**ACI Assignment 2 – Q7 - Alexa**

**Assignment Group – 75**

|  |  |  |
| --- | --- | --- |
| **Name** | **BITS Id** | **Contribution** |
| Avi Krishna Srivatsava | 2020FC04492 | 100% |
| Dola Tejesh | 2020FC04459 | 100% |
| Kommajyosula VNS Kanth | 2020FC04120 | 100% |
| Nareshkumar P | 2020FC04122 | 100% |
| T Navya Madhavi | 2020FC04007 | 100% |

**Alexa Dataset:**

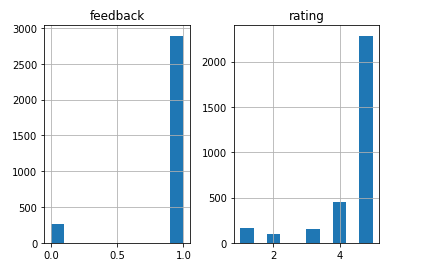
**Python Code:**

**Given Features:**

|  |  |
| --- | --- |
| **Feature Name** | **Type** |
| Rating | Categorical |
| Date (Converting to month data) | Categorical |
| Variation | Categorical |
| Feedback (Class Feature) | Categorical |

**Data Exploration and Visualization:**

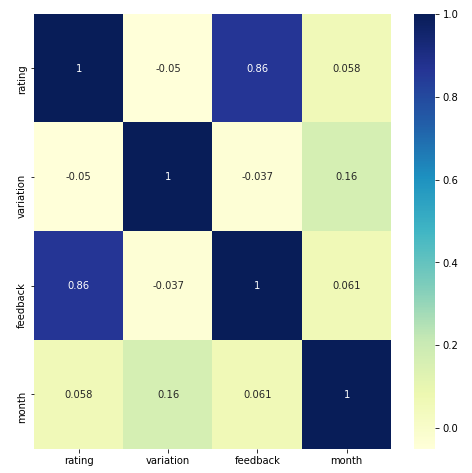
1. No Null values/noisy data present in the given dataset.
2. No outliers found in the dataset.
3. 3150 rows and 4 columns present in the dataset.
4. Assigned the unique values to the each and every value in variation feature.
5. More number of Positive feedback records present when compared to Negative feedback.
6. More number of high rated review records present when compared to the low rated review records.



1. **Changing Date column to month:**

Converted date fields to month by using below command. It will be helpful to aggregate data based on the month.

data['month'] = pd.DatetimeIndex(data['date']).month



Based on the above heat map, we are able to understand that Feedback and Rating is highly correlated and having high dependency between them.

**Calculating Conditional Probability for the features:**

Calculating the conditional probability of each and every combination of features against the feedback feature and analyze the observations to build the tree for Bayesian model.

Calculated conditional probability table for the below features combinations.

Variation vs Feedback

Month vs Feedback

Rating vs Feedback

Variation, Month vs Feedback

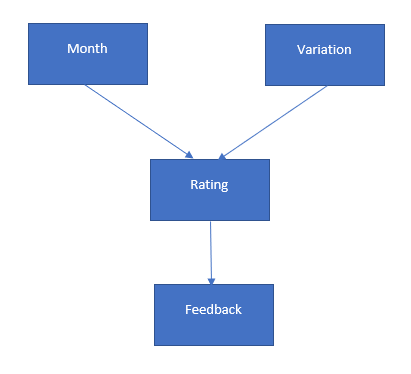
Variation, Rating vs Feedback

Month, Rating vs Feedback

Month, Rating, Variation vs Feedback

**2. Bayesian Model creation:**

Created a Bayesian model based on the below structure.



model = BayesianModel([("variation","rating"),("month","rating"),("rating","feedback")])

Rating and Feedback features are highly dependent hence both are directly connected. Different possible ratings/feedbacks are having for the different month and variation features hence they are directly connected to the rating feature.

**3. Infering the Probability for the data:**

|  |  |  |  |
| --- | --- | --- | --- |
| Rating | Month | Variation | Feedback |
| 4 | Jul | Black Plus (Alias value is 12) | 1 |

**Results:**

infer = VariableElimination(model)

print(infer.query(['feedback'],evidence={'rating': 4,'month': 7,'variation': 12}))

+-------------+-----------------+

| **feedback | phi(feedback)** |

+=============+=================+

| feedback(0) | 0.0000 |

+-------------+-----------------+

| feedback(1) | 1.0000 |

+-------------+-----------------+

It exactly identified the feedback is positive as expected.

