

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Kepribadian mempengaruhi berbagai aspek individu seperti kesehatan mental, fisik, kebugaran, dan karir. *Myers-Briggs Type Indicator* (MBTI), dikembangkan oleh Katherine Cook dan Isabel Briggs Myers berdasarkan teori Carl Gustav Jung. MBTI adalah alat psikologis yang mengelompokkan orang ke dalam 16 tipe berdasarkan preferensi dalam empat dimensi: Introversi (I) atau Ekstraversi (E), Sensing (S) atau Intuisi (N), Berpikir (T) atau Merasakan (F), dan Menilai (J) atau Merasakan (P). Tes MBTI membantu individu memahami diri, kelebihan, kelemahan, serta cara mereka melihat dunia dan mengambil keputusan, yang berguna dalam berbagai aspek kehidupan termasuk karier dan hubungan sosial. Meski reliabilitas dan validitasnya sering dikritik, MBTI tetap menjadi alat uji kepribadian terpopuler di dunia, digunakan di 115 negara, tersedia dalam 29 bahasa, dan digunakan oleh 88 perusahaan Fortune 100 [1].

Film-film dari Marvel Cinematic Universe (MCU) telah mencapai kesuksesan global dan mendapat pengakuan luas. Para aktor dan aktris dalam MCU tidak hanya diakui karena kemampuan akting mereka, tetapi juga karena kemampuan mereka membawakan karakter-karakter ikonik dalam alur cerita yang kompleks. Penggunaan media sosial, terutama Twitter, memperkuat interaksi antara para aktor/aktris dan penggemar serta memberikan pandangan lebih mendalam mengenai kepribadian mereka di luar layar. Kepribadian MBTI para pemeran Marvel, dapat memberikan wawasan tentang motivasi, perilaku, dan hubungan karakter mereka. Dengan menemukan kesamaan tipe kepribadian antara diri sendiri dan karakter Marvel yang dipilih, seseorang dapat mengevaluasi diri dan menggunakan hal itu sebagai sumber motivasi serta referensi dalam pengambilan keputusan.

Penelitian klasifikasi MBTI sudah banyak dilakukan oleh Fitri N [2], Harahap R [3], Persada P [4]. Mereka menggunakan *machine learning* untuk mengklasifikasi tipe kepribadian MBTI dengan hasil yang bervariasi. Penelitian lain oleh Alfian Hakim yang menggunakan kombinasi *Long Short Term Memory* (LSTM) dan Word2Vec untuk memprediksi kepribadian MBTI telah menunjukkan hasil yang lebih baik, dengan mencapai F1 *score* antara 0.65 dan 0.88 [5]. Selain itu kombinasi LSTM dan Word2Vec ini mendapatkan *accuracy* 0.86 pada penelitian oleh Muhammad P [6]. Meskipun demikian, penelitian ini masih dihadapkan pada tantangan ketidakseimbangan data, dimana jumlah sampel untuk setiap tipe kepribadian tidak merata. Dengan demikian, penggunaan LSTM dan Word2Vec dalam penelitian ini menunjukkan potensi untuk meningkatkan kinerja model dalam memprediksi tipe kepribadian MBTI, meskipun masih diperlukan penanganan khusus terhadap ketidakseimbangan data untuk memperoleh hasil yang lebih optimal.

Berbeda dengan penelitian lain yang menggunakan metode imbalance data-handling yaitu *Synthetic Minority Over-sampling Technique* (SMOTE) dapat meningkatkan hasil akurasi [7]. Penelitian oleh Whidiyasana yang menggunakan SMOTE juga mendapat hasil yang baik berupa peningkatan akurasi [8]. Dari penelitian tersebut SMOTE mampu meningkatkan kinerja model klasifikasi dengan cara menghasilkan sampel sintesis untuk kelas minoritas, sehingga memberikan representasi yang lebih seimbang antara kelas-kelas yang berbeda dan meningkatkan akurasi prediksi model. Studi ini menjadi salah satu acuan penting yang mendukung penggunaan SMOTE sebagai teknik penyeimbang *dataset* dalam berbagai konteks analisis data. Namun tidak disebutkan dengan pasti bagaimana data dikatakan tidak seimbang, tetapi rasio 1:10 biasanya sudah cukup dikatakan sebagai data yang tidak seimbang.

