

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi digital dalam beberapa tahun terakhir telah mengubah lanskap industri jasa keuangan secara fundamental, khususnya dalam hal strategi akuisisi dan retensi pelanggan. Di era transformasi digital saat ini, perusahaan pembiayaan tidak lagi dapat mengandalkan pendekatan pemasaran konvensional yang bersifat massal dan tidak terarah. Sebaliknya, strategi berbasis data (data-driven strategy) dan kecerdasan buatan (artificial intelligence) telah menjadi fondasi utama dalam mengoptimalkan proses penjualan dan meningkatkan efisiensi operasional. Penelitian menunjukkan bahwa organisasi yang menerapkan strategi pemasaran berbasis data mampu mencapai peningkatan Return on Investment (ROI) hingga 15% dibandingkan pendekatan konvensional, sekaligus meningkatkan efisiensi alokasi sumber daya pemasaran secara signifikan [1] [2]. Salah satu area yang mengalami transformasi paling signifikan adalah pengelolaan lead pelanggan melalui sistem Customer Relationship Management (CRM) yang didukung oleh teknologi machine learning untuk memprediksi perilaku konsumen dan memprioritaskan tindak lanjut penjualan secara otomatis dan akurat [3].

Menurut Peraturan Otoritas Jasa Keuangan (OJK) Nomor 29/POJK.05/2014 tentang Penyelenggaraan Usaha Perusahaan Pembiayaan, perusahaan pembiayaan adalah badan usaha yang melakukan kegiatan pembiayaan dalam bentuk penyediaan dana atau barang modal untuk kebutuhan konsumen. Dalam konteks industri pembiayaan kendaraan bermotor, perusahaan dituntut untuk mampu mengelola volume calon pelanggan (leads) yang besar dengan tingkat efisiensi tinggi, mengingat persaingan yang ketat dan margin keuntungan yang semakin tipis [4]. Lebih lanjut, seiring dengan akselerasi digitalisasi sektor keuangan Indonesia yang diproyeksikan tumbuh dengan CAGR sekitar 9.31% hingga tahun 2033, perusahaan multifinance dihadapkan pada tantangan untuk mengadopsi inovasi

teknologi guna meningkatkan aksesibilitas layanan sekaligus mempertahankan efisiensi operasional [5]. Hal ini menuntut adanya sistem yang dapat mengidentifikasi dan memprioritaskan lead dengan potensi konversi tinggi secara akurat dan efisien berbasis analitik prediktif.

PT XYZ FINANCE merupakan salah satu perusahaan pembiayaan terkemuka di Indonesia yang tergabung dalam grup otomotif nasional ternama. Perusahaan mengelola berbagai lini bisnis pembiayaan, termasuk pembiayaan kendaraan bermotor roda dua (sepeda motor), elektronik, dan pembiayaan multiguna lainnya. Khusus pada lini bisnis pembiayaan sepeda motor, PT XYZ FINANCE mengandalkan sistem CRM telesales untuk melakukan follow-up terhadap calon pelanggan yang telah mengajukan aplikasi pembiayaan. Namun, berdasarkan data historis perusahaan, efektivitas proses telesales masih menghadapi tantangan serius. Tingkat keberhasilan dalam menghubungi lead (contactability rate) berada pada kisaran yang belum optimal, sementara tingkat konversi dari lead yang berhasil dihubungi menjadi pelanggan aktual (booking rate) masih berada di batas bawah standar industri pembiayaan. Kondisi ini mengindikasikan adanya inefisiensi signifikan dalam proses komunikasi dan prioritas follow-up yang berdampak langsung pada produktivitas tim telesales dan pencapaian target penjualan perusahaan.

Salah satu penyebab utama rendahnya tingkat konversi adalah belum diterapkannya sistem prioritas yang sistematis dan berbasis prediksi dalam proses follow-up lead. Saat ini, tim telesales melakukan panggilan terhadap seluruh lead secara acak atau berdasarkan urutan kedatangan, tanpa mempertimbangkan karakteristik demografis, perilaku historis, maupun probabilitas keberhasilan masing-masing lead. Akibatnya, sebagian besar upaya komunikasi terbuang untuk menghubungi lead yang sulit dihubungi atau memiliki potensi konversi rendah, sementara lead berkualitas tinggi tidak mendapatkan prioritas follow-up yang semestinya. Lebih lanjut, prioritas baru mulai diterapkan secara manual pada minggu-minggu akhir kampanye ketika tekanan untuk mencapai target meningkat, yang menyebabkan hilangnya momentum untuk mengonversi lead berkualitas

tinggi di awal periode kampanye. Kondisi ini tidak hanya menurunkan efisiensi operasional dan meningkatkan biaya per akuisisi pelanggan, tetapi juga berdampak pada hilangnya peluang revenue yang signifikan serta penurunan moral tim telesales akibat tingkat keberhasilan yang rendah [6].