**4. Check the Probability of rating below (1 or 2) for the data:**

|  |  |  |
| --- | --- | --- |
| Month | Variation | Feedback |
| July | Charcoal Fabric (Alias value is 1) | 0 |

print(infer.query(['rating'],evidence={'variation': 1,'month': 7,'feedback': 0}))

**Result:**

+-----------+---------------+

| **rating | phi(rating)** |

+===========+===============+

| rating(1) | 0.3333 |

+-----------+---------------+

| rating(2) | 0.6667 |

+-----------+---------------+

| rating(3) | 0.0000 |

+-----------+---------------+

| rating(4) | 0.0000 |

+-----------+---------------+

| rating(5) | 0.0000 |

+-----------+---------------+

For the given data,

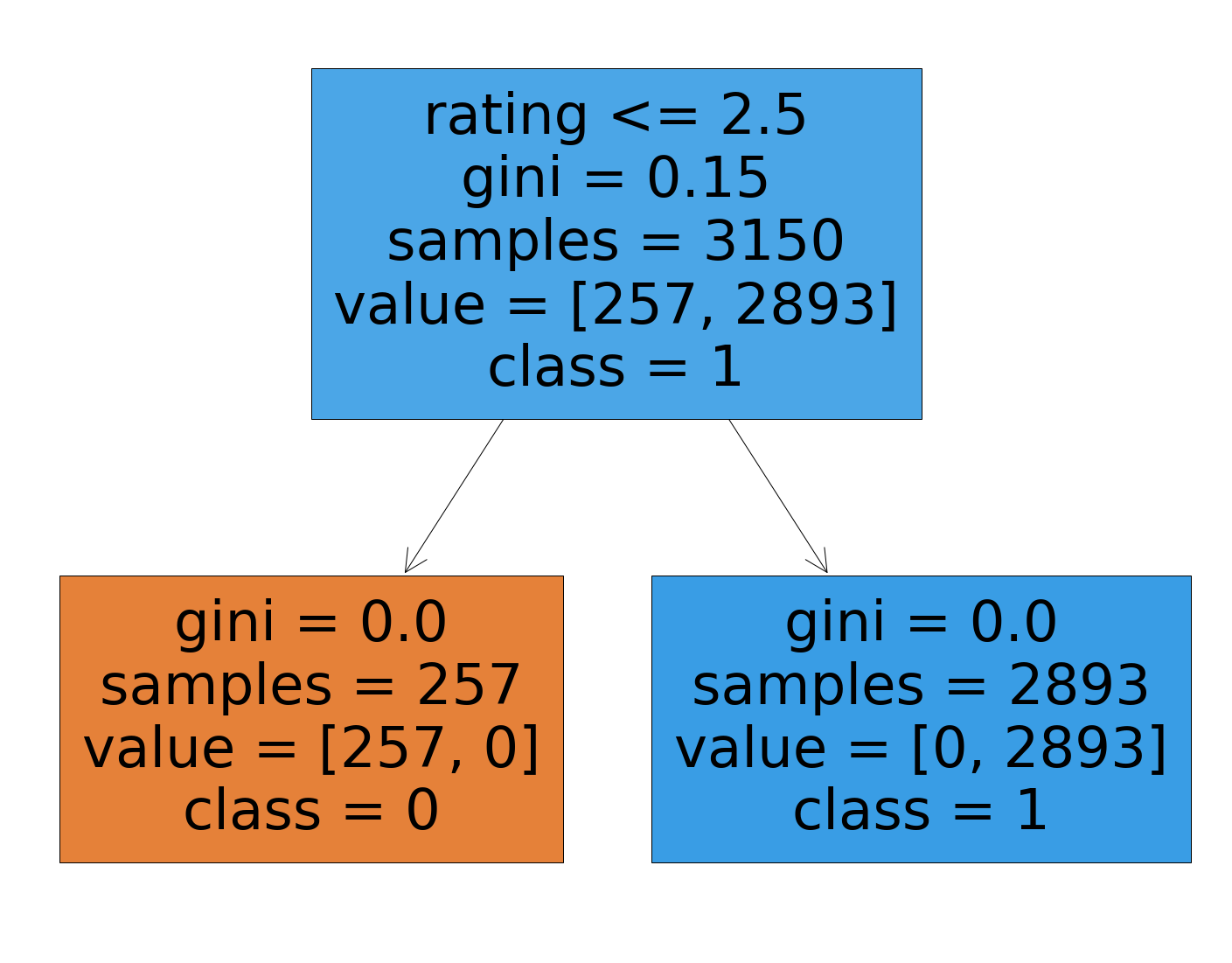
Rating 1 -> 33% Possible (Probability of Rating 1)

Rating 2 -> 67% Possible (Probability of Rating 2)

**Prolog Code:**

1. **Build Decision tree:**

Build the decision tree for the given dataset using decision tree algorithm. Please find the structure below.



