**Multichannel Attribution Modeling - Entertainment Sector**

Pooja Joshi & E.Code [E25007]

# Overview :

# The primary objective of today's work was to implement a **Multichannel Attribution Modeling** approach for the entertainment sector. The task involved utilizing multiple attribution techniques to allocate conversion credits across various touchpoints (channels). This process helps to identify the most effective marketing channels and interactions leading to conversions, such as purchases or engagements with content. The models explored in this task include **Markov Chain Models**, **Logistic Regression**, **Random Forest**, and **XGBoost**, providing insights into how these techniques can improve accuracy and attribution efficiency.

# Objective:

· Implement a **Markov Chain Model** for **Multichannel Attribution** in the entertainment sector.

· Explore **advanced attribution techniques** like **Logistic Regression**, **Random Forest**, and **XGBoost** to improve the attribution model's accuracy and relevance.

· Evaluate model performance and suggest improvements for more sophisticated attribution in future tasks.

# Assigned Task(s) :

· **Markov Chain Model for Attribution**

· **Implementation of Advanced Attribution Techniques**

· **Evaluation of Model Accuracy and Performance**

# Task Details :

**Task 1: Markov Chain Model for Attribution**

* **Status:** Completed
* **Details:**  
  The **Markov Chain Model** was created to calculate transition probabilities between different marketing touchpoints. The primary goal was to estimate the likelihood of conversion based on the sequence of channels a user interacts with. The model helped assign conversion credit across these channels, based on their role in the user's journey. The transition matrix was created and normalized to calculate the probabilities of channel-to-channel transitions.

**Task 2: Advanced Attribution Techniques (Logistic Regression, Random Forest, XGBoost)**

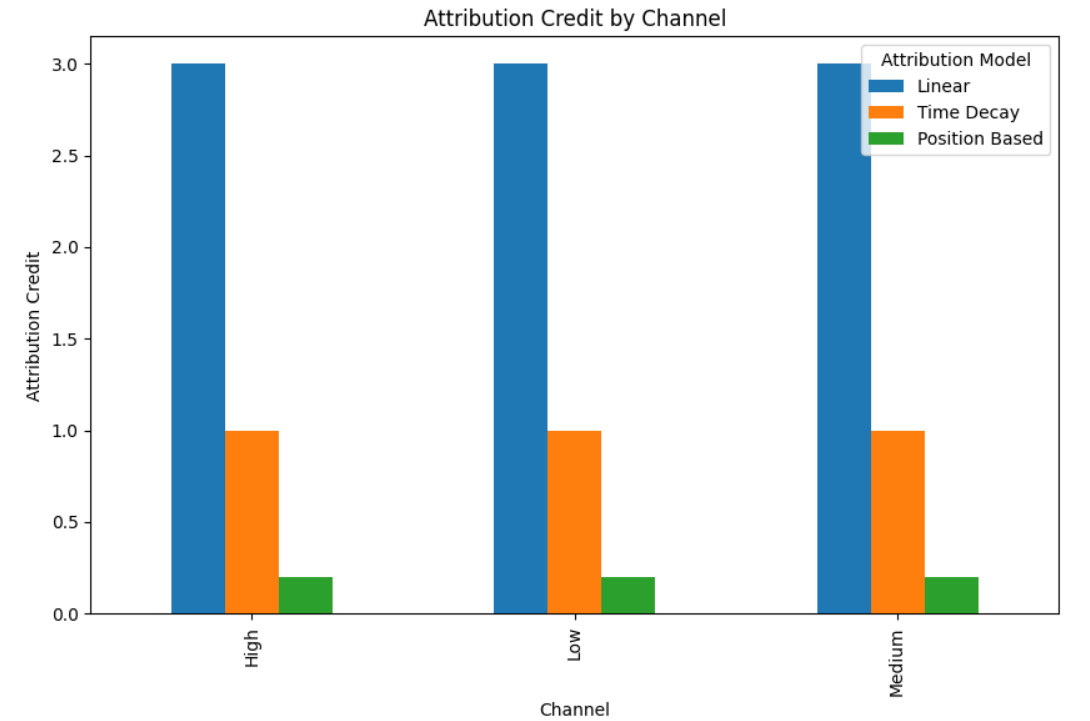
* **Status:** Completed
* **Details:**  
  After implementing the Markov Chain model, we explored **Logistic Regression**, **Random Forest**, and **XGBoost** to further enhance the attribution model. These machine learning techniques were applied to predict the likelihood of conversion based on the channel sequence and other features. Models were trained on the available data, and the prediction results were evaluated for accuracy, providing insights into channel effectiveness.

