

Investigating Bank Clients Deposit Behavior: A Descriptive and Predictive Analysis

Annisa Sekartierra Mulyanto, Marta Afifah

1. LATAR BELAKANG

Pada era globalisasi saat ini, persaingan antarperusahaan semakin ketat sehingga diperlukan strategi *marketing* yang tepat dan inovatif untuk memperluas pasar dan meningkatkan keuntungan perusahaan. Salah satu perusahaan yang perlu memperhatikan strategi *marketing* adalah perusahaan di industri perbankan. Perusahaan harus memiliki strategi *marketing* yang tepat dan efektif untuk menggaet banyak nasabah dan meningkatkan keuntungan.

Dengan perkembangan teknologi yang semakin cepat, antarperusahaan perbankan berlomba-lomba mengembangkan strategi *marketing* yang unik dan menarik untuk menggaet nasabah. Akan tetapi, strategi *marketing* yang unik dan menarik saja tidak akan cukup efektif untuk memperluas nasabah. Oleh karena itu, diperlukan analisis lebih lanjut untuk mengetahui segmentasi dan karakteristik nasabah agar strategi *marketing* yang dilakukan tepat sasaran dan efektif dalam mencapai tujuan perusahaan.

Segmentasi nasabah adalah proses membagi pasar menjadi beberapa segmen yang homogen dan mempunyai karakteristik yang sama. Dengan demikian, bank dapat lebih efektif dalam mengkomunikasikan produk dan layanan yang sesuai dengan kebutuhan dan preferensi masing-masing segmen nasabah. Berbagai penelitian sebelumnya telah menunjukkan bahwa segmentasi nasabah dapat meningkatkan kepuasan nasabah dan loyalitas mereka terhadap bank. Misalnya, penelitian oleh Takaliuang dkk. (2020) menunjukkan bahwa segmentasi pasar dapat meningkatkan nasabah sebesar 52.4%. Selain itu, penelitian oleh Purnomo dkk. (2015) juga menunjukkan bahwa segmentasi berdasarkan preferensi konsumen dapat membantu bank dalam meningkatkan keuntungan.

Oleh karena itu, dalam laporan ini, dilakukan analisis data dari Bank Portugis untuk mengetahui segmentasi dan karakteristik nasabah. Dari analisis ini, ingin diketahui teknik *marketing* yang tepat untuk menarik nasabah agar melakukan deposito ke bank. Selain itu, juga ingin diketahui teknik *marketing* yang tepat untuk menarik orang-orang yang berpotensi menjadi nasabah di bank. Hasil analisis ini diharapkan dapat membantu bank dalam meningkatkan keuntungan dan meningkatkan kepuasan nasabah.

2. DATA DAN VARIABEL

Data yang digunakan dalam analisis ini diperoleh dari *website* UCI Machine Learning dengan judul “Bank Marketing”. Data ini memuat 45211 data nasabah dan 17 variabel sebagai berikut.

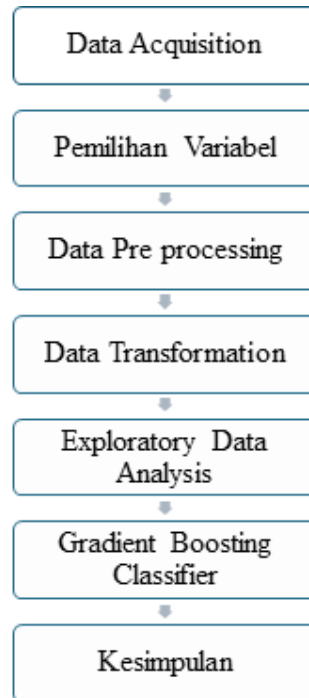
<i>Variabel</i>	Deskripsi
<i>Y</i>	apakah klien berlangganan deposito berjangka
<i>Age</i>	umur nasabah
<i>Job</i>	pekerjaan nasabah
<i>Marital</i>	status pernikahan nasabah
<i>Education</i>	status pendidikan nasabah
<i>Default</i>	apakah kredit sudah <i>default</i> ?
<i>Balance</i>	saldo rata-rata tahunan
<i>Housing</i>	memiliki pinjaman perumahan?
<i>Loan</i>	memiliki pinjaman pribadi?
<i>Contact</i>	jenis komunikasi kontak
<i>Day</i>	hari kontak terakhir dalam sebulan
<i>Month</i>	bulan kontak terakhir dalam setahun
<i>Duration</i>	durasi kontak terakhir
<i>Campaign</i>	jumlah kontak yang dilakukan selama kampanye ini dan untuk nasabah ini
<i>Pdays</i>	jumlah hari yang berlalu setelah nasabah terakhir dihubungi dari kampanye sebelumnya
<i>Before</i>	jumlah kontak yang dilakukan sebelum kampanye ini dan untuk nasabah ini
<i>Poutcome</i>	hasil kampanye pemasaran sebelumnya

