

# Problem Statement

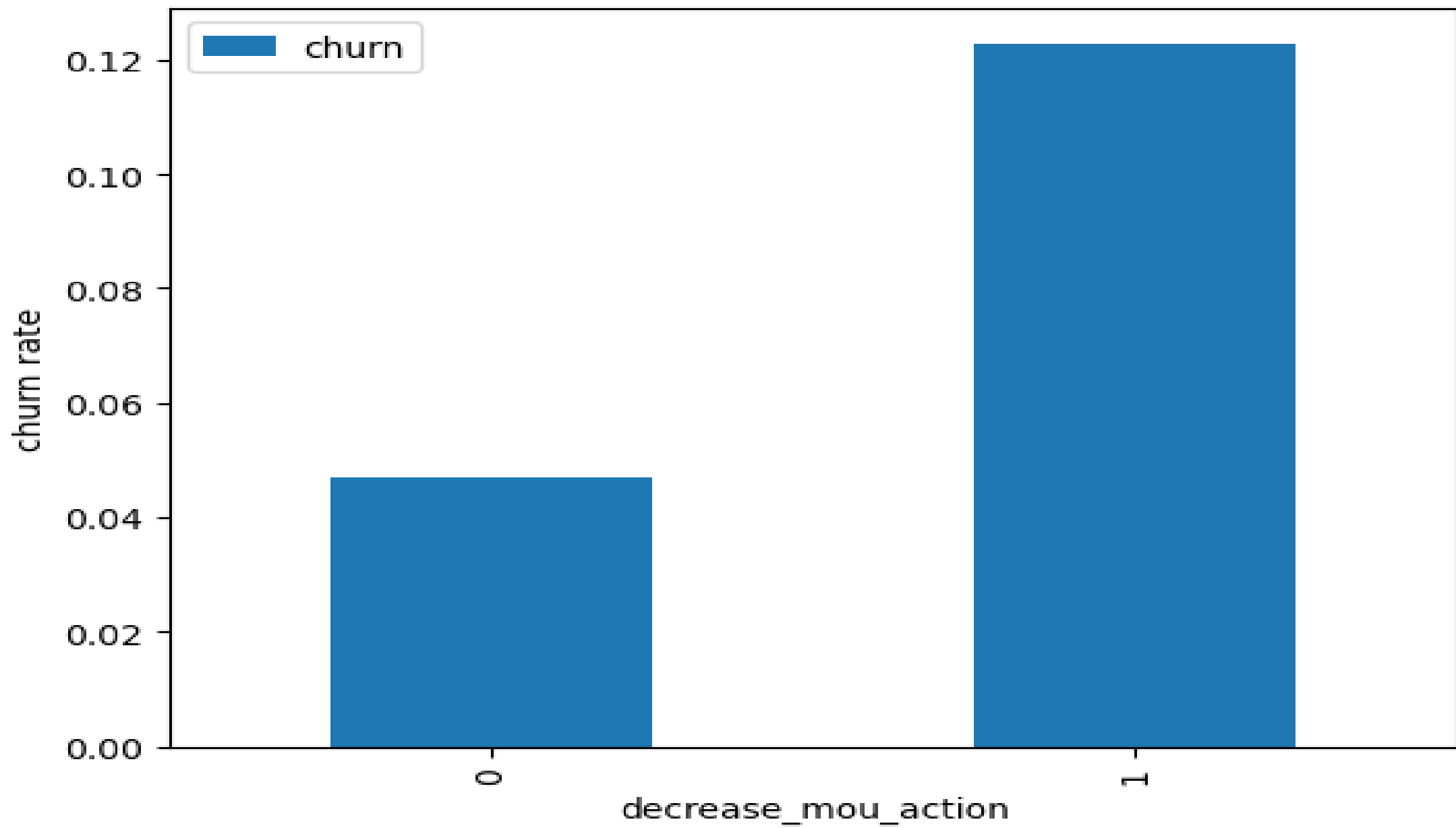
- In this highly competitive market, the telecommunications industry experiences an average of 15-25% annual churn rate.
- To reduce customer churn, telecom companies need to **predict which customers are at high risk of churn.**
- **Two ways of defining churn -1.Revenue-based churn:** Customers who have not utilised any revenue-generating facilities such as mobile internet, outgoing calls, SMS etc. over a given period of time. **2.Usage based churn-**Customers who have not done any usage, either incoming or outgoing - in terms of calls, internet etc. over a period of time.