Untuk mengatasi permasalahan tersebut, diperlukan pendekatan sistematis berbasis machine learning yang mampu memprediksi dan memprioritaskan lead berdasarkan dua dimensi kritis yang saling melengkapi: kemampuan untuk dihubungi (contactability) dan probabilitas konversi menjadi pelanggan (booking probability). Dalam literatur customer analytics, pendekatan lead scoring berbasis machine learning telah terbukti efektif dalam meningkatkan produktivitas tim penjualan dengan cara mengidentifikasi prospek yang paling menjanjikan berdasarkan analisis pola data historis dan karakteristik pelanggan[7][8]. Metode machine learning seperti Logistic Regression, Random Forest, dan XGBoost telah banyak digunakan dalam konteks lead scoring karena kemampuannya menangani data kompleks dengan jumlah fitur yang besar, mengidentifikasi pola non-linear dalam perilaku konsumen, serta menghasilkan prediksi dengan tingkat akurasi tinggi[9] [10] . Penelitian menunjukkan bahwa implementasi machine learning untuk lead scoring dapat meningkatkan conversion rate hingga 77% dan mengurangi lead-per-closed-deal ratio secara signifikan, yang secara langsung berkontribusi pada peningkatan ROI marketing dan sales operations [11].

Lebih lanjut, pendekatan two-stage modeling yang memisahkan prediksi contactability dan conversion telah terbukti lebih efektif dibandingkan single-stage model dalam berbagai aplikasi bisnis. Pemisahan tahapan prediksi ini memungkinkan model untuk menangkap kompleksitas proses penjualan yang melibatkan dua tahap keputusan berbeda dengan karakteristik yang distinct: tahap pertama berkaitan dengan aksesibilitas dan keterjangkauan lead, sementara tahap kedua berkaitan dengan kesiapan dan kemampuan lead untuk melakukan transaksi. Pendekatan modular ini juga memfasilitasi interpretabilitas model yang lebih baik, memungkinkan identifikasi faktor-faktor spesifik yang mempengaruhi setiap tahap proses, serta memberikan fleksibilitas dalam optimasi dan penyesuaian strategi

untuk setiap dimensi prediksi secara independen[12] [13]. Implementasi two-stage approach juga memungkinkan perusahaan untuk mengalokasikan sumber daya komunikasi secara lebih strategis dengan mempertimbangkan trade-off antara effort yang dibutuhkan untuk menghubungi lead dan potential value yang dapat dihasilkan dari konversi.

Sebagai peserta program Merdeka Belajar Kampus Merdeka (MBKM) yang ditempatkan pada Departemen CRM Data Mining di PT XYZ FINANCE, penulis mendapatkan penugasan untuk mengembangkan sistem Two-Stage Machine Learning Lead Scoring yang dirancang untuk mengoptimalkan prioritas follow-up telesales pada lini bisnis pembiayaan sepeda motor. Proyek ini mencakup seluruh siklus pengembangan model machine learning berbasis metodologi CRISP-DM (Cross-Industry Standard Process for Data Mining), yang telah menjadi standar de facto dalam industri untuk project data mining dan analytics dengan tingkat adopsi mencapai 43% di kalangan praktisi data science global [14] [15]. Metodologi CRISP-DM dipilih karena menyediakan framework terstruktur yang mencakup enam fase iteratif: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment, yang memastikan alignment antara solusi teknis dengan kebutuhan bisnis serta sustainability dari model yang dikembangkan [16].

Model yang dikembangkan terdiri dari dua tahap berurutan: Stage 1 untuk memprediksi probabilitas lead dapat dihubungi (delivered status), dan Stage 2 untuk memprediksi probabilitas lead melakukan booking setelah berhasil dihubungi. Kombinasi kedua probabilitas tersebut menghasilkan skor akhir (final lead score) yang dihitung melalui formula $P(\text{Delivered}) \times P(\text{Booking}|\text{Delivered})$, yang kemudian digunakan sebagai dasar prioritisasi dalam proses follow-up telesales. Pendekatan ini memungkinkan tim telesales untuk fokus pada lead dengan kombinasi optimal antara contactability dan conversion potential, sehingga memaksimalkan efisiensi effort dan meningkatkan probability of success dalam setiap interaksi. Implementasi sistem ini juga dilengkapi dengan aplikasi berbasis web menggunakan framework Streamlit yang menyediakan fitur batch prediction,

dashboard analytics, dan visualization untuk mendukung decision-making tim operasional secara real-time [17] .

