CUSTOMER CHURN PREDICTION

Abstract Customer churn analysis and prediction in telecom sector is an issue now a days because it's very important for telecommunication industries to analyze behaviors of various customer to predict which customers are about to leave the subscription from telecom company. So machine learning techniques and algorithm plays an important role for companies in today's commercial conditions because gaining a new customer's cost is more than retaining the existing ones. This project focuses on various machine learning techniques for predicting customer churn through which we can build the classification models such as Logistic Regression, Random Forest and lazy learning and also compare the performance of these models.

1. INTRODUCTION

The customer who cease a product or service for a given period is referred as churner. In a telecommunication company, the individual who has opted service from a firm is referred to as Churn. The individual who probably intends to depart from the firm in near future was predicted by the churn model. Many industries build a model like a churn as a common application for data mining technique. Mobile telephone organizations present across the globe are almost on the verge of building their own churn model. Furthermore, to retain the customers, churn results can be efficiently utilized for various other goals.Churn Management approach is actually the first step in building a model. In general, the project needs a churn model in the best way instead of taking a single method which has the best lift. So here we have built an automated application as a default for a long run.In this digital era, the client of one company may also be a consumer of one or more telecommunication firms. Some of us may use different carriers based on the distance and some others may use different carriers based on the different plans they offer. While performing the analysis using machine learning customer experience tends to provide valuable insights. Some people will change their service providers from time to time. Increase or decrease in the calling rate will also depend on different job responsibilities. Based on the availability of the data various situations may reflect**.**

Incorporating advanced machine learning techniques like ensemble models and feature engineering can indeed help improve the prediction accuracy of your customer churn model. Here's how you can integrate these techniques into your customer churn prediction using IBM Cognos:

Ensemble Models

Ensemble models combine the predictions of multiple machine learning algorithms to produce a more robust and accurate result. Popular ensemble methods include Random Forests, Gradient Boosting (e.g., XGBoost or LightGBM), and AdaBoost.

- To incorporate ensemble models into your project:
  - Train multiple base models, such as decision trees or logistic regression, on your training data.
  - Combine the predictions of these base models using techniques like bagging (for Random Forests) or boosting (for Gradient Boosting).
  - Use the ensemble model's predictions for visualization and analysis in IBM Cognos.

Feature Engineering:

- Feature engineering involves creating new features or transforming existing ones to provide more relevant information to the model. Effective feature engineering can significantly enhance model performance.
- Consider the following feature engineering techniques:
  - Create customer segmentation based on various features like transaction history, customer behavior, or demographics.
  - Calculate customer churn risk scores based on historical churn patterns and customer attributes.
  - Incorporate time-based features, such as trends, seasonality, or recency of interactions.
  - Use domain knowledge to engineer features that are specific to your industry or business.

Hyperparameter Tuning:

- Optimize the hyperparameters of your models, including those of ensemble models, to find the best configuration for your data. Grid search or random search can help automate this process.
- Utilize cross-validation to assess how well the models generalize to new data.

4. Model Stacking (Advanced Ensemble):

Consider implementing model stacking, an advanced ensemble technique. In model stacking, you train multiple diverse models and then use another model (a meta-learner) to combine their predictions.

Use IBM Cognos to visualize the predictions and results of the meta-learner, showing how different base models contribute to the final prediction.

5. Monitoring and Regular Retraining:

- Continuously monitor the performance of your ensemble models and feature engineering techniques. As your data evolves, update and retrain your models to maintain their accuracy and relevance.

6. **Explainability:**

- Use model explainability techniques to understand why certain predictions are made by your ensemble models. This can help in gaining insights into customer behavior and improving model interpretability.

7. **IBM Cognos Integration:**

- When integrating these advanced techniques into IBM Cognos, make sure your dashboards and reports reflect the enhanced predictive capabilities. Visualize the outputs of ensemble models and feature engineering to provide a comprehensive view of customer churn risk.

8. **Feedback Loop:**

- Establish a feedback loop to collect data on the outcomes of actions taken based on your churn predictions. This data can be used to further refine and improve your models and visualizations.

By combining ensemble models, feature engineering, and other advanced techniques, you can create a more accurate and powerful customer churn prediction system that leverages IBM Cognos for effective visualization and communication of insights.