Akan tetapi, pada analisis ini hanya digunakan 9 variabel, yaitu *age*, *job*, *marital*, *education*, *default*, *balance*, *housing*, *loan*, dan *y*. Selain variabel yang disebutkan, tidak digunakan dalam analisis ini karena variabel-variabel tersebut hanya terfokus pada strategi *marketing* melalui kontak telepon, sedangkan analisis kali ini ingin diketahui strategi *marketing* lain yang tepat, tidak hanya melalui kontak telepon.

3. METODE ANALISIS

3.1 Kerangka Berpikir

Berikut adalah algoritma metode analisis data yang dilakukan:



3.2 Gradient Boosting Classifier

Dikutip dari Trivusi, 2023, Gradient Boosting Classifier adalah salah satu algoritma *machine learning* yang digunakan untuk melakukan analisis klasifikasi. Algoritma ini bekerja dengan cara menggabungkan beberapa model yang lemah (*weak learners*) menjadi sebuah model yang kuat. Dalam setiap iterasi, Gradient Boosting Classifier akan menambahkan *weak learner* baru dan mengoreksi prediksi sebelumnya dengan memperhitungkan kesalahan pada prediksi tersebut. Secara matematis, Gradient Boosting mengoptimalkan suatu fungsi objektif dengan mengevaluasi gradien pada setiap titik. Salah satu fungsi objektif yang biasanya digunakan dalam algoritma ini berupa *logistic loss*, yaitu:

$$\text{Log loss} = -(y_i * \log(\text{odds}) + \log(1 - p)),$$

dengan *odds* dan p adalah probabilitas prediksi kelas pertama dan *odds* adalah $\text{odds} = \frac{p}{1-p}$.

Gradien dari *loss function* ini dihitung terhadap prediksi model. Kemudian, model diperbarui dengan menambahkan model baru yang ditimbang sesuai gradien yang dihitung.

4. HASIL DAN PEMBAHASAN

4.1 Analisis Deskriptif dengan *Exploratory Data Analysis*

Analisis deskriptif dilakukan menggunakan *dashboard* pada Tableau Public. Analisis ini dibagi menjadi dua bagian, yaitu analisis secara keseluruhan dan analisis berdasarkan kelas deposito (ya atau tidak).

Analisis secara keseluruhan

Analisis secara keseluruhan dimulai dengan menunjukkan proporsi dari kelas deposito, yaitu seberapa banyak nasabah yang melakukan deposito dan yang tidak melakukan deposito.

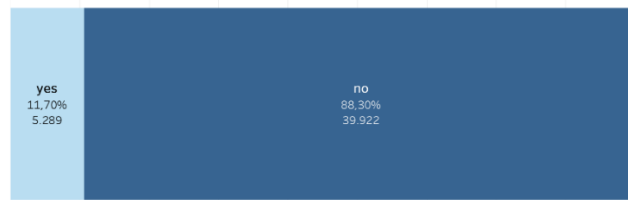


Figure 1. Overall Proportion of Depositors

Figure 1 menunjukkan bahwa secara umum proporsi nasabah bank yang melakukan deposito relatif sangatlah kecil yaitu hanyalah 11,7% dari keseluruhan nasabah. Sedangkan, nasabah yang tidak mendepositokan tabungannya sebanyak 88,3% dari keseluruhan nasabah.