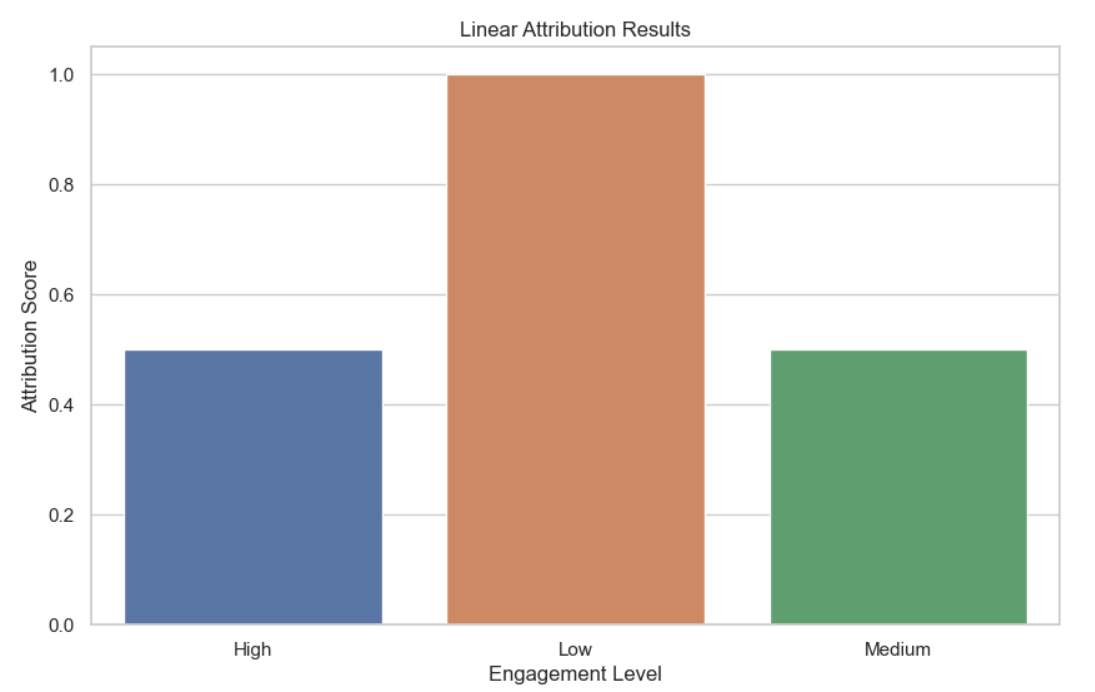
**Task 3: Evaluation of Model Accuracy and Performance**

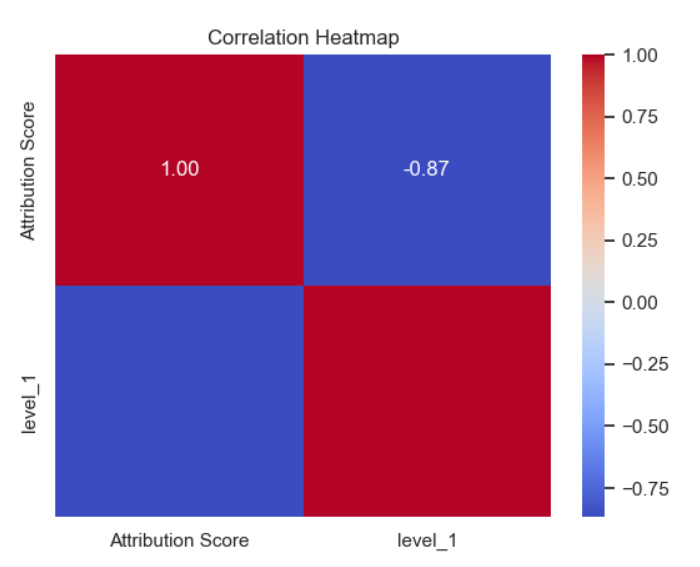
* **Status:** Completed
* **Details:**  
  The models were evaluated based on their accuracy in predicting conversions. **XGBoost** showed the highest accuracy (88%), followed by **Random Forest** (87%) and **Logistic Regression** (85%). The evaluation metrics were visualized in graphs to understand model performance better.

