International Journal of AI, Big Data, Computational and Management Studies
Noble Scholar Research Group | Volume 7, Issue 4, PP 30-43, 2026
ISSN: 3050-9416 | https://doi.org/10.63282/3050-9416.IJAIBDCMS-V7I4P104

Original Article

Predictive Lifetime Value Modeling in Multi-Channel Telecom Customer Systems

Brahmananda Naidu Dabbara
Independent Researcher, USA.

Received On: 26/08/2026 ***Revised On: 24/09/2026*** ***Accepted On: 01/10/2026*** ***Published On: 06/10/2026***

Abstract: *Telecom operators estimate customer worth from measures that look backwards. Current monthly revenue, average revenue per user, contract value, tenure, product count, payment history and recent usage are combined into high, medium and low value bands that mostly restate past spend. The result is a position, not a direction, and it says nothing about how value will move over the next one to three years. The deeper problem is not the choice of estimator. It is that the evidence is fragmented: billing, customer relationship management, digital engagement, customer service, network usage, partner channels and campaign response are each evaluated on their own, so the signals that would together indicate a change in trajectory are never combined. This paper describes a lifetime value modelling system built on a multi-channel customer profile rather than on a single system of record, and it treats the construction of that profile as the substantive engineering problem. Identity is resolved through a governed ladder: deterministic matching on stable identifiers, then controlled probabilistic matching whose confidence must clear a threshold before any merge occurs, and ambiguous records that are deliberately not merged but retained separately or held at lower confidence until stronger evidence arrives. Source-system lineage is preserved so every model feature traces to the system it came from. Channel-specific activity is normalised into one behavioural event vocabulary in which every event keeps its timestamp, source channel, product context and outcome, which is what makes cross-channel patterns visible: declining network usage, repeated contacts about one unresolved issue, an ignored retention email and a plan comparison on the website together indicate falling future value, though no single system holds the whole pattern. Features are then built from trajectory rather than snapshot, and predicted value is expressed as expected future revenue multiplied by retention probability less expected cost to serve, over twelve, twenty four and thirty six month horizons. Two governance positions are argued as design principles rather than afterthoughts. Engagement measured on the digital subset systematically understates customers who transact through stores, partners and the contact centre, so engagement is modelled across every channel. And economic prediction is separated from service entitlement: predicted value may determine which retention offer or cross-sell is appropriate, but it never determines who receives essential support, complaint resolution or network assistance, and a low predicted value never lowers service quality. Business figures are reported as pilot and implementation ranges rather than audited production results, and no model accuracy figures were available, so none are reported.*

Keywords: *Customer Lifetime Value, Multi-Channel Customer Management, Identity Resolution, Record Linkage, Data Lineage, Telecom Analytics, Churn Prediction, Cost to Serve, Model Monitoring, Algorithmic Fairness.*

1. Introduction

A telecom operator knows a great deal about what each customer paid last month and very little about what that customer will be worth in two years. The measures in routine use are backward looking by construction: current monthly revenue, average revenue per user, total contract value, tenure, number of products held, payment history and recent usage. Customers are bucketed into high, medium and low value bands largely on past spend. Every one of those measures is real and none of them answers the question that actually drives spending decisions, which is how value will evolve over the next twelve to thirty six months and what action taken now would improve it.

The consequences of measuring value this way are specific rather than abstract. Retention money is misallocated, because intervention arrives only once churn risk has become obvious, at which point the customer has usually decided. Campaigns are pitched too broadly, because they target static segments defined by historical spend. Cross-sell and upsell opportunities are missed, because the signals that reveal them sit in different systems and are never combined. High-potential customers are undervalued whenever their current average revenue per user is low, and high-revenue customers are overvalued despite declining engagement and unresolved service problems. Acquisition decisions get weighted towards initial revenue rather than long-term value and cost to serve. The shift required is a change of question, from what this customer is worth today to what they are likely to be worth across the relationship and what action now improves that.

That predictive lifetime value can be estimated with machine learning inside a customer relationship management

workflow is established, including in the author's own prior work on predictive analytics for lifetime value in telecom customer relationship management [6], and the broader modelling literature has been mapped for two decades [12]. This paper is therefore not about whether value can be predicted. It is about the layer underneath the prediction, which is where the difficulty in a real telecom estate actually lives.

A large operator does not hold one customer record. It holds many partial records, distributed across seven distinct channel and data domains, each with its own identifiers, its own schema, its own notion of an interaction and its own reasons for being incomplete. Multichannel customer management has long been understood to create exactly this problem of a customer who is one person to the business and several rows to its systems [16]. Until those rows are reconciled, a lifetime value model is not modelling a customer. It is modelling a fragment and inheriting the fragment's blind spots.

This paper makes five contributions, all of which sit on the multi-channel dimension rather than on the estimator. First, it specifies a multi-channel identity resolution ladder in which deterministic matching on stable identifiers is preferred, controlled probabilistic matching is permitted only above a governed confidence threshold, and ambiguous matches are deliberately not merged. Uncertain identity is treated as a contamination risk to the value model rather than as a coverage problem to be solved optimistically.

Second, it defines an event normalisation scheme that maps channel-specific activity into one behavioural vocabulary while preserving timestamp, source channel, product context and outcome, and it maintains source-system lineage so that every derived feature can be traced back to its originating system. Third, it shows how the unified profile exposes cross-channel patterns that no participating system can see alone, and argues that these patterns, not any single system's metrics, are what carry predictive information about future value.

Fourth, it treats channel bias as a fairness problem rather than a data-coverage inconvenience. A customer who rarely opens the app and ignores email looks disengaged in a digital-only measurement and may be heavily engaged through stores, partners and the contact centre. Modelling engagement across all channels rather than the digital subset is presented as a fairness control, not merely as a feature-engineering improvement. Fifth, it draws an explicit boundary between economic prediction and service entitlement. Predicted value is allowed to shape commercial treatment and is not allowed to shape service quality.

Sections 5 and 6 carry the first two contributions and are the core of the paper. Section 2 states plainly what was built and evaluated as against what is proposed design, because the two are not the same and the distinction matters more here than a confident tone would.

2. Scope and Evidentiary Basis

This section exists so that no reader has to guess which claims in this paper are reports of completed work and which are descriptions of proposed design. The distinction is maintained deliberately throughout, and the wording elsewhere follows it.

