Report Notes

Note down any significant findings or research to help with application development. This note is to be updated in real time as the days go by.

ADD YOUR NOTES TO YOUR SECTION ONLY!

Notes will be useful in building a finely documented project timeline which will make the report writing much easier at the end (we won't have to write the report from scratch).

Make sure to add a date to any new notes add!

Hephzibah's Section

22/01/2024

Three different potential Methods of approaches for the coursework...

Time Series Analysis

Description: Time series analysis involves examining the data to identify patterns over time. For event ticket sales, this could involve looking at trends, seasonal patterns, or cyclical fluctuations in past booking data.

Methodology: Models like ARIMA (Autoregressive Integrated Moving Average), SARIMA (Seasonal ARIMA), or more advanced machine learning models like LSTM (Long Short-Term Memory) networks can be used. These models can help forecast future ticket sales based on historical booking patterns, considering factors like time of year, day of the week, or proximity to major holidays.

Data Requirements: This method requires a time-stamped dataset of ticket sales, ideally spanning multiple events over a period of years to capture seasonal trends and long-term patterns.

Regression Analysis

Description: Regression analysis is used to predict a continuous outcome (ticket sales) based on one or more predictor variables.

Methodology: Linear regression could be a starting point but given the likelihood of non-linear relationships within the data, methods like polynomial regression, ridge regression, or lasso regression might be more appropriate. These models can incorporate various factors, such as event type, day of the week, marketing efforts, and historical attendance rates, to predict future sales.

Data Requirements: Detailed historical data on ticket sales, including factors like event characteristics (type, date, location), marketing efforts (advertising spend, promotional activities), and external factors (weather, competing events) are needed.

Classification Models for Cancellation Prediction

Description: Predicting ticket cancellations can directly impact the accuracy of ticket sales forecasts. By understanding the likelihood of cancellations, one can better estimate actual attendance.

Methodology: Machine learning classification models like logistic regression, random forests, or gradient boosting machines (like XGBoost) can be used to predict the probability of a booking being cancelled. Features can include historical cancellation rates, time between booking and event, customer demographics, and past booking behaviours.

Data Requirements: This approach requires detailed historical booking data, including whether each booking was cancelled or not, along with any available customer information and booking details (such as booking date relative to the event date).

Possible Q's for client:

- What date in each year did the event land on? (can be used to calculate time between ticket bookings and event date which gives us more data to work off for making predictions).

Using Classification models for cancellation prediction is most sensible because it is scalable and can be continuously improved by adding more data each year.

Hussnain's Section

Khadija's Section

Minoli's Section

Obaid's Section

Tebuho's Section

Waqar's Section
