

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penyebaran konten iklan judi *online* kian hari makin menakutkan. Di tahun 2020 sebanyak 14.726 ditemukan konten judi *online* mulai menyasar masyarakat pedesaan [1]. Angka tersebut sangatlah besar padahal baru berada di tingkat desa. Konten atau iklan terutama di platform digital terutama *website* sangat berdampak kepada kehidupan masyarakat kita, ditambah sebanyak 213 juta penduduk kita merupakan pengguna internet yang setiap harinya menjelajahi *website* tertentu [2]. Hal tersebut membuat efek buruk seperti *spam* dan juga dapat menambah pemain judi *online* karena mereka tertarik dengan iklan yang berdampak kepada kesehatan mental pengguna internet. Bahkan Indonesia yang mayoritas masyarakatnya beragama islam ini memiliki banyak pemain judi *online*. Padahal sudah sangat jelas larangan berjudi diatur oleh Allah SWT, dan salah satunya terdapat dalam firman Allah Qur'an Surat Al-Maidah ayat 90 tentang larangan judi dan minum khamar karena dianggap perbuatan tercela dan berdampak sia-sia yang menimbulkan penyesalan [3] [4].

Website menjadi salah satu tempat atau *platform* paling umum yang diakses oleh masyarakat, sehingga promosi judi *online* tersebar luas dan memiliki dampak besar pada orang yang membukanya [5]. Promotor akan sangat senang memasarkan iklan judi mereka agar makin banyak pemain judi *online* yang tertarik [6]. Mereka juga tidak akan segan-segan mengambil alih *website* tersebut hanya untuk menyebarkan iklan judinya. Penelitian ini akan berfokus pada *homepage website*. *Homepage website* merupakan sebuah tampilan paling awal dalam *website* yang berisikan beberapa informasi yang mana tampilan tersebut menjadi *first impression* ketika seseorang menjelajahi web tersebut untuk pertama kalinya . *Spam* iklan judi *online* akan sangat rawan terutama bagi *website* yang sudah diambil alih oleh pihak promotor judi. Dan akan sangat bahaya sekali apabila *website* tersebut merupakan milik pemerintah atau instansi lainnya Ketika ada gambar iklan judi *online* yang tertampil di dalamnya. Gambar 1.1 merupakan contoh salah satu iklan judi *online* di *homepage website*.



Gambar 1.1 Salah Satu Contoh Iklan Judi *Online* di *Homepage Website*

Maka itu, diperlukan sebuah sistem kontrol yang bisa meminimalisir tindak penyebaran citra iklan judi *online* di *platform* digital seperti *website*. Iklan yang dimaksud merupakan sebuah citra digital atau gambar yang berisikan promosi untuk mengajak para pengguna. Diperlukan sebuah sistem klasifikasi untuk memahami gambar iklan tersebut sekaligus mengurangi penyebaran iklan agar pemain judi *online* berkurang dan menghindari *spam*. Model klasifikasi tersebut dapat dibangun oleh AI (*Artificial Intelligence*) yang merupakan ilmu komputer dengan fokus pada pengembangan sistem yang dapat melakukan tugas seperti kecerdasan manusia. AI memiliki banyak struktur pembangun salah satunya adalah *Deep Learning* yang merupakan cabang dari *Machine Learning* [7] [8].

Keunggulan *Deep Learning* terletak pada kemampuannya untuk merepresentasikan data yang semakin abstrak dan kompleks seperti gambar karena memiliki vektor berdimensi tinggi yang memerlukan banyak parameter untuk menggambarkan jaringan melalui struktur *layer* dalam arsitektur *neural network*. *Neural Network* yang akan dipakai dalam penelitian ini adalah *Convolutional Neural Networks* (CNN) yang telah terbukti sangat efektif dalam berbagai tugas pengolahan gambar. CNN dapat secara otomatis mengekstraksi fitur-fitur yang relevan dari data visual, tanpa memerlukan ekstraksi fitur manual oleh manusia [9] [10] [11]. Contoh studi menggunakan citra *Chest X-ray* menunjukkan CNN memiliki akurasi lebih tinggi (0.9591) daripada KNN dan SVM (0.921 dan 0.93). Penelitian lain tentang klasifikasi citra daging sapi dan babi menemukan bahwa CNN memiliki akurasi klasifikasi 96.87%, melampaui SVM yang hanya mencapai 93.09% dengan jumlah 600 data yang diambil secara langsung di sebuah pasar tradisional. Secara keseluruhan, penelitian ini menegaskan keunggulan *Deep*

