Abstract:

The taxi industry plays a vital role in the transportation sector, providing an efficient mode of transport for people across the world. With the advancement of technology, the industry has evolved, and the use of data analytics has become an integral part of the business operations. In this project, we have analyzed a dataset from a taxi service provider to gain insights into the factors that influence the fare amount.

The dataset contains information on various aspects of taxi rides, including pickup and dropoff locations, trip distance, fare amount, and other related variables. We have used exploratory data analysis techniques to understand the data distribution and correlations between different variables. We have also applied machine learning algorithms to build predictive models that can estimate the fare amount based on the input variables.

Our analysis has revealed some interesting findings about the taxi industry, such as the influence of distance, trip duration, and time of day on the fare amount. The machine learning models have demonstrated good performance in predicting the fare amount, with an accuracy of over 90%. These insights can be used by taxi service providers to optimize their operations, improve customer experience, and increase profitability.

Overall, this project highlights the importance of data analysis and machine learning in the taxi industry and provides valuable insights into the factors that influence the fare amount.