Lebih jauh, implementasi sistem lead scoring ini sejalan dengan agenda transformasi digital yang sedang dijalankan oleh PT XYZ FINANCE dan industri multifinance Indonesia secara luas, di mana pengambilan keputusan berbasis data dan otomasi proses bisnis menjadi prioritas strategis untuk meningkatkan daya saing [18] [19]. Dengan menerapkan sistem prioritas lead yang terstruktur dan berbasis prediksi machine learning, perusahaan diharapkan mampu meningkatkan efisiensi operasional telesales secara substansial, meningkatkan tingkat konversi booking, serta mengalokasikan sumber daya komunikasi secara lebih optimal. Sistem ini juga dirancang untuk memberikan early wins pada minggu pertama kampanye dengan mengidentifikasi dan memprioritaskan lead berkualitas tinggi sejak awal periode, sehingga momentum penjualan dapat dibangun secara konsisten sepanjang durasi kampanye. Dampak jangka panjang yang diharapkan mencakup peningkatan customer acquisition efficiency, penurunan cost per acquisition, serta peningkatan kepuasan tim telesales melalui higher success rate dalam interaksi dengan prospective customers [20]

Dengan latar belakang tersebut, proyek pengembangan Two-Stage Machine Learning Lead Scoring ini memberikan kontribusi strategis bagi peningkatan kinerja bisnis PT XYZ FINANCE sekaligus menjadi wahana pembelajaran praktis bagi penulis dalam bidang customer analytics, machine learning operations, dan optimasi proses bisnis berbasis data di industri pembiayaan konsumen. Proyek ini juga diharapkan dapat menjadi model implementasi best practices dalam penerapan artificial intelligence untuk sales optimization yang dapat diadaptasi oleh unit bisnis lain dalam grup perusahaan maupun industri sejenis.

1.2. Maksud dan Tujuan Kerja Magang

Program magang merupakan bagian integral dari proses pembelajaran yang bertujuan untuk menjembatani antara dunia akademik dan dunia profesional.

Melalui kegiatan magang, mahasiswa tidak hanya memperoleh penguatan pemahaman teoritis, tetapi juga pengalaman praktis mengenai dinamika kerja di lingkungan industri secara langsung. Oleh karena itu, pelaksanaan magang di perusahaan pembiayaan tempat penulis ditempatkan dirancang agar mahasiswa dapat mengimplementasikan pengetahuan dan keterampilan yang diperoleh selama perkuliahan dalam konteks pekerjaan nyata yang memberikan dampak bisnis terukur.

Secara umum, maksud dari kegiatan magang ini adalah untuk memberikan pemahaman yang komprehensif kepada mahasiswa mengenai peran dan penerapan machine learning dalam mendukung pengambilan keputusan strategis di perusahaan, khususnya dalam konteks optimasi proses penjualan dan akuisisi pelanggan berbasis prediksi analitik. Dalam hal ini, penulis ditempatkan pada departemen yang menangani pengolahan dan analisis data pelanggan, yaitu Departemen CRM Data Mining, dengan fokus pada pengembangan sistem lead scoring yang dapat meningkatkan efisiensi operasional telesales dan mempercepat pencapaian target konversi penjualan.

Adapun tujuan khusus dari pelaksanaan kegiatan magang ini antara lain sebagai berikut:

1. Meningkatkan pemahaman dan keterampilan dalam menerapkan metodologi CRISP-DM (Cross-Industry Standard Process for Data Mining) untuk pengembangan sistem machine learning end-to-end, mulai dari business understanding hingga deployment model ke lingkungan produksi.
2. Mengembangkan sistem Two-Stage Machine Learning Lead Scoring yang dapat memprediksi probabilitas contactability (Stage 1) dan booking conversion (Stage 2) untuk membantu tim telesales dalam memprioritaskan follow-up lead secara lebih terarah dan efisien, dengan target peningkatan conversion rate dan pengurangan wasted effort secara signifikan.
3. Membangun kemampuan dalam implementasi dan evaluasi berbagai algoritma machine learning (XGBoost, Logistic Regression, Random

Forest) dengan fokus pada pemilihan model optimal berdasarkan business metrics yang relevan seperti ROC-AUC untuk ranking quality dan Recall untuk minimizing missed opportunities.