1. **Built and evaluated:** The multi-channel identity resolution approach described in Section 5 was implemented: deterministic matching on stable identifiers, controlled probabilistic matching with confidence thresholds, the deliberate decision not to merge ambiguous matches, and the retention of source-system lineage. The event normalisation scheme in Section 6 was implemented, including the requirement that every normalised event retain timestamp, source channel, product context and outcome. The signal domains and derived feature families in Section 7, the value formulation and horizon structure in Section 8, the qualitative findings about which signals carry weight in Section 9, the bias and governance controls in Section 10 including the separation of prediction from service entitlement, and the three-layer monitoring and retraining policy in Section 11 all describe work that was carried out.
2. **Proposed design and reporting guidance:** Several elements of this paper are the author's stated intent for how the work should be described and published rather than descriptions of completed measurement. Specifically, the decision to report business impact as pilot and implementation ranges, the decision not to assert an exact production customer count, the framing of scale figures as implementation or research scale, and the labelling of all quantities as experimental or pilot rather than audited are publication guidance. They are reproduced here as guidance, and this paper follows them.
3. **No audited production results were available:** No business report, finance sign-off or production measurement was available to substantiate the impact figures in Section 12, and the author was explicit that in the absence of such reports the figures should be presented as ranges rather than as audited outcomes. That is what Section 12 does. Ranges are never narrowed to a midpoint anywhere in this paper.
4. **No model accuracy figures of the author's own were available:** The framework tracks mean absolute error, root mean squared error, ranking performance, calibration and segment-level error, and Section 11 describes how. No values for any of those metrics were supplied and none are reported. There is consequently no accuracy chart, no calibration plot, no receiver operating characteristic curve and no predicted-against-actual scatter in this paper. Where a chart would require values that do not exist, the mechanism is drawn instead and labelled as a schematic.
5. **Scale:** The implementation operated at approximately one to five million customer or

subscriber records and tens of millions of behavioural and transactional records, across seven primary channel and data domains. These are implementation and research scale figures, not verified production counts, and the range on the customer figure is deliberate.

3. Related Work

Lifetime value modelling has a long and well-developed literature. The canonical treatments frame it as the present value of expected future contribution and set out the model families used to estimate it, including recency, frequency and monetary summaries, migration models, hazard and survival formulations and probability models of purchase and defection [12]. Empirical work on lifetime value has established a set of generalisations along with a set of unresolved conceptual questions, including how much of realize d value is genuinely predictable from behaviour observed so far [2]. That literature is the premise of this paper rather than its subject.

Two model families are worth distinguishing from the approach taken here. Buy-till-you-die formulations, of which the Pareto and negative binomial specification is the best known, model purchase frequency and an unobserved dropout process jointly, and deliver a probability that a customer is still active together with an expected transaction count [18]. They are elegant and parsimonious, and they are designed for settings where the observable history is a transaction stream. A telecom relationship is not primarily a transaction stream. It is a subscription with continuous usage, a service history, contract mechanics and interaction across many channels, so the informative signals extend well beyond the recency and frequency of purchase. Related work has shown that recency, frequency and monetary summaries can be mapped onto lifetime value directly, which clarifies exactly what such summaries do and do not capture [9]. What they do not capture is the direction of travel in service experience, product depth or channel engagement, which Section 9 identifies as the strongest available predictors.

Machine learning approaches to lifetime value have been deployed at industrial scale, including engagement-based systems that predict value from behavioural signals rather than from transaction history alone [19]. In telecom specifically, the closely related problem of churn prediction has been studied extensively, and the most useful strand of that work is the one that evaluates churn models by the profit consequences of acting on them rather than by classification accuracy alone [20]. That distinction between statistical performance and decision usefulness runs through Sections 11 and 13 of this paper.

The literature on acting on predictions is equally relevant. Targeting customers with the highest churn risk can be ineffective or counterproductive, because risk of leaving and responsiveness to intervention are different quantities [1]. Uplift modelling addresses this directly by estimating the incremental effect of a treatment rather than the outcome under treatment, and the field has been surveyed and benchmarked [7]. Section 13 argues that a lifetime value model which is not connected to a treatment model is an interesting number rather than an operational asset.

Where this paper departs from the lifetime value literature is in taking the construction of the customer profile as the primary problem. That places it in contact with record linkage and entity resolution. The probabilistic foundation for deciding whether two records refer to the same entity, including the formal treatment of match and non-match regions with an intermediate zone reserved for uncertainty, was established over fifty years ago [10]. Subsequent surveys catalogue the similarity measures, blocking strategies and classification approaches now available [8], and the field has been consolidated into a systematic treatment of matching concepts and techniques [5]. The intermediate zone in the classical framework is the direct ancestor of the deliberately unmerged tier in Section 5. What that literature contributes to this paper is the discipline of treating a probabilistic match as a decision with two distinct error modes rather than as a similarity score to be thresholded casually.

Because features derived from a merged profile can be wrong in ways that are invisible downstream, provenance matters. Surveys of data provenance set out why lineage is recorded, what forms it takes and what questions it can answer [13], and the requirement in Section 5 that every feature trace to its source system and to the tier at which its identity was resolved is an application of that principle to a modelling pipeline.

Finally, the monitoring and fairness literature. Concept drift, meaning change in the underlying data-generating process after deployment, has been surveyed comprehensively along with the detection and adaptation strategies available [11], and production readiness for machine learning systems has been formalised into explicit tests covering data, model, and infrastructure and monitoring [3]. On fairness, the general taxonomy of sources of bias and of fairness definitions is well established [15], and empirical work in marketing has shown that apparently neutral algorithmic delivery can produce systematically unequal exposure without any protected attribute being used as an input [14]. That result is the closest published analogue to the channel bias problem in Section 10. The self-reinforcement risk, where a model's own outputs shape the data it later learns from, has been demonstrated in recommendation settings where algorithmic confounding increases homogeneity and reduces utility [4]. Section 10 argues that a deployed lifetime value model is exposed to precisely this failure.

4. Why Fragmented Channel Data Breaks Value Estimation

Before this work, value was estimated from the measures listed in Section 1 and customers were bucketed on past spend. The estimate was not wrong so much as incomplete in a way that had direction: it systematically favoured customers who were paying well now and

systematically overlooked customers whose behaviour was moving.

The core failure was fragmentation. Billing, customer relationship management, digital engagement, customer service, network usage, partner channels and campaign response were each evaluated independently. Each of those systems is competent within its own boundary. Billing knows what was charged and whether it was paid. The service platform knows how many cases were opened and how long they took to close. The digital platform knows which sessions occurred. None of them knows what the others know, and value is a property of the whole relationship rather than of any one of its parts.



Fig 1: The Seven Primary Channel and Data Domains, What Each Contributes to the Customer Picture, and Their Resolution into a Single Profile. No Domain Can Substitute for Another, and Value Estimated from Any One of Them Inherits That Domain's Blind Spots

Six consequences followed, and they are worth stating individually because each one is a distinct commercial loss rather than a restatement of the same complaint. Retention spending was misallocated. Intervention was triggered by indicators that only became visible once churn risk was already obvious, which is late in the sequence and expensive to reverse. Campaigns were too broad. Targeting rested on static segments built from historical spend, so campaign lists were populated by customers who had been valuable rather than by customers whose value was moving in a direction that an offer could change.