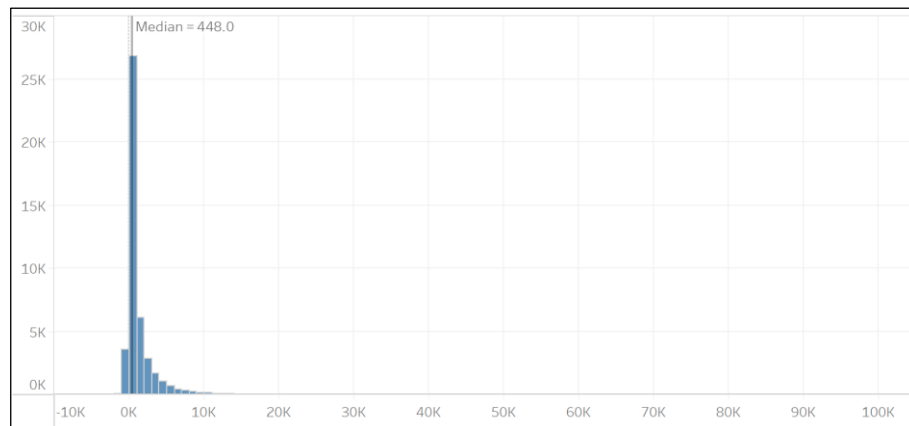


Figure 2. Client's Balance Information : Distribution

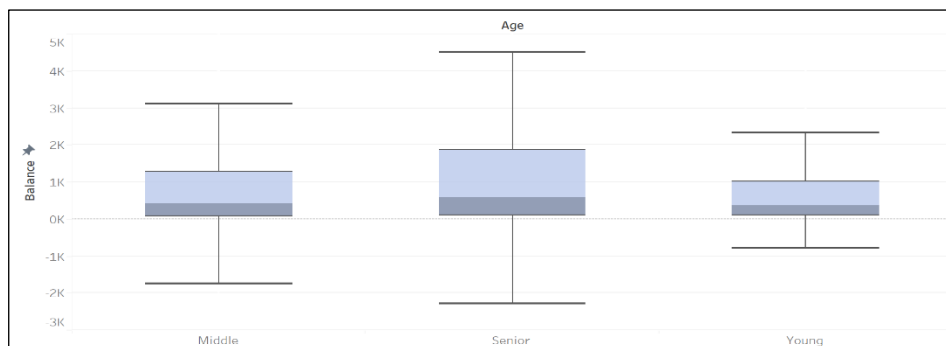


Figure 3. Balance Distribution based on Age

Variabel yang menunjukkan banyak hal menarik adalah balance. **Figure 2** memperlihatkan balance dari keseluruhan nasabah yang memiliki distribusi yang menjurai ke atas dimana balance terpusat pada nilai yang rendah dan lebih tersebar pada nilai yang tinggi. Hal menarik lainnya juga terlihat pada *boxplot* yang menunjukkan distribusi balance untuk setiap kelompok umur nasabah. **Figure 3** menunjukkan bahwa distribusi balance untuk nasabah pada kelompok umur Senior memiliki sebaran yang paling besar. Apabila dibandingkan dengan kelompok umur Young dan Middle, kelompok umur Senior memiliki balance yang terpusat di angka tinggi dimana mediannya adalah 580 dan *IQR Box* untuk Senior terletak relatif di atas *IQR Box* kedua kelompok umur lainnya.

