Decoding Student Retention and Churn of Vodafone (Telecel) in KNUST

A Survival Analysis Approach

Musah Faridu Oubda
Kassim Asana
Sarpong Linda
Torsi Edmond Collins
Asiamah Ezekiel

Kwame Nkrumah University of Science and Technology

July 28, 2024

OUTLINE

1. Introduction
2. Methodology
3. Analysis and Findings
4. Conclusion

BACKGROUND OF STUDY

- The Ghanaian telecommunications industry, particularly Vodafone (now Telecel), faces significant challenges with customer churn. Retaining customers is crucial for profitability, especially in a highly competitive market where acquiring new customers is more expensive than retaining existing ones.
- Telecel, which acquired Vodafone in early 2023, aims to enhance service offerings and improve customer retention. The study focuses on understanding and addressing student churn at KNUST using survival analysis methods to develop strategies for reducing churn and improving customer satisfaction.

PROBLEM STATEMENT

- Student churn persists due to a lack of understanding of the factors driving churn and retention.
- This research aims to identify these factors and develop strategies to improve retention rates using survival analysis models.

RESEARCH OBJECTIVES

- Main Objective: Analyze student retention and churn using survival analysis
- Specific Objectives:
  - Identify factors influencing churn
  - Analyze demographic patterns
  - Evaluate churn rates for different services
  - Assess impact of network quality on retention

METHODOLOGY

- The data was collected via a survey, capturing specific aspects relevant to the study while ensuring confidentiality and ethical consideration
- The sample size was determined through cluster sampling, targeting approximately 768 students from a population of about 85,000, with each of the 6 clusters representing different colleges within KNUST having around 128 students each.

METHODOLOGY

- The data preprocessing involved examining the dataset for missing data and handling it to ensure completeness and representativeness for analysis. Categorical variables were transformed into numeric format using label encoding with Python.
- The dataset was then organized to facilitate essential components like time duration, event indicators, and relevant covariates.

METHODOLOGY (CON'T)

- Kaplan-Meier Estimator is a non-parametric method used to estimate the survival function from lifetime data.
- Cox Proportional Hazards Model (Cox PH) is a semi-parametric model that relates the time until an event occurs to one or more covariates.
- Accelerated Failure Time (AFT) Model is estimates how covariates accelerate or decelerate the time to event.
- Akaike Information Criterion (AIC) helps in model selection by balancing model fit and complexity. A lower AIC indicates a better fit.
- Concordance Index measures a model's ability to rank predictions accurately. A higher C-index indicates better predictive ability.

DATA DESCRIPTION AND ANALYSIS

- **Demographic information:**
  Gender, college, and residence
- **Event of interest and duration**
  Churn and Level
- **Services Used:**
  Voice call, mobile data and sms texting
- **Factors influence discontinuation:**
  Multiple networks, network coverage, customer service, data allowance, high cost of services
- **Data Activity:**
  Data usage, exhaust monthly data

KAPLAN MEIER

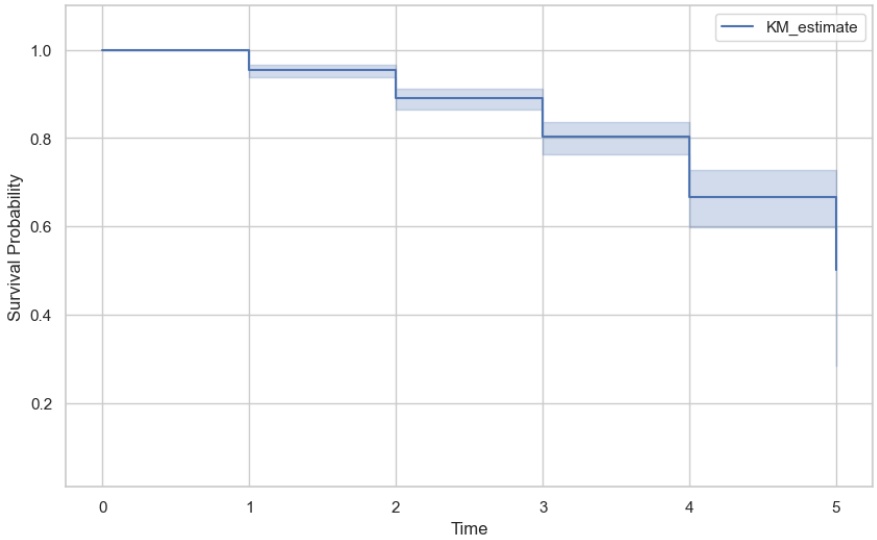

Figure: KM Curve

Event Time	Number at Risk	Number of Censored	Survival Probability	Lower Confidence Interval	Upper Confidence Interval
0	768	0	1.000000	1.000000	1.000000
1	768	35	0.954427	0.937099	0.967065
2	733	37	0.890799	0.864030	0.912565
3	696	33	0.802521	0.763671	0.835681
4	663	17	0.667443	0.597011	0.728409
5	646	2	0.500583	0.285044	0.682828