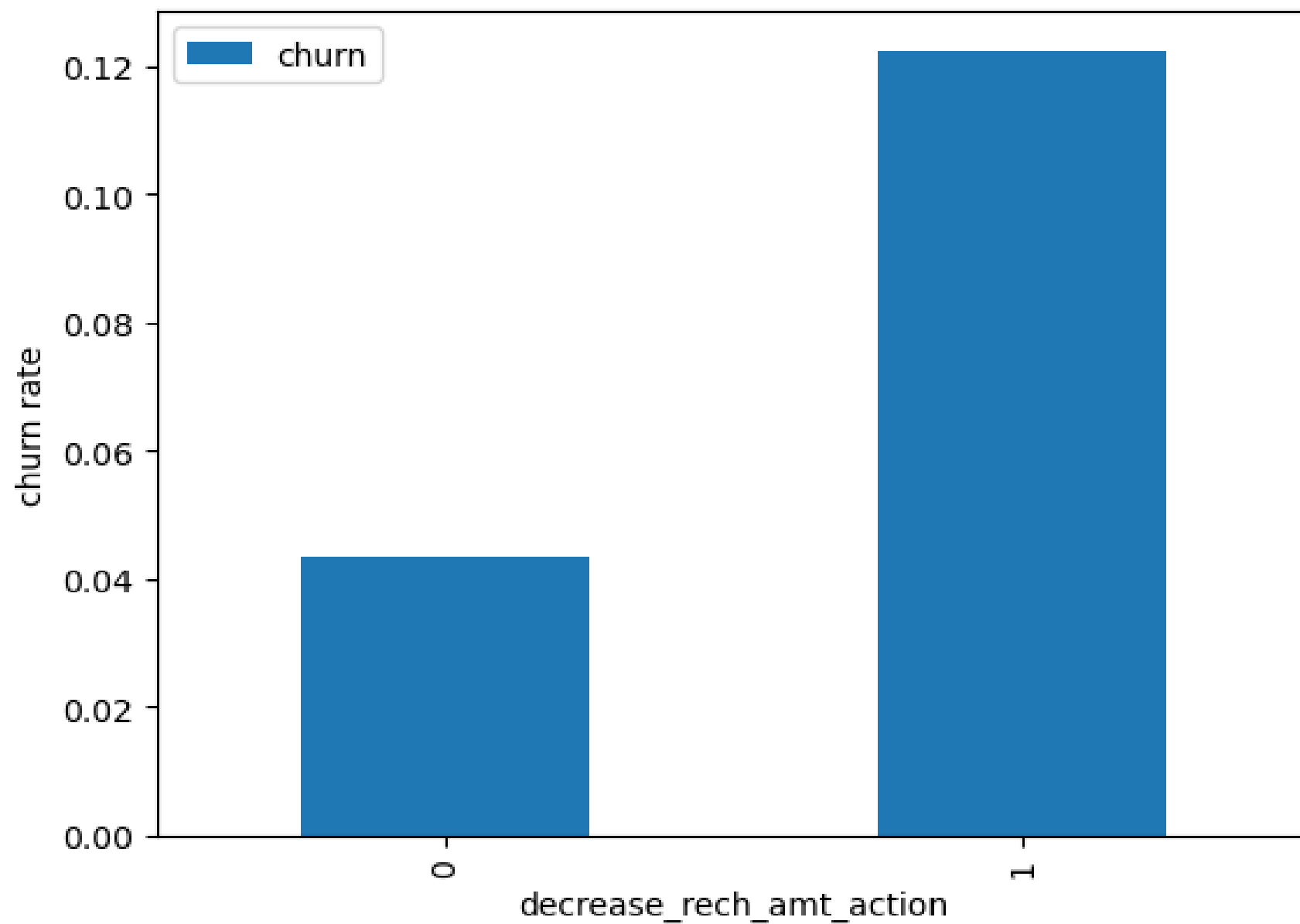
# Data Preparation

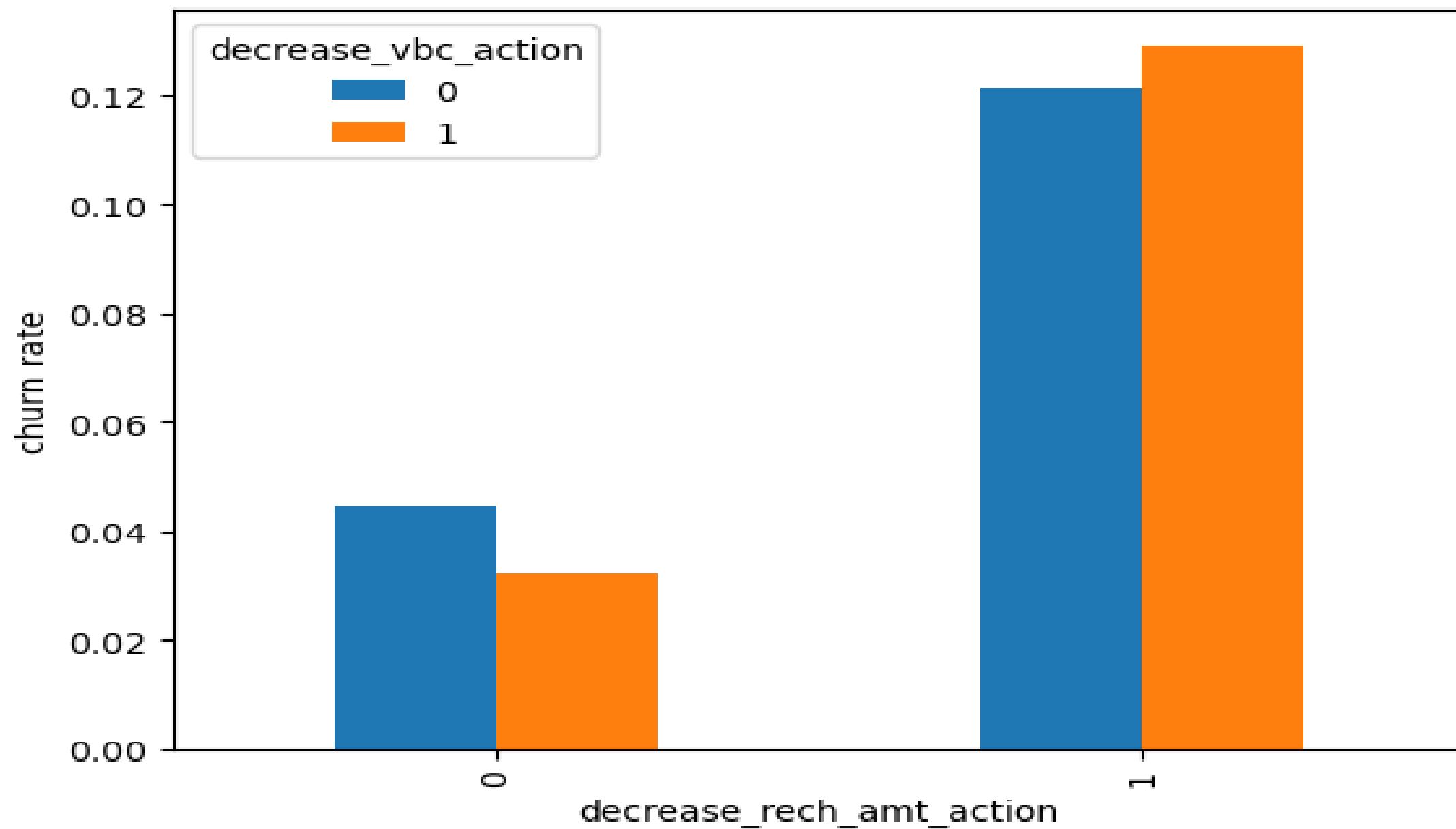
- **Filter high-value customers** -Those who have recharged with an amount more than or equal to X, where X is the **70th percentile** of the average recharge amount in the first two months (the good phase).
- **Tag churners and remove attributes of the churn phase**-Those who have not made any calls (either incoming or outgoing) AND have not used mobile internet even once in the churn phase.
- Build models to predict churn.