Kelompok umur Young memiliki balance yang lebih terfokus dan terpusat pada angka rendah dimana mediannya adalah 361. *IQR Box* dan *Whisker* pada kelompok umur ini menunjukkan bahwa balance memiliki sebaran yang paling kecil diantara kelompok umur yang lain. Kelompok umur Middle memiliki nilai balance yang berada diantara kelompok umur Senior dan Young, begitu juga dengan sebarannya yang ukurannya berada diantara kedua kelompok umur yang lain.

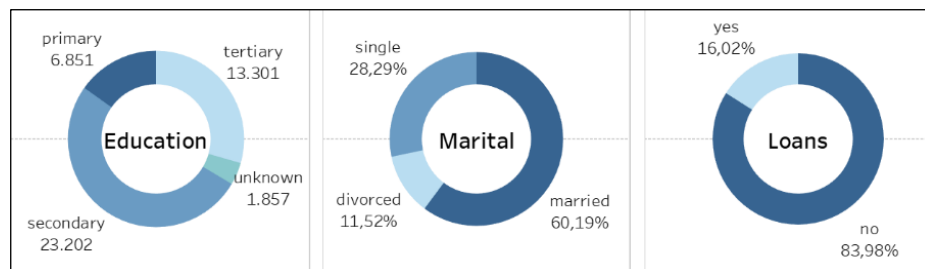


Figure 4. Client's Profiling 1

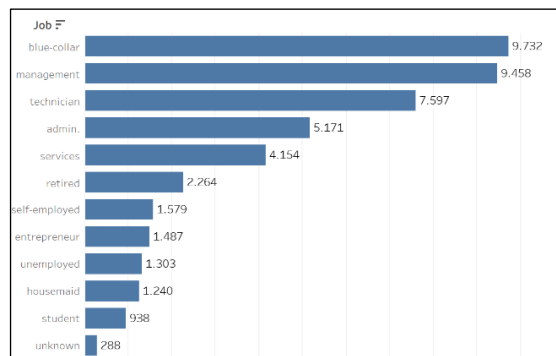


Figure 5. Client's Profiling 2

Selain Balance, penting bagi penulis untuk memperhatikan karakteristik dari nasabah. Beberapa dari karakteristik nasabah secara umum ditampilkan pada **Figure 4** dan **Figure 5**. Didapatkan bahwa persentase atau frekuensi terbesar untuk setiap variabel yaitu nasabah dengan status pendidikan *secondary*, telah menikah, tidak memiliki pinjaman, dan memiliki pekerjaan *blue-collar*.

Analisis berdasarkan kelas deposito

Penulis masih memiliki ketertarikan yang sama pada hubungan setiap variabel dan karakteristik nasabah dengan analisis secara keseluruhan yang telah dilakukan. Distribusi balance untuk kedua keterangan deposit tidak berbeda jauh tetapi tetap terlihat bahwa median untuk nasabah yang tidak deposito lebih kecil dibandingkan nasabah yang deposito ditunjukkan pada **Figure 6**.