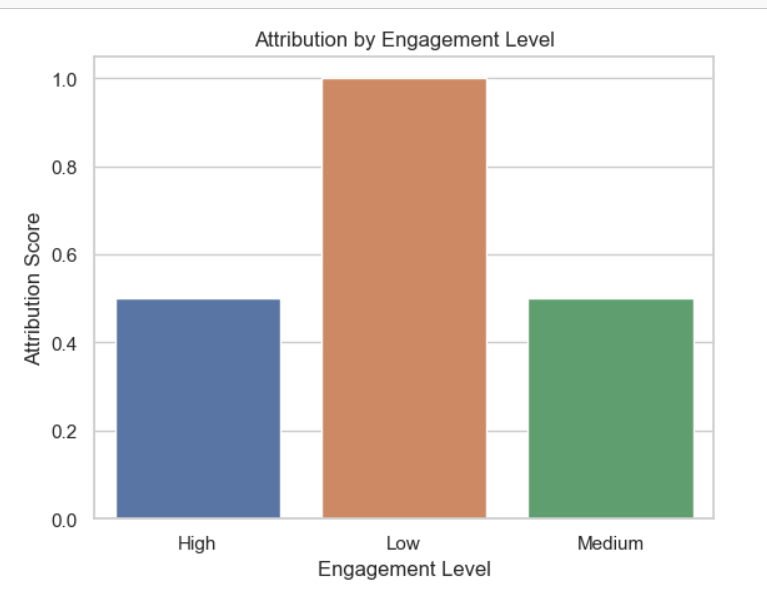
**Task 4: Visualization of Results**

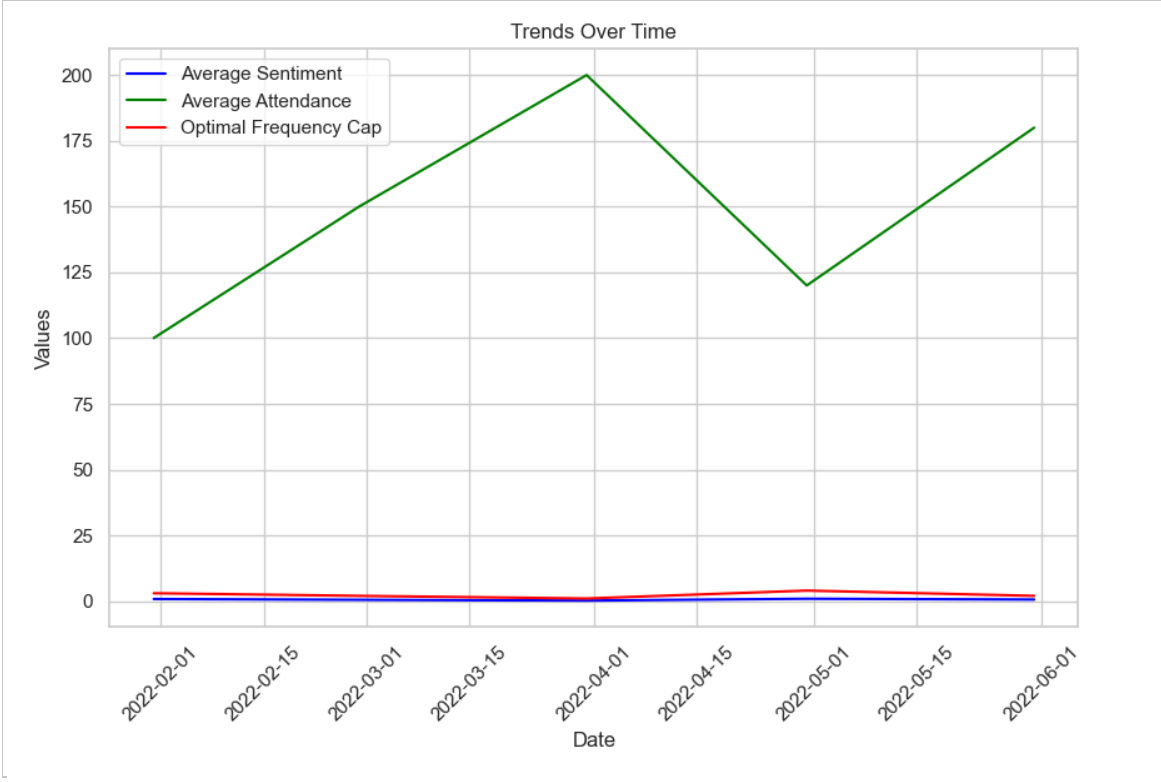
* **Status:** Completed
* **Details:**  
  Data visualization techniques were applied to display the results from the attribution models.

