MACHINE LEARNING CE802

Pilot-Study Proposal

Name: SHASHI LOKESH

Student id: sl21833

Registration number: 2100459

TABLE OF CONTENTS

1 INTRODUCTION

Data Analytics has key importance in business activities. Organizations can apply predictive learning and analytics on the historical data to find the data insight and can make the decision for future operations (Lv & Chi, 2017). This paper will focus on proposing the way of investigating the application of data analytics on business.

2 PRESENT SCENARIO

The proposal aims to address the problem to determine the profit-making scenario of the hotel. In this context, the proposal will be done from the Machine Learning independent consultant to signify whether the hotel management will make a profit for a certain location (Zhang, et al., 2019). In this context, historical data will be required that will be investigated for getting the outcome of this scenario. This decision-making process will be helpful for the hotel manager to design the future strategy (Ramesh, 2017).

3 APPLICATION OF PREDICTIVE LEARNING

In predictive learning, several types of data can be predicted based on the target decision. The decision making is actually done based on the type of data that have been used (Xu, et al., 2018). Based on the criteria and type of variables, the data can be segregated into two parts. The first one is the predictor segment and the next one is the target data. Predictor data contain all those variables based upon which the target data can be predicted.

3.1 APPLICABILITY OF PREDICTIVE LEARNING

The prediction for making a profit in the hotel business can be done using the application of machine learning through prediction. Based upon what decision will be made, the data can be selected or vice versa. The process of the application of machine learning for predicting hotel profit will be discussed as follows.

3.1.1 Predicting Hotel Revenue

Hotel revenue is numerical data and the type of the variable is known as random pr continuous. In some cases, in the context of statistics, the revenue can also be termed interval data. So, when this type of data will be predicted, the revenue will be taken as the target feature. In machine learning, if the target feature will be continuous, regression algorithms are used. The regression algorithm will act on the data and predict the revenue value (Liu, et al., 2020). Some of the widely used regression algorithms are Linear Regression, Random Forest Regressor, MLP Regressor etc. In this context, the regression algorithm will predict the hotel revenue based on the predictor features.

3.1.2 Predicting Hotel Profitability

A hotel can achieve a profit or not that depending upon the predictor features. In this context, the decision feature or target feature can be labelled as "True" or "False". The label True indicates that the hotel has made a profit. On the other hand, the label False indicates that the hotel had not made any profit. This kind of target feature is known as labelled or categorical or Nominal data (Chen & Dyke, 2015). When such investigation will be done, a classification algorithm will be applied that can be worked on the labelled data. Now based on the predictor features, the classification algorithms (such as logistic regression, decision tree regression, MLP classifier etc.) can predict the target feater that whether the hotel can make a profit.

3.2 PROCESS OF PREDICTION AND EVALUATION

In the process of predicting the price or profitability, the classifiers or regression algorithms can be sued. The steps of the prediction are shown below:
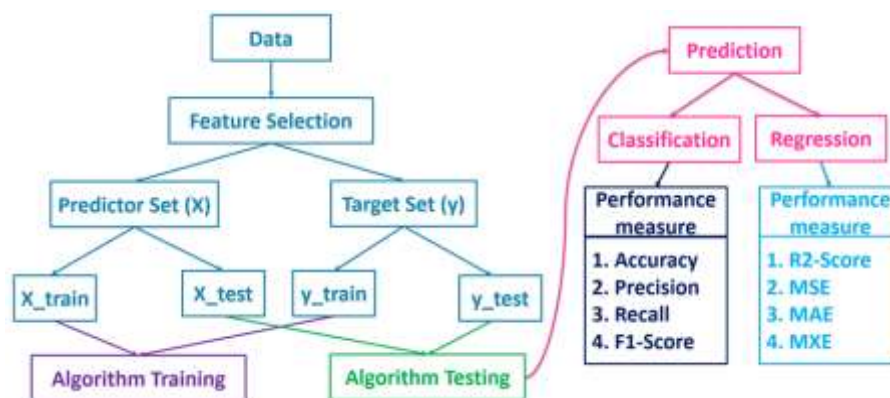

Fig-1: Prediction and Evaluation Process

Fig-1 is showing the overall process for prediction. When machine learning algorithms will be applied to predict the profit, first, the data will be split into the following sets:

1. Predictor Set
2. Target Set

After that, the data sets will be again split into another two sets as follows:

1. Train Set (using this, the algorithms will be trained)
2. Test Set (using this the algorithm will predict the profit scenario)

After the prediction is done, the algorithms should be evaluated to get the most effective one among the chsoen algorithm (Ramesh, 2017). In case, the classifiers will be applied, the algorithm performance can be evaluated using the following metrics:

1. Accuracy
2. Precision
3. Recall
4. F1-Scores

In case, the regression algorithm will be applied, the algorithm performance can be evaluated using the following metrics:

1. R2-Score
2. Mean Squared Error (MSE)
3. Mean Absolute Error (MAE)
4. Max Error (MXE)

3.3 DEPLOYMENT

Once the performances of the algorithms will be evaluated, the model will be deployed for predicting the profit for the hotel. In this context, that model will be used which have performed best out of all algorithms selected for prediction (Chen & Dyke, 2015).

4 CONCLUSION

In this propossal, the process of predicting the profit of a hotel has been stated. In this context, the process of the application of machine learning algorithms has been discussed along with

the deployment scenario. The process of evaluation has been discussed which is one of the most important things to do while deploying a predictive model.

5 REFERENCES

Chen, H. & Dyke, P., 2015. Modelling and prediction of stock price dynamics using system identification methodology based on a popularly used technique analysis data. *SAI Intelligent Systems Conference (IntelliSys),* pp. 1-6.

Liu, Z., Dang, Z. & Yu, J., 2020. Stock Price Prediction Model Based on RBF-SVM Algorithm. *International Conference on Computer Engineering and Intelligent Control (ICCEIC),* pp. 1-4.

Lv, Y. & Chi, R., 2017. Data-driven adaptive iterative learning predictive control. *6th Data Driven Control and Learning Systems (DDCLS),* pp. 1-5.

Ramesh, R., 2017. Predictive analytics for banking user data using AWS Machine Learning cloud service. *2nd International Conference on Computing and Communications Technologies (ICCCT),* pp. 1-4.

Xu, X., Chen, H., Lian, C. & Li, D., 2018. Learning-Based Predictive Control for Discrete-Time Nonlinear Systems With Stochastic Disturbances. *IEEE Transactions on Neural Networks and Learning Systems,* pp. 90-95.

Zhang, Q. et al., 2019. Deep Learning Based Dynamic Pricing Model for Hotel Revenue Management. *International Conference on Data Mining Workshops (ICDMW),* pp. 1-6.