Case Study on Decision Tree

Case Study: Predicting Customer Churn using Decision Trees  
  
Problem Statement:  
A telecom company wants to predict customer churn to improve customer retention strategies. Customer churn is defined as the event when a customer stops using the company's services. The goal is to build a predictive model using Decision Trees to identify which customers are likely to churn based on historical data.  
  
Dataset:  
The dataset includes the following features:  
- Customer ID  
- Gender  
- Age  
- Monthly Charges  
- Contract Type (Monthly, One Year, Two Year)  
- Total Tenure with the company  
- Number of Customer Service Calls  
- Churn (Yes/No)  
  
Approach:  
1. Data Preprocessing:  
 - Handle missing values.  
 - Convert categorical variables into numerical form.  
 - Split the dataset into training and testing sets.  
   
2. Model Building:  
 - Use the Decision Tree algorithm to build a classification model.  
 - Use Gini Index to determine the best split at each node.  
 - Prune the tree to avoid overfitting.  
  
3. Model Evaluation:  
 - Evaluate the performance using accuracy, precision, recall, and F1-score.  
 - Compare the model's predictions with the actual churn data.  
  
Results:  
The Decision Tree model achieved an accuracy of 85% on the test dataset. The most important features for predicting churn were Monthly Charges, Contract Type, and Tenure. Pruning helped reduce overfitting, leading to better generalization on unseen data.  
  
Conclusion:  
The Decision Tree model successfully identified customers at risk of churning, allowing the telecom company to target those customers with personalized retention offers. Decision Trees proved to be an effective tool for this classification task due to their interpretability and ability to handle both categorical and continuous data.