

BAB 1 PENDAHULUAN

1.1 Latar Belakang Masalah

Crowdfunding telah menjadi salah satu alternatif pendanaan digital yang semakin populer karena memungkinkan individu dan pelaku usaha memperoleh modal secara langsung dari masyarakat luas melalui media daring. *Crowdfunding* memberikan kesempatan bagi lembaga pendidikan dan individu untuk memperoleh pendanaan yang sebelumnya sulit didapatkan melalui jalur konvensional seperti perbankan atau bantuan pemerintah [1].

Meskipun demikian, keberhasilan kampanye *crowdfunding* tidak hanya ditentukan oleh media yang digunakan dan kebijakan yang berlaku, tetapi juga dipengaruhi oleh berbagai faktor lain, seperti kesiapan ekosistem pendukung, karakteristik proyek yang ditawarkan, strategi penyampaian informasi, serta respons dari calon pendana [2][3]. Oleh karena itu, diperlukan pemahaman yang lebih komprehensif mengenai faktor-faktor yang berperan dalam menentukan keberhasilan suatu kampanye *crowdfunding*.

Sebagai salah satu media *crowdfunding* terbesar di dunia, Kickstarter menampung jumlah kampanye yang sangat banyak dengan tingkat keberhasilan yang berbeda-beda antar kategori maupun periode pelaksanaannya. Penelitian sebelumnya menyebutkan bahwa Kickstarter mampu menghimpun pendanaan hingga “ratusan juta dolar setiap tahun”, yang menunjukkan tingginya volume transaksi pendanaan yang terjadi pada media tersebut serta besarnya tingkat partisipasi pengguna dalam mendukung berbagai proyek [4]. Hasil penelitian empiris awal juga menunjukkan bahwa tingkat keberhasilan kampanye di Kickstarter berada pada kisaran satu dari dua kampanye yang berhasil mencapai target pendanaan, yang mengindikasikan adanya peluang keberhasilan sekaligus risiko kegagalan yang cukup signifikan [5]. Selain itu, penelitian terbaru pada tahun 2024 menggunakan lebih dari 108.000 data proyek Kickstarter untuk mengevaluasi model prediksi, sehingga menunjukkan besarnya skala data yang tersedia untuk dianalisis secara komputasional [6]. Studi lainnya juga membatasi ruang lingkup penelitian pada kategori tertentu dan rentang waktu beberapa tahun, seperti pada proyek teknologi yang dianalisis dalam periode 2009 hingga 2023 [7]. Variasi cakupan data tersebut menunjukkan bahwa faktor waktu dan kategori proyek turut

memengaruhi hasil analisis, sehingga diperlukan pendekatan yang relevan dan mutakhir dalam memprediksi keberhasilan kampanye *crowdfunding*.

Meskipun jumlah kampanye terus meningkat, masih banyak proyek yang gagal mencapai target pendanaan akibat kurangnya pemahaman terhadap faktor-faktor yang memengaruhi keberhasilan kampanye. Beberapa penelitian menunjukkan bahwa kegagalan kampanye *crowdfunding* berkaitan erat dengan lemahnya indikator prediktor keberhasilan yang seharusnya dapat meningkatkan peluang pendanaan. Salah satu faktor utama yang sering ditemukan adalah rendahnya jumlah pendukung dana pada tahap awal kampanye serta lemahnya momentum penggalangan dana. Jumlah *backer* menjadi salah satu prediktor terpenting dalam menentukan keberhasilan proyek [8][9]. Selain itu, pembaruan yang rendah, penentuan durasi kampanye yang tidak optimal, serta ketiadaan elemen visual terbukti berkaitan dengan meningkatnya risiko kegagalan kampanye [9]. Kualitas deskripsi kampanye yang lemah, termasuk argumentasi proyek yang kurang kuat dan rendahnya kredibilitas sumber informasi, turut berkorelasi dengan menurunnya probabilitas pendanaan [10]. Selain itu, aspek waktu dan perhatian publik juga memengaruhi hasil kampanye. Proyek yang gagal menarik perhatian sejak awal atau diluncurkan pada periode dengan tingkat persaingan tinggi cenderung memiliki risiko kegagalan yang lebih besar [6].

