

**Convolutional Neural Network Project (CNN)**

**Mariam Haggag 52-11236**

**Jana El-Orashy 52-2384**

**Ahmed El Wassif 52-0267**

**Mohamed Sherif52-0982**

**Mostafa El feel 52-6675**

**TA: Ahmed Mazhar**

**Problem Discussion:**

Customer churn, the rate at which customers stop doing business with a company, poses challenges due to complex data, multi-faceted causes, and difficulty in prediction. It is important because it impacts revenue, customer lifetime value, brand reputation, and provides market insights. To address this problem we decided to use neural networks. Neural networks are great for figuring out why customers leave because they're really good at understanding complex patterns in data. They can pick up on important details without us having to manually tell them what to look for. Plus, they can adapt to changes over time, handle missing info well, and uncover subtle reasons behind customer churn, making them a powerful tool for making sense of diverse and evolving datasets.

**Reasoning Discussion:**

We had to choose an appropriate dataset according to several factors, we wanted a data set that is related to customer churn but still has attributes that affect customer churn and can easily track it.We decided to choose the business problem of customer churn and our data set was specifically about customer churn in Target Variable Churn: The dataset has a clear target variable, "Churn," indicating whether a customer has churned or not. This binary classification problem is well-suited for neural network models. The dataset includes a mix of categorical and numerical features such as Senior Citizen, Partner, Dependents, Tenure, Monthly Charges, etc. Neural networks can handle both types of features, making them versatile for this dataset. We chose this dataset as In customer churn prediction, there could be complex interactions between features that contribute to a customer's decision to churn, and neural networks are capable of capturing such intricate patterns.

**Importance Of Attributes:**

We have chosen the 4 attributes/columns that are most correlated with customer churning to create our model based on them.

The first attribute is paperless billing. Paperless Billing and Customer Churn had a correlation of 0.19 according to the heatmap that we have conducted in our Google Collab document. Paperless billing had the highest correlation with customer churn.

The second attribute was Monthly charges. Monthly Charges had a correlation with Customer Churn of 0.18. This means that when the monthly charges of the services increase, so does the chance of customers churning.

The third attribute was Senior citizens. Senior citizens had a churning correlation of 0.15. The chance of a customer churning increases if they happen to be a member of the senior citizens committee.

The fourth and final attribute was the Payment method. Payment Method had a correlation with customer churning of 0.11. This attribute also had a positive correlation.

We have dropped Loyalty ID and customer ID because they are unique identifiers for each customer and do not provide meaningful information for predicting churn, they can be considered irrelevant for the modeling process and even could produce noise.Another reason to drop them to reduce overfitting. Their inclusion might encourage the model to memorize the training set rather than learning meaningful patterns.

**Variation and Suitability:**

In our analysis, we chose sequential neural network (SNN) as the chosen architecture. This decision stems from the unique features of our customer churn prediction data set.

1. Time dependence in customer tenure: Our dataset includes features like “Tenure”, which represents the customer's association time. The ability of sequential neural networks to capture temporal dependencies is critical to understanding how changes in occupancy over time relate to churn.

2. Contract details and sequential patterns: Features such as 'Contract' and 'Payment method' have sequential patterns that influence customer decisions. SNN excels at recognizing and leveraging these sequential relationships, providing an advantage in capturing the nuances associated with different contract types and payment methods.

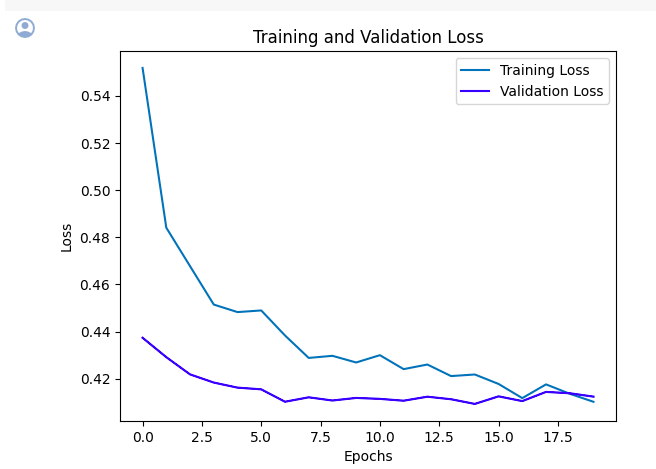
3. Sequential nature of interactive features: Customer interactions, reflected in features such as "Online Security", "Technical Support" and others, are often are sequential events. SNNs are well suited to modeling the evolving nature of customer interactions, playing a central role in predicting churn.

4. Handling varying time intervals: Sequential neural networks are capable of handling varying time intervals in sequential data. In our context, where customer interactions and behaviors may vary in frequency, SNN can effectively adapt to different time scales.

Through the use of a sequential neural network, we aim to leverage its ability to capture complex sequential patterns in our data set, thereby contributing to more accurate and robust predictions of customer Churn.

**End Results Discussion (Conclusion):**

In evaluating how well the neural network predicts customer churn in the telecom industry, the achieved accuracy of 80.77% indicates that the model accurately labeled churn status for roughly 81% of cases. Our generated model can help the telecommunication businesses and other businesses to address the problem of customer churn and Taking a closer look at precision, recall, and F1-score metrics provides a more nuanced view. Precision for non-churning customers (class 0) sits at 83%, suggesting that 83% of instances predicted as not churning were correct. Conversely, precision for churning customers (class 1) is 69%, indicating that 69% of instances predicted as churning were accurate. The recall for class 0 is 92%, showcasing the model's knack for identifying 92% of actual non-churning instances, while the recall for class 1 is 49%, indicating the model identified only 49% of actual churning instances. The F1-score, a blend of precision and recall, stands at 80%, illustrating a well-balanced and satisfying performance. The validation loss also decreased as shown in the below graph.A decreasing validation loss, when coupled with the discussed metrics such as accuracy, precision, recall, and F1-score, strengthens the overall confidence in the model's efficacy. It implies that the neural network is adapting well to the complexities of the telecommunications churn dataset since the training loss and validation loss met at some point and is becoming increasingly proficient in capturing the underlying patterns that determine whether a customer is likely to churn.This means that our model can identify customer’s with high churning probability and that companies can act fast by addressing problems, increasing loyalty and communicating with customers.



Discussing the dataset's imbalance, where non-churning instances (class 0) outnumber churning instances, suggests strategies like oversampling the minority class or considering alternative evaluation metrics. There's also an apparent precision-recall trade-off, with class 0 leaning towards precision and class 1 towards recall. Depending on the business context, tweaking the model's threshold might be necessary to prioritize precision or recall, considering the specific costs tied to false positives and false negatives.

To enhance the model, suggestions include Improving feature engineering and deepening the understanding of the business context could boost the model's predictive capabilities. Examining predicted probabilities and adjusting the decision threshold based on business needs are also crucial considerations.

In summary, while the model demonstrates solid accuracy and results are satisfactory, there's room for refinement,.A proper grasp of the business context, coupled with ongoing model tweaking, is vital for more satisfactory results. Clearly analyzing the model's strengths and limitations, discussing the impact of false positives and false negatives, and providing recommendations for further improvements complete the total evaluation of the neural network's performance in predicting customer churn.

**Dataset Link:**

https://www.kaggle.com/datasets/aadityabansalcodes/telecommunications-industry-customer-churn-dataset