

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada tahun 2024, menurut data yang ada, pengguna media sosial telah mencapai 62,3% populasi dunia [1]. Dengan ini, beragam unggahan yang dihasilkan sangatlah banyak. Unggahan tersebut dapat berupa kalimat, foto, video, maupun dokumen lainnya. Pengguna media sosial membuat unggahan untuk berbagai tujuan, mulai dari mengomentari pemerintah, memberi tanggapan pada hal-hal yang belum lama terjadi, memberikan informasi tentang keadaan sekitarnya, atau hanya untuk berkelakar.

Akan tetapi, karena kebebasan pengguna media sosial dalam membuat *post*, tak jarang muncul kata-kata maupun kalimat yang bertujuan untuk menghasut kebencian. Ujaran kebencian, atau dalam bahasa Inggris disebut *hate speech*, yang dilontarkan oleh seorang pengguna media sosial dapat menyebabkan efek psikologis yang serius pada pengguna media sosial lainnya. Hal ini disebabkan sifatnya yang menyerang dan merendahkan [2]. Merujuk pada sebuah penelitian, ditemukan bahwa ujaran kebencian berpengaruh dalam merusak kesehatan mental terhadap demografi muda [3].

Berbagai metode telah dilakukan untuk menekan penyebaran *hate speech*. Salah satunya adalah pengecekan manual oleh manusia. Pengecekan manual ini memiliki beberapa kekurangan. Pertama, yaitu karena pengecekan dilakukan setelah adanya komplain dari pengguna, yang berarti *hate speech* telah berdampak pada beberapa pengguna [2]. Kekurangan lainnya adalah potensi pembatasan kebebasan berbicara yang mungkin berlebihan oleh pihak yang melakukan pengecekan. Karena sulitnya pendeteksian, *hate speech* dapat tersebar dengan mudah. Maka dari itu, diperlukan sebuah metode untuk mendeteksi dan mencegah penyebaran *hate speech* di media sosial.

Metode yang telah ditemukan untuk mengatasi hal tersebut adalah penggunaan *deep learning* untuk pendeteksian *hate speech* secara otomatis. Dengan *deep learning* masalah seperti sulitnya pengecekan manual oleh sumber daya

manusia yang terbatas dan potensi sensor yang berlebihan dapat teratasi. Akan tetapi *deep learning* juga mempunyai permasalahannya tersendiri. Seperti perlunya *hardware* yang memadai, sehingga penggunaannya secara luas masih membutuhkan sumber daya *hardware* yang besar. Pada penelitian ini, peneliti berencana untuk mengoptimalkan *deep learning* agar penggunaan sumberdaya *hardware* dapat lebih rendah.

Selama beberapa tahun, model-model *deep learning* telah dikembangkan oleh para peneliti untuk permasalahan ini. Dengan akurasi diatas 80% [4][5][6][7]. Melihat fakta tersebut penulis berinisiatif tidak hanya mengembangkan model, tetapi juga melakukan metode yang lebih jauh, yaitu optimalisasi, dengan metode *pruning* dan *quantization*.

Selain itu, melihat beberapa penelitian sebelumnya, berbagai metode untuk mengoptimalkan model *deep learning* telah ditemukan. Song Han et al. berhasil mengurangi FLOP dari model Lenet-300-100 hingga 91,6% tanpa mengurangi akurasi dari model tersebut [8]. Mereka memangkas neuron pada sebuah neural network yang tidak memiliki koneksi masuk atau koneksi keluar. Metode tersebut disebut juga dengan *Network Pruning of Weight* yang dijelaskan oleh Sunil Vadera et al. [9].

Optimalisasi juga dapat dilakukan dengan metode *quantization*, dimana proses perhitungan matematika pada proses *inference* diubah menjadi perkalian *integer* yang normalnya dilakukan dengan perkalian *floating-point*. Benoit Jacob et al. berhasil menurunkan ukuran dan latency pada model MobileNet SSD [10].

