

BAB I

PENDAHULUAN

A. Latar Belakang

Selama beberapa dekade, pemerintah di seluruh dunia melakukan dorongan kepada masyarakat untuk meningkatkan keuangan jangka panjang akibat krisis keuangan yang terjadi tahun 2008 [1]. Demi meningkatkan keuangan jangka panjang, diperlukan literasi finansial yang memadai. Literasi finansial merupakan pemahaman akan pengelolaan keuangan yang baik untuk kehidupan sehari-hari demi meningkatkan kesejahteraan individu maupun masyarakat. Namun, masyarakat di Indonesia memiliki literasi finansial yang rendah. Hal ini dibuktikan oleh hasil *survey* Otoritas Jasa Keuangan (OJK) yang menunjukkan bahwa rata-rata literasi finansial masyarakat Indonesia hanya mencapai 49,68%, sementara inklusi keuangan mencapai 85,10%. Hal ini berarti masyarakat Indonesia mempunyai akses ke layanan keuangan, namun pengetahuan dan keterampilan dalam mengelola keuangan masih sangat kurang [2].

Data dari OJK pada tahun 2023 menunjukkan bahwa sebagian besar permasalahan finansial di Indonesia didominasi oleh rentang usia 19-34 atau generasi Z, dengan tingkat literasi rata-rata nasional sebesar 49,68% dan inklusi finansial 85,10%. Hal ini menunjukkan bahwa terdapat kesenjangan mengenai pengetahuan keuangan. Generasi Z merupakan generasi yang dilahirkan pada masa resesi ekonomi, hal ini menumbuhkan sifat kehati-hatian dan cemas terhadap risiko serta cenderung mencari suatu informasi melalui internet dan sosial media [3]. Akses terhadap internet tersebut membuat generasi ini menjadi pengguna layanan pinjaman *online* terbesar [2], dengan 80% dari mereka menggunakan aplikasi fintek (finansial teknologi) dan 60% melakukan transaksi dalam aplikasi tersebut. Penggunaan uang tunai pun menurun 40% dalam dua tahun terakhir dan 45% dari generasi Z melakukan investasi melalui saham. Hal

ini menunjukkan ketertarikan generasi tersebut terhadap investasi walaupun terdapat kesenjangan literasi keuangan yang mempengaruhi cara menabung [4].

Salah satu akibat dari rendahnya literasi finansial adalah konsumerisme. Hal ini mengakibatkan dampak merugikan bagi diri sendiri maupun negara. Salah satu contoh akibat dari konsumerisme adalah krisis keuangan pribadi. Ketika seseorang memiliki utang yang sangat tinggi, sebagian besar pendapatannya akan digunakan untuk membayar utang-utang dan bunga dari utang tersebut. Hal ini dapat mengurangi konsumsi dan investasi, akibatnya, uang yang dihasilkan menjadi tidak dapat digunakan untuk hal yang lebih produktif dan perputaran uang dalam masyarakat menjadi berkurang. Dampak ini juga menyebar hingga skala nasional, di mana menurut studi Kobayashi dan Shirai pada tahun 2017 memaparkan ketika utang menumpuk secara masif, maka perputaran uang akan terfokus untuk pembayaran utang dan bunga bukan pengeluaran untuk barang, jasa, maupun investasi. Akibatnya permintaan barang dan jasa menurun, produksi melambat dan perusahaan, mengurangi tenaga kerja [2].

Melihat pengetahuan literasi finansial Generasi Z yang rendah dan penggunaan fintek (finansial teknologi) yang kurang bijak akibat kurangnya pengetahuan literasi finansial, diperlukan pendekatan menggunakan data untuk mengetahui perilaku menabung Generasi Z. Model *machine learning* dengan metode *unsupervised learning* dipilih karena dapat melihat pola tersembunyi dalam data. Pola tersembunyi ini berupa sikap keuangan, pengetahuan keuangan, perilaku keuangan dan Materialisme. Hal ini akan mengungkapkan perilaku menabung Generasi Z dan melakukan konfirmasi survei OJK. Selain itu, hasil dari model ini juga bermanfaat bagi pemerintah dan pendidik untuk memberikan edukasi lebih dini mengenai pentingnya menabung dan literasi keuangan.