Learning dibandingkan dengan *Machine Learning* konvensional dalam hal kinerja [12].

1.2 Perumusan Masalah

Rumusan masalah yang terdapat pada penelitian ini berfokus kepada hal-hal yang menjadi sebuah pertanyaan mendalam bagaimana *Deep Learning* dengan *Convolutional Neural Network* bisa mengklasifikasikan iklan judi *online* tersebut. Fakta ditampilkan secara rinci dalam bentuk dugaan, untuk kemudian ditemukan sebuah solusi penyelesaian dari masalah tersebut. Berikut merupakan perumusan masalah yang dapat diambil untuk penelitian ini:

- a. Bagaimana menerapkan algoritma CNN untuk klasifikasi citra iklan judi *online* pada *homepage website*?
- b. Bagaimana kinerja dari algoritma CNN untuk klasifikasi citra iklan judi *online* pada *homepage website*?

1.3 Batasan Masalah

Ruang lingkup yang spesifik berisikan parameter-parameter yang akan dijelajahi seperti area geografis, waktu, populasi yang diteliti, dan juga variabel-variabel yang relevan pada penelitian ini adalah:

- a. Pengumpulan dataset untuk gambar diambil dari *Browser Chrome* saja dengan 557 data untuk 2 kelas. Dengan jumlah data citra judi *online* sebanyak 287 dan non judi *online* sebanyak 270 buah.
- b. Terdapat 2 kelas dalam penelitian ini, yaitu gambar iklan judi *online* dan gambar random apa saja yang mewakili konten pada *homepage* seperti orang, logo, dan iklan biasa.
- c. Uji coba klasifikasi dilakukan terhadap *homepage website* buatan sendiri agar tidak menyalahgunakan aturan privasi.
- d. Klasifikasi hanya bisa dilakukan terhadap bagian *homepage* saja (1 halaman).
- e. Pembangunan sistem akan dilakukan dalam bentuk model dengan *Jupyter Notebook* pada IDE *Visual Studio Code*.
- f. Sistem akan dikembangkan sampai pada tahapan pengembangan model.
- g. Model akan dibangun dengan sistem dapat menampilkan (*crawling*) seluruh gambar pada *homepage website* yang di dalamnya terdapat beberapa citra

iklan judi *online* yang telah diklasifikasi, dan akan muncul dibagian atas (*looping*).

- h. Apabila terdapat gambar yang tidak dapat ditampilkan (*crawling*), kemungkinan besar dikarenakan adanya protokol-protokol tertentu pada sebuah *website*, seperti Robots.txt, Metatags, HTTP Headers, Javascript Loading, dan lainnya.

1.4 Tujuan Penelitian

Berdasarkan perumusan masalah diatas, maka solusi yang dapat diambil untuk menyelesaikan masalah tersebut, yaitu:

- a. Menerapkan algoritma CNN (*Convolutional Neural Network*) untuk klasifikasi citra iklan judi *online* secara otomatis pada *homepage website*.
- b. Mengetahui kinerja *Deep Learning* dengan CNN pada model deteksi otomatis gambar iklan judi *online* pada *homepage website*.