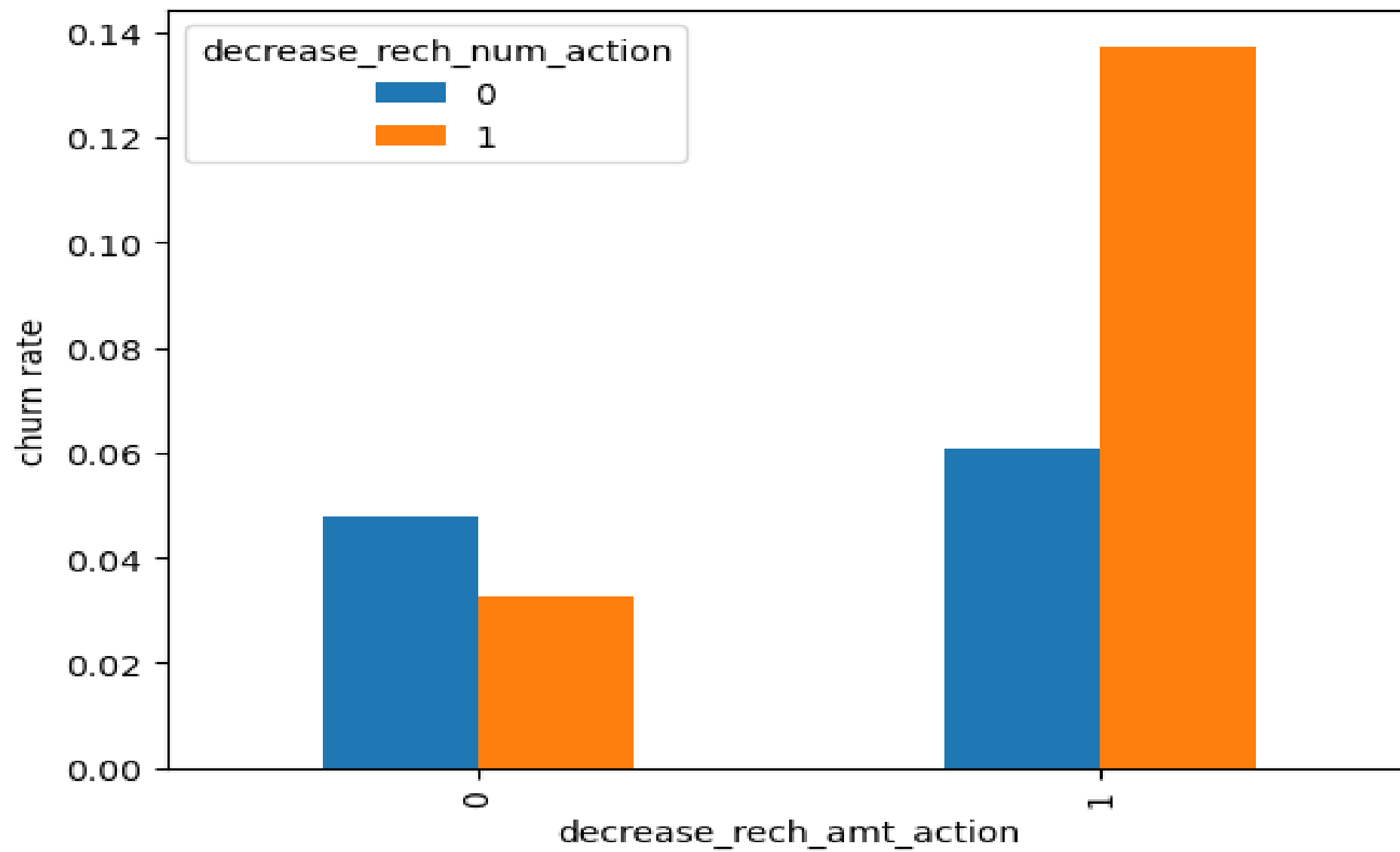
# EDA

- Cleaning Data and handling missing values.
- Insights from EDA are clearly derived and explained.
- An appropriate set of features is used to build the model.
- Class imbalance is handled using at least one of the techniques.
- The model output provides a probability score for the target variable.









# ANALYSIS

- Churn rate is more for customers whose minutes of usage decreased in the action phase than the good phase.
- Churn rate is more for customers whose number of recharge in the action phase is lesser than the number in good phase.
- The churn rate is more for the customers, whose volume based cost in action month is increased. That means the customers do not do the monthly recharge more when they are in the action phase.