**University of Petroleum and Energy Studies**



**Customer Churn in Telecommunication Company**

# PROJECT SYNOPSIS

Create a machine learning model to predict the likelihood of customer churn in a

telecommunications company.

# BACHELOR OF TECHNOLOGY

SCHOOL OF COMPUTER SCIENCE (AIML)

SUBMITTED BY

Jitender Dhariwal - R2142201723 Jatin Wadhwa - R214220562 Kartikay Tyagi - R214220603 Jatin Joshi - R214220559 Karan Taneja - R214220587

GUIDED BY

Mr. Sumit Shukla

Content Page

|  |  |  |
| --- | --- | --- |
| **Topic** | | **Page No** |
| 1. | Introduction Page | 3 |
|  |  |  |
| 2. | Literature Survey | 4 |
|  |  |  |
| 3. | Methodology | 5 |
|  |  |  |
| 4. | Facilities required for proposed work | 6 |
|  |  |  |
| 5. | References | 7 |

**Introduction**

For telecommunications firms, client churn—the process in which consumers transfer from one provider to another—raises serious concerns. Retaining clients has risen at the top of many companies' priorities due to fierce competition and changing market conditions. Telecommunications businesses are increasingly using machine learning techniques to forecast client attrition in an effort to reduce the negative effects of churn.

Machine learning models provide an effective way to examine vast amounts of consumer data and spot trends that may be used to identify which customers are most likely to leave. These models can assist telecoms firms take preventative steps to retain at-risk consumers by making use of past customer information.

In the telecoms sector, the aim of customer churn prediction using machine learning is to create models that can precisely identify consumers who are likely to churn in the near future. Companies can enhance customer happiness and loyalty by recognizing these clients in advance and taking the required actions, such as making personalized retention offers, performing outstanding customer service, or launching focused marketing efforts.

Telecommunications firms are able to more efficiently manage their resources because to the predictive capabilities of machine learning algorithms. These models help businesses to concentrate their efforts on consumers who are most likely to churn, allowing them to optimize their retention tactics and lower the expenses involved with gaining new customers.

Furthermore, customer churn prediction models can provide insights into the factors that drive customer attrition. By analyzing the most important features contributing to churn, such as poor network quality, pricing dissatisfaction, or competitor offers, companies can gain a deeper understanding of customer preferences and pain points. This information can guide strategic decision-making, product development, and customer experience improvements.

In this era of data-driven decision-making, customer churn prediction using machine learning offers telecommunications companies a competitive advantage. By harnessing the power of advanced analytics, these companies can enhance customer retention efforts, increase customer satisfaction, and ultimately improve their bottom line.

In this guide, we will explore the steps involved in building a machine learning model for customer churn prediction in a telecommunications company. We will discuss data collection, pre-processing, feature engineering, model selection, training, evaluation, and deployment. By following these steps, telecommunications companies can develop robust churn prediction models and make data-driven decisions.

**Literature Survey**

The process of building the customer churn prediction mode , its application and reasons, challenges and obstacles for designing the model are discussed in paper "Customer Churn Prediction Analysis" published in International Journal of Computer Applications 182(29):15-17. In their work, Iris Figalist, Christoph Elsner, Jan Bosch, and Helena Holmstrom Olsson develop a technique that makes it possible to create customer and end-user map data based on "customer categories," allowing the prediction model to capture all important influential aspects.[1]

Pallav Routh describes an operating approach that precisely identifies churn dangers and establishes the connection between risks and consumer behavior within the context of an unambiguous survival forest. The suggested model, in contrast to existing approaches, does not rely on a particular functional form to explain the relationship between risk and behavior, nor does it contain fundamentally different assumptions; instead, both of these assumptions are restricted to performance.[2]

Malak Fraihat describes the Selection Ensemble Model (SEM) as a well-designed churn forecaster. SEM makes a strong choice from among a collection of ML models for a combination of models to contribute to the output.[3]

According to Deepak Gupta, they attempt to compare and evaluate the effectiveness of more than 100 classifiers in the churn forecast of a telecommunications business. They have used well-known classifiers with divergent ancestry.[4]

S. W. Kim presents a model for predicting customer attrition that identifies targeted consumers and identifies industry-specific characteristics that influence customer churn through segmentation and aggregation techniques. The qualification rating filter and benefit statistics are used to choose the features. Its results show that it creates improved RF-based churn prediction classifiers.[5]