Cross-sell and upsell opportunities were missed. The signals that indicate readiness for an additional product are rarely concentrated in one system. A usage pattern shift, a service enquiry about a capability the customer does not have, and a comparison of plans on the website are three separate observations in three separate systems, and none of them is compelling alone. High-potential customers were undervalued. A customer with modest current average revenue per user but rising usage, strengthening digital engagement, growing product adoption, consistent payment and low service friction looks unremarkable in a spend-ranked segmentation.

High-revenue customers were overvalued. Declining engagement and accumulating unresolved service problems do not reduce this month's invoice, so a customer already on the path out of the relationship continues to appear in the top band until the moment they leave. Acquisition decisions were weighted to initial revenue. Judging a prospective customer on the revenue of their first contract, without regard to expected duration or cost to serve, biases

acquisition towards customers who are expensive to keep and cheap to lose.

Each of these follows from the same structural fact. Fragmented channel data does not merely make value estimation less precise. It makes it biased in a knowable direction, towards current payers and away from customers whose trajectory is the actual news.

5. Multi-Channel Identity Resolution

Nothing else in this paper works unless this section works. A lifetime value model built over a profile that has silently fused two customers, or split one customer into three, is producing confident numbers about entities that do not exist. Resolution was therefore treated as a governed process with explicit tiers and an explicit policy for uncertainty, rather than as a matching utility to be tuned for coverage.

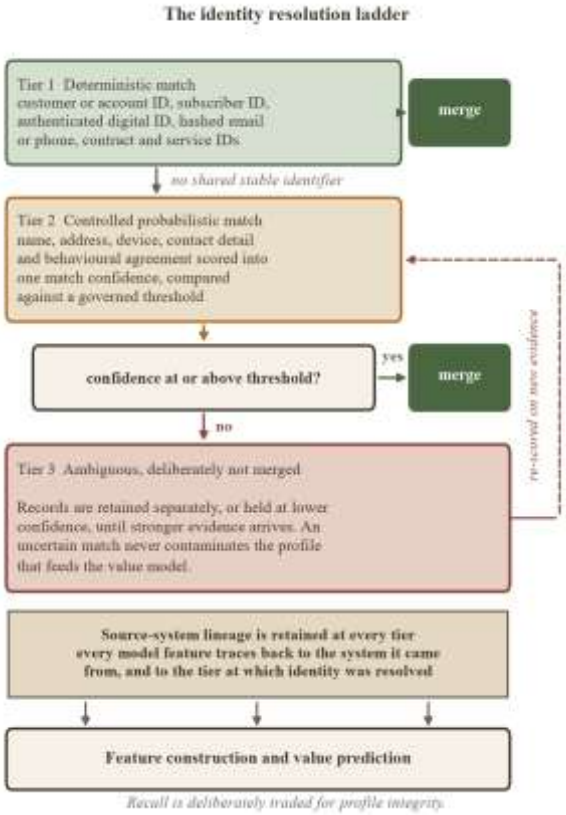


Fig 2: The Identity Resolution Ladder. Deterministic Matching on Stable Identifiers Is Preferred. Probabilistic Matching Is Permitted Only When Its Confidence Clears a Governed Threshold. Ambiguous Records Are Deliberately Not Merged, and Are Re-Scored If Stronger Evidence Later Arrives. Source-System Lineage Is Retained at Every Tier

1. **Deterministic matching comes first:** Wherever two records share a stable identifier, they are joined on it and no probabilistic reasoning is invoked. The identifiers used are the customer or account identifier, the subscriber identifier, an authenticated digital identity, a hashed email address or telephone number, and contract or service identifiers. These identifiers are not equally available across the estate, which is precisely why the later tiers exist, but where they are present they are decisive. The reason for the strict ordering is that a deterministic join fails loudly, by finding nothing, whereas a probabilistic join fails quietly, by finding something wrong.
2. **Probabilistic matching is permitted, and controlled:** Where no stable identifier is shared, agreement across name, address, device, contact details and behavioural evidence is scored into a single match confidence, which is then compared against a governed threshold. Only matches at or above the threshold are merged. The framework here is the classical one, in which the comparison of two records yields a likelihood ratio that is partitioned into a match region, a non-match region and an intermediate region reserved for cases the evidence does not settle [10], and the practical machinery of similarity measures, blocking and classification is well documented [8], [5]. The threshold is a governance parameter rather than a tuning knob, because moving it changes which customers exist in the model.
3. **Ambiguous matches are deliberately not merged:** This is the decision that most distinguishes the approach and it is the one most often made the other way. When the evidence does not clear the threshold, the records are retained separately, or held together at lower confidence, until stronger evidence arrives. They are not merged optimistically and reviewed later, because a wrongly merged profile does not announce itself: it produces a customer with contradictory usage, doubled revenue and an incoherent service history, and every feature computed from that profile is then wrong in a way that looks like signal. Two costs were accepted knowingly. Coverage is lower, so some genuine identity links are missed and some customers appear in the model as more than one entity. And the model sees less history for those customers than it could. Both costs are visible and bounded. The cost of a false merge is neither, because it propagates into features, into predictions and into the treatment decisions taken on the strength of them. Records held in the ambiguous tier are re-scored when a stronger identifier appears, which is common in telecom, where an authenticated digital login, a new contract or a verified contact detail can arrive at any point in a relationship and resolve retrospectively what was previously undecidable.
4. **Source-system lineage is maintained throughout:** Every attribute carries the identity of the system it came from, and every model feature can be traced back through the profile to that origin and to the tier at which the contributing records were resolved. This is a direct application of provenance practice [13], and it earns its cost in three ways. It makes

feature-level debugging possible, because a suspicious feature can be traced to a source rather than guessed at. It makes drift attribution possible, because a distribution shift can be localised to a system rather than attributed vaguely to customer behaviour, which matters for the retraining decisions in Section 11. And it makes explanation defensible, because a driver presented to a commercial user can be evidenced back to the system that observed it.

Table 1: Identity Resolution Tiers, the Confidence Policy Applied at Each, and the Disposition of the Record. The Intermediate Tier Is Retained by Design, Not by Omission

Tier	Rule and Confidence Policy	Disposition
Deterministic	Shared stable identifier: customer or account ID, subscriber ID, authenticated digital ID, hashed email or phone, contract or service ID	Merged into one profile. No probabilistic reasoning applied
Probabilistic, above threshold	Agreement across name, address, device, contact detail and behaviour scored into one match confidence, compared with a governed threshold	Merged, with the tier recorded in lineage
Probabilistic, below threshold	Confidence falls short of the governed threshold	Not merged. Records retained separately or held at lower confidence
Ambiguous, held	Evidence is contradictory or insufficient to settle the question	Not merged. Re-scored when a stronger identifier arrives
All tiers	Source-system lineage recorded for every attribute and every derived feature	Feature traceable to originating system and resolution tier

The governing principle is that an uncertain match must never contaminate the profile that feeds the value model. Recall is traded for profile integrity, deliberately, and the trade is recorded rather than hidden.