Kemajuan teknologi *data mining* dan *machine learning* memberikan peluang untuk memanfaatkan data historis *crowdfunding* dalam pengembangan model prediksi keberhasilan kampanye yang lebih akurat. Sejumlah penelitian menunjukkan bahwa metode berbasis *ensemble*, seperti Random Forest, memiliki kemampuan yang baik dalam mengolah data tabular yang berasal dari media Kickstarter [9][10]. Random Forest dilaporkan mampu menghasilkan tingkat akurasi yang tinggi, yaitu mencapai 87.85%, serta menunjukkan kinerja yang lebih unggul dibandingkan beberapa algoritma lain, seperti Neural Network dan Decision Trees [6]. Hasil tersebut mengindikasikan bahwa Random Forest efektif dalam menangkap hubungan yang kompleks dan non-linear antar variabel pada data *crowdfunding*. Sementara itu, Logistic Regression masih banyak digunakan sebagai model pembanding karena memiliki struktur yang sederhana serta mampu memodelkan probabilitas keberhasilan secara efektif [5]. Beberapa penelitian terdahulu juga melaporkan bahwa Logistic Regression dapat memberikan performa yang kompetitif dengan tingkat akurasi sebesar 88.2%, bahkan mencapai 94.1% pada pengujian eksternal dalam domain tertentu, seperti industri teknologi [11]. Secara umum, temuan-temuan tersebut menunjukkan bahwa performa model

prediksi tidak hanya dipengaruhi oleh algoritma yang digunakan, tetapi juga oleh karakteristik dataset, teknik *preprocessing*, dan metode evaluasi yang diterapkan. Oleh sebab itu, penggunaan dataset yang lebih mutakhir menjadi aspek penting untuk menghasilkan model prediksi yang mampu merepresentasikan kondisi aktual secara lebih akurat [5][8].

Model prediksi yang akurat dapat membantu pemilik proyek dalam merancang strategi kampanye yang lebih efektif sebelum peluncuran, seperti penentuan target dana, durasi kampanye, dan kategori proyek. Berbagai penelitian menunjukkan bahwa hasil prediksi serta analisis tingkat kepentingan fitur mampu memberikan panduan praktis dalam pengambilan keputusan desain kampanye. Model prediksi juga dapat mengidentifikasi target pendanaan yang optimal berdasarkan karakteristik proyek dan pola historis [6]. Selain itu, pengaturan durasi kampanye, dinamika waktu penggalangan dana, serta strategi pembaruan konten terbukti berpengaruh terhadap peluang keberhasilan [9][5]. Faktor lain seperti kategori proyek dan penggunaan elemen multimedia juga berkontribusi signifikan terhadap performa kampanye [10]. Lebih lanjut, model berbasis deret waktu memungkinkan deteksi dini terhadap performa kampanye, sehingga penyesuaian strategi dapat dilakukan secara lebih cepat [5]. Di sisi lain, analisis teks dapat memberikan umpan balik terkait kualitas deskripsi proyek dan daya persuasi pesan [12]. Maka dari itu, model prediksi dapat berperan sebagai alat bantu dalam merumuskan strategi kampanye yang lebih efektif dibandingkan hanya digunakan sebagai meningkatkan akurasi *forecasting*.