According to Asmin Alev Aktaş, the structure of customer data is set up in the order of customer-related data. The long-term memory model is contrasted to conventional classification techniques when used with subsequent data to measure complicated consumer categories.[6]

**Methodology**

The development of an advanced prediction model that aims to predict the likelihood of customer churn in a telecommunications company. The process of development of such a system includes the following steps:

* Data collection: Compile pertinent data about your clients, including their demographics, service usage, payment information, interactions with them, and any other pertinent data points. Make sure you have information on both churned and non-churned clients.
* Data Preprocessing: Preprocessing the data involves dealing with missing numbers, outliers, and inconsistent data. Perform operations including feature scaling, data normalization, and categorical variable encoding. Create training and test sets from the data.
* Feature engineering: Examine the gathered data and develop brand-new features that may be churn-predictive. You may, for instance, compute measures like average contact time, frequency of service issues, or client retention. Feature engineering can increase your model's ability to anticipate outcomes.
* Model Selection: Choose the best machine learning method for churn prediction in your model selection process. Logistic regression, decision trees, random forests, gradient boosting, and neural networks are examples of frequently used algorithms. When choosing an algorithm, take interpretability, complexity, and performance needs into account.
* Model training: Apply the training dataset to the chosen model. The data will be analyzed by the model to identify patterns and relationships that may be used to forecast client turnover. To enhance the performance of the model, change hyper-parameters like learning rate, regularization strength, or tree depth.
* Model Evaluation: Utilizing the testing dataset, evaluate the trained model's performance. Assess the model's churn prediction ability using evaluation metrics like accuracy, precision, recall, F1 score, or area under the ROC curve. To get a more reliable evaluation, think about applying cross-validation approaches.
* Fine-tuning the model: If the model's performance is unsatisfactory, think about iterating and improving the model. Features may be chosen, hyper parameters may be altered, various algorithms may be tried, or many models may be combined using ensemble approaches.
* Model Deployment: Deploy the model into a production environment as soon as you are happy with its performance. Install the required framework to receive forecasts of churn.

**Facilities Required for proposed work**

To develop a prediction model of customer churn in a telecommunications company you would typically require the following software and hardware:

Software Requirements:

* Python: Python is a widely used programming language. It provides various libraries and frameworks that facilitate prediction model development, such as pandas, Numpy, scikit-learn, and TensorFlow.
* Integrated Development Environment (IDE): An IDE, such as PyCharm, Visual Studio Code, or Jupyter Notebook, provides a development environment with features like code editing, debugging, and project management.
* Libraries: XGBoost is a popular library for gradient boosting, which is an ensemble technique often used for predictive modeling. It provides highly optimized implementations of boosting algorithms and can improve model performance.

Hardware Requirements:

* CPU or GPU: The hardware requirements depend on the complexity of the model and the size of the dataset. For smaller models and datasets, a modern CPU (Central Processing Unit) should suffice. However, for large-scale models and datasets, a GPU (Graphics Processing Unit) can significantly accelerate the training process.
* Sufficient RAM: Such tasks often involve processing and manipulating large amounts of text data. Having sufficient RAM (Random Access Memory) is crucial to handle the data efficiently and prevent memory limitations during training and inference.
* Storage Space: Adequate storage space is required to store the dataset, pre-trained models, and any intermediate results during the model development process.

**References**

1. Iris Figalist, Christoph Elsner, Jan Bosch, Helena Holmstrom Olsson. “Customer Churn Prediction in B2B Contexts” [2019]
2. Pallav Routh, Arkajyoti Roy, Jeff Meyer. “Estimating customer churn under competing risks”. August 2020 Journal of the Operational Research Society.

1. Ahmad Hammoudeh, Malak Fraihat, Mahmoud Almomani. Selective Ensemble Model for Telecom Churn Prediction.[2019]
2. Debjyoti Das Adhikary, Deepak Gupta. “Applying over 100 classifiers for churn prediction in telecom companies.” Multimedia Tools and Applications[2020]
3. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam and S. W. Kim, "A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector," in IEEE Access, vol. 7, pp. 60134-60149, 2019.
4. Ahmet Tuğrul Bayrak, Asmin Alev Aktaş, Orkun Susuz, Okan Tunalı. ”Churn Prediction with Sequential Data Using Long Short Term Memory”[2021]