An Explainable AI Framework for Multi-Domain Customer Churn Prediction with Generative Retention Recommendation

Sujatha Alugoju[1], Gopi Sai Dudaboina[2], Ruchitha Chintamaboina[3], Prem Naveen Abhishek Eggina[4], Dr. K. S. N. Prasad[5]

UG Students[1,2,3,4], Professor[5]

Department Of Computer Science and Engineering (Data Science)

Sasi Institute of Technology and Engineering,

Tadepalligudem, India.

Abstract

Customer churn prediction is a critical challenge across industries such as telecommunications, banking, and e-commerce, where retaining existing customers is significantly more cost-effective than acquiring new ones. This study proposes an integrated AI-driven framework that combines machine learning, Explainable Artificial Intelligence (XAI), and Generative AI to predict customer churn and generate personalized retention strategies. Multiple machine learning models, including Random Forest, CatBoost, and LightGBM, are employed to enhance predictive performance and ensure robustness across datasets, while SHAP (Shapley Additive Explanations) is utilized to improve model interpretability by identifying key factors influencing churn predictions. Furthermore, Generative AI techniques are incorporated to transform model explanations into actionable business strategies tailored to individual customer behavior, extending beyond traditional predictive systems. The proposed framework is implemented as a web-based application using Flask, enabling real-time prediction, visualization, and automated decision support. Experimental results demonstrate high predictive performance, achieving accuracy up to 95% and ROC-AUC values approaching 0.95. Overall, the system bridges the gap between predictive analytics and business decision-making by integrating prediction, explanation, and action into a unified and scalable solution for effective customer retention.

Keywords: Customer Churn Prediction, Machine Learning, Explainable Artificial Intelligence, XAI, SHAP, Generative AI, Ensemble Learning, Random Forest, CatBoost, LightGBM.

1.Introduction

Customer churn, defined as the phenomenon where customers discontinue their relationship with a company, has become a critical challenge for organizations operating in highly competitive industries such as telecommunications, banking, and e-commerce. In today's data-driven business environment, customer retention plays a crucial role in sustaining profitability and long-term growth, as acquiring new customers is significantly more expensive than retaining existing ones [5], [15]. Even a small increase in customer retention rates can lead to a substantial improvement in revenue and customer lifetime value, making churn prediction a strategic priority for modern organizations.

With the rapid growth of digital platforms and the increasing availability of customer data, machine learning techniques have been widely adopted to analyze customer behavior and predict churn with high accuracy. Traditional approaches utilize classification algorithms such as Decision Trees, Random Forest, and Gradient Boosting to identify patterns in customer interactions and service usage [16][18]. More recently, advanced boosting techniques such as XGBoost, LightGBM, and CatBoost have demonstrated superior performance due to their ability to handle large-scale data and complex feature interactions efficiently [3]. Additionally, hybrid and deep learning-based models have been proposed to further enhance predictive capabilities, particularly in large and heterogeneous datasets [4], [7]. However, despite these advancements, most existing models primarily focus on maximizing prediction accuracy while neglecting other critical aspects such as interpretability and usability in real-world scenarios.

One of the major limitations of high-performing machine learning models is their lack of transparency, as they often operate as "black-box" systems. This lack of interpretability creates significant challenges for business stakeholders, who require clear and understandable explanations to support decision-making processes. To overcome this issue, Explainable Artificial Intelligence (XAI) techniques have been introduced to improve model transparency and trust. Among these, SHAP (Shapley Additive Explanations) has emerged as a widely adopted method for interpreting model predictions by quantifying the contribution of each feature [19]. Several recent studies have successfully integrated XAI techniques into churn prediction frameworks to identify key factors influencing customer behavior and enhance decision support [9], [11]. Nevertheless, these approaches often stop at explanation and do not extend to generating actionable strategies for customer retention.

In recent years, Generative Artificial Intelligence has gained significant attention as a powerful tool for transforming analytical insights into practical business actions. Generative models can analyze patterns in customer data and generate personalized recommendations, enabling organizations to move from reactive churn prediction to proactive churn prevention [6], [14]. These models offer the capability to simulate customer behavior and suggest targeted interventions tailored to individual customer profiles. However, research in this domain is still in its early stages, and most existing solutions fail to integrate predictive modeling, explainability, and generative capabilities into a single cohesive system.