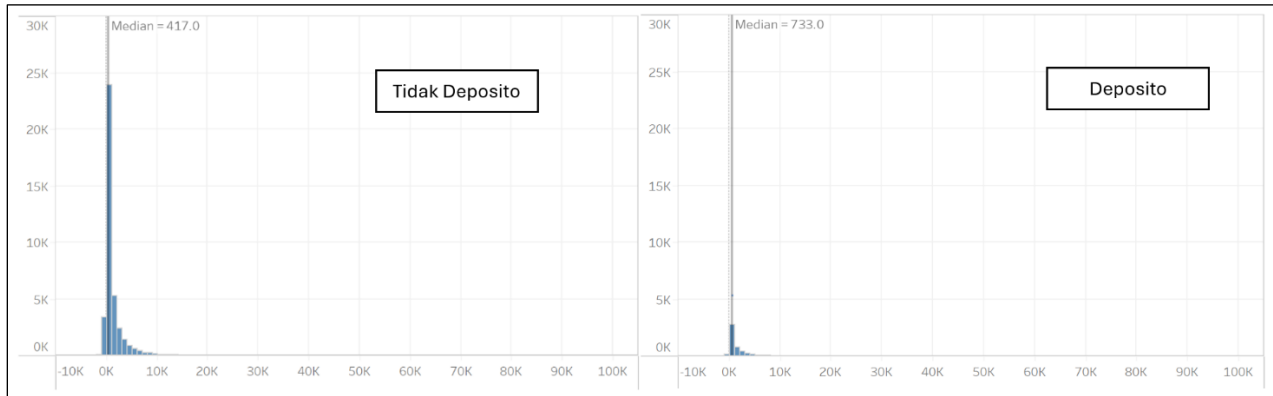


Figure 6. Client's Balance Information: Depositors and Non-Depositors

Status edukasi, pernikahan, dan pinjaman nasabah yang melakukan deposit dan yang tidak melakukan deposit tidak terdapat perbedaan yang signifikan ditunjukkan pada **Figure 7**. Namun, untuk pekerjaan terdapat beberapa perbedaan untuk kedua kejadian deposito. Hal ini terlihat pada **Figure 8** dimana frekuensi pekerjaan terbanyak untuk nasabah yang melakukan deposit adalah *blue-collar* sedangkan untuk nasabah yang tidak melakukan deposit adalah manajemen.