Namun, penelitian sebelumnya masih menggunakan dataset dengan periode yang relatif lama serta belum sepenuhnya merepresentasikan dinamika terbaru pada media *crowdfunding* yang terus berkembang. Beberapa studi menggunakan data historis yang terbatas hingga sebelum tahun 2024, bahkan hanya menganalisis kampanye hingga Februari 2019, sehingga kurang mencerminkan kondisi terkini [7]. Selain itu, penelitian sebelumnya masih memiliki keterbatasan terkait cakupan data yang digunakan, yang hanya berasal dari periode 2013–2016. Kondisi tersebut dapat menyebabkan hasil penelitian kurang representatif dalam menggambarkan karakteristik dan dinamika *crowdfunding* pada periode yang lebih mutakhir [5]. Perubahan tren *crowdfunding*, perilaku pengguna, serta perkembangan teknologi dan strategi kampanye memungkinkan adanya perbedaan pola keberhasilan proyek dari waktu ke waktu. Tidak hanya itu, sejumlah penelitian lebih menekankan pada peningkatan performa model melalui pengembangan metode atau penggabungan berbagai jenis data, seperti analisis teks dan pendekatan multimodal, tanpa

diimbangi dengan analisis faktor-faktor dominan yang memengaruhi keberhasilan kampanye secara komprehensif [5][10][13].

Di sisi lain, beberapa penelitian menunjukkan bahwa variabel *pledge* dan jumlah *backers* merupakan faktor yang sangat berpengaruh terhadap keberhasilan kampanye [14]. Namun, variabel-variabel tersebut tidak tersedia pada tahap awal sebelum peluncuran karena merepresentasikan kondisi selama atau setelah kampanye berlangsung. Selain itu, variabel *spotlight* juga memiliki keterkaitan langsung dengan hasil akhir kampanye karena menunjukkan status proyek yang telah mencapai target pendanaan. Penggunaan variabel *pledge*, *backers*, dan *spotlight* sebagai fitur input berpotensi menimbulkan *data leakage*, yaitu kondisi model mendapatkan informasi yang berkaitan langsung dengan label target. Penelitian sebelumnya menunjukkan bahwa penambahan variabel-variabel tersebut dapat meningkatkan performa model secara signifikan, namun peningkatan tersebut disebabkan oleh kebocoran informasi yang memungkinkan model mempelajari hubungan langsung dengan hasil kampanye [8].

Berdasarkan uraian tersebut, penelitian terbaru diperlukan yang berfokus pada evaluasi performa model dan mempertimbangkan pemilihan fitur yang relevan serta dapat mengidentifikasi faktor-faktor yang paling berpengaruh terhadap keberhasilan kampanye. Penelitian ini menggunakan dataset Kickstarter tahun 2025 dengan pendekatan *machine learning* untuk membangun model prediksi keberhasilan kampanye *crowdfunding*. Melalui analisis pengaruh fitur yang digunakan dalam model, penelitian ini diharapkan mendapatkan pemahaman yang lebih baik dalam mengetahui faktor-faktor yang berkontribusi terhadap keberhasilan kampanye sekaligus menghasilkan model prediksi yang relevan dengan kondisi terkini.

1.2 Rumusan Masalah

Mengacu pada latar belakang yang telah dipaparkan sebelumnya, maka permasalahan yang dikaji dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana kinerja algoritma Random Forest dan Logistic Regression dalam memprediksi keberhasilan kampanye *crowdfunding*?
2. Algoritma manakah yang memberikan tingkat akurasi dan performa terbaik berdasarkan *confusion matrix*?

3. Fitur atau variabel apa saja yang paling berpengaruh terhadap prediksi keberhasilan kampanye pada model terbaik yang dihasilkan?

1.3 Batasan Permasalahan

Untuk menjaga fokus penelitian dan mencegah pembahasan yang meluas di luar cakupan yang ditentukan, batasan-batasan dalam penelitian ini ditetapkan sebagai berikut:

1. Dataset yang digunakan dalam penelitian ini bersumber dari kampanye *crowdfunding* di media Kickstarter yang dipublikasikan sepanjang tahun 2025.
2. Penelitian difokuskan pada klasifikasi keberhasilan kampanye ke dalam dua kategori, yaitu *successful* dan *failed*.
3. Fitur yang digunakan sebagai variabel masukan dibatasi pada data numerik dan kategorikal yang tersedia sebelum periode kampanye berakhir.
4. Pendekatan *machine learning* yang digunakan dalam penelitian ini mencakup Logistic Regression dan Random Forest.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang sudah dibuat, berikut tujuan penelitian antara lain:

