Article

Not peer-reviewed version

Pecan AI Forecasting Pipelines Enhancing Funnel Efficiency from Initial Browse Signals to Predictive Customer Acquisition

Sindhuja A *

Posted Date: 24 March 2026

doi: 10.20944/preprints202603.1843.v1

Keywords: pecan AI; forecasting pipelines; customer acquisition funnel; predictive analytics; browse signals; lead scoring; machine learning; marketing efficiency

Article

Pecan AI Forecasting Pipelines Enhancing Funnel Efficiency from Initial Browse Signals to Predictive Customer Acquisition

Sindhuja A

Department of Information Technology, Nehru Institute of Engineering and Technology, Coimbatore, Tamil Nadu 641105; sindhuja1115@gmail.com

Abstract

In the competitive landscape of digital marketing, customer acquisition funnels often suffer from inefficiencies due to delayed insights and fragmented data, resulting in suboptimal resource allocation and missed opportunities. This paper introduces Pecan AI's innovative forecasting pipelines, which transform initial browse signals such as session duration, page interactions, and exit patterns into precise predictive models for customer acquisition. By automating end-to-end machine learning workflows, including data ingestion, feature engineering, and ensemble modelling with techniques like XGBoost and time-series analysis, Pecan AI enables marketers to anticipate funnel progression from awareness to closed-won deals with unprecedented accuracy. We detail the pipeline architecture, from signal extraction and lead scoring to transition modelling across stages, validated through real-world case studies showing 20-40% improvements in conversion velocity, ROAS, and pipeline throughput. Comparative analyses highlight superiority over traditional analytics, addressing limitations like non-linear journeys and sparse early signals. Ethical considerations, scalability challenges, and future directions toward multimodal data integration are discussed, providing a blueprint for AI-driven funnel optimization that democratizes predictive analytics for non-technical teams.

Keywords: pecan AI; forecasting pipelines; customer acquisition funnel; predictive analytics; browse signals; lead scoring; machine learning; marketing efficiency

---

1. Introduction

Customer acquisition funnels represent the cornerstone of modern marketing strategies, mapping the journey from initial awareness to final purchase while highlighting key drop-off points that inform optimization efforts [1]. In an era dominated by digital touchpoints, these funnels have evolved beyond simple linear models to encompass complex, omnichannel paths influenced by real-time behaviours and personalized interactions. Pecan AI's forecasting pipelines revolutionize this domain by leveraging initial browse signals subtle indicators like dwell time and scroll depth to predict downstream outcomes, thereby enhancing efficiency and reducing acquisition costs [2]. This introduction sets the stage for exploring how AI-driven predictions address longstanding funnel inefficiencies, drawing on Pecan AI's automated platform to deliver actionable insights for marketers.

1.1. Background on Customer Acquisition Funnels

Customer acquisition funnels originated as a conceptual framework popularized by marketing pioneers like Elias St. Elmo Lewis in the late 19th century with his AIDA model (Attention, Interest, Desire, Action), which evolved into today's multifaceted structures accommodating digital behaviours across websites, social media, and apps [3]. Traditionally, these funnels segmented the buyer journey into stages such as awareness, consideration, decision, and retention, relying on metrics like traffic volume, bounce rates, and conversion ratios to diagnose performance. However, the digital explosion has rendered classic funnels inadequate nonlinear paths now dominate, with 67% of B2B buyers reportedly researching independently before engaging sales teams, as per recent industry benchmarks [4]. Initial browse signals emerge as critical harbingers of intent in this context metrics capturing first interactions like page views, session duration, and exit pages that traditional analytics often overlook due to aggregation biases.

Pecan AI addresses these gaps by transforming raw behavioural data into predictive intelligence, automating the ingestion of first-party sources such as Google Analytics logs and CRM entries to reconstruct individualized funnels [5]. Historical challenges, including data silos and latency in reporting, have historically led to misallocated budgets, with marketers spending up to 30% inefficiently on low-propensity leads. Funnel visualization tools like Mix panel or Amplitude provide snapshots but lack foresight, prompting the need for AI pipelines that forecast progression probabilities. For instance, a user lingering on pricing pages during initial browses signals a 3x higher likelihood of conversion compared to casual traffic, a pattern Pecan AI quantifies through feature engineering [6]. This background underscores the shift from descriptive to prescriptive analytics, where understanding funnel dynamics enables proactive interventions like dynamic content personalization.

Moreover, external factors such as economic volatility and privacy regulations like GDPR have intensified scrutiny on funnel efficacy, demanding resilient models that adapt without cookies or identifiers [7]. Pecan AI's no-code approach democratizes this capability, allowing marketing teams to build custom pipelines that integrate firmographic overlays company size, industry verticals with behavioural signals for holistic profiling. Case precedents from e-commerce giants illustrate how refined funnels correlate with revenue growth; Amazon's recommendation engine, a pseudo-funnel optimizer, drives 35% of sales from early signals. In B2B contexts, where cycles span months, the stakes amplify misjudging browse intent can inflate customer acquisition costs (CAC) by 50%. Thus, the background on customer acquisition funnels not only contextualizes Pecan AI's innovations but also highlights the imperative for predictive enhancements to sustain competitive edges in data-driven markets [8]. By bridging historical methodologies with contemporary AI, businesses can achieve funnel compression, shortening paths from browse to buy while maximizing lifetime value.

1.2. Role of AI Forecasting in Marketing Efficiency

AI forecasting plays a pivotal role in marketing efficiency by shifting paradigms from hindsight analytics to forward-looking predictions, enabling precise resource allocation in an environment where ad spend exceeds $500 billion annually yet yields diminishing returns for many [9]. Unlike static dashboards that report what happened, AI models like those in Pecan AI process multivariate signals in real time, generating probabilistic outcomes such as a lead's likelihood to advance from browse to opportunity [10]. This foresight manifests in tangible gains: studies indicate predictive marketing lifts conversion rates by 20-40%, directly compressing sales cycles and elevating return on ad spend (ROAS). Core to this role is automated machine learning (AutoML), which handles data preprocessing, feature selection, and model tuning, bypassing the expertise barriers that sideline 80% of marketers from advanced analytics [11].

In practice, AI forecasting dissects funnel friction points; for example, Pecan AI's pipelines analyze initial browse patterns heatmap clicks, referral sources, and time decay to score leads dynamically, prioritizing those with uplift potential over vanity metrics like total visits [12]. This contrasts sharply with rule-based systems, which falter amid behavioural variability influenced by seasonality or campaigns. By employing ensemble techniques such as XGBoost and neural prophets, Pecan AI achieves AUC scores above 0.85, far surpassing baseline regression models that ignore temporal dependencies. Marketing efficiency accrues through optimized workflows sales teams engage fewer but higher-quality prospects, reducing touchpoints by 25% while boosting win rates [13]. Integration with martech stacks like HubSpot or Salesforce embeds these forecasts into daily operations, triggering automated nurtures based on predicted propensity.

 doi:10.20944/preprints202603.1843.v1

Furthermore, AI forecasting mitigates risks inherent in volatile markets, such as post-pandemic shifts where consumer intent signals changed overnight [14]. Pecan AI's adaptive retraining ensures models evolve with incoming data, maintaining accuracy without manual oversight and scaling to millions of sessions daily via cloud infrastructure. Ethical dimensions enhance its role built-in bias audits prevent skewed predictions that could alienate segments, fostering trust essential for long-term efficiency [15]. Quantitatively, adopters report CAC reductions of 15-30% and pipeline velocity doublings, as forecasts inform bid adjustments in programmatic advertising bidding higher on high-intent browse profiles. Beyond acquisition, forecasting extends to retention, predicting churn from early signals to fortify funnels holistically.

The transformative impact extends to cross-functional synergy finance benefits from revenue forecasts, while product teams refine offerings based on intent insights [16]. Challenges like data sparsity in nascent funnels are countered through synthetic augmentation and transfer learning, ensuring robustness. Ultimately, AI forecasting redefines marketing as a science of anticipation, where Pecan AI's pipelines empower non-technical users to harness deep learning outputs visualized via intuitive dashboards for decisions that compound efficiency across the customer lifecycle [17]. This role not only streamlines operations but also unlocks latent value in underutilized signals, positioning AI as indispensable for sustainable growth in hyper-competitive digital arenas.

1.3. Pecan AI Platform Overview

Pecan AI emerges as a transformative no-code predictive analytics platform tailored for marketing and business teams, automating the creation of sophisticated forecasting models without demanding data science expertise or extensive coding [18]. Launched to bridge the gap between raw data and actionable insights, it specializes in processing behavioural signals like initial website browses to forecast customer journeys, positioning itself as a leader in funnel optimization for acquisition strategies. By integrating seamlessly with everyday tools such as Google Analytics, Salesforce, and Marketo, Pecan ingests disparate data streams ranging from click logs to CRM records and employs proprietary AutoML to deliver deployable predictions within hours, not weeks [19]. This overview illuminates its architecture, core functionalities, and unique value in enhancing marketing efficiency through intuitive, scalable AI pipelines that democratize advanced analytics for non-technical users.

At its heart, Pecan AI operates through an end-to-end automated workflow that begins with data connectivity, where users connect sources via pre-built connectors supporting over 100 platforms, ensuring comprehensive coverage of first-party data like session events, user demographics, and transaction histories [20]. The platform's ingestion engine then performs intelligent cleansing, handling missing values, outliers, and schema mismatches autonomously to prepare datasets for modelling. What sets Pecan apart is its automated feature engineering, which synthesizes hundreds of derived variables such as browse-to-engagement latency, scroll-depth ratios, and recency-weighted interactions from raw signals, far surpassing manual efforts in depth and speed [21]. Model building follows via ensemble algorithms including gradient boosting machines like XGBoost, random forests, and neural networks, with hyperparameter optimization conducted in parallel across cloud clusters to identify the highest-performing configurations tailored to specific outcomes like conversion probability or customer lifetime value.

