

# Executive summary

## Situation :

Powerco is experiencing a loss of consumers, which they attribute to the sensitivity of its clients to price increases. Offering a discount of 20% to the customers who are most likely to start departing is one way.

## Complication :

- **Predictive Accuracy:** Due to the complex interaction of factors affecting consumers' activities, it may be difficult to identify customers likely to leave and give discounts to customers who wouldn't have left.
- **Unexpected Variables:** External factors may change client behaviour, making data-based estimates inaccurate.
- **Ethics:** Discounting select clients may raise problems of fairness, transparency, and discrimination among non-targeted consumers.
- **Revenue Impact:** The company's finances and growth are at risk if many people accept the deal.
- **Operational Challenges:** Targeted discounts need accurate identification, customised offers, and logistical changes, which complicate operations and may affect user experience.

## Machine Learning Modeling :

After Data cleaning, EDA and Feature engineering, I applied Random Forest Classifier. Random Forest Classifier model has been built to predict customers' churn probability, achieving an accuracy of 0.90 and Precision score of 0.85 on test set.

## Insights:

Based on the results above, we can see the performance is very bad. Although the accuracy is up to 90%, it is misleading and pointless, as we only focus on correctly predicting the positive class rather than the negative class. The model can only find out 6% of real positive samples. But luckily, in the predicted positive samples, up to 86% of them are real positive samples.