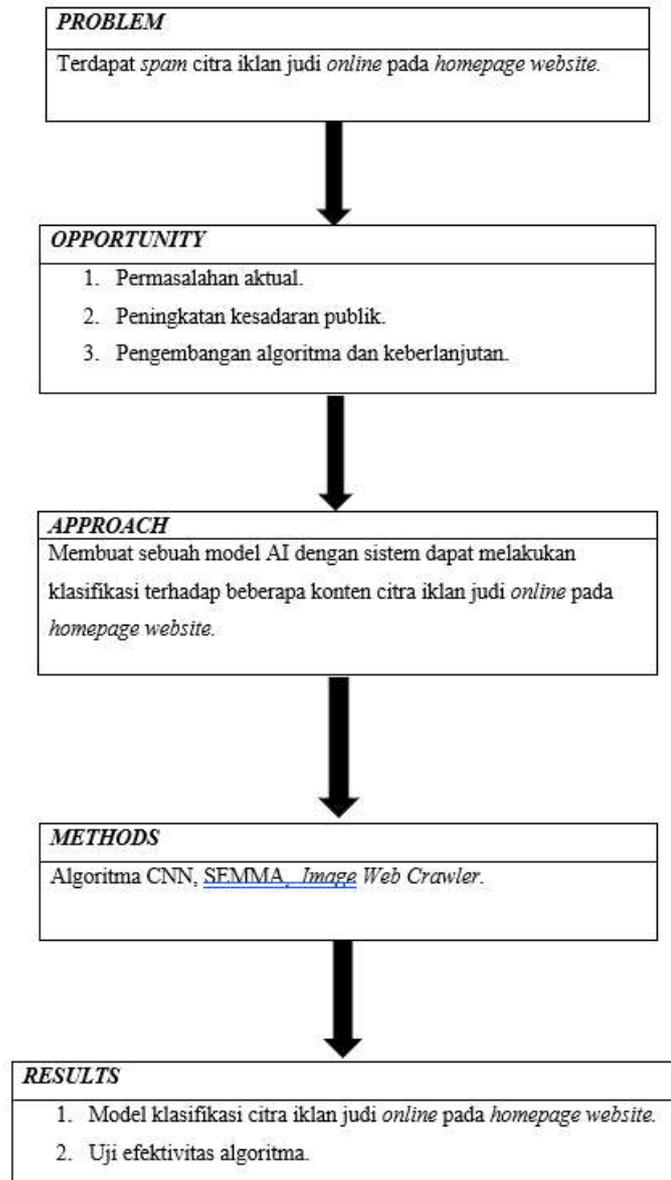
1.5 Manfaat Penelitian

Ada beberapa manfaat yang dapat diambil dari analisis dan penelitian ini, baik secara teoritis dan ilmu pengetahuan moral, etika, agama, dan untuk pengembangan model teknologi *Artificial Intelligence*, seperti:

- a. Penelitian ini berkontribusi pada pengembangan teknologi kecerdasan buatan, khususnya dalam bidang pemrosesan citra dan klasifikasi dengan CNN. Metodologi dan hasil penelitian juga dapat digunakan sebagai referensi untuk penelitian di masa depan.
- b. Dengan mengurangi eksposur terhadap iklan judi, penelitian ini membantu dalam upaya untuk mengurangi kecanduan judi dan dampak negatifnya pada masyarakat. Ini dapat berkontribusi pada kesejahteraan sosial dan kesehatan mental pengguna internet.

1.6 Kerangka Pemikiran

Sebuah sajian sistematis bagaimana proses terbentuknya sebuah pemikiran dari awal munculnya masalah, penemuan solusi dan dukungan terhadap solusi tersebut, tahapan untuk menyelesaikan solusi, juga hasil dari solusi yang telah dikerjakan untuk memecahkan masalah dituangkan dalam sebuah kerangka pikir atau pemikiran. Dalam Gambar 2.2 merupakan kerangka pemikiran yang digunakan dalam penelitian ini.



Gambar 1.2 Kerangka Pemikiran

1.7 Metode Penelitian

Metode penelitian yang digunakan pada pengembangan serta pengumpulan data dari penelitian ini adalah sebuah cara sistem kerja menyeluruh yang digunakan untuk menyelesaikan model serta mengevaluasikan kinerja dari model yang dibangun.