Figure: KM Analysis

ANALYSIS AND FINDINGS

Covariate	Weibull Coefficient	Weibull P	Cox PH Coefficient	Cox PH P
College	0.017	0.440	-0.03	0.51
Data Exhaustion	-0.176	0.102	0.41	0.14
Gender	0.218	0.006	-0.56	<0.005
High Costs Pricing	-0.06	0.423	0.16	0.36
Insufficient Data Allowance	-0.050	0.521	0.12	0.54
Mobile Data Internet	-0.118	0.335	0.32	0.24
Monthly Data Usage	0.026	0.374	-0.06	0.38
Multiple Networks	-0.104	0.546	0.21	0.62
Network Strength	-0.093	0.008	0.23	<0.005
Other Networks Better Services	-0.052	0.633	0.13	0.59
Poor Network Quality Coverage	0.071	0.378	-0.17	0.38
Residence	0.032	0.698	-0.09	0.65
SMS Text Messaging	0.039	0.600	-0.10	0.58
Unsatisfactory Customer Service	0.063	0.406	-0.15	0.40
Usage Freq	0.002	0.936	-0.00	0.96
Voice Calls	-0.080	0.421	0.16	0.48
Intercept (lambda)	2.112	0.00005	-	-
Intercept (rho)	0.906	0.00005	-	-

Figure: Cox and Weibull

ANALYSIS AND FINDINGS CON'T

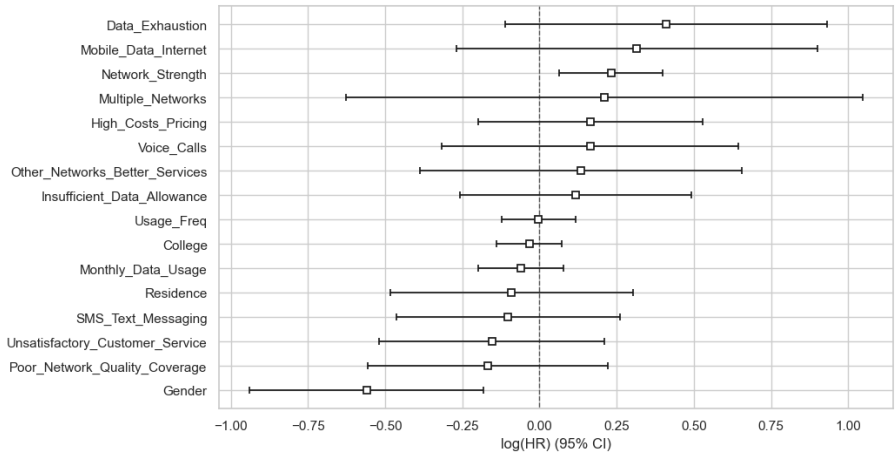

Figure: Cox Coefficients

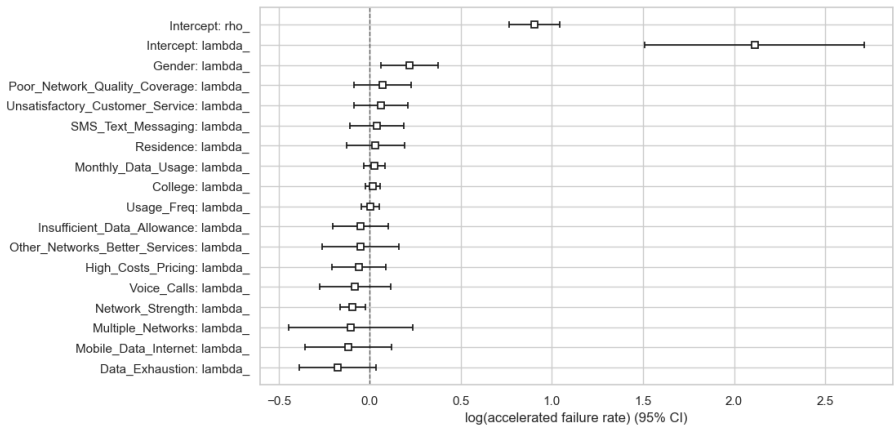

Figure: Weibull Coefficients

MODEL COMPARISON

Model	Concordance	AIC
Weibull	0.624	815.516
Cox PH	0.62	1479.47

Table: Model Concordance and AIC values

SUMMARY OF FINDINGS

- The research shows that Weibull AFT is the best model for both prediction and fit.
- High service costs, poor network coverage, and inadequate customer support drive churn. Competitive pricing, reliable network coverage, and responsive customer service improve retention.
- Data services have the highest churn rates thus indicating students' high value on reliable data. services.
- Younger students and those in their final year show higher churn rates. Gender differences are minimal and thus do not greatly affect churn rates.

RECOMMENDATIONS

- Increase the monthly data usage as most students to 10G.
- Enhance the quality and reliability of network coverage across KNUST to reduce churn rates.
- Improve the responsiveness and quality of customer support to address student concerns more effectively.
- Implement competitive pricing strategies and introduce loyalty programs to retain students.

Conclusion

THANK YOU