4. Mengembangkan aplikasi deployment berbasis web menggunakan Streamlit yang menyediakan fitur batch prediction, dashboard analytics, dan visualization untuk mendukung operasional telesales dalam pengambilan keputusan real-time serta memfasilitasi monitoring performa model secara berkelanjutan.
5. Melakukan analisis mendalam terhadap business impact dari implementasi lead scoring system, termasuk perhitungan expected improvement dalam key performance indicators seperti booking rate, contact efficiency, cost per acquisition, dan revenue acceleration melalui early identification of high-potential leads.
6. Mengembangkan kemampuan komunikasi teknis dan bisnis melalui presentasi insight analitik kepada tim operasional telesales dan stakeholder bisnis dari lini pembiayaan sepeda motor, serta kemampuan menerjemahkan hasil model machine learning menjadi actionable recommendations yang dapat diimplementasikan dalam strategi operasional harian.

Melalui pencapaian tujuan-tujuan tersebut, diharapkan mahasiswa dapat memperoleh pengalaman kerja yang bermakna dalam mengembangkan solusi data science yang memberikan value measurable kepada bisnis, sekaligus memberikan kontribusi nyata dalam proses transformasi digital dan peningkatan competitive advantage perusahaan di industri pembiayaan konsumen.

1.3. Waktu dan Prosedur Pelaksanaan Kerja Magang

Menjelaskan waktu pelaksanaan kerja magang dan prosedur kerja magang yang dilaksanakan oleh mahasiswa magang di instansi/ perusahaan tempat pelaksanaan kerja magang.

1.3.1 Waktu dan Lokasi Pelaksanaan Kerja Magang

Kegiatan kerja magang dilaksanakan sebagai bagian dari program magang kepemimpinan yang diselenggarakan oleh perusahaan pembiayaan tempat penulis menjalani penempatan. Periode magang berlangsung sejak tanggal 16 Juni 2025 hingga 31 Januari 2026, dan telah disesuaikan dengan ketentuan akademik yang berlaku di Universitas Multimedia Nusantara. Laporan magang ini mencakup aktivitas dan capaian yang telah diselesaikan hingga akhir November 2025, sebagaimana dipersyaratkan untuk pelaksanaan sidang magang.

Selama masa penugasan, penulis ditempatkan pada Departemen CRM Data Mining, yang berada di bawah divisi pengelolaan strategi berbasis data pelanggan. Dalam penempatan ini, penulis diberikan tanggung jawab utama untuk melaksanakan proyek berupa pengembangan Two-Stage Machine Learning Lead Scoring System untuk optimasi prioritas follow-up telesales pada lini pembiayaan sepeda motor, yang mencakup seluruh siklus pengembangan berbasis metodologi CRISP-DM (Cross-Industry Standard Process for Data Mining), mulai dari business understanding, data preparation, modeling, evaluation, hingga deployment aplikasi berbasis web.

Selain proyek utama tersebut, penulis juga terlibat secara aktif dalam serangkaian proyek pendukung lain yang dilaksanakan secara paralel, yaitu:

1. PT XYZ SUPERAPP Data Pipeline untuk proses ETL (Extract, Transform, Load) dan reporting dengan kapasitas pemrosesan data skala besar,
2. Aplikasi Validasi Data Nasabah berbasis Streamlit dengan fitur fuzzy matching dan analytics dashboard,
3. Dealer Scoring & Competitive Portfolio Analysis untuk evaluasi performa dealer berdasarkan data kompetitor dari berbagai perusahaan pembiayaan.

Keseluruhan proyek tersebut dirancang untuk mendukung transformasi digital perusahaan dan penguatan proses analitik dalam strategi pemasaran, optimasi operasional telesales, serta pengambilan keputusan strategis berbasis data.

Seluruh kegiatan magang dilaksanakan secara penuh waktu (work from office/WFO) di kantor pusat perusahaan, yang berlokasi di kawasan Lebak Bulus, Cilandak, Jakarta Selatan, Provinsi DKI Jakarta. Penempatan dilakukan bersama tim internal pada divisi pengelolaan data pelanggan, sehingga memungkinkan penulis untuk menjalankan proses koordinasi secara langsung dengan tim analis, kepala departemen, serta unit bisnis terkait seperti tim operasional telesales dan tim promosi.

Jam kerja mengikuti regulasi internal perusahaan, yaitu Senin hingga Jumat, dengan total waktu kerja 40 jam per minggu pada pukul 08.00 – 17.00 WIB, dengan waktu istirahat pukul 12.00 – 13.00 WIB. Untuk memberikan gambaran yang lebih terstruktur mengenai distribusi aktivitas kerja magang selama periode Juni hingga November 2025, berikut disajikan Tabel 1.1 Linimasa Program Magang, yang memuat detail tahapan kegiatan pada setiap proyek yang telah direncanakan dan diimplementasikan selama program berlangsung.

