Literature Review

Customer churn prediction has become a critical focus in research due to its significant impact on profitability and customer retention strategies. While various methods have been explored, existing challenges such as data imbalance, model complexity, and insufficient feature representation persist, driving the need for innovative approaches. This section synthesizes insights from recent studies, highlighting commonalities, advancements, and areas for future exploration.

Advancements in Predictive Modeling Several studies have introduced innovative models to address the limitations of traditional churn prediction methods. Ensemble-based approaches, such as the Ensemble-Fusion model, and deep learning techniques, like the BiLSTM-CNN model, have demonstrated superior predictive accuracy by leveraging hybrid architectures. The Ensemble-Fusion model, which combines multiple machine learning classifiers, achieved exceptional results, including a 95.35% accuracy and a 96.96% F1-score, outperforming other models. Similarly, the BiLSTM-CNN model enhanced accuracy (81%) and improved precision and recall metrics, underscoring the potential of hybrid approaches.

These advancements highlight the importance of combining diverse classifiers or techniques to better capture complex patterns in churn behavior. However, both models face challenges in feature representation and data availability. For instance, the Ensemble-Fusion model's reliance on large datasets is hindered by privacy concerns, while the BiLSTM-CNN model's use of limited features restricts its ability to fully capture churn dynamics. Future research could explore integrating pre-trained embeddings and incorporating richer, more diverse datasets to address these shortcomings.

Balancing Accuracy and Profitability While predictive accuracy is essential, profitability-driven approaches offer a complementary perspective on churn management. For example, the study on managing churn to maximize profits emphasizes the need to optimize retention campaign size and selection criteria. By focusing on loss function design, this research improved profitability metrics, achieving a Gini coefficient of 0.142 and a Top Decile Lift of 1.291. Similarly, the Bagging and Boosting models demonstrated their ability to rank customers effectively based on churn risk, with a 16% improvement in the Gini coefficient and a $3.2 million projected financial gain from targeted retention campaigns.

These studies collectively highlight the need to balance predictive accuracy with actionable profitability insights. Incorporating profit-driven metrics into model evaluation can ensure that retention strategies not only identify at-risk customers but also maximize economic returns. However, current approaches often neglect the long-term impact of interventions, presenting an opportunity for future research to explore dynamic models that account for changing market conditions and customer responses over time.

Addressing Data Imbalance and Feature Diversity Data imbalance remains a significant challenge in churn prediction, particularly for rare events like customer defection. The Bagging and Boosting study tackled this issue through balanced sampling schemes and bias correction techniques, which enhanced performance metrics while maintaining realistic churn probabilities. Additionally, the Ensemble-Fusion model's success underscores the value of integrating multiple classifiers to mitigate biases introduced by data imbalance.

Feature diversity is another critical factor in improving model performance. The BiLSTM-CNN model's reliance on numeric features and random embeddings limits its ability to capture nuanced customer behaviors. Similarly, the hierarchical competing risk model proposed for customer lifetime value focuses on specific contractual services but does not fully explore cross-domain feature integration. Future research could address these gaps by incorporating alternative representation models, such as pre-trained embeddings, and leveraging a broader range of behavioral, demographic, and interaction-based features.

Unifying Churn Prediction and Retention Strategies Ultimately, effective churn prediction models must integrate seamlessly with retention strategies to deliver actionable insights. Studies on profitability-driven approaches, such as managing churn to maximize profits, emphasize the importance of tailoring retention interventions to specific customer segments. Meanwhile, the hierarchical competing risk model offers a granular understanding of churn causes, enabling more targeted and effective retention campaigns. Combining these perspectives could lead to models that not only predict churn but also optimize retention strategies based on customer lifetime value and response heterogeneity.

Future Directions Despite significant advancements, challenges such as data privacy, model complexity, and dynamic market conditions persist. Future research should prioritize:

- Developing dynamic models that adapt to fluctuating churn rates and long-term intervention impacts.
- Integrating richer feature sets and alternative representations to enhance model robustness.
- Balancing predictive accuracy with profitability-driven metrics to ensure actionable insights.
- Exploring hybrid approaches that unify predictive modeling with tailored retention strategies.

By addressing these challenges, researchers can develop comprehensive frameworks that advance churn prediction and retention management, driving both accuracy and profitability in diverse industries.