Berdasarkan uraian di atas, penelitian ini akan menggunakan pengembangan model klasifikasi tipe kepribadian MBTI dari aktor dan aktris

film Marvel melalui analisis cuitan Twitter menggunakan LSTM dan Word2Vec serta menambahkan teknik *imbalance data-handling* seperti SMOTE untuk menyelesaikan permasalahan *dataset* yang tidak seimbang. Referensi-referensi terdahulu, termasuk penelitian pada pengguna Twitter dan Instagram, memberikan dasar kuat untuk melihat lebih lanjut bagaimana aktor dan aktris film Marvel, membangun identitas dan mengekspresikan tipe kepribadian mereka melalui media sosial.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah dijelaskan rumusan masalah yang didapat adalah bagaimana pengaruh teknik *imbalance data-handling*, dalam penelitian ini adalah *Synthetic Minority Over-sampling Technique* (SMOTE) terhadap kinerja model *Long-Short Term Memory* (LSTM) dalam mengklasifikasi tipe kepribadian MBTI.

## 1.3 Maksud dan Tujuan

Berdasarkan rumusan masalah yang ada penelitian ini bermaksud untuk mengimplementasikan teknik *imbalance data-handling*, khususnya SMOTE, terhadap model LSTM dalam mengklasifikasi tipe kepribadian. Kemudian penelitian ini bertujuan untuk mengukur performa SMOTE untuk mengatasi ketidakseimbangan *dataset* pada model LSTM dalam mengklasifikasi tipe kepribadian MBTI.

## 1.4 Batasan Masalah

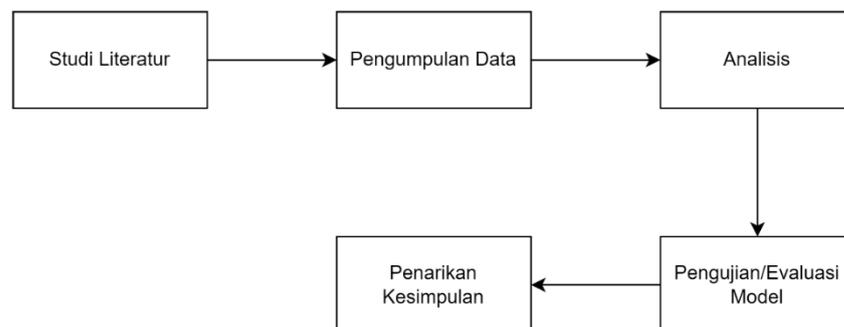
Batasan masalah penelitian ini mencakup beberapa aspek yang perlu dipertimbangkan untuk mengarahkan fokus penelitian dan membatasi ruang lingkup analisis

1. Data yang digunakan adalah twit yang diunggah oleh aktor dan aktris, data ini diambil pada tanggal 23 April 2024 s/d 25 Juni 2024.

2. Aktor dan aktris yang menjadi subjek penelitian terbatas pada mereka yang menjadi pemeran dalam Marvel Cinematic Universe (MCU) pada tahun 2008-2024.
3. Twit yang digunakan adalah twit yang berbahasa Inggris.
4. Metode penanganan *imbalanced data* menggunakan SMOTE dasar.
5. Metode pembobotan data terfokus pada penggunaan Word2Vec (CBOW).
6. Model prediksi tipe kepribadian menggunakan LSTM.
7. Analisis terbatas pada prediksi tipe kepribadian MBTI, tidak membahas atau membandingkan dengan kerangka kerja tipe kepribadian lainnya.
8. Hasil kepribadian MBTI yang dihasilkan yaitu *Extroversion-Introversion* (E/I), *Sensing-Intuition* (S/N), *Thinking-Feeling* (T/F), *Judging-Perceiving* (J/P) Berdasarkan empat domain tersebut, sebagai contoh jika hasil evaluasi psikologis individu menunjukkan cenderung pada *Introversion* (I), *Intuition* (N), *Feeling* (F), dan *Judging* (J), maka individu tersebut akan diklasifikasikan sebagai tipe kepribadian INFJ.
9. Penentuan tipe kepribadian MBTI aktor dan aktris berdasarkan data dari situs *personality-database.com* pada tanggal 23 April 2024 s/d 30 Juni 2024.