Another important limitation in existing churn prediction systems is the lack of real-world deployment and scalability. Many proposed models remain confined to experimental environments and do not address practical considerations such as real-time prediction, user interaction, and system integration. Although some studies have explored web-based implementations using frameworks such as Flask, these systems often lack explainability features and automated decision support mechanisms [12]. Furthermore, many existing approaches are domain-specific and do not provide a generalized framework applicable across multiple industries, thereby limiting their scalability and practical applicability.

To address these challenges, this study proposes an integrated AI-driven framework for customer churn prediction and retention that combines machine learning, Explainable Artificial Intelligence, and Generative AI into a unified system. The proposed approach leverages multiple machine learning models to ensure robust and accurate predictions across diverse datasets, while SHAP is employed to interpret model outputs and identify key churn drivers. These insights are further utilized by a Generative AI component to generate personalized and actionable retention strategies. In addition, the system is implemented as a web-based application to enable real-time prediction, visualization, and automated decision support. By bridging the gap between predictive analytics and business action, the proposed framework provides a comprehensive, scalable, and practical solution for effective customer retention.

2. Literature Review

Daniyal Asif, Muhammad Shoaib Arif, and Aiman Mukheimer (2025) proposed a telecom churn prediction framework using XGBoost, CatBoost, LightGBM, SHAP, and LIME [1]. Their model improves prediction accuracy and interpretability but has high computational cost and lacks deployment and automated retention strategies.

Jiaming Yuan and Hao Liu (2025) developed a telecom churn prediction model using CatBoost with SHAP-based explainability [2]. Their study improves transparency in prediction but is limited to a single telecom dataset and does not support multi-domain applications or retention strategy generation.

Hanyuan Zhang, Yankai Wang, Zexuan Li, and Xiaoyin Wang (2025) compared LightGBM, CatBoost, and XGBoost for bank customer churn prediction [3]. While their study identifies effective boosting models, it lacks explainability features and retention recommendation mechanisms.

Jiasong Ye (2025) introduced a Flask-based churn prediction system using Random Forest, SVM, XGBoost, and stacking methods. Although the system supports web deployment, it achieves only moderate accuracy and lacks XAI and intelligent retention support.

Usman Gani Joy, Kazi Ekramul Hoque, Mohammed Nazim Uddin, Linkon Chowdhury, and Seung-Bo Park (2024) proposed a hybrid churn model combining LSTM, GRU, LightGBM, and SHAP for streaming services [4]. The system improves prediction performance but requires large datasets and lacks real-time deployment and Generative AI retention features.

Awais Manzoor, Muhammad Atif Qureshi, Etain Kidney, and Luca Longo (2024) reviewed machine learning methods for churn prediction across multiple domains [5]. Their work provides theoretical insights but lacks implementation and integration of XAI with Generative AI.

Somak Saha, Chamak Saha, Md. Mahidul Haque, Md. Golam Rabiul Alam, and Ashis Talukder (2024) developed ChurnNet, a deep learning telecom churn model using CNNs and attention mechanisms [7]. Despite strong predictive accuracy, the model is less interpretable and does not provide actionable retention insights.

Arjun Sirangi (2024) explored Generative AI for predictive churn immunization using GANs, VAEs, and reinforcement learning. The approach is innovative but remains experimental, lacks real-world validation, and is not integrated with ML prediction pipelines.

Overall, existing studies improve churn prediction accuracy but fail to combine prediction, explainability, retention strategy generation, and deployment into one unified system. This gap motivates the proposed integrated framework.