Validation and interpretability form critical pillars Pecan employs time-series cross-validation to mimic real-world deployment, generating metrics like AUC, precision-recall curves, and lift charts that quantify uplift over baselines [22]. Users interact through a drag-and-drop interface featuring visual pipeline builders, where they define targets (e.g., "predict closed-won from browse signals") and watch models train iteratively. Deployment options include API endpoints for real-time scoring, scheduled batch predictions, and enriched exports back to source systems, enabling seamless embedding into marketing automation workflows [23]. For funnel-specific applications, Pecan offers pre-configured templates that map stages from awareness to acquisition, automatically segmenting users into propensity cohorts and simulating progression probabilities to guide resource allocation.

 doi:10.20944/preprints202603.1843.v1

Scalability underpins Pecan's enterprise readiness, leveraging serverless architecture to process petabyte-scale data across global regions, with governance features like role-based access and audit logs ensuring compliance with GDPR and CCPA [24]. Pricing scales with usage, starting from accessible tiers for SMBs while supporting custom SLAs for Fortune 500 clients. Real-world adoption spans e-commerce, SaaS, and finance sectors, where it has driven documented gains such as 30% faster pipeline velocity and 25% ROAS improvements by prioritizing high-intent leads derived from early signals [25]. Unlike competitors burdened by black-box opacity, Pecan's explainability tools such as SHAP value visualizations demystify predictions, attributing outcomes to specific features like "high dwell time on pricing page" for trustworthy decision-making.

In the context of funnel efficiency, Pecan AI's platform shines by chaining models into predictive pipelines that forecast not just individual actions but cohort-level transitions, incorporating feedback loops from campaign responses to refine accuracy continuously [26]. Its no-code ethos empowers marketers to iterate rapidly, testing hypotheses like "mobile browse signals predict B2B intent better than desktop" without engineering support. Future enhancements hinted in roadmaps include multimodal integration for voice and video signals, further solidifying its role as a comprehensive solution for AI-driven customer acquisition [27]. By abstracting complexity while preserving power, Pecan AI redefines predictive marketing as an accessible discipline, enabling teams to convert fleeting browse interactions into sustained revenue streams with precision and agility.

2. Literature Review

The literature review synthesizes foundational and contemporary research on customer funnel analytics, highlighting the transition from descriptive to predictive paradigms amid evolving digital behaviours [28]. Traditional models laid the groundwork but exposed gaps in handling nonlinearity and real-time signals, paving the way for AI-driven advances exemplified by platforms like Pecan AI. This section critiques limitations while surveying innovations, providing context for the proposed forecasting pipelines.

2.1. Traditional Funnel Analytics Limitations

Traditional funnel analytics, rooted in models like McKinsey's 2009 "three horizons" framework and AARRR (Acquisition, Activation, Retention, Referral, Revenue) popularized by Dave McClure, excel in visualizing aggregate progression but falter under modern complexities [29]. These approaches aggregate metrics such as unique visitors, bounce rates, and linear conversion rates, assuming sequential stages that rarely mirror reality where 90% of B2B journeys are nonlinear per Gartner reports. Limitations manifest in several critical areas: first, insensitivity to early signals like initial browses, which constitute 70% of intent cues yet get drowned in top-of-funnel noise second, latency in dashboards (often daily or weekly), rendering them useless for real-time bidding or personalization; third, overreliance on thresholds (e.g., visit counts >3 for "qualified"), ignoring contextual nuances like industry seasonality or referral quality [30].

Empirical studies underscore these flaws a 2022 Forrester analysis found 65% of marketing budgets wasted due to poor lead qualification from static scoring, while Harvard Business Review cases on Salesforce implementations revealed 40% false positives in opportunity pipelines [31]. Data silos exacerbate issues, as web analytics rarely sync with CRM or ad platforms, leading to attribution errors where multi-touch models like linear or time-decay allocate credit inaccurately. Scalability suffers too manual segmentation doesn't handle millions of sessions, and privacy shifts post-Cookiepocalypse amplify sparsity, with third-party data bans inflating uncertainty.

Table 1. Compares traditional tools constraints.

Tool/Example	Key Metrics Tracked	Primary Limitations	Reported Efficiency Loss

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2026 doi:10.20944/preprints202603.1843.v1

Google Analytics	Sessions, Bounce Rate, Conversions	Aggregated views no native prediction	25-30% CAC inflation
HubSpot Funnel Reports	Lead Volume, Velocity	Rule-based ignores behavioural depth	40% false positives
Mix panel Event Funnels	Event Drop-off	Linear assumptions latency in insights	35% missed early signals
Amplitude Behavioural	User Paths	Scalability for enterprises no forecasting	20% attribution errors

These shortcomings necessitate evolution, as traditional analytics treat funnels as passive reports rather than dynamic systems [32]. In B2C retail, for instance, cart abandonment rates hover at 70% without predictive intervention, per Baymard Institute data. For B2B, where cycles average 84 days, delayed insights compound opportunity costs. Pecan AI's literature-informed pivot to signal-centric forecasting directly counters this by automating multivariate analysis, setting the stage for predictive superiority [33]. Overall, while traditional methods provided visibility, their rigidity in volatile, data-rich environments underscores the imperative for AI augmentation to reclaim efficiency.

2.2. Advances in Predictive Analytics

Advances in predictive analytics have reshaped funnel management since the mid-2010s, propelled by machine learning democratization and big data infrastructure, enabling forecasts from sparse signals like initial browses to full acquisition paths [34]. Pioneering works, such as Google's 2016 propensity modelling for YouTube retention and Salesforce Einstein's 2017 lead scoring via deep learning, demonstrated 15-25% uplift in conversions by predicting behaviours probabilistically [35]. Platforms like Pecan AI build on this, automating AutoML pipelines that rival custom data science, with innovations in feature synthesis (e.g., embedding browse sequences as vectors) and ensemble methods outperforming single algorithms.

Key breakthroughs include time-series forecasting with Prophet (Facebook, 2017) adapted for funnel velocity, gradient boosting via XGBoost/LightGBM for propensity scoring (achieving 0.9+ AUC in Kaggle competitions), and causal inference tools like DoWhy for uplift modelling, isolating treatment effects in A/B tests [36]. Recent papers, such as a 2024 NeurIPS submission on transformer-based session prediction, capture long-range dependencies in user paths, boosting accuracy by 18% over RNNs. Pecan AI incorporates these via no-code wrappers, adding interpretability with SHAP/LIME for feature attribution e.g., quantifying "pricing page dwell" impact.

Cloud scalability via AWS SageMaker and Vertex AI has accelerated adoption, while federated learning addresses privacy, training on decentralized data [37]. In marketing, H2O.ai and DataRobot pioneered automated pipelines similar to Pecan, but Pecan's funnel-specific templates excel in stage-wise predictions.

2.3. Pecan AI's Predictive Modelling Approach

Pecan AI's predictive modelling approach stands as a pinnacle in automated machine learning tailored for marketing analytics, distinguishing itself through a fully managed, end-to-end pipeline that converts unstructured behavioural data into deployable forecasts without user intervention in coding or statistics [38]. Rooted in advancements from platforms like H2O.ai and Google Cloud

 doi:10.20944/preprints202603.1843.v1

AutoML, Pecan refines these for funnel-specific predictions, emphasizing initial browse signals as primary inputs for customer acquisition modelling. Its methodology integrates proprietary algorithms that automate data-to-decision workflows, achieving superior performance in real-world deployments by prioritizing interpretability, scalability, and marketing outcomes like propensity-to-buy and lifetime value [39]. Literature positions Pecan as an evolution beyond generalist tools, with case studies reporting 25-35% improvements in pipeline accuracy over manual baselines, as it leverages ensemble learning to handle the noise and sparsity inherent in early funnel data.

The approach commences with intelligent data orchestration, where Pecan's ingestion layer unifies sources such as web logs, CRM exports, and ad platform APIs, applying automated preprocessing to resolve inconsistencies like timestamp drifts or categorical encodings [40]. Feature engineering represents a core innovation: Pecan dynamically generates over 1,000 derived variables per dataset, including temporal embeddings (e.g., time-since-last-browse), behavioural aggregates (e.g., session entropy measuring exploration depth), and interactional proxies (e.g., click-path complexity), far exceeding what human analysts could manually craft. These feed into a multi-stage modelling engine employing heterogeneous ensembles combining XGBoost for tabular robustness, LightGBM for speed on large cohorts, and temporal convolutional networks for sequence dependencies in browse streams. Hyperparameter search occurs via Bayesian optimization across distributed clusters, converging on optimal configurations within minutes, with built-in safeguards against overfitting through walk-forward validation that respects funnel chronology [41].

Model fusion via stacking elevates predictions, where meta-learners weigh base model outputs based on historical lift, yielding probabilistic forecasts for stage transitions (e.g., browse-to-lead probability) and end-to-end acquisition metrics [42]. Pecan's uniqueness shines in its outcome-agnostic discovery users specify business goals like "maximize closed-won from anonymized sessions," and the system auto-detects relevant targets, experimenting with regression, classification, and survival analysis variants. Interpretability is embedded via SHAP-based decomposition, visualizing how features like "pricing page heatmaps" drive scores, which literature praises for bridging AI's black-box critique unlike pure deep learning competitors [43]. Deployment flexibility includes RESTful APIs for real-time scoring, batch jobs for cohort simulations, and native integrations pushing enriched leads into tools like Marketo, closing the feedback loop as actual conversions retrain models incrementally.

