

Amazon EC2 Spot Price Prediction using Regression Random Forests

Abstract— Spot instances were introduced by Amazon EC2 in December 2009 to sell its spare capacity through auction based market mechanism. Despite its extremely low prices, cloud spot market has low utilization. Spot pricing being dynamic, spot instances are prone to out-of bid failure. Bidding complexity is another reason why users today still fear using spot instances. This work aims to present Regression Random Forests (RRFs) model to predict one-week-ahead and one-day-ahead spot prices. The prediction would assist cloud users to plan in advance when to acquire spot instances, estimate execution costs, and also assist them in bid decision making to minimize execution costs and out-of-bid failure probability. Simulations with 12 months real Amazon EC2 spot history traces to forecast future spot prices show the effectiveness of the proposed technique. Comparison of RRFs based spot price forecasts with existing non-parametric machine learning models reveal that RRFs based forecast accuracy outperforms other models. We measure predictive accuracy using MAPE, MCPE, OOBError and speed. Evaluation results show that $MAPE \leq 10\%$ for 66 to 92% and $MCPE \leq 15\%$ for 35 to 81% of one-day-ahead predictions with prediction time less than one second. $MAPE \leq 15\%$ for 71 to 96% of one-week-ahead predictions. Index Terms— Amazon EC2, Compute instances, One-day-ahead prediction, One-week-ahead prediction, Regression Random Forests, Spot instances, Spot price prediction

ExistingSystem:

Spot pricing in particular is a pricing model targeted for divisible computing jobs that can shift the time of processing to when the computing resources are available at low cost. The primary requirement is that the applications must be time flexible, do not have a steep completion deadline and should be interrupt tolerant. Spot instances are also required for executing certain sudden tasks which do not need reserved instances.

.Disadvantages:

The ability to predict spot price lends itself to a variety of execution and Dynamic pricing model is not followed by any other cloud service provider. However, we do not see any issues in using our approach for forecasting spot prices of service providers who follow dynamic pricing policy; accuracy is very less.

ProposedSystem:

The objective of this work is to present and evaluate a predictive model for spot price prediction that can predict future prices with increased accuracy and speed, minimize forecasting errors and predict spot prices sufficiently far in advance to assist cloud spot users in bid decision making process with increased reliability. We compare prediction accuracy of the state of the art non-parametric supervised machine learning algorithms with Regression Random Forests (RRFs) model.

Advantages:

1. The purpose of proposed system is An analysis of the length of time epoch durations when spot price is less than on-demand price to raise users confidence level in opting for spot instances.

2. Spot price forecasting. We resort to machine learning based ensemble method namely RRFs for one-week-ahead and one-day-ahead spot price prediction. The approach focuses on both prediction accuracy and speed is high.

Modules:

We have used almost all of the same libraries which are used in normal ML/DL problems like pandas, numpy, matplotlib, sklearn etc.

Architecture:

Predicting spot prices for AWS EC2 Instances



SOFTWARE REQUIREMENTS:

OS	:	Windows
Python IDE	:	python 3.x and above

Jupyter Notebook,

Anaconda 3.5

Setup tools and pip to be installed for 3.6.x and above

HARDWARE REQUIREMENTS:

RAM : 4GB and Higher

Processor : Intel i3 and above

Hard Disk : 500GB: Minimum