**Customer Churn Prediction**

A project Report Submitted by

Meetkumar Patel

Srey Patel

Mridu

Under the Guidance of

**Prof. Dr. Farid Afzali**

In fulfilment for the award of the degree Of

Masters in System and Technology

(Automation and Smart System)



McMaster University

Hamilton, ON

2023 - 2024

**1. Introduction:**

In today's highly competitive business landscape, retaining customers is as crucial as acquiring new ones. Customer churn, the phenomenon where customers cease their relationship with a company, poses a significant challenge to businesses across various industries. Identifying the factors leading to customer churn and predicting potential churners in advance can empower organizations to implement proactive strategies, mitigate losses, and enhance customer satisfaction.

This project focuses on the development and deployment of a Customer Churn Prediction System using a Deep Learning approach, specifically an Artificial Neural Network (ANN) model.

Understanding and predicting customer churn is paramount for businesses aiming to maintain sustainable growth. The ability to anticipate and address customer dissatisfaction before it leads to churn enables companies to tailor retention strategies, enhance customer experiences, and ultimately drive profitability.

* 1. **Key Challenges:**
     1. **Data Volume and Variety:** The project deals with extensive and diverse datasets, including customer demographics, transaction histories, and usage patterns.
     2. **Model Generalization:** Designing a model that generalizes well to unseen data is crucial for the success of the prediction system.
     3. **Ethical Considerations:** Addressing ethical concerns related to privacy and data usage is of utmost importance in developing a responsible and compliant solution.
     4. **Overcoming the challenges:** The project employs a comprehensive methodology encompassing data preprocessing, model development, evaluation, and deployment. Leveraging the power of deep learning, specifically artificial neural networks, we aim to create a model that can learn intricate patterns and dependencies within the data. The aim is to determine which combination of hyperparameter in the model provides the highest accuracy.

**2. Problem Statement:**

In the contemporary business landscape, where customer acquisition costs are escalating, retaining existing clients has become a strategic imperative for companies. Customer Churn Prediction refers to the crucial task of identifying and forecasting which customers are likely to terminate or cancel a service. The financial repercussions of customer churn are profound, as the expenses associated with acquiring new clients often exceed those of retaining the existing ones.

The primary challenge faced by businesses is the need to accurately predict customer churn and implement timely measures to mitigate it effectively. The overarching goal is to proactively identify individuals or entities within a customer base that exhibit signs of dissatisfaction or an intention to discontinue their engagement with a product or service.

**Importance of Customer Churn Prediction:**

1. **Retaining Valuable Customers:** Identifying potential churners enables companies to focus retention efforts on high-value customers, ensuring that the most valuable segments of the customer base are retained.
2. **Enhancing Customer Experience:** Predictive analytics allows businesses to address customer concerns and dissatisfaction proactively, thereby improving overall customer experience and satisfaction.
3. **Strategic Decision-Making:** Accurate churn prediction empowers businesses with actionable insights. By understanding the factors contributing to churn, companies can make informed decisions to enhance products, services, or customer engagement strategies.
4. **Competitive Advantage:** Companies that excel in predicting and mitigating customer churn gain a competitive edge. They can position themselves as customer-centric organizations, attracting and retaining a loyal customer base.

**3. Proposed Solution**

In response to the critical business challenge of customer churn, we propose the development of a proactive approach utilizing Artificial Neural Networks (ANNs), a powerful deep learning technique. ANNs are particularly adept at uncovering intricate patterns within large and complex datasets, making them an ideal tool for predicting customer churn in advance.

**3.1 Data Collection and Preprocessing**

Below mentioned is the description of the used dataset:

Dataset used: [Telco Customer Churn (kaggle.com)](https://www.kaggle.com/datasets/blastchar/telco-customer-churn) .

This dataset consists of 7043 examples and below mentioned 21 features:

