NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY

(AN AUTONOMOUS INSTITUTION, AFFILIATED TO VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELGAUM, APPROVED BY AICTE & GOVT.OF KARNATAKA

LA PROPOSAL
Big Mart sales prediction Analysis Using Machine Learning Algorithms

Submitted in partial fulfilment of the requirement for the LA Component

Introduction to Machine Learning Course (18CSE71) *of 7th Semester*

Bachelor of Engineering

in

Computer Science and Engineering

Submitted by:

Abhay R	1NT18CS003
Suhas KM	1NT18CS168

Department of Computer Science and Engineering

ABSTRACT

We have collected Big Mart's 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, we are trying to understand the properties of BigMart's products and stores which play a key role in increasing sales.

INTRODUCTION

Machine learning is a subfield of artificial intelligence (AI). The goal of machine learning generally is to understand the structure of data and fit that data into models that can be understood and utilized by people. Although machine learning is a field within computer science, it differs from traditional computational approaches. In traditional computing, algorithms are sets of explicitly programmed instructions used by computers to calculate or problem solve. Machine learning algorithms instead allow for computers to train on data inputs and use statistical analysis in order to output values that fall within a specific range. Because of this, machine learning facilitates computers in building models from sample data in order to automate decision-making processes based on data inputs.

The machine learning methods that we will be using to perform the analysis of BigMart's sales are Linear Regression, Lasso, Ridge, Decision Tree Regressor, Random Forest Regressor and Extra Tree Regressor.

DATASET

Characteristics of the dataset:

Variable	Description
Item_identifier	Unique product id
Item_Weight	Weight of product
Item_Fat_Content	Whether the product is low fat or not
Item_Visibility	The % of total display area of all products in a store allocated to the particular product
Item_Type	The category to which the product belongs
Item_MRP	Maximum retail price of the product
Outlet_identifier	Unique store id
Outlet_Establishment_year	The year in which the store was established
Outlet_size	The size of the store in terms of ground area covered
Outlet_Location_Type	The type of city in which the store is located
Outlet_Type	Whether the outlet is just a grocery store or some sort of supermarket
Item_Outlet_Sales	Sales of the product in the particular store. This is the ourtcome variable to be predicted

Machine Learning Methods

1. Linear Regression
2. Ridge
3. Lasso
4. Decision Tree
5. Random forest
6. Extra Trees

Libraries

1. Pandas
2. Matplotlib
3. Seaborn
4. Scikit-learn

Presentation and Visualization

A graphical representation showing the analysis of the sales using the different machine learning methods.

Bibliography

https://github.com/aswintechguy/Machine-Learning-Projects/tree/master/Bigmart%20Sales%20Prediction%20Analysis%20-%20Regression