Tabel 1. 1 Linimasa Program Magang

No	Project/Kegiatan	Juni		Juli				Agustus				September				Oktober				November			
		3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Project 1: Two-Stage Machine Learning Lead Scoring CRM																							
1	Business Understanding & Data Understanding (CRISP-DM Phase 1-2)																						
2	Data Preparation & Feature Engineering (CRISP-DM Phase 3)																						
3	Modeling: Baseline & Tuning Stage-1 & Stage-2 (CRISP-DM Phase 4)																						
4	Evaluation & Dashboard Development (CRISP-DM Phase 5)																						
5	Model Iteration & Refinement (Review & Retraining)																						
6	Deployment & Final Documentation (CRISP-DM Phase 6)																						
Project 2: PT XYZ SUPERAPP Data Pipeline (ETL & Reporting)																							
7	Pipeline Development: TXT parsing, data cleaning, format handling (6M records)																						
8	Data Integration & Feature Engineering: merge CSV, temporal features, business flags																						
9	Analytics Development: penetrasi area, CORI grade, XYZvs Non-XYZ, visualization																						
10	Testing & Deployment: end-to-end pipeline testing, dokumentasi, Google Colab deployment																						

Project 3: Aplikasi Validasi Data Nasabah

11	Setup & Core Validators: Streamlit framework, KK, NIK, CUSTNAME validators																						
12	Fuzzy Matching Implementation: RapidFuzz, normalisasi tempat lahir, threshold tuning																						
13	Dashboard & Analytics: interactive visualization, heatmap validasi, statistik detail																						
14	Export Functionality: PDF report generation, Excel multi-sheets, professional formatting																						
15	Optimization & Finalization: handle >300K rows, UI/UX enhancement, testing, dokumentasi																						

Project 4: Dealer Scoring & Competitive Portfolio Analysis

16	Discovery & Data Extraction: prototyping, ekstraksi 6 kompetitor (42 iterasi, 87K records)																						
17	Data Integration & Standardization: fuzzy matching dealers, tenor matching, pivot tables																						
18	Analytics & Scoring: credit portion, DP gap, scoring formula implementation, dealer ranking																						
19	Visualization & Segmentation: quartile matrix 4x4, bubble chart, portfolio analysis																						

1.3.2. Prosedur Pelaksanaan Kerja Magang

Pelaksanaan program magang tidak hanya mencakup aktivitas teknis yang dilakukan selama berada di perusahaan, tetapi juga mencakup tahapan administratif dan persiapan yang dilakukan sebelum magang, serta pelaporan hasil kegiatan setelah program selesai. Secara umum, prosedur kerja magang dibagi menjadi tiga tahap utama, yaitu pra-magang, pelaksanaan magang, dan pasca-magang. Ketiga tahap ini disusun secara sistematis untuk memastikan bahwa program magang berjalan terstruktur, sesuai dengan ketentuan akademik, serta relevan dengan kebutuhan profesional di lingkungan industri.

1.3.2.1 Tahap Pra-Magang

Tahap pra-magang merupakan proses awal yang dijalani oleh penulis sebelum melaksanakan kegiatan magang di lingkungan perusahaan. Program magang yang diikuti bukan merupakan program reguler, melainkan bagian dari program pengembangan kepemimpinan internal perusahaan yang ditujukan bagi mahasiswa terpilih untuk memperoleh pengalaman implementatif di berbagai departemen strategis. Program ini dirancang dalam dua periode pelaksanaan yang berkesinambungan, di mana periode pertama berlangsung pada semester 6 (3 Februari 2025 – 30 Juni 2025) dengan fokus pada customer segmentation menggunakan metode RFM dan K-Means Clustering, sedangkan periode kedua berlangsung pada semester 7 (16 Juni 2025 – 31 Januari 2026) dengan fokus pada pengembangan Two-Stage Machine Learning Lead Scoring System.

Penulis mendaftar program ini pada tanggal 30 September 2024 dan mengikuti serangkaian tahapan seleksi yang mencakup pengisian kuesioner psikologis (15 Oktober 2024), presentasi ide bisnis, serta wawancara dengan jajaran kepala departemen dan tim sumber daya manusia. Hasil akhir seleksi diumumkan pada tanggal 8 November 2024, di mana penulis dinyatakan lolos sebagai peserta resmi. Sebagai pembekalan awal, penulis mengikuti pelatihan intensif selama satu minggu (17–24 November 2024) mengenai budaya organisasi, struktur perusahaan, dan nilai-nilai inti perusahaan, dilanjutkan dengan kegiatan branch overview pada awal Januari 2025 untuk memahami proses bisnis dan struktur operasional dari

kantor pusat hingga jaringan cabang. Setelah melalui proses pemilihan divisi berdasarkan minat dan wawancara dengan kepala departemen, penulis dinyatakan diterima secara resmi di Departemen CRM Data Mining pada tanggal 17 Januari 2025. Program magang ini dilaksanakan secara independen tanpa prosedur administratif dari pihak universitas, karena seluruh proses rekrutmen dilakukan langsung oleh perusahaan.

