**Data Analysis of Sales Prediction of Big Mart Data**



A project paper is submitted for the degree of B.Sc. Examination 2018 in Computer Science and Engineering

**Project supervisor**

Dr. Shamim Ahmad

Professor

Dept. of Computer Science and Engineering,

University of Rajshahi

**Submitted BY**

Name: Md. Ismail Hossain

ID: 1510276139

Session: 2014-15

Exam: B.Sc. part-4 even semester

Dept. of Computer Science and Engineering,

University of Rajshahi

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

UNIVERSITY OF RAJSHAHI

**CHAPTER ONE**

**INTRODUCTION**

* 1. **Introduction**

The objective of this project is to analyze sales information of some popular e-commerce sites like amazon, e-bay, Alibaba, Coursera and Udacity. This analysis will produce some report on their sales’ pattern, relation between sales’ rate and price of the product etc. This type report may help to make a good business planning. Here I have chosen BigMart data for analysis.

The data scientists at BigMart have collected 2013 sales data for 1559 products across 10 stores in different cities. Also, certain attributes of each product and store have been defined. The aim is to build a predictive model and find out the sales of each product at a particular store.

Using this model, BigMart will try to understand the properties of products and stores which play a key role in increasing sales

* 1. **Why this project is being chosen**

This project will analyze the BigMart sales data and help to take better decision for future. This project will help to make better sales plan for future

* 1. **Objective**

The aim is to build a predictive model and find out the sales of each product at a particular store.

Using this model, BigMart will try to understand the properties of products and stores which play a key role in increasing sales.

* 1. **Organization of the project**

This project is organized in details in this project paper. The hypothesis generation and coding issue is illustrated in chapter 2. This chapter describe about the python 3.5, numpy and panda (Data centric python packages). In chapter 3 elaboration of analysis

1. Hypothesis generation
2. Data exploring
3. Data cleaning
4. Feature engineering
5. Model building
   1. **Conclusion**

The data analysis of online business site, social media etc. are drawing much attention of data scientists in recent days.

So it’s the time to learn data analysis for making better decision about the future.