

# ISYE 6501 Course Project

## 1 SELECTED PROJECT

The project that was selected to be analyzed was "Turner Broadcasting uses their data to provide targeted advertising". This project was released by INFORMS (Institute for Operations Research and the Management Sciences). You can find the project [here](#).

## 2 PROJECT SUMMARY

Turner Broadcasting developed this project to improve the effectiveness of their TV ads by targeting specific audiences. By applying analytics models, they could identify distinct viewer groups, predict how each group would respond to different ads, and quickly adjust ad delivery. This data driven approach enabled them to serve more personalized ads, increase viewer engagement, and get a better return on their advertising investment.

## 3 MODELS USED

There are three different analytics models that may have been used in this project.

K-Means clustering could have been used to group viewers based on their behavior, like viewing times, preferred genres, and how often they watch content. This approach might have helped identify distinct viewer segments for targeted actions.

Logistic regression might have been used to predict the chances of a viewer interacting with or clicking on a particular advertisement. It would also offer clear probabilities, factoring in variables like demographics and viewing habits.

Random forests could have helped model complex, non-linear relationships between input features and ad responses. Additionally, they would work as a method to strengthen predictions and minimize overfitting.

## 4 HOW MODELS WERE COMBINED

The models could have been combined sequentially.

First, K-Means could have been used to create groups of similar users based on their behavior.

Second, Logistic Regression could have taken these K-Means-generated groups, along with other data fields, to predict the likelihood that users will engage with an ad.

Third, Random Forests could have refined the Logistic Regression predictions, along with other data fields, to improve accuracy and capture complex patterns in user behavior.

## **5 DATA NEEDED FOR THE MODELS**

Various different types of data would be needed to create the analytics models.

User behavior data would be needed. This data could include fields such as: Viewing history, watch time, time of day, and preferred genres.

Demographic data would also be needed. This data could include fields such as: gender, age, and location.

## **6 HOW DATA WAS COLLECTED**

Various methods could have been used to obtain the data needed for the models.

The TV providers likely already log all user behavior habits. The time a user begins watching, the time a user stops watching, and viewing history are all fields that would be logged. These could have been easily provided to Turner Broadcasting.

Similarly, when a user signs up with a TV provider, they likely already give them demographic data, such as age, gender, and location. These could have been easily provided to Turner Broadcasting.

## **7 HOW OFTEN MODELS NEED TO BE REFRESHED**

The three analytics models would likely have to be refreshed at different intervals.

The K-Means clustering model would likely have to be refreshed monthly. This is because user trends shift, but not as quickly as other data.

The Logistic Regression model and Random Forest model would likely have to be refreshed weekly or even daily. Engagement patterns shift frequently, and not refreshing the model could lead to inaccuracies.

## 8 CONCLUSION

Turner Broadcasting's advertising system demonstrates how combining multiple models can create a powerful, flexible solution. By using K-Means Clustering for grouping users, Logistic Regression for predictions, and Random Forests for identifying complex patterns, Turner Broadcasting was able to deliver more intelligent, personalized ads. This project highlights the value of integrating complementary models, supported by high-quality, regularly updated data, to generate real business impact in media and advertising.