3.Research Gap

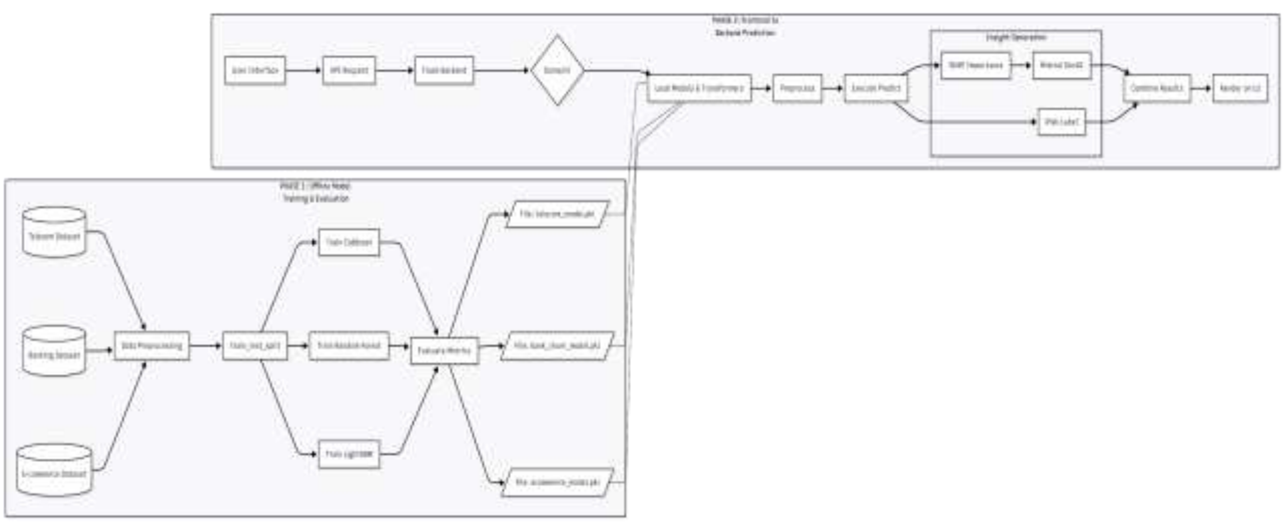
Despite significant advancements in customer churn prediction, existing research still presents several limitations that restrict its practical applicability. Most studies focus primarily on improving predictive accuracy using advanced machine learning and ensemble techniques such as XGBoost, Random Forest, and Gradient Boosting, while giving limited attention to interpretability and real-world usability [3], [16]–[18]. Although Explainable Artificial Intelligence (XAI) techniques such as SHAP have been introduced to improve transparency and identify key churn drivers, these approaches are generally limited to explanation and do not provide actionable insights for customer retention [9], [11], [19]. Furthermore, recent research exploring Generative Artificial Intelligence demonstrates the potential to generate personalized retention strategies; however, these systems are often experimental and lack strong integration with traditional machine learning pipelines [6], [14]. In addition, many existing solutions rely on single-domain datasets, limiting their generalizability across industries such as telecommunications, banking, and e-commerce [3], [5]. Several studies also highlight that current systems lack real-world deployment capabilities and user-friendly interfaces, reducing their effectiveness in practical business environments [4], [12]. Moreover, most approaches address prediction, interpretability, and strategy generation independently rather than as part of a unified framework, creating a significant gap in developing an end-to-end system that integrates accurate prediction, explainable insights, and actionable decision-making [1], [15]. Therefore, there is a clear need for a comprehensive and scalable solution that bridges this gap by combining predictive modeling, explainability, and intelligent recommendation within a single integrated system.

4. Methodology

4.1 System Architecture

The proposed system follows a multi-stage architecture integrating machine learning, Explainable Artificial Intelligence (XAI), and Generative AI to provide a complete solution for customer churn prediction and retention. The architecture consists of five major components: data preprocessing, predictive modeling, explainability module, generative recommendation system, and web-based deployment.

Initially, customer data from multiple domains, including telecommunications, banking, and e-commerce, is collected and preprocessed. The processed data is then passed through domain-specific machine learning models trained to predict churn probability. The output predictions are further interpreted using SHAP-based explainability techniques to identify key factors influencing churn. Finally, a Generative AI module transforms these insights into actionable retention strategies. The overall design aligns with recent research emphasizing integrated AI-driven churn prediction frameworks [1], [4].



4.2 Data Preprocessing and Feature Engineering

Data preprocessing is performed separately for each domain to ensure data quality and model efficiency. The preprocessing steps include handling missing values, encoding categorical variables, and feature transformation. In the telecom dataset, additional engineered features such as average call duration, service call ratio, and day-to-night usage ratio are created to capture behavioral patterns. These features enhance the model's ability to detect churn-related patterns.

For the banking dataset, categorical features such as geography and gender are encoded using label encoders, ensuring compatibility with machine learning models. Numerical features such as credit score, balance, and salary are directly utilized.