## 1.5 Metode Penelitian

Metode penelitian dapat dilihat pada Gambar 1.1



Gambar 1.1 Metode Penelitian

### **1.5.1. Studi Literatur**

Studi literatur merupakan teknik pengumpulan data dengan cara melakukan peninjauan terhadap beberapa *e-book*, makalah, tesis, skripsi dan jurnal terdahulu yang berkaitan dengan penelitian

### **1.5.2. Pengumpulan Data**

Data twit aktor dan aktris diambil dengan memanfaatkan *library* yang tersedia pada bahasa pemrograman Python. Penulis membuat daftar nama aktor dan aktris dengan nama penggunanya di Twitter terlebih dahulu sebelum mengambil data twit. Data yang sudah diperoleh kemudian diberi label tipe kepribadian MBTI secara manual berdasarkan situs *personality-database.com*.

### **1.5.3. Analisis**

Analisis dalam penelitian ini meliputi beberapa tahapan yang meliputi analisis masalah, analisis data, gambaran umum sistem, dan analisis metode. Tahapan analisis masalah akan mengidentifikasi masalah utama yang perlu diselesaikan, yaitu ketidakseimbangan *dataset* dalam mengklasifikasi tipe kepribadian. Analisis data akan mencakup pemahaman mendalam tentang sumber data yang digunakan, termasuk kondisi data, distribusi kelas, dan karakteristik lainnya. Gambaran umum sistem akan menjelaskan struktur dan fungsi keseluruhan dari sistem yang dikembangkan, termasuk algoritma, arsitektur model, dan langkah-langkah yang akan diambil dalam proses pengolahan data dan analisis. Sedangkan analisis metode akan membahas contoh data, praproses data dengan logika proses dari setiap tahapan.

### **1.5.4. Pengujian/Evaluasi Model**

Evaluasi dilakukan dengan menguji data yang sudah dilakukan SMOTE terhadap model untuk mengukur seberapa akurat model bekerja.

### 1.5.5. Penarikan Kesimpulan

Penarikan kesimpulan dari hasil proses dan evaluasi model yang telah dilakukan pada penelitian.

## 1.6 Sistematika Penulisan

### BAB 1 PENDAHULUAN

Bab 1 menguraikan tentang latar belakang yang terdiri dari permasalahan yang berkaitan dengan penelitian, rumusan masalah, maksud dan tujuan penelitian, batasan masalah, metode penelitian, dan sistematika penulisan.

### BAB 2 LANDASAN TEORI

Bab 2 membahas dan menjelaskan mengenai landasan teori yang mendukung penulisan skripsi, mulai dari *machine learning*, tipe kepribadian MBTI, Word2vec, SMOTE, LSTM.

### BAB 3 ANALISIS

Bab 3 berisikan tentang analisis-analisis terhadap model yang meliputi analisis masalah, analisis data, gambaran umum, analisis metode dan kebutuhan non fungsional.

### BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab 4 berisikan hasil implementasi dari analisis dan pengujian terhadap model yang telah dibuat.

### BAB 5 KESIMPULAN DAN SARAN

Bab 5 berisi kesimpulan mengenai hasil penelitian dan saran yang diharapkan akan berguna untuk penelitian lebih lanjut pada kasus yang sama.