

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Belakangan ini penelitian mengenai ekstraksi makna sebuah kalimat untuk diambil sentimennya dan diklasifikasikan secara otomatis menjadi sangat marak. Metode analisis sentimen dinilai efektif untuk mendapatkan informasi apakah orang lain memiliki pandangan positif atau negatif terhadap sebuah topik. Analisis sentimen merupakan proses memahami, mengekstrak, dan mengolah data tekstual secara otomatis untuk mendapatkan informasi [1]. Namun, banyak sekali orang tidak secara langsung menyampaikan apa yang dirasakannya seperti menggunakan ujaran bersifat ironi dengan tujuan mengejek atau mengkritik. Ironi adalah bentuk wicara yang pintar di mana pembicara atau penulis mengatakan atau menulis kebalikan dari apa yang mereka maksud [2]. Dengan adanya ujaran bersifat ironi, metode analisis sentimen sulit untuk memahaminya.

Penelitian tentang identifikasi ujaran bersifat ironi dalam konteks analisis sentimen telah banyak dilakukan. Diantaranya, Rilof [3] mendeteksi ironi jika memperlihatkan perbedaan yang jelas antara sentimen positif terhadap situasi yang negatif. Lunando [4] mengusulkan metode *unigram+negativity+interjection* dimana kalimat bersifat ironi tersebut biasanya diawali kata-kata yang mengandung *interjection words* seperti “aha”, “bah”, “nah”, “wew”, “wow”, “yay”, “uh”, dll. Anupam [6] mengusulkan metode *Historical Tweet-based Predictor* yang mengidentifikasi ironi jika sentimen yang diekspresikan terhadap entitas dalam target tweet setuju dengan sentimen yang diekspresikan oleh penulis terhadap entitas itu dimasa lalu.

Penelitian ini menyajikan pendekatan baru terhadap ekstraksi makna kalimat bersifat ironi, mirip dengan Anupam [6]. Namun, saat mereka melihat kesejarahan ujaran seseorang untuk dicocokkan dengan entitas target yang sama, penelitian ini bertujuan untuk mengklasifikasikan seseorang berdasarkan ujaran-ujarannya dimasa lalu.

Sejarah ujaran seseorang mempengaruhi penilaian sebuah kalimat memiliki kecenderungan ironi atau tidak. Misalnya ketika seseorang sering berbicara negatif, tiba-tiba memberikan opini positif terhadap sebuah topik sementara orang lain memberikan opini negatif terhadap topik tersebut. Apabila ketika proses pengklasifikasian sejarah seseorang tersebut sudah jelas mayoritas nilai positif atau negatifnya, maka dapat dengan mudah diklasifikasikan. Namun, masalah akan muncul ketika nilai sentimen positif dan negatif hampir seimbang sehingga menghasilkan nilai mayoritas sentimen tidak jelas, maka dari itu perlu digunakan metode *fuzzy logic* dari *soft computing* untuk menilai kecenderungan sejarah ujaran seseorang.

*Soft computing* adalah kumpulan metodologi yang bertujuan untuk mengeksplorasi toleransi ketidaktepatan dan ketidakpastian untuk mencapai traktabilitas, ketahanan, dan biaya solusi rendah. Konstituen utamanya adalah logika *fuzzy* [7]. Dan *fuzzy logic* yang dimaksud adalah “Logika fleksibel yang berkaitan dengan kalkulus yang berpredikat yang tidak tentu di dunia ini.” [8]. Logika *fuzzy* umumnya diterapkan pada masalah-masalah yang mengandung unsur ketidak pastian, ketidak tepatan, dan sebagainya. Logika *fuzzy* erat kaitannya dengan *Natural Language Proccessing* (NLP) karena dikembangkan berdasarkan bahasa manusia yang menjembatani bahasa mesin yang presisi dengan bahasa manusia yang menekankan pada makna atau arti. Logika *fuzzy* mampu membuat kriteria berdasarkan himpunan data yang batasnya tidak tegas seperti himpunan data ujaran seseorang yang terus bertambah dan mengakibatkan batas ukuran penilaian terus berubah. Maka penelitian ini bermaksud untuk membuat membuat suatu model pengklasifikasian sejarah ujaran seseorang dengan menggunakan logika *fuzzy* sehingga membuat penghitungan nilai relatif lebih stabil ketika nilai mayoritas sentimen tidak jelas.

## 1.2 Identifikasi Masalah

Pendeteksian kalimat ironi menggunakan logika *fuzzy* ini mengacu kepada ujaran positif seseorang yang berkebalikan dari mayoritas sentimen terhadap sebuah topik dan sejarah ujaran orang tersebut. Pada setiap topik harus dinilai terlebih dahulu nilai dari mayoritas sentimennya lalu disandingkan dengan satu kalimat yang diduga bermakna ironi. Setelah itu sejarah ujaran seseorang dikumpulkan dan dinilai sentimennya oleh metode analisis sentimen, lalu diklasifikasikan oleh logika *fuzzy* untuk memberi nilai kecenderungan positif dan negatifnya. Apabila ujaran yang diduga bermakna ironi tersebut dimiliki oleh seseorang dengan sejarah ujaran yang negatif, sementara orang lain memberikan opini negatif terhadap topik tersebut maka dapat ditentukan nilai potensi kalimat tersebut bermakna ironi.

