Hartej Kathuria Data Analyst Intern - Oracle Retail Bengaluru, Karnataka - Email me on Indeed: indeed.com/r/Hartej-Kathuria/04181c5962a4af19 Willing to relocate to: Delhi - Bangalore, Karnataka - Gurgaon, Haryana WORK EXPERIENCE Data Analyst Intern Oracle Retail - Bengaluru, Karnataka - June 2017 to Present Job Responsibilities: o As an intern part of the Global Retail Insights team at Oracle Retail, work involved creating a data oriented buisness case based using high level trends for various retailers using Excel and SQL. o Forecasting Sales with use of various statistical Modelling Methods using SQL and R o Market Basket Analysis using transactional data of retailers using SQL and R EDUCATION Statistics and Probability Manipal University May 2018 B. Tech in Electrical and Electronics in Embedded Systems MIT, Manipal University May 2016 SKILLS Python (2 years), SQL. (1 year), NOSQL (1 year), R (2 years), Machine Learning (2 years) PUBLICATIONS Post-operative life expectancy in lung cancer patients The objective of the project was to build an efficient predictive model based on a predefined dataset to predict whether the patient survives or dies within one year of the operation. The dataset given has 17 variables: 12 nominal, 2 ordinal and 3 numerical. The target variable has value true if the patient dies within one year of the operation else false if he survives. Tool used: R https://www.indeed.com/r/Hartej-Kathuria/04181c5962a4af19?isid=rex-download&ikw=download-top&co=IN   Predict the Happiness (Sentimental Analysis) The objective of this project was to build a binary classifcation model for the data provided by TripAdvisor consisiting of a sample of hotel reviews provided by customers.The model built can be used by them to understand the hotels listed by them.Tool Used: R Predict Network attacks The objective of this project was to build a multi-class classification model to predict the type of attack for an internet network company in Japan which has been facing huge losses due to malicious server attacks.The train dataset has 18 numerical features and 23 categorical features.The target variable has three classes.Tool Used: Python ADDITIONAL INFORMATION TECHNICAL SKILLSET \u2022 Languages & Technologies: Python, R, SQL, NoSQL, Predictive Modelling, Market Basket Analysis, Sentimental Analysis, Clustering, Bash Scripting (Preliminary), Socket Programming, Java (Preliminary) \u2022 Tools: R Studio, Jupyter, GIT, Sublime, MATLAB, Linux, KVM, Virtual Box, Open VZ, Oracle SQL Developer, MySQL, MongoDB, Excel