You should provide a discussion about the problem of choice, The challenges why this problem is important, and the methodology used in assessing such a problem.

Customer turnover, often referred to as customer churn or attrition, is the rate at which customers stop doing business with a particular company or service. In the context of banks, this refers to clients closing their accounts or moving their financial services to another institution.

Challenges:

1. Increased Competition: With numerous banking options available, customers can easily switch to competitors, especially if enticed by better interest rates, superior services, or innovative products.

2. Service Quality Concerns: A single bad experience, like errors in account management or unsatisfactory customer service, can push a customer to switch banks.

3. Financial Product Dissatisfaction: Unsatisfactory terms and conditions, high fees, or uncompetitive rates on loans or credit cards can motivate customers to seek alternatives.

Importance of the Problem:

1. Revenue Impact: Retaining existing customers is often cheaper than acquiring new ones. A high churn rate can significantly impact a bank's revenue.

2. Brand Reputation: High turnover rates can harm a bank's reputation, making it harder to attract new customers.

3. Operational Costs: Constantly having to attract and onboard new customers due to high churn can increase marketing and operational costs.

Methodology in Assessing Customer Turnover:

1. Data Collection: Compile data on how many customers are leaving and at what point in their customer journey. Use data analytics tools to segment this data, identifying patterns and high-risk demographics.

2. Feedback Surveys: Encourage departing customers to provide feedback. Understand their reasons for leaving and areas of dissatisfaction.

3. Predictive Analytics: Use machine learning and predictive analytics to forecast which customers will most likely churn based on behavioral patterns.

You should provide a discussion about the reasoning behind the choices made for the dataset of Kaggle, in order for it to be taken as the dataset for this case study.

We chose our dataset for multiple reasons such as:

1. Relevance: The dataset was relevant to our topic and it was easy to read.
2. Size: The dataset was large enough to provide accurate results, while also being practical.
3. Format and accessibility: The format of the dataset was in a form that we could easily analyze and manipulate using our tools.
4. Documentation: Our dataset was well-documented as it explained each variable easily.

You should provide a discussion about the importance of the attributes which are chosen as the attributes upon which the predictive model is based, alongside the attributes to be predicted.

We chose an attribute called satisfaction score, as an input, and exited, as an output. The attribute satisfaction score reflects how the customer feels about the solution to their complaint on a scale of 1-5 and five being the highest score. Subsequently, if the customers are not happy with the way their complaints are dealt with they most probably will leave the bank which is represented by the attribute exited which is a binary attribute 0 meaning they didn't leave, and 1 meaning they left.

In case of a variation of the Neural Network is utilized (CRR, RNN, or LSTM for example), an introduction to the variation used should be provided, alongside an argument about the suitability of the variation to the problem nature.

We used Artificial Neural Networks (ANNs) as they are powerful tools used in understanding and predicting complex patterns. They are especially useful in solving problems like customer churn, where businesses try to figure out why customers stop using their services. ANNs are great for this because they can handle large amounts of data and find hidden patterns that might show why customers are leaving. By using ANNs to analyze past data, companies can identify which customers might leave and understand the reasons behind their decision. This information is crucial for businesses to keep their customers and improve their services.

You should provide a discussion about the conclusions documenting the end results and probabilities produced by the constructed Neural Network showcasing whether the achieved results are satisfactory or not.

The neural network has to train on a certain number of epochs to improve the accuracy over time. So we decided the nb_epoch = 100. So when you run this code, you can see the accuracy in each epoch. The first epoch, the accuracy was 95.42%. 50th epoch, the accuracy was 99.86%. In the 100th epoch, our final accuracy was 99.85%.

Due to us having a huge dataset, we used the confusion matrix to help us understand the accuracy of our model easily. The model accuracy was 99.87%. After adding multiple independent variables, we believe that the results are satisfactory as the model has a high chance of predicting if a customer is leaving or not in regards to used independent variables.

The independent variables added were:

1. Credit Score: Can have an effect on customer churn, since a customer with a higher credit score is less likely to leave the bank.
2. Age: This is certainly relevant, since older customers are less likely to leave their bank than younger ones.
3. Tenure: Refers to the number of years that the customer has been a client of the bank. Normally, older clients are more loyal and less likely to leave a bank.
4. Balance: A very good indicator of customer churn, as people with a higher balance in their accounts are less likely to leave the bank compared to those with lower balances.
5. Num of Products: Refers to the number of products that a customer has purchased through the bank. If a customer purchases a lot of products through the bank then it is more likely they will stay.
6. HasCard: denotes whether or not a customer has a credit card. This column is also relevant, since people with a credit card are less likely to leave the bank
7. Is Active Member: active customers are less likely to leave the bank
8. Estimated salary: as with balance, people with lower salaries are more likely to leave the bank compared to those with higher salaries
9. Complain: customer has complaint or not
10. Points Earned: the points earned by the customer for using a credit card. People sometimes want to save their points for something big so they are less likely to leave if they have a lot of points.
11. Card Type Diamond: type of card hold by the customer
12. Card Type Gold: type of card hold by the customer
13. Card Type Platinum: type of card hold by the customer
14. Card Type Silver: type of card hold by the customer

As for the card types the higher the card type the less likely they will leave. And of course Satisfaction score was also used as mentioned before.

Dataset Link:

https://www.kaggle.com/datasets/radheshyamkollipara/bank-customer-churn/data