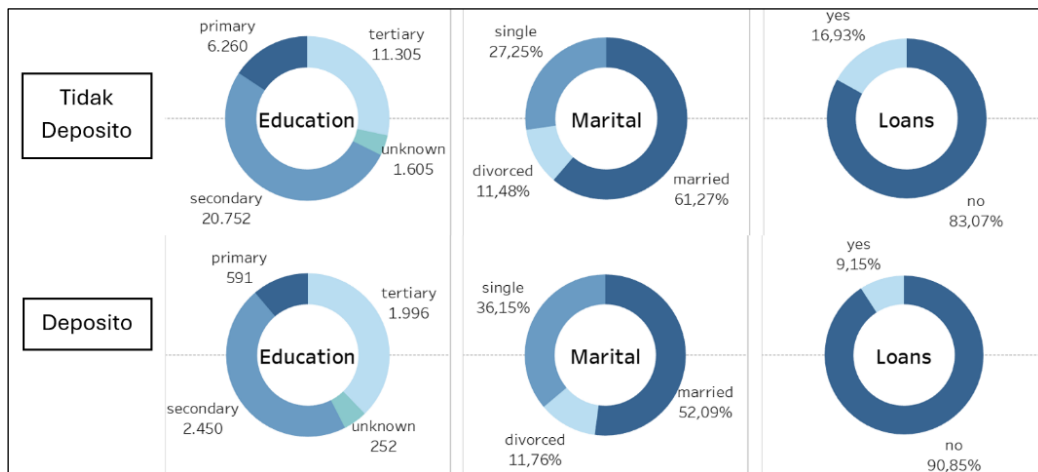


Figure 7. Client's Profiling for each Deposit Class 1

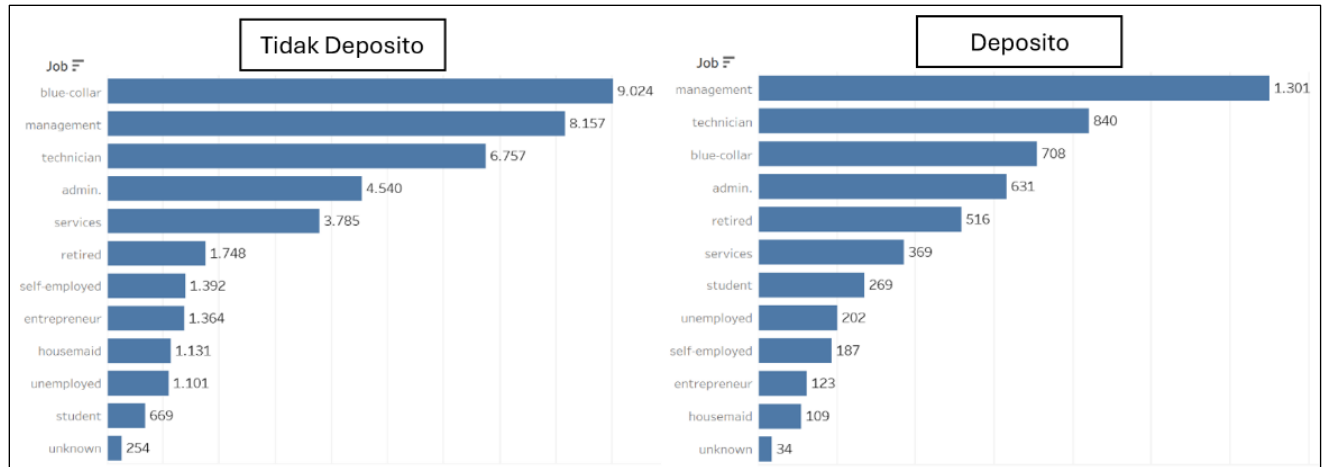


Figure 8. Client's Profilling for each Deposit Class 2

4.2 Analisis Prediktif dengan Klasifikasi: Gradient Boosting Classifier

Analisis ini menggunakan pemrograman Python dan menghasilkan model dengan akurasi sebagai berikut:

<i>Metrik Evaluasi</i>	<i>Nilai</i>
<i>Accuracy</i>	88,43%
<i>F1 Score</i>	0,76%
<i>Precision</i>	80%
<i>Recall</i>	0,38%

Model ini mampu memprediksi dengan benar, nasabah yang akan melakukan deposito atau tidak dengan akurasi sebesar 88,43%. Dengan kemampuan model dalam memprediksi benar untuk kelas positif (ya diprediksi ya) sebesar 80%.

Hasil analisis klasifikasi ini dapat digunakan untuk memprediksi nasabah baru yang kemungkinan besar melakukan deposito berdasarkan data historis nasabah-nasabah sebelumnya. Nasabah yang diprediksi akan melakukan deposito inilah yang dapat dijadikan target *marketing* utama perusahaan bank, sedangkan nasabah yang diprediksi tidak akan melakukan deposito dapat diberikan penawaran produk lain dari bank, sesuai dengan preferensi nasabah tersebut.

5. ACTIONABLE INSIGHT

Dari analisis yang telah dilakukan terdapat beberapa informasi yang didapatkan dari data. Berikut adalah informasi-informasi yang telah didapatkan dan aksinya.

1. Proporsi nasabah yang tidak melakukan deposit sangatlah kecil. Artinya, perlu adanya strategi pemasaran yang lebih kuat dalam mempromosikan deposito pada bank.
2. Karakteristik nasabah

- a. Kelompok umur senior memiliki saldo rekening yang lebih besar diantara yang lain. Artinya, bank perlu menggaet lebih banyak calon nasabah kelompok umur Senior sebagai nasabah.
 - b. Kelompok umur Young memiliki saldo rekening yang lebih kecil. Bank dapat membuat produk baru yang lebih sesuai untuk nasabah dengan kelompok umur Young.
3. Karakteristik depositor
 - a. Memiliki saldo rekening yang cenderung lebih tinggi. Bank dapat meningkatkan dan memfokuskan pemasaran mengenai deposito ke nasabah dengan saldo akun tinggi
 - b. Mayoritas memiliki pekerjaan *blue-collar*. Bank dapat melakukan pemasaran pada kantor-kantor pekerja *blue-collar*.
4. Karakteristik non-depositor
 - a. Memiliki saldo rekening yang cenderung lebih rendah. Bank dapat membuat strategi atau produk baru yang lebih sesuai untuk nasabah yang memiliki saldo rendah dikarenakan nasabah dengan saldo rendah memiliki kemungkinan kecil untuk menjadi depositor.
 - b. Mayoritas memiliki pekerjaan manajemen. Bank bisa mengurangi pemasaran pada kantor pekerjaan manajemen untuk mengurangi biaya dan efektifitas pemasaran bank.
5. Hubungan antara balance dan age memberi informasi bahwa bank perlu melakukan pemasaran yang lebih giat pada kelompok umur senior demi menggaet nasabah yang memiliki kemungkinan tinggi untuk menjadi depositor.
6. Hasil algoritma klasifikasi dapat digunakan untuk mengidentifikasi nasabah yang cenderung akan menjadi depositor dan yang tidak. Apabila nasabah memiliki kecenderungan menjadi depositor maka bank dapat memfokuskan pemasaran ke nasabah-nasabah tersebut.

