write a report on bank customer churn prediction using artificial neural networks

. It should consist of the following

broad topics:

1. Introduction about the topic you have chosen and a description of its main applications.

2. A description of the challenges of bank customer churn prediction using artificial neural networks.

3. Compare and contrast different approaches that are used to solve the topic that you have

selected. Compare at least two approaches.

1. What is the task for this assessment?

4. Choose one approaches from the step 3 that you think is better to solve your problem and

describe it in detail. [5 Marks]

5. Choose a publicly available data set that has a minimum of 2000 data points and solve the

problem. You can use the following solution approach. [30 Marks]

a. Describe the raw data set.

b. Apply pre-processing or data cleaning steps. Describe the benefit of each preprocessing technique that you have applied.

c. Extract features or attributes that you will be using in your application. Give

details of all the features that you are using in your application/project.

d. Apply basic data visualisation techniques on the selected dataset that are relevant

to your topic or data set.

e. Apply a machine learning, deep learning, or artificial intelligence approach to

solve the problem. Give a reason why you have selected the particular approach.

(Note that for some topics one or more of the following steps are not required. Hence, you

have to provide justification if you are excluding any of the step mentioned above).

6. Contribution made by each team member to this project should be clearly specified.

7. Bonus marks for choosing a challenging project. If a group selects a challenging project,

then these marks will be awarded to them. Clearly specify why you think your chosen

project is challenging. [10 Marks]

8. Report Quality: Over quality of the report will be checked, including details about each

step taken, clarity of writing, appropriate referencing, etc., and the inclusion of related

code and development documentation. [10 Marks]

Please note: All the group members will be marked equally because it is expected that all the

members of the group will be contributing to the project. However, if there is a conflict among the

group members and they complain that one or more members are not contributing to the project

then each member of the group will be required to submit an individual report and they will be

marked based on the quality of their submitted report.

The report should be 4,000 words.

Here's an outline for your report, including key points and considerations for each section:

**1. Introduction to Bank Customer Churn Prediction**

* **Definition of Customer Churn**: Explain what churn is, specifically in the context of banking.
* **Importance of Churn Prediction**: Discuss why predicting churn is critical for banks, including its impact on revenue and customer relationship management.
* **Applications of Churn Prediction**: Describe various applications such as personalized marketing strategies, customer retention programs, and improving customer service.

**2. Challenges of Bank Customer Churn Prediction**

* **Data Complexity**: Discuss the challenges related to the vast amounts of data banks handle, which can be heterogeneous and unstructured.
* **Data Privacy**: Highlight issues concerning the legal and ethical aspects of handling customer data.
* **Predictive Accuracy**: Address the difficulty in achieving high predictive accuracy due to the dynamic nature of customer behaviors and the economy.
* **Integration of Solutions**: Talk about the challenges associated with integrating predictive solutions into existing banking systems.

**3. Comparison of Different Approaches to Churn Prediction**

* **Logistic Regression**: Describe how logistic regression has been traditionally used in churn prediction, noting its interpretability and simplicity.
* **Decision Trees and Random Forests**: Explain how these methods handle categorical and continuous data, and their effectiveness in identifying significant predictors of churn.
* **Support Vector Machines (SVM)**: Outline the use of SVM in churn prediction, particularly its ability to model non-linear relationships.
* **Artificial Neural Networks (ANNs)**: Discuss the advantages of using ANNs due to their ability to learn complex patterns through deep learning techniques.
* **Comparison**: Contrast these methods in terms of accuracy, scalability, interpretability, and operational complexity.

**4. Description of the Selected Approach: Artificial Neural Networks**

* **Why ANNs?**: State the reasons for choosing ANNs over other methods, particularly focusing on their superiority in handling complex and non-linear data relationships.
* **Architecture of ANNs**: Describe the typical architecture used for churn prediction, including input, hidden, and output layers.
* **Data Preprocessing**: Explain necessary preprocessing steps like normalization and missing data handling specific to ANNs.
* **Training the Network**: Discuss the process of training ANNs, including backpropagation, learning rate adjustments, and overfitting prevention.
* **Evaluation**: Detail how the model's performance is evaluated, including metrics like accuracy, precision, recall, and AUC-ROC curve.

**5. Project Challenges**

* **Complexity of Model Tuning**: Elaborate on the difficulties in tuning hyperparameters for ANNs, which requires extensive experimentation.
* **Data Quality and Availability**: Discuss how the quality and completeness of data affect the model's performance.
* **Adapting to New Patterns**: Explain the challenge of keeping the model relevant under changing customer behaviors and market conditions.
* **Interpretability Issues**: Address the "black box" nature of ANNs, which can be a barrier in regulatory and compliance environments.

**Conclusion**

Summarize the key points discussed in the report, reiterating the importance of effective churn prediction in banking and the potential of artificial neural networks to revolutionize this field.

**Additional Considerations**

* **Word Count**: Ensure each section is detailed enough to meet the 4,000-word requirement while maintaining clarity and conciseness.
* **References**: Include academic and industry sources to bolster your points and provide credibility to your discussion.
* **Visuals**: Consider incorporating figures and tables to illustrate concepts and comparisons clearly.

This structure should guide you in creating a comprehensive report that not only addresses the technical aspects of using ANNs for churn prediction but also contextualizes the discussion within the broader challenges and needs of the banking industry.