In the e-commerce dataset, features such as customer tenure, satisfaction score, complaint history, and transaction-related attributes are used. Missing features are handled by assigning default values to maintain consistency with the trained model.

The datasets used in this study are summarized in Table 1.

Table 1: Dataset Description

Domain	Features Used	Target Variable	Dataset Size
Telecom	Call minutes, calls, service calls, plans, usage ratios	Churn (0/1)	534 samples
Banking	Credit score, geography, gender, age, balance, tenure, salary	Churn (0/1)	2000 samples
E-commerce	Tenure, satisfaction, complaints, cashback, last order days	Churn (0/1)	789 samples

The diversity of datasets improves the generalizability of the proposed system across multiple domains [5], [15].

4.3 Predictive Modeling

The system employs domain-specific machine learning models to improve prediction accuracy and generalization:

• Telecom Domain: CatBoost classifier is used due to its efficiency in handling categorical and numerical data [3].

• Banking Domain: Random Forest classifier is implemented for its robustness and ability to handle imbalanced datasets [17].

• E-commerce Domain: LightGBM model is used for its scalability and high performance on large datasets [3], [16].

The models are trained using supervised learning techniques and evaluated using performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC.

The machine learning models used across domains are summarized in Table 2.

Table 2: Machine Learning Models Used

Domain	Model Used	Reason for Selection
Telecom	CatBoost	Handles categorical data efficiently
Banking	Random Forest	Robust and handles imbalance well
E-commerce	LightGBM	Fast, scalable, high performance

The performance of these models is presented in Table 3.

Table 3: Model Performance Comparison

Domain	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Telecom	0.95	0.96	0.68	0.80	0.88
Banking	0.85	0.63	0.61	0.62	0.85
E-commerce	0.93	0.80	0.82	0.81	0.95

The results demonstrate that ensemble and boosting methods provide strong predictive performance, consistent with existing literature [16][18].

4.4 SHAP

To address the black-box nature of machine learning models, SHAP (Shapley Additive Explanations) is integrated into the system for interpretability. A TreeExplainer is used to compute SHAP values for each feature, allowing identification of the most influential factors contributing to churn predictions.

The system extracts the top contributing features based on their impact values and generates human-readable explanations indicating whether a feature increases or decreases churn risk. Additionally, SHAP summary plots are generated and visualized within the web application to provide both global and local interpretability.
This approach improves transparency and supports decision-making, as highlighted in prior research [9], [11], [19].

4.5 Generative AI-Based Retention Strategy

To enhance decision-making capabilities, a Generative AI module is incorporated to generate personalized retention strategies. The system uses a large language model (Mistral-7B-Instruct) via an inference API to generate recommendations based on churn explanations.
The model takes as input the key churn drivers identified by SHAP and produces actionable strategies tailored to the specific domain (telecom, banking, or e-commerce). In cases where the generative model fails, predefined fallback strategies are used.
This component aligns with recent advancements in AI-driven customer retention strategies [6], [14].

4.6 Web-Based Deployment

The entire system is deployed as a web-based application using the Flask framework. The application allows users to input customer data, obtain churn predictions, and view explanations and recommendations in real time. The backend handles preprocessing, model inference, and SHAP computations, while the frontend displays results such as churn probability, risk level, feature importance visualizations, and retention strategies. The system also supports batch processing and PDF report generation, improving usability in real-world applications.