6. DEPLOYMENT

Deployment dari hasil analisis berfokus pada bagian pemasaran. Bank dapat melakukan pemasaran yang lebih efektif dan efisien baik dari segi waktu maupun biaya. Untuk nasabah atau calon nasabah dengan kelompok umur Senior, bank dapat menggencarkan pemasaran via iklan TV, koran, atau majalah dimana kemungkinan besar masyarakat kelompok umur Senior melihat dan membacanya.

Algoritma klasifikasi juga dapat selalu diperbarui menggunakan data-data terkini yang diperoleh bank dari nasabah baru sehingga akan terus *update* dan relevan dengan keadaan masyarakat atau calon nasabah atau nasabah pada waktu tersebut. Algoritma ini juga dapat

diintegrasikan dengan *m-banking* untuk iklan terpersonalisasi berdasarkan identitas dari nasabah yang telah didaftarkan pada *m-banking*.

7. KESIMPULAN

Analisis terhadap data nasabah Bank dilakukan secara deskriptif dan prediktif. Hasil menunjukkan bahwa mayoritas nasabah memilih deposito, dengan proporsi yang kecil tidak melakukan deposit. Kelompok Senior cenderung memiliki saldo lebih besar, sementara nasabah muda memiliki saldo lebih kecil. Depositor cenderung memiliki saldo yang lebih tinggi dan mayoritas bekerja sebagai pekerja *blue-collar*, sementara non-depositor cenderung memiliki saldo rendah dan bekerja dalam bidang manajemen. Bank dapat meningkatkan pemasaran kepada nasabah senior dan *blue-collar*, serta fokus pada nasabah potensial yang memiliki kecenderungan menjadi depositor. Dengan strategi pemasaran yang lebih tepat sasaran, bank dapat meningkatkan daya tarik produk dan layanannya.

Didapatkan juga informasi melalui analisis prediktif dimana telah dibentuk suatu model klasifikasi yang akan mengelompokkan nasabah menjadi debitur dan non-debitur. Hasil analisis klasifikasi ini dapat memprediksi nasabah baru yang kemungkinan besar akan melakukan deposito berdasarkan data historis. Nasabah yang diprediksi akan melakukan deposito menjadi target utama pemasaran bank, sedangkan yang tidak diprediksi untuk deposito dapat ditawarkan produk bank lain sesuai preferensi mereka.

Selain itu, karakteristik nasabah yang telah diperoleh melalui analisis deskriptif dapat digunakan untuk memperluas nasabah. Informasi ini dapat dijadikan dasar bagi bank untuk melakukan strategi pemasaran yang lebih tepat sasaran untuk mendapatkan nasabah baru.