1.2 Rumusan Masalah

Dengan latar belakang yang dipaparkan di atas ditemukan masalah-masalah yang akan dijawab dan diselesaikan pada penelitian ini. Beberapa masalahnya yaitu.

1. Metode *pruning* apa yang dapat menurunkan FLOP pada model klasifikasi *hate speech*?

2. Metode apa yang dapat menurunkan ukuran model dan *inference time* pada model klasifikasi *hate speech*?
3. Seberapa baik performa model yang dihasilkan penelitian ini dalam mengklasifikasikan *hate speech*?
4. Apakah model dapat melakukan klasifikasi dan deteksi unggahan yang ada di website X secara langsung?

1.3 Batasan Masalah

Beberapa batasan yang diberikan oleh peneliti kepada penelitian ini adalah.

1. Penelitian hanya akan melakukan klasifikasi pada unggahan berbentuk teks.
2. Penelitian hanya akan dilakukan pada media sosial X.
3. Metode optimalisasi yang akan diteliti akan dibatasi sebanyak 3 metode, yaitu *Structured Pruning*, *Unstructured Pruning* dan *Quantization*.
4. Peneliti tidak menggunakan metode Hyperparameter dalam pembuatan model.

1.4 Tujuan Penelitian

Tujuan yang ingin diraih oleh peneliti adalah mengembangkan *deep learning* yang mampu mengklasifikasikan *hate speech*. Peneliti juga ingin membandingkan metode pruning yang dapat mengurangi komputasi yang diperlukan untuk mengklasifikasikan ujaran kebencian pada media sosial X serta mencari manfaat metode quantization. Peneliti juga bertujuan untuk menghasilkan aplikasi berupa ekstensi *web browser* yang dapat mengklasifikasikan unggahan pada media sosial X secara langsung, memanfaatkan model *deep learning* yang dihasilkan dari penelitian ini.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah pengetahuan terhadap optimalisasi model *deep learning* sehingga peneliti berikutnya dapat melakukan optimalisasi pada model yang mereka kembangkan. Penelitian juga dapat mengkomparasi keefektifan metode-metode optimalisasi baru pada masa yang akan datang dengan metode-metode yang diteliti pada penelitian ini, terutama pada nilai-nilai yang berhubungan

dengan kebutuhan sumberdaya model. Seperti ukuran, komputasi yang dibutuhkan, dan *inference time* oleh model.

1.6 Sistematika Penulisan

Sistematika penulisan dalam skripsi ini terdiri dari lima bab yang diuraikan secara garis besar sebagai berikut.

1. Bab 1, Pendahuluan

Pada bab ini, peneliti membahas mengenai latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan. Bab ini memberikan landasan utama mengapa penelitian dilakukan dan apa yang ingin dicapai.

2. Bab 2, Tinjauan Pustaka

Bab ini memuat tinjauan pustaka yang berhubungan dengan penelitian. Dasar-dasar teori yang digunakan dalam penelitian dijelaskan secara rinci untuk memberikan pemahaman mendalam mengenai masalah yang dibahas, termasuk literatur dan referensi yang mendukung.

3. Bab 3, Metode Penelitian

Pada bab ini, dijelaskan tentang objek penelitian, alur penelitian, serta alat dan bahan yang digunakan dalam penelitian. Metode penelitian yang dipakai dijabarkan secara sistematis agar proses penelitian dapat diulang atau dikembangkan lebih lanjut.

4. Bab 4, Hasil dan Pembahasan

Bab ini berisi hasil-hasil penelitian yang diperoleh dari data yang telah diolah serta pembahasan yang dilakukan. Hasil tersebut dianalisis untuk menjawab tujuan penelitian.

5. Bab 5, Penutup

Pada bab terakhir ini, berisi kesimpulan dari penelitian yang telah dilakukan, sesuai dengan tujuan dan rumusan masalah. Bab ini juga memberikan saran untuk penelitian selanjutnya atau untuk pengembangan lebih lanjut dari hasil penelitian yang sudah ada.