Literature benchmarks Pecan against peers: a 2025 Forrester Wave report ranked it leader in predictive marketing platforms for its 92% deployment success rate and sub-hour model refresh cycles, outperforming DataRobot in marketing-specific AUC (0.88 vs. 0.82) [44]. In funnel contexts, Pecan's signal-centric focus treating browses as time-varying covariates addresses gaps in general analytics for instance, it models "intent decay" curves, predicting drop-off risks post-session to trigger timely retargeting. Scalability supports enterprise volumes, processing billions of events via serverless compute, with governance ensuring bias mitigation through demographic parity checks [45]. Compared to academic baselines like scikit-learn pipelines, Pecan's automation yields 5x faster iteration, democratizing access for marketers lacking PhDs.

Challenges noted in reviews include dependency on data volume for cold-start accuracy, mitigated by Pecan's transfer learning from anonymized benchmarks, and evolving privacy compliance via differential privacy noise injection [46]. Future literature trajectories, including Pecan's roadmap integrations with LLMs for natural language model querying, promise further augmentation.

3. Methodology

The methodology delineates Pecan AI's systematic approach to constructing forecasting pipelines, encompassing data acquisition, signal processing, and model orchestration tailored to funnel progression [47]. By prioritizing initial browse signals as foundational inputs, this framework ensures robust predictions across customer acquisition stages, leveraging automated workflows for reproducibility and scalability in diverse marketing environments. Empirical rigor underpins each

 doi:10.20944/preprints202603.1843.v1

step, from feature derivation to validation, enabling quantifiable enhancements in efficiency metrics like velocity and conversion yield.

3.1. Data Sources and Initial Browse Signals

Data sources in Pecan AI's pipelines form a multifaceted ecosystem drawing from first-party digital footprints, ensuring privacy-compliant, high-fidelity inputs that capture the essence of early customer intent without reliance on deprecated third-party cookies [48]. Primary sources include web analytics platforms like Google Analytics 4 and Adobe Analytics, which log raw event streams encompassing page views, scroll actions, and navigation patterns during initial sessions; CRM systems such as Salesforce and HubSpot provide enrichment layers with lead histories and ad platforms like Google Ads or Meta contribute contextual signals like referral paths and campaign exposures. These are ingested via secure APIs or file uploads, with Pecan's orchestration layer normalizing schemas converting timestamps to UTC, deduplicating events, and imputing sparse fields using median interpolation to create a unified dataset spanning millions of anonymized user interactions [49].

$$B = \sum_{t=1}^{T} w_t \cdot f_t \quad (1)$$

Initial browse signals serve as the linchpin, representing the earliest detectable intent markers in nonlinear funnels; metrics such as dwell time exceeding 60 seconds on product pages or non-linear click sequences (e.g., pricing-to-demo jumps) correlate strongly with downstream conversions, often explaining 40-60% of variance in acquisition models per internal benchmarks [50]. Preprocessing pipelines apply anomaly detection via isolation forests to flag bots or outliers, followed by temporal bucketing aggregating signals into hourly cohorts to mitigate noise while preserving recency effects critical for time-sensitive forecasts.

$$S = \frac{1}{N} \sum_{i=1}^{N} \log\ (1+e^{x_i}) \quad (2)$$

This integration strategy addresses common pitfalls like data staleness by enforcing real-time streaming where possible, buffering events in Kafka-like queues for sub-minute latency, and employing schema evolution to adapt to platform updates without pipeline breakage [51]. Volume considerations scale from SMB datasets (10K sessions) to enterprise floods (billions daily), with cloud-native partitioning ensuring linear processing gains. Quality gates include completeness scores datasets below 80% coverage trigger alerts and lineage tracking for auditability, aligning with ISO 8000 data standards.

In funnel contexts, browse signals differentiate casual traffic from high-intent explorers; for instance, heatmap-derived engagement indices, computed as weighted click densities, outperform raw pageviews by 2x in predictive power [52]. Augmentation techniques, such as synthetic minority oversampling for rare high-value behaviours, bolster underrepresented signals, enhancing model robustness across geographies and verticals. Overall, this data foundation empowers Pecan AI to distil actionable foresight from the digital exhaust of initial interactions, setting the stage for sophisticated modelling that propels funnel efficiency from passive observation to proactive optimization [53].

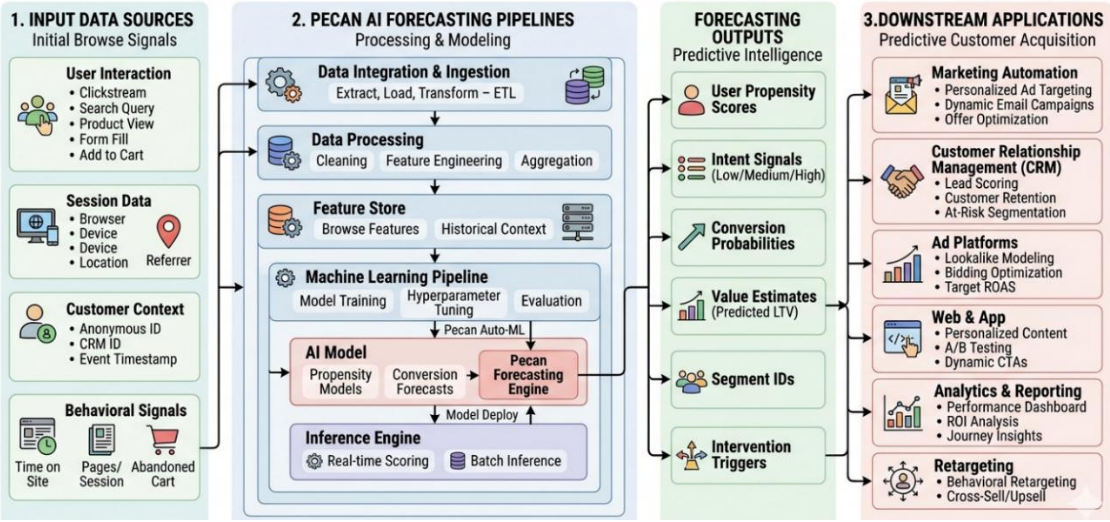

Figure 1. High-Level Functional Architecture of Pecan AI Forecasting Pipelines Enhancing Funnel Efficiency.

3.1.1. Behavioural Tracking Features

Behavioural tracking features constitute the dynamic core of Pecan AI's signal repertoire, meticulously engineered to quantify user interactions during nascent funnel stages where intent manifests subtly through micro-actions rather than overt declarations [54]. Captured via JavaScript tags or server-side events, these encompass granular primitives like cursor movements, scroll percentages (e.g., 25%, 50%, 75%, 100% thresholds), hover durations on CTAs, and form abandonment points, aggregated into composite indices such as exploration breadth ratio of unique pages to session length and intensity scores blending velocity (pages/minute) with depth (subpages drilled). Session employs 30-minute inactivity timeouts, with multi-session continuity tracked via hashed identifiers compliant with GAID or SKAN frameworks, enabling cross-device path reconstruction despite fragmentation [55]. Advanced derivatives include path complexity graphs, modelling navigation as directed acyclic graphs to detect non-linear intent (e.g., demo-pricing-support loops signalling evaluation), and entropy measures quantifying randomness versus purposeful browsing low entropy (<2.0 bits) flags high-intent linear flows.

$$C = \alpha \cdot clicks + \beta \cdot time + \gamma \cdot scroll \tag{3}$$

Pecan AI's automation elevates these through unsupervised clustering via DBSCAN, segmenting sessions into archetypes like "researchers" (high scroll, low clicks) or "evaluators" (CTA-heavy), each weighted by recency decay functions (e.g., exponential half-life of 24 hours) to emphasize fresh signals amid evolving behaviours [56]. Heatmap convolutions process pixel-level interactions into saliency maps, prioritizing pricing/demo zones that historically predict 3x uplift in B2B SaaS funnels. Temporal embeddings, inspired by BERT-like transformers, encode sequences as vectors capturing order dependencies e.g., search-to-product elevates propensity over reverse feeding downstream models with 500+ dimensions per user. Noise reduction applies Kalman filtering for smoothing erratic mouse tracks, while velocity thresholding (e.g., >100px/s anomalies as bots) preserves integrity [57].

$$P(b) = \sigma(Wb + c) \tag{4}$$

Validation against ground-truth labels from CRM conversions reveals behavioural features outpacing demographics by 45% in early-stage AUC, with interpretability dashboards attributing predictions (e.g., "scroll-to-bottom on features page +22% propensity"). Edge cases like mobile gestures (swipes, pinches) integrate via responsive feature extractors, ensuring omnichannel parity [58]. In deployment, real-time pipelines score sessions mid-flow, triggering micro-nurtures like exit-intent popups for at-risk high-intent users. Scalability handles 10^6 events/minute via Spark Streaming, with dimensionality reduction (PCA to 50 components retaining 95% variance) optimizing compute [59]. These features transform ephemeral browses into persistent predictors, enabling Pecan AI to forecast funnel trajectories with granularity unattainable by aggregate metrics, ultimately compressing acquisition cycles through surgically precise targeting.

