As The World Churns Customer Data, Business Models, and Data Science

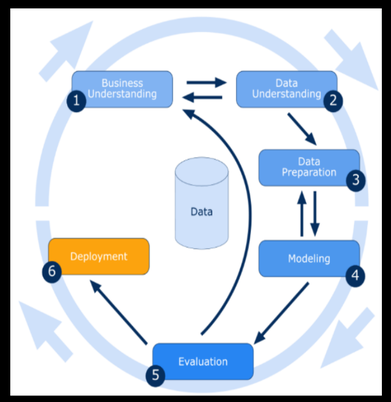

Figure 1. **The Crisp-DM Process**

WHY IS THIS DATA SCIENCE?

"DATA IS THE NEW OIL FOR ALL INDUSTRIES, AND DATA SCIENCE IS THE POWER THAT DRIVES THE INDUSTRY."

DATA SCIENCE TRANSFORMS RAW DATA INTO USEFUL INFORMATION. INDUSTRIES NEED DATA TO HELP THEM MAKE CAREFUL DECISIONS AND ARE USED IN ALMOST EVERY INDUSTRY. SOME OF THE MAIN SECTORS ARE HEALTH, FINANCE, BANKING, BUSINESS, STARTUPS, ETC. COMPANIES USE THE DATA TO ANALYZE THEIR MARKETING STRATEGIES AND CREATE BETTER ADS. THE INDUSTRY NEEDS DATA SCIENTISTS TO HELP THEM MAKE SMARTER DECISIONS.

LET US UNDERSTAND THE IMPORTANCE OF DATA SCIENCE IN OUR LIVES. GETTING A RIDE WITH UBER IS EASY. WE SIMPLY OPEN THE APP, SET YOUR PICK-UP AND DROP-OFF LOCATION, BOOK A TAXI, GETS PICKED UP AND PAY WITH YOUR PHONE. EACH TIME YOU BOOK A TAXI THROUGH UBER, YOU WILL RECEIVE AN ESTIMATED FARE AND THE TIME IT TAKES TO TRAVEL THE ROUTE. HOW CAN THESE APPLICATIONS DISPLAY ALL OF THE INFORMATION THEY DO? THE ANSWER IS THE DATA SCIENCE. USING DATA SCIENCE PREDICTIVE ANALYTICS, UBER CAN DETERMINE THE PICK-UP, DROP-OFF LOCATION AND ARRIVAL TIME.

TECHNOLOGY GIANTS SUCH AS FACEBOOK, AMAZON AND GOOGLE ARE CONSTANTLY WORKING IN THE FIELD OF MACHINE LEARNING AND DATA SCIENCE. DATA SCIENCE ENCOMPASSES PROCESSES SUCH AS PURGING, PROCESSING, AND ANALYZING DATA. A DATA SCIENTIST COLLECTS DATA FROM MULTIPLE SOURCES, E.G. FROM SURVEYS AND PHYSICAL DATA PLOTS. THEN, DATA WAS PASSED THROUGH STRICT ALGORITHMS TO EXTRACT IMPORTANT INFORMATION FROM THE DATA AND CREATE A RECORD. THIS RECORD COULD ALSO BE USED TO PARSE ALGORITHMS TO MAKE MORE SENSE.

DELIVERABLE

MAIN GOAL: REDUCE CUSTOMER CHURN BY 50%
GOALS HAS BEEN DIVIDED INTO 3 MAIN STATES NAMELY SHORT, MEDIUM- AND SHORT-TERM GOALS.
- THE SHORT-TERM GOAL IS TO REDUCE CHURN BY 20%
- THE MEDIUM-TERM GOAL IS TO REDUCE CHURN BY 20%
- THE LONG-TERM GOAL IS TO REDUCE CHURN BY 10%

BUILD MODELS USING VARIOUS MACHINE LEARNING ALGORITHMS LIKE NEURAL NETWORKS, DECISION TREE, K-MEANS, RANDOM FOREST AND LOGISTIC REGRESSION TO UNDERSTAND ROOT CAUSES OF CHURN AND IDENTIFY AT-RISK CUSTOMERS.
CREATE DASHBOARDS AND VISUALIZATIONS ON THE FLY TO ANALYZE CUSTOMER DATA WITH THE TABLEAU TO MAKE EXPLANATION EASY TO THE STAKE HOLDERS
TEST AND ANALYZE IDEAS GATHERED TO ENABLE COMPANY FOCUS RESOURCES ON CUSTOMERS THAT ARE HIGH LEAVING. ADOPT CUSTOMER FOCUS STRATEGIES TO RETAIN AND REDUCE THE RATE OF CUSTOMER CHURN.
REFINE AND REPEAT THE PROCESS QUARTERLY TO MEASURE PROGRESS

