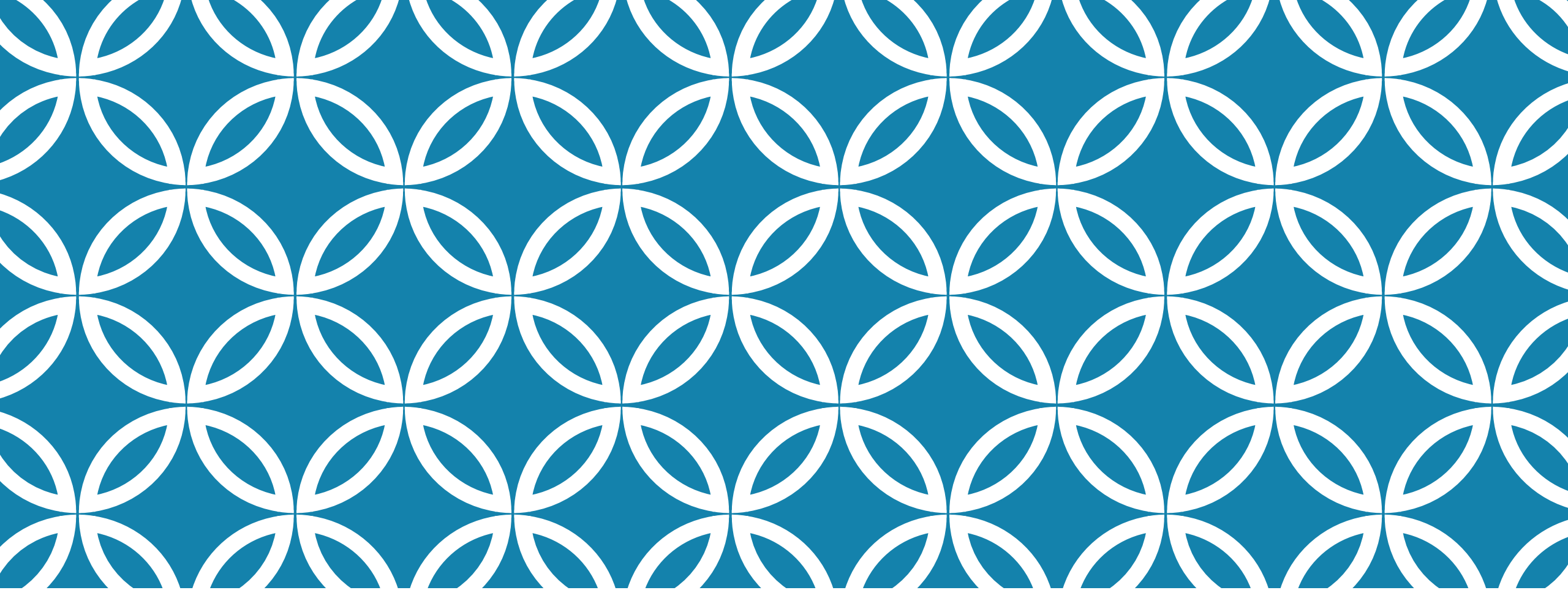

LEAD SCORING CASE STUDY

SUBMITTED BY – CHANDA SINGH

CONTENT

PROBLEM STATEMENT

AN EDUCATION COMPANY NAMED X EDUCATION SELLS ONLINE COURSES TO INDUSTRY PROFESSIONALS. ON ANY GIVEN DAY, MANY PROFESSIONALSWHO ARE INTERESTED IN THE COURSE LAND ON THEIR WEEBSITE AND BROWSE FOR COURSES. THEY HAVE PROCESSS OF FORM FILLING ON THEIR WEBSITE AFTER WHICH THE COMPANY TAKES IT AS A LEAD.

ONCE THE LEADS ARE AQUIRED, EMPLYOEES FROM THE SALES TEAM START MAKING CALL, WRITING E-MAILS, ETC.. THROUGH THIS PROCESS, SOME OF THE LEADS GET CONVERTED WHILE MOST DO NOT.

THE TYPECAL LEAD CONVERSION RATE AT X EDUCATION IS AROUND 30%. NOW . THIS MEANS IF, SAY, THEY ACQUIRE 100 LEADS IN A DAY, ONLY ABOUT 30%. OF THEM ARE CONVERTED . TO MAKE THIS PROCESS MORE EFFICIENT, THE COMPANY WISHES TO IDENTIFY THE MOST POTENTIAL LEADS, ALSO KNOWN AS HOT LEADS.

IF THEY SUCCESSFULLY IDENTIFY THIS SET OF LEADS, THE LEASD CONVERSION RATE SHOULD GO UP AS THE SALES TEAM WILL NOW BE FOCUSING MORE ON COMMUNICATING WITH THE POTENTIAL LEADS RATHER THAN MAKINGCALLS TO EVERYONE.

BUSINESS OBJECTIVE

LEAD X WANTS US TO BUILD A MODEL TO GIVE EVERY LEAD A LEAD SCORE BETWEEN 0-100. SO THAT THEY CAN IDENTIFY THE HOT LEADS AND INCREASE THEIR CONVERSION RATE AS WELL.

THE CEO WANT TO ACHIEVE A LEAD CONVERSION RATE OF 80%.

THEY WANT THE MODEL TO BE ABLE TO HANDLE FUTURE CONSTRAINTS AS WELL LIKE PEAK TIME ACTIONS REQUIRED, HOW TO UTILISE THE FULL MAN POWER AND AFTER ACHIEVING TARGET WHAT SHOULD BE THE APPROACHES.

PROBLEM APPROACH

IMPORTING THE DATA AND INSPECTING THE DATA FRAME.

DATA PREPARATION

EDA

DUMMY VARIABLE CREATION

TEST-TRAIN SPLIT

FEATURE SCALING

CORRELATION

MODEL BUILDING

MODEL EVALUATON

MAKING PREDICTIONS ON THE TEST SET