‘konten’,’mutu’,’sayang’,’suara’,’kurang’,’jelas’] |

**2.6 Word Embedding**

**2.7 Fasttext**

**2.8 Deep Learning**

**2.9 Recurrent Neural Network**

**2.10 Long Short Term Memory**

**2.11 Validasi dan Pengujian**

**2.12 Studi Pustaka**