Kerumitan untuk menyatukan nilai dari mayoritas sentimen terhadap sebuah topik dan hasil klasifikasi logika *fuzzy* dari sejarah ujaran seseorang merupakan masalah yang hendak diselesaikan pada penelitian ini. Pemikiran-pemikiran yang muncul pada penelitian ini adalah sebagai berikut :

1. Suatu ujaran positif diantara mayoritas ujaran negatif patut diduga sebagai ironi apabila sejarah ujaran seseorang tersebut negatif.
2. Diperlukan teknik yang dapat menentukan tingkat kecenderungan sentimen berdasarkan sejarah ujaran seseorang tanpa batasan yang kaku.

### 1.3 Premis & Hipotesis

Berdasarkan rumusan masalah di atas, diformulasikan premis-premis sebagai berikut.

1. Suatu ujaran positif diantara mayoritas ujaran negatif patut diduga sebagai ujaran ironi. Sejarah ujaran seseorang dapat dijadikan dasar penentuan tingkat kecenderungan ironi menggunakan teknik *Historical Tweet-based Predictor* belum dapat menentukan tingkat kecenderungan sentiment seseorang.
2. Teknik logika *fuzzy* mampu membuat kriteria berdasarkan himpunan data yang batasnya tidak tegas.

Berdasarkan premis-premis diatas maka diajukan hipotesis, yaitu penentuan suatu kalimat ironi dapat dibuat lebih akurat dan sederhana berdasarkan gabungan teknik *Historical Tweet-based Predictor* dan *Fuzzy Logic* atas himpunan data sejarah ujaran seseorang.

### 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah membuat sebuah tehnik baru yang menggabungkan *Historical Tweet-based Predictor* dan *Fuzzy Logic* atas himpunan data ujaran seseorang. Pembahasan dalam penelitian ini dibatasi agar tidak menyimpang dari tujuan yang ingin dicapai, adapun batasannya adalah:

1. Dataset yang digunakan merupakan model yang telah dibuat dimana nilai mayoritas sentimen atas himpunan sejarah ujaran seseorang nilai positif dan negatifnya mirip.
2. Menggunakan dua kelas sentimen yaitu positif dan negatif.
3. Ujaran target yang akan dinilai merupakan ujaran positif diantara mayoritas ujaran negatif.
4. Ujaran target yang akan dinilai merupakan ujaran dari seseorang yang mempunyai himpunan sejarah ujaran negatif.

## 1.5 Kontribusi

Pada sisi keilmuan, kontribusi utama yang dihasilkan dari penelitian ini adalah dihasilkannya teknik penggabungan *Historical Tweet-based Predictor* dan *Fuzzy Logic* atas himpunan data ujaran seseorang untuk memberi nilai kecenderungan ironi terhadap sebuah ujaran.

Bagi kegiatan penelitian bidang analisis sentimen, penelitian ini memberikan kontribusi mengenai ukuran kualitas dari analisis sentimen. Dan bagi kegiatan penelitian bidang analisis sentimen wilayah pendeteksian ironi adalah pembuka topik-topik penelitian berikutnya.

## 1.6 Metodologi Penelitian

Metodologi yang dilakukan pada penelitian ini adalah sebagai berikut.

(1) Identifikasi, perumusan konsep dan penentuan spesifikasi

Pada tahap ini dilakukan identifikasi masalah pada penentuan sebuah ujaran bersifat ironi, kemudian dilakukan perumusan konsep tehnik yang akan dikembangkan dan penentuan tehnik baru yang diusulkan.

(2) Perancangan sistem dan pengembangan metoda

Pada tahap ini dilakukan perancangan sistem dan subsistem yang dibangun, meliputi perangkat lunak dan model data yang dibangun.

(3) Pengujian metoda menggunakan komputer

Pada tahap ini dilakukan pengujian untuk setiap metoda yang dikembangkan melalui perangkat lunak komputer.

(4) Analisis dan evaluasi

Tahap terakhir dari kegiatan penelitian ini adalah melakukan analisis hasil pengujian yang telah dilakukan. Selanjutnya evaluasi dilakukan untuk melihat kekurangan dari sistem yang dibuat.

## **1.7 Sistematika Penelitian**

Laporan penelitian ini disusun dengan sistematika sebagai berikut:

### **Bab 1 Pendahuluan**

Bab 1 menjelaskan latar belakang penelitian, identifikasi masalah, premis dan hipotesis, tujuan, metoda yang digunakan dan sistematika penelitian.

### **Bab 2 Landasan Teori**

Bab 2 membahas berbagai konsep dasar dan teori-teori yang berkaitan dengan analisis sentimen dan logika fuzzy. Mulai dari algoritma yang digunakan, teknik penilaian sentimen, hingga teknik evaluasi dari hasil analisis sentimen.

### **Bab 3 Metode Penelitian**

Bab 3 membahas alur kerja Fuzzy Historical Classifier, data penelitian yang digunakan, dan alur pengujian.

### **Bab 4 Hasil Penelitian**

Bab 4 membahas hasil pengujian data yang telah dilakukan. Dalam bab ini akan diuraikan hasil potensi ironi menggunakan Fuzzy Historical Classifier berdasarkan alur pengujian yang telah dibuat.

### **Bab 5 Kesimpulan dan Saran**

Bab 5 menyimpulkan hasil-hasil yang diperoleh dari penelitian penelitian ini dan tindak lanjut untuk penelitian tahap berikutnya.