6. Event Normalisation and the Unified Behavioural Profile

Resolving identity produces one customer. It does not produce one history. Each contributing system records interactions in its own vocabulary, at its own granularity, with its own idea of what constitutes an event: a web session, a store task, a ticket with a severity code, a campaign delivery receipt, a partner order in a partner schema, a usage counter reading. Left in that state, the only features that can be computed across channels are counts, and counts discard the information that matters.

Activity from every channel was therefore normalised into a common behavioural model with nine event types: product inquiry, offer response, service request, complaint, purchase, upgrade, cancellation intent, digital engagement and support interaction. The normalisation is deliberately coarse, because a vocabulary that preserves every source-system nuance is not a common vocabulary. What matters is that the nine types mean the same thing regardless of which channel produced them.

Four attributes are retained on every event without exception. The timestamp, because sequence and interval carry most of the predictive content. The source channel, because which channel observed a behaviour is itself informative and because Section 10 depends on being able to measure engagement per channel. The product context, because an upgrade enquiry about one service and a complaint about another are not interchangeable. And the outcome, because an offer that was declined, an offer that was ignored and an offer that was accepted are three different facts that a naive contact count renders identical.

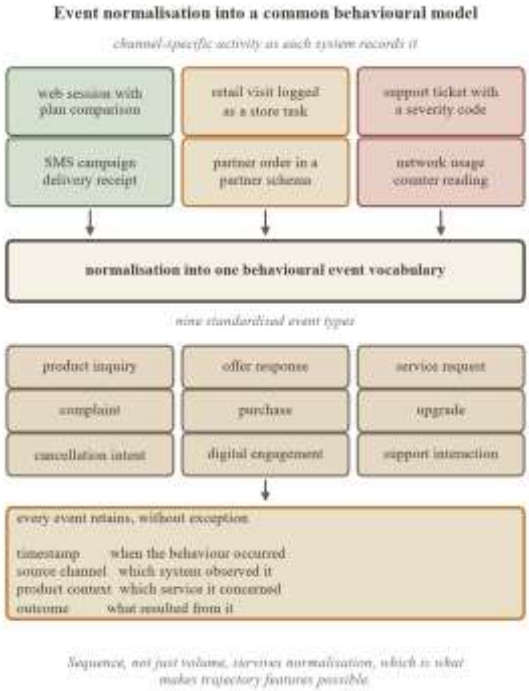


Fig 3: Event Normalisation. Channel-Specific Activity Is Mapped Into One Behavioural Event Vocabulary of Nine Standardised Types, and Every Event Retains Its Timestamp, Source Channel, Product Context and Outcome

The payoff is that patterns spanning several channels become computable. The example that makes the case concrete is one that occurs regularly in a telecom base and is invisible to every system that participates in it. Network usage declines gradually over several weeks. The customer

contacts support three times about the same unresolved issue. A retention email is sent and never opened. Shortly afterwards, the customer compares plans on the website.

Read individually, each of these is unremarkable. Usage varies seasonally and for many benign reasons. Support contacts happen. Emails go unopened constantly. Plan comparison is a normal browsing behaviour. Read together, in sequence, with timestamps and outcomes attached, they describe a customer whose future value is falling: consumption is dropping, the service experience has failed repeatedly, the retention channel has lost their attention, and they are now examining alternatives. No single source system holds this pattern. The network platform sees a usage curve. The service platform sees three tickets. The marketing platform sees a non-open. The digital platform sees a session. The pattern exists only in the unified profile, and it is the kind of pattern that carries most of the predictive information about future value.

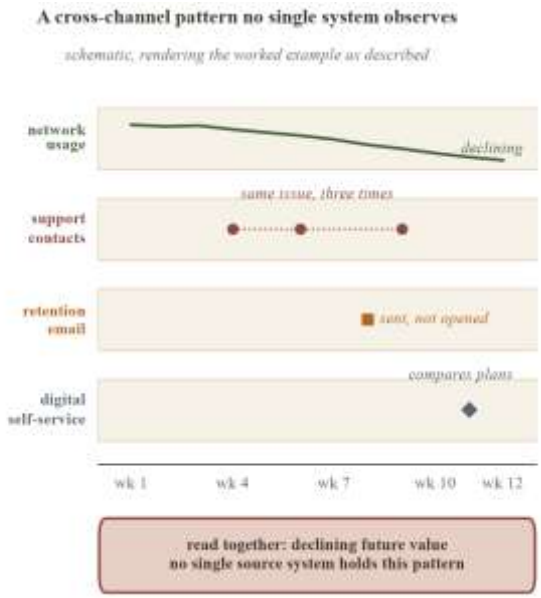


Fig 4: A Cross-Channel Pattern No Single System Observes, Rendered across Channel Lanes Over a Twelve Week Window. Schematic. The Lanes Show the Mechanism Described Above Rather Than Measured Data from a Particular Customer

This is also why the multichannel management literature treats the customer, rather than the channel, as the unit of analysis [16]. Once the profile is unified and events are normalised, the analytical question changes from what happened in this channel to what this customer is doing across all of them.

7. Feature Construction from Trajectory Rather Than Snapshot

Features were built from seven signal domains. Within each domain, the raw inputs are of secondary interest; what matters is the family of derived features constructed from them, and in particular whether those features describe a level or a change.

Table 2: Signal Domains, Their Representative Inputs, and the Derived Feature Families Built from Them

Signal Domain and Representative Inputs	Derived Feature Families
Revenue and billing. Average revenue per user, recurring charges, historical revenue, payment behaviour, outstanding balances, discounts, revenue trend	Monetary level and revenue trajectory, payment reliability, discount dependence, balance and arrears behaviour
Product and subscription. Service count and type, plan changes, upgrades and downgrades, add-on adoption, tenure, renewal dates, product combinations	Product depth, adoption and removal trend, plan movement direction, contract position relative to renewal, product affinity
Network and usage. Voice and data consumption, usage pattern change, roaming, device activity, significant declines or increases	Usage level and usage trajectory, volatility, roaming and device change, magnitude and persistence of decline
Customer relationship management and sales. Opportunity history, prior offers, campaign responses, sales interactions, segment, acquisition channel	Offer and campaign responsiveness, opportunity progression, acquisition-channel context
Service experience. Support cases, complaint frequency, case severity, resolution time, repeat contacts, escalations, service disruptions	Service friction, repeat-contact intensity on a single issue, escalation and unresolved-issue

	exposure, resolution latency
Digital and multi-channel engagement. Web and app activity, email and short message engagement, contact-centre interactions, partner-channel activity, responses to prior recommendations	Engagement recency and frequency per channel, engagement intensity across all channels, channel mix and channel migration
Churn and loyalty indicators. Tenure, engagement decline, contract expiry proximity, historical retention activity, behaviour change	Churn probability, engagement decline slope, renewal-window exposure, behavioural change magnitude

Across the domains, the features express eight underlying quantities: recency, frequency, monetary value, behavioural trend, engagement intensity, service friction, and product affinity and churn probability. Three of those eight, and arguably four, are inherently temporal. That is the design decision the section title refers to.

