**SUMMER INTERNSHIP AT MIDHANI**

**2019**

A report submitted to

Mishra Dhatu Nigam Ltd (MIDHANI)

ON

**Text Mining and Sentimental Analysis of Social Media**

Presented By

A.N. VAISHNAVI

B-Tech (CSE) IIIT, Sricity

Mentor: LAKSHMI PRASANNA

**ABSTRACT**

This project includes Text Mining and Sentiment Analysis of Social Media. Firstly Sentiment Analysis is a procedure of computationally identifying and categorising opinions expressed in a piece of text, especially in order to determine whether the writer’s attitude towards a particular topic is positive, negative or neutral. The applications of sentiment analysis are broad and powerful. The ability to extract insights from social data is a practice that is being widely used by organisations across the world. It can also be an essential part of market research and customer service approach. Subsequently, Text Mining is carried out to filter the text to generate the polarity of the text (positive, negative or neutral).

In general, it can be great importance and can provide insights that can:

1. Improve campaign success
2. Improve product messaging
3. Improve customer service

**INSIGHTS OF THE PROJECT**

**Title:** Text Mining and Sentiment Analysis of Social Media

**Dataset**: Twitter, Boston Dataset-Linear Regression (Prediction of Housing Price)

**Technology Used**: Python language, Tweepy library, Text Blob, Pandas, Matplotlib, Scikit- learn.

**Procedure**: Firstly, exploring the Boston data set using .DESCR and then renaming its column names. Given a parameter or more the goal is to predict the outcome correctly. For instance here goal is to predict the housing prices using the given features.

Secondly, Scikit learn is used to fit linear regression to the entire data set and calculate the mean squared error. Then splitting the dataset into training and testing dataset and calculating mean squared error for the training data and test data.

Lastly, plotting the residuals for the training and test datasets.

Input:

.