B. Rumusan Masalah

Berdasarkan uraian pada bagian latar belakang, dapat dirumuskan permasalahan mengenai bagaimana merancang model *machine learning* yang

mampu mengelompokkan perilaku menabung generasi Z secara optimal, dengan menghasilkan klaster-klaster yang memiliki tingkat homogenitas internal yang tinggi?

C. Batasan Masalah

Batasan masalah yang akan diungkapkan dalam tugas akhir ini sebagai berikut:

1. Model ini akan menggunakan *machine learning* dengan metode *unsupervised learning*. *Unsupervised learning* digunakan untuk melihat pola tersembunyi dalam data.
2. Data berasal dari survei yang diberikan kepada Generasi Z. Data tersebut berupa kuesioner mengenai tingkat pengetahuan, perilaku, sikap finansial serta materialisme dari generasi Z.
3. Generasi Z secara umum memiliki rentang 13 hingga 28 tahun. Namun, dalam konteks ini berumur 18 hingga 23 atau menempuh pendidikan perkuliahan di semester 1 hingga 8.
4. Hanya menggunakan satu model *machine learning* dengan tingkat separasi terbaik. Model menggunakan *agglomerative clustering* dengan *linkage* berjenis Ward.
5. Model hanya mendeteksi perilaku Generasi Z, tidak dapat melakukan prediksi dengan masukan dari pengguna.
6. Data bersifat tetap dan berfokus terhadap deteksi perilaku menabung generasi Z.

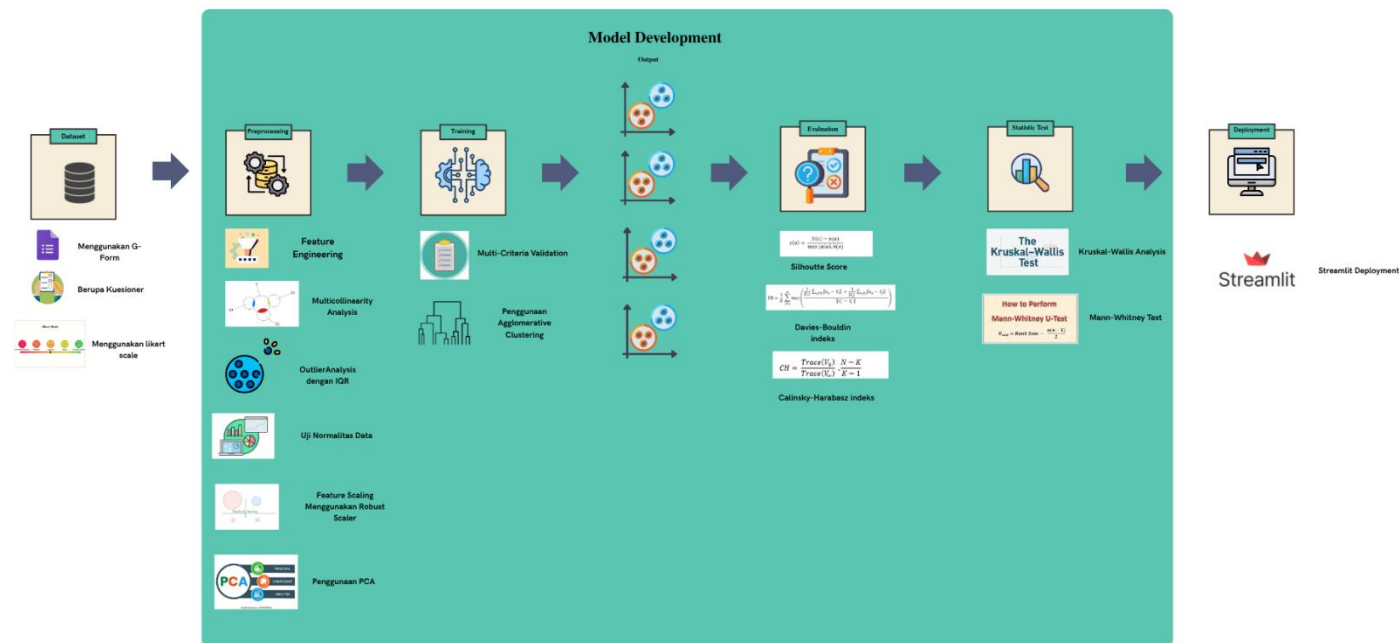
D. Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Mengembangkan model *machine learning* dengan metode *unsupervised learning* guna mengelompokkan perilaku menabung generasi Z