1.7.1 Pengumpulan Data dan Sumber

Beberapa sumber informasi untuk penelitian diambil melalui studi literatur, baik itu dari artikel, jurnal, maupun skripsi terkait penelitian. Untuk data dalam

pembangunan model, dibutuhkan sebanyak minimal 250 data gambar untuk setiap kelasnya. Data diperoleh dari Google Image untuk kemudian dilakukan observasi. Ada 2 kelas yang akan dibuat, yaitu foto iklan judi *online* dan foto random apa saja yang mewakili konten pada *homepage website*. Maka dari itu total untuk data diperkirakan mencapai 500 data foto, melihat dari sebuah penelitian yang dilakukan oleh Ocktavia Nurima Putri di tahun 2020 mengenai klasifikasi gambar jamur. Ia memakai data gambar jamur dengan dua jenis genus masing-masing berjumlah 208 gambar untuk dilatih [11].

1.7.2 Pengembangan Sistem

Pengembangan dilakukan untuk membangun sistem model secara terstrukturisasi dengan metode SEMMA (*Sample, Explore, Modify, Model, Access*). Urutan ini berisi tahap-tahap secara umum proses pengolahan data gambar iklan judi *online* dan non judi *online* serta pengembangan model. Kelebihan SEMMA ialah metode yang mudah dipahami dan dapat digunakan sebagai referensi untuk proyek data mining karena lebih terfokus pada aspek pengembangan model data *mining* yang lebih terstruktur [13]. Berikut merupakan beberapa tahapan proses pengembangan model dengan metode SEMMA:

a. *Sample*

Pada tahap ini seluruh data dari 2 kelas dikumpulkan untuk kemudian di analisis pada tahap eksplorasi.

b. *Explore*

Menganalisis dan melakukan eksplorasi untuk menemukan anomali atau keanehan pada data.

c. *Modify*

Memanipulasi dan memodifikasi data agar memenuhi syarat untuk masuk ke pemodelan.

d. *Model*

Membuat sebuah model yang akan digunakan untuk klasifikasi dengan menggunakan DNN (*Deep Neural Network*) untuk kemudian melakukan integrasi dengan *image web crawler*.

e. Access

Melakukan evaluasi dengan menilai kualitas, membuat keputusan hasil prediksi, dan menguji efisiensi terhadap *image web crawler*.

1.8 Sistematika Penulisan

Sistematika penulisan merupakan sebuah susunan seluruh bagian dalam penulisan tugas akhir ini dengan penyusunan informasi secara teratur. Berikut merupakan sistematika penulisan dalam menyusun tugas akhir ini:

BAB I PENDAHULUAN

Bagian ini berisi latar belakang, perumusan masalah, batasan masalah, tujuan, manfaat, kerangka pemikiran, metodologi pemikiran, dan sistematika penulisan yang menjadi cikal bakal terbentuknya penelitian tugas akhir ini.

BAB II KAJIAN LITERATUR

Pada bagian kajian literatur terdapat beberapa penelitian terdahulu serta dasar teori yang dijadikan pedoman untuk menyusun tugas akhir ini. Pedoman tersebut berasal dari jurnal, artikel ilmiah, dan buku terkait penelitian.

BAB III METODOLOGI PENELITIAN

Bagian ini berisi eksekusi dalam pengembangan sistem model dengan metode SEMMA yang dimulai dari tahapan awal yaitu *sample* sampai dengan *Access* untuk evaluasi model.

BAB IV HASIL DAN PEMBAHASAN

Hasil dan pembahasan berisi pemaparan bagian *evaluation* serta kinerja dari sistem yang dibangun.

BAB V KESIMPULAN DAN SARAN

Bagian ini berisi seluruh kesimpulan dari hasil penelitian serta saran untuk penelitian berikutnya agar bisa selalu terbarukan dan bermanfaat bagi umat.

DAFTAR PUSTAKA

Daftar Pustaka merupakan daftar sumber-sumber yang dipakai, dikutip dan dijadikan acuan atau pedoman untuk penulisan tugas akhir ini.

