CAPSTONE PROJECT SUBMISSION (TELECOM CHURN)

MEHAK SETHI (JIG19401)

APPROACH FOLLOWED

- Generating Data Quality Report
- Dropping variables with more than 15% missing values
- Variable Type Conversions (some variables were read in numeric form but were actually factor, so converted them)
- EDA of continuous and categorical variables
  - Continuous Variables
    - Decile binning viz a viz Churn and dropping variables with less than 4 deciles
    - Created two dummy variables and deciled them
  - Categorical Variables
    - Calculated event rate for variables and profiled them
- Exported the continuous variables and categorical variables into separate tables.

APPROACH FOLLOWED (CONTD.)

- Treated Outliers
- Missing value treatment – Deletion, Mean Imputation and Separate Category creation for Categorical variables
- I tried Factor Analysis approach but because of explainability issues I dropped it.
- Built a Logistic Regression model to shortlist significant variables and significant levels of categorical variables.
- On the basis of the Logistic Regression Summary, created dummies for significant levels of categorical variables.
- Created another iteration of the model with significant variables including dummies created previously
- Checked summary to shortlist significant variables. Created another iteration.

APPROACH FOLLOWED (CONTD.)

- Finally the 7th iteration had all significant variables and VIFs <4. Finalised this model.
- Obtained predictions on the final model.
- Arrived at a Kappa Value of 0.2 based on Kappa Value.
- Transformed the predictions to binary values based on the cutoff value.
- Model Validation

Performance Metrics	Train Data	Test Data
Accuracy	70.33%	71%
AUC	59.06%	60.57%
True Positives	2618	1095

BUSINESS QUESTION 1

- What are the top five factors driving likelihood of churn at Mobicom?
  - The top factors leading to churn (in Final Model) are as below:
  - These are arranged according to VarImp and Beta Coefficients

Variable Name	Importance Score (according to VarImp)	Variable Name	According to Beta Coefficient (abs value)
Eqpdays	18.104	Ethnic_C	1.0549 (-)
Retdays_1	12.754	retdays_1	0.7340
months	10.247	Ethnic_Z	0.5259 (-)
Rev_Range	8.493	Unq_5	0.4083
Asl_flag_Y	8.173	models4	0.3303

Since VarImp yields normalized scores and doesn't have negative values, so Beta Coefficient values are being considered for the analysis as they give a complete picture of direction and magnitude of impact.

BUSINESS QUESTION 1 (CONTD.)

According to Beta Coefficients, the following factors are of peak importance

- Ethnic_C **– a unit increase in** Ethnic_C **leads to** 105.49% **decrease in Churn (beta coefficient -**1.0549**) [ethnic_C = Ethnicity code C]**
- **Retdays_1 – a unit increase in retdays_1 leads to 73.4% increase in churn (beta coefficient 0.73406) [retdays_1 = Number of days since last retention call]**
- Ethnic_Z **– a unit increase in** Ethnic_Z **leads to** 52.59% **decrease in churn (beta coefficient -0.5259) [ethnic_C = Ethnicity Code Z]**
- Unq_5 **– a unit increase in** Unq_5 **leads to** 40.83% **increase in churn (beta coefficient** 0.4083**) [unq_5 = Family of 5 unique subscribers]**
- models4 **– a unit increase in** models4 **leads to 33.03% increase in churn (beta coefficient 0.3303) [models4 = number of models issued]**

BUSINESS QUESTION 2

- Validation of survey findings. a) Whether "cost and billing" and "network and service quality" are important factors influencing churn behaviour. b) Are data usage connectivity issues turning out to be costly? In other words, is it leading to churn?

- 2.a. Factors comprising of cost and billing and their coefficients (in Final Model)

Variable Name	Beta Coefficient	Impact
Totmrc_Mean (Mean total monthly recurring charge)	-0.00277	1 unit increase of totmrc_Mean decreases churn by 0.277%
Rev_Range (Range of revenue – charge amount)	**0.00139**	1 unit increase of Rev_Range increases churn by 0.139%
Asl_Flag_Y (Account spending limit = Yes)	-0.3067	1 unit increase in Account Spending limit decreases churn by 30.67%

BUSINESS QUESTION 2 (CONTD.)

- Factors comprising Network and service quality

Variable Name	Beta Coefficient	Impact
Change_mou (Percantage change in monthly minutes of use compared to the average of last 3 months)	-0.00025	I unit increase in change_Mou decreases churn by 0.025%
Drop_blk_Mean (Mean number of dropped/ blocked calls)	0.0033	I unit increase in drop_blk_Mean increases churn by 0.33%
Drop_vce_Range (Range of number of dropped/ failed voice calls)	0.0079	I unit increase in drop_vce_Range increases churn by 0.79%
Retdays_I (Number of days of last retention call)	0.7340	I unit increase in retdays_I increases churn by 73.40%
Callwait_Mean (Mean number of waiting calls)	-0.0056	I unit increase in Callwait_Mean decreases churn by 0.56%
Plcd_attempt_Mean (Mean number of attempted voice and data calls)	-0.0006	I unit increase in plcd_attempt_Mean decreases churn by 0.06%

BUSINESS QUESTION 2 (CONTD.)