The construction rule is that wherever a level can be computed, its trajectory is computed alongside it. Average revenue per user becomes average revenue per user plus its slope over recent months plus the magnitude and persistence of any deviation from the customer's own baseline. Product count becomes product count plus the direction of recent movement, so that a customer at three services after removing two is distinguished from a customer at three services after adding one. Support case volume becomes case volume plus repeat-contact intensity on a single unresolved issue, which is a different quantity from total case count and, as Section 9 records, a considerably more informative one.

Two properties of this construction are worth noting. First, trajectory features are only computable because event normalisation preserved timestamps and outcomes; a normalisation that collapsed events into monthly counts would have made most of them unavailable. Second, trajectory features are the ones most sensitive to identity errors, because a false merge fabricates apparent changes in level where there was only the arrival of a second customer's history. The strictness of Section 5 and the trajectory emphasis of this section are therefore connected requirements rather than two independent choices.

8. The CLV Formulation and Prediction Horizons

Predicted value is expressed in a form deliberately kept legible to commercial users:

Predicted CLV = Expected Future Revenue x Retention Probability - Expected Cost-to-Serve

Three properties of this formulation matter more than its algebra. It is multiplicative in retention, so a high revenue expectation attached to a customer unlikely to stay does not produce a high value estimate. It carries cost to serve as an explicit negative term rather than as an unstated assumption, which means two customers with identical revenue are not assumed to be economically identical: a customer requiring frequent assisted support, repeated truck rolls or heavy contact-centre time is worth measurably less than a customer generating the same revenue without them. Measuring customer profitability rather than customer revenue is long-established practice in the relationship management literature [17], and this term is where that practice enters the model. And each of the three components is separately estimable, separately monitorable and separately explainable, which matters operationally: when a value estimate moves, the question of which component moved has an answer.

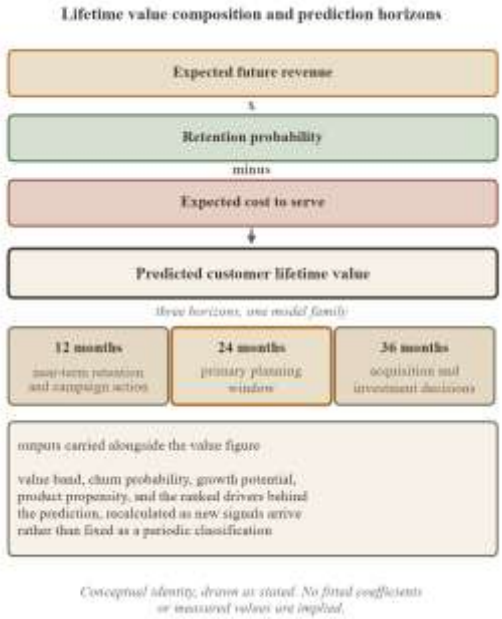


Fig 5: Value Composition and the Three Prediction Horizons, with the Outputs Carried Alongside the Value Figure. The Identity Is Conceptual, Drawn as Specified. No Fitted Coefficients or Measured Values Are Implied

Three horizons are produced from the same model family. The twenty four month horizon is the primary planning window, on the grounds that it is long enough for trajectory to express itself and short enough that the prediction remains actionable. A twelve month horizon supports near-term retention and campaign decisions, where the relevant question is what will happen inside the current budget cycle. A thirty six month horizon supports acquisition and investment decisions, where the relevant question is whether a customer segment repays its acquisition cost across a full relationship rather than a single contract term. Running all three from one feature base means the horizons are consistent with one another, and it makes the horizon itself a reporting choice rather than a modelling commitment.

The system does not output a single monetary figure. Alongside predicted value it carries a value band, a churn probability, a growth-potential assessment, product propensity scores, and the ranked drivers behind the individual prediction. The drivers are not decoration. A value figure without an explanation of what produced it cannot be

acted on, because the commercial user's next question is always what to do, and that question is answered by the drivers rather than by the number.

Finally, value is recalculated as new signals arrive rather than assigned as a periodic classification. This is a substantive difference from a quarterly scoring cycle. A customer whose usage began falling in week two of a quarter should not be carried at their old band for another ten weeks, and the events that most reliably precede value loss, service failure and engagement withdrawal, are exactly the ones that occur between scoring dates.

9. Which Signals Carry the Most Weight?

The strongest predictors were behavioural changes over time rather than static attributes. Stated in the order of their observed contribution, the groups were as follows:

1. **Revenue and usage trajectory:** The direction and consistency of movement in revenue and consumption mattered more than their level. A sustained decline was informative even when the absolute figures remained comfortable.
2. **Product depth and adoption:** Additional services, add-on adoption and upgrades were positive. Downgrades and add-on removal were strong negatives, and were among the clearest single indicators in the feature set, because removing a service is a deliberate act that customers rarely perform casually.
3. **Service experience.** Repeat cases, escalations, unresolved issues, long resolution times and repeated contact about the same problem were consistently associated with value deterioration.
4. **Engagement recency and frequency:** How recently and how often a customer engaged, measured across all channels rather than the digital ones, tracked future value more closely than any single-channel engagement measure.
5. **Tenure and contract position:** Tenure alone was weak. Tenure interacting with renewal proximity, recent product changes and engagement direction was informative, which is the expected result: the same length of relationship means something different for a customer six weeks from renewal who has just removed a service than for a customer mid-term who has just added one.
6. **Churn risk indicators:** Established churn signals contributed, as would be expected from the telecom churn literature [20], but they arrive relatively late in the sequence and are better understood as confirming a trajectory already visible in the preceding groups.

Two findings were more surprising than the ordering, and both cut against the intuitions that shaped the previous approach. The first is that service-interaction patterns mattered more than expected. Repeated contacts about the same unresolved issue flagged value deterioration even while revenue remained high. This is a case where the financially healthiest looking customers were, on a behavioural reading, already leaving. The mechanism is not mysterious: an unresolved problem contacted three times is a customer who has told the operator three times that the relationship is not working, and no billing record captures that.

The second is that high current revenue did not imply high predicted lifetime value. The converse also held. Some moderate-revenue customers scored highly on the strength of rising usage, strong digital engagement, growing product adoption, consistent payment behaviour and low service friction. A value model that reproduces the existing revenue ranking has not added anything; the disagreements between predicted value and current revenue are where the information is, and those disagreements were substantial.

