





Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

token. Eksperimen model klasifikasi token untuk tahapan ekstraksi menunjukkan bahwa model terbaik dengan akurasi 81.49% didapatkan dengan menggunakan algoritma pembelajaran IBk (*Instance-based learning*).

Berdasarkan fakta dan penelitian terkait yang telah diuraikan, maka akan dilakukan penelitian dengan judul “**Named Entity Recognition pada Tweet Iklan Menggunakan Metode *k-Nearest Neighbor***”. Proses yang akan dilakukan untuk mengidentifikasi token menggunakan encoding *Beginning, Inside, Outside* (BIO) kemudian dilakukan klasifikasi *Named Entity Recognition* menggunakan metode *k-Nearest Neighbor* yang dikelompokkan menjadi empat kategori yaitu nama penjual, nama produk, alamat dan kontak. Nama penjual adalah akun twitter yang mempromosikan produk, nama produk adalah nama barang yang dipromosikan, alamat adalah lokasi atau tempat produk dijual dan kontak adalah sarana untuk melakukan interaksi antar konsumen dan penjual barang seperti nomor telepon. Pengenalan nama entitas pada *tweet* dilakukan untuk mendapatkan informasi terstruktur dari teks yang tidak terstruktur dan dapat memberikan manfaat untuk para peneliti ekstraksi informasi khususnya dalam mengekstraksi informasi pada twitter.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka dibuat suatu rumusan masalah yaitu bagaimana menerapkan metode *k-Nearest Neighbor* untuk klasifikasi *Named Entity Recognition* pada *tweet* iklan dan bagaimana tingkat akurasinya.

## 1.3 Batasan Masalah

Adapun yang menjadi batasan masalah dalam tugas akhir adalah sebagai berikut:

1. *Tweet* yang digunakan adalah *tweet* iklan.
2. Nama entitas yang diidentifikasi yaitu nama penjual, nama produk, alamat dan kontak.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Penerapan metode *k-Nearest Neighbor* untuk *Named Entity Recognition* pada *tweet* iklan
2. Menguji tingkat akurasi pada klasifikasi *Named Entity Recognition* pada *tweet* iklan.

## 1.5 Sistematika Penulisan

Demi kemudahan dalam pembahasan penelitian ini disusun pada bab secara berurutan. Adapun bab yang terdapat pada laporan ini adalah:

### **BAB I PENDAHULUAN**

Pada Bab I menguraikan tentang latar belakang yang mendasari penelitian tugas akhir, rumusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan laporan.

### **BAB II LANDASAN TEORI**

Pada Bab II menjelaskan tentang teori-teori umum tentang Twitter, iklan, *Text Mining*, *Named Entity Recognition*, *k-Nearest Neighbor* dan teori sistem yang akan dibangun.

### **BAB III METODOLOGI PENELITIAN**

Pada Bab III ini menjelaskan tentang metodologi penelitian yang terdiri dari perumusan masalah, pengumpulan data, analisa, perancangan, implementasi, pengujian serta kesimpulan dan saran dalam penelitian.

### **BAB IV ANALISA DAN PERANCANGAN**

Pada Bab IV ini menjelaskan tentang analisa data, analisa proses dan metode serta melakukan perancangan sistem yang akan dibangun.

### **BAB V IMPLEMENTASI DAN PENGUJIAN**

Pada Bab V ini berisi tentang implementasi serta melakukan pengujian terhadap sistem *Named Entity Recognition* pada *tweet* menggunakan metode *k-Nearest Neighbor*.

## BAB VI PENUTUP

Bagian ini berisi kesimpulan hasil penelitian beserta saran-saran yang berkaitan dengan penelitian ini



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.