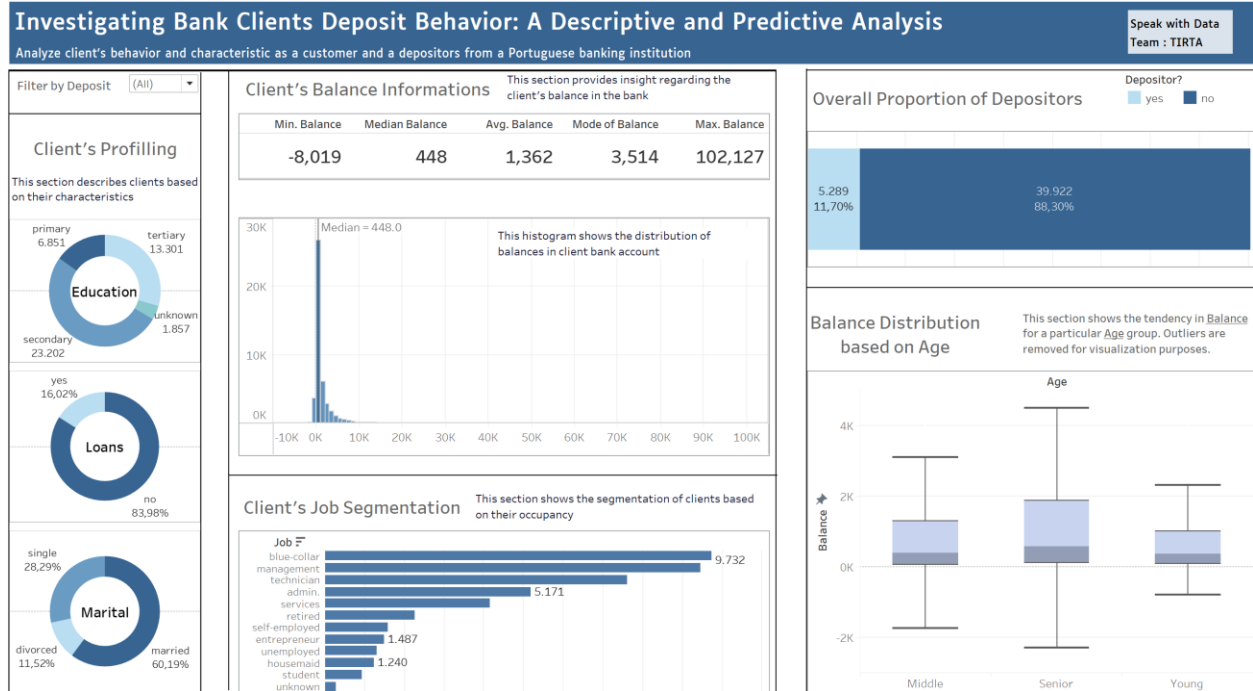
DAFTAR PUSTAKA

<https://dcb.co.id/perbankan/cara-kerja-marketing-bank-untuk-mencapai-target>

- Purnomo, H. (2010). ANALISIS PREFERENSI KONSUMEN SEBAGAI DASAR PENENTUAN SEGMENTASI DI PERBANKAN SYARIAH (Studi Kasus PT Bank Muamalat Indonesia Cabang Solo). *Jurnal Ekonomi dan Kewirausahaan Vol. 10, No. 1*, 22-31. Retrieved from Neliti.
- S. Moro, P. R. (2012, 2 13). *Bank Marketing*. Retrieved from UCI Machine Learning Repository: <https://archive.ics.uci.edu/dataset/222/bank+marketing>
- Sofie C. Takaliuang, T. M. (2020). Pengaruh Segmentasi Pasar Terhadap Peningkatan Nasabah PT. Bank SulutGo Cabang Utama Manado. *E-Journal UNSRAT*.
- Tomonori, M. (2022, Februari 8). *All You Need to Know about Gradient Boosting Algorithm – Part 2. Classification*. Retrieved from Towards Data Science: <https://towardsdatascience.com/all-you-need-to-know-about-gradient-boosting-algorithm-part-2-classification-d3ed8f56541e>
- Trivusi. (2023, Maret 12). *Gradient Boosting: Pengertian, Cara Kerja, dan Kegunaannya*. Retrieved from Trivusi: https://www.trivusi.web.id/2023/03/algoritma-gradient-boosting.html#google_vignette

LAMPIRAN

Dashboard Tableau Public



Dashboard Tableau Public dapat diakses pada link berikut : <https://bit.ly/DashboardTIRTA>