10. Fairness, Channel Bias and the Service-Entitlement Boundary

Bias control was treated as an architectural and governance matter rather than as a validation step performed once before release. Four elements were in place:

1. **Inputs:** Protected characteristics were not used as direct predictors. Potential proxies were reviewed both for whether they were necessary to the model and for whether they produced disparate impact across groups, and the feature set was kept on behavioural and service signals rather than demographics. The general taxonomy of bias sources and fairness criteria informs this design [15], and the review of proxies is the part that requires judgement rather than a metric, because a feature can be predictive, non-demographic and still function as a proxy.
2. **Segment-level performance testing:** Performance was tested across segments rather than in aggregate. The segments examined were geography, acquisition channel, product type, tenure band, and digital as against non-digital engagement. Within each, the quantities examined were prediction error, false positive and false negative behaviour, the shape of the predicted value distribution, churn classification behaviour and recommendation outcomes. Aggregate accuracy was explicitly not accepted as sufficient, because a model can be accurate overall and systematically wrong about a segment that is small in the training data and important in the business.
3. **Explainability:** Users are shown the main factors behind a prediction. This serves fairness as well as usability: an unexplained value score cannot be challenged, and an unchallengeable score is one whose biases persist.

10.1. Channel Bias

This was the most consequential finding of the fairness work and it does not appear in the general fairness literature in this form. Engagement measured on digital channels alone systematically misrepresents customers who do not use those channels. A customer who rarely opens the mobile application and does not respond to email can appear disengaged while interacting heavily through retail stores,

partner channels or the contact centre. Because engagement is among the strongest inputs to predicted value, this measurement artefact translates directly into lower predicted value, then into reduced retention investment, then into a worse commercial outcome for a customer who was in fact highly engaged. The pattern is structurally similar to the finding that neutral-looking algorithmic delivery can produce systematically unequal exposure without any protected attribute being used [14]. The correction is to model engagement across all channels rather than the digital subset. It is a fairness control first and a feature-engineering improvement second, and it is not a small one, because the customers most likely to be misread this way are older customers, customers in regions with lower digital adoption and customers who simply prefer assisted service.

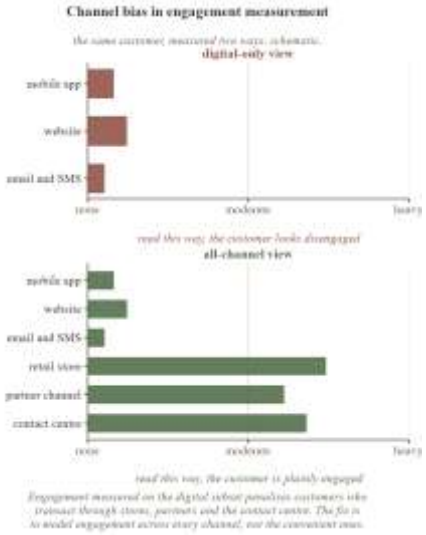


Fig 6: Channel Bias in Engagement Measurement. The Same Customer Measured on the Digital Subset and across All Channels. Schematic, Illustrating the Mechanism Rather Than a Measured Customer.

10.2. The Separation of Economic Prediction from Service Entitlement

This is stated as a design principle because it is one. Predicted lifetime value may determine which retention offer is appropriate, which cross-sell is worth making and how commercial attention is prioritised. It was never used to determine who receives essential support, whose complaint is resolved or who gets network assistance. A low predicted value never lowered service quality. The distinction is between commercial treatment, where differentiation by expected value is legitimate and ordinary, and service entitlement, where it is not. The reason for writing the boundary into the design rather than trusting operational good sense is that the pressure to cross it is continuous and always has a plausible efficiency argument attached, and a model score that is available in the service desk will eventually be used there unless something prevents it.

10.3. Post-Deployment Monitoring for Feedback Effects

The final element is a warning the author states directly, and it is the sharpest fairness problem in the whole system. A lifetime value model can bias itself over time. If customers predicted to be high value consistently receive the best offers, the best service attention and the most retention investment, their realised value grows because of the treatment they received, while customers who received less appear less valuable in the next training set. The model then learns its own past decisions as though they were customer behaviour. This is the mechanism demonstrated in recommendation settings, where algorithmic confounding progressively narrows what the system can see and degrades the utility of its own recommendations [4]. Section 11 describes the monitoring that watches for it. Monitoring mitigates the problem. It does not remove it.

11. Monitoring, Drift and Retraining Policy

Monitoring operates at three levels, and the third is the one most often omitted.

Table 3: The Three Monitoring Layers, what Each Measures, and What Triggers a Retrain

Layer	What It Measures	Retraining Trigger
Data drift	Current feature distributions against the training baseline for average revenue per user, usage, product mix, service-contact frequency, digital engagement, payment behaviour and channel preference, using population stability index and distribution-distance tests	Major drift in features the model depends on, after the cause is investigated
Prediction drift	Distributions of predicted value, churn probability and value segments. Once outcomes land, comparison with realised revenue and retention using mean absolute error, root mean squared error, ranking performance, calibration and segment-level error	Sustained high prediction error or declining ranking performance
Business drift	Whether high-value cohorts continue to deliver higher realised value, and whether model-driven treatment continues to outperform existing segmentation or holdout control groups	Business key performance indicators falling below baseline, or a material business event

1. **Data drift:** Current feature distributions are compared against the training baseline across average revenue per user, usage, product mix, service-contact frequency, digital engagement, payment behaviour and channel preference. Population stability index and distribution-distance tests are the instruments. This is standard practice in the concept drift literature [11] and it is the cheapest of the three layers to run, which is why it is often mistaken for sufficient.
2. **Prediction drift:** The distributions of predicted lifetime value, churn probability and value segment

assignment are monitored directly, because a model can receive stable inputs and still shift its outputs. Once actual outcomes become observable, predictions are compared against realised revenue and retention using mean absolute error, root mean squared error, ranking performance, calibration and segment-level error. No values for these metrics were available for this paper and none are reported here; what is described is the measurement design. Ranking performance is tracked separately from error magnitude for a practical reason: most decisions taken on a value model are ordering decisions, so a model whose absolute predictions have drifted but whose ordering holds is in a different condition from one whose ordering has broken.

