**MINOR PROJECT 1**

**SYNOPSIS ON**

**Machine Learning Pipeline on Cloud**

**Submitted By:**

|  |  |  |  |
| --- | --- | --- | --- |
| Aditya Kumar | Anuj Verma | Dhananjai Kalra | Mayank Joshi |
| 500066319 | 500069910 | 500070904 | 500070105 |

**Under the guidance of**

**Gagan Deep Singh**

**Assistant Professor**

**Department of Computer Science**

**Department of Computer Application**

**UNIVERSITY OF PETROLEUM AND ENERGY STUDIES**

**Dehradun-248007**

**2020**

****

**Department of Computer Application**

**University of Petroleum and Energy Studies, Dehradun**

**Project Proposal Approval Form (2020 - 21)**

**Minor Project 1**

**Project Title: Predictive Analysis using Machine Learning**

**1. Abstract**:

The need for new technology such as Machine Learning and Data Science is on a rise today. Whilst technology was just starting to grow, a decade or 2 earlier, the methods, tools and algorithms developers created then are still being used today. And although these techniques were effective then, today it is a completely different scenario. Machines of this era have reached their limited potential. They can only get involved when they start to think for themselves. With this project we are embarking on this colossal feat of helping machines learn.

Our project is about developing a reselling platform for cars and hosting it on the cloud. This platform will be developed using python and will use machine learning algorithms to suggest best recommendations to the user. The project also consists of the elements of website development which will provide a user interface to our web app. It will ask for a significant amount of data regarding the selling points and price range of a vehicle. Based on this data and certain results through statistical and graphical analysis, the machine learning algorithm will predict the required output.

**1.1 Keywords**: Machine Learning, Decision Tree, Regression, Python Programming, Predictive Analysis, Data Science, Feature Importance, Feature Engineering, Visualization, Flask Framework, Cloud Technology, Cloud Platform, Web App, Vehicles, Reselling, Website, Version Control System.

**2. Introduction**:

The prospect of technology, society and change has been subjected to many contradictions over the years. Needless to say, over the past, these orthodox differences have been replaced with acceptance and harmony. Finally, with that outreach, we realize that the social climate of dependence over the machines today is not just a result of the changing belief but also a result of opportunities that they can bring to the table. Now, arguing over their social effects is something left for the philosophers to discuss, we as the practitioners of the field thrive to bring new and more advanced contraptions to achieve something greater. And one such vibrant call is the pursuit of Machine Learning

Machine Learning or ML, as it is known commonly, is the future. It is exactly as it sounds, giving machines the ability to learn, adapt and change accordingly. It might sound easy, but machine learning is quite a feat to accomplish. For years now developers and researchers have been busting their heads on this topic. And although we have found some success, we still have a long way to go.

One of the applications of machine learning is predictive analysis. Certain modules and algorithms implemented through Python Programming make this task comprehensible. The machine can learn the users’ previous choices and tends to make recommendations based upon what is learned. It’s a simple mechanism on the outside, but when one looks at the big picture it shows us that you are granting the ability to think and act to an entity that has been constructed out of screws, wires and chips. It’s almost magical and yet so attainable. Machine learning is the future of science, the future of computers, and the future of our modern vibrant society.

**3. Problem Statement**:

People require vehicles and they want them cheap and in the best condition possible. Simply because of that need, they go to dealers and websites trying to purchase second-hand vehicles for cheaper prices. Where the dealers will make a fool out of these customers on the face, websites do it by running a restless algorithm which instead of best prices and new products shows the same results again and again to different users. Now that is a problem because who doesn’t like variety? Who doesn’t want excessive choices leading to descriptive purchase? With these questions in mind, the need for a new and efficient practice based on machine learning seems of importance.

**4. The Dataset**: