

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Dalam beberapa tahun terakhir, fenomena meningkatnya jumlah individu yang beralih menjadi *content creator* telah menjadi tren global. Berdasarkan data yang dipublikasikan oleh Howarth (2024), ekonomi kreator terus mengalami pertumbuhan pesat dengan nilai pasar diperkirakan mencapai \$480 miliar pada tahun 2028. Survei yang dilakukan oleh ConvertKit terhadap lebih dari 2.700 kreator mengungkapkan bahwa 46,7% kreator merupakan kreator penuh waktu (*full-timers*), 42,7% kreator paruh waktu (*part-timers*), dan hanya 10,6% yang menganggap diri mereka sebagai kreator hobi (*hobbyists*). Platform seperti YouTube, TikTok, dan Instagram menjadi pendorong utama pertumbuhan ini, dengan YouTube sendiri memiliki lebih dari 51 juta channel aktif. Kondisi ini menciptakan persaingan yang semakin ketat di antara para kreator, terutama bagi mereka yang baru memulai, untuk dapat memperoleh perhatian audiens di tengah melimpahnya konten yang tersedia.

Salah satu tantangan utama bagi kreator, khususnya kreator pemula adalah mengembangkan kemampuan desain visual mereka secara mandiri agar dapat menghasilkan karya yang tidak hanya menarik secara visual, tetapi juga bermakna dan orisinal (Alam & Haikal, 2024). Sebuah survey dari RevoU (2024) yang melibatkan lebih dari 1.000 partisipan menemukan bahwa 67% responden lebih cenderung mengklik video dengan thumbnail yang menarik perhatian. Namun, banyak kreator, khususnya kreator pemula, masih menghadapi kesulitan dalam memahami karakteristik audiens mereka. Informasi yang tersedia melalui fitur analitik platform sering kali hanya memberikan gambaran umum mengenai demografi dan tingkat interaksi audiens, sehingga kreator masih kesulitan memahami alasan di balik ketertarikan audiens terhadap suatu konten serta menerjemahkan informasi tersebut menjadi strategi pembuatan konten yang tepat (Choi dkk, 2025). Selain itu, kreator pemula juga menghadapi tantangan dalam mengidentifikasi audiens dan menentukan strategi komunikasi yang sesuai ketika membangun akun media sosial dari awal (Sanjaya dkk, 2026). Kondisi tersebut dapat menyebabkan konten yang dibuat menjadi

terlalu umum dan kurang sesuai dengan kebutuhan atau minat audiens, sehingga berpotensi memengaruhi tingkat eksposur dan performa konten (Damayanti, 2024). Desain *thumbnail* yang lebih baik, misalnya melalui penggunaan warna cerah, teks singkat yang informatif, dan elemen visual yang membangkitkan rasa penasaran, telah terbukti dapat meningkatkan performa video youtube secara signifikan (Rasjid, 2024). Algoritma platform seperti YouTube cenderung memprioritaskan konten dengan performa tinggi dari kreator mapan. Akibatnya, kreator baru (pemula) sering kali tidak mendapatkan eksposur yang memadai dalam sistem rekomendasi (Amu dkk, 2024). Oleh karena itu, penting untuk memahami elemen yang menjadi daya tarik hingga memengaruhi performa video, seperti jumlah penayangan (*view count*), judul, serta *thumbnail* yang dapat menjadi indikator awal keberhasilan konten dalam menjangkau audiens.

Sebagai solusi atas permasalahan tersebut, proyek ini menawarkan penerapan model prediksi berbasis algoritma *Convolutional Neural Network* (CNN) untuk menganalisis elemen visual pada *thumbnail*. Algoritma CNN dipilih karena kemampuannya dalam mengekstrak fitur visual secara otomatis dari gambar seperti warna, tata letak, ekspresi wajah, dan komposisi visual (Masna, 2025). Namun, daya tarik video tidak hanya ditentukan oleh aspek visual semata. Menurut Cui dkk. (2024), judul yang menarik dan mengandung emosi positif atau negatif dapat meningkatkan daya tarik dan klik video. Oleh karena itu, proyek ini juga mengintegrasikan pendekatan *Natural Language Processing* (NLP) untuk menganalisis teks judul video secara bersamaan dengan gambar *thumbnail*. Dengan menggabungkan analisis visual dan non-visual dalam satu model prediksi multimodal, proyek ini bertujuan untuk memprediksi performa video secara lebih baik.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, maka perumusan masalah dalam proyek ini adalah:

1. Kreator pemula masih mengalami kesulitan dalam menentukan *thumbnail* dan judul video yang menarik dan sesuai.
2. Kreator pemula belum memahami karakteristik *thumbnail* dan judul yang berkaitan dengan performa video YouTube.

1.3 Batasan Masalah

Adapun batasan masalah dalam pembuatan proyek akhir ini adalah:

1. Adapun batasan masalah dalam pembuatan proyek akhir ini adalah:
2. Penelitian ini hanya memfokuskan pada klasifikasi performa video YouTube berdasarkan dua elemen utama, yaitu desain visual thumbnail dan teks judul video, tanpa mempertimbangkan elemen lain seperti deskripsi, tag, durasi video, waktu unggahan, dan isi video.
3. Data yang digunakan terbatas pada metadata publik dan gambar thumbnail dari video YouTube, serta tidak mencakup data dari platform lain seperti TikTok atau Instagram.
4. Dataset yang digunakan terdiri dari 2.620 video yang diperoleh dari 14 channel YouTube, yaitu Deddy Corbuzier, Najwa Shihab, Raditya Dika, Gita Wirjawan, Ferry Irwandi, Raymond Chin, Bennix, Ricis Official, Nihongo Mantappu, Timothy Ronald, Rans Entertainment, Narasi, Trifellas, dan MALAKA.
5. Video yang digunakan dalam penelitian dibatasi pada video yang dipublikasikan dalam rentang waktu 9 Juli 2024 hingga 8 Juli 2025.
6. Penelitian ini hanya dilakukan pada video YouTube yang termasuk dalam kategori *People & Blogs*, *News & Politics*, dan *Entertainment*.
7. Model yang digunakan menggunakan pendekatan deep learning berbasis algoritma CNN untuk pemrosesan gambar thumbnail dan embedding layer yang dikombinasikan dengan CNN untuk pemrosesan teks judul, tanpa menggunakan teknik lanjutan seperti Transformer atau pretrained language model.
8. Judul video dianalisis berdasarkan representasi semantik menggunakan teknik embedding tanpa mempertimbangkan konteks emosional, gaya bahasa, maupun makna bahasa yang lebih kompleks.
9. Performa video dikelompokkan ke dalam beberapa kategori berdasarkan proses pelabelan menggunakan algoritma K-Means Clustering.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

Adapun tujuan dari penelitian ini adalah:

1. Mengembangkan model klasifikasi multimodal (visual dan teks).
2. Menerapkan CNN untuk analisis thumbnail dan judul secara bersamaan.
3. Menghasilkan model untuk mengidentifikasi kategori performa video berdasarkan thumbnail dan judul.

1.4.2 *Manfaat*

Manfaat yang diharapkan dari penelitian ini adalah

1. Menghasilkan model klasifikasi multimodal yang telah terbukti mampu mengenali performa video berdasarkan karakteristik thumbnail dan judul.
2. Memberikan gambaran mengenai perbandingan dua metode pelabelan performa video, yaitu berbasis *view ratio* dan berbasis *K-Means Clustering*, yang dapat menjadi pertimbangan dalam pengembangan model klasifikasi performa konten digital berikutnya.