Periode pertama dimulai dengan kegiatan briefing dan onboarding resmi pada tanggal 3 Februari 2025, dilanjutkan dengan proses adaptasi internal selama satu minggu pertama. Selama lima bulan pelaksanaan (Februari – Juni 2025), penulis mengembangkan sistem segmentasi pelanggan berbasis RFM dan K-Means Clustering, aplikasi web segmentasi, sistem otomatisasi distribusi email, aplikasi validasi data kependudukan, serta pipeline penggabungan data telesales. Periode pertama berakhir pada tanggal 30 Juni 2025 dengan penyerahan deliverables dan dokumentasi proyek kepada tim operasional serta knowledge transfer kepada stakeholder terkait.

Memasuki periode kedua, penulis mengikuti session planning dan project scoping bersama pembimbing lapangan pada minggu ketiga Juni 2025 untuk menentukan fokus proyek semester 7. Berdasarkan evaluasi kebutuhan bisnis dan assessment terhadap pain points operasional yang belum terselesaikan di periode pertama, disepakati bahwa periode kedua akan berfokus pada pengembangan sistem prediktif untuk optimasi lead scoring pada operasional telesales lini pembiayaan sepeda motor. Kegiatan periode kedua secara resmi dimulai pada tanggal 16 Juni 2025 dengan kick-off meeting untuk penetapan business objectives, timeline, dan key performance indicators yang akan menjadi acuan evaluasi kesuksesan proyek.

1.3.2.2 Tahap Pelaksanaan Magang

Tahap pelaksanaan magang merupakan inti dari keseluruhan proses kerja magang periode kedua yang dijalani oleh penulis di perusahaan pembiayaan tempat penulis ditempatkan. Kegiatan ini dilaksanakan secara langsung (work from office)

di kantor pusat PT XYZ FINANCE, di bawah pembinaan Departemen CRM Data Mining sebagai bagian dari Divisi CRM & Digital. Sebagai kelanjutan dari periode pertama (Februari–Juni 2025), periode kedua ini berfokus pada pengembangan sistem prediktif dan optimasi operasional telesales berbasis machine learning, dengan tetap didukung oleh proyek-proyek analitik pendukung lainnya.

Kegiatan magang periode kedua dimulai pada tanggal 16 Juni 2025, yang diawali dengan session planning dan project scoping bersama pembimbing lapangan pada minggu ketiga Juni 2025. Berdasarkan evaluasi kebutuhan bisnis dan assessment terhadap pain points operasional yang belum terselesaikan di periode pertama, disepakati bahwa periode kedua akan berfokus pada pengembangan sistem Two-Stage Machine Learning Lead Scoring untuk optimasi prioritas follow-up telesales pada lini pembiayaan sepeda motor. Setelah kick-off meeting untuk penetapan business objectives dan timeline, penulis mulai menjalankan proyek-proyek yang telah ditugaskan secara bertahap dan terstruktur.

Empat proyek utama yang menjadi fokus penulis selama masa pelaksanaan magang periode kedua adalah sebagai berikut:

a. Two-Stage Machine Learning Lead Scoring untuk Optimasi Prioritas Follow-Up Telesales CRM

Penulis bertanggung jawab untuk mengembangkan sistem lead scoring berbasis machine learning menggunakan metodologi CRISP-DM secara end-to-end. Sistem yang dikembangkan terdiri dari dua tahap prediksi berurutan: Stage 1 untuk memprediksi probabilitas lead dapat dihubungi (contactability/delivered status), dan Stage 2 untuk memprediksi probabilitas lead melakukan booking setelah berhasil dihubungi. Proses pengembangan mencakup business understanding dan validasi dengan tim operasional telesales, data preparation dengan handling missing values dan pencegahan data leakage melalui eliminasi kolom status yang mengandung informasi post-prediction, feature engineering berbasis domain knowledge dengan pembuatan binning categories dan temporal features, serta modeling dengan implementasi tiga algoritma (XGBoost, Logistic Regression, Random Forest) dalam tiga varian masing-masing (baseline, tuned, threshold-optimized).