1. **customerID**: Customer ID
2. **gender**: Whether the customer is a male or a female
3. **Senior Citizen**: Whether the customer is a senior citizen or not (1, 0)
4. **Partner**: Whether the customer has a partner or not (Yes, No)
5. **Dependents**: Whether the customer has dependents or not (Yes, No)
6. **tenure**: Number of months the customer has stayed with the company
7. **Phone Service**: Whether the customer has a phone service or not (Yes, No)
8. **Multiple Lines**: Whether the customer has multiple lines or not (Yes, No, No phone service)
9. **Internet Service**: Customer’s internet service provider (DSL, Fiber optic, No)
10. **Online Security**: Whether the customer has online security or not (Yes, No, No internet service)
11. **Online Backup**: Whether the customer has online backup or not (Yes, No, No internet service)
12. **Device Protection**: Whether the customer has device protection or not (Yes, No, No internet service)
13. **Tech Support**: Whether the customer has tech support or not (Yes, No, No internet service)
14. **StreamingTV**: Whether the customer has streaming TV or not (Yes, No, No internet service)
15. **Streaming**: Whether the customer has streaming movies or not (Yes, No, No internet service)
16. **Contract**: The contract term of the customer (Month-to-month, One year, Two year)
17. **Paperless Billing**: Whether the customer has paperless billing or not (Yes, No)
18. **Payment Method**: The customer’s payment method (Electronic check, Mailed check, Bank transfer (automatic), Credit card (automatic))
19. **Monthly Charges**: The amount charged to the customer monthly
20. **Total Charges**: The total amount charged to the customer
21. **Churn Label**: Whether the customer churned or not (Yes or No)

Before the collected data is used for training and testing, a thorough preprocessing routine ensures its suitability for the waste classification system. This includes:

1. **Handling Missing values:** There are multiple ways to handle Null values, here we are going to use the impute with Mean Method.
2. **Handling Categorical data:** Applying Encoding to convert the categorical variables as numerical/binary vectors, here we have used Onehot Encoding to convert the categorical data to binary vectors.
3. **Standardization:** Ensuring uniformity in the data for consistency, we have used Standard Scaling menthod.
4. **Data balancing:** Imbalanced datasets, where one class significantly outnumbers the other, can pose challenges in machine learning, especially when training models for classification tasks. SMOTE, or Synthetic Minority Over-sampling Technique, is a popular method for addressing class imbalance. SMOTE works by creating synthetic examples for the minority class.

**3.2 Model Training**

We employed a deep learning approach using TensorFlow and Keras to develop a neural network model for a specific task.

**Neural Network Architecture:**

We designed a neural network architecture using Keras, a high-level neural networks API running on top of TensorFlow. The architecture consists of three layers: an input layer with 40 neurons, a hidden layer with 40 neurons using Leaky ReLU activation, another hidden layer with 20 neurons and Leaky ReLU activation, and an output layer with a single neuron using the sigmoid activation function.

**Optimization and Training:**

To optimize the model's performance, we used the Adam optimizer with a learning rate of 0.001. The model was trained on a labeled dataset (xin\_train and yout\_train) for 50 epochs with a batch size of 32. The fit method was employed for training, and it automatically incorporates backpropagation to update the model's weights and biases.

**Batch Normalization:**

Batch normalization layers were added to the model after each dense layer. Batch normalization helps stabilize and accelerate the training process by normalizing the inputs of each layer.

**3.7 Model Evaluation**

Following the training phase, the model's performance was evaluated using a separate test dataset (xin\_test and yout\_test). The evaluation includes calculating the test loss and accuracy.

Cross-validation results are considered to ensure the robustness of each model's performance across different subsets of the data.

**4. Expected Outcomes**

1. **Deep Understanding of ANN Architectures:** Through the training process, the model gains a deep understanding of the underlying patterns in the data. The layers and neurons within the architecture learn to capture complex relationships and features associated with customer behaviour, demographics, and transaction history.
2. **Effective Hyperparameter Tuning:** Hyperparameter tuning involves finding the optimal configuration for parameters like learning rates, batch sizes, and layer sizes. By observing the model's performance over epochs, we can identify effective hyperparameter values that contribute to faster convergence and better accuracy.
3. **Robust Performance via Cross-Validation:** Cross-validation helps assess the model's robustness by training and evaluating it on different subsets of the data. It provides a more reliable estimate of the model's performance, reducing the risk of overfitting to a specific training-validation split.
4. **Improved Feature Identification:** The ANN, through its layers, automatically identifies relevant features and relationships within the input data. The model learns to weigh different features based on their significance in predicting customer churn, providing valuable insights into the factors influencing customer behaviour.
5. **Accurate Evaluation Metrics:** The iterative nature of training, where the model is exposed to the dataset over multiple epochs, allows for continuous refinement. The goal is to observe a gradual improvement in accuracy, indicating that the model is learning and adapting to the underlying patterns in the data. The test accuracy provides an estimate of how well the model generalizes to unseen data. The loss should decrease during training, indicating that the model is learning the patterns in the data.
6. **Actionable Insights for Informed Customer Retention Strategies:** By understanding the factors contributing to customer churn, the model provides actionable insights for informed decision-making. These insights can guide the development and implementation of targeted customer retention strategies, enhancing the overall effectiveness of customer relationship management.

**5. Discussion**

**6. Future Directions**