DATA MINING SECTION
DATA MINING USING TABLEAU(SHOWN WITH EXAMPLES AND GRAPHS/CHARTS)
WORKING WIT ALIASES
ADDING REFERENCE LINE
LOOKING FOR ANOMALIES
VALIDATING APPROACH
SOME ADVANCE TOPICS IN TABLEAU
MODELING

INTRODUCTION

AS OUR ECONOMY AND CORPORATIONS BEGIN TO OPERATE IN A GLOBAL CONTEXT, THERE HAVE BEEN INCREASING EFFORTS TO RETAIN CUSTOMERS. FREQUENT ACQUISITION AND LOSS OF CUSTOMERS IS DEFINED AS CUSTOMER CHURN AND HAS BEEN A PARTICULAR AREA OF FOCUS IN DATA SCIENCE. IT IS IMPORTANT TO BUSINESSES AS "IT IS DIRECTLY TIED TO FIRM PROFITABILITY" [4]. THE COSTS OF KEEPING A CUSTOMER ARE USUALLY LESS THAN THE COSTS OF RECRUITING NEW CUSTOMERS [21]. THIS IS WHY IT IS BECOMING INCREASINGLY IMPORTANT TO USE DATA SCIENCE TECHNIQUES AND ADVANCED ANALYTICS TO PREDICT WHICH CUSTOMERS ARE VULNERABLE TO LEAVING. IT CAN BE DIFFICULT TO DIFFERENTIATE BETWEEN CUSTOMERS WHO WILL RESPOND TO INTERVENTIONS AND THOSE WHO WILL NOT[4]. IN ADDITION, EXCESSIVE CUSTOMER TURNOVER CAN BE A SIGN OF POTENTIAL FRAUDULENT ACTIVITY. THIS IS COMPLICATED BY THE FACT THAT TECHNOLOGY CAN SERVE TWO PURPOSES TO BECOME CLOSER TO CUSTOMERS AS WELL AS ALIENATE THEM [21]. THERE IS ALSO THE RISK OF CUSTOMER CHURN ON CUSTOMERS THAT WERE WON BACK AFTER CHURNING ORIGINALLY. THIS MAKES THE SITUATION EVEN MORE COMPLEX TO FURTHER ANALYZE [16]. THERE ARE MULTIPLE FIELDS TO STUDY IN THIS; SOME ORGANIZATIONS USE PREDICTIVE MODELING BY STUDYING CUSTOMER BEHAVIOR WHILE OTHERS FOCUS ON MORE TRADITIONAL DEMOGRAPHICS (BEHAVIORAL ATTRIBUTES AND FINANCIAL CHURN PREDICTION SOURCE). TRADITIONALLY, THE DATA SCIENCE TECHNIQUE OF K-MEANS CLUSTERING IS USED TO DETERMINE RISK OF CUSTOMER CHURN (CLUSTERING PREDICTION TECHNIQUES SOURCE). HOWEVER, THERE ARE OTHER METHODS AVAILABLE TO HELP PREDICT THIS RISK. IN SOME INSTANCES, OTHER METHODS LIKE DECISION TREE ANALYSIS ARE MORE VALID AND THE FIELD CONTINUES TO BECOME MORE DIVERSE.

THE RISK CAN BE MORE THAN FINANCIAL; IN CERTAIN INSURANCE INDUSTRIES, CUSTOMER CHURN CAN SIGNIFY LOSS OF CRITICAL HEALTHCARE COVERAGE AND CAN SIGNIFICANTLY IMPACT A PERSON'S HEALTH. IN FACT, DATA SCIENCE TECHNIQUE AND PREDICTIVE ANALYTICS IN PARTICULAR ARE BEING APPLIED TO TREAT CANCER AND IMPACT HEALTHCARE OUTCOMES [17,20]. THEREFORE, IT

