BIG DATA AND DATA SCIENCE

IDEA SELECTION AND ABSTRACT SUBMISSION

NAME: JHANSI BALLA

STUDENT ID: 1101681

Project Idea: Forecasting Customer Churn in Subscription-based Businesses

Customer churn poses a significant obstacle for services that rely on subscriptions, impacting revenue and profitability. This project's goal is to utilize machine learning models to forecast customer turnover and pinpoint the main behavioral and demographic elements influencing it. The study will concentrate on constructing a forecasting model using customer information like age, gender, location (demographics), length of subscription, type of plan (subscription details), payment behavior, failed transactions (billing history), frequency of login, feature usage (usage patterns), and number of support requests, resolution time (customer support interactions). Information for this project will be gathered from a variety of sources, such as publicly accessible datasets on platforms such as Kaggle, API data from customer relationship management (CRM) systems and product analytics tools, and custom surveys for more customer feedback. Machine learning methods such as Logistic Regression, Random Forest, and XGBoost will be utilized to create the churn prediction model, with the assistance of data mining techniques to reveal underlying patterns that impact customer churn. These methods will help companies take preemptive actions to keep customers who are at risk, enhance customer lifetime value, and decrease churn rates. The anticipated result of this project is a scalable, data-driven solution that predicts customer churn accurately and offers actionable insights for retention strategies aimed at specific targets. Focusing on high-risk customers can assist subscription-based businesses in minimizing churn, boosting customer engagement, and increasing profitability. The solution will be tailored to manage extensive data by utilizing big data tools such as Apache Spark, ensuring its suitability for practical uses.