1. Mengimplementasikan algoritma Random Forest dan Logistic Regression untuk membangun model prediksi keberhasilan kampanye.
2. Menentukan nilai akurasi terbaik yang dihasilkan oleh algoritma Random Forest dan Logistic Regression.
3. Mengidentifikasi fitur-fitur yang paling berpengaruh terhadap prediksi keberhasilan kampanye.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Menyediakan informasi mengenai algoritma klasifikasi yang menunjukkan kinerja paling optimal dalam memprediksi keberhasilan kampanye *crowdfunding* melalui perbandingan antara Random Forest dan Logistic Regression.
2. Menghasilkan model prediksi yang dapat dimanfaatkan untuk mengestimasi tingkat keberhasilan kampanye *crowdfunding* berdasarkan karakteristik proyek yang tersedia.
3. Memberikan pemahaman mengenai faktor-faktor yang memiliki pengaruh signifikan terhadap keberhasilan kampanye *crowdfunding*, sehingga dapat menjadi bahan pertimbangan bagi kreator dalam merancang strategi kampanye yang lebih efektif.
4. Menambah referensi ilmiah yang dapat digunakan sebagai dasar bagi penelitian selanjutnya terkait prediksi keberhasilan *crowdfunding* dengan pendekatan *machine learning*.

1.6 Sistematika Penulisan

Sistematika susunan penulisan penelitian ini adalah sebagai berikut:

- Bab 1 PENDAHULUAN
Bab ini menjelaskan dasar pelaksanaan penelitian, yang berisi pembahasan latar belakang masalah mengenai perkembangan *crowdfunding* sebagai salah satu alternatif pendanaan digital dan permasalahan keberhasilan kampanye pada media Kickstarter. Selain itu, bab ini juga berisi rumusan masalah, batasan penelitian, tujuan penelitian, manfaat penelitian, serta sistematika penulisan skripsi.
- Bab 2 LANDASAN TEORI
Bab ini berisi landasan teori yang menjadi dasar dari penelitian, meliputi konsep *crowdfunding*, penggunaan media Kickstarter, metode klasifikasi, serta algoritma Logistic Regression dan Random Forest. Bab ini juga membahas metrik evaluasi model dan juga terdapat kajian penelitian-penelitian terdahulu yang relevan dengan topik penelitian.
- Bab 3 METODOLOGI PENELITIAN
Bab ini menjelaskan metodologi penelitian yang diterapkan, mulai dari proses

pengumpulan data, tahap *preprocessing*, pemilihan dan transformasi fitur, pembagian data menjadi data latih dan data uji, hingga proses pembangunan model menggunakan algoritma *machine learning*. Selain itu, dijelaskan pula prosedur evaluasi yang digunakan untuk menilai performa model.

- Bab 4 HASIL DAN DISKUSI

Bab ini menyajikan hasil yang telah diimplementasikan serta diujikan pada model yang telah dikembangkan. Pembahasan meliputi hasil evaluasi masing-masing algoritma pemodelan, perbandingan performa model, serta analisis berdasarkan metrik evaluasi yang digunakan. Selain itu, bab ini juga menguraikan faktor-faktor yang paling berpengaruh terhadap keberhasilan kampanye *crowdfunding* berdasarkan hasil analisis fitur.

- Bab 5 KESIMPULAN DAN SARAN

Bab ini menyajikan simpulan yang diperoleh dari hasil penelitian yang telah diterapkan serta menjawab rumusan masalah dan tujuan penelitian. Bab ini juga memberikan saran dan rekomendasi yang dapat digunakan sebagai acuan bagi peneliti-peneliti selanjutnya.