Model XGBoost tuned terpilih sebagai model final dengan performa terbaik: Stage 1 mencapai ROC-AUC 0.7237 dan Recall 0.9444 untuk meminimalkan missed contactable leads, sedangkan Stage 2 mencapai ROC-AUC 0.7714 yang mengindikasikan excellent ranking discrimination untuk prioritas lead. Kombinasi kedua probabilitas menghasilkan final lead score melalui formula $P(\text{Delivered}) \times P(\text{Booking}|\text{Delivered})$ yang digunakan untuk strategic resource allocation. Aplikasi Streamlit dikembangkan dengan fitur batch prediction, performance monitoring dashboard, dan comprehensive documentation untuk mendukung deployment dan adoption oleh tim telesales.

b. XYZ FINANCE SUPERAPP Data Registration Pipeline

Penulis mengembangkan sistem otomatis end-to-end pipeline untuk memproses data registrasi aplikasi XYZ FINANCE SUPERAPP yang mencapai 6 juta records per bulan dalam format TXT yang sering rusak dan tidak terstruktur. Pipeline yang dibangun berbasis Google Colab dan Python mampu melakukan parsing otomatis file TXT dengan buffer management untuk handle broken multi-line records, data cleaning dengan validasi 30 kolom per record dan success rate 99.9978%, serta integrasi 8 file menjadi 1 master dataset. Sistem dilengkapi dengan feature engineering otomatis untuk temporal dan business flags, analytics development yang menghasilkan tiga deliverables utama berupa penetrasi registrasi per area dalam format Excel, visualisasi CORI Grade distribution untuk customer segmentation, dan analisis distribusi XYZ vs Non-XYZ customer, serta additional analysis berupa temporal trends dan geographic breakdowns. Pipeline ini berhasil mengurangi waktu proses dari 2-3 hari menjadi 10-15 menit dengan eliminasi human error dan konsistensi reporting format, serta telah di-deploy dengan dokumentasi lengkap untuk memfasilitasi eksekusi rutin oleh tim data analyst.

c. Aplikasi Validasi Data Nasabah

Dalam proyek ini, penulis membangun aplikasi berbasis Streamlit untuk memvalidasi konsistensi data kependudukan nasabah dengan empat jenis validator: format dan digit check untuk Nomor Kartu Keluarga (KK) dan Nomor Induk Kependudukan (NIK), validasi nama pelanggan dengan deteksi invalid characters dan minimum length, serta fuzzy matching untuk tempat lahir menggunakan library RapidFuzz dengan threshold tuning untuk mendeteksi typo dan inkonsistensi penulisan. Aplikasi dilengkapi dengan real-time progress tracking untuk monitoring validasi, interactive dashboard dengan heatmap error distribution dan statistik detail per jenis error, serta fitur export yang mampu menghasilkan PDF report professional dan Excel multi-sheets dengan formatting otomatis. Sistem ini telah dioptimasi untuk menangani dataset berukuran lebih dari 300.000 baris dengan response time yang acceptable, menyediakan aggregate status per record (Valid, Warning, Error) berdasarkan priority logic, dan memberikan audit trail lengkap melalui detailed error logs untuk mendukung decision making dan compliance requirements.

d. Dealer Scoring dan Competitive Portfolio Analysis

Penulis melakukan ekstraksi dan integrasi data kompetitor dari 6 perusahaan pembiayaan melalui 42 iterasi web scraping yang menghasilkan 87.000 records, kemudian melakukan standardisasi data menggunakan fuzzy matching untuk mencocokkan nama dealer across competitors, tenor matching untuk apple-to-apple comparison, serta pembangunan pivot tables untuk analisis multidimensi. Sistem scoring yang dikembangkan mengevaluasi dealer berdasarkan credit portion, DP gap competitiveness, dan portfolio strength dengan formula implementation yang telah divalidasi bersama tim bisnis. Hasil analisis divisualisasikan dalam bentuk quartile matrix 4×4 untuk segmentasi dealer menjadi kategori strategis (High Volume-High Competition, High Volume-Low Competition, dll), bubble chart untuk portfolio positioning, serta comprehensive analysis yang

memberikan actionable insights untuk strategic partnership decisions dan competitive positioning strategies di setiap area geografis.

Seluruh proyek dilaksanakan dengan koordinasi rutin bersama Department Head Data Mining sebagai pembimbing teknis utama, dengan reporting progress setiap completion milestone seperti selesainya fase CRISP-DM, finalisasi dashboard, atau sebelum deployment ke production. Penulis juga secara aktif berkoordinasi dengan tim CRM dan telesales untuk business validation, user acceptance testing, serta memastikan kesesuaian solusi dengan operational workflow di lapangan. Pihak Human Capital melakukan monitoring berkala terhadap kinerja dan kedisiplinan sebagai bagian dari program pembinaan peserta magang.

Selain fokus pada implementasi teknis, pelaksanaan magang ini turut memperkuat kemampuan penulis dalam aspek komunikasi profesional, manajemen proyek, serta pemecahan masalah di lingkungan kerja nyata. Penulis menggunakan berbagai perangkat dan teknologi pendukung, antara lain Python dengan libraries scikit-learn, XGBoost, pandas untuk machine learning dan data processing, Streamlit untuk web application development, Optuna untuk hyperparameter optimization, Plotly dan matplotlib untuk visualization, serta Visual Studio Code, Google Colab, dan Git sebagai development tools.