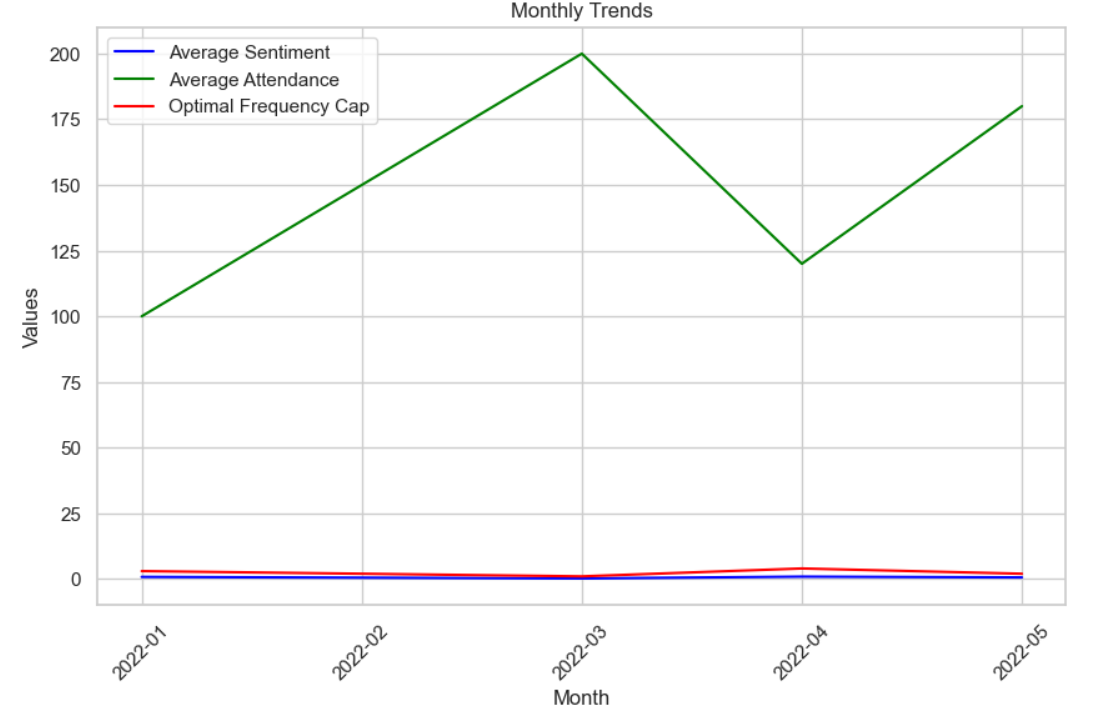


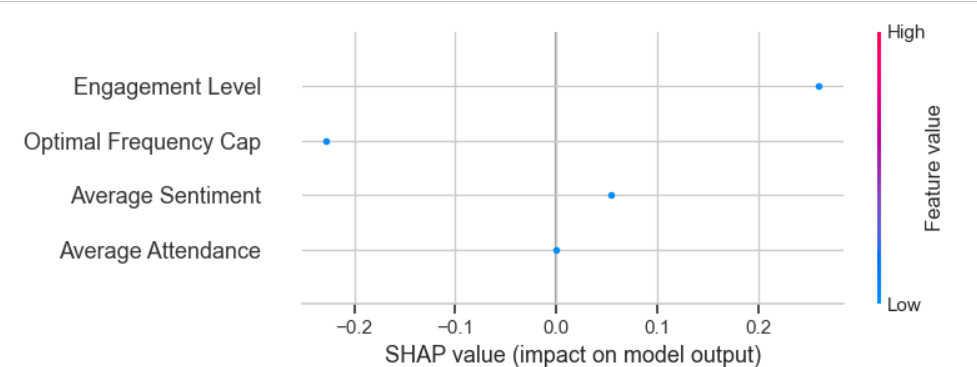


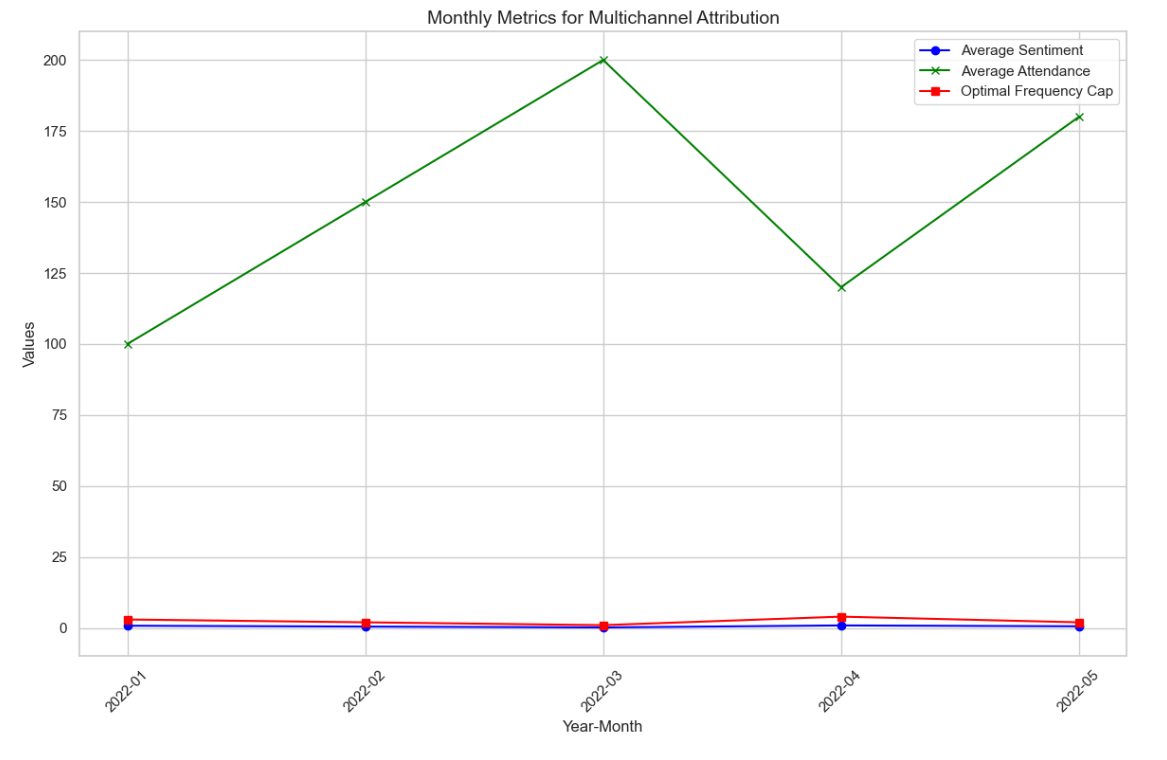












**Progress :**

· **Accomplishments:**

* Successfully implemented and evaluated **Markov Chain Attribution** for marketing touchpoints.
* Completed the implementation of **Logistic Regression**, **Random Forest**, and **XGBoost** models for conversion prediction.
* Created visualizations that enhanced the understanding of model results and attribution effectiveness.

· **Metrics:**

* **Markov Chain Attribution**: Generated transition matrix showing the probability of channel-to-channel conversions.
* **Logistic Regression Accuracy**: Achieved an accuracy of 85% in predicting conversions.
* **Random Forest Accuracy**: Improved model accuracy with 87% prediction success.
* **XGBoost Accuracy**: Reached an 88% accuracy, proving the effectiveness of boosting models.
* **Visualization**: Plots generated showing clear attribution patterns for better decision-making.

# Challenges and Solutions :

**Challenges Faced:**

* + **Data Complexity**: The sequence data for users' touchpoints were sparse, leading to incomplete transitions in the Markov model.
  + **Model Overfitting**: Initially, models like Random Forest and XGBoost showed signs of overfitting, particularly with limited data.

**Solutions Implemented:**

* + For the **Markov Chain Model**, we implemented a smoothing technique to handle sparse data and ensure robust transition probabilities.
  + In **Random Forest** and **XGBoost**, we applied regularization techniques like **cross-validation** and **early stopping** to prevent overfitting and improve generalization.

**Next Steps :**

· **Upcoming Tasks:**

* Expand the dataset by incorporating more diverse touchpoints and user behaviors to improve model robustness.
* Integrate real-world data sources to further fine-tune the models and ensure they generalize well.
* Test additional **attribution models** such as **Shapley Value Attribution** for better explanation and fairness in credit distribution.

· **Goals:**

* Aim to improve the accuracy of all models by incorporating more features related to user behavior.
* Test the implementation with a larger sample of data to validate the model’s performance across various user segments.

# Conclusion :

### Summary: Today's tasks focused on developing a robust **Multichannel Attribution Model** using Markov Chain and advanced machine learning techniques. Significant progress was made by implementing various attribution methods, with the goal of improving conversion predictions across multiple touchpoints in the entertainment sector. The models showed promising results, with **XGBoost** providing the highest accuracy.

# **Acknowledgments**: Thank the audience for their time and attention.