5. Results and Discussion

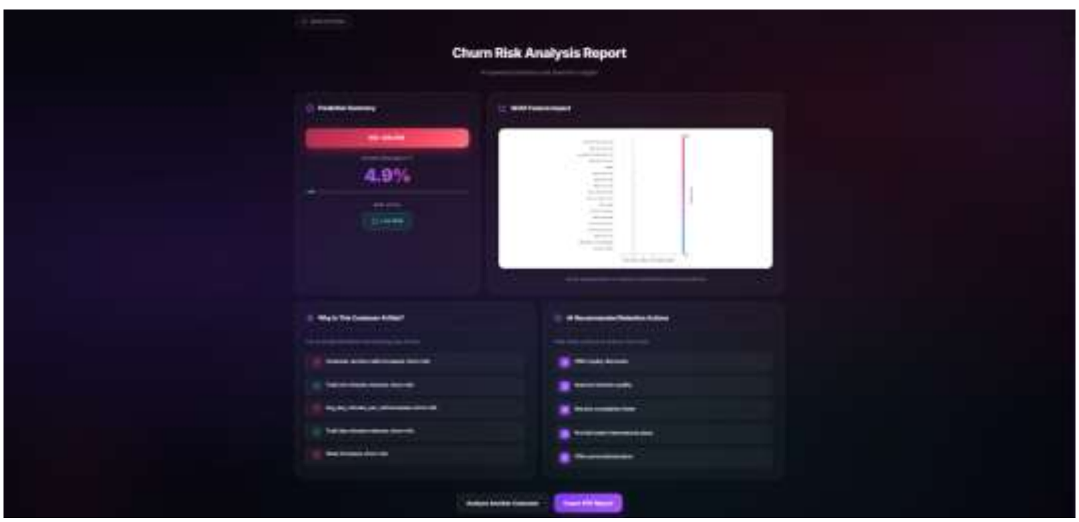


Fig. 1. Churn Risk Analysis Report for Individual Customer

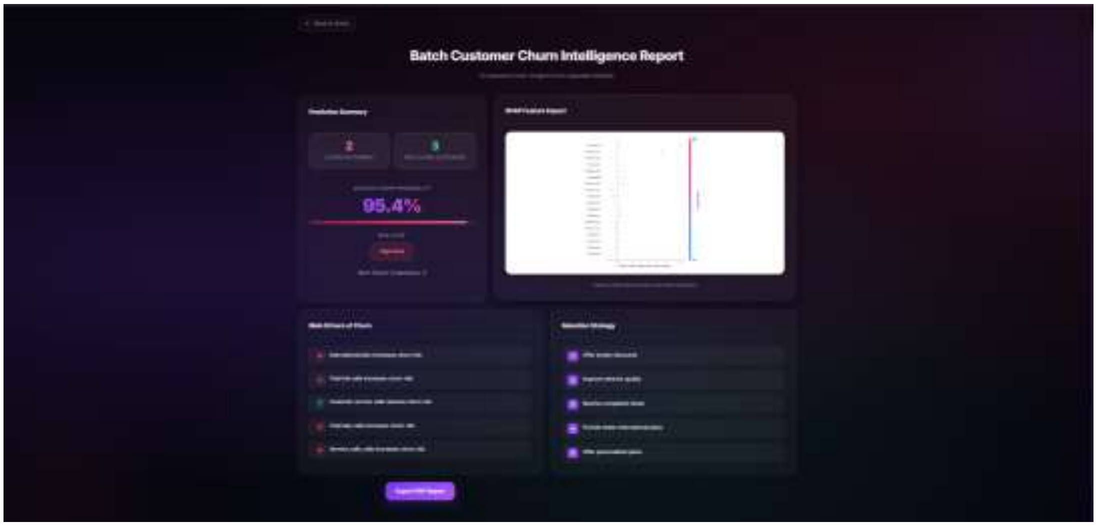


Fig. 2. Batch Customer Churn Intelligence Report

6. Conclusion

In this study, an integrated AI-driven system for customer churn prediction and retention was developed by combining machine learning, Explainable Artificial Intelligence (XAI), and Generative AI. The proposed approach effectively predicts customer churn using models such as Random Forest, CatBoost, and LightGBM, while SHAP enhances transparency by providing clear insights into the factors influencing churn. Furthermore, the integration of Generative AI enables the system to generate personalized retention strategies, bridging the gap between prediction and actionable business decisions. The deployment of the system as a web-based application ensures real-time usability and practical implementation. Overall, the proposed framework provides a scalable, interpretable, and efficient solution for customer churn management across multiple domains, contributing to improved customer retention and business performance.

7. Future Enhancements

Integrate **real-time data streaming** to enable continuous and dynamic churn prediction.
Apply **advanced deep learning models** (RNN, LSTM, Transformers) to capture sequential customer behavior.
Fine-tune **domain-specific Generative AI models** for more accurate and personalized retention strategies.
Incorporate **reinforcement learning** to optimize retention actions based on customer responses.
Utilize **multi-source data integration** (social media, feedback, transaction logs) for improved insights.
Deploy the system using **cloud-based and microservices architecture** for scalability.
Enhance the system with **interactive dashboards and visualization tools** for better user experience.
Implement **automated decision support systems** for real-time business recommendations.

