Literature Survey

Project Title **Predictive Retention**: An Explainable AI Framework for Subscription Forecasting & Proactive Churn Prevention

Paper ID	Paper Title & Year	Research Focus / Main Contribution	Methodology / Models Used	Key Insights / Techniques	Limitations & Research Gaps
[1]	Comparative Study of Resampling for Telecom (2024)	Analyzes the impact of different sampling methods on churn prediction.	XGBoost, RF, LR	Uses SMOTEENN to handle telecom data imbalance.	Focuses purely on accuracy; lacks interpretability for business decisions.
[2]	SHAP-Guided Tuning for Retention (2024)	Uses XAI to guide the hyperparameter tuning process.	XGBoost, Bayesian Opt.	Proves that SHAP values can improve model reliability.	High computational cost; lacks real-time prescription logic.
[3]	Efficiency of XAI in FinTech (2023)	Measures the trade-off between model speed and explanation depth.	LIME, SHAP, LightGBM	Balances the speed of LIME with the accuracy of SHAP.	Explanations are often too technical for marketing implementation.
[4]	Impact of Tuning on XAI Stability (2023)	Investigates how hyperparameter shifts affect feature importance.	SVM, K-NN, SHAP	Reveals that tuning shifts SHAP values significantly.	Fails to address data imbalance (no SMOTE used), leading to biased values.
[5]	Multi-dimensional Banking Analytics (2024)	Analyzes customer behavior across multiple digital touchpoints.	Hybrid Ensemble	Integrates digital banking logs for churn detection.	Results are sector-specific and do not transfer to telecom features.
[6]	Online Gaming Player Attrition (2024)	Detects patterns leading to voluntary player loss.	Transparency AI	Identifies player frustration/cheating as churn triggers.	Focuses on behavioral sentiment rather than billing/contractual drivers.
[7]	Exploration of Leverage Points (2023)	Maps global feature importance to customer segments.	XAI Framework	Identifies critical "points of no return" in the customer journey.	Diagnostic only; lacks a logic layer to convert drivers into actions.
[8]	Broadband Service Risk Prediction (2023)	Predicts churn based on technical service performance logs.	SHAP, Random Forest	Explains specific service risks (e.g., slow internet).	Ignores financial/contractual factors, looking only at tech logs.
[9]	Business Intelligence & XAI (2024)	Integrates XAI into existing executive dashboards.	XAI for BI	Maps AI output to standard BI decision-making flows.	Theoretical approach; lacks a functional prototype for live agents.
[10]	Graph-based Inter-dependency (2025)	Models customer relationships to detect "social" churn.	GNN, XGBoost	Identifies "influencer" customers whose churn affects others.	Extremely complex architecture; requires massive server resources.
[11]	E-commerce Ecosystem Growth (2024)	Focuses on community retention in digital marketplaces.	GenAI, Social Comp.	Uses generative agents to simulate retention scenarios.	Focuses on community growth rather than individual user churn.
[12]	Domain-Driven Feature Engineering (2024)	Enhances model performance via expert-defined features.	CatBoost, SMOTE	Uses explainability to validate engineered features.	High dependency on domain experts; difficult to automate.
[13]	Integrated AutoML Analytics (2025)	Streamlines the machine learning pipeline for telecom.	AutoML, XAI	Proves AutoML can match manual feature selection.	"Black box" nature of AutoML makes logic less transparent.
[14]	Retail Banking Interpretability (2023)	Examines local explanations for high-net-worth clients.	Interpretable ML	Identifies personalized reasons for bank churn.	Small dataset size limits the generalizability of the findings.
[15]	Competitive Enterprise Strategies (2024)	Uses AI to predict churn caused by competitor offers.	ML Benchmarking	Focuses on external market drivers for churn.	Lacks local interpretability (cannot explain individual churners).
[16]	Time-series B2B Churn (2024)	Predicts churn in enterprise-level contracts over time.	LSTM, GRU	Uses sequential data to find "churn-heavy" seasons.	Requires multi-year historical data which is rarely accessible.
[17]	Bank Churn Using SHAP (2023)	Provides visual explanations for retail banking customers.	RF, SHAP	Visualizes local factors (e.g., low balance) for agents.	Models show the "Why" but offer no "How-to-Fix" strategy.
[18]	Credit Card Optimization (2024)	Optimizes churn thresholds using meta-heuristic algorithms.	Sine Cosine Alg.	Mathematically optimizes the precision-recall trade-off.	Optimization is purely theoretical; ignores business constraints.
[19]	General Telecom Benchmarking (2023)	Baseline comparison of classic algorithms on telco data.	SVM, NB, KNN	Establishes standard performance metrics for the industry.	Outdated algorithms compared to modern 2025 XGBoost versions.
[20]	Financial Life Moment Segments (2024)	Identifies life events (e.g., moving house) that trigger churn.	K-means, SVM	Uses seasonal segmentation to find churn clusters.	Unsupervised nature makes it difficult to measure prediction accuracy.
[21]	**Base Paper:** In-depth XAI Analysis (2025)	Comprehensive benchmark of XAI tools for telecom retention.	RF, KNN, XGBoost	Uses SMOTEENN & SHAP/LIME; defines the "Actionability Gap."	Explicitly states that converting SHAP to "Actions" is a major future need.