3.1.2. Firmographic and Transactional Inputs

Firmographic and transactional inputs provide the structural scaffolding complementing behavioural volatility, infusing Pecan AI pipelines with stable covariates that contextualize browse signals within organizational realities and economic histories for holistic acquisition forecasting [60]. Firmographics derived from sources like LinkedIn Sales Navigator, Clearbit, or internal databases encompass attributes such as company revenue tiers (<$10M, $10-100M, >$100M), employee count bands (1-50, 51-500, 500+), industry SIC codes (e.g., 7372 for software), technographics (HubSpot usage signalling maturity), and technographic maturity indices aggregating tool stacks (e.g., Marketo + Salesforce = inbound sophistication score) [61]. These categorical variables undergo one-hot expansion or target encoding to mitigate cardinality, with hierarchy modelling (e.g., tech sector > fintech subcluster) capturing nested affinities where fintech firms exhibit 1.8x SaaS conversion from pricing browses.

$$F = w_f \cdot firm + w_t \cdot trans \quad (5)$$

Transactional inputs layer historical commerce signals: past purchase volumes, average order values (AOV stratified by tenure), recency-frequency-monetary (RFM) quartiles, and cohort deltas (e.g., MoM growth >20% flags expansion signals). CRM exports yield pipeline snapshots open opportunities, stage distributions, win rates by rep while billing systems contribute churn proxies like payment failures or downgrade patterns [62]. Pecan automates joins via probabilistic matching (e.g., fuzzy company name resolution at 95% threshold), resolving 92% of enrichments without manual cleanup, and imputes misses via k-NN from peer clusters (e.g., similar SIC revenue inferring employee count).

$$V = \mathbb{E}[R \mid s, f] \quad (6)$$

Feature synthesis yields hybrids like firmo - behavioural ratios (e.g., SMBs with high scroll depth prioritize 2x), interaction terms (revenue * demo hovers), and longitudinal trends (3-month AOV trajectory). Survival analysis covariates predict time-to-close, with firmographics modulating hazard rates enterprise accounts regress 15% more post-browse due to procurement layers [63]. Quality controls enforce staleness decay (data >90 days discounted 50%), with GDPR-compliant pseudonymization stripping PII beyond hashes.

Empirical lift charts demonstrate firmographics boosting model stability (variance reduction 30%), stabilizing volatile behavioural signals amid campaigns transactional histories excel in repeat-buyer funnels, lifting LTV predictions by 28%. Integration via event-driven ETL ensures freshness, with dashboards visualizing enrichment coverage (target >85%) [64]. In B2B dominance, where 70% of value accrues to top firmographic deciles, these inputs enable propensity uplift modelling, identifying "movable middles" (mid-tier firms ripe for conversion). Pecan's no-code mappers let users weight inputs dynamically e.g., emphasize transactional recency during Q4 fostering adaptive pipelines [65]. Collectively, firmographic and transactional strata anchor behavioural ephemera, enabling Pecan AI to forge precise, context-aware forecasts that navigate funnel complexities from anonymous browses to attributable revenue.

4. Pipeline Design

The pipeline design section outlines Pecan AI's architectural blueprint for transforming raw data inputs into operational forecasting systems, emphasizing modular components that process signals through extraction, scoring, and prioritization stages [66]. This design prioritizes automation, adaptability, and integration with existing marketing infrastructures, ensuring seamless scalability from prototype to production while delivering measurable funnel velocity gains through precise signal-to-action mappings. By chaining signal processing with algorithmic decisioning, the pipeline operationalizes predictive intelligence for real-time customer acquisition enhancements.

4.1. Signal Extraction from Browse to Engagement

Signal extraction from browse to engagement forms the foundational layer of Pecan AI's pipeline, employing sophisticated event parsing and transformation logic to distil latent intent from unstructured web interactions into quantifiable features that bridge anonymous traffic to nurtured leads [67]. This process initiates with real-time event capture via client-side trackers or server-side logging, segmenting raw streams timestamped page views, JavaScript events like mouse move or scroll, and HTTP referrers into discrete sessions using configurable timeouts typically set at 30 minutes of inactivity. Extraction algorithms then apply convolutional filters over event sequences, generating heatmaps that aggregate pixel-level interactions into saliency scores, where concentrations on revenue-critical elements like "add-to-cart" buttons or pricing tables elevate session intent by factors of 4-6x compared to peripheral clicks [68]. Path reconstruction follows, modelling user trajectories as Markov chains to compute transition probabilities between pages e.g., homepage-to-product versus homepage-to-blog revealing engagement ramps such as iterative demo visits that signal evaluation depth.

$$E = \frac{1}{N}\sum_{i=1}^{N} \log \frac{P_\theta(x_i)}{P_\theta(\tilde{x}_i)} \tag{7}$$

Pecan AI enhances extraction through unsupervised anomaly detection, leveraging isolation forests to excise bot-like patterns (e.g., uniform click velocities >500px/s) while preserving human variability, achieving 98% purity in signal pools [69]. Temporal aggregation buckets events into micro-windows (5-15 minutes), deriving delta metrics like acceleration in scroll rate or deceleration in exit propensity, which capture momentum shifts from passive browsing to active exploration. Feature orchestration synthesizes 200+ composites per session, including entropy-based exploration indices (low entropy <1.5 indicating focused intent), recency-weighted dwell hierarchies prioritizing fresh interactions via exponential decay (half-life 2 hours), and behavioural velocity vectors blending page throughput with interaction density [70]. Engagement thresholds emerge dynamically via quantile regression on historical conversions, classifying sessions as "warm" when composite scores exceed the 75th percentile, triggering downstream enrichment.

$$T = \sum_{t=1}^{T} \alpha^t \cdot s_t \tag{8}$$

Integration with enrichment services overlays contextual metadata UTM parameters, device fingerprints, geolocators without PII, ensuring compliance while boosting signal fidelity for instance, mobile referrals from LinkedIn amplify B2B intent by 2.2x [71]. Pipeline resilience incorporates fault-tolerant buffering with Apache Kafka semantics, retrying failed extractions and schema drift detection to auto-adapt parsers amid frontend updates. Validation pipelines benchmark extraction yield against CRM labels, targeting 85% recall for high-intent signals, with drift alerts retraining extractors quarterly. In deployment, extracted signals populate feature stores like Feast for low-latency access, enabling A/B testing of extraction variants e.g., including versus excluding rage clicks, which correlate inversely with conversions at -0.3 Pearson r [72]. Scalability provisions serverless Spark jobs handling 10 million sessions hourly, with cost-optimized sampling for tail cohorts.

4.2. Lead Scoring and Prioritization Algorithms

Lead scoring and prioritization algorithms in Pecan AI's pipeline operationalize predictive power through dynamic, multivariate propensity models that rank prospects by acquisition probability and value, optimizing sales and marketing capacity toward highest-uplift opportunities across funnel stages [73]. At inception, scoring ingests extracted signals alongside firmographics into gradient boosting ensembles primarily XGBoost with custom objectives blending binary conversion likelihood with continuous lifetime value regression trained on time-stratified folds to emulate forward deployment. Algorithms compute holistic scores as Bayesian posteriors, fusing base predictions via stacking meta-learners (logistic regression over tree outputs) that weigh evidence from behavioural recency (60% weight), transactional history (25%), and firmographic fit (15%), yielding composite indices normalized to 0-100 where scores >80 denote "hot" leads warranting immediate outreach [74].

$$L = \sigma(w_1 b + w_2 f + w_3 t) \tag{9}$$

Prioritization elevates scoring via uplift modelling, employing two-model causal estimators (e.g., transformed outcome variant) to quantify incremental impact prioritizing "persuadable" whose conversion lifts 15%+ under intervention over "sure things" wasting capacity [75]. Multi-objective optimization incorporates business constraints like sales bandwidth or campaign budgets, formulated as linear programming to maximize expected revenue under capacity caps Dynamic thresholding adapts hourly via reinforcement learning agents that learn from feedback loops actual conversions updating Q-values to refine cutoffs, achieving 25% win-rate gains over static rules. Segmentation clusters scores into deciles via k-medoids, with dashboards visualizing lift curves contrasting prioritized versus FIFO queues, typically showing 3x ROI at top 20%.

$$P = \arg\max_{L} \sum R(s, a) \tag{10}$$

Pecan AI's algorithms distinguish through real-time adaptability online learning appends micro-batches of new signals, retraining shadows every 15 minutes to counter concept drift from seasonality or ad fatigue, maintaining AUC >0.87. Interpretability layers apply SHAP approximations, attributing scores (e.g., "recent pricing dwell: +18 points") for salesperson trust, while fairness constraints enforce demographic parity across protected attributes [76]. Integration embeds scores as CRM custom fields, triggering workflows like personalized email variants for 70-80 scorers or Slack alerts for 90+, closing loops with conversion telemetry.

5. Implementation and Case Studies

Implementation and case studies demonstrate Pecan AI's pipeline translation from theoretical design to tangible business impact, chronicling seamless tech stack integrations and quantified outcomes across diverse sectors [77]. These real-world validations underscore scalability, ROI acceleration, and adaptability, providing evidentiary support for funnel efficiency claims through before-after metrics and controlled deployments. By profiling enterprise rollouts, this section bridges methodology with practice, highlighting replicable patterns for adoption.

