

## **BAB V**

### **SIMPULAN SARAN**

#### **5.1 Simpulan**

Pembelajaran mendalam telah menandai era baru dalam pemahaman dan penafsiran opini publik dalam skala besar. Melalui penggunaan metode ini, yang merupakan sub-bidang dari kecerdasan buatan yang mengutamakan penggunaan jaringan saraf buatan dengan banyak lapisan, telah membawa dampak yang signifikan dalam pemrosesan data yang kompleks, terutama dalam konteks analisis teks. Ketangguhan pembelajaran mendalam dalam menangani volume dan kompleksitas data teks telah membukukan keberhasilan yang mencolok dalam berbagai aplikasi, termasuk analisis sentimen. Penekanan pada pentingnya pengumpulan data yang memadai juga menjadi sorotan dalam penelitian ini. Dalam analisis sentimen menggunakan pembelajaran mendalam, ketersediaan data yang cukup dan representatif menjadi kunci utama untuk memastikan persiapan dan pemodelan yang optimal. Melalui penekanan ini, penelitian menegaskan bahwa kualitas dan kuantitas data memainkan peran krusial dalam keberhasilan analisis sentimen menggunakan metode pembelajaran mendalam. Dengan demikian, penelitian ini memperkuat pemahaman tentang pentingnya sumber daya data yang kuat dalam konteks pengembangan model analisis sentimen yang akurat dan andal.

Analisis sentimen yang menggunakan machine learning, terutama dalam konteks pemilihan umum presiden, seringkali bersumber dari data Twitter. Terdapat 17 jurnal yang menggunakan dataset dari media sosial ini, dengan total data mencapai 253.959 data. Tahapan umum dalam penelitian meliputi preprocessing dan pengumpulan data, sentiment analysis, dan tahap labeling untuk data tanpa label, menegaskan pentingnya persiapan dan analisis yang cermat. Proses preprocessing umumnya mencakup *case folding*, *stemming*, *tokenization*, *stopword removal*, *filtering*, transformasi, dan pembersihan data, menyesuaikan data untuk kebutuhan algoritma. Algoritma yang sering digunakan adalah Naïve Bayes, Support Vector Machine (SVM), dan K-Nearest Neighbor (KNN), dengan beberapa jurnal juga menggunakan Adaboost dan K-Means untuk memperluas

pendekatan analisis. Evaluasi model sering menggunakan metrik precision dan recall, diikuti oleh f1-score dan akurasi, dengan precision dan recall mencatat hasil tertinggi hingga 99% dan 100%. Tantangan utama dalam analisis sentimen bahasa Indonesia meliputi kurangnya dataset terannotasi, keberagaman bahasa, preprocessing yang rumit, klasifikasi multi-kelas, dan pemahaman konteks budaya lokal, yang memerlukan pendekatan holistik dan kerja sama lintas disiplin untuk mengatasinya.

## 5.2 Saran

Dalam penelitian ini, ditemukan bahwa pengumpulan data yang lebih luas dapat meningkatkan hasil analisis sentimen secara signifikan. Oleh karena itu, saran yang diberikan adalah sebagai berikut:

1. Perlunya upaya untuk mengumpulkan lebih banyak data guna memastikan hasil analisis sentimen yang optimal. Dengan memperluas cakupan data, penelitian ini tidak hanya memberikan kontribusi yang berharga bagi pemahaman tentang penggunaan pembelajaran mesin mendalam dalam analisis sentimen, tetapi juga memberikan arahan untuk penelitian masa depan dalam area ini.
2. Memperkaya dataset. Peneliti dapat mengeksplorasi lebih banyak pola dan tren dengan menggunakan alat pencarian lainnya seperti Web of Science, IEEE Xplore, PubMed, dan Microsoft Academic untuk memperkaya dataset jurnal sehingga jumlah jurnal yang didapat lebih banyak dengan menggunakan kombinasi kata kunci yang lebih umum sehingga pencariannya lebih luas.
3. Memperbanyak ragam keywords yang akan digunakan berdasarkan *Research Question* yang telah ditentukan, seperti sumber data, algoritma machine learning, research flow, preprocessing, dan evaluasinya. Sehingga keywords yang digunakan tidak hanya ditentukan berdasarkan jurnal yang didapatkan saja.