A novel classifcation algorithm for customer churn prediction based on hybrid Ensemble-Fusion model

- What problem does it address?
  - Customer Churn
- Why do existing methods for solving the problem not work?
  - Too many machine learning models to keep track of
- What is innovative about the article (what are the contributions)?
  - Compares the most popular models and finds the winning one
- What are the results?
  - Ensemble-Fusion model reaches 95.35% accuracy, AUC score is 91% and F1-Score reaches 96.96%
  - Ensemble-Fusion model outperforms
- What are shortcomings of the article or future research opportunities?
  - Limitations:
    - hard to gather relevant data due to data privacy issues
    - Hard to label customer churn due to lack of business knowledge
  - Future research opportunities:
    - Similar ensemble-fusion classification algorithm that substitues the baseline classifiers with reinforcement learning model-related algorithms
    - Obtain more data from industry and combining different feature data
    - Relax strict algorithmic constraints

Customer churn prediction using composite deep learning technique

- What problem does it address?
  - Composite deep learning techniques
- Why do existing methods for solving the problem not work?
  - Lacks in delivering promising results for detecting client churn
- What is innovative about the article (what are the contributions)?
  - Hybrid deep learning model termed BiLSTM-CNN
  - Increase churn prediction process's accuracy
- What are the results?
  - BiLSTM-CNN model attained a remarkable accuracy of 81%
  - Enhanced accuracy (81%), precision (66%), recall (64%), and f-score (65%)

- What are shortcomings of the article or future research opportunities?
    - Limitations:
        - Binary classifications
        - Did not explore other composite Deep Learning models
        - Only use numeric features to predict customer attrition
        - Random feature embedding and did not consider alternative representation models such as “pre-trained”
        - Only use 20 features and should of used more
    - Future research opportunities:
        - Similar ensemble-fusion classification algorithm that substitues the baseline classifiers with reinforcement learning model-related algorithms
        - Obtain more data from industry and combining different feature data
        - Relax strict algorithmic constraints

Managing Churn to Maximize Profits

- What problem does it address?
    - Creating profitable retention campaigns
- Why do existing methods for solving the problem not work?
    - Neglects optimizing target size of a retention campaign
- What is innovative about the article (what are the contributions)?
    - Pay attentioned to the choice of a loss function
- What are the results?
    - Gini coefficient of 0.142 and Top Decile Lift of 1.291
- What are shortcomings of the article or future research opportunities?
    - Limitations:
        - Rank-order customers according to the profit lift they products, in response to the specific retention campaign
        - Does not consider long-run impact of retention interventions
    - Future research opportunities:
        - Explore variations in customer responses depending on the type and depth of retention interventions and then determine the cost at which each response is maximized

Modeling Customer Lifetimes with Multiple Causes of Churn

- What problem does it address?
    - Customer churn by customer lifetime value
- Why do existing methods for solving the problem not work?
    - Extant retention models find the time till customer churn but not the heterogeneous patterns in causes of churn

- What is innovative about the article (what are the contributions)?
    - Focuses on contractual services
    - Uses a hierarchical competing risk model
- What are the results?
    - economic return on retention management activities depends on variation in customer characteristics and customer tenure with the firm, as well as specific tactics the firm wishes to deploy
    - competing risk model is necessary if managers care about the heterogeneous patterns in causes of churn
- What are shortcomings of the article or future research opportunities?
    - Limitations:
        -
    - Future research opportunities:
        - Explore marketing campaign efforts in relation to retention results

Bagging and Boosting Classification Trees to Predict Churn

What problem does it address?

- Customer churn prediction in the telecommunications industry, where churn is a rare but financially significant event.

- Why do existing methods for solving the problem not work?
  Traditional methods, such as binary logit models:
- Struggle with the imbalance in churn data (low churn rates vs. high non-churn rates).
- Fail to capture complex, nonlinear relationships between predictors and churn likelihood.
- Provide poor rankings of customers by churn risk, limiting their effectiveness for targeted retention strategies.

What is innovative about the article (what are the contributions)?

- Introduces bagging and boosting techniques to improve churn prediction accuracy, rarely applied in marketing research.
- Proposes a balanced sampling scheme and bias correction methods to address churn data imbalance.
- Quantifies financial gains from these advanced techniques in retention marketing campaigns.
- Provides diagnostic insights into churn drivers using variable importance and partial dependence plots.

What are the results?

- Bagging and boosting outperformed the logit model, with a 16% higher Gini coefficient and 26% higher top-decile lift.
- Balanced sampling with intercept correction produced the best results for churn prediction.

- Bagging led to significant financial gains, adding $3.2 million in profits in a hypothetical retention campaign.

What are shortcomings of the article or future research opportunities?
Limitations:
- The intercept correction method may not perform well in dynamic markets with fluctuating churn rates.
- Lacks theoretical frameworks for generalizing findings to other contexts or industries.
- Boosting’s complexity may limit its practical adoption.
- Future research opportunities:
- Adapt methods for dynamic markets with changing churn rates.
- Develop theoretical models to validate bias correction techniques.
- Test advanced machine-learning approaches like deep learning.
- Apply findings to other industries, such as subscription services or financial services.
- Investigate long-term impacts of retention strategies guided by machine learning.