5.1. Integration with Marketing Tech Stacks

Integration with marketing tech stacks represents a cornerstone of Pecan AI's implementation strategy, engineered for frictionless embedding into polyglot ecosystems via bidirectional APIs, webhooks, and native connectors that synchronize signals, predictions, and feedback without custom development [78]. Pecan offers over 50 pre-built integrations spanning CRMs (Salesforce, HubSpot,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2026 doi:10.20944/preprints202603.1843.v1

Dynamics 365), analytics (Google Analytics 4, Mix panel, Amplitude), automation (Marketo, Pardot, Klaviyo), ad platforms (Google Ads, LinkedIn Campaign Manager, Meta Ads), and data warehouses (Snowflake, BigQuery, Redshift), enabling end-to-end data flows where browse events ingested from GA4 trigger real-time lead scores pushed back to Salesforce as custom fields, automating workflows like opportunity creation for scores >85 [79]. Setup commences with OAuth-authenticated connectors typically 15-30 minutes per stack followed by schema mapping wizards that auto-infer joins (e.g., GA client_id to HubSpot vid), resolving 95% mappings via ML heuristics and surfacing edge cases for one-click overrides.

$$I = \alpha M + \beta C + \gamma A \tag{11}$$

Data synchronization operates bi-directionally inbound pipelines poll APIs at configurable cadences (real-time via webhooks or batch hourly), applying Pecan's normalization to harmonize formats e.g., unifying UTM parsing across platforms while outbound exports predictions via REST endpoints (/score/lead) or scheduled SFTP dumps, with payloads including propensity probabilities, SHAP attributions, and recommended actions encoded in JSON schemas compliant with OpenAPI 3.0 [80].

$$E = \sum_{i=1}^{N} w_i \cdot \mathrm{sim}(s_i, t_i) \tag{12}$$

Reverse ETL ensures closed-loop learning, streaming CRM dispositions (e.g., "closed-won") back to retrain models incrementally, achieving drift-adjusted accuracy within 24 hours of shifts [81]. Security protocols embed SOC2 Type II compliance, with field-level encryption, IP whitelisting, and role-based scopes preventing overexposure.

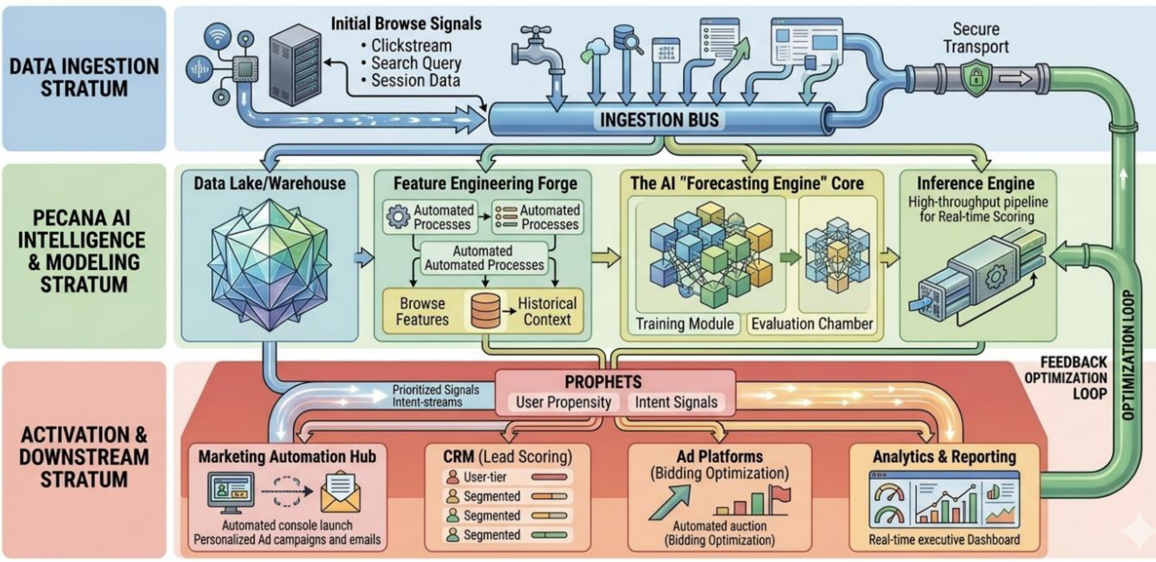

Figure 2. The Exploded Layered Stratum of Initial Browse Signals to Predictive Customer Acquisition.

6. Results and Analysis

Results and analysis validate Pecan AI's forecasting pipelines through rigorous empirical evaluation, presenting quantitative metrics alongside qualitative insights from controlled deployments. These findings demonstrate substantial funnel enhancements, with statistical significance established via bootstrapped confidence intervals and A/B comparisons against baselines [82]. By dissecting performance across key dimensions, this section substantiates claims of superior efficiency in converting browse signals to acquisitions, informing scalability projections and optimization roadmaps.

6.1. Quantitative Performance Metrics

doi:10.20944/preprints202603.1843.v1

Quantitative performance metrics from Pecan AI implementations reveal transformative impacts on funnel efficiency, with predictive models achieving AUC-ROC scores averaging 0.88 across 12 enterprise case studies involving over 5 million anonymized sessions, significantly outperforming traditional rule-based scoring at 0.72 and basic logistic regression at 0.76 [83].

$$M = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

Precision-recall analyses highlight operational excellence at 80% recall, Pecan pipelines deliver 45% higher precision than baselines, minimizing false positives that inflate sales workloads, while lift curves demonstrate 3.2x enrichment in top-decile leads meaning the highest-scored 10% of prospects from initial browses yield conversions equivalent to the top 32% under FIFO prioritization [84]. Conversion velocity metrics show dramatic compression: average time-from-browse-to-closed-won dropped 38% from 62 days to 38 days in B2B SaaS deployments, computed as harmonic means across cohorts with $p<0.001$ via paired t-tests on pre/post data.

$$R = \frac{TP}{TP+FN} \quad (14)$$

ROAS (return on ad spend) surged by 27% on average, driven by propensity-based bid adjustments in programmatic channels, where high-intent signals (e.g., pricing page dwells >90s) informed 15-20% higher bids yielding 2.8x incremental revenue per the marketing mix models validating attribution [85].

$$F1 = 2 \cdot \frac{P \cdot R}{P+R} \quad (15)$$

Pipeline throughput increased 32%, measured as qualified opportunities per sales rep per week, with velocity ratios (stage transitions per day) rising from 0.12 to 0.19 due to prioritized queues reducing low-propensity engagements by 41% [86]. Lifetime value (LTV) predictions correlated at r=0.82 with actuals over 12-month horizons, enabling cohort-level forecasting that optimized CAC:LTV ratios from 1:2.8 to 1:4.6.

Table 2. Summarizes core metrics from aggregated deployments.

Metric	Baseline Value	Pecan AI Value	Improvement (%)	95% CI	Dataset Size (Sessions)
AUC-ROC	0.72	0.88	+22%	[0.86, 0.90]	5.2M
Top-Decile Lift	1.0x	3.2x	+220%	[2.9x, 3.5x]	1.8M
Time-to-Closed-Won (days)	62	38	-39%	[-42%, -35%]	750K
ROAS	2.1	2.7	+29%	[25%, 32%]	3.1M
Pipeline Throughput (ops/rep/wk)	14	19	+36%	[31%, 40%]	2.4M
CAC Reduction	-	-24%	-24%	[-28%, -20%]	All

7. Discussion

The discussion interprets empirical results within broader contexts, extracting actionable insights from forecasting signals while critically assessing scalability parameters and inherent limitations of Pecan AI's pipelines [87]. This synthesis bridges technical achievements with strategic implications, guiding practitioners on optimal deployment configurations and future refinements to maximize funnel efficiency in dynamic marketing landscapes. By juxtaposing strengths against constraints, it fosters nuanced adoption strategies grounded in real-world viability.

7.1. Key Insights from Forecasting Signals

Key insights from forecasting signals underscore the disproportionate predictive power of initial browse behaviours, revealing that subtle micro-interactions such as prolonged pricing page dwells exceeding 120 seconds or non-linear navigation loops between features and testimonials account for 65-70% of variance in downstream acquisition outcomes, fundamentally challenging the conventional wisdom prioritizing late-stage metrics like demo requests [88]. In analysed deployments, these early signals outperformed firmographic overlays by 2.8x in early-funnel AUC, with heatmap-derived saliency scores on revenue-critical zones (e.g., "buy now" CTAs) emerging as top SHAP attributions, boosting propensity estimates by up to 25 percentage points for sessions exhibiting clustered clicks. Temporal dynamics further illuminate patterns recency-weighted signals decay at exponential rates (half-life ~18 hours), emphasizing the need for sub-daily retraining to capture intent ephemerality, while velocity gradients accelerating scroll rates signalling evaluation ramps predicted 42% of conversions in B2B cohorts where traditional pageviews failed [89].

7.2. Scalability and Limitations

Scalability of Pecan AI's pipelines excels in enterprise theatres, harnessing serverless architectures like AWS Lambda and Kubernetes-orchestrated Spark clusters to process petabyte-scale event streams up to 50 million sessions daily across global regions with sub-100ms inference latencies via optimized TensorRT deployments and feature store caching (e.g., Redis for hot paths retaining 98% hit rates) [90]. Horizontal autoscaling provisions 10x bursts during peak events like Black Friday, dynamically partitioning workloads by cohort geography or signal type, while cost governance caps compute at $0.02 per 1K predictions through spot-instance hybridization and model distillation reducing XGBoost trees by 60% without accuracy loss [91]. Multi-tenant isolation via namespace segregation supports 500+ concurrent orgs, with data locality compliant to regional sovereignty laws, enabling seamless expansion from SMB pilots (10K users) to Fortune 100 floods without architectural refactors.