REFERENCES

ABBASIMEHR, H., SETAK, M., & SOROOR, J. (2013). A FRAMEWORK FOR IDENTIFICATION OF HIGH-VALUE CUSTOMERS BY INCLUDING SOCIAL NETWORK-BASED VARIABLES FOR CHURN PREDICTION USING NEURO-FUZZY TECHNIQUES. INTERNATIONAL JOURNAL OF PRODUCTION RESEARCH, 51(4), 1279–1294. HTTPS://DOI-ORG.EZPROXY.BELLEVUE.EDU/10.1080/00207543.2012.707342

AL-SHBOUL, B., FARIS, H., & GHATASHEH, N. (2015). INITIALIZING GENETIC PROGRAMMING USING FUZZY CLUSTERING AND ITS APPLICATION IN CHURN PREDICTION IN THE TELECOM INDUSTRY. MALAYSIAN JOURNAL OF COMPUTER SCIENCE, 28(3), 213–220. HTTPS://DOI-ORG.EZPROXY.BELLEVUE.EDU/10.22452/MJCS.VOL28NO3.3

AMIN, ADNAN & ANWAR, SAJID & ADNAN, AWAIS & NAWAZ, MUHAMMAD & ALOUFI, KHALID & HUSSAIN, AMIR & HUANG, KAIZHU. (2016). CUSTOMER CHURN PREDICTION IN TELECOMMUNICATION SECTOR USING ROUGH SET APPROACH. NEUROCOMPUTING. 10.1016/J.NEUCOM.2016.12.009.HTTPS://JOURNALOFBIGDATA.SPRINGEROPEN.COM/ARTICLES/10.1186/S40537-019-0191-6.RIS

ASCARZA, E. (2018). RETENTION FUTILITY: TARGETING HIGH-RISK CUSTOMERS MIGHT BE INEFFECTIVE. JOURNAL OF MARKETING RESEARCH, 55(1), 80–98. DOI: 10.1509/JMR.16.0163

BRUGNOLI-ENSIN, I., & MULLIGAN, J. (2018). INSTABILITY IN INSURANCE COVERAGE: THE IMPACTS OF CHURN IN RHODE ISLAND, 2014-2017. RHODE ISLAND MEDICAL JOURNAL.

DELLA TORRE, E., ZATZICK, C. D., SIKORA, D., & SOLARI, L. (2018). WORKFORCE CHURNING, HUMAN CAPITAL DISRUPTION, AND ORGANISATIONAL PERFORMANCE IN DIFFERENT TECHNOLOGICAL CONTEXTS. HUMAN RESOURCE MANAGEMENT JOURNAL, 28(1), 112–127. HTTPS://DOI-ORG.EZPROXY.BELLEVUE.EDU/10.1111/1748-8583.12167

FARQUHAR, J. D. (2005). RETAINING CUSTOMERS IN UK FINANCIAL SERVICES: THE RETAILERS TALE. THE SERVICE INDUSTRIES JOURNAL, 25(8), 1029–1044. DOI: 10.1080/02642060500237478

CONCLUSION

CUSTOMER CHURN IN THE FINANCIAL AND INSURANCE SECTOR IS HIGH. COMPANIES STRUGGLE TO IDENTIFY CUSTOMERS WHO ARE LIKELY TO LEAVE BEFORE THEY HAVE LEFT. SURVEYS ARE INFREQUENT AND A POOR SERVICE MIGHT NOT SHOW UP ON A SURVEY. TO INCREASE CUSTOMER LIFETIME VALUE (CLTV), ORGANIZATIONS NEED TO UNDERSTAND THE CORRECT BEHAVIORAL ATTRIBUTES AND BUILD PREDICTIVE MODELS USING NEW AND TRADITIONAL DATA SCIENCE TECHNIQUES LIKE K-MEANS OR SPATIO-TEMPORAL ALGORITHMS [14]. THIS HELPS IN SELECTING THE CORRECT BEHAVIORAL TRAITS BASED ON TRANSACTIONS AND OTHER DEMOGRAPHIC BEHAVIORS TO IDENTIFY CUSTOMER CHURN AND DETERMINE IF A CUSTOMER IS A GOOD CANDIDATE TO BE RETAINED. ONCE IDENTIFIED, MEASURES CAN BE UNDERTAKEN TO PREVENT CUSTOMER CHURN BEFORE IT IS TOO LATE.

%

%