**INTRODUCTION**

In the era which we live in, technology has a big impact on our lives. Artificial intelligence [6], knowledge engineering, Machine learning, Deep learning [4][5], Natural language processing[7][8] are emerging technologies which plays an important role in the leading projects of today's world. Artificial intelligence is an area or branch which aims or emphasizes on creating machine that works intelligently and their reactions is similar to that of human.

In Artificial Intelligence, Machine learning is an essential and core part providing the ability of learning and improving by itself. The focus of this technique is on creation of programs which can pick the data and learn from it by itself. Earlier, statistician and developers worked together for predicting success, failure, future etc. of any product. This process led to delay of the product development and launch. Maintenance of such product in the changing technology and data is also one of the major challenges.

Machine learning made this process easier and faster. There are various Machine learning algorithms broadly categorized into four paradigms:

* Supervised learning [7] [9] [10]: This learning algorithm provides a function so as to make predictions for output values, where process starts from analysis of a known training dataset. This algorithm can be applied to the past learned data to new data using labels so as to predict future events.
* Unsupervised learning: This algorithm is used on training dataset and informs which is neither classified nor labeled. It also studies to infer a function from a system to describe a hidden structure from unlabeled data. Clustering is an approach of unsupervised learning.
* Semi supervised learning [6] [11]: It takes the characteristics of both unsupervised learning and supervised learning. These algorithms uses small amount of labeled data and large amount of unlabeled data.
* Reinforcement [12]: In this algorithm, interaction is made to environment by actions and discovering errors. It allows machines and software agents in determining ideal behavior in a specific context such that performance could be maximized.

Regression and Classification problems are types of problems in supervised learning. In classification, conclusion is drawn using values which are obtained by observation. A discrete output variable say y is approximated by this problem using a mapping function say f on input variables say x. The output of classification is generally discrete but it can also be continuous for every class label in the form of probability. A regression problem has output variable as a real or continuous value. A continuous output variable say y is approximated by this problem using a mapping function say f on input variables say x. The output of regression is generally continuous but it can also be discrete for any class label in the form of an integer. A problem with many output variables is referred to multivariate regression problem.

In this paper we will be focusing on a problem picked from hackerrank where a company is trying to launch a new car modified on the basis of the popular features of their existing cars. The popularity will be predicted using machine learning approach. It can be classified as regression problem especially a multivariate regression problem and the problem can be classified under supervised learning. Thus various supervised learning algorithms will be used for this prediction.