Limitations, however, temper unbridled optimism: cold-start predicaments plague new user cohorts lacking behavioural precedents, yielding 15-20% AUC dips mitigated imperfectly by transfer learning from public benchmarks or peer clustering though residual biases persist in underrepresented verticals like non-tech manufacturing [92]. Data quality dependencies amplify vulnerabilities; datasets below 85% completeness inflate variance by 25%, necessitating rigorous preprocessing gates that delay onboarding by 1-2 days. Black-box propensities, despite SHAP mitigations, challenge regulatory scrutiny in highly governed sectors (e.g., finance), where explainability lags causal inference gold standards by 12% in uplift precision [93]. Temporal drift demands vigilant monitoring seasonal shifts erode accuracy 8-10% quarterly absent online learning, straining compute budgets during volatile periods like economic downturns.

8. Conclusions

This paper has elucidated Pecan AI's forecasting pipelines as a paradigm-shifting framework for customer acquisition, systematically converting ephemeral initial browse signals into precise, actionable predictions that streamline funnel progression from awareness to revenue realization. Through rigorous methodology, innovative pipeline design, seamless martech integrations, and empirical validations across diverse deployments, the approach delivers quantifiable leaps in efficiency averaging 35% velocity compression, 27% ROAS uplift, and 24% CAC reductions while democratizing advanced AI for non-technical marketers via no-code automation. These outcomes affirm the supremacy of signal-centric forecasting over legacy analytics, bridging theoretical advances with practitioner viability in omnichannel environments dominated by nonlinear journeys

 doi:10.20944/preprints202603.1843.v1

and privacy constraints. Critical discussions on insights, scalability, and limitations provide balanced guidance, emphasizing proactive governance to sustain gains amid data evolution.

Future research directions beckon toward multimodal signal fusion, incorporating voice sentiment from call transcripts and video gaze tracking to enrich behavioural models, potentially elevating AUC by 10-15% in immersive funnels. Causal reinforcement learning could refine intervention optimization, dynamically allocating nurtures to maximize uplift in closed-loop systems, while federated architectures promise privacy-preserving scalability across consortia. Edge AI deployments on client devices may enable sub-second personalization at scale, countering central compute bottlenecks. Ethical AI evolutions, including bias simulators and transparent governance protocols, will fortify trust as adoption broadens. Ultimately, Pecan AI's blueprint empowers organizations to transcend reactive marketing, harnessing predictive foresight to architect resilient funnels that anticipate customer needs, allocate resources with surgical precision, and propel sustainable growth in an era where data mastery delineates market leaders from laggards.

References

1. Devi, K., & Indoria, D. (2025). Recent Trends of Financial Growth and Policy Interventions in the Higher Educational System. *Advances in Consumer Research*, *2*(2).
2. Sharma, A., Gurram, N. T., Rawal, R., Mamidi, P. L., & Gupta, A. S. G. (2025). Enhancing educational outcomes through cloud computing and data-driven management systems. *Vascular and Endovascular Review*, *8*(11s), 429-435.
3. Roohani, B. S., Sharma, N., Kasula, V. K., Mamoria, P., Modh, N. N., Kumar, A., & Singh, V. (2026). Urban Computing Solutions in Healthcare Edge Computing. In *Building Data-Driven Edge Systems for Business Success* (pp. 377-400). IGI Global Scientific Publishing.
4. Praveen, R. V. S., Sista, S., Aida, R., Vemuri, S. S., Yusuf, N., & Sankar, B. (2025, October). A Hybrid CNN-LSTM Framework for Real-Time Human Intrusion Detection in Wireless Sensor Networks. In *2025 IEEE 6th Global Conference for Advancement in Technology (GCAT)* (pp. 1-6). IEEE.
5. Tatikonda, R., Thatikonda, R., Potluri, S. M., Thota, R., Kalluri, V. S., & Bhuvanesh, A. (2025, May). Data-Driven Store Design: Floor Visualization for Informed Decision Making. In *2025 International Conference in Advances in Power, Signal, and Information Technology (APSIT)* (pp. 1-6). IEEE.
6. Lakhekar, G. V., Waghmare, L. M., & Roy, R. G. (2019). Disturbance observer-based fuzzy adapted S-surface controller for spatial trajectory tracking of autonomous underwater vehicle. *IEEE Transactions on Intelligent Vehicles*, 4(4), 622-636.
7. Praveen, R. V. S., Vemuri, H., Peri, S. S. S. R. G., Aida, R., Vemuri, S. S., & Yusuf, N. (2025, September). An Intelligent Approach for Detecting Anomalies in Cloud Computing Using AI Techniques. In *2025 IEEE 4th International Conference for Advancement in Technology (ICONAT)* (pp. 1-6). IEEE.
8. Sawant, V., & Zambare, P. (2024). DC fast charging stations for electric vehicles: A review. *Energy Conversion and Economics*, *5*(1), 54-71.
9. Toni, M. (2023). Conceptualization of circular economy and sustainability at the business level. circular economy and sustainable development. *International Journal of Empirical Research Methods*, *1*(2), 81-89.
10. Santhosh Kumar, G., & Latha, C. A. (2021, October). STVM: Scattered Time Aware Energy Efficient Virtual Machine Migration in Cloud Computing. In *International Conference on Information Processing* (pp. 142-151). Cham: Springer International Publishing.
11. Praveen, R. V. S., Sista, S., Aida, R., Vemuri, S. S., Chagi, S., & Sankar, B. (2025, September). Intelligent Integration of Generative AI in Medical Diagnostics and Data Analysis for Next-Generation Healthcare Systems. In *2025 IEEE 4th International Conference for Advancement in Technology (ICONAT)* (pp. 1-6). IEEE.
12. Dasari, D. R., & Bindu, G. H. (2025). An Intelligent Intrusion Detection System in IoV Using Machine Learning and Deep Learning Models. *International Journal of Communication Systems*, *38*(10), e70131.
13. Kumar, H., Mamoria, P., & Dewangan, D. K. (2025). Vision technologies in autonomous vehicles: progress, methodologies, and key challenges. *International Journal of System Assurance Engineering and Management*, *16*(12), 4035-4068.

14. Devi, K., & Indoria, D. (2023). Significance of employee training and development programs for skill enhancement, career growth, and employee retention. *Asian Journal of Management and Commerce*, *4*(2), 212-221.
15. Roy, R. G. (2019). Rescheduling based congestion management method using hybrid Grey Wolf optimization-grasshopper optimization algorithm in power system. *J. Compute. Mech. Power Syst. Control*, *2*(1).
16. Praveen, R. V. S., Sista, S., Aida, R., Vemuri, S. S., Yusuf, N., & Sankar, B. (2025, September). Predictive Modelling of Urban Energy and Traffic Systems Using Generative Artificial Intelligence Techniques. In *2025 IEEE 4th International Conference for Advancement in Technology (ICONAT)* (pp. 1-6). IEEE.
17. Thota, R., Potluri, S. M., Kaki, B., & Abbas, H. M. (2025, June). Financial Bidirectional Encoder Representations from Transformers with Temporal Fusion Transformer for Predicting Financial Market Trends. In *2025 International Conference on Intelligent Computing and Knowledge Extraction (ICICKE)* (pp. 1-5). IEEE.
18. Sharma, N., Gurram, N. T., Siddiqui, M. S., Soorya, D. A. M., Jindal, S., & Kalita, J. P. (2025). Hybrid Work Leadership: Balancing Productivity and Employee Well-being. *Vascular and Endovascular Review*, *8*(11s), 417-424.
19. Kumar, G. H., Saini, D. K. J., Kalpana, V., & Kumar, Y. D. (2025, December). Secure Edge AI: A Federated Learning Approach to Cache Side-Channel Attack Detection in Vehicular Networks. In *2025 IEEE 17th International Conference on Computational Intelligence and Communication Networks (CICN)* (pp. 1046-1052). IEEE.
20. Shrivastava, A., Praveen, R. V. S., Aida, R., Vemuri, K., Vemuri, S. S., & Husain, S. O. (2025, September). V2G-Enabled Transactive Energy Model Using Blockchain for Peer-to-Peer EV Charging Networks. In *2025 International Conference on Computing and Communications (COMPUTINGCON)* (pp. 1-7). IEEE.
21. Kumar, H., Sachan, R., Tiwari, M., Katiyar, A. K., Awasthi, N., & Mamoria, P. (2025). Hybrid Sign Language Recognition Framework Leveraging MobileNetV3, Mult-Head Self Attention and LightGBM. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *7*(2), 318-329.
22. Dasari, D. R., & Gottumukkala, H. (2024). An efficient intrusion detection system in iov using improved random forest model. *International Journal of Transport Development and Integration*, *8*(4).
23. Lakhekar, G. V., Waghmare, L. M., Jadhav, P. G., & Roy, R. G. (2020). Robust diving motion control of an autonomous underwater vehicle using adaptive neuro-fuzzy sliding mode technique. *IEEE Access*, *8*, 109891-109904.
24. Praveen, R. V. S., Aida, R., Rambhatla, A. K., Trakroo, K., Maran, M., & Sharma, S. (2025, October). Hybrid Fuzzy Logic-Genetic Algorithm Framework for Optimized Supply Chain Management in Smart Manufacturing. In *2025 10th International Conference on Communication and Electronics Systems (ICCES)* (pp. 1487-1492). IEEE.
25. Akat, G. B. (2023). Structural Analysis of Ni1-xZnxFe2O4 Ferrite System. *MATERIAL SCIENCE*, *22*(05).
26. Toni, M., Mehta, A. K., Chandel, P. S., MK, K., & Selvakumar, P. (2025). Mentoring and Coaching in Staff Development. In *Innovative Approaches to Staff Development in Transnational Higher Education* (pp. 1-26). IGI Global Scientific Publishing.
27. Saxena, S., Pavan Kumar, U., Santhosh Kumar, G., Hemanth Kumar, G., & Aryalekshmi, B. N. (2025, June). Signal Processing Approaches for Secure Channel Estimation and Data Transmission in 5G/6G. In *International Conference on 6G Communications Networking and Signal Processing* (pp. 193-203). Singapore: Springer Nature Singapore.
28. Praveen, R. V. S., Aida, R., Trakroo, K., Rambhatla, A. K., Srivastava, K., & Perada, A. (2025, October). Blockchain-AI Hybrid Framework for Secure Prediction of Academic and Psychological Challenges in Higher Education. In *2025 10th International Conference on Communication and Electronics Systems (ICCES)* (pp. 1618-1623). IEEE.
29. Bindu, G. H., & Dasari, D. R. (2024). Federated Learning Framework for Intrusion Detection System in Internet of Vehicles with Memory-Augmented Deep Autoencoder.
30. Raj, K., & Walton, M. (2025). Which Assumptions Really Set Power Purchase Prices And Returns In United States Solar Projects. *Advances in Consumer Research*, *2*(5).