E. Metode Penelitian

Metode penelitian yang akan digunakan dalam tugas akhir ini adalah:



Gambar 1.1. Metode Penelitian

1. Dataset

Pada Gambar 1.1 terdapat tahap dataset, tahap ini merupakan tahap pengumpulan data untuk memperoleh *dataset* yang akan digunakan untuk pelatihan model. *Dataset* diambil dari survei yang disebar ke Generasi Z. *Dataset* ini diambil melalui kuesioner yang dibuat berdasarkan arahan dari OECD (Organization for Economic Co-operation and Development). Kuesioner ini menggunakan jawaban berdasarkan *likert-scale* dan akan menilai tingkat pengetahuan, perilaku dan sikap finansial serta Materialisme dari Generasi Z. Untuk mendapatkan responden, kuesioner ini akan di sebar via *google form*.

2. Preprocessing Data

Tahap *preprocessing data* dilakukan untuk meningkatkan kualitas data. Proses ini meliputi:

a. Feature engineering

Tahap ini merupakan tahap yang digunakan untuk membuang fitur yang kurang berguna pada dataset. Penelitian ini menggunakan teknik *feature selection* untuk mengambil fitur yang penting seperti *fa1* hingga *15*, *fb1* hingga *fb13*, *fk1* hingga *fk9*, dan *m1* hingga *m10*.

b. Reverse coding pada pilihan likert

Pilihan *likert* memiliki pertanyaan yang memiliki makna terbalik dibandingkan dengan pertanyaan lainnya. Hal ini ditujukan agar semua pertanyaan memiliki interpretasi yang konsisten. *Reverse coding* dilakukan untuk membalik kembali skor dari pertanyaan yang dilakukan *reverse coding*.

c. Pengecekan normalitas data

Uji normalitas data digunakan untuk mendapatkan *insight* mengenai bentuk dari data. Digunakan metode Kolmogorov-Smirnov untuk mengecek normalitas data.

d. Pengecekan multicollinearity

Pengecekan ini menggunakan metode *pairwise* yang digunakan untuk mengecek korelasi antar fitur. Bila korelasi antar fitur tinggi, akan menyebabkan informasi yang redundan dan mempengaruhi keakuratan dari model.

e. *Scaling Data*

Scaling data merupakan teknik transformasi nilai-nilai pada fitur numerik pada *dataset* agar berada pada rentang yang sama. Contoh dari metode *scaling data* adalah *Z-score Normalization*, *Min-Max Normalization* dan *Robust scaling*. Pada penelitian ini menggunakan teknik *Robust scaler* karena terdapat banyak *outlier* yang dideteksi.

f. Penggunaan PCA (*Principal Component Analysis*)

Penggunaan PCA digunakan untuk melakukan reduksi dimensi dari data. Reduksi ini digunakan untuk mengurangi noise pada data.

2. *Training*

Pelatihan merupakan tahap untuk melakukan pembangunan model pada *machine learning*. Pada tahap ini, algoritma *agglomerative clustering* digunakan untuk menganalisis pola menabung Generasi Z.

3. *Evaluasi*

Tahap ini merupakan proses untuk mengevaluasi kualitas hasil pengelompokan yang telah dihasilkan oleh model. Evaluasi dilakukan dengan menggunakan metrik internal seperti *Silhouette Score*, *Davies-Bouldin Index*, dan *Calinski-Harabasz Index*, yang bertujuan untuk mengukur homogenitas data dan keterpisahan antar data.