**2. Rules for the Decision tree:**

Based on the decision tree results, we are created below rules in the prolog code.

If Rating =1 and Date in (list\_of\_months) and Variation in (list\_of\_variations) then Feedback=0 (Negative Feedback)

If Rating =2 and Date in (list\_of\_months) and Variation in (list\_of\_variations) then Feedback=0 (Negative Feedback)

If Rating =3 and Date in (list\_of\_months) and Variation in (list\_of\_variations) then Feedback=1 (Positive Feedback)

If Rating =4 and Date in (list\_of\_months) and Variation in (list\_of\_variations) then Feedback=1 (Positive Feedback)

If Rating =5 and Date in (list\_of\_months) and Variation in (list\_of\_variations) then Feedback=1 (Positive Feedback)

**3. Rules Code:**

is\_feedback\_positive(X,Y,Z) :- rating(X) = rating(3), member(Z,[january,february,march,april,may,june,july,august,september,october,november,december]), member(Y,[CharcoalFabric,WalnutFinish,HeatherGrayFabric,SandstoneFabric,OakFinish,Black,White,BlackSpot,WhiteSpot,BlackShow,WhiteShow,BlackPlus,WhitePlus,Configuration:FireTVStick,BlackDot,WhiteDot]).

is\_feedback\_positive(X,Y,Z) :- rating(X) = rating(4), member(Z,[january,february,march,april,may,june,july,august,september,october,november,december]), member(Y,[CharcoalFabric,WalnutFinish,HeatherGrayFabric,SandstoneFabric,OakFinish,Black,White,BlackSpot,WhiteSpot,BlackShow,WhiteShow,BlackPlus,WhitePlus,Configuration:FireTVStick,BlackDot,WhiteDot]).

is\_feedback\_positive(X,Y,Z) :- rating(X) = rating(5), member(Z,[january,february,march,april,may,june,july,august,september,october,november,december]), member(Y,[CharcoalFabric,WalnutFinish,HeatherGrayFabric,SandstoneFabric,OakFinish,Black,White,BlackSpot,WhiteSpot,BlackShow,WhiteShow,BlackPlus,WhitePlus,Configuration:FireTVStick,BlackDot,WhiteDot]).

**4. Test Results:**

After giving different values for the rating, date and variation as input, we got below results as expected.

?- myinput.

Enter DetailsEnter Rating: <1/2/3/4/5>?4.

Enter Month: <january/februaury/march/.../december>?|: july.

Enter variation: <CharcoalFabric/WalnutFinish/HeatherGrayFabric/SandstoneFabric/OakFinish/Black/White/BlackSpot/WhiteSpot/BlackShow/WhiteShow/BlackPlus/WhitePlus/ConfEnter variation: White.

**true .**

?- myinput.

Enter DetailsEnter Rating: <1/2/3/4/5>?3.

Enter Month: <january/februaury/march/.../december>?|: august.

Enter variation: <CharcoalFabric/WalnutFinish/HeatherGrayFabric/SandstoneFabric/OakFinish/Black/White/BlackSpot/WhiteSpot/BlackShow/WhiteShow/BlackPlus/WhitePlus/ConfEnter variation: BlackSpot.

**true .**

?- myinput.

Enter DetailsEnter Rating: <1/2/3/4/5>?2.

Enter Month: <january/februaury/march/.../december>?|: september.

Enter variation: <CharcoalFabric/WalnutFinish/HeatherGrayFabric/SandstoneFabric/OakFinish/Black/White/BlackSpot/WhiteSpot/BlackShow/WhiteShow/BlackPlus/WhitePlus/ConfEnter variation: BlackShow.

**false.**

?- myinput.

Enter DetailsEnter Rating: <1/2/3/4/5>?1.

Enter Month: <january/februaury/march/.../december>?|: december.

Enter variation: <CharcoalFabric/WalnutFinish/HeatherGrayFabric/SandstoneFabric/OakFinish/Black/White/BlackSpot/WhiteSpot/BlackShow/WhiteShow/BlackPlus/WhitePlus/ConfEnter variation: WhiteDot.

**false.**