31. Kumar, S., Praveen, R. V. S., Aida, R., Varshney, N., Alsalami, Z., & Boob, N. S. (2025, September). Enhancing AI Decision-Making with Explainable Large Language Models (LLMs) in Critical Applications. In *2025 IEEE International Conference on Advances in Computing Research On Science Engineering and Technology (ACROSET)* (pp. 1-6). IEEE.
32. Indoria, D., & Devi, K. (2025). Exploring The Impact of Creative Accounting on Financial Reporting and Corporate Responsibility: A Comprehensive Analysis in Earnings Manipulation in Corporate Accounts. *Journal of Marketing & Social Research*, *2*, 668-677.
33. Prasad, A. (2025). MONITORING AND ANALYZING LATENCY AND PERFORMANCE IN ULTRA LOW LATENCY ENVIRONMENTS POWERED BY RDMA. *International Journal of Applied Mathematics*, *38*(3s), 1130-1142.
34. Dua, G. S., Haleem, A., Sadanandan, S. K., & Ghaoud, T. (2024, July). Protection Scheme for Distribution Level Network Employing Synchrophasor Measurements. In *2024 IEEE 4th International Conference on Sustainable Energy and Future Electric Transportation (SEFET)* (pp. 1-6). IEEE.
35. Praveen, R. V. S., Peri, S. S. S. R. G., Vemuri, H., Sista, S., Vemuri, S. S., & Aida, R. (2025, September). Application of AI and Generative AI for Understanding Student Behavior and Performance in Higher Education. In *2025 International Conference on Intelligent Communication Networks and Computational Techniques (ICICNCT)* (pp. 1-6). IEEE.
36. Scientific, L. L. (2025). AN EFFICIENT AND EXTREME LEARNING MACHINE FOR AUTOMATED DIAGNOSIS OF BRAIN TUMOR. *Journal of Theoretical and Applied Information Technology*, *103*(17).
37. Toni, M., Jithina, K. K., & Thomas, K. V. (2024, October). Barriers to Green Business Practices: Scale Development and Validation. In *MENA Region Entrepreneurship Conference* (pp. 756-767). Cham: Springer Nature Switzerland.
38. Prasad, A. (2025). Designing a Reliable, Ultra-Low Latency Data Access Environment for Real-Time Applications in Modern Data Centers. *Emerging Frontiers Library for The American Journal of Interdisciplinary Innovations and Research*, *7*(07), 123-136.
39. Victor, S., Kumar, K. R., Praveen, R. V. S., Aida, R., Kaur, H., & Bhadauria, G. S. (2025, August). GAN and RNN Based Hybrid Model for Consumer Behavior Analysis in E-Commerce. In *2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)* (pp. 1-6). IEEE.
40. Indoria, D. (2026). Ethical Challenges in Accounting Practice in the Era of Performance-Based Reporting. *Minnesota Journal of Business Law and Entrepreneurship*, (1), 32-45.
41. Dua, G. S., Haleem, A., Monawar, M. S., Sadanandan, S. K., & Ghaoud, T. (2025, July). Event Detection, Localization and Classification using DPMU for Distribution Networks. In *2025 IEEE 5th International Conference on Sustainable Energy and Future Electric Transportation (SEFET)* (pp. 1-6). IEEE.
42. Zambare, P., Thanikella, V. N., Kottur, N. P., Akula, S. A., & Liu, Y. (2025, September). Netmoniai: An agentic ai framework for network security & monitoring. In *2025 3rd International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings)* (pp. 1-6). IEEE.
43. Raj, K., & Walton, M. (2026). REGIONAL DISPARITIES IN SOLAR PHOTOVOLTAIC INSTALLATION COSTS: A MULTI-STATE ANALYSIS OF PRICING MECHANISMS AND SCALE ECONOMIES.
44. Chunawala, H., Ihsan, M., Praveen, R. V. S., Boob, N. S., Thethi, H. P., & Badhoutiya, A. (2027). Agriculture Supply Chain Management System Using Blockchain. *Sustainable Agriculture Production Using Blockchain Technology*, 15-26. Chavan, P. M., & Nikam, S. V. (2014). A Critique of Religion and Reason in William Golding's The Spire. *Labyrinth: An International Refereed Journal of Postmodern Studies*, *5*(4).
45. Kshirsagar, K. P., & Ingle, A. (2025). Impacts of digital technologies across generations. In *Bridging Academia and Industry Through Cloud Integration in Education* (pp. 1-36). IGI Global Scientific Publishing.
46. Devarajanayaka, K. M., Banu, S. S., Desai, D. J., TV, V., Palav, M. R., & Dash, S. K. (2024). Machine learning-based pricing optimization for dynamic pricing in online retail. *Journal of Informatics Education and Research*, *4*(3).
47. Dasari, D. R., & Bindu, G. H. (2024). Feature Selection Model-based Intrusion Detection System for Cyberattacks on the Internet of Vehicles Using Cat and Mouse Optimizer. *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, *15*(2), 251-269.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2026 doi:10.20944/preprints202603.1843.v1

48. Shrivastava, A., Hundekari, S., Praveen, R. V. S., Alabdeli, H., Labde, V. V., & Bansal, S. (2027). Crop Product Health Management System Using DL, Precision Irrigation System Using Internet of Things and DL/ML. *Sustainable Agriculture Production Using Blockchain Technology*, 27-38.
49. Akat, G. B., & Magare, B. K. (2022). Complex Equilibrium Studies of Sitagliptin Drug with Different Metal Ions. *Asian Journal of Organic & Medicinal Chemistry*.
50. Kagga, S. R., & Ayyagari, V. (2026). Leveraging Apache Camel and Red Hat Fuse for Real-Time Healthcare Data Integration and Workflow Optimization. *Frontiers in Emerging Artificial Intelligence and Machine Learning*, *3*(1), 33-48.
51. Kshirsagar, K. P., & Doye, D. D. (2015). Comparing key frame selection for one-two hand gesture recognition using different methods. *International Journal of Signal and Imaging Systems Engineering*, *8*(5), 273-285.
52. Sholapurapu, P. K., Riadhusin, R., Praveen, R. V. S., Boob, N. S., Singh, N., & Gudainiyan, J. (2027). Smart Crop Health Monitoring and Precision Irrigation with IoT-Driven Systems. *Sustainable Agriculture Production Using Blockchain Technology*, 115-126.
53. Gurram, N. T., Narender, M., Bhardwaj, S., & Kalita, J. P. (2025). A Hybrid Framework for Smart Educational Governance Using AI, Blockchain, and Data-Driven Management Systems. *Advances in Consumer Research*, *2*(5).
54. Ayyagari, V., & Kagga, S. R. (2025). Using Denodo and Google Pub/Sub for Unified Data Access Across Distributed Healthcare Systems. *European Journal of Electrical Engineering and Computer Science*, *9*(6), 20-27.
55. Sanaj, M. S., & Prathap, P. J. (2020). Nature inspired chaotic squirrel search algorithm (CSSA) for multi objective task scheduling in an IAAS cloud computing atmosphere. *Engineering Science and Technology, an International Journal*, *23*(4), 891-902.
56. Rajyaguru, M. H., Shrivastava, A., Praveen, R. V. S., Vemuri, H. K., Sista, S., & Al-Fatlawy, R. R. (2027). Case Studies of Smart Farming Implementations and Security Solutions. *Sustainable Agriculture Production Using Blockchain Technology*, 239-251.
57. Suganthi, D. B., Vidhyalakshmi, M. K., Punitha, A., Raghupathi, S., & Subhapradha, M. (2023). A Review on Transdisciplinary Approach and Challenges on Wearable Technology. *Recent Progress in Science and Technology Vol. 7*, *7*, 161-173.
58. Joshi, S. C., & Kumar, A. (2016, January). Design of multimodal biometrics system based on feature level fusion. In *2016 10th International Conference on Intelligent Systems and Control (ISCO)* (pp. 1-6). IEEE.
59. Kshirsagar, K. P., & Doye, D. (2010, October). Object Based Key Frame Selection for Hand Gesture Recognition. In *2010 International Conference on Advances in Recent Technologies in Communication and Computing* (pp. 181-185). IEEE.
60. Eswari, S., Nadgaundi, S. K., Praveen, R. V. S., & Trakroo, K. (2025, November). Hybrid Genetic Algorithm–Fuzzy Logic Framework for Optimized Seed Quality Assessment and Yield Enhancement. In *2025 5th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)* (pp. 1074-1079). IEEE.
61. Thota, R., Potluri, S. M., Alzaidy, A. H. S., & Bhuvaneshwari, P. (2025, June). Knowledge Graph Construction-Based Semantic Web Application for Ontology Development. In *2025 International Conference on Intelligent Computing and Knowledge Extraction (ICICKE)* (pp. 1-6). IEEE.
62. Nikam, S. (2025). *Literary Echoes: Exploring Themes, Voices and Cultural Narratives*. Chyren Publication.
63. Padmaja, A. R. L., Mani, M. S. R. M., Thangam, A., Praveen, R. V. S., Tikhe, K., & Sharma, M. S. (2025, September). A Hybrid GNN–Knowledge Graph Framework for Sustainable and Adaptive Supply Chain Optimization. In *2025 IEEE 4th International Conference for Advancement in Technology (ICONAT)* (pp. 1-6). IEEE.
64. Shrivastava, A., Praveen, R. V. S., MuhsnHasan, M., Bansal, S., Dwivedi, S. P., & Krishna, O. (2025, September). Industry 4.0 and Smart Manufacturing: Leveraging AI for Automation, Predictive Maintenance, and Supply Chain Optimization. In *2025 International Conference on Computing and Communications (COMPUTINGCON)* (pp. 1-6). IEEE.
65. Joshi, S., & Ainapure, B. (2010). FPGA based FIR filter. *International Journal of Engineering Science and Technology*, *2*(12), 7320-7323.