3. **Business drift:** The third layer asks whether the model is still commercially useful, which is not answered by either of the first two. Two questions are tracked: whether cohorts predicted to be high value continue to deliver higher realised value, and whether model-driven treatment continues to outperform the previous segmentation approach or a holdout control group. A statistically stable model can lose practical value, and this is the layer that detects it. It is also, not incidentally, the layer that would surface the self-reinforcement problem described in Section 10, because a model that is grading its own past decisions tends to look excellent on prediction accuracy while its incremental effect against a control group decays. The distinction between statistical performance and decision value is the same one made in the profit-oriented churn literature [20], and comparison against a control group rather than against the model's own historical outputs is what uplift-oriented evaluation requires [7].
4. **Retraining is scheduled and event-driven:** A quarterly retraining cycle runs regardless. Event-based retraining is triggered by major drift in features the model depends on, sustained high prediction error, declining ranking performance, business indicators falling below baseline, or business events including new pricing, portfolio changes, acquisitions and the addition of a significant new channel. The last of these is specific to the multi-channel design: a new channel changes the meaning of the engagement features, because a customer's channel mix is measured against a set of channels that has just changed.
5. **The model is deliberately not retrained on every fluctuation:** A drift signal is investigated before it is acted on, and the investigation asks which of three things it is: a seasonal effect, a data-quality problem in a source system, or a genuine behavioural shift. Only the third warrants retraining. The other two warrant fixing something else, and retraining on them teaches the model an artefact. Source-system lineage from Section 5 does real work here, because it allows a distribution shift to be localised to a specific system rather than attributed to customers in general, which is usually enough to distinguish a broken feed from a change in behaviour.
6. **Promotion is not automatic:** Each retrained model is evaluated against the deployed version on a holdout set and, where appropriate, in champion-challenger form. Promotion requires predictive performance, stability, fairness across the segments listed in Section 10, and business impact. A candidate that improves aggregate error while degrading performance on a segment is not promoted. This gate is where the fairness testing and the monitoring architecture meet, and it reflects the production-readiness view that a model without monitoring and rollback discipline is not deployable regardless of its offline scores [3].

12. Pilot Results

This section reports outcomes in the form the author specified: as pilot and implementation ranges, not as audited production results. Every figure below is a range and is left as a range. None has been substantiated by a business report or a finance review, and none should be read as a verified production outcome.

Table 4: Pilot and Implementation Figures. These Are Pilot Ranges and Implementation Scale Figures, Not Audited Production Results

Indicator	Reported Range	Status
Retention effectiveness within prioritised high-value or high-risk segments	Improvement of about 8 to 12 per cent	Pilot range
Targeting efficiency	Improvement of about 20 to 30 per cent	Pilot range
Expected value or incremental revenue uplift within targeted cohorts	About 10 to 18 per cent	Pilot range
Acquisition-cost efficiency for higher-value cohorts	About 10 to 15 per cent	Pilot range
Unnecessary retention incentives and marketing contacts	Reduced. Not quantified	Directional, unquantified
Customer or subscriber records	Approximately 1 to 5 million	Implementation scale
Behavioural and transactional records	Tens of millions	Implementation scale
Primary channel and data domains	Seven	Implementation scale

The seven domains are customer relationship management and sales, billing and payments, network and service usage, contact centre and support, web and mobile digital engagement, marketing channels including email, short message and campaign response, and partner or retail channels.

Two clarifications are necessary and neither is a formality.

On the impact figures: They are pilot and implementation ranges. The range is the finding; narrowing any of them to a point estimate would assert a precision the underlying evidence does not support, and no such narrowing appears anywhere in this paper. The customer count is likewise given as a range because an exact production count could not be substantiated.

On accuracy: No model accuracy figures were available and none are reported. There is no value in this paper for mean absolute error, root mean squared error, mean absolute percentage error, coefficient of determination, precision, recall, F-score or area under the curve. Those metrics appear in Section 11 as the quantities the framework tracks, with no values attached. Published churn studies in telecom report their own discrimination figures under their own data conditions and evaluation protocols [20]; nothing here is a claim to those results, and no external benchmark should be read as standing in for a measurement of this system. The honest position is that the modelling and governance design is documented and the empirical validation is not, and stating that plainly is preferable to importing a comparable figure from elsewhere.

13. What Lifetime Value Models Get Wrong

Eight recurring errors are worth naming, because each has a specific consequence and most are still standard practice.

- Treating lifetime value as a financial calculation rather than evolving behavior: A discounted revenue projection applied to a static customer record is arithmetic performed on a snapshot. Value is a property of a relationship that is moving.
- Overweighting current revenue: This is the most common failure and the most expensive. It systematically overvalues customers who are paying well while disengaging, and undervalues customers whose behaviour is improving from a modest base. Section 9 found both errors in the data.
- Fragmented channel data: A value model reading one system inherits that system's blind spots, as Section 4 sets out. This is the failure this paper is principally about.
- Treating value as static: A score computed quarterly is stale for most of the quarter, and it is stale precisely during the periods when the behaviour that changes value is occurring.
- Predicting value without explaining it: An unexplained score is operationally useless, because the user's next question is what to do about it and a number does not answer that. It is also unchallengeable, which means its errors survive.
- Ignoring cost to serve: Two customers with equal revenue are not equally valuable. Treating revenue as value overstates the worth of expensive-to-serve customers and misdirects retention spending towards them. Customer profitability rather than customer revenue has been the established measurement position for two decades [17], and it remains widely ignored in practice.
- Confusing prediction with treatment: Identifying a high-value customer does not indicate what action would raise that value. The two questions are separate, and the evidence is that targeting on predicted risk alone can be ineffective because risk and responsiveness to treatment are different quantities [1]. A value model must therefore connect to churn, propensity, next-best-action and treatment-response models; uplift-style estimation of incremental effect is the relevant machinery [7]. A lifetime value score that terminates in a dashboard is a measurement, not a capability.
- Creating feedback loops that reinforce historical decisions: The self-reinforcement mechanism in Section 10 is the subtlest of the eight and the hardest to detect, because the model's accuracy metrics improve while its actual usefulness decays [4]. It requires deliberate monitoring against control groups rather than against the model's own history.

14. Limitations

The limitations of this work are substantial and specific.

- Business figures are pilot and implementation ranges, not audited production results: No business report or finance validation was available to substantiate them. They are reported as ranges because point estimates could not be substantiated, and they should be read as indicative of the scale of effect observed in pilot conditions rather than as measured production outcomes.
- Scale figures are implementation scale, not verified production counts: The record volumes in Section 12 describe the scale at which the implementation operated. The customer figure is deliberately given as a range because an exact production count could not be verified, and no exact count is asserted anywhere.
- No model accuracy figures of the author's own were available and none are reported: The framework tracks error, ranking and calibration metrics, and this paper describes that measurement design without reporting values. A reader who requires quantitative accuracy evidence will not find it here, and should not infer it from the published figures of other studies.
- Probabilistic identity matching carries irreducible risk: The ladder in Section 5 reduces false merges by refusing uncertain ones, and in doing so it accepts missed matches. Both error modes remain present at some rate. The rate itself was not

quantified in this work, which is a real gap: a full treatment would report false-match and missed-match rates against a labelled evaluation set, and no such set was available. The confidence threshold is therefore a governance judgement operating without a measured error curve beneath it.

- Simulated or projected value uplift is not causal evidence: The improvements in Section 12 were observed in pilot conditions alongside other operational changes. Establishing that the model caused them would require controlled comparison against holdout groups sustained over a full prediction horizon, which is what the business drift layer in Section 11 is designed to support but not what the figures reported here rest on.
- The feedback-loop risk is live rather than resolved: Any deployed lifetime value model that influences treatment will influence the data it is later trained on. The monitoring in Section 11 is designed to detect this and the entitlement boundary in Section 10 limits its worst consequences. Neither eliminates it. Reporting it as a managed risk rather than a solved problem is the accurate position.
- Generalisation is bounded by estate structure: The seven-domain structure, the identifier availability that determines how much work each resolution tier does, and the channel mix that makes channel bias severe are all properties of a particular kind of telecom estate. Operators with different channel structures will find the same principles applying with different weights.