References

[1] D. Asif, M. S. Arif, and A. Mukheimer, “A Data-Driven Approach with Explainable Artificial Intelligence for Customer Churn Prediction in the Telecommunications Industry,”Results in Engineering, vol. 26, 2025.DOI: 10.1016/j.rineng.2025.104629

[2] J. Yuan and H. Liu, “Telecom Customer Churn Prediction with Explainable Machine Learning,” in CMNLP 2025, 2025.DOI: 10.1145/3757110.3757152

[3] H. Zhang, Y. Wang, Z. Li, and X. Wang, "Machine Learning Models for Bank Customer Churn Prediction," in BDAIE 2025, 2025.DOI: 10.1145/3767052.3767054

[4] U. G. Joy, K. E. Hoque, M. N. Uddin, L. Chowdhury, and S.-B. Park, "A Big Data- Driven Hybrid Model for Enhancing Streaming Service Customer Retention,"IEEEAccess, 2024.DOI: 10.1109/ACCESS.2024.3401247

[5] A. Manzoor, M. A. Qureshi, E. Kidney, and L. Longo,"A Review on Machine Learning Methods for Customer Churn Prediction,"IEEE Access, 2024.
DOI: 10.1109/ACCESS.2024.3402092

[6] M. Shahabikargar, A. Beheshti, W. Mansoor, X. Zhang, E. J. Foo,A. Jolfaei, A. Hanif, and N. Shabani,"ChurnKB: A Generative AI-Enriched Knowledge Base," Algorithms, 2025.DOI: 10.3390/a18040238

[7] S. Saha, C. Saha, M. M. Haque, M. G. R. Alam, and A. Talukder,"ChurnNet," IEEE Access, 2024. DOI: 10.1109/ACCESS.2024.3349950

[8] K. Peng, Y. Peng, and W. Li,"Customer Churn Prediction and Model Interpretability Analysis,"PLOS ONE, 2023.DOI: 10.1371/journal.pone.0289724

[9] A. El Attar and M. El-Hajj,"Explainable AI-Driven Customer Churn Prediction," Frontiers in AI, 2026. DOI: 10.3389/frai.2026.1748799

[10] P. Hooda, P. Mittal, P. K. Shukla, P. K. Shukla, and A. Pandey,"Telecom Churn Analysis," Scientific Reports, 2025.DOI: 10.1038/s41598-025-34624-w

[11] H. A. Rasool, K. K. Aljawaheri, and A. A. M. Karram,"Explainable Customer Churn Prediction Using SHAP," MMEP, 2025.DOI: 10.18280/mmep.121226

[12] J. S. M. Sihombing and M. Zarlis, "Customer Churn Prediction for Retention Strategies," ICEDBC, 2026. DOI: 10.2991/978-94-6239-604-3_52

[13] T. R. Noviandy, G. M. Idroes, I. Hardi, M. Afjal, and S. Ray,"Model-Agnostic Interpretability for Churn," Infolitika Journal, 2024.DOI: 10.60084/ijds.v2i1.199

[14] A. Al Mamun, M. F. S. Tanoor, A. K. M. Masum, T. Bhuiyan,and M. M.Hassan,"Explainable AI for Customer Churn Prediction," EAI Transactions, 2026.DOI: 10.4108/airo.9813

[15] D. Freire, D. S. M. Sanchez, J. L. C. Sequera, and D. F. Moncayo,"Factors and Explainability of Telecom Churn," IEEE Access, 2024 DOI: 10.1109/ACCESS.2024.3443318

[16] T. Chen and C. Guestrin,"XGBoost: A Scalable Tree Boosting System," KDD, 2016.
DOI: 10.1145/2939672.2939785

[17] L. Breiman,“Random Forests,” Machine Learning, 2001.
DOI: 10.1023/A:1010933404324

[18] J. H. Friedman,“Greedy Function Approximation: A Gradient Boosting Machine Annals of Statistics, 2001.DOI: 10.1214/aos/1013203451

[19] S. Lundberg and S.-I. Lee,“A Unified Approach to Interpreting Model Predictions,” NeurIPS, 2017.DOI: 10.48550/arXiv.1705.07874

[20] J. Bergstra and Y. Bengio,“Random Search for Hyper-Parameter Optimization JMLR, 2012.DOI: 10.5555/2188385.2188395