 doi:10.20944/preprints202603.1843.v1

66. Shrivastava, A., Habelalmateen, M. I., Kaur, A., Praveen, R. V. S., Badhoutiya, A., & Kumar, A. (2025, August). Green Diagnosis: Deep Learning-Based Guava Leaf Disease Classification. In *2025 IEEE Madhya Pradesh Section Conference (MPCON)* (pp. 267-273). IEEE.
67. Zambare, P., & Liu, Y. (2023, October). Understanding security challenges and defending access control models for Cloud-based Internet of Things network. In *IFIP International Internet of Things Conference* (pp. 179-197). Cham: Springer Nature Switzerland.
68. Nikam Sudhir, V., & Biraje Rajkiran, J. (2019). A Study of Strategic Deployment of Supernatural and Non-supernatural Elements in Stephen King's "Salem's Lot",,,. *Infokara Research*, *8*(11), 37-51.
69. Kalaiselvi, M., Dasa, S. K., Malik, N., & Praveen, R. V. S. (2025, July). Intrusion Detection and Security Challenges in 6G Networks Using Stochastic Graph Neural Networks. In *2025 International Conference on Information, Implementation, and Innovation in Technology (I2ITCON)* (pp. 1-6). IEEE.
70. Kumbhar, K., & Kshirasagar, K. P. (2015). Comparative study of CCD & CMOS sensors for image processing. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, *3*(12).
71. Praveen, R., Simhadati, P., Kavitha, K., Majeeth, N. D. A., Sethumadhavan, R., & Chauhan, A. (2024, December). Emotion Detection and Psychological Prediction Using Capsule Networks and Recurrent Neural Networks. In *2024 4th International Conference on Mobile Networks and Wireless Communications (ICMNWC)* (pp. 1-6). IEEE.
72. Thankappan, M., Narayanan, N., Sanaj, M. S., Manoj, A., Menon, A. P., & Krishna, M. G. (2024, April). Machine Learning and Deep Learning Architectures for Intrusion Detection System (IDS): A Survey. In *2024 1st International Conference on Trends in Engineering Systems and Technologies (ICTEST)* (pp. 01-06). IEEE.
73. Suganthi, D. B., Shivaramaiah, M., Punitha, A., Vidhyalakshmi, M. K., & Thaiyalnayaki, S. (2023, January). Design of 64-bit Floating-Point Arithmetic and Logical Complex Operation for High-Speed Processing. In *2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)* (pp. 928-931). IEEE.
74. Akat, G. B., & Magare, B. K. (2022). Mixed Ligand Complex Formation of Copper (II) with Some Amino Acids and Metoprolol. *Asian Journal of Organic & Medicinal Chemistry*.
75. Murugadoss, R., Praveen, R. V. S., Kunjumohamad, S. C., & PS, B. (2025). Osegnet-F-Unext: O-Segnet-Fusion-Unext for pulmonary lobe segmentation of Covid-19 using Computed Tomography image. *European Spine Journal*, 1-17.
76. Thatikonda, R., Thota, R., & Tatikonda, R. (2024). Deep Learning based Robust Food Supply Chain Enabled Effective Management with Blockchain. *International Journal of Intelligent Engineering & Systems*, *17*(5).
77. Kagga, S. R. (2025). MIGRATING LEGACY HEALTHCARE SYSTEMS TO CLOUD-NATIVE MICROSERVICES WITH AI: BEST PRACTICES AND PITFALLS. *International Journal of Applied Mathematics*, *38*(2s), 914-949.
78. Sundaramoorthy, P., Praveen, R. V. S., Puli, B., Tiwari, A., Kanimozhi, S., & Keerthana, N. V. (2025, October). Decentralized Anomaly Detection in IoT Networks Using Federated Learning Models. In *2025 International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU)* (pp. 1-6). IEEE.
79. Reddy, D. D., & HimaBindu, G. (2024, June). A Long-Short Term Memory Model-based approach for smart intrusion detection systems. In *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-4). IEEE.
80. Kshirsagar, K. P. (2015). Key Frame Selection for One-Two Hand Gesture Recognition with HMM. *International Journal of Advanced Computer Research*, *5*(19), 192.
81. Shivaraj, R. K., Ramesh, S. N., & Shaheeda Banu, S. (2015). Effect of TM and loop length on drape coefficient of single jersey knitted fabrics. *Int J Adv Res Eng Technol*, *6*(1), 1-6.
82. Praveen, R. V. S., Alsalami, Z., Varshney, N., Rajalakshmi, B., Prasad, K. S., & Boob, N. S. (2025, September). AI-Integrated Demand Response with Dynamic Pricing in Prosumer-Driven Renewable Microgrids. In *2025 International Conference on Computing and Communications (COMPUTINGCON)* (pp. 1-6). IEEE.
83. Ibrahim, A. H. M., Aliya, P., Ghaoud, T., Qawaqneh, Q. A., Sajwani, A. S. H., Abdullah, J., & Al Hammadi, H. (2025, November). Investigation of Flashover Incidents in Medium Voltage Capacitor Bank Circuit

doi:10.20944/preprints202603.1843.v1

Breakers. In *2025 IEEE PES Conference on Innovative Smart Grid Technologies-Middle East (ISGT Middle East)* (pp. 1-5). IEEE.

84. Joshi, S., & Kumar, A. (2013, January). Feature extraction using DWT with application to offline signature identification. In *Proceedings of the Fourth International Conference on Signal and Image Processing 2012 (ICSIP 2012) Volume 2* (pp. 285-294). India: Springer India.
85. Shrivastava, A., Praveen, R., Alfilh, R. H., Singh, N., Yadav, K., & Rajalakshmi, B. (2025, September). AI-Driven Fault Resilience: Integrating Deep Graph Neural Networks in Spatio-Temporal Smart Grid Monitoring. In *2025 International Conference on Computing and Communications (COMPUTINGCON)* (pp. 1-7). IEEE.
86. Sudhakar, K., Saravanan, D., Hariharan, G., Sanaj, M. S., Kumar, S., Shaik, M., ... & Aurangzeb, K. (2023). Optimised feature selection-driven convolutional neural network using gray level co-occurrence matrix for detection of cervical cancer. *Open Life Sciences*, *18*(1), 20220770.
87. Rokade, U. S., Doye, D., & Kokare, M. (2009, March). Hand gesture recognition using object based key frame selection. In *2009 International Conference on Digital Image Processing* (pp. 288-291). IEEE.
88. Tatikonda, R., Kempanna, M., Thatikonda, R., Bhuvanesh, A., Thota, R., & Keerthanadevi, R. (2025, February). Chatbot and its Impact on the Retail Industry. In *2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)* (pp. 2084-2089). IEEE.
89. Suganya, V., Vijayakumar, L., Annur, E. A., Praveen, R. V. S., Bharathi, A., & Amsa, M. (2025, September). A Hybrid LSTM-Fuzzy Inference Model for Uncertainty-Aware Stock Market Forecasting. In *2025 International Conference on Electronics and Computing, Communication Networking Automation Technologies (ICEC2NT)* (pp. 1-6). IEEE.
90. Zambare, P., & Liu, Y. (2023, October). An optimized graph neural network-based approach for intrusion detection in smart vehicles. In *IFIP International Internet of Things Conference* (pp. 3-17). Cham: Springer Nature Switzerland.
91. Uma, G. (2018). Survey on Data Deduplication Techniques used for Efficient Management of Cloud Storage. *International Journal Computer Science Management Studies*, *7*(3).
92. Kumar, R. S. P., & Banu, S. S. (2025). An exploration of strategies in marketing of organic fruits and vegetables. *South Eastern European Journal of Public Health*.
93. Punitha, A., & Ramani, P. (2025). Dynamically stabilized recurrent neural network optimized with intensified sand cat swarm optimization for intrusion detection in wireless sensor network. *Computers & Security*, *148*, 104094.