15. Conclusion

The question a telecom operator needs answered is not what a customer is worth today but what they are likely to be worth across the relationship and what action now would improve that. The obstacle to answering it is rarely the estimator. It is that the evidence sits in seven places and is evaluated in seven places, so the patterns that indicate a change in trajectory are never assembled.

This paper has treated the assembly as the substantive problem. Identity is resolved through a ladder that prefers deterministic matching, permits probabilistic matching only above a governed confidence threshold, and deliberately refuses to merge ambiguous records rather than merging them optimistically, on the argument that a false merge propagates silently into features, predictions and decisions while a missed match is visible and bounded. Source-system lineage makes every feature traceable to its origin, which is what makes drift attributable and explanation defensible. Channel-specific activity is normalised into one behavioural vocabulary that preserves timestamp, channel, product context and outcome, which is what makes cross-channel patterns computable and trajectory features possible. Value is then expressed as expected future revenue multiplied by retention probability less expected cost to serve, over three horizons with the twenty four month window primary, and recalculated as signals arrive rather than assigned quarterly.

Two governance positions are argued as design principles. Engagement must be modelled across every channel, because measuring it on the digital subset penalises customers who transact through stores, partners and the contact centre, and that is a fairness failure rather than a data gap. And economic prediction must be separated from service entitlement: predicted value may shape which offer a customer receives and must never shape whether their problem gets fixed.

The evidence position is stated plainly rather than dressed up. The architecture and the governance design are documented. The business figures are pilot and implementation ranges, kept as ranges. No model accuracy figures were available and none are claimed. What the work establishes is that in a multi-channel telecom estate the quality of a lifetime value prediction is set largely by decisions made before the model is fitted, in how identity is resolved, how behaviour is normalised and which channels are permitted to count as engagement.

References

1. E. Ascarza, "Retention futility: Targeting high-risk customers might be ineffective," Journal of Marketing Research, vol. 55, no. 1, pp. 80-98, 2018. doi:10.1509/jmr.16.0163
2. R. C. Blattberg, E. C. Malthouse, and S. A. Neslin, "Customer lifetime value: Empirical generalizations and some conceptual questions," Journal of Interactive Marketing, vol. 23, no. 2, pp. 157-168, 2009. doi:10.1016/j.intmar.2009.02.005
3. E. Breck, S. Cai, E. Nielsen, M. Salib, and D. Sculley, "The ML test score: A rubric for ML production readiness and technical debt reduction," in Proc. 2017 IEEE International Conference on Big Data, 2017, pp. 1123-1132. doi:10.1109/bigdata.2017.8258038
4. A. J. B. Chaney, B. M. Stewart, and B. E. Engelhardt, "How algorithmic confounding in recommendation systems increases homogeneity and decreases utility," in Proc. 12th ACM Conference on Recommender Systems, 2018, pp. 224-232. doi:10.1145/3240323.3240370
5. P. Christen, Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection. Berlin, Germany: Springer, 2012. doi:10.1007/978-3-642-31164-2
6. B. N. Dabbara, "AI-driven predictive analytics in telecom CRM for customer lifetime value enhancement," Journal of Advances in Developmental Research, vol. 16, no. 2, 2025. doi:10.71097/ijaidr.v16.i2.1639
7. F. Devriendt, D. Moldovan, and W. Verbeke, "A literature survey and experimental evaluation of the state-of-the-art in uplift modeling: A stepping stone toward the development of prescriptive analytics," Big Data, vol. 6, no. 1, pp. 13-41, 2018. doi:10.1089/big.2017.0104
8. A. K. Elmagarmid, P. G. Ipeirotis, and V. S. Verykios, "Duplicate record detection: A survey," IEEE Transactions on Knowledge and Data Engineering, vol. 19, no. 1, pp. 1-16, 2007. doi:10.1109/tkde.2007.250581

9. P. S. Fader, B. G. S. Hardie, and K. L. Lee, "RFM and CLV: Using iso-value curves for customer base analysis," Journal of Marketing Research, vol. 42, no. 4, pp. 415-430, 2005. doi:10.1509/jmkr.2005.42.4.415
10. I. P. Fellegi and A. B. Sunter, "A theory for record linkage," Journal of the American Statistical Association, vol. 64, no. 328, pp. 1183-1210, 1969. doi:10.1080/01621459.1969.10501049
11. J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," ACM Computing Surveys, vol. 46, no. 4, pp. 1-37, 2014. doi:10.1145/2523813
12. S. Gupta, D. Hanssens, B. Hardie, W. Kahn, V. Kumar, N. Lin, N. Ravishanker, and S. Sriram, "Modeling customer lifetime value," Journal of Service Research, vol. 9, no. 2, pp. 139-155, 2006. doi:10.1177/1094670506293810
13. M. Herschel, R. Diestelkämper, and H. Ben Lahmar, "A survey on provenance: What for? What form? What from?," The VLDB Journal, vol. 26, no. 6, pp. 881-906, 2017. doi:10.1007/s00778-017-0486-1
14. A. Lambrecht and C. Tucker, "Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads," Management Science, vol. 65, no. 7, pp. 2966-2981, 2019. doi:10.1287/mnsc.2018.3093
15. N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," ACM Computing Surveys, vol. 54, no. 6, pp. 1-35, 2021. doi:10.1145/3457607
16. S. A. Neslin, D. Grewal, R. Leghorn, V. Shankar, M. L. Teerling, J. S. Thomas, and P. C. Verhoef, "Challenges and opportunities in multichannel customer management," Journal of Service Research, vol. 9, no. 2, pp. 95-112, 2006. doi:10.1177/1094670506293559
17. L. Ryals, "Making customer relationship management work: The measurement and profitable management of customer relationships," Journal of Marketing, vol. 69, no. 4, pp. 252-261, 2005. doi:10.1509/jmkg.2005.69.4.252
18. D. C. Schmittlein, D. G. Morrison, and R. Colombo, "Counting your customers: Who are they and what will they do next?," Management Science, vol. 33, no. 1, pp. 1-24, 1987. doi:10.1287/mnsc.33.1.1
19. A. Vanderveld, A. Pandey, A. Han, and R. Parekh, "An engagement-based customer lifetime value system for e-commerce," in Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 293-302. doi:10.1145/2939672.2939693
20. W. Verbeke, K. Dejaeger, D. Martens, J. Hur, and B. Baesens, "New insights into churn prediction in the telecommunication sector: A profit driven data mining approach," European Journal of Operational Research, vol. 218, no. 1, pp. 211-229, 2012. doi:10.1016/j.ejor.2011.09.03