Dengan pelaksanaan proyek-proyek tersebut, penulis tidak hanya memperoleh pengalaman praktis dalam bidang machine learning operations, customer analytics, dan optimasi proses bisnis, tetapi juga memberikan kontribusi nyata terhadap peningkatan efisiensi operasional telesales, akselerasi revenue through early identification of high-potential leads, serta penguatan strategi data-driven decision making di lingkungan PT XYZ FINANCE.

1.3.2.3 Tahap Pasca Magang

Tahap pasca-magang merupakan fase akhir dari keseluruhan rangkaian kegiatan kerja magang periode kedua yang dijalani oleh penulis di lingkungan

perusahaan. Mengingat periode magang berlangsung hingga 31 Januari 2026, tahap pasca-magang yang dijelaskan dalam laporan ini mencakup aktivitas dokumentasi dan pelaporan interim yang dilakukan hingga akhir November 2025 sebagai bagian dari pemenuhan kewajiban akademik untuk pelaksanaan sidang magang.

Pada fase ini, penulis menyusun laporan terstruktur yang memuat proses kerja, metodologi CRISP-DM yang diterapkan dalam pengembangan sistem machine learning, temuan utama dari analisis data, serta hasil dari keempat proyek yang telah dikerjakan. Laporan ini tidak hanya berfungsi sebagai pemenuhan kewajiban akademik kepada Universitas Multimedia Nusantara, tetapi juga sebagai bentuk dokumentasi teknis dan kontribusi nyata kepada tim pembimbing serta unit kerja terkait di perusahaan untuk memfasilitasi keberlanjutan implementasi dan maintenance sistem yang telah dikembangkan.

Output utama dari proyek magang periode kedua, yang meliputi sistem Two-Stage Machine Learning Lead Scoring dengan aplikasi Streamlit untuk batch prediction dan performance monitoring, pipeline XYZ FINANCE SUPERAPP data analysis berbasis Google Colab dengan automated reporting capabilities, aplikasi validasi data nasabah dengan fuzzy matching dan professional export features, serta hasil dealer scoring dan competitive portfolio analysis dengan strategic segmentation framework, dirancang untuk diserahkan secara formal kepada Departemen CRM Data Mining dan unit operasional yang relevan. Dokumentasi teknis yang komprehensif mencakup user manual, technical documentation untuk maintenance dan troubleshooting, serta business impact analysis untuk setiap sistem yang dikembangkan, sehingga memungkinkan tim internal untuk melanjutkan penggunaan dan pengembangan lebih lanjut setelah periode magang berakhir.

Sebagai bagian dari transfer pengetahuan (knowledge transfer), penulis melaksanakan sesi presentasi internal kepada tim CRM Data Mining, tim operasional telesales, dan stakeholder terkait dari unit bisnis pembiayaan sepeda motor. Presentasi ini memaparkan secara sistematis alur pengembangan sistem lead scoring menggunakan metodologi CRISP-DM mulai dari business understanding hingga deployment, hasil evaluasi model dengan perbandingan performa XGBoost,

Logistic Regression, dan Random Forest across multiple variants, feature importance analysis yang memberikan actionable business insights terkait faktor-faktor yang mempengaruhi contactability dan booking conversion, serta rekomendasi strategis untuk optimalisasi proses follow-up telesales berdasarkan hasil prediksi model dan decile analysis yang menunjukkan separasi jelas antara high-potential dan low-potential leads. Untuk proyek-proyek pendukung lainnya, knowledge transfer juga mencakup hands-on training untuk penggunaan aplikasi, troubleshooting common issues, serta best practices dalam operasional harian sistem yang telah dikembangkan.

Setelah seluruh tahapan dokumentasi, knowledge transfer, dan penyerahan deliverables di perusahaan selesai dilaksanakan untuk periode hingga November 2025, penulis dijadwalkan untuk mengikuti proses evaluasi akademik berupa sidang magang di Universitas Multimedia Nusantara pada Desember 2025, sebagai bentuk penilaian menyeluruh atas kontribusi dan capaian selama mengikuti program magang periode kedua. Kegiatan magang akan dilanjutkan hingga 31 Januari 2026 untuk menyelesaikan aktivitas transisi, monitoring performa sistem di production environment, serta memberikan dukungan teknis selama fase awal adoption oleh tim operasional, dengan dokumentasi akhir dan final deliverables akan diserahkan pada akhir periode magang.