- 2.b. Are data usage connectivity issues turning out to be costly?

  - In the final model, there was only one variable "drop_blk_Mean" which includes two data variables, viz a viz, "blck_dat_Mean" and "drop_dat_Mean".
  - Drop_blk_Mean had a beta coefficient of 0.0033 which indicates that 1 unit increase in drop_blk_Mean increases churn by 0.33% which is not that big a number.
  - All the data related variables got dropped in the earlier models because of insignificance/ multicollinearity even after multiple iterations of including them separately.
  - Thus, it is safe to conclude that data usage connectivity does not have a significant impact on churn.

BUSINESS QUESTION 3

- Would you recommend rate plan migration as a proactive retention strategy?

Factors explaining rate plan migration are:

- Totmrc_Mean
- Change_mou

Totmrc_Mean – Beta coefficient -0.00277 (1 unit increase in totmrc_Mean decreases churn by 0.277%)
Change_Mou – Beta Coefficient -0.00025 (1 unit increase in change_mou decreases churn by 0.025%)

Considering the extremely small beta coefficient values, it is safe to conclude that rat plan migration might not bring out significant decrease in churn. So it is not advisable to adopt this strategy.

INSIGHTS SUMMARY WITH SOME RECOMMENDATIONS

- Higher the retdays_1 (number of days since last retention call), 73.40% more chances of churn. **The company should aim to minimise the gap between two subsequent retention calls in order to improve their customer retention**. (Please note that this variable has a lot of missing value imputations so a lot of importance should not be given here)
- Churn decreases in Ethnicities Z (**African American**) and C (not given)100% (~105.49%) and 52% respectively.These ethnicities are **loyal customers** and very likely to stick with the company.
- Families with **5 unique subscribers** are 40.83% more likely to churn.The company should do in depth analysis to find the cause of this **behaviour** and try to close the gap.
- People with **4 models** issued are 33% more likely to churn.The company should find out more about this category.
- Those customers who have **account spending limits** enabled are 30% more likely to stick to the company.The company should try to **give account spending limit**s to those customers who are not availing it at present, this could help increase retention of customers.
- **Contrary to the belief that poor services or network or billing/ cost issues are the causes, it seems like the demographics of the customer base are the key reasons for customer churn.This proves that the company's offerings are not the problem.**
- **So the company should try to find out ways to appeal to more ethnicities, families, models. On the product and service side it should give more account spend limits and retention calls, these factors although are on the service side but give insights into the consumer behaviour, consumers want credit (maybe they are working professionals who get their telecom bills reimbursed and then pay) and retention calls (maybe they feel less valued because even though they have a problem but since the company didn't follow up often by not calling them sooner).**

BUSINESS QUESTION 4

- What would be your recommendation on how to use this churn model for prioritisation of customers for a proactive retention campaigns in the future?

- Churn model can be used to segment customers on their churn likeliness (probability of churning).
- The customers who have high churning probabilities can be targeted for retention strategies. This way Mobicom can be successful in retaining customers. After all retaining customers is easier and more cost effective as compared to acquiring new customers.
- However, based on the insights derived from this analysis, I would advise the company to analyse the demographics, spending habits, credit history etc on the consumer behaviour part. If the company can crack that code, then it can effectively improve retention.

BUSINESS QUESTION 5

- What would be the target segments for proactive retention campaigns? Falling ARPU forecast is also a concern and therefore, Mobicom would like to save their high revenue customers besides managing churn. Given a budget constraint of a contact list of 20% of the subscriber pool, which subscribers should prioritized if "revenue saves" is also a priority besides controlling churn. In other words, controlling churn is the primary objective and revenue saves is the secondary objective.
- Based on Test Data (19416 observations), the matrix looks like:

Revenue (Low, Medium, High)				
Churn (Low, Medium, High)		Low Revenue	Medium Revenue	High Revenue
	Low Churn	1036	2080	2772
	Medium Churn	2847	5682	**4980***
	High Churn	0	**4***	**15***

Note:
The target segment is marked in bold and indicated by star mark.
Here, low churn is <0.20, medium churn >=0.20 to <0.60, high churn is >0.60
Here, low revenue is <455.890, medium revenue >=455.890 to <949.9, high revenue >949.9

FINAL WORDS

- The target segment as given in the previous slide should be the immediate priority for the company. Losing them to the competition will result in major revenue loss for the comnay which will impact its growth adversely thus opening the floodgates to pleothra of problems like share value, liquidity problems and even insolvency (worst case scenario).
- I would highly urge the company to look into the explore factors like ethnicity, credit limit (like OLA gives OLA Money), retention calls (make them feel valued), households with 5 members (conduct survey to find problems), 4 issued models (study).
- The key to retention lies in the emotional appeal here according to the data at hand.

Thank You Jigsaw Academy for this project. It did open the horizons of my thinking, logic, coding and interpretation for insights generation. ☺