4. *Statistic test*

Tahap ini adalah proses analisis lanjutan setelah proses evaluasi selesai dilakukan. Pada tahap ini, dilakukan interpretasi lebih mendalam terhadap hasil klaster, termasuk mengidentifikasi karakteristik unik pada setiap klaster, serta menganalisis pola-pola perilaku yang muncul. Tahap ini menggunakan algoritma *Kruskal-Wallis* dan *Man-Whitney U test*. Kedua algoritma ini

digunakan karena tidak memperhatikan distribusi normal data dan tahan terhadap *outlier*.

5. *Deployment*

Web deployment adalah *platform* yang digunakan untuk melihat analisa dari setiap model. *Web deployment* ini berisi :

a. *Data Explorer*

Data Explorer merupakan halaman pada web yang berfungsi untuk melihat isi dari *dataset*. Halaman ini juga berisi *insight* dari *dataset* yang akan digunakan dalam analisis. *Insight* berupa grafik *bar chart*, *line chart* maupun *Area Chart*.

b. *Preprocessing*

Halaman *preprocessing* merupakan halaman yang digunakan untuk melakukan persiapan data sebelum dilakukan pelatihan data. Persiapan data ini berguna agar data dapat diproses dalam halaman *training*. Halaman ini juga digunakan untuk melihat variabel apa saja yang memiliki *multicollinearity*.

c. *Training*

Halaman *Training* berisi mengenai penggunaan algoritma *unsupervised learning*. Halaman ini juga digunakan untuk melakukan evaluasi jumlah kluster yang terbentuk, visualisasi kluster dan *spider-plot* untuk mengevaluasi pola antar fitur. Halaman ini juga terdapat analisa kluster berdasarkan domain. Tujuan analisa domain ini adalah untuk menganalisis perbedaan kluster berdasarkan domain secara jelas.

d. *Post Hoc*

Halaman *Post Hoc* merupakan halaman yang digunakan untuk melakukan analisis data yang telah dilakukan pelatihan. Halaman ini berisikan analisa Kruskal-Wallis dan Man-Whitney U test.

e. *Artificial Intelligence (AI) Analysis*

Halaman *AI Analysis* merupakan halaman yang digunakan untuk mencari *insight* dari data dengan melakukan integrasi dengan *gemini AI*. Halaman ini dibuat hanya sebagai *state of art* dari halaman web.

F. Sistematika Penulisan

Penelitian ini disusun secara sistematis untuk memberikan pemahaman mengenai pengembangan model *machine learning* dengan metode *unsupervised learning* guna menganalisis perilaku menabung generasi Z :

BAB I PENDAHULUAN

Pendahuluan berisi latar belakang dari pembuatan model *machine learning* dengan metode *unsupervised learning* guna menganalisis perilaku menabung generasi Z, Permasalahan yang ada, dan pemberian solusi dari penelitian ini. Bab ini juga berisi rumusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Tinjauan pustaka berisi berbagai literatur maupun penelitian terdahulu. Bab ini digunakan sebagai pedoman maupun referensi untuk melakukan penelitian. Bab ini berisi mengenai penelitian terdahulu, penggunaan algoritma *unsupervised*, dan penelitian yang mempelajari perilaku menabung gen Z.

BAB III LANDASAN TEORI

Landasan Teori berisi teori-teori mengenai literasi finansial, definisi generasi z , dan definisi *machine learning*. Teori-teori ini bermanfaat sebagai pedoman untuk membuat penelitian ini.

BAB IV *DATASET* DAN PENGEMBANGAN MODEL

Dataset dan pengembangan model berisi mengenai masalah yang lebih teknis. Bab ini akan menjelaskan mengenai isi dan analisa dari *dataset* yang digunakan. Bab ini juga akan menjelaskan mengenai pengembangan model *unsupervised learning*.

BAB V IMPLEMENTASI MODEL DAN PENGEMBANGAN SISTEM

Tahap ini mencakup proses implementasi model yang telah dikembangkan ke dalam sistem. Implementasi model ini akan berisi hasil analisa perilaku menabung generasi Z.

BAB VI PENUTUP

Penutup merupakan bagian dari kesimpulan penelitian yang telah dibuat. Bagian ini juga berisi saran terkait dari bagaimana penelitian yang telah dilakukan.

